

## Multi-Index Attribution of Extreme Winter Air Quality in Beijing, China

Christopher W. Callahan<sup>1,2</sup> , Jordan L. Schnell<sup>3</sup>, and Daniel E. Horton<sup>3</sup> 

## Key Points:

- Extreme air quality conditions present in Beijing during January 2013 represented by three indices of air quality meteorology
- Natural variability found to have a larger influence on extreme air quality events than anthropogenic forcing
- Anthropogenic forcing drove long-term positive trends in two out of three indices, despite small effect on individual events

## Correspondence to:

C. W. Callahan,  
christophercallahan2018@u.northwestern.edu

## Citation:

Callahan, C. W., Schnell, J. L., & Horton, D. E. (2019). Multi-index attribution of extreme winter air quality in Beijing, China. *Journal of Geophysical Research: Atmospheres*, 124, 4567–4583. <https://doi.org/10.1029/2018JD029738>

Received 27 SEP 2018

Accepted 2 APR 2019

Accepted article online 9 APR 2019

Published online 30 APR 2019

## Author Contributions:

**Conceptualization:** Christopher W. Callahan, Daniel E. Horton

**Formal analysis:** Christopher W. Callahan

**Methodology:** Christopher W. Callahan, Jordan L. Schnell, Daniel E. Horton

**Supervision:** Jordan L. Schnell, Daniel E. Horton

**Writing - original draft:** Christopher W. Callahan

**Writing - review & editing:** Jordan L. Schnell, Daniel E. Horton

<sup>1</sup>Program in Environmental Sciences, Northwestern University, Evanston, IL, USA, <sup>2</sup>Program in Ecology, Evolution, Ecosystems, and Society, Dartmouth College, Hanover, NH, USA, <sup>3</sup>Institute for Sustainability and Energy at Northwestern and Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL, USA

**Abstract** High-impact poor air quality events, such as Beijing's so-called "Airpocalypse" in January 2013, demonstrate that short-lived poor air quality events can have significant effects on health and economic vitality. Poor air quality events result from the combination of the emission of pollutants and meteorological conditions favorable to their accumulation, which include limited scavenging, dispersion, and ventilation. The unprecedented nature of events such as the 2013 Airpocalypse, in conjunction with our nonstationary climate, motivate an assessment of whether climate change has altered the meteorological conditions conducive to poor winter air quality in Beijing. Using three indices designed to quantify the meteorological conditions that support poor air quality and drawing on the attribution methods of Diffenbaugh et al. (2017, <https://doi.org/10.1073/pnas.1618082114>), we assess (i) the contribution of observed trends to the magnitude of events, (ii) the contribution of observed trends to the probability of events, (iii) the return interval of events in the observational record, preindustrial model-simulated climate and historical model-simulated climate, (iv) the probability of the observed trend in the preindustrial and historical model-simulated climates, and (v) the relative influences of anthropogenic forcing and natural variability on the observed trend. We find that anthropogenic influence has had a small effect on the probability of the January 2013 event in all three indices but has increased the probability of a long-term positive trend in two out of three indices. This work provides a framework for both further understanding the role of climate change in air quality and expanding the scope of event attribution.

## 1. Introduction

Beijing has recently suffered a series of poor air quality episodes, including the January 2013 "Airpocalypse," so-called due to the unprecedented observed concentrations of PM<sub>2.5</sub> (Beech, 2013; Kaiman, 2013; Zhang et al., 2013). These episodes have dramatically impacted economic vitality and human mortality and morbidity; high particulate matter concentrations have been linked to hundreds of thousands of deaths and millions of respiratory tract infections across China between 2010 and 2015 (Li et al., 2016; Yin et al., 2017). Episodes of poor air quality, or haze, occur when pollutants and their precursors enter the atmosphere from anthropogenic and/or natural sources and the overlying meteorology promotes their accumulation (Dawson et al., 2007; Jacob & Winner, 2009; Leung & Gustafson, 2005). Researchers have hypothesized that the recent haze episodes in Beijing have been significantly influenced by stagnant conditions, characterized by a weakening of the East Asian trough and anomalous southerly winds that advect moisture and pollutants north into Beijing (Chen & Wang, 2015; Guo et al., 2014). Combined with near-surface temperature inversions which also limit vertical ventilation, stagnant meteorology can produce extreme winter air quality even without an increase in pollutant emissions (Hou & Wu, 2016; Sun et al., 2014; Wang et al., 2014; Zhang et al., 2018; Zheng et al., 2015). In particular, meteorological extremes tend to co-occur with extremes in pollutant concentrations, in part driven by stagnant synoptic-scale circulation (Schnell & Prather, 2017). Finally, researchers have projected potential increases in these stagnant meteorological conditions due to anthropogenic climate change, both in China (Cai et al., 2017; Zou et al., 2017) and globally (Dawson et al., 2014; Horton et al., 2012; Horton et al., 2014; Mickley et al., 2004; Murazaki & Hess, 2006).

Given the attention placed on recent extreme poor air quality events and the projected role of climate change in worsening the meteorology that shapes these events, it is appropriate to ask if climate change has played a role in recent occurrences of extreme air quality events in China. For example, Zou et al. (2017) linked recent declines in Arctic sea ice and increases in Eurasian snow cover to the

intensification of stagnant wind speeds and near-surface temperature inversions. Wang et al. (2015) described a possible mechanism for this link: Reductions in Arctic sea ice produce positive sea-level pressure anomalies over Asia, which shift the track of regional cyclone activity to the north and produce stagnant conditions defined by a stable atmosphere and low cyclonic activity in eastern and northeastern China. Climate-change-driven decreases in Arctic sea ice may intensify this link (Zou et al., 2017). Other analyses, however, have argued that climate change has had only a negligible impact on air quality in Beijing (e.g., Shen et al., 2018), increasing the need for further examination of the link between anthropogenic forcing and stagnant meteorology.

In order to further understand the link between observed climate change and recent air quality extremes, we draw on a set of methods from an area of climate science known as detection and attribution. Detection and attribution studies aim to identify changes in the historical record (detection) and isolate the causes of those changes (attribution). Foundational detection and attribution studies aimed to determine the degree to which anthropogenic forcing contributed to twentieth century global temperature increases (Allen et al., 2007). These studies used a climate model, or an ensemble of models, to compare the observed climate with a hypothetical world without anthropogenic greenhouse gas emissions and determine the degree to which those emissions have contributed to recent trends (Allen et al., 2007; Stott et al., 2000). These studies have been essential in attributing the majority of twentieth century temperature increases to human influence rather than natural variations like solar variability or volcanic eruptions (Allen et al., 2007; Hegerl & Zwiers, 2011; Stott et al., 2000; Stott et al., 2010).

The focus of detection and attribution has expanded to examine extreme events in the historical record, a subfield known as “extreme event attribution” or “single-event attribution.” This research has increased in salience as policymakers and stakeholders desire more specific and quantifiable information on the relationship between global changes and local extremes, to better inform adaptation decision-making (Sippel et al., 2015). Event attribution is also an important component of “loss and damage” calculations as part of efforts to determine legal liability for the damage climate change has already caused (Allen, 2003; Allen et al., 2007). Much of the framework for extreme event attribution is drawn from the statistics of epidemiology, where distributions of the relevant variables in different scenarios are compared to determine how the likelihood of a given event changes between scenarios (Stone & Allen, 2005). This framework is now used in climate change event attribution studies to compare variables in the observational record and between preindustrial and anthropogenically forced climate model scenarios (e.g., Hegerl, 2015; King, van Oldenborgh, et al., 2015).

At a global level, studies have attributed some component of recent extreme temperature occurrences to anthropogenic warming, though the exact contribution of anthropogenic influence to each event varies by methodology (Christidis et al., 2011; Diffenbaugh et al., 2017; Zwiers et al., 2011). Recent event attribution studies have also examined other local extreme events and found a significant role for anthropogenic warming, and among others, the flood-producing precipitation in Louisiana in 2016 (van der Wiel et al., 2017), extreme precipitation in England in 2000 (Pall et al., 2011), the hot and dry conditions that produced the 2012–2015 California drought (Diffenbaugh et al., 2015), and the 2003 European heat wave (Stott et al., 2004). There is now an annual special issue of the *Bulletin of the American Meteorological Society* dedicated to single-event attribution studies regarding the previous year’s extreme events (e.g., Herring et al., 2017). Recent advances in computing power and attribution methodologies have even allowed “fast-track” attribution of extreme events, where events are examined and attributed almost as soon as they occur, as opposed to months or years afterward (Christidis et al., 2015; Hausteine et al., 2016; Stott et al., 2016).

Single-event attribution studies were initially focused on first-order environmental variables such as temperature and precipitation, and many of the reports in the annual *Bulletin of the American Meteorological Society* issue still focus on various extreme heat or flooding events. However, the science is advancing and now includes other, more complex environmental conditions, such as the meteorology conducive to extremely poor air quality events (e.g., Li et al., 2018; Vautard et al., 2017). Our study therefore provides a useful test case for this expansion of extreme event attribution to more complex, multi-ingredient extreme climate events, while also allowing us to answer the fundamental question: How has anthropogenic climate change influenced recent extreme air quality meteorology in Beijing?

## 2. Data and Methods

### 2.1. Data

We use reanalysis data from the R1 Reanalysis project from the National Center for Environmental Protection (NCEP) and National Center for Atmospheric Research (NCAR), herein referred to as the NCEP/NCAR R1 reanalysis, to represent the observational record (Kalnay et al., 1996). This reanalysis data set has a horizontal resolution of approximately  $2.5^\circ \times 2.5^\circ$ . We select data only from 1979 to 2016, which includes data from the more robust “satellite era,” i.e., the time period in which satellite data has strengthened and provided support for other observations (Kistler et al., 2001). All further mentions of “observations” refer to this reanalysis data.

We use observations of local  $\text{PM}_{2.5}$  concentrations from the U.S. embassy in Beijing (San Martini et al., 2015) to quantify daily air quality in Beijing from 2009 to 2016, shown in Figure 2.

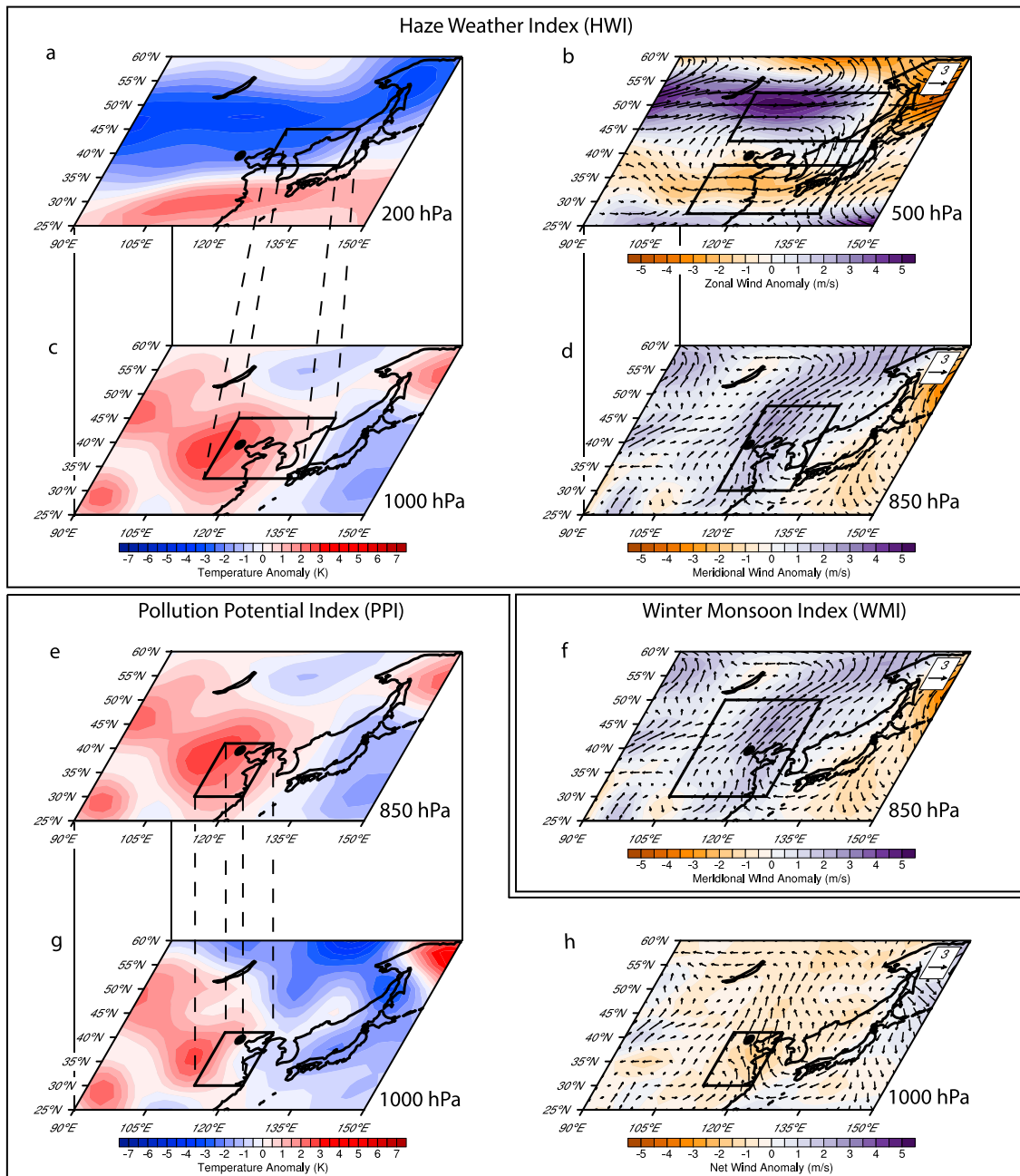
To simulate preindustrial and historical (late 20th century and early 21st century) climates, we use the Community Earth System Model (CESM) Large Ensemble (CESM-LE). CESM-LE contains a single 1,800-year-long preindustrial control run with 1850 initial conditions and constant forcing, a simulation of the historical forcing from 1920 to 2005, and a simulation of the projected future from 2005 to 2100 using the CMIP5 RCP8.5 emissions scenario (Kay et al., 2015; Taylor et al., 2012). CESM-LE uses one model (CESM with CAM5.2) with the exact same parameters, forcings, and boundary conditions but includes 30 different realizations of the historical and future climates. Each realization, branching from 1850 in the case of the first realization, and 1920 in the case of the rest, is initialized with slightly different atmospheric conditions, on the order of  $10^{-14}$  K differences in air temperature (Kay et al., 2015). To simulate the 1979–2016 period, we combine the last 27 years of each historical realization with the first 11 years of the corresponding RCP8.5 realization, and henceforth, we refer to this combination as the “historical simulation.” This version of CESM has a horizontal resolution of approximately  $1^\circ \times 1^\circ$ .

Climate model simulations inherently involve three types of uncertainty: *model uncertainty*, meaning the uncertainty associated with model construction and its response to forcing, *scenario uncertainty*, meaning the uncertainty associated with predicting the future of a given forcing (e.g., greenhouse gas emissions), and *irreducible uncertainty*, meaning the uncertainty due to the natural chaos of the climate system, otherwise known as “internal variability” (Hawkins & Sutton, 2009). Small differences in the initial conditions used to initialize climate model simulations can create cascading and chaotic effects, so using a spread of realizations from a single model with different initial conditions provides an opportunity to assess the influence of internal variability on projections of climate change (Deser et al., 2012; Deser et al., 2014). A single-model multirealization ensemble such as CESM-LE has unique benefits for extreme event detection and attribution. Climate model simulations of extreme events are especially sensitive to internal variability, since they attempt to represent a rare set of events that occur on small temporal and spatial scales (Deser et al., 2012; Diffenbaugh et al., 2017; Fischer & Knutti, 2015), and internal variability dominates the sources of uncertainty in simulations of the historical climate (Hawkins & Sutton, 2009).

### 2.2. Air Quality Indices

We use three peer-reviewed indices to quantify the meteorological conditions that shape air quality in Beijing (Figure 1). These indices take advantage of an ingredients-based meteorological approach, which seeks to identify the fundamental physical components of a phenomenon and combine them to develop a fuller picture of the causes of that phenomenon (Wetzel & Martin, 2001). Given the short timescales and small spatial scales involved, event attribution studies tend to be sensitive to the exact methodology used to quantify an event, so using multiple indices provides a more robust estimate of the influence of anthropogenic forcing (Angelil et al., 2017). We use normalized data in all indices, so all data is presented in units of standard deviations; the normalization process is detailed below for each index. These indices are as follows:

1. Haze Weather Index (HWI). We quantify haze conditions, modifying the approach of Cai et al. (2017), as the sum of the temperature gradient between 200 hPa over  $37.5\text{--}45^\circ\text{N}$  and  $122.5\text{--}137.5^\circ\text{E}$  and 850 hPa over  $32.5\text{--}45^\circ\text{N}$  and  $112.5\text{--}132.5^\circ\text{E}$  (Figures 1a and 1c), the strength of 850 hPa meridional winds over  $30\text{--}47.5^\circ\text{N}$  and  $115\text{--}130^\circ\text{E}$  (Figure 1d), and the latitudinal difference in 500 hPa zonal winds between  $42.5\text{--}52.5^\circ\text{N}$ ,  $110\text{--}137.5^\circ\text{E}$  and  $27.5\text{--}37.5^\circ\text{N}$ ,  $110\text{--}137.5^\circ\text{E}$  (Figure 1b). The HWI can be denoted as



**Figure 1.** January 2013 composite of temperature and wind speed at multiple atmospheric levels. The Haze Weather Index (HWI) is calculated as the difference between the temperature anomaly at 850 hPa and 200 hPa over the boxed regions (a, c) plus the difference between the northern zonal wind anomaly and the southern zonal wind anomaly at 500 hPa over the boxed regions (b) plus the meridional wind speed anomaly at 850 hPa over the boxed region (d). The Winter Monsoon Index (WMI) is calculated as the meridional wind anomaly at 850 hPa over the boxed region (f). The Pollution Potential Index (PPI) is calculated as the difference between the temperature anomaly at 850 hPa and 1,000 hPa (e, g) plus net wind speed anomaly at the surface (h) over the boxed regions. Black dot denotes Beijing. Data for all panels is from NCEP/NCAR R1 reanalysis. All panels show the same domain; latitude and longitude labels are shown in panel a but removed from the other panels for readability purposes.

$$\text{HWI} = \Delta T + V + \Delta U$$

where

$$\Delta T = T_{850} - T_{200}$$

$$V = V_{850}$$

$$\Delta U = U_{500,N} - U_{500,S}$$

where  $T$  refers to temperature (K),  $V$  refers to meridional wind (m/s),  $U$  refers to zonal wind (m/s), and subscripts denote pressure levels in millibars. A positive temperature gradient  $\Delta T$  represents a midtroposphere temperature inversion that prevents the vertical dispersion of pollutants, a positive meridional wind anomaly  $V$  describes increased southerly flow that transports pollutants and moisture from regions south of Beijing, and a positive zonal wind gradient  $\Delta U$  represents weakened northwesterly midtropospheric flow (Cai et al., 2017). Each quantity in the initial equation ( $\Delta T$ ,  $V$ , and  $\Delta U$ ) is calculated as a normalized anomaly before being combined into the final HWI, and the HWI is then itself normalized. Observational data is normalized by subtracting the observed mean and dividing by the observed standard deviation, and the preindustrial and historical simulations are normalized by subtracting the preindustrial mean and dividing by the preindustrial standard deviation. Cai et al. (2017) use 250 hPa temperature instead of 200 hPa temperature, but we use 200 hPa temperature since daily scale data is only available from the CESM Large Ensemble at 200 hPa and not 250 hPa. Because the temperature anomalies that Cai et al. (2017) identify at 250 hPa also exist at 200 hPa (see Figure 1 of Cai et al., 2017), this choice is unlikely to substantially alter the meaning of the HWI.

2. Western Monsoon Index (WMI). We define an index of the East Asian winter monsoon following the approach of Pei et al. (2018). The weakness of the monsoon is defined as the area-averaged meridional wind anomaly at 850 hPa over 105–125 °E and 30–50 °N (Figure 1f). The WMI can be denoted as:

$$\text{WMI} = V_{850}$$

where  $V_{850}$  refers to meridional wind (m/s) at 850 mb. Stronger southerly (positive) meridional winds are associated with increased southerly flow, indicating a weaker monsoon and increased transport of pollutants and moisture from the regions south of Beijing (Pei et al., 2018). We normalize the observed WMI by subtracting the observed mean and dividing by the observed standard deviation, and we normalize both the preindustrial and historical simulations by subtracting the preindustrial mean and dividing by the preindustrial standard deviation.

3. Pollution Potential Index (PPI). We define “pollution potential” conditions, modifying the approach of Zou et al. (2017), using a weighted index of near-surface inversions (air temperature gradient index (ATGI)) and wind speeds (wind speed index (WSI)). The ATGI is calculated by taking the area-averaged (112–122 °E and 30–41 °N) temperature difference between 850 and 1,000 hPa at each grid point (Figures 1e and 1g) and the WSI is constructed by calculating the area-averaged (112–122 °E and 30–41 °N) wind speed anomalies at 1,000 hPa (Figure 1h). The PPI is constructed by multiplying each of the ATGI and WSI by its Pearson correlation coefficient with historical air quality measurements, summing the two weighted indices, and dividing by the sum of the absolute value of the correlation coefficients. We use weights of 0.7 for the ATGI and  $-0.73$  for the WSI, from Zou et al. (2017). The PPI can be denoted as:

$$\text{PPI} = \frac{(r_1 \times \text{ATGI} + r_2 \times \text{WSI})}{(|r_1| + |r_2|)}$$

where :

$$\text{ATGI} = T_{850} - T_{1000}$$

$$\text{WSI} = \text{WS}_{1000}$$

$$r_1 = 0.7, r_2 = -0.73$$

where  $T$  refers to temperature (K),  $WS$  refers to the wind speed (m/s) calculated from the square root of the sum of squared meridional and zonal winds,  $r_1 = 0.7$ ,  $r_2 = -0.73$ , and subscripts denote pressure levels in millibars. A positive temperature gradient (ATGI) represents a near-surface temperature anomaly which prevents the vertical lifting of pollutants and a negative surface wind speed (WSI) represents weak surface winds that prevent the horizontal ventilation of those pollutants (Zou et al., 2017). The ATGI and WSI are normalized independently before construction of the PPI and then the PPI itself is normalized. In all cases, we normalize observed quantities subtracting the observed mean and dividing by the standard deviation, and we normalize the preindustrial and historical simulations by the preindustrial mean and standard deviation. We calculate the PPI using the weights derived by Zou et al. (2017), but we also recalculated the PPI without weighting either index ( $r_1 = 1$  and  $r_2 = -1$ ) and found similar results. An Anderson-Darling test indicates that the CESM-LE historical simulations of the weighted and unweighted PPI are statistically indistinguishable, indicating that our results are not sensitive to the choice of weights. For further details on the Anderson-Darling test, see section 2.3.

Figure 1 shows the January 2013 composite in each of the relevant variables and indices. January 2013 was defined by a strong warm anomaly in the lower troposphere and a cold anomaly aloft (Figures 1a and 1c); the warm anomaly in the lower troposphere created a stable atmosphere which prevented the vertical dispersion of pollutants. At the surface and 850-hPa level, southerly wind anomalies (Figures 1d, 1f, and 1h) transported pollutants and moisture conducive to pollutant formation into Beijing from the regions south of the city. At the 500-hPa level, January 2013 was characterized by weakened northwesterly flow and a stronger latitudinal difference in zonal wind between the north and south of Beijing, favorable for pollutant buildup (Figure 1b). Finally, a temperature inversion was present from 1,000 hPa to 850 hPa over Beijing (Figures 1e and 1g), which further prevented the vertical dispersion of pollutants.

These indices are not necessarily independent. For example, the HWI and WMI both include an area-averaged measure of meridional winds at 850 hPa, and the HWI and PPI both include a measure of temperature inversions centered at 850 hPa. However, each has benefits for our event attribution methodology. The HWI is synoptically oriented, incorporating measures of circulation gradients across much of northeast China, and includes the entire atmospheric column from 850 hPa to 200 hPa. Alternatively, the PPI measures only near-surface temperature inversions and surface wind speeds over a smaller domain than the HWI, allowing a closer look at the local conditions shaping air quality in Beijing. Finally, the WMI measures only the strength of the East Asian winter monsoon, which is the dominant synoptic feature in the region during winter. While the HWI includes a similar measure of midtroposphere meridional winds, it is useful to also test the East Asian winter monsoon in isolation using the WMI, given previous research demonstrating its significant individual effect on regional stagnation (e.g., Jia et al., 2015; Pei et al., 2018). Using indices that group multiple ingredients together (e.g., weak monsoon and near-surface inversion) is particularly important for air quality meteorology, as interactions between ingredients create compounding effects that exceed the sum of individual effects (Dawson et al., 2007; Dawson et al., 2009).

These three indices, however, are not an exhaustive list of metrics used to describe the meteorology that shapes air quality in China. Yang et al. (2016), for example, proposed the Parameter Linking Air-quality to Meteorology, and Wang et al. (2016), Wang et al. (2018), and Huang et al. (2018) have proposed alterations to the Air Stagnation Index used in Wang and Angell (1999) and Horton et al. (2014). In this study, we present three indices for which daily reanalysis and climate model data are available, and our detection and attribution data screening thresholds are met. However, we attempted to apply our attribution methodology to two other air quality meteorology indices known to influence pollutant accumulation: The Air Stagnation Index, or ASI (Wang & Angell, 1999), and the Siberian High Position Index, or SHPI (Jia et al., 2015). In both cases, we could not determine that the preindustrial simulation and detrended observational record were drawn from the same distribution, so we did not continue with our analysis (see section 2.3). As model physics and resolution improve and additional models are brought to bear on questions of attribution, future studies can test more additional, complex, and/or fine-scaled indicators of stagnant meteorology in Beijing.

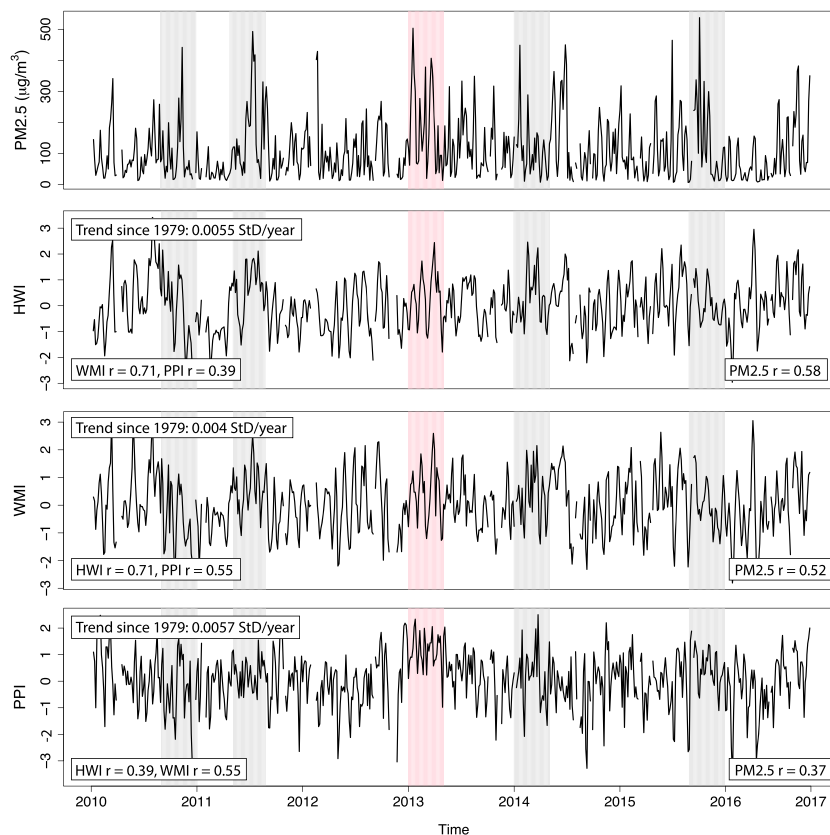
We calculate each index using daily data during boreal winter (December–February). To define “the event” to which our methods will be applied, we take the 5-day running mean of each index and treat the maximum value in January 2013 as the event. We focus on January 2013 both due to the unprecedented concentrations of pollutants during that month and the well-developed science that has described the extreme air quality

meteorology of that month (e.g., Gao et al., 2015; Kajino et al., 2017; Li et al., 2018; Sun et al., 2014; Wang et al., 2014; Ye et al., 2016; Zou et al., 2017). Such conditions, however, have also been found to support poor air quality in other months and years (Figure 2). In all following analyses, a  $p$  value of less than 0.05 is required for statistical significance, and all mentions of the “correlation coefficient” refer to the Pearson correlation coefficient.

### 2.3. Event Attribution

Following Diffenbaugh et al. (2017), we assess five attribution metrics from each index in order to determine the influence of anthropogenic forcing on the January 2013 “airpocalypse” in Beijing. We modify the approach of Diffenbaugh et al. (2017) by adding a fifth metric, drawn from a set of other studies that used CESM-LE to compare anthropogenic forcing to internal variability (e.g., Deser et al., 2012; Mankin et al., 2017), and we include it in this study to more thoroughly examine the role of anthropogenic influence on extreme air quality. The first two metrics focus solely on trends and events in the observational record (using NCEP/NCAR R1 reanalysis data), while metrics 3–5 use climate model data (CESM-LE) to explicitly distinguish anthropogenic forcing from natural variability.

1. *The contribution of the observed trend to the magnitude of the event* (Table 1 herein and Figure 1a from Diffenbaugh et al. (2017)). We first remove the linear trend from the observed time series, then calculate the observed event magnitude minus the detrended event magnitude, divided by the observed event magnitude minus the detrended mean. The resulting value represents the percent of the magnitude of the event attributable to the observed trend.
2. *The contribution of the observed trend to the probability of the event* (Figure 3a herein and Figure 1c from Diffenbaugh et al. (2017)). To generate a sample of return intervals for the given event in both the observed and detrended timeseries, we fit a Gumbel distribution, a variation of the Generalized Extreme Value distribution, to a bootstrapped sample of each timeseries 200 times. From each sample, we determine the return interval of an event in that distribution by taking the inverse of the frequency of the given event. We then take the ratio of the return intervals in the detrended timeseries to the return intervals in the observed timeseries; if the median ratio is greater than (less than) one, the return interval is longer (shorter) in the detrended timeseries and the probability of the event is higher (lower) in the observed timeseries (Singh, Tsiang, et al., 2014; Swain et al., 2014). We use a moving block bootstrap to control for autocorrelation and sample uncertainties in the recurrence interval of the relatively short observational period (Singh, Horton, et al., 2014). Note that return period ratio is a proxy for probability change; a return period ratio of 2 between the detrended timeseries and original timeseries is equivalent to determining that the observed trend increased the probability of the event by a factor of 2.
3. *The probability of the observed trend in the historical and preindustrial climates* (Figure 4 herein and Figure 1e from Diffenbaugh et al. (2017)). Once each meteorological index has been calculated from its components, we bias-correct the preindustrial control realization by shifting its mean to equal the mean of the detrended observations and to account for systematic model bias. Then, we apply an Anderson-Darling test to determine if the interannual variability of the preindustrial simulation and detrended timeseries are statistically drawn from the same population. If the result of the Anderson-Darling test is greater than 0.05, we determine that the simulations and observations are drawn from the same distribution, and we continue the analysis. If the result is less than 0.05, we determine that the simulations do not accurately represent the observations, and we do not continue. Because comparisons between the preindustrial and historical simulations do not require direct comparison to the observational record, we only bias-correct the preindustrial simulation when performing the Anderson-Darling test; we do not bias-correct the preindustrial and historical simulations when calculating the attribution metrics themselves. We find similar results if we bias-correct the individual components of the HWI and PPI as opposed to bias-correcting the final indices, though only if the final indices are normalized according to the process described above. Next, we take the linear trend from each CESM-LE 38-year historical realization (1979–2016) to generate a sample of historical trends, and randomly sample 30 different 38-year periods from the entire 1,800-year-long preindustrial control to generate a corresponding sample of preindustrial trends. If significantly more of the historical trends than the preindustrial trends are in the direction of the observed trend, we determine that climate change has contributed to the likelihood of the observed trend (Singh et al., 2016). We evaluate statistical significance using a two-tailed binomial test, with the null hypothesis that positive and negative trends should occur only as many times in the



**Figure 2.** Timeseries plots of daily winter (December–February) air quality and meteorological indices from 2010 to 2016. Air quality data is daily average  $PM_{2.5}$  concentrations from the U.S. embassy in Beijing, China. Indices are calculated from National Center for Environmental Predictions R1 reanalysis. Correlation coefficients between each index and the other two indices, along with the  $PM_{2.5}$  timeseries, are depicted in the insets. Pink box depicts January 2013; gray boxes depict other peak  $PM_{2.5}$  events (December 2010, February 2011, January 2014, and December 2015). HWI = Haze Weather Index; WMI = Winter Monsoon Index; PPI = Pollution Potential Index.

historical simulation as they do in our random sample from the preindustrial simulation (Singh et al., 2016). A p value below 0.05 in Figure 4 therefore represents a significantly higher number of positive trends occurring in the historical simulation than would be expected from the preindustrial simulation.

4. *The probability of the event in the historical and preindustrial climates* (Figure 3b herein and Figure 1f from Duffenbaugh et al. (2017)). We define the sample of return intervals of the event in the preindustrial climate to be equal to the return intervals in the detrended observed timeseries (Swain et al., 2014, 2018). We calculate the sample of event magnitudes in the preindustrial simulation that are associated with each of those return intervals. Then, for each event magnitude derived from the preindustrial simulation, we determine the return interval of that event magnitude in the historical simulation (Swain et al., 2014, 2018). Finally, we take the ratio of the return intervals between the preindustrial simulation and historical simulation (Swain et al., 2014, 2018). Similar to metric 2, if the median ratio is greater than (less than) one, the return interval is longer (shorter) in the preindustrial simulation and the probability of the event is higher (lower) in the historical simulation.
5. *The relative importance of anthropogenic forcing and internal variability to the observed trend* (Figure 5). Following Deser et al. (2014), we first define the anthropogenically forced component of the simulated historical climate as the ensemble mean of the CESM-LE historical realization. We then verify that the trend in this component is in the same direction as the trend in the observed timeseries. If so, we subtract the ensemble mean from each individual CESM-LE realization to isolate the component of that realization due to internal variability and take the ratio of the trend in the ensemble mean to the trends in the



**Table 1**  
*Attribution Statistics for Each Index*

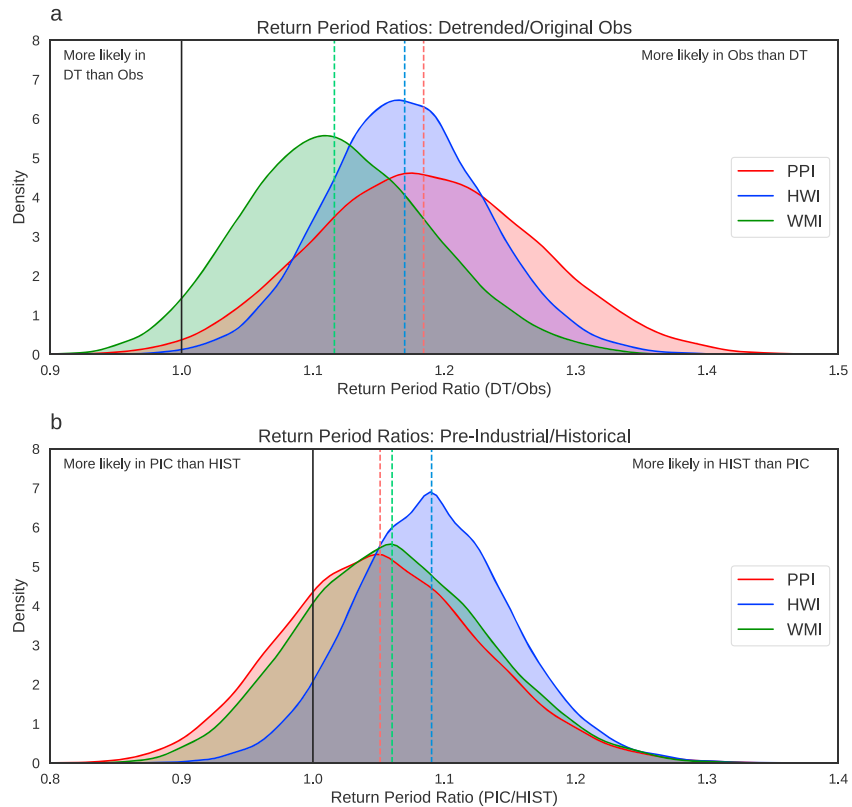
	HWI	WMI	PPI
January 2013 Magnitude (Standard Deviations)	1.53	1.58	1.81
Observed Trend (Standard Deviations/Year)	0.0055	0.004	0.0057
Observed Trend <i>p</i> Value	$2.1 \times 10^{-6}$	$1.2 \times 10^{-4}$	$1.6 \times 10^{-9}$
$\Delta$ Mag. From Trend (Metric 1)	11.5%	8.4%	10.2%
Median Return Period Ratio (DT/Obs) (Metric 2)	1.17	1.12	1.18
GCM-Obs A-D <i>p</i> Value	0.56	0.94	0.051
Hist. Simulation Trend Agreement (Metric 3)	76.7%	56.7%	76.7%
Median Return Period Ratio (PI/HIST; Metric 4)	1.09	1.06	1.05
Signal-to-Noise Ratio (Metric 5)	0.772	0.09	0.91
Median Forced/Internal Var. Ratio (Metric 5)	0.56	0.08	0.56

*Note.* The magnitude of the January 2013 event; the linear trend in the observations over the course of the winter (December–February) months from December 1979 to February 2016; the *p* value of the observed trend; the change in magnitude due to the observed trend; the median return period ratio between the detrended observations and original observations; and the *p* value of the Anderson-Darling test between the preindustrial simulation and detrended observations; the median return period ratio between the preindustrial and historical climate; the percent of historical realizations with trends in the same direction as the observed trend; signal-to-noise ratio; and the median ratio of forced component to internal variability in each historical realization (median from distributions in Figure 5).

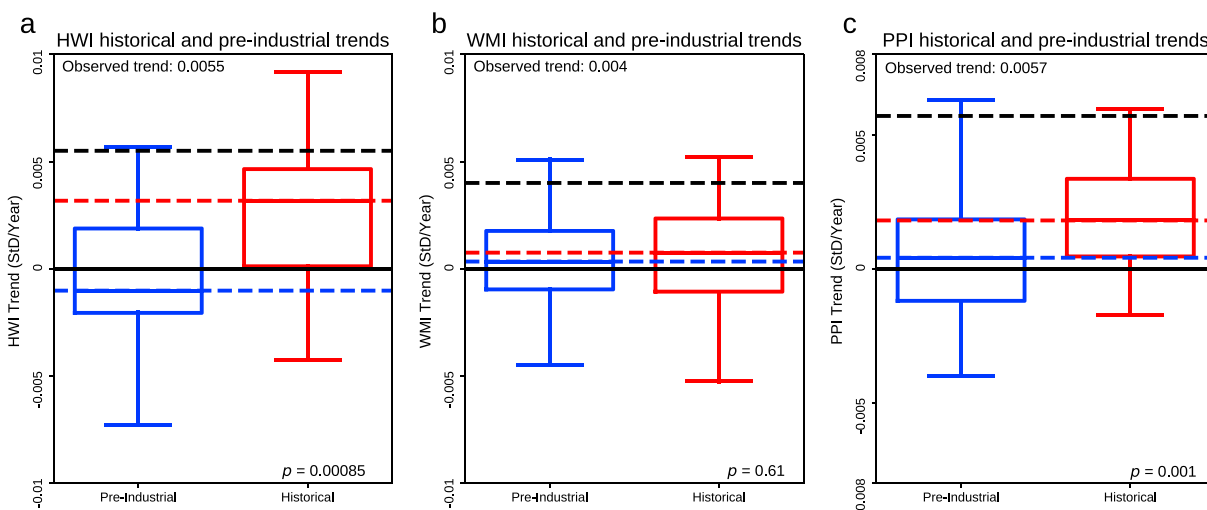
internal-variability-only components. Following Hawkins and Sutton (2009) and Mankin et al. (2017), we then quantify the signal-to-noise ratio of the historical climate by dividing the magnitude of the trend in the ensemble mean by the standard deviation in the magnitude of the trends across all realizations. We determine that anthropogenic forcing exerts a robust influence on the system if the signal-to-noise ratio is greater than or equal to 1 (Hawkins & Sutton, 2009). Signal-to-noise ratio provides a single metric for the entire set of CESM-LE historical simulations, while the ratio of the forced trend to internal variability provides a similar metric for each individual realization (Figure 5).

The first two metrics benefit from sole focus on the statistics of the observational record, thus avoiding model bias. However, they assume that a linear fit accurately describes the trend in the observations, and they do not account for changes in variance or higher-order statistical moments (Differbaugh et al., 2017). In addition, attributing the magnitude or probability of an event to an observed trend does not equate to attributing that event to climate change. The presence of a linear trend alone does not necessarily imply anthropogenic forcing, since trends may still be the result of natural variation (i.e., the difference between Figure 3a, which shows the effect of the observed trend on the event probability, and Figure 3b, which shows the effect of anthropogenic forcing on the event probability). Model simulations of the historical and preindustrial climates (i.e., the third, fourth, and fifth metrics) are therefore necessary to determine the origin of a given trend and distinguish anthropogenic forcing from underlying variability. The presence of inherent model bias requires validation that the simulations sufficiently represent the observed variability, making the Anderson-Darling test an essential step before continuing with the last three attribution metrics (Mitchell et al., 2017). The ASI and SHPI were discarded precisely because the Anderson-Darling test indicated that the model simulations did not accurately simulate observed variability (ASI  $p = 5 \times 10^{-6}$  and SHPI  $p = 0.0003$ ).

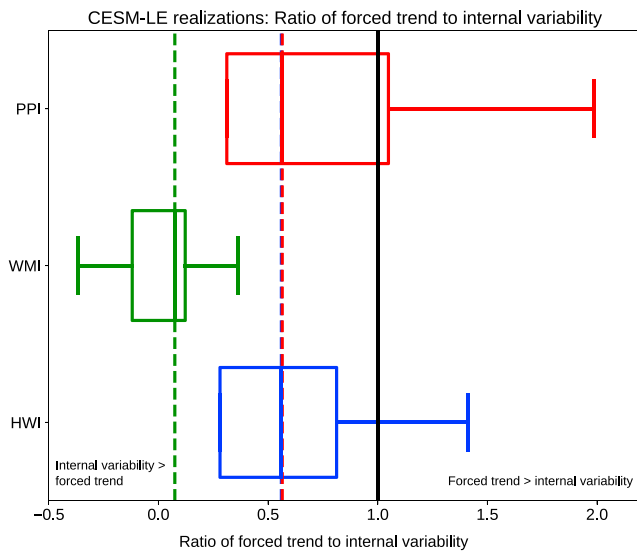
Analyzing the contribution of the observed trend to both the magnitude and probability of an event is important, as climate change may have different effects on the magnitude and probability of extreme events, and the results of attribution studies are sensitive to the exact question asked. Attribution studies of the 2010 Russian heat wave, for example, reached different conclusions about anthropogenic influence, with Dole et al. (2011) reporting that the event was primarily caused by natural variability and Rahmstorf and Coumou (2011) reporting that anthropogenic warming was the most significant contributor to the event. It is possible to reconcile these results by determining that climate change did not contribute to the magnitude of the event but did contribute to the probability of an event of a certain magnitude, demonstrating the value of partitioning magnitude and probability into separate attribution questions (Otto et al., 2012).



**Figure 3.** Return period changes for the three meteorological indices. (a) Return period ratios between the preindustrial and historical simulated climates for the HWI (blue), WMI (green), and PPI (red). (b) Return period ratios between the detrended observed climate and the original observed climate for the HWI, WMI, and PPI. Vertical dashed lines denote the median of each distribution. Note the slightly different x axis scales between the subplots. HWI = Haze Weather Index; WMI = Winter Monsoon Index; PPI = Pollution Potential Index.



**Figure 4.** Changes in trend probability between the preindustrial (blue) and historical (red) climate for the HWI (a), WMI (b), and PPI (c). Each boxplot represents the sample of trends drawn from subsets of the preindustrial control and from each of the historical realizations, respectively. Blue line denotes the median pre-industrial trend, red line denotes the median historical trend, and black line denotes the observed trend.  $P$  value represents the result of a two-tailed binomial test for the significance of the number of historical trends above 0, with the null hypothesis being that positive and negative trends are as likely in the historical simulation as they are in the random sample from the preindustrial simulation. We exclude outliers in all boxplots. HWI = Haze Weather Index; WMI = Winter Monsoon Index; PPI = Pollution Potential Index.



**Figure 5.** Distribution of ratios of forced trend (ensemble mean trend) relative to internal variability (each realization minus ensemble mean trend) for each set of CESM-LE historical realizations. Dotted vertical lines correspond to the medians of each boxplot; PPI (red) and HWI (blue) may be difficult to distinguish as they lie on top of one another. Ratios of 1 are considered robust, as 1 represents the level at which the forced component is equal in magnitude to the internal-variability-driven component. Negative ratios are possible if a realization contains a negative trend. We exclude outliers in all boxplots. CESM-LE = Community Earth System Model Large Ensemble.

### 3. Results

The HWI, WMI, and PPI all have positive and statistically significant trends in the observational record (Table 1). In addition, based on Anderson-Darling goodness-of-fit testing, CESM-LE simulates all three indices effectively, suggesting that the model is an adequate tool for representing the underlying climate variability of each index (Table 1). However, our results indicate mixed signals amongst the meteorological indices with regard to the degree of anthropogenic influence (Table 1). We find that anthropogenic forcing contributed to a small increase in the probabilities of the January 2013 event in all three indices, but signal-to-noise ratio and the degree of long-term anthropogenic influence vary widely among the indices. In all of the following, we report median values of return period ratio and forced/component internal variability ratios in the text but show the full distributions of return period ratios in Figure 3 and forced component/internal variability ratios in Figure 5. The exact magnitude of the January 2013 event in each index is indicated in Table 1; the range of these values, from 1.53 standard deviations for the HWI to 1.81 standard deviations for the PPI, demonstrates the sensitivity of air quality meteorology analyses to the exact method by which pollutant-conducive meteorology is quantified.

#### 3.1. Index Correlations

While the authors that originated each of these indices confirmed that they correlate with poor air quality events in the Beijing region, we also verify that there is some relationship between the winter values of each index and  $PM_{2.5}$  concentration data gathered by the United States

embassy in Beijing (Figure 2). The HWI has a correlation coefficient of 0.58 with  $PM_{2.5}$  concentrations in winter from 2010 to 2016, while the WMI has a correlation coefficient of 0.52 and the PPI has a correlation coefficient of 0.37. Due to differences in temporal application or data sources, these values differ from the correlation coefficients found in the papers in which these indices were originated. We exclude the  $PM_{2.5}$  timeseries prior to 2010 due to missing data, for example, which results in a slightly different value than that found by Cai et al. (2017) when testing the HWI against observed  $PM_{2.5}$  concentrations. Pei et al. (2018), in addition, correlated the WMI with daily visibility at individual weather stations as opposed to  $PM_{2.5}$  concentrations, and Zou et al. (2017) tested January monthly mean values of the PPI against monthly mean  $PM_{10}$  values over several decades as opposed to daily  $PM_{2.5}$  values from 2010 to 2016.

Given that these correlation coefficients range from 0.37 to 0.58, none of these indices are able to fully explain the variability associated with  $PM_{2.5}$  concentrations in Beijing. However, each index's correlation is statistically significant, and there are associations between peak  $PM_{2.5}$  events and peaks in some or all indices. January 2013 (Figure 2, pink highlighting) is associated with peaks in  $PM_{2.5}$  and all three indices, as are February 2011 and January 2014. Other peak  $PM_{2.5}$  events, like December 2015, are not associated with peaks in all indices. The meteorological variables included in these indices do influence air quality over Beijing but not exclusively; there are likely other meteorological phenomena or nonmeteorological variables (e.g., emissions changes) that contribute to changes in  $PM_{2.5}$  concentrations. The clear peaks associated with January 2013 in the  $PM_{2.5}$  data and all three meteorological indices support our choice to use January 2013 as an exemplar “event” for this study.

Note that these indices are also not independent from each other; the HWI and WMI are correlated with a coefficient of 0.71, likely in large part because the WMI measures the meridional wind anomaly at 850 hPa in a very similar manner to the HWI (Figures 1d and 1f). The PPI is correlated with the HWI at a coefficient of 0.39 and the HWI at a coefficient of 0.55, consistent with the way the PPI measures slightly different quantities than the HWI or WMI (Figure 1). Because these indices are not statistically independent, since they measure many of the same or similar quantities, our results do not represent the full spread of possible responses of extreme air quality events to anthropogenic forcing.

### 3.2. Haze Weather Index

There is a positive trend of 0.0055 standard deviations per year between 1979 and 2016 in the HWI, which measures midtropospheric temperature inversions, anomalous southerlies conducive to regional pollutant transport, and weak northwesterlies that allow pollutant accumulation in Beijing. This positive trend is statistically significant ( $p = 2.1 \times 10^{-6}$ ) and contributed 11.5% of the magnitude of the January 2013 event. The observed trend also increased the probability of the occurrence of the January 2013 event by a factor of almost 1.2 (Figure 3a). In contrast, using the historical and preindustrial simulations from CESM-LE, we find that anthropogenic forcing only increased the probability of the occurrence of the January 2013 event in the HWI by a factor of 1.09 (Figure 3b). While there is a significant trend in the observational record that has contributed to the likelihood of the January 2013 HWI, we can only attribute some of that contribution to anthropogenic forcing. The divergence between the contributions of the observed trend and the historical forcing suggests that the occurrence of the extreme January 2013 value in the HWI was driven more by natural variability than anthropogenic forcing.

The 76.7% of the trends in the CESM-LE historical realizations are positive, the same direction as the observed trend (Figure 4a). With a  $p$  value of  $8.5 \times 10^{-4}$ , we determine that there are significantly more positive trends in the historical simulation than in the random sample of the preindustrial simulation (Figure 4a). Anthropogenic forcing has therefore had a discernible effect on the likelihood of a decadal-scale increase in the HWI. However, natural variability has had a larger effect within the CESM-LE framework. The signal-to-noise ratio in the CESM-LE historical realizations is 0.772, below the critical value of 1, indicating that natural variability outweighed the effect of anthropogenic forcing on historical change in the HWI. Additionally, we calculate the median ratio of the forced component to internal variability in each CESM-LE realization to be 0.56 (Figure 5), which also demonstrates the greater role of natural variability than anthropogenic forcing in controlling HWI variability within each CESM-LE realization.

### 3.3. Winter Monsoon Index

The WMI, which represents southerly wind anomalies conducive to pollutant maintenance in Beijing, has a positive trend of 0.004 standard deviations per year in reanalysis data from 1979 to 2016. This trend is statistically significant ( $p = 1.2 \times 10^{-4}$ ) and contributed 8.4% of the magnitude of the January 2013 event in the WMI. This trend also increased the probability of the event's occurrence by a factor of 1.12 (Figure 3a). From the historical and preindustrial simulations, we find that anthropogenic forcing increased the likelihood of the January 2013 event by only a factor of 1.06 (Figure 3b). In addition, almost 20% of the preindustrial/historical return period ratios in Figure 3b are less than 1; a simple bootstrap was enough to alter the sign of the effect of anthropogenic forcing, demonstrating the relatively small effect of that forcing on extreme southerly wind anomalies. In a similar fashion to the HWI, the observed trend had a meaningful effect on the magnitude and probability of the January 2013 event in the WMI, but we can attribute very little, if any, of that effect to anthropogenic forcing.

While 56.7% of the historical realizations contain a positive trend, which matches the observed trend (Figure 4b), this set of trends is not statistically distinguishable from a random sample of trends from the preindustrial simulation ( $p = 0.61$ ). Therefore, we cannot conclude that anthropogenic forcing had any effect on the likelihood of a long-term increase in the WMI (and thus a long-term weakening of the East Asian winter monsoon) between 1979 and 2016. In addition, we calculate a signal-to-noise ratio of 0.09 and a forced component/internal variability ratio of 0.08 (Figure 5), indicating that in both the overall CESM-LE ensemble and in every individual realization, the influence of natural variability on WMI change was more than 10 times as large as the influence of anthropogenic forcing.

### 3.4. Pollution Potential Index

The PPI, measuring near-surface temperature inversions and weak surface winds, has a statistically significant positive trend of 0.0057 standard deviations per year in the observational record ( $p = 1.6 \times 10^{-9}$ ). This positive trend contributed 10.2% of the magnitude of the January 2013 event and increased the probability of an event of that magnitude by around 1.2 times (Figure 3a). From the historical and preindustrial simulations, we find that anthropogenic forcing made the January 2013 event in the PPI about 1.05 times as likely. Consistent with the other two indices, we find that anthropogenic forcing slightly increased the risk of the January 2013 PPI, though not as much as the other two indices (Figure 3b). As in the case of the WMI,

almost 25% of the return period ratios for the PPI in Figure 3b are less than one, meaning that statistical noise was enough to mask the influence of anthropogenic forcing in driving extreme temperature anomalies and weak surface winds.

We find that 76.7% of historical realizations of the PPI contain a positive trend (Figure 4c), consistent with the observed trend and statistically distinguishable from the random preindustrial sample ( $p = 0.001$ ). We therefore conclude that anthropogenic forcing had a significant effect on decadal-scale increases in the PPI, independent of its influence on individual events like the January 2013 event. We also calculate a signal-to-noise ratio of 0.91, indicating that anthropogenic forcing had almost as large of an effect as natural variability in shaping PPI change in the CESM-LE historical realizations. The median ratio of the forced component to internal variability was 0.56, but 26.7% of CESM-LE historical realizations had a ratio above 1, indicating that some realizations contained a discernibly greater effect from anthropogenic forcing relative to internal variability on historical change in the PPI (Figure 5).

Note that the Anderson-Darling  $p$  value for the PPI is 0.051, which is above our threshold of 0.05 but still significantly less than that of the other two indices, meaning that CESM-LE less effectively represents the underlying variability of the PPI than that of the HWI or WMI. This is consistent with the finding of Zou et al. (2017) that CESM may underestimate historical changes in the PPI. While we consider CESM-LE's representation of important meteorological features sufficient for this analysis, care should be taken in future research to explore the factors influencing model representation of observed meteorology.

#### 4. Discussion

We evaluated the effect of historical forcing on the probability and magnitude of Beijing's January 2013 air-pocalypse in multiple indices of air quality meteorology. Anthropogenic forcing had a small effect on the probability of the January 2013 event in the HWI (return period ratio of 1.09), WMI (return period ratio of 1.06), and PPI (return period ratio of 1.05), an effect that was largely outweighed by natural variation. In the WMI, anthropogenic forcing had a small or insignificant effect on the probabilities of both the January 2013 event and the long-term positive trend. However, in both the HWI and PPI, there is some divergence between the degree of anthropogenic influence on the observed trend and on the January 2013 event; anthropogenic forcing had a significant effect on the probability of the observed trend in both but a small effect on the probability of the January 2013 event. This divergence can be understood as the distinction between the effect of anthropogenic forcing on the full distribution of a variable versus its effect on extreme events at the tails of that distribution. A strong relationship between anthropogenic forcing and interdecadal change in the HWI and PPI is separate from a strong relationship between anthropogenic forcing and extremes in those indices, and our results demonstrate the importance of attribution methodologies which explicitly distinguish between those relationships.

The divergence in results between the three indices may be partially due to the different mechanisms by which climate change affects each of their constituent variables. These physical mechanisms are the subject of ongoing research. Cai et al. (2017), when developing the HWI, attributed potential future increases in HWI values to a combination of lower-troposphere warming and near-surface temperature inversions creating a more stable atmosphere, faster warming on land than the ocean creating a weakened East Asian winter monsoon and reduced northwesterly winds, and a long-term shallowing of the East Asian trough associated with weaker circulation in northeast China. In developing the PPI, Zou et al. (2017) attributed its long-term increase to decreasing Arctic sea ice and increasing Eurasian snow cover, which produce anticyclonic anomalies in the Arctic and northeast Asia and enhance the stagnant conditions over Beijing. They also found that CESM may underestimate the recent decreases in Arctic sea ice and increases in Eurasian snow, meaning our results may underestimate the influence of climate change on the PPI distribution and its extremes (Zou et al., 2017). While we cannot conclude that anthropogenic warming has affected historical change in the WMI, Pei et al. (2018) hypothesized that positive SST anomalies in the northwest Pacific create anomalously high-pressure systems off the eastern coast of China, which produce anomalous southerly winds and weaken the East Asian winter monsoon.

Note, as before, that these three indices are not statistically independent, so our results do not span the full range of uncertainty associated with anthropogenic influence on extreme air quality events in Beijing. In addition, these indices are not perfectly correlated with the physical concentration of pollutants (Figure 2),

meaning that our attribution methodology can only assess anthropogenic influence on the meteorological indices themselves not air quality as such. Further research is necessary to explicitly link anthropogenic influence to the physical concentration of pollutants. In addition, our analysis cannot disentangle the direct effects of meteorology from pollutant-meteorology feedbacks, such the role of black carbon in depressing the planetary boundary layer (Ding et al., 2016), the weakening effect of black carbon on the East Asian winter monsoon (Lou et al., 2019), or role that dust plays in suppressing wind speeds (Yang et al., 2017). However, given the complex relationships between anthropogenic and biogenic emissions of pollutants, atmospheric chemistry, and meteorology, it is reasonable to focus our analysis on direct meteorological influences to reduce the uncertainty associated with connecting all of these factors to human influence.

Our results diverge from those of Li et al. (2018), who attributed 40% of the probability of the January 2013 event in the HWI to anthropogenic forcing. A divergence in results is expected, as they define the January 2013 event according to the synoptic conditions that produce high-HWI values instead of the HWI timeseries itself. Li et al. (2018) explicitly distinguish between these high-HWI days and the background synoptic conditions that promote these extreme days, so our results and theirs are not directly comparable. In addition, our use of the CESM-LE single-model multirealization ensemble is subject to inherent model biases, as is Li et al.'s use of the MIROC5 model. For example, CESM is known to have a lower historical climate sensitivity than much of the CMIP5 ensemble, a factor that highlights the need for large initial-condition perturbation ensemble development from multiple modelling groups.

In all three indices, the contribution of the observed trend to the likelihood of the event was larger than the model-simulated contribution of anthropogenic forcing to that likelihood. This divergence demonstrates the importance of partitioning attribution methods into metrics that draw solely from the observed record and methods that use climate model simulations that represent the historical and preindustrial climates. While anthropogenic warming may be a component of the observed trend, it may only be a small component relative to natural variability. The short timescale of the observational record means it is not possible to rule out natural interannual or interdecadal variability as a significant contributor to observed trends.

By definition, detection and attribution investigations of past events restrict their analysis to the historical record. However, several investigators have found that certain meteorological conditions that support extremely poor air quality are likely to increase in frequency and/or magnitude in a high-emission future. Differences in extreme air quality meteorology between the historical record and model projections raise the question of “time of emergence,” or the period when an anthropogenic signal begins to distinguish itself from underlying climate variability (e.g., King, Donat, et al., 2015; Tan et al., 2018). Similar to detection and attribution investigations, time of emergence studies are beginning to be applied to more complex environmental variables. The expanse of both extreme event attribution and time of emergence to novel environmental conditions therefore promises to advance a more complete understanding of the changing influence of climate change in the past, present, and future.

#### Acknowledgments

Our work was supported by U.S. National Science Foundation grant CBET-1848683 to D. E. H., the Ubben Program for Climate and Carbon Science postdoctoral fellowship to J. L. S., and grants from the Weinberg College of Arts and Science and Office of the Provost at Northwestern University and the American Geophysical Union to C. W. C. Computational resources were provided by Northwestern University Information Technology and the Quest high-performance computing cluster. We thank the CESM Large Ensemble Community Project and supercomputing resources provided by NSF/CISL/Yellowstone for developing and hosting CESM-LE. We thank N. Diffenbaugh for sharing the attribution code and D. Singh, H. Chen, and Y. Suo for discussion of our results. The authors declare that they have no conflicts of interest. All raw data is publicly available; NCEP/NCAR R1 reanalysis at <https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html>, CESM-LE at <http://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html>, and Beijing embassy air quality data at <http://www.stateair.net/web/historical/1/1.html> website. Processed data for each index and final data for each figure are available at [www.github.com/ccallahan45](http://www.github.com/ccallahan45) website.

#### References

- Allen, M. (2003). Liability for climate change. *Nature*, 421(6926), 891–892. <https://doi.org/10.1038/421891a>
- Allen, M., Pall, P., Stone, D., Stott, P., Frame, D., Kim, S.-K., et al. (2007). Scientific challenges in the attribution of harm to human influence on climate. *University of Pennsylvania Law Review*, 155(6), 1353–1400.
- Angelil, O., Stone, D., Wehner, M., Paciorek, C., Krishnan, H., & Collins, W. (2017). An independent assessment of anthropogenic attribution statements for recent extreme temperature and rainfall events. *Journal of Climate*, 30, 5–16. <https://doi.org/10.1175/JCLI-D-16-0077.1>
- Beech, H. (2013). Beijing chokes on record pollution, and even the government admits there's a problem. Time.
- 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 visible data from 1960 to 2012. *Journal of Geophysical Research: Atmospheres*, 120, 5895–5909. <https://doi.org/10.1002/2015JD023225>
- Christidis, N., Stott, P., & Zwiers, F. (2015). Fast-track attribution assessments based on pre-computed estimates of changes in the odds of warm extremes. *Climate Dynamics*, 39(6), 1259–1274. <https://doi.org/10.1007/s00382-011-1184-0>
- Christidis, N., Stott, P., Zwiers, F., Shiogama, H., & Nozawa, T. (2011). The contribution of anthropogenic forcings to regional changes in temperature during the last decade. *Climate Dynamics*, 45(5–6), 1547–1564. <https://doi.org/10.1007/s00382-014-2408-x>
- Dawson, J., Adams, P., & Pandis, S. (2007). Sensitivity of PM<sub>2.5</sub> to climate in the Eastern U.S.: A modeling case study. *Atmospheric Chemistry and Physics*, 7(16), 4295–4309. <https://doi.org/10.5194/acp-7-4295-2007>
- Dawson, J., Bloomer, B., Winner, D., & Weaver, C. (2014). Understanding the meteorological drivers of U.S. particulate matter concentrations in a changing climate. *Bulletin of the American Meteorological Society*, 521–532. <https://doi.org/10.1175/BAMS-D-12-00181.1>

- Dawson, J., Racherla, P., Lynn, B., Adams, P., & Pandis, S. (2009). Impacts of climate change on regional and urban air quality in the eastern United States: Role of meteorology. *Journal of Geophysical Research*, *114*, D05308. <https://doi.org/10.1029/2008JD009849>
- Deser, C., Phillips, A., Alexander, M., & Smoliak, B. (2014). Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *Journal of Climate*, *27*(6), 2271–2296. <https://doi.org/10.1175/JCLI-D-13-00451.1>
- Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, *38*(3–4), 527–546. <https://doi.org/10.1007/s00382-010-0977-x>
- Diffenbaugh, N., Singh, D., Mankin, J., Horton, D., Swain, D., Touma, D., et al. (2017). Quantifying the influence of global warming on unprecedented extreme climate events. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(19), 4881–4886. <https://doi.org/10.1073/pnas.1618082114>
- Diffenbaugh, N., Swain, D., & Touma, D. (2015). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(13), 3931–3936. <https://doi.org/10.1073/pnas.1422385112>
- Ding, A. J., Huang, X., Nie, W., Sun, J. N., Kerminen, V. M., Petäjä, T., et al. (2016). Enhanced haze pollution by black carbon in megacities over China. *Geophysical Research Letters*, *43*, 2873–2879. <https://doi.org/10.1002/2016GL067745>
- Dole, R., Hoerling, M., Perlwitz, J., Eischeid, J., Pegion, P., Zhang, T., et al. (2011). Was there a basis for anticipating the 2010 Russian heat wave? *Geophysical Research Letters*, *38*, L06702. <https://doi.org/10.1029/2010GL046582>
- Fischer, E., & Knutti, R. (2015). Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nature Climate Change*, *5*(6), 560–564. <https://doi.org/10.1038/nclimate2617>
- Gao, M., Guttikunda, S., Carmichael, G., Wang, Y., Liu, Z., Stanier, C., et al. (2015). Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area. *Science of the Total Environment*, *511*, 553–561. <https://doi.org/10.1016/j.scitotenv.2015.01.005>
- Guo, S., Hu, M., Zamora, M., Peng, J., Shang, D., Zheng, J., et al. (2014). Elucidating severe urban haze formation in China. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(49), 17,373–17,378. <https://doi.org/10.1073/pnas.1419604111>
- Hausstein, K., Otto, F., Uhe, P., Schaller, N., Allen, M., Hermanson, L., et al. (2016). Real-time extreme weather attribution with forecast seasonal SSTs. *Environmental Research Letters*, *11*(6), 064006. <https://doi.org/10.1088/1748-9326/11/6/064006>
- Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108. <https://doi.org/10.1175/2009BAMS2607.1>
- Hegerl, G. (2015). Use of models and observations in event attribution. *Environmental Research Letters*, *10*(7), 071001. <https://doi.org/10.1088/1748-9326/10/7/071001>
- Hegerl, G., & Zwiers, F. (2011). Use of models in detection and attribution of climate change. *WIREs Climate Change*, *2*(4), 570–591. <https://doi.org/10.1002/wcc.121>
- Herring, S., Christidis, N., Hoell, A., Kossin, J., Schreck, C. III, & Stott, P. (Eds) (2017). 2017: Explaining extreme events of 2016 from a climate perspective. *Bulletin of the American Meteorological Society*, *99*(1), S1–S157. <https://doi.org/10.1175/BAMS-ExplainingExtremeEvents2016.1>
- Horton, D., Harshvardhan, & Diffenbaugh, N. (2012). Response of air stagnation frequency to anthropogenically enhanced radiative forcing. *Environmental Research Letters*, *7*(4). <https://doi.org/10.1088/1748-9326/7/4/044034>
- Horton, D., Skinner, C., Singh, D., & Diffenbaugh, N. (2014). Occurrence and persistence of future atmospheric stagnation events. *Nature Climate Change*, *4*(8), 698–703. <https://doi.org/10.1038/nclimate2272>
- Hou, P., & Wu, S. (2016). Long-term changes in extreme air pollution meteorology and the implications for air quality. *Scientific Reports*, *6*, 23792. <https://doi.org/10.1038/srep23792>
- Huang, Q., Cai, X., Wang, J., Song, Y., & Zhu, T. (2018). Climatological study of the Boundary-layer air Stagnation Index for China and its relationship with air pollution. *Atmospheric Chemistry and Physics*, *18*, 7573–7593. <https://doi.org/10.5194/acp-18-7573-2018>
- Jacob, D., & Winner, D. (2009). Effect of climate change on air quality. *Atmospheric Environment*, *43*, 51–63. <https://doi.org/10.1016/j.atmosenv.2008.09.051>
- 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. *Atmospheric Chemistry and Physics*, *15*, 11,919–11,929. <https://doi.org/10.5194/acp-15-11919-2015>
- Kaiman, J. (2013). Chinese struggle through “airpocalypse” smog. *The Guardian*.
- Kajino, M., Ueda, H., Han, Z., Kudo, R., Inomata, Y., & Kaku, H. (2017). Synergy between air pollution and urban meteorological changes through aerosol-radiation-diffusion feedback—A case study of Beijing in January 2013. *Atmospheric Environment*, *171*, 98–110. <https://doi.org/10.1016/j.atmosenv.2017.10.018>
- 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:TNYRYP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRYP>2.0.CO;2)
- Kay, J., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., et al. (2015). The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333–1349. <https://doi.org/10.1175/BAMS-D-1300255.1>
- King, A., Donat, M., Fisher, E., Hawkins, E., Alexander, L., Karoly, D., et al. (2015). The timing of anthropogenic emergence in simulated climate extremes. *Environmental Research Letters*, *10*(9), 094015. <https://doi.org/10.1088/1748-9326/10/9/094015>
- King, A., van Oldenborgh, G., Karoly, D., Lewis, S., & Cullen, H. (2015). Attribution of the record high central England temperature of 2014 to anthropogenic influences. *Environmental Research Letters*, *10*(5), 054002. <https://doi.org/10.1088/1748-9326/10/5/054002>
- Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., et al. (2001). The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, *77*, 437–471. [https://doi.org/10.1175/1520-0477\(2001\)082<0247:TNNYRM>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2)
- Leung, L., & Gustafson, W. (2005). Potential regional climate change and implications to U.S. air quality. *Geophysical Research Letters*, *32*, L16711. <https://doi.org/10.1029/2005GL022911>
- 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>
- Li, L., Lei, Y., Pan, D., Yu, C., & Si, C. (2016). Economic evaluation of the air pollution effect on public health in China’s 74 cities. *Springerplus*, *5*, 402–418. <https://doi.org/10.1186/s40064-016-2024-9>
- Lou, S., Yang, Y., Wang, H., Smith, S. J., Qian, Y., & Rasch, P. J. (2019). Black carbon amplifies haze over the North China Plain by weakening the East Asian winter monsoon. *Geophysical Research Letters*, *46*, 452–460. <https://doi.org/10.1029/2018GL080941>
- Mankin, J., Viviroli, D., Mekonnen, M., Hoekstra, A., Horton, R., Smerdon, J., & Diffenbaugh, N. (2017). Influence of internal variability on population exposure to hydroclimatic changes. *Environmental Research Letters*, *12*, 044007. <https://doi.org/10.1088/1748-9326/aa5efc>
- Mickley, L., Jacob, D., & Field, B. (2004). Effects of future climate change on regional air pollution episodes in the United States. *Geophysical Research Letters*, *31*, L24103. <https://doi.org/10.1029/2004GL021216>

- Mitchell, D., Davini, P., Harvey, B., Massey, N., Hausteine, K., Woollings, T., et al. (2017). Assessing mid-latitude dynamics in extreme event attribution systems. *Climate Dynamics*, *48*(11–12), 3889–3901. <https://doi.org/10.1007/s00382-016-3308-z>
- Murazaki, K., & Hess, P. (2006). How does climate change contribute to surface ozone change over the United States? *Journal of Geophysical Research*, *111*, D05301. <https://doi.org/10.1029/2005JD005873>
- Otto, F., Massey, N., van Oldenborgh, G., Jones, R., & Allen, M. (2012). Reconciling two approaches to attribution of the 2010 Russian heat wave. *Geophysical Research Letters*, *39*, L04702. <https://doi.org/10.1029/2011GL050422>
- Pall, P., Aina, T., Stone, D., Stott, P., Nozawa, T., Hilberts, G., et al. (2011). Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature*, *470*(7334), 382–385. <https://doi.org/10.1038/nature09762>
- Pei, L., Yan, Z., Sun, Z., Miao, S., & Yao, Y. (2018). Increasing persistent haze in Beijing: Potential impacts of weakening East Asian winter monsoons associated with northwestern Pacific sea surface temperature trends. *Atmospheric Chemistry and Physics*, *18*(5), 3173–3183. <https://doi.org/10.5194/acp-18-3173-2018>
- Rahmstorf, S., & Coumou, D. (2011). Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(44), 17,905–17,909. <https://doi.org/10.1073/pnas.1101766108>
- San Martini, F., Hasenkopf, C., & Roberts, D. (2015). Statistical analysis of PM<sub>2.5</sub> observations from diplomatic facilities in China. *Atmospheric Environment*, *110*, 174–185. <https://doi.org/10.1016/j.atmosenv.2015.03.060>
- Schnell, J., & Prather, M. (2017). Co-occurrence of extremes in surface ozone, particulate matter, and temperature over eastern North America. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(11), 2854–2859. <https://doi.org/10.1073/pnas.1614453114>
- Shen, L., Jacob, D. J., Mickley, L. J., Wang, Y., & Zhang, Q. (2018). Insignificant effect of climate change on winter haze pollution in Beijing. *Atmospheric Chemistry and Physics*, *18*(23), 17,489–17,496. <https://doi.org/10.5194/acp-18-17489-2018>
- Singh, D., Horton, D., Tsiang, M., Haugen, M., Ashfaq, M., Mei, R., et al. (2014). Severe precipitation in Northern India in June 2013: Causes, historical context, and changes in probability in “Explaining extreme events of 2013 from a climate perspective”. *Bulletin of the American Meteorological Society*, *95*(9), S1–S104. <https://doi.org/10.1175/1520-0477-95.9.S1.1>
- Singh, D., Swain, D., Mankin, J., Horton, D., Thomas, L., Rajaratnam, B., & Diffenbaugh, N. (2016). Recent amplification of the North American winter temperature dipole. *Journal of Geophysical Research: Atmospheres*, *121*, 9911–9928. <https://doi.org/10.1002/2016JD025116>
- Singh, D., Tsiang, M., Rajaratnam, B., & Diffenbaugh, N. (2014). Observed changes in extreme wet and dry spells during the south Asian summer monsoon season. *Nature Climate Change*, *4*, 456–461. <https://doi.org/10.1038/nclimate2208>
- Sippel, S., Walton, P., & Otto, F. (2015). Stakeholder perspectives on the attribution of extreme weather events: An explorative enquiry. *Weather Climate and Society*, *7*(3), 224–237. <https://doi.org/10.1175/WCAS-D-14-00045.1>
- Stone, D., & Allen, M. (2005). The end-to-end attribution problem: From emissions to impacts. *Climatic Change*, *71*(3), 303–318. <https://doi.org/10.1007/s10584-005-6778-2>
- Stott, P., Christidis, N., Otto, F., Sun, Y., Vanderlinden, J.-P., van Oldenborgh, G., et al. (2016). Attribution of extreme weather and climate-related events. *WIREs Climate Change*, *7*(1), 23–41. <https://doi.org/10.1002/wcc.380>
- Stott, P., Stone, D., & Allen, M. (2004). Human contribution to the European heat wave of 2003. *Nature*, *432*, 610–613. <https://doi.org/10.1038/nature03089>
- Stott, P., Tett, S., Jones, G., Allen, M., Mitchell, J., & Jenkins, G. (2000). External control of 20th century temperature by natural and anthropogenic forcings. *Science*, *290*, 2133–2137. <https://doi.org/10.1126/science.290.5499.2133>
- Stott, P. A., Gillett, N. P., Hegerl, G. C., Karoly, D. J., Stone, D. A., Zhang, X., & Zwiers, F. (2010). Detection and attribution of climate change: A regional perspective. *WIREs Climate Change*, *1*, 192–211. <https://doi.org/10.1002/wcc.34>
- Sun, Y., Jiang, Q., Wang, Z., Fu, P., Li, J., Yang, T., & Yin, Y. (2014). Investigation of the sources and evolution process of severe haze pollution in Beijing in January 2013. *Journal of Geophysical Research: Atmospheres*, *119*, 4380–4398. <https://doi.org/10.1002/2014JD021641>
- Swain, D., Langenbrunner, B., Neelin, J., & Hall, A. (2018). Increasing precipitation volatility in twenty-first-century California. *Nature Climate Change*, *8*(5), 427–433. <https://doi.org/10.1038/s41558-018-0140-y>
- Swain, D., Tsiang, M., Haugen, M., Singh, D., Charland, A., Rajaratnam, B., & Diffenbaugh, N. (2014). The extraordinary California drought of 2013/2014: Character, context, and the role of climate change in “Explaining extreme events of 2013 from a climate perspective”. *Bulletin of the American Meteorological Society*, *95*(9), S1–S104. <https://doi.org/10.1175/1520-0477-95.9.S1.1>
- Tan, X., Gan, T., & Horton, D. (2018). Projected timing of perceivable changes in climate extremes for terrestrial and marine ecosystems. *Global Change Biology*, *24*, Advance Online Publication(10), 4696–4708. <https://doi.org/10.1111/gcb.14329>
- Taylor, K., Stouffer, R., & Meehl, G. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- van der Wiel, K., Kapnick, S., van Oldenborgh, G., Whan, K., Philip, S., Vecchi, G., et al. (2017). Rapid attribution of the August 2016 flood-inducing extreme precipitation in south Louisiana to climate change. *Hydrology and Earth System Sciences*, *21*(2), 897–921. <https://doi.org/10.5194/hess-21-897-2017>
- Vautard, R., Colette, A., van Meijgaard, E., Meleux, F., van Oldenborgh, G., Otto, F., et al. (2017). Attribution of wintertime anticyclonic stagnation contributing to air pollution in Western Europe in “Explaining extreme events of 2016 from a climate perspective”. *Bulletin of the American Meteorological Society*, *99*(1), S70–S75. <https://doi.org/10.1175/BAMS-D-17-0113.1>
- Wang, H., Chen, H., & Liu, J. (2015). Arctic sea ice decline intensified haze pollution in eastern China. *Atmospheric and Oceanic Science Letters*, *8*, 1–9. <https://doi.org/10.3878/AOSL20140081>
- Wang, J., & Angell, J. (1999). Air stagnation climatology for the United States (1948–1998). NOAA/Air Resources Laboratory ATLAS No. 1 1999.
- Wang, X., Dickinson, R. E., Su, L., Zhou, C., & Wang, K. (2018). PM<sub>2.5</sub> pollution in China and how it has been exacerbated by terrain and meteorological conditions. *Bulletin of the American Meteorological Society*, *99*, 105–119. <https://doi.org/10.1175/BAMS-D-16-0301.1>
- Wang, X., Wang, K., & Su, L. (2016). Contribution of atmospheric diffusion conditions to the recent improvement in air quality in China. *Scientific Reports*, *6*(1), N36404. <https://doi.org/10.1038/srep36404>
- Wang, Y.-S., Yao, L., Wang, L.-L., Liu, Z.-R., Ji, D.-S., Tang, G.-Q., 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>
- Wetzel, S., & Martin, J. (2001). An operational ingredient-based methodology for forecasting midlatitude winter season precipitation. *Weather and Forecasting*, *16*, 156–167. [https://doi.org/10.1175/15200434\(2001\)016<0156:AOIBMF>2.0.CO;2](https://doi.org/10.1175/15200434(2001)016<0156:AOIBMF>2.0.CO;2)



- Yang, Y., Russell, L. M., Lou, S., Liao, H., Guo, J., Liu, Y., et al. (2017). Dust-wind interactions can intensify aerosol pollution over eastern China. *Nature Communications*, *8*, 15333. <https://doi.org/10.1038/ncomms15333>
- Yang, Y. Q., Wang, J. Z., Gong, S. L., Zhang, X. Y., Wang, H., Wang, Y. Q., et al. (2016). PLAM—A meteorological pollution index for air quality and its application in fog-haze forecasts in North China. *Atmospheric Chemistry and Physics*, *16*(3), 1353–1364. <https://doi.org/10.5194/acp-16-1353-2016>
- Ye, X., Song, Y., Cai, X., & Zheng, H. (2016). Study on the synoptic flow patterns and boundary layer process of the severe haze events over the North China Plain in January 2013. *Atmospheric Environment*, *124*, 129–145. <https://doi.org/10.1016/j.atmosenv.2015.06.011>
- Yin, P., He, G., Fan, M., Chiu, K., Fan, M., Liu, C., et al. (2017). Particulate air pollution and mortality in 38 of China's largest cities: time series analysis. *British Medical Journal*, *356*, j667. <https://doi.org/10.1136/bmj.j667>
- 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>
- Zhang, R., Jing, J., Tao, J., Hsu, S.-C., Wang, G., Cao, J., et al. (2013). Chemical characterization and source apportionment of PM<sub>2.5</sub> in Beijing: Seasonal perspective. *Atmospheric Chemistry and Physics*, *13*(14), 7053–7074. <https://doi.org/10.5194/acp-13-7053-2013>
- Zheng, G., Duan, F., Su, H., Ma, Y., Cheng, Y., Zheng, B., et al. (2015). Exploring the severe winter haze in Beijing: The impact of synoptic weather, regional transport and heterogenous 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*, e1602751. <https://doi.org/10.1126/sciadv.1602751>
- Zwiers, F., Zhang, X., & Feng, J. (2011). Anthropogenic influence on long return period daily temperature extremes at regional scales. *Journal of Climate*, *24*, 881–892. <https://doi.org/10.1175/2010JCLI3908.1>