

Dear Dr. Huffman,

We would like to thank you and the referees for their helpful comments on our manuscript entitled: “Spatial and temporal variability of sources of ambient fine particulate matter (PM_{2.5}) in California” (MS No.: acp-2014-587). Attached is a document in which we provide detailed responses to the comments.

We also addressed your additional comments and corrections in the revised versions of the manuscript (highlighted in red) and supplemental text.

We hope that the revised manuscript is now in an acceptable form for publication in *Atmospheric Chemistry and Physics*.

Sincerely,

Sina Hasheminassab

PhD candidate, Environmental Engineering

University of Southern California

3710 McClintock Ave, RTH 414

Los Angeles, CA 90007

Tel: 213-821-5960

USC aerosol lab: www.usc.edu/aerosol

Personal website: www.sinahasheminassab.wordpress.com

Anonymous referee #1

This manuscript described the spatial and temporal variability of sources of ambient PM_{2.5} in California. It is a nice summary about PM source apportionment using a comprehensive dataset collected at multiple ambient monitoring stations across the state. The results will provide useful information for future epidemiological studies and help us improve the current understanding of the relationship between PM sources and health effects. The publication in Atmospheric Chemistry and Physics is recommended after the following comments are addressed.

We would like to thank the reviewer for his/her helpful feedback. The reviewer's comments are properly noted and his/her major concern about further clarification on the PMF analyses has been carefully addressed. Please find below detailed responses to each comment.

Major Comments:

Comment 1: *On page 20048, the authors state that previous studies are impacted by the limited number of sampling locations and short measurement periods. The readers may not know the background information about those studies. Could you please select 2 or 3 out of 10 papers cited here and briefly describe their limitations? How do they compare to your data (8 sites and 6 years)?*

Response: The reviewer's comment is noted. The following part is added to the revised manuscript to provide some succinct examples of the previous source apportionment studies conducted in California and describe their limitations in terms of spatial variability and/or studied time period.

Lines 76-84: "For instance, Kim et al. (2010) analyzed the PM_{2.5} speciation data collected between 2003 and 2005 at two sampling sites in southern California (i.e. Los Angeles (LA) and Rubidoux) to identify and quantify major PM_{2.5} sources, by application of a PMF model. Using similar source apportionment approach, Hwang and Hopke (2006) evaluated the sources of ambient PM_{2.5} at two sampling sites in San Jose, using the STN data collected between 2000 and 2005. In a more comprehensive study, Chen et al. (2007) applied several receptor models to the chemically speciated PM_{2.5} measurements collected for one year (between 2000 and 2001) at 23 sites, all located in California's San Joaquin Valley (SJV), to estimate PM_{2.5} source contributions."

Comment 2: Did you run the PMF model separately for 8 sites? Or you used all daily-averaged data in one input file for the PMF model? Please clarify.

Response: The PMF model was performed for each sampling site separately. To clarify this in the revised manuscript, the following sentence under the “source apportionment” section has been updated.

Lines 167-169: “In this study, the EPA PMF receptor model (version 3.0.2.2) was performed at each sampling site separately to identify the major sources of ambient PM_{2.5} and quantify their relative contributions to total PM_{2.5} mass.”

Comment 3: On page 20050, the authors state that PMF analyses were conducted by a larger dataset between 2002 and 2013 for LA and Rubidoux, but average source contributions between 2002 and 2007 were calculated. Did you use any data between 2008 and 2013 for LA and Rubidoux in the input file when running the PMF model? If so, does that affect the PMF results as the measurement period is not consistent among different sites? Also, the PM sampling interval varies across 8 sites (every third day versus every sixth day). How did you prepare the concentration and uncertainty profiles in terms of data structure alignment? Did you use any values to replace those missing data points caused by different measurement period and sampling interval? Blank cells are not accepted by EPA PMF model 3.0.

Response:

1. The reviewer’s comment is well taken. In the companion paper of this study (Hasheminassab et al., 2014), for LA and Rubidoux we used a combined chemical dataset from 2002 to 2013 as the input file when running the PMF model. For this paper, in order to be able to compare the results with those obtained for the rest of sampling sites, we calculated the average source contributions between 2002 and 2007, using the output of the same PMF runs, which were originally conducted using the 2002-2013 chemical dataset. Hasheminassab et al. (2014) performed a rigorous sensitivity analysis to evaluate whether the source contributions between 2002 and 2007 are biased when the PMF model is performed on the combined chemical dataset from 2002 to 2013. To that end, at each sampling site, the PMF model was first conducted using the entire chemical dataset (i.e. 2002-2013), and then the results were compared to the outputs of the PMF model

performed separately on the pooled chemical datasets before and after 2007 (i.e. 2002-2006 and 2008-2012, respectively). A similar number of sources with almost identical profiles were obtained from both approaches. In addition, the PMF outcome indicated comparable source contributions from the two approaches (less than 18% difference in average source contributions among all sources). The similarity between the source contributions obtained from the two approaches was also supported by Mann-Whitney rank sum tests, which showed that the differences between the daily-resolved source contributions are not statistically significant ($p>0.05$).

The below statements are added to the revised manuscript to address the reviewer's concern:

Lines 137-150: “Time-integrated 24 h PM_{2.5} samples were collected between 2002 and 2007 at all sampling sites, except for LA and Rubidoux, at which the STN data collected from 2002 to 2013 was used as the input file when running the PMF model (Hasheminassab et al., 2014). In the present study, in order to compare the results with those obtained for the rest of sampling sites, we calculated the average source contributions between 2002 and 2007 from the output of the same PMF runs which were originally conducted using the 2002-2013 chemical dataset. By performing a sensitivity analysis, Hasheminassab et al. (2014) showed that the results of the PMF model performed on the entire chemical dataset (i.e. 2002-2013) is comparable to the output of the PMF model conducted separately on 2002-2006 and 2008-2012 datasets, in terms of the sources identified (similar number of sources with almost identical compositions) and the absolute source contributions (less than 18% difference in average source contributions among all sources). The outcome of the sensitivity analysis thus indicated that the daily-resolved source contributions between 2002 and 2007 are not significantly biased when the chemical data between 2008 and 2013 are also included into the PMF input file.”

2. As mentioned in response to the previous comment, PMF model was performed independently for each sampling site. Therefore, we had a separate input matrix of concentration and uncertainty for each site. Obviously, the dimension of the input matrices varies among all 8 sites, due to the differences in the sampling intervals and the

number of species included in the model, following the screening procedure. We did not intend to align the structure and/or dimension of the input matrices among all 8 sites. PMF model was performed separately for each sampling site and eventually the seasonal variability in the contributions of the identified sources was evaluated among all sites. In each individual input dataset, however, missing values of single species were treated following the method of Polissar et al. (1998), as mentioned in the manuscript.

Comment 4: On page 20051, could you please describe where PM filter weighing and chemical analysis were conducted? By EPA staff? On-site? Or samples were transported to other laboratories?

Response: Filter handling and custody procedures as well as protocols for the chemical analyses are thoroughly described in the U.S. EPA Quality Assurance Project Plan (QAPP) (EPA-454/R-01-001) for the field operations involved in the PM_{2.5} Speciation Trends Network (STN). Briefly, the state and local air monitoring agencies are responsible for the operation of the STN sites. They are also responsible for training the personnel performing sampling, QA/QC, data handling, and other duties related to the PM_{2.5} chemical speciation network. For STN sampling, filters are tested, equilibrated, and weighted in the U.S. EPA contract laboratories, and then they are shipped to the field. After sampling, filters bearing PM_{2.5} deposits are promptly shipped back to the laboratories for weight determination and other chemical analyses.

The following statements are added to the revised manuscript to further clarify this:

Lines 154-159: “Filter weighing and chemical analyses were performed according to the U.S. EPA Quality Assurance Project Plan (QAPP) (EPA-454/R-01-001) adopted for the STN field sampling. According to the QAPP, filters were tested, equilibrated, and weighted in the U.S. EPA contract laboratories, and then they were shipped to the field. After sampling, filters bearing PM_{2.5} deposits were promptly shipped back to the laboratories for weight determination and other chemical analyses.”

Minor Comments:

In addition to the comments mentioned above, the following changes in the manuscript are suggested.

Comment 1: Page 20046 Line 2: $dp \rightarrow$ subscript p

Response: The correction was made in the revised manuscript.

Comment 2: Page 20048 Line 29: replace CAL EPA with Cal EPA (to be consistent with the abbreviation used in acknowledgements)

Response: “CAL EPA” was replaced with “Cal EPA” in the revised manuscript.

Comment 3: Page 20050 Line 7: remove (SoCAB) as it is not used in the following sections

Response: In accordance with the reviewer’s comment, “(SoCAB)” was removed from the revised manuscript.

Comment 4: Page 20050 Line 8: LA should be defined when Los Angeles appears for the first time.

Response: The reviewer’s comment is noted. In the revised manuscript we defined LA when Los Angeles appeared for the first time.

Comment 5: Page 20051 Line 13: replace positive matrix factorization (PMF) with PMF (it has been defined previously in the introduction section).

Response: In the revised manuscript we replaced “positive matrix factorization (PMF)” with “PMF”.

Comment 6: Page 20051 Line 26 and 20052 Line 2: BDL is defined twice and differently (w/ and w/o the).

Response: The reviewer’s comment is noted. In the revised manuscript we first defined “below detection limit (BDL)” (without the), and then used the acronym (BDL).

Comment 7: Page 20053 Line 12, 16, 18: replace relative humidity with RH as it has been defined previously.

Response: “relative humidity” was replaced with RH in the revised manuscript.

Comment 8: Page 20058 Line 3: remove (CMB) as it is not used in the following sections.

Response: “(CMB)” was removed from the revised manuscript.

Comment 9: Page 20058 Line 18: replace relative humidity with RH.

Response: “relative humidity” was replaced with RH in the revised manuscript.

Comment 10: Page 20060 Line 15, 19: replace relative humidity with RH.

Response: “relative humidity” was replaced with RH in the revised manuscript.

Anonymous referee #2

The paper provides receptor-based PM_{2.5} source apportionment at 8 California sites that covers 6 year period. Although receptor-oriented techniques have been widely used for PM source apportionment, previous studies were limited particularly by a smaller data set. In this sense, the paper provides some new information that is relevant to the readers of this journal. I recommend the paper be published after the authors address the following comments.

We would like to thank the reviewer for his/her helpful and constructive comments on the manuscript. Below are detailed replies to each of the reviewer's comments. The manuscript has been also revised accordingly.

***Comment 1:** Uncertainties. Somewhere in Section 2.3.1 or 2.3.2, the authors should more clearly mention how uncertainties for different elements were estimated. In addition to analytical uncertainty, what other factors were included in calculating the uncertainties.*

Response: To address the reviewer's comment, the following statements are added to the revised manuscript:

Lines 192-201: "The uncertainties used in the PMF model were the estimated uncertainties reported in the Air Quality System (AQS) for the PM_{2.5} chemical speciation network. The uncertainties reported by STN include both the analytical uncertainties and uncertainties associated with the field sampling component (Flanagan et al., 2006). The uncertainties of elements, measured by the ED-XRF method, go through a comprehensive calculation procedure that harmonizes the uncertainties between different instruments and accounts for filter matrix effect, in addition to the field sampling and handling uncertainty (Gutknecht et al., 2010). For the other species, uncertainty is estimated as the analytical uncertainty of the instrument, augmented by 5% of the calculated concentration, assuming that this 5% is representing the total "field" variability (Flanagan et al., 2006)."

***Comment 2:** OC artifact correction. Table S2 shows corrections various quite significantly among different sites, supposedly caused by different levels of OC and the detailed chemical composition and volatility. More importantly, at a number of sites, the OC concentration might*

have decreased significantly during the six year period. Should the correction factors also be time dependent?

Response: We acknowledge the fact that OC artifact may have time variability. In our study, however, the variations of OC artifact among different years were negligible and for this reason for each sampling site we used an average OC artifact value for the entire study period (the approach similar to most of the past long-term PMF investigations). To further clarify this point and address the reviewer's concern, we picked the sampling site with the maximum estimated OC artifact (i.e. Los Angeles) and evaluated the year-to-year variability of the artifact values. At this sampling site, the annual average concentration of uncorrected OC (i.e. comprising the artifact) decreased from near $7\mu\text{g}/\text{m}^3$ in 2002 to about $5\mu\text{g}/\text{m}^3$ in 2007 (Figure (a)), in agreement with the reviewer's surmise. Figure (b) shows the scatter plot of OC TOT versus measured $\text{PM}_{2.5}$ mass concentration, in which the data points corresponding to each year are specified with a distinct marker/color. As described in the paper, OC artifact is estimated using the intercept of the linear regression of OC against $\text{PM}_{2.5}$ mass concentration, following the method of Kim et al. (2005). Table (a) presents the estimated OC artifacts (\pm standard errors), segregated by year, in Los Angeles. As can be seen, the OC artifact values do not show a significant year-to-year variability. This can be inferred from the levels of significance corresponding to the two-tailed t tests performed on the OC artifact values between each two consecutive years, as shown in Table (a) (p values ranging from 0.23 to 0.69), indicating that there is no statistically significant difference between OC artifacts among different years. Furthermore, the maximum difference between the individual estimated OC artifact and the average value calculated over all 6 years (i.e. $\sim 3.7\mu\text{g}/\text{m}^3$) is less than 17% of the average value. Considering these observations, for each sampling site we estimated one OC artifact, using the intercept of OC TOT vs $\text{PM}_{2.5}$, over the entire study period (i.e. 2002-2007), similarly to the approach taken in many other studies in the literature, using the $\text{PM}_{2.5}$ chemical speciation data over a long period of time (Hwang and Hopke, 2006; Kim and Hopke, 2008a, b). To address the reviewer's comment, the discussion provided above as well as the following figures and tables are included in the supplemental text.

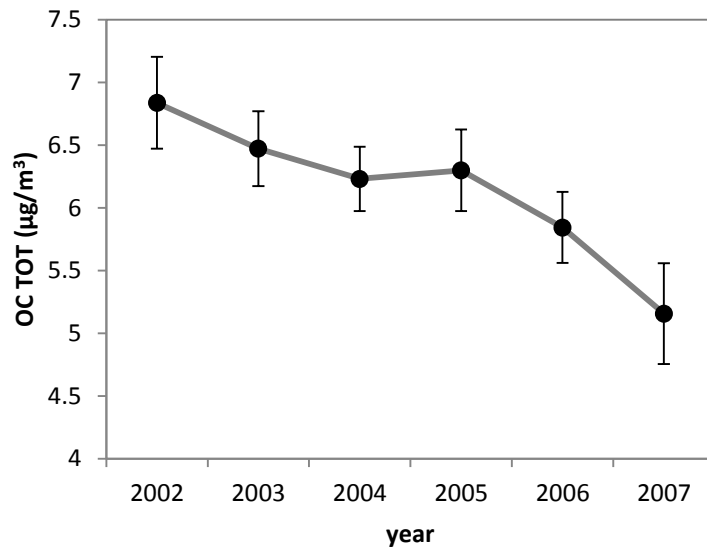


Figure (a). Annual average concentration ($\mu\text{g}/\text{m}^3$) of uncorrected organic carbon (OC) from 2002 to 2007 in Los Angeles. Error bars correspond to one standard error.

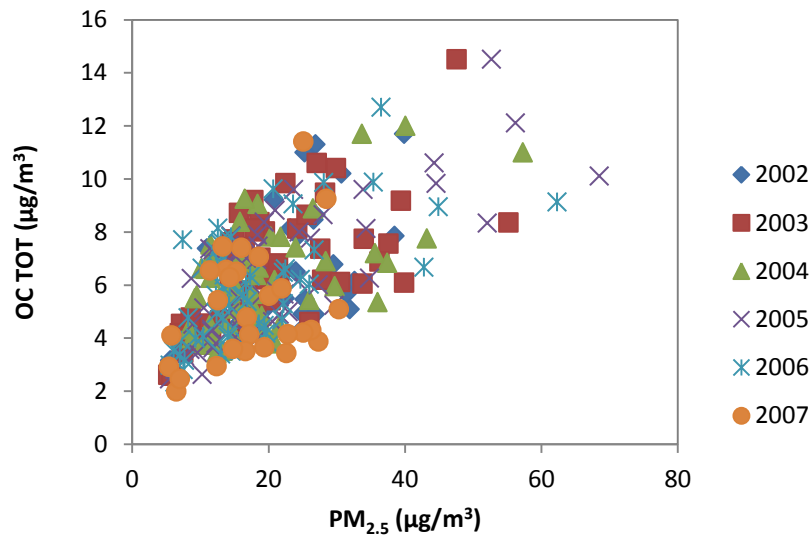


Figure (b). Scatter plot of OC mass concentration, obtained from Thermal Optical Transmittance (TOT) NIOSH 5040 method, versus $\text{PM}_{2.5}$ mass concentration in Los Angeles, segregated by year.

Table (a). Estimated OC artifacts (\pm standard errors) and the p values corresponding to the two-tailed t tests between OC artifact values in each two consecutive years. Errors correspond to one standard error.

Year	Estimated OC artifact ($\mu\text{g}/\text{m}^3$)	p value
2002	4.37 \pm 1.08	0.51
2003	3.69 \pm 0.48	0.57
2004	4.08 \pm 0.47	0.23
2005	3.32 \pm 0.41	0.53
2006	3.70 \pm 0.45	0.69
2007	3.32 \pm 1.04	

Comment 3: Is the estimated PM_{2.5} mass (as used to generate Table 3 and Figure 1) reconstructed from the resolved chemical components? If so, what is (are) the equation(s) used? Particularly, any consideration of possible seasonal variation of OM/OC when the mass was reconstructed?

Response: Since the total PM mass was included in the data matrix input to the PMF model as "total variable", the PMF model directly apportioned PM mass to each identified factor (Reff et al., 2007). Therefore, the estimated PM_{2.5} mass concentrations in Table 3 and Figure 1 are simply the sum of PM mass apportioned to each factor. Since PM mass was not reconstructed from the resolved chemical components, possible seasonal variations of OM/OC ratio would not affect the estimated PM_{2.5} mass concentrations. To clarify this, the following sentence has been updated in the revised manuscript:

Lines 282-284: "Table 3 presents the slope, intercept, and R² of the linear regressions between daily-resolved measured ambient PM_{2.5} and estimated PM_{2.5} mass concentrations, calculated by the sum of PM mass apportioned to each identified factor."

Comment 4: "standard error". In Figures 2-8, it is stated that "error bars correspond to one standard error". It is unclear, however, the exact meaning of one standard error. Are these based on all the daily concentration data (2002-2007) that fall in a given season? If so, it represents variation of the source contributions over the years. Or, are the standard errors based on error propagation of the uncertainties in the PMF resolved species concentrations using source profile uncertainties from the boot strap runs?

Response: The standard errors shown in Figures 2-8 were calculated based on all daily-resolved source contributions (between 2002 and 2007) that fall within a given season. We did not use the output of the bootstrap analysis to calculate the standard errors accompanying the seasonal averages. To clarify this, we added the following sentence in the revised manuscript:

Lines 226-227: “The standard errors accompanying the seasonal averages were calculated based on all daily-resolved source contributions that fall within a given season.”

Comment 5: Year-to-year variation. On page 20055, the authors stated that lack of year-to-year variation in source contributions can be deduced from the small standard errors in the 6-year average. However, these are absolute concentrations. I am surprised at the small standard errors in Table S2. PM2.5 decreased quite significantly from 2002-2007 at a number of sites, for example, PM2.5 in LA decreased from ~22 to ~15ug/m3 from 2000 to 2006. Should the year-to-year variation be much more significant than ~1.0ug/m3 as shown in the Table?

Response: The reviewer’s comment is duly noted. As mentioned in the manuscript and response to the previous comment, in this study we calculated the seasonal averages and the standard errors over the 6-year period. Given the extent of the present study (i.e. 2002 to 2007) and the sampling frequency at each site (every third or sixth day), seasonal-average concentrations were calculated over a large number of data points (ranging from 90 to 180) at each site and season. These numbers are presented in the following table:

Table (b). Total number of STN sampling days at each sampling site, between 2002 and 2007, that fall within a given season.

	Spring	Summer	Fall	Winter
El Cajon	147	131	118	125
Rubidoux	174	141	141	159
Los Angeles	72	78	72	68
Simi Valley	91	95	91	95
Bakersfield	146	126	109	126
Fresno	174	176	169	160
San Jose	126	119	134	128
Sacramento	164	178	178	173

It should be noted that the standard error quantifies how precisely we know the true mean of the samples by taking into account both the values of the standard deviation and the sample size:

$$\text{Standard error} = \frac{\text{Standard deviation}}{\sqrt{\text{Sample size}}}$$

Considering the large number of data points at a given site and season (refer to Table (b)), obtaining relatively small standard error is not unreasonable. However, to address the reviewer's comment and for further clarity, we added the sample size at each site and season to Table S3 and included a short discussion on the definition of the standard error to the supplemental text. These numbers would give the readers a better perception that over how many data points the reported averages and standard errors were calculated.

References

- Chen, L. W. A., Watson, J. G., Chow, J. C., and Magliano, K. L.: Quantifying PM_{2.5} Source Contributions for the San Joaquin Valley with Multivariate Receptor Models, *Environmental Science & Technology*, 41, 2818-2826, 2007.
- Flanagan, J. B., Jayanty, R. K. M., Rickman, E. E., Jr., and Peterson, M. R.: PM_{2.5} Speciation Trends Network: Evaluation of Whole-System Uncertainties Using Data from Sites with Collocated Samplers, *Journal of the Air & Waste Management Association*, 56, 492-499, 2006.
- Gutknecht, W., Flanagan, J., McWilliams, A., Jayanty, R. K. M., Kellogg, R., Rice, J., Duda, P., and Sarver, R. H.: Harmonization of Uncertainties of X-Ray Fluorescence Data for PM_{2.5} Air Filter Analysis, *Journal of the Air & Waste Management Association*, 60, 184-194, 2010.
- Hasheminassab, S., Daher, N., Ostro, B. D., and Sioutas, C.: Long-term source apportionment of ambient fine particulate matter (PM_{2.5}) in the Los Angeles Basin: A focus on emissions reduction from vehicular sources, *Environmental Pollution*, 193, 54-64, 2014.
- Hwang, I., and Hopke, P. K.: Comparison of source apportionments of fine particulate matter at two San Jose speciation trends network sites, *Journal of the Air & Waste Management Association* (1995), 56, 1287-1300, 2006.
- Kim, E., Hopke, P. K., and Qin, Y.: Estimation of Organic Carbon Blank Values and Error Structures of the Speciation Trends Network Data for Source Apportionment, *Journal of the Air & Waste Management Association*, 55, 1190-1199, 2005.
- Kim, E., and Hopke, P. K.: Source characterization of ambient fine particles at multiple sites in the Seattle area, *Atmospheric Environment*, 42, 6047-6056, 2008a.
- Kim, E., and Hopke, P. K.: Characterization of Ambient Fine Particles in the Northwestern Area and Anchorage, Alaska, *Journal of the Air & Waste Management Association*, 58, 1328-1340, 2008b.
- Kim, E., Turkiewicz, K., Zulawnick, S. A., and Magliano, K. L.: Sources of fine particles in the South Coast area, California, *Atmospheric Environment*, 44, 3095-3100, 2010.
- Polissar, A. V., Hopke, P. K., Paatero, P., Malm, W. C., and Sisler, J. F.: Atmospheric aerosol over Alaska: 2. Elemental composition and sources, *Journal of Geophysical Research: Atmospheres*, 103, 19045-19057, 1998.
- Reff, A., Eberly, S. I., and Bhave, P. V.: Receptor Modeling of Ambient Particulate Matter Data Using Positive Matrix Factorization: Review of Existing Methods, *Journal of the Air & Waste Management Association*, 57, 146-154, 2007.