Supporting Information for:

Assessment of reductions in emission-driven air pollution during the Beijing Olympic Games, Shanghai World Expo, Guangzhou Asian Games and Wuhan COVID-19 Lockdown

Ying Li[1,2], Haoxiang Xu[1]

[1] *Department of Ocean Science and Engineering, the Southern University of Science and Technology, Shenzhen, China*

[2] *Center for Oceanic and Atmospheric Science at susTech (COAST), Southern University of Science and Technology, Shenzhen, China*

Correspondence to: Ying Li (liy66@sustc.hk)

This file includes:

Page S1-S7

Additional data and method description

Table S1

Fig. S1

1. Data

Specifically, NO_x, and $PM_{2.5}$ pollution data were measured at Peking University for the period 1 July to 31 August in 2007 and 2008, used for Beijing Olympic Games; SO_2, NO_2, and PM_{10} pollution data of multiple stations for the period 1 May to 30 November in 2009 and 2010 were analyzed for the Shanghai World Expo; Five pollutants (SO_2, NO_2, PM_{10}, O_3, O_x) at Luhu Park station in Guangzhou for the period 1 October to 31 November of 2009 and 2010 were analyzed for the Guangzhou Asian Games; Pollutant $PM_{2.5}$, NO_2, SO_2, O_3 were measured for the period 23 January to 10 March in 2019 and 2020, at Ganghua station and Moon Lake station, for the analysis of Wuhan COVID-19 Lockdown. The types of pollutant used, period of event, and period of analysis, name of air quality and meteorology stations, and location (longitude and latitude coordinates) of all stations were summarized in Table S1. The central part of Fig. S1 gives the location of four mega-cities in China. The corresponding stations (air quality and meteorology stations) applied in each event are also shown in the four subfigures (Fig. S1 (a), (b), (c), and (d)) which are satellite image taken from Amap. Satellite image provide a cleaner view to distinguish the urban, suburban, rural region, and vegetation. Within each satellite images in Fig. S1, red dot represent the air quality station while green dot is the meteorology station.

By taking the benefit from the air quality monitoring network in Shanghai city, a more comprehensive analysis can be carried out for the Expo event. To meet the requirement of sample size in the decomposition method, our study period (1 May to 30 November) is set one month longer than the Expo period (1 May to 31 October). Therefore, SO_2, NO_2, and PM_{10}

pollution data of multiple stations (Pudong, DianShan Lake, city average) for the period 1 May to 30 November of 2009 and 2010 are analyzed for the Shanghai World Expo. PuDong (PD) station located at the urban area close to the city center is referred as a representative urban site of Shanghai. DianShan Lake (DSL) station is a rural station located at rural area by the DianShan River near the boundary of Shanghai and Zhejiang province. However, since relative high pollution levels were observed at DSL station than the other urban stations in Shanghai, it is denoted as a regional downwind station for YRD area (Zhao et al., 2015; Li et al., 2015). DSL reflects the emission impact from Shanghai, Zhejiang, Jiangsu and long-range transport due to its location, therefore, DSL is select to represent the regional background situation for the Shanghai Expo event. City average represents the average values of all the 9 national control stations of Shanghai. The wind data are taken from the closest Global Telecommunications System (GTS) stations of the World Meteorology Organization (WMO), which representing the regional meteorology. Satellite map and location of stations are provided in Fig. S1.

2. Methodology

Li et al. (2014) proposed a method to differentiate the effects of wind and non-wind (e.g., emissions) contributions by connecting the pollution changes to the changes of wind and pollution roses at two different time periods. Here, we briefly summarize the method to make this paper self-contained.

The time average of any pollutant concentration over the period T can be written as the dot product of the wind rose (normalized frequency distribution of wind as a function of speed and direction, $\boldsymbol{F}_{s\theta}$) and pollution rose (the average pollutant concentration as a function of wind speed and direction, $\boldsymbol{C}_{s\theta}$) as

$$\overline{c} = \boldsymbol{F}_{s\theta} \cdot C_{s\theta}, \tag{1}$$

where s denotes wind speed and θ denotes the wind direction. For two measurements at two time periods T_1 and T_2 (hereinafter, the subscripts 1 and 2 denote two time periods), by Taylor expansion, the average pollution concentrations at period T_2 can be written as,

$$\overline{c}_{F_2,C_2} = \overline{c}_{F_1,C_1} + \left.\frac{\partial \overline{c}}{\partial \boldsymbol{F}}\right|_{T_1} \cdot \Delta \boldsymbol{F} + \left.\frac{\partial \overline{c}}{\partial \boldsymbol{C}}\right|_{T_1} \cdot \Delta \boldsymbol{C} + \Delta \boldsymbol{F}^T \left.\frac{\partial^2 \overline{c}}{\partial \boldsymbol{F} \partial \boldsymbol{C}}\right|_{T_1} \Delta \boldsymbol{C} + \ldots . \tag{2}$$

We can assume that the *changes* in $\boldsymbol{F}$ and $\boldsymbol{C}$ are generally independent, which means the factors causing $\boldsymbol{C}$ (e.g., anthropogenic emissions) variation are generally not those that cause the $\boldsymbol{F}$ (wind) variation, and vice versa. Then we obtain $\frac{\partial \overline{c}}{\partial \boldsymbol{F}} = \boldsymbol{C}$, $\frac{\partial \overline{c}}{\partial \boldsymbol{C}} = \boldsymbol{F}$, $\frac{\partial^2 \overline{c}}{\partial \boldsymbol{F} \partial \boldsymbol{C}}$ is an identity matrix, and all of the higher order derivatives of $\overline{c}$ with respect to $\boldsymbol{F}$ and $\boldsymbol{C}$ are 0. Then equation (2) can be written as

$$\overline{c}_{F_2,C_2} - \overline{c}_{F_1,C_1} = (\mathbf{F}_{s_2\theta_2} - \mathbf{F}_{s_1\theta_1}) \cdot \mathbf{C}_{s_1\theta_1} + (\mathbf{C}_{s_2\theta_2} - \mathbf{C}_{s_1\theta_1}) \cdot \mathbf{F}_{s_1\theta_1} + (\mathbf{F}_{s_2\theta_2} - \mathbf{F}_{s_1\theta_1}) \cdot (\mathbf{C}_{s_2\theta_2} - \mathbf{C}_{s_1\theta_1}) \tag{3}$$

The first term on the right hand side of equation (3) explains the change in pollutant concentration in period T_2 as a result of the change in the wind rose if there is no change in the pollution rose (it is defined as the *wind effect,* C_w). The second term explains the change in pollutant concentration due to the change in the pollutant rose if there is no change in the wind rose (it is defined as the *non-wind effect* C_{nw}). The last term is the non-linear residue (it is defined as the *nonlinear effect* C_{nl}). We can explicitly write them as

$$\begin{aligned} C_w &= (\mathbf{F}_{s_2\theta_2} - \mathbf{F}_{s_1\theta_1}) \cdot \mathbf{C}_{s_1\theta_1} = \bar{c}_{F_2,C_1} - \bar{c}_{F_1,C_1} \\ C_{nw} &= (\mathbf{C}_{s_2\theta_2} - \mathbf{C}_{s_1\theta_1}) \cdot \mathbf{F}_{s_1\theta_1} = \bar{c}_{F_1,C_2} - \bar{c}_{F_1,C_1} \\ C_{nl} &= (\mathbf{F}_{s_2\theta_2} - \mathbf{F}_{s_1\theta_1}) \cdot (\mathbf{C}_{s_2\theta_2} - \mathbf{C}_{s_1\theta_1}) = \bar{c}_{F_2,C_2} - \bar{c}_{F_2,C_1} - \bar{c}_{F_1,C_2} + \bar{c}_{F_1,C_1} \end{aligned} \tag{4}$$

An index L is defined to quantify the magnitude of the non-linear term with the wind and non-wind effects,

$$L = 1 - \frac{\left|\bar{c}_{F_2,C_2} - \bar{c}_{F_1,C_2} - \bar{c}_{F_2,C_1} + \bar{c}_{F_1,C_1}\right|}{\left|\bar{c}_{F_2,C_1} - \bar{c}_{F_1,C_1}\right| + \left|\bar{c}_{F_1,C_2} - \bar{c}_{F_1,C_1}\right| + \left|\bar{c}_{F_2,C_2} - \bar{c}_{F_1,C_2} - \bar{c}_{F_2,C_1} + \bar{c}_{F_1,C_1}\right|}. \tag{5}$$

These terms are simply the dot product combinations of the wind rose ***F*** and pollution rose ***C*** from the two periods. The range of L is between 0 and 1. Monte Carlo assessment shows the linear decomposition satisfies 90%, 95%, and 99% confidence level (*C-L*) with L index equal or larger than 0.87, 0.95, and 0.98, respectively. However, if L is significantly smaller than 1 (e.g. <0.87), there is considerable interaction between the wind and non-wind factors. This nonlinear residue term includes both the unresolved effects occurring in time-scales shorter than the sampling time (e.g., hourly), and other potential non-linear effects. More details about this approach can be found in *Li et al.* (2014).

In this study, to construct wind and pollution roses based on an actual observational dataset, we use wind direction and speed intervals of 10 degrees and 1 m/s, respectively. The roses are 36 x 10 (treat wind speed $\geq$ 9 m/s as one bin) in size. We do not include hours with missing wind or pollution data in constructing either the wind rose or the pollution rose to minimize uncertainty. The correction of the missing data causes that the annual mean value calculated by the dot product of the wind and pollution roses could be slightly different with the average value calculated by the time average. More details of the methodology can be referred to Li et al. (2014).

Table S1. Information and data used in the decomposition analysis for the four events

Events	Beijing Olympic	Shanghai expo	Guangzhou Asian game	Wuhan Lockdown
Event period	2008-8-8 2008-8-24	2010-5-1 2010-10-31	2010-11-12 2010-11-27	2020-1-23 2020-4-8
Data available for period 1	2007-7-1 2007-8-31	2009-5-1 2009-11-30	2010-10-1 2010-11-31	2019-1-23 2019-3-10
Data available for period 2	2008-7-1 2008-8-31	2010-5-1 2010-11-30	2011-10-1 2011-11-31	2020-1-23 2020-3-10
Pollutants	NO_x $PM_{2.5}$	SO_2 NO_2 PM_{10}	SO_2 NO_2 PM_{10} O_3 O_x	$PM_{2.5}$ NO_2 SO_2 O_3 O_3(8h)
Air quality measurement station	Peking University station	DianShan Lake station (DSL), PuDong station (PD), city average	Luhu Park station	Ganghua station, Moon Lake station
Wind station	ZGGG	ZSSS	ZBAA	ZHHH
Coordinator of air quality measurement station	[39.99, 116.32]	DSL: [31.22,121.53] PD: [31.09, 120.98]	[23.15, 113.28]	Ganghua: [30.61, 114.43] Moon Lake: [30.55, 114.25]
Coordinator of GTS	[39.93, 116.28]	[31.20, 121.33]	[23.40,113.3]	[30.62, 114.12]

Note: The two date inside period rows (row 2-3) are the starting and end of time in format of year-month-day. Content of row 8-9 are geographic coordinate system (in format of [latitude, longitude]). O_3(8h) in the Pollutant row is the mean ozone concentration for the previous 8 hour.

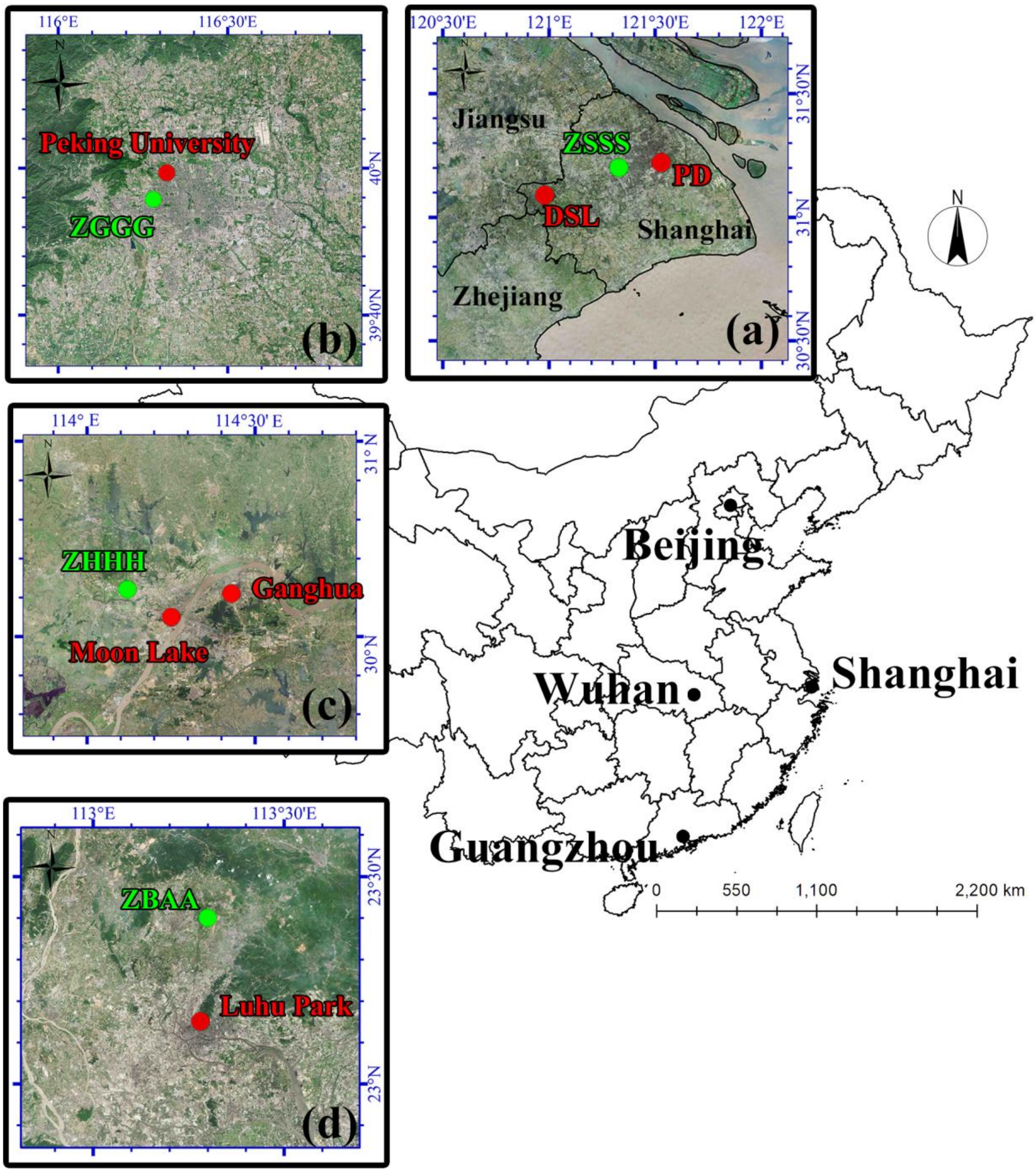

Fig. S1. The boundary silhouette map shows the location of four mega cities in China. Solid lines represent the provinces boundary. The four satellite images are (a) Shanghai; (b) Beijing; (c) Wuhan; and (d) Guangzhou. For the satellite images: red dot represents the air quality station and green dot is meteorology station; black solid line in (a) is the provinces boundary between Shanghai, Jiangsu Province, and Zhejiang province.