AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2015JC011173

Long-term variations in global sea level extremes

Marta Marcos[1], Francisco M. Calafat[2], Ángel Berihuete[3], and Sönke Dangendorf[4]

[1]IMEDEA(UIB-CSIC), Esporles, Spain, [2]National Oceanography Centre, Southampton, UK, [3]Department of Statistics and Operations Research, University of Cádiz, Cádiz, Spain, [4]Research Institute for Water and Environment, University of Siegen, Siegen, Germany

Key Points:
- A state space approach is used to estimate changes in sea level extremes
- Intensity and frequency of extreme sea levels vary coherently at most sites
- Changes in extremes unrelated to mean sea level are regionally consistent

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6
- Figure S7
- Figure S8
- Figure S9
- Figure S10
- Figure s11
- Figure S12
- Figure S13
- Figure S14
- Figure S15
- Figure S16
- Figure S17
- Figure S18

Correspondence to:
M. Marcos,
marta.marcos@uib.es

Citation:
Marcos, M., F. M. Calafat, Á. Berihuete, and S. Dangendorf (2015), Long-term variations in global sea level extremes, *J. Geophys. Res. Oceans, 120*, 8115–8134, doi:10.1002/2015JC011173.

Received 27 JUL 2015
Accepted 17 NOV 2015
Accepted article online 19 NOV 2015
Published online 18 DEC 2015

Abstract Decadal to multidecadal variations in sea level extremes unrelated to mean sea level changes have been investigated using long tide gauge records distributed worldwide. A state space approach has been applied that provides robust solutions and uncertainties of the time evolving characteristics of extremes, allowing for data gaps and uneven sampling, both common features of historical sea level time series. Two different models have been formulated for the intensity and for the occurrence of extreme sea level events and have been applied independently to each tide gauge record. Our results reveal two key findings: first, the intensity and the frequency of occurrence of extreme sea levels unrelated to mean sea level vary coherently on decadal scales in most of the sites examined (63 out of 77) and, second, extreme sea level changes are regionally consistent, thus pointing toward a common large-scale forcing. This variability of extremes associated with climate drivers should be considered in the framework of climate change studies.

1. Introduction

Extreme sea levels resulting from the combined effects of mean sea level (MSL), tidal oscillations, and storm surges are a serious potential threat for coastal environments and assets. Changing climatic and topographic conditions may alter one or more of these contributions to extreme sea levels, thus leading to increased or decreased coastal vulnerability at the seaside. MSL varies in a wide range of spatiotemporal scales due to ocean warming/cooling, water redistribution, and water mass variations associated with changes in the land ice volume and in land hydrology. There is scientific consensus on the rise of MSL since the beginning of the twentieth century [*Church and White*, 2011; *Jevrejeva et al*., 2008; *Hay et al*., 2015] at both global and regional scales, as a consequence of an increased ocean warming and enhanced land-based ice melting. In some regions, the increase in MSL may induce significant changes in tidal ranges as well [*Arns et al*., 2015; *Mawdsley et al*., 2015]. Conversely, long-term changes in storm surges, which are associated with changes in the intensity, occurrence, and paths of storms and low-pressure systems, are more unclear [e.g., *Von Storch*, 2014; *Dangendorf et al*., 2014]. Indeed, the IPCC AR5 [*Church et al*., 2013] establishes only "low confidence" in the long-term trends for these phenomena.

The uncertainty in long-term variability of sea level extremes is partly due to the inherent definition of an extreme as a rare event; the number of extreme events is small (typically a few every year), and this hinders our ability to compute robust statistics on their temporal distribution. Also, remarkable is the scarcity of high-frequency sea level observations, especially when long time series are concerned. The relatively low number of such long records is a constraint for the study of the spatial distribution of changes in extreme events. Nevertheless, long and good quality sea level time series with high-frequency (hourly) sampling do exist and have been explored in earlier works. Global studies have been carried out, for example, by *Menéndez and Woodworth* [2010] who investigated changes in extremes in recent decades using more than 250 tide gauge records, and *Merrifield et al*. [2013] who used 145 good quality stations longer than 10 years to estimate the contributing factors to high waters. Many other works have a more regional or even local focus, mostly because of the geographical variability of extreme sea levels, especially when they are considered in combination with MSL and tides, and sometimes also due to the availability of particular series of data. Examples of recent regional studies are *Marcos et al*. [2009] in Southern Europe; *Thompson et al*. [2013], *Ezer and Atkinson* [2014], and *Wahl and Chambers* [2015] along the U.S. coasts; *McInnes et al*. [2009]

and *Haigh et al*. [2014] in Australia; examples of local studies are *Woodworth and Blackman* [2002] in Liverpool, *Talke et al*. [2014] in New York, and *Dangendorf et al*. [2014] in Cuxhaven, among many others.

Extreme sea levels are characterized by a strong nonstationary nature. There are different approaches to address nonstationarity in extremes: some authors have used parametric methods that assume statistical distributions of sea level extremes varying as linear trends and/or following determined climate proxies [*Menéndez and Woodworth*, 2010; *Letetrel et al*., 2010; *Marcos et al*., 2011; *Cid et al*., 2015]. Other works are based on nonparametric approximations that analyze changes in extremes using sliding time windows [*Mudersbach and Jensen*, 2010; *Talke et al*., 2014; *Wahl and Chambers*, 2015]. In the same line, *Butler et al*. [2007] adopted a locally varying model for the parameters of the distribution. It must be noted that in all these cases a priori constraints are imposed on the way the extremes vary, either because they must follow a given temporal pattern or because the size of the time window conditions the temporal scale of variability. Furthermore, these methods rely on Maximum Likelihood Estimation (MLE) to adjust the observed empirical distribution of extreme values to a theoretical one, typically Generalized Extreme Value (GEV) or Generalized Pareto Distribution (GPD). However, the convergence of MLE is not necessarily ensured, especially when only a few measurements are available. In order to avoid this limitation, we have chosen to use state space models for analyzing changes in both the intensity and the frequency of occurrence of extreme sea levels. The major advantages of this methodology, besides that no assumptions are made regarding the temporal evolution of the extremes, are that uneven data sampling and data gaps are allowed in the time series and that estimates at each time step are derived using all observations available.

Earlier studies have reported long-term changes in extremes that are in most cases driven by MSL variations or changes in tidal constituents [*Zhang et al*., 2000; *Woodworth and Blackman*, 2004; *Marcos et al*., 2009; *Menéndez and Woodworth*, 2010; *Weisse et al*., 2014]. That is, linear trends on the storm surge component are generally nonsignificant. This does not mean that variability does not exist in this contribution, but rather that a linear model is not suitable to describe it. This fact calls for less limited methods to determine the time evolution of the rate of sea level rise than just the fitting of linear and quadratic curves to long time spans of observations. Our purpose here is to examine and describe the observed spatial and temporal variations in both the intensity and the frequency of sea level extremes that are unrelated to MSL changes and tides. To do so, we will base our analyses on long and high-frequency tide gauge measurements covering most regions of the world. We concentrate our assessment on the variability of the storm surge component of extreme sea levels, which, for simplicity, throughout this paper, will be referred to as sea level extreme.

In the following, we describe the data set we use to derive the series of extremes at the tide gauge sites (section 2). Section 3 is devoted to the explanation of the methods used, paying special attention to the state space models that have been applied in the context of extremes. The results are presented for the intensity (section 4) and frequency (section 5) of sea level extremes, followed by a general discussion in section 6. Finally, some concluding remarks summarizing the most important findings are provided in section 7.

2. Sea Level Tide Gauge Records

Hourly sea level observations compiled in the Global Extreme Sea Level Analysis (GESLA) tide gauge data set [*Menéndez and Woodworth*, 2010] have been used. This data set consists of 675 stations of variable quality and length, obtained from the international databases at the University of Hawaii Sea Level Center and the Global Sea Level Observing System, and complemented with additional observations from national data providers. Two more stations not included in GESLA were added: Marseille time series, recently recovered, analyzed, and distributed [*Woppelmann et al*., 2014] and Cuxhaven [*Dangendorf et al*., 2014]. Those stations that were available at the international databases were updated with the most recent years of observations. We initially selected the time series with at least 50 years of data and 70% of completeness, resulting in 122 stations.

Tidal oscillations were removed from each of this initial set of time series using the matlab UTide software [*Codiga*, 2011] applied to the entire time series. One of the advantages of this package with respect to other freely available utilities is that it allows for unevenly distributed temporal sampling in the sea level records. This is a rather important issue when analyzing long time series, as it is common to find time shifts and

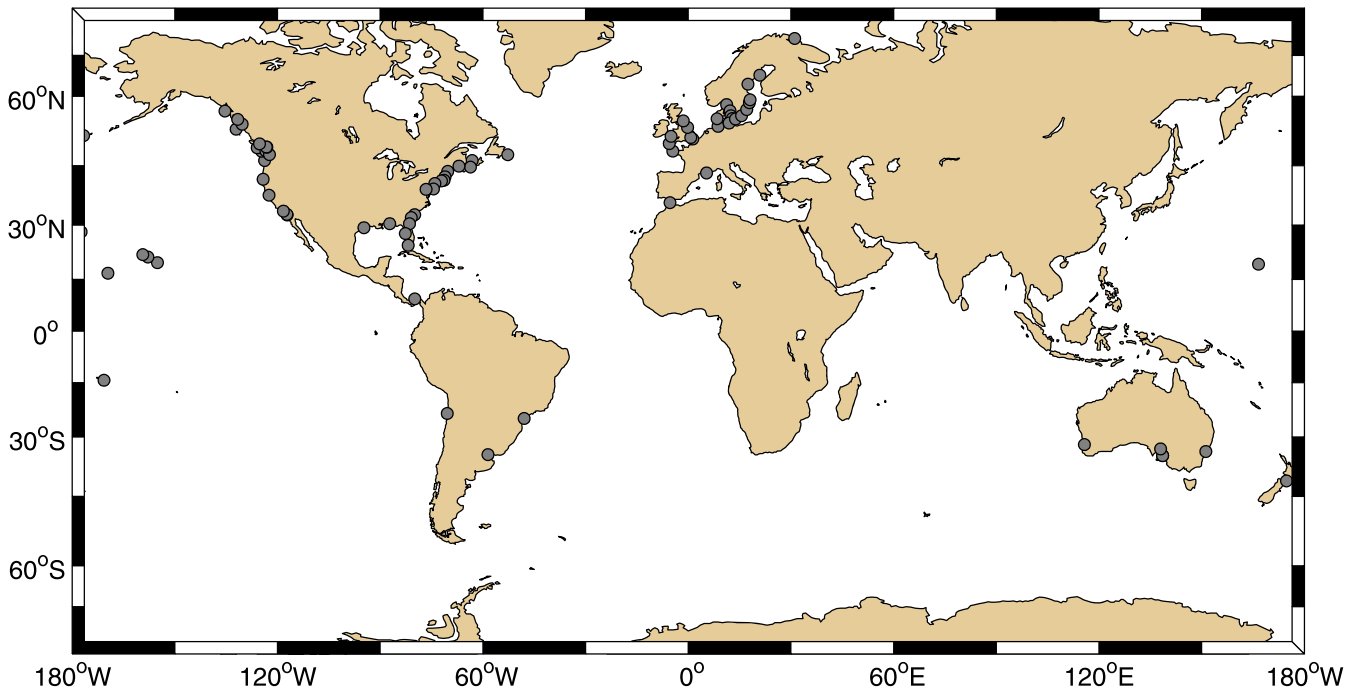

Figure 1. Map with the location of tide gauge stations.

changes in the internal clock of the instrument that would not be accounted for otherwise. Nevertheless, it must be remarked that tidal residuals (i.e., observed sea level minus the tidal components) often contain remaining tidal signals, indicating that the timing of the observations was not perfectly estimated in the historical records. This may represent a shortcoming, especially in regions where tidal oscillations are much larger than the storm surges. Tidal residuals for each record were then visually inspected in order to detect and remove isolated spikes and unrealistic jumps. In some tide gauge stations from the Pacific Ocean, there were clear signatures of tsunamis; although real sea level extremes, these were removed since our focus is on storm surges. There were tide gauge records for which the storm surges were indistinguishable from the residual tides (Balboa station is an example of large tides of $\sim$5 m and small surges of only a few cm) and therefore these were withdrawn from our analysis. The final set resulted in 77 tide gauge stations for which we could reliably compute the storm surge component. Their spatial distribution is mapped in Figure 1 and they are listed in Table 1 together with their location and period of operation. There is a geographical bias of the tide gauge stations, with the majority of them being located along the European and North American coasts. The bias is not the result of our stringent selection criteria, but rather related to the limitation of the historical sea level data set [*Holgate et al.*, 2013].

3. Methods

Hourly tidal residuals were deseasoned by removing the mean annual and semiannual components estimated by the means of a harmonic analysis. As our focus is on changes in storminess that are unrelated to MSL variations, we have also removed this contribution. To do so, MSL at each station was computed applying a Butterworth low-passed filter of order 2 and 1 year cutoff period to the detided and deseasoned time series. Hourly surge records were built by removing the MSL signal from the detided and deseasoned time series and were then used to generate the two complementary data sets described below.

The first data set is aimed at representing the changes in the intensity of the extreme sea levels. It consists of selecting the five highest values per year for each hourly surge record. To ensure the independence of events, a separation of at least 72 h between two successive extremes was imposed. Years were defined as starting on 1 April in order to avoid splitting either boreal or austral winters into subsequent blocks. This is especially relevant for stations located at mid and high latitudes, for which extreme sea level episodes

Table 1. Tide Gauge Data Set, Location, and Period of Operation

Name	Latitude (°N)	Longitude (°E)	Period
Marseille	43.2790	5.3540	1849–2012
Ceuta	35.9000	−5.3167	1944–2010
Brest	48.3830	−4.5000	1846–2011
Newlyn	50.1000	−5.5400	1915–2009
Milfordhaven	51.7023	−5.0143	1953–2006
Dover	51.1144	1.3225	1934–2006
Sheerness	51.4431	0.7500	1952–2006
Immingham	53.6329	−0.1869	1953–2006
North Shields	55.0074	−1.4398	1946–2006
Cuxhaven	53.8667	8.7167	1918–2014
Esbjerg	55.4667	8.4333	1950–2001
Smogen	58.3667	11.2167	1910–2008
Varberg	57.1000	12.2167	1886–1982
Hornbaek	56.1000	12.4667	1891–2011
Klagshamn	55.5167	12.9000	1929–2008
Gedser	54.5667	11.9333	1891–2011
Ystad	55.4167	13.8167	1886–1987
Kungsholmsfort	56.1000	15.5833	1886–2008
Olandsnorraudde	57.3667	17.1000	1886–2008
Landsort	58.7500	17.8667	1886–2006
Stockholm	59.3200	18.0800	1889–2011
Draghallan	62.3333	17.4667	1898–1966
Ratan	64.0000	20.9167	1891–2008
Vardo	70.3333	31.1000	1947–2003
St. Johns	47.5667	−52.7167	1935–2008
Charlottetown	46.2333	−63.1167	1911–2008
Halifax	44.6800	−63.6100	1920–2010
Eastport	44.9100	−66.9900	1929–2011
Portland	43.6600	−70.2500	1910–2011
Boston	42.3500	−71.0500	1921–2011
Newport	41.5100	−71.3330	1930–2011
New London	41.3550	−72.0867	1938–2012
New York	40.7000	−74.0200	1920–2011
Atlantic City	39.3500	−74.4200	1911–2011
Baltimore	39.2667	−76.5833	1902–2006
Charleston	32.7820	−79.9250	1921–2011
Fort Pulaski	32.0330	−80.9020	1935–2011
Mayport	30.3950	−81.4317	1928–2000
Key West	24.5530	−81.8080	1913–2011
St. Petersburg	27.7600	−82.6267	1946–2012
Pensacola	30.4030	−87.2120	1923–2011
Galveston	29.3300	−94.7400	1904–2011
Cristobal	9.3670	−79.8830	1907–2011
Cananeia	−25.0167	−47.9250	1954–2005
Buenos Aires	−34.6667	−58.5000	1905–1961
Antofagasta	−23.6500	−70.4000	1945–2002
San Diego	32.7130	−117.1730	1906–2011
La Jolla	32.8670	−117.2500	1924–2011
Los Angeles	33.7170	−118.2720	1923–2011
San Francisco	37.8000	−122.4670	1897–2011
Crescent City	41.7450	−124.1833	1933–2012
Astoria	46.2200	−123.7670	1925–2011
Neah Bay	48.3683	−124.6167	1934–2012
Tofino	49.1540	−125.9130	1909–2008
Victoria	48.4200	−123.3700	1909–2011
Seattle	47.6000	−122.3333	1901–2006
Vancouver	49.2870	−123.1100	1909–2008
Point Atkinson	49.3370	−123.2530	1914–2008
Campbell River	50.0420	−125.2470	1967–2008
Queen Charlotte City	53.2520	−132.0720	1957–2008
Prince Rupert	54.3170	−130.3240	1908–2008
Ketchikan	55.3333	−131.6250	1918–2012
Sitka	57.0517	−135.3417	1938–2012
Adak	51.8633	−176.6317	1950–2012
Hilo	19.7333	−155.0667	1927–2012
Honolulu	21.3000	−157.8600	1905–2011
Nawiliwili Bay	21.9667	−159.3500	1954–2012
Johnston Is.	16.7383	−169.5300	1947–2012

mostly occur during the winter season [*Menéndez and Woodworth*, 2010]. Only those years with at least 7 months of data were considered in the analyses.

The second data set accounts for changes in the occurrence of extreme sea level events. It was constructed selecting the number of events per year over a given threshold. For this purpose, three different thresholds were used, namely 99th, 99.5th, and 99.9th percentiles computed over the entire period of the record, representing from moderate to strong episodes. The same event-independence criterion as in the first data set was applied.

In the following, we use the state space approach for extreme sea level time series defining suitable models for each of the two data sets. The state space model of a continuous-time dynamic system is a useful framework for treating complex physical problems in the time domain defined by a first-order differential equation and a relation between the observations and the unobserved states. The first-order differential equation is called the state equation and is generally written as:

$$\dot{\theta} = G(\theta, t) + \omega, \quad \omega \sim \mathcal{N}_d(0, W(t)) \quad (1)$$

and the relationship between the observations and states is the observation equation:

$$y(t) = F(\theta, t) + v, \quad v \sim \mathcal{N}_m(0, V(t)) \quad (2)$$

where $y(t)$ are the observations at time, θ are the unobserved states, and v and w are zero mean Gaussian errors with covariance matrices V and W, respectively. In order to deal with a tractable problem, usually functions F and G are discretized and the differential equation in (1) is rewritten as a difference equation according to:

$$\theta_{t+1} = G_t(\theta_t) + \omega_t, \quad \omega_t \sim \mathcal{N}_d(0, W_t) \quad (3)$$

$$y_t = F_t(\theta_t) + \upsilon_t, \quad \upsilon_t \sim \mathcal{N}_m(0, V_t) \quad (4)$$

In general, our goal will be to estimate the posterior probability density function of the state θ given all observations $p(\theta_t|y_{1:n})$, which is provided by the

Table 1. (continued)

Name	Latitude (°N)	Longitude (°E)	Period
Midway Is.	28.2167	−177.3667	1947–2012
Wake Is.	19.2833	166.6167	1950–2012
Pago Pago	−14.2833	−170.6833	1948–2012
Wellington Harbour	−41.2833	174.7833	1944–2013
Fort Denison	−33.8550	151.2256	1914–2004
Port Adelaide_inner	−34.9258	138.5997	1933–1999
Port Adelaide_outer	−34.9258	138.5997	1940–2004
Port Pirie	−33.1770	138.0102	1941–2004
Fremantle	−32.0531	115.7459	1897–2004

Bayesian methodology. When analyzing climatic time series, there are two major advantages of space state modeling: first, no assumptions are made regarding the temporal variability of the observations and, second, uneven data sampling and gaps in the time series are allowed (although the price to pay is an increase in the uncertainties).

A well-known model is the Gaussian state space model with continuous states and discrete time (GSSM) that corresponds to the case in which G and F are linear operators. Then equations (3) and (4) adopt in this case the following form:

$$\theta_{t+1}=G_t\theta_t+\omega_t,\ \ \omega_t\ \sim\ \mathcal{N}_d(0,W_t) \tag{5}$$

$$y_t=F_t\theta_t+\upsilon_t,\ \ \upsilon_t\ \sim\ \mathcal{N}_m(0,V_t) \tag{6}$$

The GSSM will be completely described by G_t, F_t, W_t, V_t, and the initial state value θ_0. Under these assumptions, the Kalman Filter method [*Kalman*, 1960] provides efficient recursive formulas for calculating the posterior probability density functions of the state θ_t given the history of observations up to the time step t $p(\theta_t|y_{1:t})$, which are also Gaussian. The Kalman Filter uses past and current observations to predict the current states; an extension of this is the Kalman Smoother, which uses all available observations at all times to generate the posterior density given all observations $p(\theta_t|y_{1:n})$. Despite its simplicity, this formulation has many applications in real environmental problems [e.g., *Lee and Berger*, 2003; *Kurtenbach et al*., 2012; *Laine et al*., 2014]. However, all the strength of this methodology relies on the linear-Gaussian properties that do not always hold, especially when extreme events are concerned.

Our aim is to determine the latent nonstationary time evolution of the intensity and occurrence of sea level extremes. To do so, we will assume the following state equations:

$$\alpha_{t+1}=\alpha_t+\beta_t+\omega_{t,1},\ \ \omega_{t,1}\ \sim\ \mathcal{N}(0,W_t) \tag{7}$$

$$\beta_{t+1}=\beta_t+\omega_{t,2},\ \ \omega_{t,2}\ \sim\ \mathcal{N}(0,W_t) \tag{8}$$

That is, the time series is changing as a Markovian local level model α_t affected by β_t which is regarded as the velocity of those changes. Equations (7) and (8) are related to equation (5) by:

$$\theta_t=\begin{pmatrix}\alpha_t\\ \beta_t\end{pmatrix},\ \ G_t=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}.$$

The statistical properties of the data of each set of observations considered here are clearly differentiated and therefore the observation equation (6) will take different forms in each case according to:

$$p(y_t|\theta_t)=p\left(y_t|G_t(\alpha_t\ \ \beta_t)^T\right) \tag{9}$$

While the highest extremes per year follow a Generalized Extreme Value (GEV) Distribution, the number of extremes per year can be represented by Poisson distributions. We have therefore analyzed the two data sets separately, as described below.

3.1. Space State Model of Extremes Intensity

When assessing the changes in the intensity of the extremes, the observation equation (9) is defined as:

$$y_t|\mu_t\ \sim\ GEV(\mu_t,\sigma,\zeta) \tag{10}$$

That is, the intensities of the extremes follow a GEV distribution with location μ_t, scale σ, and shape ζ parameters. Among these three parameters, only the location, which controls the magnitude of the extremes, is considered to be time varying. The shape parameter has been shown to remain essentially

constant [*Davison and Ramesh*, 2000]. Some tests were carried out allowing temporal variations in the scale parameter, but these showed that it did not change more than 3% with respect to its averaged value and in turn increased significantly the confidence intervals; therefore, it was kept constant too. The state equations are defined thus for the location parameter, that is, in equation (7), we define $\alpha_t = \mu_t$. Here the location parameter is modeled by means of a nearly constant velocity model, with the 2 × 2 covariance matrix defined as:

$$W_t = q\begin{pmatrix} \Delta t^3/3 & \Delta t^2/2 \\ \Delta t^2/2 & \Delta t \end{pmatrix} \tag{11}$$

Note that the observations y_t here are five-variate observations in time, corresponding to the five extremes per year of the data set.

Once the probability distribution function of the location parameter is known, the results can be expressed in terms of the return levels for a given period. The M-year return levels, defined as the probability of a given value to be exceeded once every M years, at each time step t are computed as:

$$RL = \mu_t - \frac{\sigma}{\varsigma}\left[1 - \left(-log\left(1 - \frac{1}{5M}\right)\right)^{-\varsigma}\right].$$

Therefore, return levels will follow the same temporal variability as the location parameter.

As commented above, our aim here is to make inferences about the states conditional on all observations. Note, however, that the scale factor q of the error covariance matrix W_t in equation (11) is unknown and thus needs to be determined. Note also that our observation equation is not Gaussian and thus the Kalman Filter cannot be used to solve the model. Here we opt to use a novel technique named particle Markov chain Monte Carlo (PMCMC) sampler to solve our state space model. A PMCMC sampler is a method that combines a particle filter with a Markov chain Monte Carlo sampler to make statistical inferences [*Andrieu et al*., 2010], and is especially fitted to address the problem of simulating state trajectories in state space models. One of the key features of the PMCMC is that it samples directly from the joint posterior density of the states and the parameters. In our case, this means that the PMCMC allows us to sample from $p(\theta_t, q|y_{1:n})$, which implies that the scale factor q is updated together with the states. In particular, here we use a state-of-the-art PMCMC sampler referred to as particle Gibbs with ancestor sampling (PGAS) [*Lindsten et al*., 2014]. PGAS shows good mixing of the Markov kernel as compared to other particle Gibbs sampler enabling us to use a smaller number of particles and thus to reduce the computational cost of the algorithm.

The PGAS algorithm that we use is the same as in section 5.1 (algorithm 3) of *Lindsten et al*. [2014]. Here we use 500 particles and the number of iterations for the MCMC is set to 20,000 with a burning period of 2000. Here we use multinomial resampling as the resampling scheme. Finally, we ascribe an inverse Wishart prior distribution (one dimensional) to the parameter q. The inverse Wishart distribution is convenient because it is a conjugate prior for the likelihood function in our state equation, making it easy to sample from the posterior $p(q|\theta_{1:n}, y_{1:n})$ in our PGAS implementation.

In order to assess the skill of the PGAS method, we have tested its performance on synthetic time series in which the location parameter has been predefined. We have used random walks to generate two 150 years long time series of the location parameter, showing three and one oscillations, respectively (black curves in Figure 2). For each case, a GEV distribution with realistic scale and shape parameters was used to generate 100 time series of five extremes per year. Each of these series was then analyzed using the PGAS algorithm to estimate the time-varying location parameter. The comparisons between the true predefined location parameter and the estimations for the 100 synthetic time series are plotted in Figure 2, when the five extremes per years were used in the model (red curves). The 95% probability envelopes are also plotted for the time series having the highest correlation with the true values. The fitting between the calculated and the true location parameters is very satisfactory in both cases with very little dispersion in the 100 realizations. Note also that most curves fall within the 95% confidence interval, indicating that the estimated confidence intervals provide a realistic measure of the robustness of our estimates. Also, the correspondence among the estimated curves is consistent, showing average correlations of 0.87 and 0.81, respectively. For completeness, we have also computed the location parameters applying the PGAS method to one extreme

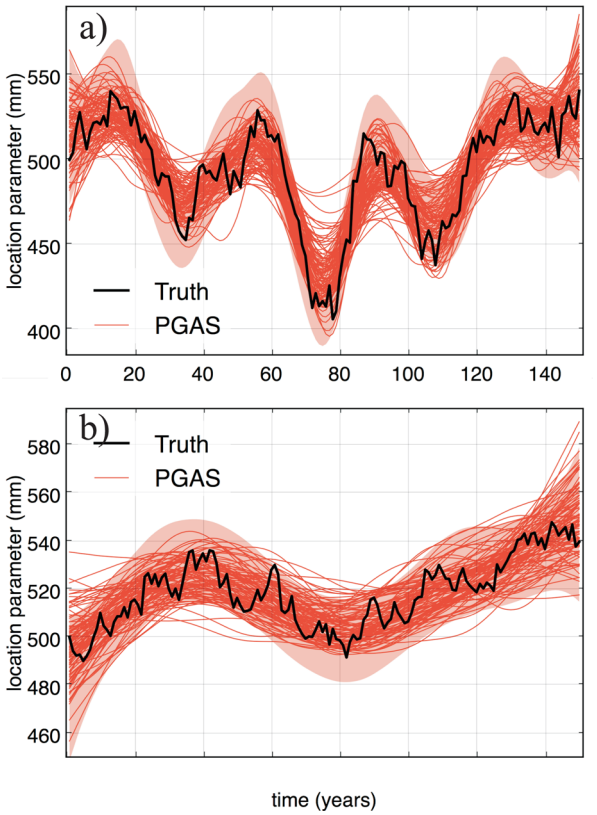

Figure 2. Synthetic time series of the location parameter generated with a random walk (black curves), and estimations of the location parameter using PGAS for 100 time series (red lines) using five extremes per year. Shading corresponds to the 95% probability envelopes of series that is most correlated with the true values.

per year. Results (supporting information Figure S1) show weaker correspondence and consequently lower correlations of 0.59 and 0.61. Therefore, in our analyses, we have chosen five extremes per year.

3.2. Space State Model of Extremes Occurrence

Occurrences of extreme events follow a Poisson distribution, where the Poisson counts are assumed to be conditionally independent over time given the latent process. Therefore, the observational density in equation (9) is defined by:

$$y_t|\vartheta_t \sim Poisson\left(u_t \cdot e^{\vartheta_t}\right) \qquad (12)$$

where u_t is the exposure, namely the population depending on time, which in our case is constant and has no impact on the computations ($u_t = 1$). And $\vartheta_t = log(\lambda_t)$, with λ_t being the intensity value in the Poisson distribution, i.e., the mean number of extreme events per unit of time. In other words, $(y_t|\vartheta_t) = u_t e^{\vartheta_t} = u_t \lambda_t$. The state equation (7) is defined in this case for the intensity value as:

$$\vartheta_t = \alpha_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}\left(0, \sigma_\epsilon^2\right)$$

with ε_t being an additional white noise component which captures the extra variation of the series. Furthermore, in this case, the errors $w_{t,i}$ in equations (7) and (8) are uncorrelated normal distributions, implying that the covariance matrix W_t is diagonal and takes the form:

$$W_t = \begin{pmatrix} \sigma_\alpha^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

The Poisson state space model is solved by using the methodology applied to the exponential family of state space models. The algorithm used here is the KFAS (Kalman Filter and Smoother for Exponential Family State Space Models) algorithm developed by *Helske* [2015] and available at http://cran.r-project.org/package=KFAS. The Poisson model is approximated by a Gaussian model by replacing the observations by pseudoobservations using the log-link function $\vartheta_t = log(\lambda_t)$. Estimates for ϑ_t are computed by using Kalman filtering and smoothing from the approximating Gaussian model, based on the simulation smoothing algorithm by *Durbin and Koopman* [2002].

In order to ensure the suitability of the model, the standardized recursive residuals were computed for each time series. These were checked to have approximately zero mean and unit variance as well as nonsignificant autocorrelations, which were found to be true for all considered time series.

4. Changes in Extreme Intensity

Results of the PGAS algorithm are illustrated in Figure 3 for a selected set of tide gauge stations representative of different regions. Supporting information Figure S2 contains the results for the rest of the stations of our data set. Changes in the distribution of the location parameter are represented with their 95% probability envelope (blue curves). Likewise, changes in 50 years return levels (RL50) that are derived from the time-

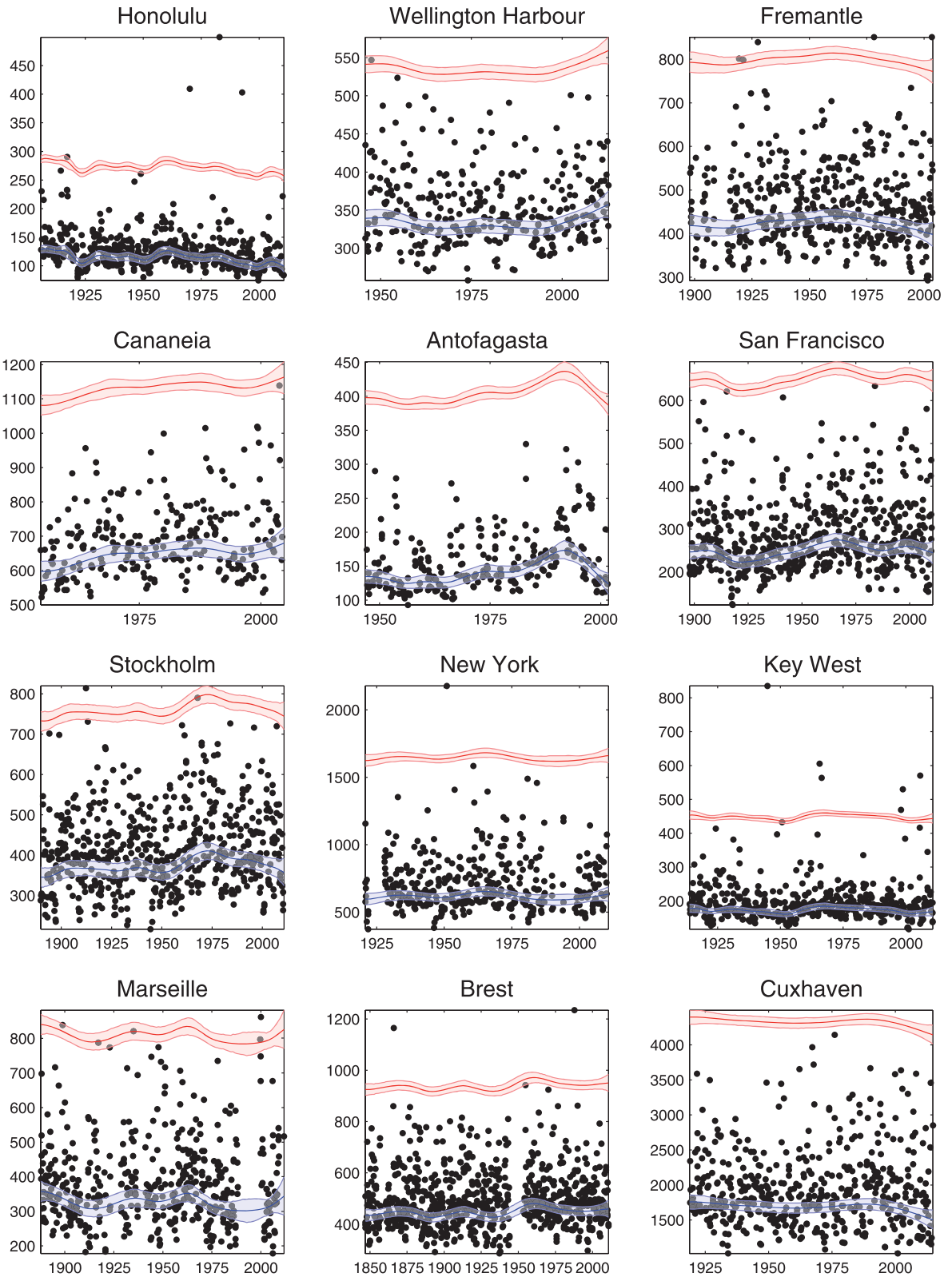

Figure 3. Five highest extremes per year (black dots) and their corresponding time-varying location parameter (blue) and 50 years return level (red) with the 95% confidence interval. Units are in mm.

varying probability distribution functions are also shown with their 95% probability (red curves). Note that the PMCMC is able to estimate the parameters of the distribution even in absence of observations, although with larger confidence intervals. The results for each individual station highlight the nonstationary nature of the extreme distributions, with changes that are often above their 95% uncertainty ranges. Decadal variability becomes evident in most stations, as in San Francisco or Marseille (Figure 3), two of the longest high-frequency tide gauge records. Gradual centennial trends are, on the contrary, more difficult to identify and can be detected in only a few locations. For instance, this is the case at Honolulu and Cananeia stations (Figure 3).

Values of RL50 shown in Figure 3 and supporting information Figure S2 have very different ranges, from less than 40 cm in Honolulu, La Jolla, Nawiliwili Bay, and Pago Pago to more than 4 m in Cuxhaven. The amplitudes of the extreme surges are largely controlled by the intensity of the storms hitting the coast and by morphological features such as continental shelves. Highest extremes are found at high-latitude stations, as in the North and the Baltic Seas, in agreement with previous studies that attributed sea level extreme intensity with atmospheric activity (e.g., *Dangendorf et al*. [2013], in the German Bight; *Merrifield et al*. [2013], globally). However, atmospheric storminess is not the only element that impacts on the extreme intensity, as demonstrated the fact that nearby stations that are subject to the same atmospheric forcing may have significantly different averaged return levels. Other factors, such as local bathymetry and orientation also play a role in the amplitude of the surges, as well as the shape of the ocean in front of the coast (i.e., in shallow shelf seas surges become much higher than in open ocean regions).

Given the large differences of RL50 among stations, even at the regional scale, the consistency of the observed temporal variability in extreme intensity is explored on the basis of the changes in RL50 relative to the mean value at each station. The results are plotted in Figure 4, in which stations have been sorted following the coastlines (the same order as in Table 1, see also Figure 5 for approximate location). Once the tide gauge stations are ordered regionally, coherent decadal to multidecadal variations of extreme intensities become evident. For example, around the mid-twentieth century a decrease in extreme intensity (in RL50) is observed in northern Europe, whereas an increase is found in Southern Europe. Interestingly, during the same period, positive values are also observed along the North American coasts in the Atlantic and the Pacific. From Figure 3, it is evident that a linear trend is not able to describe the long-term behavior in the intensity of sea level extremes. This is further demonstrated by computing the averaged changes for the entire period, plotted a squares in Figure 4.

A complementary way of highlighting the decadal variability in sea level extremes is presented in Figure 5. Averaged changes in RL50 for 20 year periods have been computed at each station by fitting a linear trend for the selected period and multiplying the slope by the period length. Three periods have been chosen, based on the variability shown in Figure 4, namely 1940–1960, 1970–1990, and 1990–2010. Only those stations with at least 15 years of valid data during the corresponding period have been included. According to the maps shown in Figure 5, 1990–2010 has been a period of enhanced storminess in southern Europe and the North American coasts, when compared to previous decades; however, this decadal increase in storminess is comparable to the changes recorded during the mid-twentieth century in the same regions.

Our results contrast with some earlier works. For example, *Mudersbach and Jensen* [2010] analyzed the tide gauge in Cuxhaven using annual maxima fitted to a GEV distribution with time-varying location and scale parameters. They found relatively large values of the location parameter prior to 1950 and recently. However, they did not remove MSL, which has been shown to drive to a large extent the magnitude of extremes [*Marcos et al*., 2009; *Menéndez and Woodworth*, 2010], explaining thus the discrepancies with our results. *Butler et al*. [2007] analyzed decadal variations in storm surges in the North Sea. At Dover station, their results show a distinct behavior in terms of return level variability. However, they used the 20 largest values per year, thus focusing on relatively moderate extreme events. Furthermore, they used a likelihood approach, which as explained below may also impact the results. *Talke et al*. [2014] analyzed a long reconstructed sea level time series at New York and obtained the same evolution as that depicted by our results. In a recent paper, *Wahl and Chambers* [2015] analyzed extreme sea levels unrelated to MSL along the U.S. coastlines. We found similar temporal evolution in return levels in some of the common stations (their Figure 8), such as Key West, Mayport, Boston, and Portland; but we also found very distinct behavior in other stations such as Victoria and Astoria. There are three major differences between our approach and that followed by *Wahl and Chambers* [2015]: the separation of their analyses into seasons, the number of extremes per year selected, and the methodology applied to extract return levels which in their case

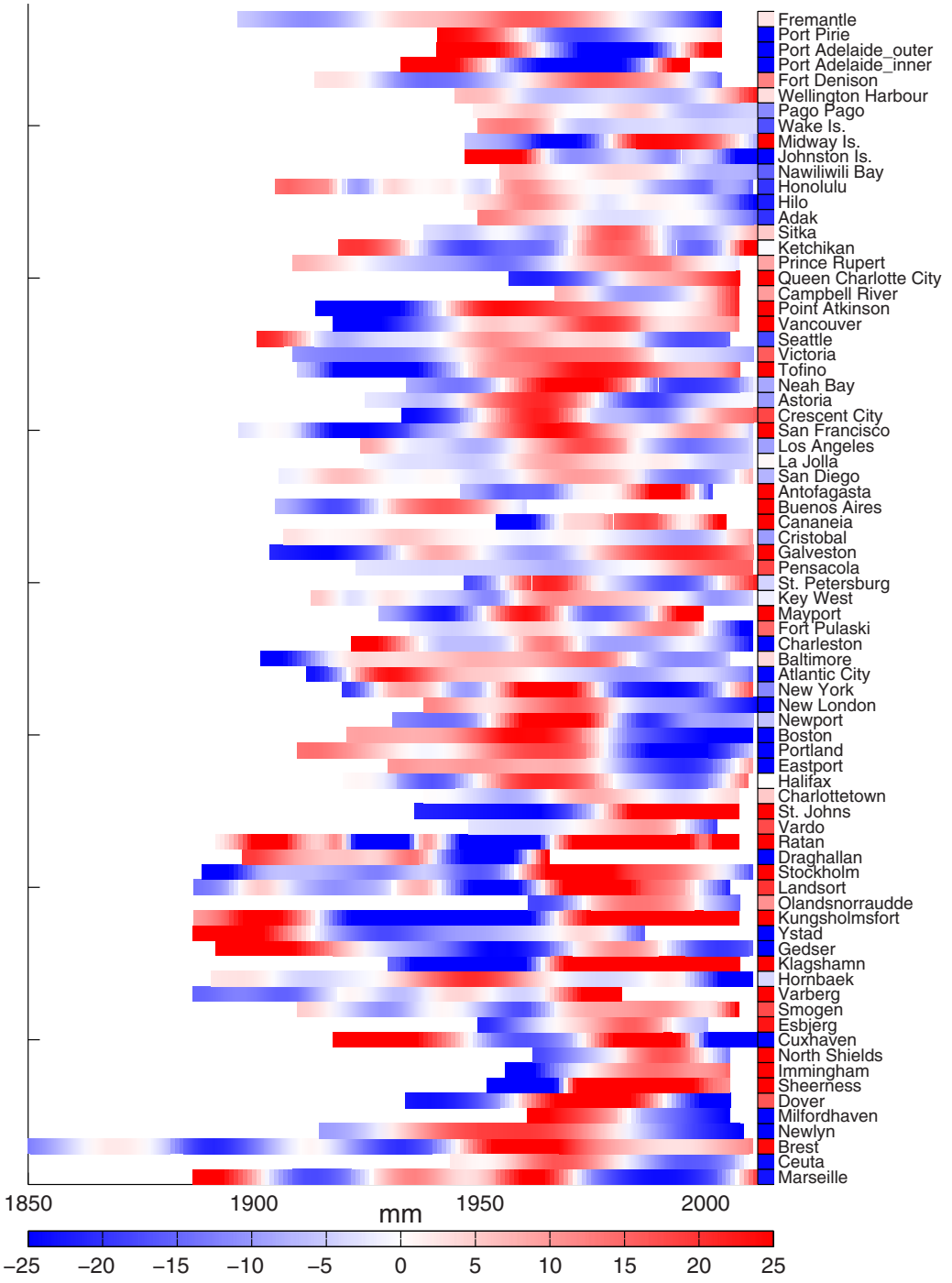

Figure 4. Changes in RL50 relative to their mean values at each tide gauge station. Squares on the right represent the changes in RL50 fitting a linear trend. Units are in mm.

corresponds to a MLE of the GEV parameters for time windows of 37 years. The first factor may be relevant for low-latitude stations, in which seasonal extremes are equally important in summer and winter and they are originated by different mechanisms. The two latter can be critical. To illustrate this, we analyzed the synthetic series described in section 3.1 using MLE with sliding time windows of 37 years with five and one extremes per year. The results are shown in supporting information Figure S3 and demonstrate that this method has limited ability to capture the time-varying changes in the location parameter, especially when only one extreme per year is considered (supporting information Figures S3c and S3d). Average correlations

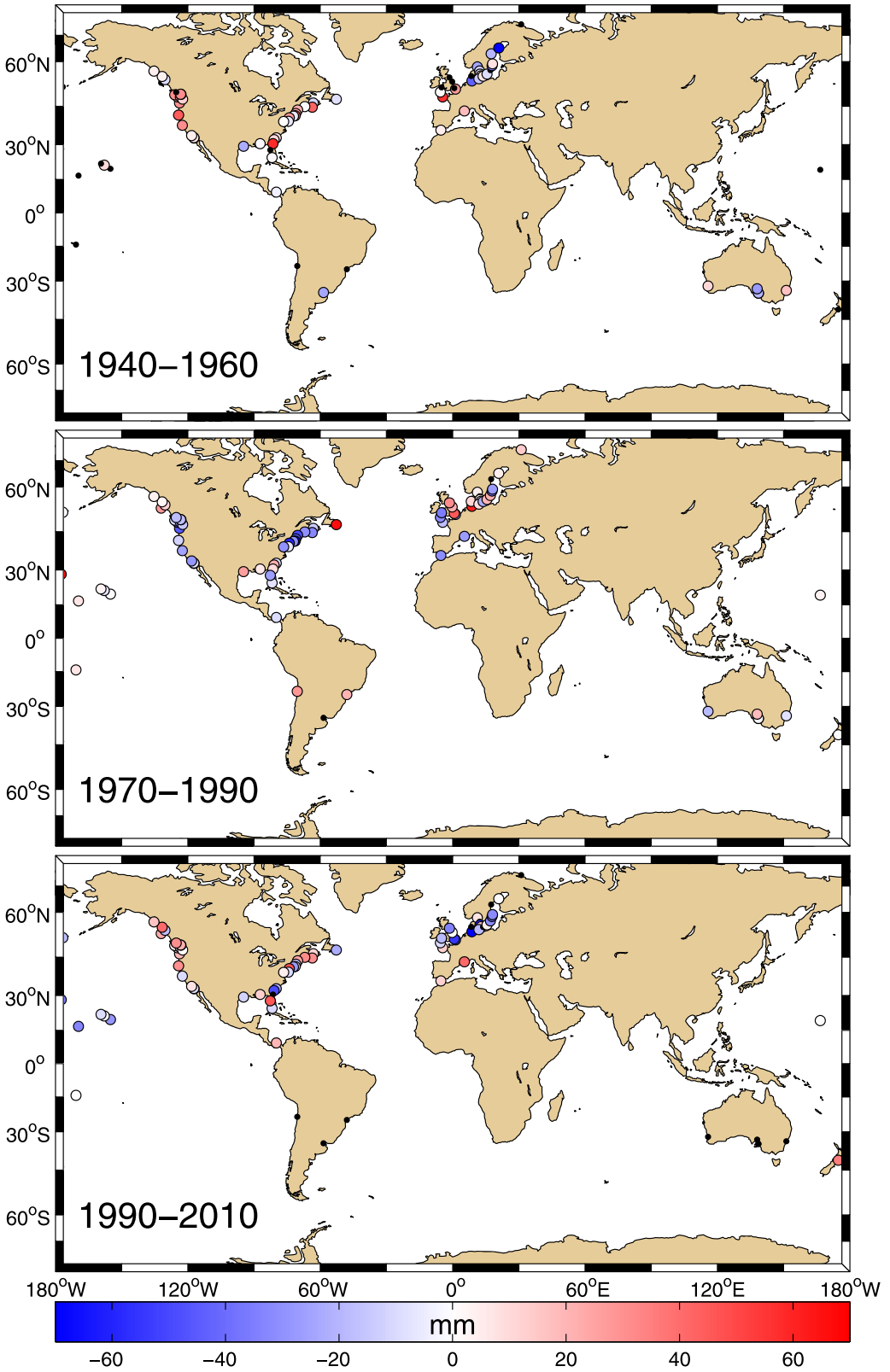

Figure 5. Averaged changes of RL50 (in mm) at each tide gauge station for the periods quoted. Those stations with less than 15 years for the selected period are indicated by black dots.

for the two cases are 0.29 and 0.36 with one extreme per year and 0.44 and 0.59 with five extremes per year. Results can be satisfactory only when the variations in return levels are longer than the time window used and the number of extremes is large enough (e.g., supporting information Figure S3b).

5. Changes in Extreme Occurrence

The variability in the occurrence of extremes is shown in Figure 6 for the same selected set of representative stations as in Figure 3. Results for the rest of locations can be found in the supporting information (Figure S4). The variations in the number of events per year, together with their 95% confidence bounds are shown for events exceeding the 99th, 99.5th, and 99.9th percentiles. In the same line as with the changes in intensity, the extreme episodes often present decadal to multidecadal variations in their frequency that are larger than the associated uncertainties. In order to improve the representation of the decadal variability of the number of events, the changes are plotted with respect to their average value and the stations have been sorted regionally in the same manner as in Figure 4. The results for the number of events exceeding the 99.5th percentile are plotted in Figure 7 and reveal, once again, a clear spatial pattern. This spatial structure moreover matches the pattern obtained with the intensity of extremes (Figure 4), indicating coherent changes in strength and frequency in sea level extremes. This point will be further discussed in the next section. Overall, the picture is the same when the 99th or 99.9th percentiles are used as thresholds, indicating that in most cases moderate and strong events vary accordingly at decadal time scales. However, there are exceptions to this rule, especially when the strongest events are concerned, as discussed below.

We have explored the similarities and differences between changes in moderate and strong events using the frequency of episodes exceeding the 99th, 99.5th, and 99.9th percentiles as resulting from our analysis. As already shown by *Dangendorf et al*. [2014] the use of different thresholds may lead to different definition of extreme sea levels. We found that in 67 tide gauge stations (87% of the total number) the three curves have correlations higher than 0.8, after detrending. However, if the frequencies of occurrence are approximated by a linear fit, only 50 stations display the same sign for all three curves. An example of this is the case of Marseille (Figure 6), where the number of events exceeding the 99.9th percentile progressively increases (from an average value of 0.7–0.8 events per year during its operation period), whereas the events exceeding 99th and 99.5th percentiles decrease by a larger amount (from 3.2 to 2.7 extremes per year over the 99.5th percentile). Another interesting case is San Diego (supporting information Figure S4), where moderate events exceeding the 99th and 99.5th percentiles present decadal variations with an overall negative trend enhanced after 1975, while the strongest extremes show a gradual increase in the number of episodes (from 1 to 1.3 events per year on average).

6. Consistency Between Extreme Intensity and Frequency

Decadal to multidecadal variability in extreme sea levels unrelated to MSL has been found to show coherent changes in both intensity and frequency. In other words, for the majority of the locations within this study, when the intensity of extremes increases, so do their probability of occurrence, and vice versa. This common behavior of the intensity and the frequency of extremes is hinted from Figures 4 and 7, which display coincident temporal patterns. In order to better characterize these similarities, the correlation coefficients between the RL50 and the number of events exceeding the 99.5th percentile for each tide gauge station have been computed and are plotted in Figure 8. Equivalent results are obtained when the other two percentiles are used as thresholds. For representation purposes, the correlations of those stations exhibiting negative values (4 out of 77) have been set to zero. The statistical significance of the correlations has been computed following *Ebisuzaki* [1997], a nonparametric method in which a large (10,000 in our case) set of synthetic series is built from the original one using its Fourier transform with the amplitudes of the Fourier components being preserved but their phases taking random values. Stations showing significant correlations at the 90% confidence level are filled in Figure 8. As a result, in 63 stations (82% of the total), the correlations between RL50 and the occurrence of extremes are significant at the 90% confidence level, with a mean value of 0.86. When the same computation is done for the number of extremes exceeding the 99.9th threshold, the number of stations with significant correlations is 53 (69%), in agreement with the differences in occurrence between moderate and strong events reported above. There is no geographical grouping of stations for which intensity and frequency of extremes are uncorrelated, indicating that these effects are likely local.

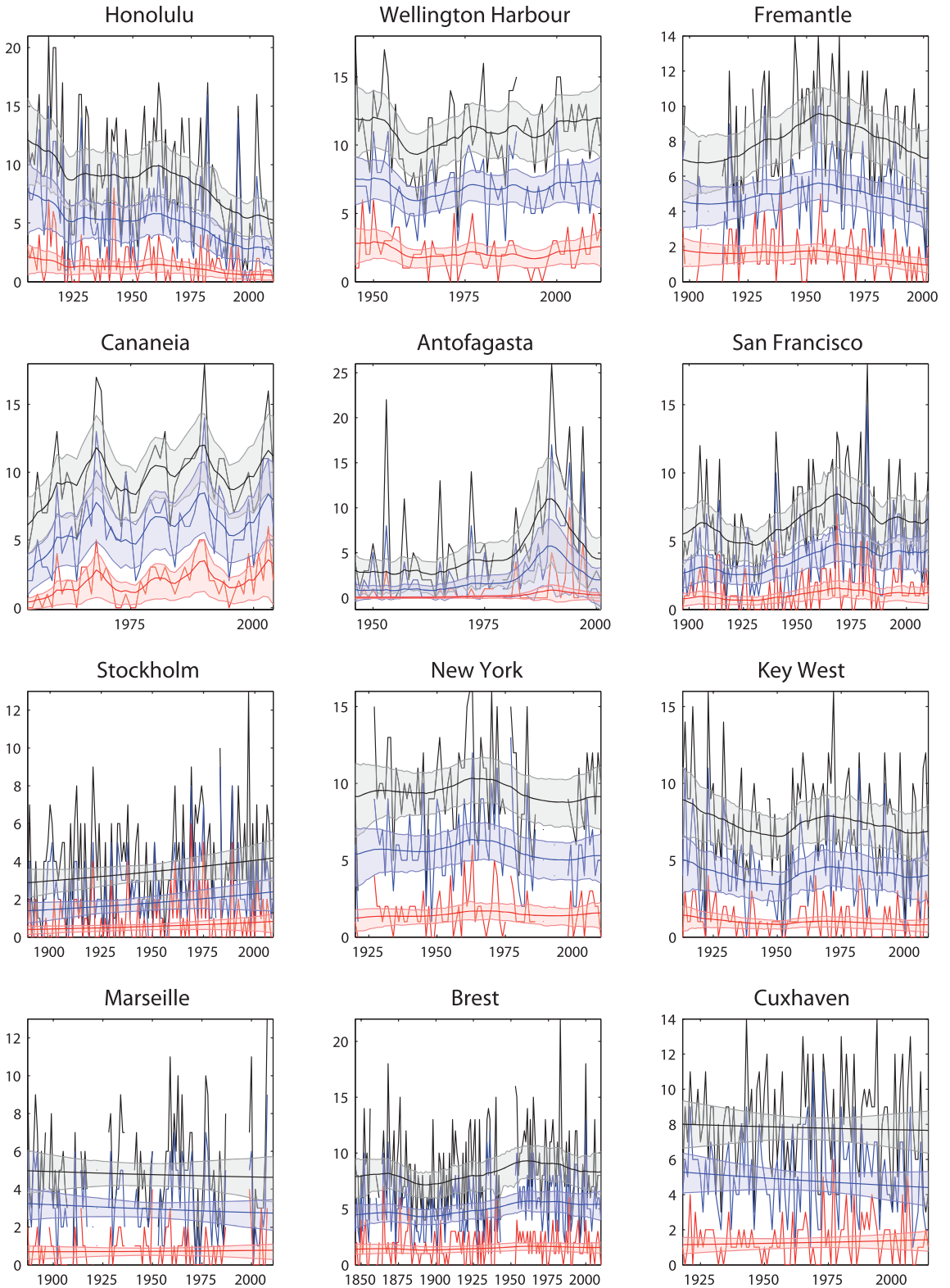

Figure 6. Extreme occurrences (in numbers per year) above the 99th (black), 99.5th (blue), and 99.9th (red) percentiles and their smoothed curves together with the 95% confidence intervals.

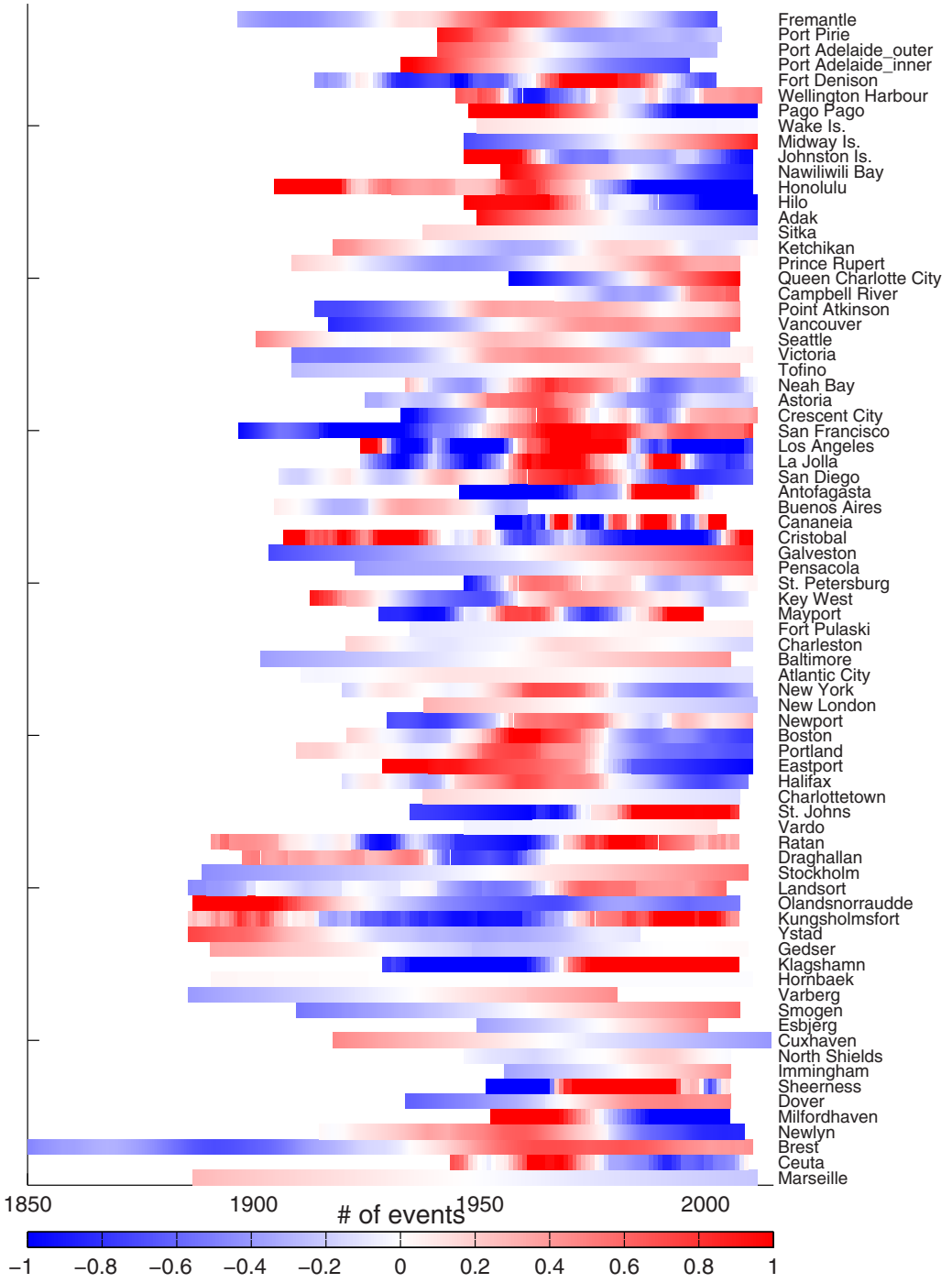

Figure 7. Changes in extreme occurrence over the 99.5th percentile at each tide gauge station relative to the corresponding average value.

7. Multidecadal Changes in Extremes and Climate Indices

An interesting feature of the observed changes in extreme sea levels is the regional consistency that they exhibit in Figures 4 and 7, when stations are grouped geographically. This evidence suggests that the decadal variability in extremes is controlled by large-scale atmospheric forcing. This is in line with previous studies with a regional focus [e.g., *Zhang et al.*, 2000; *Sweet and Zervas*, 2011; *Dangendorf et al.*, 2014; *Wahl and Chambers*, 2015]. We have therefore explored this mechanism seeking relationships between extreme variations and large-scale climate indices. For simplicity and given the strong link between changes in extreme intensity and frequency, we

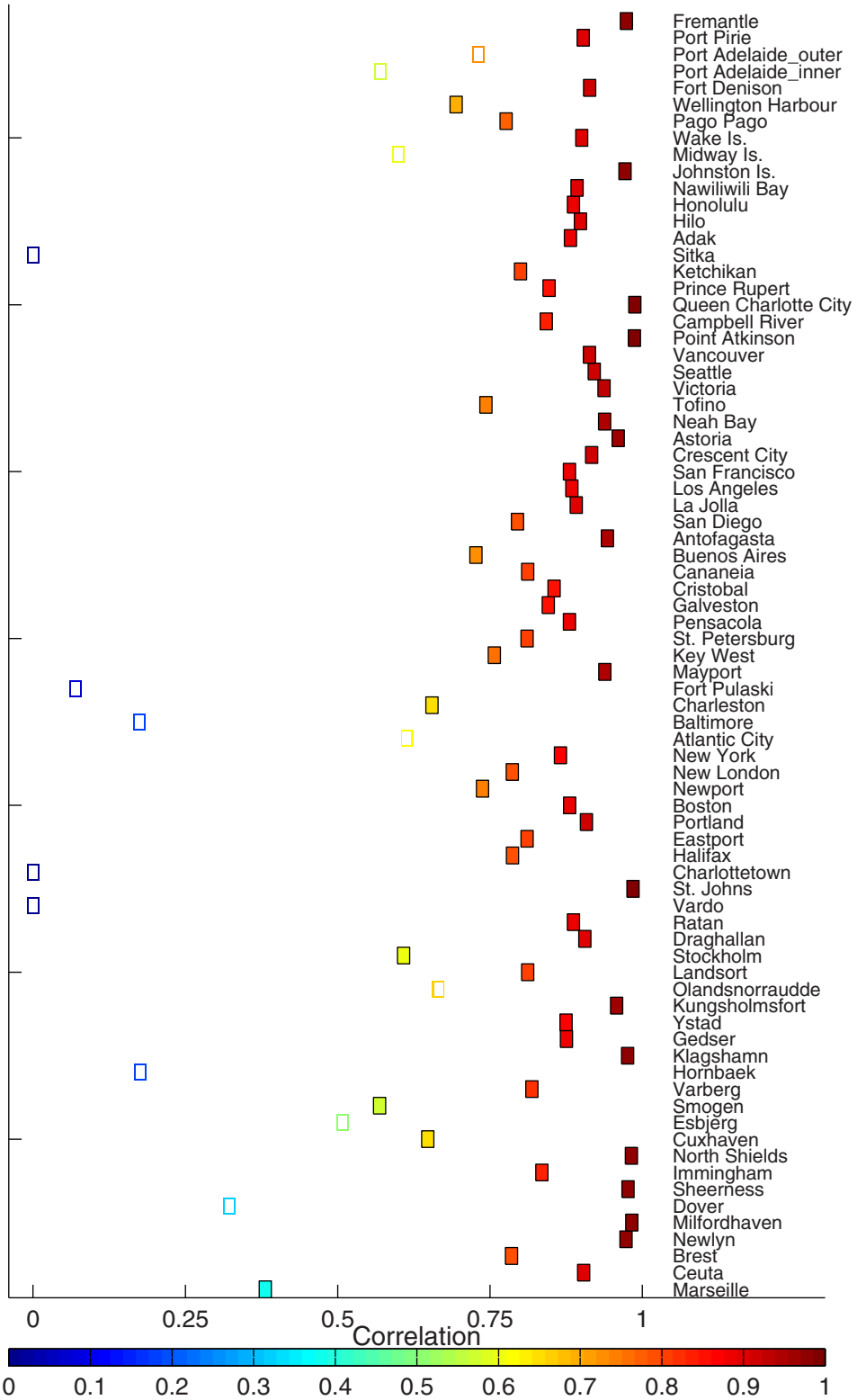

Figure 8. Correlation coefficients between RL50 and smoothed number of events over the 99.5th percentile. Filled squares denote statistically significant correlations at the 90% confidence limits.

Table 2. Selected Climate Indices Used in This Study, Starting Year, and Source

North Atlantic Oscillation (NAO)	1899	https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-pc-based
Atlantic Multidecadal Oscillation (AMO)	1861	http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/AMO/
Arctic Oscillation (AO)	1871	http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/ao20thc.long.data
Southern Oscillation Index (SOI)	1882	http://www.cpc.ncep.noaa.gov/data/indices/
El Niño 12	1870	http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino12/
El Niño 34	1870	http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino34/
Pacific Decadal Oscillation (PDO)	1900	http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/PDO/
Southern Annular Mode (SAM)	1957	http://www.nerc-bas.ac.uk/icd/gjma/sam.html

will focus on the former, that is, on RL50 curves. The climate indices used are listed in Table 2 with indications of their sources. Indices series, commonly provided on a monthly basis, are first yearly averaged and then smoothed using a Butterworth low-pass filter of order 2 with a cutoff frequency of 20 years in order to be consistent with slow variations in RL50. To focus on the more robust results, we considered RL50 series with at least 80 years of data, which resulted in 41 tide gauge stations. The statistical significance of the correlations was estimated once again following *Ebisuzaki* [1997], as described above. The resulting correlations that are significant at the 90% confidence level are plotted in Figure 9. It is noticeable that significant correlations appear between stations and indices that describe remote patterns. For example, extremes along the Pacific coast of North America (from San Francisco to Point Atkinson) seem to be correlated with the NAO; also, most of the stations display significant (although low) correlations with the SAM, even when they are located in high latitudes of the Northern Hemisphere. Although actual teleconnections have been reported between remote climate indices even at decadal time scales [e.g., *Li et al.*, 2015], these features do not necessarily respond to realistic forcing mechanisms, as they can be consequence of using smoothed climate indices that are often significantly correlated to each other. This is also a reason why only a few of the most relevant climate indices have been included in this analysis. Thus, it is more likely that Pacific North American stations respond to changes in the PDO, which for the considered period seems to be in phase with the NAO at multidecadal time scales. In this particular region, it is worth noting that the northernmost stations, from San Francisco to Point Atkinson, are correlated with the major climate modes of the Pacific Ocean, namely the PDO and the indices characterizing El Niño; however, the lower-latitude stations, Los Angeles, La Jolla and San Diego, do not show significant correlations with any of the indices that characterize the climate in the Pacific Ocean. This is in agreement with results shown in Figure 5 that evidence the different evolution of extremes in these southern stations of the U.S. west coast.

The European stations reflect the well-known NAO pattern with its characteristic dipole: negative correlations in Southern Europe (Marseille in this case) and positive in northern stations, following the same spatial pattern as the correlations with MSL [e.g., *Woolf et al.*, 2003; *Tsimplis et al.*, 2005; *Tsimplis and Shaw*, 2008]. The impact of the NAO on sea level extremes in Europe has already been pointed out in earlier works [e.g., *Marcos et al.*, 2009; *Menéndez and Woodworth*, 2010; *Cid et al.*, 2015] and explained by the northward (southward) shift in the storm tracks when the NAO is in its positive (negative) phase. Notably, the stations along the North American Atlantic coast also display significant and high negative correlations with NAO, unlike MSL in this region. This was already pointed out by *Talke et al.* [2014] for the New York tide gauge and by *Ezer and Atkinson* [2014] along the U.S. East coast. The substantial match between both sides of the Atlantic Ocean and the link with the NAO is highlighted in Figure 10, in which RL50 changes with respect to their mean values have been plotted for Marseille and New York tide gauges, together with the low-pass filtered NAO index (scaled for the sake of clarity). At multidecadal time scales, sea level extremes in New York and in Marseille follow the same temporal pattern, controlled by climate variations that in the Atlantic are predominantly described by the NAO. To further explore this pattern, we have computed the correlations between RL50 in these two stations and mean sea level atmospheric pressure (MSLPr) obtained from the Twentieth Century Reanalysis [*Compo et al.*, 2011]. As with the climate indices, the MSLPr at each grid point has been low-pass filtered using a Butterworth filter with a cutoff frequency of 20 years. The resulting maps are included in Figure 10 (right) and show a clear dipole pattern in the North Atlantic with negative correlations at low and midlatitudes and positive at higher latitudes that resembles the spatial fingerprint of the NAO.

The relationships between variations in RL50 and climate indices reveal that the long-term evolution of sea level extremes is indeed controlled by large-scale climate variations. However, these relations do not identify the physical mechanisms responsible for such changes. Explaining in detail the driving mechanisms of

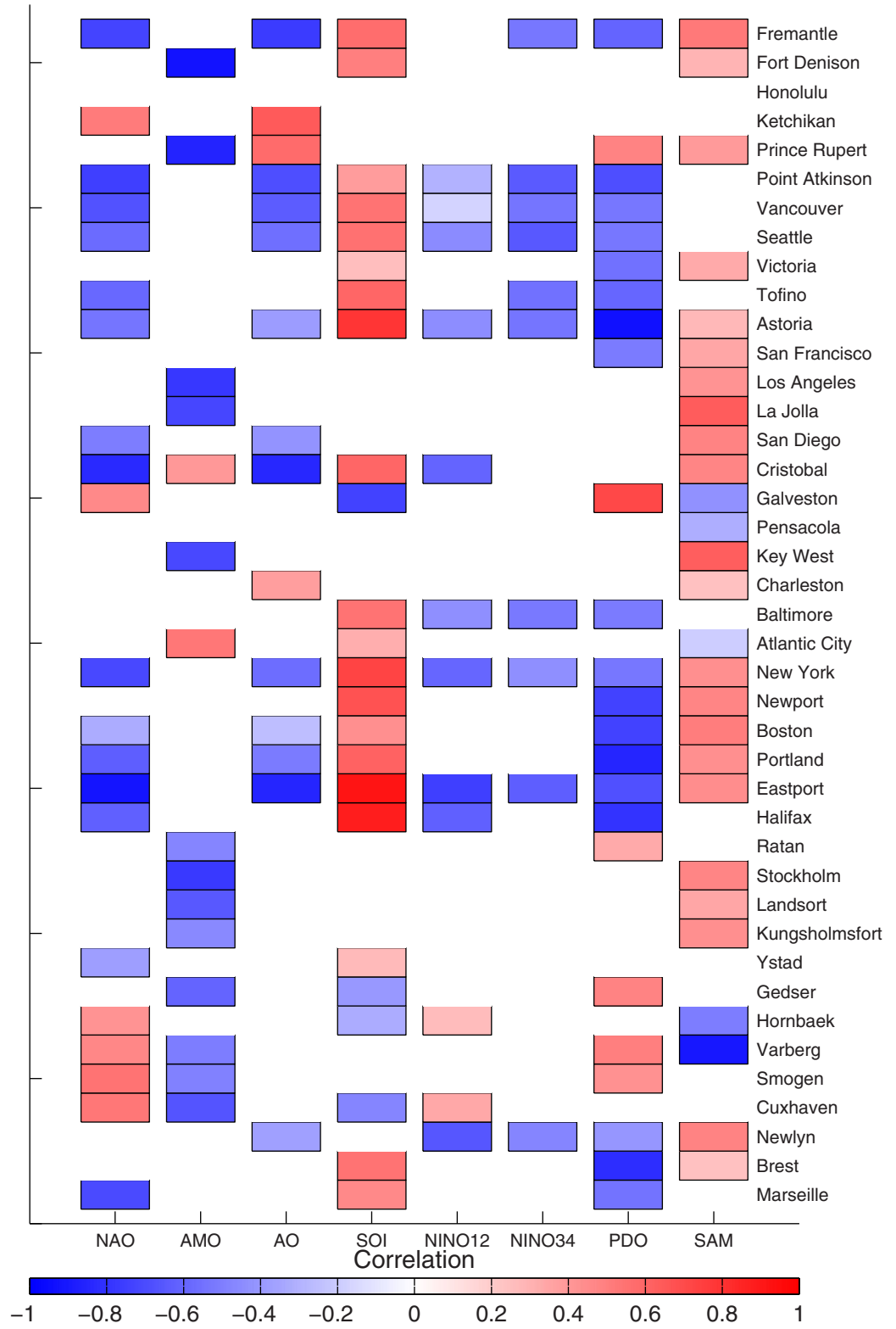

Figure 9. Correlations between climate indices and RL50 at stations longer than 80 years. Only correlations that are significant at the 90% confidence level are plotted.

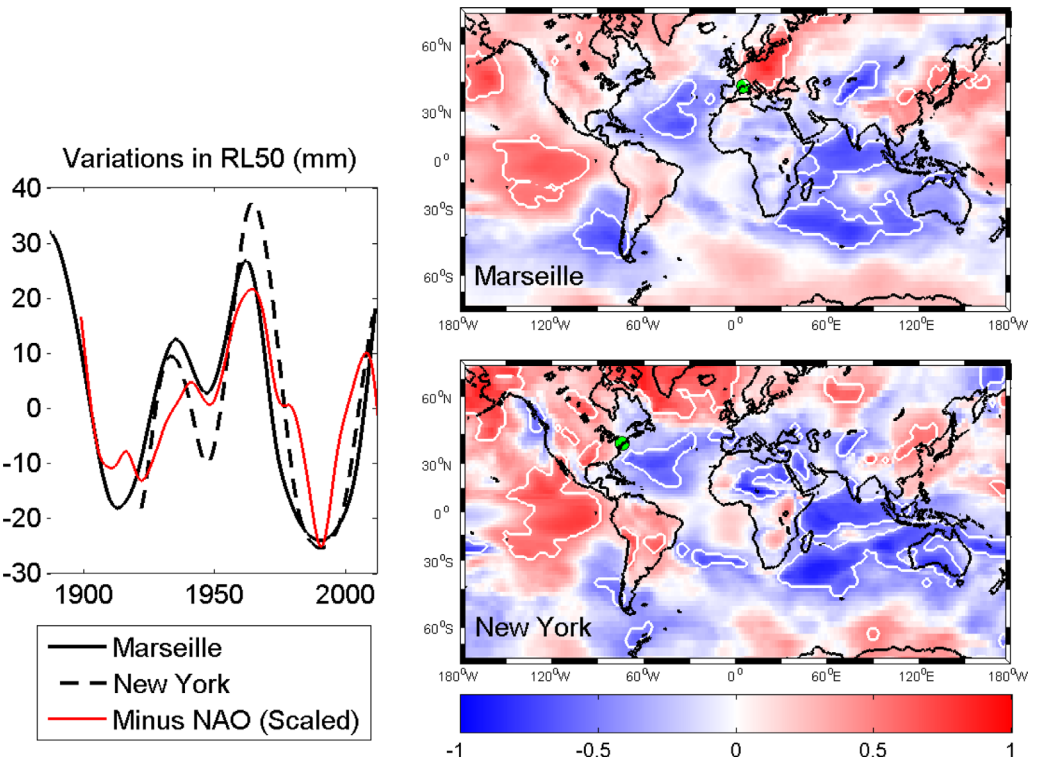

Figure 10. (left) Variations in RL50 (wrt the corresponding mean value) in Marseille and New York tide gauges and the low-pass filtered NAO index. (right) Correlations between RL50 in (top) Marseille and (bottom) New York with low-pass filtered mean sea level pressure. White contours denote correlations of 0.5 and the locations of the tide gauges are represented with a green dot.

the variability in extremes is not within the scope of the present paper, as it would probably require dedicated studies for different regions that are affected by rather distinct atmospheric and oceanographic regimes (for instance with more regional indices such as defined in *Dangendorf et al.* [2014]).

8. Concluding Remarks

In this paper, we have used the largest high-frequency tide gauge sea level record data set available so far to explore long-term (decadal to multidecadal) changes in extreme episodes of the storm surge component. Datum shifts and timing errors are very common in historical sea level records; therefore, a careful and detailed quality control is unavoidable for them to be useful for extreme studies. Otherwise, there is a serious risk of identifying fake extremes from tidal residuals, for example. Long and good quality time series are geographically biased toward the European and North American coasts, reflecting the historical limitations of the worldwide tide gauge data set. Thus, our analyses cover the oceans basin to the extent as possible, but there is a substantial lack of information especially in the Southern Hemisphere and the Indian Ocean. This inherent limitation could be overcome using the output of high-frequency storm surge modeling. In this case, similar analyses could be carried out for the global coastlines.

Our choice was to use state space models to unmask the long-term behavior of time series of sea level extremes. This methodology is more complex than the most extended Maximum Likelihood Estimator used to fit an empirical distribution to a theoretical distribution describing the extremes; however, it is also more robust as it ensures convergence of the solutions and it also allows data gaps in the observations, something usual in historical time series. Besides its robustness, a major motivation of using this approach is that it avoids constraining the temporal variability of extremes (for example, assuming proxies, or linear or parabolic adjustments). Another major advantage over other simpler approaches, such as those that use shifting time windows, is that estimates of the state of the system at each time step are based on all observations available since our method samples from the joint posterior density of the states and the parameters conditioned on all observations. Also, the resulting time series has the same length of the original data set, as no data are lost in presence of gaps or at the beginning/end of the record, although uncertainty increases due to the use of

less information. Based on these models, we have presented evidences of decadal to multidecadal variations in sea level extremes that are unrelated to MSL changes. We note here that, at some stations both extratropical and tropical cyclones affect extreme sea levels, so the selection of yearly extreme events may comprise samples associated with these two different forcing mechanisms and their statistics would be probably better represented by distinct distribution functions. We proved that, in almost all cases studied (82%), the intensity and the frequency of extreme episodes vary accordingly, implying that in epochs of stronger events these are also more likely to occur. This observed variability cannot, in general, be described by a simple linear trend and thus the question of whether the extreme sea level events associated with storminess are increasing or not becomes very difficult to address without knowledge of the causes of this variability.

In any extreme analysis, a decision has to be made regarding the definition of the extreme events, either in the form of the number of extremes (when a GEV distribution is used) or of a threshold value (if a Generalized Pareto Distribution is preferred). In our case, we chose five extremes per year, since our PGAS simulations demonstrated that uncertainties with a smaller number could be, in some cases, too large to drive any conclusion on the temporal variability of the extremes. However, it must be mentioned that most extreme events (e.g., yearly maxima) may be associated to weather patterns that are different from those leading to moderate extremes. This is a point that should be addressed locally and used to define the level of uncertainty allowed in each case.

Based on our approach, we conclude that there exists regional coherence in the variability of sea level extremes, which points toward large-scale climate drivers of the long-term changes in storminess. This coherency is observed in both the intensity and the occurrence of extremes, thus reinforcing the reliability of the state space models in this context, provided that we applied two models to two differentiated sets of data. The physical mechanisms responsible for the observed long-term variations in extremes at regional scales should be further explored on a regional basis, paying attention to the characteristic atmospheric and oceanic regimes in each area. This would allow exploring long-term gradual changes associated with external forcing factors and ultimately would facilitate the detection of external signals in extreme events, something very relevant in a climate change context that may impact the intensity and the frequency of extreme episodes [*Church et al*., 2013].

Acknowledgments

M. Marcos acknowledges a "Ramon y Cajal" contract funded by the Spanish Ministry of Economy. This work was supported by the research project CLIMPACT (CGL2014-54246-C2-1-R) funded by the Spanish Ministry of Economy. Twentieth Century Reanalysis V2 data are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. The authors are grateful to Thomas Wahl and another anonymous reviewer for their helpful comments. The GESLA data set was assembled by Philip Woodworth (National Oceanography Centre, UK), John Hunter (University of Tasmania, Australia), and Melisa Menendez (University of Cantabria, Spain) from national and international sea level databanks for the study of extreme sea level change and variability. An updated data set called GESLA-2 assembled by Woodworth, Hunter, and Menendez together with Marta Marcos (University of the Balearic Islands, Spain) and Ivan Haigh (University of Southampton, UK) is currently in preparation. Depending on the conditions the supplying authorities specified, this new data set will eventually be made available upon request (plw@noc.ac.uk).

References

Andrieu, C., A. Doucet, and R. Holenstein (2010), Particle Markov chain Monte Carlo methods, *J. R. Stat. Soc., Ser. B*, *72*, 269–342.

Arns, A., T. Wahl, S. Dangendorf, and J. Jensen (2015), The impact of sea level rise on storm surge water levels in the northern part of the German Bight, *Coastal Eng.*, *96*, 118–131.

Butler, A., J. E. Hefferman, J. A. Tawn, R. A. Flather, and K. J. Horsburgh (2007), Extreme value analysis of decadal variations in storm surge elevations, *J. Mar. Syst.*, *67*, 189–200, doi:10.1016/j.jmarsys.2006.10.006.

Cid, A., M. Menéndez, S. Castanedo, A. Abascal, F. Méndez, and R. Medina (2015), Long-term changes in the frequency, intensity and duration of extreme storm surge events in southern Europe, *Clim. Dyn.*, 1–14, doi:10.1007/s00382-015-2659-1.

Church, J., et al. (2013), Sea level change, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 1137–1216, Cambridge University Press, Cambridge, U. K.

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

Codiga, D. L. (2011), Unified tidal analysis and prediction using the UTide Matlab functions, *Tech. Rep. 2011-01*, 59 pp., Grad. Sch. of Oceanogr., Univ. of R. I., Narragansett. [Available at ftp://www.po.gso.uri.edu/pub/downloads/codiga/pubs/2011Codiga-UTide-Report.pdf]

Compo, G. P., et al. (2011), The twentieth century reanalysis project, *Q. J. R. Meteorol. Soc.*, *137*, 1–28, doi:10.1002/qj.776.

Dangendorf, S., C. Mudersbach, J. Jensen, A. Ganske, and H. Heinrich (2013), Seasonal to decadal forcing of high water level percentiles in the German Bight throughout the last century, *Ocean Dyn.*, *63*, 533–548.

Dangendorf, S., S. Müller-Navarra, J. Jensen, F. Schenk, T. Wahl, and R. Weisse (2014), North Sea storminess from a novel storm surge record since AD 1843, *J. Clim.*, *27*, 3582–3595.

Durbin, J., and S. J. Koopman (2002), A simple and efficient simulation smoother for state space time series analysis, *Biometrika*, *89*, 603–615.

Ebisuzaki, W. (1997), A method to estimate the statistical significance of a correlation when the data are serially correlated, *J. Clim.*, *10*, 2147–2153, doi:10.1175/1520-0442(1997)010 < 2147:AMTETS>2.0.CO;2.

Ezer, T., and L. P. Atkinson (2014), Accelerated flooding along the U.S. East Coast: On the impact of sea-level rise, tides, storms, the Gulf Stream, and the North Atlantic Oscillations, *Earth's Future*, *2*, 362–382, doi:10.1002/2014EF000252.

Haigh, I. D., E. M. S. Wijeratne, L. R. MacPherson, C. B. Pattiaratchi, M. S. Mason, R. P. Crompton, and S. George (2014), Estimating present day extreme water level exceedance probabilities around the coastline of Australia: Tides, extra-tropical storm surges and mean sea level, *Clim. Dyn.*, *42*, 121–138, doi:10.1007/s00382-012-1652-1.

Hay, C. C., E. Morrow, R. E. Kopp, and J. X. Mitrovica (2015), Probabilistic reanalysis of twentieth-century sea-level rise, *Nature*, *517*, 481–484.

Helske, J. (2015), KFAS: Kalman Filter and Smoother for Exponential Family State Space Models, R package version 1.1.2. [Available at http://cran.r-project.org/package=KFAS.]

Holgate, S. J., A. Matthews, P. L. Woodworth, L. J. Rickards, M. E. Tamisiea, E. Bradshaw, P. R. Foden, K. M. Gordon, S. Jevrejeva, and J. Pugh (2013), New data systems and products at the permanent service for mean sea level, *J. Coastal Res.*, *29*(3), 493–504, doi:10.2112/JCOASTRES-D-12-00175.1.

Jevrejeva, S., J. C. Moore, A. Grinsted, and P. L. Woodworth (2008), Recent global sea level acceleration started over 200 years ago?, *Geophys. Res. Lett.*, *35*, L08715, doi:10.1029/2008GL033611.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, *82*(1), 35–45.

Kurtenbach, E., A. Eicker, T. Mayer-Gürr, M. Holschneider, M. Hayn, M. Fuhrmann, and J. Kusche (2012), Improved daily GRACE gravity field solutions using a Kalman smoother, *J. Geodyn.*, *59–60*, 39–48, doi:10.1016/j.jog.2012.02.006.

Laine, M., N. Latva-Pukkila, and E. Kyrölä (2014), Analysing time-varying trends in stratospheric ozone time series using the state space approach, *Atmos. Chem. Phys.*, *14*, 9707–9725, doi:10.5194/acp-14-9707-2014.

Lee, J., and J. O. Berger (2003), Space-time modeling of vertical ozone profiles, *Environmetrics*, *14*, 617–639, doi:10.1002/env.608.

Letetrel, C., M. Marcos, B. M. Míguez, and G. Woppelmann (2010), Sea level extremes in Marseille (NW Mediterranean) during1885–2008, *Cont. Shelf Res.*, *30*, 267–1274, doi:10.1016/j.csr.2010.04.003.

Li, W., L. Li, and Y. Deng (2015), Impact of the interdecadal pacific oscillation on tropical cyclone activity in the North Atlantic and Eastern North Pacific, *Sci. Rep.*, *5*, 12358, doi:10.1038/srep12358.

Lindsten, F., M. I. Jordan, and T. B. Schön (2014), Particle Gibbs with ancestor sampling, *J. Mach. Learning Res.*, *15*, 2145–2184.

Marcos, M., M. N. Tsimplis, and A. G. P. Shaw (2009), Sea level extremes in southern Europe, *J. Geophys. Res.*, *114*, C01007, doi:10.1029/2008JC004912.

Marcos, M., G. Jordà, D. Gomis, and B. Pérez (2011), Changes in storm surges in Southern Europe from a regional model under climate change scenarios, *Global Planet. Change*, *77*, 116–128, doi:10.1016/j.gloplacha.2011.04.002.

Mawdsley, R. J., I. D. Haigh, and N. C. Wells (2015), Global secular changes in different tidal high water, low water and range levels, *Earth's Future*, *3*, 66–81, doi:10.1002/2014EF000282.

McInnes, K., I. Macadam, G. D. Hubbert, and J. G. O'Grady (2009), A modelling approach for estimating the frequency of sea level extremes and the impact of climate change in southeast Australia. *Nat. Hazards*, *51*(1), 115–137.

Menéndez, M., and P. L. Woodworth (2010), Changes in extreme high water levels based on a quasi-global tide-gauge data set, *J. Geophys. Res.*, *115*, C10011, doi:10.1029/2009JC005997.

Merrifield, M. A., A. S. Genz, C. P. Kontoes, and J. J. Marra (2013), Annual maximum water levels from tide gauges: Contributing factors and geographic patterns, *J. Geophys. Res. Oceans*, *118*, 2535–2546, doi:10.1002/jgrc.20173.

Mudersbach, C., and J. Jensen (2010), Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline, *J. Flood Risk Manage.*, *3*, 52–62, doi:10.1111/j.1753-318X.2009.01054.x.

Sweet, W. V., and C. Zervas (2011), Cool-season sea level anomalies and storm surges along the U.S. East Coast: Climatology and comparison with the 2009/10 El Niño, *Mon. Weather Rev.*, *139*, 2290–2299, doi:10.1175/MWR-D-10-05043.1.

Talke, S. A., P. Orton, and D. A. Jay (2014), Increasing storm tides in New York Harbor, 1844–2013, *Geophys. Res. Lett.*, *41*, 3149–3155, doi:10.1002/2014GL059574.

Thompson, P. R., G. T. Mitchum, C. Vonesch, and J. Li (2013), Variability of winter storminess in the eastern United States during the twentieth century from tide gauges, *J. Clim.*, *26*, 9713–9726, doi:10.1175/JCLI-D-12-00561.1.

Tsimplis, M. N., and A. G. P. Shaw (2008), The forcing of mean sea level variability around Europe, *Global Planet. Change*, *63*, 196–202.

Tsimplis, M. N., et al. (2005), Towards a vulnerability assessment of the UK and northern European coasts: The role of regional climate variability, *Philos. Trans. R. Soc. A*, *363*(1831), 1329–1358, doi:10.1098/rsta.2005.1571.

Von Storch, H. (2014), Storm surges: Phenomena, forecasting and scenarios of change, *Proc. IUTAM*, *10*, 356–362, doi:10.1016/j.piutam.2014.01.030.

Wahl, T., and D. P. Chambers (2015), Evidence for multidecadal variability in US extreme sea level records, *J. Geophys. Res. Oceans*, *120*, 1527–1544, doi:10.1002/2014JC010443.

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

Woodworth, P. L., and D. L. Blackman (2002), Changes in high waters at Liverpool since 1768, *Int. J. Climatol.*, *22*, 697–714, doi:10.1002/joc.761.

Woodworth, P. L., and D. L. Blackman (2004), Evidence for systematic changes in extreme high waters since the mid-1970s, *J. Clim.*, *17*, 1190–1197.

Woolf, D. K., A. G. P. Shaw, and M. N. Tsimplis (2003), The influence of the North Atlantic Oscillation on sea-level variability in the North Atlantic region, *J. Atmos. Ocean Sci.*, *9*(4), 145–167.

Wöppelmann, G., M. Marcos, A. Coulomb, B. Martín Míguez, P. Bonnetain, C. Boucher, M. Gravelle, B. Simon, and P. Tiphaneau (2014), Rescue of the historical sea level record of Marseille (France) from 1885 to 1988 and its extension back to 1849–1851, *J. Geod.*, *88*, 869–885, doi:10.1007/s00190-014-0728-6.

Zhang, K., B. C. Douglas, and S. P. Leatherman (2000), Twentieth-Century Storm Activity along the U.S. East Coast, *J. Climate*, *13*, 1748–1761, doi:10.1175/1520-0442(2000)013<1748:TCSAAT>2.0.CO;2.