Supplemental materials for

Inter-comparison of NIOSH and IMPROVE protocols for OC and EC determination: Implications for inter-protocol data conversion

Cheng Wu¹, X. H. Hilda Huang¹, Wai Man Ng², Stephen M. Griffith³, Jian Zhen Yu^{1, 2, 3}

[1] Division of Environment, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

[2] Environmental Central Facility, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

[3] Department of Chemistry, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

This document contains four supporting tables, twenty-four supporting figures, and description of two additional reconstruction approaches other than the two described in the main text.

Descriptions of M2-1 and M3 reconstruction approaches

M2 can be further improved if chemical composition data is available. As discussed in the main text, the abundance of K^+ and Fe can affect EC discrepancy. To reflect the contributions from these factors, PC_{NSH_TOR} , K^+ and Fe are included to approximate PC_{IMP_TOR} by multiple linear regression (MLR), then equation 3 can be rewritten as

$$\mathbf{M2 - 1:} \quad EC_{IMP_TOR} = AEC_{NSH} + OC4_{NSH} - (a_1 \times PC_{NSH_TOR} + a_2 \times K^+ + a_3 \times Fe) \quad (S1)$$

a_1 , a_2 and a_3 are coefficients in multiple linear regression (MLR). K^+ is measured by ion chromatography and Fe is detected by X-ray fluorescence.

Another reconstruction approach (M3) also utilizes the relationship of PC. Combining Equations 3 and 4 yields

$$EC_{IMP_TOR} = EC_{NSH_TOT} + OC4_{NSH} + (PC_{NSH_TOT} - PC_{IMP_TOR}) \quad (S2)$$

In Equation S2, PC_{IMP_TOR} is the only unknown term on the RHS, and empirical regression is used to replace it with known variables leading to the third method (M3) for EC reconstruction

$$\mathbf{M3:} \quad EC_{IMP_TOR} = EC_{NSH_TOT} + OC4_{NSH} + a \times (PC_{NSH_TOT} - PC_{NSH_TOR}) + b \quad (S3)$$

Based on the linear relationship (Figure S11) between $(PC_{NSH_TOT} - PC_{NSH_TOR})$ and $(PC_{NSH_TOT} - PC_{IMP_TOR})$, EC_{IMP_TOR} can be estimated as shown in Equation S3.

Results of M2-1 and M3 reconstruction approaches

Reconstruction results on EC by M2-1 are shown in Figure S12a~c. The performance are similar between the three scenarios, and the R^2 by M2-1 is the highest among all four reconstruction approaches, which proves that including K^+ and Fe can improve the reconstruction comparing to M2. EC frequency distributions of M2-1 are shown in Figure S15 and Figure S1, which are similar to M2. OC reconstruction by M2-1 is shown in Figure S13, all of which exhibit an underestimation trend (slope 0.82~0.85). The OC distributions (Figure S17 and S19) by M2-1 is sharper than M2, confirming the advantage of using MLR. The reconstructed OC/EC ratios by M2-1 are underestimated from 50% to 64% as shown in Figure S14a~c. The bias is associated with the sharper OC/EC distribution by reconstruction compared to the measurements (Figure S19g~i)

The EC reconstruction results of M3 are shown in Figure S12d~f. M3 reconstruction by season (Figure S12e) or by site (Figure S12f) yield higher R^2 than single parameter reconstruction, but the degree of overestimation is also increased slightly. EC frequency distributions (Figures S15 and S16) also confirm that M3 can provide higher precision (narrower distributions) than M1, but with a higher bias (asymmetric distributions). OC reconstruction by M3 is underestimated as shown in Figure S13d~f by more than 12% (slope 0.85~0.88). The OC distributions (Figures S17j~l and S18j~l) indicate that season and site parameter scenarios yield little improvement compared to the single parameter scenario. The reconstructed OC/EC ratios by M3 are underestimated from 44% to 70% as shown in Figure S14d~f. The bias is due to a sharper OC/EC distribution by reconstruction compared to the measurements (Figures S19j~l and S20j~l).

SOC estimation by M2-1 and M3

The usability of reconstructed EC_{IMP_TOR} and OC_{IMP_TOR} for SOC estimation using M2-1 and M3 are investigated. To account for the temporal variations of $(OC/EC)_{pri}$, seasonal $(OC/EC)_{pri}$ are calculated using OC and EC reconstructed by M2-1, and M3 (Table S2). These $(OC/EC)_{pri}$ values are then subject to SOC estimation following Equation 10. It is very clear that the frequency distribution of reconstructed SOC deviate from the native SOC (Figure S24a and S24c). SOC by M2-1 and M3 are both underestimated by 32~34% by the regression slope. The moderate R^2 (Figure S24b and S24d) also suggests the SOC by reconstructed EC_{IMP_TOR} and OC_{IMP_TOR} are poorly correlated with SOC by measured EC_{IMP_TOR} and OC_{IMP_TOR} . Similar to the results shown in the main text, the significant bias and moderate correlations suggest that reconstructed EC_{IMP_TOR} and OC_{IMP_TOR} are not suitable for SOC estimation.

References

- Chow, J. C., Watson, J. G., Chen, L. W. A., Chang, M. C. O., Robinson, N. F., Trimble, D., and Kohl, S.: The IMPROVE-A temperature protocol for thermal/optical carbon analysis: maintaining consistency with a long-term database, *J. Air. Waste. Manage.*, 57, 1014-1023, Doi 10.3155/1047-3289.57.9.1014, 2007.
- Lin, P., Engling, G., and Yu, J. Z.: Humic-like substances in fresh emissions of rice straw burning and in ambient aerosols in the Pearl River Delta Region, China, *Atmos. Chem. Phys.*, 10, 6487-6500, 10.5194/acp-10-6487-2010, 2010.
- Tang, B., Chiang, Y., Baldwin, A., and Yeung, C.: Study of the integrated rail-property development model in Hong Kong, Research Centre for Construction & Real Estate Economics, Department of Building & Real Estate, Faculty of Construction & Land Use, the Hong Kong Polytechnic University, 2004.

Table S1. Temperature programs of the IMPROVE and NIOSH protocols.

Carrier Gas	Carbon Fraction		NIOSH		IMPROVE ²	
			T (°C) ¹	RT (s) ¹	T (°C)	RT (s)
Helium	AOC	OC1	310	80	120	150-580
		OC2	475	60	250	150-580
		OC3	615	60	450	150-580
		OC4	870	90	550	150-580
2% Oxygen in helium	AEC	EC1	550	45	550	150-580
		EC2	625	45	700	150-580
		EC3	700	45	800	150-580
		EC4	775	45		
		EC5	850	45		
		EC6	870	45		

¹ T: temperature (°C) and RT: Residence time (seconds).

² The IMPROVE temperature program was used for measurements reported in this work. Another related temperature protocol, termed IMPROVE_A, is typically adopted on DRI Model 2001 carbon analyzers. The IMPROVE_A temperature protocol defines temperature plateaus of 140 °C for OC1, 280 °C for OC2, 480 °C for OC3, and 580 °C for OC4 in a helium (He) carrier gas and 580 °C for EC1, 740 °C for EC2, and 840 °C for EC3 in a 98% He/2% oxygen (O₂) carrier gas (Chow et al., 2007). These temperatures used with the new hardware in DRI Model 2001 better match the sample temperatures experienced in the analysis using IMPROVE protocol on the previous models of OGC analyzers.

Table S2. OC_{4NSH}/TC_{NSH} dependency on K^+/EC_{NSH} by independent t-test using SPSS

Group Statistics					
		N	Mean	Std. Deviation	Std. Error Mean
OC4 _{NSH} /TC _{NSH}	low K ⁺	119	0.1407	0.05983	0.00548
	high K ⁺	122	0.2686	0.11182	0.01012

OC_{4NSH}/TC_{NSH}	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	31.32	0.000	-11.03	239	0.000	-0.12786	.01159	-0.15	-0.105
Equal variances not assumed			-11.11	186	0.000	-0.12786	.01151	-0.15	-0.105

Table S3. $(OC/EC)_{pri}$ calculation using MRS for individual seasons

Season	Measured OC and EC		Reconstructed OC and EC			
	NIOSH TOT	IMPROVE TOR	IMPROVE TOR			
			M1	M2	M2-1	M3
Spring	1.31	0.58	0.47	0.84	0.64	0.70
Summer	1.15	0.55	0.47	0.68	0.62	0.62
Fall	1.26	0.71	0.46	0.62	0.80	0.59
Winter	1.85	0.93	0.81	1.18	0.95	0.94

Table S4. Parameters for OC and EC reconstruction in Hong Kong.

Approach*			2011-2013		2011-2012		2011		2012		2013	
			a1,a2,a3	b	a1,a2,a3	b	a1,a2,a3	b	a1,a2,a3	b	a1,a2,a3	b
M2-1	by season	Spring	0.44		0.38		0.82		0.00		0.53	
			3.61	0.00	3.73	0.00	2.60	0.10	2.71	0.00	2.98	0.10
			0.00		0.00		0.42		0.42		0.00	
		Summer	0.02		0.00		0.23		0.00		0.06	
			2.68	0.02	2.74	0.00	2.49	0.00	3.32	0.00	1.94	0.00
			2.23		2.30		1.99		1.00		2.58	
		Fall	0.13		0.35		0.09		0.81		0.20	
			2.40	0.02	2.02	0.00	2.64	0.00	2.08	0.10	3.94	0.10
			2.42		2.07		0.00		0.87		0.00	
		Winter	1.04		0.96		1.01		0.72		1.15	
			1.29	0.10	1.22	0.10	1.13	0.10	2.51	0.00	1.37	0.10
			1.37		2.00		1.88		1.47		0.70	
	by site	Roadside MK	0.57		0.47		0.63		0.17		0.71	
			2.46	0.00	2.64	0.00	2.02	0.00	3.12	0.00	2.25	0.00
			0.27		0.52		1.35		0.49		0.00	
		TW	1.04		1.22		1.29		1.00		0.73	
			1.63	0.00	1.33	0.00	1.06	0.00	1.70	0.00	2.02	0.00
			1.19		0.74		1.31		0.69		2.48	
		YL	0.76		0.81		0.48		1.07		0.66	
			1.89	0.00	1.75	0.06	1.64	0.01	2.09	0.00	2.09	0.00
			1.33		0.96		3.14		0.00		2.02	
		Urban CW	0.98		0.95		0.86		1.13		0.99	
			1.18	0.00	1.09	0.00	0.61	0.00	2.25	0.00	1.38	0.00
			2.45		2.61		4.11		0.00		2.32	
		TC	0.95		0.93		0.87		1.03		0.94	
			1.63	0.00	1.54	0.06	1.38	0.10	2.29	0.00	1.75	0.00
			1.37		1.13		1.48		0.32		1.90	
		Urban sites combined	0.94		0.98		0.88		1.02		0.85	
			1.60	0.00	1.49	0.02	1.25	0.50	2.18	0.00	1.79	0.00
			1.47		1.19		2.10		0.26		2.12	
	Suburban WB		1.42		1.60		1.51		1.44		1.14	
			1.13	0.10	0.90	0.10	0.60	0.10	1.98	0.00	1.21	0.10
			1.27		0.87		2.12		0.25		2.91	
M3	by season	Spring	a	b	a	b	a	b	a	b	a	b
			0.91	-0.17	1.02	-0.18	0.78	-0.12	1.04	-0.23	0.92	-0.23
			0.77	0.00	0.47	0.00	0.73	0.00	0.57	-0.11	1.30	-0.07
		Summer	0.71	-0.25	0.60	-0.19	0.92	-0.18	0.47	-0.39	0.97	-1.21
			0.60	-0.22	0.77	-0.37	1.67	-2.38	0.76	-0.24	0.49	-0.14
			0.60	-0.22	0.77	-0.37	1.67	-2.38	0.76	-0.24	0.49	-0.14
		Fall	1.02	0.09	0.97	0.08	0.90	0.07	1.13	0.10	1.11	0.14
			0.55	-0.01	0.48	0.00	0.55	0.01	0.40	0.00	0.85	-0.06
			0.62	-0.01	0.56	-0.01	0.63	0.00	0.77	-0.23	0.75	-0.02
		Winter	0.49	0.01	0.42	0.01	0.46	0.01	0.42	-0.10	0.69	-0.02
			0.50	0.00	0.44	0.00	0.47	0.00	0.79	-0.19	0.81	-0.06
			0.54	0.00	0.48	0.00	0.53	0.00	0.56	-0.12	0.76	-0.04
	by site	Urban sites combined	0.33	-0.09	0.23	-0.07	0.16	-0.06	0.63	-0.21	0.48	-0.13

*Reconstruction equations:

$$\mathbf{M2-1:} \quad EC_{IMP_{TOR}} = AEC_{NSH} + OC4_{NSH} - (a_1 \times PC_{NSH_{TOR}} + a_2 \times K^+ + a_3 \times Fe + b)$$

$$\mathbf{M3:} \quad EC_{IMP_{TOR}} = EC_{NSH_{TOT}} + OC4_{NSH} + a \times (PC_{NSH_{TOT}} - PC_{NSH_{TOR}}) + b$$

Figure S1 Data from six Hong Kong Air Quality Monitoring Sites (AQMS) are used in this study, including a roadside site Mong Kok (MK), four urban sites Central/Western (CW), Tsuen Wan (TW), Tung Chung (TC) and Yuen Long (YL), and a suburban site suburban Clear Water Bay (WB). The lines in different color represent the mass transit railway (MTR) in Hong Kong, which covers most urban areas (41% of population live within 500 m of MRT station (Tang et al., 2004))

Figure S2 Comparison of total carbon quantification by IMPROVE TOR and NIOSH TOT protocols. (a) Scatter plot with weighted orthogonal distance regression (WODR). (b) Histogram of NIOSH TC/IMPROVE TC ratio.

Figure S3 Seasonal variations of EC discrepancy attribution.

Figure S4 Inter-site and seasonal variations of NIOSH carbon fractions in HK samples.

Figure S5 Scatter plots with linear regressions between IMPROVE TOR EC (EC_{IMP_TOR}) and NIOSH TOT EC (EC_{NSH_TOT}). The slopes represent the degree of discrepancy between the two protocols on EC determination. The color coding represent K^+ / EC_{NSH} ratio, higher values are more likely associated with biomass influence. (a)~(c) EC_{IMP_TOR} in y axis and EC_{NSH_TOT} in x axis. (a) All samples. (b) Samples with lowest 10% K^+ / EC_{NSH} ratio. (c) Samples with highest 10% K^+ / EC_{NSH} ratio. (d)~(f) EC_{IMP_TOR} in y axis and $EC_{NSH_TOT} + OC4_{NSH}$ in x axis. (d) All samples. (e) Samples with lowest 10% K^+ / EC_{NSH} ratio. (f) Samples with highest 10% K^+ / EC_{NSH} ratio.

Figure S6 Frequency distributions of $OC4_{NSH}/TC$ ratio. (a) All samples. (b) Samples with lowest 10% K^+/EC_{NSH} ratio. (c) Samples with highest 10% K^+/EC_{NSH} ratio.

Figure S7 Scatter plots with linear regressions between IMPROVE TOR EC (EC_{IMP_TOR}) and NIOSH TOT EC (EC_{NSH_TOT}). The slopes represent the degree of discrepancy between the two protocols on EC determination. The color coding represent Fe/EC_{NSH} ratio. (a)~(c) EC_{IMP_TOR} in y axis and EC_{NSH_TOT} in x axis. (a) All samples. (b) Samples with lowest 10% Fe/EC_{NSH} ratio. (c) Samples with highest 10% Fe/EC_{NSH} ratio. (d)~(f) EC_{IMP_TOR} in y axis and $EC_{NSH_TOT} + OC4_{NSH}$ in x axis. (d) All samples. (e) Samples with lowest 10% Fe/EC_{NSH} ratio. (f) Samples with highest 10% Fe/EC_{NSH} ratio.

Figure S8 Scatter plots with linear regressions between IMPROVE TOR EC (EC_{IMP_TOR}) and NIOSH TOT EC (EC_{NSH_TOT}). The slopes represent the degree of discrepancy between the two protocols on EC determination. The color coding represent Al/ EC_{NSH} ratio. (a)~(c) EC_{IMP_TOR} in y axis and EC_{NSH_TOT} in x axis. (a) All samples. (b) Samples with lowest 10% Al/ EC_{NSH} ratio. (c) Samples with highest 10% Al/ EC_{NSH} ratio. (d)~(f) EC_{IMP_TOR} in y axis and $EC_{NSH_TOT}+OC4_{NSH}$ in x axis. (d) All samples. (e) Samples with lowest 10% Al/ EC_{NSH} ratio. (f) Samples with highest 10% Al/ EC_{NSH} ratio.

Figure S9 Linear regression between EC_{NSH_TOT} and EC_{IMP_TOR} for M1 reconstruction.

Figure S10 Linear relationship between PC_{NSH_TOR} and PC_{IMP_TOR} for M2 reconstruction.

Figure S11 Linear relationship between $PC_{NSH_TOT} - PC_{NSH_TOR}$ and $PC_{NSH_TOT} - PC_{IMP_TOR}$ for M3 reconstruction.

Figure S12 Comparison of reconstructed EC_{IMP_TOR} and measurement EC_{IMP_TOR} in year 2013. (a) Regression by single parameter for whole dataset using M2-1. (b) Regression by season specific parameters using M2-1. (c) Regression by site specific parameters using M2-1. (d) Regression by single parameter for whole dataset using M3. (e) Regression by season specific parameters using M3. (f) Regression by site specific parameters using M3.

Figure S13 Comparison of reconstructed OC_{IMP_TOR} and measurement OC_{IMP_TOR} in year 2013. (a) Regression by single parameter for whole dataset using M2-1. (b) Regression by season specific parameters using M2-1. (c) Regression by site specific parameters using M2-1. (d) Regression by single parameter for whole dataset using M3. (e) Regression by season specific parameters using M3. (f) Regression by site specific parameters using M3.

Figure S14 Comparison of reconstructed OC/EC_{IMP_TOR} and measurement OC/EC_{IMP_TOR} in year 2013. (a) Regression by single parameter for whole dataset using M2-1. (b) Regression by season specific parameters using M2-1. (c) Regression by site specific parameters using M2-1. (d) Regression by single parameter for whole dataset using M3. (e) Regression by season specific parameters using M3. (f) Regression by site specific parameters using M3.

Figure S15 Frequency distributions of reconstructed EC (in blue) compared with measured EC (in red). The two distributions in each plot are overlaid, the darker color represent the overlapped areas of the two distributions.

Figure S16 Frequency distributions of difference between reconstructed EC and measured EC.

Figure S17 Frequency distributions of reconstructed OC (in blue) compared with measured EC (in red). The two distributions in each plot are overlaid, the darker color represent the overlapped areas of the two distributions.

Figure S18 Frequency distributions of difference between reconstructed OC and measured OC.

Figure S19 Frequency distributions of reconstructed OC/EC ratio (in blue) compared with measured OC/EC ratio (in red). The two distributions in each plot are overlaid, the darker color represent the overlapped areas of the two distributions.

Figure S20 Frequency distributions of difference between reconstructed OC/EC and measured OC/EC

Figure S21 Monthly time series of measured and reconstructed EC and OC at 5 sites in HK (four urban sites Central/Western (CW), Tsuen Wan (TW), Tung Chung (TC) and Yuen Long (YL), and a suburban site suburban Clear Water Bay (WB)). The reconstruction uses M2 site-specific parameters obtained from 2011-2012 data as shown in Table 2. Sample size of Jan 2012 is too small and not included in the plot. Samples of Oct 2013 analyzed by NIOSH protocol did not pass QA/QC check, so the corresponding reconstructed results are not included in the plot.

Figure S22 Primary ratio of WSOC/Sugars from a emission study in PRD region (Lin et al., 2010).

Figure S23 Linear regression between SWSOC and SOC at WB site. (a) SOC by NIOSH TOT (b) SOC by IMPROVE TOR

Figure S24 Histogram comparison of original $\text{SOC}_{\text{IMP_TOR}}$ (in red) with reconstructed $\text{SOC}_{\text{IMP_TOR}}$ (in blue): (a) by M2-1. (c) by M3. Scatter plot comparison of original $\text{SOC}_{\text{IMP_TOR}}$ (in x axis) with reconstructed $\text{SOC}_{\text{IMP_TOR}}$ (in y axis): (b) by M2-1, (d) by M3