LONG-TERM TRENDS
IN EXTREME ENVIRONMENTAL EVENTS
WITH CHANGEPOINT DETECTION

by

Mintaek Lee

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Computing

Boise State University

August 2022

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Mintaek Lee

Dissertation Title: Long-Term Trends in Extreme Environmental Events with Changepoint Detection

Date of Final Oral Examination: 10 June 2022

The following individuals read and discussed the dissertation submitted by student Mintaek Lee, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Jaechoul Lee, Ph.D.	Chair, Supervisory Committee
Jodi L. Mead, Ph.D.	Member, Supervisory Committee
Edoardo Serra, Ph.D.	Member, Supervisory Committee

The final reading approval of the dissertation was granted by Jaechoul Lee, Ph.D., Chair of the Supervisory Committee. The dissertation was approved by the Graduate College.

나의 할머니께 이 논문을 헌정합니다

ACKNOWLEDGMENTS

My sincere appreciation extends to my family. Especially, I extend my deepest gratitude to my parents, Yong Son Lee and Kwanseon Yoo, who have always encouraged me to be the better version of myself and sacrificed a great deal in supporting me, my sister, Suhyun Lee, for her constant support and encouragement, and my late grandfather, James Bowen, for giving me one more reason to return to Boise. I also recognize our four-legged companions, Nabi and Goguma, for giving me joy in my life and another reason to return home everyday.

My special appreciation now goes to my dear friends Dr. John Clemens and Jaimos Skriletz for being there for me when I needed someone to talk to during my most stressful times. Our weekly excursions to downtown for lunch were certainly a highlight of each week, despite our inability to decide on the location.

A very special acknowledgment goes to my wife, Mary Lee, who I will forever be grateful to. She has been extremely supportive of me throughout my studies and has made countless sacrifices for me. She did not sign up for graduate studies and all the struggles that come with it, but she still chose to endure them with me nonetheless. This dissertation is as much hers as it is mine.

Last but not least, my sincere gratutide extends to everyone who believed in me, even when I wouldn't believe in myself.

ABSTRACT

This dissertation examines long-term trends in extreme environmental events with considerations for changepoints and autocorrelation. Due to changes in measurement location, observer, instrument, sampling protocol, local ecosystem, etc., many environmental time series often contain inhomogeneous changes in their distributions. If ignored in the modeling process, these inhomogeneities could produce misleading estimation of the long-term trends in these environmental extremes. Because documentations for these changepoint-inducing events could be incomplete or missing in many cases, those changepoints need to be estimated from the data. Here, we use a genetic algorithm to estimate the number and times of changepoints in the environmental extremes as a data homogenization procedure before estimating their long-term trends and return levels. We illustrate our methods using two different extreme environmental series: monthly maximum coastal sea levels and weekly maximum ozone concentrations.

Increase in extreme sea levels can bring disastrous outcomes to people living in coastal regions by increasing flood risk or inducing stronger storm surges. With a substantial portion of the global community living in low elevation regions, it is crucial to understand how extreme sea levels have been changing over time. Therefore, we first study long-term trends in monthly maximum sea levels from coastal regions around the world. As strong periodicity and autocorrelation are pertinent to the sea level data, bootstrap techniques are used to obtain more realistic confidence intervals to the estimated trends and return levels. We find that the consideration of changepoints

changed the estimated long-term trends of 89 tide gauges (approximately 30% of tide gauges considered) by more than 20 cm century^{-1}.

Next, we examine another, but equally important environmental extreme event: extreme ozone concentrations. Specifically, we study long-term trends in weekly maximum ozone concentrations from the contiguous United States. Because exposure to an unhealthy level of ozone (even for a short period of time) can adversely affect one's health, understanding how extreme ozone events have changed over time can be of public health interest. Whereas monthly maximum sea levels in many locations exhibit weak autocorrelation, the weekly maximum ozone concentrations in many locations show non-ignorable autocorrelation even at two distant time points. For more accurate estimation of changepoints in the presence of long-memory autocorrelation, we develop a genetic algorithm based changepoint detection method for extreme value series with long-memory autocorrelation. This method is subsequently applied to detect changepoints in extreme ozone series. We find that the consideration of changepoints changed the long-term trend estimates in 78 counties (approximately 20% of the counties considered) by more than 0.03 ppm century^{-1}. Lastly, we find that almost all counties considered in the study are projected to experience unhealthy levels of ozone concentrations exceeding the EPA threshold at least once within 10 years.

Our results for these two extreme environmental events are summarized in maps with estimated long-term linear trends and return levels.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF MAPS

CHAPTER 1

INTRODUCTION

1.1   Motivation

Extreme environmental events like blizzards, heat waves, hurricanes, ozone episodes, storm surges, wildfires, etc., often profoundly impact human lives, economies, and ecosystems. As a result of climate change, extreme weather and climate events are projected to become more frequent and intense, thereby increasing its impact on many areas of life (USGCRP, 2018). Quantifying long-term trends in the environmental extremes can thus serve as a useful indicator to track and better understand the Earth's changing climate. However, many environmental time series contain inhomogeneous changes in their distributions induced by various undocumented outside factors such as observer changes, instrument changes, measurement location changes, or regime shifts in local ecosystems. Ignoring those changepoints can result in misleading long-term trend estimates for the given data.

To illustrate the possible issues with trend estimation in the presence of changepoints, we explore the extreme sea level time series data from Miyakejima, Japan, shown at Figure 1.1. Upon visual inspection, there appears to be a potential changepoint with an upwards mean shift to the Miyakejima series around September 2000, denoted by the purple dashed line. We first estimate the trend line using the ordinary least squares regression ignoring this potential changepoint, which estimates an

increasing trend of 0.78 m century^{-1}. We estimate the trend line again, but with a mean shift at September 2000 allowed. The least squares estimates a decreasing trend of -0.57 m century^{-1} with a mean shift of 0.59 m at September 2000. The estimated trend lines with and without changepoints are also imposed by red and blue lines respectively. The estimated trend line with a changepoint appears to better explain the overall trends of the Miyakejima series, suggesting that the extreme sea levels of Miyakejima likely have decreased over time, rather than gradually increasing over time. This example illustrates the possible consequences of ignoring changepoints in trend estimation. We discuss this issue in detail using extreme value methods in the following chapters.
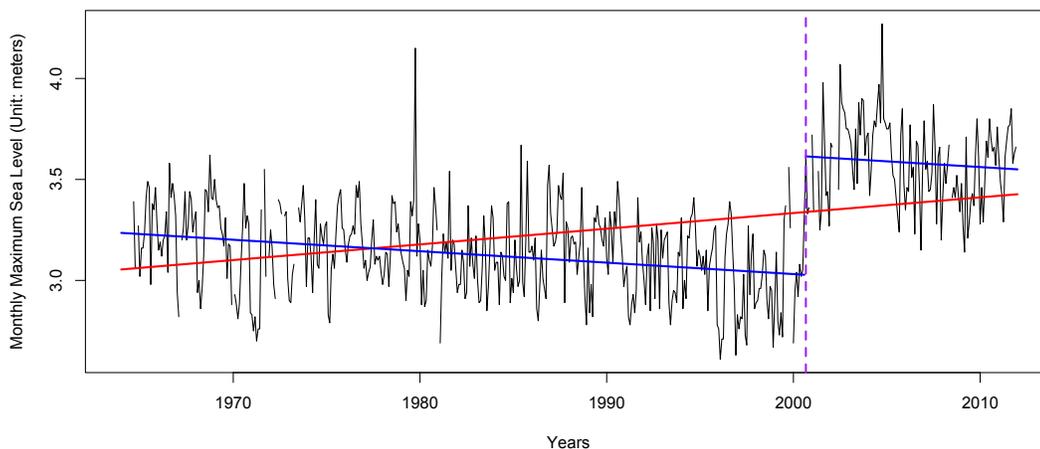
Figure 1.1: Least squares estimated trend lines for the extreme sea levels in Miyakejima, Japan (blue solid lines, trends with a September 2000 changepoint; red solid line, trends without allowing changepoints)

1.2 Overview

In this dissertation, we aim to rigorously quantify long-term linear trends in two different environmental extremes: monthly maximum coastal sea levels and weekly maximum ozone concentrations. These extremes have the potential to tremendously

impact both human lives and ecosystems. Extreme sea level events can dislodge people living in low-elevation coastal regions by inducing more intense floods or storm surges and possibly rendering those areas uninhabitable. Extreme ozone events can exacerbate asthma and other respiratory symptoms in vulnerable populations and adversely affect sensitive vegetation and ecosystems. Therefore, it is important to understand how these extremes have changed over time. To do so, we use a genetic algorithm to estimate the number and times of changepoints in these extremes. These estimated changepoints are then used as a part of the data homogenization procedure before estimating their long-term linear trends using extreme value methods.

In addition to changepoints, autocorrelation also forms an important aspect in analysis of environmental extremes. Most extreme value analysis techniques are developed under the assumption that data are independent. With many environmental time series exhibiting some levels of autocorrelation, this independence assumption is often not realistic. Naïve use of these techniques in the presence of autocorrelation could result in substantial estimation bias in model parameters and underestimation of standard errors for the estimates. There are many different approaches in dealing with autocorrelated data in extreme value studies and the choice of method is dependent on the characteristics of data.

The first approach we consider to address this autocorrelation issue is incorporating the extremal index in the extreme value model. The extremal index has been considered by many past authors (cf. Acero, García, and Gallego, 2011; Li et al., 2016; Rueda et al., 2016) to analyze extreme series with mild autocorrelation. This approach requires the raw data set where extreme observations are drawn from to be nearly independent if they are far enough away from each other in time (cf. Leadbetter, Lindgren, and Rootzen, 1983, pp. 52–54), which in turn implies that

the autocorrelation in extreme value series need to be weak. The implementation of this approach is fairly straightforward and are generally computationally inexpensive. We use this extremal index approach to address autocorrelation issues in monthly maximum coastal sea level analysis in Chapter 2.

However, some extreme value series exhibit non-ignorable autocorrelation, rendering the extremal value approach insufficient in correcting estimation bias due to autocorrelation. In this case, we can consider directly modeling the autocorrelation structure in the extreme value series via copula transformation. The copula transformation has been considered by some past authors (cf. Ribatet and Sedki, 2013; Zhu, Liu, and Lund, 2019) to analyze extreme series with non-ignorable autocorrelation between short time points. However, if the autocorrelation in extreme value series at two distant time points cannot be ignored, thus exhibiting long-memory autocorrelation, implementation of this copula transformation approach poses some challenges and is computationally expensive. In the weekly maximum ozone concentration analysis in Chapter 3, where it was observed that the extreme ozone series exhibit long-term autocorrelation, we use this copula transformation approach with considerations for long-memory autocorrelation.

This remainder of this dissertation proceeds as follows: Chapter 2 summarizes our analysis of monthly maximum coastal sea levels. Chapter 3 details our weekly maximum ozone concentration analysis. In Chapter 4, we conclude with a summary discussion and avenues for future research.

CHAPTER 2

MONTHLY MAXIMUM COASTAL SEA LEVELS[1]

2.1 Introduction

With over 10% of the world's population living in low elevation coastal areas (Mc-Granahan, Balk, and Anderson, 2007), the increase in sea levels poses a great threat to our society. In particular, the increase in extreme coastal sea levels could bring devastating outcomes by inducing more intense floods or storm surges. Furthermore, unlike most extreme weather events that bring devastating, but at least partially recoverable damage, extreme sea level rise could critically impact coastal regions by making low elevation regions permanently uninhabitable.

Many past authors agree that the global mean sea levels have been increasing over time. Jevrejeva et al. (2008) reconstructed mean global sea level from 1700 using the tide gauge data and found that global mean sea level has increased by 6 cm during the 19th century and another 19 cm in the 20th century. Church and White (2011) found that the mean global sea level has increased by about 21 cm during 1880–2009 based on sea level records from both tide gauges and satellite altimeter. From the satellite altimeter data, which first became available in 1993, many authors also comment that

---

[1]This chapter is a reprint of an article in the Journal of the Royal Statistical Society: Series C (Applied Statistics). The original citation of the article is as follows:
Lee, M. and Lee, J. (2021) Long-term trend analysis of extreme coastal sea levels with changepoint detection. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 70(2):434–458
© 2021 Royal Statistical Society

the sea level has been gradually rising over time (Watson et al., 2015; Chen et al., 2017; Nerem et al., 2018). Levermann et al. (2013) projected that the global mean sea level will continue to rise for centuries based on their ice-sheet and climate models.

To quantify long-term trends in extreme sea level data, extreme value methods should be used, since mean and extreme statistics are statistically independent for a large sample under some minor regularity conditions (McCormick and Qi, 2000). Therefore, one should not necessarily assume that extreme sea levels would exhibit the same features as mean sea levels. Recent authors applied extreme value analysis techniques to study extreme coastal sea levels and quantify their long-term trends in a regional scale. Marcos and Woodworth (2017) used tide gauge records from the North Atlantic Ocean and the Gulf of Mexico to study relationships between mean and extreme sea levels. They found overall increasing trends of annual 99th percentile of total sea level for most locations, except the Baltic Sea which showed decreasing trends. Wahl and Chambers (2015) analyzed extreme sea levels of the contiguous United States coastlines and found significant increasing long-term trends in 99.5th percentile of observed sea levels between 1929 and 2013 at almost all locations. Wang and Zhou (2017) applied the peaks-over-threshold method to sea level observations from five tide gauges in Macao and Hong Kong, concluding that there are no significant linear trends in extreme sea level in the Pearl River Estuary, China.

On a global scale, only a few researchers have examined extreme coastal sea levels. Menéndez and Woodworth (2010) used time-dependent generalized extreme value (GEV) distribution to estimate decadal variation and long-term variations in monthly maximum sea levels. Merrifield et al. (2013) estimated contributing factors to annual maximum sea levels above the annual mean from 145 tide gauges and decomposed

them into multiple components. Marcos et al. (2015) used a state-space approach to explicate decadal to multidecadal variations in sea level extremes from 122 tide gauges. Wahl et al. (2017) used 20 different extreme value models on 510 tide gauges to investigate inter-model uncertainties in extreme sea level rise projections.

Two critical issues tackled in this study are changepoints and temporal correlation. First, sea level observations naturally contain inhomogeneous changes in their means for various reasons, such as changes in instruments, relocation of tide gauges, earthquakes, land reclamation, dredging, etc. (cf. Becker et al., 2009). If these mean changes are ignored, the long-term trend for the data can be erroneously estimated (cf. Lund and Reeves, 2002). Second, sea levels also exhibit strong temporal correlation by nature. Since most extreme value methods are developed under the assumption that data are independent, a naïve use of extreme value methods based on independence, including decorrelation techniques, could result in substantial estimation bias.

In this paper, we study long-term linear trends in monthly maximum coastal sea levels by applying extreme value methods with changepoint and temporal correlation considerations. We use a genetic algorithm to estimate the number and locations of mean shifts induced by changepoints to rigorously quantify the long-term trends in the extreme coastal sea levels. Also, as sea level data exhibit strong temporal correlation, we incorporate the extremal index parameter to correct estimation bias in return levels due to temporal correlation. Bootstrap techniques are used to construct more realistic confidence intervals for the estimated linear trends and return levels. As demand increases among practitioners in various disciplines who want to apply extreme value methods and changepoint detection techniques, we thoroughly illustrate our methodology and plan to aid practitioners by making our R programming codes available online.

This paper proceeds as follows: Section 2.2 describes the data set used in our study. Section 2.3 discusses the extreme value methods and changepoint detection techniques applied to the sea level data. Section 2.4 describes the simulation study on changepoint detection and summarizes the performance of the changepoint method. Section 2.5 illustrates our methods using a case study for the sea level series from Fishguard, UK. Section 2.6 summarizes our trend and return level analysis for the entire sea level data set. In Section 2.7, we conclude with further remarks.

2.2   The Data

This study uses the Global Extreme Sea Level Analysis Version 2 (GESLA-2) data set, published by Woodworth et al. (2016). The GESLA data set provides a semi-global coverage on high frequency (mostly hourly) sea level observations from 1276 tide gauge locations around the world, compiled from 27 public sources, including the University of Hawaii Sea Level Center and British Oceanographic Data Centre (BODC) that collectively account for about 70% of the tide gauge records. Sea level observations made by a tide gauge represent the vertical height from the sea surface to a vertical datum. The GESLA sea level observations are recorded in meters and accompanied by a quality control value. We considered only those observations marked as either "correct value" or "interpolated value" and treated the ones marked as "doubtful value", "isolated spike or wrong value", or "missing value" as missing values. These missing values, accounting for 9.03% of all GESLA data, were discarded from our analysis.

Although the GESLA data set went through some quality control checks provided by the data providers, it still requires preprocessing for our analysis. Because the GESLA data set is a compilation of sea level data sets from multiple data providers,

there are often more than one sea level series pertaining to the same geographic location. Nearly all of these multiple records are identical to each other except for having slightly different temporal coverages. For these locations, only the series with the most non-missing observations was used in our analysis. In the case that these duplicate records offer some non-overlapping sea level observations, they were merged to attain the longest possible temporal coverage for that location. In addition, some sea level series in the data set contain temporal gaps in observations without documentation. We treated this record as missing values. For our extreme value analysis, if more than 7% of sea level observations from a single calendar month are missing, we did not extract the monthly maximum and treated that month as missing. We then considered all monthly maximum sea level series with at least 30 years of non-missing data (at least 360 non-missing monthly maxima).
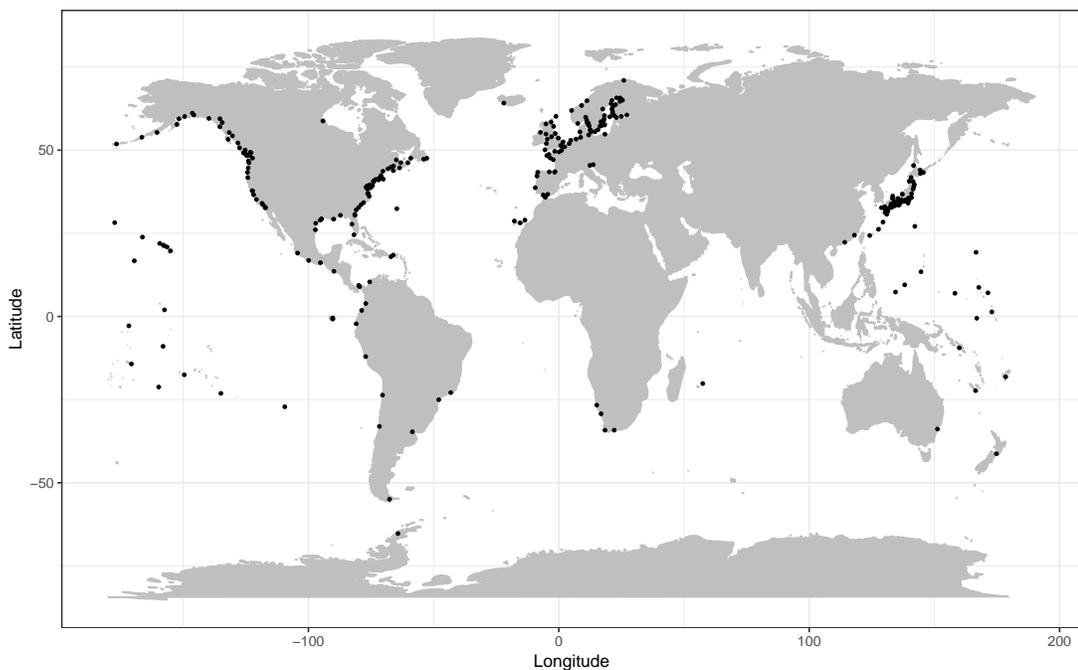
Map 2.1: Spatial location of the 300 tide gauges after preprocessing

Our preprocessing procedures resulted in 300 monthly maximum sea level series for our analysis. Map 2.1 shows the spatial location of the selected 300 tide gauges used in our analysis. About 82% of the 300 tide gauges are located in Europe, North America, and Northeast Asia, providing a sufficient coverage to the North Pacific Ocean and North Atlantic Ocean. About 15% of the tide gauges are sparsely located in Pacific Islands and South America. Most of the remaining tide gauges are located in Africa and Oceania.

Although the GESLA data set provides a near global coverage, some regions are underrepresented. Since we mainly consider locations with at least 30 years of non-missing monthly maximum sea level data, the limited availability of data impacts the scope of our studies. Specifically, except for a single tide gauge in Mauritius, no tide gauges from the coasts of East Africa, Middle East, South Asia, Southeast Asia, and Western Australia are considered, which essentially leaves out the entire coast of the Indian Ocean from our analysis. The western side of Africa and the northeastern portion of South America are also left out, limiting the coverage of the coasts of the South Atlantic Ocean in our study.

2.3   Methods

2.3.1   Block maxima methods with periodic, trend, and changepoint features

Suppose $X_1, \ldots, X_m$ are independent and identically distributed (IID) random variables with a common distribution function $F(\cdot)$. Define $Y_{(m)} = \max\{X_1, \ldots, X_m\}$ as the maximum statistic of these m random variables. If there are sequences of constants $\{a_m\}$ and $\{b_m > 0\}$ that scale the maximum statistic $Y_{(m)}$ such that

$$\lim_{m \to \infty} P\left(\frac{Y_{(m)} - a_m}{b_m} \leq y\right) = G(y)$$

and $G(\cdot)$ is a non-degenerate distribution function, then the limiting distribution function $G(\cdot)$ is the following generalized extreme value (GEV) distribution:

$$G(y) = \exp\left\{-\left[1 + \xi\left(\frac{y - \mu}{\sigma}\right)\right]_+^{-1/\xi}\right\},$$

where $[x]_+ = \max\{x, 0\}$. The parameters $\mu \in (-\infty, \infty)$, $\sigma \in (0, \infty)$, and $\xi \in (-\infty, \infty)$ are called location, scale, and shape parameters, respectively.

Block maxima methods use a sequence of these maximum statistics. To elaborate, we reexpress the raw GESLA hourly sea level series $\{X_1, \ldots, X_N\}$ at a gauge site as a set of n blocks with block size m:

$$\{(X_1, \ldots, X_m), (X_{m+1}, \ldots, X_{2m}), \ldots, (X_{(n-1)m+1}, \ldots, X_{nm})\},$$

where $N = nm$. We then compute the maximum statistic from each block and denote the t-th block maximum statistic as $Y_t = \max\{X_{(t-1)m+1}, \ldots, X_{tm}\}$ for $t = 1, \ldots, n$. The extreme value theorem states that if block size m is sufficiently large, the GEV distribution is a proper probability distribution for $\{Y_1, \ldots, Y_n\}$ regardless of the distribution function $F(\cdot)$ from which the raw data $\{X_1, \ldots, X_N\}$ is taken. Therefore, we assume that the monthly maximum sea level series $\{Y_t\}$ follows a GEV distribution.

Important features pertinent to monthly maximum sea level series should be considered for the GEV parameters. First, monthly maximum sea levels naturally exhibit strong periodicity due to tides caused by gravitational attraction from the sun and the moon and other external factors. This periodicity needs to be taken

into account in GEV models. Second, if monthly maximum sea levels are changing, the rate of change should be considered. The model without trend, when the series in fact has a trend, will result in erroneous estimation of the model parameters. Third, monthly maximum sea level series features changepoints for many reasons, including instrument changes, location changes, and changes in times at which the measurements are made. Such changepoints can result in misleading outcomes if not addressed properly.

In this study, the location parameter for the monthly maximum sea level series $\{Y_t\}$ is posited to include a periodic function, which we use a cosine wave with two harmonics, a long-term linear trend, and an unknown number of changepoints, each including a mean shift in $\{Y_t\}$ as follows:

$$\mu_t = \beta_0 + \sum_{j=1}^{2} \left\{ \beta_{2j-1} \cos\left(\frac{2j\pi t}{T}\right) + \beta_{2j} \sin\left(\frac{2j\pi t}{T}\right) \right\} + \alpha\left(\frac{t}{100T}\right) + \delta_t. \qquad (2.1)$$

Here, $T = 12$ is the period for the monthly maximum series, α is the long-term linear trend representing the expected change in maximum sea level over a century, and δ_t is the mean shift term, parameterizing the magnitude of mean shifts due to c changepoints at times $\tau_1, \ldots, \tau_c$ as:

$$\delta_t = \Delta_0 I(1 \leq t < \tau_1) + \Delta_1 I(\tau_1 \leq t < \tau_2) + \cdots + \Delta_c I(\tau_c \leq t \leq n),$$

where $I(E)$ is an indicator function returning 1 if E is true and 0 otherwise. The baseline mean shift term Δ_0 is set to be zero for parameter identifiability. The scale parameter is also parameterized as a cosine wave with two harmonics:

$$\sigma_t = \omega_0 + \sum_{j=1}^{2} \left\{ \omega_{2j-1} \cos\left(\frac{2j\pi t}{T}\right) + \omega_{2j} \sin\left(\frac{2j\pi t}{T}\right) \right\}. \qquad (2.2)$$

The seasonal dependence in extreme sea levels is mainly due to astronomical (spring tides) and meteorological (storminess season) influences (Menéndez and Woodworth, 2010). To capture the possibility of annual and semi-annual periodic features in monthly maximum sea levels, we parameterize the GEV location and scale parameters using two sinusoidal harmonics as shown in (2.1) and (2.2). This parameterization is previously used by other authors, including Menéndez and Woodworth (2010) and Weisse et al. (2014). This said, the number of harmonics in the location and scale parameter expressions can be varied, but we found that two work well with most GESLA monthly maximum sea levels. In addition, we assume that the GEV shape parameter ξ is constant over time, because the shape parameter estimate can be numerically unstable (Smith, 2014) and likely cause undesirable convergence issues when a complicated model is used for ξ, especially over short segments. This constant ξ assumption is often physically and numerically supported by other researchers (Zhang, Zwiers, and Li, 2004; Parey et al., 2007; Hoang, Parey, and Dacunha-Castelle, 2009).

For parameter estimation, if the changepoint number c is known and these c changepoints are also known to occur at the times $\tau_1, \ldots, \tau_c$, the unknown GEV model parameters are denoted by $\boldsymbol{\eta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \alpha, \Delta_1, \ldots, \Delta_c, \omega_0, \omega_1, \omega_2, \omega_3, \omega_4, \xi)^{\mathrm{T}}$. The log-likelihood function of the $\mathrm{GEV}(\mu_t, \sigma_t, \xi)$ distribution with μ_t in (2.1) and σ_t in (2.2) can be written as

$$
\ell(\boldsymbol{\eta}) =
\begin{cases}
-\sum_{t=1}^{n} \ln \sigma_t - \sum_{t=1}^{n} \left[1 + \xi \left(\dfrac{y_t - \mu_t}{\sigma_t} \right) \right]^{-1/\xi} \\
\quad - \left(1 + \frac{1}{\xi} \right) \sum_{t=1}^{n} \ln \left[1 + \xi \left(\dfrac{y_t - \mu_t}{\sigma_t} \right) \right], & \text{if } \xi \neq 0; \\
-\sum_{t=1}^{n} \ln \sigma_t - \sum_{t=1}^{n} \left(\dfrac{y_t - \mu_t}{\sigma_t} \right) - \sum_{t=1}^{n} \exp \left[-\left(\dfrac{y_t - \mu_t}{\sigma_t} \right) \right], & \text{if } \xi = 0.
\end{cases}
$$

$$(2.3)$$

We then use a numerical optimizer to find the maximum likelihood estimates of GEV parameters $\boldsymbol{\eta}$ by maximizing this log-likelihood function.

In practice, however, the number of changepoints c and changepoint times $\tau_1, \ldots, \tau_c$ are all unknown and need to be estimated. We explain the estimation method for these unknown changepoint parameters in the following subsection.

2.3.2 Changepoint detection using a genetic algorithm

As the GESLA data set lacks metadata other than basic geographical information, documented changepoint information is greatly limited. For this reason, we use a genetic algorithm (GA) to detect any significant mean shifts due to changepoints present in the monthly maximum sea levels. Our GA method is based on the approaches in Li and Lund (2012) and Lee, Li, and Lund (2014) with modifications specific to monthly maximum sea levels. We implement the GA as follows.

1. An initial generation of $L = 200$ changepoint configurations (called "chromosomes") is randomly generated. Each chromosome is expressed as $(c; \tau_1, \ldots, \tau_c)$, where c is the number of changepoints, and τ_j is the time (month) at which the j-th changepoint occurs.

2. The 200 chromosomes are probabilistically crossed as follows. Each chromosome is ranked based on its fitness value, where the fittest chromosome is assigned the highest rank L. One mother and one father are then selected from the 200 chromosomes. The i-th chromosome is selected as a mother with probability $R_i / \sum_{l=1}^{L} R_l$, where R_l refers to the l-th chromosome's rank. To select a father, ranks for the remaining 199 chromosomes are reassigned and the selection process is repeated. Once mother and father are chosen, a

child is probabilistically generated. To elaborate, suppose $(a; \kappa_1, \ldots, \kappa_a)$ and $(b; \zeta_1, \ldots, \zeta_b)$ are chosen as parents. These two chromosomes are first merged, resulting in $(a + b; \tau_1, \ldots, \tau_{a+b})$. Next, after eliminating duplicate times, each τ_i is removed from the child with probability 0.5. Each remaining τ_i then remains unchanged with probability 0.4, moves one month forward with probability 0.3, or moves one month backward with probability 0.3.

3. Every non-changepoint time location is assigned a probability of $p_{\text{mut}} = 0.002$ to be selected as an additional changepoint. Changepoint times chosen from this process are called mutations, which form an important aspect of the GA to avoid falling into local optima.

4. Once a generation of 200 chromosomes is generated, the crossing and mutation are repeated to obtain new generations until we reach the 300-th generation. The fittest chromosome from all of these 300 generations is then chosen as the estimated changepoint configuration.

Our GA method differs from those of Li and Lund (2012) and Lee, Li, and Lund (2014) in a way that we directly apply the GA to monthly maximum sea level series with the following modifications for better performance. From our preliminary simulation study, we found that the GA tends to overfit chromosomes when the target series does not have linear trends. To address this problem, we introduced the elitist selection. The two fittest chromosomes from the previous generation are kept without any alterations. In addition, these two "elite" chromosomes are crossed with each other and mutated with the same mutation probability as others, producing a new elite chromosome. These three elite chromosomes are then passed over to form the next generation along with 197 other new chromosomes.

We use the minimum description length (MDL) as the fitness function to estimate the changepoint number and times (cf. Lu, Lund, and Lee, 2010). For a chromosome of c changepoints at times $\tau_1, \ldots, \tau_c$, the MDL is calculated as:

$$\mathrm{MDL}(\boldsymbol{\eta}, c, \tau_1, \ldots, \tau_c) = -\ell_{\mathrm{opt}}(\boldsymbol{\eta}) + P(c; \tau_1, \ldots, \tau_c). \tag{2.4}$$

Here, $\ell_{\mathrm{opt}}(\boldsymbol{\eta})$ is the optimized value of the GEV log-likelihood function in (2.3) calculated at the maximum likelihood estimates of $\boldsymbol{\eta}$ for a given changepoint configuration $(c; \tau_1, \ldots, \tau_c)$. The penalty term $P(c; \tau_1, \ldots, \tau_c)$ is expressed as

$$P(c; \tau_1, \ldots, \tau_c) = \ln(c+1) + \frac{1}{2} \sum_{j=2}^{c+1} \ln(\tau_j - \tau_{j-1}) + \sum_{j=2}^{c+1} \ln \tau_j,$$

where $\tau_{c+1} = n + 1$. If missing observations exist in the j-th segment on times $\tau_{j-1}, \ldots, \tau_j - 1$, then $\tau_j - \tau_{j-1}$ in the penalty term is replaced with the number of non-missing observations in that segment. A changepoint configuration with a smaller MDL value is preferred. The performance of the GA method for changepoint estimation is assessed via simulation in Section 2.4.

We selected the GA technique because GAs are less restrictive to use in a long-term trend study than many other multiple changepoint detection methods. To be specific, the target MDL function in (2.4) cannot be reexpressed as $\sum_{j=1}^{c+1} C_j$, where C_j is a cost function associated with the j-th segment. Search algorithms that use this type as the target function for optimization, including WBS (Fryzlewicz, 2014) and PELT (Killick, Fearnhead, and Eckley, 2012), will not work, because the parameter estimate of the long-term trend α, for example, depends on all the data points, not only the data in any single segment. The model parameters α, β's, ω's, and ξ can be poorly

estimated from small segments. Next, the standard errors for the GEV estimates with GA are estimated using a moving block bootstrapping method as illustrated in Section 2.3.4.

2.3.3 Return levels with non-stationary and dependent series

Return levels are an important aspect in extreme value analysis. The return level associated with the return period z years is the expected value that is to be exceeded once every z years on average. Since temporal correlation in the data can seriously affect the accuracy of return levels (cf. Fawcett and Walshaw, 2012; Reich, Shaby, and Cooley, 2014), past authors often considered an additional parameter, called extremal index. The extreme value theorem described in Section 2.3.1 then can hold true for dependent series by incorporating the extremal index.

Suppose $X_1^*, \ldots, X_m^*$ are IID random variables with a common marginal distribution $F(\cdot)$ and $X_1, \ldots, X_m$ are correlated and stationary random variables with the same marginal distribution $F(\cdot)$. Define $Y_{(m)}^* = \max\{X_1^*, \ldots, X_m^*\}$ and $Y_{(m)} = \max\{X_1, \ldots, X_m\}$. Under the assumption that $X_1, \ldots, X_m$ satisfy the $D(u_m)$ condition (cf. Leadbetter, Lindgren, and Rootzen, 1983, pp. 52–54),

$$\lim_{m \to \infty} P\left(\frac{Y_{(m)}^* - a_m}{b_m} \leq y\right) = G(y)$$

for some sequences $\{a_m\}$ and $\{b_m\}$, if and only if

$$\lim_{m \to \infty} P\left(\frac{Y_{(m)} - a_m}{b_m} \leq y\right) = G^\theta(y)$$

for a constant $\theta \in (0, 1]$. Here, the parameter θ is the extremal index, and the limiting

distribution function $G^\theta(\cdot)$ has the following GEV distribution expression:

$$G^\theta(y) = \exp\left\{-\left[1 + \xi\left(\frac{y - \mu_\theta}{\sigma_\theta}\right)\right]_+^{-1/\xi}\right\},$$

where $\mu_\theta = \mu - \frac{\sigma}{\xi}(1 - \theta^\xi)$ and $\sigma_\theta = \sigma\theta^\xi$ (cf. Coles, 2001, pp. 92–97). When $\theta = 1$, the limiting distribution of $Y_{(m)}$ is the same as that of $Y^*_{(m)}$. In short, this result implicates that if two far enough sets of correlated random variables are nearly independent (so that the $D(u_m)$ condition is satisfied), then the GEV distribution still can be an approximate distribution with aid of θ.

The conventional definition of the return levels with stationary data assumes the probability of exceedance to be constant over time. Since this assumption is not satisfied under non-stationarity as is the case with GESLA monthly maximum sea levels, we instead use the method of Parey et al. (2007) and Parey, Hoang, and Dacunha-Castelle (2010). Specifically, we now estimate the level r_z for which the expected number of exceedances in z years ($12z$ months) is one. The z-year return level of non-stationary monthly maxima is then the solution to the following non-linear equation

$$1 = \sum_{t=t_I}^{t_I+12z-1} (1 - G_t^\theta(r_z)), \tag{2.5}$$

where $G_t^\theta(\cdot)$ is the time-dependent GEV distribution function in month t, and $t_I \geq n$ is a predetermined starting time for return level computation. We set t_I to be January 2020 for all tide gauges, interpreting r_z as the lowest monthly maximum sea level that is expected to be exceeded once in a z-year period starting from January 2020. For the stationary case, this definition is equivalent to the conventional definition of return levels (Parey, Hoang, and Dacunha-Castelle, 2010).

To accurately compute the return level estimate, we therefore need to estimate the extremal index θ. Although we do not have a strong preference on a particular estimation method, we use the semiparametric maxima estimator (Northrop, 2015). This method does not need parametric modeling for raw hourly sea level data, which would be a very challenging task for the raw hourly sea level data in this study. In addition, unlike many others, this estimator does not require an arbitrary selection of threshold but uses the relationship between the distribution of block maxima and the marginal distribution of the raw data. Northrop (2015) showed that this estimator is competitive compared to other existing estimators in a simulation study.

The Northrop's method proceeds as follows. Suppose $X_1, \ldots, X_N$ are strictly stationary random variables with a common distribution function $F(\cdot)$ and extremal index θ. Let Y_t be the maximum statistic for the t-th block (month) of size m for $t = 1, \ldots, n$. Define $V_t = -m \log F(Y_t)$, which follows an exponential distribution with mean $1/\theta$. Since $F(\cdot)$ is often unknown, it is empirically estimated:

$$
\hat{F}_{-t}(y) = \begin{cases} \frac{1}{N-m} \sum\limits_{X_s \in A_{-t}} I(X_s \leq y), & \text{if } y \geq \min A_{-t}; \\ \frac{1}{N}, & \text{if } y < \min A_{-t}, \end{cases}
$$

where $A_{-t} = \{X_1, \ldots, X_N\} \setminus \{X_{(t-1)m+1}, \ldots, X_{tm}\}$. Then, V_t can be estimated as $\hat{V}_t = -m \ln \hat{F}_{-t}(Y_t)$ for $t = 1, \ldots, n$. Then, Northrop (2015) derives $\hat{\theta} = n / \sum_{t=1}^{n} \hat{V}_t$.

To use this method, we need to convert the raw hourly sea level series to a stationary series. Our stationary conversion procedure is illustrated in Section 2.5.2. Once the extremal index θ is estimated, we use the estimate $\hat{\theta}$ to adjust the GEV distribution function for temporal correlation as done in (2.5). The z-year return level at a GESLA station is then estimated using its most recent changepoint with

the assumption that its estimated trend would persist into the future for the z-year return period. Since the closed form expression for r_z in (2.5) is not available for the non-stationary case, we use a grid search algorithm to numerically estimate r_z.

2.3.4 Bootstrap confidence intervals for trends and return levels

Now, we quantify the uncertainty in the parameter estimates. The standard errors of the parameter estimates would be calculated based on the observed information matrix, which is computed using software. The asymptotic $100(1 - \alpha)\%$ confidence intervals of the parameter estimates are typically computed as: (parameter estimate) $\pm z_{\alpha/2}$ (standard error), where $z_{\alpha/2}$ is the upper $\alpha/2$-th quantile from the standard normal distribution. However, if the data are correlated, the standard errors computed from the observed information matrix can be biased, implicating that the asymptotic confidence intervals would not produce the intended confidence level. To obtain more realistic standard errors and confidence intervals for the maximum likelihood estimates of GEV parameters, long-term linear trend α, and return levels, we use a bootstrap method.

Classical bootstrap methods assume that the data are IID, so the dependence structure of the population distribution cannot be fully preserved in the resampling process. Hence, IID bootstrap methods would fail to adequately approximate the true distributions of GEV parameter estimates if applied to dependent data. Out of those approaches in dealing with dependent data, we use a moving block bootstrap (Künsch, 1989) to obtain our bootstrap samples. The moving block bootstrap method proceeds as follows. From the monthly maximum sea level series $\{Y_1, \ldots, Y_n\}$ at a gauge site, $n - k + 1$ overlapping blocks of size k can be formed. Out of these blocks, n/k blocks

are randomly selected with replacement to generate a b-th bootstrap sample of size n. This process is repeated until B bootstrap samples are obtained.

Once B bootstrap samples are obtained, conventional percentile bootstrap methods use the upper and lower $\alpha/2$-th quantiles of the bootstrap estimates of GEV parameters to construct a $100(1 - \alpha)\%$ bootstrap confidence interval. However, if the distribution of bootstrap estimates is skewed, percentile bootstrap methods often fail to reach the desired coverage probabilities. Our preliminary analysis suggests that the shape parameter for the GEV model of monthly maximum sea level series is negative for most locations we considered. If the shape parameter is negative, the GEV distribution is left skewed, indicating that the distribution of a return level is also left skewed. To correct the bias due to skewness, we use the bias-corrected and accelerated (BCa) bootstrap method (Efron, 1987) to compute the confidence interval from bootstrap samples.

2.4 A Simulation Study

A simulation study was performed to evaluate the effectiveness of the GA changepoint method described in Section 2.3.2. We focus on the following two questions: how well does the method accurately estimate the number of changepoints and how well does the method correctly detect the changepoint times.

We consider the eight scenarios as summarized in Table 2.1. The first two consider a no-changepoint model without linear trends (Scenario 1) and with linear trends (Scenario 2), estimating the false positive rates of the GA when there are in fact no changepoints. For those series with no changepoints, a low false detection rate is desired. Scenarios 3 and 4 have one changepoint in month $\tau_1 = 200$ to estimate the

true positive rates for one changepoint. Scenarios 5 and 6 assume two changepoints each under a different linear trend and changepoint setting. Specifically, Scenario 5 represents gradually increasing mean shifts without linear trends, making changepoint detection harder, because the increasing mean shifts can be easily confounded with a linear trend for some changepoint techniques. This issue can be problematic as it would suggest a spurious linear trend when there is in fact no linear trend. Scenario 6 has one temporary mean shift occurring at $\tau_1 = 200$, then it reverts back to pre-changepoint level at $\tau_2 = 300$. This scenario tests if the GA can correctly detect a temporary mean shift in means lasting a relatively short time. The last two scenarios assume three changepoints. Scenario 7, as an extension of Scenario 5, considers gradually increasing mean shifts without linear trends. In Scenario 8, motivated from the Fishguard series in Section 2.5, we assess the GA in another challenging setting where the true linear trend is decreasing with positive mean shifts, because no-changepoint models could incorrectly identify the decreasing trend as an increasing trend.

Table 2.1: Linear trend and changepoint configurations for the simulation study

Scenario	Linear trend	Mean shifts	Changepoint configuration
1	0		$(c = 0)$
2	2.0		$(c = 0)$
3	0	0.4	$(c = 1;\ \tau_1 = 200)$
4	2.0	0.4	$(c = 1;\ \tau_1 = 200)$
5	0	0.4, 0.8	$(c = 2;\ \tau_1 = 200, \tau_2 = 400)$
6	2.0	−0.4, 0	$(c = 2;\ \tau_1 = 200, \tau_2 = 300)$
7	0	0.4, 0.8, 1.2	$(c = 3;\ \tau_1 = 150, \tau_2 = 300, \tau_3 = 450)$
8	−0.5	0.4, 0.8, 1.2	$(c = 3;\ \tau_1 = 100, \tau_2 = 300, \tau_3 = 500)$

Long-lasting cyclical autocorrelation is present in raw hourly sea level series at some GESLA locations. To reflect this type of correlation, we used the following data generating scheme.

1. An hourly series of length $N = mn = 720 \times 600$ (600 months of hourly observations) is generated from a stationary Gegenbauer process with $u = 0.875$ and a given value of $\lambda \in (0, 0.5)$. The generated series is then standardized to have a mean of 0 and a standard deviation of 1. This standardized stationary hourly series, denoted by $\{\tilde{X}_1, \ldots, \tilde{X}_N\}$, serves as our base series to mimic temporal correlation in hourly sea levels.

2. Next, we mimic periodic fluctuations in means and variances of the raw hourly data by using the following periodic means and standard deviations: for time (hour) $s = 1, \ldots, N$,

$$\ddot{\mu}_s = 2.6 + 0.001 \sin(2\pi s/T_1) - 0.015 \sin(4\pi s/T_1) - 0.002 \cos(4\pi s/T_1)$$
$$+ 0.001 \cos(2\pi s/T_2) + 0.001 \sin(4\pi s/T_2) + 0.001 \cos(4\pi s/T_2),$$
$$\ddot{\sigma}_s = 0.5 + 0.001 \cos(2\pi s/T_1) - 0.002 \sin(4\pi s/T_1) - 0.002 \cos(4\pi s/T_1)$$
$$- 0.001 \sin(2\pi s/T_2) + 0.002 \cos(2\pi s/T_2) - 0.01 \sin(4\pi s/T_2) - 0.2 \cos(4\pi s/T_2),$$

where $T_1 = 24.838$ and $T_2 = 708.734$ represent a tidal lunar day and synodic lunar month, respectively. These model specifications are chosen from our analysis of the hourly sea level series to imitate its periodicity and variability in hourly sea levels. We consider $\{\ddot{X}_1, \ldots, \ddot{X}_N\}$, where $\ddot{X}_s = \ddot{\mu}_s + \ddot{\sigma}_s \tilde{X}_s$, to be a simulated series of hourly sea levels without linear trends and mean shifts induced by changepoints.

3. For a given scenario in Table 2.1, an hourly sea level series $\{X_1, \ldots, X_N\}$ is generated by incorporating long-term trend and mean shifts occurring in months τ_j's into $\{\ddot{X}_1, \ldots, \ddot{X}_N\}$.

4. A monthly maximum series $\{Y_1, \ldots, Y_n\}$ is obtained with a block size $m = 720$. We consider this simulated maximum series to reflect temporal dependence in hourly sea levels.

Following these steps, we generated 1000 monthly maximum series of $n = 600$ for each of the eight simulation scenarios. Fig. A.1 (in the supplementary materials) shows an exemplary time plot from each scenario along with true changepoint times marked by red vertical lines (Scenarios 3–8). We chose the Gegenbauer process to reflect long-lasting cyclical autocorrelation present in the hourly sea level series (Woodward, Gray, and Elliott, 2017). The periodicity and autocorrelation of a Gegenbauer process are governed by the parameters u and λ. We set u to be 0.875 for a periodic cycle of 12.433 ($= 2\pi/\cos^{-1}(0.875)$) hours, mimicking a tidal pattern present in some GESLA stations. Two values of λ are selected: $\lambda = 0.25$ for a weak but long-lasting temporal correlation case and $\lambda = 0.375$ for a moderately strong and long-lasting temporal correlation case, therefore we assess how our GA performs under different levels of temporal correlation.

Using the MDL in (2.4) as the fitness function to minimize, our GA method is applied to these simulated monthly maximum series $\{Y_1, \ldots, Y_n\}$ to estimate the number and times of changepoints. Table 2.2 shows the estimated detection rates of the correct changepoint numbers and the estimated accuracy rates of detected changepoint times with three different tolerance bands (± 3, ± 6, and ± 9 months) for the weak but long-lasting correlation case with $\lambda = 0.25$. The GA appears to perform well in estimating the changepoint number for all eight scenarios since it correctly estimated c in nearly all of the repetitions. The GA changepoint times are also close to the true changepoint times for all simulation scenarios. About 87–93% of all repetitions correctly estimated the changepoint times within a three-month margin.

If an error is allowed up to nine months, the GA correctly estimates the changepoint times nearly 99% of the time. A frequency histogram of the estimated changepoint times is shown for Scenarios 3–8 in Fig. A.2 (in the supplementary materials). Most of the estimated changepoint times are clustered around the true changepoint times with a minimal variability. Our GA method performs well in estimating changepoint times for these six scenarios with weak but long-lasting temporal correlation.

Table 2.2: Detection counts and accuracy rates for the GA method with $\lambda = 0.25$

Scenario	$c = 0$	$c = 1$	$c = 2$	$c = 3$	$c = 4$	$c = 5$	± 3 mos.	± 6 mos.	± 9 mos.
1	**999**	1	0	0	0	0			
2	**991**	9	0	0	0	0			
3	0	**993**	7	0	0	0	90.6%	97.7%	98.7%
4	0	**993**	7	0	0	0	92.8%	97.6%	98.7%
5	0	0	**995**	5	0	0	90.6%	97.4%	99.4%
6	0	0	**995**	4	1	0	87.2%	96.7%	98.5%
7	0	0	0	**997**	3	0	89.0%	96.1%	98.4%
8	0	0	0	**993**	7	0	88.3%	97.5%	98.8%

Table 2.3 summarizes the estimated detection rates of the changepoint number and the estimated accuracy rates of changepoint times under the moderately strong and long-lasting correlation case with $\lambda = 0.375$. Our GA correctly estimates the changepoint numbers about 87–93% of the time for all eight scenarios. Within a three-month margin, the GA correctly estimates the changepoint times about 72%–78% of the time. The accuracy rates increase to around 85–90% with a nine-month margin. Fig. 2.1 shows a frequency histogram of the estimated changepoint times for Scenarios 3–8. Although the estimated times are slightly more spread out than those of the case with $\lambda = 0.25$, they are still well clustered around the true changepoint times with a minimal variability. These findings, combined with our results from the weaker correlation case, suggest that our GA performs well in estimating the

number and locations of mean shifts induced by changepoints even for a series with moderately strong, long-lasting temporal correlations.

Table 2.3: Detection counts and accuracy rates for the GA method with $\lambda = 0.375$

Scenario	$c=0$	$c=1$	$c=2$	$c=3$	$c=4$	$c=5$	± 3 mos.	± 6 mos.	± 9 mos.
1	**913**	52	35	0	0	0			
2	**913**	57	27	2	1	0			
3	0	**878**	85	36	1	0	75.6%	82.9%	85.8%
4	0	**875**	98	25	2	0	73.8%	81.5%	85.2%
5	0	0	**922**	69	9	0	77.7%	86.4%	89.5%
6	0	0	**906**	84	10	0	73.0%	83.8%	86.4%
7	0	0	0	**927**	68	5	75.6%	85.4%	90.0%
8	5	4	7	**897**	81	6	72.2%	82.9%	87.0%

2.5 Case Study: Fishguard, UK

2.5.1 Long-term trend estimation with changepoints

We use the sea level data of Fishguard, UK to illustrate our methods. Once the monthly maximum sea level series was extracted from the raw GESLA hourly data at this gauge site by following our preprocessing procedures as explained in Section 2.2, we applied the GA method to the Fishguard monthly maximum series using the MDL in (2.4) as the GA's fitness function to optimize. The GA estimates three changepoints at times $\tau_1 = 141$ (September 1974), $\tau_2 = 355$ (July 1992), and $\tau_3 = 592$ (April 2012).

Although metadata are not available for most GESLA stations, BODC, the provider of the Fishguard data, informed us that the Fishguard tide gauge had experienced two location changes in 1975 and June 1988. The gauge also had a new instrument installed in June 1988, set to record data in 15 minute intervals. However, it took a few more years to fully change over, so hourly data were recorded until December
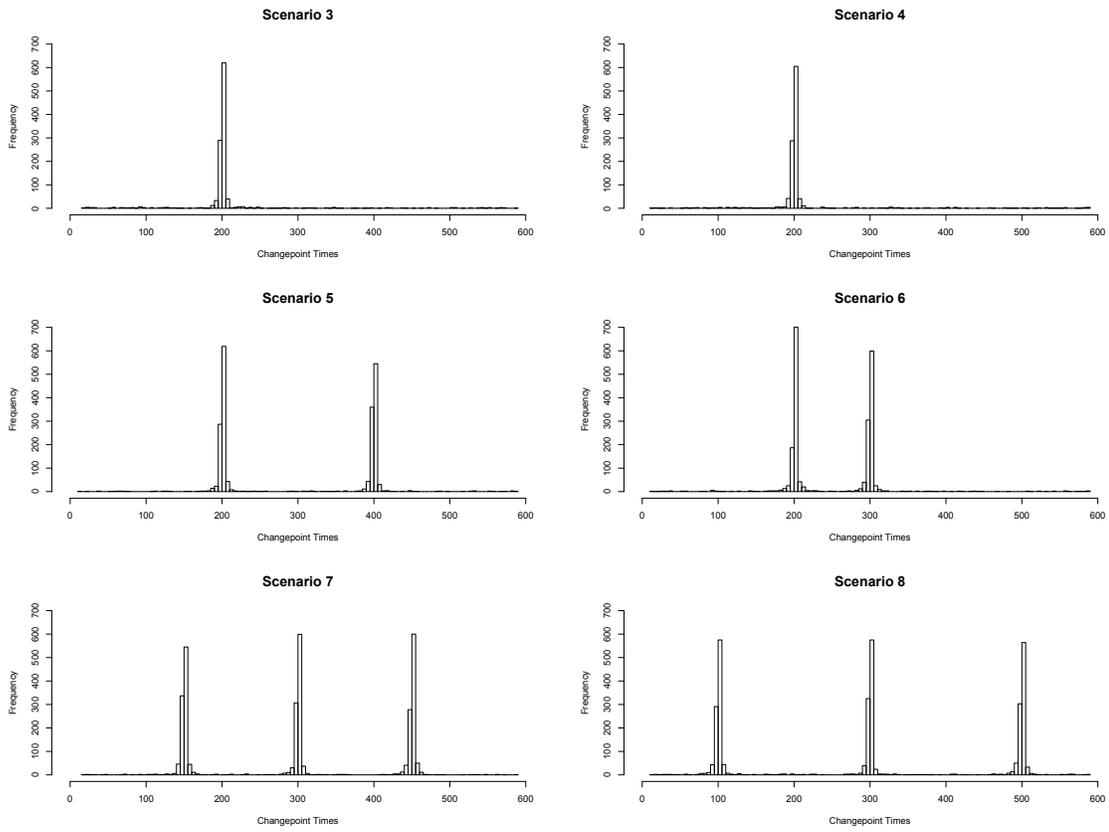
Figure 2.1: Histograms of detected changepoint times from the GA method with $\lambda = 0.375$ for Scenarios 3–8

1992. These two location and instrument changes in 1975 and 1992 appear to concur with the first two GA changepoints on September 1974 and July 1992. After then, the Fishguard tide gauge had been experiencing siltation issues, where tide sensors get buried in silt causing a back pressure, from as early as March 2013. This siltation issue could be a factor contributing to the third Fishguard GA changepoint in April 2012.

Using these three GA changepoints in the GEV model with the parameter specifications in (2.1) and (2.2), we estimate the model parameters by maximizing the likelihood function in (2.3). For comparison, we also estimate the parameters without allowing changepoints. The standard errors for the parameter estimates are computed from 10,000 bootstrap samples by using the moving block bootstrap method with a block size of $k = 12$ as indicated in Section 2.3.4.

Table 2.4 summarizes our GEV model parameter estimates along with their corresponding bootstrap standard errors in parentheses. MDL is substantially improved in the GEV model with the three changepoints considered, suggesting that the GEV model with GA changepoints offers a better fit to the Fishguard monthly maximum series than the GEV model with changepoints ignored. Whereas most estimates are similar for the two models, the estimated long-term trend parameter α, one of our main interests, has oppositely changed from positive to negative with a larger magnitude. To be specific, in the changepoints-ignored model, the estimated linear trend is 34 cm century^{-1} with the 95% BCa bootstrap confidence interval of $(24, 48)$. However, with the three changepoints included in the model, the estimated linear trend drastically changes to -80 cm century^{-1} with the 95% BCa bootstrap confidence interval of $(-112, -44)$. This finding indicates that the monthly maximum sea level series in Fishguard, UK has in fact significantly decreased by about 80 cm for the last century, rather than gradually increasing over the record period. Fig. 2.2

shows a time plot of the Fishguard monthly maximum sea levels along with the three estimated changepoint times marked by purple vertical lines. The estimated trends with and without changepoints are also imposed by red and blue lines, respectively. The estimated trend line with those three changepoints appears to be appropriate for the Fishguard series.

Table 2.4: Fishguard GEV parameter estimates and bootstrap standard errors in parentheses (units: meters for β's, Δ's, and ω's; m century^{-1} for α)

	Three changepoints	No changepoints
β_0	5.038 (0.020)	4.998 (0.020)
β_1	0.108 (0.012)	0.109 (0.012)
β_2	-0.027 (0.011)	-0.033 (0.012)
β_3	-0.154 (0.015)	-0.156 (0.015)
β_4	0.014 (0.015)	0.015 (0.016)
α	-0.800 (0.175)	0.342 (0.063)
Δ_1	0.199 (0.033)	
Δ_2	0.418 (0.062)	
Δ_3	0.651 (0.089)	
ω_0	0.163 (0.006)	0.173 (0.006)
ω_1	0.026 (0.006)	0.023 (0.007)
ω_2	0.009 (0.007)	0.011 (0.008)
ω_3	0.007 (0.008)	0.012 (0.008)
ω_4	0.008 (0.006)	0.007 (0.007)
ξ	-0.250 (0.027)	-0.237 (0.033)
θ	0.128	0.128
MDL	-245.886	-234.857

A goodness-of-fit for the GEV model is also performed to check if our GEV model fits well to the Fishguard monthly maximum sea level series. Fig. 2.3 shows the Gumbel scaled quantile-quantile plots for our GEV models, visually comparing the two GEV models with the three GA changepoints considered and without allowing changepoints. Overall, the changepoints-included GEV model is an improved fit to the Fishguard series, since the residuals from the changepoints-included GEV model

in the plot (left) nearly form a straight line. This supports our claim that changepoints should be considered in the GEV analysis of extreme sea levels.
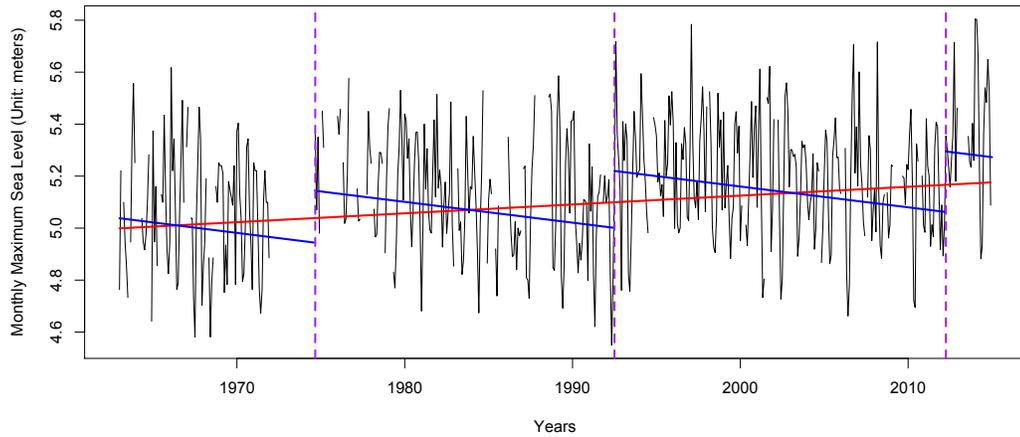
Figure 2.2: **Estimated trend lines for the Fishguard monthly maximum sea levels (blue solid lines, the trend with three GA estimated changepoint times; red solid line, the trend without allowing changepoints)**
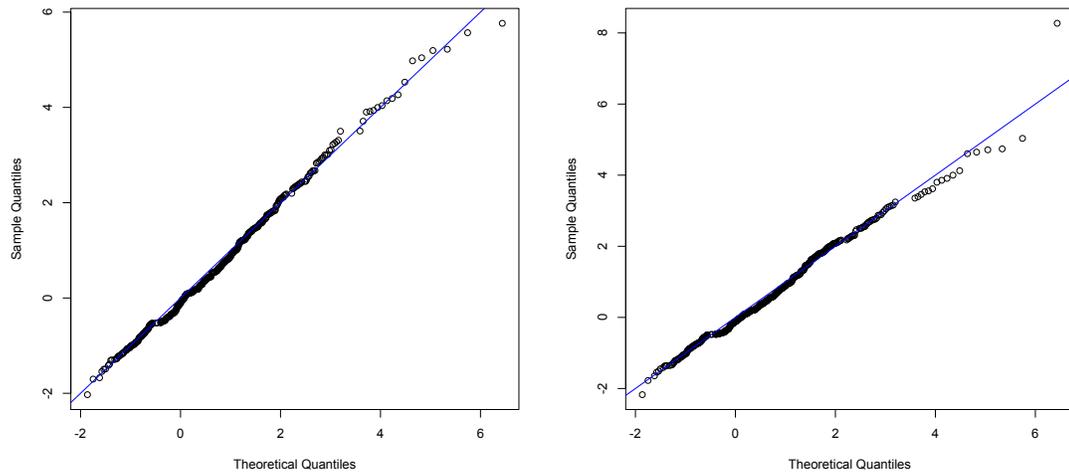
Figure 2.3: **Gumbel scaled quantile-quantile plots for the GEV models with three GA estimated changepoints included (left) and without allowing changepoints (right) for the Fishguard monthly maximum sea levels**

2.5.2 Return level estimation with non-stationarity and temporal correlation

Since the Fishguard hourly sea level series is strongly autocorrelated, we estimate the extremal index θ and use it for accurate return level estimation as explained in Section 2.3.3. However, as Northrop's estimator is developed with a stationary series, we first remove non-stationary features, such as mean shifts, linear trend, and periodicity, from the raw data. We used the following stationary conversion procedure. Other approaches can be used as an alternative.

1. We convert the non-stationary raw hourly sea level data $\{X_1, \ldots, X_N\}$ to the changepoint-adjusted and detrended hourly data $\{\ddot{X}_1, \ldots, \ddot{X}_N\}$, where $\ddot{X}_s = X_s - \psi_s - \gamma s$, and ψ_s and γ are mean shift and linear trend parameters for the raw data. However, there are some challenges in this changepoint and trend estimation. First, a single tide gauge typically contains around a half million sea level observations, which is too large for effective changepoint estimation using existing changepoint methods. Second, most sea level series from the raw data exhibit strong serial correlation and do not follow the Gaussian distribution, limiting the use of many useful Gaussian-based changepoint methods available. Third, the marginal distribution of $\{X_1, \ldots, X_N\}$ is unknown to us, which leads to the use of non-parametric changepoint methods. To the best of our knowledge, there are no methods developed under this challenging and complicated case.

2. To overcome these issues, we approximate changepoints in the raw data via changepoints in monthly median series. We chose monthly median over monthly mean, since median is more robust than mean if the distribution is skewed or

outliers are present in the data as is the case with sea levels. We apply the E-divisive method to the monthly median series and use the estimated E-divisive changepoints as a proxy for the changepoints in the raw data. Matteson and James (2014) showed that the E-divisive method is an effective non-parametric method in a simulation study. Our preliminary simulation results also suggest that the E-divisive method performs satisfactorily in detecting changepoints in a correlated raw series when applied to monthly median series without linear trends.

3. For the Fishguard monthly median series, the E-divisive method estimates five changepoints in February 1973, August 1997, September 2000, March 2007, and September 2011. Two of the three GA changepoints for the monthly maximum series mostly concur with the E-divisive changepoints from the monthly median series. We use the E-divisive changepoint times to estimate the mean shift sizes and linear trend in the raw hourly data by the least squares method. Subtracting these estimates from $\{X_1, \ldots, X_N\}$ produces $\{\dddot{X}_1, \ldots, \dddot{X}_N\}$.

4. The changepoint-adjusted and detrended hourly series $\{\dddot{X}_1, \ldots, \dddot{X}_N\}$ is then transformed to the stationary series $\{\tilde{X}_1, \ldots, \tilde{X}_N\}$ by taking $\tilde{X}_s = (\dddot{X}_s - \dddot{\mu}_s)/\dddot{\sigma}_s$, where $\dddot{\mu}_s$ and $\dddot{\sigma}_s$ are hourly mean and standard deviation of the $\dddot{X}_s$ series and are calculated by a similar method to Woody, Wang, and Dyer (2016). To elaborate, for each hour ν of the synodic lunar month (approximately 708.734 hours or 29.5 days), we compute the means and standard deviations

$$\bar{\bar{X}}_\nu = \frac{1}{n_\nu} \sum_{j=0}^{\lfloor N/T_L \rfloor} \dddot{X}_{[jT_L+\nu]}, \quad \dddot{S}_\nu = \sqrt{\frac{1}{n_\nu - 1} \sum_{j=0}^{\lfloor N/T_L \rfloor} (\dddot{X}_{[jT_L+\nu]} - \bar{\bar{X}}_\nu)^2},$$

where n_ν is the number of non-missing values during hour $\nu = 1, \ldots, 708$ of the synodic lunar month, $\lfloor a \rfloor$ denotes the largest integer smaller than or equal to a, $[a]$ returns the index of $\ddot{X}_s$ closest to a, and $T_L = 708.734$. Next, a regression model with four harmonics of periods $T_1 = 24.838/2$, $T_2 = 24.838$, $T_3 = 708.734/2$, and $T_4 = 708.734$ is fitted to smooth these means and standard deviations. Here, T_1 and T_2 account for a tidal lunar day, and T_3 and T_4 account for a synodic lunar month. These smoothed values are then extended to all hourly times $s = 1, \ldots, N$, resulting in the periodic mean $\ddot{\mu}_s$ and periodic standard deviation $\ddot{\sigma}_s$ over the entire record period. Finally, we obtain $\{\tilde{X}_1, \ldots, \tilde{X}_N\}$, where $\tilde{X}_s = (\ddot{X}_s - \ddot{\mu}_s)/\ddot{\sigma}_s$, and consider this series to be stationary.

Now, we estimate the extremal index θ by applying the Northrop's semiparametric maxima estimator to the stationary hourly series $\{\tilde{X}_1, \ldots, \tilde{X}_N\}$. The return levels for 25, 50, 75, and 100 years for the Fishguard monthly maximum sea levels are then computed from the GEV model with the three GA changepoints. To adjust the return levels for non-stationarity and temporal correlation, we apply the method in (2.5) with $\hat{\theta}$ used in $G_t^\theta(\cdot)$ as illustrated in Section 2.3.3. To make the return levels more informative, we subtract the median of all sea level observations pertaining to the last 12 months' records. The resulting return levels are then the expected maximum amount of exceedance from the typical current sea level for a given time period starting from January 2020. This, in turn, normalizes the return sea level estimates across the sites and allows us to make meaningful comparisons between tide gauges. Table 2.5 summarizes the estimated 25, 50, 75, and 100-year Fishguard sea level return exceedances with the three GA changepoints considered and without allowing changepoints. Their associated 95% BCa bootstrap confidence intervals are also reported. We find that the consideration of changepoints noticeably changes the

return levels. For example, the estimated 50-year return level is lowered by about 22 cm when the GA changepoints are considered.

Table 2.5: Fishguard monthly maximum sea level return exceedances and their 95% BCa bootstrap confidence intervals in parentheses (unit: meters)

Return exceedances	Three changepoints	No changepoints
r_{25}	2.899 (2.820, 3.048)	3.014 (2.955, 3.086)
r_{50}	2.902 (2.818, 3.048)	3.121 (3.041, 3.219)
r_{75}	2.902 (2.818, 3.046)	3.210 (3.108, 3.339)
r_{100}	2.902 (2.818, 3.047)	3.296 (3.169, 3.456)

2.6 GESLA data analysis

As illustrated in Section 2.5, our methods were applied to the monthly maximum sea level series at all 300 gauge sites. Out of the 300 sites, 47.7% (143 sites) are identified to have no significant mean shifts in their monthly maximum series. Most of these sites are either in North America or Europe. The other 52.3% (157 sites) are flagged for one or more changepoints. Of those sites with at least one GA changepoints, 89.2% (140 sites) have experienced one to four changepoints. Five or more changepoints are also found in the other 10.8% (17 sites), of which 11 sites are located in the Pacific Ocean and the other six are scattered around in Japan, Central America, and North America. With these GA changepoints incorporated into our GEV model, the GEV parameters were estimated via the maximum likelihood method for all 300 locations. The GEV parameters without allowing changepoints were also estimated for comparison purposes.

To illustrate how changepoints can influence the long-term trend estimation, we select six gauge sites with a noticeable difference in trend estimates. The selected
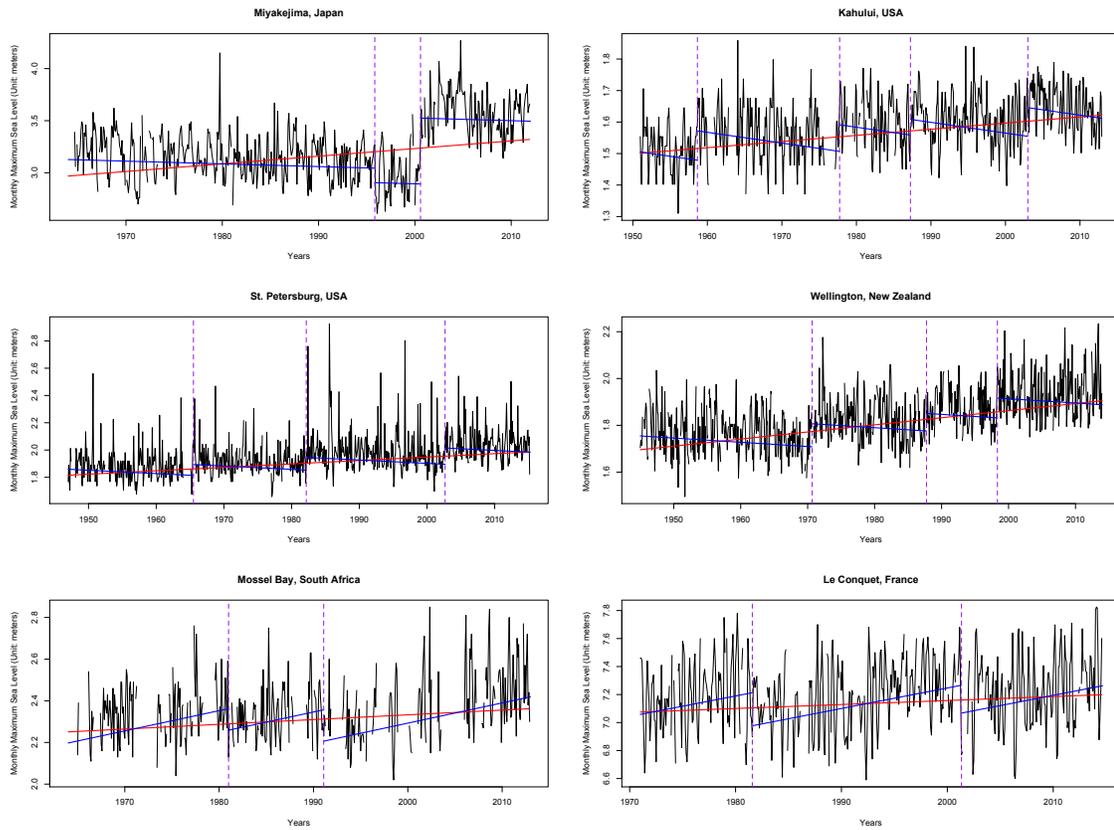
Figure 2.4: Estimated trend lines for the monthly maximum sea levels (blue solid lines, the trend with GA estimated changepoint times; red solid line, the trend without allowing changepoints)

gauge sites are Miyakejima, Japan; Kahului, USA; St. Petersburg, USA; Wellington, New Zealand; Mossel Bay, South Africa; and Le Conquet, France. Fig. 2.4 shows the monthly maximum sea levels at these sites along with their estimated linear trends with and without changepoints. The purple vertical lines in this figure denote the GA estimated changepoint times. The Miyakejima series has experienced two changepoints, including a substantial upward mean shift in September 2000. It is uncertain to us what caused these mean shifts, because the GESLA data set contains no quality control flags for these times. Whereas the trend estimate was

positive (73 cm century^{-1}) when changepoints are not allowed, these two changepoints considerably decrease the trend estimate to a negative trend (-26 cm century^{-1}). Similar to the Miyakejima series, the estimated trends at Kahului, St. Petersburg, and Wellington oppositely change to negative when changepoints are taken into account, suggesting that monthly maximum sea levels in these sites have decreased rather than increased. Consideration of changepoints, however, does not necessarily decrease trend estimates. For example, the monthly maximum sea levels at Mossel Bay and Le Conquet show a substantial increase to their estimated trends when changepoints are considered.

The impact of changepoints on the long-term trend estimation is further examined for those 157 gauge sites with at least one changepoints. Fig. 2.5 displays the estimated linear trends for these sites from the GEV models with changepoints considered and without allowing changepoints. Overall, the distributions of the trend estimates appear to be slightly left-skewed after few outliers are ignored. However, consideration of changepoints has increased variability and reduced the average estimated trends. Whereas the trend estimates without allowing changepoints have a mean of 25 cm century^{-1} and a standard deviation of 28 cm century^{-1}, the estimated trends with GA changepoints considered have a mean of 20 cm century^{-1} and a standard deviation of 45 cm century^{-1}. Fig. 2.6 shows a scatter plot of the trend estimates with changepoints included against those with changepoints ignored for the 157 sites. Many sites in the figure show substantial changes in their trend estimates, further supporting that changepoints should be considered to accurately quantify trends in monthly maximum sea levels.

A geographical map of the estimated long-term trends from the GEV model with changepoints considered for the 300 sites is presented in Map 2.2. Overall, monthly
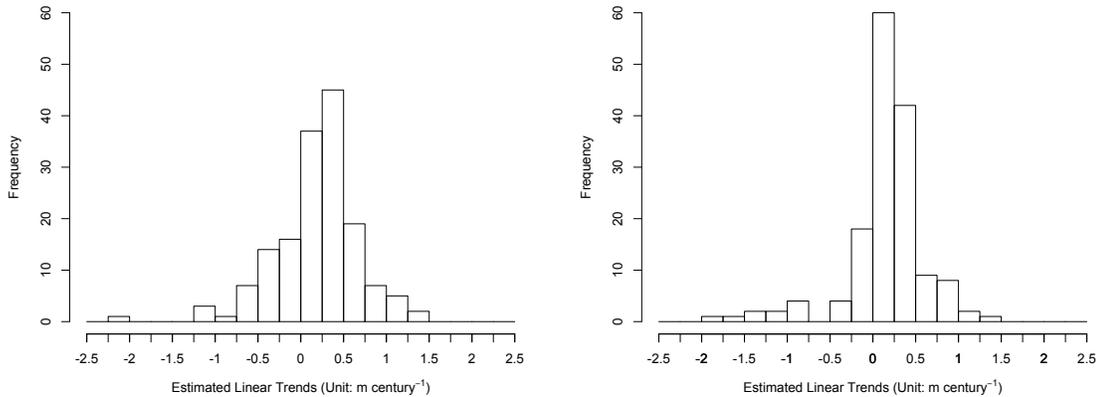
Figure 2.5: Histograms of the estimated linear trends with changepoints considered (left) and without allowing changepoints (right) for the 157 sites with at least one GA changepoints (unit: m century^{-1})

maximum sea levels have increased in the coasts of the North Atlantic Ocean, whereas the Gulf of Alaska and Baltic Sea have experienced decreasing monthly maximum sea levels. We also find that although most of Pacific Islands had overall increasing estimated linear trends when changepoints are ignored, estimated trends become negative for many locations once changepoints are considered.

Next, we calculate the 50-year return levels as in (2.5) for the 300 sites by using the extremal index estimated via Northrop's method. The median of all sea level observations pertaining to the last 12 months' records at each gauge site is subtracted from the return level estimates. To illustrate how temporal correlation in sea levels impact the return level estimation, we revisit the Miyakejima, Kahului, St. Petersburg, Wellington, Mossel Bay, and Le Conquet stations. The 50-year return level exceedances from the median sea level for these sites are estimated using (i) the GEV model with both changepoints and extremal index considered, and for comparison, also using (ii) the GEV model with changepoints included but temporal correlation ignored and (iii) the GEV model that ignores both changepoints and
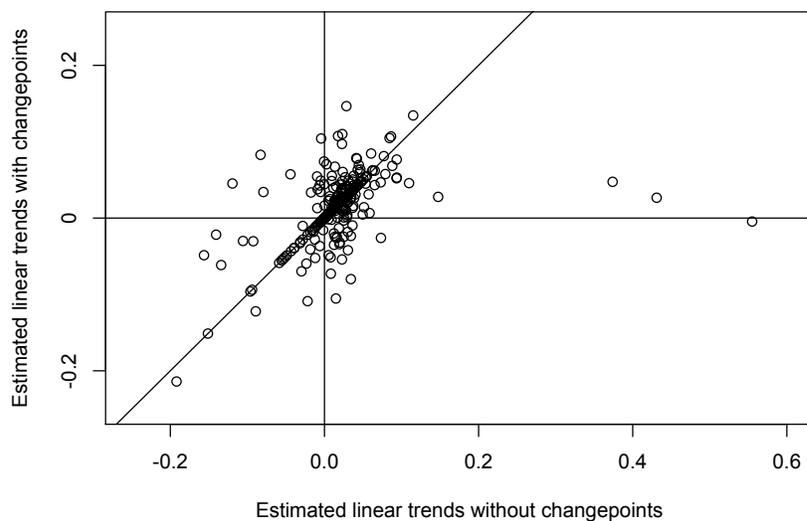
Figure 2.6: **A scatter plot of the estimated trends with changepoints considered against the estimated trends without allowing changepoints for the 157 sites with at least one GA changepoints (unit: m century^{-1})**
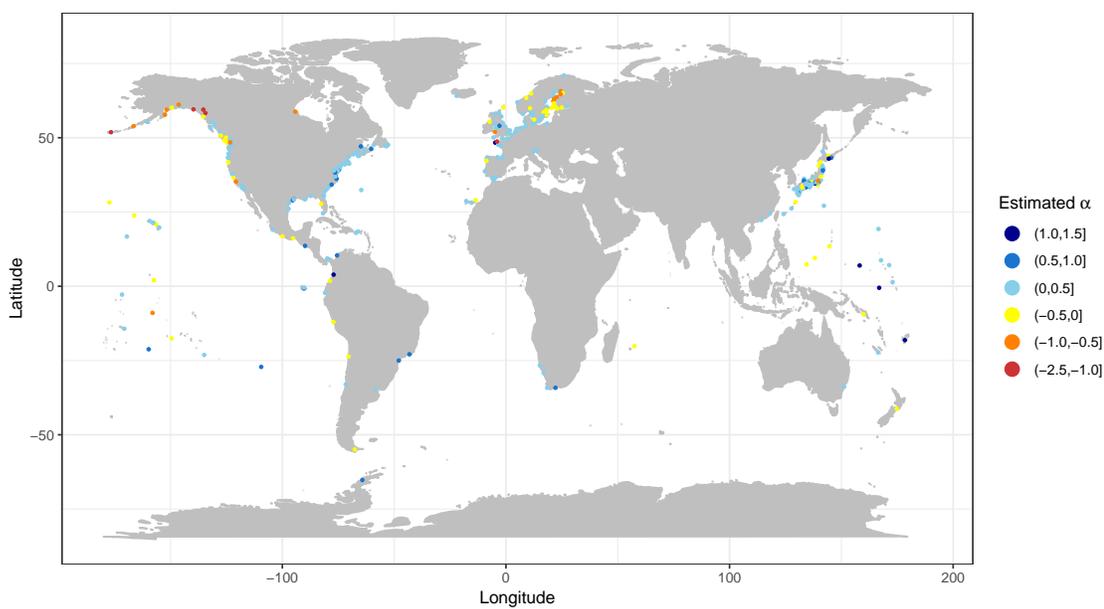
Map 2.2: **Estimated linear trends from the GEV model with changepoints considered (unit: m century^{-1})**

temporal correlation. Fig. 2.7 shows the estimated 50-year return level exceedances from the three GEV models. Once changepoints were included, return exceedance estimates decrease for those four stations with their trend estimates changing to a negative trend (Miyakejima, Kahului, St. Petersburg, and Wellington), and return exceedance estimates increase for the two stations that experienced a more positive trend estimate (Mossel Bay and Le Conquet). This is expected, since a larger long-term trend implicates a higher risk in return sea level exceedances. On the other hand, the consideration of extremal index results in lower return sea level exceedances for all six stations, implying that these six sites are in fact under a lower risk of extreme sea level events once temporal correlation is considered. In short, the temporal correlation in sea levels must be quantified and included in the modeling process to obtain accurate return level estimates of monthly maximum sea levels.
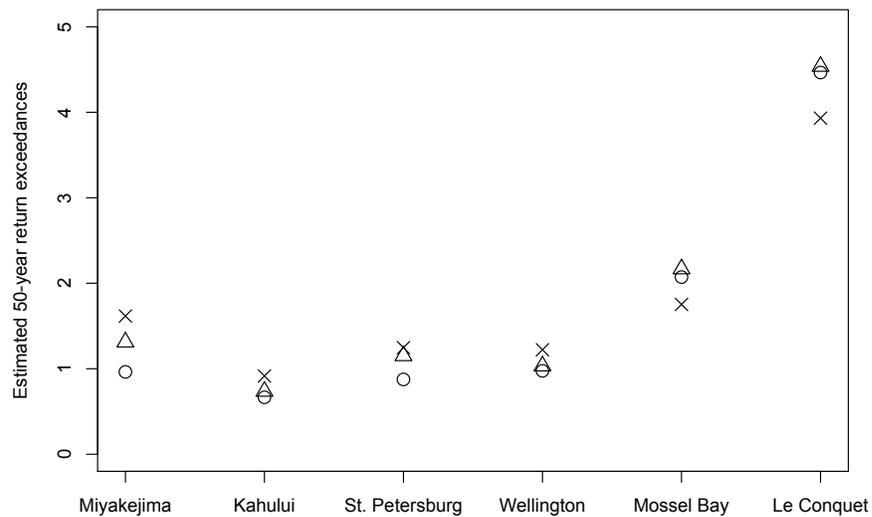
Figure 2.7: **Monthly maximum sea level 50-year return exceedances ($\circ$, change-points and extremal index considered; $\triangle$, changepoints only; $\times$, changepoints and temporal correlation ignored; unit: meters)**

Map 2.3 presents the estimated 50-year return sea level exceedances from the median for all 300 sites from the GEV model with GA estimated changepoints and temporal correlation considered. Coasts of the Northwestern Europe and the Gulf of Alaska appear to have the highest risk of extreme sea level events, with their typical current sea levels expected to exceed approximately 3 to 4 meters on average once between January 2020 and December 2069.
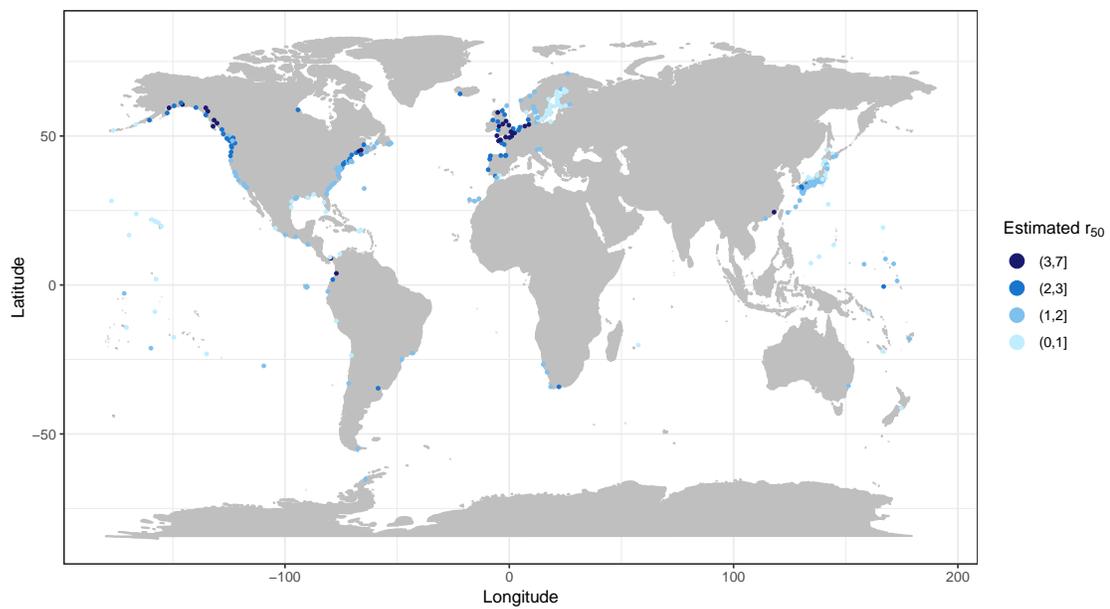
Map 2.3: Estimated monthly maximum sea level 50-year return exceedances (unit: meters)

2.7 Closing Comments

Sea level observations often contain undocumented changes in their means due to instrument changes, location changes, earthquakes, land reclamation, dredging, and so on. The number of changepoints and changepoint times are unknown if a metadata

of recording such changes does not exist or is not complete, as is often the case with sea level data. Trend analysis methods that ignore such changepoint features can result in erroneous estimates. We developed a GA method that uses a GEV-based likelihood and MDL penalty to detect the changepoint number and times in the monthly maximum sea level series at the 300 gauge sites selected for our study. The estimated changepoint number and times were then used to rigorously quantify long-term trends in the monthly maximum sea levels.

We find that monthly maximum sea levels have overall increased in many regions, including the Gulf of Mexico, the American coast of the North Atlantic Ocean, the Northwestern European coasts, and the Pacific Coast of Japan. This finding is consistent with those of Menéndez and Woodworth (2010), Wahl and Chambers (2015), and Marcos and Woodworth (2017). However, some regions have experienced a decrease in monthly maximum sea levels. Our decreasing trend for the Baltic Sea also appears in Menéndez and Woodworth (2010) and Marcos and Woodworth (2017). These decreasing trends for the Northwestern European coasts and the Baltic Sea could be due to the post-glacial land uplift in that region (The BACC II Author Team, 2015).

Our GA method has found one or more significant mean shifts in monthly maximum sea levels at 157 sites. Although the consideration of changepoints does not necessarily result in drastic changes to trend estimates at all gauge sites, trend estimates could be greatly influenced by detected changepoints. In particular, when changepoints are ignored, the monthly maximum sea levels in Pacific Islands show overall increasing trends, which is consistent with the finding of Menéndez and Woodworth (2010). However, we find that once changepoints are taken into account, the estimated trends for many Pacific Islands have decreased, revealing that the monthly

maximum sea levels of Pacific Islands are not uniformly increasing over time. Overall, 38 sites (12.67% of 300 sites), mostly scattered around the Pacific Ocean, had their estimated trends change signs after the consideration of changepoints. In addition, 29.7% (89 sites) of the 300 sites have their estimated trends changed by more than 20 cm century^{-1} after changepoints are considered. Among these 89 substantially impacted sites, 52 of them further show changes of more than 40 cm century^{-1}. Map 2.4 depicts the spatial location of these impacted sites. Many of these sites are spread out around the equator from Central Pacific Ocean to the western side of the South America, and some are clustered around the southern coast of Japan. There are also several such changepoint-influenced sites in the coasts of North America and Europe.
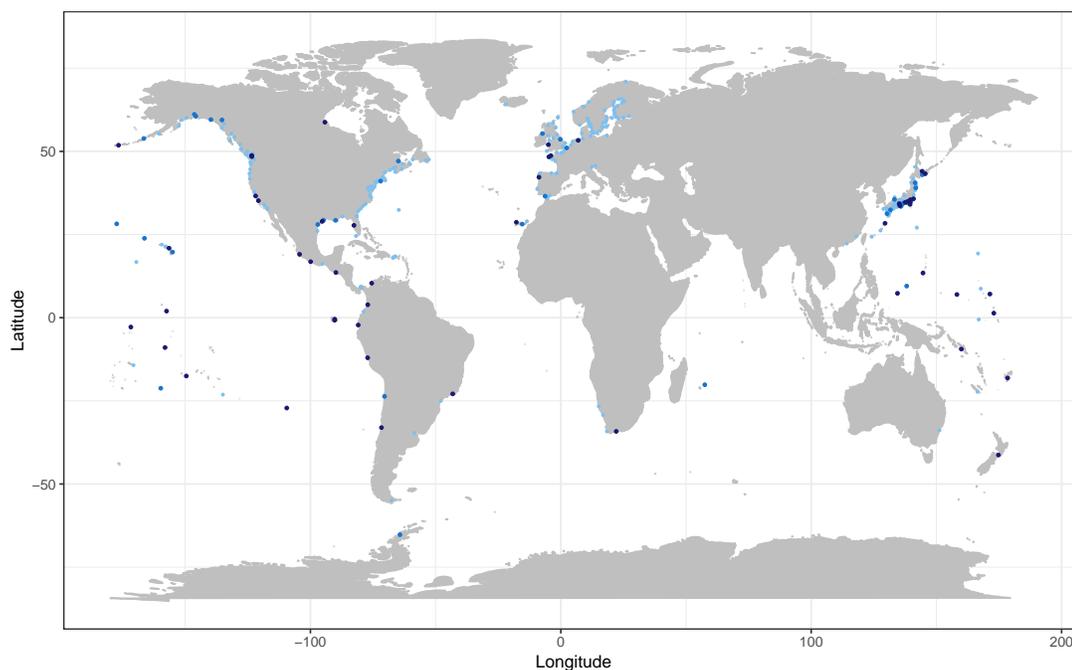
Map 2.4: **Spatial location of the tide gauges where the consideration of change-points changed the estimated trends by more than 40 cm century^{-1} (52 sites, dark blue), between 20 and 40 cm century^{-1} (37 sites, blue), and less than 20 cm century^{-1} (211 sites, light blue)**

Strong temporal correlation is pertinent in most raw sea level data. The extremal index θ, a measure for the strength of temporal correlation, was estimated between 0.005 to 0.501, with a median of 0.111. These small values of $\hat{\theta}$ indicate that many GESLA sea level series exhibit strong temporal correlations. For this reason, return sea level estimates could be greatly affected by the estimated extremal index. Therefore, we incorporate the extremal index to the GEV distribution for accurate return level estimation for all 300 sites. We find that 26 sites have their estimated 50-year return levels changed by more than 70 cm when extremal index is considered, and there are another 47 sites with changes between 30 and 70 cm. Map 2.5 shows the spatial location of these 73 gauge sites along with other tide gauges considered in this study. A large number of sites in the European coasts, particularly the Baltic Sea, have their estimated return exceedances severely impacted by the temporal correlation. Many sites in the Japanese coasts and the American coast of the North Atlantic Ocean, particularly the Gulf of Mexico, are also substantially affected. Bootstrap methods were used to compute the standard errors of GEV parameter estimates and bias-corrected confidence intervals for return levels. The GA estimated changepoints also affect these return levels when changepoints are detected.

There are some avenues for future research. First, our GEV model considers possible changes in the mean level of monthly maximum sea level series, while the GEV scale and shape parameters and the extremal index are not influenced by changepoint-inducing events. This model specification is due to our understanding that typical changepoint-inducing events for sea levels, such as relocation of tide gauges, repairs and/or changes in measuring equipment, or change in the elevation due to natural disasters, would affect the mean level of sea level series the most while having a lesser impact on long-term trends, variability, shape, and temporal correla-
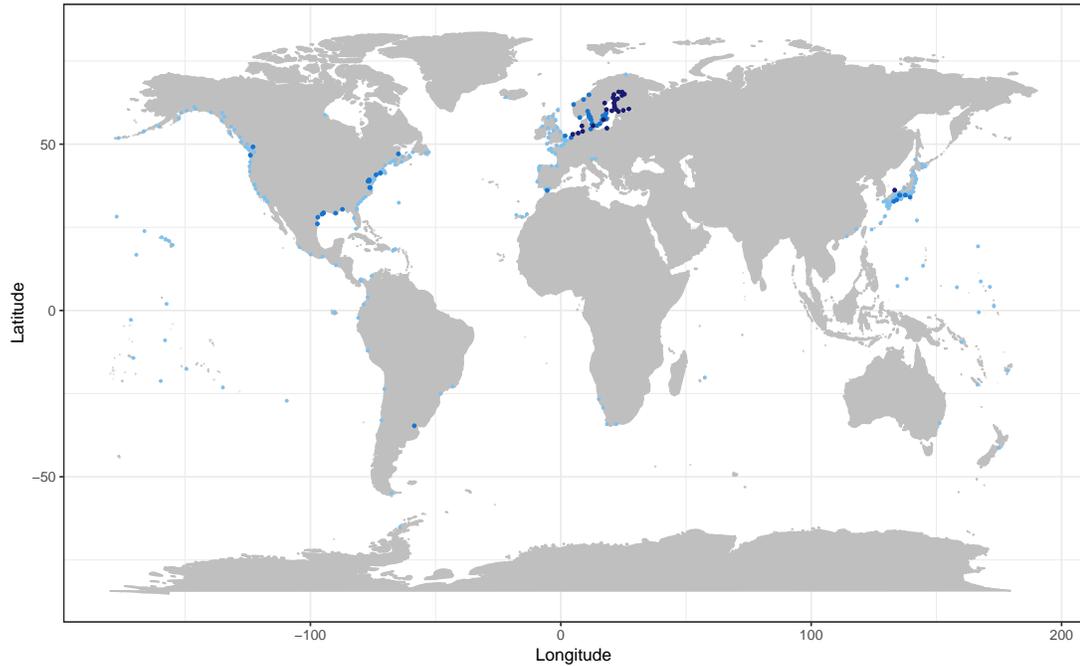
Map 2.5: Spatial location of the tide gauges where the consideration of extremal index changed the estimated 50-year return sea level exceedances from the median level by more than 70 cm (26 sites, dark blue), between 30 and 70 cm (47 sites, blue), and less than 30 cm (227 sites, light blue)

tion of sea levels. Although our model specification satisfactorily offers a good balance between goodness-of-fit and parsimoniousness, one could consider a GEV model with more complicated parameterization to capture more of the varying climate signals in sea levels. Second, the performance of our GA could be further assessed under more diverse simulation scenarios, including model mis-specification and different data generating schemes, to provide more robust evidence for the effectiveness of the GA method in different environments. Finally, our stationary conversion process applied to hourly sea level series as illustrated in Section 2.5.2 is ad-hoc. Developing an alternative procedure that reflects more diverse non-stationary features in the hourly sea level data could be further considered.

Acknowledgments

CHAPTER 3

WEEKLY MAXIMUM OZONE CONCENTRATIONS

3.1 Introduction

Extreme ground-level ozone pollution influences vast aspects of our lives, including human health, vegetation, and ecosystems. Unlike the naturally-occurring stratospheric ozone that filters the Sun's harmful radiation, ground-level ozone is a secondary pollutant formed via chemical reactions between nitrogen oxides (NOx) and volatile organic compounds (VOCs) under sunlight. Exposure to ground-level ozone along with other air pollutants are associated with a higher risk of death from cardiopulmonary causes (Jerrett et al., 2009), such as cardiovascular and respiratory diseases (Lim et al., 2019) and neurological diseases (Zhao et al., 2021). Further, ozone pollution damages vegetation and impacts agricultural production. For example, McGrath et al. (2015) found that elevated ozone concentrations reduced maize and soybean productions in the U.S. from the rain-fed fields by about 10% and 5% respectively, causing a total loss of approximately $9 billion each year. In the United States, the Environmental Protection Agency (EPA) sets the limit for ground-level ozone pollution along with other airborne pollutants via the National Ambient Air Quality Standards (NAAQS). As of 2015, the NAAQS regulates the daily maximum 8 hour average concentration at a standard of 0.07 ppm for most regions in the U.S.

Many authors examined how ground-level ozone levels have been changing over

time. To name a few, Jhun et al. (2015) found significant decreases in daytime ozone concentrations in the U.S. during warm seasons (May–October) and at peak levels ($\geq$75-th percentile). But they also found significant increases at non-peak levels ($<$75-th percentile) mostly due to increases during nighttime and cold seasons (November–April). Simon et al. (2015) evaluated ozone trends across the U.S. and found decreasing trends in the summer, less urbanized areas, and at peak levels (95-th percentiles), whereas they found increasing trends in winter, more urbanized areas, and at 5-th percentiles. They found decreased variability in overall ozone emission distribution. Yan, Lin, and He (2018) also found that the U.S. daytime ozone during summer has decreased from 1990 to 2014 due to reductions in anthropogenic emissions.

Although these authors found decreasing trends in the peak ozone levels, extreme value methods should be used to accurately estimate the long-term trends in extreme ozone. There are a few authors who applied extreme value methods to examine extreme ozone concentrations at a local scale. Dupuis (2005) analyzed weekly maximum ozone from four locations in southern Ontario, Canada and found that the area is projected to experience high ozone episodes (defined as ozone concentrations exceeding 0.08 ppm) with 0.55 probability each year. Chan and So (2018) applied their copula-based spatial generalized extreme value model to weekly maximum ozone along with other airborne pollutants from Pearl River Delta region in China. Hazarika, Borah, and Prakash (2019) analyzed daily maximum ozone in Delhi, India using generalized extreme value distribution.

At a regional scale, Shen, Mickley, and Gilleland (2016) developed a hybrid extreme value model to examine how climate change influences the number of days with high ozone concentrations (ozone episodes) in the U.S. They found that, assuming constant anthropogenic emissions at the present level, there will be as much as

3–9 more days of ozone episodes each year in the Northeastern, Midwestern, and Southwestern U.S., and 0–2 days elsewhere by the 2050s. Phalitnonkiat et al. (2016) analyzed 95-th percentile of the summertime ozone in the continental U.S. using generalized Pareto distribution with Hill estimator. They found that although the overall distributions of the extreme ozone in the Eastern U.S. have decreased in recent years, their tails became heavier, suggesting there may be more intense extreme ozone events in the future.

We aim to quantify long-term linear trends in U.S. county-level weekly maximum ozone concentrations using extreme value methods. However, there are two critical issues to consider to obtain accurate trend estimates. First, due to the events inducing changes in the ozone concentration observations such as changes in observer, instrument, measurement location, data collection method, etc., weekly maximum ozone concentration series can contain inhomogeneous changes in their distribution. If not appropriately considered in the modeling process, such changes could lead to misleading results, erroneously concluding that the extreme ozone series has been increasing over time although it has in fact been decreasing, or vice versa. The times at which these changes occur are often undocumented and therefore need to be estimated from the data. For this, we use a genetic algorithm (GA) to detect changepoints in the extreme ozone concentrations. Second, many weekly maximum ozone concentration time series often show long-memory autocorrelation, implying that the autocorrelation of the weekly maximum ozone concentrations at two distant time points is not ignorable. The methods that do not effectively take such long-memory autocorrelation into account could mistakenly detect spurious changepoints in the extreme ozone data, consequently resulting in inaccurate estimation of long-term trends. For more accurate estimation of changepoints with long-memory

autocorrelation present in weekly maximum ozone series, we extend the short-memory copula-GEV likelihood (Zhu, Liu, and Lund, 2019) to a long-memory copula-GEV likelihood, further developing a new GA-based changepoint detection method for an autocorrelated extreme series with long memory.

The rest of this paper proceeds as follows. Section 3.2 describes the extreme ozone concentration data set used in this study along with our data preprocessing procedures. Section 3.3 develops our extreme value model for weekly maximum ozone series with considerations for long-memory autocorrelation by incorporating the copula transformation and also illustrates the genetic algorithm for changepoint detection in a long-memory autocorrelated block maximum series. Section 3.4 summarizes the simulation study to assess the performance of our changepoint detection technique under different levels of the long-memory autocorrelation. Section 3.5 illustrates the application of our methods by making an in-depth exploration of the weekly maximum ozone series from Clark County in Nevada and Doña Ana County in New Mexico. Section 3.6 summates the long-term trend and return level analysis for the U.S. county-level weekly maximum ozone data. In Section 3.7, we conclude with further comments.

3.2 The Extreme Ozone Data

The EPA's Air Quality System (AQS) offers ambient air sample data collected by the EPA, state, local, and tribal air pollution control agencies. The AQS data offers high frequency observations on various air pollutants, such as ozone, nitrogen dioxide, PM 2.5, etc. We used hourly ozone concentration measurements to extract weekly maximum ozone concentration series from 1980 to 2021. The AQS hourly

ozone records, measured in ppm (parts-per-million), were downloaded from the EPA's website at `https://aqs.epa.gov/aqsweb/airdata/download_files.html#Raw`.

Extracting a county-level weekly maximum ozone concentration series from hourly ozone series requires the following data preprocessing procedures. The AQS ozone concentration observations are typically reported from multiple sites located in their respective county. For an analysis of extreme ozone concentrations at a county level, we need to combine ozone records from multiple sites to form one aggregated ozone series as a representative of the county. To aggregate for a county, we computed the mean of all non-missing hourly ozone observations at their measurement sites for each hour. We then use this hourly ozone mean series to be the hourly ozone concentration series for the county. Once an aggregated hourly series is obtained for each county, we proceeded to the weekly maximum extraction. Specifically, a weekly maximum was extracted only if the missing rate for the week is less than 15%. Otherwise, the weekly maximum ozone was not extracted and we treated that week as missing. Once a weekly maximum ozone concentration series was extracted for each of all available counties, we consider only those counties with at least 780 weeks ($\approx$ 15 years) of non-missing weekly maxima and an overall missing rate of less than 30%.

The preprocessing procedures above selected 395 U.S. counties for our weekly maximum ozone series analysis. Alaska, Hawaii, and other outlying U.S. territories were not considered in this study. Map 3.1 depicts the spatial location of the selected 395 counties. The Southwestern and Northeastern U.S. are fairly well represented, accounting for about 40% of the selected counties. Coastal counties in the Southeastern U.S. are also well represented, providing sufficient coverage on that region. However, counties in the Northwestern and Midwestern U.S. are sparsely represented. Many non-coastal counties in the Southern U.S. are also not included in our study.
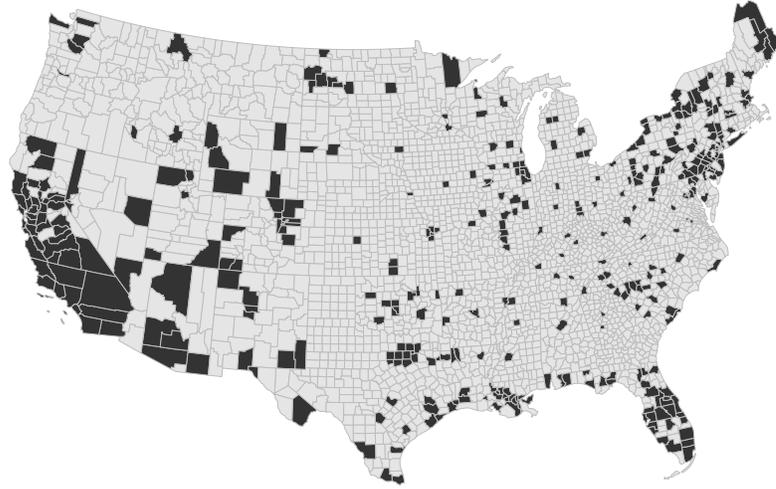
Map 3.1: Spatial location of the selected 395 counties after data preprocessing

3.3 Methods

3.3.1 Long-memory copula-GEV likelihood

As an extreme value counterpart to the central limit theorem, the extreme value theorem offers a limiting distribution for extreme statistics (cf. de Haan and Ferreira, 2006). Suppose $X_1, \ldots, X_m$ are independent and identically distributed (IID) random variables and $Y_{(m)} = \max\{X_1, \ldots, X_m\}$ is the maximum statistic of these m random variables. The extreme value theorem states that if the limiting distribution of $Y_{(m)}$ exists after appropriate normalization, then it takes the following generalized extreme value (GEV) distribution:

$$G(y) = \exp\left\{ -\left[1 + \xi\left(\frac{y - \mu}{\sigma}\right) \right]_+^{-1/\xi} \right\}, \tag{3.1}$$

where $[y]_+ = \max\{y, 0\}$. Here, $\mu \in (-\infty, \infty)$, $\sigma \in (0, \infty)$, and $\xi \in (-\infty, \infty)$ are GEV location, scale, and shape parameters, respectively.

Extreme environmental time series studies often analyze a sequence of these block maximum statistics $\{Y_1, \ldots, Y_n\}$ extracted from an initial environmental time series $\{X_1, \ldots, X_m, \ldots, X_{(n-1)m+1}, \ldots, X_N\}$ with $N = nm$. By the extreme value theorem, the GEV distribution is an appropriate probability distribution to those maximum statistics as long as the block size is set sufficiently large. For maximum likelihood estimation with independent block maxima, the GEV log-likelihood function is obtained by adding the logarithm of the marginal GEV probability density function evaluated at each Y_t. For an autocorrelated block maximum series, Zhu, Liu, and Lund (2019) used a copula transformation to construct a likelihood function that retains the GEV marginal distribution in each time t and short-memory autocorrelation $\sum_{h=-\infty}^{\infty} |\rho(h)| < \infty$ appearing in a time series model such as ARMA(p, q). However, when applied to a block maximum series with long-memory autocorrelation $\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty$ as in ARFIMA(p, d, q), the short-memory copula-GEV likelihood could produce biased model parameter estimates with underestimated standard errors for the estimates.

To be specific, Zhu, Liu, and Lund (2019) constructed the short-memory copula-GEV likelihood as follows. Suppose a base process $\{Z_1, \ldots, Z_n\}$ is a sequence of stationary random variables with a standard normal marginal distribution function $\Phi(\cdot)$. For $t = 1, \ldots, n$, define

$$Y_t = G^{-1}(\Phi(Z_t)), \tag{3.2}$$

where $G^{-1}(\cdot)$ is the inverse of the GEV distribution function. By the probability integral transformation, $\Phi(Z_t)$ follows a standard uniform distribution, which enforces Y_t via $G^{-1}(\cdot)$ to marginally follow a GEV distribution with the same correlation structure as Z_t for each t (cf. Nelsen, 2006, pp. 14–24). They derived the joint

density function of $Y_1, \ldots, Y_n$ using the Jacobian transformation as:

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = f_{Z_1,\ldots,Z_n}\left(\Phi^{-1}(G(y_1)), \ldots, \Phi^{-1}(G(y_n))\right)|\boldsymbol{J}|,$$

where $f_{Z_1,\ldots,Z_n}(\cdot)$ is the multivariate normal probability density function of the autocorrelated base process $\{Z_1, \ldots, Z_n\}$ and $\boldsymbol{J}$ is the Jacobian of the transformation with

$$|\boldsymbol{J}| = \prod_{t=1}^{n} \frac{\partial Z_t}{\partial Y_t} = \prod_{t=1}^{n} \frac{\partial \Phi^{-1}(G(Y_t))}{\partial Y_t} = \prod_{t=1}^{n} \frac{g(Y_t)}{\phi(\Phi^{-1}(G(Y_t)))}.$$

Here, $g(\cdot)$ is the GEV probability density function and $\phi(\cdot)$ is the standard normal probability density function. This method expresses the joint density function of $Y_1, \ldots, Y_n$ as a multivariate normal distribution function (Gaussian copula) as a proxy to characterize the correlation structure in Y_t while keeping GEV marginal distribution. Zhu, Liu, and Lund (2019) also proved the consistency and asymptotic normality of the maximum likelihood estimators for the short-memory copula-GEV likelihood when the base process follows an autoregressive moving-average (ARMA) process.

Now, to include long-memory autocorrelation in block maximum series in the modeling process, we develop the long-memory copula-GEV likelihood for a block maximum series with long-memory autocorrelation. For this, we assume that the base process $\{Z_1, \ldots, Z_n\}$ is a long-memory Gaussian process specified as the autoregressive fractionally integrated moving-average ARFIMA(p, d, q) model as follows:

$$\phi(B)(1 - B)^d Z_t = \theta(B)\varepsilon_t.$$

Here, $\phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$ is a causal p-th order AR polynomial with ϕ_j's as AR coefficients, $\theta(B) = 1 - \sum_{j=1}^{q} \theta_j B^j$ is an invertible q-th order MA polynomial with

θ_j's as MA coefficients, B is a backshift operator defined as $BZ_t = Z_{t-1}$, and $\{\varepsilon_t\}$ is a Gaussian white noise process with mean 0 and variance σ_ε^2.

Extension of this copula-GEV likelihood to a long-memory process is not straightforward and poses some challenges. Notably, the parameter estimation for the copula-GEV model with long-memory base process is much more computationally expensive than the one for short-memory process, necessitating the use of a more efficient computing algorithm. Our parameter estimation approach is illustrated in Section 3.3.4.

3.3.2   Non-stationarities with changepoints, trend, and periodicity

Changepoints, trends, and periodicities are important non-stationary features that commonly appear in many environmental time series. Because models that ignore these features could produce erroneous results, these non-stationary features must be considered in the modeling process. Our long-memory copula-GEV likelihood function is parameterized to capture the effects of changepoints and non-stationarity to accurately estimate a long-term trend in an extreme value series with long-memory autocorrelation.

The GEV location parameter for block maximum series $\{Y_1, \ldots, Y_n\}$ is parameterized to include a mean shift term induced by changepoint events, a periodic function composed of a sinusoidal wave with $K^{(0)}$ harmonics, and a long-term linear trend term. To elaborate, if there are c changepoints at times $\tau_1, \ldots, \tau_c$, we model the location parameter as

$$\mu_t = \beta_0 + \delta_t^{(0)} + \sum_{j=1}^{K^{(0)}} \left\{ \beta_{2j-1} \cos\left(\frac{2j\pi t}{T}\right) + \beta_{2j} \sin\left(\frac{2j\pi t}{T}\right) \right\} + \alpha\left(\frac{t}{LT}\right). \qquad (3.3)$$

Here, T is the period for the block maximum series. The long-term linear trend

parameter α is interpreted as the expected change in the block maximum series over L years, because

$$E(Y_{t+LT}) - E(Y_t) = \left[\mu_{t+LT} + \frac{\sigma_{t+LT}}{\xi} (\Gamma(1-\xi) - 1) \right] - \left[\mu_t + \frac{\sigma_t}{\xi} (\Gamma(1-\xi) - 1) \right] = \alpha,$$

if $\xi < 1$ and no changepoints have occurred between times t and $t + LT$. Similarly, we model the GEV scale parameter as

$$\ln \sigma_t = \omega_0 + \delta_t^{(1)} + \sum_{j=1}^{K^{(1)}} \left\{ \omega_{2j-1} \cos \left(\frac{2j\pi t}{T} \right) + \omega_{2j} \sin \left(\frac{2j\pi t}{T} \right) \right\}. \tag{3.4}$$

The shifts in the location parameter μ_t and in the scale parameter σ_t are

$$\delta_t^{(s)} = \begin{cases} 0, & \text{if } t = 1, \ldots, \tau_1 - 1; \\ \Delta_1^{(s)}, & \text{if } t = \tau_1, \ldots, \tau_2 - 1; \\ \vdots & \\ \Delta_c^{(s)}, & \text{if } t = \tau_c, \ldots, n, \end{cases}$$

for $s \in \{0, 1\}$. We assume that the the GEV shape parameter ξ is constant, although a more complex parameterization of ξ could be used, to lessen potential numerical instability issues in its estimate (cf. Zhang, Zwiers, and Li, 2004; Rust, Maraun, and Osborn, 2009).

Return levels for extreme environmental data form an important aspect of the extreme value analysis in quantifying the future risk of extreme events. In the stationary block maxima model, the z-year return level is the high quantile such that each block maximum has probability $1/(Tz)$ of exceeding that quantile. However, if a non-stationary time-dependent extreme value model is used, the probability of

exceedance is no longer constant over time. Instead, the method of Parey et al. (2007) and Parey, Hoang, and Dacunha-Castelle (2010) can be used to estimate return levels for non-stationary block maximum series. This method interprets the return levels as the value such that the expected number of exceedances during the given period is one. Specifically, the z-year return level r_z of a block maximum series is the solution to the following non-linear equation:

$$1 = \sum_{t=t_I}^{t_I+Tz-1} (1 - G_t(r_z)),$$

where $G_t(\cdot)$ is the time-dependent GEV distribution function, which is expressed as (3.1) with μ_t in (3.3) and σ_t in (3.4), and $t_I \geq n$ is a predetermined initial time for return levels. We use a grid search algorithm to numerically estimate r_z.

3.3.3 Changepoint detection using a genetic algorithm

Many environmental time series feature changepoints for many reasons, including instrument changes, location changes, changes in the data collection method, regime shifts in local ecosystem, etc. Such changepoints, if not addressed properly, can produce misleading results. However, detecting changepoints is a computationally expensive process. In an exhaustive search for optimal changepoints, there are $\binom{n-1}{c}$ different configurations for c changepoints for a series with size n. With potential changepoint number c in $\{0, 1, \ldots, n-1\}$, there are 2^{n-1} distinctive changepoint configurations to consider, a practically impossible task even with a moderately sized data set.

Due to this computational issue in changepoint detection, some past authors (cf. Davis, Lee, and Rodriguez-Yam, 2006; Lu, Lund, and Lee, 2010; Li and Lund, 2012; Lee, Li, and Lund, 2014; Hewaarachchi et al., 2017) used the genetic algorithm (GA),

an evolutionary algorithm that effectively searches for an optimal solution over a large parameter space using the principle of natural evolution: selection, crossover, and mutation. Their GA methods often encode each changepoint configuration (chromosome) as $(c; \tau_1, \ldots, \tau_c)$, where c is the number of changepoints and τ_j is the time at which the j-th changepoint occurs, and successfully found undocumented changeppoints in climate and environmental time series.

For our long-memory copula-GEV model with a block maximum series, we implement the GA for changepoint detection in a similar manner as in Lee, Li, and Lund (2014), except we directly apply the GA to the block maximum series without using a reference series. Also, we set each generation to consist of 180 distinct chromosomes and the mutation probability to 0.001. To improve the efficiency of the selection process, we use the elitist selection (cf. Bhandari, Murthy, and Pal, 1996) so that the two fittest chromosomes from a current generation are passed over to the next generation without any alterations. These two chromosomes are also crossed with each other and mutated with the same probability of mutation to create a new chromosome, the third elite one to be passed over to next generation. If the GA reaches the 200-th generation, we complete the GA and select the fittest chromosome from all 200 generations as the GA estimated changepoint configuration for the block maximum series.

To evaluate the model fitness of each chromosome, we use a penalized likelihood approach with the minimum description length (MDL) as the fitness function for the GA (cf. Davis, Lee, and Rodriguez-Yam, 2006; Lu, Lund, and Lee, 2010). However, the MDL expression should consider our GEV model specification. For this, we denote a vector of the GEV and ARFIMA parameters as

$$\boldsymbol{\eta} = (\beta_0, \ldots, \beta_4, \alpha, \Delta_1^{(0)}, \ldots, \Delta_c^{(0)}, \omega_0, \ldots, \omega_4, \Delta_1^{(1)}, \ldots, \Delta_c^{(1)}, \xi, d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)^{\mathrm{T}}.$$

The MDL for the model fit with a chromosome $(c; \tau_1, \ldots, \tau_c)$ is expressed as

$$\mathrm{MDL}(\boldsymbol{\eta}, c, \tau_1, \ldots, \tau_c) = -\ell_{\mathrm{opt}}(\boldsymbol{\eta}|c; \tau_1, \ldots, \tau_c) + P(c; \tau_1, \ldots, \tau_c).$$

Here, $\ell_{\mathrm{opt}}(\boldsymbol{\eta}|c; \tau_1, \ldots, \tau_c)$ is the optimized value of the copula-GEV log-likelihood calculated at the maximum likelihood estimates of the GEV and ARFIMA parameters $\boldsymbol{\eta}$ for a given changepoint configuration $(c; \tau_1, \ldots, \tau_c)$. Our penalty term $P(c; \tau_1, \ldots, \tau_c)$ for a changepoint configuration $(c; \tau_1, \ldots, \tau_c)$ under our GEV model specification is calculated as

$$P(c; \tau_1, \ldots, \tau_c) = \sum_{j=2}^{c+1} \ln(\tau_j - \tau_{j-1}) + \ln(c+1) + \sum_{j=2}^{c+1} \ln \tau_j,$$

where $\tau_{c+1} = n + 1$. In the first penalty term, the two shift parameters per each changepoint ($\Delta_j^{(0)}$ and $\Delta_j^{(1)}$ for $j = 1, \ldots, c$) are collectively charged the penalty of $\sum_{j=2}^{c+1} \ln(\tau_j - \tau_{j-1})$, because these shift parameters are real-valued parameters to be estimated over the j-th series segment $\{Y_{\tau_{j-1}}, \ldots, Y_{\tau_j - 1}\}$. If there is at least one missing observation in the j-th segment, then $\tau_j - \tau_{j-1}$ is replaced with the number of non-missing observations in the segment. The second term denotes the penalty for the number c of changepoints as an unknown integer-valued parameter. The third term penalizes the changepoint times $\tau_1, \ldots, \tau_c$, because each τ_j is an unknown integer-valued parameter bounded above by τ_{j+1}.

There is a non-ignorable risk of detecting spurious changepoints if even a short-memory autocorrelation is present in data and changepoint methods do not consider the autocorrelation (Tang and MacNeill, 1993). This spurious changepoint detection

issue could be more prevalent in time series with long memory (Nunes, Kuan, and Newbold, 1995; Kuan and Hsu, 1998; Krämer and Sibbertsen, 2002). In fact, Varneskov and Perron (2018) found that a short-memory based changepoint method tends to overestimate the number of true changepoints when the method is applied to an autocorrelated series with long memory. These findings indicate that if a block maximum series exhibits long-memory autocorrelation, changepoint methods that do not appropriately consider long-memory autocorrelation could detect spurious changepoints. We explore this spurious detection issue in Section 3.4 with simulated data.

3.3.4   Numerical computation and implementation

Computation of the joint probability density function for an autocorrelated extreme series with long memory can be computationally intensive even with a moderately sized sample size n. The Durbin-Levinson algorithm or innovations algorithm (cf. Brockwell and Davis, 2002, pp. 69–71), which have the numerical complexity $\mathcal{O}(n^2)$ for a linear stationary process, can be used to calculate the log-likelihood function of a stationary Gaussian process. Whereas Zhu, Liu, and Lund (2019) considered low-order ARMA(p, q) models as a base process in their methods, we consider the ARFIMA(p, d, q) model to use for a long-memory extreme series. With our GA using 200×180 evaluations of the log-likelihood function until termination, the Durbin-Levinson algorithm is not desirable as long as a long-memory base process is considered.

That said, we consider approximating the log-likelihood function of $Z_1, \ldots, Z_n$ using the method developed by Chan and Palma (1998). Palma and Chan (1997) then extended the method to deal with missing data. Their method uses an m-truncated moving-average expansion of the ARFIMA process to approximate its likelihood

function. Further, this algorithm has a numerical complexity of $\mathcal{O}(n \times m^2)$, where $m \ll n$ is a predetermined order for the moving-average truncation and does not increase with sample size n. This approach is often more efficient than the Durbin-Levinson algorithm (or other algorithms with comparable computational complexity, such as the innovations algorithm) particularly for a large sample size as long as m is set reasonably small. Technical details of this method are illustrated in the supplementary materials. We apply this moving-average truncation to approximate the log-likelihood function of a base process $\{Z_1, \ldots, Z_n\}$ and implement it in C++ to further improve the computation time.

Using the Kalman recursion, Chan and Palma (1998) and Palma and Chan (1997) expressed the joint probability density function of $\{Z_1, \ldots, Z_t\}$ as:

$$f_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) = (2\pi)^{-n/2} \left(\prod_{t=1}^{n} \Delta_t \right)^{-1/2} \exp\left\{ -\frac{1}{2} \sum_{t=1}^{n} \frac{(z_t - \hat{z}_t)^2}{\Delta_t} \right\},$$

where $\hat{z}_t = E(z_t | z_1, \ldots, z_{t-1})$ is the one-step ahead prediction of z_t and $\Delta_t = Var(z_t - \hat{z}_t)$ is the one-step predictor error variance. If the changepoint number c and their time locations $\tau_1, \ldots, \tau_c$ are given, the log-likelihood function for the GEV and ARFIMA parameters $\boldsymbol{\eta}$ is calculated by

$$
\begin{aligned}
\ell(\boldsymbol{\eta}) = &-\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{n} \ln \Delta_t - \frac{1}{2} \sum_{t=1}^{n} \frac{(z_t - \hat{z}_t)^2}{\Delta_t} \\
&+ \sum_{t=1}^{n} \ln g_t(y_t) - \sum_{t=1}^{n} \ln \phi(\Phi^{-1}(G_t(y_t))),
\end{aligned}
\tag{3.5}
$$

with an observed extreme series $\{y_1, \ldots, y_n\}$. Here, $z_t = \Phi^{-1}(G_t(y_t))$ from (3.2). We use a numerical optimizer to find the maximum likelihood estimates for $\boldsymbol{\eta}$ by maximizing the log-likelihood function.

However, the maximum likelihood estimator of the fractional differencing parameter d in the ARFIMA(p, d, q) model is often negatively biased, particularly when AR or MA components are involved (Smith, Taylor, and Yadav, 1997). Underestimation of d can produce biased standard errors of the parameter estimates which in turn would make the asymptotic confidence intervals based on these biased standard errors fail to maintain the intended coverage probability. Therefore, we use a bias-corrected and accelerated (BCa) bootstrap method (Efron, 1987) with moving block resampling (Künsch, 1989) to obtain more realistic standard errors and confidence intervals for the maximum likelihood estimates of the GEV and ARFIMA parameters $\boldsymbol{\eta}$ and return levels r_z.

3.4   A Simulation Study

Our simulation study evaluates the effect of long-memory autocorrelation in block maximum series on changepoint detection by assessing the performance of the long-memory copula-GEV likelihood in Section 3.3.1 under different long-memory autocorrelation settings. To reflect realistic features in a block maximum series with long-memory autocorrelation, we generate a simulated series from an autocorrelated GEV model based on the weekly maximum ozone series of Clark County in Nevada. The Clark County GEV model is discussed in detail in the next section. Two different changepoint scenarios are considered: no changepoints (Scenario A) and two changepoints (Scenario B). Scenario A estimates the false positive rates of the GA with long-memory copula-GEV likelihood when no changepoints are assumed. The true positive rates for the long-memory copula-GEV likelihood are estimated in Scenario B under a two-changepoint setting.

The data generation scheme consists of the following two stages. In the first stage, we generate a stationary standard normal ARIFIMA$(0, d, 1)$ base process $\{Z_1, \ldots, Z_n\}$ with $n = 1825$ (≈ 35 years) under the four different long-memory autocorrelation cases: (a) $(d = 0.1, \theta = 0)$ for a weak long-memory autocorrelation, (b) $(d = 0.25, \theta = 0)$ for moderate long-memory autocorrelation, (c) $(d = 0.4, \theta = 0)$ for a strong long-memory autocorrelation, and (d) $(d = 0.4, \theta = 0.4)$ for a case where a strong long-memory autocorrelation is present with a moderate short-memory moving-average autocorrelation. We note that for the copula-GEV model in Section 3.3.1, the white noise variance for $\{Z_t\}$ was parameterized so that the base process has a unit variance. In the second stage, a long-memory autocorrelated copula-GEV series $\{Y_1, \ldots, Y_n\}$ is obtained by the transformation $Y_t = G_t^{-1}(\Phi(Z_t))$ in (3.2) with the following non-stationary GEV parameters:

$$\mu_t = \beta_0 + \delta_t^{(0)} - 0.017 \cos\left(\frac{2\pi t}{T}\right) - 0.001 \sin\left(\frac{2\pi t}{T}\right) - 0.002 \cos\left(\frac{2\pi t}{T/2}\right) + 0.01\left(\frac{t}{100T}\right),$$

$$\ln \sigma_t = \omega_0 + \delta_t^{(1)} - 0.28 \cos\left(\frac{2\pi t}{T}\right) - 0.08 \sin\left(\frac{2\pi t}{T}\right) - 0.04 \cos\left(\frac{2\pi t}{T/2}\right) + 0.05 \sin\left(\frac{2\pi t}{T/2}\right),$$

and $\xi = -0.1$. Here, t denotes time in week and $T = 365.25/7$ is selected for the periodicity in weekly maximum ozone series. The shift terms for location and scale parameters are set as follows. For Scenario A, we set $\beta_0 = 0.05$, $\omega_0 = -5.0$, and $\delta_t^{(0)} = \delta_t^{(1)} = 0$. For Scenario B, we set $\beta_0 = 0.07$, $\omega_0 = -4.0$, and assume two changepoints at $t \in \{100, 600\}$ with their associated location and scale shifts as

$$\delta_t^{(0)} = \begin{cases} 0, & \text{if } t = 1, \ldots, 99; \\ -0.015, & \text{if } t = 100, \ldots, 599; \\ -0.020, & \text{if } t = 600, \ldots, 1825, \end{cases} \qquad \delta_t^{(1)} = \begin{cases} 0, & \text{if } t = 1, \ldots, 99; \\ -0.8, & \text{if } t = 100, \ldots, 599; \\ -1.4, & \text{if } t = 600, \ldots, 1825. \end{cases}$$

Figure 3.1 shows generated ARFIMA copula-GEV series for each of the four differ-ent autocorrelation cases under Scenario A (no changepoints) and Scenario B (two changepoints) with the true changepoint times displayed by purple vertical lines. As expected, the generated log-memory copula-GEV series could exhibit spurious changepoints more often with larger d.

For each generated ARFIMA copula-GEV series, we estimate the number of changepoints and their associated time locations via the GA. To understand how long-memory autocorrelation in the generated ARFIMA coupla-GEV series affects the accuracy of changepoint detection, we compare the following three different autocor-relation models: our long-memory ARFIMA copula-GEV model, the short-memory AR(1) copula-GEV model in Zhu, Liu, and Lund (2019), and the autocorrelation-ignored copula-GEV model.

Table 3.1 shows the estimated number of changepoints under Scenario A (no changepoints). When a copula-GEV series features weak long-memory autocorrela-tion with $d = 0.1$, the GAs using the autocorrelated ARFIMA and AR(1) copula-GEV models perform comparably, correctly identifying no changepoints 96–97% of the times. As d gets larger, however, the AR(1) copula-GEV GA method produces a higher false positive rate, detecting spurious changepoints more frequently. We find that the ARFIMA copula-GEV GA method performs well across all four cases, con-sistently maintaining 92–99% accuracy. On the contrary, the autocorrelation-ignored copula-GEV GA method performs the worst, detecting a large number of spurious changepoints with larger d.

Table 3.2 summarizes the estimated number of changepoints for the Scenario B (two changepoints). As in Scenario A, both long-memory and short-memory GAs perform similarly when the long-memory autocorrelation is weak with $d = 0.1$,
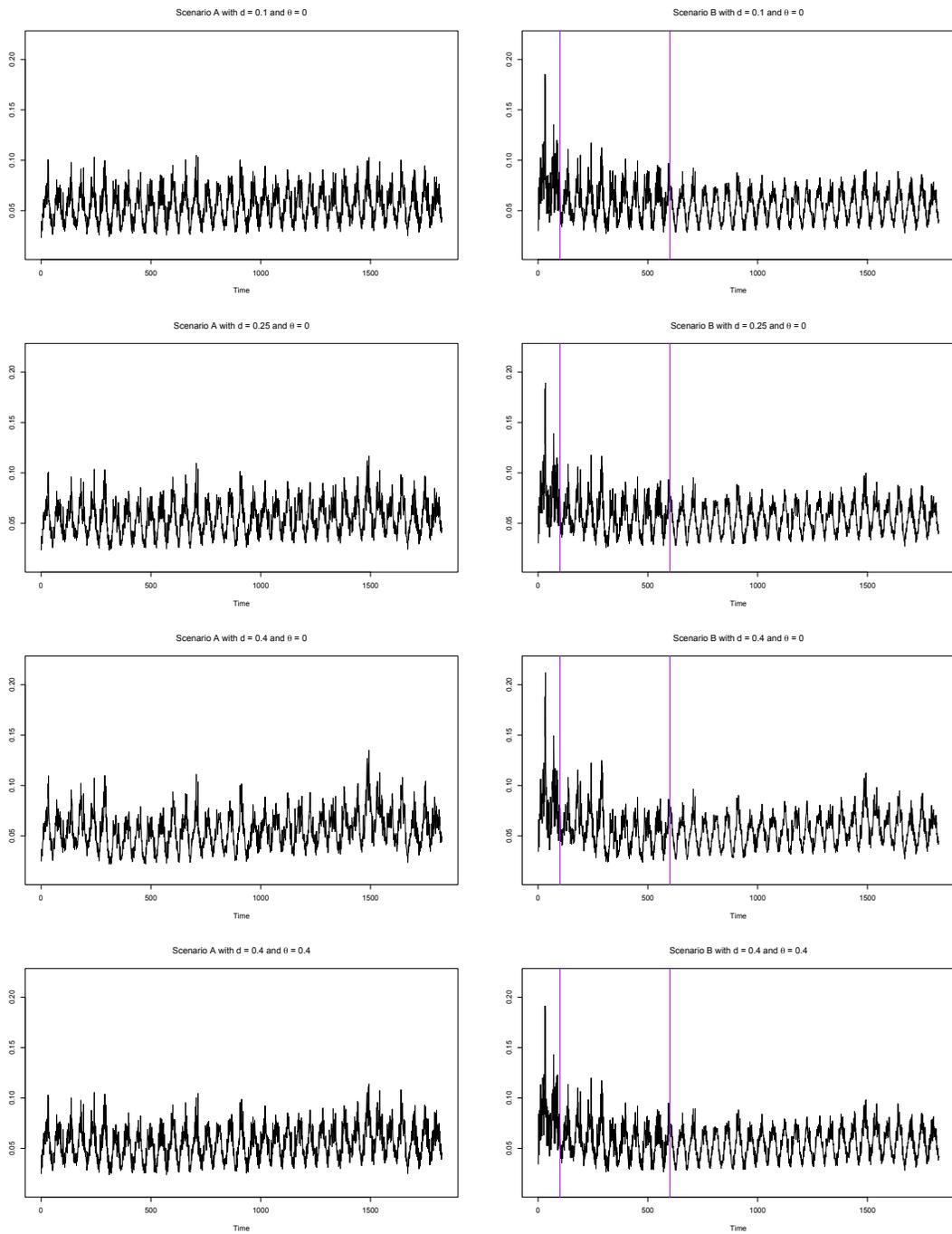
Figure 3.1: Autocorrelated copula-GEV series generated with four different autocorrelation cases under Scenario A (left: no changepoints) and Scenario B (right: two changepoints)

Table 3.1: GA detection summary for three copula-GEV models on Scenario A

Case	Base process	$\bar{\hat{c}}$	$\hat{c}=0$	$\hat{c}=1$	$\hat{c}=2$	$\hat{c}=3$	$\hat{c}=4$	$\hat{c}=5+$
$d=0.1$ $\theta=0$	ARFIMA$(0,d,0)$	0.04	**96**	4	0	0	0	0
	AR(1)	0.03	**97**	3	0	0	0	0
	Autocorr. Ignored	0.15	**86**	13	1	0	0	0
$d=0.25$ $\theta=0$	ARFIMA$(0,d,0)$	0.01	**99**	1	0	0	0	0
	AR(1)	0.44	**69**	18	13	0	0	0
	Autocorr. Ignored	1.60	**18**	29	34	13	6	0
$d=0.4$ $\theta=0$	ARFIMA$(0,d,0)$	0.08	**92**	8	0	0	0	0
	AR(1)	1.39	**25**	29	30	14	2	0
	Autocorr. Ignored	6.91	**0**	2	2	2	8	86
$d=0.4$ $\theta=0.4$	ARFIMA$(0,d,1)$	0.04	**96**	4	0	0	0	0
	AR(1)	2.26	**7**	20	31	28	10	4
	Autocorr. Ignored	2.36	**5**	19	38	20	10	8

Table 3.2: GA detection summary for three copula-GEV models on Scenario B

Case	Base process	$\bar{\hat{c}}$	$\hat{c}\leq 1$	$\hat{c}=2$	$\hat{c}=3$	$\hat{c}=4$	$\hat{c}=5$	$\hat{c}=6+$
$d=0.1$ $\theta=0$	ARFIMA$(0,d,0)$	2.01	0	**99**	1	0	0	0
	AR(1)	2.03	0	**97**	3	0	0	0
	Autocorr. Ignored	2.15	0	**86**	13	1	0	0
$d=0.25$ $\theta=0$	ARFIMA$(0,d,0)$	2.01	0	**99**	1	0	0	0
	AR(1)	2.44	0	**67**	24	7	2	0
	Autocorr. Ignored	3.45	0	**21**	38	23	11	7
$d=0.4$ $\theta=0$	ARFIMA$(0,d,0)$	2.01	0	**99**	1	0	0	0
	AR(1)	3.30	0	**25**	41	19	11	4
	Autocorr. Ignored	7.59	0	**2**	1	6	8	83
$d=0.4$ $\theta=0.4$	ARFIMA$(0,d,1)$	2.01	0	**99**	1	0	0	0
	AR(1)	3.71	0	**15**	37	23	15	10
	Autocorr. Ignored	4.11	0	**9**	30	20	26	15

correctly identifying two changepoints in nearly all repetitions. However, as the long-memory autocorrelation in a simulated copula-GEV series gets stronger, the AR(1) GA substantially overestimates c. We find that our ARFIMA copula-GEV GA performs very well in all cases, consistently attaining 99% accuracy in identifying two changepoints. Along with the results from Scenario A, this finding suggests that the long-memory copula-GEV likelihood should be used to accurately detect changepoints in autocorrelated copula-GEV series with long memory.

3.5   Case Study Analysis

This section illustrates our methods by analyzing two county-level weekly maximum ozone concentration series in Clark County, Nevada and Doña Ana County, New Mexico. Clark County is an urban county where most of its residents live in the Las Vegas area, one of the fastest growing U.S. cities in the last few decades. Doña Ana County is adjacent to El Paso, Texas in the U.S. and Ciudad Juárez in Mexico, a region where air pollution is of particular concern due to its frequent exceedance of the NAAQS standards and the trans-boundary flow of pollutants between the U.S. and Mexico (cf. Shi, Fernando, and Yang, 2009).

We apply the GA method with the long-memory copula-GEV likelihood, as illustrated in Section 3.3.1, to the weekly maximum series of these two counties. We parameterize the GEV location and scale parameters as described in Section 3.3.2 to include a periodic functions with two harmonics ($K^{(0)} = K^{(1)} = 2$). The period of each series is selected to be $T = 365.25/7$ to explain annual cycles in the extreme ozone series. The long-term linear trend α is posited to represent the expected change in the weekly maxima over $L = 100$ years. In addition, our preliminary analysis found long-memory autocorrelation characteristics with sample autocorrelations decaying

slowly and short-memory autocorrelation features with week-to-week autocorrelations. Therefore, we consider the ARFIMA$(0, d, 1)$ model to characterize long and short memory autocorrelations present in weekly maximum ozone series. Specifically, the differencing parameter d captures long-memory autocorrelation and the first-order MA parameter θ models short-memory autocorrelation. Although a higher order for AR and MA components can be considered, we found the ARFIMA$(0, d, 1)$ model to offer a good balance between parsimony and model complexity in characterizing the long and short memory autocorrelations in weekly maximum ozone series.

Now, we estimate changeppoints in the weekly maximum ozone series. For the Clark County series, the GA method estimates two changepoints at the 18-th week of 1982 and 40-th week of 1991. Using these two changepoints, we fit the long-memory copula-GEV model with an ARFIMA$(0, d, 1)$ base process using the Kalman truncation with $m = 52$. The standard errors of the GEV and ARFIMA parameter estimates are calculated using the moving block bootstrap of Künsch (1989) with a block size of 157 (≈ 3 years). Table 3.3 summarizes the GEV and ARFIMA model parameter estimates for the Clark County weekly maximum ozone series along with their corresponding bootstrap standard errors. The fractional differencing parameter d is estimated to be 0.396, indicating that the Clark County series exhibits strong long-memory autocorrelation. Our estimate for the long-term linear trend α is 0.019 ppm century^{-1} with a 95% BCa bootstrap confidence interval of $(0.011, 0.034)$. This result suggests that the weekly maximum ozone series has been significantly increasing over time in Clark County.

For the Doña Ana County weekly maximum ozone series, the GA method detects four changepoints at the 17-th week of 1981, 28-th week of 1985, 32-nd week of 1987, and 14-th week of 1997. Using these four GA estimated changepoints, we fit the long-

Table 3.3: Estimated GEV and ARFIMA parameters and their associated bootstrap standard errors for the weekly maximum ozone series in Clark County (units: ppm for β's and Δ's; ppm century^{-1} for α)

Parameters	Estimates (SE)	Parameters	Estimates (SE)
β_0	0.0711 (0.0023)	ω_0	-3.9846 (0.0961)
β_1	-0.0170 (0.0009)	ω_1	-0.2978 (0.0354)
β_2	-0.0003 (0.0004)	ω_2	-0.0768 (0.0236)
β_3	-0.0019 (0.0002)	ω_3	-0.0534 (0.0239)
β_4	0.0003 (0.0003)	ω_4	0.0528 (0.0290)
$\Delta_1^{(0)}$	-0.0144 (0.0028)	$\Delta_1^{(1)}$	-0.7869 (0.0987)
$\Delta_2^{(0)}$	-0.0206 (0.0028)	$\Delta_2^{(1)}$	-1.3552 (0.0992)
α	0.0191 (0.0066)	d	0.3957 (0.0404)
ξ	-0.1189 (0.0161)	θ	0.0561 (0.0625)

Table 3.4: Estimated GEV and ARFIMA parameters and their associated bootstrap standard errors for the weekly maximum ozone series in Doña Ana County (units: ppm for β's and Δ's; ppm century^{-1} for α)

| Parameters	Estimates (SE)	Parameters	Estimates (SE)
β_0	0.0500 (0.0021)	ω_0	-4.1648 (0.0786)
β_1	-0.0122 (0.0003)	ω_1	-0.3095 (0.0354)
β_2	-0.0002 (0.0002)	ω_2	-0.1649 (0.0264)
β_3	-0.0022 (0.0003)	ω_3	-0.0777 (0.0221)
β_4	0.0005 (0.0002)	ω_4	0.0370 (0.0254)
$\Delta_1^{(0)}$	0.0053 (0.0021)	$\Delta_1^{(1)}$	-0.7486 (0.0776)
$\Delta_2^{(0)}$	0.0048 (0.0026)	$\Delta_2^{(1)}$	-0.2163 (0.0859)
$\Delta_3^{(0)}$	0.0096 (0.0026)	$\Delta_3^{(1)}$	-0.6877 (0.0858)
$\Delta_4^{(0)}$	0.0082 (0.0038)	$\Delta_4^{(1)}$	-1.1221 (0.0868)
α	-0.0043 (0.0120)	d	0.2034 (0.0315)
ξ	-0.1780 (0.0182)	θ	0.0450 (0.0267)
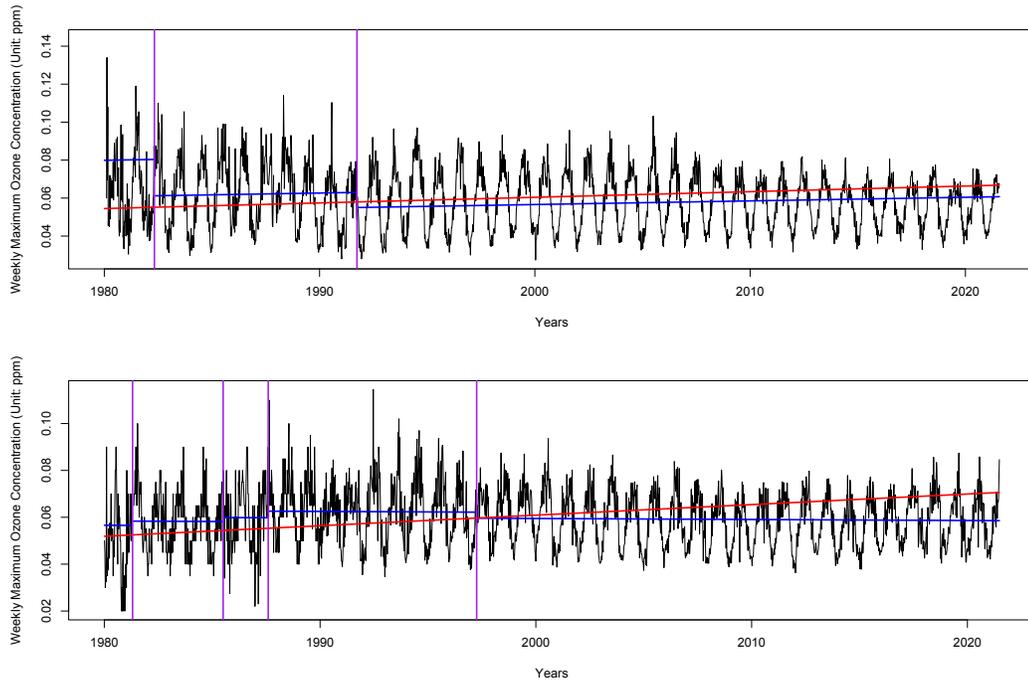
Figure 3.2: Estimated trend lines for the weekly maximum ozone series in Clark County (top) and Doña Ana County (bottom) (blue solid lines, the long-memory copula-GEV trend with corresponding GA estimated changepoint times; red solid line, the long-memory copula-GEV trend without allowing changepoints)

memory copula-GEV model to the Doña Ana County series. The estimated model parameters along with their corresponding bootstrap standard errors are summarized in Table 3.4. The long-memory parameter d is estimated to be 0.203, indicating that there exists moderately persisting long-memory autocorrelation. The long-term linear trend α is estimated to be -0.004 ppm century^{-1} with a 95% BCa bootstrap confidence interval of $(-0.024, 0.024)$, showing an insignificant long-term trend.

To demonstrate the necessity of considering changepoints in extreme ozone analysis, we now compare the estimated trend lines from the long-memory copula-GEV model with changepoints considered and with changepoints ignored. Figure 3.2 shows the time plots of the Clark County and Doña Ana County weekly maximum ozone

series with their corresponding GA estimated changepoints denoted in purple vertical lines. The estimated long-term linear trends with GA changepoints considered and with changepoints ignored are displayed in blue and red solid lines respectively for each series. The estimated trend line with changepoints appears to better explain long-term trend in the series than the trend line estimated with changepoints ignored. Further, the MDL is substantially improved when changepoints were considered for both series. To elaborate, the long-memory copula-GEV model with changepoints ignored returned the MDL of -9774.947 for the Clark County series and -9560.231 for the Doña Ana County series, whereas the long-memory copula-GEV model with GA changepoints yielded the MDL of -9869.718 and -9745.098, respectively. These results suggest that consideration of changepoints could significantly improve a model fit for the weekly maximum ozone series.

Table 3.5: **The MDL and estimated long-term trend α with 95% BCa bootstrap confidence interval (unit: ppm century^{-1} for α) from three autocorrelated copula-GEV models considering changepoints for Clark County and Doña Ana County**

	Clark County		Doña Ana County	
	MDL	Long-term trend	MDL	Long-term trend
ARFIMA$(0, d, 1)$	-9869.718	0.0191 (0.0110, 0.0339)	-9745.098	0.0043 (-0.0244, 0.0238)
AR(1)	-9819.809	-0.0060 (-0.0126, 0.0016)	-9725.684	0.0062 (-0.0132, 0.0356)
Autocorr. Ignored	-7717.019	-0.0055 (-0.0134, 0.0008)	-7716.396	-0.0280 (-0.0427, -0.0030)

We revisit the three different autocorrelation models in Section 3.4 and compare their performances with Clark County and Dona Ana County. For the Clark County series, the AR(1) copula-GEV GA estimates three changepoints and the autocorrelation-ignored GA estimates five changepoints. For the Doña Ana County

series, the AR copula-GEV GA estimates five changepoints and the autocorrelation-ignored GA estimates six changepoints. Table 3.5 summarizes the MDL and the estimated long-term linear trends from the three autocorrelated GEV models with their corresponding GA changepoints. For both series, the MDL from the ARFIMA copula-GEV model is the most desirable, suggesting that the copula-GEV model with an ARFIMA$(0, d, 1)$ base process offers a better fit to those series than the other two copula-GEV models. Combined with the earlier finding that both extreme ozone series exhibit moderate to strong long-memory autocorrelation, these results suggest that additional changepoints estimated by the AR copula-GEV GA and autocorrelation-ignored GA are likely to be spurious, consequently suggesting that the estimated trends from these two copula-GEV models are likely inaccurate. These findings support our claim that the long-memory copula-GEV model should be used to analyze a weekly maximum ozone series with long-memory autocorrelation.
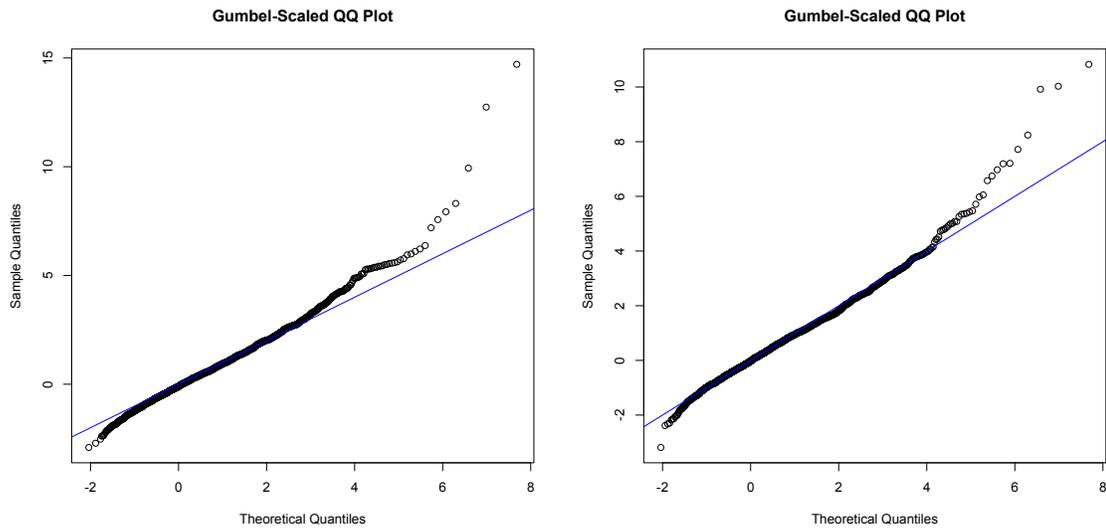

Figure 3.3: **Gumbel-scaled quantile-quantile plots for the ARFIMA copula-GEV model with corresponding GA estimated changepoints considered for the weekly maximum ozone series in Clark County (left) and Doña Ana County (right)**

Goodness-of-fit procedures for the copula-GEV model are performed to assess if the ARFIMA copula-GEV model with GA changepoints is an appropriate model to the Clark County and Doña Ana County series. Figure 3.3 shows the Gumbel-scaled quantile-quantile plots for the ARFIMA copula-GEV model to both series with their corresponding GA changepoints. Overall, the ARFIMA copula-GEV model with GA changepoints is a reasonable fit to the weekly maximum ozone series in both counties, because most sample quantiles match with their theoretical quantiles and only a very small portion ($< 0.4\%$) of the sample quantiles is greater than their theoretical counterparts in the upper tail.
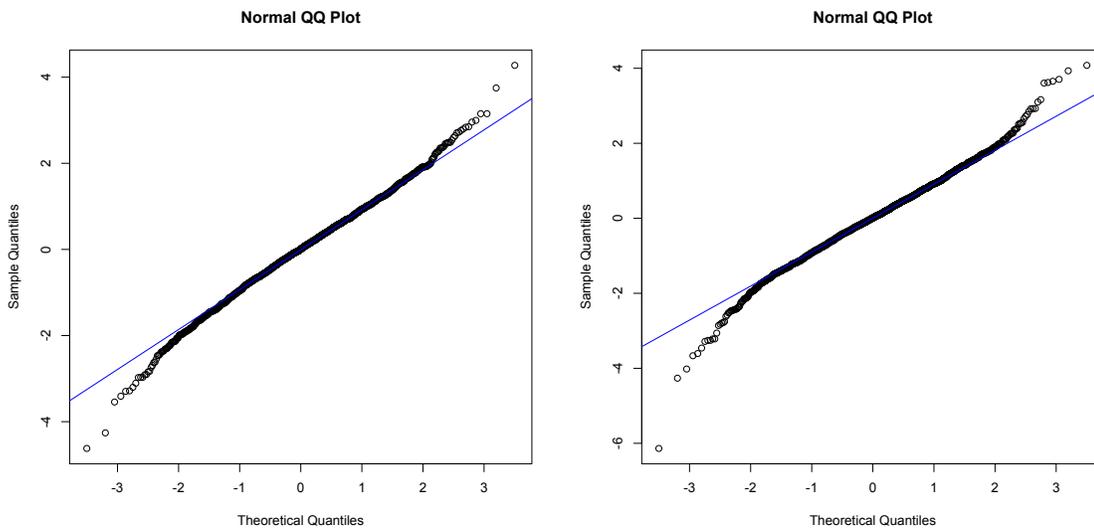
Figure 3.4: Normal quantile-quantile plots for the estimated white noise process for weekly maximum ozone series in Clark County (left) and Doña Ana County (right)

Next, we assess our time series model choice for the base process $\{Z_t\}$. If the ARFIMA$(0, d, 1)$ model adequately characterizes the long-memory autocorrelation in a weekly maximum ozone series, the estimated white noise process $\{\hat{\varepsilon}_t\}$ should be approximately normal with zero mean and nearly uncorrelated. Figure 3.4 shows

the normal quantile-quantile plots for the estimated white noise processes, showing that $\{\hat{\varepsilon}_t\}$ appears to be approximately normal for the weekly maximum ozone series in Clark County and Doña Ana County, with only a small portion ($< 0.2\%$) of the sample quantiles indicating slight heavy tails. The sample autocorrelations of the estimated white noise process $\{\hat{\varepsilon}_t\}$ are shown in Figure 3.5. The autocorrelations are fairly weak over all the lags, except a very small negative lag-1 autocorrelation for the Clark County series. This indicates that the ARFIMA$(0, d, 1)$ model is an effective choice for the long-memory autocorrelation in the Clark County and Doña Ana County extreme ozone series without a serious issue for the model fit.
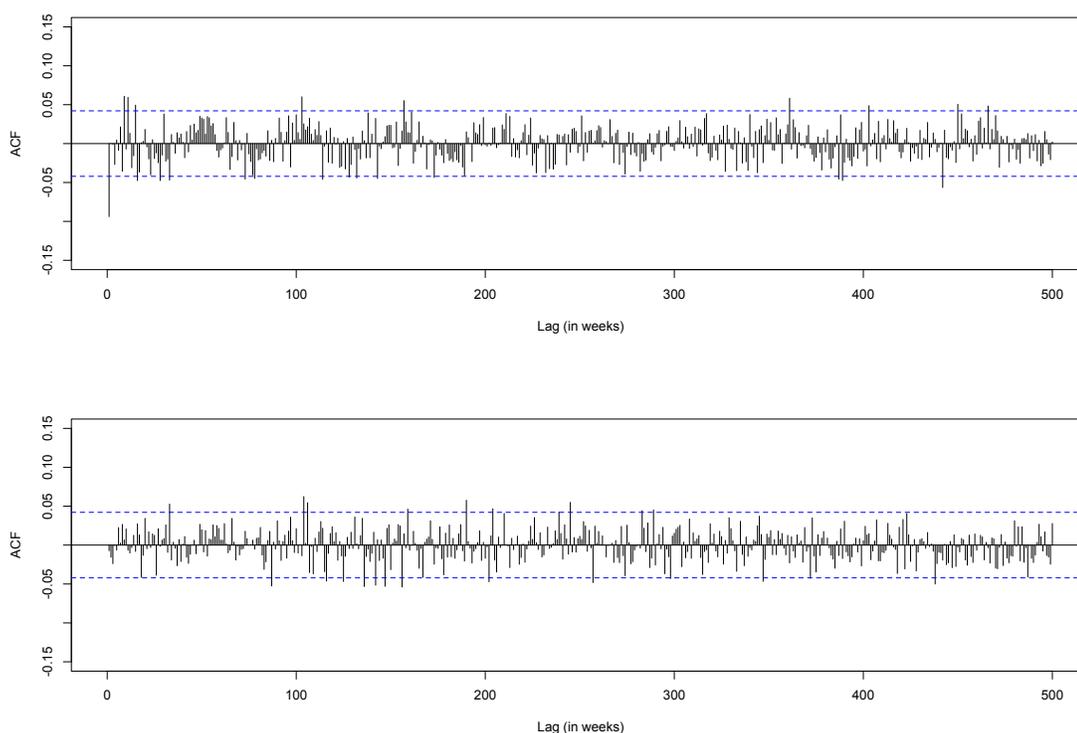

Figure 3.5: Sample autocorrelation function of the estimated white noise process for the weekly maximum ozone series in Clark County (top) and Doña Ana County (bottom)

We now estimate the 5-, 10-, and 20-year return levels for the weekly maximum ozone concentrations in Clark County and Doña Ana County. We selected the initial period for return levels t_I to be the first week of 2023, so that the z-year return level is interpreted as the lowest weekly maxima that is expected to be exceeded once in a z-year period starting from 2023. We then calculated 95% confidence intervals for return level estimates using a BCa bootstrap method with moving-block resampling and delete-1 jackknife. Table 3.6 summarizes the estimated weekly maximum ozone return levels and their 95% BCa bootstrap confidence intervals. Weekly maximum ozone concentrations in Clark and Doña Ana Counties are expected to exceed 0.08–0.10 ppm at least once in 5–10 years. With the EPA recommended threshold for ozone pollution being 0.07 ppm, we project that both counties are under a risk of unhealthy levels of extreme ozone events in the near future.

Table 3.6: Weekly maximum ozone return level and their 95% BCa bootstrap confidence intervals in parentheses for Clark County and Doña Ana County (unit: ppm)

Return level	Clark County	Doña Ana County
r_5	0.0943 (0.0874, 0.1000)	0.0862 (0.0826, 0.0911)
r_{10}	0.0973 (0.0900, 0.1037)	0.0882 (0.0841, 0.0938)
r_{20}	0.1007 (0.0927, 0.1082)	0.0899 (0.0847, 0.0966)

3.6 U.S. Extreme Ozone Long-term Trends

Our methods were applied to the weekly maximum ozone concentration series at all 395 counties selected for this study. The GA method with our long-memory copula-GEV likelihood estimated no changepoints in 112 counties (28.35%), one changepoint in 171 counties (43.29%), two changepoints in 84 counties (21.27%), and three changepoints in 18 counties (4.56%), and four or more changepoints in 10

counties (2.53%) with a mean of 1.109 changepoints. Although many counties in the Midwestern and Northeastern U.S. have experienced no changepoints, nearly all coastal counties have had at least one changepoint. The GA using short-memory AR(1) copula-GEV likelihood and autocorrelation-ignored copula-GEV likelihood were also applied for comparison purposes. The short-memory copula-GEV GA detected a mean of 1.828 changepoints, and the autocorrelation-ignored GA found on average 2.851 changepoints across all 395 counties. The summary of results from these three GAs are in the supplementary materials.
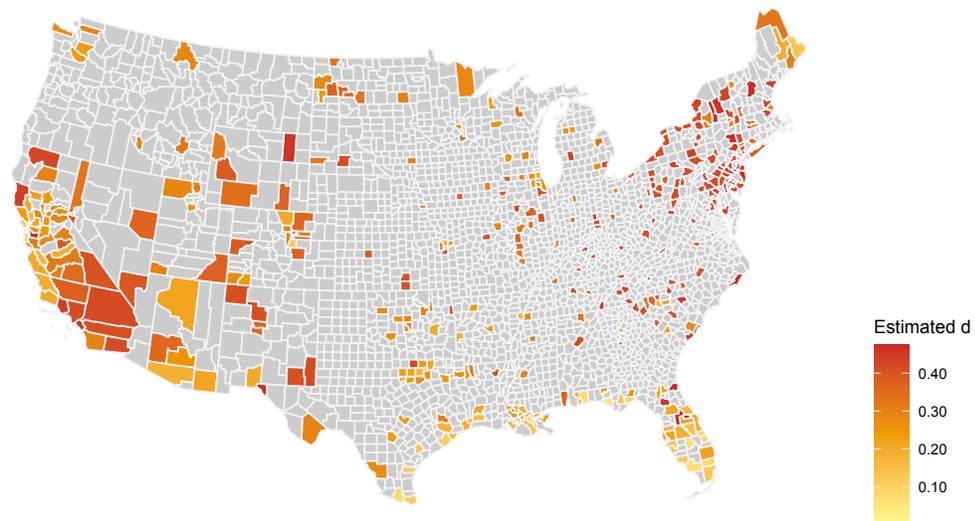
Map 3.2: Spatial patterns of the long-memory parameter estimates in the ARFIMA copula-GEV model

With the changepoints estimated by the long-memory copula-GEV GA, we estimate the GEV and ARFIMA parameters by maximizing the long-memory copula-GEV likelihood in (3.5) for all weekly maximum series in the selected 395 counties. The maximum likelihood estimates of the long-memory parameter d range from 0.009

to 0.475 with a median of 0.332. Map 3.2 depicts a spatial map of these long-memory parameter estimates. Although the weekly maximum ozone series overall exhibit moderate to strong long-memory autocorrelation, we find different patterns across the regions. For instance, most counties in the Gulf Coast show relatively weak long-memory autocorrelation with most d estimates between 0.1 and 0.2. However, those counties located in the Northeastern U.S. and southern California exhibit the strongest long-memory autocorrelation with the estimated d around 0.4. Because the changepoint methods based on the short-memory or autocorrelation-ignored copula-GEV likelihood could find spurious changepoints when the long-memory autocorrelation is strong, our results on the d estimates can be informative for identifying those counties with spurious changepoints and inaccurate long-term trend estimates.
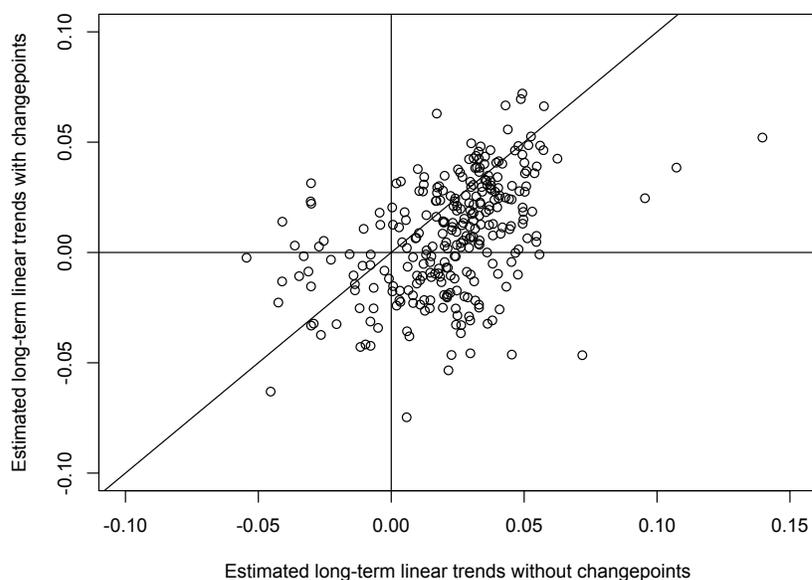

Figure 3.6: A scatter plot of the estimated long-term trends with changepoints considered against the estimated trends with changepoints ignored for the 283 counties with at least one GA estimated changepoint (unit: ppm century^{-1})

We now examine how the consideration of changepoints influences the long-term trend estimates in U.S. weekly maximum ozone series. When changepoints are not considered in the modeling process, the long-term trend estimates across all 395 counties have a mean of 0.020 ppm century^{-1} and a standard deviation of 0.024 ppm century^{-1}. However, our trend estimates with changepoint consideration have a mean of 0.009 ppm century^{-1} and a standard deviation of 0.025 ppm century^{-1}, showing that the consideration of changepoints has reduced the average estimated trends. Figure 3.6 shows a scatter plot of the long-term trends estimated with and without changepoint consideration for the 283 counties that have at least one GA estimated changepoint. Further, it was observed that about a third of the counties had their estimated trends change signs once changepoints were considered. This finding supports our claim that changepoints must be considered to accurately quantify long-term trends in weekly maximum ozone concentrations.
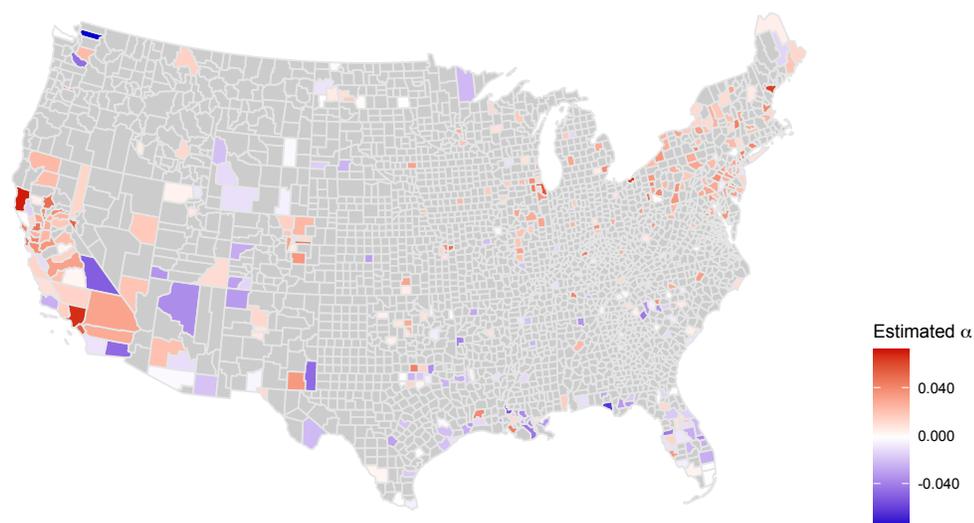
Map 3.3: Spatial map of the estimated long-term trends in U.S. weekly maximum ozone concentrations (unit: ppm century^{-1})
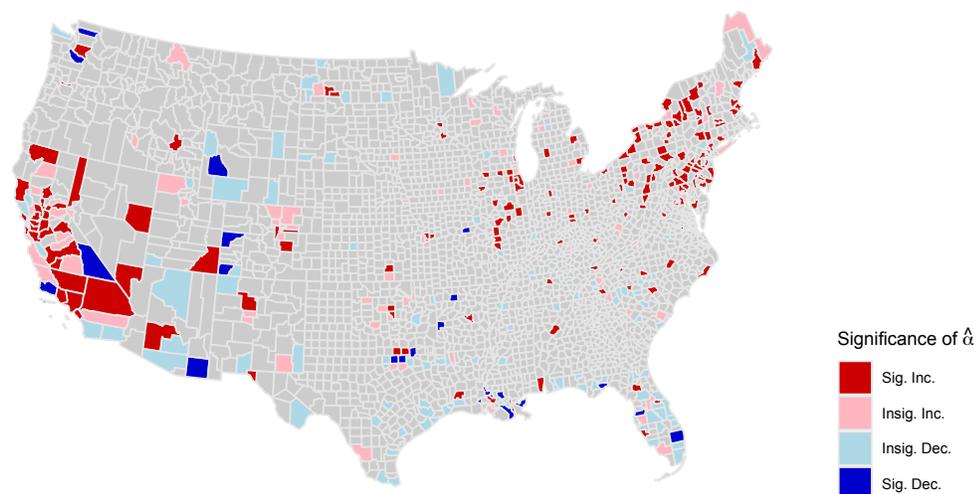
Map 3.4: Significance map of the estimated U.S. long-term trends based on their 95% BCa bootstrap confidence intervals. The U.S. counties are colored with red if their estimated long-term trends are significantly larger than zero, light red if not significantly different from zero but positive, light blue if not significantly different from zero but negative, and blue if significantly smaller than zero.

Map 3.3 shows a geographical map of the estimated long-term trends from U.S. weekly maximum ozone series. Significance of the all long-term trend estimates based on their 95% BCa bootstrap confidence intervals is displayed in Map 3.4. Overall, the weekly maximum ozone concentrations in the Midwestern and Northeastern U.S. have significantly increased over time. Nearly all counties in California and Nevada also show increasing trends, with many of these trends being statistically significant. On the contrast, many counties in the southern parts of Wyoming and South Dakota, the Gulf Coast, Southeastern U.S., and U.S.-Mexico border show insignificantly decreasing trends.

Lastly, we estimate the weekly maximum ozone concentration 10-year return level r_{10} for all 395 counties and summarize the results in Map 3.5. Many counties in the Northeastern U.S. and coastal counties in Texas and Louisiana appear to be under the highest risk of future extreme ozone events with their weekly maximum ozone expected to exceed over 0.15 ppm, more than two times higher than the EPA threshold 0.07 ppm, at least once between 2023 and 2032. On the other hand, many counties in Florida, Arizona, and Wyoming along with the Californian coasts and U.S.-Canada border have relatively lower estimated 10-year return levels with extreme ozone events projected to be around 0.07–0.10 ppm at least once in the 10-year period. Lastly, although the estimated r_{10} varies greatly across the regions, ranging from 0.066 to 0.195 ppm, we find that almost all counties are projected to experience unhealthy level of ozone concentrations in the near future.
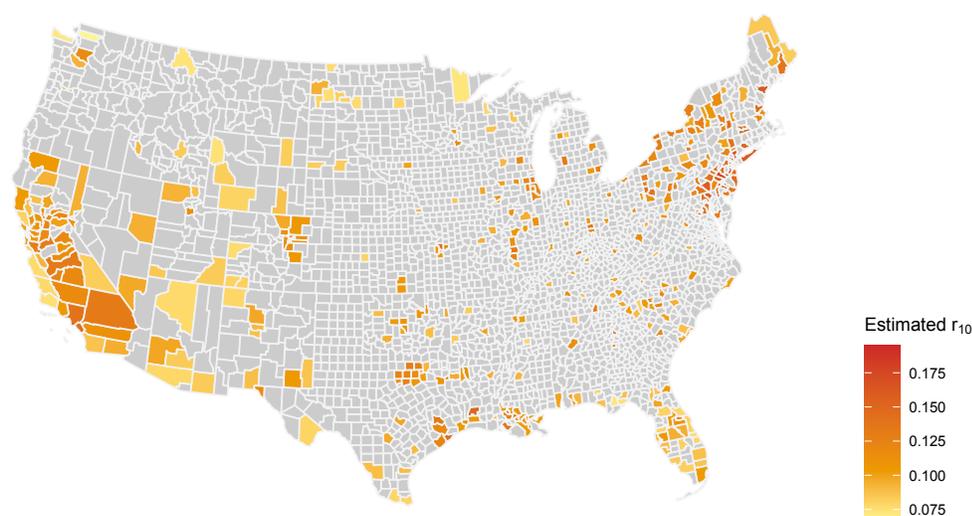

Map 3.5: The estimated 10-year return ozone concentration level r_{10} (unit: ppm)

3.7   Closing Comments

In this paper, we proposed extending the short-memory copula-GEV likelihood developed by Zhu, Liu, and Lund (2019) to block maximum series with long-memory autocorrelation to be used in the genetic algorithm for changepoint detection in long-memory autocorrelated block maximum series. Our extension allows the autocorrelation in block maximum series to be modeled using a long-memory time series model, such as ARFIMA(p, d, q) model. In a simulation study, we showed that if a block maximum series has long-memory autocorrelation, GA methods for changepoint detection without long-memory consideration tend to detect spurious changepoints. On the other hand, our GA method using long-memory copula-GEV likelihood performed quite well in correctly identifying changepoints (or lack thereof) in all levels of long-memory autocorrelation considered. In a case study with applications to weekly maximum ozone concentrations from Clark County in Nevada and Doña Ana County in New Mexico, our long-memory copula-GEV model offers a reasonable fit to the extreme ozone data in those counties as long as changepoints are considered.

We find that weekly maximum ozone concentrations have increased in many parts of the contiguous U.S., particularly the Northeastern, Midwestern, and Southwestern U.S. However, many coastal counties in the Southeastern U.S. had their weekly maximum ozone concentrations decrease over time. These are overall consistent with the findings of Phalitnonkiat et al. (2016) and Shen, Mickley, and Gilleland (2016). Lastly, increasing trends were found in most of California, metropolitan areas of Chicago and Denver, and coastal regions in New England. We note that these areas are also designated by the EPA to have not met the NAAQS ozone emissions requirements ("nonattainment areas") in recent years (US EPA, 2022).

Our GA method detected at least one changepoint in 283 counties (71.65% of 395 counties) and consideration of these GA estimated changepoints greatly impacted the long-term trend estimates in many of these counties. Map 3.6 shows the location of counties where the consideration of changepoints had substantial impacts to their long-term trend estimates. Of those 283 counties with at least one changepoint, 27.56% (78 counties) had their trend estimates changed by more than 0.03 ppm century^{-1}. Most of these counties appear to be fairly scattered around the country, except for a cluster around the Texas-Louisiana coastline.
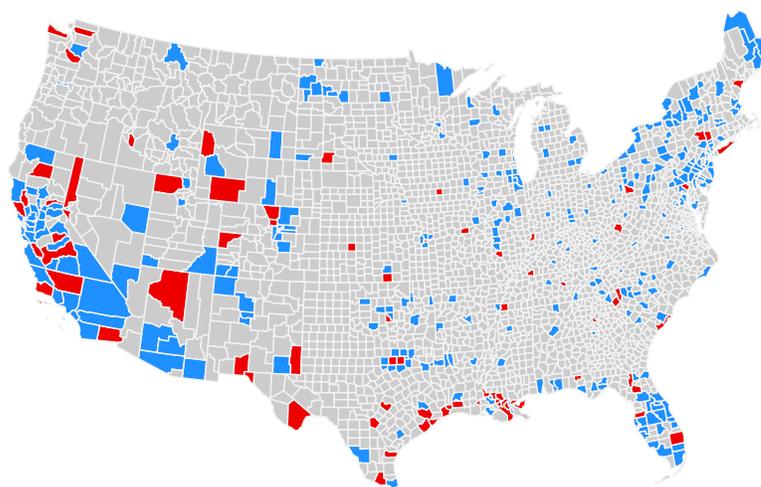
Map 3.6: Counties where the consideration of changepoints changed the long-term trend estimates by more than 0.03 ppm century^{-1} (red) and less than 0.03 ppm century^{-1} (blue)

The copula-GEV likelihood approach requires an appropriate time series model to be determined for base process. Although the ARFIMA(p, d, q) model was used to illustrate our long-memory copula-GEV likelihood, a different long-memory time series model can be considered. In addition, although we applied our methods to weekly maximum ozone concentrations, other environmental extremes with long-

memory autocorrelation can be analyzed using the long-memory copula-GEV likelihood. Lastly, our method could be further modified for other long-memory autocorrelated environmental series (not necessarily extremes) by substituting the marginal GEV distribution in (3.2) with another suitable probability distribution.

Although many areas in the contiguous U.S. were well represented in our study, some portions of the U.S. were sparsely represented due to limited data availability, particularly the Northwestern and Midwestern U.S. This limited spatial coverage restricts the scope of our analysis. It would be possible to have improved coverage in currently underrepresented regions when more data becomes available in the future.

Acknowledgments

CHAPTER 4

CONCLUDING REMARKS

4.1   Summary Discussion

This dissertation rigorously estimated long-term trends and return levels of two different extreme environmental events: monthly maximum coastal sea levels and weekly maximum ozone concentrations. Due to changes in measuring location, instrument, observer, sampling protocol, local ecosystem etc., these extreme environmental data were found to often contain undocumented inhomogenous shifts in their distributions. For accurate modeling of these extreme environmental events, we used a genetic algorithm to estimate the number and times of these changepoints from the data. These GA estimated changepoints were then used to quantify their long-term trends and return levels.

In Chapter 2, we examined monthly maximum sea levels from coastal regions around the world. Although the monthly maximum sea level series show weak autocorrelation in most locations, raw hourly sea level series were found to exhibit strong and long-lasting autocorrelation. In this case, naïve application of extreme value methods with independence assumption could result in substantial estimation bias in parameter estimates. Therefore, we incorporated the extremal index parameter, measuring the strength of autocorrelation in the raw hourly series, in our extreme value model to correct the estimation bias in parameter estimates. Our analysis

found that the consideration of changepoints changed the estimated linear trends of 89 tide gauges (approximately 30% of tide gauges considered) by more than 20 cm century^{-1}. The consideration of autocorrelation via the incorporation of extremal index substantially influenced return level estimates of 73 tide gauges (approximately 24% of tide gauges considered) by changing their estimated 50-year return levels by more than 30 cm.

In Chapter 3, we analyzed weekly maximum ozone concentrations from the contiguous United States at a county level. Unlike monthly maximum sea level series which exhibited weak autocorrelation in most locations, many weekly maximum ozone series were found to exhibit long-memory autocorrelation, requiring a new model for long-memory extreme value series. To effectively consider this long-memory autocorrelation in the modeling process, we extended the short-memory copula-GEV likelihood developed by Zhu, Liu, and Lund (2019) to a long-memory copula-GEV likelihood. We then used this long-memory copula-GEV likelihood to further develop a genetic algorithm based changepoint detection method for extreme value series with long-memory autocorrelation. In a simulation study, we found that our GA method with long-memory copula-GEV likelihood performs well in correctly identifying changepoint numbers in block maximum series with varying levels of long-memory autocorrelation, whereas the method without long-memory consideration were found to overestimate changepoint numbers in the presence of long-memory autocorrelation. Next, we used our GA method with long-memory copula-GEV likelihood to accurately estimate the number and times of changepoints in the weekly maximum ozone series. Overall increasing trends were found in weekly maximum ozone concentrations across many regions of the U.S. We also found that the consideration of changepoints in the modeling process changed long-term trend estimates of 78 counties (approximately

20% of counties considered) by more than 0.03 ppm century^{-1}. Lastly, our analysis revealed that nearly all counties are projected to experience unhealthy levels of ozone concentrations in the near future, exceeding the EPA recommended threshold of 0.07 ppm for ozone pollution.

4.2   Future Work

There are several future research topics that can arise from our work in this dissertation. First, the effectiveness of our GA methods for changepoint detection were assessed via simulation study under the assumption that model specification was done correctly. It would be interesting to assess the performance of our methods under simulation scenarios when the model is mis-specified.

Second, we focused on analyzing environmental extremes at an individual site level and our methods were subsequently developed to address issues arising from temporal correlation in the extreme value series. Extension of our methods to jointly model environmental extremes at multiple sites by incorporating a spatiotemporal method can be considered. This extension could make spatially out-of-sample forecasting possible and offer useful insights on possible geographical effects on extreme environmental events.

Third, extending our methods to extreme value series with clumping-at-zero can be beneficial because this extension could make our methods applicable to more diverse settings where zero-valued observations are prevalent. A discrete-continuous mixture model (cf. Weglarczyk, Strupczewski, and Singh, 2005; Couturier and Victoria-Feser, 2010; Hautsch, Malec, and Schienle, 2014; Harvey and Ito, 2020) can be considered in modifying our extreme value methods to allow for frequent zero-valued observations.

Last, the main focus of this dissertation was the analysis of extreme environmental data using statistical methods based on extreme value theory. Due to the interdisciplinary nature of environmental science, it would be interesting to consider collaborating with researchers from other disciplines for future studies. For extreme sea levels, some possible interdisciplinary collaborations could be projecting future extreme sea level events based on different climate change scenarios, examining effectiveness and making comparisons of different greenhouse gas reduction policies on extreme sea level rise, and quantifying economic or environmental damages of projected future extreme sea level events. For extreme ozone concentrations, possible interdisciplinary collaborations include considering additional air pollutants as covariates for ground-level ozone formation in the modeling process, examining joint relationships between the ozone concentration and its other covariates under long-memory autocorrelation, investigating long-term implications and effectiveness of air pollution control policies, and exploring epidemiological and economic impacts of extreme ozone events.

REFERENCES

Acero, F. J., García, J. A., and Gallego, M. C. (2011). Peaks-over-threshold study of trends in extreme rainfall over the Iberian Peninsula. *Journal of Climate*, 24:1089–1105.

Becker, M., Karpytchev, M., Davy, M., and Doekes, K. (2009). Impact of a shift in mean on the sea level rise: Application to the tide gauges in the Southern Netherlands. *Continental Shelf Research*, 29(4):741–749.

Bhandari, D., Murthy, C. A., and Pal, S. K. (1996). Genetic algorithm with elitist model and its convergence. *International Journal of Pattern Recognition and Artificial Intelligence*, 10(6):731–747.

Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer Science & Business Media, New York, NY, 2nd edition.

Chan, N. H. and Palma, W. (1998). State space modeling of long-memory processes. *The Annals of Statistics*, 26(2):719–740.

Chan, R. K. S. and So, M. K. P. (2018). Multivariate modelling of spatial extremes based on copulas. *Journal of Statistical Computation and Simulation*, 88(12):2404–2424.

Chen, X., Zhang, X., Church, J. A., Watson, C. S., King, M. A., Monselesan, D., Legresy, B., and Harig, C. (2017). The increasing rate of global mean sea-level rise during 1993-–2014. *Nature Climate Change*, 7:492–495.

Church, J. A. and White, N. J. (2011). Sea-level rise from the late 19th to the early 21st century. *Surveys in Geophysics*, 32:585–602.

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

Couturier, D.-L. and Victoria-Feser, M.-P. (2010). Zero-inflated truncated generalized Pareto distribution for the analysis of radio audience data. *The Annals of Applied Statistics*, 4(4):1824–1846.

Davis, R. A., Lee, T. C. M., and Rodriguez-Yam, G. A. (2006). Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association*, 101(473):223–239.

de Haan, L. and Ferreira, A. F. (2006). *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, NY.

Dupuis, D. J. (2005). Ozone concentrations: A robust analysis of multivariate extremes. *Technometrics*, 47(2):191–201.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82(397):171–185.

Fawcett, L. and Walshaw, D. (2012). Estimating return levels from serially dependent extremes. *Environmetrics*, 23(3):272–283.

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Annals of Statistics*, 42(6):2243–2281.

Harvey, A. and Ito, R. (2020). Modeling time series when some observations are zero. *Journal of Econometrics*, 214(1):33–45.

Hautsch, N., Malec, P., and Schienle, M. (2014). Capturing the zero: a new class of zero-augmented distributions and multiplicative error processes. *Journal of Financial Econometrics*, 12(1):89–121.

Hazarika, S., Borah, P., and Prakash, A. (2019). The assessment of return probability of maximum ozone concentrations inan urban environment of Delhi: A generalized extreme value analysis approach. *Atmospheric Environment*, 202:53–63.

Hewaarachchi, A. P., Li, Y., Lund, R., and Rennie, J. (2017). Homogenization of daily temperature data. *Journal of Climate*, 30(3).

Hoang, T. T. H., Parey, S., and Dacunha-Castelle, D. (2009). Multidimensional trends: The example of temperature. *The European Physical Journal Special Topics*, 174:113–124.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.

Jerrett, M., Burnett, R. T., 3rd, C. A. P., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., and Thun, M. (2009). Long-term ozone exposure and mortality. *The New England Journal of Medicine*, 360(11):1085–1095.

Jevrejeva, S., Moore, J. C., Grinsted, A., and Woodworth, P. L. (2008). Recent global sea level acceleration started over 200 years ago? *Geophysical Research Letters*, 35(8):L08715.

Jhun, I., Coull, B. A., Zanobetti, A., and Koutrakis, P. (2015). The impact of nitrogen oxides concentration decreases on ozone trends in the USA. *Air Quality, Atmosphere & Health*, 8(3):283–292.

Killick, R., Fearnhead, P., and Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107(500):1590–1598.

Krämer, W. and Sibbertsen, P. (2002). Testing for structural changes in the presence of long memory. *International Journal of Business and Economics*, 1(3):235–242.

Kuan, C. and Hsu, C. (1998). Change-point estimation of fractionally integrated processes. *Journal of Time Series Analysis*, 19(6):693–708.

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3):1217–1241.

Leadbetter, M. R., Lindgren, G., and Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York, NY, 1st edition.

Lee, J., Li, S., and Lund, R. (2014). Trends in extreme U.S. temperatures. *Journal of Climate*, 27(11):4209–4225.

Levermann, A., Clark, P. U., Marzeion, B., Milne, G. A., Pollard, D., Radic, V., and Robinson, A. (2013). The multimillennial sea-level commitment of global warming. *Proceedings of the National Academy of Sciences*, 110(34):13745–13750.

Li, S. and Lund, R. (2012). Multiple changepoint detection via genetic algorithms. *Journal of Climate*, 25(2):674–686.

Li, Z., Wang, Y., Zhao, W., Xu, Z., and Li, Z. (2016). Frequency analysis of high flow extremes in the Yingluoxia watershed in Northwest China. *Water*, 8:215.

Lim, C. C., Hayes, R. B., Ahn, J., Shao, Y., Silverman, D. T., Jones, R. R., Garcia, C., Bell, M. L., and Thurston, G. D. (2019). Long-term exposure to ozone and cause-specific mortality risk in the United States. *American Journal of Respiratory and Critical Care Medicine*, 200(8):1022–1031.

Lu, Q., Lund, R., and Lee, T. C. M. (2010). An MDL approach to the climate segmentation problem. *The Annals of Applied Statistics*, 4(1):299–319.

Lund, R. and Reeves, J. (2002). Detection of undocumented changepoints: A revision of the two-phase regression model. *Journal of Climate*, 15(17):2547–2554.

Marcos, M., Calafat, F. M., Berihuete, A., and Dangendorf, S. (2015). Long-term variations in global sea level extremes. *Journal of Geophysical Research: Oceans*, 120(12):8115–8134.

Marcos, M. and Woodworth, P. L. (2017). Spatiotemporal changes in extreme sea levels along the coasts of the North Atlantic and the Gulf of Mexico. *Journal of Geophysical Research: Oceans*, 122(9):7031–7048.

Matteson, D. S. and James, N. A. (2014). A nonparametric approach for multiple change point analysis of multivariate data. *Journal of the American Statistical Association*, 109(505):334–345.

McCormick, W. P. and Qi, Y. (2000). Asymptotic distribution for the sum and maximum of Gaussian processes. *Journal of Applied Probability*, 37(4):958–971.

McGranahan, G., Balk, D., and Anderson, B. (2007). The rising tide: Assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization*, 19(1):17–37.

McGrath, J. M., Betzelberger, A. M., Wang, S., Shook, E., Zhu, X.-G., Long, S. P., and Ainsworth, E. A. (2015). An analysis of ozone damage to historical maize and soybean yields in the United States. *Proceedings of the National Academy of Sciences*, 112(46):14390–14395.

Menéndez, M. and Woodworth, P. L. (2010). Changes in extreme high water levels based on a quasi-global tide-gauge data set. *Journal of Geophysical Research*, 115(C10):C10011.

Merrifield, M. A., Genz, A. S., Kontoes, C. P., and Marra, J. J. (2013). Annual maximum water levels from tide gauges: Contributing factors and geographic patterns. *Journal of Geophysical Research: Oceans*, 118(5):2535–2546.

Nelsen, R. B. (2006). *An Introduction to Copulas.* Springer Science & Business Media, New York, NY.

Nerem, R. S., Beckley, B. D., Fasullo, J. T., Hamlington, B. D., Masters, D., and Mitchum, G. T. (2018). Climate-change–driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the National Academy of Sciences*, 115(9):2022–2025.

Northrop, P. J. (2015). An efficient semiparametric maxima estimator of the extremal index. *Extremes*, 18:585–603.

Nunes, L. C., Kuan, C.-M., and Newbold, P. (1995). Spurious break. *Econometric Theory*, 11(4):736–749.

Palma, W. and Chan, N. H. (1997). Estimation and forecasting of long-memory processes with missing values. *Journal of Forecasting*, 16(6):395–410.

Parey, S., Hoang, T. T. H., and Dacunha-Castelle, D. (2010). Different ways to compute temperature return levels in the climate change context. *Environmetrics*, 21(7-8):698–718.

Parey, S., Malek, F., Laurent, C., and Dacunha-Castelle, D. (2007). Trends and climate evolution: Statistical approach for very high temperatures in France. *Climatic Change*, 81:331–352.

Phalitnonkiat, P., Sun, W., Grigoriu, M. D., Hess, P., and Samorodnitsky, G. (2016). Extreme ozone events: Tail behavior of the surface ozone distribution over the U.S. *Atmospheric Environment*, 128:134–146.

Reich, B. J., Shaby, B. A., and Cooley, D. (2014). A hierarchical model for serially-dependent extremes: A study of heat waves in the western US. *Journal of Agricultural, Biological, and Environmental Statistics*, 19(1):119–135.

Ribatet, M. and Sedki, M. (2013). Extreme value copulas and max-stable processes. *Journal de la Société Française de Statistique*, 154(1):138–150.

Rueda, A., Camus, P., Tomás, A., Vitousek, S., and Méndez, F. J. (2016). A multivariate extreme wave and storm surge climate emulator based on weather patterns. *Ocean Modelling*, 105:242–251.

Rust, H., Maraun, D., and Osborn, T. (2009). Modelling seasonality in extreme precipitation. *The European Physical Journal Special Topics*, 174(1):99–111.

Shen, L., Mickley, L. J., and Gilleland, E. (2016). Impact of increasing heat waves on U.S. ozone episodes in the 2050s: Results from a multimodel analysis using extreme value theory. *Geophysical Research Letters*, 43(8):4017–4025.

Shi, C., Fernando, H. J. S., and Yang, J. (2009). Contributors to ozone episodes in three US/Mexico border twin-cities. *Science of the Total Environment*, 407(18):5128–5138.

Simon, H., Reff, A., Wells, B., Xing, J., and Frank, N. (2015). Ozone trends across the United States over a period of decreasing NOx and VOC emissions. *Environmental Science & Technology*, 49(1):186–195.

Smith, J., Taylor, N., and Yadav, S. (1997). Comparing the bias and misspecification in ARFIMA models. *Journal of Time Series Analysis*, 18(5):507–527.

Smith, R. L. (2014). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90.

Tang, S. M. and MacNeill, I. B. (1993). The effect of serial correlation on tests for parameter change at unknown time. *The Annals of Statistics*, 21(1):552–575.

The BACC II Author Team (2015). *Second Assessment of Climate Change for the Baltic Sea Basin*. Springer, Cham, Heidelberg, Germany.

U.S. Environmental Protection Agency (EPA) (2022). *Nonattainment Areas for Criteria Pollutants (Green Book)*. https://www.epa.gov/green-book.

USGCRP (2018). Impacts, risks, and adaptation in the United States. In Reidmiller, D. R., Avery, C. W., Easterling, D. R., Kunkel, K. E., Lewis, K. L. M., Maycock, T. K., and Stewart, B. C., editors, *Fourth National Climate Assessment, Volume II*. U.S. Global Change Research Program, Washington, DC.

Varneskov, R. T. and Perron, P. (2018). Combining long memory and level shifts in modelling and forecasting the volatility of asset returns. *Quantitative Finance*, 18(3):371–393.

Wahl, T. and Chambers, D. P. (2015). Evidence for multidecadal variability in US extreme sea level records. *Journal of Geophysical Research: Oceans*, 120(3):1527–1544.

Wahl, T., D. Haigh, I., J. Nicholls, R., Arns, A., Dangendorf, S., Hinkel, J., and B. A. Slangen, A. (2017). Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nature Communications*, 8:16075.

Wang, W. and Zhou, W. (2017). Statistical modeling and trend detection of extreme sea level records in the Pearl River Estuary. *Advances in Atmospheric Sciences*, 34:383–396.

Watson, C. S., White, N. J., Church, J. A., King, M. A., Burgette, R. J., and Legresy, B. (2015). Unabated global mean sea-level rise over the satellite altimeter era. *Nature Climate Change*, 5:565–568.

Weglarczyk, S., Strupczewski, W. G., and Singh, V. P. (2005). Three-parameter discontinuous distributions for hydrological samples with zero values. *Hydrological Processes: An International Journal*, 19(15):2899–2914.

Weisse, R., Bellafiore, D., Menéndez, M., Méndez, F., Nicholls, R. J., Umgiesser, G., and Willems, P. (2014). Changing extreme sea levels along European coasts. *Coastal Engineering*, 87:4–14.

Woodward, W. A., Gray, H. L., and Elliott, A. C. (2017). *Applied Time Series Analysis*. CRC Press, Boca Raton, FL, 2nd edition.

Woodworth, P. L., Hunter, J. R., Marcos, M., Caldwell, P., Menéndez, M., and Haigh, I. (2016). Towards a global higher-frequency sea level dataset. *Geoscience Data Journal*, 3(2):50–59.

Woody, J., Wang, Y., and Dyer, J. (2016). Application of multivariate storage model to quantify trends in seasonally frozen soil. *Open Geosciences*, 8(1):310–322.

Yan, Y., Lin, J., and He, C. (2018). Ozone trends over the United States at different times of day. *Atmospheric Chemistry and Physics*, 18(2):1185–1202.

Zhang, X. F., Zwiers, F. W., and Li, G. (2004). Monte Carlo experiments on the detection of trends in extreme values. *Journal of Climate*, 17(10):1945–1952.

Zhao, N., Pinault, L., Toyib, O., Vanos, J., Tjepkema, M., and Cakmak, S. (2021). Long-term ozone exposure and mortality from neurological diseases in Canada. *Environment International*, 157:106817.

Zhu, L., Liu, X., and Lund, R. (2019). A likelihood for correlated extreme series. *Environmetrics*, 30(4):e2546.

APPENDIX A

SUPPLEMENTARY MATERIALS FOR CHAPTER 2

A.1   Additional Graphics

In Section 2.4, we considered the eight scenarios as summarized in Table 2.1. Fig A.1 shows an exemplary time plot of the simulated monthly maximum sea level series from each scenario with true changepoint times marked by red vertical lines (Scenarios 3–8).

Fig A.2 shows a frequency histogram of the GA estimated changepoint times for a weak but long-lasting temporal correlation case with $\lambda = 0.25$. For all scenarios considered, most of the GA estimated changepoint times are clustered around the true changepoint times with a minimal variability. This result supports our finding that our GA method performs well in estimating changepoint times for Scenarios 3–8 in a weak but long-lasting temporal correlation case.
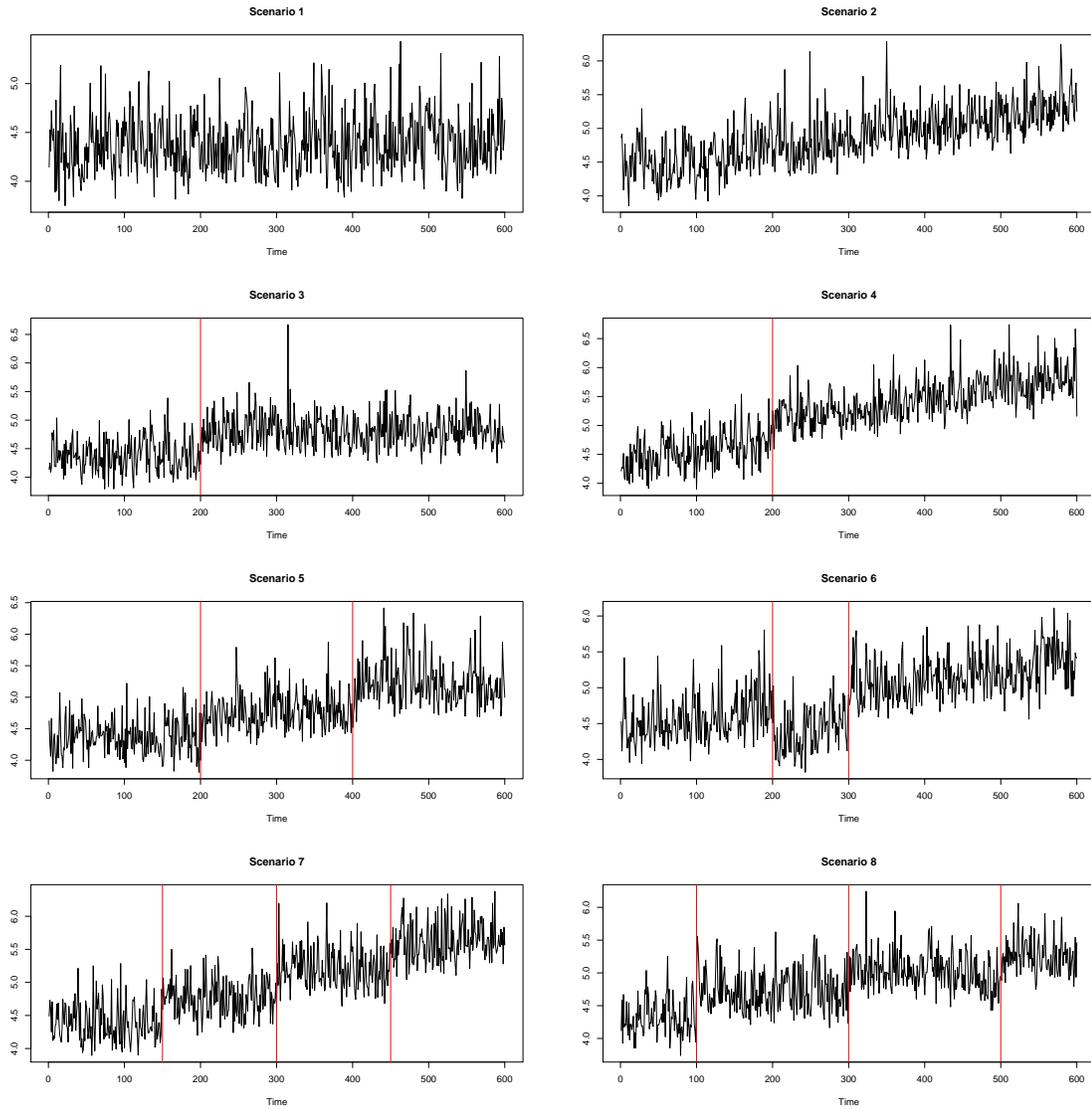
Figure A.1: Time plot of the simulated monthly maximum series under Scenarios 1–8
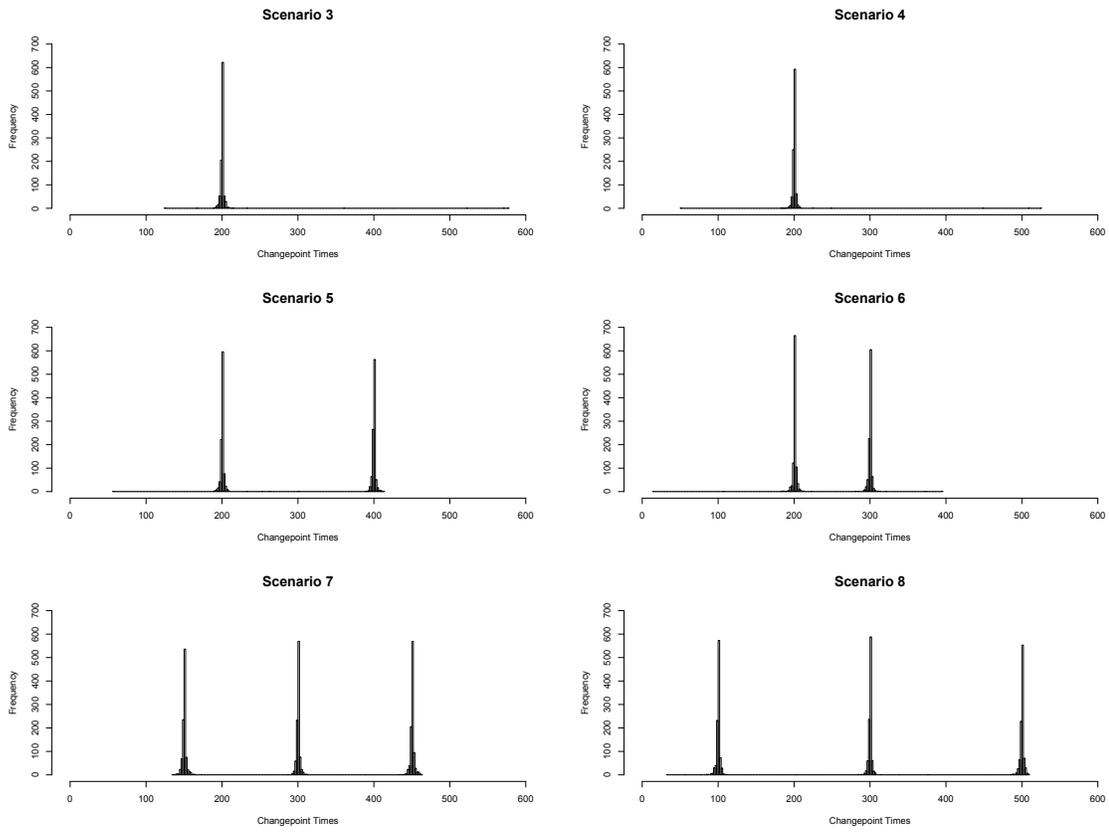
Figure A.2: Histograms of detected changepoint times from the GA method with $\lambda = 0.25$ for Scenarios 3–8

APPENDIX B

SUPPLEMENTARY MATERIALS FOR CHAPTER 3

B.1 Technical Details of the Truncated m-Dimensional State-Space Method for ARFIMA(p,d,q) Process

This section illustrates the technical details of the truncated m-dimensional state-space method of Chan and Palma (1998) and Palma and Chan (1997) applied to the ARFIMA(p, d, q) process in Section 3.3.4 as follows.

Since the ARFIMA(p, d, q) process Z_t is stationary and causal, it can be represented as an infinite moving-average process:

$$Z_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} = \left(1 - \sum_{j=1}^{p} \phi_j B^j\right)^{-1} (1 - B)^{-d} \left(1 - \sum_{j=1}^{q} \theta_j B^j\right) \varepsilon_t, \qquad \text{(B.1)}$$

where ψ_t is the coefficients of the infinite moving-average process (cf. Hosking, 1981). The ψ_t and white noise variance ε_t are set such that $V(Z_t) = 1$ to ensure the marginal distribution of each Z_t is standard normal. This infinite moving-average representation of Z_t in (B.1) can be approximated by truncating the infinite summation after m terms,

$$Z_t = \sum_{k=0}^{m} \psi_k \varepsilon_{t-k}. \qquad \text{(B.2)}$$

This truncated model can then be represented as an m-dimensional state-space

system. For $t = 1, \ldots, n$,

$$\boldsymbol{x_{t+1}} = F\boldsymbol{x_t} + H\varepsilon_t,$$

where $\boldsymbol{x_t} = (z(t|t-1), z(t+1|t-1), \cdots, z(t+m-1|t-1))^{\mathrm{T}}$ with $z(t|t-1) = E(z_t|z_{t-1}, z_{t-2}, \ldots)$ and

$$z_t = G\boldsymbol{x_t} + \varepsilon_t.$$

The system matrices are as follows.

$$F_{m \times m} = \begin{pmatrix} 0 & I_{m-1} \\ 0 & 0 \end{pmatrix},$$

$$H_{m \times 1} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_m \end{pmatrix},$$

and

$$G_{1 \times m} = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}.$$

Following the approach by Palma and Chan (1997), the one-step predictions can be recursively produced using the following Kalman equations:

$$\Delta_t = G\Omega_t G^{\mathrm{T}} + \sigma_\varepsilon^2,$$

$$\Theta_t = F\Omega_t G^{\mathrm{T}} + \sigma_\varepsilon^2 H,$$

$$\Omega_{t+1} = \begin{cases} F\Omega_{t+1}F^{\mathrm{T}} + \sigma_\varepsilon^2 HH^{\mathrm{T}} - \Omega_t\Delta_t^{-1}\Theta_t^{\mathrm{T}}, & \text{if } z_t \text{ is known;} \\[2ex] F\Omega_{t+1}F^{\mathrm{T}} + \sigma_\varepsilon^2 HH^{\mathrm{T}}, & \text{if } z_t \text{ is missing,} \end{cases}$$

$$\hat{\boldsymbol{x}}_{t+1} = \begin{cases} F\hat{\boldsymbol{x}}_{\boldsymbol{t}} + \Theta_t\Delta_t^{-1}(z_t - G\hat{\boldsymbol{x}}_{\boldsymbol{t}}), & \text{if } z_t \text{ is known;} \\[2ex] F\hat{\boldsymbol{x}}_{\boldsymbol{t}}, & \text{if } z_t \text{ is missing,} \end{cases}$$

and

$$\hat{z}_t = G\hat{\boldsymbol{x}}_{\boldsymbol{t}}.$$

Based on the Kalman recursive equations, the joint probability density function of $Z_1, \ldots, Z_t$ can then be written as follows.

$$f_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) = (2\pi)^{-n/2} \left(\prod_{t=1}^n \Delta_t\right)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^n \frac{(z_t - \hat{z}_t)^2}{\Delta_t}\right\},$$

where $\hat{z}_t = E(z_t|z_1, \ldots, z_{t-1})$ is the one-step ahead prediction of z_t where $\hat{z}_1 = 0$ and $\Delta_t = Var(z_t - \hat{z}_t)$ is the one-step predictor error variance.

B.2 GA Detection Counts on U.S. County-Level Weekly Maximum Ozone Concentrations

This section summarizes changepoint detection results on weekly maximum ozone concentrations using three different autocorrelation models in the GA: long-memory ARFIMA$(0, d, 1)$ copula-GEV model, short-memory AR(1) copula-GEV model, and the autocorrelation-ignored copula-GEV model. Table B.1 summarizes the estimated number of changepoints in weekly maximum ozone series from all 395 counties. Maps B.1–B.3 depict spatial maps of the estimated changepoint numbers by

the ARFIMA$(0, d, 1)$ copula-GEV GA, AR(1) copula-GEV GA, and autocorrelation-ignored copula-GEV GA, respectively. As expected, the AR(1) copula-GEV GA and autocorrelation-ignored GA overall estimate larger number of changepoints than the ARFIMA$(0, d, 1)$ copula-GEV GA in most counties.

Table B.1: Detection counts for the GA method using three copula-GEV models on weekly maximum ozone concentrations from 395 U.S. counties
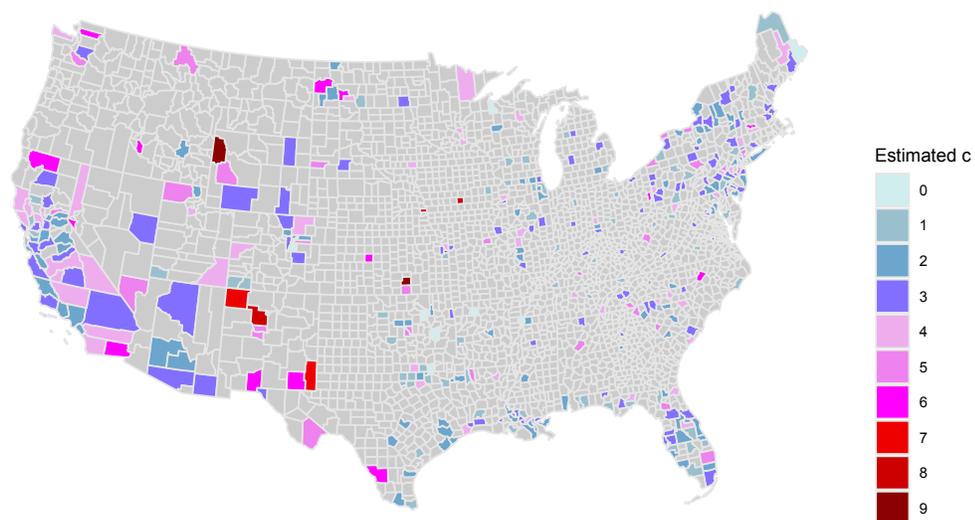
Base process	$c = 0$	$c = 1$	$c = 2$	$c = 3$	$c = 4$	$c = 5$	$c = 6$	$c = 7$	$c = 8$	$c = 9$
ARFIMA$(0, d, 1)$	112	171	84	18	6	3	1	0	0	0
AR(1)	31	149	120	64	19	19	9	1	2	0
Autocorr. Ignored	8	70	100	109	52	32	16	2	3	3
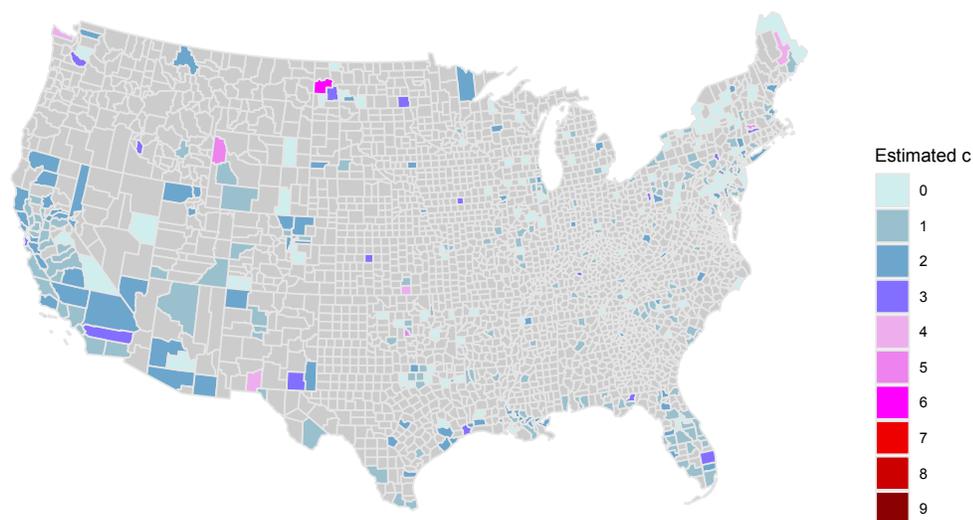
Map B.1: Spatial patterns of the estimated changepoint numbers from the GA using long-memory copula-GEV
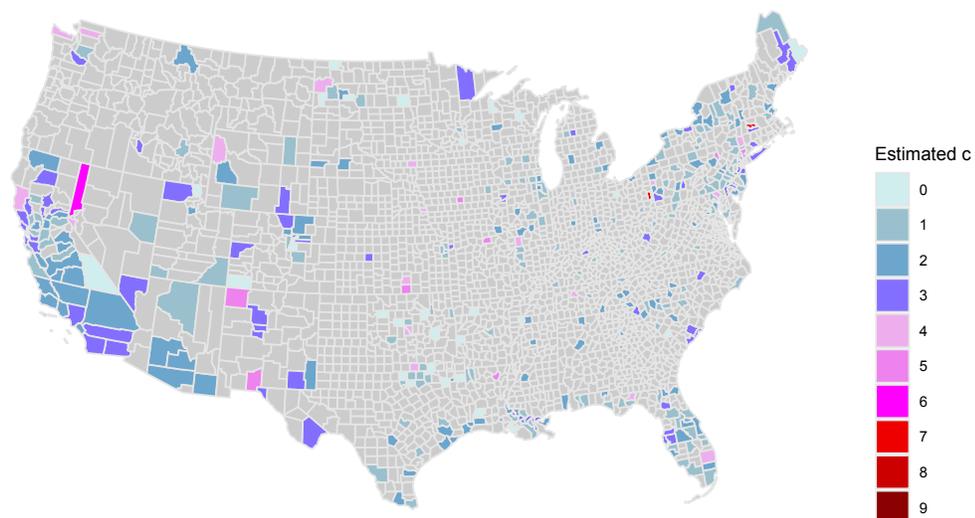
Map B.2: Spatial patterns of the estimated changepoint numbers from the GA using short-memory copula-GEV

Map B.3: Spatial patterns of the estimated changepoint numbers from the GA using autocorrelation-ignored copula-GEV