# Predicting the impact of climate change on severe winter haze pollution events in Beijing using extreme value theory

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

[1]Harvard College, Cambridge, Massachusetts, USA.

[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA.

Corresponding author: Lu Shen (lshen@fas.harvard.edu)

**Key Points:**

- Likelihood of extreme winter haze pollution in Beijing can be fit to meridional wind and relative humidity using a point process model
- Application of the model to climate projections shows that the frequency of extreme haze events is most likely to decrease by midcentury
- Disagreement with previous work projecting an increase in haze frequency is explained by omission of RH from their models

## Abstract

We use extreme value theory to develop point process (PP) statistical models relating the probability of extreme winter 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 US embassy. We find that 850 hPa meridional wind velocity (V850) and relative humidity (RH) successfully predict days when daily mean $PM_{2.5}$ will exceed 300 μg $m^{-3}$ as well as higher thresholds. We apply the PP models to mid-21$^{st}$ century climate model projections from the CMIP5 ensemble under two forcing scenarios (RCP8.5 and RCP4.5). We find that the frequency of haze events is most likely to decrease because of climate change, driven mainly by a decrease in RH. Our results contrast with previous studies that projected increases in haze frequency but did not include RH as a meteorological predictor.


## 1 Introduction

Beijing in winter experiences very severe air pollution events, commonly referred to as 'winter haze.' Concentrations of fine particulate matter ($PM_{2.5}$) can exceed 300 μg $m^{-3}$ on a 24-h 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 μg $m^{-3}$ in Beijing. They then used climate simulations from the IPCC CMIP5 multimodel ensemble driven by the RCP8.5 (business-as-usual) emissions scenario to project a significant increase in this index over the 21$^{st}$ 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 21$^{st}$-century climate change would worsen winter haze. On the other hand, *Horton et. al.* [2014] using the same CMIP5 RCP8.5 archive find no change between 1986-2005 and 2046-2065 in the frequency of stagnation events over Beijing.

EVT offers a statistical framework to predict rare occurrences [*Coles*, 2001]. Here we model the relationship of extreme haze days in Beijing with 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 Switzerland, the US, and South America [*Rieder et. al.*, 2010; *Shen et. al.*, 2016; *Martins et. al.*, 2017], but has not been applied to Beijing haze to our knowledge.

## 2 Data and Methods

We use 2009-2017 wintertime (DJF) time series of 24-h average $PM_{2.5}$ in Beijing, starting in December 2009 and ending in February 2017, measured at the US embassy and available from (http://www.stateair.net/web/historical/1/1.html). Previous studies have shown that this dataset is representative of $PM_{2.5}$ in Beijing [*San Martini et. al.*, 2015; *Mukherjee and Toohey*, 2016]. *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 (NCEP/NCAR) 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 500hPa zonal wind (δU500) north of Beijing (42.5˚–52.5˚ N, 110˚–137.5˚ E) vs. south of Beijing (27.5˚–37.5˚ N, 110˚–137.5˚ E), and (3) the vertical temperature difference ($\delta T_{850\text{-}250}$) 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). We also consider (4) the relative humidity (RH) at Beijing International airport obtained from the 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 [*Wang et. al.*, 2014; Zheng et. al., 2015; *Lueng et. al.*, 2018; *Shen et. al.*, 2018]. The Beijing airport data is more reliable for local RH than the NCEP/NCAR reanalysis [Shen et al., 2018]. All meteorological variables are used as 24-h averages matched to the $PM_{2.5}$ observations.

Figure 1 shows the PDF of wintertime 24-h average $PM_{2.5}$ in Beijing for 2009-2017. We fit the high tail of the PDF to a Poisson point process (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-hour average $PM_{2.5}$ in Beijing above a $u = 300$ μg m$^{-3}$ threshold, representing the 95$^{th}$ percentile of the data (Figure 1).

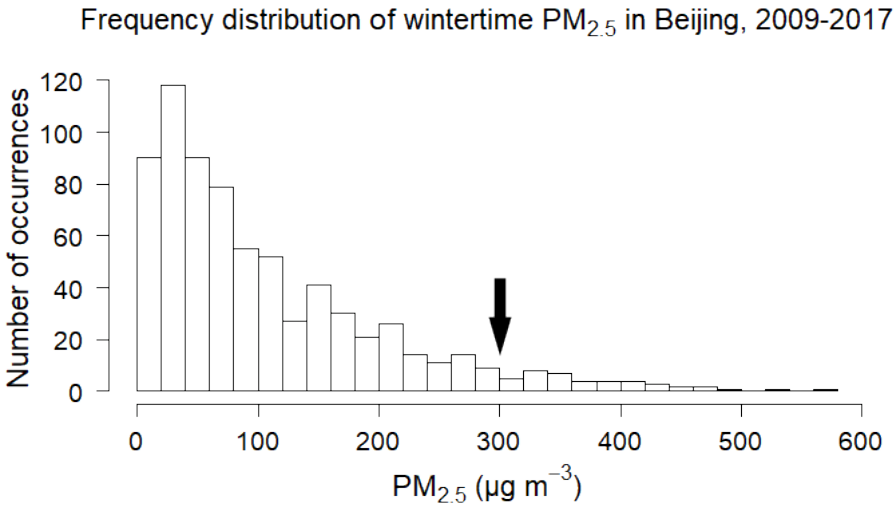


**Figure 1.** Frequency distribution of daily wintertime (DJF) 24-hour average $PM_{2.5}$ in Beijing for 2009-2017, starting in December 2009 and ending in February 2017. Data are from the US embassy site. The arrow indicates the 300 μg m$^{-3}$ threshold used in our

84 definition of extreme events and representing the 95$^{th}$ percentile of the frequency
85 distribution.

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

90 $$P(y > u \mid v, r) = \frac{1}{n_a}\left(1 + \xi\left(\frac{u - \mu_{v,r}}{\phi_v}\right)\right)^{-\frac{1}{\xi}} \quad (1)$$

91 with

92 $$\mu_{v,r} = av + br + c \quad (2)$$

93 $$\phi_v = e^{dv+f} \quad (3)$$

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

100 $$L(\mu_{v,r}, \phi_v, \xi) = \underbrace{\exp\left(-\frac{1}{n_a}\sum_{t=1}^{n}\left(1 + \frac{\xi(u - \mu_{v,r})}{\phi_v}\right)^{-\frac{1}{\xi}}\right)}_{(a)} \underbrace{\prod_{1}^{n}\left(\frac{1}{\phi_v}\left(1 + \frac{\xi(y_t - \mu_{v,r})}{\phi_v}\right)^{-\frac{1}{\xi}-1}\right)^{I(y_t>u)}}_{(b)} \quad (4)$$

101 where the summation is done over all observations $(y_t, v_t, r_t)$ at times $t \in [1, n]$ from 2009 to
102 2017. Here $I(y_t > u)$ is the indicator function equal to 1 if $y_t > u$ and 0 otherwise. *L* is a
103 product of two factors: factor (b) represents exceedances $(y_t > u)$ explicitly as a product of
104 independent generalized Pareto densities, while factor (a) incorporates information from all
105 observations, even those that are not extreme. *L* is maximized when $\mu_{v,r}$, $\phi_v$, and $\xi$ best explain
106 the data. The optimization is performed with extRemes, an EVT package in the statistical
107 software R [*Gilleland and Katz,* 2011].The forms of $\mu_{v,r}$ and $\phi_v$ in equations (2) and (3) were
108 chosen after testing a range of possible forms (such as $\phi = e^{dv+fr+g}$) to maximize the value of
109 *L* subject to the Akaike Information Criterion (AIC). The AIC penalizes models with a large
110 number of optimized parameters *p* and rewards those with a high maximum likelihood value *L*:

111 $$\mathrm{AIC} = 2p - 2\ln L \quad (5)$$

112 The model with the lowest AIC is preferred [*Aikake,* 1974].

113 **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. The model matches the observed occurrences. Maximum probabilities are for near-zero wind (stagnant conditions) and high RH (conducive to PM formation; *Wang et al.* [2014]).

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 $\mu_{v,r}$, $\phi_v$, 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.

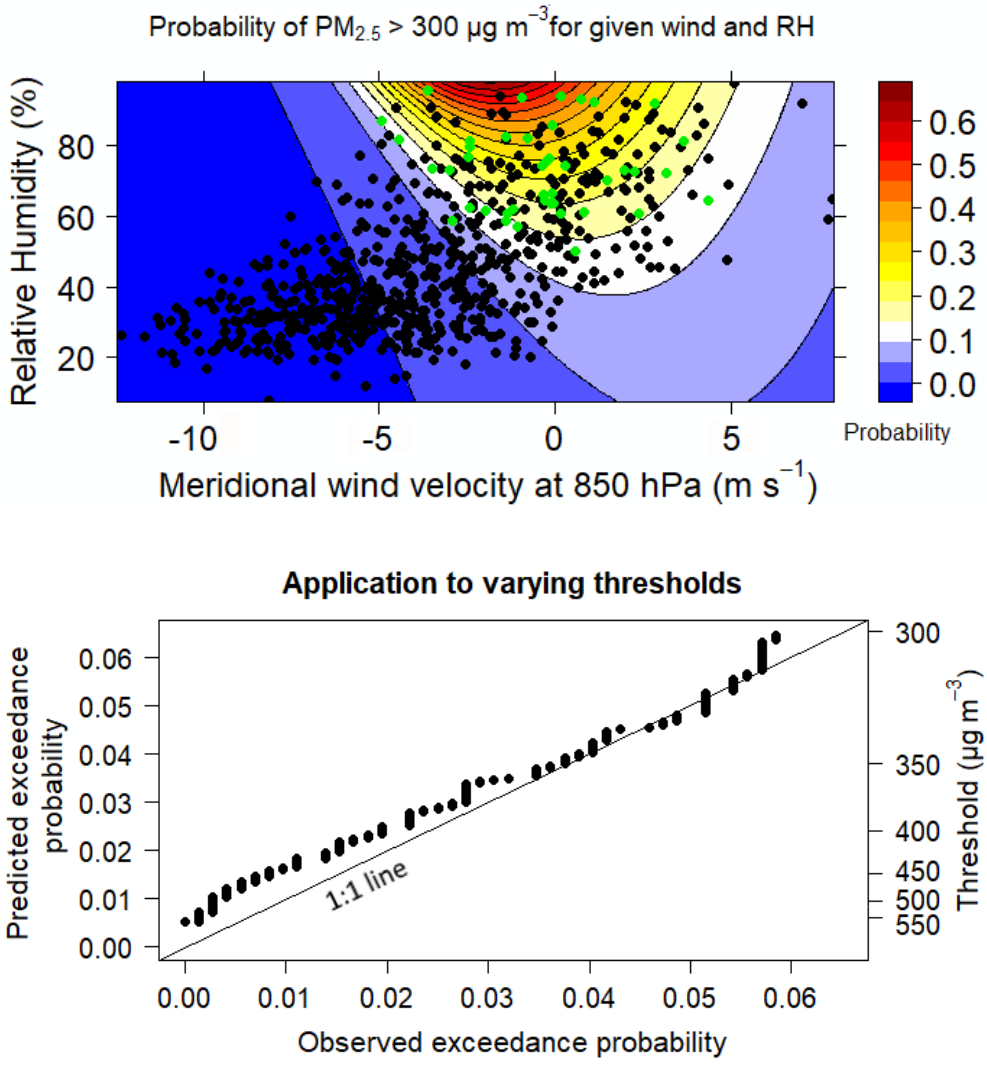


**Figure 2.** Probabilities of exceedance of extreme daily mean PM2.5 thresholds in wintertime (DJF) Beijing haze, as described by an extreme value point process model PP(V850, RH) conditioned on 850 hPa meridional wind velocity (V850, positive northward) and relative humidity (RH). Colored contours in the top panel shows the model probability of PM2.5 exceeding 300 μg m$^{-3}$ as a function of V850 and RH. Symbols show the 2009-2017 observations, with observed exceedances of the 300 μg m$^{-3}$

threshold indicated in green. Each symbol represents average V850 and RH values on one day. The bottom panel shows the ability of the same model to reproduce the observed probability of occurrence of higher thresholds. The identity line is plotted for reference.

We also attempted to fit our EVT model to all four meteorological covariates identified in **Section 2** for Beijing haze: V850, RH, δU500, and $\delta T_{850\text{-}250}$. However, this was unsuccessful because of strong correlations between the variables. To get around this problem, we performed a principal component analysis (PCA) 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 $\delta T_{850\text{-}250}$, 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, $\delta T_{850\text{-}250}$) model are given in **Figure S1.**

For direct comparison with *Cai et. al.*, [2017], we fit yet a third model using the three covariates identified in that paper (V850, δU500, and $\delta T_{850\text{-}250}$). A PCA shows that the first principal component accounts for 84% of the variance. We fit our PP(V850, δU500, $\delta T_{850\text{-}250}$) 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 ($\mathrm{AIC}_i$ 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

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

where $\mathrm{AIC}_{\min}$ = min ($\mathrm{AIC}_i$). 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, $\delta T_{850\text{-}250}$), and PP(V850, δU500, $\delta T_{850\text{-}250}$). 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 [*Zheng et. al*., 2015; *Shen et. al.*, 2018].

We can now project the effect of 2006-2060 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 18 global climate models participating in the IPCC Coupled Model Intercomparison Project (CMIP5, see **Table S1** for details) that report daily V850 and RH data. To obtain a timeseries for future RH, we average spatially over the same region used to calculate V850 (30˚–47.5˚ N, 115˚–130˚ E). Present-day RH statistics in these CMIP5 models often differ significantly from the Beijing airport data used to fit our PP model. For consistency in application of the PP model, we adjust 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.

**Figure 3** shows the two-dimensional frequency 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 2006-15 period to 2051-60 for each model and then average these changes over the models. Results are shown for the RCP4.5 and RCP8.5 radiative scenarios, where the RCP4.5 scenario describes a more moderate climate forcing than the RCP8.5 business-as-usual scenario. We use 18 models from the RCP8.5 scenario and 16 from RCP4.5, reflecting the variables available.

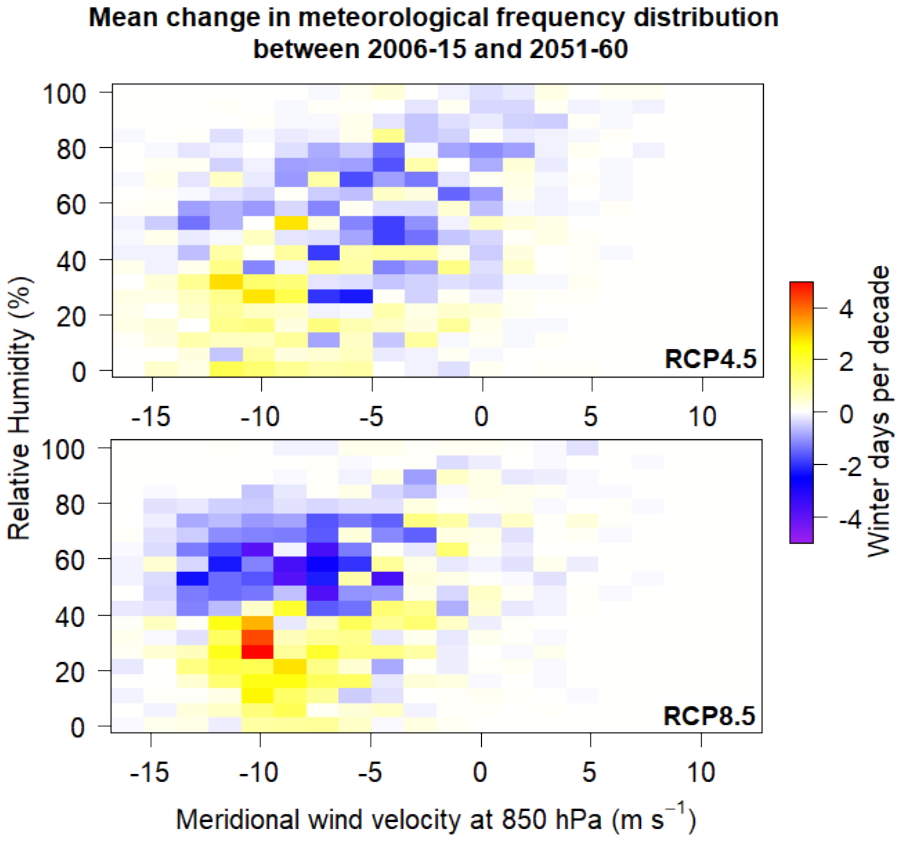


**Figure 3.** Projected changes in the wintertime (DJF) meridional wind velocity at 850 hPa (V850) and relative humidity (RH) over Beijing from 2006-2015 to 2051-2060. Results are from the ensemble of CMIP5 climate models reporting daily V850 and RH, and for two climate forcing scenarios (RCP4.5 and 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.

We find in Figure 3 a significant decrease in RH from 2006-2015 to 2015-2060 but no significant change in V850. These trends are broadly consistent with trends in eastern China as a whole [*Shen et al.*, 2018]. The decrease in RH under greenhouse warming may be expected from increased land-ocean temperature contrast and poleward expansion of

the Hadley cell [Byrne et al., 2013; Lau et al., 2015]. There is however substantial variability between models including in the sign. For the RCP4.5 scenario 63% of the CMIP5 models project a decrease in mean RH, and 75% project a decrease in the frequency of days where RH is greater than 80%. For the RCP8.5 scenario 78% project a decrease in mean RH, and 61% project a decrease in the frequency of RH greater than 80%.

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 the PP(V850, RH) model, which we have shown above to be most reliable. 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 the decrease in the RCP4.5 scenario is significant at the 5% level and amounts to $0.58 \pm 0.39$ fewer 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. In the RCP8.5 scenario, 56% of the models show a decrease.

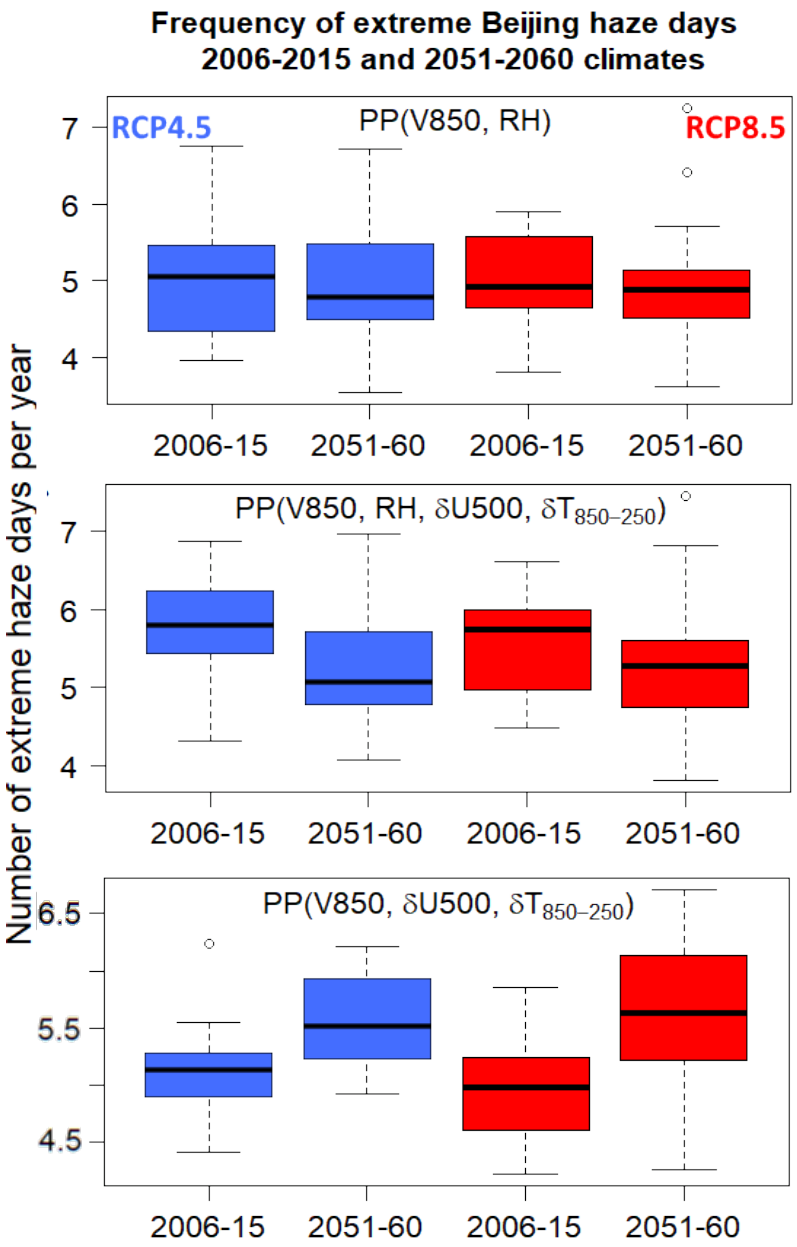


**Figure 4.** Change in the number of extreme Beijing haze days per year for the 2051-2060 versus 2006-2015 climate as predicted from extreme value theory using climate projections from 18 CMIP5 models. Extreme haze days are defined by 24-hour average $PM_{2.5}$ > 300 μg $m^{-3}$ and are assumed to occur in winter (DJF) 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 (δU500), and vertical temperature difference between 850 and 250 hPa ($\delta T_{850-250}$). The boxplots show statistics for the ensemble of CMIP5 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)

The middle and bottom panels of Figure 4 show results from the alternative PP models. The PP(V850, RH, δU500, $\delta T_{850\text{-}250}$) model shows significant decreases in the frequency of haze days by 2060 for both the RCP4.5 and RCP8.5 scenarios, by $-0.87 \pm 0.25$ days (10%) and $-0.54 \pm 0.20$ days (10%) respectively for the two scenarios. By contrast, the PP(V850, δU500, $\delta T_{850\text{-}250}$) model shows a significant increase for both scenarios, by $0.87 \pm 0.27$ 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 μg $m^{-3}$ between the 1950-1999 and 2050-2099 climates (as compared to 2006-2015 vs. 2051-2060 in our work). As pointed out above, a major flaw of this PP(V850, δU500, $\delta T_{850\text{-}250}$) model is that it does not include RH as a predictor variable.

We repeated the statistical comparison of haze frequency between the 2006-2015 and 2051-2060 climates for two higher thresholds, 350 and 400 μg $m^{-3}$, 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.

In summary, we find from extreme value theory that the probability of severe wintertime haze events in Beijing (24-h average $PM_{2.5}$ > 300 μg $m^{-3}$) can be successfully represented by a point process (PP) model with meridional wind velocity (V850) and relative humidity (RH) as the two meteorological predictors. The probability of extreme events is highest under near-zero V850 and high RH, as would be expected from stagnation and conditions conducive to aqueous aerosol formation. Alternative PP models including as additional meteorological variables the gradient in 500 hPa zonal wind (δU500) and the vertical temperature gradient ($\delta T_{850\text{-}250}$) are not as successful. RH is a particularly important predictor. 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 extreme events given constant emissions; the decrease is driven mainly by decreasing RH. The RCP8.5 scenario shows no significant trend. A PP(V850, RH, δU500, $\delta T_{850\text{-}250}$) model shows significant decreases in haze frequency in both scenarios. In contrast, a PP(V850, δU500, $\delta T_{850\text{-}250}$) model not including RH as predictor variable shows an increase in haze days for both the RCP4.5 and RCP8.5 scenarios, consistent with the results of Cai et al. [2017]. However, we find in evaluating model success that RH needs to be included as a predictor variable since extreme haze events are strongly associated with high RH. Increasing greenhouse forcing may be expected to cause a decrease of RH over China due to increased land-ocean temperature contrast and the poleward expansion of the Hadley circulation. We conclude that 21st-century climate change should not increase the frequency of extreme haze events in Beijing and is more likely to decrease it.

## Acknowledgments

This work was funded by the Harvard Global Institute (HGI), with additional funding to DCP by a Harvard College Research Program (HCRP) fellowship and a Harvard College Program for Research in Science and Engineering (PRISE) fellowship. All data is freely available from the sources noted in the **Data and Methods** section. We thank the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups listed in **Table S1**. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating

support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

We thank Eric Gilleland for helpful discussion.

## 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

Burnham, K. P., & Anderson, D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York, NY. https://doi.org/10.1007/978-0-387-22456-5_1

Byrne, M. P., & O'Gorman, P. A. (2016). Understanding Decreases in Land Relative Humidity with Global Warming: Conceptual Model and GCM Simulations. Journal of Climate, 29(24), 9045–9061. https://doi.org/10.1175/JCLI-D-16-0351.1

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

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), 2015JD023225. https://doi.org/10.1002/2015JD023225

Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values. Springer London.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, *7*(1), 1–26. https://doi.org/10.1214/aos/1176344552

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, 698–703. https://doi.org/10.1038/nclimate2272

Jia, B., Wang, Y., Yao, Y., & Xie, Y. (2015). A new indicator on the impact of large-scale circulation on wintertime particulate matter pollution over China. Atmos. Chem. Phys., 15(20), 11919–11929. https://doi.org/10.5194/acp-15-11919-2015

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., … Joseph, D. (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. Atmos. Chem. Phys., 18(9), 6733–6748. https://doi.org/10.5194/acp-18-6733-2018

Martins, L. D., Wikuats, C. F. H., Capucim, M. N., de Almeida, D. S., da Costa, S. C., Albuquerque, T., … Martins, J. A. (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

Mukherjee, A., & Toohey, D. W. (2016). A study of aerosol properties based on observations of particulate matter from the U.S. Embassy in Beijing, China. Earth's Future, 4(8), 2016EF000367. https://doi.org/10.1002/2016EF000367

Ouyang, Y. (2013). China wakes up to the crisis of air pollution. The Lancet Respiratory Medicine, 1(1), 12. https://doi.org/10.1016/S2213-2600(12)70065-6

Peplow, M. (2014). Beijing smog contains witches' brew of microbes. Nature News. https://doi.org/10.1038/nature.2014.14640

Rieder, H. E., Staehelin, J., Maeder, J. A., Peter, T., Ribatet, M., Davison, A. C., … Holawe, F. (2010). Extreme events in total ozone over Arosa – Part 1: Application of extreme value theory. Atmos. Chem. Phys., 10(20), 10021–10031. https://doi.org/10.5194/acp-10-10021-2010

San Martini, F. M., Hasenkopf, C., & Roberts, D. C. (2015). Statistical analysis of PM2.5 observations from diplomatic facilities in China. Atmospheric Environment, 110. https://doi.org/10.1016/j.atmosenv.2015.03.060

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(8), 2016GL068432. https://doi.org/10.1002/2016GL068432

Shen, L., Jacob, D. J., Mickley, L. J., Wang, Y., Zhang, Q. (2018). Uncertain effect of climate change on winter haze pollution in Beijing. Nature Communications. Submitted.

Wang, Y., Yao, L., Wang, L., Liu, Z., Ji, D., Tang, G., … Xin, J. (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

Xu, P., Chen, Y., & Ye, X. (2013). Haze, air pollution, and health in China. The Lancet, 382(9910), 2067. https://doi.org/10.1016/S0140-6736(13)62693-8

Zhang, R., Jing, J., Tao, J., Hsu, S.-C., Wang, G., Cao, J., … Shen, Z. (2013). Chemical characterization and source apportionment of PM2.5 in Beijing: seasonal perspective. Atmos. Chem. Phys., 13(14), 7053–7074. https://doi.org/10.5194/acp-13-7053-2013

Zheng, G. J., Duan, F. K., Su, H., Ma, Y. L., Cheng, Y., Zheng, B., … He, K. B. (2015). Exploring the severe winter haze in Beijing: the impact of synoptic weather, regional transport and heterogeneous reactions. Atmos. Chem. Phys., 15(6), 2969–2983. https://doi.org/10.5194/acp-15-2969-2015

Zhong, J., Zhang, X., Dong, Y., Wang, Y., Liu, C., Wang, J., … Che, H. (2018). Feedback effects of boundary-layer meteorological factors on cumulative explosive growth of

342 PM2.5 during winter heavy pollution episodes in Beijing from 2013 to 2016. Atmos.
343 Chem. Phys., 18(1), 247–258. https://doi.org/10.5194/acp-18-247-2018

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