HHS Public Access

Author manuscript

Environ Res. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

Environ Res. 2017 October ; 158: 54–60. doi:10.1016/j.envres.2017.06.001.

Evaluation of a data fusion approach to estimate daily $PM_{2.5}$ levels in North China

Fengchao Liang[a,c], **Meng Gao**[b], **Qingyang Xiao**[c], **Gregory R. Carmichael**[b], **Xiaochuan Pan**[a], and **Yang Liu**[c]

[a]Department of Occupational and Environmental Health, School of Public Health, Peking University, Beijing 100191, China

[b]Center for Global and Regional Environmental Research, the University of Iowa, Iowa City, IA 52242, USA

[c]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA

Abstract

$PM_{2.5}$ air pollution has been a growing concern worldwide. Previous studies have conducted several techniques to estimate $PM_{2.5}$ exposure spatiotemporally in China, but all these have limitations. This study was to develop a data fusion approach and compare it with kriging and Chemistry Module. Two techniques were applied to create daily spatial cover of $PM_{2.5}$ in grid cells with a resolution of 10 km in North China in 2013, respectively, which was kriging with an external drift (KED) and Weather Research and Forecast Model with Chemistry Module (WRF-Chem). A data fusion technique was developed by fusing $PM_{2.5}$ concentration predicted by KED and WRF-Chem, accounting for the distance from the central of grid cell to the nearest ground observations and daily spatial correlations between WRF-Chem and observations. Model performances were evaluated by comparing them with ground observations and the spatial prediction errors. KED and data fusion performed better at monitoring sites with a daily model R^2 of 0.95 and 0.94, respectively and $PM_{2.5}$ was overestimated by WRF-Chem (R^2=0.51). KED and data fusion performed better around the ground monitors, WRF-Chem performed relative worse with high prediction errors in the central of study domain. In our study, both KED and data fusion technique provided highly accurate $PM_{2.5}$. Current monitoring network in North China was dense enough to provide a reliable $PM_{2.5}$ prediction by interpolation technique.

Keywords

$PM_{2.5}$; KED; WRF-Chem; Data fusion; Spatiotemporal model

Correspondence to: Xiaochuan Pan; Yang Liu.

Appendix A. Supporting information: Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.envres.2017.06.001.

1. Introduction[1]

The North China region, home to 350 million people and the Beijing Metropolitan Area, is the cultural and political center of China and one of its economic hubs. For the past three decades, rapid urbanization, population growth and expansion of industrial land use have severely affected the air quality in this region, specifically fine particulate matter ($PM_{2.5}$). $PM_{2.5}$ levels in most cities in the North China remain above both the WHO guideline and China's National Ambient Air Quality Standard (NAAQS, http://kjs.mep.gov.cn/) (Hu et al., 2014; Zhang and Cao, 2015). Numerous studies worldwide have linked exposure to $PM_{2.5}$ to various adverse health effects, including respiratory and cardiovascular mortality and morbidity, birth outcomes, as well as diabetes (Brook et al., 2013; Gauderman et al., 2004; Lepeule et al., 2012; Miller et al., 2007; Stieb et al., 2016).

Previously, most studies in China have typically characterized the exposures to $PM_{2.5}$ using ground observations (Zhang and Cao, 2015). Since 2013, China has invested heavily in the development of a national regulatory monitoring network, with now over 1500 monitors in operation nationwide. Similar to the network operated by the U.S. Environmental Protection Agency, the coverage of China's regulatory network is mostly concentrated in urban areas. Since major emission sources such as coal-fired power plants and biomass burning are located in rural areas, this network alone is insufficient to fully characterize the spatial and temporal variability of $PM_{2.5}$ levels.

For the past 15 years, various techniques have been developed to fill the data gaps left by ground monitors. At one end of this spectrum of techniques, geostatistical interpolation methods such as inverse distance weighting (IDW), spline interpolation and kriging have been applied to obtain $PM_{2.5}$ concentration surfaces using ground observations (Li and Liu, 2014; Li et al., 2016; Ramos et al., 2016; Sampson et al., 2013). For example, a study in Korea used multiple spatial interpolation methods to estimate air pollutant exposure, including average values from all monitors, nearest neighbor, IDW and Ordinary kriging (OK), which were all based on 13 ground monitors. It showed that OK provided the most accurate estimated exposures (Son et al., 2010). Since kriging and similar techniques are heavily dependent on ground data support, it works poorly in regions or during the time periods with few monitors. In China, very little $PM_{2.5}$ data exist before 2013 that can be used for spatial interpolation. At the other end of this spectrum, techniques independent from ground observations such as atmospheric chemical transport models (CTMs) can predict $PM_{2.5}$ concentrations with complete spatial and temporal coverage (Gao et al., 2016b; Hu et al., 2016). A well-known limitation of CTMs is that the errors in pollutant emissions inventory and meteorological fields can introduce substantial errors in model predictions (Hogrefe et al., 2015; Zhang et al., 2016). To take advantage of the high accuracy of kriging and full coverage of CMT predictions, a fusion approach has been proposed that OK fused with Community Multiscale Air Quality (CMAQ) model to predict 12 pollutants in Georgia, United States (Friberg et al., 2016; Puttaswamy et al., 2014).

[1] $PM_{2.5}$, fine particulate matter; NAAQS, National Ambient Air Quality Standard; IDW, inverse distance weighting; OK, Ordinary kriging; CTMs, chemical transport models; CMAQ, Community Multiscale Air Quality; WRF-Chem, Weather Research and Forecast Model with Chemistry Module; KED, kriging with external drift; DOY, day-of-year; MOSAIC, Model for Simulating Aerosol Interactions and Chemistry; CV, cross-validation; OBS, observation.

Given that data fusion technique for kriging and CTMs has not been applied in China, our objective is to develop such a fusion approach and evaluate the accuracy of spatiotemporally resolved ambient $PM_{2.5}$ concentration predicted by kriging, Weather Research and Forecast Model with Chemistry Module (WRF-Chem) and data fusion techniques in North China.

2. Data and methods

2.1. Study domain, modeling grid and ground monitoring data

Our study domain covers Beijing, Tianjin, Hebei, Shanxi, and Shandong province as well as part of Liaoning, Inner Mongolia, Shaanxi, Henan, Anhui, and Jiangsu province (Fig. 1). The air mass over the densely populated and highly industrialized low lands is isolated by the Inner Mongolian Plateau to the north, the Taihang Mountains to the west with elevations above 1500 m, and the Bohai Sea to the east. Strong local emissions and poor dispersion conditions are major contributors to the severe $PM_{2.5}$ pollution levels (Gao et al., 2016b; Li and Liu, 2011). Daily mean $PM_{2.5}$ concentrations in 2013 from 410 monitors were obtained from China's Ministry of Environmental Protection (http://english.mep.gov.cn/) and local Environmental Protection Bureaus in the study domain, of which 362 are operated by the national government and 48 by local governments; 365 monitors are located in our study domain and the others were used to make reliable predictions at the border areas. Daily mean number (N) of operating monitors in this work was 283, with a range of 59–349. We created a 10 km resolution modeling grid in this domain in order to compare different $PM_{2.5}$ modeling results, a total of 13,326 grid cells contributed in daily prediction.

2.2. Kriging model

Kriging is a linear weighted combination of observed values that uses spatial autocorrelation among data to determine the weights with or without taking account for an external trend or trend model. Under suitable assumptions on the priors, kriging gives the best linear unbiased prediction of the intermediate values (Cressie, 1990). We developed a kriging with external drift (KED) model to spatially interpolate daily $PM_{2.5}$ concentration observations to the entire modeling grid. KED is a particular case of universal kriging, where the prediction for nonstationary processes is performed by taking into account both local trends of the target variable and external drift (a spatial trend) when minimizing the estimation variance (Chauvet and Galli, 1982; Hengl et al., 2003). Previous studies have showed that elevation is a good indicator of $PM_{2.5}$ levels prediction (Pearce et al., 2009; Tunno et al., 2016). We used the 30 m resolution elevation data from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 2 (GDEM V2: https://asterweb.jpl.nasa.gov/gdem.asp). In this study, the square root of elevation in grid cells was used as the external drift term. Matern semi-variogram model was selected when fitting the relationship between distance of observations and semi-variances, which is a special case of exponential and Gaussian semi-variogram based on the parameter which controls the smoothness (Fuentes, 2001). Because of its flexibility, Matern semi-variogram has been commonly used in Kriging models (Pardo-Iguzquiza and Chica-Olmo, 2008). The semi-variograms were fitted at the seasonal level instead of the daily level to maintain a steady relationship between spatial lag and semi-variances in case of extreme episodes.

Summary statistics of semi-variogram parameters fitted in different seasons were shown in Table S1. The KED model in our study can be written as follows:

$$\mathrm{PM}_{2.5(s,t)}^{\mathrm{KED}}=\sum_{m=1}^{N}\omega_{(s,t)}^{\mathrm{KED}}\times \mathrm{PM}_{2.5(m,t)} \quad (1)$$

where $\mathrm{PM}_{2.5(s,t)}^{\mathrm{KED}}$ is KED-interpolated $PM_{2.5}$ concentration at grid cell *s* on day-of-year (DOY) *t*; $\omega_{(s,t)}^{\mathrm{KED}}$ is the KED weight for $PM_{2.5}$ observation taken at monitoring site *m* at grid cell *s* on DOY *t*; $PM_{2.5(m,t)}$ is observation value of $PM_{2.5}$ at monitoring site *m* on DOY *t*.

2.3. WRF-Chem simulated $PM_{2.5}$ Concentrations

The chemistry version of the Weather Research and Forecasting model (WRF-Chem) simulates meteorology and chemistry simultaneously and considers the interactions between them (Grell et al., 2005). It has been adopted to study particle pollution in China in recent years (Gao et al., 2016a, 2016b). In this study, we used the WRF-Chem model version 3.5.1 with two nested domains (81 km and 27 km), covering most areas of east Asia and east China for the 2013 model year. The domains have 27 vertical layers, up to a minimum pressure of 50 hPa. The 8-bin sectional Model for Simulating Aerosol Interactions and Chemistry (MOSAIC) with aqueous-phase chemistry were used to simulate sulfate, nitrate, chloride, ammonium, sodium, black carbon, primary organic mass, liquid water and other inorganic mass and the range of each bin were 0.039–0.078 μm, 0.078–0.156 μm, 0.156–0.312 μm, 0.312–0.625 μm, 0.625–1.25 μm, 1.25–2.5 μm, 2.5–5.0 μm, 5.0–10 μm (Zaveri et al., 2008). The Lin scheme (Lin et al., 1983), Rapid Radiative Transfer Model (RRTM) longwave (Mlawer et al., 1997), Goddard shortwave radiation (Kim and Wang, 2011), Noah Land Surface Model (Chen and Dudhia, 2001) and Yonsei University Planetary Boundary Layer (Hong et al., 2006) were used as physical modules in the simulations. Anthropogenic emissions are taken from the MIX emission inventory, which is a MOSAIC Asian anthropogenic emission dataset for the Model Inter-Comparison Study for Asia and the Hemispheric Transport of Air Pollution projects (http://www.meicmodel.org/dataset-mix.html) (Li et al., 2015). This inventory considers emissions of sulfur dioxide, nitrogen oxides, carbon monoxide, non-methane volatile organic compounds, ammonia, black carbon, organic carbon, $PM_{2.5}$, PM_{10}, and carbon dioxide by five sectors, namely power generation, industry, residential, transportation, and agriculture. Biogenic emissions are predicted hourly by the Model of Emissions of Gases and Aerosols from Nature (MEGAN) algorithm (Guenther et al., 2006). As shown in the grid emission maps for gaseous and aerosol species in Li's study (Li, 2017), emissions in North China are much higher than in other areas. Enhanced emissions are mainly located in southern Hebei and Shandong, except that high NH_3 emissions are mostly located in Henan province. Detailed model description and evaluation is provided elsewhere (Gao et al., 2016b; Yu et al., 2012). Simulated daily $PM_{2.5}$ concentrations at 27 km resolution were re-sampled to the 10 km grid so that they could be compared with other $PM_{2.5}$ exposure estimates at the same resolution.

2.4. Data fusion approach

Based on the work of Friberg et al. (2016), we developed a three-step approach to fuse the ground monitoring data of $PM_{2.5}$ and WRF-Chem simulation. First, we conducted daily domain-wide calibration of the KED and WRF-Chem estimated $PM_{2.5}$ concentrations in order to align the outputs of these two models (hence the fusion model outputs) on any given day to the domain-average $PM_{2.5}$ concentration observed by the monitors. A simple linear regression model between the observed and KED-estimated daily $PM_{2.5}$ concentrations was developed at monitor locations daily. The regression coefficients were used to adjust the daily KED predictions in each grid cell as in Eq. (2):

$$C_{1(s,t)}=\alpha_{1,t}+\beta_{1,t}\times \mathrm{PM}^{\mathrm{KED}}_{2.5(s,t)} \quad (2)$$

where $C_{1(s,t)}$ is daily adjusted KED $PM_{2.5}$ fields at grid cell s on DOY t; $\alpha_{1,t}$ and $\beta_{1,t}$ are the fitted linear regression intercept and slope on DOY t using KED predictions matched with observations. Similarly, a linear regression between the observed $PM_{2.5}$ concentrations and matched 10 km WRF-Chem predictions was developed at monitor locations daily. The regression coefficients were used to adjust the WRF-Chem predictions in each grid cell as in Eq. (3):

$$C_{2(s,t)}=\alpha_{2,t}+\beta_{2,t}\times \mathrm{WRF}_{10\mathrm{km}(s,t)} \quad (3)$$

where $C_{2(s,t)}$ is daily adjusted WRF-Chem fields (10 km resolution) at grid cell s on DOY t; $\alpha_{2,t}$ and $\beta_{2,t}$ fitted linear regression intercept and slope on DOY t using WRF-Chem predictions matched with observations.

In the second step, we calculated the correlation coefficients of $C_{1(s,t)}$ and $C_{2(s,t)}$ in a given grid cell with observations in order to assign the weights of the KED and WRF-Chem predictions in the daily fused $PM_{2.5}$ concentrations. The correlation coefficients between $PM_{2.5}$ observations from all paired monitors during the entire study period were first calculated, then the relationship of the inter-monitor correlation coefficients and their distances was fitted with an exponential function (correlogram showed in Fig. S1). It is known that kriging error increases as predictions are made further away from ground monitors. Building this relationship using paired observations instead of directly calculating the correlation between KED predictions (i.e., $C_{1(s,t)}$) and observed $PM_{2.5}$ concentrations from the nearest monitor avoids the potential interference of the changing kriging error. The fitting parameters r_0 and D were used to predict the correlation coefficient between calibrated-KED predictions and observations in Eq. (4). Distance to the nearest ground monitor for each grid cell varied daily since observations may be missing on some days. Spatial distribution of annual mean distance was showed in Fig. S2.

$$r_{1(s,t)}=r_0 e^{-d/D} \quad (4)$$

where $r_{1(s,t)}$ is the predicted correlation coefficient between $C_{1(s,t)}$ in grid cell s on DOY t and the nearest monitoring site at a distance d.

The correlation coefficient $r_{2(t)}$ between WRF-Chem predicted and observed daily $PM_{2.5}$ concentrations on DOY t was calculated using matched observations and WRF-Chem predictions at all monitor locations in the study domain.

In the final step, we constructed the weights of the KED ($w_{(s,t)}$) and WRF-Chem predictions ($1-w_{(s,t)}$) in the fused product as in Eq. (5):

$$w_{(s,t)}=\frac{r_{1(s,t)} \times (1-r_{2(t)})}{r_{1(s,t)} \times (1-r_{2(t)})+r_{2(t)} \times (1-r_{1(s,t)})} \quad (5)$$

The final fused daily $PM_{2.5}$ concentration in grid cell s on DOY t ($Fusion_{(s,t)}$) is calculated as:

$$\text{Fusion}_{(s,t)}=w_{(s,t)} \times C_{1(s,t)}+(1-w_{(s,t)}) \times C_{2(s,t)} \quad (6)$$

2.5. Model performance evaluation and cross-validation

Simple linear regression models were performed for the paired data at ground monitor locations to evaluated model fitting and cross-validation (CV) performances of KED, WRF-Chem and data fusion models. Coefficients of determination (R^2) at both daily and annual levels were compared. Correlation coefficient (r) between WRF-Chem simulated $PM_{2.5}$ and observations at monitor sites were also calculated to compare the WRF-Chem model performances between our study and previous studies in the eastern Asia. Ten-fold CV was used to assess potential model over-fitting and prediction accuracy for KED and the fusion model.

To show the model fitting performances at different $PM_{2.5}$ levels, 365 days of 2013 were classified to six levels based on the daily domain-averaged $PM_{2.5}$ concentrations observed by ground monitors. The grading standard was consistent with that published by China's Ministry of Environmental Protection (http://datacenter.mep.gov.cn/index). Model R^2 at daily level, mean biases (Eq. (7)) and mean errors (Eq. (8)) of KED, WRF-Chem and data fusion models were calculated.

$$\text{Meanbias}=\frac{1}{N_2}\sum_{m=1}^{N_2}\frac{\sum_{t=1}^{N_1}(\text{OBS}_{(m,t)}-\text{Pred}_{(m,t)})}{N_1} \quad (7)$$

$$\text{Meanerror}=\frac{1}{N_2}\sum_{m=1}^{N_2}\frac{\sum_{t=1}^{N_1}|\text{OBS}_{(m,t)}-\text{Pred}_{(m,t)}|}{N_1} \quad (8)$$

where N_1 and N_2 are the counts of days and monitors in each $PM_{2.5}$ grade level, respectively; $OBS_{(m,t)}$ is daily $PM_{2.5}$ concentration observed at monitoring site m on DOY t; $Pred_{(m,t)}$ is daily $PM_{2.5}$ concentration predicted by KED, WRF-Chem or data fusion models at monitoring site m on DOY t.

To evaluate model prediction uncertainty, we calculated the prediction error of $\text{PM}^{\text{KED}}_{2.5(s,t)}$ and $C_{1(s,t)}$ as the square root of KED expected variance using the R package "geoR". For $WRF_{10\text{km}(s,t)}$ and $C_{2(s,t)}$, we first calculated the squared difference between observations and $WRF_{10\text{km}(s,t)}$ or $C_{2(s,t)}$ predictions at monitoring sites, respectively, then spatially interpolated these squared differences to all grid cells; finally, square root of the interpolated results were presented as the prediction errors. Prediction error of data fusion was performed as the weighted sum of the errors of $C_{1(s,t)}$ and $C_{2(s,t)}$. All data processing and modeling were conducted in R.

3. Results

Summary statistics of seasonal and annual mean observed and estimated $PM_{2.5}$ concentrations by KED, WRF-Chem and the fusion method at ground monitor locations in the study domain are shown in Table 1. The annual mean $PM_{2.5}$ concentration of the 365 ground monitors was 90.9 μg/m^3, similar with the $PM_{2.5}$ fields modeled by KED and data fusion at monitoring sites, but lower than that simulated by WRF-Chem. For all estimation methods, $PM_{2.5}$ concentrations in winter (Jan., Feb. and Dec.) were much higher than that in the other seasons both at site and domain level (spring is defined as Mar., Apr. and May, summer is Jun., Jul. and Aug. and autumn is Sep., Oct. and Nov).

Table 2 shows model performance statistics. When compared with ground observations, both the KED and fusion model performed well at both the daily and annual levels, with model fitting and CV R^2 values above 0.9. The agreement between uncalibrated WRF-Chem simulations and ground observations is worse. Fig. 2 indicates that KED predictions at monitoring locations are in overall good agreement with observations with a slight underestimation at high $PM_{2.5}$ levels. WRF-Chem had a substantial overestimation across the entire data range especially at the lower end of $PM_{2.5}$ levels. The combined effect is reflected in the fused $PM_{2.5}$ data which shows good overall agreement with observations as indicated by the high R^2 values, a slight over-estimation at low $PM_{2.5}$ levels, and a slight underestimation at high $PM_{2.5}$ levels. Models performances at different $PM_{2.5}$ grade levels were shown in Table 3. KED and data fusion performed better at all $PM_{2.5}$ levels with daily R^2 of 0.83–0.93, mean biases of −1.1 to 1.8 μg/m^3 and mean errors of 4.5–31.3 μg/m^3. For WRF-Chem simulation, relative poor model fitting R^2 were shown at all $PM_{2.5}$ levels 0.26–0.39) compared to an overall R^2 of 0.51 in the whole year of 2013.

Time series analysis of ground observations, KED, WRF-Chem and fusion estimates of daily $PM_{2.5}$ concentrations showed that all three models were able to demonstrate the seasonal variability of $PM_{2.5}$ levels, with higher concentrations in winter and lower concentrations in the summer. However, WRF-Chem substantially overestimated $PM_{2.5}$ levels, especially in the summer, resulting in 320 days in exceedance of the Chinese NAAQS (75 μg/m^3), as compared to 193 days observed by the ground network (Fig. S3). The spatial distribution patterns of annual mean $PM_{2.5}$ modeled by the KED, WRF-Chem and fusion model in 2013 were shown in Fig. 3A-C. The KED and fused PM2.5 fields both have the highest $PM_{2.5}$ level (> 155 μg/m^3) in south Hebei and the lowest level in north Hebei and southcentral Inner Mongolia region (20–25 μg/m^3). The WRF-Chem predicted high $PM_{2.5}$ levels span south Hebei, north Henan and west Shandong provinces (greater than 175 μg/m^3), and low $PM_{2.5}$ levels in the northeastern part of Inner Mongolia (20–25 μg/m^3). Seasonally, the spatial patterns of the three models are more similar in winter than the rest of the year when WRF-Chem overestimates $PM_{2.5}$ levels in much of Hebei, Henan and Shandong provinces (Fig. S4).

The lower panel of Fig. 3 showed the spatial distribution of the annual mean prediction errors from the three models in our modeling grid. As expected, the KED errors are the smallest in the grid cells near the monitors, and grow larger in grid cells further away from the monitors. Given the relatively dense coverage of monitors in our study domain, the KED errors are below 35 μg/m^3, and they are generally below 25 μg/m^3 at areas where the nearest monitor is within 100 km away (Fig. S2). The WRF-Chem model error is spatially smoother with the highest error in north Henan, south Shanxi and southeast Hebei provinces (65–75 μg/m^3) and the lowest in part of Hebei and Jiangsu provinces (25–35 μg/m^3). The fusion prediction errors share similar spatial patterns with KED in much of the domain with prediction errors around 20–30 μg/m^3. To show the contributions of KED or WRF-Chem to the fusion results, we plotted the annual mean weights of WRF-Chem ($1 - \omega$) fitted in the data fusion model (Fig. 4). The annual mean weights of WRF-Chem in grid cells within 10–50 km of a monitor is generally below 0.25, meaning that the contribution of WRF-Chem simulations to the fused $PM_{2.5}$ concentrations in these grid cells is less than 25%, whereas that of KED is more than 75%. The contribution of WRF-Chem simulations grows larger away from the monitors and eventually surpasses that of the KED estimates in grid cells that are more than 200 km away from any monitor. At the far north of our study domain, WRF-Chem simulations contribute most to the fused $PM_{2.5}$ levels. Seasonally, WRF-Chem contributed slightly less to the fused $PM_{2.5}$ fields in autumn compared that in other seasons (Fig. S5).

4. Discussion

Our results suggested that KED provided reliable temporal $PM_{2.5}$ predictions with both daily and annual R^2 of 0.95 at locations with ground monitors and daily CV R^2 of 0.92 at locations without monitors (Table 2). Also, KED performed well at all $PM_{2.5}$ levels with model fitting R^2 ranging from 0.85 to 0.93 (Table 3). Compared with previous studies, our KED model performed better than similar models in North America, e.g., annual CV R^2 of 0.88 in U.S. and daily CV R^2 of 0.46–0.89 in Montreal (Ramos et al., 2016; Sampson et al., 2013). Given the fact that the accuracy of KED interpolated $PM_{2.5}$ concentrations varied

with the distribution of ground monitors, KED performed better around ground monitors, especially in monitor-dense areas, whereas the prediction error was larger in rural and less-developed areas, such as Inner Mongolia (Fig. 3D). Compared with the US and Canada, the ground $PM_{2.5}$ monitoring network in the Beijing-Tianjin-Hebei city-cluster and surrounding areas are significantly denser. To explore the practicality of our KED model in regions with fewer ground monitors, we conducted two-fold CVs for both the KED and data fusion models. The model performances statistics essentially remain unchanged as compared to the 10-fold CV results (0.90 vs. 0.92 for KED and 0.89 vs. 0.91 for data fusion at daily level). Therefore, we believe that KED and data fusion models developed in our study will be also suitable for regions with fewer ground air quality monitors.

As WRF-Chem was run without any calibration against the observations, both daily and annual mean $PM_{2.5}$ were overestimated. As shown in Fig. S3, the overall overestimation of $PM_{2.5}$ is mainly due to overestimation in warmer months. This is very likely related to the errors in model wet deposition. In summer, weather prediction models have difficulties in accurately predicting precipitation, which may lead to high predicted aerosol concentration. Besides, it may also be caused by uncertainties in emissions inventory. Still, compared with other modeling studies in China, our WRF-Chem model shows a comparable performance. For daily comparison, the correlation coefficient (r) between observations and our model simulations is 0.7. Zhou et al. (2017) forecasted 24-h $PM_{2.5}$ levels by an operational WRF-Chem model with an r value of 0.67 at daily level in eastern China in 2014 and 2015. Guo et al. (2016) estimated air pollutants during Asia-Pacific Economic Cooperation (APEC) China 2014 in Beijing, the r value between daily observed and WRF-Chem simulated $PM_{2.5}$ was 0.81. Liu et al. (2016) simulated $PM_{2.5}$ in East Asia, compared to daily observations at four sites in North China, the r values in January and July of 2008 were 0.5 and 0.8, respectively. Previous studies showed that WRF-Chem underestimated $PM_{2.5}$ levels in wintertime (Gao et al., 2015) and severe haze days (Zhou et al., 2017). A slight underestimation (annual mean bias was 5.3 $\mu g/m^3$) was also seen in our study during the days with daily $PM_{2.5}$ concentrations higher than 251 $\mu g/m^3$ (Table 3). Aerosol formation and depositions varied with seasons and pollution levels, which may lead to the differences of model performance in different studies.

Data fusion method present here is a hybrid approach of KED and WRF-Chem to predict the spatial distribution of daily $PM_{2.5}$ concentrations. Previous applications of fusion techniques in China are very limited. Lv et al. (2016) fused an OK-based spatial interpolator with satellite remote sensing aerosol optical depth (AOD) to improve model performance when retrieving $PM_{2.5}$ concentrations over North China. As a result, daily CV R^2 increased from 0.48 to 0.61. Both Friberg's and our data fusion approach (Friberg et al., 2016) blended the outputs of a geostatistical model with CTM simulations, however, our spatial annual R^2 (0.94) was substantially higher than that in U.S. (0.63). Compared with Friberg et al., firstly, we applied kriging with an external drift of root elevation instead of OK to capture the spatial variation introduced by geography. Secondly, we had many more ground monitors (compared to only 42 $PM_{2.5}$ monitors in Georgia, U.S.), which allowed us to make a daily domain-wide calibration for both KED and WRF-Chem results rather than at the annual level (Eqs. (3)–(4)). The high density of monitors also helped to minimize the prediction error of KED. Thirdly, in the second step of our fusion work, $r_{2(t)}$ was presented as daily

spatial correlation between WRF-Chem simulation and observations rather than just the average temporal correlation at monitoring sites, thus the weights of WRF-Chem simulation can change daily. This model flexibility made the $PM_{2.5}$ prediction in the temporal level more reliable.

This study indicated that data fusion is an effective approach to improve the accuracy of WRF-Chem simulation spatiotemporally (Figs. 2 and 3) by fusing with KED interpolation. The fused $PM_{2.5}$ concentrations in grid cells near ground monitors resemble the KED result more closely, i.e., $(1-\omega) < 0.5$ (Fig. 4). Further away, the WRF-Chem simulations play an increasingly more important role in shaping the fused results with $(1-\omega)$ approaching or surpassing 0.5 in the areas with fewer ground monitors, e.g. Northern Inner Mongolia. Therefore, it is reasonable to expect model simulations to contribute more substantially to the fused $PM_{2.5}$ concentrations elsewhere in China. An important limitation of interpolation methods is that the $PM_{2.5}$ gradient may be overly smoothed between observations, whereas WRF-Chem allows more spatial variation by employing chemical transport and emissions information in the simulation. The data fusion approach in our study had a comparable performance with KED (Figs. 2 and 3), but it may contain more spatial variation at grid cells without observations, which can only be identified by using independent ground measurements.

Our study has a few limitations. WRF-Chem simulated $PM_{2.5}$ concentrations with a resolution of 10 km was resampled from 27 km, hence spatial variation of $PM_{2.5}$ was underestimated and may introduce some errors when compared with observations and KED result with a resolution of 10 km. Model simulations at a higher spatial resolution may potentially reduce WRF-Chem prediction errors but the associated computational cost will increase significantly. In addition, we interpolated WRF-Chem prediction errors at monitoring locations to each grid cell since no well-established methods were available to compare the prediction errors for both observation-based (i.e., KED) and observation-independent techniques (i.e., WRF-Chem). Further research on the optimal evaluation method is needed but beyond the scope of this work.

5. Conclusion

To our knowledge, this is the first study in China that developed a $PM_{2.5}$ fusion model with full coverage and explored the differences of $PM_{2.5}$ concentrations predicted by KED interpolation, WRF-Chem simulation and data fusion techniques. Data fusion approach had a comparable performance with KED and performed better than WRF-Chem simulations both temporally and spatially. KED predicted daily $PM_{2.5}$ pollution with high accuracy in the North China Plain, especially in the vicinity of ground monitors. By assigning the weights dynamically, data fusion techniques fused ground observations with chemical transport model simulated values with high spatiotemporal accuracy, which provides a methodological framework to merge $PM_{2.5}$ concentrations estimated by different approaches. Our future work will focus on introducing higher resolution model simulations and expanding our modeling domain with a more heterogeneous distribution of ground monitors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work of Y. Liu was partially supported by the NASA Applied Sciences Program (grant NNX14AG01G) and Jet Propulsion Laboratory (contract # 1558091). The work of X. Pan was supported by the Natural Sciences Foundations of China (No. 81273033 and No. 81372950). The work of F. Liang at Emory University as a visiting student was supported by the China Scholarship Council (CSC). The work of M. Gao and G. R. Carmichael were supported by the NASA Applied Sciences Program (grant NNX11AI52G) and EPA STAR program (RD-83503701).

References

Brook RD, et al. Long-term fine particulate matter exposure and mortality from diabetes in Canada. Diabetes Care. 2013; 36:3313–3320. [PubMed: 23780947]

Chauvet, P., Galli, A. Universal Kriging. Centre de géostatistique et morphologie mathématique; 1982.

Chen F, Dudhia J. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system Part I: model implementation and sensitivity. Mon Weather Rev. 2001; 129:569–585.

Cressie N. The origins of kriging. Math Geol. 1990; 22:239–252.

Friberg MD, et al. Method for fusing observational data and chemical transport model simulations To estimate spatiotemporally resolved ambient air pollution. Environ Sci Technol. 2016; 50:3695–3705. [PubMed: 26923334]

Fuentes M. A high frequency kriging approach for non-stationary environmental processes. Environmetrics. 2001; 12:469–483.

Gao M, et al. Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area. Sci Total Environ. 2015; 511:553–561. [PubMed: 25585158]

Gao M, et al. Response of winter fine particulate matter concentrations to emission and meteorology changes in North China. Atmos Chem Phys. 2016a; 16:11837–11851.

Gao M, et al. Modeling study of the 2010 regional haze event in the North China Plain. Atmos Chem Phys. 2016b; 16:1673–1691.

Gauderman WJ, et al. The effect of air pollution on lung development from 10 to 18 years of age. N Engl J Med. 2004; 351:1057–1067. [PubMed: 15356303]

Grell GA, et al. Fully coupled “online” chemistry within the WRF model. Atmos Environ. 2005; 39:6957–6975.

Guenther A, et al. Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). Atmos Chem Phys. 2006; 6:3181–3210.

Guo JP, et al. Impact of various emission control schemes on air quality using WRF-Chem during APEC China 2014. Atmos Environ. 2016; 140:311–319.

Hengl T, et al. Comparison of kriging with external drift and regression-kriging. ITCpp. 2003 51 Technical note.

Hogrefe C, et al. Annual application and evaluation of the online coupled WRF-CMAQ system over North America under AQMEII phase 2. Atmos Environ. 2015; 115:683–694.

Hong SY, et al. A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Weather Rev. 2006; 134:2318–2341.

Hu JL, et al. Spatial and temporal variability of PM2.5 and PM10 over the North China Plain and the Yangtze River Delta, China. Atmos Environ. 2014; 95:598–609.

Hu JL, et al. One-year simulation of ozone and particulate matter in China using WRF/CMAQ modeling system. Atmos Chem Phys. 2016; 16:10333–10350.

Kim HJ, Wang B. Sensitivity of the WRF model simulation of the east asian summer monsoon in 1993 to shortwave radiation schemes and ozone absorption. Asia-Pac J Atmos Sci. 2011; 47:167–180.

Lepeule J, et al. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities Study from 1974 to 2009. Environ Health Perspect. 2012; 120:965–970. [PubMed: 22456598]

Li L, Liu DJ. Study on an air quality evaluation model for Beijing city under haze-fog pollution based on new ambient air quality standards. Int J Environ Res Public Health. 2014; 11:8909–8923. [PubMed: 25170682]

Li LJ, Liu Y. Space-borne and ground observations of the characteristics of CO pollution in Beijing, 2000–2010. Atmos Environ. 2011; 45:2367–2372.

Li LX, et al. Spatiotemporal interpolation methods for the application of estimating population exposure to fine particulate matter in the contiguous US and a real-time web application. Int J Environ Res Public Health. 2016; 13

Li M. MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. Atmos Chem Phys Discuss. 2017; 15:34813–34869.

Li M, et al. MIX: a mosaic Asian anthropogenic emission inventory for the MICS-Asia and the HTAP projects. Atmos Chem Phys Discuss. 2015; 15:34813–34869.

Lin YL, et al. Bulk parameterization of the snow field in a cloud model. J Clim Appl Meteorol. 1983; 22:1065–1092.

Liu XY, et al. Application of online-coupled WRF/Chem-MADRID in East Asia: model evaluation and climatic effects of anthropogenic aerosols. Atmos Environ. 2016; 124:321–336.

Lv BL, et al. Improving the accuracy of daily PM2.5 distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in North China. Environ Sci Technol. 2016; 50:4752–4759. [PubMed: 27043852]

Miller KA, et al. Long-term exposure to air pollution and incidence of cardiovascular events in women. N Engl J Med. 2007; 356:447–458. [PubMed: 17267905]

Mlawer EJ, et al. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J Geophys Res-Atmos. 1997; 102:16663–16682.

Pardo-Iguzquiza E, Chica-Olmo M. Geostatistics with the Matern semivariogram model: a library of computer programs for inference, kriging and simulation. Comput Geosci. 2008; 34:1073–1079.

Pearce JL, et al. Characterizing the spatiotemporal variability of PM2.5 in Cusco, Peru using kriging with external drift. Atmos Environ. 2009; 43:2060–2069.

Puttaswamy SJ, et al. Statistical data fusion of multi-sensor AOD over the continental United States. Geocarto Int. 2014; 29:48–64.

Ramos Y, et al. Spatio-temporal models to estimate daily concentrations of fine particulate matter in Montreal: kriging with external drift and inverse distance-weighted approaches. J Expo Sci Environ Epidemiol. 2016; 26:405–414. [PubMed: 26648248]

Sampson PD, et al. A regionalized national universal kriging model using Partial Least Squares regression for estimating annual PM2.5 concentrations in epidemiology. Atmos Environ (1994). 2013; 75:383–392. [PubMed: 24015108]

Son JY, et al. Individual exposure to air pollution and lung function in Korea Spatial analysis using multiple exposure approaches. Environ Res. 2010; 110:739–749. [PubMed: 20832787]

Stieb DM, et al. A national study of the association between traffic-related air pollution and adverse pregnancy outcomes in Canada, 1999–2008. Environ Res. 2016; 148:513–526. [PubMed: 27155984]

Tunno BJ, et al. Spatial variation in inversion-focused vs 24-h integrated samples of PM2.5 and black carbon across Pittsburgh, PA. J Expo Sci Environ Epidemiol. 2016; 26:365–376. [PubMed: 25921079]

Yu M, et al. Sensitivity of predicted pollutant levels to urbanization in China. Atmos Environ. 2012; 60:544–554.

Zaveri RA, et al. Model for Simulating Aerosol Interactions and Chemistry (MOSAIC). J Geophys Res-Atmos. 2008; 113

Zhang Y, et al. Application of WRF/Chem over East Asia: part I. Model evaluation and intercomparison with MM5/CMAQ. Atmos Environ. 2016; 124:285–300.

Zhang YL, Cao F. Fine particulate matter (PM2.5) in China at a city level. Sci Rep. 2015; 5

Zhou G, et al. Numerical air quality forecasting over eastern China: an operational application of WRF-Chem. Atmos Environ. 2017; 153:94–108.

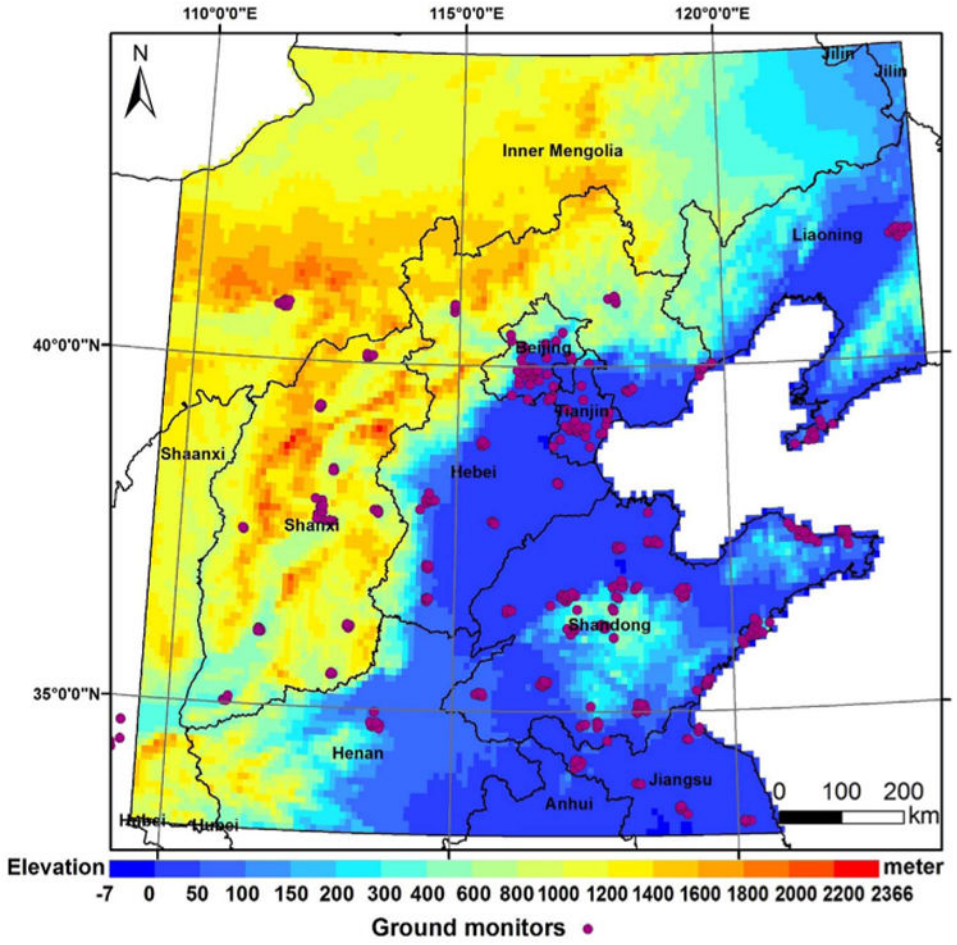

Fig. 1.
Study domain and distribution of ground monitors in North China.

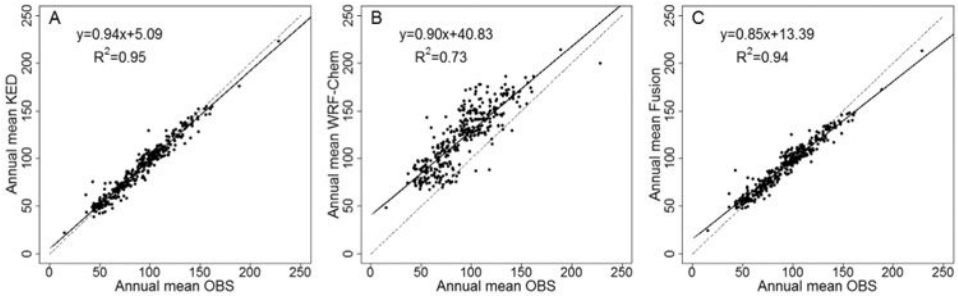

Fig. 2.

Scatter plots of annual mean $PM_{2.5}$ levels estimated by the KED, WRF-Chem and data fusion model compared with ground observations (OBS, N = 365).

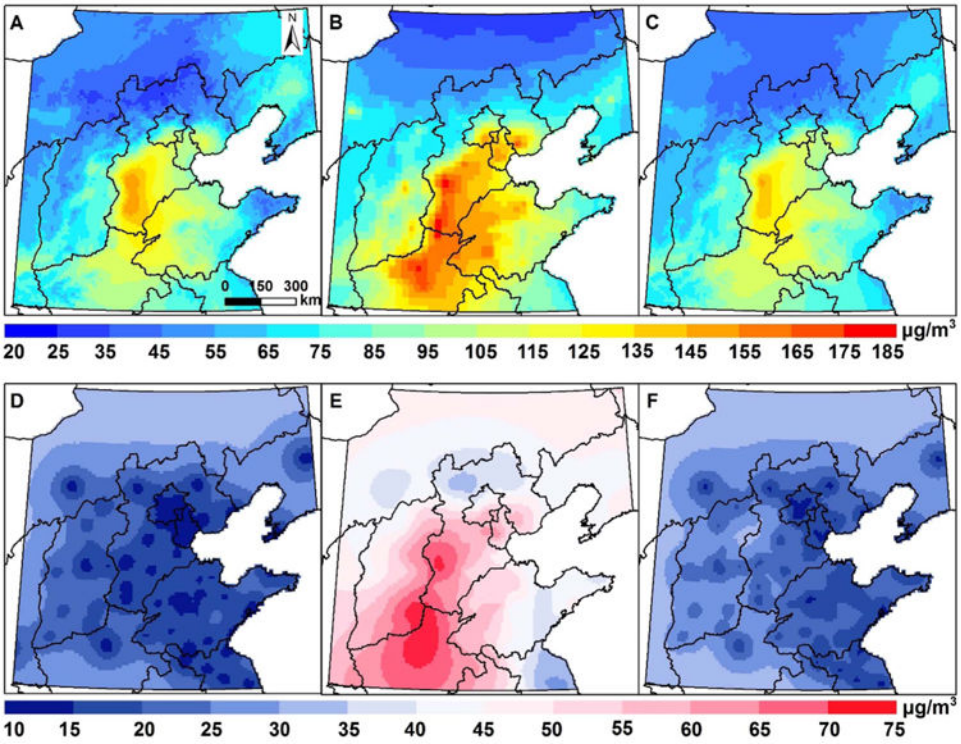

Fig. 3.

Spatial distributions annual mean estimated PM2.5 levels by (A) KED, (B) WRF-Chem and (C) Fusion model. D, E and F are annual mean prediction error of the KED, WRF-Chem and fusion model, respectively.

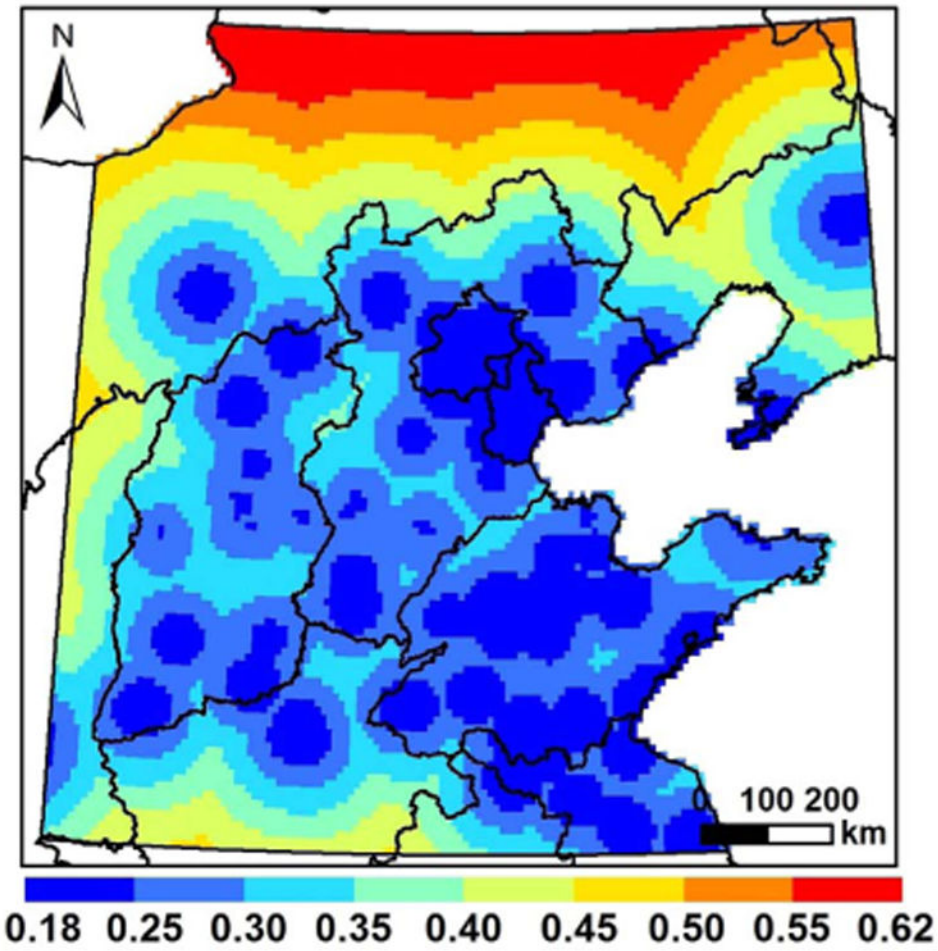

Fig. 4.

Spatial distribution of the annual mean contributions of the WRF-Chem $(1-\omega)$ to the fused $PM_{2.5}$ fields.

Table 1

Annual and seasonal averaged $PM_{2.5}$ concentration (μg/m^3) of ground observations, KED, WRF-Chem and data fusion at monitoring sites.

	Winter	Spring	Summer	Autumn	Annual
N of *OBS*	364	360	359	360	365
$\overline{\mathrm{OBS}_{(m)}}$	138.3	74.6	68.3	83.5	90.9
$\overline{\mathrm{KED}_{(m)}}$	136.9	74.2	68.0	83.0	90.3
$\overline{\mathrm{WRF}_{(m)}}$	165.4	95.8	108.3	118.0	121.6
$\overline{\mathrm{Fusion}_{(m)}}$	138.3	74.5	68.4	83.6	91.0

Table 2

Model performances of KED, WRF-Chem and data fusion.

Item	KED	WRF-Chem	Fusion	N
Daily R^2 at monitoring sites	0.95	0.51	0.94	103,445
Annual R^2 at monitoring sites	0.95	0.73	0.94	365
Daily CV R^2	0.92	–	0.91	103,445
Annual CV R^2	0.92	–	0.92	365

Table 3

Model fitting R^2, mean biases and mean errors (μg/m^3) of KED, WRF-Chem and data fusion at different $PM_{2.5}$ levels.

Levels	R^2 at daily level			Mean bias			Mean error			N
	KED	WRF-Chem	Data fusion	KED	WRF-Chem	Data fusion	KED	WRF-Chem	Data fusion	
0–35	0.85	0.29	0.83	0.1	−22.4	0.0	4.5	24.5	5.0	4155
36–75	0.89	0.28	0.88	0.4	−38.2	0.0	7.4	43.1	8.3	45,258
76–115	0.91	0.34	0.89	0.6	−29.4	0.0	9.8	42.5	11.7	31,289
116–150	0.93	0.39	0.91	1.1	−22.9	0.1	13.7	52.2	17.3	10,531
151–250	0.92	0.39	0.90	1.8	−23.1	−0.5	19.4	68.1	23.6	11,325
≥251	0.92	0.26	0.91	1.8	5.3	−1.1	27.5	88.6	31.3	887