

**Special Section:**

Atmospheric PM<sub>2.5</sub> in China: indoor, outdoor, and health effects

**Key Points:**

- We developed accurate and fast emulators to predict air quality in China from emission changes
- Winter ambient fine particulate matter concentrations were primarily sensitive to residential, industrial, and agricultural emissions
- The National Air Quality Target is unlikely to be met in winter without large emission reductions in the residential and industrial sectors

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

L. Conibear,  
[L.A.Conibear@leeds.ac.uk](mailto:L.A.Conibear@leeds.ac.uk)

**Citation:**

Conibear, L., Reddington, C. L., Silver, B. J., Chen, Y., Knote, C., Arnold, S. R., & Spracklen, D. V. (2021). Statistical emulation of winter ambient fine particulate matter concentrations from emission changes in China. *GeoHealth*, 5, e2021GH000391. <https://doi.org/10.1029/2021GH000391>

Received 19 JAN 2021

Accepted 17 APR 2021

**Author Contributions:**

**Conceptualization:** Luke Conibear, Stephen R. Arnold, Dominick V. Spracklen

**Data curation:** Luke Conibear, Ben J. Silver, Ying Chen

**Formal analysis:** Luke Conibear

© 2021. The Authors. *GeoHealth* published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Statistical Emulation of Winter Ambient Fine Particulate Matter Concentrations From Emission Changes in China

Luke Conibear<sup>1</sup> , Carly L. Reddington<sup>1</sup> , Ben J. Silver<sup>1</sup> , Ying Chen<sup>2</sup>, Christoph Knote<sup>3</sup>, Stephen R. Arnold<sup>1</sup> , and Dominick V. Spracklen<sup>1</sup>

<sup>1</sup>Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, UK, <sup>2</sup>College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK, <sup>3</sup>Faculty of Medicine, University of Augsburg, Augsburg, Germany

**Abstract** Air pollution exposure remains a leading public health problem in China. The use of chemical transport models to quantify the impacts of various emission changes on air quality is limited by their large computational demands. Machine learning models can emulate chemical transport models to provide computationally efficient predictions of outputs based on statistical associations with inputs. We developed novel emulators relating emission changes in five key anthropogenic sectors (residential, industry, land transport, agriculture, and power generation) to winter ambient fine particulate matter (PM<sub>2.5</sub>) concentrations across China. The emulators were optimized based on Gaussian process regressors with Matern kernels. The emulators predicted 99.9% of the variance in PM<sub>2.5</sub> concentrations for a given input configuration of emission changes. PM<sub>2.5</sub> concentrations are primarily sensitive to residential (51%–94% of first-order sensitivity index), industrial (7%–31%), and agricultural emissions (0%–24%). Sensitivities of PM<sub>2.5</sub> concentrations to land transport and power generation emissions are all under 5%, except in South West China where land transport emissions contributed 13%. The largest reduction in winter PM<sub>2.5</sub> exposure for changes in the five emission sectors is by 68%–81%, down to 15.3–25.9 μg m<sup>-3</sup>, remaining above the World Health Organization annual guideline of 10 μg m<sup>-3</sup>. The greatest reductions in PM<sub>2.5</sub> exposure are driven by reducing residential and industrial emissions, emphasizing the importance of emission reductions in these key sectors. We show that the annual National Air Quality Target of 35 μg m<sup>-3</sup> is unlikely to be achieved during winter without strong emission reductions from the residential and industrial sectors.

### 1. Introduction

Air pollution exposure is a leading public health problem in China (GBD 2017 Risk Factor Collaborators, 2018; Yin et al., 2020). Despite recent improvements in the air quality of China, air pollution exposure remains high, requiring further emission reductions to improve public health (Silver et al., 2020; Zhao et al., 2018). The impacts of potential emission changes on air quality can be determined through chemical transport model simulations. However, the large computational demands of these models limit the number of simulations and scenarios that can be feasibly investigated.

Machine learning models have been used for many different applications in air quality research (Bellinger et al., 2017; Karpatne et al., 2019; Reichstein et al., 2019; Rybarczyk & Zalakeviciute, 2018; Weichenthal et al., 2018). Specifically, machine learning models have been used as computationally efficient emulators of explicit chemical transport models, and used to explore uncertainties and sensitivities (Aleksankina et al., 2019; Carslaw et al., 2013; Lee et al., 2011, 2012, 2016; Ryan et al., 2018), to replace gas-phase chemistry schemes (Keller & Evans, 2019), to predict aerosol mixing states (Hughes et al., 2018), and to predict air quality in India (Y. Chen et al., 2020) and the United Kingdom (Beddows et al., 2017). Y. Chen et al. (2020) trained machine learning models on simulation data from chemical transport models to predict changes in Indian air quality from emission changes, enabling extensive sensitivity analyses to be undertaken. Other machine learning approaches concerning Chinese air quality have been used to decouple the effects of meteorology and policies (Y. Zhang et al., 2020), to fuse model simulations with ground observations (Lyu et al., 2019), to optimize economic pathways to achieve air quality goals (Huang et al., 2020), and for the prediction of air pollution concentrations (G. Chen et al., 2018; Q. Li et al., 2018; Ma et al., 2019; Wei et al., 2020; Zhan et al., 2017).

**Funding acquisition:** Stephen R. Arnold, Dominick V. Spracklen  
**Investigation:** Luke Conibear  
**Methodology:** Luke Conibear, Ying Chen, Christoph Knote  
**Project Administration:** Stephen R. Arnold, Dominick V. Spracklen  
**Resources:** Stephen R. Arnold, Dominick V. Spracklen  
**Software:** Luke Conibear  
**Validation:** Luke Conibear  
**Visualization:** Luke Conibear  
**Writing – original draft:** Luke Conibear  
**Writing – review & editing:** Luke Conibear, Ben J. Silver, Ying Chen, Christoph Knote, Stephen R. Arnold, Dominick V. Spracklen

Our aim was to develop computationally efficient emulators to predict winter ambient fine particulate matter (PM<sub>2.5</sub>) concentrations from emission changes in China. The emulators were trained on simulated data for January 2015 from chemical transport models, where emissions were varied in the following five key anthropogenic sectors: residential, industry, land transport, agriculture, and power generation. We used these emulators to explore how emission changes impacted winter PM<sub>2.5</sub> exposure. To our knowledge, this is the first study using emulators to predict air quality in China from emission changes.

## 2. Methods

### 2.1. Simulator

Simulations were conducted using the Weather Research and Forecasting model online-coupled with the Chemistry (WRFChem) version 3.7.1 (Grell et al., 2005; Skamarock et al., 2008). We refer to the chemical transport model (WRFChem) as the simulator. Simulations were for January 2015 with 1-month spin-up. The simulator domain covered China at 30 km (0.3°) horizontal resolution. The simulator setup is provided in Table S1 and described fully in our previous work (Reddington et al., 2019; Silver et al., 2020).

Anthropogenic emissions for China were provided by the Multi-resolution Emission Inventory for China (MEIC) emission inventory for 2015 at 0.25° × 0.25° horizontal resolution (M. Li et al., 2017; MEIC Research Group & Tsinghua University, 2019; Zheng et al., 2018). Emissions were speciated for black carbon (BC), organic carbon (OC), PM<sub>2.5</sub>, coarse particulate matter (PM<sub>10</sub>), carbon monoxide (CO), ammonia (NH<sub>3</sub>), nitrogen oxides (NO<sub>x</sub>), sulfur dioxide (SO<sub>2</sub>), and non-methane volatile organic compounds (VOC, Figures S1–S3). Anthropogenic emissions of methane inside China, and all anthropogenic emissions outside of China, were from the Emission Database for Global Atmospheric Research with the Task Force on Hemispheric Transport of Air Pollution (EDGAR-HTAP) version 2.2 for 2010 at 0.1° × 0.1° horizontal resolution (Janssens-Maenhout et al., 2015). Sectoral emissions were provided for land transport, industry, residential energy use, power generation, shipping, aircraft, and agriculture. A diurnal cycle was applied to the anthropogenic emissions (Qi et al., 2017; Zheng et al., 2017).

Biomass burning emissions were from the Fire Inventory from National Center for Atmospheric Research (FINN) version 1.5 with a horizontal resolution of 1 km (Wiedinmyer et al., 2011). Biomass burning emissions were vertically distributed evenly throughout the model boundary layer. Emissions from vegetation were calculated online using the Model of Emissions of Gases and Aerosol from Nature (MEGAN, Guenther et al., 2006). Dust emissions were calculated online using the Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) with the Air Force Weather Agency modifications (Legrand et al., 2019).

Gas phase chemistry was simulated using the extended Model for Ozone and Related Chemical Tracers (MOZART, Emmons et al., 2010; Hodzic & Jimenez, 2011; Knote et al., 2014). Aerosol physics and chemistry was simulated using the updated Model for Simulating Aerosol Interactions and Chemistry (MOSAIC) scheme with aqueous chemistry and the following four sectional discrete size bins: 0.039–0.156, 0.156–0.625, 0.625–2.5, and 2.5–10 μm (Hodzic & Knote, 2014; Zaveri et al., 2008). The secondary organic aerosol formation was based on an updated volatility basis set mechanism (Knote et al., 2015).

Microphysics were simulated using the Morrison two-moment scheme (Morrison et al., 2009). Chemical initial- and boundary-conditions were taken from operational simulations of the MOZART global chemistry transport model driven by the Goddard Earth Observing System Model (GEOS, National Center for Atmospheric Research, 2016). Meteorological initial- and boundary-conditions were taken from the European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA)-Interim global product (Dee et al., 2011), on a N256 (~35 km at the equator) grid at the surface and on a N128 (~70 km at the equator) grid above the surface, and were updated every 6 h. WRF meteorology was nudged to these fields above the boundary layer.

### 2.2. Simulator Evaluation

Simulator evaluation against measurements was conducted in our previous work (Reddington et al., 2019; Silver et al., 2020). Measurement data was taken from over 1,600 sites across China, Hong Kong, and Taiwan as detailed in Silver et al. (2018). The normalized mean bias factor (NMBF) and the normalized mean absolute error factor (NMAEF) were used to evaluate the simulator (Yu et al., 2006). For January 2015, the

simulator slightly overestimated  $PM_{2.5}$  concentrations across China (NMBF = 0.13 and NMAEF = 0.4), and the simulator slightly underestimated  $PM_{2.5}$  concentrations within the Guangdong-Hong Kong-Macau Greater Bay Area (GBA) in South Central China (NMBF = -0.04 and NMAEF = 0.11). Our evaluation shows that the simulator is able to predict the spatial pattern and magnitude of  $PM_{2.5}$  concentrations across China.

### 2.3. Emulator

Machine learning models can predict outputs based on statistical associations with inputs. We refer to the machine learning model developed here as the emulator. The aim of the emulators was to create computationally efficient proxies of the simulator that represent the relationships between changes in anthropogenic emissions and air quality. The emulators make predictions specific to their training data without explanatory knowledge (Deutsch, 2012; Pearl, 2019).

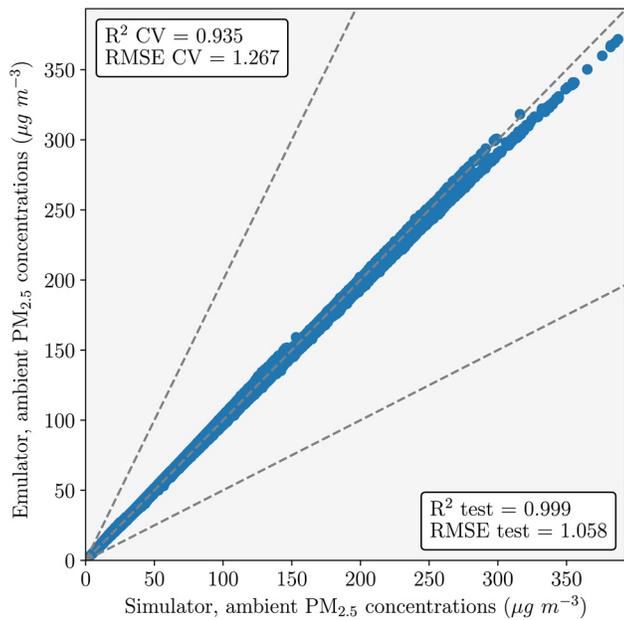
We developed emulators to predict how  $PM_{2.5}$  concentrations change as emissions from the residential (RES), industrial (IND), land transport (TRA), agricultural (AGR), and power generation (ENE) sectors change within mainland China. The emulator inputs were simulator data from 50 training runs and 5 test runs (Tables S2 and S3, respectively). The number of training runs (50) was determined as 10 times the number of inputs (five emission sectors) (Loepky et al., 2009). Both the training and test simulator runs were designed from separate maxi-min Latin hypercube space-filling designs for each of the five inputs over 100,000 iterations (McKay et al., 1979). All anthropogenic species emitted for that sector were scaled by the corresponding factor between 0% and 150% for a given run. A different emulator was developed for each grid cell within China (15,278 grid cells in total) to capture the spatial distribution of the pollutants.

The emulator design was optimized using the Tree-based Pipeline Optimization Tool (TPOT) (Le et al., 2020; Olson et al., 2016; Tran et al., 2016). The TPOT considered a range of emulator designs, evaluated each using 10-fold cross validation, and optimized them for accuracy (Banzhaf et al., 1998; Fortin et al., 2012). Accuracy was measured as the coefficient of determination ( $R^2$ ) and precision was measured as the root mean squared error (RMSE). The emulators focused on a Gaussian process regressor based on previous studies (Aleksankina et al., 2019; Beddows et al., 2017; Bellinger et al., 2017; Carslaw et al., 2013; Y. Chen et al., 2020; Lee et al., 2011, 2012, 2016; Rasmussen & Williams, 2006; Ryan et al., 2018; Rybarczyk & Zalakeviciute, 2018; Wild et al., 2020). The TPOT was used on 50 grid cells within China, selected using a reproducible random seed. The optimized emulator had a 10-fold cross-validation  $R^2$  value of 0.9993. The optimized emulator design included input preprocessors (Yeo & Johnson, 2000), output preprocessors (zero-mean and unit variance), and a Gaussian process regressor with a Matern 5/2 kernel (Conibear, 2020).

The emulators were used to predict output concentrations for configurations across the five emission sectors covering all permutations across a 0%–150% matrix of emission scaling factors at 10% increments. Exposures were estimated as population-weighted concentrations using population count data for 2015 at  $0.25^\circ \times 0.25^\circ$  resolution obtained from the Gridded Population of the World, Version 4.11 (Center for International Earth Science Information Network & NASA Socioeconomic Data and Applications Center, 2018). National and regional analyses were conducted according to the following groupings (Figure S4): North China (Beijing, Tianjin, Hebei, Shanxi, and Inner Mongolia), North East China (Liaoning, Jilin, and Heilongjiang), East China (Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, and Shandong), South Central China (Henan, Hubei, Hunan, Guangdong, Guangxi, Hainan, Hong Kong, and Macau) including the GBA, South West China (Chongqing, Sichuan, Guizhou, Yunnan, and Tibet), and North West China (Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang), and the GBA individually.

### 2.4. Emulator Evaluation

The emulators for all grid cells were evaluated on the training data using  $k$ -fold cross validation (where  $k = 10$ ) and separately on the held-out test data (Figure 1).  $K$ -fold cross validation randomly splits the training data into  $k$  smaller sets, trains the emulators on each of the  $k-1$  smaller data sets, and then validates the emulators on the final held-out data. The performance measures were then averaged across all the  $k$ -fold cross validations.



**Figure 1.** Emulator evaluation for ambient fine particulate matter (PM<sub>2.5</sub>) concentrations on the test (generalization) data. Evaluation metrics used were the coefficient of determination ( $R^2$ ) and the root mean squared error (RMSE). Evaluation metrics given for both the test and training (cross-validation, CV) data.

Across all emulators, the mean 10-fold cross validation  $R^2$  values was 0.935 and the mean RMSE value was  $1.267 \mu\text{g m}^{-3}$ . For the test data, the  $R^2$  value was 0.999 and the RMSE value was  $1.058 \mu\text{g m}^{-3}$ . This means that the emulator generalized well to new data and predicted 99.9% of the variance in PM<sub>2.5</sub> concentrations for a given emission configuration.

### 2.5. Sensitivity Analysis

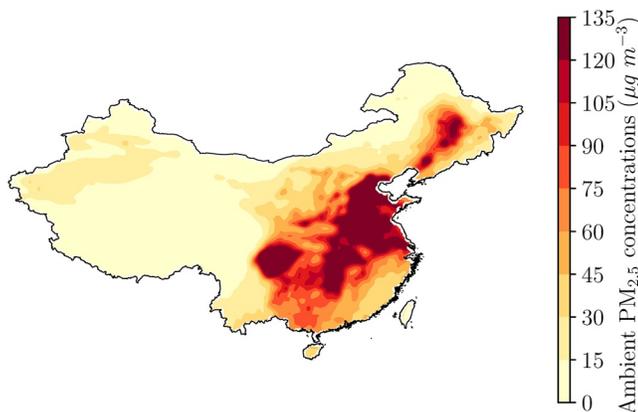
Global sensitivity analyses were performed on the emulators using a Saltelli sampler and a Sobol analyzer (Iooss & Lemaitre, 2015; Kennedy & O’Hagan, 2000; Saltelli et al., 2010; Sobol, 2001). First-order, second-order, and total sensitivity indices were calculated for each input and output. First-order sensitivity indices measure the contribution to the output variance by a single input individually. Absolute first-order sensitivity indices are the first-order sensitivity indices multiplied by the baseline PM<sub>2.5</sub> concentrations. Second-order sensitivity indices measure the contribution to the output variance caused by the interaction of two inputs. Total sensitivity indices measure the contribution to the output variance caused by an input, including both first-order effects and all higher-order interactions. The sensitivity indices for all emission sectors sum to 1 for each grid cell and were shown as percentages. The sensitivity indices were estimated based on 12,000 emulator runs per grid cell, decided upon through Equation 1, where  $N$  was 1,000 and  $D$  was the number of inputs (Kennedy & O’Hagan, 2000; Saltelli et al., 2010).

$$\text{Emulator runs} = N \times (2D + 2) \tag{1}$$

## 3. Results and Discussion

### 3.1. Baseline Ambient PM<sub>2.5</sub> Concentrations

The emulator results and discussion are monthly-means for January 2015. PM<sub>2.5</sub> concentrations from the emulator baseline (i.e., all emission sectors at 100%) are shown in Figure 2 and average PM<sub>2.5</sub> exposures are given in Table 1. PM<sub>2.5</sub> exposures are above  $100 \mu\text{g m}^{-3}$  in China overall, driven by high exposure over North, South Central, South West, and East China, with lower exposure in the GBA, North West China, and North East China.



**Figure 2.** Emulator predictions of monthly-mean (January 2015) ambient fine particulate matter (PM<sub>2.5</sub>) concentrations for the baseline scenario with all emission sectors at 100% across China.

### 3.2. Key Sensitivities of Ambient PM<sub>2.5</sub> Concentrations to Emission Changes

Table 2 shows the first-order sensitivities for PM<sub>2.5</sub> concentrations for each of the five emission sector inputs. Across China, PM<sub>2.5</sub> concentrations are most sensitive to residential emissions (64%), then approximately equally sensitive to industrial (16%) and agricultural emissions (14%). In all regions, PM<sub>2.5</sub> concentrations are primarily sensitive to residential emissions (51%–94%). North West China has approximately equal sensitivity to industrial (23%) and agricultural emissions (22%). In South West China, agricultural emissions (24%) dominate over land transport (13%) and industrial emissions (7%). In the GBA, North China, North East China, East China, and South Central China industrial emissions (5%–31%) dominate over agricultural emissions (0%–11%). In all regions except in South West China, sensitivities of PM<sub>2.5</sub> concentrations to land transport emissions are under 5%. Sensitivities of PM<sub>2.5</sub> concentrations to power

**Table 1**  
*Emulator Baseline Monthly-Mean (January 2015) Ambient Fine Particulate Matter (PM<sub>2.5</sub>) Exposure*

Baseline	China	GBA	North China	North East China	East China	South Central China	South West China	North West China
Ambient PM <sub>2.5</sub> exposure (μg m <sup>-3</sup> )	104.4	54.8	122.2	80.8	104.1	111.5	116.6	50.2

generation emissions are all under 3%. Second-order sensitivities, which estimate sensitivities from input interactions, are all less than 1%.

Figure 3 shows the absolute first-order sensitivities for PM<sub>2.5</sub> concentrations, which are the first-order sensitivities multiplied by the baseline PM<sub>2.5</sub> concentrations. The absolute first-order sensitivities are largest for residential and industrial emissions. The first-order sensitivities, which do not account for the magnitude of PM<sub>2.5</sub> concentrations, are higher for land transport and agricultural emissions in South West China (Figure S5) where PM<sub>2.5</sub> concentrations are relatively low. All anthropogenic emission sources are lower in South West China compared to other regions, though there are large emissions from the land transport and agricultural sectors in the neighboring regions (Figures S2 and S3) explaining the higher sensitivities.

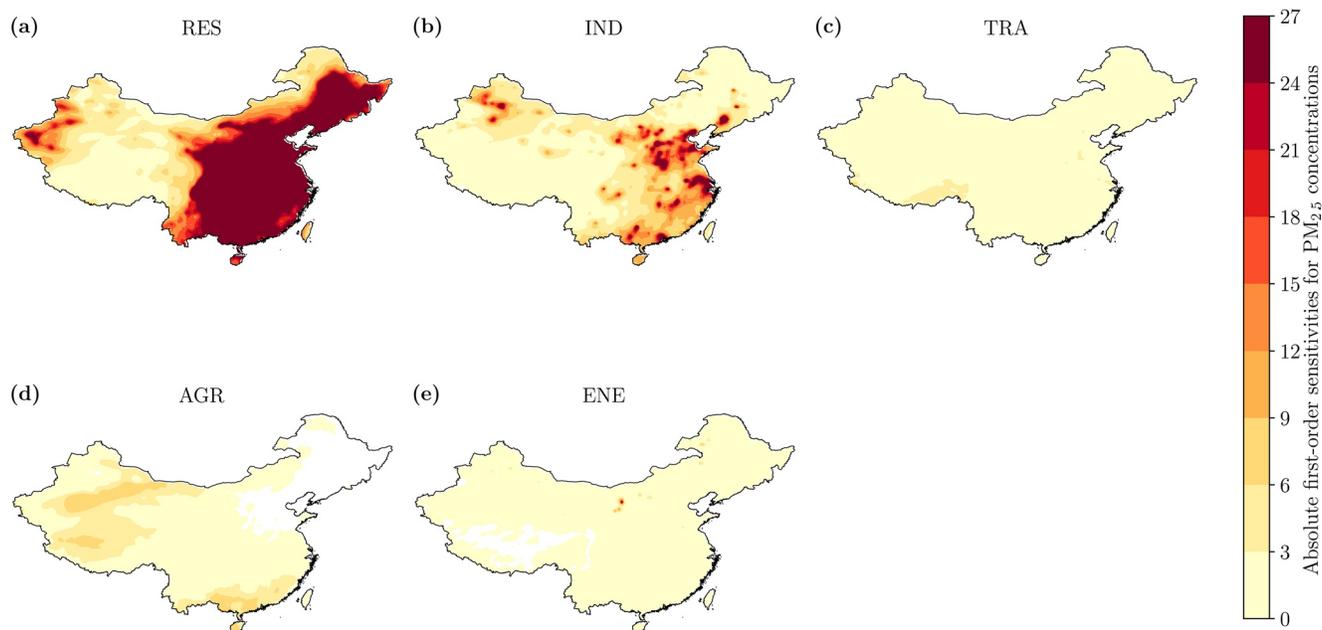
### 3.3. Impact of Changes in Individual Emission Sectors on Ambient PM<sub>2.5</sub> Exposure

The impact of individual changes in emissions, whilst holding other emissions sectors constant at the baseline, on PM<sub>2.5</sub> exposure are shown in Figure 4. PM<sub>2.5</sub> exposure can decrease by 31%–64%, down to 26.7–50.3 μg m<sup>-3</sup>. The National Air Quality Target of 35 μg m<sup>-3</sup> can only be achieved in North West and North East China by reductions in one sector alone. For all regions, reductions in residential emissions produced the largest reduction in PM<sub>2.5</sub> exposure, followed by industrial emissions. Across China, a 30% reduction in residential emissions reduces PM<sub>2.5</sub> exposure by 18.8 μg m<sup>-3</sup> (18%), compared to a 6.3 μg m<sup>-3</sup> (6%) reduction for a 30% reduction in industrial emissions. Individual changes in land transport, agricultural, and power generation emissions produce smaller impacts on PM<sub>2.5</sub> exposure. In all regions except the GBA, reductions in land transport emissions produce a larger reduction in PM<sub>2.5</sub> exposure compared to reductions in agricultural emissions. In the GBA, reductions in agricultural emissions produce a stronger reduction in PM<sub>2.5</sub> exposure than reductions in land transport emissions. The response in PM<sub>2.5</sub> exposure from emission changes is approximately linear for all emission sectors.

At the annual timescale, industrial (30%) and residential (26%) emissions dominate the contribution to PM<sub>2.5</sub> concentrations across China, while there are smaller contributions from agricultural (16%), power generation (14%), and land transport (7%) emissions (Reddington et al., 2019). In January only, the contribution from residential emissions increases to 57% of PM<sub>2.5</sub> concentrations across China, the contributions from industrial (21%) and land transport (11%) emissions remain similar, and the contributions from agricultural and power generation emissions decrease to 3% and 5%, respectively. Even in the highly urbanized GBA, the contribution from land transport emissions to January PM<sub>2.5</sub> concentrations is 9% and from agricultural emissions is 17%. These findings suggest that land transport and power generation emissions are not dominant sources of PM<sub>2.5</sub> concentrations in China. The relatively small contributions of land transport and power generation emissions to regional PM<sub>2.5</sub> concentrations across China has been found in other

**Table 2**  
*The First-Order Sensitivities (%) for Ambient Fine Particulate Matter (PM<sub>2.5</sub>) Concentrations From the Emulator in January 2015 Per Region and Emission Sector of Residential (RES), Industry (IND), Land Transport (TRA), Agriculture (AGR), and Power Generation (ENE)*

		China	GBA	North China	North East China	East China	South Central China	South West China	North West China
First order (%)	RES	64	56	73	94	77	81	52	51
	IND	16	31	20	5	18	13	7	23
	TRA	4	2	1	1	1	1	13	3
	AGR	14	11	4	0	3	4	24	22
	ENE	1	0	3	1	0	0	1	1



**Figure 3.** Absolute first-order sensitivities of monthly-mean (January 2015) ambient fine particulate matter (PM<sub>2.5</sub>) concentrations across China (first-order sensitivities multiplied by the baseline PM<sub>2.5</sub> concentrations). The five key emission sectors are (a) residential (RES), (b) industry (IND), (c) land transport (TRA), (d) agriculture (AGR), and (e) power generation (ENE).

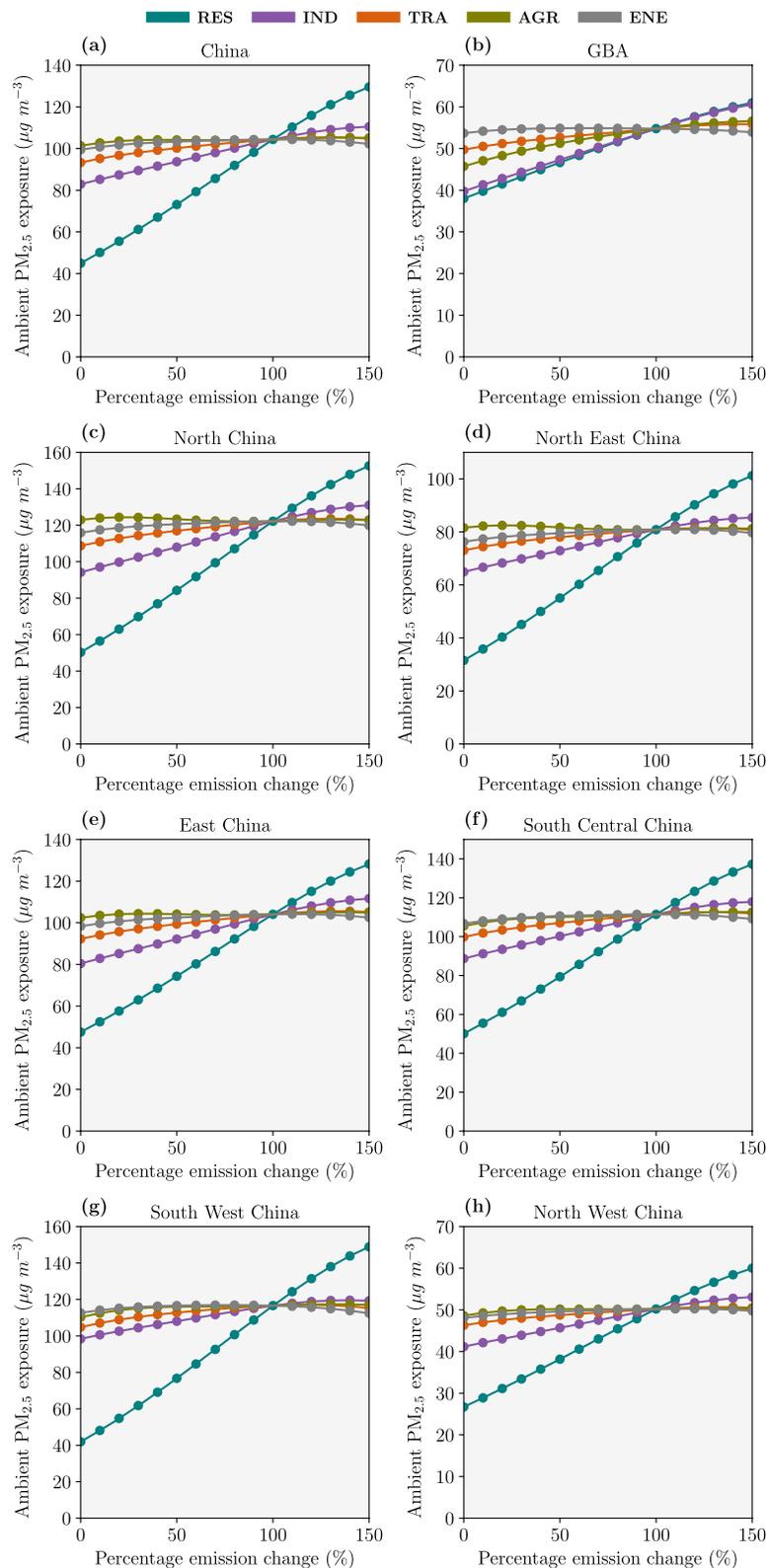
modeling studies (GBD MAPS Working Group, 2016; Gu et al., 2018; Hu et al., 2017; Karagulian et al., 2017; Reddington et al., 2019; Shi et al., 2017; Silva et al., 2016) and was further confirmed during the Coronavirus Disease 2019 (COVID-19) lockdown, when the substantial decreases in land transport and power generation activity only reduced PM<sub>2.5</sub> concentrations by 10% across China (Silver et al., 2020).

### 3.4. Impact of Changes in Multiple Emission Sectors on Ambient PM<sub>2.5</sub> Exposure

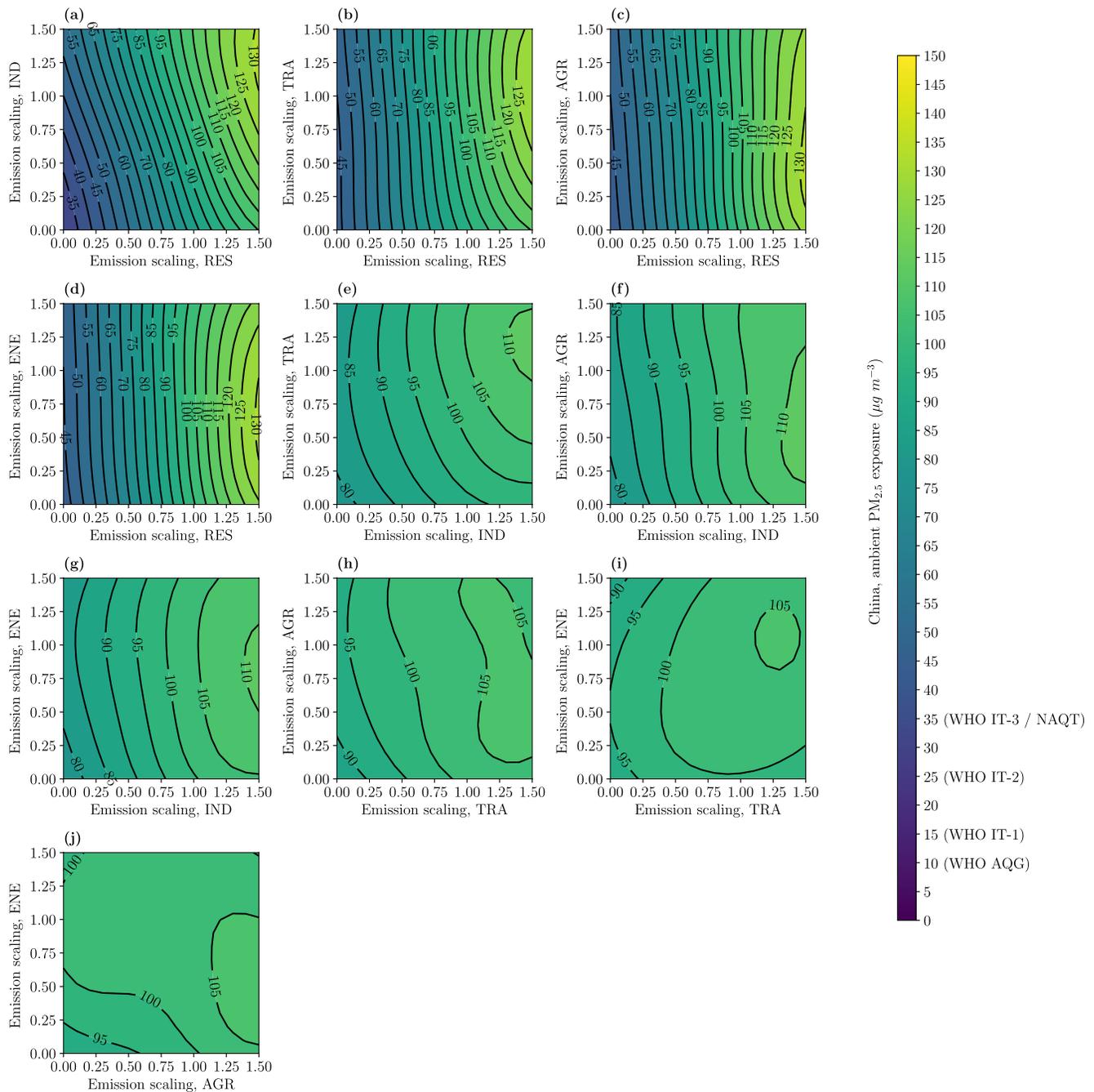
The combined impacts of variations in two emission sectors on PM<sub>2.5</sub> exposure for China are shown in Figure 5, and the impacts for specific regions are shown in Figures S6–S12. The greatest reductions in PM<sub>2.5</sub> exposure are achieved by reducing residential and industrial emissions, with PM<sub>2.5</sub> exposure going below the National Air Quality Target of 35  $\mu\text{g m}^{-3}$ . The combined impacts of reducing residential and industrial emissions on reducing PM<sub>2.5</sub> exposures are dominated by the reductions in residential emissions for most regions across China (Figure 5), except for the GBA (Figure S6) where there are approximately equal contributions from residential and industrial emissions. Without reducing residential and industrial emissions, reductions in PM<sub>2.5</sub> exposure across China are limited, with exposure remaining above 90  $\mu\text{g m}^{-3}$  (Figure 5). The exception was for the GBA, where sizable reductions in PM<sub>2.5</sub> exposure are also obtained from reducing land transport and agricultural emissions together.

Reductions in anthropogenic emissions from the five sectors reduce PM<sub>2.5</sub> exposure by 68%–81%, down to 15.3–25.9  $\mu\text{g m}^{-3}$ . Even when anthropogenic emissions in China from the residential, industrial, land transport, agricultural, and power generation sectors are completely removed, PM<sub>2.5</sub> exposure remains above the World Health Organization annual guideline of 10  $\mu\text{g m}^{-3}$ . In these scenarios, emissions remain from the shipping and aviation sectors, from outside of China, outdoor biomass burning (agricultural and forest fires), and from natural sources such as dust and biogenics. A multi-model estimate suggests that outdoor biomass burning contributes 4% of annual-mean PM<sub>2.5</sub> exposure across China (Reddington et al., 2019). This demonstrates the challenges of reducing winter PM<sub>2.5</sub> exposure, even under strong national emission reductions. Future work is required to explore the contribution of these different sources.

Our results highlight the importance of reducing residential and industrial emissions in order to reduce winter PM<sub>2.5</sub> exposure across China, confirming the suggestions of previous studies using different approaches (GBD MAPS Working Group, 2016; Lelieveld et al., 2015; Reddington et al., 2019). The 2018–2020



**Figure 4.** The impact of individual changes in emissions per sector on the monthly-mean (January 2015) ambient fine particulate matter (PM<sub>2.5</sub>) exposure for (a) China, (b) the Guangdong-Hong Kong-Macau Greater Bay Area (GBA), (c) North China, (d), North East China, (e) East China, (f) South Central China, (g) South West China, and (h) North West China for each of the five key emission sectors: residential (RES), industry (IND), land transport (TRA), agriculture (AGR), and power generation (ENE).



**Figure 5.** The combined impact of variations in two emission sectors on ambient fine particulate matter (PM<sub>2.5</sub>) exposure for China from (a) residential (RES) and industry (IND), (b) RES and land transport (TRA), (c) RES and agriculture (AGR), (d) RES and power generation (ENE), (e) IND and TRA, (f) IND and AGR, (g) IND and ENE, (h) TRA and AGR, (i) TRA and ENE, and (j) AGR and ENE emissions. Air quality targets shown for the World Health Organization's (WHO) Air Quality Guideline (AQG), Interim Target 1 (IT-1), Interim Target 2 (IT-2), Interim Target 3 (IT-3), and China's National Air Quality Target (NAQT).

3-year plan introduced specific policies for the residential sector in North China in winter to achieve 70% clean heating by 2021 under the Clean Heating Plan (Ministry of Environmental Protection of China, 2017; National Development and Reform Commission of China, 2017). This policy aims to reduce solid fuel use for heating and has the potential to improve air quality and public health in North China (Liu et al., 2019; Meng et al., 2019, 2020; Qin et al., 2017; Zhao et al., 2018). To date there are no specific policies for tackling residential cooking and heating from solid fuels in South China, highlighting a key policy gap that could provide major air quality improvements.

The emulators were trained on simulator data from January 2015. We focused on this period due to the high  $PM_{2.5}$  concentrations experienced in winter across China. As the emulators were trained on data from one particular year, they do not account for interannual variability in meteorology, which can impact  $PM_{2.5}$  concentrations (Hou et al., 2019; Zhai et al., 2019; Q. Zhang et al., 2019). In future work, we will extend these emulators for long-term air pollution exposure and the associated chronic disease burden. Here, our focus is on the sensitivity of air pollution to changes in emissions. Different emulators could be developed that explore the sensitivity of air pollution to meteorological variables.

#### 4. Conclusion

Air pollution exposure is a leading public health problem in China. Chemical transport models are often used to quantify the impacts of emission changes on air quality. However, the number of sensitivity analyses that can be explored is limited by their large computational demands. Machine learning models can emulate chemical transport models to provide computationally efficient predictions of air quality based on statistical associations with emission changes. Our aim was to develop novel emulators to predict winter ambient  $PM_{2.5}$  concentrations from emission changes in five key anthropogenic sectors (residential, industry, land transport, agriculture, and power generation) in China. We used these emulators to explore how  $PM_{2.5}$  exposure in January 2015 varied as emissions from the different sectors were varied.

The emulators were based on Gaussian process regressors with Matern kernels. The emulators were optimized to apply Yeo-Johnson parametric transformations on the inputs, and zero-mean and unit variance to the transformed outputs. The emulators predicted 99.9% of the variance in  $PM_{2.5}$  concentrations for a given input configuration of emission changes. Emulators were developed for each grid cell across China (15,278 grid cells in total) to capture the spatial distribution of  $PM_{2.5}$  concentrations.

Global sensitivity analyses were performed on the emulators using a Saltelli sampler and a Sobol analyzer. First-order sensitivity indices were calculated to measure the contribution that each emission sector makes to the variance within  $PM_{2.5}$  concentrations.  $PM_{2.5}$  concentrations are primarily sensitive to residential (51%–94% of first-order sensitivity index), industrial (7%–31%), and agricultural emissions (0%–24%). Sensitivities of  $PM_{2.5}$  concentrations to land transport emissions are mostly under 5%, except in South West China where they are 13%. Sensitivities of  $PM_{2.5}$  concentrations to power generation emissions are all under 3%.

The emulators were used to predict  $PM_{2.5}$  concentrations for every emission configuration of the five anthropogenic emission sectors within a 0%–150% matrix (100% representing January 2015 emissions) at 10% increments. The largest reduction in  $PM_{2.5}$  exposure is by 68%–81%, down to 15.3–25.9  $\mu\text{g m}^{-3}$  across China. Even under these emission reduction scenarios,  $PM_{2.5}$  exposure exceeds the World Health Organization annual guideline of 10  $\mu\text{g m}^{-3}$ . The greatest reductions in  $PM_{2.5}$  exposure are all driven by reducing residential and industrial emissions. The annual National Air Quality Target of 35  $\mu\text{g m}^{-3}$  is unlikely to be achieved during winter without stringent emission reductions from the residential and industrial sectors. China is implementing ambitious strategies to reduce emissions in the energy generation, industrial, and land transportation sectors, but until recently there has been less focus on the residential sector (Liu et al., 2016, 2019; Meng et al., 2019, 2020; Zhao et al., 2018). Our work provides further evidence that new policies targeting emission reductions in the residential sector are required if  $PM_{2.5}$  exposure across China is to be substantially reduced.

Future work is required to develop these emulators to predict the chronic public health impacts from long-term air pollution exposure from multiple air pollutants in order to determine the optimum emission reduction strategy to provide the largest improvement in public health.

#### Conflict of Interest

The authors declare no conflict of interest relevant to this study.

## Data Availability Statement

Code to setup and run WRFChem (using WRFotron version 2.0) is available through Conibear and Knote (2020). Emulator code and data is available through Conibear (2020). The trained emulators per grid cell in China that support the findings of this study are available at [doi.org/10.5518/953](https://doi.org/10.5518/953).

## Acknowledgments

The authors gratefully acknowledge support from the AIA Group Limited, a European Research Council Consolidator Grant (771492), and the Natural Environment Research Council (NE/S006680/1). This work was undertaken on Advanced Research Computing, part of the High-Performance Computing facilities at the University of Leeds, UK. This work used the WRFotron version 2.0, a tool to automatize WRFChem runs with re-initialized meteorology (Conibear & Knote, 2020). The authors acknowledge the use of the WRFChem preprocessor tools mozbc, fire\_emiss, anthro\_emiss, and bio\_emiss provided by the Atmospheric Chemistry Observations and Modeling (ACOM) of the National Center for Atmospheric Research (NCAR). We acknowledge the use of the Model for Ozone and Related Chemical Tracers (MOZART) global model output available at <http://www.acom.ucar.edu/wrf-chem/mozart.shtml>. The authors acknowledge the use of the emission pre-processor available at [github.com/douglowe/WRF\\_UoM\\_EMIT](https://github.com/douglowe/WRF_UoM_EMIT). The authors thank Qiang Zhang and Meng Li for providing the MEIC data. The authors acknowledge the Python Software Foundation, Python Language Reference, available at [python.org](https://python.org). The authors are particularly grateful to the Python libraries NumPy (Harris et al., 2020), Pandas (McKinney, 2010), Matplotlib (Hunter, 2007), SciPy (Virtanen et al., 2020), xarray (Hoyer & Hamman, 2017), Cartopy (Met Office, 2015), GeoPandas (Jordahl et al., 2020), Salem (Mausson et al., 2019), Jupyter (Kluyver et al., 2016), Scikit-learn (Pedregosa et al., 2011), TPOT (Olson et al., 2016), SALib (Herman & Usher, 2017), pyDOE, Seaborn (Waskom et al., 2020), Rasterio (Gillies, 2013), Affine, xESMF (Zhuang et al., 2020), Dask (Rocklin, 2015), and the Pangeo project (Abernathy et al., 2017). The boundaries shown on any maps in this work do not imply any judgment concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

## References

- Abernathy, R., Paul, K., Hamman, J., Rocklin, M., Lepore, C., Tippet, M., et al. (2017). *Pangeo NSF earthcube proposal*. <https://doi.org/10.6084/m9.figshare.5361094.v1>
- Aleksankina, K., Reis, S., Vieno, M., & Heal, M. R. (2019). Advanced methods for uncertainty assessment and global sensitivity analysis of an Eulerian atmospheric chemistry transport model. *Atmospheric Chemistry and Physics*, 19(5), 2881–2898. <https://doi.org/10.5194/acp-19-2881-2019>
- Banzhaf, W., Nordin, P., Keller, R. E., & Francone, F. D. (1998). *Genetic programming: An Introduction on the automatic evolution of computer programs and its applications*. dpunkt.verlag and Morgan Kaufmann Publishers Inc.
- Beddows, A. V., Kitwiroon, N., Williams, M. L., & Beevers, S. D. (2017). Emulation and sensitivity analysis of the community multiscale air quality model for a UK ozone pollution episode. *Environmental Science & Technology*, 51(11), 6229–6236. <https://doi.org/10.1021/acs.est.6b05873>
- Bellinger, C., Shazan, M., Jabbar, M., Zaïane, O., & Osornio-vargas, A. (2017). A systematic review of data mining and machine learning for air pollution epidemiology. *BMC Public Health*, 17(907), 1–19. <https://doi.org/10.1186/s12889-017-4914-3>
- Carslaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., et al. (2013). Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, 503(7474), 67–71. <https://doi.org/10.1038/nature12674>
- Center for International Earth Science Information Network, & NASA Socioeconomic Data and Applications Center. (2018). *Gridded population of the world, version 4 (GPWv4): Population count, revision 11*. Columbia University. <https://doi.org/10.7927/H4PN93PB>
- Chen, G., Wang, Y., Li, S., Cao, W., Ren, H., Knibbs, L. D., et al. (2018). Spatiotemporal patterns of PM10 concentrations over China during 2005–2016: A satellite-based estimation using the random forests approach. *Environmental Pollution*, 242, 605–613. <https://doi.org/10.1016/j.envpol.2018.07.012>
- Chen, Y., Wild, O., Ryan, E., Sahu, S. K., Lowe, D., Archer-Nicholls, S., et al. (2020). Mitigation of PM2.5 and ozone pollution in Delhi: A sensitivity study during the pre-monsoon period. *Atmospheric Chemistry and Physics*, 20, 499–514. <https://doi.org/10.5194/acp-20-499-2020>
- Conibear, L. (2020). *lukeconibear/emulator: Emulator for WRF-Chem (Version v0.1-alpha)*. Zenodo. <https://doi.org/10.5281/zenodo.3813461>
- Conibear, L., & Knote, C. (2020). *wrfchem-leeds/WRFotron: WRFotron 2.0*. Zenodo. <https://doi.org/10.5281/zenodo.3624087>
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. <https://doi.org/10.1002/qj.828>
- Deutsch, D. (2012). *The beginning of infinity: Explanations that transform the world*. Penguin.
- Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., et al. (2010). Description and evaluation of the model for ozone and related chemical tracers, version 4 (MOZART-4). *Geoscientific Model Development*, 3, 43–67. <https://doi.org/10.5194/gmd-3-43-2010>
- Fortin, F. A., De Rainville, F. M., Gardner, M. A., Parizeau, M., & Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research*, 13, 2171–2175.
- GBD 2017 Risk Factor Collaborators. (2018). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study. *The Lancet*, 392, 1923–1994. [https://doi.org/10.1016/S0140-6736\(18\)32225-6](https://doi.org/10.1016/S0140-6736(18)32225-6)
- GBD MAPS Working Group. (2016). *Burden of disease attributable to coal-burning and other air pollution sources in China* (Special Report 20). Health Effects Institute.
- Gillies, S. (2013). *Rasterio: Geospatial raster I/O for Python programmers*. <https://doi.org/github.com/mapbox/rasterio>.
- Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled “online” chemistry within the WRF model. *Atmospheric Environment*, 39(37), 6957–6975. <https://doi.org/10.1016/j.atmosenv.2005.04.027>
- Gu, Y., Wong, T. W., Law, C. K., Dong, G. H., Ho, K. F., Yang, Y., et al. (2018). Impacts of sectoral emissions in China and the implications: Air quality, public health, crop production, and economic costs. *Environmental Research Letters*, 13(8), 084008. <https://doi.org/10.1088/1748-9326/aad138>
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., & Geron, C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmospheric Chemistry and Physics*, 6(11), 3181–3210. <https://doi.org/10.5194/acp-6-3181-2006>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature*, 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Herman, J., & Usher, W. (2017). SALib: An open-source Python library for sensitivity analysis. *Journal of Open Source Software*, 2(9). <https://doi.org/10.21105/joss.00097>
- Hodzic, A., & Jimenez, J. L. (2011). Modeling anthropogenically controlled secondary organic aerosols in a megacity: A simplified framework for global and climate models. *Geoscientific Model Development*, 4(4), 901–917. <https://doi.org/10.5194/gmd-4-901-2011>
- Hodzic, A., & Knote, C. (2014). WRF-Chem 3.6.1: MOZART gas-phase chemistry with MOSAIC aerosols (Vol. 7). Atmospheric Chemistry Division (ACD), National Center for Atmospheric Research (NCAR).
- Hou, X., Zhu, B., Kumar, K. R., & Lu, W. (2019). Inter-annual variability in fine particulate matter pollution over China during 2013–2018: Role of meteorology. *Atmospheric Environment*, 214, 116842. <https://doi.org/10.1016/j.atmosenv.2019.116842>
- Hoyer, S., & Hamman, J. J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5, 1–6. <https://doi.org/10.5334/jors.148>
- Hu, J., Huang, L., Chen, M., Liao, H., Zhang, H., Wang, S., et al. (2017). Premature mortality attributable to particulate matter in China: Source contributions and responses to reductions. *Environmental Science & Technology*. <https://doi.org/10.1021/acs.est.7b03193>

- Huang, J., Zhu, Y., Kelly, J. T., Jang, C., Wang, S., Xing, J., et al. (2020). Large-scale optimization of multi-pollutant control strategies in the Pearl River Delta region of China using a genetic algorithm in machine learning. *Science of the Total Environment*, 722, 137701. <https://doi.org/10.1016/j.scitotenv.2020.137701>
- Hughes, M., Kodros, J. K., Pierce, J. R., West, M., & Riemer, N. (2018). Machine learning to predict the global distribution of aerosol mixing state metrics. *Atmosphere*, 9(15), 1–18. <https://doi.org/10.3390/atmos9010015>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9, 90–95. <https://doi.org/10.1109/mcse.2007.55>
- Iooss, B., & Lemaitre, P. (2015). A review on global sensitivity analysis methods. In *Uncertainty management in simulation optimization of complex systems: Algorithms and applications*. [https://doi.org/10.1007/978-1-4899-7547-8\\_5](https://doi.org/10.1007/978-1-4899-7547-8_5)
- Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., et al. (2015). HTAP\_v2.2: A mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution. *Atmospheric Chemistry and Physics*, 15(19), 11411–11432. <https://doi.org/10.5194/acp-15-11411-2015>
- Jordahl, K., Bossche, J. V. den, Wasserman, J., McBride, J., Fleischmann, M., & Gerard, J. (2020). *geopandas/geopandas: v0.7.0 (version v0.7.0)*. Zenodo. <https://doi.org/10.5281/zenodo.3669853>
- Karagulian, F., Dingenen, R. V., Belis, C. A., Janssens-Maenhout, G., Crippa, M., Guizzardi, D., et al. (2017). Attribution of anthropogenic PM<sub>2.5</sub> to emission sources. A global analysis of source-receptor model results and measured source-apportionment data. *European Commission JRC Technical Reports*, 49, 104676. <https://doi.org/10.2760/344371>
- Karpatne, A., Ebert-Uphoff, I., Ravela, S., Bubaie, H. A., & Kumar, V. (2019). Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8), 1544–1554. <https://doi.org/10.1109/tkde.2018.2861006>
- Keller, C. A., & Evans, M. J. (2019). Application of random forest regression to the calculation of gas-phase chemistry within the GE-OS-Chem chemistry model v10. *Geoscientific Model Development*, 12, 1209–1225. <https://doi.org/10.5194/gmd-12-1209-2019>
- Kennedy, M., & O'Hagan, A. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1), 1–13. <https://doi.org/10.1093/biomet/87.1.1>
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., et al. (2016). *Jupyter notebooks—A publishing format for reproducible computational workflows*. <https://doi.org/10.3233/978-1-61499-649-1-87>
- Knote, C., Hodzic, A., & Jimenez, J. L. (2015). The effect of dry and wet deposition of condensable vapors on secondary organic aerosols concentrations over the continental US. *Atmospheric Chemistry and Physics*, 15(1), 1–18. <https://doi.org/10.5194/acp-15-1-2015>
- Knote, C., Hodzic, A., Jimenez, J. L., Volkamer, R., Orlando, J. J., Baidar, S., et al. (2014). Simulation of semi-explicit mechanisms of SOA formation from glyoxal in aerosol in a 3-D model. *Atmospheric Chemistry and Physics*, 14(12), 6213–6239. <https://doi.org/10.5194/acp-14-6213-2014>
- Le, T. T., Fu, W., & Moore, J. H. (2020). Scaling tree-based automated machine learning to biomedical big data with a feature set selector. *Bioinformatics*, 36(1), 250–256. <https://doi.org/10.1093/bioinformatics/btz470>
- Lee, L. A., Carslaw, K. S., Pringle, K. J., & Mann, G. W. (2012). Mapping the uncertainty in global CCN using emulation. *Atmospheric Chemistry and Physics*, 12(20), 9739–9751. <https://doi.org/10.5194/acp-12-9739-2012>
- Lee, L. A., Carslaw, K. S., Pringle, K. J., Mann, G. W., & Spracklen, D. V. (2011). Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters. *Atmospheric Chemistry and Physics*, 11, 12253–12273. <https://doi.org/10.5194/acp-11-12253-2011>
- Lee, L. A., Reddington, C. L., & Carslaw, K. S. (2016). On the relationship between aerosol model uncertainty and radiative forcing uncertainty. *Proceedings of the National Academy of Sciences of the United States of America*, 113(21), 5820–5827. <https://doi.org/10.1073/pnas.1507050113>
- Legrand, S. L., Polashenski, C., Letcher, T. W., Creighton, G. A., Peckham, S. E., & Cetola, J. D. (2019). The AFWA dust emission scheme for the GOCART aerosol model in WRF-Chem v3.8.1. *Geoscientific Model Development*, 12, 131–166. <https://doi.org/10.5194/gmd-12-131-2019>
- Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525(7569), 367–371. <https://doi.org/10.1038/nature15371>
- Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., et al. (2017). Anthropogenic emission inventories in China: A review. *National Science Review*, 4(6), 834–866. <https://doi.org/10.1093/nsr/nwx150>
- Li, Q., Wang, Y.-y., Guo, Y., Zhou, H., Wang, X., Wang, Q., et al. (2018). Effect of airborne particulate matter of 2.5 μm or less on preterm birth: A national birth cohort study in China. *Environment International*, 121, 1128–1136. <https://doi.org/10.1016/j.envint.2018.10.025>
- Liu, J., Kieseewetter, G., Klimont, Z., Cofala, J., Heyes, C., Schöpp, W., et al. (2019). Mitigation pathways of air pollution from residential emissions in the Beijing-Tianjin-Hebei region in China. *Environment International*, 125, 236–244. <https://doi.org/10.1016/j.envint.2018.09.059>
- Liu, J., Mauzerall, D. L., Chen, Q., Zhang, Q., Song, Y., Peng, W., et al. (2016). Air pollutant emissions from Chinese households: A major and underappreciated ambient pollution source. *Proceedings of the National Academy of Sciences of the United States of America*, 113(28), 7756–7761. <https://doi.org/10.1073/pnas.1604537113>
- Loeppky, J. L., Sacks, J., & Welch, W. J. (2009). Choosing the sample size of a computer experiment: A practical guide. *Technometrics*, 51(4), 366–376. <https://doi.org/10.1198/tech.2009.08040>
- Lyu, B., Hu, Y., Zhang, W., Du, Y., Luo, B., Sun, X., et al. (2019). Fusion method combining ground-level observations with chemical transport model predictions using an ensemble deep learning framework: Application in China to estimate spatiotemporally-resolved PM<sub>2.5</sub> exposure fields in 2014–2017. *Environmental Science & Technology*, 53(13), 7306–7315. <https://doi.org/10.1021/acs.est.9b01117>
- Ma, J., Cheng, J. C. P., Lin, C., Tan, Y., & Zhang, J. (2019). Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques. *Atmospheric Environment*, 214, 116885. <https://doi.org/10.1016/j.atmosenv.2019.116885>
- Maussion, F., TimoRoth, T., Dusch, M., & Landmann, J. (2019). *fmaussion/salem: v0.2.4 (Version v0.2.4)*. Zenodo. <https://doi.org/10.5281/zenodo.2605265>
- McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2), 239–245. <https://doi.org/10.1080/00401706.1979.10489755>
- McKinney, W. (2010). *Data structures for statistical computing in python* (pp. 51–56). Proceedings of the 9th Python in Science Conference.
- MEIC Research Group, & Tsinghua University. (2019). *Multi-resolution emission inventory for China (MEIC) version 1.3*. Retrieved from <http://www.meicmodel.org/>
- Meng, W., Shen, H., Yun, X., Chen, Y., Zhong, Q., Zhang, W., et al. (2020). Differentiated-rate clean heating strategy with superior environmental and health benefits in Northern China. *Environmental Science & Technology*, 54, 13458–13466. <https://doi.org/10.1021/acs.est.0c04019>

- Meng, W., Zhong, Q., Chen, Y., Shen, H., Yun, X., Smith, K. R., et al. (2019). Energy and air pollution benefits of household fuel policies in northern China. *Proceedings of the National Academy of Sciences of the United States of America*, 201904182. <https://doi.org/10.1073/pnas.1904182116>
- Met Office. (2015). *Cartopy: A cartographic python library with a matplotlib interface*. <https://doi.org/scitools.org.uk/cartopy>
- Ministry of Environmental Protection of China. (2017). *Beijing-tianjin-hebei and surrounding areas 2017 air pollution prevention and control work plan*. Retrieved from [http://dqhj.mee.gov.cn/dtxx/201703/t20170323\\_408663.shtml](http://dqhj.mee.gov.cn/dtxx/201703/t20170323_408663.shtml)
- Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*, 137(3), 991–1007. <https://doi.org/10.1175/2008mwr2556.1>
- National Center for Atmospheric Research. (2016). *ACOM MOZART-4/GEOS-5 global model output*. UCAR. Retrieved from <http://www.acom.ucar.edu/wrf-chem/mozart.shtml>
- National Development and Reform Commission of China. (2017). *Work plan for clean heating in winter in northern China (2017–2021)*. Retrieved from [http://www.gov.cn/xinwen/2017-12/20/content\\_5248855.htm](http://www.gov.cn/xinwen/2017-12/20/content_5248855.htm)
- Olson, R. S., Bartley, N., Urbanowicz, R. J., & Moore, J. H. (2016). *Evaluation of a tree-based pipeline optimization tool for automating data science* (pp. 485–492). GECCO 2016—Proceedings of the 2016 Genetic and Evolutionary Computation Conference. <https://doi.org/10.1145/2908812.2908918>
- Pearl, J. (2019). The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3), 54–60. <https://doi.org/10.1145/3241036>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in python Fabian. *Journal of Machine Learning Research*, 12, 2825–2830.
- Qi, J., Zheng, B., Li, M., Yu, F., Chen, C., Liu, F., et al. (2017). A high-resolution air pollutants emission inventory in 2013 for the Beijing-Tianjin-Hebei region, China. *Atmospheric Environment*, 170, 156–168. <https://doi.org/10.1016/j.atmosenv.2017.09.039>
- Qin, Y., Wagner, F., Scovronick, N., Peng, W., Yang, J., Zhu, T., et al. (2017). Air quality, health, and climate implications of China's synthetic natural gas development. *Proceedings of the National Academy of Sciences of the United States of America*, 114(19), 4887–4892. <https://doi.org/10.1073/pnas.1703167114>
- Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. The MIT Press, Massachusetts Institute of Technology.
- Reddington, C. L., Conibear, L., Knote, C., Silver, B. J., Li, Y. J., Chan, C. K., et al. (2019). Exploring the impacts of anthropogenic emission sectors on PM2.5 and human health in South and East Asia. *Atmospheric Chemistry and Physics*, 19, 11887–11910. <https://doi.org/10.5194/acp-19-11887-2019>
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat, fnm (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
- Rocklin, M. (2015). *Dask: Parallel computation with blocked algorithms and task scheduling* (pp. 126–132). Proceedings of the 14th Python in Science Conference, (SCIPY). <https://doi.org/10.25080/majora-7b98e3ed-013>
- Ryan, E., Wild, O., Voulgarakis, A., & Lee, L. (2018). Fast sensitivity analysis methods for computationally expensive models with multi-dimensional output. *Geoscientific Model Development*, 11(8), 3131–3146. <https://doi.org/10.5194/gmd-11-3131-2018>
- Rybarczyk, Y., & Zalakeviciute, R. (2018). Machine learning approaches for outdoor air quality modelling: A systematic review. *Applied Sciences*, 8, 2570. <https://doi.org/10.3390/app8122570>
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2), 259–270. <https://doi.org/10.1016/j.cpc.2009.09.018>
- Shi, Z., Li, J., Huang, L., Wang, P., Wu, L., Ying, Q., et al. (2017). Source apportionment of fine particulate matter in China in 2013 using a source-oriented chemical transport model. *Science of the Total Environment*, 601–602, 1476–1487. <https://doi.org/10.1016/j.scitotenv.2017.06.019>
- Silva, R. A., Adelman, Z., Fry, M. M., & West, J. J. (2016). The impact of individual anthropogenic emissions sectors on the global burden of human mortality due to ambient air pollution. *Environmental Health Perspectives*, 124(11), 1776–1784. <https://doi.org/10.1289/ehp177>
- Silver, B., Conibear, L., Reddington, C. L., Knote, C., Arnold, S. R., & Spracklen, D. V. (2020). Pollutant emission reductions deliver decreased PM2.5-caused mortality across China during 2015–2017. *Atmospheric Chemistry and Physics*, 20, 11683–11695. <https://doi.org/10.5194/acp-20-11683-2020>
- Silver, B., He, X., Arnold, S. R., & Spracklen, D. V. (2020). The impact of COVID-19 control measures on air quality in China. *Environmental Research Letters*, 15(8). <https://doi.org/10.1088/1748-9326/aba3a2>
- Silver, B., Reddington, C. L., Arnold, S. R., & Spracklen, D. V. (2018). Substantial changes in air pollution across China during 2015 to 2017. *Environmental Research Letters*.
- Skamarock, W. C., Klemp, J. B., Dudhi, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). *A description of the advanced research WRF version 3* (NCAR Tech 468+STR, p. 113).
- Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1–3), 271–280. [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6)
- Tran, C. T., Zhang, M., Andreea, P., & Xue, B. (2016). A wrapper feature selection approach to classification with missing data. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 9597, pp. 685–700). [https://doi.org/10.1007/978-3-319-31204-0\\_44](https://doi.org/10.1007/978-3-319-31204-0_44)
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*.
- Waskom, M., Botvinnik, O., Ostblom, J., Gelbart, M., Lukauskas, S., & Hobson, P. (2020). *mwaskom/seaborn: v0.10.1 (April 2020) (Version v0.10.1)*. Zenodo. <https://doi.org/10.5281/zenodo.3767070>
- Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., et al. (2020). Improved 1 km resolution PM2.5 estimates across China using enhanced space-time extremely randomized trees. *Atmospheric Chemistry and Physics*, 20(6), 3273–3289. <https://doi.org/10.5194/acp-20-3273-2020>
- Weichenthal, S., Hatzopoulou, M., & Brauer, M. (2018). A picture tells a thousand...exposures: Opportunities and challenges of deep learning image analyses in exposure science and environmental epidemiology. *Environment International*, 0–1. <https://doi.org/10.1016/j.envint.2018.11.042>
- Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., et al. (2011). The Fire INventory from NCAR (FINN)—A high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development*, 4, 624–641. <https://doi.org/10.5194/gmd-4-625-2011>

- Wild, O., Voulgarakis, A., O'Connor, F., Lamarque, J.-F., Ryan, E. M., & Lee, L. (2020). Global sensitivity analysis of chemistry-climate model budgets of tropospheric ozone and OH: Exploring model diversity. *Atmospheric Chemistry and Physics*, *20*, 4047–4058. <https://doi.org/10.5194/acp-20-4047-2020>
- Yeo, I.-K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, *87*(4), 954–959. <https://doi.org/10.1093/biomet/87.4.954>
- Yin, P., Brauer, M., Cohen, A. J., Wang, H., Li, J., Burnett, R. T., et al. (2020). The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990–2017: An analysis for the Global Burden of Disease Study 2017. *The Lancet Planetary Health*, *5*(196)(20), 1–13. [https://doi.org/10.1016/S2542-5196\(20\)30161-3](https://doi.org/10.1016/S2542-5196(20)30161-3)
- Yu, S., Eder, B., Dennis, R., Chu, S.-H., & Schwartz, S. E. (2006). New unbiased symmetric metrics for evaluation of air quality models. *Atmospheric Science Letters*, *7*(1), 26–34. <https://doi.org/10.1002/asl.125>
- Zaveri, R. A., Easter, R. C., Fast, J. D., & Peters, L. K. (2008). Model for Simulating Aerosol Interactions and Chemistry (MOSAIC). *Journal of Geophysical Research*, *113*(D13204), 1–29. <https://doi.org/10.1029/2007jd008782>
- Zhai, S., Jacob, D. J., Wang, X., Shen, L., Li, K., Zhang, Y., et al. (2019). Fine particulate matter (PM<sub>2.5</sub>) trends in China, 2013–2018: Contributions from meteorology. *Atmospheric Chemistry and Physics Discussions*, 1–19. <https://doi.org/10.5194/acp-2019-279>
- Zhan, Y., Luo, Y., Deng, X., Chen, H., Grieneisen, M. L., Shen, X., et al. (2017). Spatiotemporal prediction of continuous daily PM<sub>2.5</sub> concentrations across China using a spatially explicit machine learning algorithm. *Atmospheric Environment*, *155*, 129–139. <https://doi.org/10.1016/j.atmosenv.2017.02.023>
- Zhang, Q., Zheng, Y., Tong, D., Shao, M., Wang, S., Zhang, Y., et al. (2019). Drivers of improved PM<sub>2.5</sub> air quality in China from 2013 to 2017. *Proceedings of the National Academy of Sciences of the United States of America*, 1–7. <https://doi.org/10.1073/pnas.1907956116>
- Zhang, Y., Vu, T. V., Sun, J., He, J., Shen, X., Lin, W., et al. (2020). Significant changes in chemistry of fine particles in wintertime Beijing from 2007 to 2017: Impact of clean air actions. *Environmental Science & Technology*, *54*(3), 1344–1352. <https://doi.org/10.1021/acs.est.9b04678>
- Zhao, B., Zheng, H., Wang, S., Smith, K. R., Lu, X., Aunan, K., et al. (2018). Change in household fuels dominates the decrease in PM<sub>2.5</sub> exposure and premature mortality in China in 2005–2015. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(49), 12401–12406. <https://doi.org/10.1073/pnas.1812955115>
- Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., et al. (2018). Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. *Atmospheric Chemistry and Physics*, *18*(19), 14095–14111. <https://doi.org/10.5194/acp-18-14095-2018>
- Zheng, B., Zhang, Q., Tong, D., Chen, C., Hong, C., Li, M., et al. (2017). Resolution dependence of uncertainties in gridded emission inventories: A case study in Hebei, China. *Atmospheric Chemistry and Physics*, *17*(2), 921–933. <https://doi.org/10.5194/acp-17-921-2017>
- Zhuang, J., Dussin, R., Jüling, A., & Rasp, S. (2020). *JiaweiZhuang/xESMF: v0.3.0 adding ESMF.LocStream capabilities (version v0.3.0)*. Zenodo. <https://doi.org/10.5281/zenodo.3700105>