Received: 15 February 2022 | Revised: 19 July 2022 | Accepted: 8 September 2022

DOI: 10.1111/jfr3.12862

SPECIAL ISSUE

CIWEM Chartered Institution of Water and Environmental Management | Journal of Flood Risk Management | WILEY

Extreme skew surge estimation combining systematic skew surges and historical record sea levels on the English Channel and North Sea coasts

Laurie Saint Criq[1] | **Eric Gaume**[2] | **Yasser Hamdi**[3] | **Taha B. M. J. Ouarda**[1]

[1]Centre Eau-Terre-Environnement, Institut national de la recherche scientifique, Quebec, Canada

[2]Géotechnique Environnement Risques Naturels et Sciences de la Terre, Université Gustave Eiffel—Campus de Nantes, Bouguenais, France

[3]Bureau d'expertise en hydrogéologie et sur les risques d'inondation, météorologiques et géotechniques, IRSN, Fontenay-aux-Roses, France

Correspondence
Laurie Saint Criq, Centre Eau-Terre-Environnement, Institut national de la recherche scientifique, Quebec, Canada.
Email: lauriesaintcriq@gmail.com

Abstract

Coastal planning implies the estimation of extreme sea levels. As the distribution of astronomical high tides can be predicted, most recent publications suggest focusing on the estimation of extreme skew surges. Historical information, record sea levels observed before the beginning of systematic tide gauge recordings, can improve estimations. The corresponding skew surges can be estimated but are not necessarily exhaustive. Indeed, some historical extreme skew surges can remain unnoticed if they are combined with low or moderate tides, or for a variety of reasons. To deal with this exhaustiveness issue, a previous publication proposed an unbiased method for combining systematic period skew surges with historical period extreme sea levels. This method appeared more reliable than previously proposed approaches. The present study aims at presenting a broader evaluation of this method, based on its application to nine sites located on the English Channel and North Sea coasts. The method is also improved to consider several historical periods and various types of historical information. Results confirm the method to be reliable, useful, and relevant. A number of recommendations is also formulated for the selection and use of historical information for sea level frequency analyses.

KEYWORDS

coastal, extreme events, extreme value statistics, statistical methods

1 | INTRODUCTION

Coastal planning needs to consider the risk of marine submersion especially in the present context of climate change. Then, estimating extreme sea levels, and especially extreme sea levels at high tide (maximum sea levels) which are the most dangerous, is essential for design purposes. Maximum sea levels can be defined as the combination of the predicted astronomical high tide and the skew surge. The astronomical high tide is a deterministic and predictable variable caused by gravitational forces. The skew surge is the difference between the observed maximum sea level and the predicted astronomical high tide during a tidal cycle and it is caused by atmospheric phenomena. The statistical distribution of maximum sea levels (or sea levels) is usually obtained after a convolution of the statistical distribution of predicted astronomical high tides (or predicted astronomical tides) and the statistical distribution of skew surges (or surges) which needs to be estimated (Batstone et al., 2013; Dixon & Tawn, 1999; Liu et al., 2010; Mazas et al., 2014; Pugh & Vassie, 1978; J. Tawn et al., 1989; J. A. Tawn, 1992; Tomasin & Pirazzoli, 2008). The independence between astronomical high tides and skew surges can be reasonably assumed (Williams et al., 2016) compared to the independence between astronomical tides and surges. Hence, the use of skew surges instead of surges makes the convolution easier because the interactions between both components of maximum sea levels can be neglected (Mazas et al., 2014). Moreover, according to (Arns et al., 2015; Batstone et al., 2013; Wahl & Chambers, 2015), the skew surge is a better indicator of the meteorological impact on sea level than the surge.

The integration of historical information allows to increase the sample size, to enrich the data set in the right tail of the distribution and then, to reduce the estimation uncertainties of the statistical distribution (Benito et al., 2004; Gaál et al., 2010; Hamdi et al., 2015; Ouarda & Ashkar, 1998; Payrastre et al., 2011; Reis & Stedinger, 2005). In the context of coastal analysis, historical events are the record maximum sea levels that occurred before the beginning of the systematic observation period—that is, period during which tide gauge recordings are available—and that may be retrieved from archives or field surveys. The skew surges corresponding to these records can be estimated, but some extreme skew surges can easily remain unnoticed if they are combined with low or moderate astronomical high tides and do not generate extreme maximum sea levels (Outten et al., 2020). This is visible with the measured systematic data sets: some extreme skew surges, even the largest, can be missed when extreme maximum sea levels are sampled (see Figure 1 for an example). Then, the exhaustiveness of historical information for the calibration of a skew surge statistical method, which is essential for an unbiased statistical inference (Gaume, 2018), cannot be guaranteed in the case of skew surges. Frau et al. (2018) and Hamdi et al. (2015) attempt to overcome the issue of the historical skew surges exhaustiveness but respectively present some positive and negative bias demonstrated through Monte Carlo simulations in Saint Criq et al. (2022), hereinafter referred to as SQ22. A large number of random samples was simulated with the same characteristics based on four real case studies. Then the maximum likelihood estimates and the posterior credibility intervals were analyzed with boxplots and rank histograms (Nguyen et al., 2014).

SQ22 propose a new statistical inference procedure including the historical record maximum sea levels, instead of their corresponding skew surges. Indeed, the historical record maximum sea levels larger than a high enough threshold can reasonably be considered exhaustive during the historical period. For the calibration of the statistical distribution of skew surges, this method combines, in a single Bayesian inference procedure, information of two different natures: skew surges for the systematic period and extreme maximum sea levels for the historical period. The systematic skew surges are sub-sampled to a peaks-over-threshold (POT) sample and are modeled by a Generalized Pareto (GP) distribution. The statistical distribution of the historical maximum sea levels is replaced in the likelihood by a numerical approximation which allows the implementation of this new inference procedure. This method, which will be called the historical sea levels (HSLs) method, is based on the hypotheses of independence between astronomical high tides and skew surges.

The HSL method appears reliable for the quantiles and parameters estimated by the maximum likelihood, as well as for the posterior credibility intervals of quantiles estimated in the Bayesian framework. The reduction of the averaged widths of the posterior credibility intervals (compared to the analysis of the systematic data only) shows the added value of the historical record maximum sea levels. The HSL method is applied to some case studies selecting a historical threshold that is not exceeded during the historical period according to the available information. This deterioration of the historical information is only moderately observable on the obtained posterior credibility intervals. This means that knowing that a sea level has not been exceeded—for example, a coastal infrastructure has not been submerged—during a

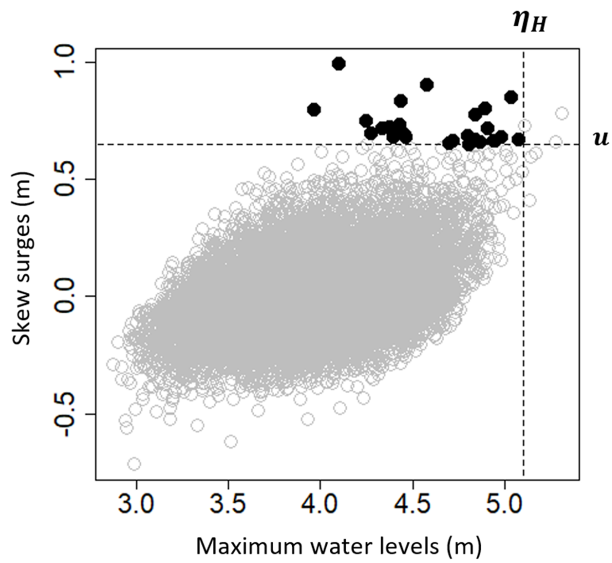

FIGURE 1 Systematic events at Aberdeen (UK). The black points represent the extreme skew surges (larger than u) when sampling the extreme maximum sea levels (larger than η_H).

CIWEM Chartered Institution of Water and Environmental Management Journal of Flood Risk Management—WILEY

historical period is a valuable information from a statistical point of view. This is an interesting result, especially at sites without available historical information.

The present study aims to challenge the HSL method proposed in SQ22 and its conclusions with new case studies. The HSL method is summarized in Section 2.1, the likelihood formulation is adapted to consider various uncertainty levels affecting historical information—that is, in case of uncertainties on historical information. Section 2.2 presents the evaluation methodology. The application case studies are presented in Section 2.3. The application results are illustrated and discussed in Section 3.

2 | METHODS

2.1 | Combination of skew surges and sea levels in the same likelihood

SQ22 propose to combine in a single Bayesian inference, $X_{sys} = \{x_{sys,1}, x_{sys,2}, ..., x_{sys,m}\}$, the systematic POT sample of m skew surges exceeding the threshold u during the systematic period of w_S years and $Z_{hist} = \{z_{hist,1}, z_{hist,2}, ..., z_{hist,h_z}\}$, the h_z historical record maximum sea levels exceeding the threshold η_H during the historical period of duration w_H years. The global likelihood of the systematic skew surges and the historical record maximum sea levels $L(X_{sys}, Z_{hist}|\theta)$ is given by the product of the likelihood of the systematic skew surge sample $L(X_{sys}|\theta)$ and the likelihood of the HSL sample $L(Z_{hist}|\theta)$:

$$L(X_{sys}, Z_{hist}|\theta) = L(X_{sys}|\theta).L(Z_{hist}|\theta) \quad (1)$$

with $\theta = (\sigma, \xi, \lambda)$ the parameters to estimate. Skew surges exceeding the threshold u are chosen to be modeled by a GP distribution of scale parameter $\sigma > 0$ and shape parameter $\xi \in \mathbb{R}$. The number of skew surges exceeding the threshold u per year follows a Poisson process of intensity λ (Coles, 2001).

$L(X_{sys}|\theta)$, the likelihood of the systematic skew surge sample is defined, $\forall\ i \in \{1, ..., m\}$, $x_{sys,i}$ independent and identically distributed (i.i.d), by:

$$L(X_{sys}|\theta) = \mathbb{P}_\theta(M = m). \prod_{i=1}^{m} f_\theta(x_{sys,i}) \quad (2)$$

with $\mathbb{P}_\theta(M = m) = \frac{(\lambda w_s)^m}{m!} \exp(-\lambda w_S)$ the probability of observing m skew surges exceeding u during a systematic period of w_S years and f_θ the GP probability density function.

On average, there are 706 high tidal levels during a year so, there are $N_{w_H} = 706 \times w_H$ high tidal levels during the historical period. Here, two consecutive events are assumed to be independent, but a storm can last 2 or 3 days, especially on the French Atlantic coast (Bulteau et al., 2015; Kergadallan et al., 2014). Then, to consider the temporal dependence, the number of high tidal levels per year should be reduced. $L(Z_{hist}|\theta)$, the likelihood of the historical maximum sea levels, describes the $N_{w_H} - h_z$ maximum sea levels that do not exceed η_H and the h_z maximum sea levels that exceed η_H during the historical period of duration w_H years. $\forall\ i \in \{1, ..., h_z\}$, $z_{hist,i}$ i.i.d, $L(Z_{hist}|\theta)$ is defined by:

$$L(Z_{hist}|\theta) = \widetilde{G}_\theta(\eta_H)^{N_{w_H} - h_z}. \left[1 - \widetilde{G}_\theta(\eta_H)\right]^{h_z}. \prod_{i=1}^{h_z} \frac{\widetilde{g}_\theta(z_{hist,i})}{1 - \widetilde{G}_\theta(\eta_H)} \quad (3)$$

with $\widetilde{g}_\theta$, $\widetilde{G}_\theta$ respectively the probability density and cumulative distribution functions of high tidal levels which can be numerically computed given the statistical distributions of astronomical high tides, skew surges lower than u and skew surges exceeding u (for more details, see Appendix A in SQ22).

Likelihood (3) is valid if the historical maximum sea levels z_{hist} are considered to be accurately known. However, uncertainties affecting z_{hist} can be accounted for (see Equation 4): (a) range data, the historical maximum sea levels are known with uncertainties, they are comprised between the perception threshold η_H and a value ϕ_H with $\phi_H > \eta_H$, (b) binomial censored data, it is known that the threshold η_H has been exceeded h_z times over the historical period but the corresponding maximum sea levels are unknown, (c) lower limit data, a limit value ϕ_H has not been exceeded during the historical period. Note that the range data likelihood (4a) could be more precise taking into account uncertainties on each HSL such as $\forall$ $i\ \{1, ..., h_z\}$, $z_{hist,i} \in \left[\eta_H^i, \phi_H^i\right]$.

$$L(Z_{hist}|\theta) = \begin{cases} \widetilde{G}_\theta(\eta_H)^{N_{w_H} - h_z}. \left[\widetilde{G}_\theta(\phi_H) - \widetilde{G}_\theta(\eta_H)\right]^{h_z} \\ \quad \text{if (a) range data,} \\ \widetilde{G}_\theta(\eta_H)^{N_{w_H} - h_z}. \left[1 - \widetilde{G}_\theta(\eta_H)\right]^{h_z} \\ \quad \text{if (b) binomial censored data,} \\ \widetilde{G}_\theta(\phi_H)^{N_{w_H}} \\ \quad \text{if (c) lower limit data.} \end{cases} \quad (4)$$

2.2 | Methods

The RStan package was used to conduct Bayesian Markov Chain Monte Carlo inferences based on the formulated likelihood with non-informative priors. The results of the inference procedure consist in the posterior densities for the calibrated parameters $\theta = (\sigma, \xi, \lambda)$ and of the corresponding skew surge quantiles. Note that Stan model related to the likelihood based on systematic skew surges and historical maximum sea levels (Equation 3) was made available within the scope of SQ22 at https://doi.org/10.5281/zenodo.6260203.

The HSL method is challenged with different evaluation criteria based on several tide gauges covering a variety of situations by their systematic and historical characteristics (see Section 2.3). First, the HSL method is implemented on observed systematic skew surges and exact historical maximum sea levels to measure the possible reduction of uncertainties by the integration of historical record maximum sea levels. The exact historical maximum sea levels are supposed to be perfectly known without any uncertainty. Then, uncertainties on historical maximum sea levels have to be accounted for to be in a more realistic context (see Equation 4). Then, the HSL method is implemented on observed systematic skew surges and historical information according to its accuracy (exact data, range data, binomial censored data and lower limit data) to evaluate the sensitivity of the accuracy historical information on the inference results, especially of the width of the posterior credibility intervals. In the cases of range data and lower limit data, the limit value ϕ_H is chosen as the maximum of the observed HSL added to 0.01 ($\phi_H = \max(z_{hist}) + 0.01$). Finally, the objective is to evaluate the most relevant descriptive parameters of the historical information. Then, from the observed data sets, a sensitivity analysis is conducted to the length of the historical duration w_H and to the limit value ϕ_H on the posterior credibility intervals. In the case of exact data (likelihood (3)), w_H is taken as $w_H + \alpha$ with α the duration per historical record event. There are uncertainties to estimate w_H and in the worst case, the bias on w_H is around α years. In the cases of range data (likelihood (4a)) and lower limit data (likelihood (4c)), ϕ_H is taken as $\max(z_{hist}) + \beta$, with $\max(z_{hist})$ the maximum record HSL and β varying in $\{0.01, 0.5, 1, 2\}$.

2.3 | Case study

Within the scope of this study, a database is developed, selecting tide gauges whose historical information is available or with long systematic series that can be split, for sake of illustration, into two periods (systematic and historical) to create artificial historical information as was done for Brest and Saint Nazaire in SQ22. This database contains nine tide gauges: Calais, Cherbourg, Le Havre and Saint Malo on the English Channel coast, Nieuport and Oostende on the Belgium North coast, Aberdeen on the UK North coast, and Delfzijl and Vlissingen on the Netherlands North coast (Bernardara et al., 2011; Haigh et al., 2010, 2016; Sterl et al., 2009; Woodworth et al., 2007). Figure 2 shows their geographic location.

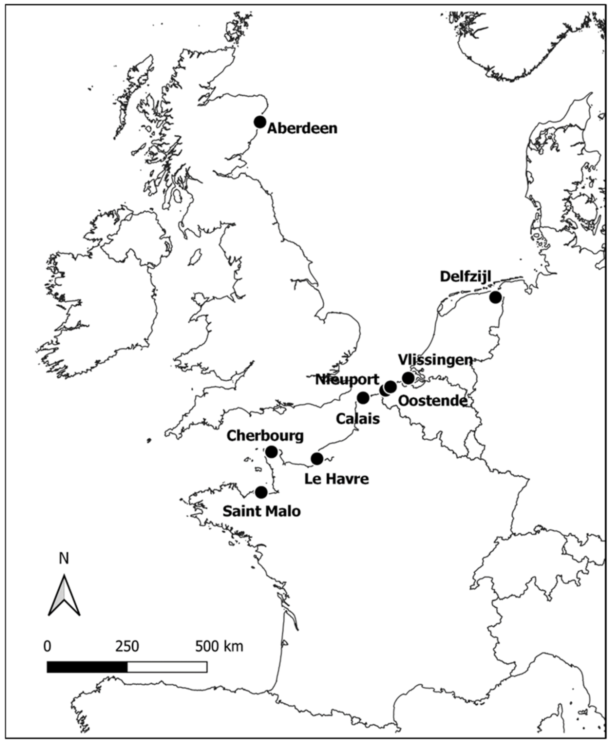

FIGURE 2 Geographic location of selected tide gauges

2.3.1 | Systematic data sets

The French Hydrographical and Oceanographical Service (data.shom.fr), the Flemish Agency for Maritime and Coastal Services (meetnetvlaamsebanken.be), the British Oceanographic Data Centre (bodc.ac.uk) and the Ministery of Infrastructure and Water Management of the Netherlands (rijkswaterstaat.nl) make available the hourly tide gauge data at the mentioned sites. These data are processed with a harmonic analysis through the R package *TideHarmonics* (Stephenson, 2015). As tidal predictions are calculated for the present time, to obtain the actual surges of past periods, the data must be corrected for a possible eustatism. If there is a significant trend in the temporal evolution of the annual mean of sea levels (calculated according to the Permanent Service for Mean Sea Level recommendations), then the series of observed sea levels is adjusted so that the corresponding series of annual means is stationary. The hourly astronomical tidal levels are computed, then, the series of astronomical

TABLE 1 Characteristics of the systematic data sets, the lowest astronomical high tide (LAHT), and the highest astronomical high tide (HAHT)

Site	Tide gauge recording period	w_S (years)	u (m)	n	Tide surge ratio[a]	LAHT (m)	HAHT (m)
Aberdeen	1990–2021	24.61	0.65	26	12.50	3.05	4.85
Calais	1941–2021	42.56	0.60	43	25.20	5.16	7.84
Cherbourg	1943–2020	46.40	0.44	45	28.26	4.38	7.13
Delfzijl	1987–2021	34.53	1.71	33	2.05	0.81	1.84
Le Havre	1938–2021	50.72	0.60	69	25.59	5.92	8.55
Nieuport	2000–2020	20.18	0.71	26	13.65	3.28	5.43
Oostende	2000–2020	20.26	0.75	23	12.63	3.19	5.26
Saint Malo	1986–2021	26.64	0.50	30	47.62	7.81	13.52
Vlissingen	1987–2021	34.54	1.00	29	5.41	0.89	2.92

Note: The systematic duration w_S does not account for the period of missing values.
[a]Ratio of the 98% astronomical high tide to the 98% skew surge quantile (Dixon & Tawn, 1999).

high tides and skew surges can be extracted. Finally, the GP parameter stability criterion—based on fitting the model at a range of thresholds (Coles, 2001)—helps to select the POT threshold u (see Table 1) and to build the POT sample. The nine POT samples obtained are verified to be i.i.d. The independence between skew surges and astronomical high tides, which is a necessary hypothesis to apply the HSL method, is tested for the nine tide gauges through the approach proposed by Arns et al. (2020) This independence can be assumed between the extreme skew surges and their associated astronomical high tides, but not for astronomical high tides and skew surges leading to the extreme maximum sea levels (see Appendix A). The systematic data sets for the nine selected tide gauges have different characteristics (see Table 1). The systematic period is more or less well documented (the systematic duration varies between 20.18 and 50.72 years and the POT sample size varies between 23 and 69), the orders of magnitude for the skew surges are different (the POT threshold varies between 0.44 and 1.71 m) and there are various patterns for the astronomical high tide component. For the nine tidal gauges, a tide surge ratio higher than 2 confirms the need to model separately the astronomical high tides and the skew surges as shown in (Haigh et al., 2010). Saint Malo has a very high tide surge ratio and its highest astronomical high tide is very important compared to the other tide gauges, then the astronomical high tide appears to be the dominant component for record maximum sea levels. In this case, the HSL method is particularly interesting since the historical skew surges are certainly non-exhaustive and also particularly challenged because of the limited weigh of the extreme skew surges on the historical extreme sea level distribution. Calais, Cherbourg, Le Havre and Brest (SQ22) have similar values of tide surge ratio, lowest and highest astronomical high tide which testify of the relative importance of the high tide compared to the skew surge. At Aberdeen, Nieuport and Oostende, the astronomical high tide component is even more moderate, the tide surge ratio is about 13. Finally, at Delfzijl and Vlissingen, the maximum sea levels are clearly driven by the skew surge. Indeed, their tide surge ratios are reduced and the POT threshold is higher than the lowest astronomical high tide which is especially visible at Delfzijl. Then, at these stations, record maximum sea levels are related to record skew surges and the use of HSL method would be less essential than at Saint Malo since historical record skew surges would tend to be (or very closed to be) exhaustive and the method proposed by Hamdi et al. (2015) could then be used with probable limited bias.

2.3.2 | Historical data sets

Historical record maximum sea levels are provided by a database developed by the French Institute for Radiological Protection and Nuclear Safety (Giloy et al., 2018) for Belgium and French tide gauges and by a UK database developed by the British Oceanographic Data Centre from 1915 (Haigh et al., 2017) for Aberdeen. The systematic series of Delfzijl and Vlissingen start respectively in 1879 and 1863, present a measuring gap between 1960 and 1987. Then, they are split into a systematic sample (after 1987) and a historical sample (before 1960). For both Dutch tide gauges, the historical record maximum sea levels above 4 m are extracted. The hypothesis of stationary is accepted for the complete series of Delfzijl (from 1879 to 2021), but rejected for the complete series

TABLE 2 Characteristics of the historical data sets

Site	Range	w_H (years)	η_H (m)	ϕ_H (m)	h_z	$\mathbb{P}(\mathrm{A}\,\vert\,X_{sys})$	$\mathbb{P}(\mathrm{B}\,\vert\,X_{sys})$	w_H/w_S
Aberdeen	1915–1990	80	5.10	5.10	2	1.00	1.00	3.25
Calais	1905–1941	50	7.70	7.70	1	1.00	1.00	1.17
Cherbourg	1821–1943	130	7.17	7.96	2	1.00	0.00	2.80
Delfzijl	1879–1960	28.15	4.34	4.34	1	0.96	0.96	0.82
Le Havre	1882–1938	70	8.40	8.55	3	1.00	1.00	1.38
Le Havre	1938–2021	33.26	9.28	9.28	1	0.00	0.00	0.66
Nieuport	1932–2000	80	6.08	6.73	5	0.99	0.00	3.96
Oostende	1932–2000	80	6.01	6.66	4	0.82	0.00	3.95
Saint Malo	1877–1986	100	12.96	12.96	1	1.00	1.00	3.75
Vlissingen	1863–1960	17.84	4.05	4.54	3	0.00	0.00	0.52

Note: η_H and ϕ_H are the observed minimum and maximum historical maximum sea levels. $\mathrm{A}\,|\,X_{sys}$ (respectively $\mathrm{B}\,|\,X_{sys}$) represents the event "η_H (respectively ϕ_H) is exceeded at least h_z (respectively 1) time(s) during the historical period knowing the systematic data set." Then, $\mathbb{P}(\mathrm{A}\,|\,X_{sys}) = 1 - \sum_{i=0}^{h_z-1} \tilde{G}_{\widehat{\theta}}(\eta_H)^{N-i}$ and $\mathbb{P}(\mathrm{B}\,|\,X_{sys}) = 1 - \tilde{G}_{\widehat{\theta}}(\phi_H)^{N}$ are computed with the parameters $\widehat{\theta}$ estimated on systematic data sets by the maximum likelihood.

of Vlissingen (from 1863 to 2021). Indeed, the two highest skew surges occurred during the historical period and are very large compared to the other skew surges of the complete series. The complete series contains 43 skew surges at Delfzijl and 47 at Vlissingen.

The nine tide gauges also present various situations for the historical period (see Tables 2 and B1): the length of the historical period (from 17.84 to 130 years), the number of historical record maximum sea levels (from 1 event to 5 events) and the height of the record maximum sea levels (η_H varies between 4.05 and 12.96 m) which is directly related to the astronomical high tidal range (see Table 1). The historical duration w_H is chosen larger than the time laps between the first historical record and the beginning of the systematic period. $\mathbb{P}(\mathrm{A}|X_{sys})$ and $\mathbb{P}(\mathrm{B}|X_{sys})$ are computed to test the consistency of historical samples according to the systematic data sets, on the assumption that skew surges are following a GP distribution. At Aberdeen, Calais, Delfzijl, Le Havre and Saint Malo, the historical data set seems to be consistent with the distribution adjusted on the systematic skew surges ($\mathbb{P}(\mathrm{A}|X_{sys})$ and $\mathbb{P}(\mathrm{B}|X_{sys})$ close to one). In these cases, the systematic and historical data sets seem consistent and the systematic data sets do not seem to impose strong constraints on the skew surge distribution. At Cherbourg, Nieuport and Oostende, according to the systematic data sets, the threshold η_H can be easily exceeded h_z times during the historical period ($\mathbb{P}(\mathrm{A}\,|\,X_{sys})$ high), but the maximum historical observed sea level can hardly be exceeded during the historical period ($\mathbb{P}(\mathrm{B}|X_{sys}) = 0$). At Vlissingen, the sea level distribution estimated only on systematic skew surges can hardly allow the realization of the record maximum sea levels observed during the historical period ($\mathbb{P}(\mathrm{A}|X_{sys}) = \mathbb{P}(\mathrm{B}|X_{sys}) = 0$). The statistical consistency between systematic and historical data sets is highly questionable if we assume that the skew surges follow a GP distribution. Mixing both datasets in the same statistical inference procedure is likewise questionable and may result in problematic outcomes as will be illustrated hereafter.

Tide gauges can sometimes be defective as it was the case at Le Havre in 1984 when a major event occurred but could not be recorded (see Table B1). This observation is integrated as historical information in the HSL method. The 1984 event is very extreme and it appears reasonable to suppose that, if another event would have exceeded it during the systematic period but not measured, this event would have been retrieved. Then, the historical duration associated to this event is taken as the sum of the gap measures between 1938 and 2021. Consequently, at Le Havre, two historical periods have been defined (see Table 2).

3 | APPLICATION OF THE HSL METHOD TO OBSERVED DATA SETS

The HSL method is applied to observed data sets. In Section 3.1, the historical record maximum sea levels retrieved from archives (see Table B1) are supposed to be precisely known.

3.1 | Exact data

At Delfzijl and Vlissingen, as the complete skew surge series are available (see Section 2.3.2), it is possible to compare the results of the HSL method with the ideal case—that is, perfect knowledge of the historical skew surges. The adjusted credibility intervals with the HSL method are very similar to those obtained in the ideal case, and even if narrower, they are consistent with the data sets since they contain the complete skew surges series (see Figure 3). This result is very satisfactory since a large number of historical events is processed in the inference for the ideal case compared to the HSL historical samples: 43 (respectively 18) historical skew surges versus 1 (respectively 3) historical maximum sea levels at Delfzijl (respectively Vlissingen). This and similar results obtained in SQ22 for the French tide gauges of Brest and Saint Nazaire allow to validate the HSL method.

The inclusion of the exact historical record maximum sea levels leads to a significant reduction of the width of the posterior credibility intervals for all tide gauges

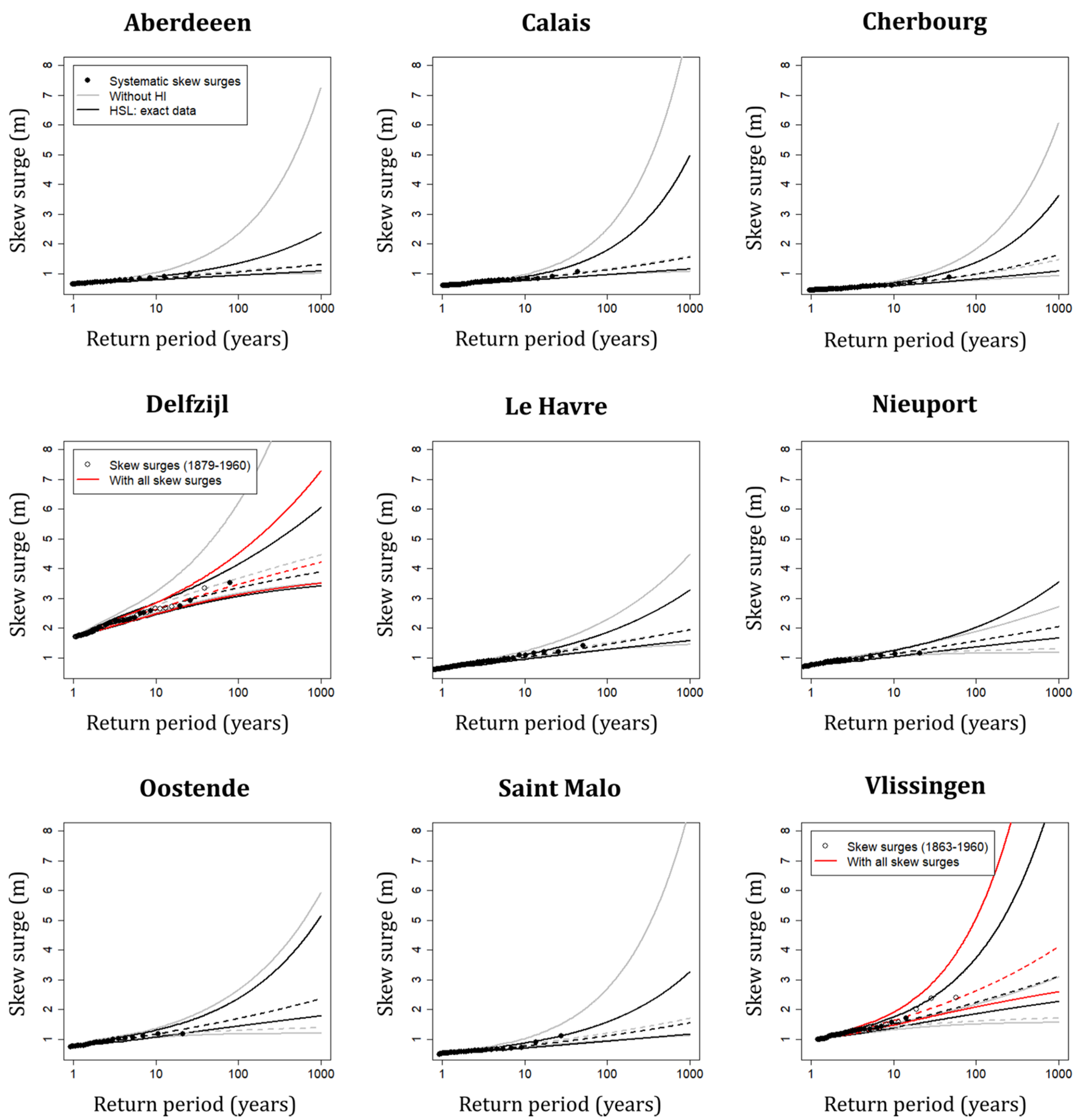

FIGURE 3 Ninety percent skew surge posterior credibility intervals without historical information (HI), based on the systematic data (gray line), integrating the historical record sea levels as exact data (black line) and in the ideal case, when all skew surges are known during the systematic and historical periods (black line) for Delfzijl and Vlissingen. The dotted lines represent the median estimates of each method.

except for Nieuport and Vlissingen, but to a redirection of the posterior credibility intervals, and for the last case, a significant increase of the estimated quantile (see Figure 3). At Nieuport and Vlissingen, this result was expected because of statistical inconsistency of the systematic and historical data sets. The reduction of uncertainties for the skew surge quantiles—that is, reduction of the posterior credibility interval width—is discussed in light of the consistency between systematic and historical samples (see $\mathbb{P}(A|X_{sys})$ and $\mathbb{P}(B|X_{sys})$ in Table 2). For all tide gauges and all methods, the median adjustments are in the lower part of the posterior credibility intervals. It is interesting to notice that the median adjustments obtained with the HSL method and without historical information are almost superposed.

At Aberdeen, Calais, Delfzijl and Saint Malo, the systematic and historical data sets appear statistically consistent and the posterior credibility intervals are significantly narrower in comparison to the other sites. Moreover, at Aberdeen and Saint Malo, the historical duration is at least three times longer than the systematic duration which helps to significantly lower the upper bound of the posterior credibility interval. At Calais, due to equivalent lengths of the systematic and historical periods, the uncertainties could not be as reduced as for Aberdeen and Saint Malo. At Delfzijl, the historical duration is slightly lower than the systematic duration, but the strong reduction of uncertainties (as for Aberdeen and Saint Malo) may be allowed because of the very high initial uncertainties—that is, the posterior credibility interval based on the systematic skew surges is very large. At Cherbourg and Oostende, the historical sample appears to contain, at least, one exceptional event clearly not consistent with the systematic sample ($\mathbb{P}(B|X_{sys}) = 0$ in Table 2), so the integration of historical maximum sea levels only allows to slightly reduce the uncertainties. A similar phenomenon is observed at Le Havre, for the second historical period (1938–2021).

As a first conclusion, these analyses indicate that some factors have an influence on the reduction of the credibility intervals, like the consistency between systematic and historical samples, the initial uncertainties and the length of the historical duration compared to the systematic duration. Note that the most important reductions (at Aberdeen, Calais, Delfzijl and Saint Malo) are obtained with only one or two record maximum sea levels exceeding the historical perception threshold, whereas the least reduced posterior credibility intervals (at Le Havre, Cherbourg and Oostende) are obtained with three or four record maximum sea levels (see Table 2 and Figure 3).

The inconsistency between systematic and historical data sets at Vlissingen and Nieuport leads to a redirection of the posterior statistical distributions and credibility intervals. Indeed, at Vlissingen, three very extreme maximum sea levels (see Table 2) are reported during a very short historical period compared to the systematic duration ($w_H \approx \frac{w_S}{2}$). At Nieuport, extreme events are also observed, but during a relatively long historical period ($w_H \approx 4 \times w_S$). Their empirical frequency is more in accordance with the systematic observations, then the redirection of the posterior credibility intervals is more limited than at Vlissingen.

3.2 | Sensitivity analysis of historical information quality

Accounting for the possible uncertainties affecting the historical information (range data, binomial censored data or lower limit data instead of exact data) has an impact that is only moderately observable on the posterior credibility intervals except in some particular cases and an almost invisible impact on the median adjustments (see Figure 4). This is consistent with the conclusions drawn by Payrastre et al. (2011) on river discharge series: the length of the historical period has more impact on the inference results than the number and accuracy of documented historical events.

At Aberdeen, Calais and Saint Malo, the posterior credibility intervals produced with any type of historical information are mostly superposed which means that no matter the accuracy of historical records, the information content is almost the same. Indeed, there is only one exceedance at Calais and Saint Malo and two exceedances of same value at Aberdeen. At Delfzijl, there is also one exceedance, the historical posterior credibility intervals are also very close except for binomial censored data, for which the upper bound is quite higher than the other ones. Indeed, this type of historical information is the least restrictive for the magnitude of historical maximum sea levels and leads to the highest upper bound as it is observed at most of the tide gauges (Aberdeen, Calais, Delfzijl, Le Havre, Nieuport, Oostende, and Vlissingen). There are also some particularities directly related to the samples.

At Le Havre and Nieuport, the four posterior credibility intervals are quite close. The case of Oostende is similar with a higher scattering. At Cherbourg, the posterior credibility intervals are superposed except the one obtained with binomial censored data. Indeed, in this case, the highest HSL seems highly improbable according

CIWEM Chartered Institution of Water and Environmental Management Journal of Flood Risk Management—WILEY

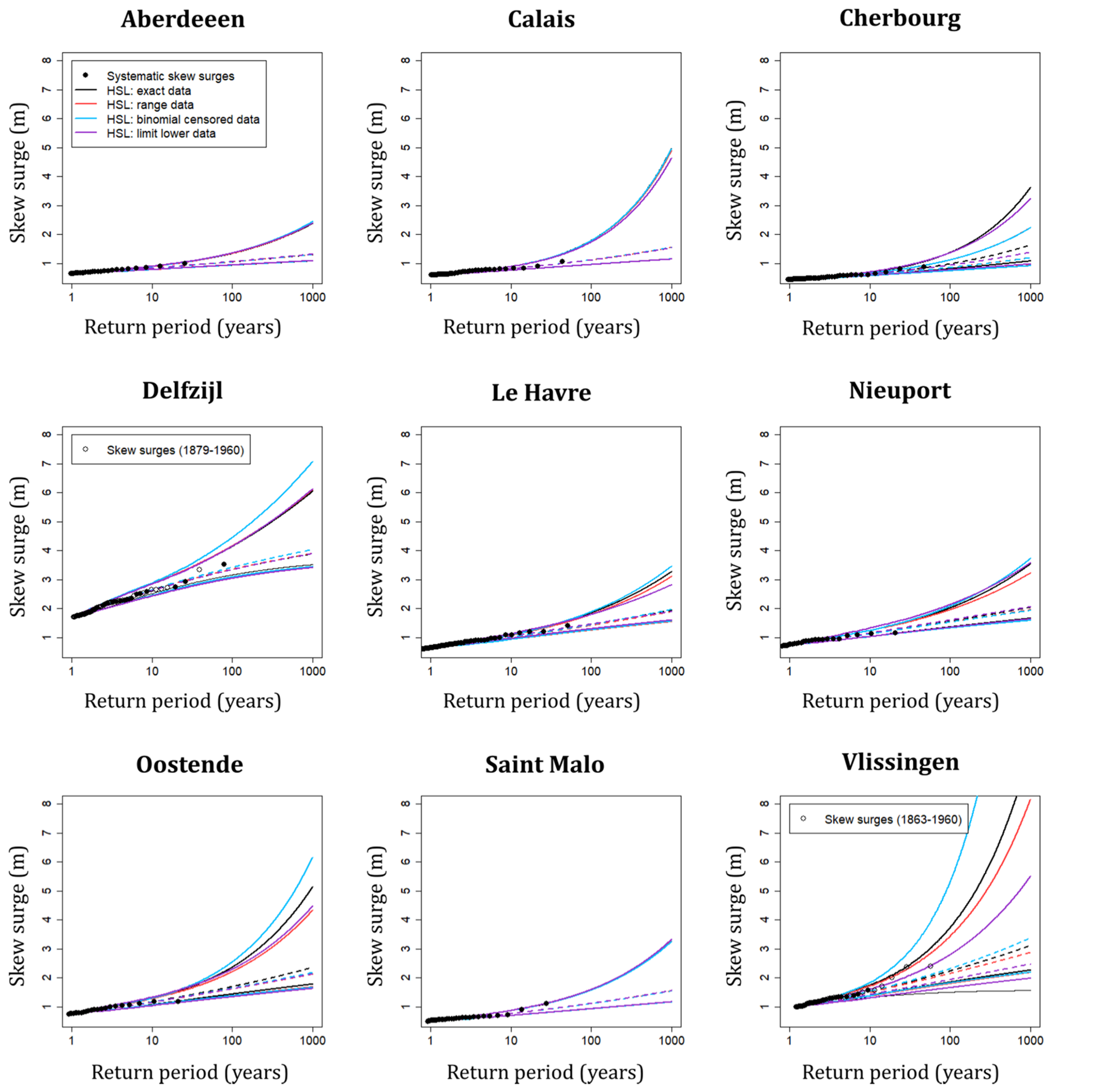

FIGURE 4 Ninety percent posterior skew surge posterior credibility intervals based on the systematic data and on historical maximum sea levels as exact data (black line), range data (red line), binomial censored data (blue line) or lower limit data (purple line). The dotted lines represent the median estimates of each method.

to the systematic data set (see Table 2), pulling up the calibrated statistical distribution. The binomial censored data method relaxes the constraint imposed to the calibration by this extreme historical record.

The case of Vlissingen is very special compared to the other ones since the systematic and historical data sets appear extremely inconsistent (see Section 2.3.2). The strong sensitivity to the uncertainties on the historical data, and then to the weight given to the historical data in the adjustment, reflects this statistical inconsistency.

3.3 | Complementary analysis

Some complementary analyses are conducted to evaluate the sensitivity of the inference outcomes to (1) the estimated historical duration w_H, (2) the estimated limit value ϕ_H for historical record maximum sea levels, when the "range data" likelihood is used, or (3) when the "lower limit data" likelihood is used. Similar results are obtained for all nine sites. For the sake of simplification, only the results of Aberdeen are presented herein (see

Figure 5). This data set has been selected because of the statistical consistency between systematic and historical data sets.

The historical duration w_H has to be estimated. It is, at least, the time laps between the first historical record and the last one or between the first historical record and the beginning of the systematic period. But, it is common to select a slightly larger value for w_H to avoid introducing biases as discussed by Schendel and Thongwichian (2017). According to these considerations, at Aberdeen, the historical duration has been estimated to $w_H = 80$ years. The historical inventory covers the period 1915 – 1990 and there are two reported record maximum sea levels, then the empirical return period of these records is about 30 years. Figure 5a shows the posterior credibility intervals, in the case of exact data, when the historical duration w_H varies in a reasonable range of 80 ± 30 years. The historical duration appears to have a moderate impact on the statistical inference results, the longer estimated historical duration, the lower are the uncertainties. Yet, the additional information content increases with the length of the estimated historical duration.

As expected, in the case of range data, the posterior credibility intervals are relatively insensitive to uncertainties affecting the estimated limit value ϕ_H for historical record maximum sea levels (see Figure 5b).

Finally, in the case of lower limit data, the posterior credibility intervals are very sensitive to the estimated limit value ϕ_H (see Figure 5c). The lower the limit value, the more informative is the added historical content. Yet, the possible range of the historical maximum sea levels is reduced leading to narrower posterior credibility intervals. When the lower limit value ϕ_H tends to infinity, the added value of the historical information is very reduced and the obtained adjustments tend to the ones obtained with systematic data only. Yet, a very high ϕ_H is not informative at all in the case of lower limit data.

Based on all previous analyses and Payrastre et al. (2011), the most determining information included in the historical data set appears to be the years without observed records. Then, a particular attention must be given to the determination of the value w_H. This conclusion is valid provided that the historical and systematic data sets appear consistent at the light of the calibrated statistical distribution. If it is not the case, it does not mean that the historical data is useless. On the contrary, it questions the adequacy of the statistical distribution and hence the basics of the risk assessment method. But this last issue goes far beyond the objectives of this paper.

4 | CONCLUSION

SQ22 proposed a new statistical inference procedure to properly integrate historical information about extreme maximum sea levels in skew surge statistical analyses, to overcome the issue of non-exhaustiveness of the skew surge variable. The procedure consists in combining data sets of different natures: skew surges for the recent period and maximum sea levels for the historical period. This new method is valid under the assumption of independence between astronomical high tides and skew surges and it is implemented by replacing the analytic expression of the probability density and cumulative functions of maximum sea levels by numerical approximations in the likelihood formulations. The flexibility of the likelihood formulation allows to consider different levels of uncertainties affecting the historical information (see Section 2.1) and several historical periods.

The method proposed in SQ22 was implemented herein on nine additional case studies with different

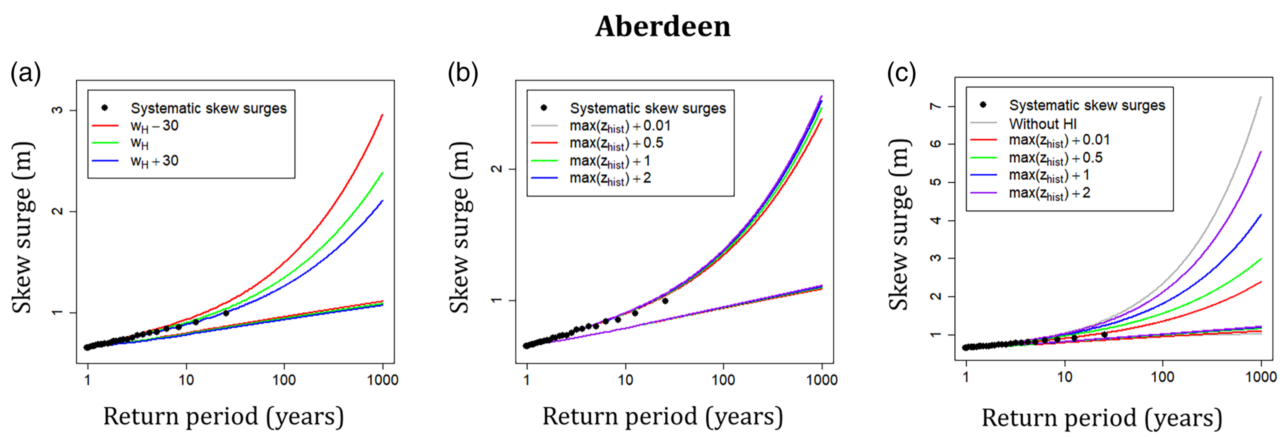

FIGURE 5 Ninety percent skew surge posterior credibility intervals based on the systematic data and the historical record maximum sea levels as (a) exact value adjusting the historical duration (likelihood (3)), (b) range value adjusting the upper bound of the interval ϕ_H (likelihood (4a)) and (c) lower limit data adjusting the value not exceeded ϕ_H (likelihood (4c)).

characteristics. The comparison with results based on a perfect knowledge of historical skew surges, when this information was available, confirm the relevance of the proposed approach. Indeed, the posterior credibility intervals appear unbiased and very close to those obtained with a perfect knowledge of the historical skew surges. The results of the present study indicate that all data, being recorded by tide gauges or not, are informative, deserve to be integrated in statistical analyses, on the condition of being exhaustive. The integration of historical information should lead to reduce the estimation uncertainties on condition that the systematic and historical data sets appear statistically consistent, at the light of the calibrated distribution. Otherwise, the historical information, if criticized and confirmed, still should be considered as of crucial importance. It will not help reducing statistical inference uncertainties but will clearly question the adequacy of the statistical distribution and hence the basics of the risk assessment method. In any case, even if less accurate than systematic measurements, historical information should be considered in risk assessment procedures.

The sensitivity analyses conducted reveal some key features for useful historical record series. First, the duration of the period covered by the historical inventory w_H appears to be of primary importance. The level of accuracy of the historical record maximum sea levels has little influence on the statistical inference results. Second, the limit value ϕ_H, which it has not been exceeded during the historical period, is decisive. Its underestimation will introduce significant biases in the inference procedure and should be avoided. But its overestimation, for sake of prudence, will reduce the historical information content. It is of course possible to consider several historical periods differing by their content and accuracy. Future efforts should focus on the integration of expert knowledge such as coastal structure with known altitude which has not been submerged during a considered historical period. This is very promising for sites with high uncertainties.

The HSL method presents good properties, but should be improved by considering the skew surge—high tide dependence and the temporal dependence between two consecutive events. One weakness of this method is that historical information is only available at a few tide gauges as historical data collection is difficult and expensive. Consequently, the use of expert knowledge and the integration of other types of information should be highly considered.

DATA AVAILABILITY STATEMENT

The French Hydrographical and Oceanographical Service (data.shom.fr), the Flemish Agency for Maritime and Coastal Services (meetnetvlaamsebanken.be), the British Oceanographic Data Centre (bodc.ac.uk) and the Ministery of Infrastructure and Water Management of the Netherlands (rijkswaterstaat.nl) make available the hourly tide gauge data at the mentioned sites. Historical record sea levels are provided by a database developed by the French Institute for Radiological Protection and Nuclear Safety (Giloy et al., 2018) for Belgium and French tide gauges and by a UK database developed by the British Oceanographic Data Centre from 1915 (Haigh et al., 2017) for Aberdeen.

ORCID

Laurie Saint Criq https://orcid.org/0000-0002-0404-7186

REFERENCES

Arns, A., Wahl, T., Haigh, I. D., & Jensen, J. (2015). Determining return water levels at ungauged coastal sites: A case study for northern Germany. *Ocean Dynamics*, *65*(4), 539–554. https://doi.org/10.1007/s10236-015-0814-1

Arns, A., Wahl, T., Wolff, C., Vafeidis, A. T., Haigh, I. D., Woodworth, P., Niehüser, S., & Jensen, J. (2020). Non-linear interaction modulates global extreme sea levels, coastal flood exposure, and impacts. *Nature Communications*, *11*(1), 1918. https://doi.org/10.1038/s41467-020-15752-5

Batstone, C., Lawless, M., Tawn, J., Horsburgh, K., Blackman, D., McMillan, A., Worth, D., Laeger, S., & Hunt, T. (2013). A UK best-practice approach for extreme sea-level analysis along complex topographic coastlines. *Ocean Engineering*, *71*, 28–39. https://doi.org/10.1016/j.oceaneng.2013.02.003

Benito, G., Lang, M., Barriendos, M., Llasat, M. C., Francés, F., Ouarda, T., Thorndycraft, V., Enzel, Y., Bardossy, A., Coeur, D., & Bobée, B. (2004). Use of systematic, palaeoflood and historical data for the improvement of flood risk estimation. Review of scientific methods. *Natural Hazards*, *31*(3), 623–643. https://doi.org/10.1023/B:NHAZ.0000024895.48463.eb

Bernardara, P., Andreewsky, M., & Benoit, M. (2011). Application of regional frequency analysis to the estimation of extreme storm surges. *Journal of Geophysical Research: Oceans*, *116*, C02. https://doi.org/10.1029/2010JC006229

Bulteau, T., Idier, D., Lambert, J., & Garcin, M. (2015). How historical information can improve estimation and prediction of extreme coastal water levels: Application to the Xynthia event at La Rochelle (France). *Natural Hazards and Earth System Sciences*, *15*(6), 1135–1147. https://doi.org/10.5194/nhess-15-1135-2015

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer Science & Business Media.

Dixon, M. J., & Tawn, J. A. (1999). The effect of non-stationarity on extreme sea-level estimation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *48*(2), 135–151. https://doi.org/10.1111/1467-9876.00145

Frau, R., Andreewsky, M., & Bernardara, P. (2018). The use of historical information for regional frequency analysis of extreme skew surge. *Natural Hazards and Earth System Sciences*, *18*(3), 949–962. https://doi.org/10.5194/nhess-18-949-2018

Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., & Viglione, A. (2010). Inclusion of historical information in flood frequency analysis using a Bayesian MCMC technique: A case study for the power dam Orlík, Czech Republic. *Contributions to Geophysics and Geodesy*, *40*, 121–147. https://doi.org/10.2478/v10126-010-0005-5

Gaume, E. (2018). Flood frequency analysis: The Bayesian choice. *Wiley Interdisciplinary Reviews: Water*, *5*(4), e1290. https://doi.org/10.1002/wat2.1290

Giloy, N., Duluc, C.-M., Frau, R., Ferret, Y., Bulteau, T., Mazas, F., & Sauzeau, T. (2018). La base de données TEMPETES: un support pour une expertise collégiale et interdisciplinaire des informations historiques de tempêtes et de submersions. *XVèmes Journées Nationales Génie Côtier—Génie Civil*, 823–832. https://doi.org/10.5150/jngcgc.2018.093

Haigh, I. D., Nicholls, R., & Wells, N. (2010). A comparison of the main methods for estimating probabilities of extreme still water levels. *Coastal Engineering*, *57*(9), 838–849. https://doi.org/10.1016/j.coastaleng.2010.04.002

Haigh, I. D., Ozsoy, O., Wadey, M. P., Nicholls, R. J., Gallop, S. L., Wahl, T., & Brown, J. (2017). *An improved database of coastal flooding in the United Kingdom from 1915 to 2016*. British Oceanographic Data Centre, Natural Environment Research Council.

Haigh, I. D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., Horsburgh, K., & Gouldby, B. (2016). Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK. *Scientific Data*, *3*(1), 160107. https://doi.org/10.1038/sdata.2016.107

Hamdi, Y., Bardet, L., Duluc, C.-M., & Rebour, V. (2015). Use of historical information in extreme-surge frequency estimation: The case of marine flooding on the La Rochelle site in France. *Natural Hazards and Earth System Sciences*, *15*(7), 1515–1531. https://doi.org/10.5194/nhess-15-1515-2015

Kergadallan, X., Bernardara, P., Benoit, M., & Daubord, C. (2014). Improving the estimation of extreme sea levels by a characterization of the dependence of skew surges on high tidal levels. *Coastal Engineering Proceedings*, *1*, 48. https://doi.org/10.9753/icce.v34.management.48

Liu, J. C., Lence, B. J., & Isaacson, M. (2010). Direct joint probability method for estimating extreme sea levels. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, *136*(1), 66–76. https://doi.org/10.1061/(ASCE)0733-950X(2010)136:1(66)

Mazas, F., Kergadallan, X., Garat, P., & Hamm, L. (2014). Applying POT methods to the revised joint probability method for determining extreme sea levels. *Coastal Engineering*, *91*, 140–150. https://doi.org/10.1016/j.coastaleng.2014.05.006

Nguyen, C. C., Gaume, E., & Payrastre, O. (2014). Regional flood frequency analyses involving extraordinary flood events at ungauged sites: Further developments and validations. *Journal of Hydrology*, *508*, 385–396. https://doi.org/10.1016/j.jhydrol.2013.09.058

Ouarda, T. B. M. J., & Ashkar, F. (1998). Effect of trimming on LP III flood quantile estimates. *Journal of Hydrologic Engineering*, *3*(1), 33–42. https://doi.org/10.1061/(ASCE)1084-0699(1998)3:1(33)

Outten, S., Wolf, T., Mangini, F., Chen, L., & Nilsen, J. E. (2020). Reassessing extreme sea level events through interplay of tides and storm surges. EGU General Assembly 2020, EGU2020-8000. https://doi.org/10.5194/egusphere-egu2020-8000

Payrastre, O., Gaume, E., & Andrieu, H. (2011). Usefulness of historical information for flood frequency analyses: Developments based on a case study. *Water Resources Research*, *47*, W08511. https://doi.org/10.1029/2010WR009812

Pugh, D. T., & Vassie, J. M. (1978). Extreme sea levels from tide and surge probability. *Coastal Engineering Proceedings*, *1*(16), 52. https://doi.org/10.9753/icce.v16.52

Reis, D., & Stedinger, J. (2005). Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology*, *313*, 97–116. https://doi.org/10.1016/j.jhydrol.2005.02.028

Saint Criq, L., Gaume, E., Hamdi, Y., & Ouarda, T. B. M. J. (2022). Extreme sea level estimation combining systematic observed skew surges and historical record sea levels. *Water Resources Research*, *58*(3), e2021WR030873. https://doi.org/10.1029/2021WR030873

Schendel, T., & Thongwichian, R. (2017). Considering historical flood events in flood frequency analysis: Is it worth the effort? *Advances in Water Resources*, *105*, 144–153. https://doi.org/10.1016/j.advwatres.2017.05.002

Stephenson. (2015). *TideHarmonics: Harmonic analysis of tides*. R package (software).

Sterl, A., van den Brink, H., de Vries, H., Haarsma, R., & van Meijgaard, E. (2009). An ensemble study of extreme storm surge related water levels in the North Sea in a changing climate. *Ocean Science*, *5*(3), 369–378. https://doi.org/10.5194/os-5-369-2009

Tawn, J., Vassie, J., & Gumbel, E. (1989). Extreme sea levels; the joint probabilities method revisited and revised. *Proceedings of the Institution of Civil Engineers*, *87*(3), 429–442. https://doi.org/10.1680/iicep.1989.2975

Tawn, J. A. (1992). Estimating probabilities of extreme sea-levels. *Applied Statistics*, *41*(1), 77. https://doi.org/10.2307/2347619

Tomasin, A., & Pirazzoli, P. A. (2008). Extreme sea levels in the English Channel: Calibration of the joint probability method. *Journal of Coastal Research*, *4*, 1–13. https://doi.org/10.2112/07-0826.1

Wahl, T., & Chambers, D. P. (2015). Evidence for multidecadal variability in US extreme sea level records. *Journal of Geophysical Research: Oceans*, *120*(3), 1527–1544. https://doi.org/10.1002/2014JC010443

Williams, J., Horsburgh, K. J., Williams, J. A., & Proctor, R. N. F. (2016). Tide and skew surge independence: New insights for flood risk. *Geophysical Research Letters*, *43*(12), 6410–6417. https://doi.org/10.1002/2016GL069522

Woodworth, P. L., Flather, R. A., Williams, J. A., Wakelin, S. L., & Jevrejeva, S. (2007). The dependence of UK extreme sea levels and storm surges on the North Atlantic Oscillation. *Continental Shelf Research*, *27*(7), 935–946. https://doi.org/10.1016/j.csr.2006.12.007

How to cite this article: Saint Criq, L., Gaume, E., Hamdi, Y., & Ouarda, T. B. M. J. (2025). Extreme skew surge estimation combining systematic skew surges and historical record sea levels on the English Channel and North Sea coasts. *Journal of Flood Risk Management*, *18*(1), e12862. https://doi.org/10.1111/jfr3.12862

CIWEM Chartered Institution of Water and Environmental Management Journal of Flood Risk Management—WILEY

APPENDIX A: INDEPENDENCE BETWEEN SKEW SURGES AND ASTRONOMICAL HIGH TIDES

The independence between skew surges and astronomical high tides is verified, as in (Arns et al., 2020) through the scatter plot of predicted astronomical high tides versus the corresponding skew surges and a Kendall test (see Figure A1) If the Kendall's value τ is equal to 0, there is no correlation between astronomical high tides and skew surges. On the other hand, if $|\tau| = 1$, astronomical high tides and skew surges are perfectly dependent.

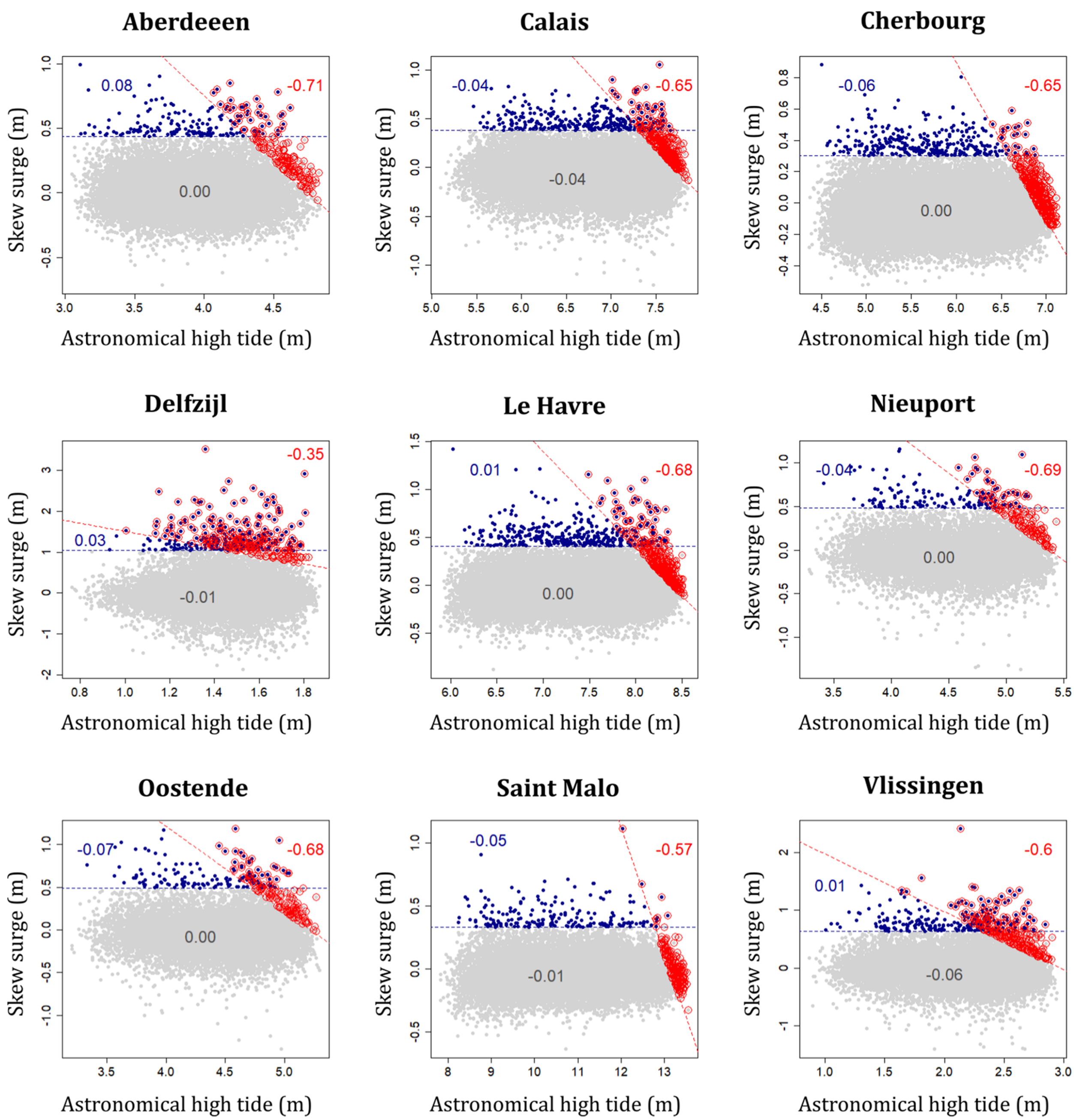

FIGURE A1 Scatter plot of predicted astronomical high tides versus observed skew surges and the Kendall's τ values. The gray points represent all the maximum sea levels, the red points represent the 1% largest maximum sea levels and the blue points represent the maximum sea levels associated with the 1% largest skew surges. The colored numbers are the Kendall's τ values.

APPENDIX B: HISTORICAL INFORMATION

TABLE B1 Record historical maximum sea levels with their corresponding skew surges

Site	Date	Maximum sea levels (m)	Skew surges (m)
Cherbourg	13-09-1821	7.96	1.28
Cherbourg	01-01-1877	7.17	0.71
Saint Malo	01-01-1877	12.96	0.43
Le Havre	27-10-1882	8.42	0.30
Le Havre	22-01-1890	8.55	0.55
Le Havre	06-12-1896	8.40	0.54
Calais	07-01-1905	7.70	0.56
Vlissingen	12-03-1906	4.05	1.49
Vlissingen	31-12-1910	4.33	2.37
Delfzijl	13-01-1916	4.34	3.34
Oostende	23-11-1930	6.19	1.64
Oostende	26-04-1944	6.01	1.81
Oostende	01-03-1949	6.04	1.57
Nieuport	01-02-1953	6.73	2.17
Oostende	01-02-1953	6.66	2.22
Vlissingen	01-02-1953	4.54	2.41
Nieuport	21-03-1961	6.10	1.39
Aberdeen	28-02-1967	5.10	0.49
Aberdeen	29-09-1969	5.10	0.64
Le Havre	23-11-1984	9.28	1.18
Nieuport	15-11-1993	6.14	0.95
Nieuport	28-01-1994	6.08	1.18
Nieuport	02-01-1995	6.08	1.13