



Supplement of

Impact of western Pacific subtropical high on ozone pollution over eastern China

Zhongjing Jiang et al.

Correspondence to: Jing Li (jing-li@pku.edu.cn)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

1. Data quality control procedure

As raw ozone data released by the China National Environmental Monitoring Centre (www.cnemc.cn) may suffer from various problems like instrumental errors and systematic bias, a series of quality control was applied to remove outliers and invalid measurements. We describe the data quality control protocol in the following five consecutive steps.

(1) Setting up lower and upper boundaries:

We set lower and upper limit values for the valid range of ozone concentrations of $[0,1000] \ \mu g \ m^{-3}$. Although ozone concentration can reach several hundred in some extreme cases, data larger than 1000 $\ \mu g \ m^{-3}$ was found to always show a discontinuity with adjacent data. We thus considered it erroneous and marked them as NaN.

(2) Removing a string of data with the same value:

If in 5 consecutive hours, at least 4 are not NaN but having the same value, these data were eliminated.

(3) Removing abrupt jumps in data:

As the product of a stationary chemical reaction, it is unreasonable to find huge spikes in the ozone time series. Here we used two tests to remove these unrealistic changes. Bilateral test (at least 7 consecutive data in 7 hours) is used to remove cases like Fig.S1 a and unilateral test (at least 4 consecutive data in 4 hours) is used to remove cases when the right side or left side is a NaN value, examples of which are shown in Fig.S1 b&c.

- (4) Continuity check:
 - In 12 consecutive hours, if there is at most 1 valid data, we set this series of data as NaN.
 - In 24 consecutive hours, if there is at most 3 valid data, we set this series of data as NaN.
 - 3) In 48 consecutive hours, if there is at most 5 valid data, we set this series of data as NaN.
- (5) Checking the number of valid data:

We abandoned the sites which have more than half missing values in a month.

These quality control procedures removed approximately 2.35% of the raw data from 2014 to 2018. For each site, 84% to 100% (average 98% for all sites) of the hourly measurements were retained. After completing these data quality control procedures, although the amount of available data is reduced, the analysis results should be more reliable since the apparently unreasonable values are excluded. In addition, as the invalid measurement appears more frequently at night and ozone concentration always reach its peak in the afternoon, daily maximum 8 h average (MDA8) ozone metrics used in this study is less affected by the quality-control related data reduction.



Figure S1. Three the examples of invalid data.

2. Dry deposition



Figure S2. The changes in dry deposition flux and dry deposition velocity at the surface level in GEOS-Chem model. The first row shows the differences between strong and normal WPSH conditions, and the second row shows those between weak and normal WPSH conditions.

3. Correlation analysis



Figure S3. Correlation coefficients, between simulated daily MDA8 ozone concentrations and meteorological variables including SLP, precipitation, relative humidity, cloud cover, solar radiation, 2 m temperature, wind speed, 10 m U wind, and 10m V wind calculated for the summer periods from 2014 to 2018.

4. Model evaluation



Figure S4. Observed (symbols) and simulated (filled contours) seasonal mean surface ozone MDA8 concentrations (ppbv) over China for summer from 2014 to 2018 (a-e).

4. Model evaluation



Figure S5. Normalized mean bias (%) between simulated and observed seasonal mean surface ozone MDA8 concentrations (ppbv) over China for summer from 2014 to 2018 (a-e).