



OPEN

A database of global storm surge reconstructions

DATA DESCRIPTOR

Michael Getachew Tadesse  & Thomas Wahl

Storm surges are among the deadliest coastal hazards and understanding how they have been affected by climate change and variability in the past is crucial to prepare for the future. However, tide gauge records are often too short to assess trends and perform robust statistical analyses. Here we use a data-driven modeling framework to simulate daily maximum surge values at 882 tide gauge locations across the globe. We use five different atmospheric reanalysis products for the storm surge reconstruction, the longest one going as far back as 1836. The data that we generate can be used, for example, for long-term trend analyses of the storm surge climate and identification of regions where changes in the intensity and/or frequency of storms surges have occurred in the past. It also provides a better basis for robust extreme value analysis, especially for tide gauges where observational records are short. The data are made available for public use through an interactive web-map as well as a public data repository.

Background & Summary

Understanding the stochastic nature of extreme storm surges and how they are affected by climate change and variability is important for efficient design of coastal defense structures and planning of future coastal adaptation¹. A major hurdle in assessing long-term trends in the storm surge climate and performing robust statistical analyses is the lack of sufficiently long data records^{2,3}.

Taking extreme value analysis as an example, the rule of thumb is that extrapolation of extreme water level events should be limited to return periods not longer than four times the available record length⁴. This means in order to derive an estimate of the 100-year storm surge level, at least 25 years of data are needed (to derive a 500-year event, 125 years of data would be needed). However, tide gauge records, which are the main data source for storm surge information, often cover much shorter time periods. From the global dataset of extreme sea level observations we used, 45% of the 882 tide gauges included in this study have less than 25 years of data. Moreover, analyzing long-term trends and multi-decadal variability in the storm surge climate is also hampered by short record lengths. To assess the long-term trend in global mean sea level it has been estimated, for example, that a minimum of 60 years of data are required⁵, which is the case for only 12% of the tide gauges analyzed here; given the larger variability, even longer records are needed to reliably explore long-term changes in the storm surge climate.

Hydrodynamic storm surge models have recently been developed and applied to simulate storm surges globally for the past and future, but the computational burden of running these models is high, limiting the possibility to create very long hindcasts (currently simulations only go back to 1979)^{6,7}. In light of these challenges and to overcome this limitation of large-scale hydrodynamic models, data-driven statistical approaches have been used to develop long-term storm surge reconstructions (as far back as 1866) at the regional⁸ and global scale⁹. The statistical models are trained with predictors from climate reanalysis and observed water levels at tide gauge locations. Both paradigms, data-driven methods as well as hydrodynamic models, have their own advantages and disadvantages (see Tadesse *et al.*¹⁰ for more information).

We present here the Global Storm Surge Reconstruction (GSSR) database making use of already developed data-driven models¹⁰ and multiple satellite-era as well as longer-term atmospheric reanalysis products, going as far back as 1836. The long-term surge reconstruction we present here can be used for more robust extreme value analysis, especially in locations where observational records are short, as well as to better understand the trends and longer-term variations in the storm surge climate (e.g., intensity and frequency of storm surge events) from the mid-nineteenth century until present (please see Usage Notes).

Civil, Environmental, and Construction Engineering & National Center for Integrated Coastal Research, University of Central Florida, Orlando, USA. ✉e-mail: michaeltadesse@knights.ucf.edu

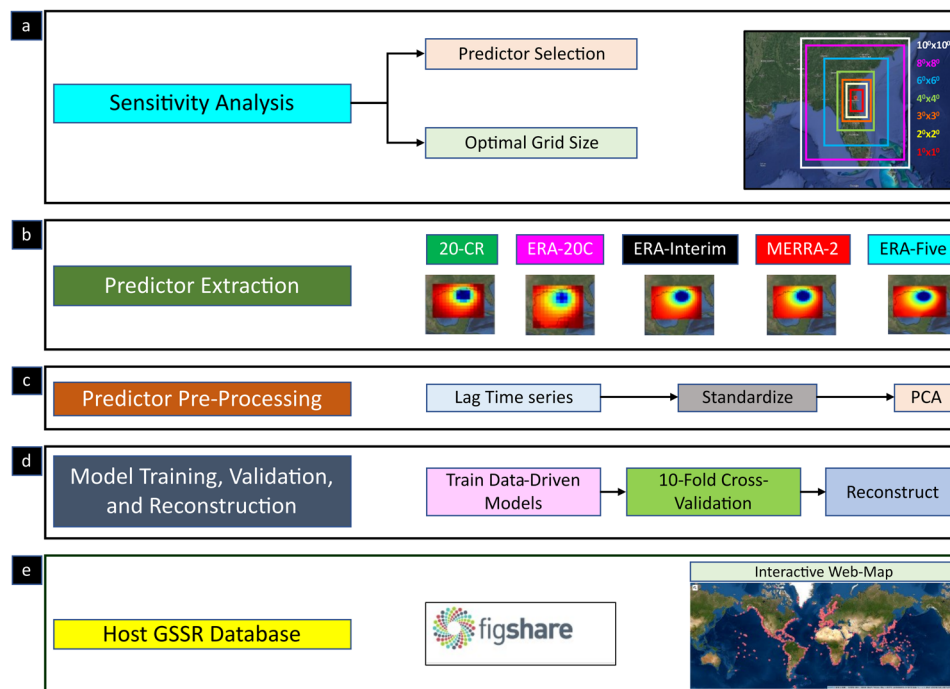


Fig. 1 Schematic diagram of the overall workflow implemented to produce the GSSR database. The different steps include sensitivity analysis (a), predictor extraction (b), predictor pre-processing (c), model training, validation and reconstruction (d) and hosting the GSSR database (e).

Name	Source	Spatial resolution	Temporal resolution	Coverage
ERA-20C	ECMWF	1.125° × 1.125°	3 hourly	1900–2010
20-CR	NOAA-CIRES-DOE	1.0° × 1.0°	3 hourly	1836–2015
ERA-Interim	ECMWF	0.75° × 0.75°	6 hourly	1979–2019
MERRA-2	NASA	0.625° × 0.5°	1 hourly	1980–2019
ERA5	ECMWF	0.25° × 0.25°	1 hourly	1979–2019

Table 1. Metadata for the climate reanalysis datasets used for surge reconstruction.

The data-driven models employed for the surge reconstruction were developed by Tadesse *et al.*¹⁰ (hereinafter referred to as T20) by using multiple linear regression and Random Forest at each tide gauge location to derive daily maximum surge values. The overall workflow for the database development is outlined in Fig. 1 and the individual steps are discussed in more detail below.

Methods

Data sources. The training component of our storm surge reconstruction methodology employs two major input datasets, namely predictors and predictands. Predictors in T20 are atmospheric and oceanographic variables, including zonal (u_{10}) and meridional (v_{10}) wind speeds, mean sea-level pressure (SLP), sea-surface temperature (SST) and precipitation. We utilize five medium to long-term climate reanalysis products (Table 1) in order to extract relevant predictors (see Sensitivity Analysis for details). The necessary pre-processing steps are applied on predictors obtained from each reanalysis product (see Predictor Pre-processing for details).

The atmospheric reanalyses we use to produce GSSR (Table 1) can be categorized into two categories. The first category comprises satellite-era reanalyses including ERA-Interim¹¹ and ERA5¹² from the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Modern Era Retrospective analysis for Research and Applications (MERRA-2)¹³ from the National Aeronautics and Space Administration (NASA). The second category comprises centennial reanalyses including the 20th Century Reanalysis (20-CRV3)¹⁴ – a joint product from the National Oceanic and Atmospheric Administration (NOAA), Cooperative Institute of Research in Environmental Sciences (CIRES), and U.S. Department of Energy (DOE) – as well as ERA-20C¹⁵ from ECMWF. The centennial reanalyses have a coarser spatial (see Fig. 1b; the sea-level pressure from Hurricane Katrina) and in most cases lower temporal resolution than the satellite-era reanalyses, but they allow the reconstruction of longer storm surge time series. The web links to the data repositories from which we downloaded the reanalysis datasets are provided in a README file pertaining to the resulting surge reconstructions (see Table 2 for the DOIs to access the different surge reconstructions).

Data	DOI
20-CR Surge Reconstruction ²³	https://doi.org/10.6084/m9.figshare.12971075
ERA-20C Surge Reconstruction ²³	https://doi.org/10.6084/m9.figshare.12971054
ERA-Interim Surge Reconstruction ²³	https://doi.org/10.6084/m9.figshare.12971090
MERRA-2 Surge Reconstruction ²³	https://doi.org/10.6084/m9.figshare.12971009
ERA5 Surge Reconstruction ²³	https://doi.org/10.6084/m9.figshare.12970931
All scripts used for GSSR ²³	https://doi.org/10.6084/m9.figshare.12978191
Common Period Validation ²³	https://doi.org/10.6084/m9.figshare.13416293
Online Interactive Web-Map	http://gssr.info

Table 2. Information on the repository where each reconstruction is stored.

The predictand are daily maximum storm surge values at 882 tide gauges. Observed sea level data for individual tide gauges were obtained from the Global Extreme Sea Level Analysis (GESLA-2) database¹⁶. We remove annual mean sea-level from the hourly sea-level data. Daily maximum surge values (i.e., the predictand) are extracted by applying a harmonic tidal analysis to detrended hourly sea level, and deriving the maximum daily values of the residuals after the tidal signal was removed.

Sensitivity analysis. The methodology used to develop the GSSR database is the same as outlined in T20. However, we carry out an additional two-phased sensitivity analysis (Fig. 1a) with the goal of simplifying the modelling approach and making it more efficient when reconstructing surge time series over long time scales with predictor data sets from multiple reanalysis products with varying spatial and temporal resolution.

First, we explore a possible simplification by reducing the number of predictors used in T20 (10 m u10 and v10, SLP, precipitation, and SST). Conventionally, SLP and wind speed are the two main predictors used to model storm surges with statistical^{18,9,17} or numerical^{18–20} approaches. Hence, we explore whether dropping some of the predictors used in T20 has a significant impact on the results. We first establish a baseline scenario that only uses SLP, u10, and v10 as predictors (the training and validation procedure is the same as in T20). Then, we add precipitation as an additional predictor and compare the validation results against the baseline scenario (Fig. 2a). We repeat the same with SST (Fig. 2b). All predictors for this experiment come from the ERA-Interim reanalysis. At more than 92% of the tide gauges adding precipitation as a predictor changes the model accuracy, expressed as Root Mean Square Error (RMSE), by less than 5 mm; for SST, the same is true for 99% of the tide gauges. However, adding precipitation lowers RMSE more substantially for some tide gauges (e.g., New York – The Battery has a 7.3 mm reduction in RMSE, which is the highest reduction we find); such larger changes are very localized and could be the result of freshwater flows affecting the recorded surges at the tide gauges, which is not accounted for in our analysis when excluding precipitation. Similarly, there are tide gauges (mostly in the tropics) where larger RMSE reductions are found when including SST (e.g., at Puerto Deseado, Argentina, we find a 14 mm reduction). Nevertheless, the changes in model accuracy are negligible in the majority of locations, but removing SST and precipitation predictors leads to significant improvements in the efficiency of the model; hence, we chose the baseline model-setup to generate the GSSR database.

In the second phase of the sensitivity analysis, we identify the optimal area around tide gauges from where predictor information is considered to train/validate the data-driven models; T20 used a $10^\circ \times 10^\circ$ box around every tide gauge. This approach, while feasible with the ERA-Interim reanalysis used in T20, leads to increased computational cost for MERRA-2 and especially ERA5 which have much higher spatial resolution (and in the case of ERA5 also higher temporal resolution). This significantly increases the amount of data that is considered to define the predictor time series. We explore here the optimal size of the area around a tide gauge that reconciles a tradeoff between model accuracy and computation time. To this end, we test box sizes of $1^\circ \times 1^\circ$, $2^\circ \times 2^\circ$, $3^\circ \times 3^\circ$, $4^\circ \times 4^\circ$, $6^\circ \times 6^\circ$, and $8^\circ \times 8^\circ$ and train/validate the data-driven models and compare model accuracies with the baseline case that uses a $10^\circ \times 10^\circ$ box. The results (Fig. 2c,d) indicate that an area of $6^\circ \times 6^\circ$ is the optimal choice (i.e., shortest distance to upper left and lower left corners) considering model accuracy (measured in RMSE and Pearson's Correlation) and computational time (measured in hours). The latter includes the time for extracting the predictors and training and validating the model. Results are shown for ERA-Interim; differences in computation time for different box sizes are much larger for MERRA-2 and ERA5. Hence, we use a $6^\circ \times 6^\circ$ box around each tide gauge when developing the GSSR database.

Predictor extraction. Based on the findings from the sensitivity analysis, predictors (u10, v10, and SLP) are extracted from within a $6^\circ \times 6^\circ$ box around each tide gauge from the netcdf files of the atmospheric reanalyses and stored in a comma-separated value (.csv) file. The satellite-era reanalyses have a higher spatial resolution and hence a much larger number of grid points. For instance, ERA5 contains around 576 grid points in a $6^\circ \times 6^\circ$ box, whereas 20-CR contains only 36 grid points (see also Fig. 1b).

Predictor pre-processing. Following the extraction of the three predictors, we concatenate all extracted files per predictor corresponding to each year. For instance, for ERA-Interim and ERA5 with 40 years of data we concatenate 40 years of predictor data to form one long time series for each predictor and tide gauge (see Table 1 for time periods covered by the different reanalyses). Next, the three predictor time series (u10, v10, and SLP) are combined into one matrix. Following T20 we lag this matrix of multiple predictors as far back as 30 hours in order to include the delay effect of predictors on daily maximum surge. After the time-lagging of predictors,

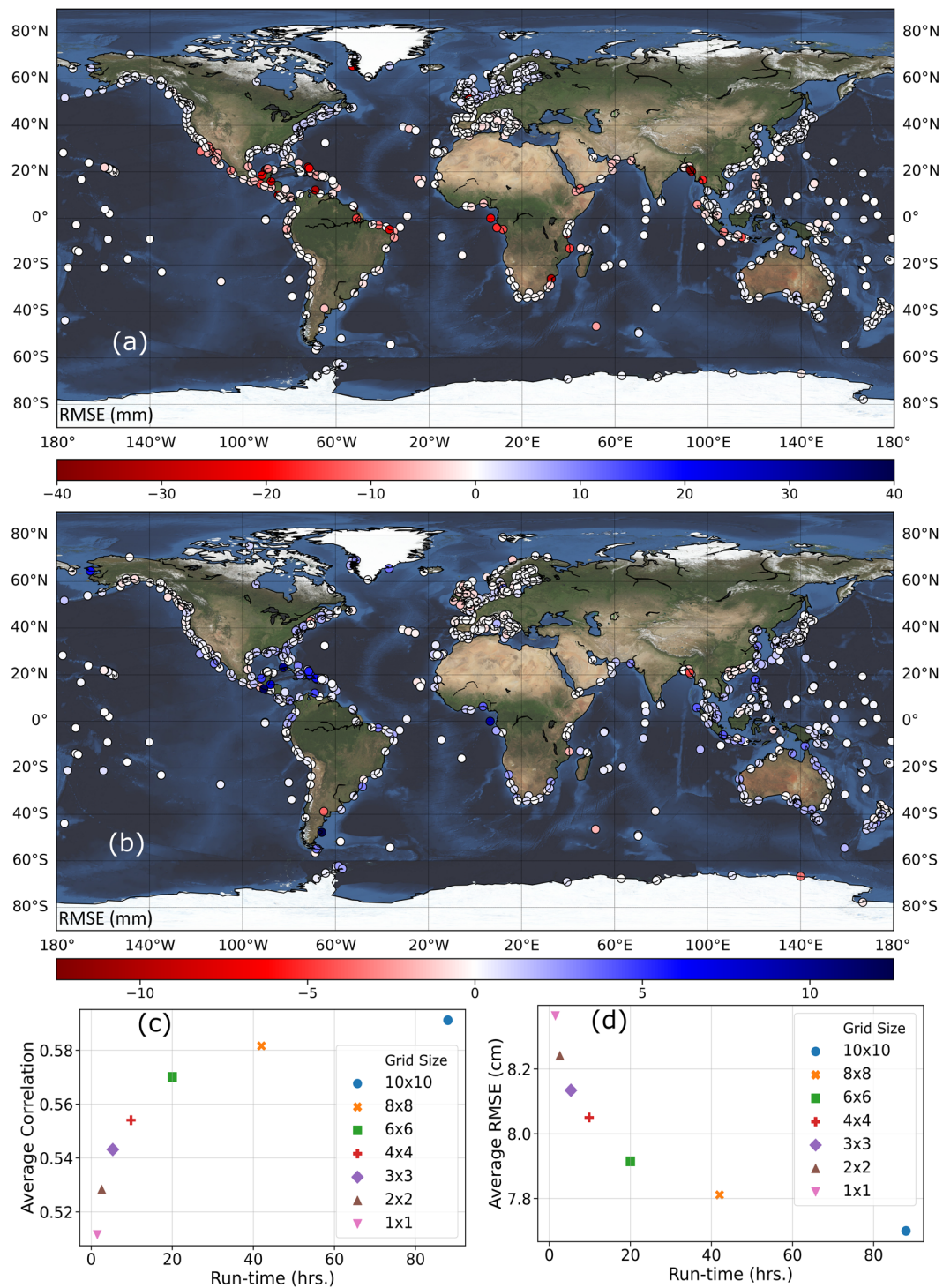


Fig. 2 Sensitivity analysis to explore model simplifications. Changes in model accuracy, expressed as RMSE (mm), when adding precipitation (a) and SST (b) to the baseline model. Negative RMSE values (shown in red) indicate lower model accuracy when adding either precipitation or SST and positive RMSE values (shown in blue) indicate higher model accuracy when adding either precipitation or SST. Changes in model accuracy when using varying box sizes to derive predictor data expressed as correlation (c) and RMSE (d).

we standardize the predictor matrix to account for the different units used by each predictor. At this stage, the predictor matrix is very large and it becomes a daunting task to apply further matrix operations, especially for the satellite-era reanalyses with high spatial and temporal resolution. For instance, the predictor matrix for ERA5 at this stage includes more than 24,000 features, whereas for 20-CR it includes approximately 2,700 features. In order to reduce dimensionality, Principal Component Analysis (PCA) is applied to extract only the principal components that explain 95% of the variance within the predictor matrix. This reduces the feature size of the

predictor matrix to approximately 250, making further matrix operations more convenient. This step concludes the pre-processing stage (see Fig. 1c) and the predictor matrix can then be used for training and validation of the data-driven models.

Model training, validation, and reconstruction. We use the pre-processed predictor matrix to train and validate the data-driven models following T20 (see Fig. 1d), using both multiple linear regression and Random Forest to link predictor and predictand data. Note, that only a portion of the predictor matrix is used for training/validation as the predictand (observed daily maximum surge) almost always has shorter coverage than the predictor data from the reanalyses. The 10-Fold cross-validation technique is used for validation. For each reanalysis dataset, we provide the validation of the data-driven models at each tide gauge location in terms of Pearson's correlation coefficient as well as the RMSE:

$$\text{Pearson's Correlation } (r_{xy}) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}}$$

where x and y represent the values of the observed and modeled surges and \bar{x} and \bar{y} , the mean of observed and modeled surges.

For the Technical Validation (see below) we also use the Relative RMSE (RRMSE) as an additional metric, following T20. RRMSE normalizes RMSE by the maximum surge variability observed at a specific tide gauge (difference between the maximum and minimum surge). This validation information can be found in the metadata section of the repository where the surge reconstructions are stored (refer Table 2). Once the training and validation stage is complete, the full predictor matrix is used to develop the surge reconstruction over the entire period where predictor information is available from the reanalysis datasets. As a measure of uncertainty, we provide the 95% prediction intervals for all simulated daily surge values.

Data Records

The DOIs of the repositories for the surge reconstructions (derived with different reanalysis products) are shown in Table 2; included in these repositories are README files, folders that contain the surge reconstructions for all tide gauges, and metadata folders with model validation results for all tide gauges (both for daily maximum surges and extreme surges above the 95th percentile). Daily max surge time series (along with 95% prediction intervals) are stored, for individual tide gauges, in comma-separated value (.csv) format.

Technical Validation

We validate the surge reconstructions for the different reanalysis products and individual tide gauges using a 10-fold cross-validation. As the five reanalysis products have different start and end times, the validation periods (or fold sizes) for which the performance metrics are derived are also different (surge reconstructions derived with the longer-term reanalyses have longer validation periods). The validation results are provided in the metadata section of the corresponding surge reconstruction (refer Table 2 for the respective DOIs).

In addition to the separate validation for the different surge reconstructions, we also assess the variability in model accuracy among the five surge reconstructions for a specific tide gauge. This allows to investigate if there are any spatial patterns where certain reanalysis products lead to better results than others. To allow for a more direct comparison across all surge reconstructions, we select a common period (1980–2010) which is covered by all five reanalysis products. We validate the reconstructions during this period in terms of Pearson's correlation, RMSE, and RRMSE, and show the variability (measured as standard deviation) of model accuracy (refer Table 2).

Figure 3, for instance, illustrates on the left panes the variability in the different performance metrics across the five surge reconstructions and depicts the respective reanalysis product that leads to the highest model accuracy. On the right panes, we show the spatial distribution (summarized in 10-degree bands) of the validation results for the surge reconstructions that lead to the highest model accuracy. This is in agreement with validation results reported in T20, including overall higher model accuracy in higher latitudes and lower accuracy in the tropics. We find an average RMSE of 9.9 cm (std = 4.1) for extratropical regions and 6.3 cm (std = 3.1) for tropical regions. Average correlation is 0.6 (std = 0.13) for extratropical regions and 0.33 (std = 0.16) for tropical regions. Moreover, at the vast majority of tide gauges the surge reconstructions derived with ERA-Interim or ERA5 lead to the best results (for 33% and 30% of tide gauges respectively). Notable exceptions are the east coast of South America, east coast of Australia, and parts of southeast Asia, where MERRA-2 leads to the best results at many locations. Overall, MERRA-2 leads to the best results at 18% of the tide gauges. At the remainder of locations ERA-20C (12%) and 20-CR (7%) lead to the best results.

We also compare the surge reconstructions with observations from the perspective of extremes. For each tide gauge, observed surges above the 95th percentile threshold are compared with the simulated values (Fig. 4). When considering the 95th percentile threshold, some tide gauges with short records have only very few data points leading to insignificant correlation coefficients; these are not shown in Fig. 4 and as a result the number of tide gauges in Fig. 4 is slightly lower than that of Fig. 3. We find an average RMSE of 22 cm (std = 9.9) for extratropical regions and 15 cm (std = 8.5) for tropical regions. Average correlation is 0.32 (std = 0.16) for extratropical regions and 0.31 (std = 0.25) in tropical regions. For 65% of tide gauges the same reanalysis that gives the best validation results for daily maximum surges also gives the best validation results for the extremes, whereas in 35% of the cases a different reanalysis leads to better results when only focusing on extremes (therefore, as outlined above, we provide validation results for both in the metadata).

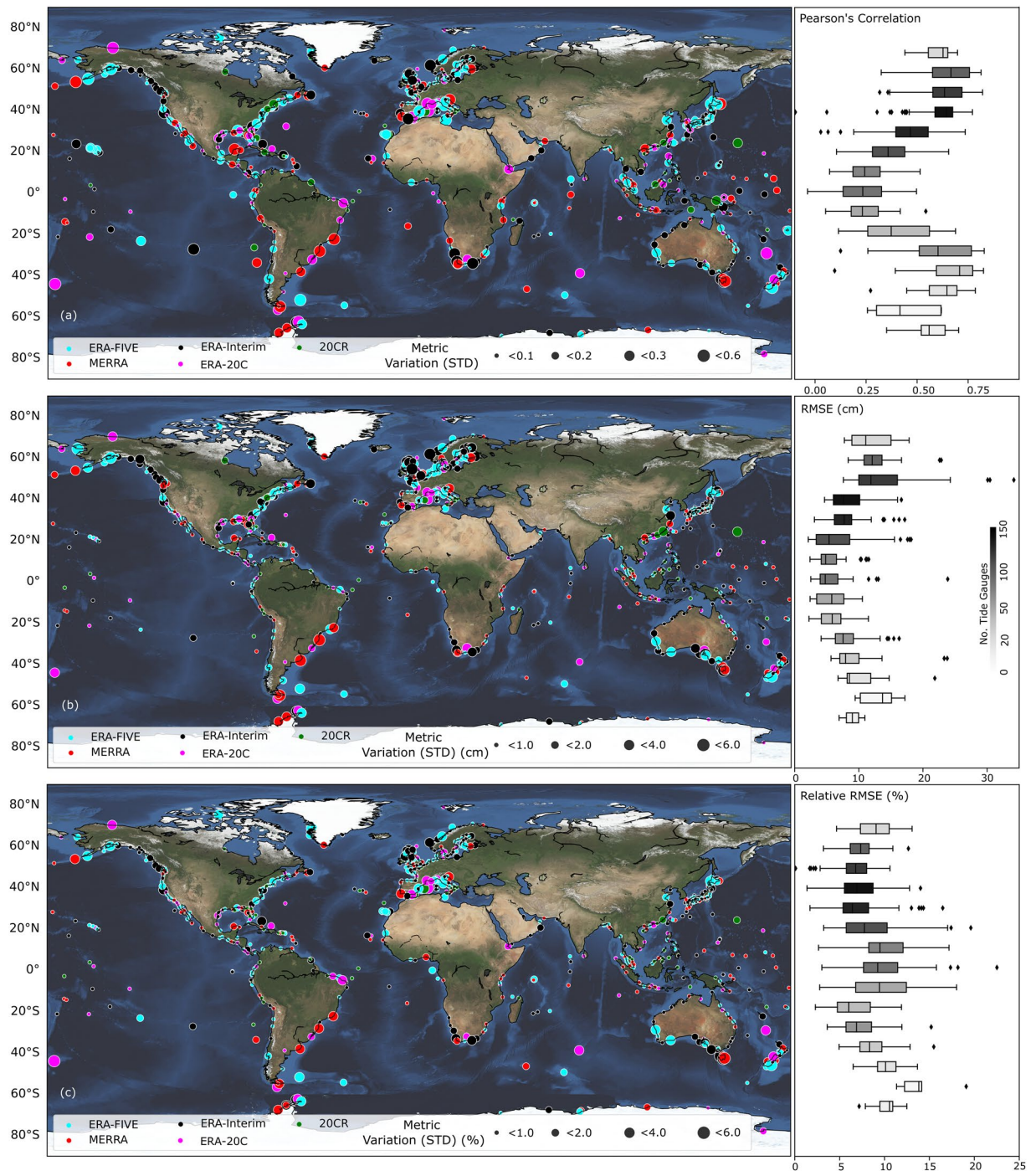


Fig. 3 Model validation results. Validation of five surge reconstructions during 1980–2010 in terms of Pearson's correlation (a), RMSE (b), and RRMSE (c). Color-coding indicates which reanalysis led to the best surge reconstruction and circle sizes indicate the magnitude of the standard deviation of the metrics across all reconstructions. Right panes show the variability of the metrics for the best surge reconstruction over 10 degree latitude bands. Tide gauges at latitudes greater than 65°N and lower than 65°S are included in the first and last boxplots respectively for each plot in the right panel, due to the small amount of tide gauges in these regions.

When comparing the validation results across surge reconstructions at individual tide gauges, higher variability is found in higher latitudes in both hemispheres (Figs. 3 and 4, left panes), compared to tropical regions. This means that in high latitudes some reanalysis products lead to much better results than others, while in the tropics all reanalyses lead to similar results. This can be corroborated by the findings in T20 that there is overall more variance in wind speed and SLP in higher latitudes and this is reflected better in some reanalyses, leading to higher variability of model accuracy across reanalyses at individual tide gauge locations. In tropical regions,

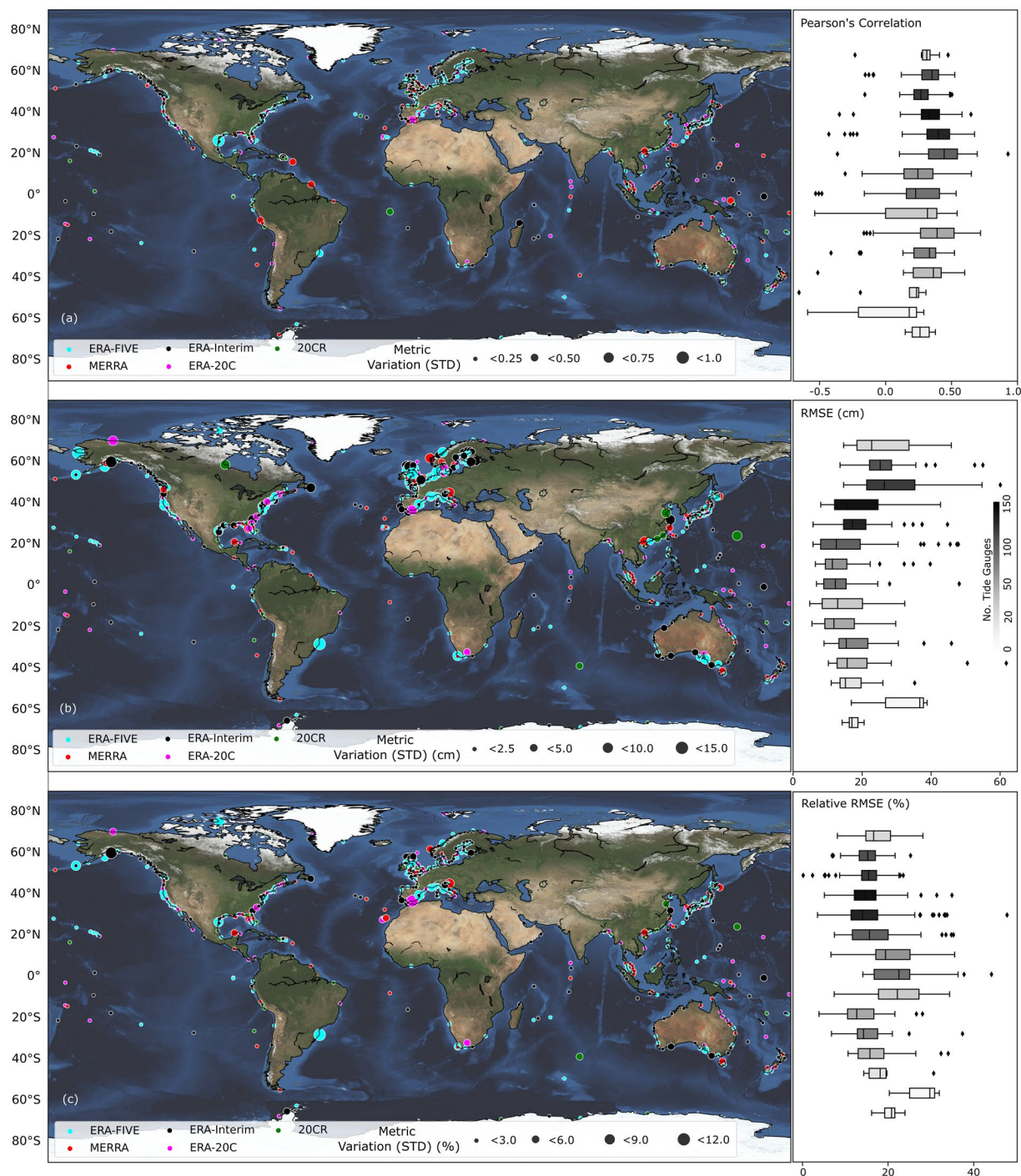


Fig. 4 Model validation results for extreme events. Validation of five surge reconstructions during 1980–2010 in terms of Pearson's correlation (a), RMSE (b), and RRMSE (c) for extreme surge events (above the 95% percentile). Color coding indicates which reanalysis led to the best surge reconstruction and circle sizes indicate the magnitude of the standard deviation of the metrics across all reconstructions. Right panes show the variability of the metrics for the best surge reconstruction over 10 degree latitude bands. Tide gauges at latitudes greater than 65°N and lower than 65°S are included in the first and last boxplots respectively for each plot in the right panel, due to the small amount of tide gauges in these regions.

where predictor variability is smaller and model performance overall poorer, the differences across reanalyses are less pronounced.

This is further elaborated in Fig. 5 where we show a spatial distribution (summarized in 30-degree bands) of the validation of all five surge reconstructions in addition to the ensemble mean. The ensemble mean is computed by taking the average of the five surge reconstructions at a daily time step. This time series is then validated against

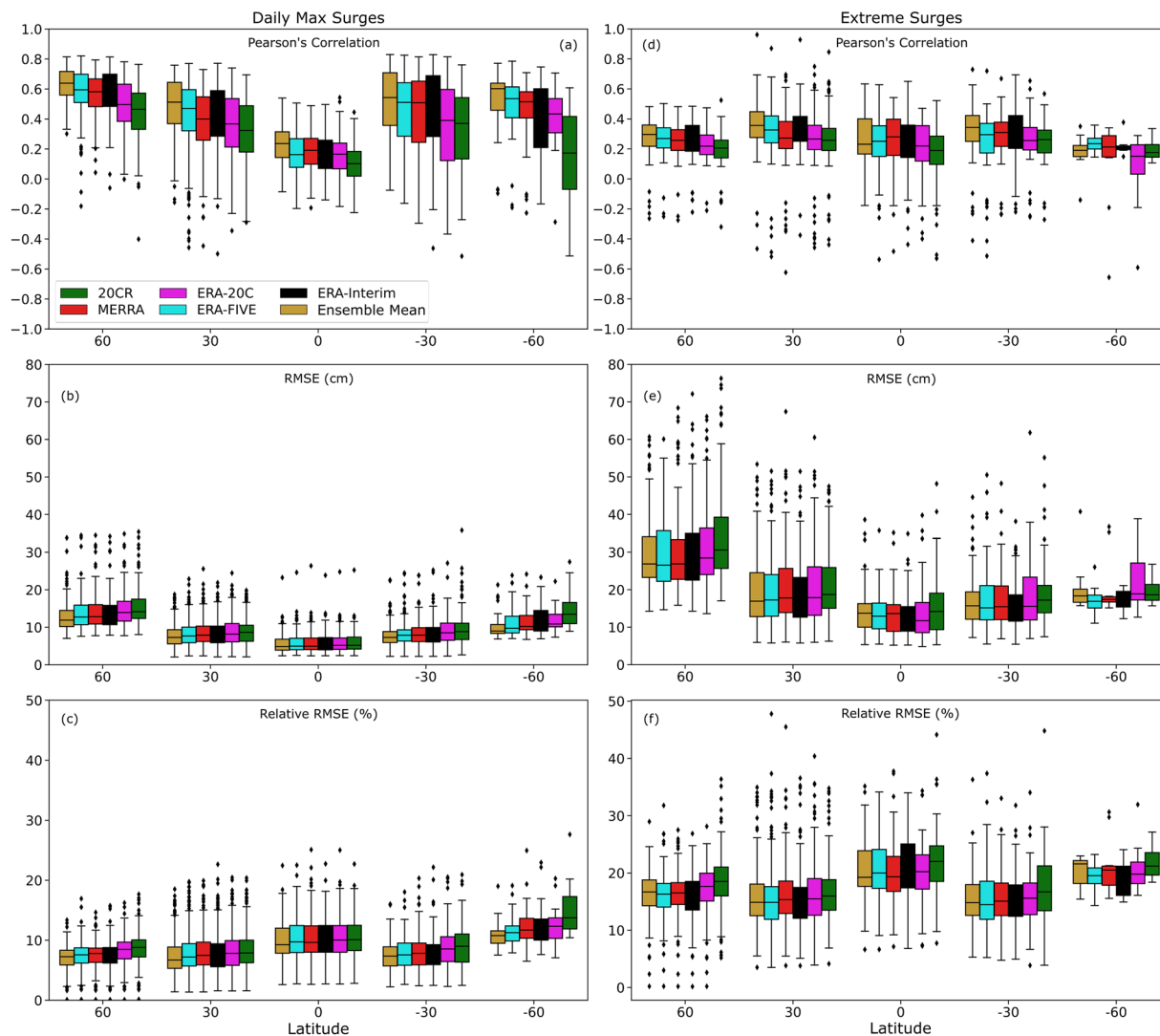


Fig. 5 Spatial variation of model performance for different reanalysis products. Validation results for five surge reconstructions and the ensemble mean when using all daily values (**a–c**) and for extremes above the 95th percentile (**d–f**); color coding denotes different reanalysis products as shown in the legends. (**a**) and (**d**) show results for Pearson's correlation coefficient, (**b**) and (**e**) for RMSE, and (**c**) and (**f**) for RRMSE.

the observed surge. The ensemble mean has relatively higher model accuracy in most places especially in higher latitudes. For future work, a weighted mean ensemble approach²¹ (by weighting the five reconstruction time series based on their validation results) could be tested as opposed to the simple mean ensemble approach used here.

In order to test the difference in performance of our surge reconstructions before and after the satellite era, we selected tide gauges that have data from 1949–2009 and used 1949–1978 and 1979–2009 (with 75% completeness) to represent the pre-satellite era and satellite era respectively. We selected these periods to match overlapping periods of 20-CR and ERA-20C. We validated the surge reconstructions from 20-CR and ERA-20C during these periods and find that out of 49 tide gauges, 31 have a lower RMSE during the satellite era (1979–2009) with an average 4.8% (std = 5%) reduction in RMSE. The remaining 18 tide gauges have an average increase in RMSE of 3.4% (std = 3.9%). RMSE results for these tide gauges are shown in Online-only Table 1.

As outlined above, one advantage of the data-driven models is the ability to not just reconstruct storm surges from multiple reanalysis products, but also over long time periods, which allows more robust analysis of trends, variability, and assessing extreme values. If we only compare results from the centennial 20-CR and ERA-20C reconstructions, ERA-20C leads to better results at approximately 70% of the tide gauges. This is likely due to the fact that 20-CR does not correct for biases in surface pressure observations from ships and buoys²² as is done in ERA-20C. If we were to use these two reanalysis products only, we would see a maximum of 20% reduction in model accuracy (compared to using the best of all five reanalyses) for 87% of the tide gauges and a maximum of 40% reduction for 98% of the tide gauges. On average the best of these two reconstructions has 0.8 cm (std = 1.12 cm) higher RMSE than the best reconstruction. When focusing only on extreme events larger than the 95th percentile, the values remain similar; a maximum of 20% reduction for 85% of the tide gauges and a

maximum of 40% reduction for 98% of the tide gauges. The global mean increase in RMSE is 2 cm (std = 3 cm) compared to the best surge reconstruction for the extremes.

Usage Notes

The five surge reconstructions based on the five reanalyses are available for public use at the specified DOIs in Table 2. Users can download the full reconstruction data for all tide gauges for the respective reconstruction. For users interested only in one tide gauge (or few tide gauges), we provide an interactive web-map where users can navigate to download the data manually. An option for bulk download of the full reconstruction is also available through the web-map. The web-map can be accessed through <http://gssr.info>.

This first version of GSSR provides useful data for certain types of analyses, especially in regions/locations where the model performs well but little observational data is available; this could include, for example, the analysis of longer-term trends, assessing variability and links to large-scale climate patterns, or performing extreme value analysis. However, users should be aware of certain shortcomings and use and interpret the results derived with the GSSR data accordingly. For example, previous work has shown that the quality of centennial reanalysis products declines when going back to the early 20th and 19th centuries¹⁷. The time periods during which the longer-term surge reconstructions are (more) reliable will be assessed in a follow-up study by comparing our reconstructions against historic sea level data using change point detection techniques, among others. Hence, we recommend that users who use the full centennial surge reconstructions interpret results with caution at this point, especially when assessing non-stationarity (as it may stem from changes in the quality of the reanalysis data over time). Furthermore, our validation shows that the model, in line with previously used global storm surge models, performs poorly in the tropical regions where predictor data shows very little variability and other processes which are not included here may be more relevant (e.g., related to wind-wave and swell activity). However, extreme value analysis as part of flood risk assessments or coastal design and adaptation studies is usually performed on the still water levels and not just the surge component. In T20 we show that the performance metrics increase substantially when our surge reconstructions are combined with predicted tide levels and then compared to observed still water levels, including high correlation and low RMSE for locations in the tropics (see Fig. 8 in T20). Bias correction techniques are also widely applied to derive more robust results from extreme value analyses when using model output data, and have not been explored with the GSSR data. We plan to make regular updates to the web-tool by including new data and products directly derived from the surge reconstructions, including more detailed information on how far back centennial reconstructions can be trusted in different regions, as well as results from assessing trends and performing extreme value analysis on still water levels.

Code availability

All scripts used for predictor extraction, predictor pre-processing, model training/validation, and surge reconstruction are available for download under the DOI listed in Table 2.

Received: 28 October 2020; Accepted: 23 March 2021;

Published online: 04 May 2021

References

- Arns, A., Wahl, T., Haigh, I. D., Jensen, J. & Pattiaratchi, C. Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise. *Coast. Eng.* **81**, 51–66 (2013).
- Haigh, I. D. *et al.* Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK. *Sci. Data* **3**, 1–14 (2016).
- Marcos, M. & Calafat, F. M. Long-term variations in global sea level extremes. *J. Geophys. Res. Ocean.* **175**, 8115–8134 (2015).
- Pugh, D. & Woodworth, P. *Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes. Oceanography* vol. 28 (Cambridge University Press, 2014).
- Douglas, B. C. Global sea rise: A redetermination. *Surv. Geophys.* **18**, 279–292 (1997).
- Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H. & Ward, P. J. A global reanalysis of storm surges and extreme sea levels. *Nat. Commun.* **7**, 1–11 (2016).
- Muis, S. *et al.* A High-Resolution Global Dataset of Extreme Sea Levels, Tides, and Storm Surges, Including Future Projections. *Front. Mar. Sci.* **7**, 1–15 (2020).
- Cid, A., Wahl, T., Chambers, D. P. & Muis, S. Storm Surge Reconstruction and Return Water Level Estimation in Southeast Asia for the 20th Century. *J. Geophys. Res. Ocean.* **123**, 437–451 (2018).
- Cid, A., Camus, P., Castanedo, S., Méndez, F. J. & Medina, R. Global reconstructed daily surge levels from the 20th Century Reanalysis (1871–2010). *Glob. Planet. Change* **148**, 9–21 (2017).
- Tadesse, M., Wahl, T. & Cid, A. Data-Driven Modeling of Global Storm Surges. *Front. Mar. Sci.* **7**, 1–19 (2020).
- Berrisford, P. *et al.* The ERA-Interim Archive. ERA Report Series No.1 Version 2.0 (European Centre for Medium Range Weather Forecasts, 2011).
- Hersbach, H. *et al.* Global reanalysis: goodbye ERA-Interim, hello ERA5. *ECMWF*. <https://doi.org/10.21957/vf291hehd7> (2019).
- Bosilovich, M. G., Lucchesi, R. & Suarez, M. MERRA-2: File Specification. **9** (2016).
- Compo, G. P. *et al.* The Twentieth Century Reanalysis Project. 1–28. <https://doi.org/10.1002/qj.776> (2011).
- Poli, P., Hersbach, H. & Dee, D. P. ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. 4083–4097. <https://doi.org/10.1175/JCLI-D-15-0556.1> (2016).
- Woodworth, P. L. *et al.* Towards a global higher-frequency sea level dataset. *Geosci. Data J.* **3**, 50–59 (2017).
- Dangendorf, S. *et al.* North sea storminess from a novel storm surge record since AD 1843. *J. Clim.* **27**, 3582–3595 (2014).
- Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H., Ward, P. J. & Muis, S. A global reanalysis of storm surges and extreme sea levels. *Nat. Commun.* **7**, 1–11 (2016).
- Vousdoukas, M. I., Voukouvalas, E., Annunziato, A., Giardino, A. & Feyen, L. Projections of extreme storm surge levels along Europe. *Clim. Dyn.* **47**, 3171–3190 (2016).
- Dullaart, J. C. M., Muis, S., Bloemendaal, N. & Aerts, J. C. J. H. Advancing global storm surge modelling using the new ERA5 climate reanalysis. *Clim. Dyn.* <https://doi.org/10.1007/s00382-019-05044-0> (2019).

21. Lin, N., Marsooli, R. & Colle, B. A. Storm surge return levels induced by mid-to-late-twenty-first-century extratropical cyclones in the Northeastern United States. *Clim. Change* **154**, 143–158 (2019).
22. Poli, P. & N. C. for A. R. S. (Eds). The Climate Data Guide: ERA-20C: ECMWF's atmospheric reanalysis of the 20th century (and comparisons with NOAA's 20CR). (2017).
23. Tadesse, M. G. & Wahl, T. A Database of Global Storm Surge Reconstruction (GSSR). *figshare* <https://doi.org/10.6084/m9.figshare.c.5124878> (2021).

Acknowledgements

M.G.T. and T.W. were funded by the National Aeronautics and Space Administration (NASA) under the New (Early Career) Investigator Program in Earth Science (grant number: 80NSSC18K0743) and the NASA Sea Level Science Team (grant number: 80NSSC20K1241). Support for the Twentieth Century Reanalysis Project version 3 dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), by the National Oceanic and Atmospheric Administration Climate Program Office, and by the NOAA Physical Sciences Laboratory. We also would like to thank ECMWF and NASA as well as GESLA for making the reanalysis datasets and the sea-level data available free of charge.

Author contributions

M.G.T. carried out the preparation and analysis of the data for the work and wrote the manuscript. T.W. provided continuous supervision, guided the research, and co-wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.G.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2021