

1 **Meteorological Change and Impacts on Air Pollution –**
2 **Results from North China**

3 **Ziping Xu¹, Song Xi Chen², Xiaoqing Wu³**

4 ¹Yuanpei College, Peking University; Department of Statistics, University of Michigan, Ann Arbor, MI,
5 USA

6 ²Guanghua School of Management and

7 ²Center for Statistical Science, Peking University, Beijing 100871, China

8 ³Department of Geological & Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA

9 **Key Points:**

- 10 • Meteorological changes led to 1.9 percent to 2.7 percent reduction in annual PM2.5
11 averages over North China 2014 to 2016 driven by temperature warming.
12 • Significant increases are detected in the surface temperature, boundary layer height
13 and dissipation and decreases in relative humidity.
14 • The meteorological change should not be held responsible for the regional air pol-
15 lution problem in North China.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which

may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1029/2020JD032423](https://doi.org/10.1029/2020JD032423)

Abstract

There have been speculations that the severe air pollution experienced in North China was the act of meteorological change in general and a decreasing northerly wind in particular. We conduct a retrospective analysis on 1979-2016 reanalysis data from ERA-Interim of ECMWF over a region in North China to detect meteorological changes over the 38 years. No significant reduction in the northerly wind within the mixing layer is detected. Statistically significant increases are detected in the surface temperature, boundary layer height and dissipation, and significant decreases in relative humidity in the region between the first and second 19-year periods from 1979 to 2016. We build regression models of $PM_{2.5}$ on the meteorological variables using data in 2014, 2015 and 2016 to quantify effects of the meteorological changes between the two 19 years periods on $PM_{2.5}$ under the emission scenarios of 2014-2016. It is found that despite the warming, dew point temperature had been largely kept under control as the region had gotten dryer. This made the effects of temperature warming largely favorable to $PM_{2.5}$ reduction as it enhances boundary layer height and dissipation. It is found that the meteorological changes would lead to 1.29% to 2.76% reduction in annual $PM_{2.5}$ averages with January, March and December having more than 4% reduction in the three years. Thus, the meteorological change in North China had helped alleviate $PM_{2.5}$ to certain extent and should not be held responsible for the regional air pollution problem.

1 Introduction

A substantial part of China has experienced severe air pollution, and the region around the North China Plain (NCP) is the most severe. Air pollution is a major challenge faced by China as it tries to find a balance between economic growth and environmental sustainability. $PM_{2.5}$ [airborne particular matters (PM) with aerodynamic diameters less than $2.5\mu m$] has been the main air pollutant in China, whose main components are sulfate, nitrate, ammonium, organic carbon and elemental carbon. Epidemiological evidence shows that exposure to $PM_{2.5}$ can cause lung morbidity (Donaldson et al., 1998) and serious respiratory and cardiovascular diseases, and even death (Schwartz, 2000; Pope III et al., 2002; R. Chen et al., 2013).

The relationships between $PM_{2.5}$ and meteorological variables have been well studied in literature, such as in Tai et al. (2010, 2012) and Jacob and Winner (2009) in the context of climate change. Tai et al. (2012) revealed strong correlations between main species of $PM_{2.5}$ and temperature and relative humidity, respectively, using either the US EPA ground observations and simulated Geos-Chem grid data over the continental US. Jacob and Winner (2009) reviewed results from nine studies on the impacts of meteorological variables on $PM_{2.5}$. They used the perturbation analysis in Chemical Transportation Models (CTM) driven by General Circulation Model simulation (GCM) that mimicked various climate change scenarios. The review found that temperature, boundary layer height (BLH) and relative humidity had significant effects for $PM_{2.5}$, although the effects were weaker than those for the ozone, due to different components of the PM reacting differently to the meteorological conditions. Shen et al. (2018) and Pendergrass et al. (2019) revealed that winter $PM_{2.5}$ in Beijing was strongly correlated with the 850 hPa meridional wind velocity and the relative humidity. Stable synoptic and near-surface meteorological conditions were found to be strongly and positively associated with $PM_{2.5}$ concentration (Zheng et al., 2015; H. Chen & Wang, 2015; Miao et al., 2015). By analyzing data from the US Embassy, Liang et al. (2015) found that about 75% of $PM_{2.5}$'s variation in Beijing was driven by the meteorological factors. L. Chen et al. (2018) found the dew point temperature was more significant on $PM_{2.5}$ than the temperature and pressure based on air quality monitoring data in North China.

Climate change and its impacts on air pollution has been much investigated extensively in the literature. As summarized by Jacob and Winner (2009), existing studies can

67 be categorized as three approaches: (i) empirical correlation between PM and meteorological variables; (ii) perturbation analysis that perturb meteorological variables in regional Chemical Transport Models (CTM) so as to identify sensitive variables; and (iii) the General Circulation Model (GCM)-CTM simulation under various climate change scenarios. In particular, Tagaris et al. (2007) and Avise et al. (2009) projected declined $\text{PM}_{2.5}$ in year 2050 relative to that of year 2000 in the annual mean (-10%) and the July mean ($-1 \mu\text{g}/\text{m}^3$), respectively, in the US.

74 For the effects of climate change on China's air pollution, the results were mixed. There were studies projecting worsening air pollution under scenarios of climate change. H. Wang et al. (2015) found significant correlation between the number of winter haze days in eastern China with decreasing Arctic sea ice in the proceeding autumn based on 30 years' data from 1980s, and projected a worsening winter haze situation due to a northward shift of winter cyclone activity in East China (not focused on North China). Cai et al. (2017) found the winter haze days in Beijing were associated with northerly wind at 850 hPa, a dipole circulation pattern at 500 hPa, and temperature difference between the stratosphere and lower troposphere. By applying a statistical model to the future climate projections in the high emission scenario (Taylor et al., 2012), they found the circulation change induced by greenhouse gas emissions would contribute to more frequent haze days in Beijing from 2050 to 2100. In contrast, there were studies which predicted that the climate change was unlikely to significantly increase Beijing's $\text{PM}_{2.5}$ (Yin & Wang, 2016, 2017; Shen et al., 2018; Leung et al., 2018; Pendergrass et al., 2019). The number of winter haze days derived from the visibility observations in northern China displayed a decreasing trend from 1980 to 2012 and then this trend reverses after 2012 (Yin & Wang, 2016, 2017). Conclusions of both Shen et al. (2018) and Pendergrass et al. (2019) were based on the first principal component (PC) of the meteorological variables, which was dominated by the 850 hPa meridional wind velocity and the relative humidity, and had a strong explanation power on Beijing's $\text{PM}_{2.5}$. Shen et al. (2018) found that Beijing's $\text{PM}_{2.5}$ was directly dependent on a dipole structure in Arctic sea ice rather than the amount of sea ice, and the climate change effect was unable to reverse the effect of the current emission control effort in reducing Beijing's $\text{PM}_{2.5}$. And Pendergrass et al. (2019) indicated that climate change was unlikely to increase the frequency of severe $\text{PM}_{2.5}$ pollution events. Also based on the PC analysis, Leung et al. (2018) found that $\text{PM}_{2.5}$ in Beijing-Tianjin-Hebei region was highly dependent on the frequency of springtime Siberian highs and the relative humidity, and they projected a $0.46 \pm 0.28 \mu\text{g}/\text{m}^3$ reduction in 2050's annual $\text{PM}_{2.5}$ level relative to that of 2000's.

102 Our study contributes to the study on the effects of the climate changes on China's air pollution from a different perspective. It was motivated to empirically verify a speculation that the severe air pollution in North China was partly due to a weakening northerly wind as results of the meteorological change and wind energy development (Yan, 2015; Cheng, 2016). Two papers published in 2018 added more relevance to our study. X. Wang et al. (2018) found high aerosol loading as well as aerosols containing different elements encouraged tree growth. In another study, by simulations from the WRF-Chem model with emission inventory of 2013 and a module for the impacts of the surface roughness on the wind speed, Long et al. (2018) showed that the afforestation in the mountainous regions north and west of Beijing had contributed to a 6% increase in $\text{PM}_{2.5}$ in Beijing-Tianjin-Hebei region over the two-month period from December 2013 to January 2014.

113 Unlike most of the aforementioned studies on the climate changes, which are based on simulated projections via relevant models under certain emission and climate change and climate scenarios, this study takes a retrospect approach by obtaining realized meteorological changes and $\text{PM}_{2.5}$ -meteorological relationship (rather than scenarios) embedded in the observations. Our study consists of three steps. In the first step, the realized pattern of the meteorological change are acquired by conducting statistical testing for changes in the ECMWF's reanalysis data between the two 19-year periods from

120 1979 to 2016 over 261 spatial grids of size 0.5×0.5 degree over a region in North China
121 that encounters the most severe air pollution in the nation. In the second step, the emis-
122 sion scenarios of 2014 to 2016 are estimated by regression models between $PM_{2.5}$ and
123 the meteorological variables based on data of the three years. Finally, we evaluate the
124 impacts of the meteorological changes that happened from 1979 to 2016 on $PM_{2.5}$ con-
125 centration from 2014 to 2016 by substituting the meteorological means of the two 19 year
126 periods to the fitted regression models of 2014-2016, respectively.

127 Our study cannot confirm diminishing northerly wind in the region from 1979 to
128 2016 as speculated, but finds instead significant temperature warming, drying, and in-
129 creased boundary layer height and dissipation over the 38 year period from 1979 to 2016
130 in North China, which led to better vertical ventilation and a reducing effect on $PM_{2.5}$
131 as confirmed from the three year (2014-2016)'s regression models on the $PM_{2.5}$. Our cal-
132 culation shows that the meteorological changes had actually decreased $PM_{2.5}$ in the ma-
133 jority of the months for most cities of the study region. Our results are in tune with the
134 results of Shen et al. (2018); Leung et al. (2018); Pendergrass et al. (2019) confirming
135 the "climate benefit", although we are focused on the past effects on 2014-2016's $PM_{2.5}$
136 while those three studies were focused on the projected effects in the future.

137 2 Data and Models

138 2.1 Study Region

139 The study region shown in Figure 1 ranges from 32.75° N to 42.75° N latitude and
140 from 112.75° E to 119.75° E longitude excluding area over the sea. Located at the heart
141 of North Central China that encompasses NCP and its surrounding areas, it covers a land
142 area of $310,000 \text{ km}^2$ with a population exceeding 320 million which accounts for about
143 $1/4$ of China's population. In the last two decades, the region witnessed rapid industri-
144 alization and urbanization as well as chronic air pollution (Xu et al., 2013; Liang et al.,
145 2015; Zhang et al., 2017).

146 2.2 Meteorological Data

147 We utilized a six-hourly gridded dataset, ERA-Interim of Global Reanalysis from
148 ECMWF (Dee et al., 2011), which is produced with a data assimilation scheme advanc-
149 ing forward in time using 12-hourly analysis cycles. Starting from year 1979, the dataset
150 has a spatial resolution of 0.5° latitude by 0.5° longitude, which results in 261 equal-size
151 grids over the study region, excluding areas over the sea. The temporal frequency of the
152 data is four times per day at 0200, 0800, 1400 and 2000 local time (LT). We choose ERA-
153 Interim dataset because it contains the years 2014 to 2016 that cover the range of the
154 available air quality data. We considered a time span up to the year 2016, which gives
155 rise to 38 years of data. We evenly divide it into two periods: years 1979-1997 as Period
156 1 and years 1998-2016 as Period 2. Period 1 serves as the "control" period and Period
157 2 as the "treatment" period for the meteorological change. Incidentally, the year 1979
158 marked the first year of the economic reform in China, and the first period between 1979
159 to 1997 corresponded to a relatively slow pace of economic growth and industrialization,
160 while the second period witnessed more rapid economic expansion as shown by the gross
161 domestic product (GDP) and energy consumption statistics in Figure S1 in the suppl-
162 ementary information (SI).

163 Ten basic variables in the ECMWF dataset are considered. They include four surface-
164 level variables: ground surface pressure (SP) in *hPa*, two-meter air temperature (T2M)
165 and dew-point temperature (D2M) both in *Kelvin* and total precipitation (PREC) in
166 *meter*; and two variables on the intensity of turbulent processes and vertical structure
167 of the boundary layer: boundary layer dissipation (BLD) in J/m^3 and height (BLH) in
168 *meter*.

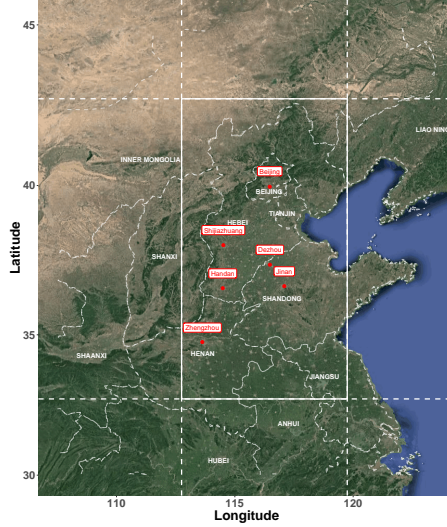


Figure 1. The study region marked within the white square (32.75° N to 42.75° N in latitude and from 112.75° E to 119.75° E in longitude excluding sea area) with the red dots marking six selected cities where we will conduct further analysis.

The choice of these variables is due to their roles in defining the basic meteorological condition within the boundary layer: the secondary generation (D2M and T2M), removal (PREC) and vertical dispersion (SP, BLH and BLD) of $\text{PM}_{2.5}$.

The remaining four variables are wind speed vector (u, v) in the eastward and northward components due to their effects on the horizontal removal or transportation of pollutants, temperature (T) in *Kelvin* and specific humidity (SH) in kgkg^{-1} . The four are available over vertical pressure layers from 1000 hPa to 300 hPa, with a 25 hPa increment from 1000 to 750 hPa (11 layers) and a 50 hPa increment above 750 hPa (9 layers to 300 hPa). We did not consider data above 300 hPa as those pressure levels are well above the boundary layer in all seasons as shown in Figure S2 in the SI.

The following variables at each pressure layer are deduced from the ECMWF dataset. They include potential temperature (PT), relative humidity (RH), and wind speed $\text{WS} = \sqrt{u^2 + v^2}$. Furthermore, we categorize wind to four broad directions NE, NW, SE and SW according to NE in $[0^{\circ}, 90^{\circ})$, NW in $[90^{\circ}, 180^{\circ})$, SW in $[180^{\circ}, 270^{\circ})$ and SE in $[270^{\circ}, 360^{\circ})$. Non-zero wind speed is assigned to the prevailing wind direction at pressure level p at time t , while the other three directions' are set to zero.

As air pollution is affected by meteorological conditions at all pressure levels within the boundary layer, we integrate each variable over the pressures levels beneath the boundary layer at a time, and obtain integrated wind speed at the four directions, denoted as (INE, INW, ISE, ISW), integrated potential temperature (IPT), integrated specific humidity (ISH) and integrated relative humidity (IRH). Indeed, the cleaning or increasing effects of $\text{PM}_{2.5}$ are the act of the winds within the air column rather than at a single pressure level, which has been used in Wuerch et al. (1972), Egan and Mahoney (1972) and Holzworth (1972). Figure S2 in the SI provides distributions of the numbers of pressure layers inside the BL for each season and time of the day from the years 1979 to 2016.

The assimilated data are subject to errors as reported in Travis et al. (2016), since they depend on the quality of the models used for the underlying meteorological process. To gain information on the quality of the assimilated meteorology by ECMWF in the study region, we have computed monthly Pearson's correlation coefficients between the

198 assimilated surface pressure (SP), temperature (T2M) and relative humidity (RH) with
 199 the corresponding site measurements from 2010 to 2016 over 67 grid points where site
 200 observations are available. The same correlations are also computed for another assim-
 201 ilated data dataset, NASA’s MERRA2 (Modern-Era Retrospective Analysis for Research
 202 and Applications Version 2). The correlations are displayed in Figures S3 of SI and are
 203 summarized in Tables S1 of SI. Table S1 shows that ECMWF had substantially higher
 204 correlations with the site measurements than MERRA2 in T2M (9.3% higher) and RH
 205 (55.1% higher). Although ECMWF’s correlations were 5% lower than those of MERRA2
 206 for SP, the monthly averaged correlations were all higher than 0.898. This lends support
 207 to the assimilated meteorology by ECMWF relative to MERRA2 for the study region.

208 2.3 Air Pollution Data

209 The air pollution data are $PM_{2.5}$ observations at the hourly frequency from 161
 210 air-quality monitoring sites at 32 cities in the study region. These sites are the so-called
 211 Guokong (nationally controlled) sites, which are directly administrated by the Ministry
 212 of Ecology and Environment with the data transmitted instantaneously to a data cen-
 213 ter in Beijing to avoid potential local interference. Regional monitoring data on $PM_{2.5}$
 214 were not available before Year 2013, the year China’s air quality monitoring network was
 215 established. We did not use $PM_{2.5}$ data prior to 2014 as they endured high proportions
 216 of missing values in the study region. To match the six-hourly frequency of the mete-
 217 orological data, we choose $PM_{2.5}$ data at 0200, 0800, 1200 and 2000LT. The data were
 218 available from January 2013 when China first established a national air quality moni-
 219 toring network. The quality of the data was evaluated in Liang et al. (2016) in a com-
 220 parative study to data from the US diplomatic posts in China, and they found the two
 221 data sources were highly consistent.

222 Considering the air quality data is to establish a relationship between the mete-
 223 orological variables and $PM_{2.5}$ concentration so that the impact of meteorological change
 224 on air pollution can be evaluated. The air quality sites are matched to the center of the
 225 nearest grid in the ECMWF data. The locations of the 161 sites with the grids overlaid
 226 are shown in Figure S5 of SI.

227 2.4 Models and Assumptions

228 Let X denote one of the meteorological variables (SP, T2M, D2M, BLH, BLD, IPT,
 229 IRH, INW, INE, ISW and ISE), and $X_{iajt}(s)$ denote the observed value at a grid loca-
 230 tion s at Period $i \in \{1, 2\}$ (one of the two 19 year periods), year $a \in \{1, \dots, 18\}$, month
 231 $j \in \{1, \dots, 12\}$, and a six-hourly $t \in \{1, \dots, n_j\}$. Namely, it is the t -th observations
 232 in month j that has a total of n_j observations.

233 The model for the meteorological variable at a grid s is

$$234 X_{iajt}(s) = \mu_j^F(t; s) + \mu_{ij}^C(t; s) + \epsilon_{iajt}(s), \quad (1)$$

235 where $\mu_j^F(t; s)$ denotes the pre-1998 average at each time t of month j , $\mu_{ij}^C(t; s)$ repre-
 236 sents the meteorological change term and $\epsilon_{iajt}(s)$ is the random error for natural inter-
 237 nal variability in the system. The fixed effect $\mu_j^F(t; s)$ describes the underlying pre-1998
 238 climatological regime and is indexed to the time t of month j , reflecting the underlying
 239 meteorological pattern at grid location s and the particular time t of the month.

The meteorological change term $\mu_{ij}^C(t; s)$ measures the change between the two 19-
 years periods. To make the model identifiable, we assume in Period 1

$$\mu_{1j}^C(t) \equiv 0 \quad \text{for all } j \text{ and } t.$$

240 Moreover, we assume the meteorological change in Period 2, if any, is time-homogeneous
 241 within a month j so that

$$242 \mu_{2j}^C(t) = \Delta_j. \quad (2)$$

243 The random errors $\{\epsilon_{iajt}(s)\}$ are serially dependent satisfying the α -mixing condition
 244 (Bosq, 1998; Fan & Yao, 2008) to reflect the temporal persistence of the meteorologi-
 245 cal processes. It is reasonable to assume

$$\begin{aligned} \epsilon_{iajt}(s_1) \text{ and } \epsilon_{ia'jt'}(s_2) \text{ are independent for any } t \text{ and } t' \\ \text{and any two grid locations } s_1 \text{ and } s_2 \text{ as long as } a \neq a', \end{aligned} \quad (3)$$

246 namely the residual series in different years are independent.

247 The null and alternative hypotheses we want to test for month j at grid s are, re-
 248 spectively,

$$249 \quad H_{0js} : \mu_{2j}^C(t; s) = 0 \text{ versus } H_{1js} : \mu_{2j}^C(t; s) \neq 0. \quad (4)$$

To remove the fixed effect $\mu_j^F(t; s)$, which represents the underlying meteorological pat-
 250 tern, we take differences between the two periods to attain

$$Y_{aj}(t; s) \hat{=} X_{2aj}(t; s) - X_{1aj}(t; s) = \mu_{2j}^C(t) + \epsilon_{2ajt}(s) - \epsilon_{1ajt}(s).$$

250 The test statistic for the meteorological change hypothesis (4) is

$$251 \quad T_j(s) = \frac{1}{19n_j} \sum_{a=1}^{19} \sum_{t=1}^{n_j} Y_{aj}(t; s). \quad (5)$$

It can be shown that $T_j(s)$ is asymptotic normally distributed under H_{0js} such that

$$\sqrt{19n_j}T_j(s) \rightarrow N(0, \sigma_j^2(s)) \text{ in distribution as } n_j \rightarrow \infty,$$

252 where $\sigma_j^2(s)$ is the asymptotic variance of $T_j(s)$. As the residuals are temporally depen-
 253 dent, we use the spectral density approach (Brockwell & Davis, 2013; S. X. Chen & Tang,
 254 2005) to obtain a consistent estimator $\hat{\sigma}_j(s)$ of $\sigma_j(s)$. See Appendix A.1 for more details.

255 2.5 Control False Discovery Rate

256 The proposed testing for meteorological changes over the two 19-year periods is con-
 257 ducted by obtaining the p-value on each grid box first followed by combining the p-values
 258 from all the 261 grid boxes with the controlled false discovery rate (FDR). Specifically,
 259 for each meteorological variable, we conduct 261 tests based on the statistics $\{T_j(s)\}_{s=1}^{261}$
 260 that correspond to the 261 grids in the study region at each month $j = 1, \dots, 12$. Due
 261 to the repeated testing over the 261 grids, there will be an accumulation of false rejec-
 262 tions (false discoveries), namely rejecting no-meteorological change simply due to the type-
 263 I testing error. To control the false discoveries, we carry out a multiple testing proce-
 264 dure that controls the FDR (Benjamini & Hochberg, 1995). Specifically, for each vari-
 265 able to be tested for meteorological changes between the two periods, let $\eta_1, \eta_2, \dots, \eta_{261}$
 266 be the p-values of the tests for the hypotheses H_{0js} , $s = 1, \dots, 261$, respectively; and
 267 let $\eta_{(1)} \leq \eta_{(2)} \leq \dots \leq \eta_{(m)}$ be the p-values in ascending order. The multiple testing
 268 procedure that controls the FDR at an α -level rejects all $H_{(0js)}$ with $\eta_{(i)} \leq (i\alpha)/261$.
 269 In our study, we set $\alpha = 0.05$.

270 It is noted that the variation in the data and the test statistic at each grid-box is
 271 assessed in the context of the 261 grid-boxes via the FDR control. The FDR control leads
 272 to smaller significance levels $(i\alpha)/261$ instead of α at the grid level, which would reduce
 273 the influence of outliers on the conclusion of the analysis.

274 3 Detecting Meteorological Changes from 1979 to 2016

275 We present testing results on statistically significant changes in the meteorologi-
 276 cal variables between the two 19-year periods from 1979 to 2016, which will be used as

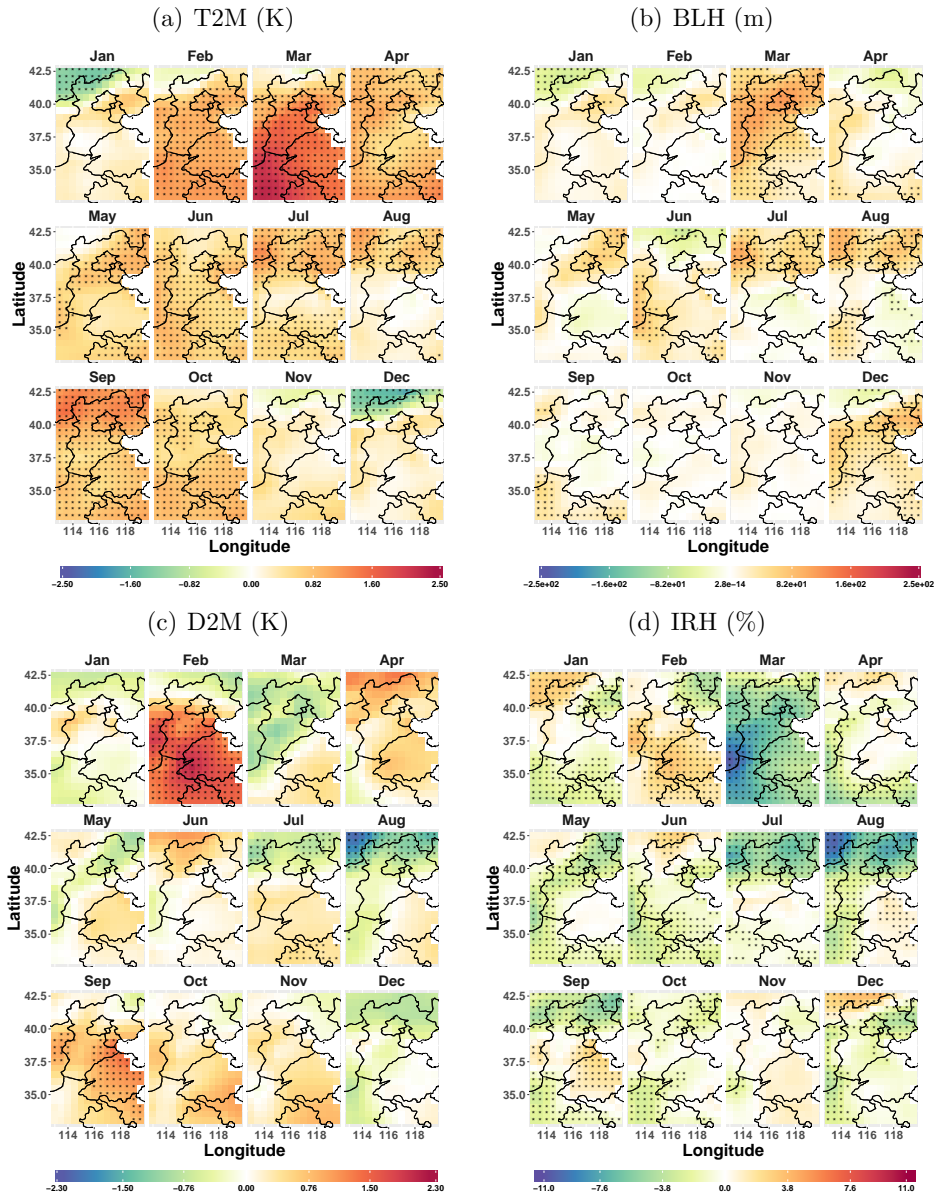


Figure 2. Differences between two 19-year periods' average values for two meter temperature (in *Kelvin*) (a), boundary layer height (BLH in meter) (b), two meter dew point temperature (in *Kelvin*) (c) and integrated relative humidity (in %) (d) for the twelve months at the 261 grids between the two time periods. Redness (blueness) indicate an increase (decrease) in a variable while stars mark statistically significance change by controlling the FDR at 5%. Black lines mark the provincial borders.

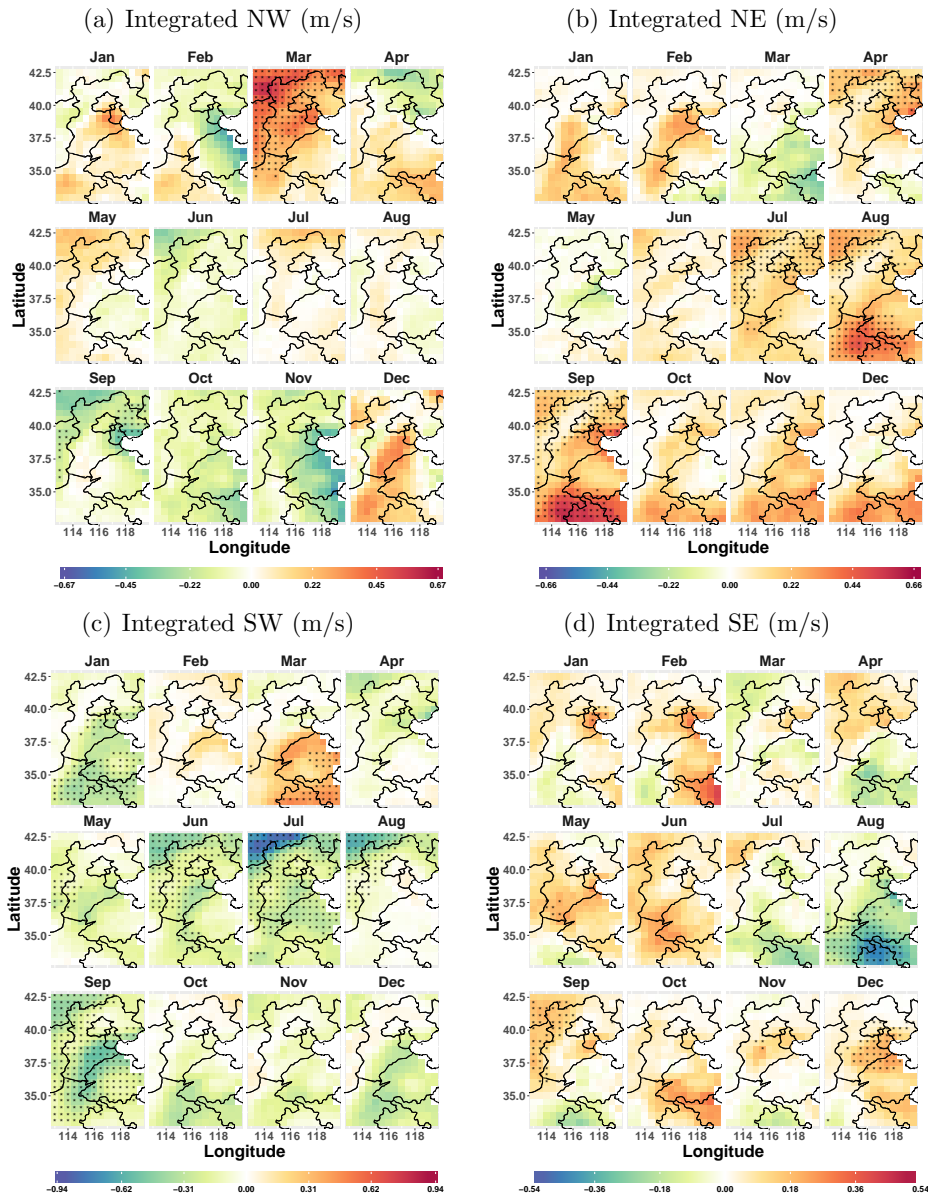


Figure 3. Differences between two 19-year periods' averages for the integrated wind speed (meter/second) at four directions: NW (a), NE (b), SW (c) and SE (d) for the twelve months at the 261 grids between the two time periods. Redness (blueness) indicate increase (decrease) in the variable while stars mark significance change by controlling the FDR at 5%.

277 the meteorological inputs in Section 5 when we quantify their effects on the $PM_{2.5}$ under
278 the emission profiles of the three years from 2014 to 2016. Figure 2 displays results
279 on T2M, BLH, D2M, and IRH for the twelve months at the 261 grids.

280 The most pronounced changes between the two 19-year periods were in the two-
281 meter temperature (T2M) whose change was evident in the nine months from February
282 to October over a large area of the study region. This confirmed surface warming over
283 a vast portion of the region by about 0.54 K (0.2% of the average T2M in Period 1) in
284 the nine months. There were some significant decreases for the surface temperature in
285 the mountainous north-west corner in December and January. From November to Jan-
286 uary, decreased temperature (statistically significant or not) only occurred in north-
287 west mountain areas. For the vast areas in the NCP, there was no grid that had decreased
288 T2M from November to January, and six of the nine months between February and Oc-
289 tober witnessed significant increases over 70% of the grids.

290 BLH was significantly increased in March at 88% of the grids, followed by Decem-
291 ber at around 50% of the grids. The average increase in March over the entire region was
292 57m (representing 9% of the average in Period 1), and the average increase in Decem-
293 ber was 27m (7% of the average in Period 1). BLH also increased in July and August
294 in the northern part of the region, and it only significantly decreased in an area that con-
295 sisted of 11% of the grids toward the northwest corner in January, and no more than 9
296 grids in June and December toward the northern edge of the region. Figure 2 indicates
297 that the change in T2M was consistent with the change of BLH in most months because
298 the temperature is a determining factor of BLH.

299 Consistent with the BLH, the average BLD in Period 2 (shown in Figure S9) had
300 significantly increased over quite a large area (more than 50% of grids) in March and De-
301 cember representing 25% and 15% increase over the average in Period 1, respectively.
302 It decreased over the north and east area of the region in June and no much change in
303 the other nine months.

304 Much less average changes were detected for D2M. The two-meter dew point tem-
305 perature was increased in February in a vast area south of Beijing over more than 80%
306 of grids. D2M was lower in Period 2 at 31 and 75 grids in July and August, respectively.
307 It is noted that the significant surface warming in Period 2 had not translated to much
308 increase in D2M as North China became drier by a decrease in IRH as shown in panel
309 (d) of Figure 2. This was because the warming was much canceled by the decrease in
310 IRH that led to a rather stable dew point temperature. An increase in D2M would be
311 bad news as it is positively correlated with $PM_{2.5}$ as shown in the next section.

312 The changes between the two periods for the vertically integrated potential tem-
313 perature (IPT) and specific humidity (ISH), as shown in Figure S6 of SI, had similar pat-
314 terns to those of T2M and D2M, respectively, which suggests that the situations within
315 the BL are highly correlated with those at the surface level. Figure S7 also displays the
316 changes in the precipitation (PREC), which shows that the amount of precipitation was
317 significantly reduced as was the case for the RH in the northern part of the regions over
318 the summer, which was part of the reason that prevented D2M to raise despite the warm-
319 ing.

320 As will be shown later on the relationship between $PM_{2.5}$ and the meteorological
321 variables, increased BLH and BLD, and lower dew point temperature tended to reduce
322 the level of $PM_{2.5}$. The lower dew point in the northern edge of the region (just north
323 of Beijing) also helped reduce $PM_{2.5}$.

324 Figure 3 presents the average changes in the integrated wind speed in the four di-
325 rections. The most significant changes occurred in the NE and SW winds, with the NE
326 significantly increased from July to September and SW decreased from June to Septem-
327 ber over a large portion of the region. For INW and ISE, the changes were much sub-

328 dued. INW increased in March over a large area in the north and west part of the re-
 329 gion and decreased in the central west and a northeast area in September. There were
 330 no significant changes in the other ten months, and especially no significant changes in
 331 three winter months. ISE decreased in the southern area of Henan province in August,
 332 and increased in Shanxi and northwest of Beijing in September and an area of eastern
 333 Hebei and Tianjin in December. The increase in December was negative news for air qual-
 334 ity in Beijing as it led to the transportation of pollution from the heavily industrialized
 335 area south and east of Beijing.

336 Overall, we did not see a significant reduction in the average NW wind in the re-
 337 gion and the area around Beijing in particular, except over two areas in September. Septem-
 338 ber is a month with relatively low air pollution. The significant increase in the NE wind
 339 and the reduction in SW wind were favorable to relief the air pollution as shown in the
 340 next section. For three winter months, when the air pollution was the severest, the ver-
 341 tically integrated wind speeds of the four directions showed no obvious change, which
 342 suggested that the pollution in winter could not be attributed to a decrease in the northerly
 343 wind.

344 To justify for the choice of the 19-year bin, we have implemented the same test over
 345 a smaller temporal bin of 9 years. Specifically, we replicated the testing for meteorolog-
 346 ical changes over 9-year bins within the first and the second 19-year periods with a mid-
 347 dle gap year, respectively, and also the changes between the second 9-year of the first
 348 19 year period and the first 9-year of the second 19-year period. The test results are dis-
 349 played in Figure S12 in the SI. When the size of the temporal bin was reduced to 9 years,
 350 more significant changes were detected which indicated more internal variability and a
 351 pattern that might differ from the pattern using the 19-year bin. This was largely due
 352 to the short term temporal variation of the data. However, by averaging over the two
 353 19-years bins, we detected more stable meteorological changes of longer-term, which are
 354 the changes we are interested in this study.

355 4 Impacts of Meteorological variables on PM_{2.5} in Winter

356 To establish a quantitative relationship between the air quality and the meteorolo-
 357 gical variables, we conduct a regression analysis that links PM_{2.5} concentration with
 358 the nine meteorological variables in six major cities in the region: Beijing, Dezhou, Han-
 359 dan, Jinan, Shijiazhuang and Zhengzhou, which are well separated spatially as shown
 360 in Figure 1. Beijing's PM_{2.5} data were from nine sites after excluding one background
 361 measurement site and two sites in the outlying northern districts. The numbers of sites
 362 for the other five cities were three in Dezhou and Handan, respectively, six in Jinan, five
 363 in Shijiazhuang and seven in Zhengzhou. As air pollution is the most severe in the win-
 364 ter heating season for the study region, we focus on two heating seasons from Novem-
 365 ber 15, 2014/2015 to March 15, 2015/2016, respectively. Next section will consider other
 366 months of the year in all 32 cities in a larger scale analysis. Our purpose here is to build
 367 yearly meteorology to PM_{2.5} profiles from 2014 to 2016, which will be used to quantify
 368 effects of the meteorological changes happened from 1979 to 2016 on 2014-2016's PM_{2.5}
 369 in the next section.

370 Let $Y_{at}(c, s)$ be the logarithm of PM_{2.5} at site s of city c , year a ($a = 1, 2$ for two
 371 heating seasons) and time t , and $X_{at}(c, s)$ consists of the nine meteorological variables:
 372 SP, logarithm of BLD (LogBLD), logarithm of BLH (LogBLH), T2M, D2M, INE, INW,
 373 ISE, ISW and lagged PM_{2.5} in the previous six hours. The log-transform is to reduce
 374 the skewness in PM_{2.5} and those meteorological variables to make them more symmet-
 375 rically distributed. We did not include precipitation (PREC) as the region is very dry
 376 in winter, which would result in a vast number of zero values. Neither we considered IPT
 377 or IRH as they are highly correlated with T2M and D2M, respectively. The models with

378 D2M replaced by IRH and ISH respectively, and T2M replaced by IPT all had lower R^2 's
379 as shown in Tables S2, S4 and S5 of the SI, which lends support for the use of D2M.

The regression model is

$$Y_{at}(c, s) = \alpha_a(c) + \beta_a^T(c)X_{at}(c, s) + \gamma_a(c)Y_{a(t-1)}(c, s) + \eta_a^T(c)I_t + \epsilon_{at}(c, s), \quad (6)$$

380 where $\alpha_a(c)$ represents the fixed effect in each city and $\beta_a(c)$ is the regression coefficient
381 to the meteorological variables, $I_t := (I_8(t), I_{14}(t), I_{20}(t))^T$ are indicators for 08:00, 14:00
382 and 20:00 local times respectively, $\eta_a(c)$ are the fixed time effects at the three hours, and
383 $\epsilon_{at}(c, s)$ are residuals with zero mean and a finite variance. Specifically, $I_8(t) = 1$ when
384 t corresponds to 08:00, and 0 otherwise; and the two other indicators $I_{14}(t)$ and $I_{20}(t)$
385 can be defined similarly. As the study utilizes the six-hourly data, I_t reflects the diurnal
386 emission pattern at 08:00, 14:00 and 20:00, respectively, with that at the 02:00 be-
387 ing the baseline.

Model (6) has an auto-regressive (AR) component via the lagged term $\gamma_a(c)Y_{a(t-1)}(c, s)$.
Auto-regressive models have been utilized in modeling the stratospheric ozone (Bojkov
et al., 1990; Reinsel et al., 2005). The process of $PM_{2.5}$ consists of several major forces:
the underlying emission, the meteorology as well as their interaction. If we difference Model
(6) between times t and $t-1$, we attain

$$\Delta Y_{at}(c, s) = \beta_a^T(c)\Delta X_{at}(c, s) + \eta_a^T(c)\Delta I_t + \gamma_a(c)\Delta Y_{a(t-1)}(c, s) + \Delta \epsilon_{at}(c, s),$$

388 where Δ is the difference operator with respect to t . The difference model implies that
389 the changes in $PM_{2.5}$ at time t is driven by a change of meteorological variable at time
390 t and the change of $PM_{2.5}$ at the previous time $t-1$, which reflects the changes in the
391 underlying emission in the previous time period.

392 Another physical interpretation of the auto-regressive Model (6) is attained if we
393 substitute $Y_{a(t-1)}(c, s)$ by $t-1$ meteorological variables and $t-2$ pollution concentra-
394 tion $Y_{a(t-2)}(c, s)$ according to (6), and repeat it recursively for L times to attain

$$\begin{aligned} Y_{at}(c, s) = & \alpha_a(c) \sum_{l=0}^L \gamma_a^l(c) + \sum_{l=0}^L \gamma_a^l(c) [\beta_a^T(c)X_{a(t-l)}(c, s) + \eta_a^T(c)I_{t-l}] \\ & + \gamma_a^L(c)Y_{a(t-L)}(c, s) + \sum_{l=0}^L \gamma_a^l(c)\epsilon_{a(t-l)}(c, s). \end{aligned} \quad (7)$$

395 This form implies that $PM_{2.5}$ at time t depends on a cumulative effects of the meteo-
396 rological variables from times $t-L+1$ to t and the "initial" $PM_{2.5}$ value $Y_{a(t-L)}(c, s)$
397 at $t-L$, in the presence of the city fixed effect and the diurnal effect. The cumulative
398 meteorological effects are transferred iteratively from the lagged $PM_{2.5}$ term.

399 Model (6) is a yearly-stratified model, proposed largely because emission data at
400 the six-hourly frequency are not available and yet a much changing emission profile be-
401 tween years 2014 and 2016 due to the pollution mitigation initiative in the study region.
402 A model that is based on the entire data (un-stratified) with a yearly dummy variable
403 would still have heterogeneity due to the changing emission profiles, which cannot be fully
404 explained by the yearly dummy variable. The impacts of the emission are reflected in
405 the intercept parameter $\alpha_a(c)$, the three 6-hourly dummy variables and the regression
406 coefficients $\beta_a(c)$.

407 In fitting the regression model, data from all monitoring sites in a city were pooled
408 together to estimate the regression coefficients. This would alleviate to certain extent
409 the sample size issue when fitting the yearly models. It is noted from Table 1 that de-
410 spite we conducted the yearly regression, the sign and the magnitude of the significant
411 regression coefficients over the two heating seasons in each city were largely consistent,

412 indicating a certain degree of model robustness. To make the estimated regression co-
 413 efficients directly comparable, we standardized all meteorological variables in (6) by their
 414 respective means and standard deviations. The coefficients $\alpha_a(c)$, $\eta_a(c)$ and $\beta_a(c)$ were
 415 estimated with the ordinary least squares. The standard deviations of the estimates were
 416 obtained by the AR-Sieve bootstrap (Bühlmann, 2002), which led to the p-values for the
 417 significance of the regression coefficients.

Table 1. Estimated regression coefficients of model (6) and their statistical significance for two heating seasons of the six cities.

(a) November 15th 2014 to March 15th 2015

	Beijing	Dezhou	Handan	Jinan	SJZ	Zhengzhou
SP	-0.06***	-0.1131***	-0.036	-0.08***	-0.0102	-0.009
T2M	0.01	-0.0365	-0.003	-0.09**	0.0114	0.186***
D2M	0.08***	0.178***	0.082**	0.23***	-6e-04	-0.013
LogBLD	-0.14***	-0.1547***	-0.064*	-0.17***	-0.0121	-0.117***
LogBLH	-0.19***	-0.1737**	0.022	-0.11**	-0.1233***	0.013
INW	-0.01	0.113*	-0.124***	0.2***	-0.2029***	-0.048
INE	-0.07***	0.1244***	-0.009	0.17***	-0.0365*	-0.003
ISW	0.06***	0.0723*	-0.077**	0.09***	0.0737***	-0.049**
ISE	0.02*	0.0365	-0.01	0.08***	0.0231**	-0.05***
hour8	-0.02	0.0166	-0.212***	0.02	-0.1004***	0.062***
hour14	0.08***	2e-04***	-0.233	0.09***	-0.1913	-0.204
hour20	0.16	0.0972	-0.147	0.18**	-0.0442	0.072***
Lagged Y	0.61***	0.5474***	0.586**	0.57***	0.5724	0.682
R^2	0.66	0.60	0.49	0.55	0.60	0.64

(b) November 15th 2015 to March 15th 2016

SP	-0.058***	-0.07*	-0.09***	-0.0446**	-0.002	-0.026*
T2M	-0.034*	-0.15**	-0.19***	-0.0729**	-0.074**	0.007
D2M	0.128***	0.16**	0.15***	0.1175***	0.105***	0.025*
LogBLD	-0.077***	-0.16***	-0.06*	-0.124***	-0.002	-0.145***
LogBLH	-0.202***	-0.12**	0.08**	-0.0423	-0.11***	0.044**
INW	-0.029*	0.05	-0.35***	-0.0763**	-0.233***	-0.097***
INE	-0.039***	-0.01	-0.09***	-0.0574*	-0.097***	-0.091***
ISW	0.025**	0.06**	-0.13***	-0.0474**	0.007	-0.032**
ISE	-0.038***	0.03	-0.13***	-8e-04	-0.027***	-0.058***
hour8	0.002	0.05*	-0.05*	0.0505***	-0.113***	0.032**
hour14	0.179***	0.05*	-0.08***	0.0441**	-0.096	-0.11*
hour20	0.167*	0.11**	0.13***	0.1152**	0.095**	0.099
Lagged Y	0.644***	0.64**	0.6***	0.6928***	0.599***	0.762*
R^2	0.74	0.68	0.64	0.716	0.662	0.75

Note: *p<0.05; **p<0.01; ***p<0.001

418 Table 1 (a) & (b) report the estimated regression coefficients for the two heating
 419 seasons of the six cities with marked levels of significance. Table 2 provides the orders
 420 of meteorological covariables being selected in a forward selection procedure based on
 421 the Akaike Information Criterion (AIC). There is no surprise to see that the lagged PM_{2.5}
 422 was the most significant variable in all cities with both the largest coefficient values and

Table 2. Variable ranks in the forward variable selection based on the AIC criterion, which denote the orders in which the variables were selected with the number on the left (right) based on the 2014-2015 (2015-2016) heating season of the six cities. A ”-” indicates the selection procedure was ended before the variable, whose rank is set as 13.

	Beijing	Dezhou	Handan	Jinan	SJZ	Zhengzhou	Average Rank
Lagged Y	01, 01	01, 01	01, 01	01, 01	01, 01	01, 01	1.0, 1.0
Hour20	03, 02	06, 04	05, 04	03, 05	08, 04	08, 07	5.5, 4.3
LogBLD	04, 03	05, 03	02, 05	05, 04	-, -	05, 06	5.7, 5.7
LogBLH	08, 05	02, 02	-, 12	10, 02	06, 02	-, 02	8.7, 4.2
INW	02, 12	-, -	07, 02	07, 11	02, 03	02, 04	5.5, 7.5
Hour14	09, 06	08, -	03, 03	11,07	03, 07	03, 05	6.2, 6.8
D2M	06, 07	03, 05	06, 11	02, 03	-, 08	-, 11	7.2, 7.5
INE	05, 09	07, 05	-, 08	07, 13	05, 05	-, 03	8.3, 7.2
SP	07, 04	04, 07	09, 09	06, 10	-, -	-, 09	8.7, 8.7
Hour8	12, -	-, 09	04, 13	-, 06	04, 06	09, 10	9.2, 9.5
T2M	-, 11	-, 06	-, 10	04, 09	-, 09	04, 08	10, 8.8
ISE	11,10	-, -	-, 06	08, 08	09, 10	06, 08	10, 9.2
ISW	10,08	-, 10	08, 07	09, 12	07, -	07, 12	9.0, 10.3

the leading rank being selected, as it reflected the strong temporal persistence nature of $PM_{2.5}$. It is noted that as all the covariables had been standardized by their respective means and standard deviation, the magnitude of the estimated regression coefficients in Table 1 (a) & (b) and the ranking in Table 2 were highly consistent to each other in that those variables with larger and significant estimated coefficients (in absolute value) tended to rank higher in Table 2. The diurnal effect was significant in all six cities with at least one (two) significant hour in 2015 (2016) heating season, relative to the baseline 02:00, which also showed an increased diurnal effect between the two heating seasons.

As shown in Table 1 (a) & (b), both BLD and BLH were the top-ranked meteorological variables as reflected by both the numeric size of the estimated coefficients, the range of the p-values and the rank being selected (Table 2). The largely negative coefficients to BLD and BLH indicated their reducing effects as both variables reflect the vertical ventilation. INW ranked next in Table 2, which also had significant reducing effects on $PM_{2.5}$ in 7 out of 12 heating season and city combinations. INE had the same reducing effects, while the ranking was lower than INW. These largely confirmed the benefits of the northerly wind in the study region.

Among the three surface variables (T2M, D2M and SP), D2M was the most significant (at 5% level) in 10 out of the 12 city and season combinations with a significant $PM_{2.5}$ enhancing effect, followed by SP which had a significant reducing effect in 8 out of the 12 season and city combinations. T2M had a mixed and weaker effect on $PM_{2.5}$. This was because T2M is positively correlated with both dew point temperature and BLH. However, the dew point was $PM_{2.5}$ enhancing while BLH was $PM_{2.5}$ reducing, as respectively shown by the signs of the regression coefficients in Table 1. These conflicting effects as well as the added diurnal effects via the time dummy variables made the net effect of T2M mixed and weaker.

Regarding the four integrated wind speeds, the ones at the NW and NE directions were more influential than those at SE and SW. INW and INE significantly reduced $PM_{2.5}$

451 in Beijing and Shijiazhuang with relatively larger coefficients and smaller p-values. The
452 signs of the coefficients to the integrated wind speeds of the four directions reflected the
453 geographical emission profiles surrounding a city. It is noted that INW and INE's co-
454 efficients were positive for Dezhou and Jinan in 2015 because the two cities are located
455 south of Hebei with heavy industrial emission, and therefore, the northerly wind would
456 bring in transported pollution. For the same reason, coefficients to the two southerly winds
457 had more positive signs than the two northerly winds, which was especially the case for
458 Beijing, Dezhou, and Shijiazhuang. The three cities had more sources of emission to their
459 south. Furthermore, from Table 1 (a) & (b), we can find that except for the integrated
460 wind speeds, nearly all the regression coefficients of the meteorological variables had stable
461 signs between the two winter heating seasons, which implied that those variables' im-
462 pacts on the air pollution were largely stable over time. Figures S7 and S8 in SI display
463 the pairwise regressions between $PM_{2.5}$ and the meteorological variables without the lagged
464 $PM_{2.5}$ term, whose coefficients largely had the same signs as those in Table 1 with the
465 lagged term.

466 5 Meteorological Change's Effects on $PM_{2.5}$

467 We evaluate the effects of the meteorological change between the two 19-year pe-
468 riods from 1979 to 2016 on $PM_{2.5}$ under three-yearly emission regimes from 2014 to 2016,
469 respectively. This is carried out by a two-step procedure. It first builds the regression
470 function between $PM_{2.5}$ and the meteorological variables via Model (6) for each season
471 (not just the heating season) of the three years from 2014 to 2016, respectively. Then,
472 the monthly averages of the meteorological variables in the two 19-year periods are sub-
473 stituted, respectively, to the estimated regression functions to attain a pair of the fitted
474 values, and their differences at each month of 2014 to 2016. The differences are the ef-
475 fects of the meteorological change that happened between 1979 and 2016 on $PM_{2.5}$ un-
476 der the emission regimes from 2014 to 2016, respectively.

477 It is also noted that the three years' $PM_{2.5}$ and meteorological data from 2014 to
478 2016 were for establishing the relationship between $PM_{2.5}$ and the meteorological vari-
479 ables, rather than the meteorological changes as the latter were assessed based on 38 year's
480 ECMWF data from 1979 to 2016.

481 The rationale of our method has similarity to other studies (H. Wang et al., 2015;
482 Shen et al., 2018; Leung et al., 2018) for the climate change effects on air pollution, which
483 are usually consists of four steps: (i) meteorology and PM correlation/regression mod-
484 eling, (ii) scenarios of future climate change, (iii) scenarios of emission, (iv) obtaining
485 the climate change effects on air pollution by substituting the climate change and emis-
486 sion scenarios under (ii) and (iii) to the established model in (i). The analogues of our
487 approach to the existing approaches are in (i) we use the three year's data (2014-2016)
488 to obtain the meteorology to pollution's models; in (ii) we use the meteorological changes
489 observed between the two 19-year periods from the 38-year's ECMWF reanalysis data
490 as the meteorological change scenario. However, the regression models built in (i) also
491 served as the emission scenarios for 2014-2016 in (iii) to attain the study objective. There-
492 fore, in (iv), we only substitute the meteorological changes attained in (ii) to the regres-
493 sion model in (i).

494 Comparing with the regression analysis in the last section, the analysis is extended
495 beyond the six cities and the heating seasons to the 32 cities that we have air quality data
496 and the entire year, respectively. Since most air quality sites are located in the center
497 of a city, we match each city to one grid which is nearest to the city center, and use the
498 grid's reanalysis data as the meteorological variables in the regression analysis for the
499 city.

500 The regression model (6) that measures the meteorological effect on $\text{PM}_{2.5}$ is estimated on four seasons: (i) Heating season ($q = 1$): November 15th last year to March 501 15th, (ii) Spring ($q = 2$): March 16th to May 31st; (iii) Summer ($q = 3$): June 1st to 502 August 31st and (iv) Fall ($q = 4$): September 1st to November 14th. Having the regression 503 model built over the four seasons instead of each month is to allow enough data since 504 we work on six-hourly instead of hourly observations. 505

506 The regression model (6) is re-expressed as

$$Y_{aqt}(c, s) = \alpha_{aq}(c) + \beta_{aq}^T(c)X_{aqt}(c) + \gamma_{aq}(c)Y_{aq(t-1)}(c, s) + \eta_{aq}^T(c)I_t + \epsilon_{aqt}(c, s), \quad (8)$$

507 where $a = 2014, 2015$ and 2016 for the three years, $q = 1, 2, 3$ and 4 for the four seasons, 508 and t is still the 6-hourly index, and s denotes a site in city c . The model is fitted 509 using the same approach as that for model (6).

Let $\bar{X}_{1j}(c)$ and $\bar{X}_{2j}(c)$ be the empirical averages of the meteorological variables for month j at a city c in the two 19-years periods, respectively. The estimated effects of meteorological change on $\log(\text{PM}_{2.5})$ at month j of year a are

$$\hat{\beta}_{aq}^T(c)(\bar{X}_{2j}(c) - \bar{X}_{1j}(c))/\{1 - \hat{\gamma}_{aq}(c)\}, \quad (9)$$

510 where $\hat{\beta}_{aq}(c)$ and $\hat{\gamma}_{aq}(c)$ are the estimated regression coefficients for the q -th quarter that 511 covers month j . The division of $(1 - \hat{\gamma}_{aq}(c))$ is due to the iterative lagged effect by as- 512 signing $L = \infty$ in (7). Hence, the meteorological effect is accumulated like a geomet- 513 ric progression which results in the factor $(1 - \hat{\gamma}_{aq}(c))^{-1}$ in (9). It is noted that as the 514 effect is measured for each year from 2014 to 2016, as there is no need to pool with other 515 year's data. The uncertainty associated with the relative small sample size is alleviated 516 as we merge the multiple air quality monitoring sites in each city, which allows more data 517 in building the yearly regression model in the paper. It is noted that the air quality sites 518 in a city are close to each other relative to the size of the grid boxes for the ECMWF 519 reanalysis data, which supports the site merge.

520 We transform the meteorological effects on $\log \text{PM}_{2.5}$ back to $\text{PM}_{2.5}$ by the expo- 521 nential transform to attain $\hat{\nu}_{2aj}(c) - \hat{\nu}_{1aj}(c)$ where $\hat{\nu}_{iaj}(c) = \exp\{(\hat{\alpha}_{aq}^T(c) + \hat{\beta}_{aq}^T(c)\bar{X}_{ij}(c) + 522 \hat{\eta}_{aq}^T(c)\bar{I})/(1 - \hat{\gamma}_{aq}(c))\}$ is the average $\text{PM}_{2.5}$ (in $\mu\text{g}/\text{m}^3$) at month j of year a in a city 523 c , and $\bar{I} = (0.25, 0.25, 0.25)^T$ is the average indicator values over the two periods. The 524 standard errors can be obtained as outlined in Appendix A.2. For March and November, 525 we average the effects for $q = 1$ and 2 , and $q = 1$ and 4 , respectively, since March 526 (November) falls to heating ($q = 4$) and the spring (fall) seasons, respectively.

527 Figure 4 presents the monthly meteorological change effects on $\text{PM}_{2.5}$ for each city, 528 which displays significant decreases in $\text{PM}_{2.5}$ in January, March, and December in all 529 three years from 2014 to 2016. November was the only month that had increased $\text{PM}_{2.5}$ 530 in some cities.

531 Figure 5 aggregates the results in Figure 4 by providing the raw and relative me- 532 teorological change effects for each month on $\text{PM}_{2.5}$. It shows that the overall effects of 533 the meteorological change in the 32 cities were largely consistent among the three years. 534 Figure 5 (a) shows that the percentages of cities that experienced unfavorable effects (in- 535 creased $\text{PM}_{2.5}$) were far less than those with favorable effects (reducing $\text{PM}_{2.5}$). In par- 536 ticular, the percentages of favorable effects were all above 47% in January, March and 537 December across the three years. June was the month that had the largest unfavorable 538 percentage changes at 11.6% on average, followed by November. The confidence inter- 539 vals given in panel (b) shows that the unfavorable effects in June and November were 540 not significant at the 5% level. In contrast, the favorable effects in January, March, and 541 December implicated in Figure 4 were confirmed at 5% level of significance. There was 542 evidence of the favorable effects of the meteorological changes on $\text{PM}_{2.5}$ in the five months

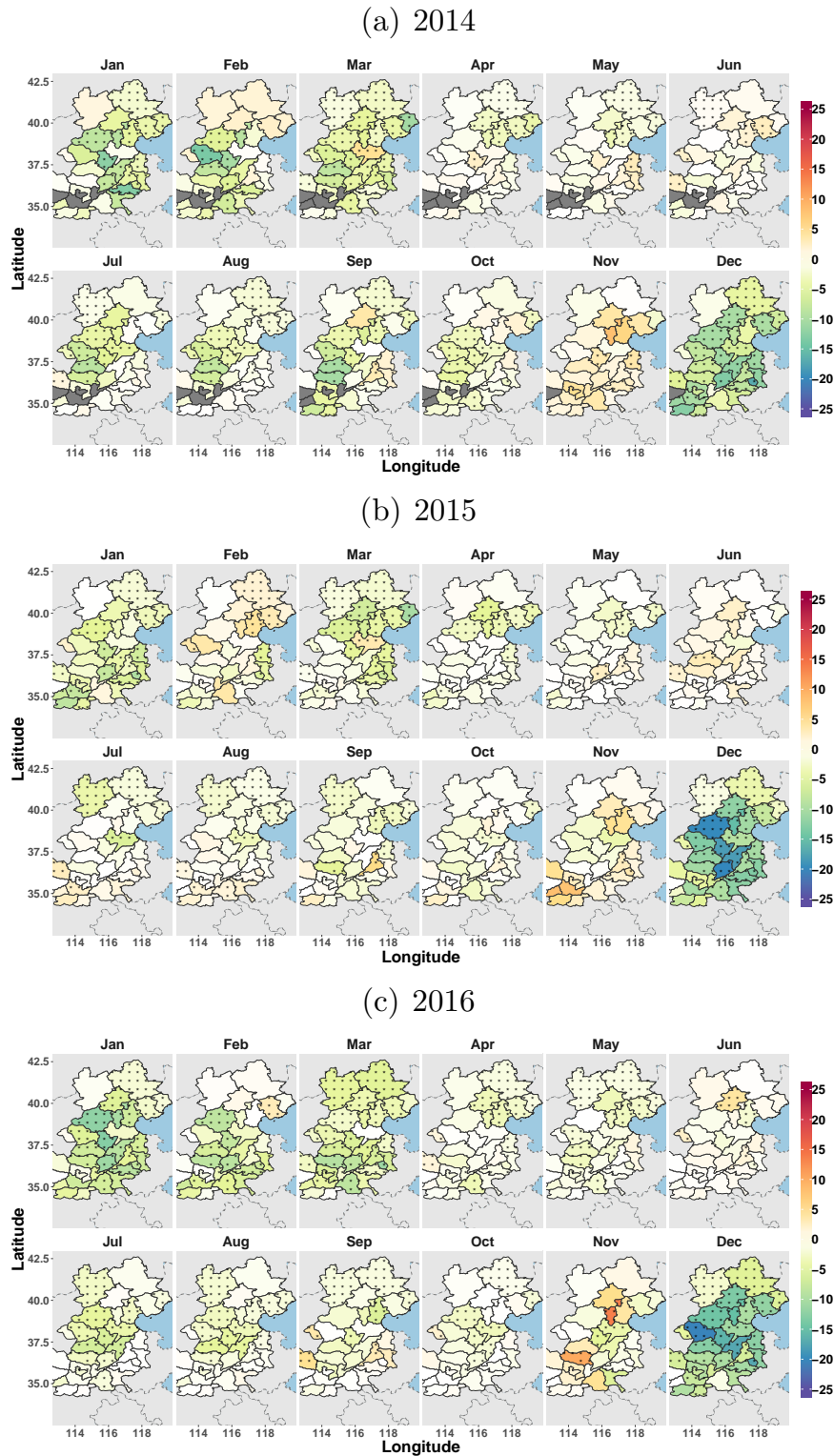


Figure 4. Estimated meteorological change effects on $PM_{2.5}$ (in $\mu g/m^3$) for the 32 cities. The effects are shown for the twelve months between years of 2014 and 2016. We mark all the grids with stars for cities with significant changes (at 5% level). (In 2014, some cities are missing because of the high data missing rate in the sites.)

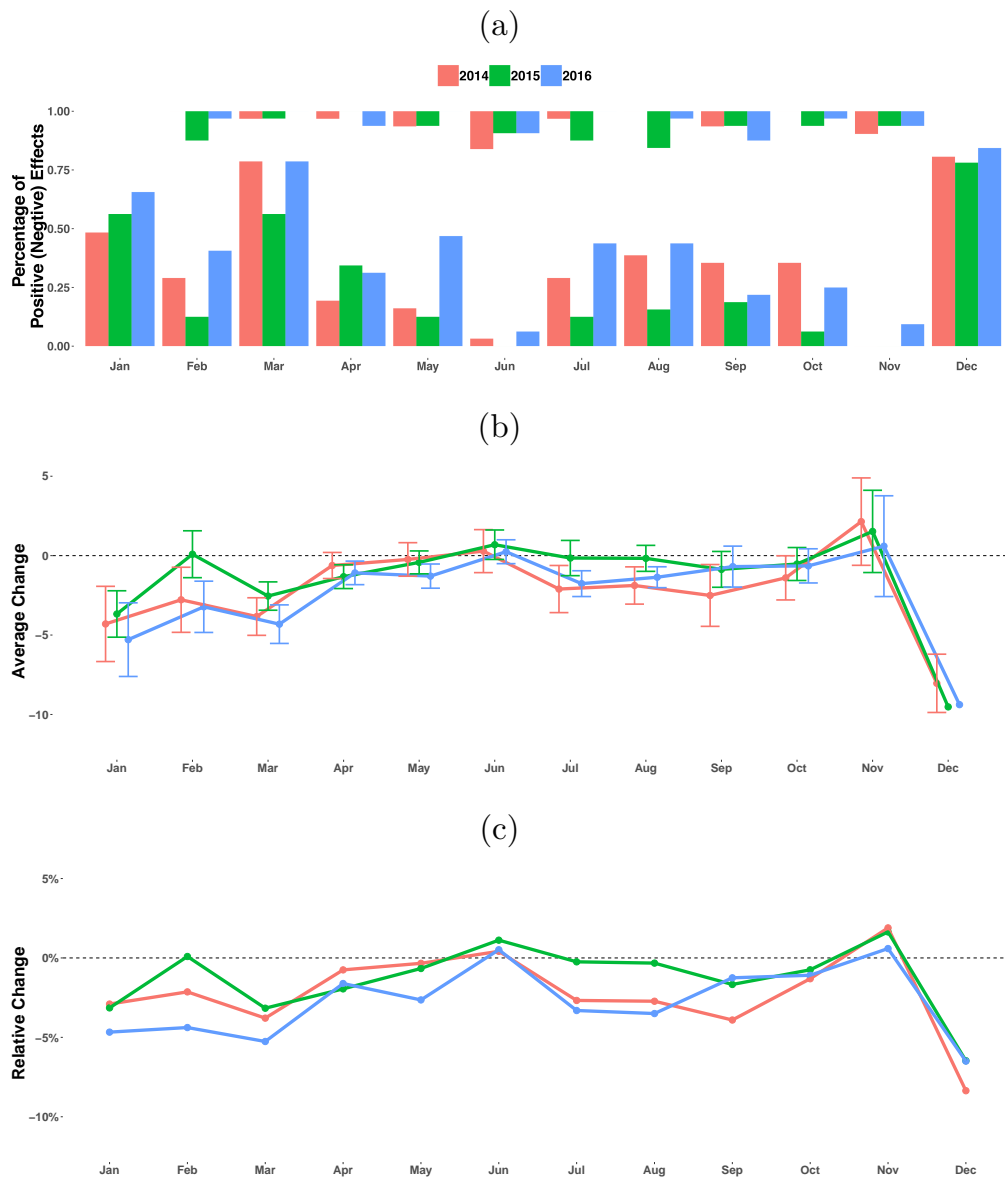


Figure 5. Percentages of cities that had significant meteorological change effects (a), and the absolute (b) and relative (c) (related to original average) average PM_{2.5} changes (in $\mu\text{g}/\text{m}^3$) in the 32 cities in the three years. In (a), bars from bottom (top) indicate percentages of the cities which had significant favorable (unfavorable) meteorological change effects on PM_{2.5}; (b), the average meteorological change effects and the 95% confidence intervals over all 32 cities; (c), relative meteorological change effects over the 32 cities.

of April, July, August, September, and October, where the overall estimated regional effects were negative. Among the 15 month-year combinations, there were 8 of them being statistically significant at 5% level and the other 7 were not significant. But, still, we can see the quite strong evidence of the favorable effects of the meteorological changes.

The months of January, March, and December experienced the most favorable and significant meteorological effects with the average $PM_{2.5}$ reduced between $2.55 \sim 9.52 \mu g/m^3$. The amount of reduction in January and December, the two worst months for air quality, amounted to about 3.57% and 7.10% of the average $PM_{2.5}$, respectively and were highly significant. Despite the meteorological change effect were significant statistically, the average annual percentages of reduction for 2014-2016's $PM_{2.5}$ ranged from 1.29% to 2.76% that mounted to $1.41 \sim 2.35 \mu g/m^3$ reduction per year. Thus, they accounted for only a small percentage of $PM_{2.5}$ concentration as shown in Figure 5 (c), which means that the alleviating effects of the meteorological change were rather limited and could not be relied upon for solving the air pollution problem in the region.

6 Discussion

The observed percentage range of effects due to the meteorological change was consistent with some existing studies. Tagaris et al. (2007) projected continental US's annual average $PM_{2.5}$ concentration to decline 10% in 2050 vs the 2000 level, and Avise et al. (2009) projected the US July average $PM_{2.5}$ to be down $-1 \mu g/m^3$ over the same time frame as Tagaris et al. Our results showing some limited amount of meteorological "benefit" on $PM_{2.5}$ levels in 2014-2016 were comparable with the results of Shen et al. (2018); Leung et al. (2018) and Pendergrass et al. (2019) on the climate change effects on future Beijing or eastern China's $PM_{2.5}$ level under various climate change and emission scenarios.

Our study differs from existing studies on the meteorological change's effects on air pollution in several aspects. One is that our study is based on the historical data without a need for providing explicit scenarios in emission and meteorological changes. Our study implicitly utilized emission information in the years from 2014 to 2016, and explicitly the meteorological change information from 1979 to 2016. The second difference is that our study region is much smaller spatially than some of the studies which allows us to better control hidden spatial effects shared in the region.

Our analysis reveals that the region in North China during the 38 years between 1979 and 2016 had witnessed significant changes in several meteorological variables, which include an increased surface temperature by 0.57 degree, the boundary layer height and dissipation, and the vertically integrated NE wind speed; reduced RH and vertically integrated SW wind speed. Overall speaking, the region has become warmer and dryer. The warming has led to increases in both boundary layer height and dissipation, which are the two main drivers for the favorable effect on $PM_{2.5}$. Our analysis did not confirm the widespread speculation of reduced northerly wind. In contrast, the NE wind had increased while NW wind had no sustained trend. Hence, the speculation was very much caused by increased public anxiety on air pollution as people's expectation for its improvement had increased.

The overall effects of the meteorological change had contributed to net reductions of $PM_{2.5}$ in many months of the three years after controlling for emission at each grid and month in the study region. It is noted that despite the warming, dew point temperature had been largely kept under control as the region had got dryer. This made the effects of temperature warming largely favorable to $PM_{2.5}$ as it enhances boundary layer height and dissipation. A conclusion from the analysis is that the meteorological change in North China had helped alleviate the air pollution, and was certainly not responsible for the regional air pollution problem.

Acknowledgments

The research was partially supported by China's National Key Research Special Program Grant no. 2016YFC0207701, National Natural Science Foundation of China key Grant 71532001, Center for Statistical Science and LMEQF at Peking University. The ERA-Interim reanalysis data used in this study were collected from the ECMWF website <https://apps.ecmwf.int/>. The MERRA2 dataset can be accessible at <https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/data.access/>. We archived our pollution data in the repository <https://archive.ics.uci.edu/ml/datasets/Beijing+Multi-Site+Air-Quality+Data>. A temporary copy of the data was in SI.

References

- Avise, J., Chen, J., Lamb, B., Wiedinmyer, C., Guenther, A., Salathé, E., & Mass, C. (2009). Attribution of projected changes in summertime us ozone and PM 2.5 concentrations to global changes. *Atmospheric Chemistry and Physics*, 9(4), 1111–1124. doi: <https://doi.org/10.5194/acp-9-1111-2009>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B*, 57(1), 289–300.
- Bojkov, R., Bishop, L., Hill, W., Reinsel, G., & Tiao, G. (1990). A statistical trend analysis of revised Dobson total ozone data over the northern hemisphere. *Journal of Geophysical Research: Atmospheres*, 95(D7), 9785–9807. doi: <https://doi.org/10.1029/JD095iD07p09785>
- Bosq, D. (1998). *Nonparametric statistics for stochastic processes: estimation and prediction* (Vol. 110). New York, NY: Springer Science.
- Brockwell, P. J., & Davis, R. A. (2013). *Time series: Theory and methods*. New York, NY: Springer Science & Business Media. doi: <https://doi.org/10.1007/978-1-4419-0320-4>
- Bühlmann, P. (2002). Bootstraps for time series. *Statistical Science*, 17(1), 52–72.
- Cai, W., Li, K., Liao, H., Wang, H., & Wu, L. (2017). Weather conditions conducive to Beijing severe haze more frequent under climate change. *Nature Climate Change*, 7(4), 257.
- Chen, H., & Wang, H. (2015). Haze days in North China and the associated atmospheric circulations based on daily visibility data from 1960 to 2012. *Journal of Geophysical Research: Atmospheres*, 120(12), 5895–5909. doi: <https://doi.org/10.1002/2015JD023225>
- Chen, L., Guo, B., Huang, J., He, J., Wang, H., Zhang, S., & Chen, S. X. (2018). Assessing air-quality in beijing-tianjin-hebei region: The method and mixed tales of pm2. 5 and o3. *Atmospheric environment*, 193, 290–301.
- Chen, R., Zhao, Z., & Kan, H. (2013). Heavy smog and hospital visits in beijing, china. *American journal of respiratory and critical care medicine*, 188(9), 1170–1171.
- Chen, S. X., & Tang, C. (2005). Nonparametric inference of value-at-risk for dependent financial returns. *Journal of Financial Econometrics*, 3(2), 227–255. doi: <https://doi.org/10.1093/jjfinec/nbi012>
- Cheng, M. (2016). Beijing haze is rooted in wind attenuation (in Chinese). *China ScienceNet*. Retrieved from <http://news.sciencenet.cn>
- Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., . . . Bauer, P. (2011). The era-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. doi: <https://doi.org/10.1002/qj.828>
- Donaldson, K., Li, X., & MacNee, W. (1998). Ultrafine (nanometre) particle mediated lung injury. *Journal of aerosol science*, 29(5-6), 553–560.
- Egan, B. A., & Mahoney, J. R. (1972). Numerical modeling of advection and diffusion of urban area source pollutants. *Journal of Applied Meteorology*, 11(2),

- 646 312–322. doi: [https://doi.org/10.1175/1520-0450\(1972\)011<0312:NMOAAD>2.0](https://doi.org/10.1175/1520-0450(1972)011<0312:NMOAAD>2.0)
 647 .CO;2
- 648 Fan, J., & Yao, Q. (2008). *Nonlinear time series: Nonparametric and parametric*
 649 *methods*. New York, NY: Springer-Verlag.
- 650 Holzworth, G. C. (1972). Mixing heights, wind speeds, and potential for urban
 651 air pollution throughout the contiguous United States. In *United states envi-*
 652 *ronmental protection agency publication* (Vol. 101, p. 118). Research Triangle
 653 Park, NC: Environmental Protection Agency.
- 654 Jacob, D. J., & Winner, D. A. (2009). Effect of climate change on air quality. *Atmo-*
 655 *spheric Environment*, *43*(1), 51–63. doi: <https://doi.org/10.1016/j.atmosenv>
 656 .2008.09.051
- 657 Küünsch, H. R. (1989). The jackknife and the bootstrap for general stationary ob-
 658 servations. *The Annals of Statistics*, *17*(3), 1217–1241.
- 659 Leung, D. M., Tai, A. P., Mickley, L. J., Moch, J. M., Donkelaar, A. v., Shen, L., &
 660 Martin, R. V. (2018). Synoptic meteorological modes of variability for fine
 661 particulate matter (pm 2.5) air quality in major metropolitan regions of china.
 662 *Atmospheric Chemistry and Physics*, *18*(9), 6733–6748.
- 663 Liang, X., Li, S., Zhang, S., Huang, H., & Chen, S. X. (2016). PM2.5 data reli-
 664 ability, consistency, and air quality assessment in five Chinese cities. *Journal of*
 665 *Geophysical Research: Atmospheres*, *121*, 10,220–10,236. doi: [https://doi.org/](https://doi.org/10.1002/2016JD024877)
 666 [10.1002/2016JD024877](https://doi.org/10.1002/2016JD024877)
- 667 Liang, X., Zou, T., Guo, B., Li, S., Zhang, H., Zhang, S., ... Chen, S. X. (2015).
 668 Assessing Beijing’s PM2.5 pollution: severity, weather impact, APEC and
 669 winter heating. *Proceedings of the Royal Society A*, *471*, 20150257. doi:
 670 <https://doi.org/10.1098/rspa.2015.0257>
- 671 Long, X., Bei, N., Wu, J., Li, X., Tian, F., Li, X., ... An, Z. (2018). Does afforestation
 672 deteriorate haze pollution in Beijing–Tianjin–Hebei (BTH), China? *Atmo-*
 673 *spheric Chemistry and Physics*, *18*(15), 10869. doi: [https://doi.org/10.5194/](https://doi.org/10.5194/acp-18-10869-2018)
 674 [acp-18-10869-2018](https://doi.org/10.5194/acp-18-10869-2018)
- 675 Miao, Y., Hu, X., Liu, S., Qian, T., Xue, M., Zheng, Y., & Wang, S. (2015). Sea-
 676 sonal variation of local atmospheric circulations and boundary layer struc-
 677 ture in the Beijing-Tianjin-Hebei region and implications for air quality.
 678 *Journal of Advances in Modeling Earth Systems*, *7*(4), 1602–1626. doi:
 679 <https://doi.org/10.1002/2015MS000522>
- 680 Pendergrass, D., Shen, L., Jacob, D., & Mickley, L. (2019). Predicting the impact
 681 of climate change on severe wintertime particulate pollution events in beijing
 682 using extreme value theory. *Geophysical Research Letters*, *46*(3), 1824–1830.
- 683 Pope III, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K.,
 684 & Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and
 685 long-term exposure to fine particulate air pollution. *Jama*, *287*(9), 1132–1141.
- 686 Reinsel, G. C., Miller, A. J., Weatherhead, E. C., Flynn, L. E., Nagatani, R. M.,
 687 Tiao, G. C., & Wuebbles, D. J. (2005). Trend analysis of total ozone data for
 688 turnaround and dynamical contributions. *Journal of Geophysical Research:*
 689 *Atmospheres*, *110*(D16), 306. doi: <https://doi.org/10.1029/2004JD004662>
- 690 Schwartz, J. (2000). The distributed lag between air pollution and daily deaths. *Epi-*
 691 *demiology*, *11*(3), 320–326.
- 692 Shen, L., Jacob, D. J., Mickley, L. J., Wang, Y., & Zhang, Q. (2018). Insignifi-
 693 cant effect of climate change on winter haze pollution in beijing. *Atmospheric*
 694 *Chemistry and Physics*, *18*(23), 17489–17496.
- 695 Tagaris, E., Manomaiphibon, K., Liao, K.-J., Leung, L. R., Woo, J.-H., He, S.,
 696 ... Russell, A. G. (2007). Impacts of global climate change and emissions
 697 on regional ozone and fine particulate matter concentrations over the United
 698 States. *Journal of Geophysical Research: Atmospheres*, *112*(D14), 312. doi:
 699 <https://doi.org/10.1029/2006JD008262>
- 700 Tai, A. P., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine

- 701 particulate matter (PM_{2.5}) and meteorological variables in the United
 702 States: Implications for the sensitivity of PM_{2.5} to climate change. *At-*
 703 *mospheric Environment*, 44(32), 3976–3984. doi: [https://doi.org/10.1016/](https://doi.org/10.1016/j.atmosenv.2010.06.060)
 704 [j.atmosenv.2010.06.060](https://doi.org/10.1016/j.atmosenv.2010.06.060)
- 705 Tai, A. P., Mickley, L. J., Jacob, D. J., Leibensperger, E., Zhang, L., Fisher, J. A., &
 706 Pye, H. (2012). Meteorological modes of variability for fine particulate matter
 707 (PM_{2.5}) air quality in the United States: implications for PM_{2.5} sensitivity to
 708 climate change. *Atmospheric Chemistry and Physics*, 12(6), 3131–3145. doi:
 709 <https://doi.org/10.5194/acp-12-3131-2012>
- 710 Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and
 711 the experiment design. *Bulletin of the American Meteorological Society*, 93(4),
 712 485–498. doi: <https://doi.org/10.1175/BAMS-D-11-00094.1>
- 713 Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., ...
 714 Sulprizio, M. P. (2016). Why do models overestimate surface ozone in the
 715 southeast United States? *Atmospheric Chemistry and Physics*, 16(21), 13561–
 716 13577. doi: <https://doi.org/10.5194/acp-16-13561-2016>
- 717 Wang, H., Chen, H., & Liu, J. (2015). Arctic sea ice decline intensified haze pollu-
 718 tion in eastern china. *Atmospheric and Oceanic Science Letters*, 8(1), 1–9.
- 719 Wang, X., Wu, J., Chen, M., Xu, X., Wang, Z., Wang, B., ... Miao, G. (2018).
 720 Field evidences for the positive effects of aerosols on tree growth. *Global*
 721 *Change Biology*, 24(10), 4983–4992. doi: <https://doi.org/10.1111/gcb.14339>
- 722 Wuerch, D., Courtois, A., Ewald, C., & Ernct, G. (1972). A preliminary transport
 723 wind and mixing height climatology for St. Louis, Missouri. *National Oceanic*
 724 *and Atmospheric Administration Technical Memorandum National Weather*
 725 *Service Central Region*, 49, 13.
- 726 Xu, P., Chen, Y., & Ye, X. (2013). Haze, air pollution, and health in China. *The*
 727 *Lancet*, 382(9910), 2067. doi: [https://doi.org/10.1016/S0140-6736\(13\)62693-8](https://doi.org/10.1016/S0140-6736(13)62693-8)
- 728 Yan, L. (2015). Influence of wind power development on the heavy haze in Beijing,
 729 Tianjin and Hebei (in Chinese). *Environmental Protection and Circular Econ-*
 730 *omy*(2015.09), 67–71.
- 731 Yin, Z., & Wang, H. (2016). The relationship between the subtropical western pa-
 732 cific sst and haze over north-central north china plain. *International Journal of*
 733 *Climatology*, 36(10), 3479–3491.
- 734 Yin, Z., & Wang, H. (2017). Role of atmospheric circulations in haze pollution in de-
 735 cember 2016. *Atmospheric Chemistry and Physics*, 17(18), 11673.
- 736 Zhang, S., Guo, B., Dong, A., He, J., Xu, Z., & Chen, S. X. (2017). Cautionary
 737 tales on air-quality improvement in Beijing. *Proceedings of the Royal Society*
 738 *A: Mathematical, Physical and Engineering Sciences*, 473, 20170457. doi:
 739 <https://dx.doi.org/10.6084/m9.figshare.c.3865483>
- 740 Zheng, G., Duan, F., Su, H., Ma, Y., Cheng, Y., Zheng, B., ... Chang, D. (2015).
 741 Exploring the severe winter haze in Beijing: the impact of synoptic weather,
 742 regional transport and heterogeneous reactions. *Atmospheric Chemistry and*
 743 *Physics*, 15(6), 2969–2983. doi: <https://doi.org/10.1175/2010JCLI3850.1>

744 Appendix A Methods

745 A1 Variance Estimation for Average Climate Change.

746 Let $Y_{ajt}(s) = X_{2ajt}(s) - X_{1ajt}(s)$ be the difference of meteorological observations
 747 between two period in a -th year at month j and hour t at a grid s . To simplify the ex-
 748 pedition, as the same procedures will be replicated for each year, month and grid, we will
 749 hide the subscripts a , j and s in $Y_{ajt}(s)$ and simply write it as Y_t .

750 Let μ be the mean of Y_t . An estimator of μ is $\hat{\mu} = n^{-1} \sum_{t=1}^n Y_t$, where n is the
 751 number of observations at the month. The variance is $\text{Var}(\hat{\mu}) = n^{-1} \sigma^2(n)$, where $\sigma^2(n) =$

752 $\gamma(0) + 2 \sum_{k=1}^{n-1} (1 - k/n) \gamma(k)$ and $\gamma(k) = \text{cov}(Y_t, Y_{t+k})$ for positive integers k . A com-
 753 mon approach for variance estimation in time series is the spectral approach.

754 Let $\phi(\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma(k) \exp(-ik\lambda)$ for $\lambda \in [-\pi, \pi]$ be the spectral den-
 755 sity of $\{Y_t\}_{t=1}^n$. According to (Brockwell & Davis, 2013) (Corollary 4.3.2), we can esti-
 756 mate $\sigma^2(n)$ with estimating $\phi(0)$.

Let $I_n(\omega_j) = n^{-1} |\sum_{l=1}^n Y_l e^{-il\omega_j}|^2, j = 0, \pm 1, \dots, \pm[n/2]$ be the sample periodogram. According to Theorem 5.2.6 of (Brockwell & Davis, 2013) for any $j \in T$

$$I_n(\omega_j) = (2\pi)\phi(\omega_j)E_j + R_j,$$

757 where $\{E_j\}_{j \in T}$ are independent standard exponential random variables and $\{R_j\}_{j \in T}$ are
 758 asymptotically negligible terms.

759 Take the logarithm on both side of equation above and ignore $\{R_j\}$

$$760 \log\{I_n(\omega_j)/2\pi\} = \log\{\phi(\omega_j)\} + \log(E_j) \quad j \in T. \quad (\text{A1})$$

Note that $E\{\log(E_j)\} = -0.5$, let $\eta_j = \log(E_j) + 0.57721$, $W_j = \log\{I_n(\omega_j)/2\pi\} + 0.57721$. The equation (A1) can be approximated by the following nonparametric regres-
 sion:

$$W_j = m(\omega_j) + \eta_j \quad j \in T.$$

The Nadaraya-Waston estimator of $m(\omega)$ based on a kernel K and a smoothing band-
 width b is

$$\hat{m}_b(\omega) = \frac{\sum_{j \in T} K_1(\frac{\omega - \omega_j}{b}) W_j}{\sum_{j \in T} K_1(\frac{\omega - \omega_j}{b})}.$$

761 Hence, the kernel estimator of $\phi(0)$ is $\hat{\phi}(0) = \exp\{\hat{m}_b(0)\}$. The standard deviation of
 762 $\hat{\mu}$ is $2\pi\hat{\phi}(0)/n$. See (S. X. Chen & Tang, 2005) for more details.

763 A2 Average Climate Condition of a City

We choose the nearest grid to the average latitude and longitude of all the air qual-
 ity sites in a city c to represent average meteorological condition of the city. Let $\mu_{1j}(c)$
 and $\mu_{2j}(c)$ be the means of the nine dimensional vector of the meteorological variables
 for each month (j) for city c in the two 19-year periods, respectively. They are estimated
 by

$$\hat{\mu}_{lj}(c) = \frac{1}{19n_j} \sum_{a=1}^{19} \sum_{t=1}^{n_j} X_{iajt}(c) \quad \text{for } i = 1, 2.$$

764 According to assumption (1),

$$\begin{aligned} \hat{\mu}_{1j}(c) &= \mu_j^F(c) + \frac{1}{19n_j} \sum_{a=1}^{19} \sum_{t=1}^{n_j} \epsilon_{1ajt}(c) \\ \hat{\mu}_{2j}(c) &= \mu_j^F(c) + \mu_j^C(c) + \frac{1}{19n_j} \sum_{a=1}^{19} \sum_{t=1}^{n_j} \epsilon_{2ajt}(c). \end{aligned}$$

765 A3 Average PM_{2.5} and Its Dependence on Climate Condition

766 We represent the expectation of PM_{2.5} at a city c as a function of the average me-
 767 teorological condition μ_X of the city, which is a 9-dimensional vector, via the linear model
 768 (8). We will substitute μ_X with either $\mu_{1j}(c)$ and $\mu_{2j}(c)$ to gauge on the effects of the
 769 meteorological change on PM_{2.5}

770 Recall that $Y_{aqt}(c, s)$ in model (8) be the log PM_{2.5} level at a monitoring site s at
 771 city c in a quarter p that covers or intercept with month j and year a . Then, $Z_{aqt}(c, s) =$

772 $\exp(Y_{aqt}(c, s))$ be $\text{PM}_{2.5}$ in the original scale. Let the average of the meteorological vari-
 773 able $X_{aqt}(c)$ be a generic μ_X , and $\mu_{Y_{aq}}(c, \mu_X)$ be the expectation of $Y_{aqt}(c, s)$.

774 Then, with $\bar{I} := (0.25, 0.25, 0.25)^T$,

$$\begin{aligned}\mu_{Y_{aq}}(c, \mu_X) &= \mathbf{E}[\alpha_{aq}(c) + \beta_{aq}^T(c)X_{aqt}(c) + \eta_{aq}^T(c)I_t + Y_{aq(t-1)}(c, s)\gamma_{aq}(c)] \\ &= \mathbf{E}\left[\sum_{i=0}^{\infty}\{\alpha_{aq}(c) + \beta_{aq}^T(c)X_{aq(t-i)}(c) + \eta_{aq}^T(c)I_t\}\gamma_{aq}^i(c)\right] \\ &= \frac{1}{1 - \gamma_{aq}(c)}\{\alpha_{aq}(c) + \beta_{aq}(c)^T\mu_X + \eta_{aq}^T(c)\bar{I}\},\end{aligned}$$

775 which is depends on the average meteorological condition μ_X .

Let $\tilde{\alpha}_{aq}(c) = \alpha_{aq}(c)/(1 - \gamma_{aq}(c))$, $\tilde{\beta}_{aq}(c) = \beta_{aq}(c)/(1 - \gamma_{aq}(c))$ and $\tilde{\eta}_{aq}(c) = \eta_{aq}(c)/(1 - \gamma_{aq}(c))$. Then, an estimator for $\mu_{Y_{aq}}(c, \mu_X)$ is

$$\hat{\mu}_{Y_{aq}}(c, \mu_X) = \hat{\alpha}_{aq}(c) + \hat{\beta}_{aq}^T(c)\mu_X + \hat{\eta}_{aq}^T(c)\bar{I},$$

776 where $\hat{\alpha}_{aq}(c)$, $\hat{\eta}_{aq}(c)$ and $\hat{\beta}_{aq}(c)$ are estimators by plugging-in the least square regression
 777 estimators to $\tilde{\alpha}_{aq}(c)$, $\tilde{\beta}_{aq}(c)$ and $\gamma_{aq}(c)$ in model (8). Since $\hat{\eta}_{aq}^T(c)\bar{I}$ is independent from
 778 μ_X as it is for $\tilde{\alpha}_{aq}(c)$. We combine them together by ignoring $\hat{\eta}_{aq}^T(c)\bar{I}$ in the following
 779 analysis.

780 Put $\mu_{Z_{aq}}(c, \mu_X) = E\{Z_{aqt}(c, s)\} = E\{\exp(\hat{\mu}_{Y_{aq}}(c, s))\}$ be the average $\text{PM}_{2.5}$ in
 781 the original scale. An initial estimator of $\mu_{Z_{aq}}(c, \mu_X)$ is $\hat{\mu}_{Z_{aq}}(c, \mu_X) = \exp(\hat{\mu}_{Y_{aq}}(c, \mu_X))$.

782 A4 Meteorological Change Effects Estimation

783 We define meteorological change effect in month j and year a at city c as

$$784 \widehat{\Delta Z}_{aj}(c) = \hat{\mu}_{Z_{aq}}(c, \hat{\mu}_{2j}(c)) - \hat{\mu}_{Z_{aq}}(c, \hat{\mu}_{1j}(c)), \quad (\text{A2})$$

785 which is the difference in the expectations of raw $\text{PM}_{2.5}$ under the average meteorolog-
 786 ical condition between the two time periods. Here, p is the time segment that month j
 787 falls into or intercept with. The latter case includes March and November, which we will
 788 use the average of the two adjacent time segments to measure the meteorological change
 789 effect.

790 We carry a bias correction to the estimator in (A2) to offset the non-linearity in
 791 the exponential transformation. Let $\hat{\xi}_{aj}(c) = (\hat{\mu}_{1j}^T(c), \hat{\mu}_{2j}^T(c), \hat{\alpha}_{aq}(c), \hat{\beta}_{aq}^T(c))^T$, $\xi_{aj}(c) =$
 792 $E(\hat{\xi}(c)) = (\mu_j^{FT}(c), \mu_j^{FT}(c) + \mu_j^{CT}(c), \tilde{\alpha}_{aq}(c), \tilde{\beta}_{aq}^T(c))^T$. Let $h_1(x) = e^{x_3 + x_4^T x_1}$, $h_2(x) =$
 793 $e^{x_3 + x_4^T x_2}$ and $h(x) = h_2(x) - h_1(x)$. Then, $h(\hat{\xi}_{aj}(c)) = \widehat{\Delta Z}_{aj}(c)$.

For $\widehat{\Delta Z}_{aj}(c)$, a Taylor expansion near $\xi_{aj}(c)$ gives

$$\widehat{\Delta Z}_{aj}(c) \approx h(\xi_{aj}(c)) + \nabla h(\xi_{aj}(c))^T(\hat{\xi}_{aj}(c) - \xi_{aj}(c)) + \frac{1}{2}(\hat{\xi}_{aj}(c) - \xi_{aj}(c))^T \nabla^2 h(\xi_{aj}(c))(\hat{\xi}_{aj}(c) - \xi_{aj}(c)).$$

The expectation of $\widehat{\Delta Z}_{aj}(c)$ is approximately

$$E(\widehat{\Delta Z}_{aj}(c)) \approx h(\xi_{aj}(c)) - \frac{1}{2}\text{tr}(\nabla^2 h(\xi_{aj}(c))E((\hat{\xi}_{aj}(c) - \xi_{aj}(c))(\hat{\xi}_{aj}(c) - \xi_{aj}(c))^T)).$$

794 The bias corrected estimation of $\widehat{\Delta Z}_{aj}(c)$ is

$$795 \widehat{\Delta Z}_{aj}(c) = \hat{\mu}_{Z_{aq}}(c, \hat{\mu}_{2j}(c)) - \hat{\mu}_{Z_{aq}}(c, \hat{\mu}_{1j}(c)) + \frac{1}{2}\text{tr}[\nabla^2 h(\hat{\xi}_{aj}(c))\widehat{\text{Var}}(\hat{\xi}_{aj}(c))]. \quad (\text{A3})$$

The first and second order gradient of $h(x)$ are

$$\nabla h(x) = \begin{pmatrix} -x_3 h_1(x) \\ x_4 h_2(x) \\ h(x) \\ x_2 h_2(x) - x_2 h_1(x) \end{pmatrix} \quad \text{and}$$

$$\nabla^2 h(x) = \begin{pmatrix} -x_4 x_4^T h_1(x) & 0 & -x_4 h_1(x) & -(I + x_4 x_4^T) h_1(x) \\ 0 & x_4 x_4^T h_2(x) & x_4 h_2(x) & (I + x_4 x_4^T) h_2(x) \\ -x_4^T h_1(x) & x_4^T h_2(x) & h(x) & x_2^T h_2(x) - x_1^T h_1(x) \\ -x_1 x_4^T h_1(x) & x_2 x_4^T h_2(x) & x_2 h_2(x) - x_1 h_1(x) & x_2 x_2^T h_2(x) - x_1 x_1^T h_1(x) \end{pmatrix}.$$

For the variance of $\hat{\xi}_{aj}(c)$, it can be shown that

$$\text{Var}(\hat{\xi}_{aj}(c)) = \begin{pmatrix} \text{Var}(\hat{\mu}_{1j}(c)) & 0 & 0 \\ 0 & \text{Var}(\hat{\mu}_{2j}(c)) & 0 \\ 0 & 0 & \text{Var}((\hat{\alpha}_{aq}(c), \hat{\beta}_{aq}^T(c))) \end{pmatrix}.$$

796 Thus, its estimation can be attained by separately estimating the three diagonal covari-
797 ances, say $\widehat{\text{Var}}(\hat{\mu}_{1j}(c))$, $\widehat{\text{Var}}(\hat{\mu}_{2j}(c))$ and $\widehat{\text{Var}}((\hat{\alpha}_{aq}(c), \hat{\beta}_{aq}^T(c)))$.

798 Note that $\text{Var}(\hat{\mu}_{ij}(c)) = \frac{1}{19} \text{Var}(\frac{1}{n_j} \sum_{t=1}^{n_j} \epsilon_{iajt}(c))$, where n_j is the number of ob-
799 servations in month j of a year. Let $\tilde{\epsilon}_{iajt}(c) = X_{iajt}(c) - 18^{-1} \sum_{a' \neq a} X_{ia'jt}(c) = \epsilon_{iajt}(c) -$
800 $18^{-1} \sum_{a' \neq a} \epsilon_{ia'jt}(c)$ and $\tilde{\Sigma}_{ij}(c) = \text{Var}(\frac{1}{n_j} \sum_{t=1}^{n_j} \tilde{\epsilon}_{iajt}(c))$. Then $\tilde{\Sigma}_{ij}(c) = 19 \Sigma_{ij}(c)/18$,
801 where $\Sigma_{ij}(c) = \text{Var}(\frac{1}{n_j} \sum_{t=1}^{n_j} \epsilon_{iajt}(c))$. Let $\tilde{\Sigma}_{ij}(c)$ be the block bootstrap estimator of
802 $\tilde{\Sigma}_{ij}(c)$ (Küünsch, 1989). Then, $\widehat{\text{Var}}(\hat{\mu}_{ij}(c)) = 18 \tilde{\Sigma}_{ij}(c)/19^2$. The variance of $(\hat{\alpha}_{aq}(c), \hat{\beta}_{aq}(c))$
803 can be obtained via the sieve bootstrap (Bühlmann, 2002) for time series regression mod-
804 els.

805 A5 Estimating the Variance of Meteorological Change Effects

By ignoring the second order term in (A3), the variance of $\widehat{\Delta Z}_{aj}(c)$ is approximated by applying the delta method

$$\text{Var}(\widehat{\Delta Z}_{aj}(c)) \approx \nabla h(\hat{\xi}_{aj}(c))^T \text{Var}(\hat{\xi}_{aj}(c)) \nabla h(\hat{\xi}_{aj}(c)),$$

806 which can be estimated by plugging in $\nabla h(\hat{\xi}_{aj}(c))$ and $\widehat{\text{Var}}(\hat{\xi}_{aj}(c))$.

807 A6 Estimating Regional Average Meteorological Change Effect and its 808 Variance

The regional average meteorological change effect is given by

$$\hat{\eta}_{aj} = \frac{1}{32} \sum_{c=1}^{32} \widehat{\Delta Z}_{aj}(c).$$

809 Let $\hat{\xi}_{aj} = (\hat{\xi}_{aj}(1), \dots, \hat{\xi}_{aj}(32))^T$ and $g(x) = \frac{1}{32} \sum_{c=1}^{32} h(x_c)$. Then, $\hat{\eta}_{aj} = g(\hat{\xi}_{aj})$. Its
810 variance can be approximated by the delta method

$$\begin{aligned} \text{Var}(\hat{\eta}_{aj}) &= \text{Var}(g(\hat{\xi}_{aj})) \\ &\approx \nabla g^T(\hat{\xi}_{aj}) \text{Var}(\hat{\xi}_{aj}) \nabla g(\hat{\xi}_{aj}), \end{aligned}$$

811 where $\nabla g^T(x) = (\nabla h^T(x_1), \dots, \nabla h^T(x_{32}))/32$.

We partition $\text{Var}(\hat{\xi}_{aj})$ into 32×32 block matrices with (p, q) -th block matrix be-
ing

$$\text{Var}(\hat{\xi}_{aj})_{pq} = \text{Cov}(\hat{\xi}_{aj}(p), \hat{\xi}_{aj}(q)).$$

The cases for $p = q$ have been discussed in the previous subsection. When $p \neq q$, it can be shown that

$$\text{Cov}(\hat{\xi}_{aj}(p), \hat{\xi}_{aj}(q)) = \begin{pmatrix} \text{Cov}(\hat{\mu}_{1j}(p), \hat{\mu}_{1j}(q)) & 0 & 0 \\ 0 & \text{Cov}(\hat{\mu}_{2j}(p), \hat{\mu}_{2j}(q)) & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

812 Its estimation can be attained by estimating $\widehat{\text{Cov}}(\hat{\mu}_{1j}(p), \hat{\mu}_{1j}(q))$ and $\widehat{\text{Cov}}(\hat{\mu}_{2j}(p), \hat{\mu}_{2j}(q))$,
 813 with the block bootstrap method, the method we have used earlier.

814 It can be shown that $\hat{\eta}_{aj}$ is asymptotically distributed, which leads to the confi-
 815 dence interval in Figure 5 (b).

Figure1.

Author Manuscript

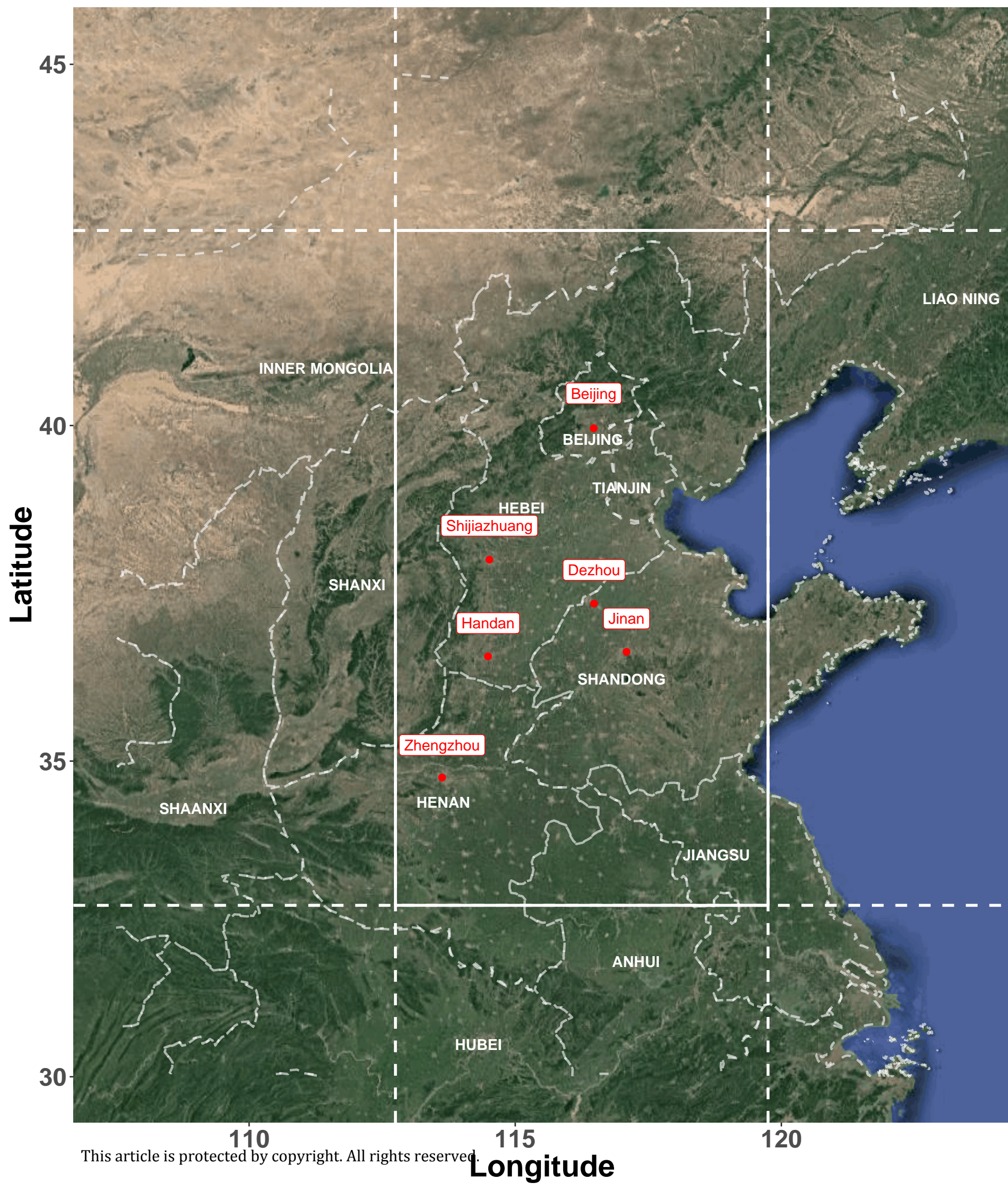
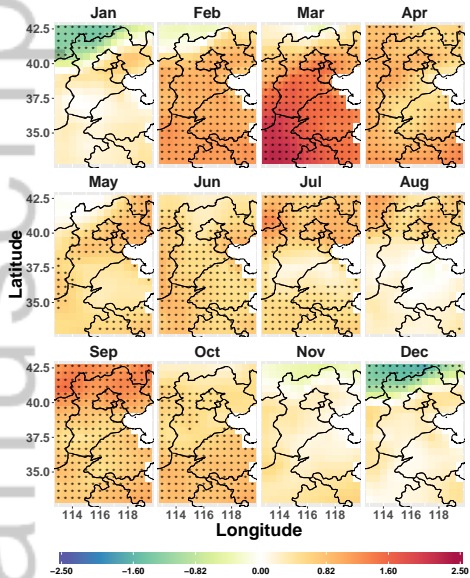


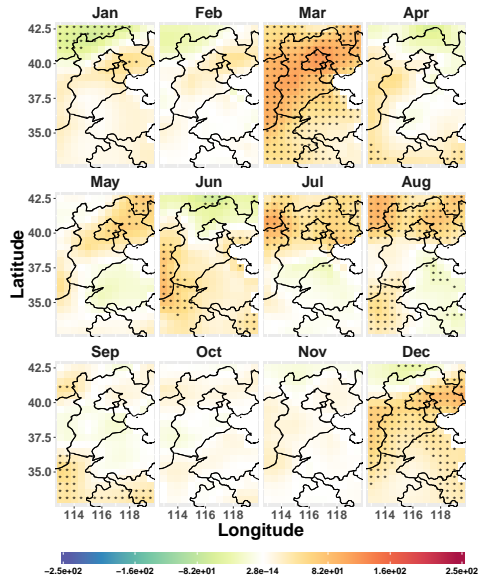
Figure2.

Author Manuscript

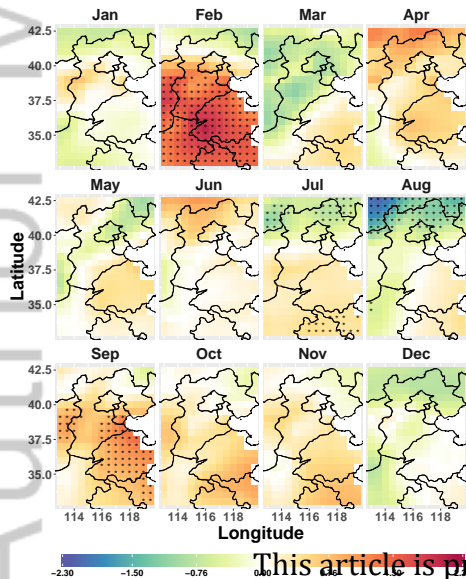
(a) T2M (K)



(b) BLH (m)



(c) D2M (K)



(d) IRH (%)

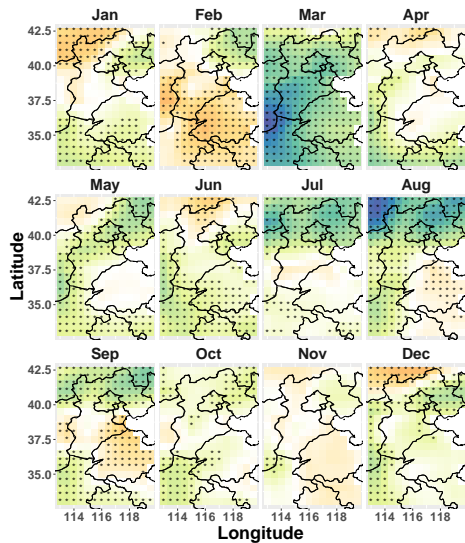
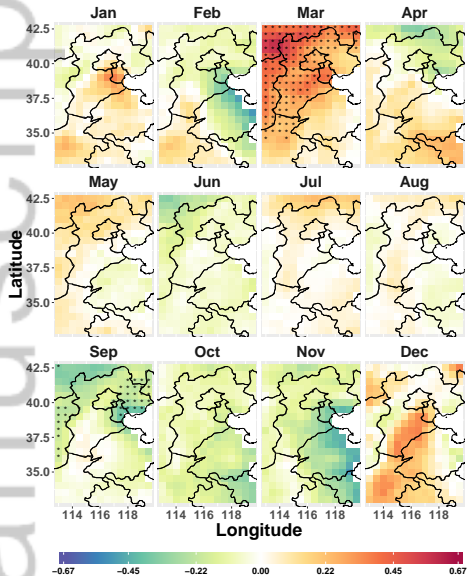


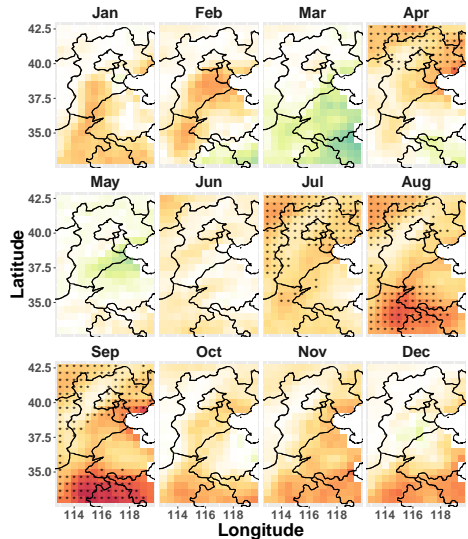
Figure3.

Author Manuscript

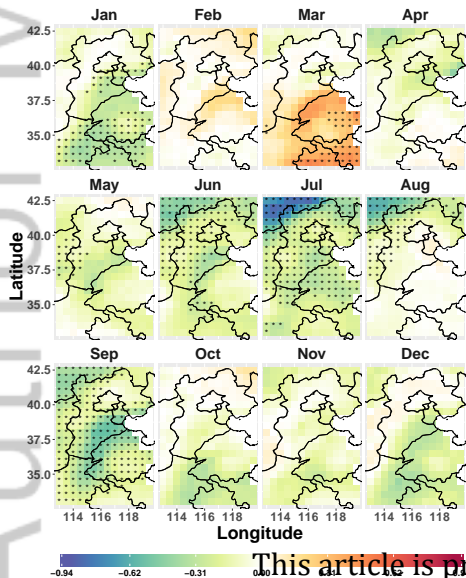
(a) Integrated NW (m/s)



(b) Integrated NE (m/s)



(c) Integrated SW (m/s)



(d) Integrated SE (m/s)

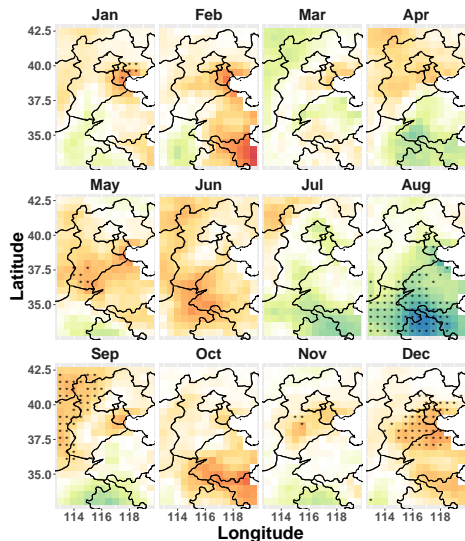
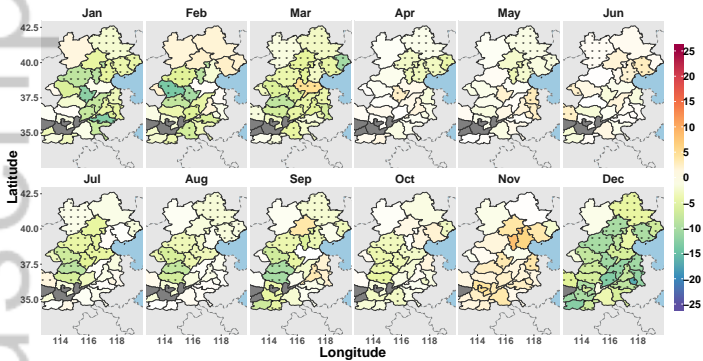


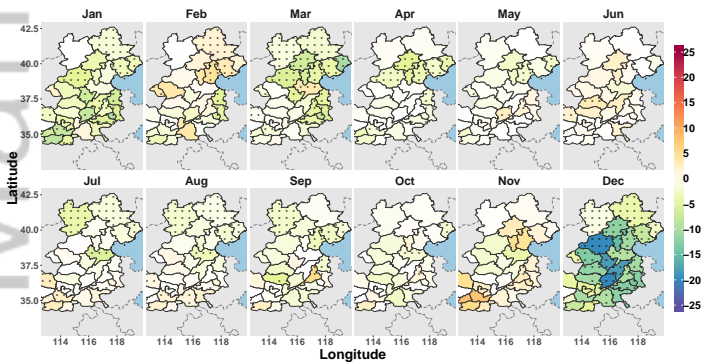
Figure4.

Author Manuscript

(a) 2014



(b) 2015



(c) 2016

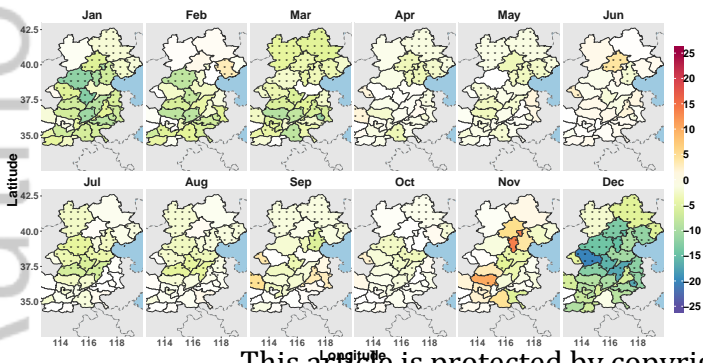
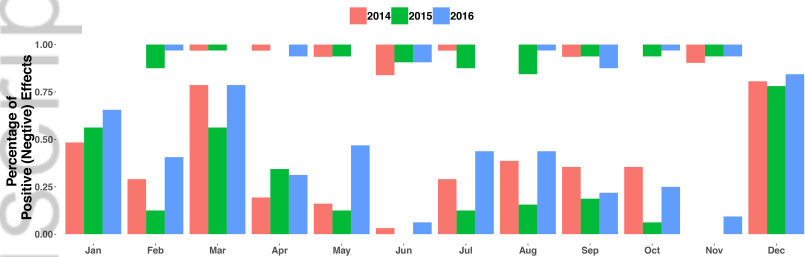


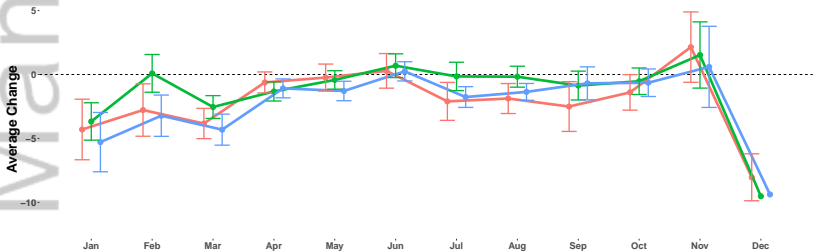
Figure5.

Author Manuscript

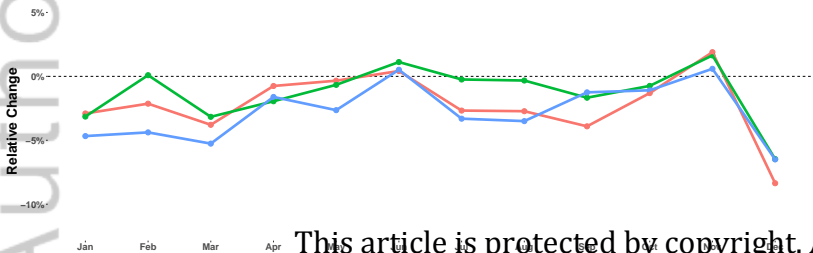
(a)

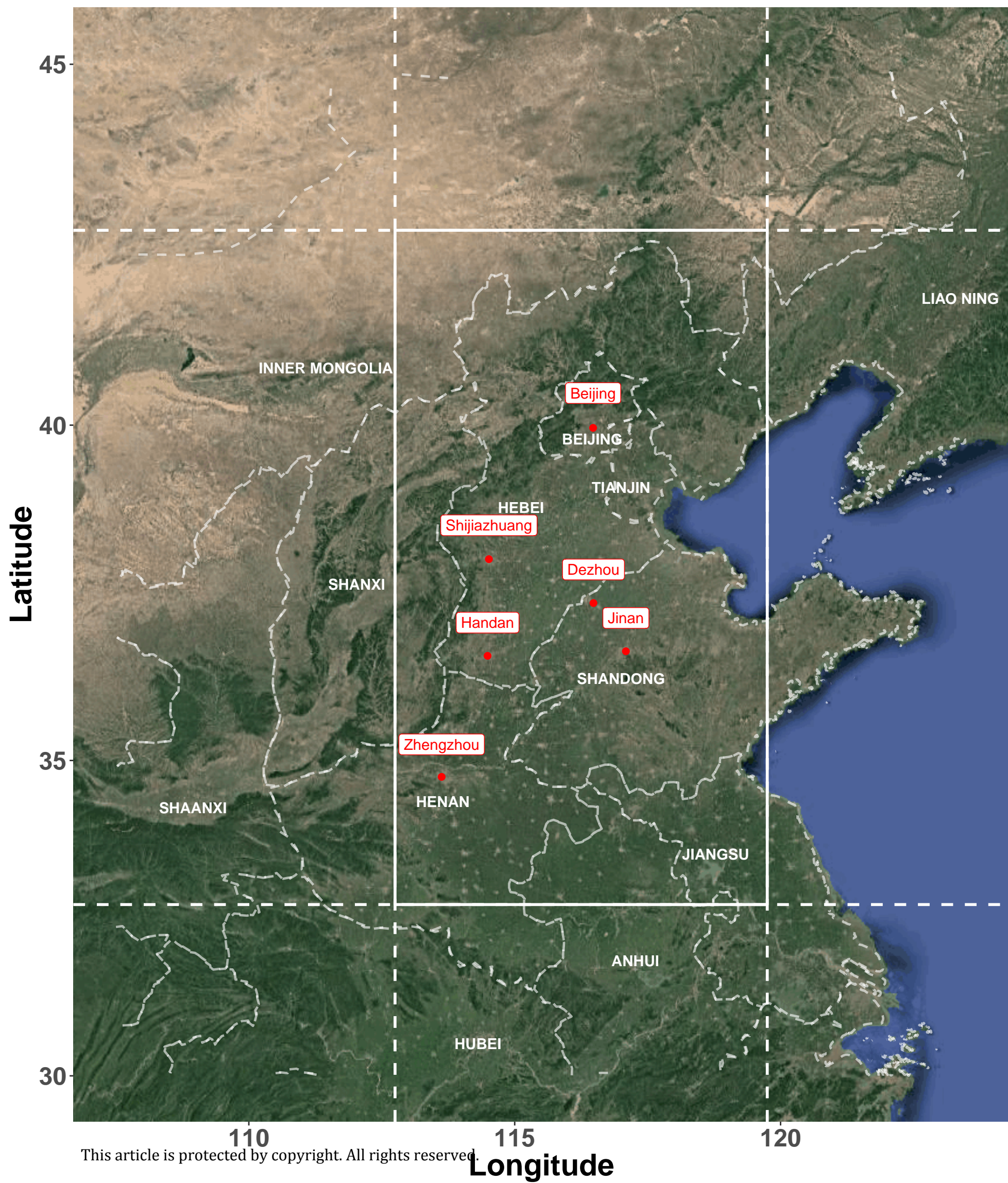


(b)

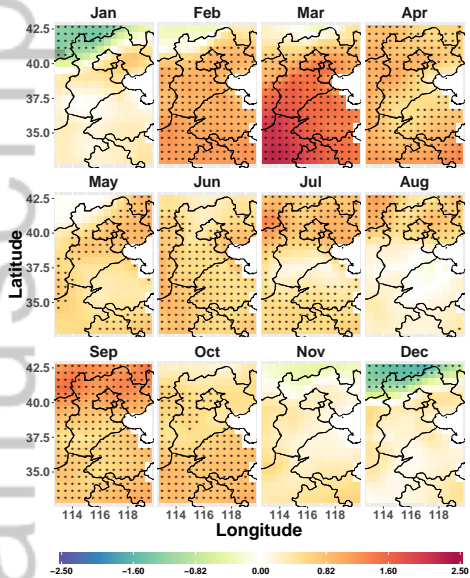


(c)

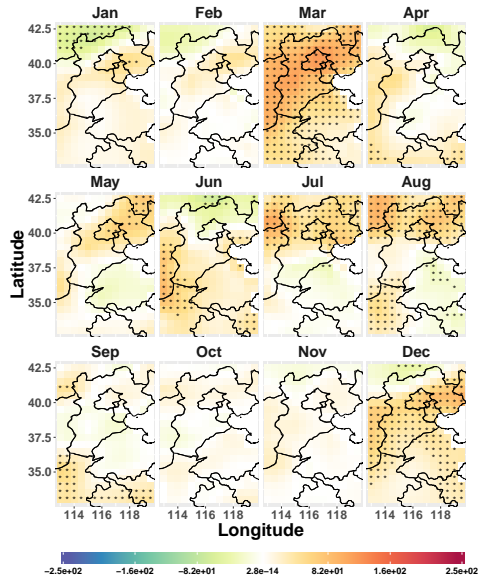




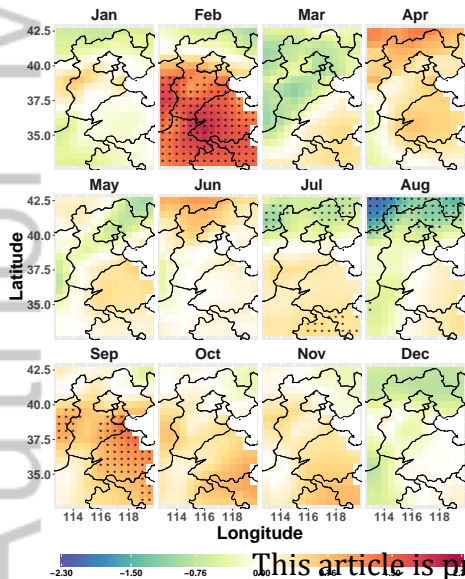
(a) T2M (K)



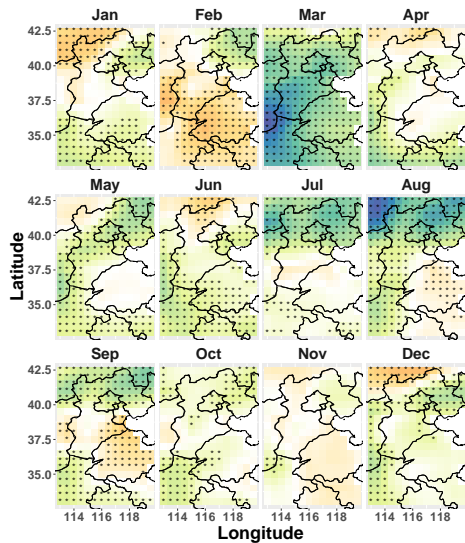
(b) BLH (m)



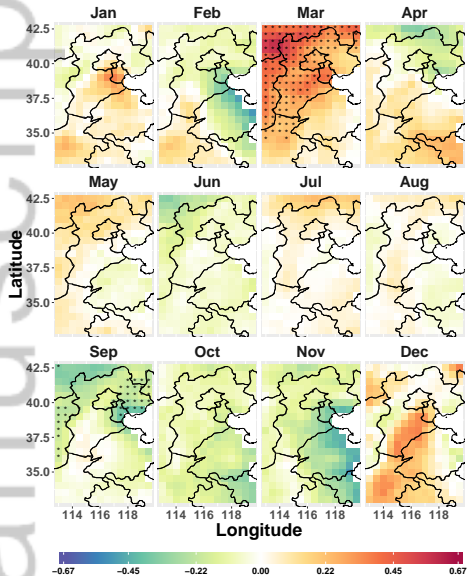
(c) D2M (K)



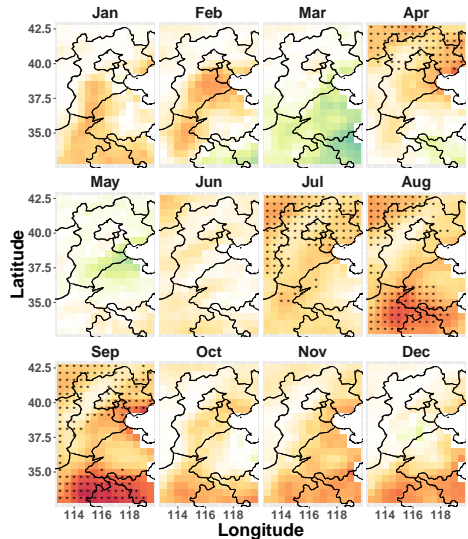
(d) IRH (%)



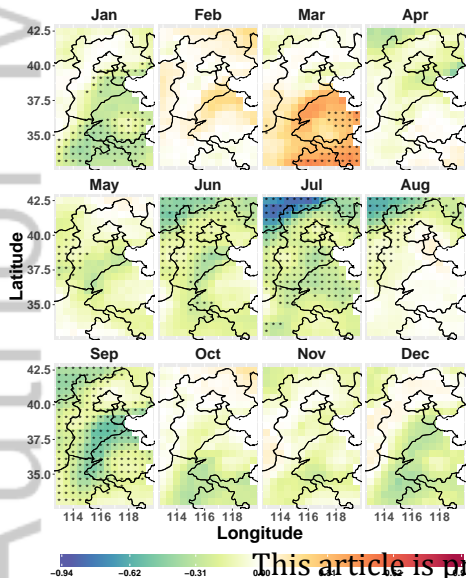
(a) Integrated NW (m/s)



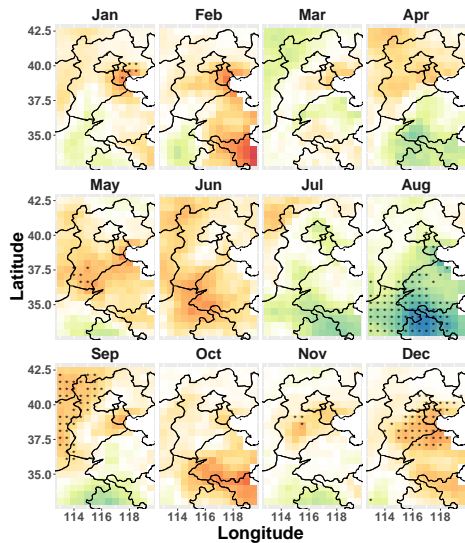
(b) Integrated NE (m/s)



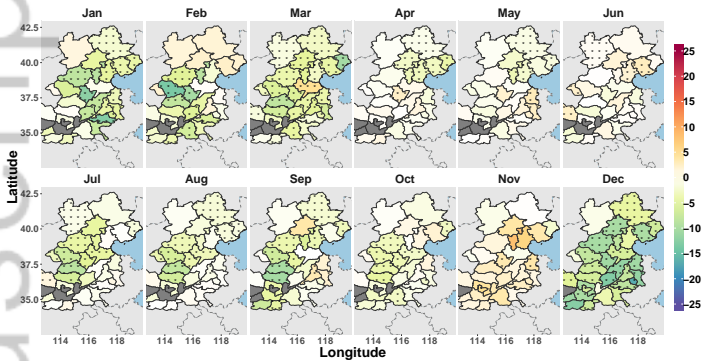
(c) Integrated SW (m/s)



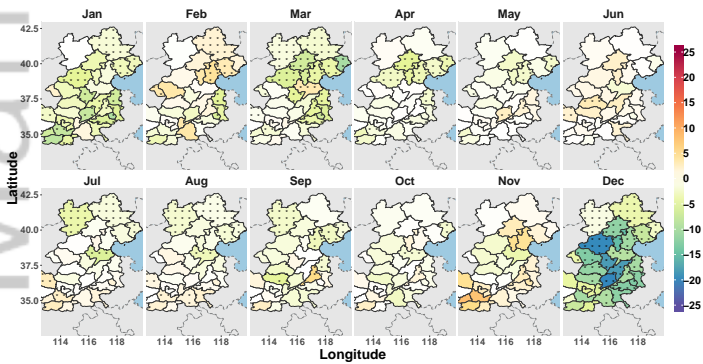
(d) Integrated SE (m/s)



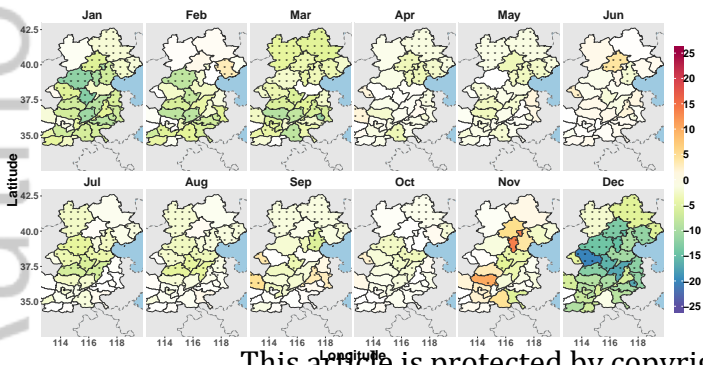
(a) 2014



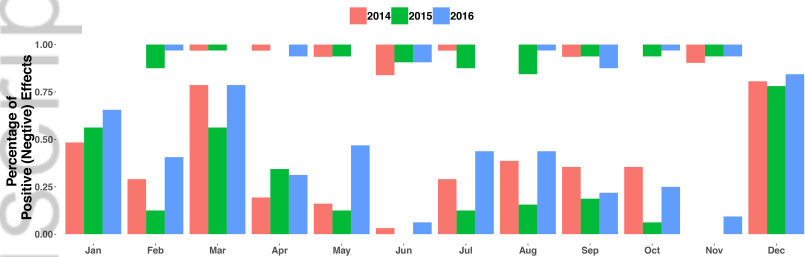
(b) 2015



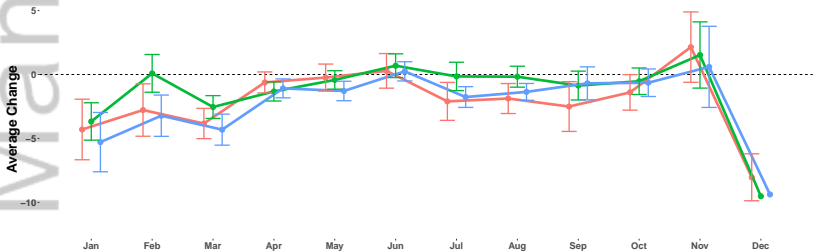
(c) 2016



(a)



(b)



(c)

