**Point by Point Response to Review Comments (ACPD-14-57-2014)**
***By Anonymous Referee #2***

**Comment**

While the manuscript does present some useful and hard won data, the analysis and interpretation is quite weak and far from convincing. As noted below, this is the case throughout the paper, but most evident in the section on the estimation of (OC/EC)vehicle. There is little or no attempt to compare and contrast the results presented with the many similar studies done by numerous researchers (including some referred to in the manuscript). Given these serious deficiencies, I cannot recommend full publication in ACP.

**Response:**

We would like to clarify that our analysis of the hourly OCEC data in our study site (a roadside location in a busy urban district) has led us to derive a $(OC/EC)_{vehicle}$, not $(OC/EC)_{primary}$. The latter ratio is the ratio that has been examined and derived from ambient OCEC measurements in many other studies. As far as we are aware, this is the first study in which hourly OCEC measurements are used to derive $(OC/EC)_{vehicle}$ in a roadside environment and subsequently the contribution of vehicular emissions to carbonaceous $PM_{2.5}$ is estimated Unlike other studies reporting OCEC measurements, we did NOT attempt to estimate POC and SOC in this work.

**Comment**

Major Concerns:

Section 3.2: The authors describe and use an empirically based method for estimating the contributions of primary and secondary carbon to their measurement data.

**Response:**

In this work we aim to estimate $(OC/EC)_{vehicle}$ using roadside ambient OC EC measurements. We did not attempt to estimate the contributions of primary and secondary carbon, although the empirical approaches in getting $(OC/EC)_{vehicle}$ in this work and $(OC/EC)_{primary}$ in many studies in the literature are similar in that EC is used as a tracer for combustion sources. In our work, we regard EC as the tracer for vehicular emissions, considering our study location being a roadside site in a busy urban district.

**Comment**

The authors make reference to earlier work by Harrison's group (Castro et al., 1999) and Turpin's group (Lim and Turpin, 2002). However, their analysis produces results that is not at all convincing and leads me to believe that the assumptions required to yield reasonable estimates for this method may not hold in this case. The Castro et al. work showed urban OC/BC ratios ranging from 1.1. to 1.3, while the Lim and Turpin work showed OC/EC ratios ranging from 1.75 to 2.09 for the data grouped OC/EC ratio (note that all intercepts in this work were positive, on average). In contrast, the authors present estimates for the (OC/EC)pri ratio in Tables 1, 2, and 3 that range from 0.41 to 1.49 (Table 1), from .53 to 1.41 (Table 2) and from 0.44 to 8.56 (!) (Table 3).

**Response:**

(1) $(OC/EC)_{primary}$ is highly location-dependent, as the mix of primary combustion sources is specific to locations. None of the measurements reported in Castro et al (1999) and Lim and Turpin (2002) were made in a roadside environment comparable to the one described in this manuscript.

(2) We mistakenly used "(OC/EC)pri" in Table 3 caption. The correct one should be "$(OC/EC)_{min}$" The (OC/EC) ratios listed in Table 3 are not $(OC/EC)_{primary}$. Instead, they are calculated $(OC/EC)_{min}$ if OC and EC measurements in select sections of time periods in a day are used for the calculation. Therefore the value 8.56 (i.e., $(OC/EC)_{min}$ calculated for the 19:00-22:00 period in April 2012) should not be interpreted as an $(OC/EC)_{primary}$ estimate.

(3) We are aware of the studies in the literature reporting semi-continuous OC and EC measurements and subsequently estimating primary and secondary OC. Besides Lim and Turpin's work, Polidori et al. (2006) conducted a one-year OC and EC measurements in Pittsburgh supersite which is located in a suburban area. They estimated the primary OC/EC ratio (1.24–3.06) from data subsets which were obtained from time periods when primary emissions were dominant. Park et al. (2005) sampled for 9 months in Baltimore supersite which is in an urban residential area. They selected several time periods expected to be dominated by primary emissions and estimated the $(OC/EC)_{primary}$ ratio as 1.70–2.98. Plaza et al. (2006) collected samples during the summer and winter of 2003 at an urban background/suburban site in Madrid, India. OC/EC ratios were then calculated for different select periods (e.g. summer, traffic period, episode, etc.) and values of 0.84 and 1.02 were obtained for summer and autumn-winter, respectively. Lin et al. (2009) conducted a year-round measurement of OC and EC on the campus of PKU, which is located in urban area of Beijing, 200 m away from a major road with heavy traffic. The sampling inlet is 15 m above ground level. They adopted the minimum (OC/EC) ratio method to derive the $(OC/EC)_{primary}$ ratio of 1.5–2.6 for daytime and 1.0–2.6 for nighttime. Yu et al. (2009) collected samples in March 2006 at two sites in Mexico City. They estimated the $(OC/EC)_{primary}$ ratio to be 0.61 for the site closer to the city and 2.26 for the site further away from the downtown Mexico City. Hu et al. (2012) conducted OC and EC measurements during the PRIDE-PRD 2006 summer campaign at a rural site that is 50 km to the northwest of Guangzhou. They gave an estimate of 1.1 as $(OC/EC)_{primary}$ during local primary emission days. They also proposed values of 1.57 and 1.42 as $(OC/EC)_{primary}$ for daytime and nighttime, respectively.

The sampling site (MK) in our study is in a roadside environment in Hong Kong, located at the junction of two major roads with heavy vehicular traffic, a significant part of which comes from the diesel-fuelled bus fleet. The dispersion of the emissions is severely hampered by the high density of tall buildings along the roads, i.e., the "street-canyon effect". As a result, the EC concentrations at MK were significantly higher than those at other general air quality monitoring stations in Hong Kong and the OC/EC ratios were quite often observed to be below 1 during morning traffic rush hours.

**Comment**

Another problem is the large number of negative intercepts in their regressions, some as large in magnitude as -2.16. What does this mean physically? There is no acknowledgement of this being a problem, and no explanation. After reading this section, I have no choice but to conclude that this analysis method is either inappropriate for this data set (most likely), or improperly applied. The results simply don't make sense!

**Response:**

We agree that it is difficult to impart physical meaning to negative intercepts. To gain an understanding of this issue, we have re-examined the regression methods in deriving the slopes and the intercepts by applying different linear regression methods to the data sets and comparing the results. Linear regressions based on ordinary least squares (OLS) only take into account of measurement uncertainty for y-variable while assume no measurement errors for x-variable. Deming regressions consider measurement errors in both regressed variables x and y, which is expected to provide a better best-fit line. There are different forms of Deming regression because of different ways of representing measurement errors in x and y, i.e. $\omega(X_i)$ and $\omega(Y_i)$ in the equation below for $S$, which is the sum of the square of the perpendicular distances between the data points and the regression line (Saylor et al., 2006).

$$S = \sum [\omega(X_i)(x_i - X_i)^2 + \omega(Y_i)(y_i - Y_i)^2]$$

In our original manuscript, the Deming regression adopts a value of 1 for the ratio of $\omega(X_i)$ and $\omega(Y_i)$ (i.e., $\lambda$ in the expression below), i.e., equal measurement uncertainties for variable $X_i$ and $Y_i$.

$$\lambda = \omega(X_i)/\omega(Y_i)$$

Saylor et al (2006) compared two forms of Deming regression, default Deming regression with $\lambda$=1 and optimal Deming regression with an accurate estimate of $\lambda$ ($\lambda = Var(\varepsilon_{OC})/Var(\varepsilon_{EC})$ (where $Var(\varepsilon)$ is the variance of the measurement errors, $\varepsilon$). Using simulated EC and OC data, they demonstrated that the optimal Deming regression provides excellent results while the default Deming regression yields a slope of 6% larger than the true value and a negative intercept of -1.28 due to inaccurate representation of error variance. We therefore carried out optimal Deming regression on the data sets in this work, with λ estimated to be the average measurement uncertainty ratio of individual X and Y measurements.

The new regression results are shown in the revised Table 1.

(Revised) **Table 1**. Deming regression results of $(OC/EC)_{min}$ (the slope) and non-combustion term b (the intercept) on a monthly, seasonal and annual basis from the one-year carbon measurements at MK AQMS.

| Time period | No. of data | λ (Y/X) | $(OC/EC)_{min}$ (slope) | Non-combustion term b (intercept)[1] | Correlation coefficient ($R^2$) |
|---|---|---|---|---|---|
| May 2011 | 34 | 1.29 | 0.70 (±0.029) | 0.35 (±0.221) | 0.96 |
| Jun. 2011 | 24 | 1.03 | 0.44 (±0.037) | 0.16 (±0.257) | 0.87 |
| Jul. 2011 | 12 | 0.94 | 0.38 (±0.022) | -0.14 (±0.089) | 0.98 |
| Aug. 2011 | 24 | 1.03 | 0.51 (±0.012) | -0.22 (±0.052) | 0.99 |
| Sep. 2011 | 36 | 1.27 | 0.50 (±0.017) | 0.82 (±0.119) | 0.91 |
| Oct. 2011 | 38 | 1.43 | 0.69 (±0.018) | 0.34 (±0.098) | 0.93 |
| Nov. 2011 | 36 | 1.43 | 0.59 (±0.035) | 0.97 (±0.202) | 0.83 |
| Dec. 2011 | 36 | 2.12 | 1.46 (±0.025) | -0.19 (±0.161) | 0.99 |
| Jan. 2012 | 38 | 1.85 | 1.43 (±0.046) | -0.66 (±0.330) | 0.97 |
| Feb. 2012 | 36 | 1.68 | 0.97 (±0.027) | 0.33 (±0.193) | 0.99 |
| Mar. 2012 | 34 | 1.41 | 0.76 (±0.025) | 0.79 (±0.169) | 0.96 |
| Apr. 2012 | 36 | 1.29 | 0.31 (±0.038) | 1.91 (±0.231) | 0.63 |
| Summer | 94 | 1.11 | 0.47 (±0.004) | -0.11 (±0.018) | 0.97 |
| Fall | 72 | 1.39 | 0.62 (±0.011) | 0.40 (±0.061) | 0.92 |
| Winter | 142 | 1.78 | 0.96 (±0.008) | 0.13 (±0.051) | 0.93 |
| Spring | 66 | 1.29 | 0.60 (±0.022) | 0.33 (±0.137) | 0.89 |
| Year | 372 | 1.45 | 0.58 (±0.002) | 0.010 (±0.009) | 0.93 |

[1] Values inside parentheses are 95% confidence interval.

The revised Table 1 shows that the optimal Deming regression yields more physically reasonable intercepts. January 2012 data had the largest negative intercept (-0.66). To understand the issue of negative intercepts, let's examine regression lines obtained with ordinary least squares (OLS), default Deming, and optimal Deming regression for the January 2012 data, as shown in Figure R1. The OLS regression results in a positive intercept (0.86) while the two Deming regressions give negative intercepts. The different regression lines are a result of difference in assigning weights to individual observations. This result suggests that the regression line intercept is fairly sensitive to weights assigned to individual observations, or in another word, error variances for X and Y variables. For actual ambient data, it is difficult to identify a subset of data that is free of secondary OC contribution or such a subset of data does not exist. In addition, multiple primary combustion sources (having different $(OC/EC)_{primary}$) co-exist and their relative strengths vary with time at a given ambient location. Both factors would contribute to scattering of the data that are used for deriving $(OC/EC)_{primary}$, which in turn could lead to a negative intercept, as illustrated by Figure R1.

On the other hand, we note the slope is much less sensitive to different regression approaches. As we stressed earlier, the focus of this paper is $(OC/EC)_{vehicle}$ in our roadside study location, which is derived from the slope values. The negative intercept is therefore not an issue of consequence for this work.

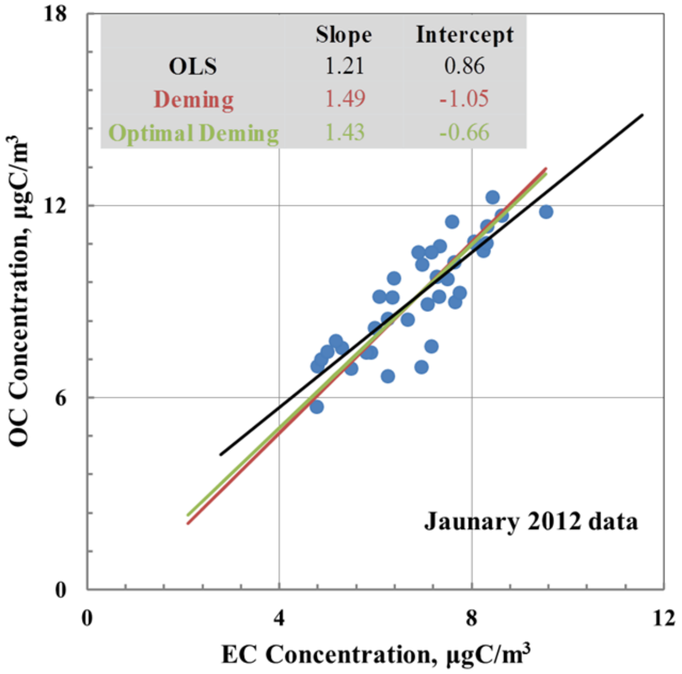

Figure R1. Regression lines of 5% lowest (OC/EC) data ($n$ = 38) for January 2011 by ordinary least squares (OLS), default Deming, and optimal Deming regression.

Tables 2 and 3 are revised to results using optimal Deming regressions.

(revised) Table 2. Estimated $(OC/EC)_{min}$ ($R^2$ in parentheses) by Deming regression of OC on EC from time periods of 7:00–11:00, 16:00–19:00 and 19:00–22:00 for individual months at MK AQMS.

| Lowest % by OC/EC | No. of data | $(OC/EC)_{min}$ (slope) | Non-combustion term b (intercept) | Correlation coefficient ($R^2$) |
|---|---|---|---|---|
| 5 | 94 | 0.47 (±0.004) | -0.11 (±0.018) | 0.97 |
| 10 | 188 | 0.54 (±0.003) | -0.13 (±0.013) | 0.96 |
| 20 | 376 | 0.62 (±0.002) | -0.06 (±0.009) | 0.94 |
| 30 | 564 | 0.72 (±0.002) | -0.17 (±0.008) | 0.93 |
| 40 | 752 | 0.77 (±0.002) | -0.11 (±0.007) | 0.93 |
| 50 | 940 | 0.84 (±0.001) | -0.14 (±0.006) | 0.93 |
| 60 | 1128 | 0.97 (±0.002) | -0.46 (±0.007) | 0.94 |
| 70 | 1316 | 1.09 (±0.002) | -0.72 (±0.007) | 0.95 |
| 80 | 1504 | 1.23 (±0.002) | -1.03 (±0.007) | 0.95 |
| 90 | 1692 | 1.32 (±0.002) | -1.05 (±0.007) | 0.95 |
| 100 | 1878 | 1.29 (±0.002) | -0.55 (±0.006) | 0.94 |

(revised) Table 3. Estimated $(OC/EC)_{min}$ ($R^2$ in parentheses) by Deming regression of OC on EC from time periods of 7:00–11:00, 16:00–19:00 and 19:00–22:00 for individual months at MK AQMS.

| Month | Time period | | |
|---|---|---|---|
| | 7:00–11:00 | 16:00–19:00 | 19:00–22:00 |
| May 2011 | 0.63 (0.85) | 1.02 (0.77) | 2.12 (0.94) |
| Jun. 2011 | 0.35 (0.68) | 0.53 (0.75) | 2.63 (0.99) |
| Jul. 2011 | 0.69 (0.80) | 0.93 (0.87) | 1.59 (0.97) |
| Aug. 2011 | 0.79 (0.84) | 0.51 (0.79) | 2.28 (0.98) |
| Sep. 2011 | 0.65 (0.69) | 0.28 (0.39) | 1.46 (0.91) |
| Oct. 2011 | 0.78 (0.69) | 0.62 (0.60) | 2.33 (0.96) |
| Nov. 2011 | 1.40 (0.89) | 1.32 (0.91) | 3.66 (0.99) |
| Dec. 2011 | 2.71 (0.97) | 2.90 (0.98) | 3.77 (0.99) |
| Jan. 2012 | 1.59 (0.91) | 1.45 (0.93) | 2.46 (0.98) |
| Feb. 2012 | 0.87 (0.89) | 0.62 (0.75) | 2.81 (0.98) |
| Mar. 2012 | 0.73 (0.86) | 0.91 (0.72) | 1.98 (0.93) |
| Apr. 2012 | 1.08 (0.79) | 1.61 (0.91) | 4.95 (0.99) |

**Comment**

Page 62, line 18: I believe the average should be 0.35 ugC/m3.

Response: Thanks the reviewer for pointing out our miscalculation and the revision was made accordingly.

**Comment**

Page 65, lines 24-28: While the different sampling periods could increase the scatter, it is not clear how or why it could cause bias!

Response: We agree with the reviewer that the explanation of the discrepancies about TC, OC and EC between semi-continuous measurements and filter-based data was not very clear. The paragraph is revised as follows,

"The comparison results showed that the TC concentrations from semi-continuous method agree fairly well with both Partisol filter measurements ($R^2$ = 0.98, $\%\overline{RB}$ = -29.6%, $\%\overline{RSD}$ = 23.4%) and high-volume filter measurements ($R^2$ = 0.99, $\%\overline{RB}$ = -16.4%, $\%\overline{RSD}$ = 15.2%). Fairly good correlations were also observed for all the OC measurements ($R^2$ = 0.97, $\%\overline{RB}$ = -33.8%, $\%\overline{RSD}$ = 27.7% for RT-OC vs. Partisol-OC and $R^2$ = 0.98, $\%\overline{RB}$ = -17.9%, $\%\overline{RSD}$ = 18.4% for RT-OC vs. HV-OC). The Y/X ratios average at 0.75±0.11 and 0.86±0.11 for RT-TC vs. Partisol-TC and RT-TC vs. HV-TC, respectively. The Y/X ratios for RT-OC vs. Partisol-OC and RT-OC vs. HV-OC were 0.72±0.14 and 0.85±0.18, respectively. These numbers suggest that in general both the TC and OC measurements from off-line filter samples were larger than those observed by the semi-continuous method. More specifically, the discrepancies were larger between RT data and Partisol data than those between RT data and HV data. In addition to the uncertainties associated with the sampling and analysis processes, another possible reason is the positive artifacts caused by organic vapor adsorption on the quartz fiber filters since no denuder was used in either Partisol or HV samplings. The amount of organic carbon adsorbed onto the quartz fiber filter in Partisol sampling was expected to be higher than that in HV sampling as the Partisol sampling face velocity is approximately half of the HV sampling face velocity (McDow et al., 1990).

The EC data comparisons showed a certain degree of scatter ($R^2$ = 0.93 for RT-EC vs. Partisol-EC and $R^2$ = 0.86 for RT-EC vs. HV-EC) while the average Y/X ratios for EC suggested that the semi-continuous data agree better with filter-based measurements (0.88±0.26 and 1.04±0.38 for RT vs. Partisol samples and RT vs. HV samples, respectively). The poor agreement between thermal EC from the semi-continuous analyzer and filter-based EC has been reported in several studies due to high detection limit or/and differences in the temperature programs (e.g. Schauer et al., 2003; Bae et al., 2004; Venkatachari et al., 2006) The discrepancies between RT-EC and filter-based EC measured at roadside in this study, on the other hand, might also be due to the different sampling durations of the measurement methods. The semi-continuous analyzer collected PM2.5 samples for a total of 1104 minutes on a daily basis, accounting for about 3/4 of the 24-hour period. The sampled air by the RT-OCEC analyzer might not be able to fully represent the 24-h integrated sampling period by the filter-based measurements because of the high carbon concentrations with large variations at MK."

**Comment**

Page 67, lines 6-19: OC observed at a roadside location is pretty much always dominated by local sources, vehicles in particular. If the OC and EC were not constantly being produced by local sources, there would be a very clear decrease in the middle of the day as the boundary layer reaches its maximum depth.

Response: It is not always the case in MK sampling area that OC is dominated by local sources. Two observations support this. First, the OC concentration levels were higher in winter (when the prevailing winds were northerlies) and transitional seasons (when winds came from both north and south) than that in summer (when southerlies were the prevailing winds). Second, OC concentrations were lower during holidays and higher during weekdays in summer while this weekday-holiday variation became less notable in other seasons. Based on these two observations, we could conclude that local emissions are important sources for OC at our sampling site while the OC level could also be influenced by regional transport of air pollutants when the winds come from the continent.

**Comment**

Page 68, lines 3-11: Figures 4 and 5 show NOx levels on the order of 100-150 ppb and ozone levels on the order of 10-15 ppb. The oxidants in this environment are totally controlled by NO2, and comparing OC to ozone alone is not all that informative.

Response: We agree with the reviewer that in the roadside environment, the ozone concentration level is much lower due to titration by NO. Nevertheless, an ozone peak appearing in the early afternoon was consistently observed in different seasons (Figure 5). Such a temporal characteristic tended to serve as an indicator of photochemical activity for investigating the SOC formation through photochemical processes.

**Comment**

Page 68, lines 12-17: The authors do not show this data. If the data is plotted with uncertainty error bars, are the "peaks" robust? The text is descriptive, but gives no physical interpretation.

Response: The diurnal profile of averaged OC/EC ratios over a year is shown in box plot in Figure R2 and this figure is now included in the supporting material document. It can be seen that three peaks of OC/EC ratio were observed during the day.

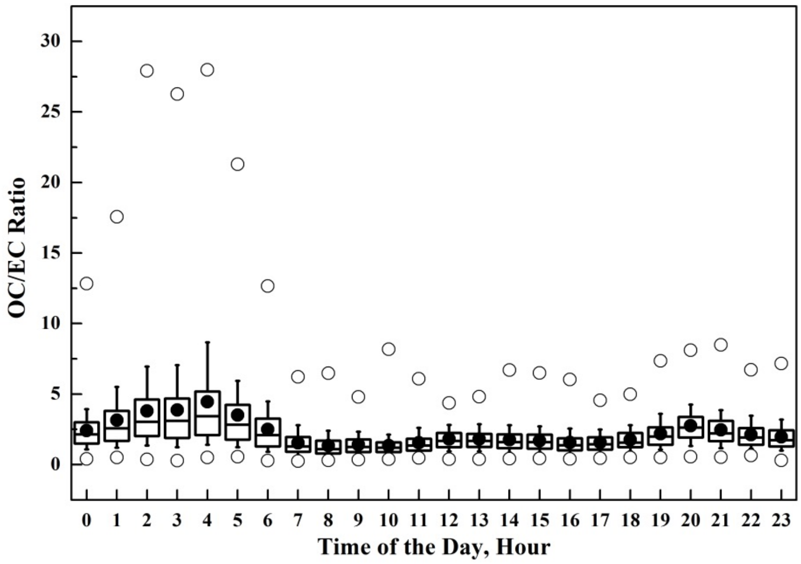

Figure R2. The diurnal variation of average OC/EC ratios. (The box length: the 25th and the 75th percentiles; the whiskers: the 10th and the 90th percentiles; the dot in the box: the average; the line in the box: the median; the circles: the minimum and maximum values).

Reference:

Saylor, R. D., E. S. Edgerton, and B. E. Hartsell: Linear regression techniques for use in the EC tracer method of secondary organic aerosol estimation, Atmos. Environ., 40, 7546-7556, 2006.