

Behind the Smoke: An Extreme Value Analysis of Air Pollution in Minnesota

Jacob Flignor, Libby Nachreiner, Yicheng Shen, Karen Wang

February 15, 2022

Poor air quality is a major environmental health threat. Even short-term exposure to poor air quality—such as during extreme pollution events—can cause severe respiratory distress. While there have been significant decreases in Minnesota air pollution levels over the past 40 years, the summer of 2021 upset this trend with Hennepin County reporting the highest particulate measure in the past 20 years. This study focuses on analyzing the extreme values of pollutant concentration levels of sulfur dioxide (SO_2) and fine inhalable particles ($\text{PM}_{2.5}$) across three Minnesota counties as collected by the Environmental Protection Agency from 1980 to 2021. We employ extreme value analysis methods to fit the pollutant data. The models find that SO_2 levels have fallen substantially since 1980 in accordance with EPA policies regulating diesel fuel and coal power plants. This dramatic decrease has made the magnitude of severe pollution incidents appear far more extreme than in earlier decades, with typical events in the 1980-1990s often equating to one in a hundred year events today. By contrast, no downward trend in $\text{PM}_{2.5}$ levels was observed over the past 20 years, an expected result given that $\text{PM}_{2.5}$ has more varied sources and is therefore harder to regulate than SO_2 . However, models show a significant seasonal trend with peaks during winter months, revealing this past ‘summer of smoke’ as particularly extreme.

Keywords: Extreme Value Analysis; Generalized Extreme Value Distribution (GEV); Generalized Pareto Distribution (GP); Air Quality

1 Introduction

The summer of 2021 was one of Minnesota’s hottest and driest summers, though it was perhaps most notable for its historic pollution (Minnesota Department of Natural Resources 2022). Called the “summer of smoke,” pollution levels around the state reached an unprecedented high, as seen in Figure 1 (Minnesota Department of Natural Resources 2021). The Minneapolis-St Paul Airport reported smoke for 25 days between June and August. During the peak of pollution, from July 28 through August 6, the Minnesota Pollution Control Agency reported that the air quality in all Minnesota counties was unhealthy for all people. The dangers of these types of events are immense. Even after visible smoke has reduced, remaining small particles in the air still pose dangers. The California Air Resources Board has observed premature mortality, increased hospital admissions for heart or lung causes, acute and chronic bronchitis, asthma attacks, emergency room visits, and respiratory symptoms from consistent exposure to these particulates. These adverse health effects have been reported primarily in infants, children, and older adults with preexisting heart or lung diseases (California Air Resources Board 2021).



Figure 1: Comparison of severe air pollution in the city of Minneapolis July 15, 2021 v.s July 29, 2021 (Ian Leonard)

In a broader, historical context however, the current levels of particulate air pollution are far from their worst. Air pollution compounded and grew increasingly dire in America starting from the 19th century, coal-driven Industrial Revolution throughout the post-WWII industrial boom (Air Quality Life Index 2022). The notably hazardous air quality—and general environmental concerns—drew public attention, culminating in a landmark 1970 amendment to the federal Clean Air Act that allowed the Environmental Protection Agency (EPA) to set federal air quality regulations (Air Quality Life Index 2022). Improvements in air pollution alone, largely due to such regulation, have added 1.5 years to the life expectancy of the average American since 1970 (Air Quality Life Index 2022).

These two narratives reveal different forces in modern day air quality: while pollution has broadly been reduced in the past 50 years, there are still extreme pollution events that affect our day-to-day lives. It is important to recognize that air pollution is often singularly measured and examined by the Air Quality Index (AQI), as demonstrated in contemporary studies (Wang et al. 2017; Liu and Zhang 2021; Varapongpisan, Ingsrisawang, et al. 2019). However, Kanchan, Gorai, and Goyal (2015) point out a significant imperfection in this artificial index: while AQI is intended to holistically represent air pollution composed of several pollutants, its value is solely determined by the highest individual pollutant level. This poses problems when examining trends of individual pollutants, with different sources, behaviors, and consequences.

To better understand the intricacies of air quality, we examined direct measurements of concentrations of sulfur dioxide (SO_2) and fine inhalable particles, with diameters that are 2.5 micrometers and smaller ($\text{PM}_{2.5}$), from the EPA’s Air Quality Metadata System. Exposure to SO_2 can harm the human respiratory system and make breathing difficult, although its emissions are common in many places. The largest source of SO_2 in the atmosphere is the burning of fossil fuels by power plants and other industrial facilities (Chen et al. 2007). Smaller sources of SO_2 emissions include industrial processes such as extracting metal from ore, natural sources such as volcanoes, and locomotives, ships and other heavy equipment that burns fuel with a high sulfur content. $\text{PM}_{2.5}$ is a broader category of pollutants. Such particulate matter can be composed of hundreds of different chemicals (Tucker 2000). Direct sources of $\text{PM}_{2.5}$ include construction sites, unpaved roads, fields, smokestacks or fires, power plants, and automobiles. In Minnesota, the majority of $\text{PM}_{2.5}$ comes from chemical reactions between other pollutants, including SO_2 , ammonia, nitrogen oxides, natural gasses, and vapors (Minnesota Pollution Control Agency 2022). Inhalation of $\text{PM}_{2.5}$ is especially dangerous, as the small particles are able to travel deep into one’s respiratory tract and lungs (New York State Department of Health 2022).

To quantify extreme events and contextualize them within a greater trend, we employ extreme value analysis (EVA). The objective of EVA is to quantify the stochastic behavior of a process at unusually large or small levels, thereby providing a framework that enables extrapolation. Focusing specifically on measures of extremity, as opposed to measures of center, is valuable in this pollution context since these events often constitute the most immediate health consequences and are most noticeable to the general population.

In this report, we first explain the theoretical foundations of EVA. Then, we discuss various EVA models of two pollutants— SO_2 and $\text{PM}_{2.5}$ —to make conclusions about Minnesota’s air pollution.

2 Methods

There are two principal approaches to EVA: the block maxima method and the threshold exceedance method. Concerns regarding independence and nonstationary may be addressed through declustering and covariates, respectively.

2.1 Block maxima method

The extremal types theorem shown below states that the only possible limit distribution of properly normalized maxima (or minima) of a sequence of independent and identically distributed random variables is the extreme value distribution.

Theorem 1.1 If there exists a sequence of constants $a_n > 0$ and b_n such that

$$Pr((M_n - b_n)/a_n \leq z) \rightarrow G(z) \text{ as } n \rightarrow \infty$$

where G is a non-degenerate distribution function, then G belongs to one of the following families:

$$I : G(z) = \exp\{-\exp[-(\frac{z-b}{a})]\}, -\infty < z < \infty;$$

$$II : G(z) = \begin{cases} 0, & z \leq b, \\ \exp-(\frac{z-b}{a})^{-\alpha}, & z > b; \end{cases}$$

$$III : G(z) = \begin{cases} \exp-[-\frac{z-b}{a}]^{\alpha}, & z < b, \\ 1, & z \geq b, \end{cases}$$

for parameters $a > 0$, b and, in the case of families II and III, $\alpha > 0$.

As Theorem 1.1 shows, the normalizing of sequences leads to a non-degenerate distribution function of a series of maxima known as the Generalized Extreme Value (GEV) distribution. The GEV distribution is composed of three families, which are distinguished by their tail behavior. The Weibull (III) distribution has an infinite lower tail but finite upper bound. The Fréchet (II) distribution, by contrast, has an infinite upper tail but finite lower bound. The Gumbel (I) distribution, unlike the Fréchet and Weibull distributions, has both an infinite lower and upper tail. The conclusions we draw about the behavior of extreme values are thus greatly impacted by our choice of extreme value distribution for modeling. This is a problem because we don't want to make a prior assumption that our data fits with one of these families.

Fortunately, this issue of varying tail behavior among the three families is resolved by reformulating them into the Generalized Extreme Value (GEV) distribution. The GEV distribution function is an approximation

of the distribution of maxima that uses three parameters. The location parameter, describes how shifted the distribution is horizontally, and σ , the scale parameter, describes how spread out the distribution is. The shape parameter, is the most important parameter because it defines the model. A negative shape parameter means that we are modeling with a Weibull distribution, a positive shape parameter indicates that we are modeling with a Fréchet distribution, and a shape parameter of 0 corresponds to a Gumbel distribution. Therefore, through inference on the shape parameter of the GEV distribution function, we can determine the best tail behavior modeled by our data and give measures of uncertainty.

The block maxima approach that follows from the extreme value theorem is the first standard approach to modeling extremes. Under the block maxima approach, data are blocked into sequences of equal length, generating a series of block maxima to which a GEV distribution can be fitted. The choice for block size has a bias-variance tradeoff. With smaller blocks, there will be more maxima included in the model, so the variance of the estimates of the parameters decreases will decrease, but bias will increase as data points at lower levels are deemed maxima. Similarly, an increase in block size would decrease bias while increasing variance. However, in practice, block sizes are often chosen to correspond to a certain familiar time period. A length of one year is a common choice, then, due to its practicality in data collection and ability to remove any seasonal fluctuations.

2.2 Threshold Exceedance Approach

A major limitation of the block maxima approach is that it excludes extreme events other than the most extreme one per block, leading to information loss. A secondary approach to modeling extreme events is that of threshold exceedance, in which the events are considered “extreme” if they exceed a certain threshold u . The distribution of the amount by which events exceed u , given that they do, can be asymptotically approximated by a distribution within the Generalized Pareto family. The distribution function of the Generalized Pareto Distribution (GPD) is as follows:

$$H(Y) = 1 - \left(1 + \frac{\xi Y}{\sigma}\right)^{-\frac{1}{\xi}}$$

ξ is the shape parameter, and σ is the scale parameter, It is notable that the shape parameter, ξ_{GPD} , is equal to the shape parameter from the corresponding GEV distribution, ξ_{GEV} . Moreover, the GPD scale parameter is a function of the location and scale parameters from the corresponding GEV distribution: $\sigma_{GPD} = \sigma_{GEV} + \xi_{GPD}(u_{GPD} - \mu_{GEV})$.

Threshold selection in threshold exceedance models—like block size selection for block maxima models—is challenged by a trade-off between bias and variance. At too low a threshold, the asymptotic basis of the

limiting distribution will not be met, leading to bias in estimations. At too high a threshold, there will be few excesses to model upon, leading to more variability in estimates.

Generally, the lowest threshold at which the limit model is reasonable is selected. There are two methods to determine what this level is, which are mean residual life plot and threshold plot.

The fit of a generalized Pareto model may be assessed through examining probability plots, quantile plots, return level plots, and density plots.

2.3 Return Levels

Return levels and return periods are useful ways of interpreting this relationship between a model and its parameters. We can quantify an event as a 1 in 1/p-year event this way, where 1/p years would be the return period for that achieved level. Another way to interpret a model is to use a return level, saying that a specified return level z_p is expected to be exceeded once every 1/p year.

A return level, z_p , for the GEV distribution can be calculated using the following formula:

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} [1 - \{-\log(1-p)\}^{-\xi}], & \xi \neq 0, \\ \mu - \sigma \log\{-\log(1-p)\}, & \xi = 0, \end{cases}$$

where $G(z_p) = 1 - p$.

Return levels for threshold exceedance models may be calculated as follows, where x_m is an m -observation return level (return level for every m observations). Moreover, $\zeta_m u = \Pr X > u$, the probability any event being considered “extreme.” This is estimated as the sample proportion of events that exceed μ : $\hat{\zeta}_\mu = \frac{k}{n}$.

$$\text{For } \xi \neq 0: x_m = u + \frac{\sigma}{\xi} [(m\zeta_u)^\xi - 1]$$

$$\text{For } \xi = 0: x_m = u + \sigma \log(m\zeta_u)$$

It is easier to interpret return levels on an annual scale. In this case, an N -year return level is equivalent to the m -observation return level where $m = N(n_y)$, and n_y is the number of observations per year.

Therefore:

$$\text{For } \xi \neq 0: Z_N = u + \frac{\sigma}{\xi} [(Nn_y\zeta_u)^\xi - 1]$$

$$\text{For } \xi = 0: Z_N = u + \sigma \log(Nn_y\zeta_u)$$

These quantile expressions enable probability models to be expressed on the scale of the data, which allows for easier interpretation of a model to its parameters.

2.4 Return Level Plots

Additionally, return level plots are helpful for visual interpretation of the relationship between a model and its parameters. The choice of scale compresses the tail of the distribution so that the effect of extrapolation is highlighted in return level plots, making them especially convenient for the interpretation and validation of a model. Within each block of a return level plot, $1/p$ th of the data are above the p -year return level.

2.5 Nonstationarity & Dependence

Both of the methods introduced so far require independent and identically distributed data. Pollution data is often temporally correlated and non-stationary. GEV models with the raw pollution data resulted in uncertain fits. Thus, we applied various statistical methods to attain data with independent extremes. We first discuss methods addressing dependence and then introduce an approach to address non-stationarity.

In the GEV context, we do not need to be all too concerned about dependence. Generally, in environmental datasets, we can safely assume long range dependence is weak, which implies the block maxima are asymptotically independent.

While we do not have to be concerned about dependence in the GEV context, we do need to need to adjust the data to fit a reasonable threshold exceedance model. Dependence is an issue for GPD models as it is likely that a given pollution event will have more than one observation above the threshold. Ideally, we want to keep the largest record exceedance from every pollution event. This way the effects of dependence can be removed, yet the amount of data kept is maximized.

We employ a declustering algorithm to remove this dependence. The declustering algorithm has three steps. Firstly, the process starts when a data point is found above the previously determined threshold. That data point opens the cluster. A cluster is then closed when there are r consecutive values found below the threshold for some chosen run length r . After a cluster is closed, the highest value in the cluster is kept; all other data points are dropped. Then, the GPD model is fit with the highest value in each cluster. Similar to when selecting a threshold, there is a bias-variance trade off when selecting a run length. Too small a value will lead to the problem of independence being unrealistic for nearby clusters; too large a value will lead to a concatenation of clusters that could reasonably have been considered as independent, and therefore to a loss of valuable data.

Different methods are required to address non-stationarity. In our analysis, we employ covariates that model the general trends in pollution data. When considering these trends, we can bet more confident that our model results in a better, and more certain, fit. A generic example of a covariate model is shown below in the GEV and GPD context

$$Z_t \sim GEV(\mu(t), \sigma_1(t), \xi(t))$$

$$X_t \sim GPD(\sigma_2(t), \xi(t))$$

where Z_t are the annual maxima, X_t are the threshold exceedances and $\mu(t), \sigma_i(t), \xi(t)$ are functions of covariates. All of the parameters are shown as functions of covariates that, in theory, model the general trends in the data. In this example, the shape parameter (ξ) in both models is generally not modeled with a covariate. EVA literature cautions against estimating the shape parameter as it is particularly hard to precisely model ξ .

3 Data

The main focus of this project is to understand the air pollution in Minnesota, therefore we employ data analysis on the extreme daily measurements of two pollutant data, SO_2 and $\text{PM}_{2.5}$, over time.

3.1 Pollutant Data

The federal Clean Air Act requires the EPA to set National Ambient Air Quality Standards (NAAQS) for principal pollutants that are harmful to public health and the environment (Driesen 1996). NAAQS includes standards that govern the amounts of ground-level ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), particulate matter (PM_{10} or $\text{PM}_{2.5}$), lead (Pb), and nitrogen dioxide (NO_2) in the ambient air. The EPA collects these air quality data from individual states counties, and this information is compiled in a publicly available database.

The **RAQSAPI** package (version 2.0.1) was used to access the interface of the EPA’s database and extract daily air pollutant records (McCrowey et al. 2021). SO_2 records were available from 1980 to 2021, monitored in Hennepin and Dakota counties. $\text{PM}_{2.5}$ records are available from 2000 to 2021, monitored in Hennepin and Saint Louis counties. We chose these counties because the EPA has kept the longest and most consistent records there. The 1 hour sample duration was selected for SO_2 , and the 24 hour sample duration was selected for $\text{PM}_{2.5}$ because these were the most consistently available measures.

Data from Hennepin County was recorded at 528 Hennepin Ave. Data in Dakota County was collected at 12821 Pine Bend Trail. The measurement site for St. Louis County was 1202 East University Circle.



Figure 2: The distribution of daily and annual sulfur dioxide records in Hennepin and Dakota counties.

3.2 SO₂ EDA & Policy Covariate

SO₂, measured by parts in billion (ppb), was monitored in Minnesota starting from 1980. Due to limitations from the EPA's data collection history, the monitoring site in Dakota county started reporting SO₂ nine years earlier (first on January 1, 1980) than that in Hennepin (first on October 6, 1989). The former has 11,049 daily observations and the latter has 16,704 daily observations. The concentration of SO₂ shows a decreasing trend at both sites over the years as shown in Figure 2.

There are several dramatic drops in the observed SO₂ in Hennepin throughout the past decades, most notably after the 21st century. For example, the annual SO₂ maximum in 2011 was 84.1% lower than that in 2010. A linear time trend is insufficient to explain these sudden drops, requiring us to incorporate more meaningful covariates that influence SO₂ in the region.

According to the U.S. Energy Information Administration, coal-fired power plants have provided the largest share of Minnesota's electricity net generation until 2020, when their contribution fell below the amount supplied by renewables and nuclear power for the first time (Goldstein, Gounaridis, and Newell 2020). In other words, regulations of traditional energy power plants would have particularly strong effects on the SO₂ emissions in Minnesota since they play such a vital role in providing statewide energy and operate on a large scale. Our policy research has thus identified several significant time points where policy implemented

may have affected SO₂ pollution in Minnesota.

- In 1984, new regulations on power plant smokestacks required utilities and factories across the country to reduce emissions of sulfur dioxide by 8 percent (Latin 1984).
- The most recent amendment to the Clean Air Act in 1990 promotes the use of clean low sulfur coal and natural gas, as well as innovative technologies to clean high sulfur coal. The amendment ultimately reduced energy waste and created enough of a market for cleaner fuels (Ferrall 1991).
- Beginning in 2006, the EPA began to phase-in more stringent regulations to lower the amount of sulfur in diesel fuel to 15 parts per million (ppm). This fuel is known as ultra-low sulfur diesel (ULSD). Since 2007, low sulfur diesel fuel (specified at 500 ppm) and ULSD fuel was also phased in for nonroad, locomotive, and marine (NRLM) diesel fuel (Ris 2007). The regulations were expanded to all vehicles in 2010.
- On June 2, 2010, the EPA revised the SO₂ standard, establishing a stricter one-hour primary standard of 75 ppb. On June 2, for the first time, the EPA proposed to regulate the disposal of coal combustion residuals from electric utilities. On December 16, 2011, the Environmental Protection Agency finalized the first national standards to reduce air pollution from coal-fired power plants (EPA, n.d.).

We observe a step-by-step enhancement of regulations on addressing SO₂ emissions nationwide from 1980 through today. In particular, our policy research indicates that established policies are usually held in place or further tightened in the next regulation period. We therefore created 5 distinct policy phases based on important regulations when were implemented. The phases are A: 1980-1984, B: 1985-1990, C: 1991-2006, D: 2007-2010, and finally E: 2011-2021. By this classification, Hennepin’s records contain four policy phases and Dakota contains five phases. Through the exploratory data analysis shown in Figure 2, we perceive each policy phase as an important indicator of major changes in SO₂ trends throughout the years.

3.3 PM_{2.5} EDA and Seasonality Component

Measures for PM_{2.5} were collected using the Andersen RAAS2.5-300 PM_{2.5} SEQ w/WINS Gravimetric method at two locations in Hennepin County (2727 10th Avenue South and 5005 Minnetonka Blvd) and one location in St. Louis County (1202 East University Circle), starting in 2000. The sample duration for these measure was 24 hours, with the unit of measure being $\mu g/m^3$ under the 2012 NAAQS PM_{2.5} 24-hour standards.

The PM_{2.5} records in both counties do not show a clear trend on PM_{2.5} levels over the past 2 decades, as shown in Figure 3. However, our data from Hennepin County does show that this past summer’s high levels of PM_{2.5} were perhaps higher than we would expect, with an annual maximum recorded on July 30th of about $61 \mu/m^3$, far higher than the average value of PM_{2.5} recorded in our Hennepin record at about 9

μ/m^3 . The average value of $PM_{2.5}$ was $5.85 \mu/m^3$ in St. Louis County, with a noticeable maximum of $50.2 \mu/m^3$ recorded in February of 2008.

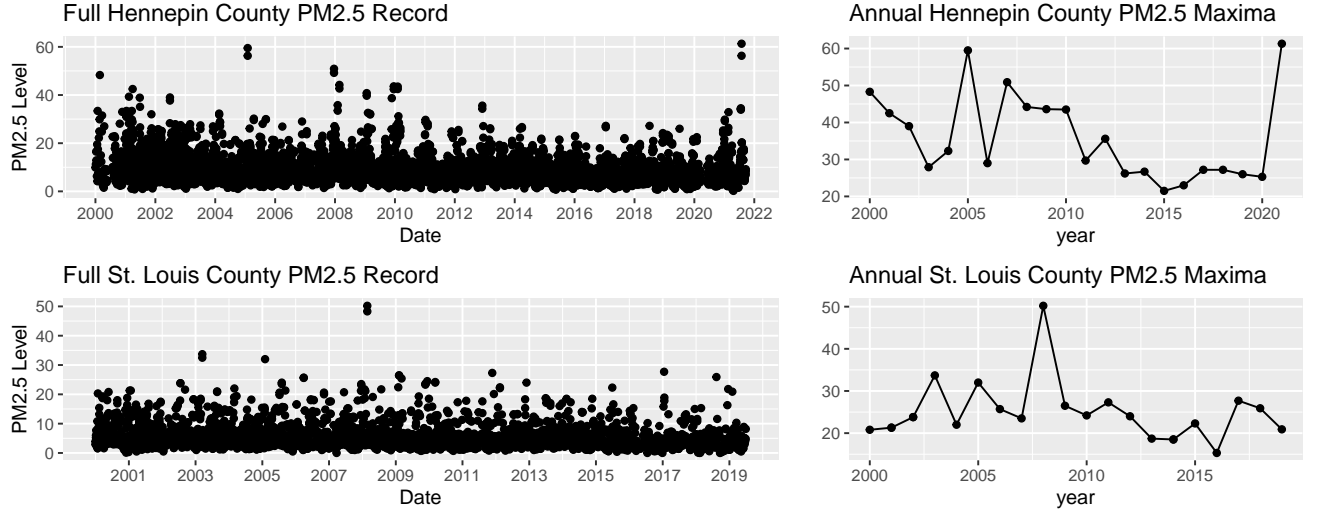


Figure 3: Exploratory data analysis for $PM_{2.5}$ data

4 Results

For both SO_2 and $PM_{2.5}$, we fit models based on GEV and GP distributions. The model structures differ by the covariate used and the consideration of seasonality. The results are shown below.

4.1 Sulfur Dioxide

Modeling Annual Maxima

The GEV distribution was initially fit to SO_2 annual maxima from both Hennepin and Dakota counties. The naïve model, along with models with a year and policy phase covariate, were tested. A year covariate seems reasonable as year proves to be a good predictor of pollution level, with years near 1980 resulting in higher SO_2 levels and more recent years resulting in a lower SO_2 record. Based on AIC and BIC, the model with the covariates was preferred.

All of these models, however, contained large standard errors for parameter estimates and hard to interpret return levels that limited the model's legitimacy. For Hennepin, the standard error for the shape parameter ($\sigma_{\xi_H} = 0.187$) was greater than the estimate itself ($\xi_H = 0.167$). For Dakota, the shape parameter was slightly more certain with an estimate of $\xi_D = 1.184$ and a standard error of $\sigma_{\xi_D} = 0.226$. Specifying the GEV models as Gumbel distributions does not remedy the issue.

The uncertain fit from these GEV models seems to be due to information loss issues. In fitting the GEV distribution, roughly 99.8% of the data is dropped as only the largest event every year is preserved. While a certain amount of information loss is required for EVA analysis, the GEV method may not be able to produce a reasonable fit with only 30 or 40 observations. Research on sample size limitations with GEV methods indicates that a sample size should be considered reasonable if the given sample size can make satisfactory statistical inferences on the 100-year return level (Yuzhi Cai 2010). Given the return level plots, it is clear that the 33 annual maxima in Hennepin County and the 42 annual maxima in Dakota County are not able to adequately infer the 100-year return level.

Modeling Threshold Exceedance

These concerns of large uncertainty and small sample size from GEV models motivated a different investigation into the SO_2 records using the GPD as threshold exceedance models. Since the GPD model does not restrict the data to exclusively block maxima, this model limits the amount of information loss relative to the GEV models. We continued to model SO_2 with the policy covariate since SO_2 records exhibit a decreasing trend across policy phases in two counties.

The SO_2 record shows clear temporal dependence in the data, meaning that the record from one day could be highly correlated with past and future days. For example, from February 26 to February 29, 2004, the records indicate that there were consecutive days having SO_2 concentrations greater than 90 ppb. As mentioned in section 2, we can reduce the influence of these clustered extremes and remove the dependence in the data by the declustering algorithm. With this data, we chose a run length of five, meaning that the fifth data point that does not exceed the threshold marks the beginning of a new cluster.

To have a reasonable threshold choice for the declustered observations, we followed the procedures for threshold selection by Coles et al. (2001) and estimated the fitted GPD model at a range of thresholds. We then compared the resulting model estimations of the shape parameter as shown in Figure 4 where the maximum likelihood estimates of the shape parameter are plotted against the percentiles of declustered distribution.

Taking into account the confidence intervals and locating the point where estimates of the shape parameter begin to stay approximately constant, we decided to choose the 90 percentile of distributions as the threshold for each policy phase in Hennepin and 85 percentile threshold for those in Dakota.

Based on our threshold selection, there are 1,258 declustered observations above the 90% threshold in Hennepin and 2,606 declustered observations above the 85% threshold in Dakota. These exceedance cases, highlighted in Figure 5, are of primary interests in the GPD models.

In the next step, we fit the GPD model on the two declustered SO_2 data sets, with the scale parameter being modeled by the policy phases as shown in Equation (1) below. $H(y)$ is the distribution function of

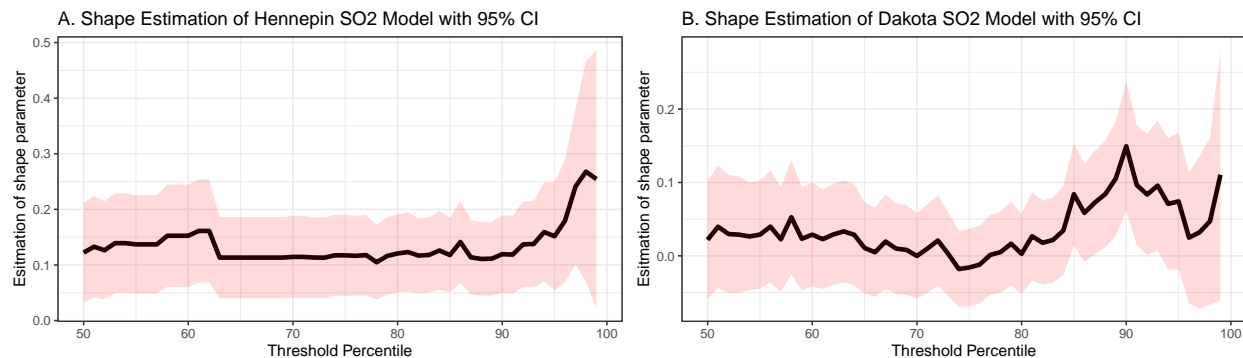


Figure 4: Estimations of the shape parameter by different threshold choices: The threshold plot relies on the idea that once a valid threshold is reached, the parameter estimates will become stable. Therefore, we generate the shape estimates with confidence intervals for models throughout a range of potential thresholds. We then select the lowest percentile threshold at which point the parameter estimates start appearing constant.

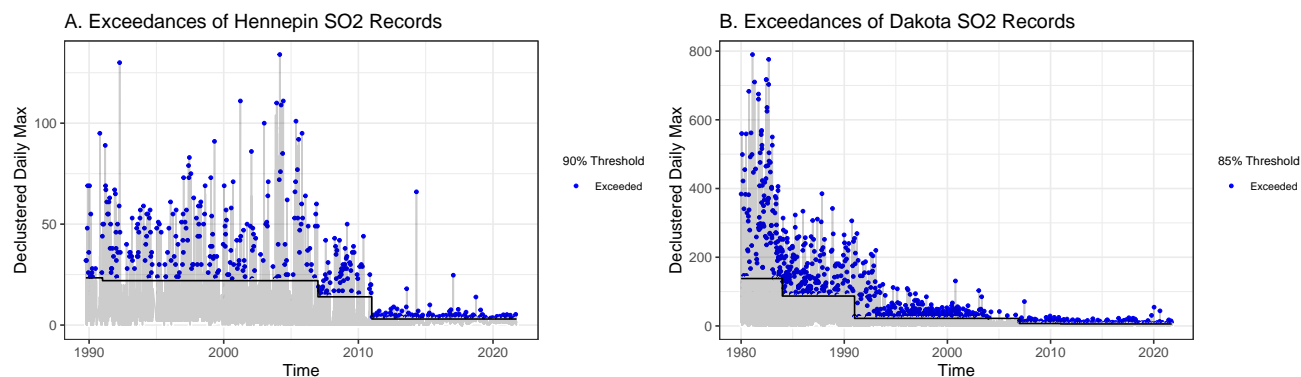


Figure 5: The declustered exceedances of daily sulfur dioxide records in Hennepin and Dakota at their thresholds respectively.

declustered exceedences. All parameters of the GPD models are estimated by maximum likelihood. The models' summaries are presented in Table 1 below. Relative to the GEV models tested previously, we now observe lower standard errors of the parameters estimated, thereby reducing the amount of uncertainty in our models.

$$H(y) = 1 - \left(1 + \frac{\xi y}{\sigma_y}\right)^{\frac{-1}{\xi}} \quad (1)$$

where

$$\sigma_y = \sigma_a + \sigma_b * (\text{Phase B}) + \sigma_c * (\text{Phase C}) + \sigma_d * (\text{Phase D}) + \sigma_e * (\text{Phase E})$$

y : declustered exceedences; ξ : shape parameter; σ_i : scale parameter by policy phase

Phase B-E : Indicator variable for current phase

The distribution plots of GPD models (seen in Appendix Figure 11 and 12) show that the exceedance observations are fairly low values in policy phase D and E, mostly contributed by an overall decreasing trend. In the most recent decade, the estimated probability of exceeding 75 ppb is very low whereas in the 1990s there was a much higher probability of exceeding this unhealthy threshold.

Dakota's situation given the early years was much worse and its SO₂ levels were way above the unhealthy levels for sensitive groups. From the 1980s to the 1990s, there was over a five percent chance of observing high SO₂ levels exceeding 185 ppb, so extreme that the air was unhealthy for all of the population. Despite that, in the most recent period, the distributions of their exceedances are also concentrated in lower values similar to Hennepin's.

Table 1: SO2 GPD Model Summary

Hennepin GPD	Estimates	SE	Dakota GPD	Estimates	SE
sigma_b (B)	14.517	4.041	sigma_a (A)	168.406	13.802
sigma_c (C vs B)	9.302	4.369	sigma_b (B vs A)	-92.618	14.721
sigma_d (D vs B)	-2.647	4.347	sigma_c (C vs A)	-139.351	13.645
sigma_e (E vs B)	-12.874	4.034	sigma_d (D vs A)	-161.272	13.780
shape (xi)	0.119	0.036	sigma_e (E vs A)	-164.260	13.764
			shape (xi)	0.084	0.036

In terms of the return values, we obtain the maximum likelihood estimates from fitted GPD models

at return periods of 2, 20 and 100 years. The numbers of exceedances over the return levels are generally what we expected to see: for Hennepin SO₂ model, the observed values exceeded the 2-year return levels, or 730-observation return levels, 8 times over the past 33 years, whereas Dakota SO₂ model's 2-year return levels were exceeded 21 times over the 41 years.

Table 2: Return Level Summary

County	Policy Phase	2-year level	20-year level	100-year level
Dakota	A	852.665	1432.984	1910.666
Dakota	B	408.620	669.780	884.751
Dakota	C	145.299	245.419	327.832
Dakota	D	37.272	61.854	82.088
Dakota	E	23.395	37.682	49.443
Hennepin	B	83.494	140.988	191.668
Hennepin	C	120.600	214.936	298.089
Hennepin	D	63.138	110.152	151.592
Hennepin	E	9.804	16.314	22.052

We also use the delta method to calculate the standard errors and subsequently 95% confidence intervals of the return levels at each return period. The return level plots and confidence intervals are presented in Figure 6, showing a step-wise distribution due to policy phases.

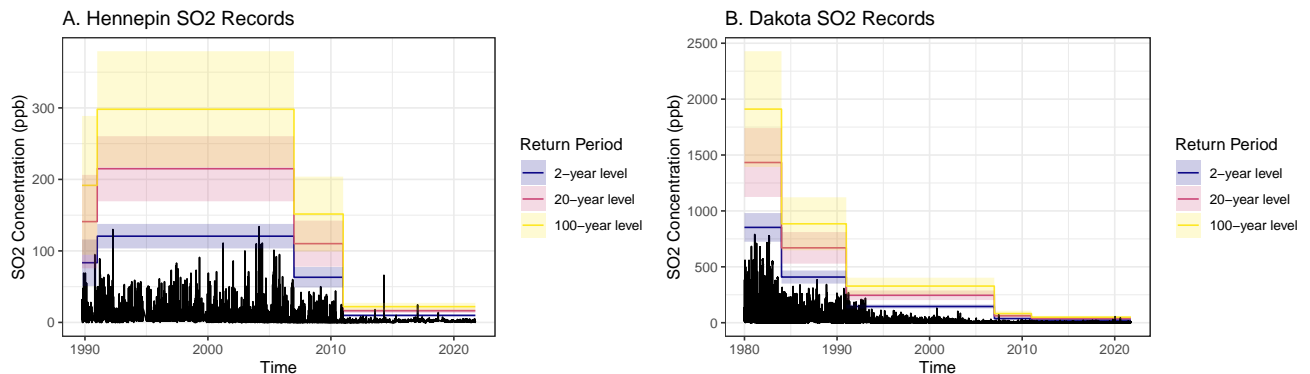


Figure 6: Return levels and their 95 percent confidence intervals of daily sulfur dioxide records in Hennepin and Dakota for return periods of 2, 20 and 100 years.

Discussion of SO₂ Model Implications

Modeling the threshold exceedances of SO₂ provides us with insights of how the level of this particular type of air pollutant has changed in Minnesota over the years. The decreasing trends of exceedance values from both SO₂ GPD models are consistent with the history of pollution regulation. This points to the strong association between policy and SO₂. As pollution regulation strengthened over the past 40 years, SO₂

pollution dramatically dropped. Interpretation from return levels also shows this dynamic. In Hennepin, the two-year return level at 1992 (during policy phase B) was roughly 115 ppb, and the two-year return level in 2021 (policy phase E) was around 25 ppb. This means that a pollution event extreme enough to be observed only every two years would have needed to be 115 ppb in 1992, but nowadays, with much stricter pollution regulation in place, that value has reduced to 25 ppb.

Further, the return levels also show the reduced health concerns from SO_2 in present day. While observing a pollution event larger than 75 ppb, the unhealthy amount of SO_2 pollution, in both Hennepin and Dakota counties was fairly likely (lower than the two-year return level) in 1980-2000, this is no longer the case in 2020. Indeed, to observe such a level in policy phase E would be more than a 1 in 500-year event.

Moreover, the interpretation of return levels helps us compare the two fitted GPD models respectively. When doing so, there is reason to believe that the Hennepin model is a bit better than the Dakota equivalent. In particular, the transition to policy phase C in Dakota includes a fair number of clustered exceedances which, more than anything else, reveals the imperfect nature of the policy phase covariate. Specifically, there are 16 exceedances in the transition to policy phase C, when the GPD model itself only expects around three exceedances over the two-year return level. Relevant to both models, however, is the fact that the extreme events recorded in the past 5 or so years are modeled to be extremely extreme. For example, the extreme event in Hennepin in 2014 is predicted to be more than a 1 in 250-year event. These remarkably unlikely events potentially indicate an imperfect shape parameter. A potential solution not assessed in this analysis would be to model the shape parameter with the policy phase covariate. While we prioritized analyzing simpler models, this change may result in a superior fit.

4.2 $\text{PM}_{2.5}$

Baseline Models

Naive GEV models were fit to the annual $\text{PM}_{2.5}$ maxima values in Hennepin and St. Louis counties. For Hennepin county, two models were fit: one including and one excluding the data from the extreme events from past summer. Return levels from these naive models fell in line with expectations, with the Hennepin model estimating that the past summer’s event was a 16- to 23-year-event. However, like with the naive models of SO_2 levels, these models suffered from unstable shape parameters. For Hennepin County, $\hat{\xi} = 0.307$ (SE = 0.263). For St. Louis County, $\hat{\xi} = 0.097$ (SE = 0.126).

Similarly, naive GPD models were fit for the two counties, with the thresholds for exceedance set at the 90th percentile of $\text{PM}_{2.5}$ levels for each respective county. In Hennepin County, this was $16.52 \mu/m^3$; in St. Louis County, this was $11.5 \mu/m^3$. For GPD models, examinations of mrl plots supported these threshold

choices.

Seasonality Model

Background research unveiled no obvious covariates for $PM_{2.5}$, though literature supporting a seasonal trend for $PM_{2.5}$ was found. $PM_{2.5}$ is generally expected to be worse in the wintertime because temperature inversions, in which there is a warm layer of air between cool layers of air, prevent air circulation and trap particulate matter near the earth's surface. Moreover, increased burning of fossil fuels for warmth during wintertime contributes to $PM_{2.5}$ production (Oceanic and Atmospheric Administration 2022) Therefore, seasonality was implemented using harmonic regression with a harmonic period of 1 year, with respect to both threshold selection and as a predictor for σ .

Quantile regression was employed to select the threshold, denoted by u , for Hennepin and St. Louis counties as follows:

$$u_d = Q^{(p)}[PM_{2.5} \text{ levels}] = \beta_{0,p} + \beta_{1,p}(\cos(2\pi \frac{d}{365.25})) + \beta_{2,p}(\sin(2\pi \frac{d}{365.25})),$$

where d = day of the year.

A declustering algorithm based on these seasonal thresholds with a run length of 5 yielded yielded 239 extreme events in Hennepin County and 155 extreme events in St. Louis County to inform further analysis (Figure 7).

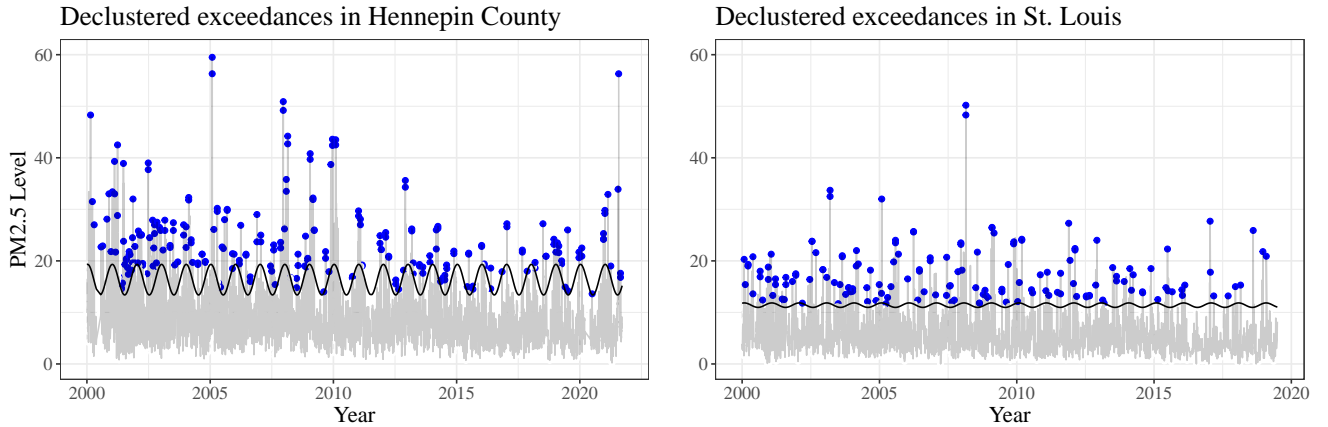


Figure 7: Declustered extreme $PM_{2.5}$ events using seasonal threshold

The following model was then assumed for the distribution of these excesses:

$$GPD(\sigma = \sigma_0 + \sigma_1(\cos(2\pi \frac{d}{365.25})) + \sigma_2(\sin(2\pi \frac{d}{365.25})), \xi),$$

Both the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) supported the inclusion of a seasonality component over naive baseline modeling for both counties.

The following parameter estimates were obtained for the models:

Hennepin County PM2.5 Model | St. Louis PM2.5 County Model

	Estimate	SE		Estimate	SE
σ_0	7.701	0.686	σ_0	7.213	0.787
σ_1	0.197	0.668	σ_1	0.821	0.734
σ_2	1.112	0.941	σ_2	1.883	0.964
ξ	0.00542	0.0599	ξ	-0.154	0.0671

Return level plots generated with 2-year, 10-year, and 50-year return levels and 95% confidence intervals calculated by the delta method confirm the expected seasonal trend—with peaks in PM_{2.5} extremes in the wintertime and dips in the summertime (Figure 8). This seasonal trend is also visible in the threshold levels, though it is amplified in the projected return levels for extreme events. Moreover, the model for Hennepin County estimates the “summer of smoke” event, which peaked on July 30, 2021 at 61.3 $\mu\text{g}/\text{m}^3$, as a once-in-50-years event. After considering seasonality, this event was revealed to be more unusual than previously thought because it occurred during the summertime, when PM_{2.5} is expected to be at its lowest.

Moreover, a comparison of return level estimates for the two counties over the course of a year show that while Hennepin County experienced higher extreme levels of PM_{2.5} when compared the St. Louis County, the seasonal trend is remarkably consistent between the two regions (Figure 9). This visualization also clearly displays an intensified seasonal effect for higher return levels; the difference between the expected winter and summer return levels is greater for the 50-year return period than the 2-year return period.

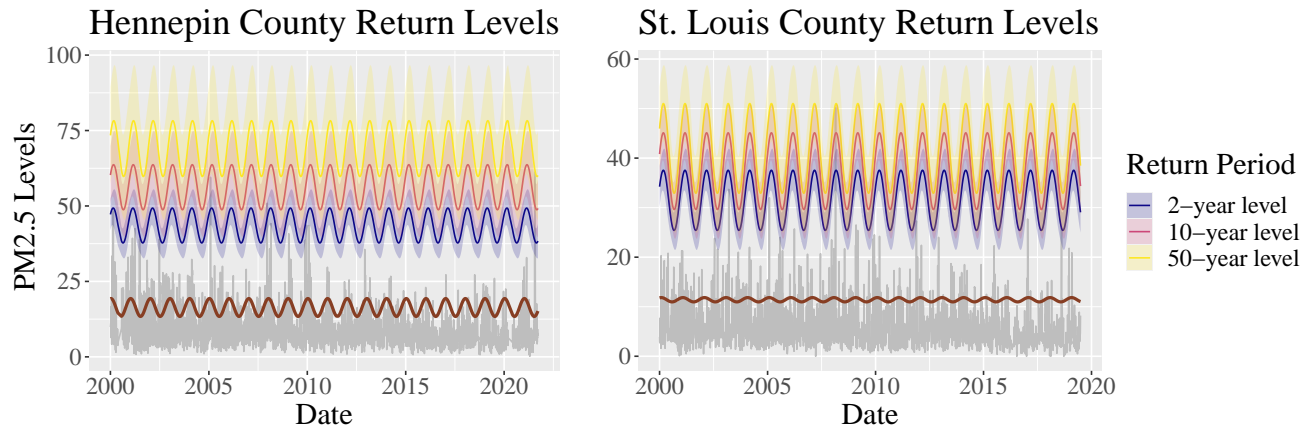


Figure 8: Return level plots for PM_{2.5} with 95 percent confidence intervals

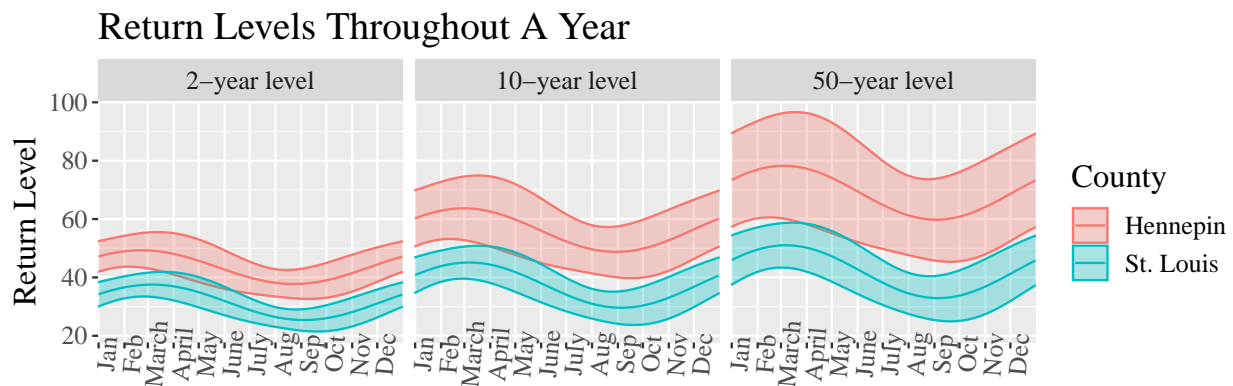


Figure 9: Return levels of PM_{2.5} with 95 percent confidence intervals over the course of a year

Density plots of the estimated GP distributions may be used to assess the likelihood of extreme PM_{2.5} values reaching certain values. Two specific values of interest for PM_{2.5} are the lower thresholds for the “Unhealthy for Sensitive Groups” ($35 \mu\text{g}/\text{m}^3$) and “Unhealthy for All” ($55 \mu\text{g}/\text{m}^3$) zones, as defined by NAAQS. Figure 10 isolates March 4 as a representative date for the winter peak of PM_{2.5} and August 3 as the counterpart for the summer dip.

PM_{2.5} Model Implications

No clear increasing or decreasing trend was found in extreme events of PM_{2.5} in the past two decades. However, a strong seasonal trend was confirmed, with peaks in the wintertime and dips in the summertime. This seasonal trend appears to be amplified in extreme values. This trend was consistent across Hennepin and St. Louis counties. Additionally, there are consistently higher particulate matter levels in Hennepin County compared to St. Louis. It is possible that this is related to differing levels of urbanization between

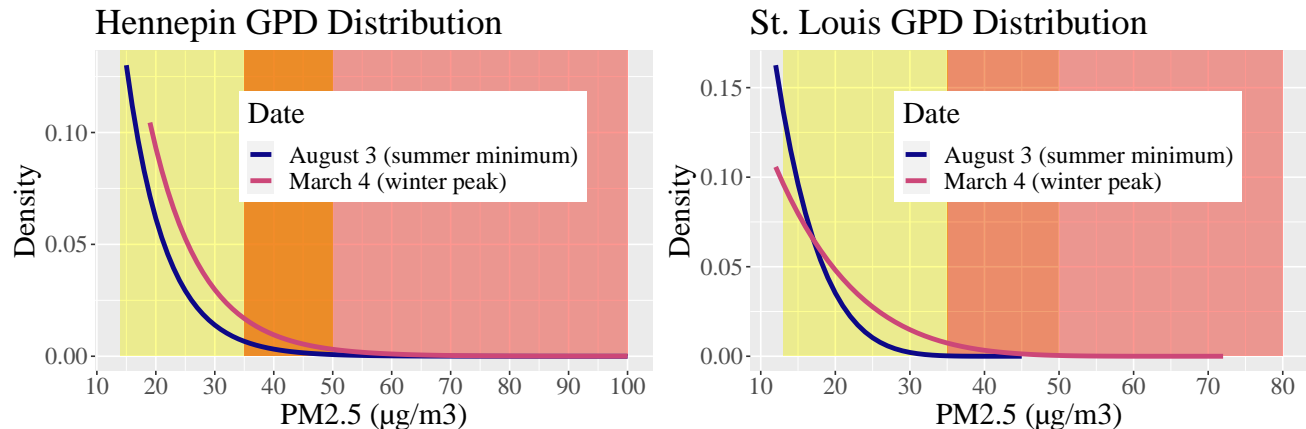


Figure 10: Density plots of estimated GP distributions of extreme PM_{2.5} levels in Hennepin and St. Louis counties

the two regions, but an in-depth examination of PM_{2.5} sources is beyond the scope of this report.

Consideration of seasonality characterizes the “summer of smoke” as more extreme than previously thought. While GEV and GPD modeling approaches estimated that the past summer’s pollution levels were between once-in-16-to-30-year events (depending on whether 2021 data was included in the modeling), the seasonal model considers it a once-in-50-year event.

5 Discussion

The limitations of our data restrain us from making more generalized conclusions to more regions. The types of pollutants and the number of monitoring sites we chose were based on the length of available records the EPA provides.

In terms of PM_{2.5}, we did not find a covariate that is significant enough to be associated with its trend. But it is reasonable to explore more factors such as wind, temperature, and population density that could help explain its variations. For SO₂, although we took the policy phase as a covariate, it is still a subjective choice based on nationwide and statewide policy history. When used on a specific county-level site, it is not perfect and could be further improved.

Despite that, the extreme value analysis shown here have provided valuable insights. The results from our EVA analysis greatly depend on the pollutant that we model. In the SO₂ context, a pollutant that is fairly easy to regulate, there is a clear decreasing trend in the record. Particularly when modeled with policy phase, there is a strong association between the two, implying that stricter regulation will result in lower SO₂ values. Further, in the present day, health concerns from SO₂ are essentially non-existent. Interestingly, the low baseline level of SO₂ in the 2020s results in extreme events that appear to be far more unlikely than

events of similar magnitude 30 years ago. Finally, the event that prompted this study, Minnesota’s summer of smoke, was not observed in the SO_2 record.

The $\text{PM}_{2.5}$ record tells a significantly different story. There was no clear or consistent trend in the $\text{PM}_{2.5}$ annual records since 2000, probably due to the fact that the pollutant itself is nearly impossible to regulate. However, $\text{PM}_{2.5}$ does have a strong seasonal cycle that peaks in the winter and is at its lowest in the summer. When viewed with this seasonal trend, the summer of smoke event is categorized as far more extreme than might have been initially suspected.

We have also discovered that one measurement is insufficient for describing air pollution. The AQI, for example, may suggest that overall levels of air pollution are relatively stable in Hennepin County since it only records the value of the pollutant that was most extreme on a given day. This analysis would somewhat align with what we learned about $\text{PM}_{2.5}$, but fails to capture the decreasing trend in SO_2 over the last 40 years and its relationship with policy regulations. Therefore, as policies for regulating air pollution are considered and the quantification of its consequences for human health is attempted, we have found that it is critical to consider multiple measures of air pollution rather than just one additive index. In other words, it is worthwhile to model multiple individual pollutants in order to make more comprehensive assessments of air pollution in a given location.

To sum up, in light of the concerning air quality in Minnesota in the summer of 2021, the usage of EVA models in the project enables us to better understand the trend of extreme values of air pollution in the region and make assessments of how extreme that a particular SO_2 or $\text{PM}_{2.5}$ pollution event is. The findings are valuable for educating the public and pave the road for more in-depth future research.

6 Acknowledgments

We want to express our sincere gratitude to Professor Andy Poppick for offering important advice, guidance and revision suggestions throughout this project.

We would also like to show appreciation to Professor Deborah Gross, Professor of Chemistry, for providing us great resources and insights into the study of air pollution.

7 Appendix

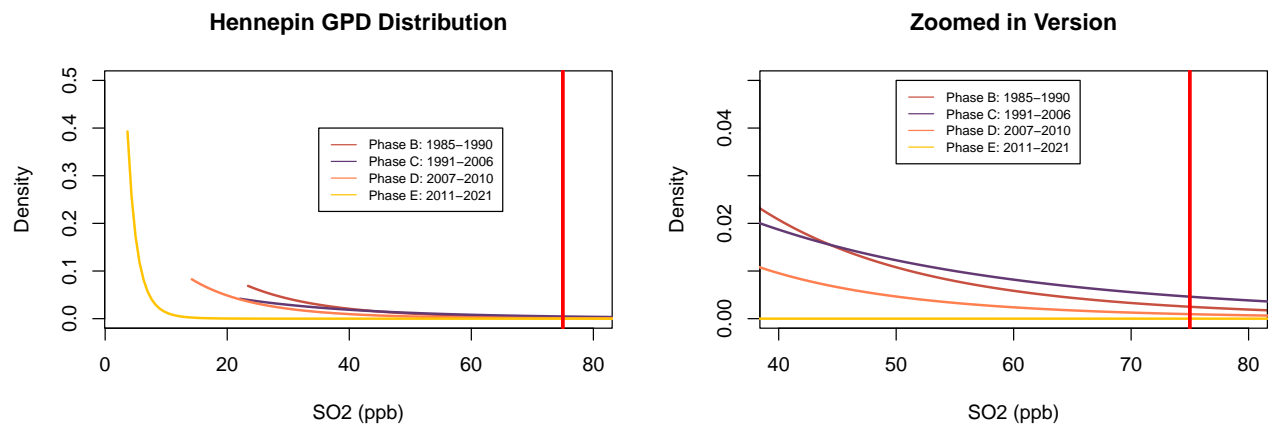


Figure 11: Hennepin SO2 GPD density plots, which show that the estimated probability of exceedances over 75 ppb is lower in recent policy phases.

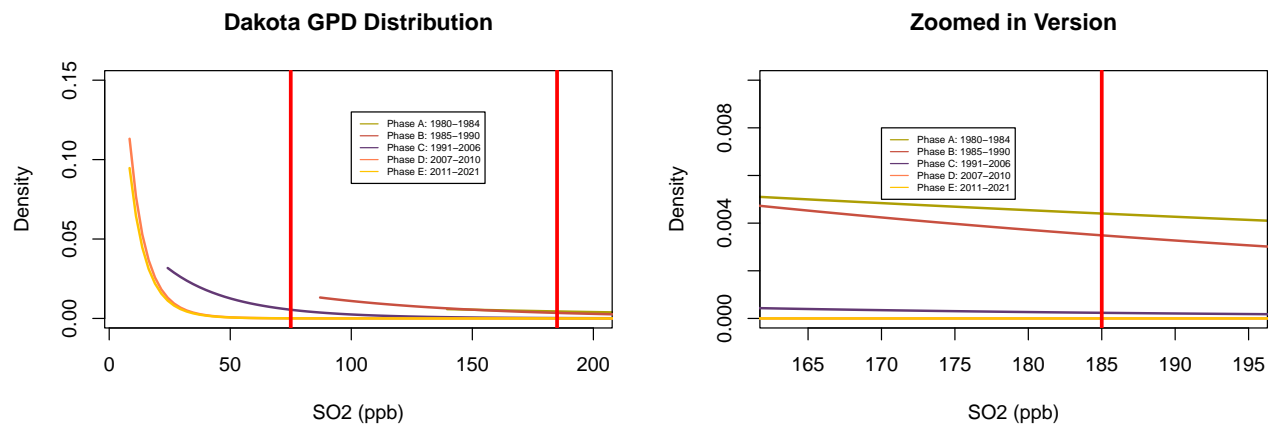


Figure 12: Dakota SO2 GPD density plots, which show that earlier policy phases had much higher probabilities of exceeding 185 ppb.

8 Reference

- Air Quality Life Index. 2022. “Policy Impact: United States: Clean Air Act (1970),” 1. %22<https://aqli.epic.uchicago.edu/policy-impacts/united-states-clean-air-act/>%22.
- California Air Resources Board. 2021. “Inhalable Particulate Matter and Health (Pm2.5 and Pm10),” 1. %22<https://www.dnr.state.mn.us/climate/journal/summer-smoke-2021.html>%22.
- Chen, Tze-Ming, Ware G Kuschner, Janaki Gokhale, and Scott Shofer. 2007. “Outdoor Air Pollution: Nitrogen Dioxide, Sulfur Dioxide, and Carbon Monoxide Health Effects.” *The American Journal of the Medical Sciences* 333 (4): 249–56.
- Coles, Stuart, Joanna Bawa, Lesley Trenner, and Pat Dorazio. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Vol. 208. Springer.
- Driesen, David M. 1996. “Five Lessons from the Clean Air Act Implementation.” *Pace Envtl. L. Rev.* 14: 51.
- EPA. n.d. “Legislative and Regulatory Timeline for Fossil Fuel Combustion Wastes.” <https://www.epa.gov/coalash/legislative-and-regulatory-timeline-fossil-fuel-combustion-wastes>.
- Ferrall, Brian L. 1991. “The Clean Air Act Amendments of 1990 and the Use of Market Forces to Control Sulfur Dioxide Emissions.” HeinOnline.
- Goldstein, Benjamin, Dimitrios Gounaridis, and Joshua P Newell. 2020. “The Carbon Footprint of Household Energy Use in the United States.” *Proceedings of the National Academy of Sciences* 117 (32): 19122–30.
- Kanchan, Kanchan, Amit Kumar Gorai, and Pramila Goyal. 2015. “A Review on Air Quality Indexing System.” *Asian Journal of Atmospheric Environment* 9 (2): 101–13.
- Latin, Howard. 1984. “Ideal Versus Real Regulatory Efficiency: Implementation of Uniform Standards and Fine-Tuning Regulatory Reforms.” *Stan. L. Rev.* 37: 1267.
- Liu, Hui, and Xinyu Zhang. 2021. “AQI Time Series Prediction Based on a Hybrid Data Decomposition and Echo State Networks.” *Environmental Science and Pollution Research*, 1–23.
- McCrowey, Clinton, Timothy Sharac, Nick Mangus, Doug Jager, Ryan Brown, Daniel Garver, Benjamin Wells, and Hayley Brittingham. 2021. *A r Interface to the US EPA Air Quality System Data Mart API*.
- Minnesota Department of Natural Resources. 2021. “Summer of Smoke, 2021,” 1. %22<https://www.dnr.state.mn.us/climate/journal/summer-smoke-2021.html>%22.
- . 2022. “The Drought of 2021,” 1. %22<https://www.dnr.state.mn.us/climate/journal/drought-2021.html#:~:text=Overview,drought%20in%2010%2D30%20years%22>.
- Minnesota Pollution Control Agency. 2022. “Fine Particle Pollution,” 1. %22<https://www.pca.state.mn.us/air/fine-particle-pollution>%22.
- New York State Department of Health. 2022. “Fine Particles (PM 2.5) Questions and Answers,” 1.

- %22https://www.health.ny.gov/environmental/indoors/air/pmq_a.htm#:~:text=Particles%20in%20the%20PM2.5%20size%20range%20are%20able%20to,nose%20and%20shortness%20of%20breath.%22.
- Oceanic, National, and Atmospheric Administration. 2022. “A New View of Wintertime Air Pollution,” 1. %22<https://research.noaa.gov/article/ArtMID/587/ArticleID/2450/A-New-View-of-Wintertime-Air-Pollution%22>.
- Ris, Charles. 2007. “US EPA Health Assessment for Diesel Engine Exhaust: A Review.” *Inhalation Toxicology* 19 (sup1): 229–39.
- Tucker, W Gene. 2000. “An Overview of Pm2. 5 Sources and Control Strategies.” *Fuel Processing Technology* 65: 379–92.
- Varapongpisan, Tunyathron, Lily Ingsrisawang, et al. 2019. “Generalized Extreme Value Model for Predicting the Occurrence of Air Pollutions.” PhD thesis, Kasetsart University.
- Wang, Deyun, Shuai Wei, Hongyuan Luo, Chenqiang Yue, and Olivier Grunder. 2017. “A Novel Hybrid Model for Air Quality Index Forecasting Based on Two-Phase Decomposition Technique and Modified Extreme Learning Machine.” *Science of The Total Environment* 580: 719–33.
- Yuzhi Cai, Dominic Hames. 2010. “Minimum Sample Size Determination for Generalized Extreme Value Distribution.” *Communications in Statistics - Simulation and Computation* 40 (1): 87–98.