Predicting the Impact of Climate Change on Severe Wintertime Particulate Pollution Events in Beijing Using Extreme Value Theory

D. C. Pendergrass[1] [iD], **L. Shen[2]** [iD], **D. J. Jacob[2]**, and **L. J. Mickley[2]** [iD]

[1]Harvard College, Cambridge, MA, USA, [2]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

Abstract We use extreme value theory to develop point process statistical models relating the probability of extreme winter particulate pollution events in Beijing ("winter haze") to local meteorological variables. The models are trained with the 2009–2017 record of fine particulate matter concentrations ($PM_{2.5}$) from the U.S. embassy. We find that 850-hPa meridional wind velocity (V850) and relative humidity successfully predict the probability for 24-hr average $PM_{2.5}$ to exceed 300 $\mu g/m^3$ (95th percentile of the frequency distribution) as well as higher thresholds. We apply the point process models to mid-21st century climate projections from the Coupled Model Intercomparison Project Phase 5 model ensemble under two radiative forcing scenarios (RCP8.5 and RCP4.5). We conclude that 21st century climate change alone is unlikely to increase the frequency of severe $PM_{2.5}$ pollution events ($PM_{2.5} > 300$ $\mu g/m^3$) in Beijing and is more likely to marginally decrease the probability of such events.

Plain Language Summary We use extreme value theory, a branch of statistics concerned with outliers and unusual events, to develop a model relating the probability of extreme pollution events in Beijing to local weather variables. Haze in Beijing is worst in the winter, so we restrict our study to December, January, and February. We train our models with the 2009–2017 record of fine particulate matter concentrations measured at the U.S. embassy, a pollutant behind many of these haze events. We find that north-south wind velocity and relative humidity successfully predict days when daily mean particulate matter concentrations will exceed a threshold of 300 $\mu g/m^3$. We apply our statistical models to mid-21st century climate projections under two scenarios: business-as-usual emissions and significant reduction in emissions. We find that the frequency of haze events is most likely to decrease because of climate change, driven mainly by a decrease in relative humidity. This result illustrates the importance of including humidity in estimates of future fine particulate matter concentrations.
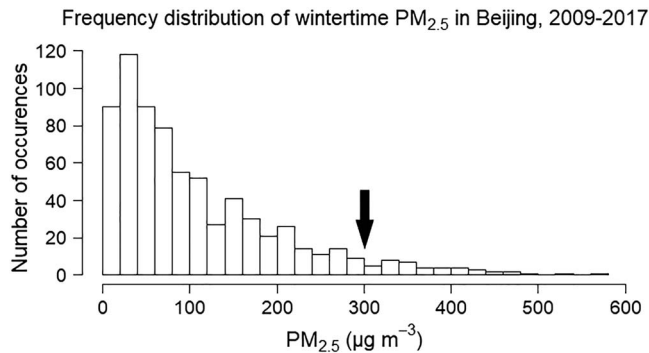
1. Introduction

Beijing in winter experiences severe air pollution events, commonly referred to as "winter haze." Concentrations of fine particulate matter ($PM_{2.5}$) can exceed 300 $\mu g/m^3$ on a 24-hr average basis. Efforts are underway to reduce emissions, but some studies have suggested that climate change may partly offset the air quality gains by favoring the meteorological conditions leading to winter haze (Cai et al., 2017; Zou et al., 2017). Here we use extreme value theory (EVT) to investigate this climate penalty as it applies to severe pollution events.

Cai et al. (2017) found that a meteorological index based on meridional wind velocity, zonal wind gradient, and vertical temperature gradient was strongly correlated with winter haze days exceeding 150 $\mu g/m^3$ in Beijing. They then used climate simulations from the Intergovernmental Panel on Climate Change (IPCC) Coupled Model Intercomparison Project Phase 5 (CMIP5) multimodel ensemble driven by the RCP8.5 (business-as-usual) emissions scenario to project a significant increase in this index over the 21st century, resulting in a 50% increase in haze event frequency. Zou et al. (2017) linked Arctic sea ice loss in the previous autumn and heavy snowfall over Eurasia to poor ventilation conditions over the East China Plain, again implying that 21st century climate change would worsen winter haze. On the other hand, Horton et al. (2014) using the CMIP5 RCP8.5 archive found no change between 1986–2005 and 2046–2065 in the frequency of stagnation events over Beijing.

Figure 1. Frequency distribution of daily wintertime (December-January-February) 24-hr average $PM_{2.5}$ in Beijing for 2009–2017, starting in December 2009 and ending in February 2017. Data are from the U.S. embassy site. The arrow indicates the 300 $\mu g/m^3$ threshold used in our definition of extreme events and representing the 95th percentile of the frequency distribution.

EVT offers a statistical framework to predict the probability of rare occurrences (Coles, 2001). Here we model the probability of extreme haze days in Beijing in relation to meteorological variables using an EVT nonstationary Poisson point process (PP) model, which simultaneously fits the rate of $PM_{2.5}$ exceedances above a certain threshold and the probability density function (PDF) of $PM_{2.5}$ above that threshold. EVT has been applied to extreme pollution episodes in Europe, the United States, and South America (Eastoe, 2009; Martins et al., 2017; Rieder et al., 2013, 2015; Shen et al., 2016) but has not been applied to Beijing haze to our knowledge.
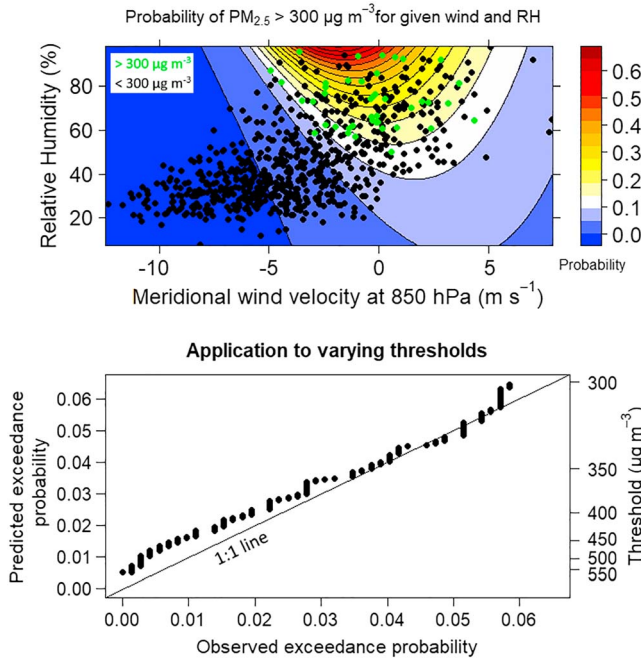
2. Data and Methods

We use the 2009–2017 wintertime (December-January-February [DJF]) time series of 24-hr average $PM_{2.5}$ measured at the U.S. embassy in Beijing, starting in December 2009 and ending in February 2017 (http://www.stateair.net/web/historical/1/1.html). The measurements are referenced to a fixed relative humidity (RH) following standard practice. Previous studies have shown that the U.S. embassy data set is representative of $PM_{2.5}$ in Beijing (Li et al., 2018). Cai et al. (2017) used it for their analysis of the relationship between Beijing haze events and meteorological variables.

We use the National Center for Environmental Prediction and National Center for Atmospheric Research reanalysis at 2.5° by 2.5° resolution (Kalnay et al., 1996) to construct the meteorological covariates of wintertime Beijing haze previously identified by Cai et al. (2017) in their Haze Weather Index. These include (1) 850-hPa meridional wind velocity (V850), averaged spatially in the area around and to the southeast of Beijing (30–47.5° N, 115–130° E), (2) the difference in the 500-hPa zonal wind (δU500) north of Beijing (42.5–52.5° N, 110–137.5° E) versus south of Beijing (27.5–37.5° N, 110–137.5° E), and (3) the vertical temperature difference ($\delta T_{850-250}$) between the lower troposphere at 850 hPa (32.5–45° N, 112.5–132.5° E) and the upper troposphere at 250 hPa (37.5–45° N, 122.5–137.5° E). V850 in particular has a well-established relationship with severe haze in Beijing reflecting the ventilation by the prevailing northerly (negative V850) wind (Li et al., 2018; Shen et al., 2018). We also consider (4) the RH at Beijing International airport obtained from the National Oceanic and Atmospheric Administration (NOAA) Integrated Surface Database (https://www.ncdc.noaa.gov/isd/data-access), since several previous studies have reported a positive correlation of Beijing haze with RH (Leung et al., 2018; Wang et al., 2014; Zhang et al., 2018; Zheng et al., 2015). High RH is not only an indicator of stagnation but also provides a high aerosol water content for secondary $PM_{2.5}$ formation to take place (Song et al., 2018; Wang et al., 2014; Woo & McNeill, 2015). RH and V850 are correlated, but the combination of the two provides a better predictor of Beijing haze than either variable alone (Shen et al., 2018). Shen et al. (2018) showed that the Beijing airport RH data are more reliable than the National Center for Environmental Prediction and National Center for Atmospheric Research reanalysis and are strongly correlated with weather station data at other sites in the North China Plain. All meteorological variables are used as 24-hr averages matched to the $PM_{2.5}$ observations.

Figure 1 shows the PDF of wintertime 24-hr average $PM_{2.5}$ in Beijing for 2009–2017. We fit the high tail of the PDF to a Poisson PP model. EVT holds that if we consider a sufficiently large number of independent, identically distributed measurements of a random variable X and restrict our attention to the codomain $[u, \infty)$ for a sufficiently large threshold u, the PDF of $X \in [u, \infty)$ converges to a PP model regardless of the original underlying distribution of X (Coles, 2001). Our PP model simulates the Poisson process limit of 24-hr average $PM_{2.5}$ in Beijing above a $u = 300$ $\mu g/m^3$ threshold, representing the 95th percentile of the data (Figure 1).

We can fit the PP model to different meteorological predictors. Consider a fit PP(V850, RH) to V850 (v) and RH (r), the two meteorological variables identified in Shen et al. (2018) as the strongest predictors of Beijing haze. Let y denote 24-hr average $PM_{2.5}$. The probability that y will exceed the threshold u given (v, r) is modeled with the marginal distribution

Figure 2. Probabilities of exceedance of extreme 24-hr average PM$_{2.5}$ thresholds in wintertime (December-January-February) Beijing haze, as described by an extreme value point process model PP(V850, RH) conditioned on 850-hPa meridional wind velocity (V850, positive northward) and relative humidity (RH). Colored contours in the top panel show the model probability of PM$_{2.5}$ exceeding 300 μg/m^3 as a function of V850 and RH. Symbols show the 2009–2017 observations of 24-hr average V850 and RH, with observed exceedances of the 300 μg/m^3 threshold indicated in green and other (nonexceedance) observations in black. The bottom panel shows the ability of the same point process model to reproduce the observed probability of occurrence of higher thresholds. Predicted exceedance probability is computed by averaging the probabilities calculated in equation (1) for each of the 719 observed days as we vary the threshold u from 300 to 577 μg/m^3. The identity line is plotted for reference.

$$P(y>u|v,r) = \frac{1}{n_a}\left(1 + \xi\left(\frac{u-u_{v,r}}{\phi_v}\right)\right)^{-\frac{1}{\xi}}, \quad (1)$$

$$\mu_{v,r} = av + br + c, \quad (2)$$

$$\phi_v = e^{dv+f}. \quad (3)$$

Here $\mu_{v,\,r}$ is the location parameter written as a linear function of v and r, ϕ_v (>0) is the scale parameter written as an exponential function of v, and ξ is the shape parameter which describes the shape of the distribution rather than shifting it (as $\mu_{v,\,r}$ does) or stretching it (as ϕ_v does). n_a represents the number of observations in a year (90, excluding leap days). a, b, c, d, f, and ξ are model parameters to be optimized. This optimization is done by maximizing the likelihood function L (Coles, 2001):

$$L(\mu_{v,r}, \phi_v, \xi) = \underbrace{\exp\left(-\frac{1}{n_a}\sum_{t=1}^{n}\left(1 + \frac{\xi(u-\mu_{v,r})}{\phi_v}\right)^{-\frac{1}{\xi}}\right)}_{(a)} \quad (4)$$

$$\underbrace{\prod_{1}^{n}\left(\frac{1}{\phi_v}\left(1 + \frac{\xi(y_t-\mu_{v,r})}{\phi_v}\right)^{-\frac{1}{\xi}-1}\right)^{I(y_t>u)}}_{(b)},$$

where the summation is done over all observations (y_t, v_t, r_t) at times $t \in [1, n]$ from 2009 to 2017. Here $I(y_t > u)$ is the indicator function equal to 1 if $y_t > u$ and 0 otherwise. L is a product of two factors: factor (b) represents exceedances ($y_t > u$) explicitly as a product of independent generalized Pareto densities, while factor (a) incorporates information from all observations, even those that are not extreme. L is maximized when $\mu_{v,\,r}$, ϕ_v, and ξ best explain the data. The optimization is performed with extRemes, an EVT package in the statistical software R (Gilleland & Katz, 2011). The forms of $\mu_{v,\,r}$ and ϕ_v in equations (2) and (3) were chosen after testing a range of possible forms (such as $\phi = e^{dv + fr + g}$) to maximize the value of L subject to the Akaike Information Criterion (AIC). The AIC penalizes models with a large number of optimized parameters p and rewards those with a high maximum likelihood value L:
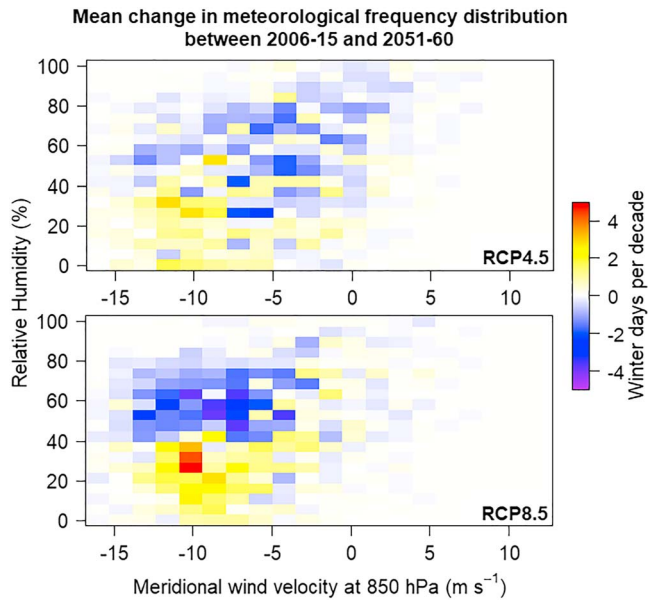
$$\text{AIC} = 2p - 2\ln L. \quad (5)$$

The model with the lowest AIC is preferred (Akaike, 1974).

In some EVT studies (e.g., Rieder et al., 2010), declustering as suggested in Davison and Smith (1990) is applied to data above a threshold. A cluster would be defined here as a series of consecutive days where PM$_{2.5}$ exceeds a threshold u. When declustering is applied, only the maximum of each cluster is used in the EVT analysis. We applied declustering as a sensitivity test to our PP (V850, RH) model and found that it applied greater weight to RH by preferring forms (μ_r, ϕ, ξ), (μ_r, ϕ_v, ξ), and (μ_r, ϕ_r, ξ) over our standard model form ($\mu_{v,\,r}, \phi_v, \xi$). This is because low-wind conditions have greater persistence than high-RH conditions. We choose not to use declustering in our standard model because of the limited size of our data set.

3. Results and Discussion

Figure 2 (top panel) shows the model probabilities for PM$_{2.5}$ to exceed the $u = 300$ μg/m^3 threshold as a function of V850 and RH. Maximum probabilities are for low wind (stagnant conditions) and high RH (conducive to PM formation). The observations show a similar pattern. We see from Figure 2 that V850 and RH are only partly correlated in the observations; low-wind conditions can be associated with low RH and

Figure 3. Projected changes in the wintertime (December-January-February) meridional wind velocity at 850 hPa (V850) and relative humidity (RH) over Beijing from 2006–2015 to 2051–2060. Results are from the ensemble of Coupled Model Intercomparison Project Phase 5 climate models reporting daily V850 and RH and for two climate forcing scenarios (RCP4.5 and RCP8.5). Values are changes in the number of winter days per decade for each (V850, RH) bin, computed for individual models and then averaged across models. The scale is the same as for the extreme haze probability map in Figure 2.

extreme events do not occur then. Although the model describes an increased probability of extreme events at high RH and V850, the probability of these events remains relatively low as expected from their extreme nature. For example, the probability of exceeding 300 μg/m^3 when RH > 60% and |V850| < 5 m/s is 24% in the observations and 25% in the model (as compared to 5% in the general population). The model probability of an extreme event increases to above 50% when RH exceeds 85% and V850 is in the range of −3 to +2 m/s. The observations are consistent with this though the sample is small.

An important property of the PP model is threshold invariance. Although we fit the model with $u = 300$ μg/m^3, EVT predicts that the same parameters $\mu_{v, r}$, ϕ_v, and ξ should yield accurate estimates for any value $u > 300$ μg/m^3 in equation (1) (Coles, 2001). This is verified in the bottom panel of Figure 2, and the implication is that the model describes not only the probability of exceeding 300 μg/m^3 but also the PDF above that threshold.

We also attempted to fit our EVT model to all four meteorological covariates identified in section 2 for Beijing haze: V850, RH, δU500, and δT$_{850-250}$. However, this was unsuccessful because of excessive correlations between the variables. To get around this problem, we performed a principal component analysis of the four variables and found that the first two principal components account for 90% of the variance. Thus, we fit the PP model to just these two components. The first principal component gives an almost-equal weighting to δU500, V850, and δT$_{850-250}$, with RH counting for slightly less; the second principal component is dominated by RH. Best performance for the EVT model takes the location parameter as a linear combination of the first two principal components and the scale parameter as an exponential function of the first principal component. The resulting probability heatmap and threshold exceedance plot for this PP(V850, RH, δU500, δT$_{850-250}$) model are given in supporting information Figure S1.
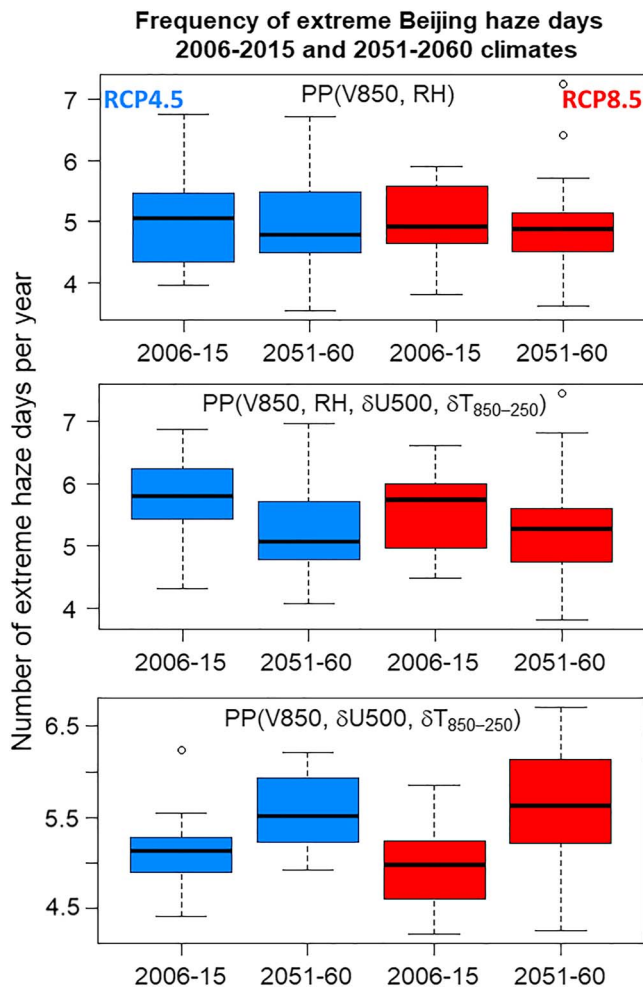
For a more direct comparison with Cai et al. (2017), we fit yet a third model using the three covariates identified in that paper (V850, δU500, and δT$_{850-250}$). A principal component analysis shows that the first principal component accounts for 84% of the variance. We fit our PP(V850, δU500, δT$_{850-250}$) model with just that vector as a covariate. The location parameter is again a linear function, and the scale parameter is an exponential.

By calculating and comparing the AIC for these three candidate PP models (AIC$_i$ with $i \in \{1, 2, 3\}$), we can determine which model i minimizes information loss. We do this by comparing Aikake weights w_i, given by

$$w_i = \frac{\exp\left(\frac{1}{2}\left(\text{AIC}_{\min} - \text{AIC}_i\right)\right)}{\sum_{j=1}^{3} \exp\left(\frac{1}{2}\left(\text{AIC}_{\min} - \text{AIC}_j\right)\right)}, \tag{6}$$

where AIC$_{\min} = \min(\text{AIC}_i)$. The Aikake weights represent the relative likelihood of a model given the data and for our three models are 0.87, 0.13, and 0.00, respectively, for PP(V850, RH), PP(V850, RH, δU500, δT$_{850-250}$), and PP(V850, δU500, δT$_{850-250}$). This provides strong evidence for preferring the PP(V850, RH) model, a conclusion supported by previous work pointing to V850 and RH as the most powerful predictors of Beijing haze (Shen et al., 2018; Zheng et al., 2015).

We can now project the effect of 21st century climate change on the frequency of extreme Beijing haze events by applying our PP(V850, RH) model to daily meteorological projections from the ensemble of global climate models participating in the IPCC Coupled Model Intercomparison Project (CMIP5, see Table S1 for details). We use the 18 models that report daily V850 and RH data for future climate scenarios. We analyze two climate change scenarios, RCP8.5 (business-as-usual) and RCP4.5 (more moderate climate forcing). RH in the CMIP5 models is sampled as a spatial average over the same region as V850 (30– 47.5° N, 115–130° E) to be

Figure 4. Change in the number of extreme Beijing haze days per year for the 2051–2060 versus 2006–2015 climate as predicted from extreme value theory using climate projections from 18 Coupled Model Intercomparison Project Phase 5 models. Extreme haze days are defined by 24-hr average $PM_{2.5} > 300 \mu g/m^3$ and are assumed to occur in winter (December-January-February) only. The three panels show predictions from different Poisson point process (PP) models fit to different combinations of local meteorological variables including meridional wind velocity at 850 hPa (V850), relative humidity (RH), meridional gradient in 500-hPa zonal wind velocity ($\delta U500$), and vertical temperature difference between 850 and 250 hPa ($\delta T_{850-250}$). The boxplots show statistics for the ensemble of Coupled Model Intercomparison Project Phase 5 models: medians, 25th and 75th percentiles, and ranges. All changes with the exception of RCP8.5 in the topmost panel are statistically significant at the 95% level (see text).

more consistent with the scale resolved by climate projections. Results are similar when RH is sampled only for the Beijing grid cells, as reported below. We analyze projections of climate change from 2006–2015 to 2051–2060 as sufficiently distant to obtain a significant signal while sufficiently near to be relevant to air quality management. We also show projections to 2090–2099 in supporting information S1.

Present-day RH statistics (mean and standard deviation) in the CMIP5 models often differ from the Beijing airport data used to fit our PP model. For consistency in application of the PP model, we adjust all four of the 2006–2015 meteorological variables for each CMIP5 model to match the observed means and standard deviations used in the PP model, and we apply the same adjustment to the CMIP5 model values through to 2060. We execute this adjustment by multiplying the model meteorological data by a constant factor to adjust the standard deviation and adding a constant factor to adjust the mean.

Figure 3 shows the two-dimensional distribution of mean projected changes of DJF V850 and RH for the ensemble of models under the two scenarios of climate forcing and on the same scale as Figure 2. Here we take the change from the 2006–2015 period to 2051–2060 for each model and then average these changes over the models. We find a significant decrease in RH from 2006–2015 to 2051–2060 but no significant change in V850. These trends are broadly consistent with projected climate trends for eastern China as a whole (Shen et al., 2018).

Previous studies of 21st century climate change have projected a general decrease in RH over land due to faster warming of the land surface relative to the ocean and poleward expansion of the Hadley cell (Byrne & O'Gorman, 2013; Fu & Feng, 2014; Lau & Kim, 2015). This effect is apparent in 1973–2016 trends from the Hadley Center observation data (Figure S3), which show a prevailing decrease of RH over China including Beijing. The CMIP5 model simulations for the same 1973–2016 period also show a decrease over most of China but an increase over parts of eastern China, with a slight decrease over Beijing. The more consistent CMIP5 projections of decreasing RH by mid-21st century likely reflect the stronger greenhouse forcing.

There is however an important difference in the pattern of RH decrease between the two scenarios. In RCP8.5, the decrease in RH does not affect the subdomain of low V850 and high RH with the highest probability of extreme haze as modeled in Figure 2, whereas it does in RCP4.5. The trends in Figure 3 computed for individual CMIP5 climate models can be combined with the extreme haze probabilities in Figure 2 to project the effect of climate change on extreme haze. For a given model, we calculate the expected number of extreme days for a given year by summing the probabilities for individual days as calculated by equation (1) over the 90

DJF days. Results are shown in Figure 4 (top panel) for the ensemble of CMIP5 models. The model medians show a decrease in both the RCP4.5 and RCP8.5 scenarios. By calculating a 2006–2060 linear regression of the annual number of haze days averaged over the ensemble of models, we find that the decrease in the RCP4.5 scenario is significant at the 5% level and amounts to 0.58 ± 0.20 fewer extreme haze days per year in 2060 than in 2006, representing a 10% decrease. The decrease in the RCP8.5 scenario is not statistically significant. In the RCP4.5 scenario, 69% of the CMIP5 models show a decrease in the frequency of haze days between 2006 and 2060. In the RCP8.5 scenario, 56% of the models show a decrease.

The middle and bottom panels of Figure 4 show results from the alternative but inferior PP models. The PP(V850, RH, $\delta U500$, $\delta T_{850-250}$) model shows significant decreases in the frequency of extreme haze days

by 2060 for both the RCP4.5 and RCP8.5 scenarios, by 0.87 ± 0.49 days (10%) and 0.54 ± 0.39 days (10%), respectively, for the two scenarios. By contrast, the PP(V850, δU500, δT$_{850-250}$) model shows a significant increase for both scenarios, by 0.87 ± 0.27 days (20%) for RCP8.5. This is consistent with Cai et al. (2017), who found from the same three variables in the RCP8.5 scenario a 50% increase in the frequency of haze days $>150\ \mu g/m^3$ between the 1950–1999 and 2050–2099 climates (as compared to 2006–2015 vs. 2051–2060 in our work). Projections to 2100 show similar but amplified patterns (see Figure S2). However, we cannot do an exact comparison of our results to Cai et al. (2017) because 26% of winter days have PM$_{2.5}$ concentrations higher than $150\ \mu g/m^3$ (Figure 1), so an extreme value model is not applicable for that threshold.

We repeated the statistical comparison of haze frequency between the 2006–2015 and 2051–2060 climates for two higher thresholds, 350 and $400\ \mu g/m^3$, to check the sensitivity of our results to threshold selection. For all three PP models, the significance and sign of the changes in the frequency of haze days was consistent at these higher thresholds. We also repeated our analysis for RH averaged over the two gridcells containing Beijing alone, rather than the larger (30– 47.5° N, 115–130° E) region, to check the sensitivity of our results to region selection. For all three PP models in both RCP scenarios, the significance and sign of the changes in the frequency of haze days were consistent, with one exception: The PP(V850, RH, δU500, δT$_{850-250}$) model does not project a significant change under the RCP8.5 scenario when RH is averaged over this smaller region, whereas there is a slight decrease projected when RH is averaged over the larger region.

In summary, we find from EVT that the probability of severe wintertime PM$_{2.5}$ pollution events in Beijing (24-hr average PM$_{2.5}$ > $300\ \mu g/m^3$, 95th percentile of the frequency distribution for winters 2009–2017) can be successfully represented by a PP model with meridional wind velocity (V850) and RH as the two meteorological predictors. The probability of extreme events is the highest under low wind and high RH, as would be expected from stagnation and chemistry. RH is an important predictor variable, independent of V850; occurrences of low V850 together with low RH do not lead to extreme PM$_{2.5}$. Alternative PP models including as additional meteorological variables the gradient in 500-hPa zonal wind (δU500) and the vertical temperature gradient (δT$_{850-250}$) are not as successful at predicting extreme events. Application of the PP(V850, RH) model to 2006–2060 CMIP5 simulations of climate change following the RCP4.5 scenario reveals a 10% decrease in the probability of meteorological conditions leading to extreme PM$_{2.5}$. The RCP8.5 scenario shows no significant trend. A PP(V850, RH, δU500, δT$_{850-250}$) model shows significant decreases in extreme PM$_{2.5}$ in both scenarios. RH is generally expected to decrease under 21st century warming land due to increased land-ocean temperature contrast and the poleward expansion of the Hadley circulation. As emissions in China decrease, our extreme events analysis will become relevant to lower PM$_{2.5}$ exceedance thresholds such as the current China Grade 2 PM$_{2.5}$ 24-hr standard of $75\ \mu g/m^3$. We conclude that 21st century climate change is unlikely to increase the frequency of severe PM$_{2.5}$ pollution events in Beijing and is more likely to decrease it.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Byrne, M. P., & O'Gorman, P. A. (2013). Link between land-ocean warming contrast and surface relative humidities in simulations with coupled climate models. *Geophysical Research Letters*, *40*, 5223–5227. https://doi.org/10.1002/grl.50971

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), 257–262. https://doi.org/10.1038/nclimate3249

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer. https://doi.org/10.1007/978-1-4471-3675-0

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B: Methodological*, *52*(3), 393–425. https://doi.org/10.1111/j.2517-6161.1990.tb01796.x

Eastoe, E. F. (2009). A hierarchical model for non-stationary multivariate extremes: A case study of surface-level ozone and NOx data in the UK. *Environmetrics*, *20*(4), 428–444. https://doi.org/10.1002/env.938

Fu, Q., & Feng, S. (2014). Responses of terrestrial aridity to global warming. *Journal of Geophysical Research: Atmospheres*, *119*, 7863–7875. https://doi.org/10.1002/2014JD021608

Gilleland, E., & Katz, R. W. (2011). New software to analyze how extremes change over time. *Eos, Transactions American Geophysical Union*, *92*(2), 13–14. https://doi.org/10.1029/2011EO020001

Horton, D. E., Skinner, C. B., Singh, D., & Diffenbaugh, N. S. (2014). Occurrence and persistence of future atmospheric stagnation events. *Nature Climate Change*, *4*(8), 698–703. https://doi.org/10.1038/nclimate2272

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Lau, W. K. M., & Kim, K.-M. (2015). Robust Hadley Circulation changes and increasing global dryness due to CO2 warming from CMIP5 model projections. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(12), 3630–3635. https://doi.org/10.1073/pnas.1418682112

Leung, D. M., Tai, A. P. K., Mickley, L. J., Moch, J. M., van Donkelaar, A., Shen, L., & Martin, R. V. (2018). Synoptic meteorological modes of variability for fine particulate matter (PM2.5) air quality in major metropolitan regions of China. *Atmospheric Chemistry and Physics*, *18*(9), 6733–6748. https://doi.org/10.5194/acp-18-6733-2018

Li, K., Liao, H., Cai, W., & Yang, Y. (2018). Attribution of anthropogenic influence on atmospheric patterns conducive to recent most severe haze over eastern China. *Geophysical Research Letters*, *45*, 2072–2081. https://doi.org/10.1002/2017GL076570

Martins, L. D., Wikuats, C. F. H., Capucim, M. N., de Almeida, D. S., da Costa, S. C., Albuquerque, T., et al. (2017). Extreme value analysis of air pollution data and their comparison between two large urban regions of South America. *Weather and Climate Extremes*, *18*, 44–54. https://doi.org/10.1016/j.wace.2017.10.004

Rieder, H. E., Fiore, A. M., Horowitz, L. W., & Naik, V. (2015). Projecting policy-relevant metrics for high summertime ozone pollution events over the eastern United States due to climate and emission changes during the 21st century. *Journal of Geophysical Research: Atmospheres*, *120*, 784–800. https://doi.org/10.1002/2014JD022303

Rieder, H. E., Fiore, A. M., Polvani, L. M., Lamarque, J.-F., & Fang, Y. (2013). Changes in the frequency and return level of high ozone pollution events over the eastern United States following emission controls. *Environmental Research Letters*, *8*(1). https://doi.org/10.1088/1748-9326/8/1/014012

Rieder, H. E., Staehelin, J., Maeder, J. A., Peter, T., Ribatet, M., Davison, A. C., et al. (2010). Extreme events in total ozone over Arosa—Part 1: Application of extreme value theory. *Atmospheric Chemistry and Physics*, *10*(20), 10021–10031. https://doi.org/10.5194/acp-10-10021-2010

Shen, L., Jacob, D. J., Mickley, L. J., Wang, Y., & Zhang, Q. (2018). Effect of climate change on winter haze pollution in Beijing: Uncertain and likely small. *Atmospheric Chemistry and Physics Discussions*, 1–14. https://doi.org/10.5194/acp-2018-932

Shen, L., Mickley, L. J., & Gilleland, E. (2016). Impact of increasing heat waves on U.S. ozone episodes in the 2050s: Results from a multimodel analysis using extreme value theory. *Geophysical Research Letters*, *43*, 4017–4025. https://doi.org/10.1002/2016GL068432

Song, S., Gao, M., Xu, W., Sun, Y., Worsnop, D. R., Jayne, J. T., et al. (2018). Possible heterogeneous hydroxymethanesulfonate (HMS) chemistry in northern China winter haze and implications for rapid sulfate formation. *Atmospheric Chemistry and Physics Discussions*, 1–26. https://doi.org/10.5194/acp-2018-1015

Wang, Y., Yao, L., Wang, L., Liu, Z., Ji, D., Tang, G., et al. (2014). Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China. *Science China Earth Sciences*, *57*(1), 14–25. https://doi.org/10.1007/s11430-013-4773-4

Woo, J. L., & McNeill, V. F. (2015). simpleGAMMA v1.0—A reduced model of secondary organic aerosol formation in the aqueous aerosol phase (aaSOA). *Geoscientific Model Development*, *8*(6), 1821–1829. https://doi.org/10.5194/gmd-8-1821-2015

Zhang, Q., Ma, Q., Zhao, B., Liu, X., Wang, Y., Jia, B., & Zhang, X. (2018). Winter haze over North China Plain from 2009 to 2016: Influence of emission and meteorology. *Environmental Pollution*, *242*(Pt B), 1308–1318. https://doi.org/10.1016/j.envpol.2018.08.019

Zheng, G. J., Duan, F. K., Su, H., Ma, Y. L., Cheng, Y., Zheng, B., et al. (2015). Exploring the severe winter haze in Beijing: The impact of synoptic weather, regional transport and heterogeneous reactions. *Atmospheric Chemistry and Physics*, *15*(6), 2969–2983. https://doi.org/10.5194/acp-15-2969-2015

Zou, Y., Wang, Y., Zhang, Y., & Koo, J.-H. (2017). Arctic sea ice, Eurasia snow, and extreme winter haze in China. *Science Advances*, *3*(3), e1602751. https://doi.org/10.1126/sciadv.1602751