Supplement of

A machine learning approach to quantify meteorological drivers of ozone pollution in China from 2015 to 2019

Xiang Weng et al.

Correspondence to: Xiang Weng (x.weng@uea.ac.uk)

The copyright of individual parts of the supplement might differ from the article licence.
Supplementary Figures

Figure S1. Monthly average MDA8 ozone during 2015-2019 in 4 megacity clusters (BTH, PRD, Sichuan and YRD) of China. BTH shows a highest monthly average MDA8 ozone in June, while PRD’s ozone is relatively low in summer but high in September.
Figure S2. An example of a station’s MDA8 ozone data (station ID: 2554A) that is considered to be unreliable. It is characterized by a relative long period (2015-Sep-15 to 2015-Nov-11) with MDA8 ozone consistently at 14 μg m⁻³. Stations in this case are not used in this study.

(a)

Figure S3. An example of a station’s ozone data (station ID: 2248A) that is considered to be unreliable. A spike of MDA8 ozone (a) and hourly ozone (b) occurred on 2016-Aug-16 with at about 300 μg m⁻³. After the spike, hourly ozone dropped to nearly 10 to 20 μg m⁻³ and the value remains stable throughout September and October of 2016.

(b)
Figure S4. Coefficient of determination ($R^2$) between deseasonalized observation ozone and deseasonalized predicted values of MDA8 in multiple linear regression (MLR) with all 11 meteorological predictors (a) and stepwise MLR with 3 most dominant local meteorological predictors as input (b).

Figure S5. relationship between deseasonalized temperature at 2m and deseasonalized MDA8 ozone from April to October during 2015 to 2019 in BTH.
Figure S6. Examples of deseasonalized ozone predicted by RR-2D (a) and RFR-2D (b) in comparison with deseasonalized observations in PRD during April to October of 2015. For the low anomaly on 2015-Oct-4 (indicated by the black arrow in the figure), RR-2D has a better prediction compared to RFR-2D, which suggests its ability of extrapolation; while the overprediction of high anomaly by RR-2D on 2015-Apr-14 (red arrow) indicates its trade-off for having a risk of over-extrapolation.
Figure S7. Time series of MDA8 ozone anomaly (ppbv) from 2015 to 2019 during April to October in BTH (a), PRD (b), Sichuan (c) and YRD (d). Monthly average ozone for each region is first calculated. Anomalies are calculated by subtracting each monthly average ozone by the corresponding month’s average throughout 2015 to 2019. The black lines represent observational anomalies. The blue lines are the anomalies predicted by ridge regression–2D (RR–2D), which indicate the trends of meteorologically driven ozone anomalies. The red lines are observational ozone anomalies subtracted by meteorologically driven ozone anomalies.

Supplementary Table

Table S1. List of stations that are not used in this study.

| Station ID | Longitude | Latitude | Station ID | Longitude | Latitude |
|------------|-----------|----------|------------|-----------|----------|
| 2702A      | 79.9117   | 37.1013  | 2288A      | 117.7442  | 30.8811  |
| 2701A      | 79.9485   | 37.1152  | 2287A      | 117.7806  | 30.9414  |
| 2697A      | 76.1861   | 39.7153  | 2286A      | 117.8078  | 30.9222  |
| 2696A      | 80.2956   | 41.1933  | 2285A      | 117.8178  | 30.9414  |
| 2695A      | 80.2828   | 41.1636  | 2271A      | 117.3605  | 32.9427  |
| 2694A      | 82.0806   | 44.8969  | 2260A      | 131.0032  | 45.7677  |
| 2691A      | 87.2997   | 44.0114  | 2248A      | 131.1638  | 46.6572  |
| 2679A      | 106.196   | 37.9723  | 2223A      | 120.3939  | 41.615   |
| 2678A      | 106.2025  | 37.9844  | 2181A      | 112.7383  | 38.4519  |
| 2677A | 106.1532 | 37.9648 | 2175A | 110.9956 | 35.0147 |
| 2676A | 97.3731  | 37.3753 | 2174A | 112.7105 | 37.7087 |
| 2674A | 100.2561 | 34.4714 | 2173A | 112.7306 | 37.7111 |
| 2655A | 100.4497 | 38.9389 | 2168A | 112.4254 | 39.3179 |
| 2654A | 100.4686 | 38.9467 | 2167A | 112.4549 | 39.3606 |
| 2648A | 104.1731 | 36.5481 | 2166A | 112.44  | 39.3514 |
| 2630A | 80.1161  | 32.5   | 2161A | 112.835  | 35.4934 |
| 2624A | 91.7608  | 29.2313 | 2054A | 114.1044 | 32.1078 |
| 2618A | 98.8601  | 25.8567 | 2010A | 112.44   | 39.3514 |
| 2604A | 100.0782 | 23.8982 | 1998A | 119.18   | 31.9555 |
| 2601A | 100.98   | 22.7633 | 1997A | 119.146  | 31.9555 |
| 2599A | 100.2497 | 26.8802 | 1983A | 121.531  | 36.913  |
| 2590A | 104.8811 | 25.0992 | 1981A | 122.038  | 37.197  |
| 2580A | 104.9544 | 26.5506 | 1958A | 86.2381  | 41.7128 |
| 2556A | 103.0001 | 29.9816 | 1957A | 86.2022  | 41.7192 |
| 2555A | 103.0109 | 29.9834 | 1953A | 85.1186  | 45.6886 |
| 2554A | 103.0013 | 29.9899 | 1952A | 84.8897  | 45.5828 |
| 2521A | 107.3476 | 22.4137 | 1951A | 84.8661  | 45.6033 |
| 2520A | 109.2317 | 23.7369 | 1945A | 98.2908  | 39.7711 |
| 2518A | 108.1009 | 24.6967 | 1943A | 102.1725 | 38.5339 |
| 2515A | 111.5622 | 24.4072 | 1801A | 118.48   | 31.6928 |
| 2498A | 111.26   | 23.4794 | 1796A | 118.3667 | 31.3139 |
| 2497A | 111.2353 | 23.415  | 1795A | 118.37   | 31.4189 |
| 2496A | 111.3178 | 23.475  | 1794A | 118.3528 | 31.3508 |
| 2495A | 111.2897 | 23.4792 | 1780A | 123.9305 | 47.3386 |
| 2491A | 111.9892 | 27.7044 | 1369A | 113.628  | 22.4251 |
| 2409A | 115.6558 | 34.429  | 1362A | 113.891  | 22.5794 |
| 2405A | 112.5003 | 32.9735 | 1279A | 117.124  | 31.8516 |
| 2400A | 114.005  | 33.568  | 1278A | 117.278  | 31.7386 |
| 2374A | 114.3703 | 27.7914 | 1277A | 117.302  | 31.7956 |
| 2369A | 114.99   | 27.1311 | 1276A | 117.336  | 31.8585 |
| 2318A | 118.7175 | 30.9431 | 1275A | 117.266  | 31.9438 |
| 2317A | 118.7386 | 30.9742 | 1274A | 117.25   | 31.8572 |
| 2316A | 118.7581 | 30.9447 | 1269A | 119.879  | 28.4231 |
| 2311A | 115.8067 | 33.8399 | 1164A | 120.613  | 31.2703 |
| 2310A | 116.5661 | 31.7712 | 1103A | 123.428  | 41.8472 |
| 2308A | 116.478  | 31.7618 | 1090A | 111.651  | 40.7579 |
| 2307A | 116.508  | 31.7371 | 1084A | 112.469  | 37.7124 |
| 2306A | 116.989  | 33.6306 |
|    |      |        |        |
|----|------|--------|--------|
| 2303A | 115.8556 | 32.8603 |
| 2299A | 118.3244 | 32.2786 |
| 2298A | 118.3094 | 32.3153 |
| 2297A | 118.1371 | 30.2756 |
| 2296A | 118.3236 | 29.7207 |
| 2295A | 118.3057 | 29.7128 |
| 2294A | 116.9896 | 30.6145 |
| 2292A | 117.0331 | 30.512  |
| 2291A | 117.0549 | 30.5103 |
| 2290A | 117.8472 | 30.9697 |
| 2289A | 117.8561 | 30.9219 |