Supplementary Materials for

Observations reveal changing coastal storms around the United States

Joao Morim[1,2]*†, Thomas Wahl[1,2], DJ Rasmussen[3], Francisco M. Calafat[4,5], Sean Vitousek[6], Soenke Dangendorf[7], Robert E. Kopp[8], Michael Oppenheimer[9]

Corresponding author. Email: *jmorimnascimento@ucf.edu

Affiliations:

[1]Deparument of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, Florida, USA
[2]National Center for Integrated Coastal Research, University of Central Florida, Orlando, Florida, USA
[3]Princeton School of Public and International Affairs, Princeton University, Princeton, NJ, USA
[4]Marine Physics and Ocean Climate, National Oceanography Centre, Liverpool, UK
[5]Department of Physics, University of the Balearic Islands, Palma, Spain.
[6]U.S. Geological Survey (USGS), Pacific Coastal and Marine Science Center, USA
[7]Tulane University, School of Science & Engineering, Department of River-Coastal Science and Engineering, New Orleans, LA, USA.
[8]Department of Earth and Planetary Sciences and Rutgers Climate & Energy Institute, Rutgers University, New Brunswick, NJ, USA
[9]Department of Geosciences and Princeton School of Public and International Affairs and High Meadows Environmental Institute, Princeton University, Princeton, NJ, USA

The PDF file includes:

Supplementary Information
Figs. S1 to S15

Supplementary Information

1. BAYEX parameter layer and priors

Here, we use a full Bayesian approach (BAYEX), and hence, all model parameters are estimated from observations. This requires that we ascribe prior distributions to all model parameters. A full description of all BAYEX model parameters has been provided elsewhere[27,28], and hence we defer to such references for extended details. Before selecting any model priors, a sensitivity analysis of BAYEX to prior choices was conducted for each subregion/model domain by comparing estimated GEV model parameters (location, scale, shape) based on different length scales (ρ_μ, $\rho_{\mu 0}$, $\rho_{\mu 00}$, ρ_σ) and standard deviations (γ_μ, $\gamma_{\mu 0}$, $\gamma_{\mu 00}$, γ_σ) of the Gaussian processes following previous work[27,28]: half-N(0,1), half-N(0,2), half-N(0,10). These hyperparameters that define the covariance functions of all Gaussian processes[27] describe the spatial scales and temporal variability of the extreme storm surge field, respectively. The mean differences between estimated GEV parameters across all tide gauge locations with respect to half-N(0,1) are found to be relatively small (< 5%), consistent with previous work[27,28]. We also compared different prior choices for BAYEX regression coefficients $\beta_{u,width}$ and $\beta_{\sigma,width}$: N(0,1.5), N(0,3) and N(0,10). These hyperparameters denote the location and scale coefficients of the shelf width covariates used to inform BAYEX[27,28]. The comparison shows that estimated GEV parameters are fairly consistent for all other prior choices for all domain. The range and variability found from our sensitive analysis for all parameters considered are consistent with previous work[278,28]. The following priors are ascribed as BAYEX parameters and rescaled:

− The parameter α (controls relative contributions of small-scale errors) is bounded to (0,1), and hence, we let α ~ U(0,1).

− The length scale of kernel functions (τ) (used to describe spatial residual dependence) is assigned a half-normal distribution: τ ~ half-N(0,0.5). A standard deviation of 0.5 corresponds to half of a synoptic scale (~500 km), which is a general measure of spatial extents associated with extra-tropical and tropical cyclone storms. Note that estimates using τ ~ half-N(0,0.5), half-N(0,1) and half- N(0,0.25) leads to very similar results.

− The standard deviations of the model Gaussian processes (Equation 4) are assumed to follow a half-normal distribution: γ_μ, $\gamma_{\mu 0}$, $\gamma_{\mu 00}$ and γ_σ ~ half-N(0,1) and γ_ξ ~ half-N(0,0.3). A half-normal distribution is one of the recommended priors for scale parameters in hierarchical models[27] since it allows us to constrain the value of a parameter from above while allowing it to be (arbitrarily) close to zero (https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations).

− The length scale parameters of the model Gaussian processes (Equation 4) are assumed to follow a half-normal distribution: ρ_μ, $\rho_{\mu 0}$, $\rho_{\mu 00}$. ρ_σ ~ half-N(0,0.5). It should be noted that there is no

information within the observations to characterize scales above the maximum distance between tide gauge sites, hence we impose a half-normal distribution.

− The length scale of ξ (ρ_ξ) is assumed to follow an inverse gamma distribution ρ_ξ ~ inv-gamma (3,1). The reason for adopting an inverse gamma distribution is to prevent ρ_ξ from defaulting to unrealistically small values, which it tends to do with less informative priors.

− The spatial mean $\bar{\xi}$ is assigned a normal distribution ~ N(0,0.6) based on lower and upper bounds have been selected based on results of individual GEV fits to TG annual maxima for each model domain (Table S1). For Alaska, a normal distribution ~ N(-0.2,0.2) is used based on individual GEV fits to TG annual maxima.

− The regression coefficients, β_u and β_σ, are assumed to follow a normal distribution: (β_u, β_σ): ~ N(0,3).

For a full explanation about BAYEX prescribed model priors, including those used for length scale parameters and regression coefficients, see refs. (*27*, 28). Note that prior distributions are basically probability distributions on model parameters that specify prior beliefs on how 'plausible' a given parameter is before considering any observational data. The posterior distribution, as from Bayes' rule, reflects a balance between a particular prior distribution and likelihood. If our prescribed data are sufficiently informative, posterior values will be largely governed by likelihood estimates.

Figures

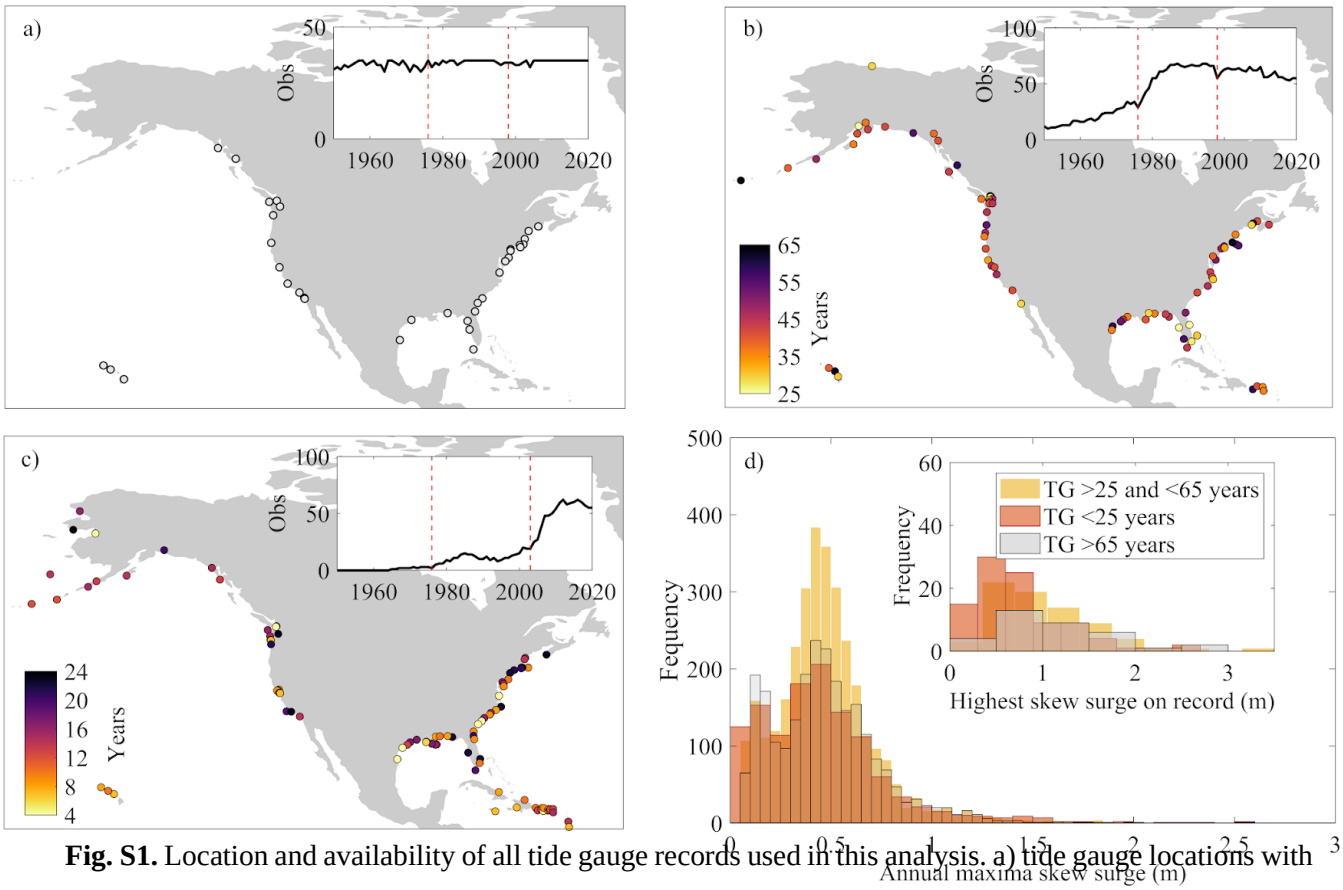


Fig. S1. Location and availability of all tide gauge records used in this analysis. a) tide gauge locations with historical record lengths of at least 65 years which have been traditionally used in at-site and spatial models. b-c) show tide gauge locations with record lengths between 25-65 years and less than 25 years, respectively. The total number of observed annual maxima skew surges per year between 1950-2020 for each sub-panel is presented. Subpanel d) shows histograms of annual maxima skew surges (m) and largest annual maxima values (m) sampled from tide gauges with records of at least 65 years (a), between 25 and 65 years (b), and less than 25 years (c) as per legend.

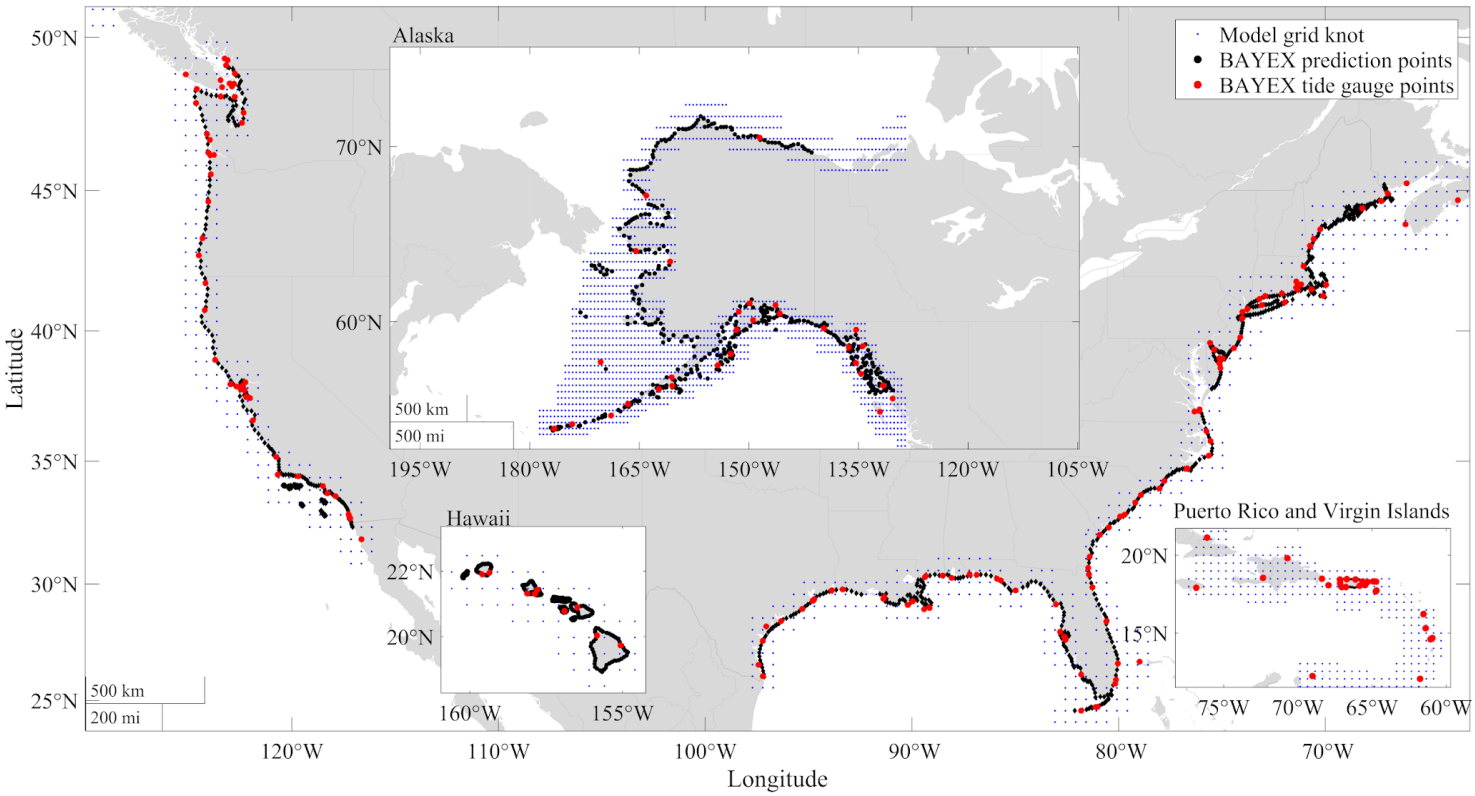


Fig. S2. The combined U.S. domain modeled using BAYEX (see Methods; Table 1). The red solid circles denote tide gauge stations, blue dots denote model grid knots (0.5 degree spacing) and black points denote model prediction points as per legend.

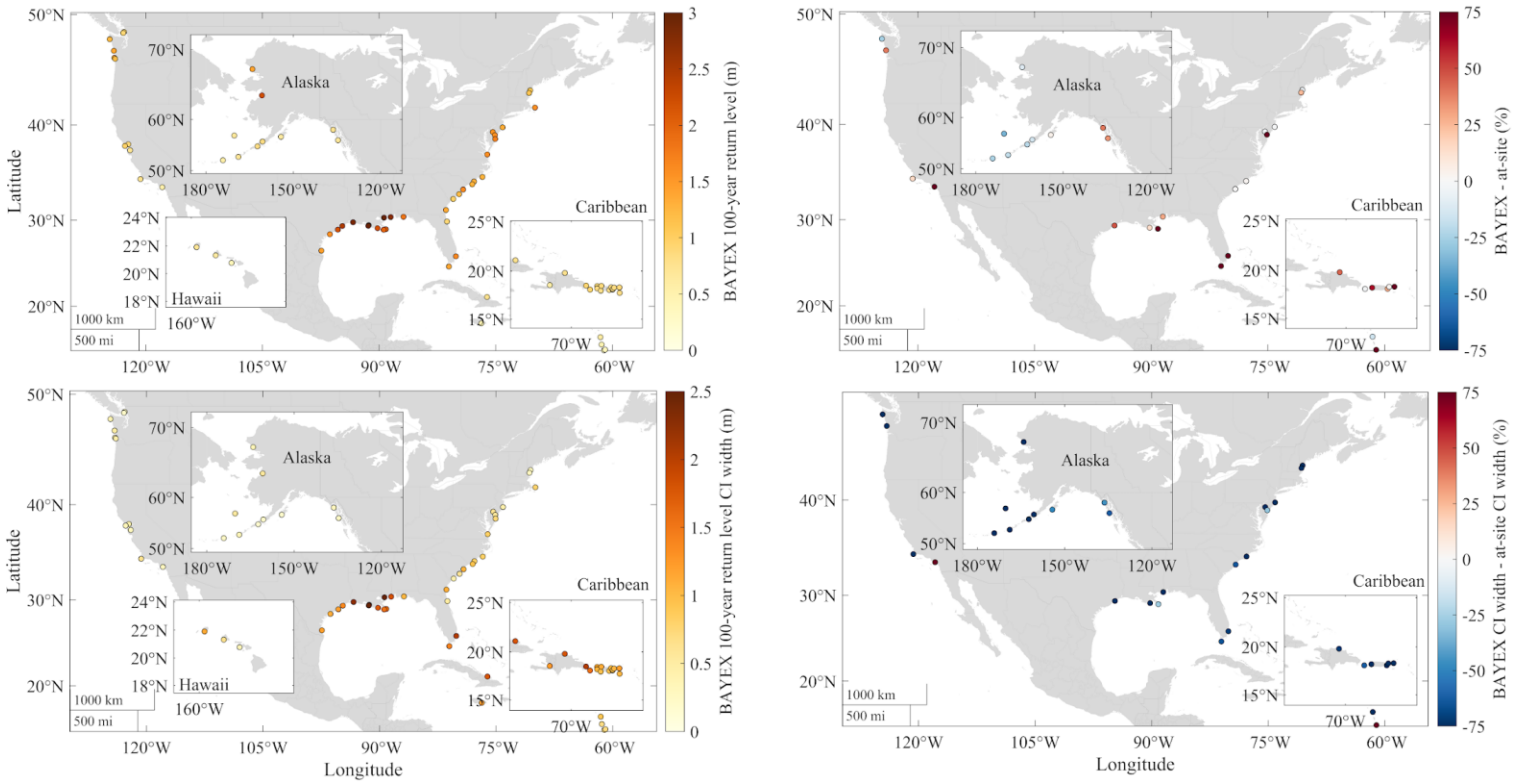


Fig. S3. Same as Fig. 1 but for tide gauge locations with records of less than 25 years. Note that some tide gauge sites are not shown within panels b) and d) since at-site GEV model estimates did not converge. Note that μ is time-varying and so we use μ from 2020 as reference.

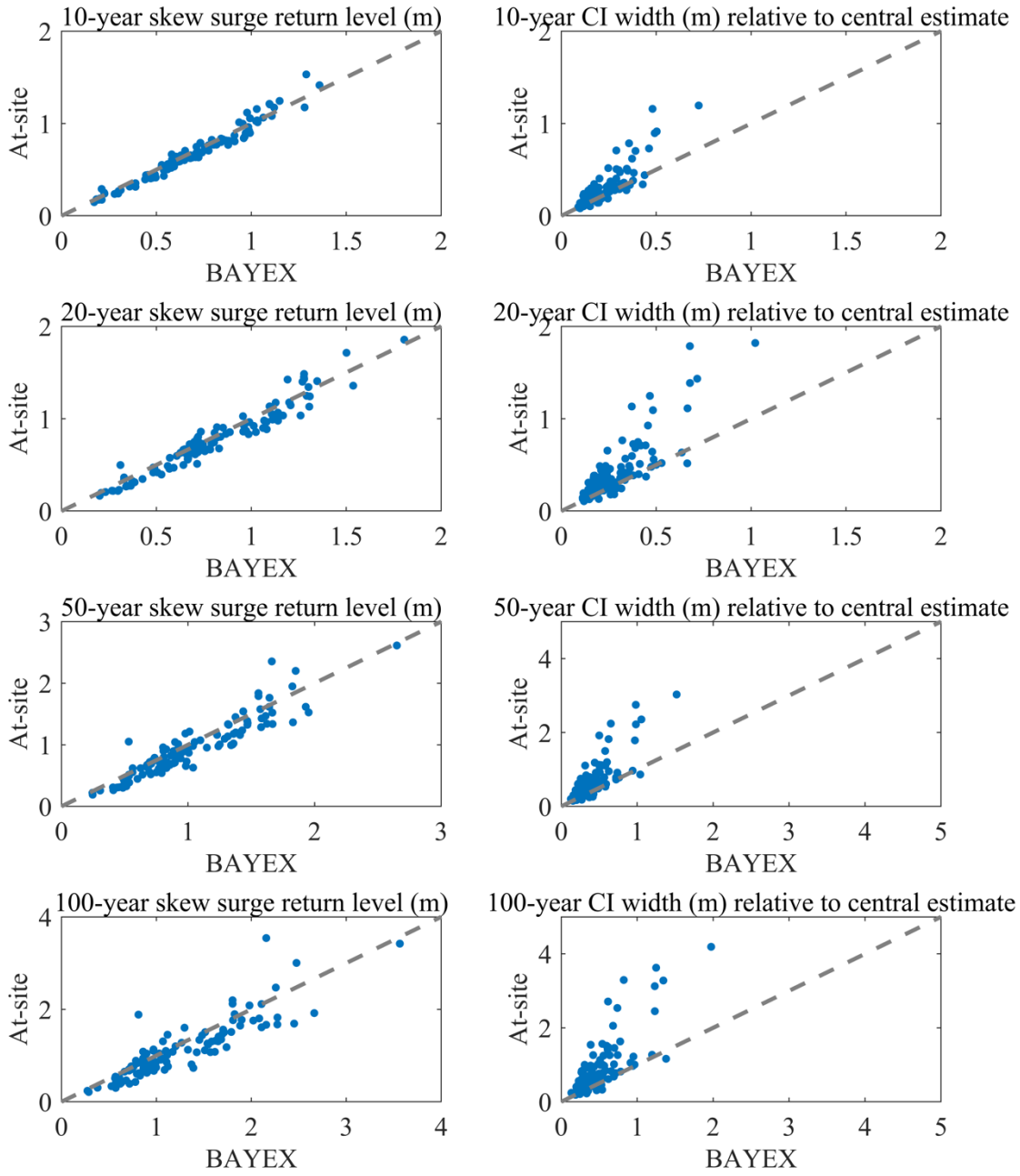
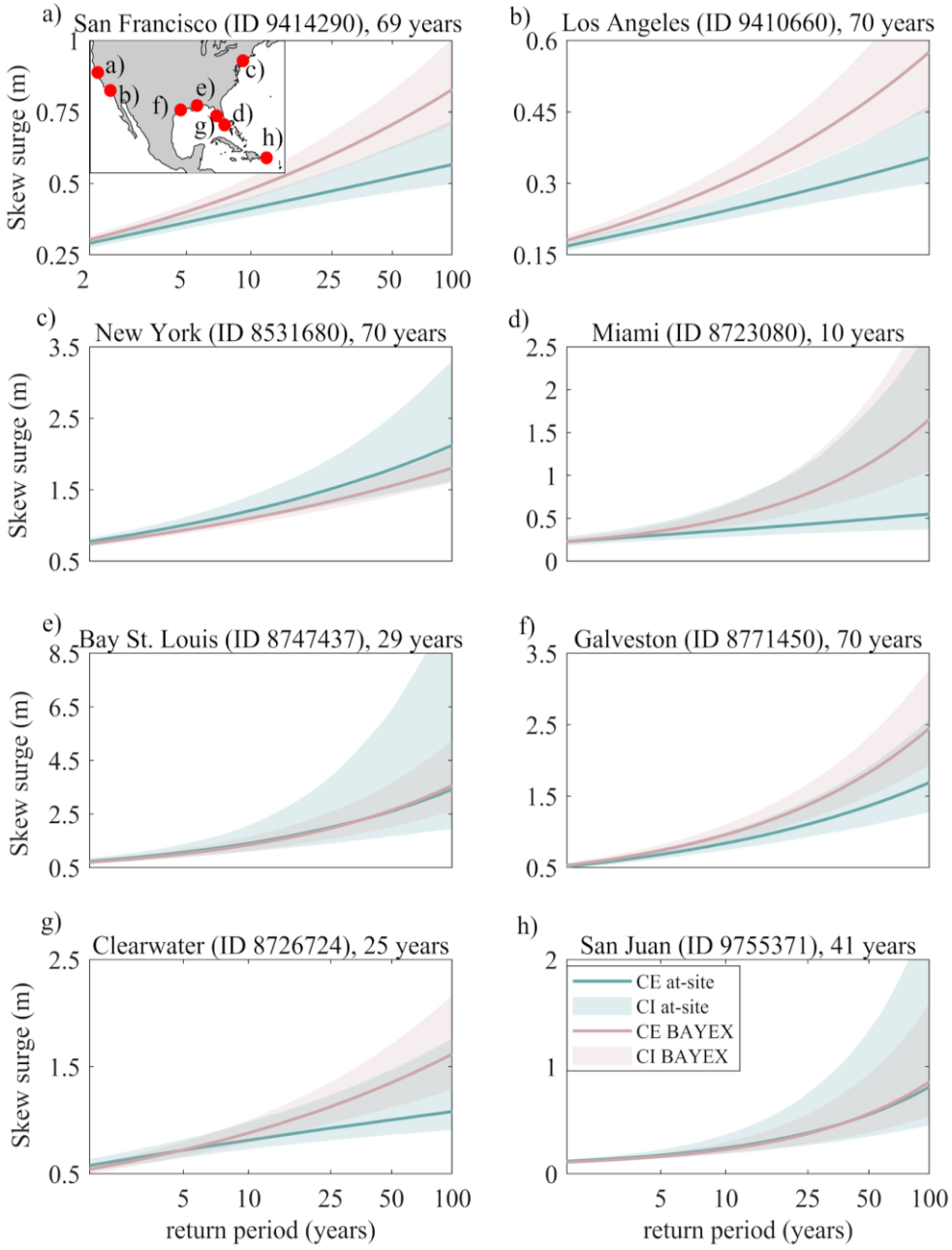


Fig. S4. Comparison of storm surge return levels estimated using BAYEX and at-site analysis for coastal tide gauge stations near major coastal cities. Solid lines represent central-median estimates (CE) and shaded bands represent 90% credible intervals (CI) as per legend. The ID code for each tide gauge station and their record lengths are shown in sub-panel titles. The locations of all tide gauges are also shown (see map inset in panel (a)). Note that μ is time-varying and so we use μ from 2020 as reference.

Fig. S4. Scatter plots of BAYEX vs at-site differences for different return levels. The left column presents



differences for 10-, 20-, 50- and 100-year return levels, respectively. The right column presents differences for their respective credible interval (CI) widths (relative to central estimates). Note that μ is time-varying and so we use μ from 2020 as reference.

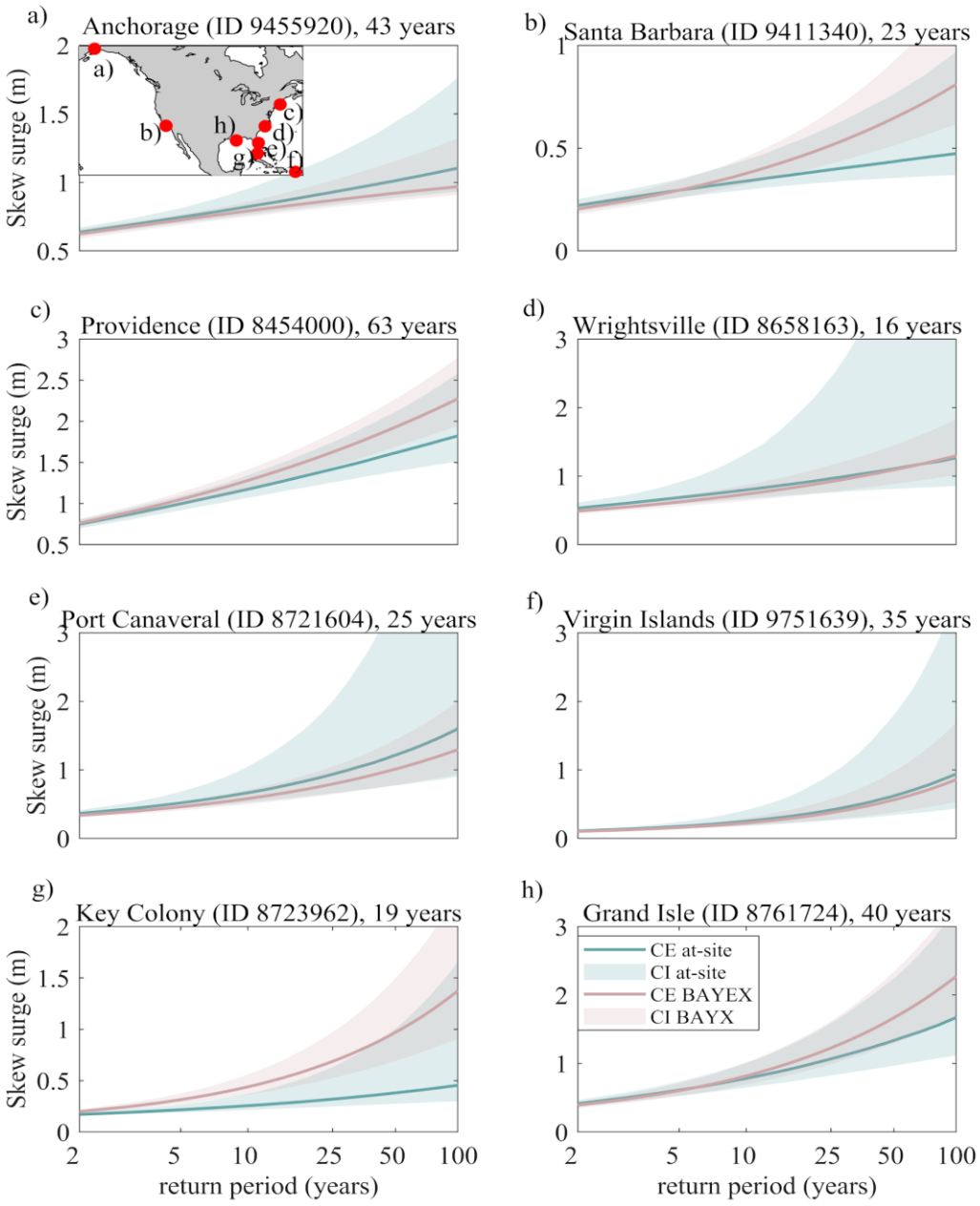


Fig. S6. Same as Fig. S5.

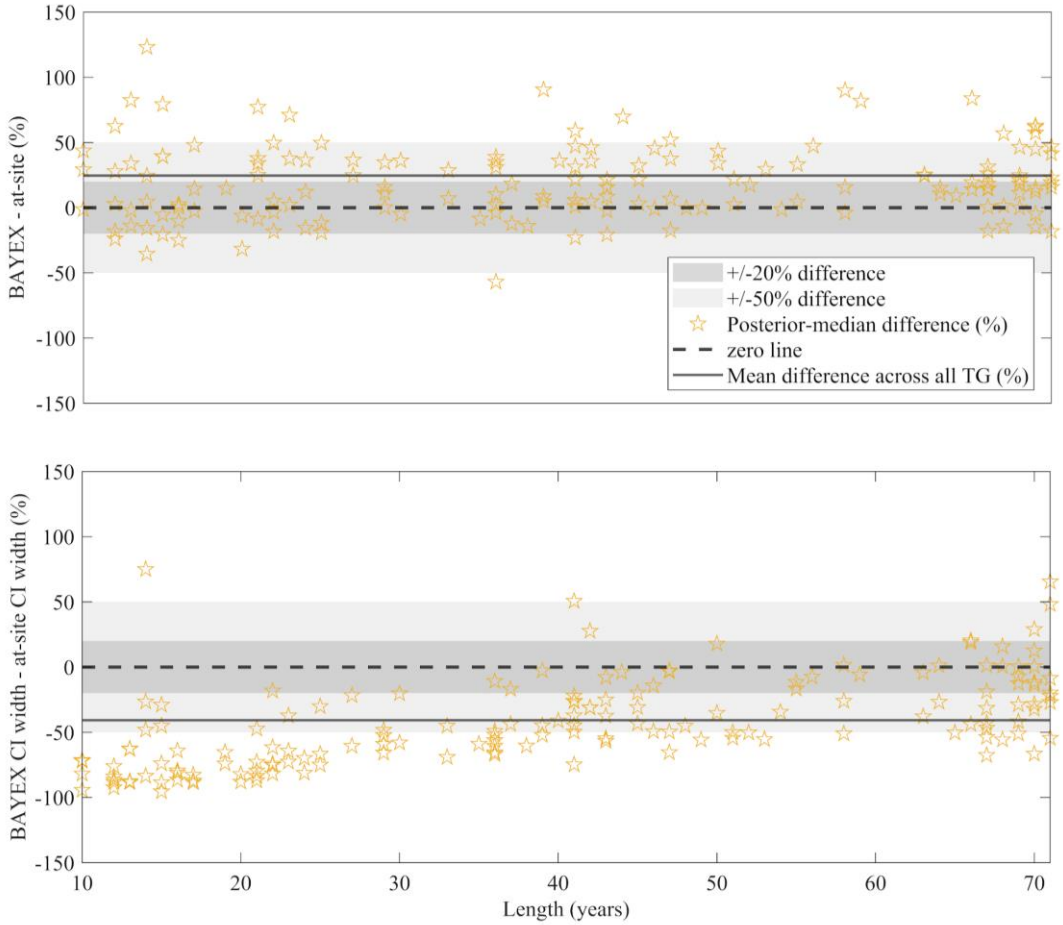


Fig. S7. Impact of tide gauge record length on differences between estimates of 100-year storm surge return levels (top panel) and associated 90% credible intervals (CI) (bottom panel) as obtained from BAYEX and at-site analysis.

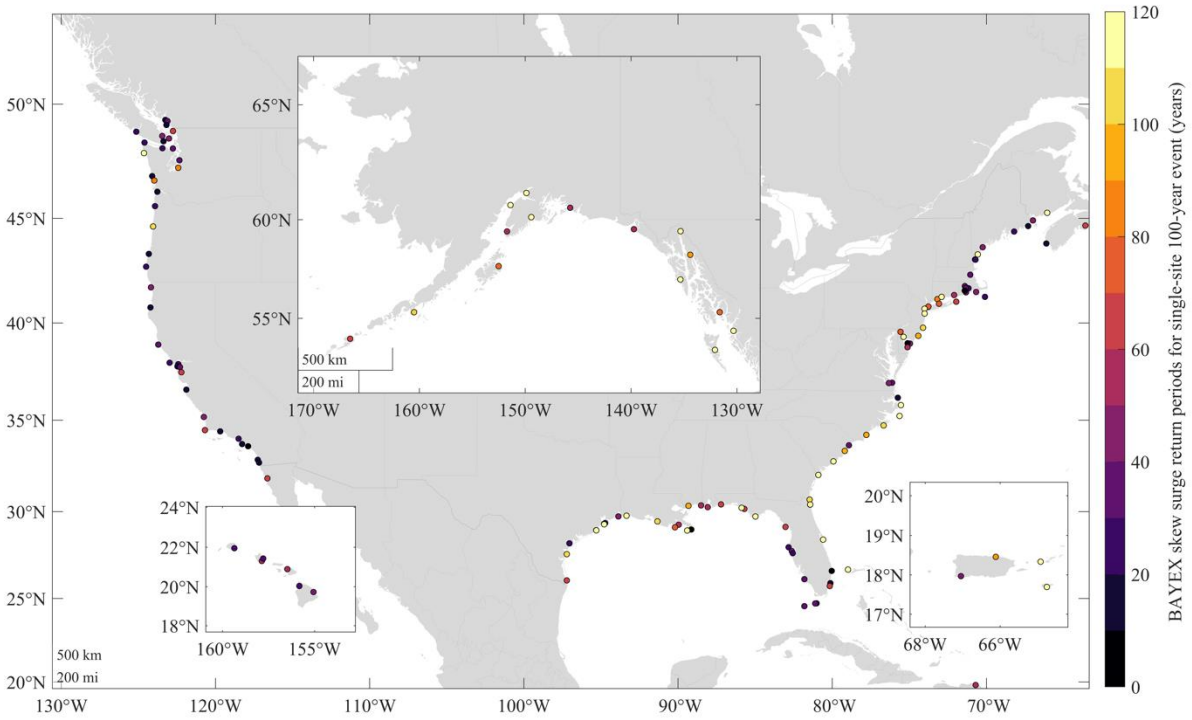


Fig. S8. Skew surge return periods (years) from BAYEX corresponding to 100-year return levels calculated based on at-site analysis for tide gauge locations with records of less than 25 years. The median across tide gauge sites is 47-year return period. Note that μ is time-varying and so we use μ from 2020 as reference.

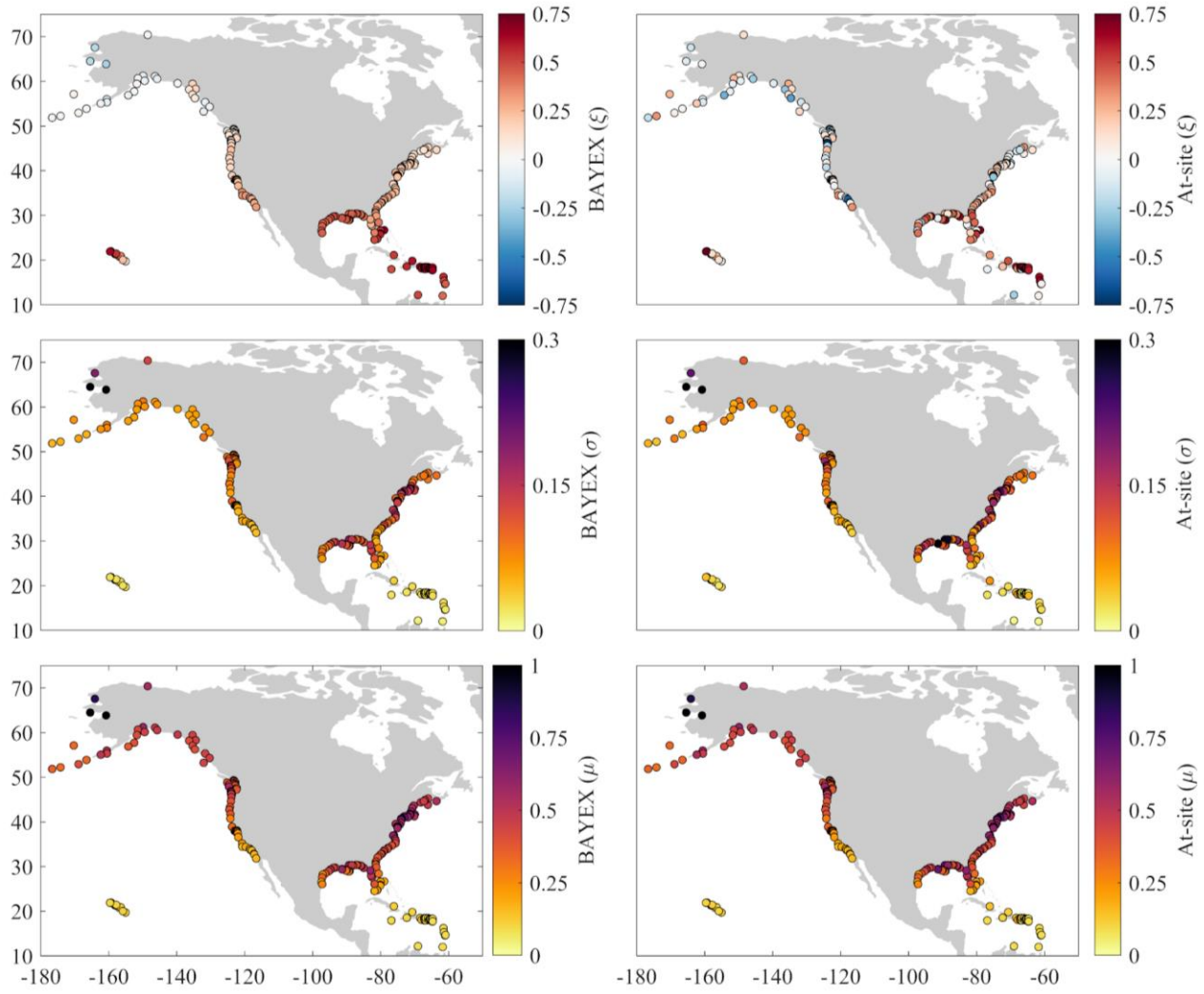


Fig. S9. Comparison of GEV model parameters from BAYEX (left) and at-site analysis for all tide gauges with records of at least 10 years. The left panels show BAYEX posterior-median values for GEV shape (ξ), scale (σ), and location (μ) parameters as per legend; right panels show results based on at-site analysis. Note that μ is time-varying and so we use μ from 2020 as reference.

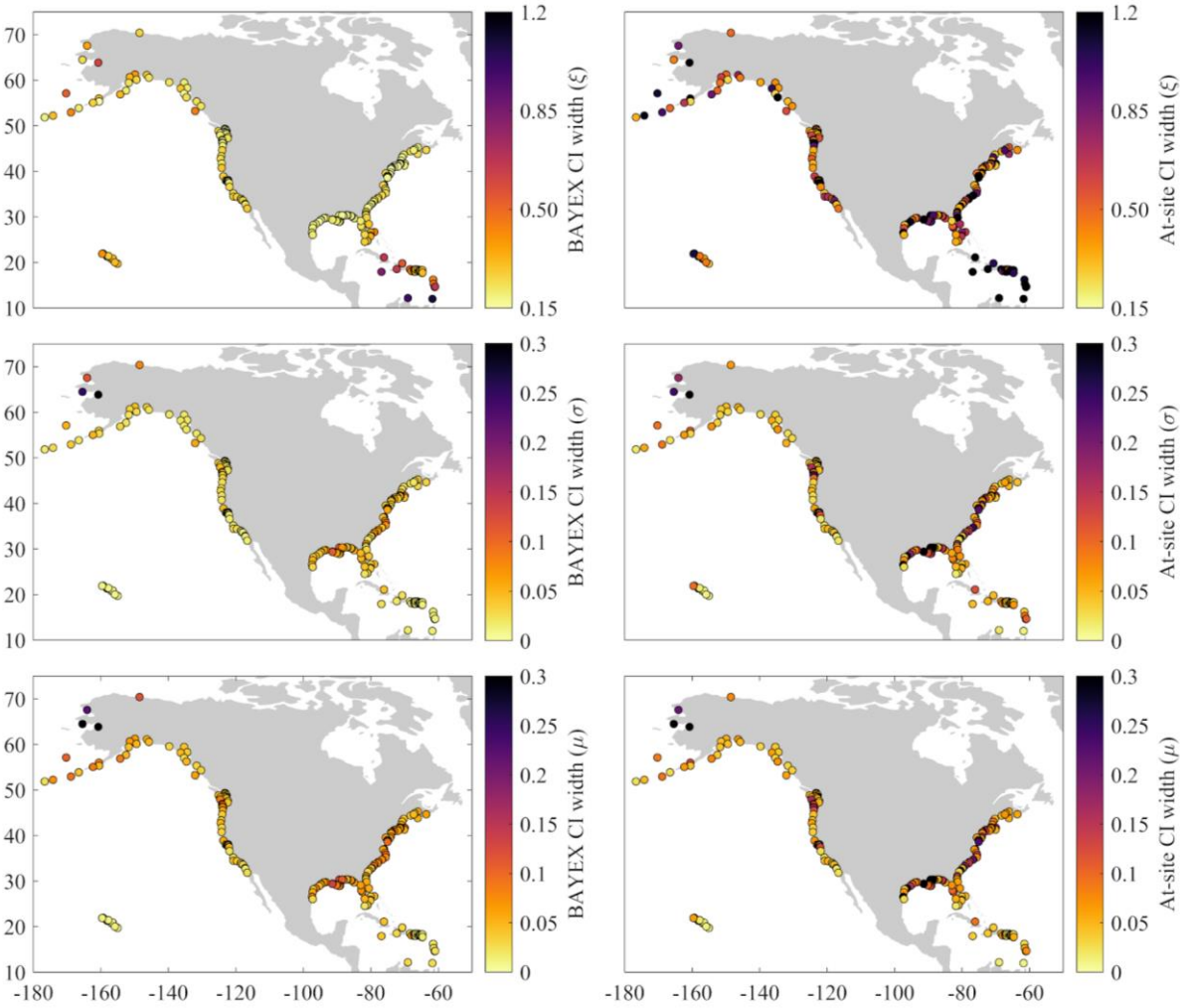


Fig. S10. Comparison of 90% CI width from BAYEX at-site analysis at tide gauge stations with records of at least 10 years. The left panels show BAYEX 90% CI width for GEV shape (ξ), scale (σ) and location (μ) parameters as per legend and right panels show results based on at-site analysis. Note that μ is time-varying and so we use μ from 2020 as reference.

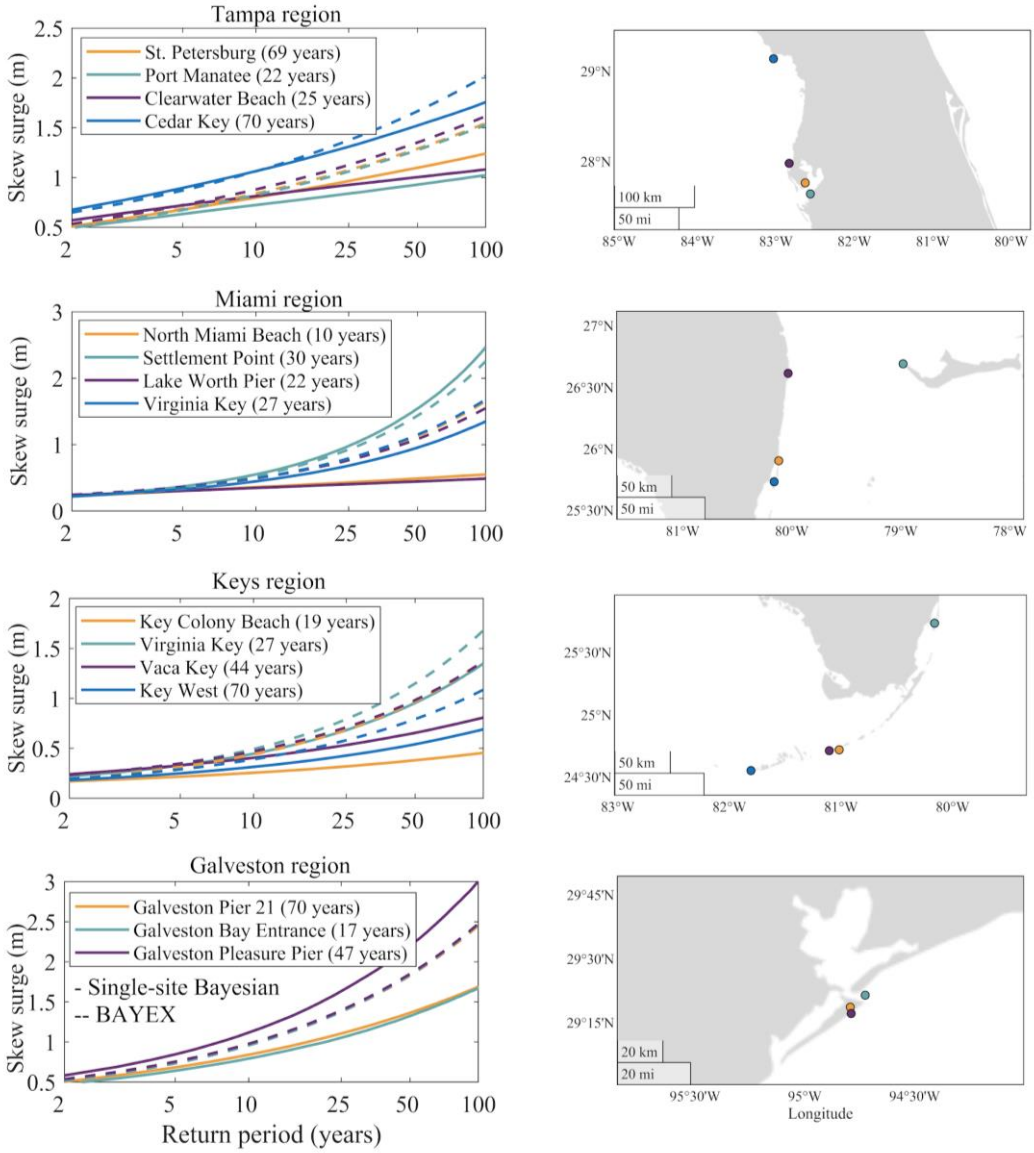


Fig. S11. Example of comparison of surge return levels from BAYEX and at-site analysis for different U.S. coastal regions (all prone to tropical cyclones) with multiple nearby tide gauges with different temporal coverage (see legend for number of years of data available from each tide gauge station). Left panels show, for each region, BAYEX posterior-median estimates at each tide gauge station (dashed lines) and respective at-site estimates (solid lines) as per legend. Right panels show geographic locations of the tide gauges (note the scales to show distances). For visualization, we only display central estimates with uncertainty provided in Fig. S6-S7. Note that μ is time-varying and so we use μ from 2020 as reference.

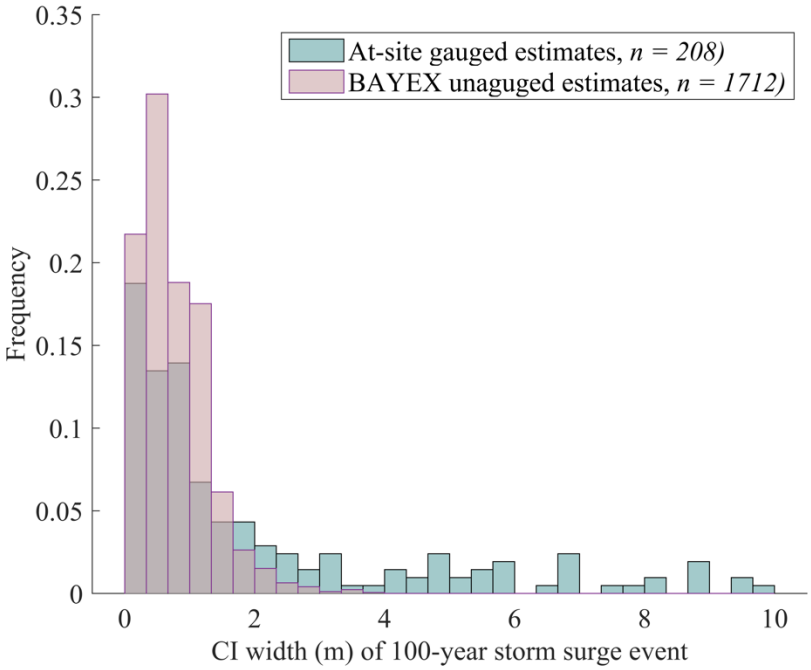


Fig. S12. Credible interval (CI) widths associated with 100-year storm surge events calculated for tide gauge sites (n = 208) using a traditional GEV model based on Bayesian inference and at ungauged locations using BAYEX (n = 1712).

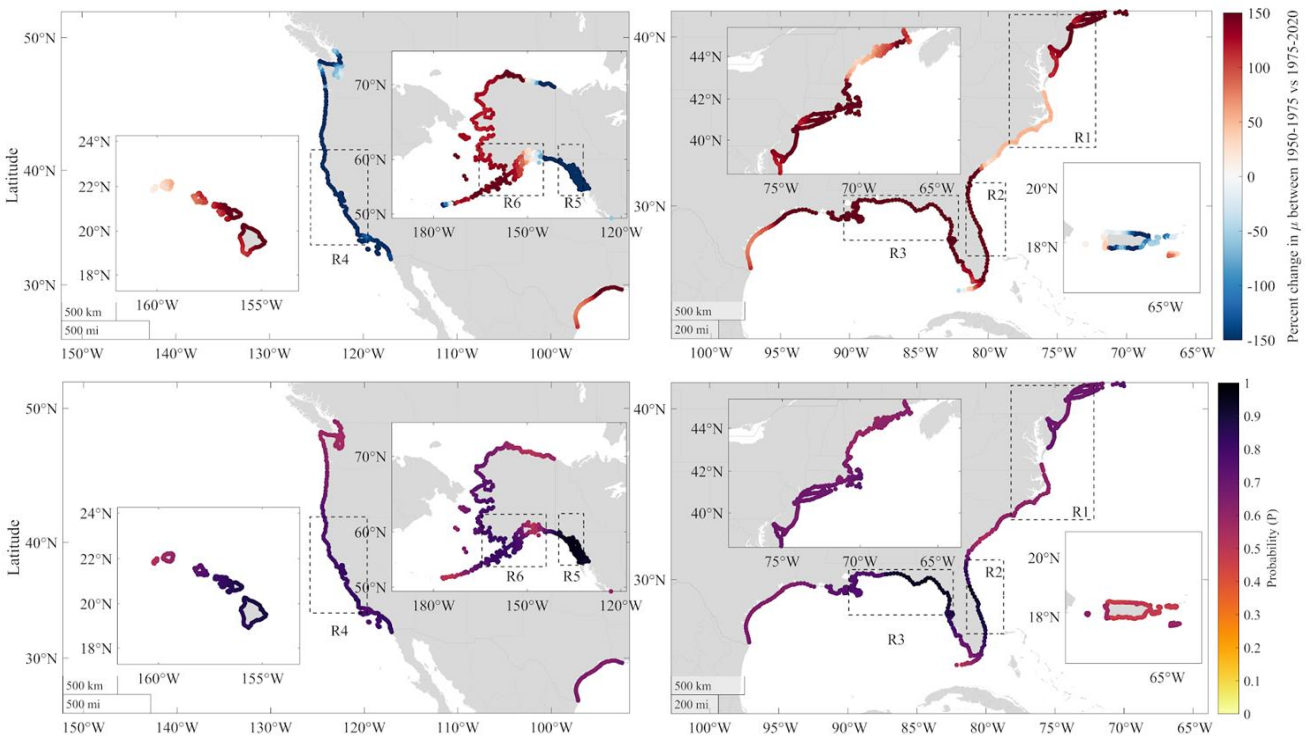


Fig. S13. Percent changes (%) in BAYEX posterior-median trends in μ (mm/year) between 1950-2020 and 1975-2020 (top panels) and associated probability (P) of μ trends during 1975-2020 being larger than 1950-2020 trends (bottom panels). The regions (R1-R6) delimited using dashed lines (--) denote hotspots of positive and negative trends in μ between 1950-2020 (Fig. 3). Note that blue color represents a percentage decrease in absolute values and red color represents a percentage increase in absolute values. The changes (%) are calculated as: $100 \times (\mu_{1975\text{-}2020} - \mu_{1950\text{-}1975}) / |\mu_{1950\text{-}1975}|$).

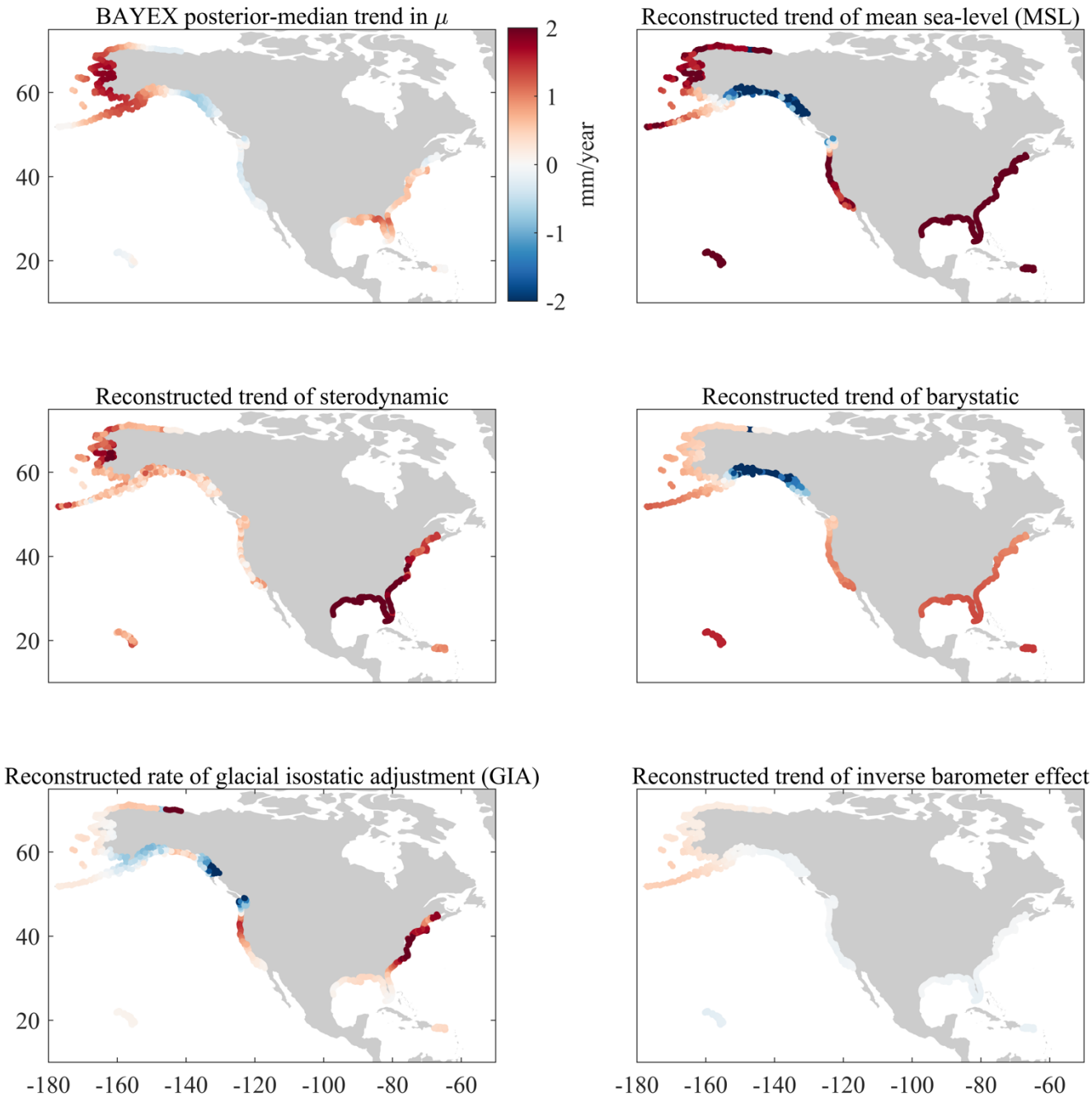


Fig. S14. Trends in μ (mm/year) from BAYEX between 1975-2020 alongside with estimated linear trends in annual mean sea level (MSL) and its main contributors between 1975-2020 (see Methods for details on the MSL reconstruction).

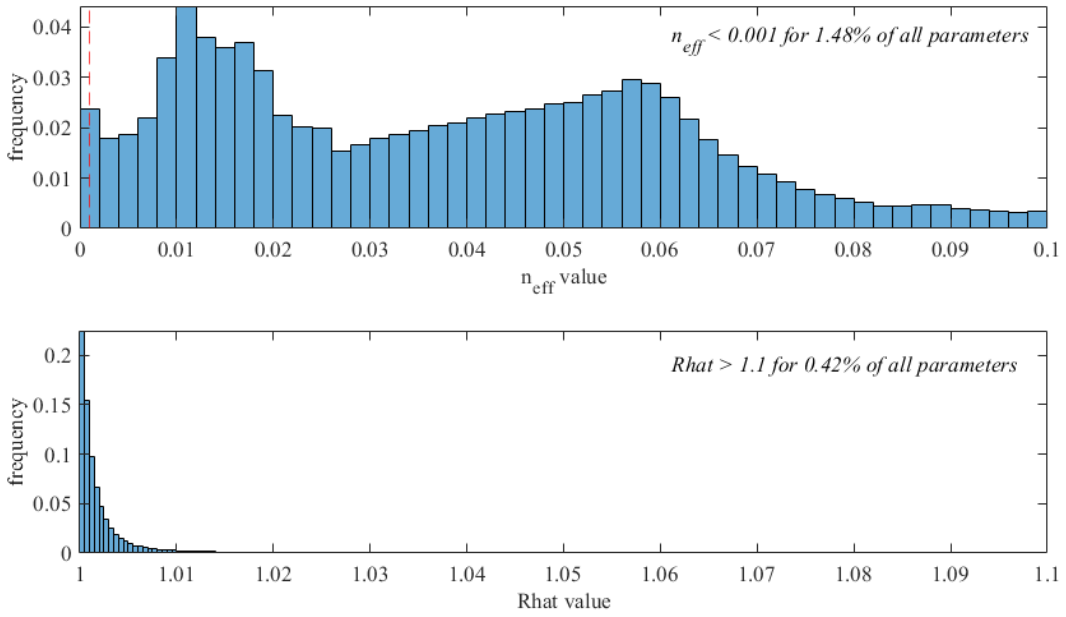


Fig. S15. MCMC diagnostics for BAYEX. Top panel shows effective sample size (n_{eff}) which is a measure of autocorrelation with values less than 0.001 indicating poor mixing chains and biased estimates. Bottom panel shows potential scale reduction statistic ($\hat{R}$) which compares sample variances both within individual chains and across multiple randomly initialized chains. At convergence, $\hat{R}$ must be close to 1 for all BAYEX parameters, whereas $\hat{R}$ >1.1 is indicative of non-convergence.

Tables

Table. S1. BAYEX regional domains used.

BAYEX domain	Region and coordinates	Number of tide gauges in region
Alaska	50.3630-71.8630 N; -129.6320 to -153.1320 W	29
Hawaiian Islands	18.4750-22.4750 N; -156.0920 to -159.0920 W	10
U.S. West Coast	30.8500- 49.8500 N; -117.1360 to -123.1360 W	46
U.S. Gulf Coast	25.5610-31.0610 N; -97.7160 to -86.2160 W	28
U.S. Southeast Coast	24.0510-32.0510; N; -83.0750 to -80.5750 W	24
U.S. East Coast	29.7320- 37.2320 N; -81.4660 to -75.4660 W	21
U.S. Northeast	36.46700-45.9670 N; -76.6130 to -62.6130 W	38
U.S. Puerto Rico/Virgin	10.00500-21.5050 N; -65.8500 to -74.8500 W	27
		208*

*total number of tide gauge sites after excluding common sites between adjacent domains.