Unveiling the optimal regression model for source apportionment of the oxidative potential of PM

Authors response

We thank the reviewers for their time and valuable comments that helped improve our manuscript's quality. We have answered the reviewers' comments in red and in *blue italic* are the changes included in the manuscript.

Anonymous Referee #1

The manuscript prepared by Thuy et al. deconvolved the source of OP^{DTT} and OP^{AA} for 1-year PM_{10} samples collected at six sites in France using receptor models (EPA PMF) coupled with eight separate regression methods. The relationship among source, composition, and OP is complex. The importance of apportionment methods on OP repartition by different PM sources was investigated in a variety of environments. Potential impacts from source variation on OP apportionment were also evaluated. The study demonstrated that different monitoring sites may have different optimal OP SA regression methods. By performing in-depth analysis of PM chemical composition and OP source, the authors further provided recommendations on how to select a suitable OP source apportionment model based on specific data characteristics and offered the possibility of real-time OP monitoring and accurate source identification in future studies. Overall, the manuscript was within the scientific scope, but there remain some unresolved issues. Therefore, I would like to recommend this manuscript to be published in Atmos. Chem. Phys. once the following concerns can be fully addressed.

Specific comments:

The authors only applied the PMF-regression methods to OP^{DTT} and OP^{AA}. There are many other acellular OP assays. The authors did not discuss the applicability of the current source apportionment method to other acellular OP assays. For example, whether the SA optimization framework proposed here can be applicable for OP^{GSH}, OP^{ESR}, OP^{DCFH}, etc. At least this should be discussed in the limitation part at the end of the manuscript.

Reply: Thanks for your suggestion. The present study only focused on the 2 most popular OP assays, mostly because this is already demonstrative of our purpose, testing the applicability of a large set of deconvolution methods for these two widely used assays. Further, we do not think that there is any data set published that present a full comparison of PM sources study with PMF for 6 sites of different typology for a full year, associated with OP measurement with more than 2 assays. Our group already published a study of OP source apportionment at one site with 5 different OP assays (OP_{AA}, OP_{DTT}, OP_{DCFH}, OP_{OH}, OP_{FOX}) by using MLP and WLS models, showing that these regression methods are applicable for all these assays (https://doi.org/10.1039/D3EA00007A). We understand that the work proposed by the reviewer would be interesting and valuable, but it represents an amount of work out of

our reach at the moment, and thus, we have updated the limitations of the study as follows (line 661 to 664):

This study only focused on the two most popular OP assays of PM10 (OP$_{DTT}$ and OP$_{AA}$). However, there are actually various OP assays, such as OP$_{DCFH}$, OP$_{OH}$, OP$_{FOX}$, OP$_{GSH}$, OP$_{ESR}$ and different sizes of PM (PM1, PM2.5, PM5). Further research should include more OP assays, which can be helpful in evaluating the performance of various regression models for different OP and different PM sizes.

Another limitation of the present work is the authors should be aware of the differences between intrinsic particle-bond PM OP measured/discussed here and in vitro ROS release. What kind of implications that the current work can help elucidate the health outcomes from PM exposure? The authors may want to add some specific discussions more about the implications of future applying such OP SA method in real-time OP monitoring and PM health impacts prediction.

Reply: Thanks for this comment. We agree with the reviewer that the OP measurement is only one aspect (the exogenous one) of the formation of the in vivo ROS in human, the other one being the endogenous reactions that take place in the body. This is clearly acknowledged in the research community working on OP measurement of PM (Brook et al., 2010; Tao et al., 2003). It is also recognized that measuring OP, with this relevant concept of the quantification of the release of oxidative species to the body, is probably one step forward for a better understanding and prediction of the health impacts compared to the measurement of the sole PM mass. Studies have shown the associations between OP and toxicological assays on epithelial cells (Daellenbach et al., 2020; Leni et al., 2020). All the more, several recent studies tend to indicate that OP may be more predictive of health effects, including several performed in our group that found positive associations between OP$_{DTT}$ and respiratory outcomes in the little childhood (L. J. S. Borlaza et al., 2023; Marsal et al., 2023) and with acute respiratory outcomes in Bolivia (Borlaza et al., 2024 under review). However, this approach needs to assess long time series better to understand the associations between OP and health outcomes.

Further, just like it is for the PM mass, unraveling the different sources of the PM that are important for the health impact due to their OP values is quite important both for a better assessment of epidemiology studies, but also for the implementation of their regulation. Future studies may apply the best deconvolution of OP of PM sources by the different methods discussed in our work in order to provide the most accurate long-term time series of OP exposure per source to evaluate their associations to specific health outcomes.

We believe that our work also paves the way for the coming applications for future near-real-time (NRT) OP monitoring and the possibilities it opens. The fact that current NRT PM chemical measurements are mainly performed for PM1 fraction is evidently something to be considered, since it changes the perception of the PM sources (cutting off most of the sources emitting the in coarse mode) and provides a more difficult framework for PMF studies. Our team already tried to tackle these difficulties, with the coupling of NRT chemical measurement with off line 4-hr OP measurements (Camman et al., 2023).

The authors developed this OP source apportionment method for PM10 samples and proposed online OP monitoring by the end of this manuscript. Most online PM chemical composition monitoring techniques are designed for smaller particles (PM1 or PM2.5). Have the authors testified the robustness

of this OP SA method with PM of smaller sizes? In other words, the authors should be aware of this and give some discussion over the potential impacts of PM size on the current OP SA method.

Reply: This comment is very similar to the end of the previous one. We agreed with the referee that the PM size could impact the OP SA since the sources were changed by the PM faction. For instance, the tracers of traffic non-exhaust and primary biogenic cannot be appropriately identified in the PM1 fraction. However, we present here different statistical models that describe the OP based on the PM sources. These kinds of models can be applied to all sorts of particle sizes, the only condition is the robustness of PMF-derived sources. Our group recently published an article about the OP SA of PM1, which showed that the MLR model works well for apportioning PM1 (Camman et al., 2023) (https://doi.org/10.5194/acp-24-3257-2024). In the present study, we have not tested the performance of the OP SA method for different PM sizes. Nevertheless, we are looking forward to test the OP SA methods developed in this study with the sets of online data (chemistry + OP) that we intend to collect in a near future. The limitation of the study is updated as follows (line 661 to 664):

This study only focused on the two most popular OP assays of PM10 (OP DTT and OP AA). However, there are actually various OP assays, such as OP_{DCFH}, OP_{OH}, OP_{FOX}, OP_{GSH}, OP_{ESR} and different sizes of PM (PM1, PM2.5, PM5). Further research should include more OP assays, which can be helpful to evaluate the performance of regression models in different OP and different PM sizes.

The division of 80% for training and 20% for testing has any references? Have the authors performed sensitivity analysis on other training vs testing percentages?

Reply: The training set is the dataset used to build and fit predictive models, while the test set is a subset of the dataset to assess the likely future performance of a model. The division of 80% for training and 20% for testing in our study is based on Patero principle (Dunford et al., 2014), which demonstrates that 80% of consequences come from 20% of the causes. The 80/20 split is widely used and the same as the split obtained in each iteration of 5-fold cross-validation. Our main concern was to ensure our results were robust to different train/test splits (with the same train/test ratio); this is why we performed 500 random 80/20 splits for each experiment and report the variability across the 500 repetitions. We did examine a 70/30 train/test split, whose results did not clearly differ when accounting for the variability across the 500 random splits.

OPv is significantly affected by PM mass concentration while OP_m (mass normalized OP) reflects chemical composition. This may be another piece of work but have the authors looked into the OP SA method on OPm for those PM10 samples? Some discussions are needed.

Reply: Thanks for your comments. Since the goal of the paper is ultimately for the quest of the best deconvolution of the PM sources for exposure, we of course focused on OP_v. In addition, the PMF-derived sources are the contribution of sources to mass concentration of PM_{10} which are normalized with the air volume sampled ($\mu g\ m^{-3}$), while the OP_m is the intrinsic OP property of 1 μg of PM. These measurements are not in the same state (one normalized by air volume and the other normalize by PM_{10} mass) and we could not perform an OP SA on the OP_m.

Line 39: Crouse et al., 2012 may not be the right citation for this statement

 Reply: Thanks for your comment. We updated the reference (line 39)

However, various research shows that the impacts of PM also depend on other factors such as chemical composition, size distribution, particle morphology, and biological mechanisms (Brook et al., 2010).

Line 42: any reference for the OP definition?

Reply: Thanks for your clarification, we added the reference (line 45).

The quantification of the PM capacity to generate ROS into a biological media is called oxidative potential (OP) (Bates et al., 2019; Daellenbach et al., 2020; Dominutti et al., 2023).

Line 46: the most prone to

Reply: Thanks for your revision, we updated the sentence (line 47):

The relationship between PM chemical components and OP activities may identify which components are the most prone to generate ROS.

Line 55: any references for this

Reply: Thanks for your comment, we added the references (line 57-58)

Regression analysis is the most common and effective way to estimate the redox activity of receptor model-derived PM sources (Borlaza et al., 2021; Dominutti et al., 2023; Fadel et al., 2023; Li et al., 2023; Liu et al., 2018; Weber et al., 2018; Zhang et al., 2023).

Line 73: the first place of OLS should be in Line 60

Reply: Thanks for your comment, we updated in the main text as follows (line 62 – 64):

Numerous regression models can be used for such OP source apportionment (SA), with multiple linear regression fitted by ordinary least squares (OLS) being the most common regression technique (Bates et al., 2015; Deng et al., 2022; Li et al., 2023; Liu et al., 2018; Shangguan et al., 2022; Verma et al., 2014; Y. Wang et al., 2020; Yu et al., 2019)

Line 185-280: the introduction of 8 regression models can be moved to supplemental information. Method 2.5 and 2.6 can be combined

Reply: Thanks for your suggestion, we thought that it was crucial to show all the methods that we applied together with the condition and assumption of the regression model. That elucidates some conditions that should be respected when using the model for OP source apportionment. Hence, we want to put the description of the methods in the main text.

We combined section 2.5 and 2.6.

Line 303: be specific on what are the 2 OPs

Reply: Thanks for your comment. We clarified in the main text (line 317-319)

Conversely, for the alpine valley sites, CHAM presents higher OP_{AA} than OP_{DTT}, while GRE-fr experiences similar levels of OP_{AA} and OP_{DTT}.

Line 349: this is a bit misleading. In the methodology the author claimed 8 regression models but here you added 3 weighted models so 11 in total actually—this should be clearly stated in the above method section and Figure 1

Reply: Thanks for your suggestion. These are always 8 regression techniques, but some of them were applied twice, with and without weighting (PLS, Ridge, Lasso). We clarified in the main text (line 101-104).

These 8 regression techniques were applied to find the relationship between OP and PM sources, however, PLS, Ridge, and Lasso were performed 2 times, with and without weighting, consequently, there are 11 results of regression techniques that will be presented.

Line 398: Any inner relationship between collinearity and heteroscedasticity?

Reply: There is no connection between collinearity and heteroscedasticity; one or both can be present. In addition, a study by Alabi et al. 2020 (http://dx.doi.org/10.4236/ojs.2020.104041) showed that statistical tests for heteroscedasticity may fail in the presence of collinearity.

Line 432: any explanations for the larger differences among small OPAA sources? Also, previous work showed that OPAA is metal sensitive (DOI: 10.1021/acs.est.8b03430), I am not sure why the dust source plays a negative role in OPAA

Reply: Thanks for your comments. We agreed that OP_{AA} is sensitive to some metals, including Cu, Fe, Pb, Zn, and Mn (Bates et al., 2019; Calas et al., 2018). Nevertheless, the dust source in our study is mainly mineral dust, which contains a high proportion of Ca^{2+}, Al, and Ti. These metals are known as low redox-active species whose correlations with OP_{AA} are lower than 0.3 (Calas et al., 2018). In addition, the negative intrinsic OP_{AA} of dust is not surprising and was found in other studies in France (Dominutti et al., 2023; Weber et al., 2018).

Figure 6 and Figure 7: it is hard to believe salt, nitrate and sulfate rich sources are significant contributors of OPAA and OPDTT given that sulfate and nitrate cannot consume AA or DTT

Reply: Thanks for your comment. Secondary inorganic aerosol (SIA) has been shown to have less effect on ROS generation (Daellenbach et al., 2020). However, the sulfate-rich and nitrate-rich in NIC were identified with 15-25% of metals (As, Cd, Mn, Mo, Ni, Pb, Ti, V, Zn) in their chemical profiles, so, they are not solely sulfate or nitrate. The presence of these metals could be the reason that makes a higher contribution of sulfate and nitrate rich to OP_{AA} and OP_{DTT}. The high intrinsic OP of sulfate rich and nitrate rich in PM_{10} is also found in previous studies (D. Wang et al., 2022; Weber et al., 2018). The NIC sampling site is located near the port area of Nice where there are many shipping activities. The shipping emissions, including EC, V, Ni and some metals, could be transported with sea salt as well as aging particles such as SIA. The chemical profiles of salt, sulfate-rich and nitrate-rich in NIC are added in Figure S.20, S.21, S.22 in the Supplement, respectively.

Line 473: Any references for this 60% threshold? It may be too rush to draw the 60% conclusion here given that the mechanisms are not fully elucidated. Some earlier work on synergistic and antagonistic OP can also be added here, https://doi.org/10.1021/acs.est.7b01272; https://doi.org/10.5194/acp-18-3987-2018

Reply: Samake et al. (2017) demonstrated that the presence of bacterial cells in aerosol decreasing the redox activity of Cu and 1,4-naphthoquinone, with a maximum decreasing of 60% compared to the oxidative reactivity considered individually. Pietrogrande et al. (2022) indicated that the mixture of Cu, Fe, 9,10-phenanthrene quinone and 1,2-naphthoquinone reduces the rate consumption of AA and DTT,

up to 50% depending on the quantity of each chemical. Wang et al. (2018) reported that the mixing of Cu and naphthalene secondary organic aerosol (SOA) and phenanthrene SOA only got half of DTT rate consumption compared to the consumption when considering separately. Xiong et al. (2017) showed the presence of antagonists in the interaction of Fe and quinones, nevertheless, much lower than those in the other studies (under 10%). These references reported that antagonistic effects of a mixture can significantly reduce the consumption rate of OP_{DTT} and OP_{AA}, and this impact varies widely from 10% to 60% depending on the type of chemical species and the quantity of each species in the mixture. We selected the maximum antagonistic effect (60%) based on these studies to describe the possibility of getting the intrinsic OP negative, although it definitely exists many kinds of mechanisms behind this effect that we cannot simulate. We added these references in the main text as follow (line 485-489):

Second, we postulate that negative intrinsic OP values are possible since previous studies have reported that total PM intrinsic OP can be modulated due to the synergetic/antagonistic effects involving, for example, soluble copper, quinones, and bacteria (Borlaza et al., 2021; Pietrogrande et al., 2022; Samake et al., 2017; S. Wang et al., 2018; Xiong et al., 2017).

Line 580: the authors pointed out that one of the limitations of the present work is the type of regression method used and recommended more methods should be applied in future studies. Can the authors give some examples of future regression methods here?

Reply: Thanks for your comments. We added more details in the limitations section (line 639-645).

This study compares eight regression models but is not exhaustive; further research could add more regression techniques to evaluate result variations across models. The potential techniques that could be applied for OP SA are gradient boosting techniques for resolving regression models, or supervised machine learning techniques which allows the investigation of linear and non-linear regression relationships. However, the consistently strong performance of ordinary linear regression across six locations in France suggests that there may be little to gain from applying more complex models in areas with similar PM_{10} sources.

Line 590: writing of this paragraph, and the whole manuscript can be improved.

Reply: Thanks for your suggestion. We rewrote this paragraph (line 655-660).

Observations ranged between 100 and 200 samples at each site, which may be insufficient to obtain a fair performance of GLM, decision trees and neural network models even though this number of samples is sufficient to address SA through the PMF model for offline analyses. Therefore, this study outlines well the limitations of GLM, RF, and MLP for offline datasets. Future investigations should be performed in an extended dataset, such as long-term or real-time measurement data, to investigate the performance of machine learning algorithms.

Anonymous Referee #2

This manuscript explores the relationship between PM10 sources and its oxidative potential (OP) at six sites in France, utilizing PMF for source apportionment and machine learning to estimate each source's intrinsic OP. By examining source characteristics, the authors devise a protocol for selecting the best

regression model for specific datasets, enhancing consistency in intrinsic OP values across sites compared to a reference model.

The study insightfully addresses the choice of the optimal machine learning model based on dataset characteristics like collinearity and heteroscedasticity. The rigorous strategy for optimal source selection and the comparison of the best and reference algorithms' intrinsic OP support the selection criteria.

However, the methodology lacks clarity on model testing and validation protocols, and intrinsic OP levels for RF and MLP models are absent. The paper does not discuss extrinsic OP, which may diminish its environmental relevance. Further discussions on algorithm selection based on volume-normalized OP could enhance the manuscript's utility for future research on source contributions to OP.

Overall, this well-conducted study merits publication with minor revisions, addressing suggested improvements for a more comprehensive presentation and analysis.

Specific comments are as follows:

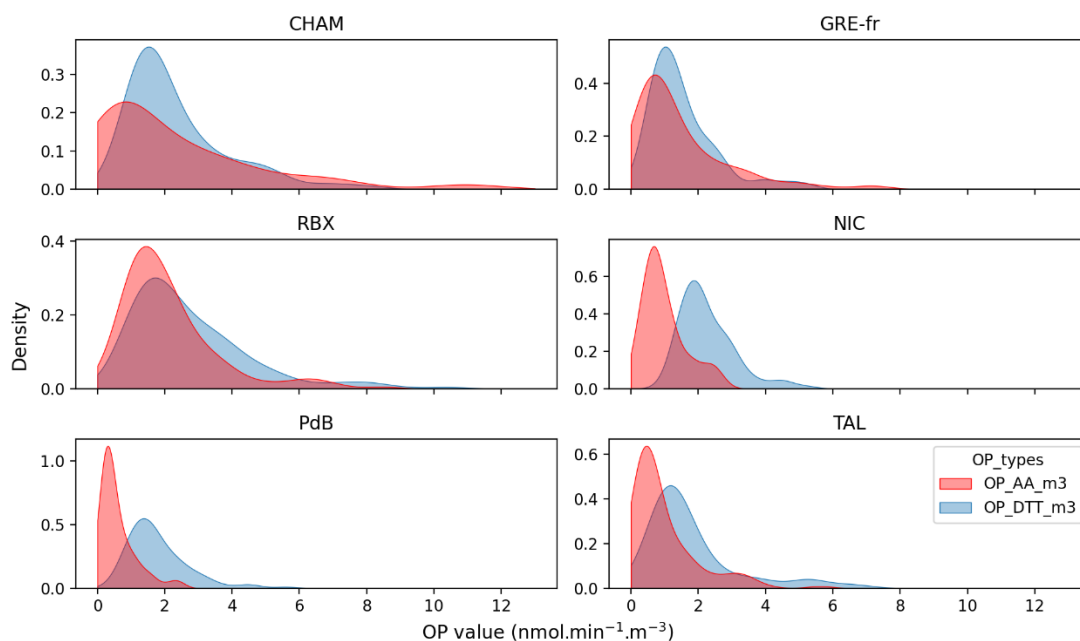
The authors mentioned an 80-20 train-test split method. For optimal hyperparameter development, a validation set is crucial. Did the authors incorporate a validation phase beyond testing?

Reply: Thanks for your question. Yes, we used 5-fold cross-validation on the training data for hyperparameter tuning. The training dataset was separated into 5 parts: 4 parts were used for training, and the remaining was used for validation. This process was repeated 5 times, and the hyperparameter value producing the lowest mean RMSE across the 5 parts was selected. This is now explained in the text as follow (line 265-270 for RF and line 288-291 for MLP):

RF is customizable via hyperparameters such as the number of trees, the size of the bootstrap sample, and the number of features to evaluate at each node. The hyperparameters tuning used 5-fold cross-validation on the training data for hyperparameter tuning. The training dataset was separated into 5 parts: 4 parts were used for training, and the remaining was used for validation. This process was repeated 5 times, and the hyperparameter value producing the lowest mean RMSE across the 5 parts was selected. The hyperparameters tuning is shown in section S1.1 Supplement.

The choice of hyperparameters to ensure the MLP model's robustness is processed by hyperparameter tuning using 5-fold cross-validation and shown in section S1.2 of the supplement. Thanks to hyperparameter tuning, the two hidden layers and a logistic sigmoid activation function were selected in this study to capture the non-linear relationships between OP activities and PM sources.

For the GLM model, a logarithmic transformation of the dependent variable showed poor predictions. Were alternative transformations, such as inverse or power functions, considered?

Reply: We explored a logarithmic transformation because the distribution of OP_v was approximately log-normal (Figure below). Given the strong performance of the (ordinary) linear regression models, we did not explore other transformations, but this could be an interesting point for future studies.

The authors noted insufficient sample numbers for robust results. Did combining similar sites improve predicted intrinsic OP?

Reply: This is a point that we have considered, in order to improve the performance of the MLP and RF. However, as shown in Table S2 Supplement, the sites with the same typology do not have the same PM sources. Combining these sites potentially risk losing some specific sources in a site, for example, HFO in PdB or Industrial in PdB, TAL, GRE, which are known to have a high redox activity in literature. However, as said in the conclusions, this is something to be tested in future research for longer time series data.

RF and MLP's poor test results might result from unoptimized hyperparameters. Consider further tuning and a dedicated validation set for improvement.

Reply: Thanks for your suggestion. The hyperparameters are shown in Table S.11, Supplement. The hyperparameters of the best performance are different across sites; however, the change in the hyperparameter did not change much the model accuracy (under 0.1 of accuracy score). Further, running the model is very time-consuming: 500 runs for each model per site take over 48 hours to get the result. In addition, given the good performance of OLS, RF is unlikely to give much better results on the dataset of this study. The extensive parameter tuning therefore could be a good direction for further studies.

Intrinsic OP might not be directly obtainable from RF and MLP but could be inferred from source importance and mass contribution.

Reply: The feature importance of RF and MLP could be assessed to understand which feature (or PM sources in our case) is vital in predicting OP. This study estimated the permutation importance of each PM source as the mean increase in the mean squared error of predicted OP when the values of the PM source were permuted. However, the feature importance does not represent the intrinsic predictive value. In addition, the presence of a source could improve the accuracy, but we cannot know precisely if the interaction between OP and this source is negative or positive, i.e. OP value increases or decreases as

the contribution of a source increases. For all these reasons, we decided not to include the feature importance result in the paper. Indeed, a SHapley Additive exPlanations analysis was performed but not included in this paper because the overall performances of RF and MLP were not good.

Adding extrinsic OP, normalized by air volume, could clarify environmental implications across models.

Reply: Thanks for this comment. We agree with the reviewer that the evaluation of volume-normalized OP could bring other insights into the environmental implications of OP. However, this includes the mass of each source, which is not similar between the sites evaluated since it depends on the total PM mass. We then focused on the intrinsic OP obtained for each PM source for comparison purposes.

The MAE and RMSE equations (line 290) are incorrect.

Reply: Thanks for your comment. We corrected the equations (line 306 and 307).

$$MAE = \frac{\sum_{i=0}^{m}|y_i - \widehat{y_i}|}{m}\ (Eq\ 13)$$

$$RMSE = \sqrt{\frac{\sum_{i=0}^{m}(y_i - \widehat{y_i})^2}{m}}\ (Eq\ 14)$$

It's recommended to display MAE and RMSE values for training data in Figure 5, alongside test data.

"OP intrinsic" (line 453) should be revised to "intrinsic OP."

Reply: Thanks for your comment. The MAE and RMSE values (mean ± std) for training and testing dataset are shown in Table S.4 and S.5 in the Supplement, respectively. We did not include the MAE and RMSE values of training dataset in Figure 5 because it is dense and invisible to put the both training and testing results on this plot. In addition, we selected the best model based on the performance of the testing, so that we wanted to show the values of testing dataset.

We updated the line as follow (line 468):

The comparison of intrinsic OP among regression models in NIC demonstrated that OP_{DTT} and intrinsic OP_{AA} values exhibit variation across different models with and without weighting, illustrating that the choice of the model significantly influences the values obtained for intrinsic OP of PM sources (A similar pattern is observed for all other sites and shown in Fig S.3 to Fig S.7 for OP_{AA} and Fig S.8 to S.12 for OP_{DTT}).

The sentence spanning lines 474-476 requires clarification.

Reply: This was already discussed in a question of the referee #1. The main text is updated as follow (line 488-499)

Samake et al. (2017) demonstrated that the presence of bacterial cells in aerosol decreasing the redox activity of Cu and 1,4-naphthoquinone, with a maximum decreasing of 60% compared to the oxidative reactivity considered individually. Pietrograndé et al. (2022) indicated that the mixture of Cu, Fe, 9,10-phenanthrene quinone and 1,2-naphthoquinone reduces the rate consumption of AA and DTT, could be up to 50% depending on the quantity of each chemical. Wang et al. (2018) reported that the mixing of Cu and naphthalene secondary organic aerosol (SOA) and phenanthrene SOA only got half of DTT rate

consumption compared to the consumption of considering separately. Xiong et al. (2017) showed the presence of antagonists in the interaction of Fe and quinones, nevertheless, much lower than those in the other studies (under 10%). These references reported that the antagonistic of a mixture reduces the consumption rate of OP_{DTT} and OP_{AA}, and this impact varies widely from 10% to 60% depending on the type of chemical species and the quantity of each species in the mixture.

Anonymous Referee #3

Multilinear regression models have been commonly used to bridge PM mass sources obtained from PMF with PM oxidative potential (OP). However, a comprehensive assessment and comparison of those models has yet to be conducted. Thuy et al. evaluated the performances of eight regression techniques in estimating the contribution of PM10 sources to PM10 OP (OP_{AA} and OP_{DTT}). From the evaluation, a flowchart was established as a guideline for regression model selection. However, a few concerns shall be addressed below:

Most of my major concerns are related to the regression models:

In Figure 6 and Figure 7, results from WLS were identical with wRidge; while OLS were the same as Ridge. From Figure S3 to Figure S12, WLS and Ridge are equal, and OLS and wRidge are the same. I felt that this is very strange. How could the regression models have the same results, and how could the model similarity trend in the main manuscript and SI be different?

Reply: Thanks for your comments. The first thing we should clarify here is in Figure 6 and Figure 7, the "wRidge" is the model incorporating the weighting, and "Ridge" is the model without weighting. From Figure S3 to Figure S12, "Ridge" is the model incorporating the weighting and "Ridge wo weight" is the model without weighting. We are sorry for the confusion, all figures (Figure S3 to Figure S12) in the Supplement were corrected to the same name as Figures 6 and 7. Overall, for all sites, the results from WLS were identical with wRidge, and the OLS were the same as Ridge.

The results for OLS and Ridge as well as for WLS and wRidge were very similar because we used a small value (0.01) for lambda, the regularization parameter of ridge regression. As shown in Eq9, the minimize term of Ridge will become OLS when $\lambda = 0$. The lambda was chosen by implementing hyperparameter tuning, with λ in (5, 1, 0.5, 0.1, 0.01, 0.005, 0.001, 0.0005, 0.0001).

In Figure S6, the dust in Lasso has some contribution, but why the contribution became 0 in wLasso? A detailed explanation of those models incorporated weighting is required.

Reply: Thanks for your comment. The different result between the model applying the weighting and without weighting is not surprising, as shown in Figures 6 and 7 for NIC and Figures S3 to S12 for remaining sites, both OP_{DTT} and OP_{AA}. The model without weighting treats every data point differently; the point with lower uncertainty in OP analysis has more effect on the model. The model without weighting treats every data point similarly. That is the reason why we get different results for Lasso and wLasso. The intrinsic OP of dust is 0 in Lasso without weighting, demonstrating that the source of dust does not help to get a robust performance for Lasso, so it is shrinked to 0. The minimized term of Ridge and Lasso with weighting was added in the main text (line 233 and 246, respectively):

Ridge:

$$Minimize: \frac{1}{2m} \sum_{i=1}^{m} w_i (y_i - \widehat{y_i})^2 + \lambda * \sum_{j=1}^{p} \beta_j^2 \ (Eq9)$$

Least Absolute Shrinkage and Selection Operator (Lasso):

$$Minimize: \frac{1}{2m} \sum_{i=1}^{m} w_i (y_i - \widehat{y_i})^2 + \lambda * \sum_{j=1}^{p} |\beta_j| \ (Eq11)$$

Lasso regression encourages 0 coefficients on the factors. From my experience, many of the factors' (components, sources, etc.) coefficients became 0 when applying Lasso. While in this study, conducting Lasso does not exclude many sources from the OP contribution. Some minor sources are still in the Lasso results (such as primary biogenic and salt in Lasso in Figure S8). Can you explain the details of doing the Lasso regression?

Reply: Thanks for your question. We implemented hyperparameter tuning, with λ in (5, 1, 0.5, 0.1, 0.01, 0.005, 0.001, 0.0005, 0.0001). The selecting parameter was run using the 5-fold cross-validation. Finally, the best solution varied between 0.005 and 0.01 for 6 sites. We selected the highest value of λ i.e., 0.01. The minor sources are still in the Lasso result because of the low amount of shrinkage. However, in our case, the best λ is very low, so few sources were excluded from the model. That could be explained by the synergistic effects can cause a source with low intrinsic OP to be important, and the model supposed every PM source is important in OP prediction, consequently, the tuning selected a value for lambda that keeps most sources in the model. Indeed, for some sites (GRE-fr, PdB) the Lasso does shrink more than 4 sources to 0 (Figure S4, S5 Supplement).

Line 460: 'Intrinsic OP values obtained in this way from the best model encompassing all six sites are called intrinsic OP of the best model, and the intrinsic OP values derived from the OLS from all six sites are called intrinsic OP of the reference model.' The authors should clearly list the best models for each site.

Reply: Thanks for your suggestion. We added this information in the main text as follow (lines 476-479):

*Intrinsic OP values obtained in this way from the best model (the best model presented in Table 3 for OP_{AA} and Table 4 for OP_{DTT}) encompassing all six sites are called **intrinsic OP of the best model,** and the intrinsic OP values derived from the OLS from all six sites are called **intrinsic OP of the reference model.***

The advantage of the best model over OLS shall be better emphasized. Some irrelevant content shall be reduced.

Reply: Thanks for your suggestion. We updated in the main text (line 514 to line 539 for the comparison of OP_{AA} and line 544 to line 581 for the comparison of OP_{DTT}.

However, the interquartile ranges (IQR) of the intrinsic OP values are consistently narrower for the best models across all sources, accounting for less divergence in intrinsic OP values across sites. However, the interquartile ranges (IQR) of the intrinsic OP values are consistently narrower for the

best models across all sources, accounting for less divergence in intrinsic OP values across sites. Moreover, the median intrinsic OP values obtained from the best model closely approximated the mean values, indicating the absence of extreme intrinsic OP values. For instance, in the case of road traffic, the mean and median values were 0.24 and 0.23 nmol min^{-1} µg^{-1}, respectively. Conversely, the reference model exhibited a large difference between the mean and median values, implying lower consistency across sites and sampling iterations. The same result was observed in biomass burning source, in which the median and mean intrinsic OP in the best model had fewer discrepancies. Further, the biomass burning intrinsic OP in GRE-fr of the best model is more consistent with those in other sites (best: 0.30 nmol min^{-1} µg^{-1}, reference: 0.35 nmol min^{-1} µg^{-1}).

When considering sources with low intrinsic OP, the variability can be larger between the two methods. As an example, for the sulfate-rich sources, the median intrinsic OP values were positive (0.002 nmol min^{-1} µg^{-1}), while the mean intrinsic OP values were negative (-0.008 nmol nmol min^{-1} µg^{-1}). The mean intrinsic OP in the best model exhibited fewer negative values in individual sites than in the reference model (for aged salt, salt, primary biogenic, MSA rich, sulfate-rich and nitrate-rich). In addition, the best model showed the less disparate intrinsic OP among individual sites for instance, the aged salt sources in GRE-fr and the primary biogenic and salt sources in CHAM. Furthermore, the best model displayed an intrinsic OP meaningful in terms of geochemical, which showed in the source of salt, primary biogenic, sulfate-rich. For instance, in the reference model, the average intrinsic OP of the primary biogenic in NIC (-0.03 nmol min^{-1} µg^{-1}), the intrinsic OP of salt in GRE-ft (-0.07 nmol min^{-1} µg^{-1}) as well as the sulfate-rich source in CHAM (-0.05 nmol min^{-1} µg^{-1}) represented a 100% reduction compared to the mean intrinsic OP of all sites. Moreover, the negative intrinsic OP was observed in NIC (Primary biogenic), and some extreme values in GRE-fr (aged salt, salt), CHAM (salt, primary biogenic, MSA-rich) (where heteroscedasticity was presented) in the OLS model, underscores that the model assumptions on data characteristics proving false could impact the accuracy of OP prediction. Consequently, these results highlight the advantage of considering the data in model selection.

Line 544 to line 581 for OP$_{DDT}$ comparison as follow:

Similar to OP$_{AA}$, for OP$_{DTT}$ the IQR of the best model is narrower for most of the sources than the IQR of the reference model (OLS). Except for the road traffic, industrial, and MSA-rich, the IQR is slightly higher in the best model (Figure 9 and Table S.9). In the two models, the mean intrinsic OP is essentially unchanged, where the traffic is the most critical source (0.27±0.10), followed by HFO (0.18±0.01), biomass burning (0.12±0.03), dust (0.12±0.07), primary biogenic (best: 0.10±0.06, reference: 0.12±0.08) and MSA rich (best: 0.11±0.09, reference: 0.09±0.09). The minimum difference between the two models in the dominant sources again confirms the conclusion in the OP$_{AA}$ comparison, demonstrating **the similar pattern of the best and the reference model in the most crucial sources of OP.** *For both best and reference, OP$_{DTT}$ activities showed sensitivity to more sources than OP$_{AA}$, as discussed in many works (Borlaza et al., 2021; Calas et al., 2019; Dominutti et al., 2023; Fadel et al., 2023).*

While the best and reference models give the same mean intrinsic OP$_{DTT}$ of all sites, the mean OP$_{DTT}$ at each individual site can vary substantially between the two models. The best model exhibited the positive intrinsic OP for all sources, while the reference model displayed negative intrinsic OP in RBX (MSA-rich and sulfate-rich). Especially in the case of sulfate-rich in RBX, the negative intrinsic OP in the

reference model passed the threshold of negative value, which presented a 110% reduction compared to the mean intrinsic OP of all sites. This is also found in the OP_AA comparison, which confirmed that the best model generates a geochemical meaningful OP intrinsic. In addition, the best model exhibited consistent intrinsic OP across sites, especially for the source of dust, salt, primary biogenic, sulfate-rich in TAL (heteroscedasticity is presented in this site), where OP intrinsic in TAL in the best model is more similar to the other sites. For instance, the reference model presented that the intrinsic OP in TAL is 0.20 nmol min⁻¹ μg⁻¹, far from the mean of all sites (0.07 nmol min⁻¹ μg⁻¹). We observed the same for OP intrinsic of nitrate-rich source in CHAM (where the heteroscedasticity is detected), which displayed the less dissimilar of CHAM with the other site in the best model. This again validates the conclusion in OP_AA comparison, demonstrating that respecting model assumption is essential to obtain a robust OP SA result.

Figure 8 and Figure 9 are difficult to understand. Box plots shall be better (just a suggestion). I also have a concern related to the calculation of the median value. For example, in Nitrate rich of Figure 8. The 3rd and 4th values in OLS are near ~-0.01, so the median value shall be close to -0.01. However, the displayed median value is ~-0.005.

Reply: Thanks for your comment. We would like to keep this plot as it is since, with the box plot, we cannot clearly see the average intrinsic OP of every site. To clarify, the points in Figure 8 represent the mean intrinsic OP but not the median. Therefore, it is reasonable to get the 3rd and 4th values are near -0.01 and the median value of all sites is -0.006.

Line 290: There is no difference in calculating MAE and RMSE.

Reply: Thanks for your comment. We corrected the equations as follow (line 306 and 307):

$$MAE = \frac{\sum_{i=0}^{m}(y_i - \widehat{y_i})}{m} (Eq13)$$

$$RMSE = \sqrt{\frac{\sum_{i=0}^{m}(y_i - \widehat{y_i})^2}{m}} (Eq14)$$

Line 460: 'The OLS model is used as a representative of usual practices that do not consider the database characteristics.' References are required to support your statements.

Reply: We added the reference for this sentence (line 473-474)

The OLS model is used as a representative of usual practices that do not consider the database characteristics (Williams et al., 2013).

There are some citation errors, such as 'Liu, & Ng. (2023). Toxicity of Atmospheric Aerosols: Methodologies & Assays.', 'Paatero, P., & Tappert, U. (1994). POSITIVE MATRIX FACTORIZATION: A NON-NEGATIVE FACTOR MODEL WITH OPTIMAL UTILIZATION OF ERROR ESTIMATES OF DATA VALUES*. In ENVIRONMETRICS (Vol. 5).', 'Wang, Wang, M., Li, S., Sun, H., Mu, Z., Zhang, L., Li, Y., & Chen, Q. (2020). Study on the oxidation potential of the water-soluble components of ambient PM2.5 over Xi'an, China: Pollution levels, source apportionment

and transport pathways. Environment International, 136(January), 105515. https://doi.org/10.1016/j.envint.2020.105515

Reply: thanks for your remark, we corrected these references in the main text

Anonymous Referee #4

Thuy et al. have performed a study to provide a guideline for researchers interested in apportioning the contribution of different PM sources from conventional PMF-based source apportionment studies and linking it to its oxidative potential (OP). They achieved this by systematically comparing several commonly employed regression models and calculating their OP predictions. They compared the intrinsic PM OP using eight different MLR models and discussed the limitations and strengths of each approach. Finally, they provided a workflow for choosing the best OP model based on the PMF data available.

Overall, the manuscript is well-written, easy to follow, and focused. However, I would appreciate clarification on the methodology section, the environmental relevance of the results, and the result interpretation section of the manuscript.

Specific comments:

Lines 95-100: The authors mention, "OP analytical errors were used in weighing." Were these analytical errors calculated based on replicate measurements of the same sample? This information is important because, typically, in a lab setting, when performing PM OP analysis, there should not be significant variations in the standard deviation (SD) if replicates are analyzed. This is because the errors should ideally be more or less constrained within a lab using the same measurement protocol. The sample replicate SD gets even more constrained, especially when using automated OP systems for PM OP measurements (SD < 10%). Consequently, I don't think weighing based on the same sample's analytical replicates would significantly impact the regression (unless there was, in fact, high variability in the replicate measurements, which is an OP experiment protocol issue and should be fixed first) since all the PM OP from a single site was measured from the same lab. I expect a limited spread in weights. I would like to know the authors' thoughts on what uncertainty to be used when running these models. The choice of uncertainty data used will also alter the workflow provided in Figure 10 when choosing the best model for the dataset?

Reply: Thanks for the interesting suggestion. Yes, the analytical errors were calculated as the standard deviation of 4 replicate measurements on the same sample, as practiced in our lab. We agree with the referee that the SD is consistently below 10% for our analytical errors and that maybe we could not get a significant impact. However, when applying analytical errors as the weighting, we aim to mark out that the OP PM_{10} gets a high analytical uncertainty and should have a lower impact on the model. In our opinion, this is a reasonable point of view. Nevertheless, from a static point of view, the weight could be assessed in different ways (Montgomery C et al., 2012) (Page 191): (1) Prior information from the theoretical model, (2) using the residual extracted from OLS model, (3) The selecting of weighting based on the uncertainty of instrument if the dependent variable measured by a different method and (4) If the

dependent variable is the average of different observations, the weighting selected based on the error of these observations (our weightings were chosen in this case). Overall, the weighting choice depends on the study's purpose. We cannot recommend the selection of the weighting in Figure 10 since we did not perform the test on this. However, the limitations have been updated for future studies (line 665-670).

This study used the analytical uncertainty as the weighting for the weighted model. However, the weighting can be selected based on different ways, as reported by Montgomery et al. (2012)*: (1) Prior information from the theoretical model, (2) Using the residual extracted from the OLS model, (3) The selecting of weighting based on the uncertainty of instrument if the dependent variable measured by a different method and (4) If the dependent variable is the average of different observations, the weighting selected based on the error of these observations.*

Regarding the sample set used in this study, since the compiled data was not provided at a single location and was spread across multiple previously published studies from the research group, it was difficult to visualize the entire dataset. However, based on the information provided, I would assume that for all the data presented, the authors must have observed a good correlation (r) between OPv and PM10 mass concentration for all sites. My question is whether the authors observe any difference in model performance for datasets with low or poor correlation between PM mass and OPv. OPv is the combined effect of PM mass and intrinsic OP; I want to know if the results will hold in cases where intrinsic OP was more important than bulk mass of the PM, in driving the OPv?

Reply: Thanks for your question. The first step we were doing before running an MLR model was looking at the relationship between PM mass concentration and OP_V as well as OP_m, to investigate the global relationship of OP and PM. The table below shows the coefficient of determination (R^2) between the PM mass concentration and OP_V (PM vs OPv) and R^2 between the observed OP_V and predicted OP_V by the best model (Note: Model performance). The OP source apportionment model's performance was not clearly related to the correlation between OP and PM mass concentration. For example, PM mass was more correlated with OP_{DTT} than with OP_{AA}, but the model performed better with OP_{AA} than OP_{DTT}. PM mass was least correlated with OP_{AA} at NIC and PdB (correlation < 0.4) but the model performed similarly at these sites and at TAL, where the correlation between PM mass and OP_{AA} was much higher.

	GRE-fr	CHAM	NIC	PdB	RBX	TAL	Note
OP^{AA}_v	0.68	0.55	0.35	0.36	0.62	0.69	PM vs OP_v
OP^{AA}_v	0.94	0.89	0.78	0.81	0.67	0.82	Model performance
OP^{DTT}_v	0.78	0.86	0.75	0.58	0.83	0.68	PM vs OP_v
OP^{DTT}_v	0.83	0.83	0.68	0.75	0.65	0.70	Model performance

I am also interested in the relative contribution of the different identified sources to overall OPv since OPv is a more health-relevant endpoint than OPm. This will also inform us if these models can identify and quantify PM10 sources that contribute differently to the PM10 mass vs. OPv. The whole objective of this exercise is to quantify the health-related impact of PM10. Are there any differences in the source-specific relative contribution based on OP source apportionment vs. PM mass source apportionment? If not for all models; you could show this comparison for the best regression-modeled OPv at each site and compare it to PM mass contribution.

Reply: Thanks for your suggestion. Yes, we indeed obtained the difference between the contribution of sources to PM mass and OPv. This was described in several of our previous papers (Borlaza et al., 2021b; Dominutti et al., 2023; Weber et al., 2021). For instance, in RBX, biomass burning contribution in PM10 is only 2 µg m^{-3} (10% to total mass), while this source contributes highest to OP$_{AA}$ (0.6 nmol min^{-1} m^{-3}, 25% to OP$_{AA}$). Alternatively, the secondary inorganic aerosol is the main contributor to PM$_{10}$ mass but has low redox activity for all sites. Additionally, the OP$_m$ of PM sources is similar across sites, but each source's contribution to OP$_v$ varies between sites because of differences between sites in source contributions to PM mass concentration. We updated the contribution of sources to PM and OP$_V$ for each site in the Supplement (Figure S.14 to Figure S.19).

Finally, is PMF followed by MLR the best approach for PM OP source apportionment? As you have also described in your introduction, OP is a complex reaction term, and specific components in PM could be driving these sources. In the conventional PMF approach, the emphasis is on mass-based apportionment, and if the contribution of a source to the mass is below a certain threshold, the source is often eliminated. One contention for such an approach is that we would be missing out on identifying sources that may have a low contribution to overall PM10 mass but are significant contributors to OP. How would the authors suggest approaching this complexity, especially considering a major application of the research to use with "European Directive 605 2008/50/CE"? I would expect one goal of the revision to include OP to be to give more insights into identifying sources with high intrinsic OP and less contribution to mass and vice versa.

Reply: Thanks for your remarks and questions. Tackling the complexity of PM sources determination as well as the OP of PM processes is indeed the aim of the authors' research, which focuses on improving the methodology of both PM source apportionment and OP deconvolution. Our group has been published various publications for 10 years which aimed to reach the limit of the PMF in determine the PM sources by incorporating the more tracer of sources (Borlaza et al., 2021a; Samaké et al., 2019; Waked et al., 2014; Weber et al., 2019) as well as developed the OP source apportionment techniques, by introduce the linear models (Weber et al., 2018, 2021) and non-linear models (Borlaza et al., 2021b; Dominutti et al., 2023).

We used to measure about 160 chemical species at least 1 year for every site study. The sources tracer is sufficient to identify from 10 to 12 PM sources, which get the slope almost to 1 and R^2 > 0.9 for all of these sites. In addition, our PMF result can identify sources that contribute from 1 to 4% in mass (secondary biogenic in GRE-fr, CHAM, BBX), and which are already very minor sources. Conversely, we are conscious that sometimes, the PMF results do not adequately represent PM mass concentration for several reasons, such as the lack of a trace species to identify a source, an insufficient sample size or the source contribution being too small to be identified or collinearity matters. For example, the both sources of traffic including exhausted and non-exhausted are rarely separated because of the strong co-variation. For these cases, we could recommend subtracting the total source contribution from PM mass concentration to get a part that PMF cannot simulate. The information in this part maybe contains the vital source.

This study demonstrated that the source has a less significant mass contribution but high redox activity, such as primary traffic in CHAM (5% in PM mass contribution), but gets the second rank in intrinsic OP$_{DTT}$ (0.17 nmol min^{-1} µg^{-1}). Conversely, sulfate-rich sources contribute 27% in PM mass

concentration, but its intrinsic OP_{AA} is 0.01 nmol min^{-1} µg^{-1}. The contribution of sources to PM_{10} and OP_V for each site are shown in the Supplement (Figure S.14 to figure S.19). Finally, we acknowledge that the assumption of assigning the relative contribution of OP from multi-linear regression may be strong if we consider the potential antagonistic and synergistic effects in OP assays. But if we were to miss important sources in terms of OP that had a minimal contribution in terms of mass and that could have been omitted, we would obtain a weak signal for OP reconstruction, yet we obtain a very good comparison between the modeled OP from MLR and the observations. For all these reasons, we are confident in the validity of the PMF + MLR approach for assigning the apportionment of oxidive potential to each source of PM emissions.

We added in the recommendation as follow (line 626-637):

Finally, these techniques of OP apportionment could not be well performed with uncertain PMF-derived sources. The PMF results sometimes do not adequately represent PM mass concentration for several reasons, such as the lack of a trace species to identify a source, an insufficient sample size, the source contribution being too small to be identified (under 1%), or collinearity matters. The important information could be missed because of these problems in PMF implementation, which is apprehended by the model's low accuracy. Our study did not encounter this problem since the PMF is harmonized and performed according to European recommendations which could well perform the regression technique and allow to obtain a very satisfactory successive OP modelled in comparison to observations after regression techniques (R^2 from 0.7 to 0.9). However, this problem could potentially happen, and for these cases, we could recommend either subtracting the total source contribution from PM mass concentration to get a part that PMF cannot simulate. The information in this part may contain vital sources. Alternatively, it is possible to re-execute the PMF to validate the result and ensure the robustness of the chemical profile and the contribution of sources.

Minor comments

Since this is a numerical model intercomparison study, sharing the code or uploading it to a public repository is important. While I understand these are standard models, the codes are still useful for the reader and reviewers to understand the specific constraints used, how the uncertainty was handled, etc.

Reply: Thanks for your suggestion, we also would like to share the model for the other researcher. The code will be shared in a repository. DOI: https://doi.org/10.5281/zenodo.11070914.

Throughout the main text, instead of using the term "OP activity" or "OP", it is more appropriate to write "PM10 OP". OP is general terminology used in different fields of science; here, you are working with PM10 OP specifically.

Reply. Thanks for your suggestion. We modified the main text (line 156-157).

To simplify the denotation of PM_{10} OP, OP is used for represented to the PM_{10} OP throughout this article.

In Tables S6 to S8, include the OP units. Also, mention in the SI tables that the data reported are for intrinsic PM10 OP. Intrinsic OP is PM size-specific, so it is important to mention the size of the PM investigated.

Reply. Thanks for your suggestion. We added the unit in these tables in the Supplement. We updated OP to PM_{10} OP in Supplement.

Bates, J. T., Fang, T., Verma, V., Zeng, L., Weber, R. J., Tolbert, P. E., Abrams, J. Y., Sarnat, S. E., Klein, M., Mulholland, J. A., & Russell, A. G. (2019). Review of Acellular Assays of Ambient Particulate Matter Oxidative Potential: Methods and Relationships with Composition, Sources, and Health Effects. *Environmental Science and Technology*, *53*(8), 4003–4019. https://doi.org/10.1021/acs.est.8b03430

Bates, J. T., Weber, R. J., Abrams, J., Verma, V., Fang, T., Klein, M., Strickland, M. J., Sarnat, S. E., Chang, H. H., Mulholland, J. A., Tolbert, P. E., & Russell, A. G. (2015). Reactive Oxygen Species Generation Linked to Sources of Atmospheric Particulate Matter and Cardiorespiratory Effects. *Environmental Science and Technology*, *49*(22), 13605–13612. https://doi.org/10.1021/acs.est.5b02967

Borlaza, L. J. S., Uzu, G., Ouidir, M., Lyon-Caen, S., Marsal, A., Weber, S., Siroux, V., Lepeule, J., Boudier, A., Jaffrezo, J.-L., Slama, R., Lyon-Caen, S., Siroux, V., Lepeule, J., Philippat, C., Slama, R., Hofmann, P., Hullo, E., Llerena, C., … group, the S. cohort study. (2023). Personal exposure to PM2.5 oxidative potential and its association to birth outcomes. *Journal of Exposure Science & Environmental Epidemiology*, *33*(3), 416–426. https://doi.org/10.1038/s41370-022-00487-w

Borlaza, L., Weber, S., Jaffrezo, J. L., Houdier, S., Slama, R., Rieux, C., Albinet, A., Micallef, S., Trébluchon, C., & Uzu, G. (2021). Disparities in particulate matter (PM10) origins and oxidative potential at a city scale (Grenoble, France) - Part 2: Sources of PM10 oxidative potential using multiple linear regression analysis and the predictive applicability of multilayer perceptron n. *Atmospheric Chemistry and Physics*, *21*(12), 9719–9739. https://doi.org/10.5194/acp-21-9719-2021

Borlaza, L., Weber, S., Uzu, G., Jacob, V., Cañete, T., Micallef, S., Trébuchon, C., Slama, R., Favez, O., & Jaffrezo, J.-L. (2021). Disparities in particulate matter (PM10) origins and oxidative potential at a city scale (Grenoble, France) - Part 1: Source apportionment at three neighbouring sites. *Atmospheric Chemistry and Physics*, *21*(7), 5415–5437. https://doi.org/10.5194/acp-21-5415-2021

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Whitsel, L., & Kaufman, J. D. (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the american heart association. *Circulation*, *121*(21), 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1

Calas, A., Uzu, G., Besombes, J. L., Martins, J. M. F., Redaelli, M., Weber, S., Charron, A., Albinet, A., Chevrier, F., Brulfert, G., Mesbah, B., Favez, O., & Jaffrezo, J. L. (2019). Seasonal variations and chemical predictors of oxidative potential (OP) of particulate matter (PM), for seven urban French sites. *Atmosphere*, *10*(11). https://doi.org/10.3390/atmos10110698

Calas, A., Uzu, G., Kelly, F. J., Houdier, S., Martins, J. M. F., Thomas, F., Molton, F., Charron, A., Dunster, C., Oliete, A., Jacob, V., Besombes, J. L., Chevrier, F., & Jaffrezo, J. L. (2018). Comparison between five acellular oxidative potential measurement assays performed with detailed chemistry on PM10 samples from the city of Chamonix (France). *Atmospheric Chemistry and Physics*, *18*(11), 7863–7875. https://doi.org/10.5194/acp-18-7863-2018

Camman, J., Chazeau, B., Marchand, N., Durand, A., Gille, G., Lanzi, L., Jaffrezo, J., Wortham, H., & Uzu, G. (2023). *Oxidative potential apportionment of atmospheric PM 1: A new approach combining high-sensitive online analysers for chemical composition and offline OP measurement technique*. *July*, 1–34.

Daellenbach, K. R., Uzu, G., Jiang, J., Cassagnes, L.-E., Leni, Z., Vlachou, A., Stefenelli, G., Canonaco, F., Weber, S., Segers, A., & Sources, al. (2020). Sources of particulate-matter air pollution and its oxidative potential in Europe of particulate-matter air pollution and its oxidative potential in Europe. *Nature*, *587*(7834). https://doi.org/10.1038/s41586-020-2902-8ï

Deng, M., Chen, D., Zhang, G., & Cheng, H. (2022). Policy-driven variations in oxidation potential and source apportionment of PM2.5 in Wuhan, central China. *Science of the Total Environment*,

853(May), 158255. https://doi.org/10.1016/j.scitotenv.2022.158255

Dominutti, P. A., Borlaza, L., Sauvain, J. J., Ngoc Thuy, V. D., Houdier, S., Suarez, G., Jaffrezo, J. L., Tobin, S., Trébuchon, C., Socquet, S., Moussu, E., Mary, G., & Uzu, G. (2023). Source apportionment of oxidative potential depends on the choice of the assay: insights into 5 protocols comparison and implications for mitigation measures. *Environmental Science: Atmospheres*. https://doi.org/10.1039/d3ea00007a

Fadel, M., Courcot, D., Delmaire, G., Roussel, G., Afif, C., & Ledoux, F. (2023). Source apportionment of PM2.5 oxidative potential in an East Mediterranean site. *Science of the Total Environment*, *900*(July). https://doi.org/10.1016/j.scitotenv.2023.165843

Leni, Z., Cassagnes, L. E., Daellenbach, K. R., Haddad, I. El, Vlachou, A., Uzu, G., Prévôt, A. S. H., Jaffrezo, J. L., Baumlin, N., Salathe, M., Baltensperger, U., Dommen, J., & Geiser, M. (2020). Oxidative stress-induced inflammation in susceptible airways by anthropogenic aerosol. *PLoS ONE*, *15*(11 November). https://doi.org/10.1371/journal.pone.0233425

Li, J., Zhao, S., Xiao, S., Li, X., Wu, S., Zhang, J., & Schwab, J. J. (2023). Source apportionment of water-soluble oxidative potential of PM 2 . 5 in a port city of Xiamen , Southeast China. *Atmospheric Environment*, *314*(June), 120122. https://doi.org/10.1016/j.atmosenv.2023.120122

Liu, W. J., Xu, Y. S., Liu, W. X., Liu, Q. Y., Yu, S. Y., Liu, Y., Wang, X., & Tao, S. (2018). Oxidative potential of ambient PM2.5 in the coastal cities of the Bohai Sea, northern China: Seasonal variation and source apportionment. *Environmental Pollution*, *236*, 514–528. https://doi.org/10.1016/j.envpol.2018.01.116

Marsal, A., Slama, R., Lyon-Caen, S., Borlaza, L. J. S., Jaffrezo, J. L., Boudier, A., Darfeuil, S., Elazzouzi, R., Gioria, Y., Lepeule, J., Chartier, R., Pin, I., Quentin, J., Bayat, S., Uzu, G., Siroux, V., Eyriey, E., Licinia, A., Vellement, A., … Slama, R. (2023). Prenatal exposure to pm2:5 oxidative potential and lung function in infants and preschool-age children: A prospective study. *Environmental Health Perspectives*, *131*(1). https://doi.org/10.1289/EHP11155

Montgomery C, D., Peck A, E., & Vining, G. G. (2012). *Introducing To Linear Regression Analysis (5th ed.)*.

Pietrogrande, M. C., Romanato, L., & Russo, M. (2022). Synergistic and Antagonistic Effects of Aerosol Components on Its Oxidative Potential as Predictor of Particle Toxicity. *Toxics*, *10*(4). https://doi.org/10.3390/toxics10040196

Samaké, A., Jaffrezo, J. L., Favez, O., Weber, S., Jacob, V., Canete, T., Albinet, A., Charron, A., Riffault, V., Perdrix, E., Waked, A., Golly, B., Salameh, D., Chevrier, F., Miguel Oliveira, D., Besombes, J. L., Martins, J. M. F., Bonnaire, N., Conil, S., … Uzu, G. (2019). Arabitol, mannitol, and glucose as tracers of primary biogenic organic aerosol: The influence of environmental factors on ambient air concentrations and spatial distribution over France. *Atmospheric Chemistry and Physics*, *19*(16), 11013–11030. https://doi.org/10.5194/acp-19-11013-2019

Samake, A., Uzu, G., Martins, J. M. F., Calas, A., Vince, E., Parat, S., & Jaffrezo, J. L. (2017). The unexpected role of bioaerosols in the Oxidative Potential of PM. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-11178-0

Shangguan, Y., Zhuang, X., Querol, X., Li, B., Moreno, N., Trechera, P., Sola, P. C., Uzu, G., & Li, J. (2022). Characterization of deposited dust and its respirable fractions in underground coal mines: Implications for oxidative potential-driving species and source apportionment. *International Journal of Coal Geology*, *258*(December 2021). https://doi.org/10.1016/j.coal.2022.104017

Tao, F., Gonzalez-Flecha, B., & Kobzik, L. (2003). Reactive oxygen species in pulmonary inflammation by ambient particulates. *Free Radical Biology and Medicine*, *35*(4), 327–340. https://doi.org/10.1016/S0891-5849(03)00280-6

Verma, V., Fang, T., Guo, H., King, L., Bates, J. T., Peltier, R. E., Edgerton, E., Russell, A. G., & Weber, R. J. (2014). Reactive oxygen species associated with water-soluble PM2.5 in the southeastern United States: Spatiotemporal trends and source apportionment. *Atmospheric Chemistry and*

Physics, *14*(23), 12915–12930. https://doi.org/10.5194/acp-14-12915-2014

Waked, A., Favez, O., Alleman, L. Y., Piot, C., Petit, J. E., Delaunay, T., Verlinden, E., Golly, B., Besombes, J. L., Jaffrezo, J. L., & Leoz-Garziandia, E. (2014). Source apportionment of PM10 in a north-western Europe regional urban background site (Lens, France) using positive matrix factorization and including primary biogenic emissions. *Atmospheric Chemistry and Physics*, *14*(7), 3325–3346. https://doi.org/10.5194/acp-14-3325-2014

Wang, D., Shen, Z., Zhang, Q., Lei, Y., Zhang, T., Huang, S., Sun, J., Xu, H., & Cao, J. (2022). Winter brown carbon over six of China's megacities: Light absorption, molecular characterization, and improved source apportionment revealed by multilayer perceptron neural network. *Atmospheric Chemistry and Physics*, *22*(22), 14893–14904. https://doi.org/10.5194/acp-22-14893-2022

Wang, S., Ye, J., Soong, R., Wu, B., Yu, L., Simpson, A. J., & Chan, A. W. H. (2018). Relationship between chemical composition and oxidative potential of secondary organic aerosol from polycyclic aromatic hydrocarbons. *Atmospheric Chemistry and Physics*, *18*(6), 3987–4003. https://doi.org/10.5194/acp-18-3987-2018

Wang, Y., Wang, M., Li, S., Sun, H., Mu, Z., Zhang, L., Li, Y., & Chen, Q. (2020). Study on the oxidation potential of the water-soluble components of ambient PM2.5 over Xi'an, China: Pollution levels, source apportionment and transport pathways. *Environment International*, *136*(January), 105515. https://doi.org/10.1016/j.envint.2020.105515

Weber, S., Salameh, D., Albinet, A., Alleman, L. Y., Waked, A., Besombes, J. L., Jacob, V., Guillaud, G., Meshbah, B., Rocq, B., Hulin, A., Dominik-Sègue, M., Chrétien, E., Jaffrezo, J. L., & Favez, O. (2019). Comparison of PM10 sources profiles at 15 french sites using a harmonized constrained positive matrix factorization approach. *Atmosphere*, *10*(6). https://doi.org/10.3390/atmos10060310

Weber, S., Uzu, G., Calas, A., Chevrier, F., Besombes, J. L., Charron, A., Salameh, D., Ježek, I., Močnik, G., & Jaffrezo, J. L. (2018). An apportionment method for the oxidative potential of atmospheric particulate matter sources: Application to a one-year study in Chamonix, France. *Atmospheric Chemistry and Physics*, *18*(13), 9617–9629. https://doi.org/10.5194/acp-18-9617-2018

Weber, S., Uzu, G., Favez, O., Borlaza, L., Calas, A., Salameh, D., Chevrier, F., Allard, J., Besombes, J. L., Albinet, A., Pontet, S., Mesbah, B., Gille, G., Zhang, S., Pallares, C., Leoz-Garziandia, E., & Jaffrezo, J. L. (2021). Source apportionment of atmospheric PM10 oxidative potential: Synthesis of 15 year-round urban datasets in France. *Atmospheric Chemistry and Physics*, *21*(14), 11353–11378. https://doi.org/10.5194/acp-21-11353-2021

Williams, M., Gomez Grajales, C. A., & Kurkiewicz, D. (2013). Assumptions of Multiple Regression: Correcting Two Misconceptions - Practical Assessment, Research & Evaluation. *Practical Assessment, Research, and Evaluation (PARE)*, *18*(11), 1–16. https://scholarworks.umass.edu/pare/vol18/iss1/11

Xiong, Q., Yu, H., Wang, R., Wei, J., & Verma, V. (2017). Rethinking Dithiothreitol-Based Particulate Matter Oxidative Potential: Measuring Dithiothreitol Consumption versus Reactive Oxygen Species Generation. *Environmental Science and Technology*, *51*(11), 6507–6514. https://doi.org/10.1021/acs.est.7b01272

Yu, S. Y., Liu, W. J., Xu, Y. S., Yi, K., Zhou, M., Tao, S., & Liu, W. X. (2019). Characteristics and oxidative potential of atmospheric PM2.5 in Beijing: Source apportionment and seasonal variation. *Science of the Total Environment*, *650*, 277–287. https://doi.org/10.1016/j.scitotenv.2018.09.021

Zhang, L., Hu, X., Chen, S., Chen, Y., & Lian, H. Z. (2023). Characterization and source apportionment of oxidative potential of ambient PM2.5 in Nanjing, a megacity of Eastern China. *Environmental Pollutants and Bioavailability*, *35*(1). https://doi.org/10.1080/26395940.2023.2175728