



**HAL**  
open science

# Dominance of the mean sea level in the high-percentile sea levels time evolution with respect to large-scale climate variability: a Bayesian statistical approach

Jeremy Rohmer, Gonéri Le Cozannet

## ► To cite this version:

Jeremy Rohmer, Gonéri Le Cozannet. Dominance of the mean sea level in the high-percentile sea levels time evolution with respect to large-scale climate variability: a Bayesian statistical approach. *Environmental Research Letters*, 2019, 14, 10.1088/1748-9326/aaf0cd . hal-02380692

**HAL Id: hal-02380692**

**<https://hal.science/hal-02380692>**

Submitted on 28 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LETTER • OPEN ACCESS

# Dominance of the mean sea level in the high-percentile sea levels time evolution with respect to large-scale climate variability: a Bayesian statistical approach

Recent citations

- [Quantifying uncertainties of sandy shoreline change projections as sea level rises](#)  
Gonéri Le Cozannet *et al*

To cite this article: Jeremy Rohmer and Gonéri Le Cozannet 2019 *Environ. Res. Lett.* **14** 014008

View the [article online](#) for updates and enhancements.

## Environmental Research Letters



## LETTER

# Dominance of the mean sea level in the high-percentile sea levels time evolution with respect to large-scale climate variability: a Bayesian statistical approach

## OPEN ACCESS

RECEIVED  
11 June 2018REVISED  
9 November 2018ACCEPTED FOR PUBLICATION  
14 November 2018PUBLISHED  
9 January 2019

Jeremy Rohmer and Gonéri Le Cozannet

BRGM, 3 av. C. Guillemin-45060 Orléans Cedex 2 - France

E-mail: [j.rohmer@brgm.fr](mailto:j.rohmer@brgm.fr)**Keywords:** mean sea level, extremes, Bayesian structure time series model, Kalman-filter, climate indicesSupplementary material for this article is available [online](#)

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



## Abstract

Changes in mean sea level (MSL) are a major, but not the unique, cause of changes in high-percentile sea levels (HSL), e.g. the annual 99.9th quantile of sea level (among other factors, climate variability may also have huge influence). To unravel the respective influence of each contributor, we propose to use structural time series models considering six major climate indices (CI) (Arctic Oscillation, North Atlantic Oscillation, Atlantic Multidecadal Oscillation, Southern Oscillation Index, Niño 1 + 2 and Niño 3.4) as well as a reconstruction of MSL. The method is applied to eight century-long tide gauges across the world (Brest (France), Newlyn (UK), Cuxhaven (Germany), Stockholm (Sweden), Gedser (Denmark), Halifax (Canada), San Francisco (US), and Honolulu (US)). The treatment within a Bayesian setting enables to derive an importance indicator, which measures how often the considered driver is included in the model. The application to the eight tide gauges outlines that MSL signal is a strong driver (except for Gedser), but is not unique. In particular, the influence of Arctic Oscillation index at Cuxhaven, Stockholm and Halifax, and of Niño Sea Surface Temperature index 1 + 2 at San Francisco appear to be very strong as well. A similar analysis was conducted by restricting the time period of interest to the 1st part of the 20th century. Over this period, we show that the MSL dominance is lower, whereas an ensemble of CI contribute to a large part to HSL time evolution as well. The proposed setting is flexible and could be applied to incorporate any alternative predictive time series such as river discharge, tidal constituents or vertical ground motions where relevant.

## 1. Introduction

Changes in extreme sea levels are recognized mainly driven by variations in mean sea level (MSL), as pointed out in the 2013 IPCC report (IPCC 2013). Previous studies have shown and extensively documented the relation between variability and trends in extremes and MSL; see for instance, Menéndez and Woodworth (2010), Woodworth *et al* (2011), Wahl and Chambers (2015). However, MSL being the major contributor does not necessarily mean that it is the only one. Today, extreme sea level events associated with storm surges are already prominent threats for populations and ecosystems on the coasts. The combination of such extreme events with MSL rise and the

related damages are a significant concern (Wahl 2017) for present day and for the future. Therefore, planning adaptation strategies (like hard protection, soft protection, accommodation, retreat) and designing protective infrastructures (like dikes, and seawalls) for the future both require improving the understanding and the predictability of such extreme events via a deeper knowledge of their drivers.

On top of the MSL, forcing related to ocean and atmosphere variations influence extreme sea levels at inter-annual and decadal time scales. To investigate such large-scale climate variability, climate indices (denoted CI) like North Atlantic Oscillation (NAO) or Atlantic Multidecadal Oscillation (AMO) have been used; see e.g. Menéndez and Woodworth (2010),

Woodworth *et al* (2011), Talke *et al* (2014), Marcos *et al* (2015), Wahl and Chambers (2015, 2016), Mawdsley and Haigh (2016), Marcos and Woodworth (2017), Wong *et al* (2018). Far from being theoretical, these research studies bring societal benefits. First, understanding and quantifying the role of each influencing factor is a prerequisite to a formal detection and attribution of coastal extreme water levels, which is a perceived important to stimulate mitigation of climate change (Schwab *et al* 2017). Second, assuming that MSL is the only driver of extreme sea levels, and neglecting the influence of other evolving factors such as CIs, might reduce confidence in current estimates and future projections of extreme coastal water levels, and might even mislead coastal adaptation practitioners in the worst case.

The most widespread method to address the problem of controls on extreme sea levels is the use of linear (Pearson) correlation coefficients (e.g. Marcos and Woodworth (2017): section 4) with possible combination with significance testing techniques. Though efficient and easy-to-implement, this approach remains global and does not account for: (1) the nature of the data, which are time series i.e. the considered time series may be correlated at different time instants; (2) multiple possible drivers (i.e. predictors), which may act in interaction. To overcome both limitations, Wahl and Chambers (2016) proposed the use of multiple regression models. Yet, this approach ignores structural uncertainty, which arises when different alternative models can be proposed to explain the data and appear ‘equally appropriate’ from a statistical perspective.

A possible option for handling this type of uncertainty is to fit a multiple regression model within a Bayesian setting, which enables to examine a large number of possible variable configurations with respect to the observations (Hoeting *et al* 1999). The Bayesian treatment of the fitting process then provides an importance measure (here the posterior inclusion probability as defined for instance by Scott and Varian 2014), which measures how often each driver is included in the regression model. This is can be used to unravel the respective influence of each driver of extreme sea levels.

In the present study, we propose to rely on this Bayesian approach to quantify the respective contribution of the different drivers to the time evolution of high-percentile sea levels (denoted HSL), such as 99 or 99.9th quantile, over inter-annual to centennial time-scales. Multiple possible predictors are accounted for, namely, the annual MSL and six large-scale CI (see further details in section 2). A more advanced method for time series modelling than multiple regression modelling is chosen, namely the technique of Bayesian structural time series model (denoted *bsts* in the following), as developed by Scott and Varian (2014, 2015), which is particularly flexible to model the complex multi-variable, correlational structure of any temporal data.

The paper is organised as follows. In a first section, we describe the century-long sea level time series used for the study, the pre-processing that we conducted and the CI. In a second section, we provide full details of the proposed procedure based on *bsts* method. In section 4, we analyse and discuss the results of the procedure applied to eight tide gauges across the world.

## 2. Data

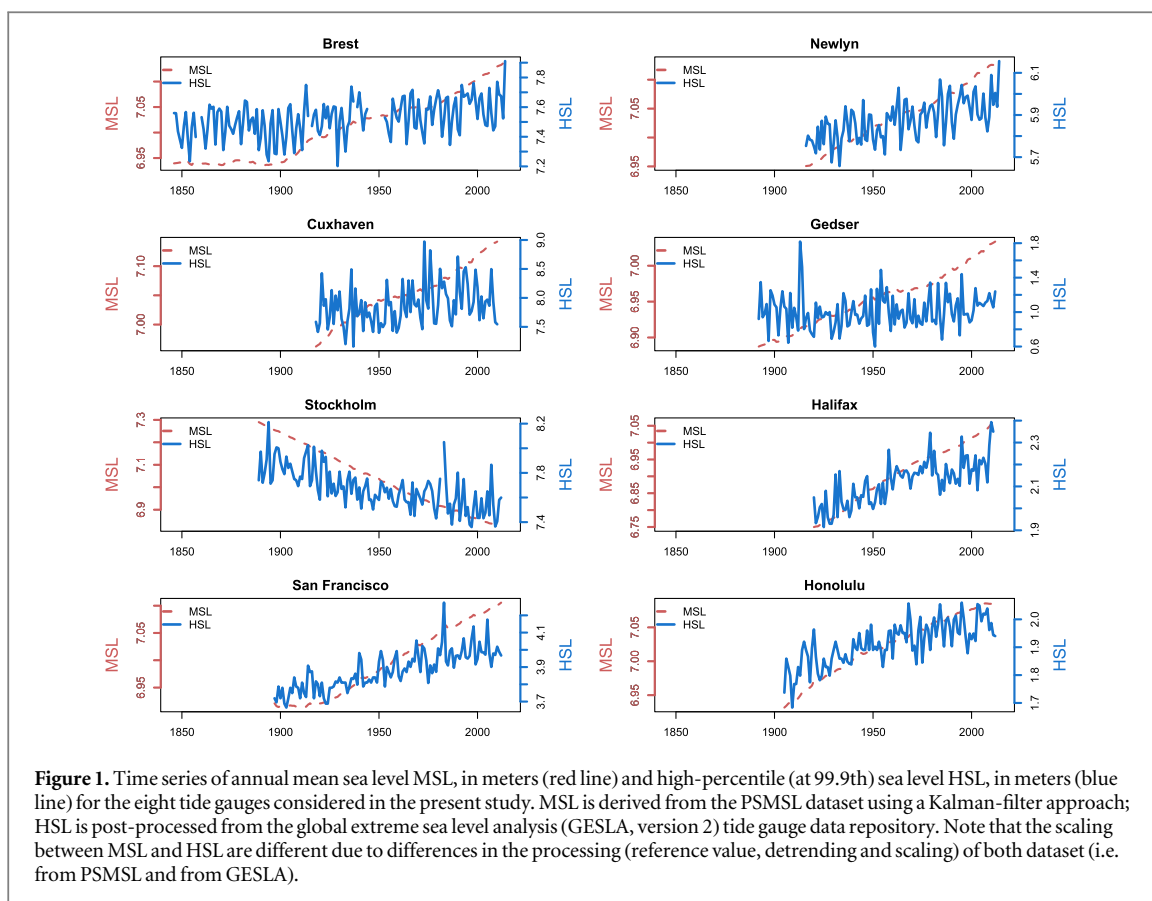
We have extracted from the Global Extreme Sea Level Analysis (GESLA, version 2<sup>1</sup>, Woodworth *et al* 2017) tide gauge data repository, eight of the longest (quasi century-long) time series of sea levels sampled at a hourly frequency, namely: Brest, France (1846–2014); Newlyn, UK (1915–2014); Cuxhaven, Germany (1918–2015); Gedser, Denmark (1891–2012); Stockholm, Sweden (1889–2012); Halifax, Canada (1920–2011); San Francisco, US (1897–2012); Honolulu US (1905–2012). See locations in supplementary materials, available online at: [stacks.iop.org/ERL/14/014008/mmedia](http://stacks.iop.org/ERL/14/014008/mmedia).

These data present very limited missing values except for Brest for the time interval between 1943–1953 and for Stockholm for the year 1982. These missing data were replaced by means of an imputation approach based on a Kalman smoothing approach (see e.g. Moritz and Bartz-Beielstein 2017). The impact of this choice on the results presented in section 4 was tested by using other imputation approaches (mean value, linear interpolation, etc), but appears to be very limited (see supplementary materials). We adopted the percentile time series analysis described by Woodworth and Blackman (2004), and we extracted the annual 99.9th percentile from the hourly SL data. Note that the annual 99.9th percentile approximately corresponds to the level of the eight highest hourly sea level values over the considered year. This constitutes the annual HSL time series used in the following. The influence of the quantile level is investigated in section 4.3.

The annual MSL was estimated at each time step by post-processing the annual data provided by PSMSL<sup>2</sup> using a forward–backward Kalman-filter (e.g. Visser *et al* 2015); see further details in supplementary materials. This procedure allows separating nonlinear changes in MSL from inter-annual modes of variability, which are part of the residual signal to be detected in HSL time series. Figure 1 shows the different time series for the eight locations. Note that the scaling between MSL and HSL are different due to differences in the processing (reference value, detrending and scaling) of both dataset (i.e. from PSMSL and from GESLA).

<sup>1</sup> <http://gesla.org/>.

<sup>2</sup> <http://psmsl.org/>.



**Table 1.** Climate indices selected in the present study.

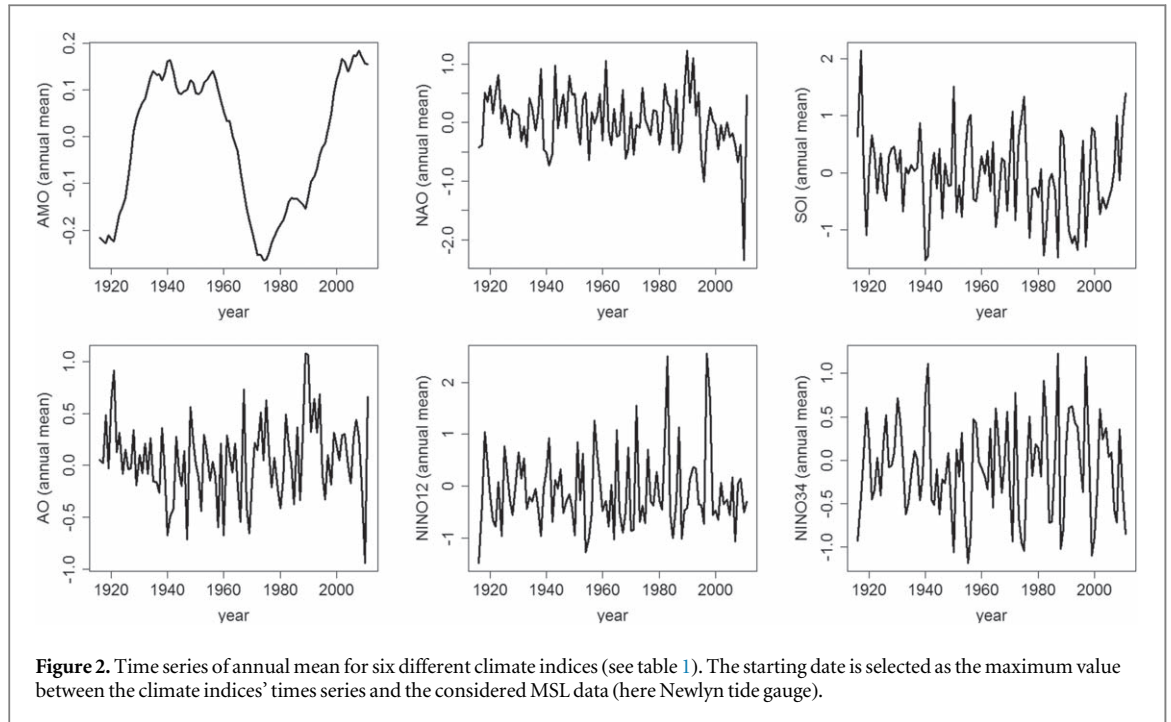
| Climate index  | Starting date | Source (last access on 19/09/2018)   |
|--|---------------|--|
| North Atlantic Oscillation (NAO)                         | 1821          | <a href="https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/NAO/">https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/NAO/</a>  |
| Atlantic Multidecadal Oscillation (AMO—smoothed version) | 1861          | Enfield <i>et al</i> (2001), Rayner <i>et al</i> (2003)  |
| Arctic Oscillation (AO)                                  | 1871          | <a href="https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/AMO/">https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/AMO/</a><br><a href="https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/ao20thc.long.data">https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/ao20thc.long.data</a>     |
| Southern Oscillation Index (SOI)                         | 1866          | Ropelewski and Jones (1987), Allan <i>et al</i> (1991), Können <i>et al</i> (1998)   |
| El Nino 12   | 1870          | <a href="https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/SOI/">https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/SOI/</a><br>Rayner <i>et al</i> 2003  |
| El Nino 34   | 1870          | <a href="https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino12/">https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino12/</a><br>Rayner <i>et al</i> 2003<br><a href="https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino34/">https://esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino34/</a> |

Table 1 provides the source of the CI selected in the present study. Six major CI were selected, namely NAO, AMO, Arctic Oscillation (AO), Southern Oscillation Index (SOI), Niño 1 + 2 SST Index (NINO12), and Niño 3.4 SST Index (NINO34). These indices were selected because previous studies have shown their role in controlling HSL variability (e.g. Marcos *et al* 2015), and because the time series currently available are in the order of 150 years (with a starting date before 1900) and therefore allow to detect the influence of predictors at low frequency (table 1). The monthly time series are converted to annual data by taking the mean value for each year. The starting date is selected as the maximum value between the CI

times series and the considered MSL data (in figure 2, it corresponds to 1915 in the Newlyn case).

### 3. Method

We aim at modelling the observed HSL series by accounting for two sources of information, namely: (1) the time series behaviour of HSL prior to the considered time instant (i.e. the trend); (2) the behaviour of other (control) time series that is known to be predictive of HSL, namely MSL and  $d$  different climate indices  $CI_{1,\dots,d}$  (here termed as the predictors). Both sources of information are combined using a



structural time series model trained within a Bayesian setting denoted *bsts* (full details are provided by Scott and Varian 2014) as follows:

$$\begin{aligned} \text{HSL}_t &= \text{trend} + \text{regressors} + \text{noise} \\ \text{HSL}_t &= \mu_t + \beta_0 + \beta_{\text{MSL}} \times \text{MSL}_t + \beta_{\text{CI}_1} \times \text{CI}_{1,t} \\ &+ \beta_{\text{CI}_2} \times \text{CI}_{2,t} + \dots + \beta_{\text{CI}_d} \times \text{CI}_{d,t} + \varepsilon_t, \end{aligned} \quad (1)$$

where  $\beta$  are the regression coefficients (with  $\beta_0$  a constant value termed as *intercept*), which are related to MSL and the  $d$  CI. The noise term  $\varepsilon_t$  follows a zero-mean Gaussian distribution with fixed standard deviation  $\sigma_\varepsilon$  and is assumed to be independent from the other parameters (i.e. the evolution is driven by an independent Gaussian random walk).

The term  $\mu_t$  at time instant  $t$  is modelled via a 1-lag autoregressive model:

$$\mu_t = \phi_1 \mu_{t-1} + \varepsilon_{\mu,t}, \quad (2)$$

where  $\varepsilon_{\mu,t}$  follows a zero-mean Gaussian distribution (with fixed standard deviation  $\sigma_\mu$ ), and is independent from the other parameters;  $\phi$  are the regression coefficients.

The model fitting is conducted within a Bayesian setting, which allows accounting for empirical priors on the afore-described parameters (i.e. the regression coefficient  $\phi_1$ , the regression coefficients  $\beta$  and the different variance parameters of the error terms  $\varepsilon$ ) and the initial states (i.e. the term  $\mu_t$ , and the predictors).

A spike-and-slab prior is assumed to express a prior belief that a sparse set of variables can explain the response, i.e. that most of the regression coefficients  $\beta$  are exactly zero (Scott and Varian 2015). Full details on the regressor priors' definition and parametrisation are provided in supplementary materials. The full model is estimated via a Bayesian model averaging

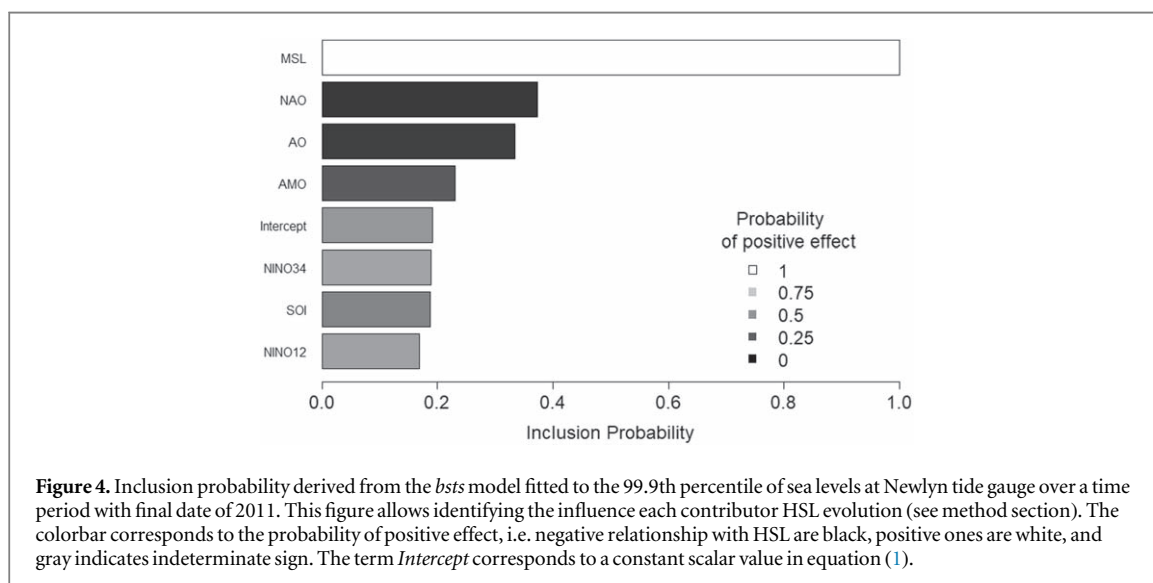
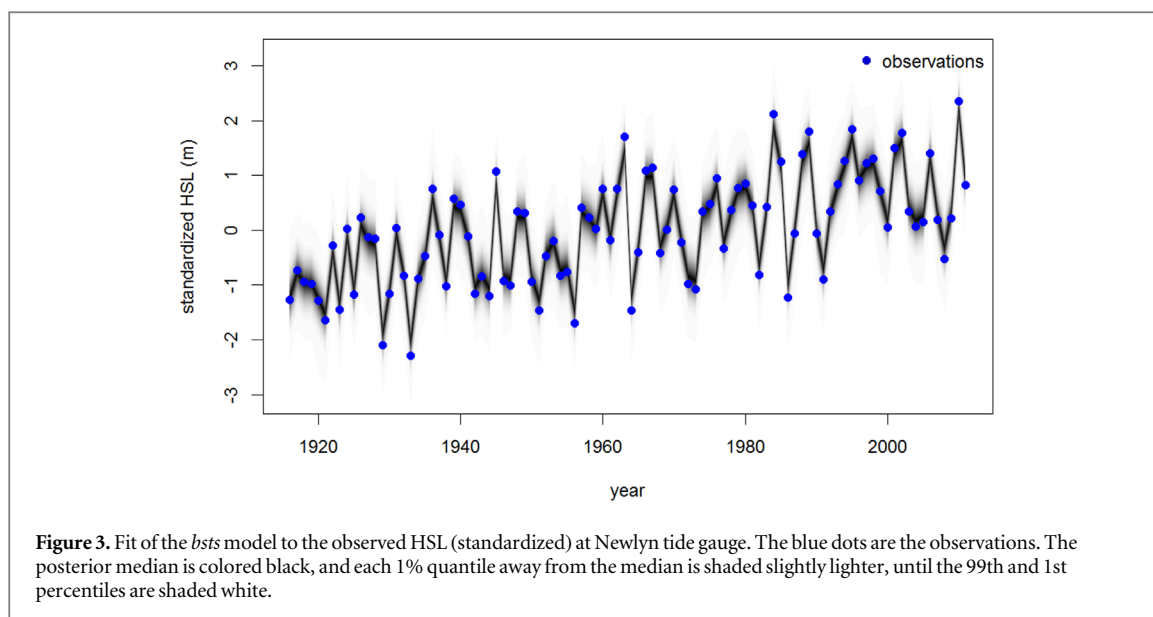
procedure (Hoeting *et al* 1999), which combines information from the priors and uses a Markov chain Monte Carlo procedure (see technical details in Scott and Varian 2014; section 4). A space of large number (typically 1000–10 000) of likely variable configurations that explain the response are computed using Gibbs sampling techniques and treated within a stochastic search variable selection framework (George and McCulloch 1993).

From the random sampling procedure, the marginal posterior inclusion probability for each predictor can be evaluated, i.e. the proportion of Monte Carlo draws with a regression coefficient  $\beta \neq 0$ . This means that the inclusion probability measures how often each predictor is selected during the Monte Carlo procedure. Thus, the larger the inclusion probability, the larger the influence of the considered predictor. By using the *bsts*-derived inclusion probability, it is then possible to measure the influence of each driver (MSL or CI) while accounting for the complex multivariable (i.e. presence of multiple predictors), correlational structure (i.e. we are dealing with time series) of the data.

The capability of the *bsts* model to reproduce the observations over the considered time period is measured using the coefficient of determination (denoted  $R^2$ ) considering the error between the posterior predictive mean for the modelled (denoted  $\widehat{\text{HSL}}^\mu$ ) and the observed HSL as follows:

$$R^2 = 1 - \frac{\sum_{t=1}^{t=T} (\text{HSL}(t) - \widehat{\text{HSL}}^\mu(t))^2}{\sum_{t=1}^{t=T} (\text{HSL}(t) - \text{HSL}^\mu)^2}, \quad (3)$$

where  $\text{HSL}^\mu$  is the average value of HSL over the considered time interval. If  $R^2$  is close to 1, this means



that the posterior predictive mean of the *bsts* model explains almost the whole HSL variability.

In practice, the *bsts* model is fitted using the R package called ‘*bsts*’ available at: <https://cran.r-project.org/web/packages/bsts/index.html>

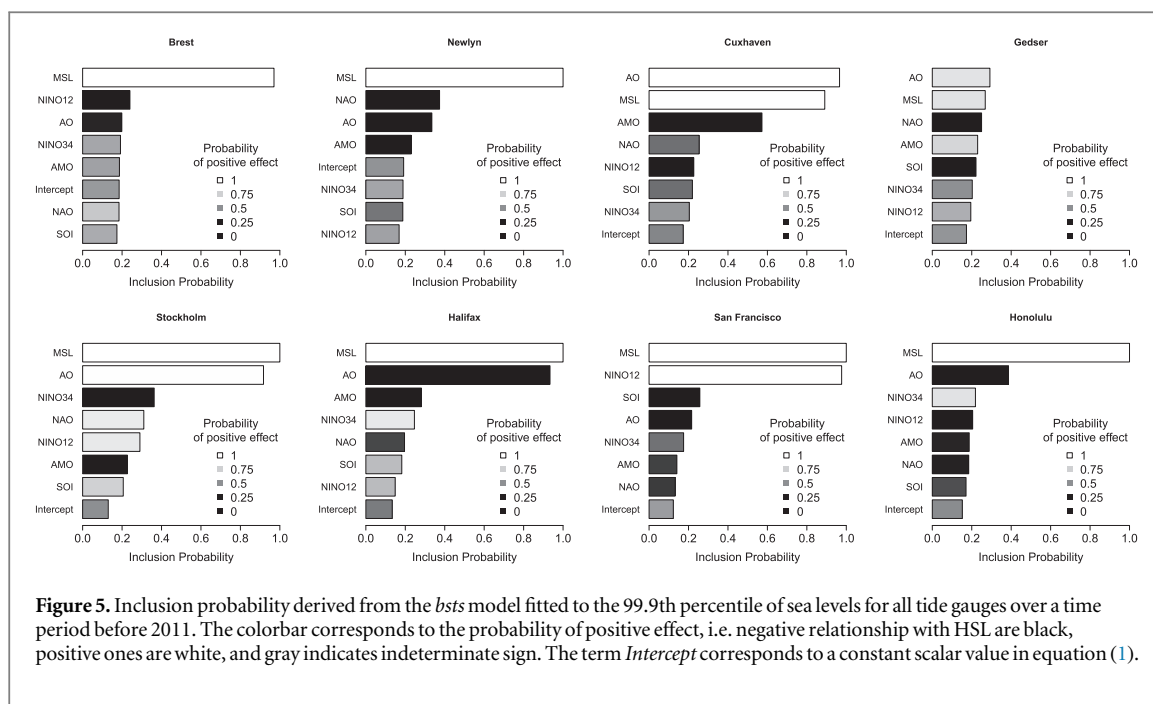
## 4. Application

### 4.1. Application to Newlyn tide gauge

To illustrate the different steps described in section 3, the procedure is first applied to the Newlyn tide gauge. All time series (HSL, MSL, and CI) are first standardized, i.e. the mean was subtracted and the time series was divided by their respective standard deviation. The evolution of the observed (standardized) HSL is modelled using the afore-described *bsts* approach by using 10 000 random draws (figure 3). The influence of the number of random samples is investigated in supplementary materials.

The coefficient of determination  $R^2$  here reaches values larger than 90% with a full coverage of the 99% posterior predictive intervals, hence yielding very satisfactory goodness of fit.

Figure 4 displays the posterior inclusion probability and shows that for all of 10 000 random draws, MSL is almost systematically selected (inclusion probability of nearly 100%), hence confirming the dominant influence of MSL to explain changes in HSL. The two other important drivers of HSL time evolution appears to be the NAO and AO with inclusion probability values of respectively 37% and 33%. The bars are shaded on a continuous [0, 1] scale in proportion to the probability of a positive coefficient, so that negative coefficients are black, positive coefficients are white, and gray indicates indeterminate sign. This shows that MSL has a positive relationship with HSL (i.e. positive effect), whereas NAO and AO have a negative one.



**Figure 5.** Inclusion probability derived from the *bsts* model fitted to the 99.9th percentile of sea levels for all tide gauges over a time period before 2011. The colorbar corresponds to the probability of positive effect, i.e. negative relationship with HSL are black, positive ones are white, and gray indicates indeterminate sign. The term *Intercept* corresponds to a constant scalar value in equation (1).

#### 4.2. Application to the set of tide gauges

The procedure described for Newlyn is performed for the seven other tide gauges using 10 000 random samples. The influence of the number of samples is tested in supplementary materials. The starting date is selected as the maximum value between the CI times series and the considered MSL data and the final date is 2011. The goodness of fit of the *bsts* model is checked for each tide gauge, which shows that the  $R^2$  coefficient reaches for all cases values of around 90% (see the fitted models in supplementary materials).

Figure 5 presents the inclusion probability for each tide gauge. Several observations can be made.

- Figure 5 clearly shows that MSL dominates the HSL time evolution of most tide gauges (except for Gedser) with an inclusion probability superior to 90% and a positive effect (indicated by the white colored bar in figure 5).
- For Brest, Newlyn and Honolulu, the CI contribution remains low-to-moderate with an inclusion probability not larger than  $\approx 40\%$ , hence indicating that MSL is the main driver.
- No particular tendency can be identified for Gedser, where all predictors (MSL and CI) appear to have a low-to-moderate inclusion probability of the order of 20%.
- For Cuxhaven, Stockholm, Halifax, and San Francisco, a combination of two drivers exist, namely MSL + AO for the three first tide gauges, and MSL + NINO12 for the latter. The inclusion probability of the CI is larger than 90%, i.e. as large as the one of MSL or even larger at Cuxhaven. The sign of the AO effect is positive for

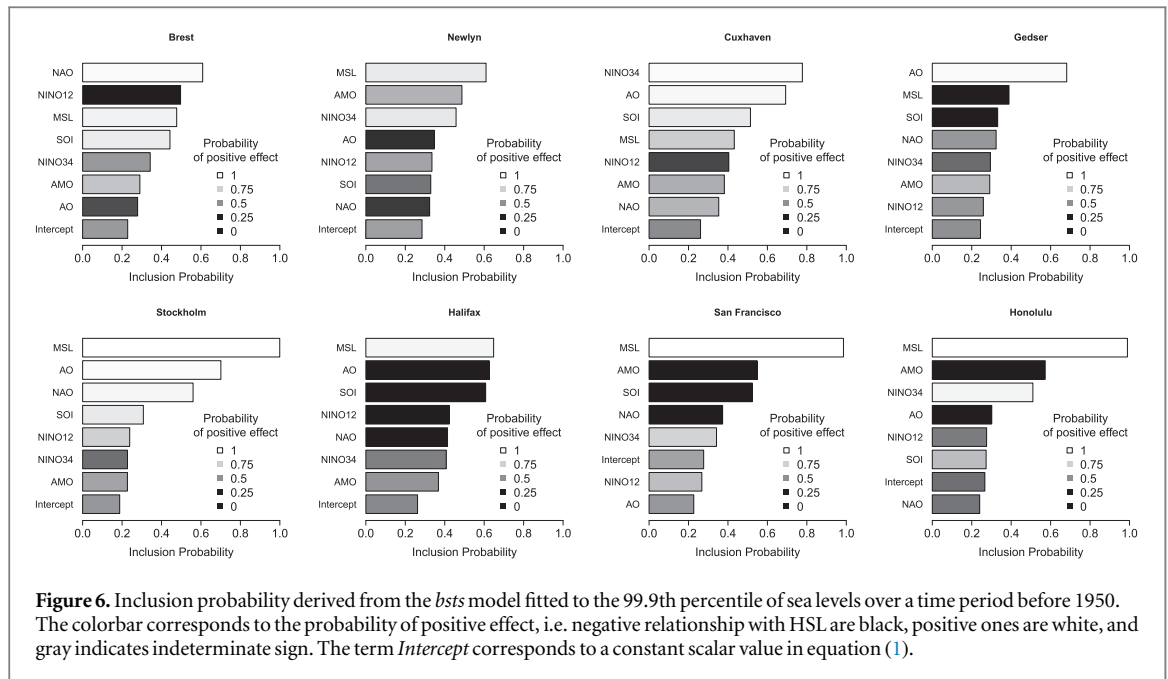
Cuxhaven and Stockholm, but negative for Halifax. The difference in the sign effect may be related to the location of each tide gauges with respect to the sea level anomalies characterizing AO in North Atlantic, which corresponds to a positive pattern at latitudes  $37^\circ\text{--}45^\circ\text{ N}$  (Halifax latitude is  $44.6^\circ\text{ N}$ ) and a negative above  $45^\circ\text{ N}$  (Cuxhaven and Stockholm latitudes are  $53.8^\circ\text{ N}$  and  $59.3^\circ\text{ N}$ ).

- Surprisingly, our study little highlights NAO as a strong driver at the exception for Newlyn contrary to previous studies, (e.g. Marcos and Woodworth 2017). A possible explanation is the strong relationship between AO and NAO (e.g. Ambaum *et al* 2001) so that the signal held by NAO may be ‘viewed’ as redundant with the one of AO by the *bsts* model.
- A clear evidence of strong Niño 1 + 2 SST influence is evidenced at San Francisco. This appears consistent with conclusions of different studies on the link between extreme coastal response and El Niño events (see e.g. Barnard *et al* 2017 and references therein).

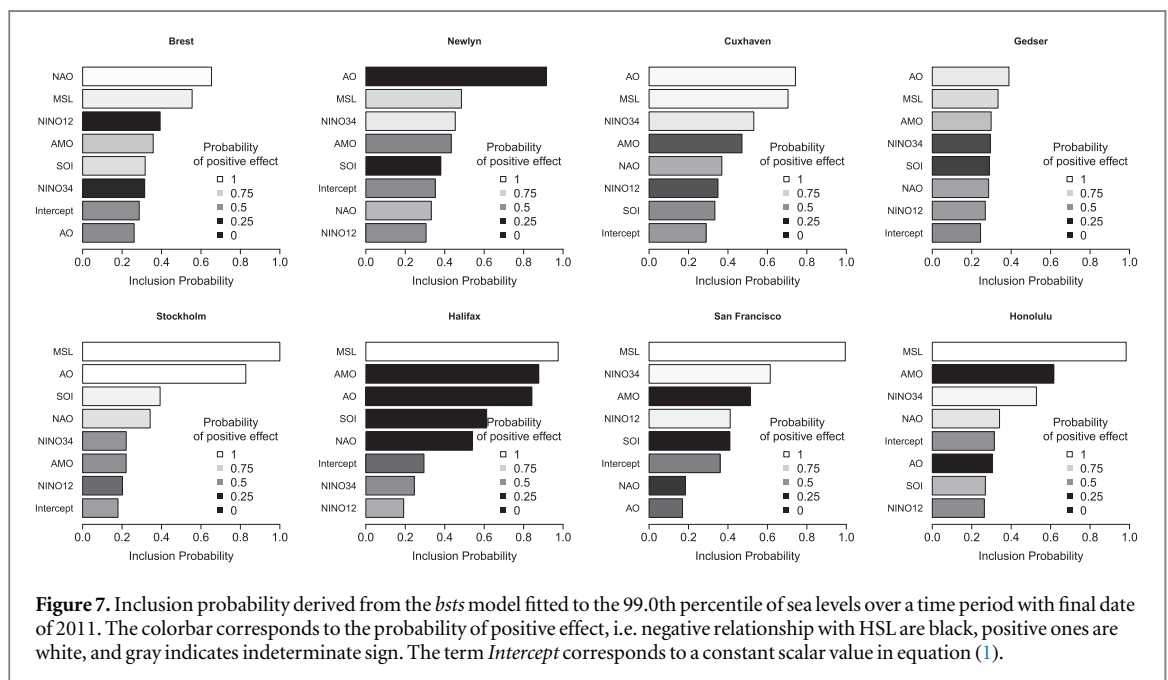
#### 4.3. Influence of the assumptions

In this section, we investigate the implications of two assumptions. First, the length of the training time period is investigated. In section 4.2, the dominance of MSL and of AO and NINO12 was identified over a time-period with a final date of 2011; we now investigate the prevalence of this influence in the first half of the 20th century, i.e. when the final date is set at 1950. Figure 6 shows that the corresponding inclusion probability values have largely changed from the ones of figure 5. Clearly, the large dominance of MSL has





**Figure 6.** Inclusion probability derived from the *bsts* model fitted to the 99.9th percentile of sea levels over a time period before 1950. The colorbar corresponds to the probability of positive effect, i.e. negative relationship with HSL are black, positive ones are white, and gray indicates indeterminate sign. The term *Intercept* corresponds to a constant scalar value in equation (1).



**Figure 7.** Inclusion probability derived from the *bsts* model fitted to the 99.0th percentile of sea levels over a time period with final date of 2011. The colorbar corresponds to the probability of positive effect, i.e. negative relationship with HSL are black, positive ones are white, and gray indicates indeterminate sign. The term *Intercept* corresponds to a constant scalar value in equation (1).

decreased for most tide gauges; in particular, the inclusion probability of MSL has lowered to 40%–60% at Brest, Newlyn, Cuxhaven, and Halifax. Contrary to the analysis on the whole time period before 2011, the HSL evolution before 1950 appears to be driven not only by two predictors, but by an ensemble of them as exemplified at Halifax, where both AO and SOI have a large influence with inclusion probability larger than 60%. For Stockholm, San Francisco, and Honolulu, the large dominance of MSL persists before 1950, but with larger influence of other CI, whose inclusion probability values are of the order of 60%. For Gedser, a larger dominance of AO is identified before 1950. A possible interpretation holds as follows. During the 20th century, MSL is rising at all stations

considered in this study, except for Stockholm. This means that extreme sea levels are more influenced by the cumulative effects of sea level rise during the 2nd half of the 20th century than during earlier periods of time. Furthermore, it can be noticed that the longest time series (Brest) includes records before the onset of sea level rise, during which the influence of NAO is expected to be prominent, as quantitatively confirmed in figure 6 (top left panel).

Second, we analyse whether the dominance of MSL still prevails when the threshold chosen to select HSL events is lowered, e.g. down to 99% (approximately corresponding to the level of the 90 highest hourly sea level values over the considered year). Figure 7 shows that the conclusions drawn from

figure 5 are no longer valid for most tide gauges at the exception of Cuxhaven (which still presents a dominant combined role of AO and MSL) and at Gedser (where hardly any tendency can be identified). The decreased influence of MSL is noticeable for Stockholm, Halifax, San Francisco and Honolulu with a clear increased influence of different CI. For instance, at Halifax, AMO and AO both have inclusion probability larger than 80%. For Brest and Newlyn, the influence of MSL appears to be diminished in comparison to the one of NAO and AO respectively.

The sensitivity to the threshold chosen to select extreme sea level can be related to a physical explanation. Since this threshold is lowered from the annual 99.9th to the annual 99th percentile, the number of events included in the analysis is increased. Consequently, HSL time series corresponding to the lowest quantile reflect more persistent storm weather conditions, which are themselves more apparent in the mean annual CI (e.g. Seierstadt *et al* 2007). Hence, although consistent with an intuitive analysis, the results presented in figures 5 and 7 confirm quantitatively that the influence of CI increases when lowering the threshold for selecting extreme sea level.

## 5. Concluding remarks and further work

The proposed approach based on structural time series modelling provides a rigorous setting for measuring the importance of MSL in the evolution of HSL. The proposed procedure in combination with a Bayesian treatment of the fitting process enables to deal with the limitations of current available procedures, namely: (1) the presence of multiple alternative drivers; (2) the temporal nature of the data and (3) the uncertainty in the model construction. By applying *bsts* modelling procedure, we evaluate the inclusion probability, which measures how often a given driver is included in the regression model to explain HSL. This shows that, for the considered tide gauges, the MSL signal dominates the HSL evolution (except for Gedser). This result is in agreement with past studies, which have outlined the key role of MSL in extreme sea levels (see for instance the recent study by Marcos and Woodworth 2017). Interestingly, the MSL influence is clearly limited when the analysis is conducted before 1950, hence indicating a more marked influence of MSL on HSL as the cumulated amount of sea level rise increases (except for Stockholm). Our study also highlights that MSL is not the only and unique contributor. This is particularly evidenced at three tide gauges where AO index appears to have strong influence, but with opposite effect, namely positive at Stockholm and Cuxhaven and negative at Halifax. A clear evidence of strong Niño 1 + 2 SST influence is also evidenced at San Francisco, which appears to be consistent with the city location on the west coast of the US. These results may also suggest a connection

between the climate driver and the considered tide gauge's location (see e.g. Woodworth *et al* 2011, Marcos *et al* 2015), i.e. Pacific North American tide gauges are more likely be dominated by indices characterizing El Niño and European tide gauges are more likely be influenced by major climate modes of the Atlantic North Ocean. Marcos *et al* 2015 also reported identification of significant correlation to 'remote' CI (e.g. correlation of San Francisco extreme to NAO). Similar 'remote' influence of CI on storminess have been identified earlier by Seierstadt *et al* (2007). To some extent, they are evidenced by the *bsts* model as well, but mostly until 1950. Overall, our results suggest that the *bsts* model behaves more robustly as it identifies only a couple of important drivers, which remain consistent with the considered location.

In the present study, we have focused on measuring the importance of MSL and CI in the HSL time signal. Yet, sea level time series at tide gauges are known to be the complex result of different phenomena in interplay (in addition to MSL) including: (1) multidecadal variability related to CI. The present study focused on six major CI, but other could be easily integrated and in particular new CI specifically developed for the purpose of HSL variability study (as Wahl and Chambers 2016 did); (2) effects of river flows (in estuaries, see e.g. Piecuch *et al* 2018) and evolving complex nonlinear shallow water processes, affecting the tidal levels and constituents (Woodworth 2010), as evidenced for instance in the northern part of the German Bight (Arns *et al* 2015) or along the US coasts (e.g. Ray 2009); (3) nonlinear vertical ground motions (e.g. Raucoules *et al* 2018) (4) technological artifacts such as modification in the protocol for data acquisition errors (e.g. Woodworth 2010).

Incorporating other predictive time series related to additional physical processes driving extreme sea levels constitutes a line for future research. This is made possible owing to the high flexibility of the proposed setting whatever the type and number of predictive time series.

## Acknowledgments

We acknowledge financial support of Project EUPHEME, which is part of ERA4CS, an ERA-NET initiated by JPI Climate, and funded by FORMAS (SE), BMBF (DE), BMWFW (AT), IFD (DK), MINECO (ES), ANR (FR) with co-funding by the European Union (Grant 690462). We thank Robert Vautard, Pascal Yiou and Marta Marcos for useful discussions on attribution techniques and extreme sea levels time series at tide gauges. We thank two reviewers for their comments that led to a significant improvement of the article.

## ORCID iDs

Jeremy Rohmer  <https://orcid.org/0000-0001-9083-5965>

Gonéri Le Cozannet  <https://orcid.org/0000-0003-2421-3003>

## References

- Allan R J, Nicholls N, Jones P D and Butterworth I J 1991 A further extension of the tahiti-darwin SOI, early SOI results and Darwin pressure *J. Clim.* **4** 743–9
- Ambaum M H P, Hoskins B J and Stephenson D B 2001 Arctic oscillation or north atlantic oscillation? *J. Clim.* **14** 3495–507
- Arns A, Wahl T, Dangendorf S and Jensen J 2015 The impact of sea level rise on storm surge water levels in the northern part of the German Bight *Coast. Eng.* **96** 118–31
- Barnard P L *et al* 2017 Extreme oceanographic forcing and coastal response due to the 2015–2016 El Niño *Nat. Commun.* **8** 14365
- Enfield D B, Mestas-Nunez A M and Trimble P J 2001 The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US *Geophys. Res. Lett.* **28** 2077–80
- George E I and McCulloch R E 1993 Variable selection via Gibbs sampling *J. Am. Stat. Assoc.* **85** 398–409
- Hoeting J A, Madigan D, Raftery A E and Volinsky C T 1999 Bayesian model averaging: a tutorial (disc: P401-417) *Stat. Sci.* **14** 382–401
- IPCC 2013 *Climate Change 2013. The Physical Science Basis. Contribution of Working Group I to the 5th Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker (Cambridge: Cambridge University Press)
- Können G P, Jones P D, Klotfen M H and Allan R J 1998 Pre-1866 extensions of the Southern Oscillation Index using early Indonesian and Tahitian meteorological readings *J. Climate* **11** 2325–39
- Marcos M, Calafat F M, Berihuete A and Dangendorf S 2015 Long-term variations in global sea level extremes *J. Geophys. Res. Oceans* **120** 8115–34
- Marcos M and Woodworth P L 2017 Spatio-temporal changes in extreme sea levels along the coasts of the North Atlantic and the Gulf of Mexico *J. Geophys. Res. Oceans* **122** 7031–48
- Mawdsley R J and Haigh I D 2016 Spatial and temporal variability and long-term trends in skew surges globally *Frontiers Mar. Sci.* **3** 29
- Menéndez M and Woodworth P L 2010 Changes in extreme high water levels based on a quasi-global tide-gauge dataset *J. Geophys. Res.* **115** C10011
- Moritz S and Bartz-Beielstein T 2017 imputeTS: time series missing value imputation in R *R. J.* **9** 207–18
- Piecuch C G, Bittermann K, Kemp A C, Ponte R M, Little C M, Engelhart S E and Lentz S J 2018 River-discharge effects on United States Atlantic and Gulf coast sea-level changes *Proc. Natl Acad. Sci.* **115** 7729–34
- Raucoules D, Le Cozannet G, de Michele M and Capo S 2018 Observing water-level variations from space-borne high (SAR) resolution Synthetic Aperture Radar image correlation *Geocarto Int.* **33** 977–87
- Ray R D 2009 Secular changes in the solar semidiurnal tide of the western North Atlantic Ocean *Geophys. Res. Lett.* **36** L19601
- Rayner N A, Parker D E, Horton E B, Folland C K, Alexander L V, Rowell D P, Kent E C and Kaplan A 2003 Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century *J. Geophys. Res.* **108** 4407
- Ropelewski C F and Jones P D 1987 An extension of the Tahiti-Darwin Southern Oscillation Index *Mon. Weather Rev.* **115** 2161–5
- Seierstad I A, Stephenson D B and Kvamstø N G 2007 How useful are teleconnection patterns for explaining variability in extratropical storminess *Tellus* **59** 170–81
- Schwab M, Meinke I, Vanderlinden J P and von Stoch H 2017 Regional decision-makers as potential users of Extreme Weather Event Attribution—Case studies from the German Baltic Sea coast and the Greater Paris area *Weather Clim. Extremes* **18** 1–7
- Scott S L and Varian H R 2015 Bayesian variable selection for nowcasting economic time series *Economic Analysis of the Digital Economy* (Chicago, IL: University of Chicago Press) pp 119–35
- Scott S L and Varian H R 2014 Predicting the present with bayesian structural time series *Int. J. Math. Modelling Numer. Optim.* **5** 4–23
- Talke S A, Orton P and Jay D A 2014 Increasing storm tides in New York Harbor, 1844–2013 *Geophys. Res. Lett.* **41** 3149–55
- Visser H, Dangendorf S and Petersen A C 2015 A review of trend models applied to sea level data with reference to the ‘acceleration-deceleration debate’ *J. Geophys. Res. Oceans* **120** 3873–95
- Wahl T 2017 Sea-level rise and storm surges, relationship status: complicated! *Environ. Res. Lett.* **12** 111001
- Wahl T and Chambers D P 2015 Evidence for multidecadal variability in US extreme sea level records *J. Geophys. Res. Oceans* **120** 1527–44
- Wahl T and Chambers D P 2016 Climate controls multidecadal variability in US extreme sea level records *J. Geophys. Res. Oceans* **121** 1274–90
- Wong T, Klufas A, Srikrishnan V and Keller K 2018 Neglecting model structural uncertainty underestimates upper tails of flood hazard *Environ. Res. Lett.* **13** 074019
- Woodworth P L 2010 A survey of recent changes in the main components of the ocean tide *Cont. Shelf Res.* **30** 1680–91
- Woodworth P L and Blackman D L 2004 Evidence for systematic changes in extreme high waters since the mid-1970s *J. Climate* **17** 1190–97
- Woodworth P L, Hunter J R, Marcos M, Caldwell P, Menendez M and Haigh I 2017 Towards a global higher-frequency sea level data set *Geosci. Data J.* **3** 50–9
- Woodworth P L, Menéndez M and Gehrels W R 2011 Evidence for century-timescale acceleration in mean sea levels and for recent changes in extreme sea levels *Surv. Geophys.* **32** 603–18