Atmospheric
Chemistry
and Physics

Open Access

Discussions

EGU

# *Interactive comment on* "Temporal evolution of main ambient PM$_{2.5}$ sources in Santiago, Chile, from 1998 to 2012" *by* Francisco Barraza et al.

**Francisco Barraza et al.**

fjbarraz@uc.cl

Received and published: 15 May 2017

General comments

The authors give thanks to the reviewer for all the comments and questions that made it possible to vastly improve the manuscripts and the robustness of the results. The main changes to the new manuscript are focused in three areas: i) an appropriate description of the sampling and the quality assurance/quality control of the chemical data, ii) new comparative discussion of our results with other Latin American studies, but mostly with other cities abroad with similar multi-year data and methodology to ours, iii) more discussion about meteorological and climatic phenomena that play an important role on Santiago's air quality. More detail about the changes and answers to the reviewer are described below

C1

Response to comments by Reviewer 2

General: The project seems to be carefully thought out. The analytical methodology (PMF 5.0 and Unmix 6.0) seems appropriate; however, a separate detailed sampling and QA/QC section is needed. Language and spellings need to be improved. Concentrations should be expressed in 3 significant figures throughout the text and in the figures and tables. The author should compare the data with other studies in urban areas. As such I recommend that it be published with major revision:

1) Page 3: "$\mu$g/m3" should be "$\mu$g/m3" - be consistent throughout the text, figures, and tables.

Answer: This was corrected in the revised manuscript

2) Page 3: "24-hour" or "24-hours" or 24 h" – be consistent with one of them.

Answer: This was corrected in the revised manuscript

3) Page 3: No mention for the sampling and analysis for PM2.5? How PM2.5 samples were obtained? Which type of filter was used? Were the filters weighed in the clean room? Which analytical balance was used? Any QA/QC?

Answer: the follow section was added with this information.

2.2 Laboratory and QA/QC analysis Filters were inspected before being used, and the particles' concentration were determined gravimetrically using a microbalance, with a resolution of 0.01 mg. All filter (blank and filter samples) were stored at constant temperature (22±3°C) and relative humidity (40% HR ±3%) for a least 24-hours before being weighed. Those filters were analyzed using X-ray fluorescence (XRF) at the Desert Research Institute, Reno, NV, USA. The Ministry for the Environment provided the database containing the elemental analyses of those filters. In order to build statistical models based on robust chemical signals, we decided to keep only those elements selected in other studies that used the same data (CMM-MMA, 2011; Koutrakis et al., 2005; Sax et al., 2007; Valdes et al., 2012), for which more than 75% of the sam-

C2

ples contained valid measurements above the detection limit. The limit of detection (LOD) was calculated for each element as three times the standard deviation of the blank (blanks represented approximately 10% of the samples). This public database (gravimetry and elemental analysis) has been used in several studies and all of them have already described the laboratory and QA/QC methodology (CMM-MMA, 2011; Jhun et al., 2013; Koutrakis et al., 2005; Sax et al., 2007). Thus, out of the 49 elements reported, we only kept 22: Na, Mg, Al, Si, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, As, Se, Br, Sr, Ba and Pb. The missing data in these twenty-two species were dealt with as follows. First, we let the receptor models' (PMF5, UNMIX6) internal algorithms deal with them, which consists of replacing missing values with the median of the complete time-series for each species. Since replacing missing data with the median can lead to severe distortions in the data, we have also used a custom-written algorithm. This method interpolates up to three consecutive missing values using MATLAB's piecewise cubic interpolation algorithm. Sections of four or more consecutive missing values are filled in by summing up a mirrored copy of equal data length on both sides of the missing records, weighted by a cos2 function, thus ensuring that no artificial frequencies or discontinuities are introduced in the signal - the data filling algorithm is a MATLAB code that is available upon request to F. Lambert. We ensured that only relatively small missing data gaps were filled by using this method, so no large data sections were artificially created. The original and interpolated data are shown in Supplementary figure S1. Since our custom algorithm does not introduce discontinuities in the time series, we used this method for our analysis. In contrast, the receptor model results using the median-based missing data replacement can lose the seasonal signal of some species. Accordingly, the model results using the median-based filling algorithm yielded more variability in Cl, Ti Cr, Ni and As (species with important number of blanks, see Supplementary figure S1).

4) Page 3: A detailed QA/QC section for XRF analysis should be included. How often were the "QC" samples run? (What % age?). No estimates of recovery. What is the limit of quantitation? What is the uncertainty? Any blank correction? Precision and

accuracy?

Answer: A new section on Laboratory and QA/QC analysis was added with the most of this information.

5) Page 4: Did the authors find selenium?

Answer: Selenium was initially considered, but finally removed, because we couldn't get a source apportionment model using Se. Near 27% of Se data were either below LOD or missing; this might explain why we couldn't get a statistically significantly model that included Se.

6) Page 4: Did the authors do the PMF analysis for the missing data? How was this handled?

Answer: The missing data were treated in two separate ways. The first one consisted in leaving them blank and letting the models use their internal algorithm to deal with them, which consists of replacing them with the median values of the complete time-series, for each element. Since replacing missing data with the median can lead to distortions in the data, we also used a custom-written algorithm. This method interpolates up to three consecutive missing values using a piecewise cubic interpolation algorithm. Sections of four or more consecutive missing values are filled by summing up a mirrored copy of equal length of the data on both sides of the empty section, weighted by a cos2 function. We ensured that only relatively small gaps were considered to fill in the missing data, to avoid creating artificial variability in the data. Although both methods yielded comparable results, we have used the custom-written algorithm in this analysis, as it does not introduce discontinuities in the time series.

This methodology is now better explained and we have also added a new Supplementary Figure 1 in the revised manuscript that shows the effect of the filling algorithm.

The missing data in those twenty-two species were dealt with as follows. First, we let the receptor models' (PMF5, UNMIX6) internal algorithms deal with them, which con-

sists of replacing missing values with the median of the complete time-series for each species. Since replacing missing data with the median can lead to severe distortions in the data, we have also used a custom-written algorithm. This method interpolates up to three consecutive missing values using MATLAB's piecewise cubic interpolation algorithm. Sections of four or more consecutive missing values are filled in by summing up a mirrored copy of equal data length on both sides of the missing records, weighted by a cos2 function, thus ensuring that no artificial frequencies or discontinuities are introduced in the signal - the data filling algorithm is a MATLAB code that is available upon request to F. Lambert. We ensured that only relatively small missing data gaps were filled by using this method, so no large data sections were artificially created. The original and interpolated data are shown in Supplementary figure S1. Since our custom algorithm does not introduce discontinuities in the time series, we used this method for our analysis. In contrast, the receptor model results using the median-based missing data replacement can lose the seasonal signal of some species. Accordingly, the model results using the median-based filling algorithm yielded more variability in Cl, Ti Cr, Ni and As (species with important number of blanks, see Supplementary figure S1).

See supplementary Figure S1

7) Page 5: The contribution of Pb from industrial emissions cannot be ruled out. Motor vehicle is not the only source of Pb.

Answer: We agree with the reviewer. More discussion about the Pb from industrial emissions is added.

8) Page 7: "artefact" should be "artifact"

Answer: This was corrected in the revised manuscript

9) Page 8: "Gramsch et al. (2013)". Missing in the reference section.

Answer: This was corrected in the revised manuscript

10) Page 8 Lines 10 – 12: Did the private cars use diesel as a fuel? Primary source of

BC are emissions from diesel engines, cook stoves, wood burning and forest fires.
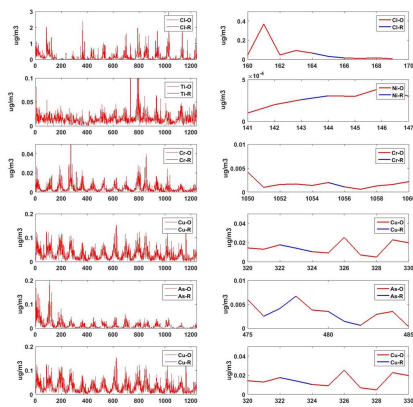
Answer: The sentence from lines 9-12 is a summary of the conclusions in (Gramsch et al, 2013). After reviewing that paper again, we have decided to drop this reference from that paragraph. The reason is that those authors compared ambient concentrations of BC in June 2005 and June 2007 in several streets (roadside sites). However, monthly precipitations were 173 and 80 mm, respectively, so the ambient BC changes reported by those authors are explained by changes in traffic emissions and meteorological conditions as well.

11) Page 12: "Boisier, J.P., . . .. . ..Mu?????oz, F.," should be corrected. Answer: This has been corrected in the revised manuscript

Figure S1 Replacement of missing data. Original data (red line) with missing data filled using a custom-written algorithm (blue line) for Cl, Ti, Cr, Ni, Cu and As are show on the left. The right hand side of the figure shows a zoom of the left panels for selected periods.



5

1

**Fig. 1.**

C7