

# Long-term trend analysis of extreme coastal sea levels with changepoint detection

Mintaek Lee | Jaechoul Lee

Boise State University, Boise, ID, USA

## Correspondence

Jaechoul Lee, Department of Mathematics,  
Boise State University, 1910 University  
Drive, Boise, ID 83725, USA.  
Email: jaechoullee@boisestate.edu

## Abstract

Sea level rise can bring disastrous outcomes to people living in coastal regions by increasing flood risk or inducing stronger storm surges. We study long-term linear trends in monthly maximum sea levels by applying extreme value methods. The monthly maximum sea levels are extracted from multiple tide gauges around the coastal regions of the world over a period of as long as 169 years. Due to instrument changes, location changes, earthquakes, land reclamation, dredging, etc., the sea level data could contain inhomogeneous shifts in their means, which can substantially impact trend estimates if ignored. To rigorously quantify the long-term linear trends and return levels for the monthly maximum sea level data, we use a genetic algorithm to estimate the number and times of changepoints in the data. As strong periodicity and temporal correlation are pertinent to the 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 linear trends of 89 tide gauges (approximately 30% of tide gauges considered) by more than 20 cm century<sup>-1</sup>. Our results are summarized in maps with estimated extreme sea level trends and 50-year return levels.

## KEY WORDS

bootstrap confidence interval, changepoints, extreme sea levels, generalized extreme value distribution, genetic algorithm, temporal correlation

## 1 | INTRODUCTION

With over 10% of the world's population living in low elevation coastal areas (McGranahan et al., 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 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 the sea level has been gradually rising over time (Chen et al., 2017; Nerem et al., 2018; Watson et al., 2015). 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 & 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) analysed 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 & 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.

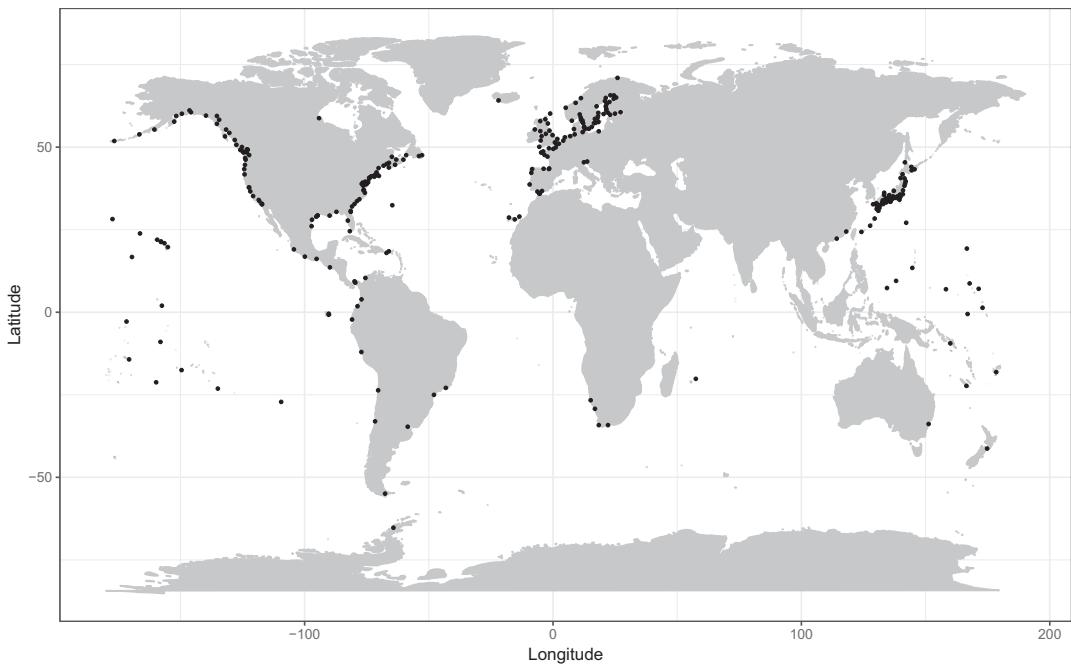
This paper proceeds as follows: Section 2 describes the data set used in our study. Section 3 discusses the extreme value methods and changepoint detection techniques applied to the sea level data. Section 4 describes the simulation study on changepoint detection and summarizes the performance of the changepoint method. Section 5 illustrates our methods using a case study for the sea level series from Fishguard, UK. Section 6 summarizes our trend and return level analysis for the entire sea level data set. In Section 7, we conclude with further remarks.

## 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 metres 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).

Our preprocessing procedures resulted in 300 monthly maximum sea level series for our analysis. Figure 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



**FIGURE 1** Spatial location of the 300 tide gauges after preprocessing

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 north-eastern portion of South America are also left out, limiting the coverage of the coasts of the South Atlantic Ocean in our study.

### 3 | METHODS

#### 3.1 | Block maxima methods with periodic, trend and changepoint features

Suppose  $X_1, \dots, X_m$  are independent and identically distributed (IID) random variables with a common distribution function  $F(\cdot)$ . Define  $Y_{(m)} = \max\{X_1, \dots, X_m\}$  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

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

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  $\{X_1, \dots, X_N\}$  at a gauge site as a set of  $n$  blocks with block size  $m$ :  $\{(X_1, \dots, X_m), (X_{m+1}, \dots, X_{2m}), \dots, (X_{(n-1)m+1}, \dots, X_{nm})\}$ , where  $N = nm$ . We then compute the maximum statistic from each block and denote the  $t$ th block maximum statistic as  $Y_t = \max\{X_{(t-1)m+1}, \dots, X_{tm}\}$  for  $t = 1, \dots, n$ . The extreme value theorem states that if block size  $m$  is sufficiently large, the GEV distribution is a proper probability distribution for  $\{Y_1, \dots, Y_n\}$  regardless of the distribution function  $F(\cdot)$  from which the raw data  $\{X_1, \dots, X_N\}$  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:

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

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

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

where  $I(E)$  is an indicator function returning 1 if  $E$  is true and 0 otherwise. The baseline mean shift term  $\Delta_0$  is set to be zero for parameter identifiability. The scale parameter is also parameterized as a cosine wave with two harmonics:

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

The seasonal dependence in extreme sea levels is mainly due to astronomical (spring tides) and meteorological (storminess season) influences (Menéndez & 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 Equations (1) and (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  $\xi$  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  $\xi$ , especially over short segments. This constant  $\xi$  assumption is often physically and numerically supported by other researchers (Hoang et al., 2009; Parey et al., 2007; Zhang et al., 2004).

For parameter estimation, if the changepoint number  $c$  is known and these  $c$  changepoints are also known to occur at the times  $\tau_1, \dots, \tau_c$ , the unknown GEV model parameters are denoted by  $\eta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \alpha, \Delta_1, \dots, \Delta_c, \omega_0, \omega_1, \omega_2, \omega_3, \omega_4, \xi)^T$ . The log-likelihood function of the  $\text{GEV}(\mu_t, \sigma_t, \xi)$  distribution with  $\mu_t$  in Equation (1) and  $\sigma_t$  in Equation (2) can be written as

$$\ell(\eta) = \begin{cases} -\sum_{t=1}^n \ln \sigma_t - \sum_{t=1}^n \left[ 1 + \xi \left( \frac{y_t - \mu_t}{\sigma_t} \right) \right]^{-1/\xi} \\ \quad - \left( 1 + \frac{1}{\xi} \right) \sum_{t=1}^n \ln \left[ 1 + \xi \left( \frac{y_t - \mu_t}{\sigma_t} \right) \right], & \text{if } \xi \neq 0; \\ -\sum_{t=1}^n \ln \sigma_t - \sum_{t=1}^n \left( \frac{y_t - \mu_t}{\sigma_t} \right) - \sum_{t=1}^n \exp \left[ -\left( \frac{y_t - \mu_t}{\sigma_t} \right) \right], & \text{if } \xi = 0. \end{cases} \quad (3)$$

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

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

### 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 et al. (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  $(c; \tau_1, \dots, \tau_c)$ , where  $c$  is the number of changepoints, and  $\tau_j$  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  $(a; \kappa_1, \dots, \kappa_a)$  and  $(b; \zeta_1, \dots, \zeta_b)$  are chosen as parents. These two chromosomes are first merged, resulting in  $(a+b; \tau_1, \dots, \tau_{a+b})$ . Next, after eliminating duplicate times, each  $\tau_i$  is removed from the child with probability 0.5. Each remaining  $\tau_i$  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 300th 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 et al. (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 et al., 2010). For a chromosome of  $c$  changepoints at times  $\tau_1, \dots, \tau_c$ , the MDL is calculated as:

$$\text{MDL}(\boldsymbol{\eta}, c, \tau_1, \dots, \tau_c) = -\ell_{\text{opt}}(\boldsymbol{\eta}) + P(c; \tau_1, \dots, \tau_c). \quad (4)$$

Here,  $\ell_{\text{opt}}(\boldsymbol{\eta})$  is the optimized value of the GEV log-likelihood function in Equation (3) calculated at the maximum likelihood estimates of  $\boldsymbol{\eta}$  for a given changepoint configuration  $(c; \tau_1, \dots, \tau_c)$ . The penalty term  $P(c; \tau_1, \dots, \tau_c)$  is expressed as

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

where  $\tau_{c+1} = n+1$ . If missing observations exist in the  $j$ th segment on times  $\tau_{j-1}, \dots, \tau_j - 1$ , 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 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 Equation (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 et al., 2012), will not work, because the parameter estimate of the long-term trend  $\alpha$ , for example, depends on all the data points, not only the data in any single segment. The model parameters  $\alpha, \beta$ 's,  $\omega$ 's, and  $\xi$  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 3.4.

### 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 & Walshaw, 2012; Reich et al., 2014), past authors often considered an additional parameter, called extremal index. The extreme value theorem described in Section 3.1 then can hold true for dependent series by incorporating the extremal index.

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

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

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

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

for a constant  $\theta \in (0, 1]$ . Here, the parameter  $\theta$  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  $\theta$ .

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 et al. (2010). Specifically, we now estimate the level  $r_z$  for which the expected number of exceedances in  $z$  years (12 $z$  months) is one. The  $z$ -year return level of non-stationary monthly maxima is then the solution to the following non-linear equation

$$1 = \sum_{t=t_l}^{t_l+12z-1} (1 - G_t^\theta(r_z)), \quad (5)$$

where  $G_t^\theta(\cdot)$  is the time-dependent GEV distribution function in month  $t$ , and  $t_l \geq n$  is a predetermined starting time for return level computation. We set  $t_l$  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 et al., 2010).

To accurately compute the return level estimate, we therefore need to estimate the extremal index  $\theta$ . 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 modelling 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  $X_1, \dots, X_N$  are strictly stationary random variables with a common distribution function  $F(\cdot)$  and extremal index  $\theta$ . Let  $Y_t$  be the maximum statistic for the  $t$ th block (month) of size  $m$  for  $t = 1, \dots, n$ . 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:

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

where  $A_{-t} = \{X_1, \dots, X_N\} \setminus \{X_{(t-1)m+1}, \dots, X_m\}$ . Then,  $V_t$  can be estimated as  $\hat{V}_t = -m \ln \hat{F}_{-t}(Y_t)$  for  $t = 1, \dots, n$ . 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 5.2. Once the extremal index  $\theta$  is estimated, we use the estimate  $\hat{\theta}$  to adjust the GEV distribution function for temporal correlation as done in Equation (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 Equation (5) is not available for the non-stationary case, we use a grid search algorithm to numerically estimate  $r_z$ .

### 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  $\alpha$ , 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  $\{Y_1, \dots, Y_n\}$  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.

## 4 | A SIMULATION STUDY

A simulation study was performed to evaluate the effectiveness of the GA changepoint method described in Section 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 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

TABLE 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	( $c = 3; \tau_1 = 100, \tau_2 = 300, \tau_3 = 500$ )

$\tau_2 = 300$ . 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 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.

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  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$ , 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)  $s = 1, \dots, N$ ,

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

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  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$ , where  $\tilde{X}_s = \tilde{\mu}_s + \tilde{\sigma}_s \tilde{Z}_s$ , 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 1, an hourly sea level series  $\{X_1, \dots, X_N\}$  is generated by incorporating long-term trend and mean shifts occurring in months  $\tau_j$ 's into  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$ .
4. A monthly maximum series  $\{Y_1, \dots, Y_n\}$  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. Figure S.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 et al., 2017). The periodicity and autocorrelation of a Gegenbauer process are governed by the parameters  $u$  and  $\lambda$ . 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  $\lambda$  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 Equation (4) as the fitness function to minimize, our GA method is applied to these simulated monthly maximum series  $\{Y_1, \dots, Y_n\}$  to estimate the number and times of changepoints. Table 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 ( $\pm 3$ ,  $\pm 6$  and  $\pm 9$  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

**TABLE 2** Detection counts and accuracy rates for the GA method with  $\lambda = 0.25$ 

Scenario	$c = 0$	$c = 1$	$c = 2$	$c = 3$	$c = 4$	$c = 5$	$\pm 3$ mos.	$\pm 6$ mos.	$\pm 9$ 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%

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 Figure S2. 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 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. Figure 2 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.

## 5 | CASE STUDY: FISHGUARD, UK

### 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, we applied the GA method to the Fishguard monthly maximum series using the MDL in Equation (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).

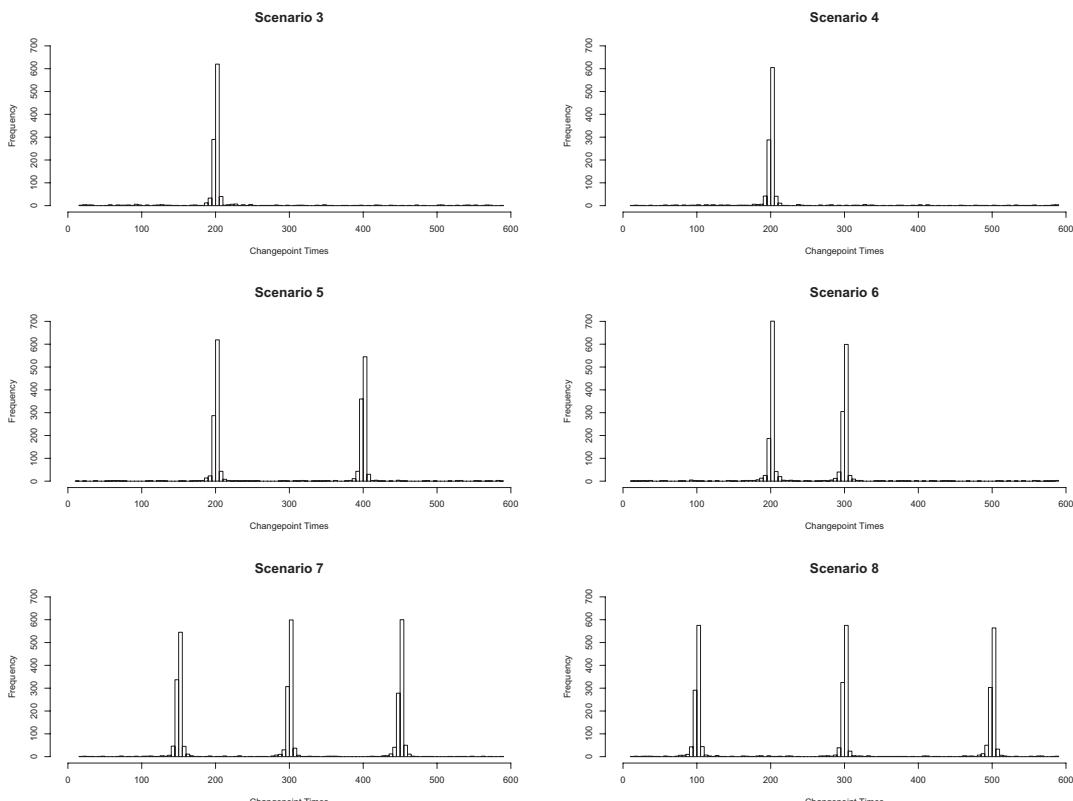
Although metadata are not available for most GESLA stations, BODC, the provider of the Fishguard data, informed us that the Fishguard tide gauge had experienced two location changes in 1975 and June 1988. The gauge also had a new instrument installed in June 1988, set to record data in 15-minute intervals. However, it took a few more years to fully change over, so hourly data were recorded until

TABLE 3 Detection counts and accuracy rates for the GA method with  $\lambda = 0.375$ 

Scenario	$c = 0$	$c = 1$	$c = 2$	$c = 3$	$c = 4$	$c = 5$	$\pm 3$ mos.	$\pm 6$ mos.	$\pm 9$ 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%

December 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 Equations (1) and (2), we estimate the model parameters by maximizing the likelihood function in

FIGURE 2 Histograms of detected changepoint times from the GA method with  $\lambda = 0.375$  for Scenarios 3–8

Equation (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 3.4.

Table 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  $\alpha$ , 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  $34 \text{ cm century}^{-1}$  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  $-80 \text{ cm century}^{-1}$  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. Figure 3 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 blue and red lines, respectively. The estimated trend line with those three changepoints appears to be appropriate for the Fishguard series.

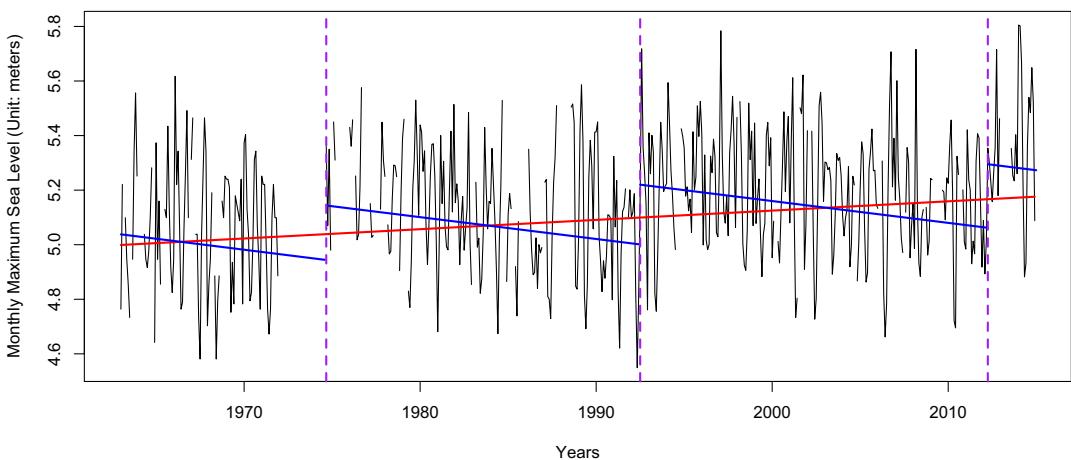
A goodness-of-fit for the GEV model is also performed to check if our GEV model fits well to the Fishguard monthly maximum sea level series. Figure 4 shows the Gumbel scaled quantile-quantile plots for our GEV models, visually comparing the two GEV models with the three GA changepoints considered and 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.

## 5.2 | Return level estimation with non-stationarity and temporal correlation

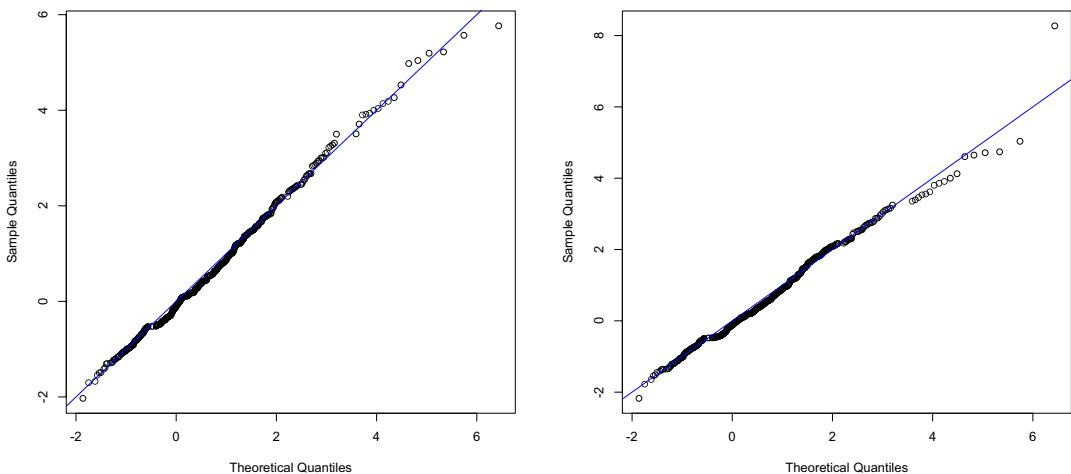
Since the Fishguard hourly sea level series is strongly autocorrelated, we estimate the extremal index  $\theta$  and use it for accurate return level estimation as explained in Section 3.3. However, as Northrop's

**TABLE 4** Fishguard GEV parameter estimates and bootstrap standard errors in parentheses (units: metres for  $\beta$ 's,  $\Delta$ 's and  $\omega$ 's; m century $^{-1}$  for  $\alpha$ )

Three changepoints	No changepoints	Three changepoints	No changepoints		
$\beta_0$	5.038 (0.020)	4.998 (0.020)	$\omega_0$	0.163 (0.006)	0.173 (0.006)
$\beta_1$	0.108 (0.012)	0.109 (0.012)	$\omega_1$	0.026 (0.006)	0.023 (0.007)
$\beta_2$	-0.027 (0.011)	-0.033 (0.012)	$\omega_2$	0.009 (0.007)	0.011 (0.008)
$\beta_3$	-0.154 (0.015)	-0.156 (0.015)	$\omega_3$	0.007 (0.008)	0.012 (0.008)
$\beta_4$	0.014 (0.015)	0.015 (0.016)	$\omega_4$	0.008 (0.006)	0.007 (0.007)
$\alpha$	-0.800 (0.175)	0.342 (0.063)	$\xi$	-0.250 (0.027)	-0.237 (0.033)
$\Delta_1$	0.199 (0.033)		$\theta$	0.128	0.128
$\Delta_2$	0.418 (0.062)				
$\Delta_3$	0.651 (0.089)		MDL	-245.886	-234.857



**FIGURE 3** 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) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** 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 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

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  $\{X_1, \dots, X_N\}$  to the changepoint-adjusted and detrended hourly data  $\{\ddot{X}_1, \dots, \ddot{X}_N\}$ , where  $\ddot{X}_s = X_s - \psi_s - \gamma s$ , and  $\psi_s$  and  $\gamma$  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  $\{X_1, \dots, X_N\}$  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  $\{X_1, \dots, X_N\}$  produces  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$ .
4. The changepoint-adjusted and detrended hourly series  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$  is then transformed to the stationary series  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$  by taking  $\tilde{X}_s = (\tilde{X}_s - \tilde{\mu}_s) / \tilde{\sigma}_s$ , where  $\tilde{\mu}_s$  and  $\tilde{\sigma}_s$  are hourly mean and standard deviation of the  $\tilde{X}_s$  series and are calculated by a similar method to Woody et al. (2016). To elaborate, for each hour  $\nu$  of the synodic lunar month (approximately 708.734 hours or 29.5 days), we compute the means and standard deviations

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

where  $n_\nu$  is the number of non-missing values during hour  $\nu = 1, \dots, 708$  of the synodic lunar month,  $[a]$  denotes the largest integer smaller than or equal to  $a$ ,  $[a]$  returns the index of  $\tilde{X}_s$  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  $s = 1, \dots, N$ , resulting in the periodic mean  $\tilde{\mu}_s$  and periodic standard deviation  $\tilde{\sigma}_s$  over the entire record period. Finally, we obtain  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$ , where  $\tilde{X}_s = (\tilde{X}_s - \tilde{\mu}_s) / \tilde{\sigma}_s$ , and consider this series to be stationary.

Now, we estimate the extremal index  $\theta$  by applying the Northrop's semiparametric maxima estimator to the stationary hourly series  $\{\tilde{X}_1, \dots, \tilde{X}_N\}$ . 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 Equation (5) with  $\hat{\theta}$  used in  $G_i^\theta(\cdot)$  as illustrated in Section 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

**TABLE 5** Fishguard monthly maximum sea level return exceedances and their 95% BCa bootstrap confidence intervals in parentheses (unit: metres)

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)

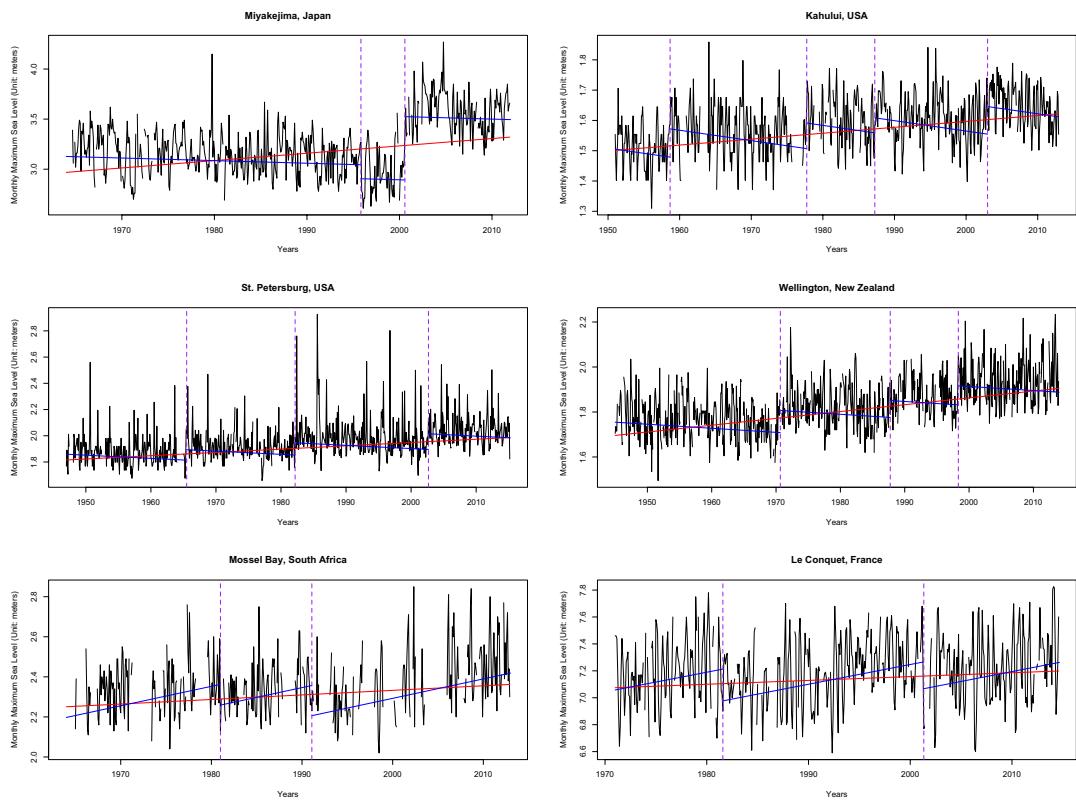
the return sea level estimates across the sites and allows us to make meaningful comparisons between tide gauges. Table 5 summarizes the estimated 25-, 50-, 75- and 100-year Fishguard sea level return exceedances with the three GA changepoints considered and without allowing changepoints. Their associated 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.

## 6 | GESLA DATA ANALYSIS

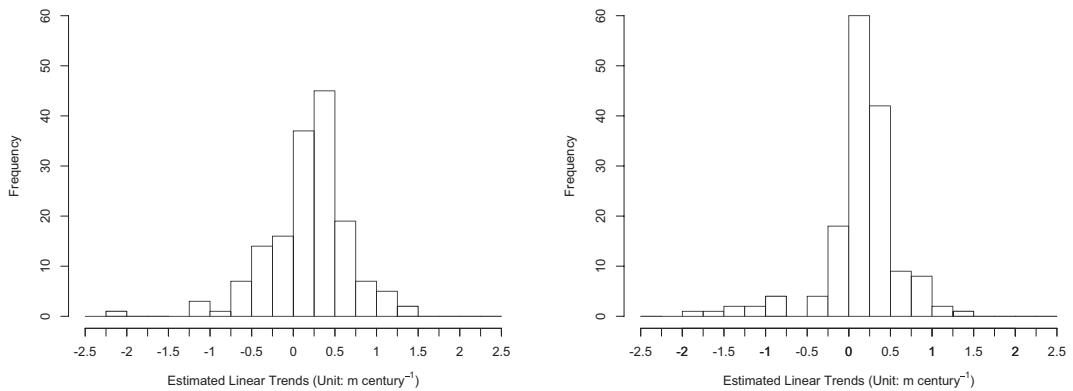
As illustrated in Section 5, our methods were applied to the monthly maximum sea level series at all 300 gauge sites. Out of the 300 sites, 47.7% (143 sites) are identified to have no significant mean shifts in their monthly maximum series. Most of these sites are either in North America or Europe. The other 52.3% (157 sites) are flagged for one or more changepoints. Of those sites with at least one GA changepoint, 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 gauge sites are Miyakejima, Japan; Kahului, USA; St Petersburg, USA; Wellington, New Zealand; Mossel Bay, South Africa; and Le Conquet, France. Figure 5 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 ( $73 \text{ cm century}^{-1}$ ) when changepoints are not allowed, these two changepoints considerably decrease the trend estimate to a negative trend ( $-26 \text{ cm century}^{-1}$ ). 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 changepoint. Figure 6 displays the estimated linear trends for these sites

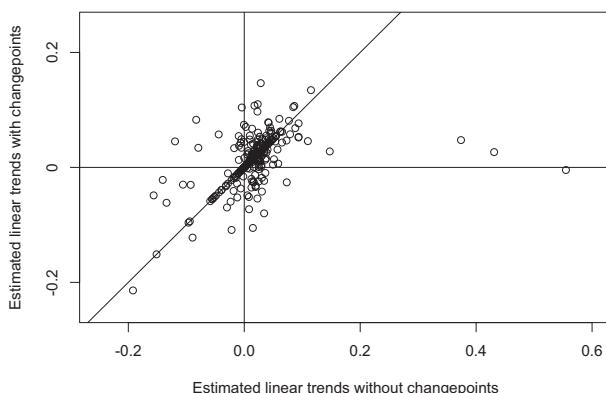


**FIGURE 5** Estimated trend lines for the monthly maximum sea levels (blue solid lines, the trend with GA estimated changepoint times; red solid line, the trend without allowing changepoints) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

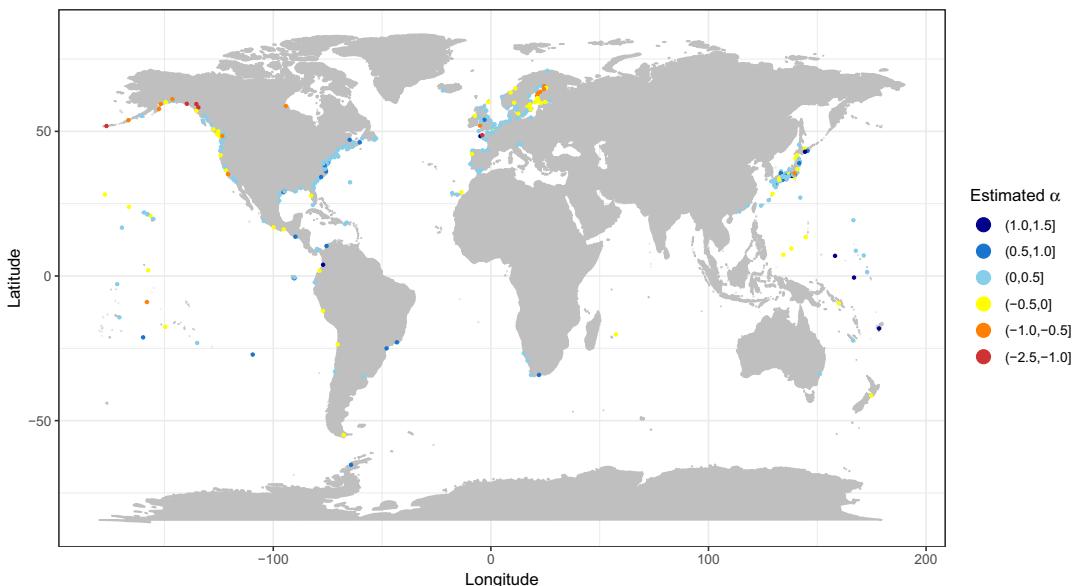


**FIGURE 6** Histograms of the estimated linear trends with changepoints considered (left) and without allowing changepoints (right) for the 157 sites with at least one GA changepoint (unit:  $\text{m century}^{-1}$ )

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  $25 \text{ cm century}^{-1}$



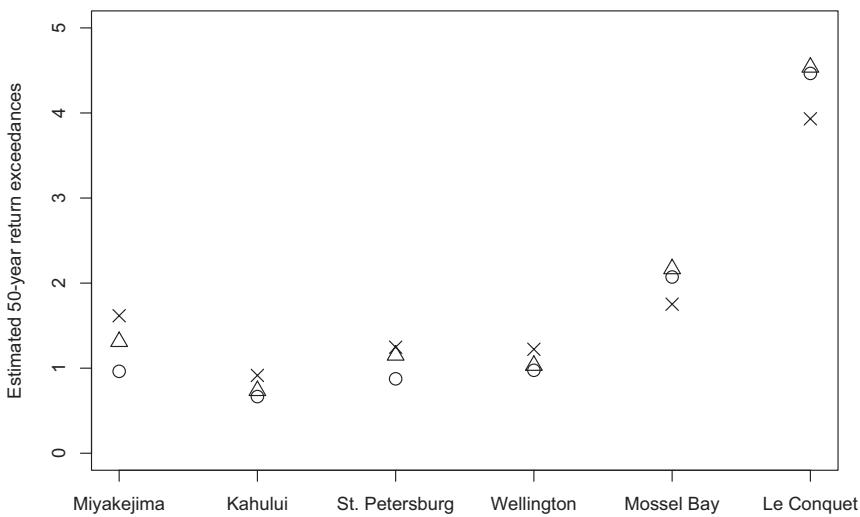
**FIGURE 7** A scatter plot of the estimated trends with changepoints considered against the estimated trends without allowing changepoints for the 157 sites with at least one GA changepoints (unit: m century<sup>-1</sup>)



**FIGURE 8** Estimated linear trends from the GEV model with changepoints considered (unit: m century<sup>-1</sup>)  
[Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and a standard deviation of 28 cm century<sup>-1</sup>, the estimated trends with GA changepoints considered have a mean of 20 cm century<sup>-1</sup> and a standard deviation of 45 cm century<sup>-1</sup>. Figure 7 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 Figure 8. Overall, monthly 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



**FIGURE 9** Monthly maximum sea level 50-year return exceedances (○, changepoints and extremal index considered; △, changepoints only; ×, changepoints and temporal correlation ignored; unit: metres)

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 Equation (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 temporal correlation. Figure 9 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 modelling process to obtain accurate return level estimates of monthly maximum sea levels.

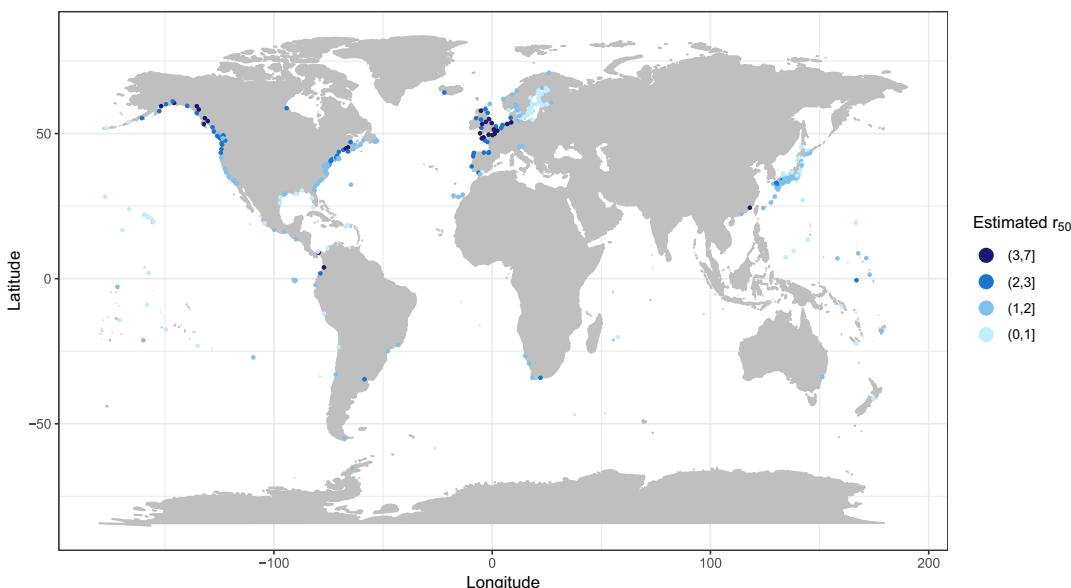
Figure 10 presents the estimated 50-year return sea level exceedances from the median for all 300 sites from the GEV model with GA estimated changepoints and temporal correlation considered. Coasts of the Northwestern Europe and the Gulf of Alaska appear to have the highest risk of extreme sea level events, with their typical current sea levels expected to exceed approximately 3–4 m on average once between January 2020 and December 2069.

## 7 | CLOSING COMMENTS

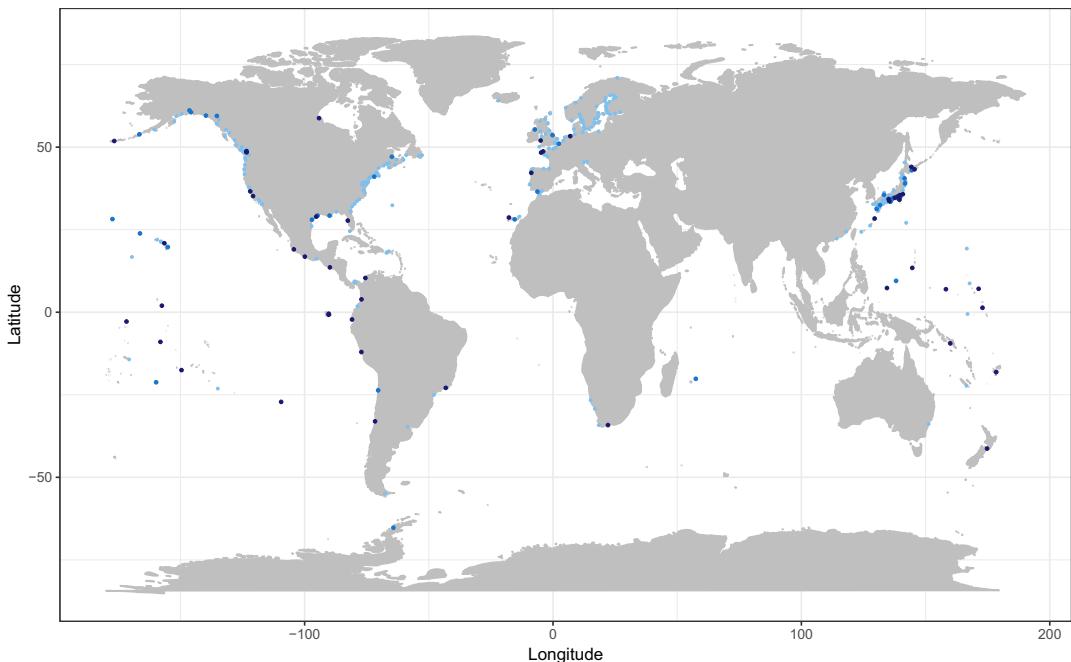
Sea level observations often contain undocumented changes in their means due to instrument changes, location changes, earthquakes, land reclamation, dredging and so on. The number of changepoints and changepoint times are unknown if a metadata of recording such changes does not exist or is not complete, 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 (Hünicke et al., 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



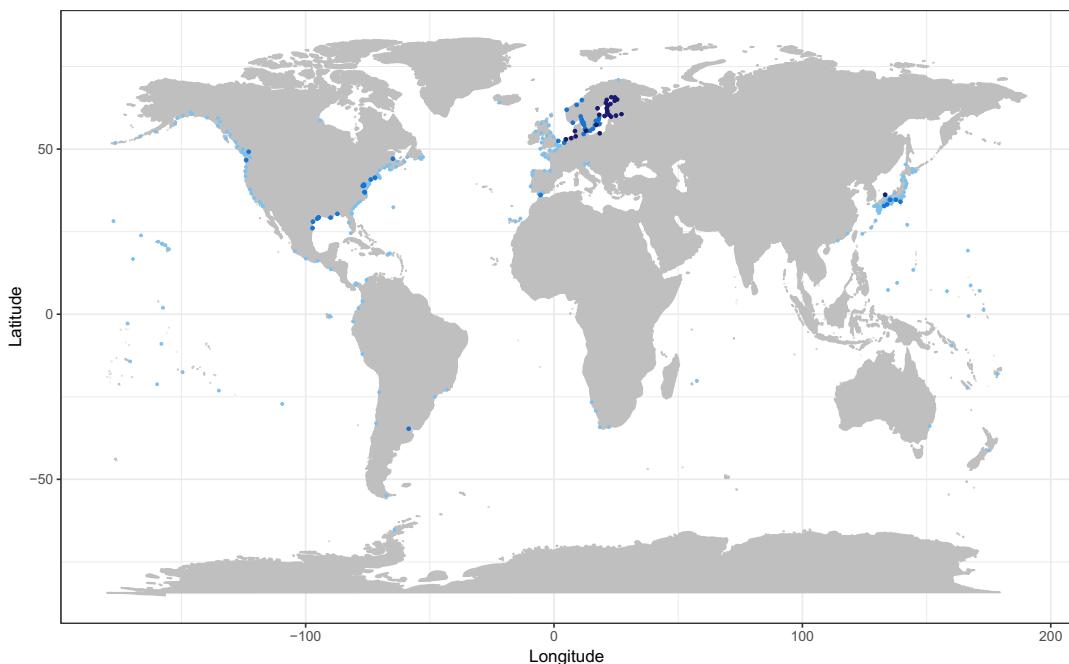
**FIGURE 10** Estimated monthly maximum sea level 50-year return exceedances (unit: metres) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 11** Spatial location of the tide gauges where the consideration of changepoints changed the estimated trends by more than  $40 \text{ cm century}^{-1}$  (52 sites, dark blue), between  $20$  and  $40 \text{ cm century}^{-1}$  (37 sites, blue), and less than  $20 \text{ cm century}^{-1}$  (211 sites, light blue) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

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  $20 \text{ cm century}^{-1}$  after changepoints are considered. Among these 89 substantially impacted sites, 52 of them further show changes of more than  $40 \text{ cm century}^{-1}$ . Figure 11 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.

Strong temporal correlation is pertinent in most raw sea level data. The extremal index  $\theta$ , a measure for the strength of temporal correlation, was estimated between 0.005 and 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. Figure 12 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.



**FIGURE 12** 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) [Colour figure can be viewed at [wileyonlinelibrary.com](https://wileyonlinelibrary.com)]

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 correlation 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 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.

## ACKNOWLEDGEMENTS

The authors thank the editor, associate editor and referees for their thoughtful reviews and valuable suggestions that significantly improved this paper. The authors also acknowledge the high-performance computing support of the R2 compute cluster (DOI: 10.18122/B2S41H) provided by Boise State University's Research Computing Department. The sea level data were downloaded from the Global Extreme Sea Level Analysis Version 2 (GESLA-2) at <https://gesla.org>.

## REFERENCES

- Becker, M., Karpytchev, M., Davy, M. & 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, 741–749.
- Chen, X., Zhang, X., Church, J.A., Watson, C.S., King, M.A., Monselesan, D. et al. (2017) The increasing rate of global mean sea-level rise during 1993–2014. *Nature Climate Change*, 7, 492–495.
- Church, J.A. & 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*. London, UK: Springer-Verlag.
- Efron, B. (1987) Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82, 171–185.
- Fawcett, L. & Walshaw, D. (2012) Estimating return levels from serially dependent extremes. *Environmetrics*, 23, 272–283.
- Fryzlewicz, P. (2014) Wild binary segmentation for multiple change-point detection. *The Annals of Statistics*, 42, 2243–2281.
- Hoang, T.T.H., Parey, S. & Dacunha-Castelle, D. (2009) Multidimensional trends: The example of temperature. *The European Physical Journal Special Topics*, 174, 113–124.
- Hünicke, B., Zorita, E., Soomere, T., Madsen, K.S., Johansson, M. & Suursaar, Ü. (2015) Recent change - sea level and wind waves. In *Second assessment of climate change for the Baltic Sea basin* (eds The BACC II Author Team), pp. 155–185. Heidelberg, Germany: Springer.
- Jevrejeva, S., Moore, J.C., Grinsted, A. & Woodworth, P.L. (2008) Recent global sea level acceleration started over 200 years ago? *Geophysical Research Letters*, 35, L08715.
- Killick, R., Fearnhead, P. & Eckley, I.A. (2012) Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107, 1590–1598.
- Künsch, H.R. (1989) The Jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17, 1217–1241.
- Leadbetter, M.R., Lindgren, G. & Rootzén, H. (1983) *Extremes and related properties of random sequences and processes*, 1st edn. New York, NY: Springer-Verlag.
- Lee, J., Li, S. & Lund, R. (2014) Trends in extreme U.S. temperatures. *Journal of Climate*, 27, 4209–4225.
- Levermann, A., Clark, P.U., Marzeion, B., Milne, G.A., Pollard, D., Radic V. et al. (2013) The multimillennial sea-level commitment of global warming. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 13745–13750.
- Li, S. & Lund, R. (2012) Multiple changepoint detection via genetic algorithms. *Journal of Climate*, 25, 674–686.
- Lu, Q., Lund, R. & Lee, T.C.M. (2010) An MDL approach to the climate segmentation problem. *The Annals of Applied Statistics*, 4, 299–319.
- Lund, R. & Reeves, J. (2002) Detection of undocumented changepoints: A revision of the two-phase regression model. *Journal of Climate*, 15, 2547–2554.
- Marcos, M. & 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, 7031–7048.
- Marcos, M., Calafat, F.M., Berihuete, A. & Dangendorf, S. (2015) Long-term variations in global sea level extremes. *Journal of Geophysical Research: Oceans*, 120, 8115–8134.
- Matteson, D.S. & James, N.A. (2014) A nonparametric approach for multiple change point analysis of multivariate data. *Journal of the American Statistical Association*, 109, 334–345.
- McCormick, W.P. & Qi, Y. (2000) Asymptotic distribution for the sum and maximum of Gaussian processes. *Journal of Applied Probability*, 37, 958–971.
- McGranahan, G., Balk, D. & Anderson, B. (2007) The rising tide: Assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization*, 19, 17–37.
- Menéndez, M. & Woodworth, P.L. (2010) Changes in extreme high water levels based on a quasi-global tide-gauge data set. *Journal of Geophysical Research*, 115, 10011.
- Merrifield, M.A., Genz, A.S., Kontoes, C.P. & Marra, J.J. (2013) Annual maximum water levels from tide gauges: Contributing factors and geographic patterns. *Journal of Geophysical Research: Oceans*, 118, 2535–2546.
- Nerem, R.S., Beckley, B.D., Fusilero, J.T., Hamlington, B.D., Masters, D. & Mitchum, G.T. (2018) Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 2022–2025.

- Northrop, P.J. (2015) An efficient semiparametric maxima estimator of the extremal index. *Extremes*, 18, 585–603.
- Parey, S., Malek, F., Laurent, C. & Dacunha-Castelle, D. (2007) Trends and climate evolution: Statistical approach for very high temperatures in France. *Climatic Change*, 81, 331–352.
- Parey, S., Hoang, T.T.H. & Dacunha-Castelle, D. (2010) Different ways to compute temperature return levels in the climate change context. *Environmetrics*, 21, 698–718.
- Reich, B.J., Shaby, B.A. & 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, 119–135.
- Smith, R.L. (2014) Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72, 67–90.
- Wahl, T. & Chambers, D.P. (2015) Evidence for multidecadal variability in US extreme sea level records. *Journal of Geophysical Research: Oceans*, 120, 1527–1544.
- Wahl, T., Haigh, I.D., Nicholls, R.J., Arns, A., Dangendorf, S., Hinkel, J. et al. (2017) Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nature Communications*, 8, 16075.
- Wang, W. & 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. & Legresy, B. (2015) Unabated global mean sea-level rise over the satellite altimeter era. *Nature Climate Change*, 5, 565–568.
- Weisse, R., Bellafiore, D., Menéndez, M., Méndez, F., Nicholls, R.J., Umgessner, G. et al. (2014) Changing extreme sea levels along European coasts. *Coastal Engineering*, 87, 4–14.
- Woodward, W.A., Gray, H.L. & Elliott, A.C. (2017) *Applied time series analysis with R*, 2nd edn. Boca Raton, FL: CRC Press.
- Woodworth, P.L., Hunter, J.R., Marcos, M., Caldwell, P., Menéndez, M. & Haigh, I. (2016) Towards a global higher-frequency sea level dataset. *Geoscience Data Journal*, 3, 50–59.
- Woody, J., Wang, Y. & Dyer, J. (2016) Application of multivariate storage model to quantify trends in seasonally frozen soil. *Open Geosciences*, 8, 310–322.
- Zhang, X.F., Zwiers, F.W. & Li, G. (2004) Monte Carlo experiments on the detection of trends in extreme values. *Journal of Climate*, 17, 1945–1952.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Lee M, Lee J. Long-term trend analysis of extreme coastal sea levels with changepoint detection. *J R Stat Soc Series C*. 2021;70:434–458. <https://doi.org/10.1111/rssc.12466>