# Global application of a regional frequency analysis on extreme sea levels

4  Authors: Thomas P. Collings[1], Niall D. Quinn[1], Ivan D. Haigh[1,2], Joshua Green[1,2], Izzy Probyn[1], Hamish
5  Wilkinson[1], Sanne Muis[3,4], William V. Sweet[5], Paul D. Bates[1,6]

7  Affiliations of Authors:

8  1. Fathom, Floor 2, Clifton Heights, Clifton, Bristol, UK. BS8 1EJ

9  2. School of Ocean and Earth Science, University of Southampton, National Oceanography Centre,
10 European Way, Southampton SO14 3ZH

11 3. Deltares, Delft, Netherlands

12 4. Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, Amsterdam, Netherlands

13 5. National Oceanic and Atmospheric Administration, National Ocean Service, Silver Spring, MD,
14 United States

15 6. School of Geographical Sciences, University of Bristol, Bristol, UK

16 Correspondence to: t.collings@fathom.global

# Abstract

Coastal regions face increasing threats from rising sea levels and extreme weather events, highlighting the urgent need for accurate assessments of coastal flood risk. This study presents a novel approach to estimating global Extreme Sea Level (ESL) exceedance probabilities, using a Regional Frequency Analysis (RFA) approach. The research combines observed and modelled hindcast data to produce a high-resolution (~1 km) dataset of ESL exceedance probabilities, including wave setup, along the entire global coastline, excluding Antarctica.

The methodology presented in this paper is an extension of the regional framework from Sweet *et al.* (2022), with innovations made to incorporate wave setup and apply the method globally. Water level records from tide gauges and a global reanalysis of tide and surge levels are integrated with a global ocean wave reanalysis. Subsequently, these data are regionalised, normalised, and aggregated, and then fit with a Generalised Pareto distribution. The regional distributions are downscaled to the local scale using the tidal range at every location along the global coastline, obtained through a global tide model. The results show 8cm of positive bias at the 1-in-10-year return level, when compared against individual tide gauges.

The RFA approach offers several advantages over traditional methods, particularly in regions with limited observational data. It overcomes the challenge of short and incomplete observational records by substituting long historical records with a collection of shorter but spatially distributed records. This spatially distributed data not only retains the volume of information but also addresses the issue of sparse tide gauge coverage in less populated areas and developing nations. The RFA process is illustrated using Cyclone Yasi (2011) as a case study, demonstrating how the approach can improve the characterisation of ESLs in regions prone to tropical cyclone activity.

In conclusion, this study provides a valuable resource for quantifying global coastal flood risk, offering an innovative global methodology that can contribute to preparing for, and mitigating against, coastal flooding.

57

## Plain language summary

59 Coastal areas are at risk of flooding from rising sea levels and extreme weather events. This study

60 uses a new way to figure out how likely coastal flooding is around the world. The method uses data

61 from observations and computer models to create a detailed map of where these floods might

62 happen at the coast. The approach can predict flooding in areas where there is little or no data. The

63 results can be used to help get ready for and prevent this type of flooding.

64

65

## 1. Introduction

Flooding provides one of the greatest threats to coastal communities globally, causing devastating impacts to affected regions. Notable events which have caused significant coastal flooding in recent years include: Cyclone Amphan (2020), which struck the Bay of Bengal producing a storm surge of up to 4.6m along the coast of Western Bengal, killing 84 people, and causing total losses over 13 billion USD (India Meteorological Department, 2020, Kumar *et al.*, 2021); Hurricane Harvey (2017), the second most costly hurricane to hit the US after Katrina (2005), which impacted 13 million people, hitting the state of Texas with a maximum storm surge of 3.8m (Amadeo, 2019); and Typhoon Jebi (2018), driving storm surges of over 3m in Osaka Bay, Japan, combined with wave action which led to flooding exceeding 5m above mean sea level (Mori *et al.*, 2019). Approximately 10% of the world's population (768 million people) live below 10m above mean sea level (Nicholls *et al.*, 2021). Coastal flooding is expected to increase dramatically into the future, predominantly caused by sea-level rise (Taherkhani *et al.*, 2020), and compounded by continued growth and development in coastal populations (Neumann *et al.*, 2015). Therefore, continuing to improve the understanding of coastal flooding is vital.

Coastal floods are driven by extreme sea levels, which arise as combinations of: (1) astronomical tides; (2) storm surges (driven by tropical and extra-tropical cyclones) and associated seiches; (3) waves, especially setup and runup; and (4) relative mean sea level changes (including sea-level rise and vertical land movement). Risk assessments of coastal flooding require high-quality and high-resolution flood hazard data, typically in the form of flood inundation maps. Inundation maps are usually derived from hydraulic models, which use high resolution extreme sea level (ESL) exceedance probabilities as a key input (e.g., Bates *et al.*, 2021; Mitchell *et al.*, 2022). The development of coastal inundation maps is reliant on coastal boundary conditions points that vary in resolution depending on application. Previous studies (e.g., Barnard *et al.*, 2019) have used 100m resolution at local sales, while regional studies (e.g., Bates *et al.*, 2021, Environment Agency, 2018) have employed resolutions between 500m and 2km.

Traditional methods for computing ESL exceedance probabilities involve extreme value analysis of measurements from individual tide gauges or wave buoys. However, long, complete records spanning numerous decades are necessary to obtain robust estimates of

97    ESL return levels (Coles, 2001). The Global Extreme Sea Level Analysis (GESLA-3) database

98    provides sea level records for over 5,000 tide gauge stations (Haigh *et al.*, 2021), but these

99    tide gauges still cover only a small fraction of the world's coastlines. Wave buoys are even

100   more sparse, largely restricted to the Northern Hemisphere and long historical records are

101   marred by discontinuities (Timmermans *et al.*, 2020). Even in areas with relatively high tide

102   gauge or wave buoy density, there are still large expanses of coastline which remain

103   ungauged. While rare extreme weather events (such as intense tropical cyclones (TCs)) are

104   often many hundreds of kilometres in size, the precise impact of the corresponding ESL can

105   often be highly localised (Irish *et al.*, 2008), meaning the peak surge occurs in an ungauged

106   location. The particular locale of peak surge for an event is determined by storm

107   characteristics, local bathymetry and coastal geography, amongst other factors (Shaji *et al.*,

108   2014). Therefore, relying on past observation-based analyses of ESL exceedance

109   probabilities to characterise return levels across a region will likely lead to the under

110   representation of rare extreme events. Finally, another limitation is that many previous

111   analyses of ESL exceedance probabilities consider the still water level component (i.e., tide

112   plus storm surge) separately from the wave set up and run up (Haigh *et al.*, 2016, Muis *et

113   al.*, 2016, Ramakrishnan *et al.*, 2022).

114   One solution to overcome sparse datasets is to use ESL hindcasts created by state-of-the-art

115   models. These include regional (e.g., (Andrée *et al.*, 2021, Siahsarani *et al.*, 2021, Tanim &

116   Akter, 2019) or global tide-surge (such as Deltares' Global Tide Surge Model v3.0 (hereafter

117   referred to as GTSM; Muis *et al.*, 2020) or wave models (e.g., Liang *et al.*, 2019).  These are

118   used to fill the spatial and temporal gaps in the observation records via historical reanalysis

119   simulation. However, their ability to accurately capture extreme events is hampered by the

120   atmospheric forcing data that is used to drive the models, as reanalysis products like ERA5

121   (Hersbach *et al.*, 2020) commonly contain biases in representing meteorological extremes

122   such as TCs (Slocum *et al.*, 2022), leading to an underestimation of event intensity.

123   Furthermore, the time period captured in reanalysis products is not adequate to represent

124   the characteristics (e.g., frequencies) of particularly rare events such as intense TCs. To

125   overcome this limitation, some studies have used synthetic event datasets representing TC

126   activity over many thousands of years (e.g., Haigh *et al.*, 2014; Dullaart *et al.*, 2021),

127   however this approach is computationally expensive.

128    An alternative and less computationally demanding solution that helps address some of the

129    problems inherent in estimating ESLs around the world's coastlines from the observational

130    record, is regional frequency analysis (RFA). The RFA methodology was originally developed

131    to estimate streamflow within a hydrological context (e.g., Hosking and Wallis, 1997), but

132    has since been used in many applications requiring extreme value analysis of meteorological

133    parameters including coastal storm surge (e.g., Bardet *et al.*, 2011; Weiss and Bernardara,

134    2013; Arns *et al.*, 2015; Calafat *et al. 2022*) and extreme ocean waves (e.g., Campos *et al.*,

135    2019, Lucas *et al.*, 2017, Vanem, 2017). The principle of an RFA is founded on the basis that

136    a homogenous region can be identified, throughout which similar meteorological forcings

137    and resultant storm surge or wave events could occur, even if the extreme events have not

138    been seen in part of that region in the historical record (Hosking and Wallis, 1997). RFA has

139    been used on a regional scale to produce coastal ESL exceedance probabilities including:

140    France (Andreevsky *et al.*, 2020, Hamdi *et al.*, 2016); the US coastline (Sweet *et al.*, 2022);

141    Northern Europe (Frau *et al.*, 2018); US coastal military sites (Hall *et al.*, 2016); and the

142    Pacific Basin (Sweet *et al.*, 2020).  However, an RFA approach has not (to our knowledge)

143    been applied globally.

144    The overall aim of this paper is to, for the first time, apply an RFA approach to estimate ESL

145    exceedance probabilities, including wave setup, along the entire global coastline. These

146    exceedance probabilities aim to better characterise ESLs driven by rare, extreme events,

147    such as those from TCs, which are poorly represented in the historical record. Uniquely, this

148    study uses both measured and hindcast datasets; includes tides, storm surges, and wave

149    setup; and calculates exceedance probabilities at high resolution (1 km) globally. The

150    specific objectives of this paper are to:

151    (1) develop and apply the RFA globally (excluding Antarctica), utilising both

152           observational tide gauge, and modelled hindcast sea level and wave records;

153    (2) illustrate  how the RFA methodology improves the representation of rare extreme

154           events in the ESL exceedance probabilities using cyclone Yasi, which impacted the

155           Australian coastline in 2011, as a case study;

156    (3) validate the RFA against exceedance probabilities estimated from the GESLA-3 global

157           tide gauge database; and

158     (4) Finally, quantify how much the RFA increases the estimation of ESL exceedance

159         probabilities in areas prone to TC activity when compared to single site analysis,

160         using hindcast datasets (Muis *et al.*, 2020 and Dullaart *et al.*, 2021).

161 This paper is laid out as follows: The datasets used are described in Section 2. The

162 methodology is detailed in Section 3, addressing objective 1. Results and validation are

163 described in Section 4, addressing objectives 2, 3, and 4. A discussion of the key findings and

164 conclusions are then given in Sections 5 and 6, respectively.

165

## 2. Data

167 We use seven primary sources of data in this study, namely: (1) still sea-level observations

168 contained in the GESLA-3 tide gauge dataset; (2) global still sea-level simulations from the

169 GTSM hindcast based on the ERA5 climate reanalysis; (3) tidal predictions from the FES2014

170 finite element hydrodynamic model; (4) significant wave heights derived from the ERA5

171 climate reanalysis; (5) mean dynamic topography from HYBRID-CNES-CLS18-CMEMS2020;

172 (6) Copernicus DEM to create a global coastline dataset; and (7) the COAST-RP dataset from

173 Dullaart *et al.*, (2021) to validate the RFA methodology. These seven datasets are described

174 in turn below.

175 Still sea level records are assembled from the GESLA-3 (Global Extreme Sea Level Analysis)

176 tide gauge dataset version 3 (Caldwell *et al.*, 2015, Haigh *et al.*, 2021). The GESLA-3 dataset

177 includes high-frequency water level time series from over 5,000 tide gauges around the

178 globe, collated from 36 international and national providers. Data providers have differing

179 methods of quality control, however each record was visually assessed by the authors of the

180 GESLA-3 dataset and graded as either: (i) no obvious issues; (ii) possible datum issues; (iii)

181 possible quality control issues; or (iv) possible datum and quality control issues. Only

182 records with no obvious issues were used in this study.

183 As discussed in Section 3, the hindcast, GTSM-ERA5 is used in all areas which are not

184 covered by tide gauge observations. GTSM is a depth-averaged hydrodynamic model built

185 using the DELFT-3D hydrodynamic model, which makes use of an unstructured, global,

186 flexible mesh with no open boundaries (Muis *et al.*, 2020). The model has a coastal

187     resolution of 2.5km (1.25km in Europe), and a deep ocean resolution of 25km. The GTSM-

188     ERA5 dataset spans the period 1979-2018, and is developed by forcing GTSM with hourly

189     fields of ERA5 10-metre wind speed and atmospheric pressure (Hersbach *et al.*, 2020).

190     GTSM-ERA5 has a 10-minute temporal resolution and provides a timeseries at locations

191     approximately every 50km along the coastline (10km in Europe). Validation carried out by

192     Muis *et al*. (2020) shows that the dataset performs well against observations of annual

193     maximum water level, exhibiting a mean bias of -0.04 m and a mean absolute percentage

194     error of 14%.

195     We use the FES2014 tidal database to generate tidal timeseries at GTSM-ERA5 locations and

196     RFA output locations. The RFA output resolution is much higher than the output resolution

197     of GTSM-ERA5, which is why FES2014 is used instead. FES2014 is a finite element

198     hydrodynamic model which combines data assimilation from satellite altimetry and tide

199     gauges (Lyard *et al.*, 2021). The model solves the barotropic tidal equations, as well as the

200     effects from self-attraction and loading. The gridded resolution of the output is 1/16°. The

201     model was extensively validated against tide gauges, satellite altimeter observations, and

202     alternative global tide models by Lyard *et al*. (2021) and was found to have an improved

203     variance reduction in nearly all areas, especially in shallow water regions. The Python

204     package distributed with the FES2014 data (https://github.com/CNES/aviso-fes) was used to

205     simulate tidal timeseries.

206     To calculate wave set up we use significant wave heights (Hs) from the ERA5 reanalysis

207     (Hersbach *et al.*, 2020), covering the period 1979 to 2020. The spatial resolution of the ERA5

208     wave model output is 0.5° x 0.5°, and the temporal resolution is hourly. Independent

209     validation of hourly Hs performed by Wang & Wang (2022) finds little bias in the dataset (-

210     0.058 m), however the authors go on to conclude that Hs of extreme waves tends to be

211     underestimated (by 7.7% in the 95% percentile), a conclusion supported by Fanti *et al.*

212     (2023).

213     We use mean dynamic topography (MDT) to convert water levels from mean sea level as

214     measured by tide gauges to mean sea level as referenced by a geoid, for use in subsequent

215     future studies involving inundation assessments using hydraulic modelling. MDT describes

216     the change in sea surface height due to the effects of the winds and currents in the ocean.

217     Digital elevation models (DEMs), a key input to hydraulic models, typically use a geoid as a

218 vertical datum. A geoid is an equipotential surface of mean sea level under the sole effect of

219 gravity, in the absence of land masses, currents and tides (Bingham & Haines, 2006). To

220 convert water levels from tide gauge mean sea level to the geoid mean sea level, the

221 HYBRID-CNES-CLS18-CMEMS2020 MDT dataset is used (Mulet *et al.*, 2021). The spatial

222 resolution of this dataset is 0.125° x 0.125°. Errors associated with this dataset are largely

223 caused by the input satellite altimetry data and can be up to 10 cm in some areas. The MDT

224 at the shoreline is illustrated in the Appendix Fig. A1.

225 The Copernicus 30m DEM (European Space Agency, 2021) is used to create a high-resolution

226 global coastline. This is used to define the RFA output points at approximately 1 km intervals

227 along the global coastline (excluding Antarctica), resulting in over 3.4 million points.

228 Finally, in addition to GTSM-ERA5, we use the COAST-RP dataset from Dullaart *et al.* (2021)

229 to validate the RFA methodology. COAST-RP uses the same hydraulic modelling framework

230 as GTSM-ERA5 but simulates extra-tropical and tropical surge events separately using

231 different forcing data. In areas prone to TC activity, synthetic TCs representing 3,000 years

232 under current climate conditions are used from the STORM dataset (Bloemendaal *et al.*,

233 2020). These synthetic TC model runs have been validated against observed IBTrACS-forced

234 model runs, and found to show differences in ESLs at the 1 in 25 year return level of less

235 than 0.1 m at 67% of the output locations in TC prone areas (Dullaart *et al.*, 2021). In regions

236 impacted only by extra-tropical storms, a 38-year timeseries of ERA5 data is used (Hersbach

237 *et al.*, 2020). The surge levels from each set of simulations are probabilistically combined

238 with tides to result in a global database of dynamically modelled storm-tides.

239

240 ## 3. Methods

241 The first objective of this study is to develop and apply an RFA approach globally,

242 encompassing still water levels and wave set up. In Section 3.1 we describe the methods

243 used to process the data used in this study. In Section 3.2 we layout the global application of

244 the RFA approach using observational and modelled data.  The methods used to validate the

245 results are explained in Section 3.3.

246 An overview of our methodology is illustrated in Fig. 1. This study broadly follows the

247 methodology of Sweet *et al.* (2022) and applies an RFA to both tide gauge and GTSM-ERA5

248 records. As such, the terms 'water level record' and 'record location' are used to describe

249 both tide gauge records and GTSM-ERA5 data. The method can be summarised in five key

250 steps: (i) collation and pre-processing of tide gauge, GTSM-ERA5, FES2014, and ERA5 Hs

251 data; (ii) spatial discretisation of water level records into regions; (iii) application of the RFA

252 to regional water level records (in areas unsuitable for an RFA (because there are less than 3

253 gauges in a region, or the regional water levels records are heterogenous), a peaks-over-

254 threshold analysis of individual GTSM-ERA5 water level records is used); (iv) conversion

255 (downscaling) of RFA exceedance levels to local exceedance levels at the output coastline

256 points, using FES2014 tidal range (in areas unsuitable for an RFA, nearest-neighbour

257 interpolation is used to assign local exceedance levels); and (v) correction of bias and

258 datums to convert water levels to geoid mean sea level, using FES2014 mean higher high

259 water and global MDT (HYBRID-CNES-CLS18-CMEMS2020). The final section of the methods

260 (vi) describes the validation techniques. These steps are described in detail below.

261

262

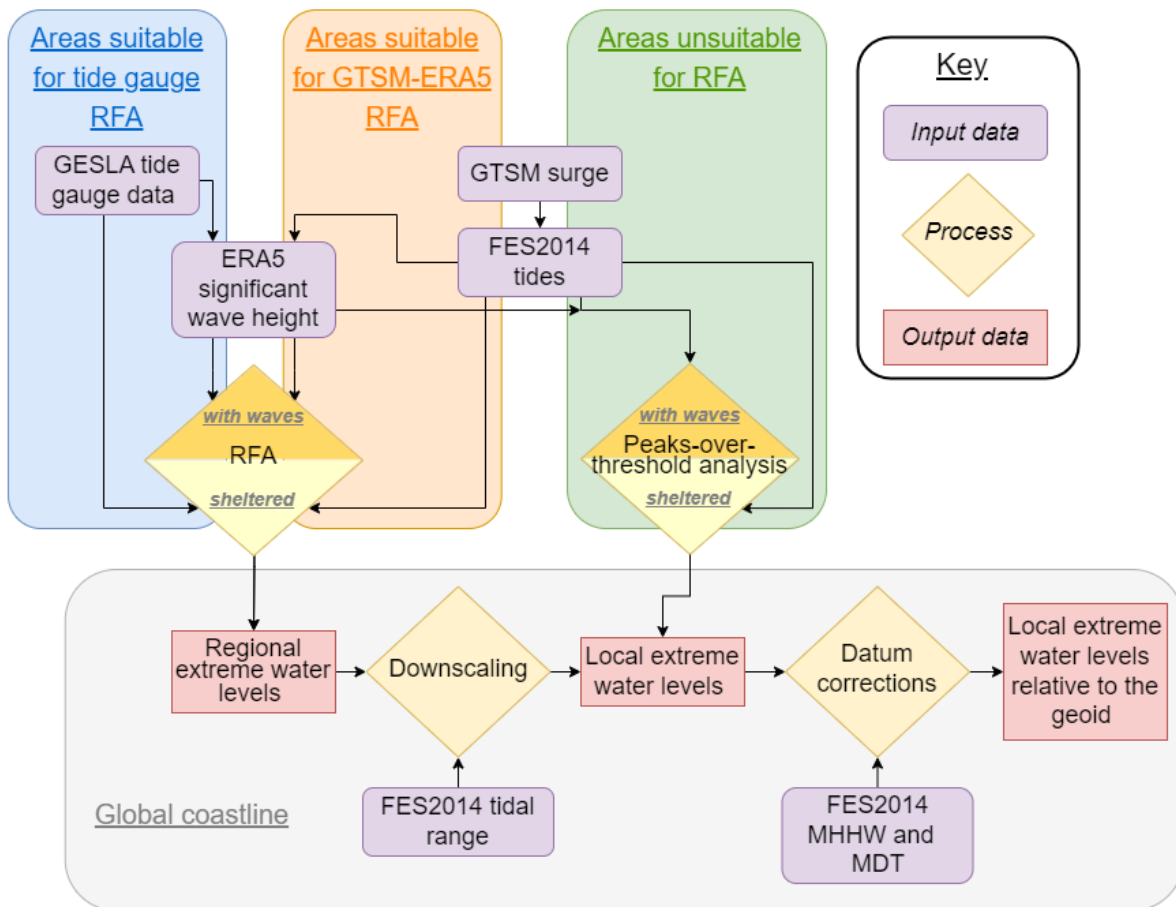



*Figure 1: Schematic flow diagram detailing the data sources and processes involved in producing a global set of extreme water levels*

## 3.1 Data processing

The GESLA-3 dataset was filtered to sample appropriate input data by removing duplicates, gauges located in rivers (away from the coast), and gauges that fail quality control checks carried out by the authors of the dataset (such as suspected datum jumps). The surge component of GTSM-ERA5 at each record location is isolated from the water level timeseries using a tide only simulation and superimposed upon a tidal timeseries created with FES2014, as the FES2014 tidal elevations performed better than those of GTSM in initial testing against in-situ observation. The decision to use tides from FES2014 is further supported by the conclusion from Muis *et al,* (2020), in which they state "It appears that biases increase in regions with a high tidal range, such as the North Sea, northern Australia, and the northwest of the United States and Canada, which could indicate that GTSM is outperformed by the FES2012 model that was used to develop the GTSR dataset." Tidal

278    timeseries were also computed at each of the coastline output locations for use in

279    downscaling the regional outputs, and in the bias and datum corrections of the local ESL.

280    Wave setup is the static increase in water level attributed to residual energy remaining after

281    a wave breaks (Dean & Walton, 2010), and therefore is only observed in areas exposed to

282    direct wave action. In this study, wave setup is approximated as 20% significant wave height

283    (Hs) from the ERA5 reanalysis, following the recommendation from the review of numerous

284    laboratory and field experiments (Dean & Walton, 2010) and previous related studies (Bates

285    *et al.*, 2021, Vousdoukas *et al.*, 2016). Wave setup is assigned to the nearest record location

286    using a nearest-neighbour approach.  Wave setup is assumed to be absent in sheltered

287    areas (e.g., bays and estuaries). To account for this, the global coastline is classified as either

288    sheltered or exposed, and the final extreme water levels are drawn from an RFA that is

289    processed with or without wave setup added in. To classify the coastline, each coastline

290    point is evaluated to determine if it is exposed from a minimum 22.5° angle over a fetch of

291    50km. A total of 16 equal angle transects are drawn, extending 50km from each coastline

292    point. If two or more adjacent transects do not intersect with land, the coastline point is

293    considered exposed. Applying wave setup using this approach is an obvious simplification

294    that has been used for the ease of global application. In reality wave setup is impacted by

295    local bathymetry and coastal geometry, as well as local wind and wave conditions. There are

296    other more complex methods for estimating wave setup that incorporate some aspects of

297    bathymetry and coastal geometry, such as Stockdon et al. (2006).

298    To process the RFA with wave setup, daily maximum wave setup is added to the daily

299    highest water levels. Where tide gauge records fall outside of the temporal range of the

300    ERA5 data, a copula-based approach was used to fit a simple statistical model between daily

301    peak water levels and daily max Hs, providing a prediction of the daily max Hs. The RFA is

302    then executed as described below. Tide gauges are assumed to be located in sheltered

303    regions, such as bays and estuaries, thus tide gauge records are not impacted by wave

304    setup.

305    ## 3.2 Spatial discretisation of water level records into regions

306

307    Water level records are spatially clustered to form a potential pool from which regional

308    exceedance levels can be characterised. To do this, the global coastline is divided into 1° by

309     1° grid cells, which are used as the regions to apply the outputs for each RFA. All record

310     locations within a 400km radius (same as Hall *et al.* (2016) and Sweet *et al.* (2022)) of the

311     grid cell centroid that have at least 10 consecutive years of good (>90% completeness) data

312     are identified (minimum of 3 water level records, maximum of 10 (same as Sweet *et al.*

313     (2022)). This step is illustrated in Fig. 2A. Record locations which are geographically within

314     range, but are separated by a large expanse of land, and thus likely forced by different

315     storm patterns are removed from the record location selection. To achieve this, a line is

316     drawn between the grid cell centroid and each record location. The land intersected by the

317     line is divided, and the areas of land on either side of the line are summed. A ratio of the

318     length of the line to the area of land segmented by the line is then calculated. A threshold of

319     100 was empirically evaluated using expert judgement based on a number of test cases,

320     above which records are removed from the grid cell analysis. This approach ensures that, for

321     example, record locations located on the east coast of Florida (e.g., Mayport) are not

322     grouped with those on the west coast (e.g., Cedar Key) when characterising regional growth

323     curves, despite the relatively short straight-line distance between them. Fig. 2A exemplifies

324     three tide gauges which have been excluded from possible selection despite lying within a

325     400km radius to the grid cell centroid as the land that separates them is considerably large

326     when compared to the distance.  This spatial discretisation of regions results in a total of

327     836 tide gauge records (with a mean record length of 17 years) and 18628 GTSM-ERA5

328     records for use in the application of the RFA.

329

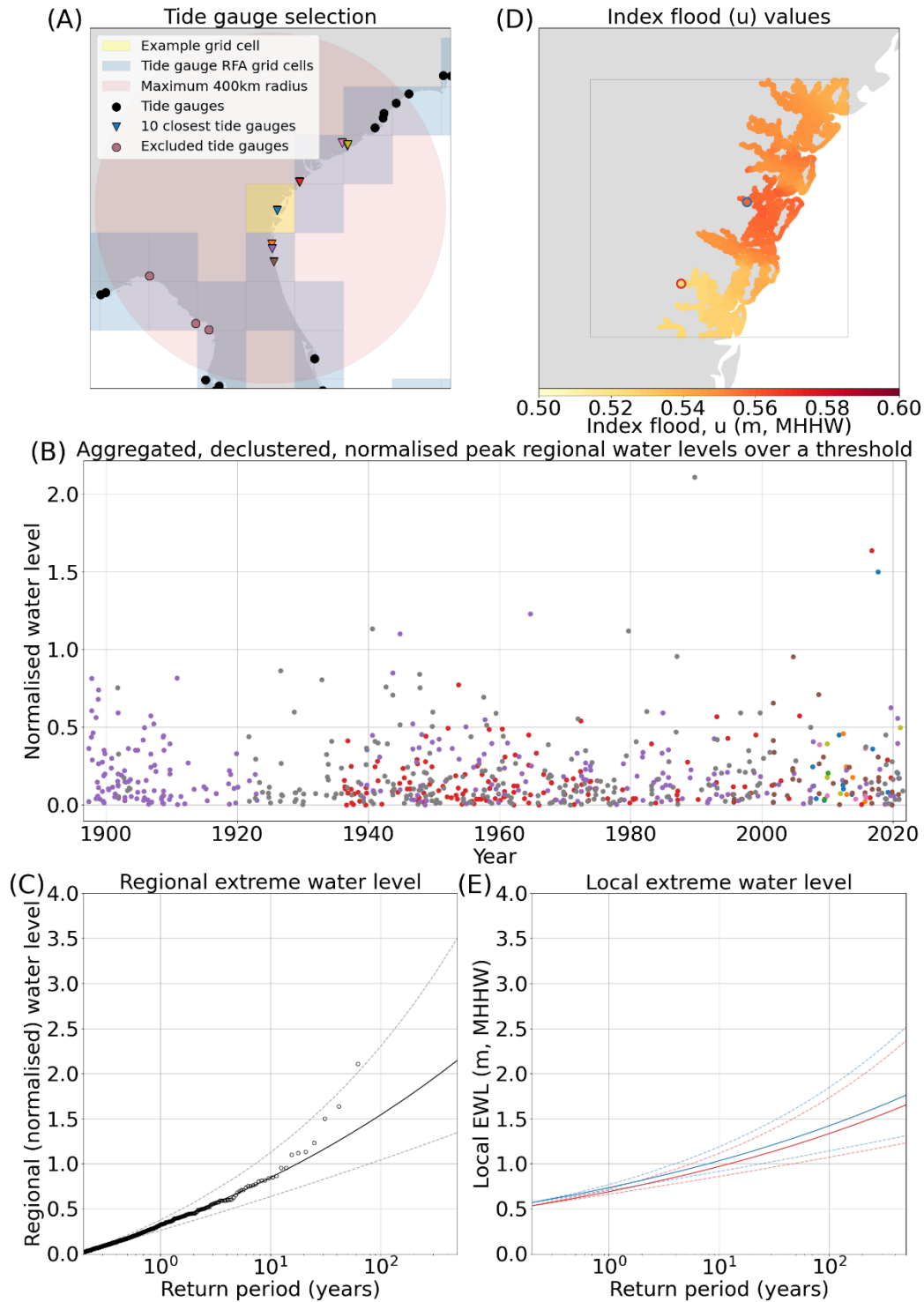



Figure 2: *Illustrating a selection of the steps through the RFA. (A) The 1° by 1° grid cells along the East Coast of the US,*
*along with the locations of the tide gauges, and the tide gauges selected for the RFA of the example grid cell. The tide*
*gauges excluded from possible selection by the distance/land area ratio are also indicated. (B) The aggregated, declustered,*
*normalised peak regional water levels over a threshold for each of the tide gauges used in the example grid cell. The colours*
*indicate peak water levels from the individual tide gauges in the region. (C) The regional extreme water levels, ascertained*
*by fitting a Generalised Pareto distribution to the data displayed in panel (B). (D) The index flood values of the example grid*
*cell, found by linearly interpolating the u value from the two closest tide gauges, and scaling by tidal range. The locations of*
*two coastline points used to produce local extreme water levels in panel E are also highlighted. (E) The local extreme water*
*level at two shoreline points inside the example grid cell, each with different index flood values as indicated in panel D.*

339  The RFA is preferentially applied to tide gauges in areas where the gauge density is sufficient

340  (minimum 3 gauges within a 400km radius, same as Hall *et al.* (2016) and Sweet *et al.*

341  (2022)). Outside of these areas, the RFA is implemented using data from GTSM-ERA5. In

342  some regions, the density of homogenous record locations from GTSM-ERA5 is also too low

343  for the RFA to function, in which case the ESL exceedance probabilities are interpolated

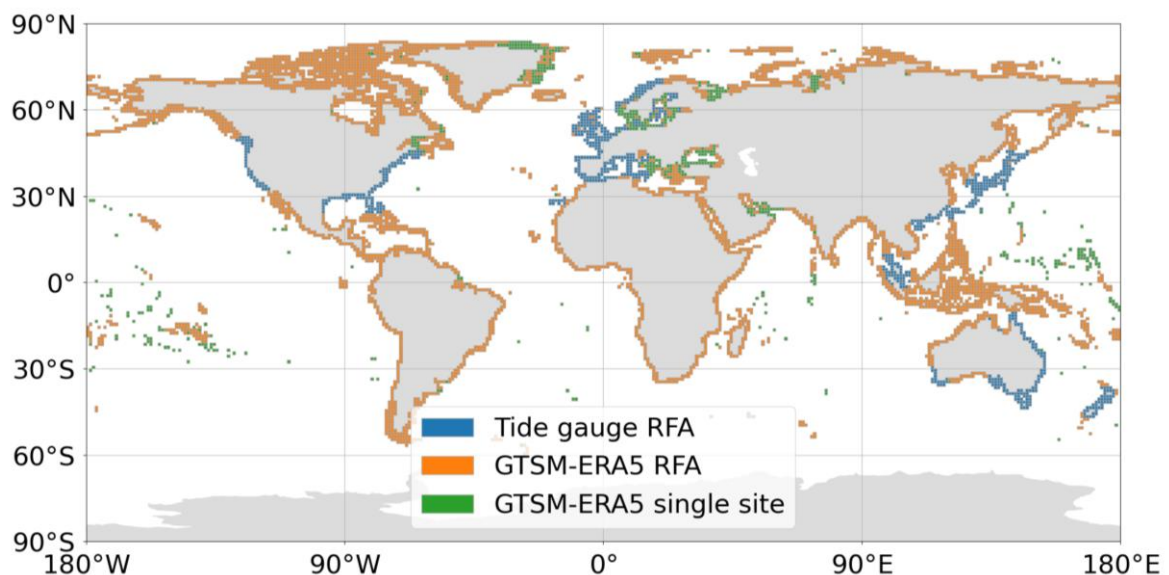
344  from a single site peaks-over-threshold analysis of the nearest GTSM-ERA5 record location.

345  The geographical locations of these areas are shown in Fig. 3. From the 5,975 global coastal

346  grid cells, ESLs at 851 are computed using tide gauge data, 4,555 are calculated using an RFA

347  of GTSM-ERA5 data, and 569 are calculated using GTSM-ERA5 data from the nearest record

348  location.

349

350  *Figure 3: This map shows the global distribution the areas in which the tide gauge RFA is used, the GTSM-ERA5 RFA is used,*
351  *and the areas which are interpolations of single site analysis from GTSM-ERA5.*

## 3.3 Application of the RFA

353

354  Tide gauge records are referenced to different vertical datums, so to ensure consistency, the

355  mean over the most recent 19-year epoch is subtract from the water level record, and the

356  timeseries is linearly detrended to the centre year of the most recent available epoch (2002-

357  2020), resulting in 2011. GTSM-ERA5 records are referenced to MSL over the period of

358  1986-2005, and so the timeseries are linearly detrended to reference the same tidal epoch

359  as the tide gauge records, centred on 2011. Within each cluster of gauge (or model) records,

360     the water level time series are resampled to hourly resolution and converted to mean

361     higher high water, defined as the mean daily highest water level over a 19-year epoch, to

362     account for differences in tidal range between record locations. In the case of records with

363     fewer than 19 years of data available the maximum continuous epoch is used instead.

364     Daily highest water level is determined from the hourly time series of each measured or

365     modelled record. The time series are then declustered using a 4-day storm window to

366     ensure event independence. This window length was used by Sweet *et al.,* 2020 and Sweet

367     *et al.,* 2022, and is a similar length to the storms that cause surge events in the UK (Haigh *et*

368     *al.*, 2016). The index flood $u$, defined as the 98th percentile of the declustered daily highest

369     water levels (Sweet *et al.,* 2022), is used as the exceedance threshold at which to normalise

370     the water level at each record location, as follows:

371     $Normalised\ water\ level = (Observed\ exceedence\ water\ level - u)\ /u$        (eq. 1)

372     The normalised datasets are then aggregated and further declustered to ensure only one

373     peak water level is retained for each regional event. This is shown in Fig. 2B for an example

374     grid cell. Following Hosking and Wallis (1997), a statistical heterogeneity test (H) is

375     undertaken to ensure the homogeneity of the region. If the H-score is less than 2, then the

376     region is considered sufficiently homogenous. If the H-score is greater than 2, then the

377     furthest water level record from the grid cell centroid is removed from the region, and the

378     test re-run. This process is repeated until the H-score is less than 2. In a minority of cases,

379     the heterogeneity test fails due to an anomalous record that lies within the closest 3

380     sampling locations to the grid cell centroid. In this instance the test is rerun, except after the

381     furthest record is removed, all the remaining records are sequentially removed and

382     replaced, until the H-score is less than 2.

383     After the region is confirmed to be homogenous, a Generalised Pareto distribution is fitted

384     to the aggregated, declustered, normalised regional water levels using a penalised

385     maximum likelihood method to estimate regional extreme water levels (REWLs). This is

386     illustrated at an example in Fig. 2C.  This is repeated for the aggregated regional water levels

387     for each 1° by 1° grid cell.  While theoretically correct, applying distribution fits to real world

388     data can sometimes give unrealistic results, particularly in the estimation of the lower

389     frequency space. In these cases, growth curve optimisation is undertaken to ensure the

output local extreme water levels are plausible in real world scenarios. To ensure

consistency, an empirical threshold of 0.35 for the shape parameter is used to determine

which curves will generate unrealistic extreme water levels. The empirical threshold of the

shape parameter is determined based on expert judgement of plausible real world

maximum surge heights in the low frequency events. To correct these curves, where this

threshold is exceeded, we use the shape and scale parameters of the nearest grid cell which

has a shape parameter less than 0.35. In total, 34 grid cells had their shape and scale

parameters adjusted, mostly concentrated in the Gulf of Mexico and Japan.

## 3.4 Downscaling to local extreme water levels

Local extreme water levels (LEWLs) are then estimated from the regional growth curves

using the following relationship:

$$LEWL = (REWL * u) + u \qquad \text{(eq. 2)}$$

for each coastal point along the coastline contained within the grid cell represented by the

REWL. The index $u$ is estimated at the coastline points using an inverse distance weighting

interpolation of the $u$ values for the two closest record locations, scaled by tidal range. This

deviates from the methodology set out by Sweet *et al.* (2022), in which they recommend

drawing $u$ values from a linear regression of $u$ against tidal range values from record

locations across a region. We found this approach led to significant differences in LEWLs at

record locations when compared to single site analysis of water level records, and hence

have modified the methodology. Fig. 2D exhibits an example of the index flood for every

shoreline point in an example grid cell. Tidal ranges are calculated as the difference

between mean higher high water and mean lower low water. Tidal harmonics from FES2014

are used to predict mean higher high water and mean lower low water at each coastline

point. The index flood, $u$, is used to downscale the REWLs, which represent the ESL

characteristics of the entire grid cell. LEWLs are output in the format of return levels for a

range of exceedance probabilities. Two example LEWL curves are shown in Fig. 2E, which

have been computed using different index flood values, as indicated in Fig. 2D.

## 3.5 Bias and datum corrections

420   The last stage of the LEWL calculation involved characterisation and removal of bias in the

421   high frequency portion of the exceedance probability curves, relative to a single site analysis

422   of water level records (within which we expect the high frequency water levels to be

423   accurately modelled). Other surge RFA studies also concluded that the approach generally

424   yields higher estimated surge heights when compared to single site analysis, because during

425   the regionalisation process an extreme event that occurred in one location is assumed to

426   have the same probability of occurring at another location within the homogeneous region.

427   (Bardet *et al., 2011; Sweet et al., 2022).* Bias is quantified based on the divergence in the 1-

428   in-1-year return period at each tide gauge/GTSM-ERA5 location and the corresponding

429   LEWL predictions. This bias is used as a correction term and is removed from the LEWLs. As

430   the density of the coastline points is much greater than the density of the tide

431   gauges/model output locations, the correction term is interpolated across all coastal LEWL

432   points based on correlation between monthly values of the 99th percentile of tidal

433   elevations produced over a 3-year period centred on 2011,  computed using FES2014 at the

434   tide gauge/GTSM-ERA5 location and neighbouring coastline points. The mean bias

435   correction across all gauges is 8 cm.

436   Datum corrections are applied to ensure the LEWLs are correctly referenced to a vertical

437   datum which can be used for hazard assessment applications, such as inundation modelling.

438   Inundation models utilise digital elevation models, which typically reference a geoid as the

439   vertical datum. The output water levels from the RFA are transformed from mean higher

440   high water to Mean Sea Level (MSL) by adding the approximation of mean higher high water

441   (above MSL) from the FES2014 simulations to each of the boundary condition points. The

442   corrected MDT dataset from (Mulet *et al.*, 2021) is applied to convert water levels from MSL

443   from the FES2014 model to the 'MSL' of a commonly used geoid, EGM08.

## 3.6 Validation methods

444
445

446   In this section we define a range of validation techniques used to address objectives 3 and 4.

447   To validate the RFA ESLs against tide gauge records from GESLA (objective 3), a comparison

448   is made against ESL exceedance probabilities calculated at the individual tide gauges used to

449   inform the RFA. To quantify the degree to which the RFA approach improves the estimation

450      of ESL exceedance probabilities compared to single site analysis (objective 4), two

451      assessments are made.

452      Firstly, the divergence between GTSM-ERA5 RFA ESL and GTSM-ERA5 single site ESL for the

453      entire global coastline are quantified. These are then contrasted against the differences

454      between return levels from GTSM-ERA5 (Muis *et al.*, 2020) and COAST-RP (Dullaart *et al.*

455      2021). The comparison can then identify regions in which the historical ESLs are poorly

456      represented due to the limited record lengths.

457      Secondly, a leave-one-out cross validation is undertaken using GTSM-ERA5 data. Leave-one

458      out-cross validation aims to address the common issues involved with validating statistical

459      models. One common method to validate models is split-sample validation, in which the

460      data is split into two groups, a training set and a validation set, which are generally 70% and

461      30% of the data respectively. The model is then trained on the larger set and validated

462      against the smaller set. The drawbacks of this method include a highly variable validation

463      error, due to the selection of the training and validation sets, as well as a validation error

464      bias caused by training the model on only 70% of the available data (James *et al.*, 2013).

465      Instead of using a 70/30 split of the data, leave-one-out cross validation uses a larger

466      proportion of the data to train the model, while validating against a smaller sub-sample, but

467      repeats this process multiple times to generate a robust validation. To do this, we identified

468      1000 grid cells which use 10 GTSM-ERA5 records for the RFA and contain 3 GTSM-ERA5

469      record locations inside the grid cell (and therefore the RFA can be used to directly estimate

470      ESLs at the record locations). One of the GTSM-ERA5 records from inside the grid cell is

471      removed from the RFA process, and the REWL is calculated using the 9 remaining gauges.

472      The LEWL is then predicted at the record location which has been left out, using the index

473      flood, u at the record location. These LEWLs are then contrasted with a single site analysis of

474      the water level record that was removed from the RFA. The process is then repeated for the

475      2 other GTSM-ERA5 record locations which lie within the grid cell. This means each of the

476      1000 models is being tested three times, against 90% of the available data, thus giving a

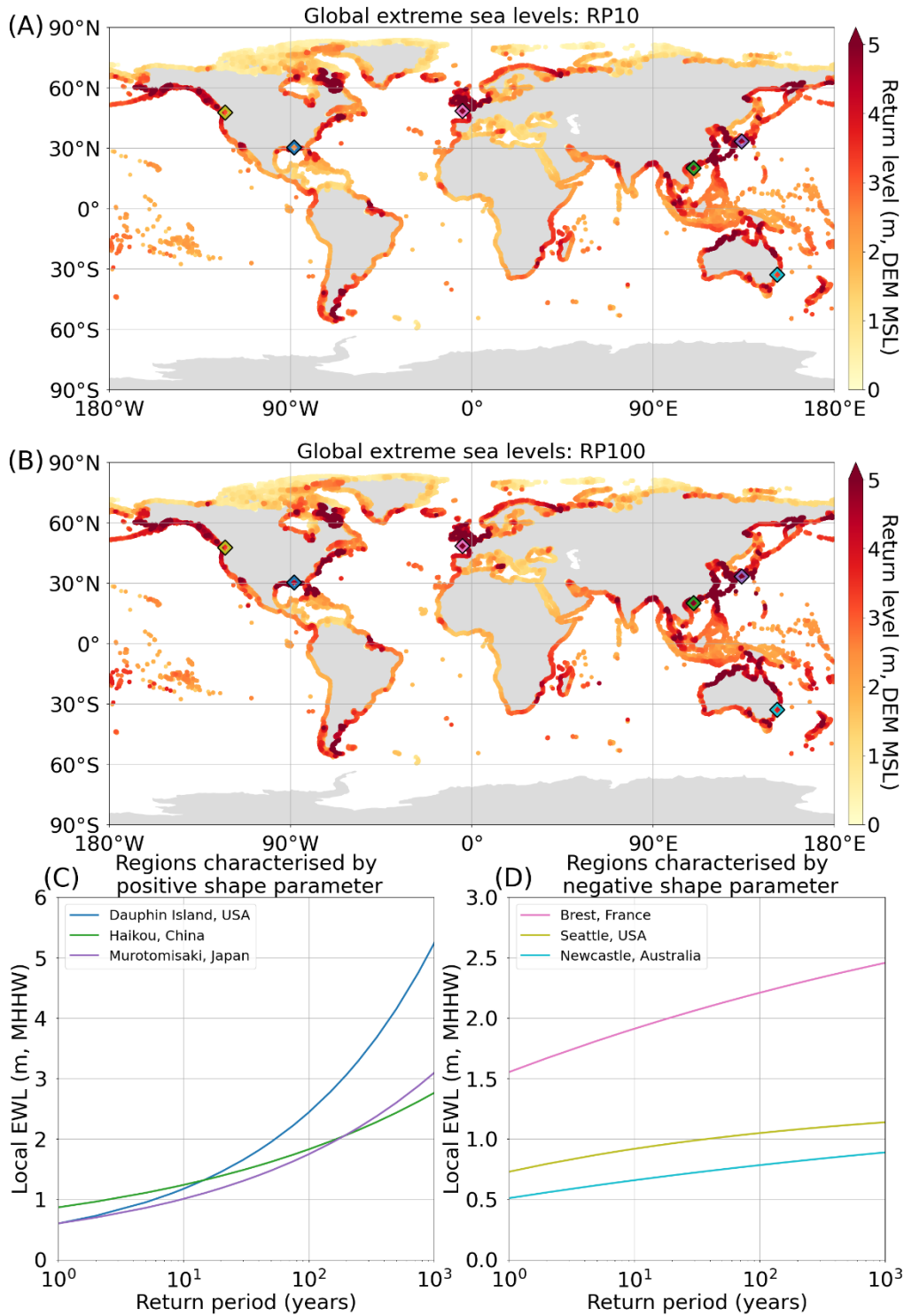
477      more robust realisation of the model when trained on 100% of the data.

478

# 4. Results

The results section is divided into four sub-sections. Section 4.1 presents the results of the global application of the RFA, showing both the global view of two return periods and the return levels for selected sites around the world. Section 4.2 illustrates how the RFA methodology improves the characterisation of rare extreme events using Cyclone Yasi (objective 2). In section 4.3 we validate the RFA against estimates of ESL from GESLA tide gauges (objective 3). Finally, in section 4.4 we quantify the improvements made by using an RFA approach when compared to a single site analysis of water levels (objective 4).

## 4.1 Global application of RFA

The final ESL exceedance probabilities (including wave setup) created at high resolution around the global coastline are displayed in Fig. 4, for the 1-in-10 and 1-in-100-year return periods. Both the 1-in-10 year (Fig. 4A) and 1-in-100 year (Fig. 4B) return periods show similar spatial patterns, with 1-in-100-year return periods exhibiting greater increases as expected in areas prone to TC activity (e.g., the Gulf of Mexico, Australia, Japan, and China). ESLs are higher in regions with large tidal ranges such as the Bay of Fundy, the Patagonia Shelf, the Bristol Channel in UK, the northern coast of France, and the northwest coast of Australia. The return levels for 6 select tide gauge locations, 3 of which are characterised by a positive and 3 of which are characterised by negative shape parameter from the Generalised Pareto distribution are shown in Fig. 4C and 4D respectively, relative to mean higher high water. The locations of the 6 tide gauges are indicated in both Fig. 4A and 4B. Regions exhibiting positive shape parameters are typically prone to TC activity and associated surge and wave events. As a result, these regions experience more significant increases in return levels at higher return periods than regions with negative shape parameters. Regions characterised by negative shape parameters have different drivers of ESL events, for instance extra-tropical storms surges or tide dominated ESLs (Sweet *et al.*, 2020).

507

508 *Figure 4: The final global RFA results output at approximately 1km resolution along the entire global coastline (excluding*
509 *Antarctica) for RP10 (A) and RP100 (B). Return levels are referenced to DEM MSL, and so represent surge, waves and tide.*
510 *Return levels (relative to mean higher high water) for 6 tide gauges in regions characterised by either positive or negative*
511 *shape parameter of the Generalised Pareto distribution are shown in panels (C) and (D) respectively. The locations of the 6*
512 *tide gauges are indicated by the diamonds plotted on both panels (A) and (B).*

513

## 4.2 Tropical Cyclone Yasi

Our second study objective is to illustrate how the RFA methodology previously described can draw on few, rare events, to provide more realistic representation of low frequency ESL exceedance probabilities across a region, using the case study of cyclone Yasi which impacted the Australian coastline in 2011. Cyclone Yasi made landfall on the North-eastern coast of Australia, in the Queensland region, between 14:00 and 15:00 UTC on the 2nd of February 2011. It is the strongest cyclone to have impacted the region since 1918, with possible windspeeds of 285km/h and minimum record pressure centre of 929 hPa (Australia Bureau of Meteorology, 2011). When it made landfall, Yasi was a category 4 storm on the Saffir-Sampson scale. The path and strength of the storm are shown in Fig. 5A.

The total water levels, relative to mean higher high water, for all the tide gauges in the region are shown in Fig. 5B. Cardwell had the highest surge, and highest total water level, by a considerable margin compared to neighbouring tide gauges, receiving a surge of over 3m above mean higher high water. Clump Point also showed a definitive but less substantial surge signal, whereas the other gauges showed much smaller surge effects or even no surge at all. The historical water level records of all the gauges in the regions are included in Fig. 5C. The tide gauges span different temporal ranges, and many have years which are incomplete. The longest record is at Townsville, which started in the late 1950s. Despite this record, the largest event is cyclone Yasi by over 1.5m (at Cardwell).

Based on this historical record, no other surge event of this magnitude has impacted this section of coastline since the records began. There are, however, records of other historic extreme events that predate tide gauges affecting the region. For example, Cyclone Mahina, which made landfall in Princess Charlotte Bay (approximately 100km north of Cooktown) in 1899, reportedly had a surge height approaching 10m (Needham *et al.*, 2015). The idea that this stretch of coastline is at risk of TC generated ESLs is further supported by STORM, a dataset of 10,000 years of synthetic hurricane tracks (Bloemendaal *et al.*, 2020). IBTrACS shows just eight category 4 and 5 hurricanes impacting this 700km stretch of coastline between 1980 and 2022 (shown in the Appendix Fig. A2; Knapp *et al.*, 2010). In contrast, the STORM dataset has 333 events affecting the area, producing a more continuous spread of landfall locations along the coastline. In addition, large surges are sometimes not captured in this region due to the lack of gauges in rural areas (Needham *et al.*, 2015).

545

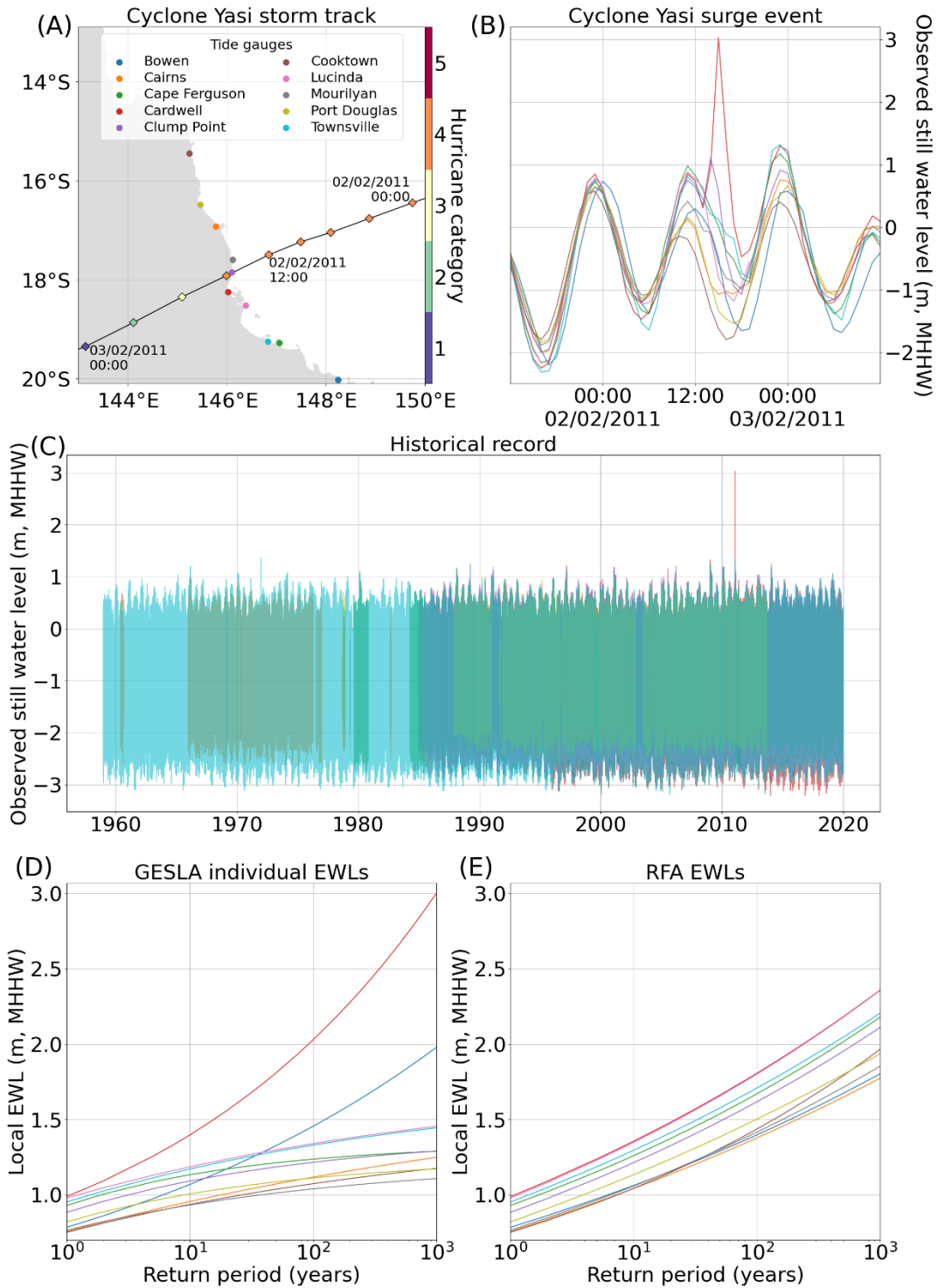



Figure 5: Tropical Cyclone Yasi: (A) The storm track of cyclone Yasi, covering a 24-hour period over the landfall event. The locations of the 10 closest tide gauges along the Queensland coast are also included. Times are in UTC. (B) The observed water level timeseries for the same 24-hour period at each of the 10 tide gauges in the region. Times are in UTC. (C)  The entire historical record of all 10 gauges in the region. (D) The return period curves of individual gauges fit with Generalised Pareto distribution. (E) The return period curves at the gauge locations from the RFA.

The return period curves, calculated by fitting a Generalised Pareto distribution to the

peaks-over-threshold water levels at each individual tide gauge, for each of the 10 gauges in

553  the region, are shown in Fig. 5D.  As expected, Cardwell has the largest return levels and the

554  steepest curve. All the other gauges, except Bowen, exhibit negative shape parameters,

555  characterised by a decreasing gradient of the return period curves. In a region which is

556  prone to TCs, this is a dangerous underestimation of the risk from cyclone induced surges. In

557  some coastal ESL studies, ESLs are calculated at each gauge, and then interpolated along the

558  coastline, such as in the UK (Environment Agency, 2018). In this case, that approach would

559  lead to a gross disparity from the actual risk of storm surges to coastal communities in the

560  area.

561  In contrast, Fig. 5E shows the return period curves estimated from the RFA at the tide gauge

562  locations. All of the curves now have positive shape parameters, characterised by increasing

563  gradients of the curves. The curves of Cardwell and Bowen have been reduced somewhat,

564  while all the other curves have been increased significantly. This demonstrates the

565  regionalisation process, by which the extreme event at Cardwell can be used to propagate

566  the risk along the coastline to areas which have not had an extreme event on record, or

567  have short, incomplete, or non-existent tide gauge records. This reinforces the key strengths

568  of the RFA, namely: (1) the ability to spatially account for rare extreme events, (2) the use of

569  short and incomplete tide gauge records to produce robust parameter fits, and (3) the

570  ability to downscale the results into regions which aren't covered by tide gauges at all.

571  ## 4.3 Comparisons with GESLA

572  The third objective is to validate ESLs calculated using our RFA against those calculated

573  directly from the measured GESLA-3 global tide gauge database. Contrasting the RFA results

574  with ESL exceedance probabilities calculated through a Generalised Pareto distribution fit at

575  individual tide gauges yields promising results. Fig. 6A shows the spatial distribution of the

576  difference at the 1-in-10-year return period for Europe, the United States, and the East

577  Pacific. In areas impacted by TCs (e.g., the Gulf of Mexico, North-Eastern Coast of Australia,

578  and Japan) we broadly see that the RFA has increasing return levels across most gauges.

579  Increases in the 1-in-10-year return level are also observed in areas usually associated with

580  extra-tropical storms (e.g., Europe), suggesting gauges in these regions also suffer from

581  under sampling of rare surge events. Extreme surge events can be undersampled for two

582  reasons. Firstly, by their very nature, they are rare and might never have occurred at a

583    specific location. Secondly, as a result of a scarcity of in-situ tide gauges, surges can occur

584    and remain unrecorded.

585    In all areas shown in Figure 6A, some gauges show decreases in the return levels. This could

586    be driven by either shape parameter limiting (to prevent unrealistically large water levels),

587    an anomalously large number of events impacting the gauge, or due to a single anomalously

588    large event impacting the gauge, which is then smoothed out through the regionalisation

589    process, as was the case in Cardwell, Australia (Fig. 5E). Of the gauges shown in the Fig. 6A,

590    only 5 had limited shape parameters, which were located in the Gulf of Mexico. The

591    distribution of the differences at RP10 is shown in Fig. 6B with a positive skew, detailing the

592    5th and 95th percentiles as -8cm and 27cm respectively. The spread of the data increases

593    across the three selected return periods (1-in-2, 1-in-10 and 1-in-100 year) presented in in

594    Fig. 6C, as well as the mean bias, which increased from 2 cm in the 1-in-2 year return level,

595    to 21cm in the 1-in-100 year return level.

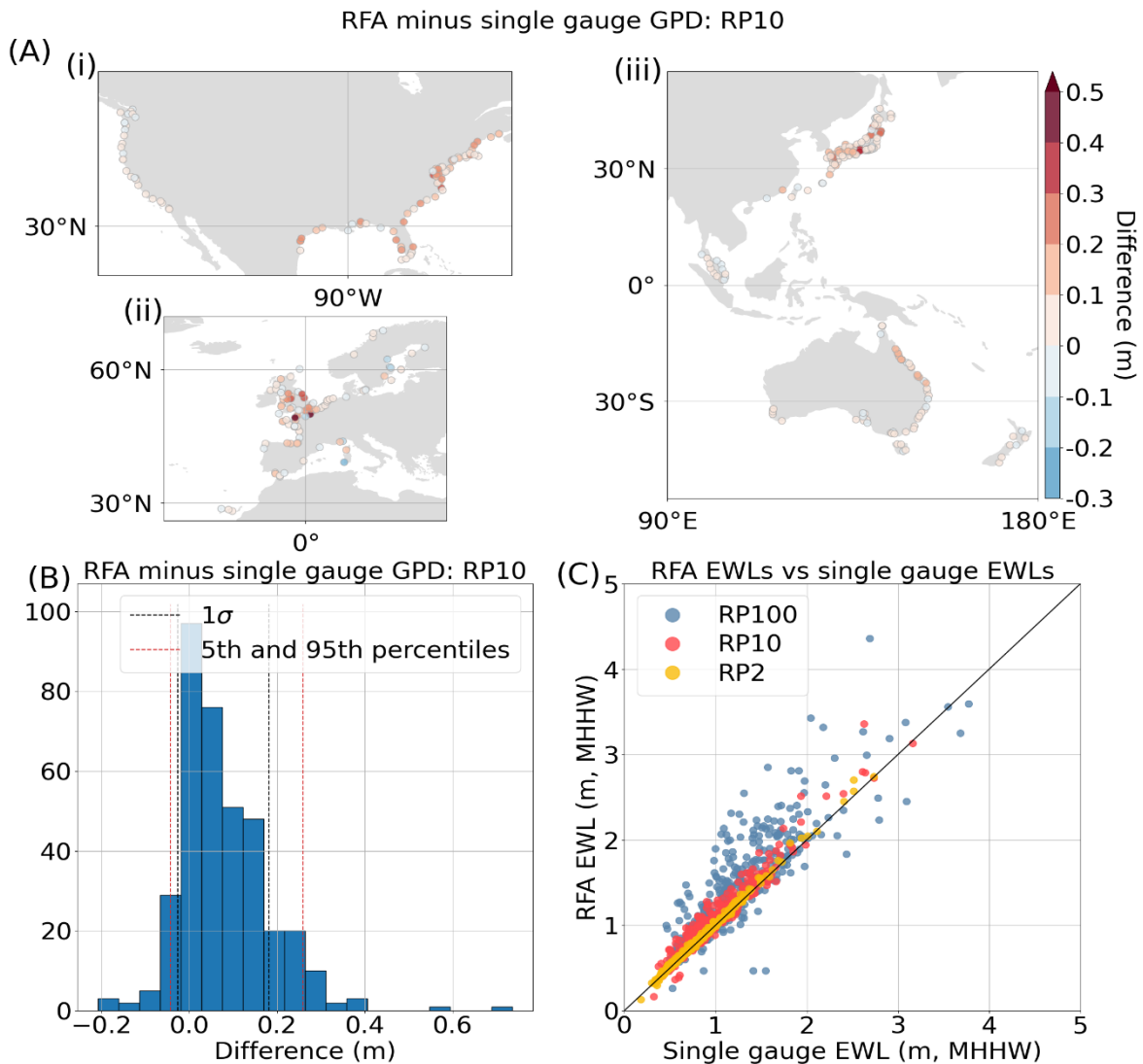



*Figure 6: Comparison of RFA water levels against extreme water levels calculated at individual gauges from GESLA by fitting a Generalised Pareto distribution to peaks-over-threshold water levels. (A) The spatial distribution of the difference at RP10 for (i) the contiguous US, (ii) Europe, (iii) Japan, Malaysia, Australia and New Zealand, (B) a histogram of the distributions of difference at RP10, including the locations of the 5th and 95th percentiles and 1 standard deviation from the mean, and (C) a scatter plot of EWLs (RP2, RP10, RP100) from the RFA and the EWLs calculated using a single site Generalised Pareto distribution fit. The black line indicates a 1:1 perfect fit.*

## 4.4 Quantifying the increases made by the RFA when compared to single site analysis

The fourth objective is to quantify the increases made to ESL exceedance probabilities in TC prone areas by the RFA, when compared to a single site analysis. Figure 7A shows the deviation in the 1-in-100-year return period between the GTSM-ERA5 RFA carried out across the global coastline, and a single site peaks-over-threshold analysis of GTSM-ERA5 water level records. Only differences greater or less than 0.25 m and -0.25 m respectively, are

611 plotted. There are evident increases to RFA ESLs in areas prone to TCs. The Gulf of Mexico,

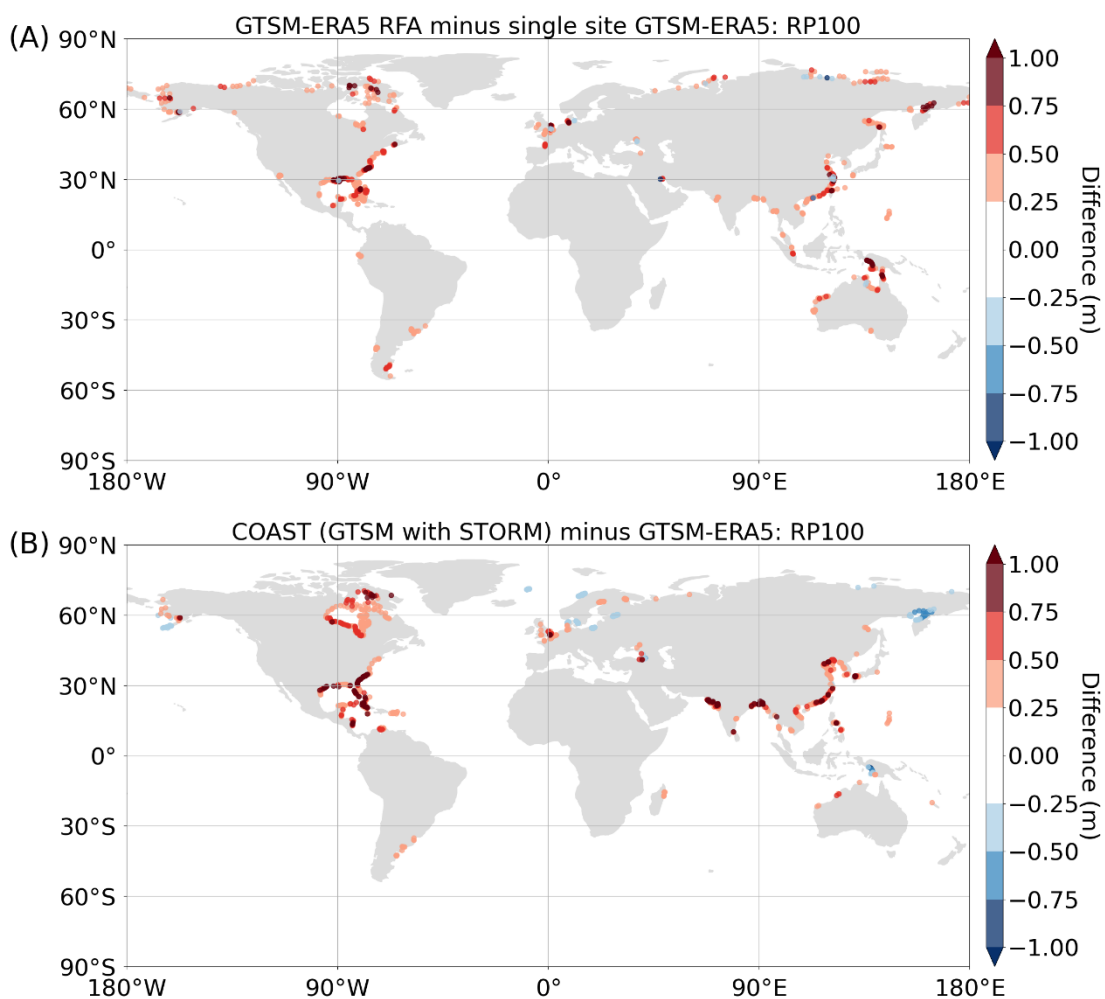
612 the East Coast of the US, Southern China, and the North-East Coast of Australia show the

613 largest increases. Sporadic negative differences are also observed in Fig. 7A, which are

614 driven by a smoothing of ESL exceedance probabilities at locations which have experienced

615 anomalously high ESL compared to the local region. From this we see that the RFA is capable

616 of incorporating the influence of TCs that were not present in the historical record, but

617 statistically could occur as indicated by the regional characteristic.

618

619 *Figure 7: The spatial distributions of: (A) the differences between the GTSM-ERA5 RFA 1-in-100-year return period (RP100)*
620 *and the RP100 of single site GTSM-ERA5 data fit with a Generalised Pareto distribution to the peaks-over-threshold water*
621 *levels; and (B) the differences in RP100 published by the COAST-RP (GTSM forced with STORM) paper (Dullaart et al., 2021)*
622 *and RP100 published by the original GTSM paper (Muis et al., 2020). Only differences greater or less than 0.25 m and -0.25*
623 *m, respectively, are plotted.*

624    These findings can be supported by the results shown in Fig. 7B, which shows the

625    differences between COAST-RP and GTSM-ERA5. COAST-RP is GTSM forced with STORM

626    (10,000 years of synthetic TCs) in areas prone to TC activity, instead of ERA5 (Dullaart *et al.*,

627    2021). The areas of positive difference highlight locations where COAST-RP is greater than

628    GTSM-ERA5, and so give an indication of the areas in which the synthetic hurricanes make

629    landfall. These patterns are broadly similar to those of the RFA, shown in Fig. 7A. However,

630    there are two areas which stand out for being poorly characterised by the RFA, namely: the

631    Bay of Bengal and the western Gujarat region of India. Large differences are also observed

632    in Hudson Bay, Canada, however we suspect these discrepancies are the result of

633    differences in the approach to modelling extra-tropical regions, as TCs do not make landfall

634    here.

635    Figure 8 shows the results of the leave-one-out cross validation of the global coastal LEWLs.

636    In general, the RFA tends to increase return levels due to the regionalisation process. These

637    findings match those of (Sweet *et al.*, 2020, Sweet *et al.*, 2022) upon which our approach is

638    based. This is evident throughout the world, with the majority of gauges exhibiting increases

639    of less than 5 cm at the 1-in-10-year return period (Fig. 8A). The central 90th percentile

640    band of the data for the 1-in-10-year return period ranges from -3 to 18 cm, as shown in Fig.

641    8B. However, the spread of the data is more pronounced at the higher return periods, as

642    shown in Fig. 8C. Some regions of the world have greater increases, in the order of 30 – 40

643    cm for the 1-in-10 year return period. These gauges are mostly concentrated in TC basins,

644    namely the Caribbean, the Gulf of Mexico, Japan, China, the Philippines, plus the East and

645    West Coasts of Australia. This demonstrates the process by which the RFA better represents

646    extreme rare events that are typically under sampled in the historical record. By drawing on

647    all the events captured by gauges across the region, the RFA reveals that there is greater risk

648    of extreme events by considering their potential occurrence in areas that, by chance, have

649    not been previously impacted as observed in historical records. Similarly, oversampling is

650    clearly evident at 1-in-100-year return periods, for which nearly a third of locations show

651    decreases in ESL exceedance probabilities compared to the single site analysis. The

652    magnitude of these decreases tend to be much smaller than the increases seen.

653

654

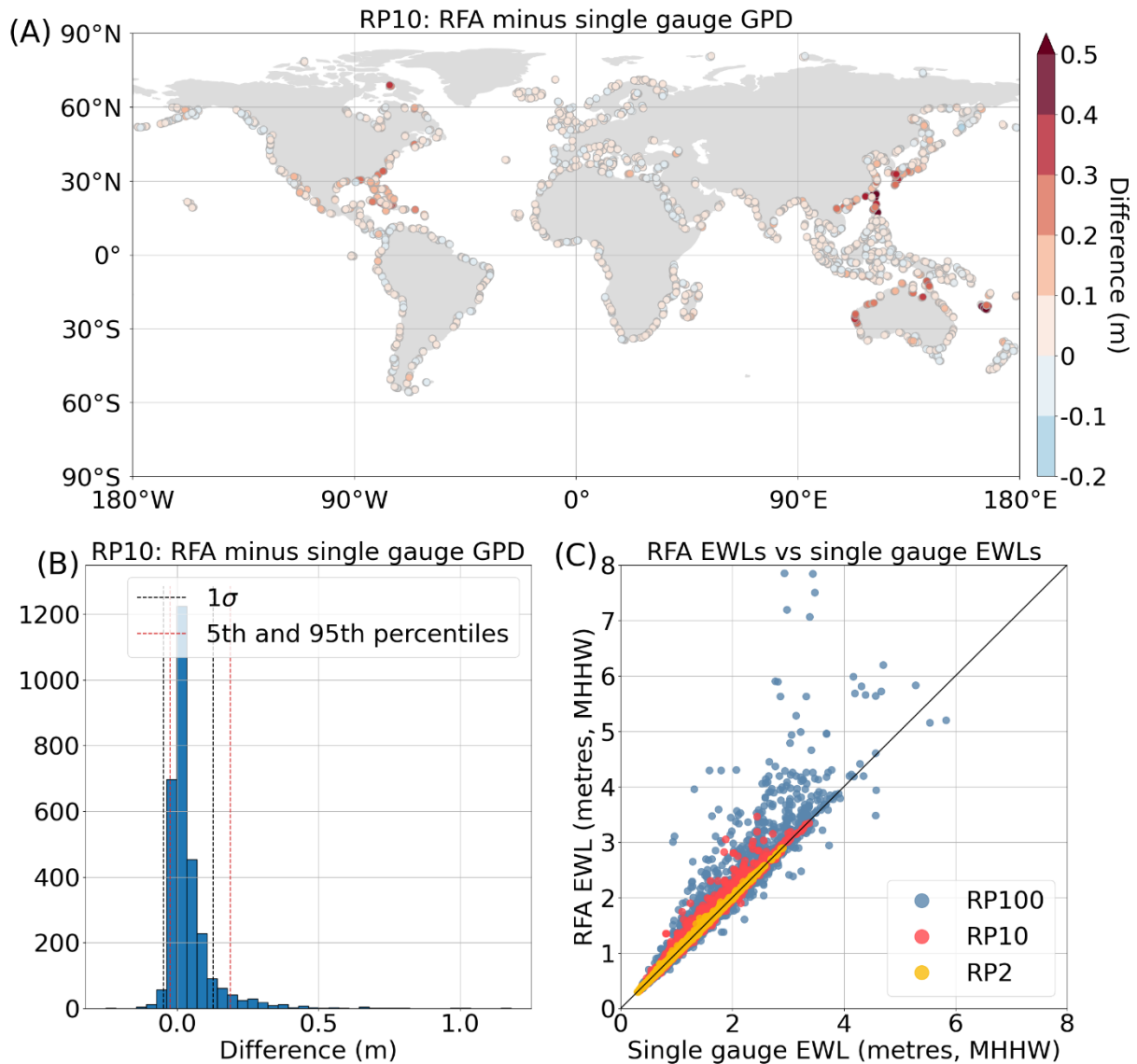



*Figure 8: The results of the leave-one-out cross validation of the RFA on GTSM-ERA5 gauges. (A) The spatial distribution of difference between the leave-one-out cross validation RFA RP10 (1 in 10-year return period) and the single site Generalised Pareto distribution RP10, (B) a histogram of the distribution of the differences in RP10 including the locations of the 5th and 95th percentiles and 1 standard deviation from the mean, and (C) a scatter plot of EWLs (RP2, RP10, and RP100) predicted using the leave-one-out cross validation RFA and the EWLs calculated using a single site Generalised Pareto distribution fit. The black line indicates a 1:1 perfect fit.*

## 5. Discussion

The ESL exceedance probabilities dataset that is presented in this paper is the first global dataset, to our knowledge, to be derived using an RFA approach, using a synthesis of observed and modelled hindcast data. The resulting data is output at high resolution (~1 km) along the entire global coastline (excluding Antarctica), includes wave setup, and better

666  captures the coastal flood risk from TCs. This approach is notable for being computationally

667  inexpensive compared to more traditional approaches for deriving ESL exceedance

668  probabilities via hydrodynamic modelling.

669  As previously discussed in the introduction section, relying solely on observational records

670  to estimate ESL exceedance probabilities can significantly bias results. To fit robust

671  parameter estimates and obtain confident exceedance probabilities sufficient for informing

672  flood risk managers, long term and consistent high quality observational records are needed

673  (Coles, 2001).  While some tide gauge and wave records span numerous decades, many

674  records only cover  a handful of recent decades (e.g., 10-30 years) or have significant gaps in

675  their historical records. This often means quality data is excluded from analyses as their

676  records are too short to produce robust parameter estimates. Furthermore, gauges are

677  relatively sparse, especially in less populated areas and developing nations. While surges

678  and waves typically impact large regions, peak water levels are usually only observed over

679  smaller areas (i.e., a single bay, estuary or beach). As a result, measured records can easily

680  miss the maximum of an extreme event, thus mischaracterising extreme water levels of the

681  event. As such, rare extreme events that characterise the upmost tails of the distributions of

682  ESLs, such as TCs, are repeatedly under sampled in the historic record, in both frequency

683  and magnitude.

684  By using an RFA approach, we demonstrate how we have improved these issues. The RFA

685  can be viewed as a space-for-time approach, where long historical records (which give

686  robust parameter estimates) are substituted for a collection of shorter records that cover a

687  larger area. The volume of data (and subsequent extreme events) is retained, but the

688  individual records can be much shorter. In this study, records as short as 10 years have been

689  utilised. Furthermore, the regionalisation process works to overcome the issues with gauge

690  density by disseminating the hazard presented by rare extreme events, as shown using the

691  Cyclone Yasi example. From the 10 gauges in the region, the only record to have captured

692  an historic extreme surge event of the magnitude observed during Cyclone Yasi was

693  Cardwell, despite this section of coastline being at known risk to TC activity. A single site

694  analysis of tide gauge data in this region would likely underpredict the real risk of ESLs

695  generated by TCs in areas which haven't had a direct impact in the observational record. On

696    the other hand, the damping of the return levels in the RFA output at Cardwell and Bowen

697    could mean an underprediction of the risk from surges in these locations.

698    Global hydrodynamic models that simulate tide and surge (e.g., GTSM) or waves have been

699    developed to substitute observational records, especially in regions not covered by tide

700    gauges. These models have been demonstrated to represent historic extreme events to a

701    high degree of accuracy when forced using historical observational data pertaining to the

702    event (Yang *et al.*, 2020). However, using these models for the characterisation of

703    exceedance probabilities is limited by the availability of long term high-quality global

704    reanalysis data, that captures the full extent of meteorological extremes that drive large

705    surge events. The RFA is aims to address this by using a space-for-time approach, however it

706    is still limited by the bounds of the GTSM-ERA5 data. As demonstrated in Fig. 7, the

707    distribution of increases to local return levels made by the RFA broadly follows the same

708    patterns globally as the differences between COAST-RP and GTSM-ERA5. As TC hazard is

709    typically underrepresented due to short records, it can be inferred that the increases

710    observed across these regions are an improvement on a single site analysis.

711    While the RFA is capable of identifying areas of increased risk from TC activity, it is still

712    constrained by the training data available. This is demonstrated in Fig. 7. Two distinct areas

713    lack increased water levels in the RFA difference plot (Fig. 7A), namely: the Bay of Bengal

714    and Northwestern coasts of India and Pakistan. ERA5, the forcing data used for GTSM-ERA5

715    has been found to consistently underestimate TC intensity in both minimum sea level

716    pressure and maximum windspeed (Dulac *et al.*, 2023). Consequently, the intensity of

717    extreme events in GTSM-ERA5 in these regions could underrepresent the potential hazard

718    from TC activity. If the maximums of extremes are not captured in the reanalysis data, then

719    the full magnitude of the surge cannot be simulated by GTSM-ERA5. As such, the RFA will

720    have smaller or fewer extremes with which to draw data from when characterising rare

721    extreme events, therefore leading to a persistent underestimation of the return levels.

722    Coastal flood hazard mapping is usually carried out using inundation models that simulate

723    the propagation of water over the coastal floodplain. To accurately capture the footprint of

724    the surge on the land, inundation models require high-resolution boundary conditions at

725    regular intervals along the coastline. The density of boundary condition points needs to be

726    sufficient to capture local variability in ESLs along a coastline, which can be caused by

727    bathymetric and topographic features such as narrow channels, enclosed bays, barrier

728    island and estuaries. The spatial resolution of tide gauges, even in the areas of highest gauge

729    density, is insufficient for direct use in inundation modelling and therefore requires some

730    form of interpolation and/or extrapolation. Similarly, while GTSM-ERA5, is run at a

731    reasonably high coastal resolution, publicly available data is only output at approximately

732    50km resolution outside of Europe, and therefore does not meet the standards necessary

733    for coastal floodplain inundation modelling. Using the RFA to downscale the regional

734    extreme water levels allows for the possibility of implementing tide gauge data and the

735    outputs from GTSM-ERA5 as boundary conditions for subsequent inundation models. In

736    addition, the downscaling process involves scaling the water levels by tidal range and thus

737    enables dynamic characteristics of the surge, such as amplification at the head of estuaries,

738    to be reproduced in the inundation models. This downscaling process is, however, limited

739    by the resolution of the tide model used to obtain the tidal range values. In the case of this

740    study, FES2014 is output at 1/16$^{th}$ of a degree (approximately 7km at the equator).

741    Ultimately, the future of delineating the flood hazard from TCs lies in multi-ensemble

742    models using 100's of 1,000's of years' worth of synthetically generated storms forcing high-

743    resolution tide-surge-wave models. However, the computational cost of running such

744    simulations is enormous when compared to the cost of running an RFA on a relatively short

745    hindcast record. In the same way, dynamically modelled waves are usually excluded from

746    global simulations that consider exceedance probabilities due to the computational

747    expense. At the same time, failing to considering the joint dependence of surge and waves

748    can lead to an underestimation of ESL exceedance levels by up to a factor of two along 30%

749    of the global coastline (Marcos *et al.*, 2019).  This reinforces the significance of the RFA

750    methodology in characterising global coastal flood risk.

751    Validating the RFA is nuanced, as assessing metrics compared with observed record is: (a)

752    validating against the data used to build the RFA in the first place; and (b) not recognising

753    the inadequacies of the tide gauge records that the RFA is attempting to mitigate. Leave-

754    one-out cross validation highlights the strengths of the RFA, without succumbing to the

755    shortfalls inherent in the observational record. The increased LEWLs in the regions prone to

756    TC activity once again demonstrates the RFA's ability to spatially disperse the hazard of low

757    probability extreme events across a region. It is worth noting that the leave-one-out cross

758     validation is the best possible representation of the RFA as only grid cells that use data from

759     10 record locations are used, so each model is trained on the maximum amount of data

760     possible. In some areas, the number of records used can be as low as three, and so the

761     ability for the RFA to reproduce water levels in these regions could be compromised.

762     Applying the RFA as done in this study does have its limitations. Firstly, changing our

763     definition of a homogeneous region would likely have a great impact on our results. In

764     future iterations of this study, we recommend carrying out a sensitivity analysis to

765     understand how using different maximum radii to select water level records impacts upon

766     estimated extreme water levels within the region. Secondly, delineating the global coastline

767     into 1° by 1° tiles and evaluating a different RFA for each tile results in some complex areas

768     of coastline being summarised by a single regional growth function. Examples of this are

769     seen in Japan, where exposed coastlines of the North Coast are contained in the same tile as

770     a sheltered bay that is open to the South Coast. A solution to this would be to classify

771     coastlines based on descriptors, as carried out by Sweet *et al.* (2020). These descriptors

772     could include characteristics such as dominant forcing type, geographic location, and/or

773     local coastal dynamics. The method used to incorporate wave setup is another constraint, as

774     it has been greatly simplified for ease of global application. Improving upon this should also

775     be a focus of future studies. Lastly, another limitation of the approach used in this study is

776     the static shape parameter limiter. It is probable that the maximum shape parameter varies

777     by location around the world, and that by implementing a fixed threshold globally we are

778     perhaps limiting some of the most extreme events in some regions. Improving this section

779     of the methodology is a high priority for future updates.

780     The outputs from the RFA should be supplemented with local knowledge wherever possible,

781     and the uncertainties in the results should be considered before the data is used. The RFA is

782     a powerful tool for estimating return levels in ungauged locations or in locations where the

783     historical records are short or incomplete, but there are risks associated with both

784     overpredicting and underpredicting surge heights. Underprediction can lead to complacency

785     among coastal managers and the potentially dangerous assumption that communities are

786     safe from surge risk. Conversely, overprediction can result in unnecessary cost for risk

787     mitigation measures and potential economic loss driven by a lack of investment in a region

788     deemed at risk. Disseminating the risk of TC generated surges over a region could lead to

789 overprediction in some locations, and so conducting sensitivity analyses to understand the

790 robustness of findings is recommended, especially in the context of coastal management

791 and safety assessments. The RFA has been developed in this study as a method for regional

792 to continental to global scale risk analyses from globally available data, and not local

793 studies. The results give a first order approximation of extreme water levels in ungauged

794 locations. It is not expected that they would be used in the design for local flood defences,

795 for example.

796 Going forward, the RFA framework developed in this study can be easily updated with the

797 availability of new data. Possible next steps could also include using GTSM simulations of

798 future climate scenarios, as well as measured wave data. To this end, a global wave dataset

799 similar to GESLA would be instrumental in collating wave data from the numerous buoys

800 globally. Future updates could also include an assessment of using different extreme value

801 distributions, perhaps following the mixed climate approach of O'Grady *et al.*, (2022).

802 In the near future, we plan to use the global exceedance probabilities derived in this paper

803 as boundary conditions for inundation modelling of the coastal floodplain of the entire

804 globe, using the 2D hydraulic model LISFLOOD-FP (Bates *et al.*, 2010). This presents an

805 exciting opportunity to provide an invaluable resource that will help to better quantify

806 global coastal flood risk.

807

## 6. Conclusions

809 In this paper we have demonstrated an RFA approach utilising both measured and modelled

810 hindcast records to estimate ESL exceedance probabilities, including wave setup, at high

811 resolution (~1 km) along the entire global coastline (with the exception of Antarctica). Our

812 methodology is computationally inexpensive and is more effective in accurately estimating

813 the low frequency exceedance probabilities that are associated with rare extreme events,

814 compared to approaches that consider data from single sites. We have demonstrated, using

815 Cyclone Yasi (2011) which impacted the Australia coast, the ability of the RFA to better

816 characterise ESLs in regions prone to TC activity. Furthermore, on the global scale we have

817 exemplified how the RFA, when trained on relatively short reanalysis data, can reproduce

818 patterns of increased water levels similar to those present in dynamic simulations of 10,000

819 years of synthetic hurricane tracks. The RFA methodology shown provides a promising

820 avenue for improving our understanding of coastal flooding and enhancing our ability to

821 prepare for and mitigate its devastating impacts. In the future, we plan to use the

822 exceedance probabilities from this study as boundary conditions for an inundation model

823 covering the global coastal floodplain.

824

## 7. References

Amadeo, K. (2019) Hurricane Harvey Facts , Damage and Costs. Retrieved from
https://www.lamar.edu/_files/documents/resilience-recovery/grant/recovery-and-
resiliency/hurric2.pdf

Andrée, E., Su, J., Larsen, M. A. D., Madsen, K. S. & Drews, M. (2021) Simulating major storm surge
events in a complex coastal region. *Ocean Model.* **162**. doi:10.1016/j.ocemod.2021.101802

Andreevsky, M., Hamdi, Y., Griolet, S., Bernardara, P. & Frau, R. (2020) Regional frequency analysis of
extreme storm surges using the extremogram approach. *Nat. Hazards Earth Syst. Sci.* **20**(6).
doi:10.5194/nhess-20-1705-2020

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 Dyn.* **65**(4). doi:10.1007/s10236-015-
0814-1

Australia Bureau of Meteorology. (2011) Severe Tropical Cyclone Yasi. Retrieved from
http://www.bom.gov.au/cyclone/history/yasi.shtml

Bardet, L., Duluc, C. M., Rebour, V. & L'Her, J. (2011) Regional frequency analysis of extreme storm
surges along the French coast. *Nat. Hazards Earth Syst. Sci.* **11**(6). doi:10.5194/nhess-11-1627-
2011

Barnard, P. L., Erikson, L. H., Foxgrover, A. C., Hart, J. A. F., Limber, P., O'Neill, A. C., Ormondt, M.
van, et al. (2019) Dynamic flood modeling essential to assess the coastal impacts of climate
change. *Sci. Rep.* **9**(1). doi:10.1038/s41598-019-40742-z

Bates, P. D., Horritt, M. S. & Fewtrell, T. J. (2010) A simple inertial formulation of the shallow water
equations for efficient two-dimensional flood inundation modelling. *J. Hydrol.* **387**(1–2).
doi:10.1016/j.jhydrol.2010.03.027

Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., Savage, J., et al. (2021) Combined
Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates.
*Water Resour. Res.* **57**(2). doi:10.1029/2020WR028673

Bingham, R. J. & Haines, K. (2006) Mean dynamic topography: Intercomparisons and errors. *Philos.
Trans. R. Soc. A Math. Phys. Eng. Sci.* **364**(1841). doi:10.1098/rsta.2006.1745

Bloemendaal, N., Haigh, I. D., Moel, H. de, Muis, S., Haarsma, R. J. & Aerts, J. C. J. H. (2020)
Generation of a global synthetic tropical cyclone hazard dataset using STORM. *Sci. Data* **7**(1).
doi:10.1038/s41597-020-0381-2

Calafat, F. M., Wahl, T., Tadesse, M. G. & Sparrow, S. N. (2022) Trends in Europe storm surge
extremes match the rate of sea-level rise. *Nature* **603**(7903). doi:10.1038/s41586-022-04426-5

Caldwell, P. C., Merrifield, M. A. & Thompson, P. R. (2015) Sea level measured by tide gauges from
global oceans — the Joint Archive for Sea Level holdings (NCEI Accession 0019568), Version 5.5.
*NOAA Natl. Centers Environ. Inf.* (18).

Campos, R. M., Guedes Soares, C., Alves, J. H. G. M., Parente, C. E. & Guimaraes, L. G. (2019)
Regional long-term extreme wave analysis using hindcast data from the South Atlantic Ocean.
*Ocean Eng.* **179**. doi:10.1016/j.oceaneng.2019.03.023

Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*. Bristol: Springer.

866 Dean, R. & Walton, T. (2010) Wave Setup. In: *Handbook of Coastal and Ocean Engineering*, 1st ed.,
867      Vol. 1–2, 1–24. World Scientific Publishing co. doi:10.1142/10353

868 Dulac, W., Cattiaux, J., Chauvin, F., Bourdin, S. & Fromang, S. (2023) Assessing the representation of
869      tropical cyclones in ERA5 with the CNRM tracker. *Clim. Dyn.* doi:10.1007/s00382-023-06902-8

870 Dullaart, J. C. M., Muis, S., Bloemendaal, N., Chertova, M. V., Couasnon, A. & Aerts, J. C. J. H. (2021)
871      Accounting for tropical cyclones more than doubles the global population exposed to low-
872      probability coastal flooding. *Commun. Earth Environ.* **2**(1). doi:10.1038/s43247-021-00204-9

873 Environment Agency. (2018) Coastal flood boundary conditions for the UK: 2018 update. Retrieved
874      from https://www.gov.uk/government/publications/coastal-flood-boundary-conditions-for-uk-
875      mainland-and-islands-design-sea-levels

876 European Space Agency. (2021) Copernicus Global Digital Elevation Model. *Open Topol.* 1.
877      doi:doi.org/10.5069/G9028PQB

878 Fanti, V., Ferreira, Ó., Kümmerer, V. & Loureiro, C. (2023) Improved estimates of extreme wave
879      conditions in coastal areas from calibrated global reanalyses. *Commun. Earth Environ.* **4**(1).
880      doi:10.1038/s43247-023-00819-0

881 Frau, R., Andreewsky, M. & Bernardara, P. (2018) The use of historical information for regional
882      frequency analysis of extreme skew surge. *Nat. Hazards Earth Syst. Sci.* **18**(3).
883      doi:10.5194/nhess-18-949-2018

884 Haigh, I D, Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Hague, B. S., Bradshaw, E., et al.
885      (2021) GESLA Version 3: Amajor update to the global higher-frequency sea-level dataset.
886      *EarthArXiv* (released), 1–34. Retrieved from https://www.gesla.org.

887 Haigh, Ivan D., MacPherson, L. R., Mason, M. S., Wijeratne, E. M. S., Pattiaratchi, C. B., Crompton, R.
888      P. & George, S. (2014) Estimating present day extreme water level exceedance probabilities
889      around the coastline of Australia: Tropical cyclone-induced storm surges. *Clim. Dyn.* **42**(1–2).
890      doi:10.1007/s00382-012-1653-0

891 Haigh, Ivan D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., Horsburgh, K., et al.
892      (2016) Spatial and temporal analysis of extreme sea level and storm surge events around the
893      coastline of the UK. *Sci. Data*. doi:10.1038/sdata.2016.107

894 Hall, J. A., Gill, S., Obeysekera, J., Sweet, W., Knuuti, K. & Marburger, J. (2016) Regional Sea Level
895      Scenarios for Coastal Risk Management : Managing the Uncertainty of Future Sea Level Change
896      and Extreme Water Levels for Department of Defense Coastal Sites Worldwide (April), 224.
897      Retrieved from https://apps.dtic.mil/sti/citations/AD1013613

898 Hamdi, Y., Duluc, C. M., Bardet, L. & Rebour, V. (2016) Use of the spatial extremogram to form a
899      homogeneous region centered on a target site for the regional frequency analysis of extreme
900      storm surges. *Int. J. Saf. Secur. Eng.* **6**(4). doi:10.2495/SAFE-V6-N4-777-781

901 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., et al.
902      (2020) The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**(730), 1999–2049. John Wiley &
903      Sons, Ltd. doi:https://doi.org/10.1002/qj.3803

904 Hosking, J. R. M. & Wallis, J. R. (1997) *Regional Frequency Analysis: An approach based on L-
905      moments, Cambridge University Press, New York.* Cambridge Universtiy Press.
906      doi:https://doi.org/10.1017/CBO9780511529443

907 India Meteorological Department. (2020) Super Cyclonic Storm Amphan over the southeast Bay of
908      Bengal: Summary. Retrieved from

909        https://internal.imd.gov.in/press_release/20200614_pr_840.pdf

910 Irish, J. L., Resio, D. T. & Ratcliff, J. J. (2008) The influence of storm size on hurricane surge. *J. Phys.*
911        *Oceanogr.* **38**(9). doi:10.1175/2008JPO3727.1

912 James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013) *An Introduction to Statistical Learning*.
913        Springer. doi:10.1007/9781461471387

914 Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J. & Neumann, C. J. (2010) The international
915        best track archive for climate stewardship (IBTrACS). *Bull. Am. Meteorol. Soc.* **91**(3).
916        doi:10.1175/2009BAMS2755.1

917 Kumar, S., Lal, P. & Kumar, A. (2021) Influence of Super Cyclone "Amphan" in the Indian
918        Subcontinent amid COVID-19 Pandemic. *Remote Sens. Earth Syst. Sci.* **4**(1–2).
919        doi:10.1007/s41976-021-00048-z

920 Liang, B., Gao, H. & Shao, Z. (2019) Characteristics of global waves based on the third-generation
921        wave model SWAN. *Mar. Struct.* **64**. doi:10.1016/j.marstruc.2018.10.011

922 Lucas, C., Muraleedharan, G. & Guedes Soares, C. (2017) Regional frequency analysis of extreme
923        waves in a coastal area. *Coast. Eng.* **126**. doi:10.1016/j.coastaleng.2017.06.002

924 Lyard, F. H., Allain, D. J., Cancet, M., Carrère, L. & Picot, N. (2021) FES2014 global ocean tide atlas:
925        Design and performance. *Ocean Sci.* **17**(3), 615–649. doi:10.5194/os-17-615-2021

926 Marcos, M., Rohmer, J., Vousdoukas, M. I., Mentaschi, L., Cozannet, G. Le & Amores, A. (2019)
927        Increased Extreme Coastal Water Levels Due to the Combined Action of Storm Surges and
928        Wind Waves. *Geophys. Res. Lett.* **46**(8). doi:10.1029/2019GL082599

929 McGranahan, G., Balk, D. & Anderson, B. (2007) The rising tide: Assessing the risks of climate change
930        and human settlements in low elevation coastal zones. *Environ. Urban.* **19**(1).
931        doi:10.1177/0956247807076960

932 Mitchell, D., Hawker, L., Savage, J., Bingham, R., Lord, N. S., Khan, M. J. U., Bates, P., et al. (2022)
933        Increased population exposure to Amphan-scale cyclones under future climates. *Clim. Resil.*
934        *Sustain.* **1**(2), 1–16. doi:10.1002/cli2.36

935 Mori, N., Yasuda, T., Arikawa, T., Kataoka, T., Nakajo, S., Suzuki, K., Yamanaka, Y., et al. (2019) 2018
936        Typhoon Jebi post-event survey of coastal damage in the Kansai region, Japan. *Coast. Eng. J.*
937        **61**(3). doi:10.1080/21664250.2019.1619253

938 Muis, S., Apecechea, M. I., Dullaart, J., Lima Rego, J. de, Madsen, K. S., Su, J., Yan, K., et al. (2020) A
939        High-Resolution Global Dataset of Extreme Sea Levels, Tides, and Storm Surges, Including
940        Future Projections. *Front. Mar. Sci.* **7**(April), 1–15. doi:10.3389/fmars.2020.00263

941 Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H. & Ward, P. J. (2016) A global reanalysis of
942        storm surges and extreme sea levels. *Nat. Commun.* **7**(1), 11969. doi:10.1038/ncomms11969

943 Mulet, S., Rio, M. H., Etienne, H., Artana, C., Cancet, M., Dibarboure, G., Feng, H., et al. (2021) The
944        new CNES-CLS18 global mean dynamic topography. *Ocean Sci.* **17**(3), 789–808. doi:10.5194/os-
945        17-789-2021

946 Needham, H. F., Keim, B. D. & Sathiaraj, D. (2015) A review of tropical cyclone-generated storm
947        surges: Global data sources, observations, and impacts. *Rev. Geophys.* **53**(2), 545–591.
948        doi:10.1002/2014RG000477

949 Neumann, B., Vafeidis, A. T., Zimmermann, J. & Nicholls, R. J. (2015) Future coastal population
950        growth and exposure to sea-level rise and coastal flooding - A global assessment. *PLoS One*

951    **10**(3). doi:10.1371/journal.pone.0118571

952    Nicholls, R. J., Lincke, D., Hinkel, J., Brown, S., Vafeidis, A. T., Meyssignac, B., Hanson, S. E., et al.
953        (2021) A global analysis of subsidence, relative sea-level change and coastal flood exposure.
954        *Nat. Clim. Chang.* **11**(4). doi:10.1038/s41558-021-00993-z

955    O'Grady, J. G., Stephenson, A. G. & McInnes, K. L. (2022) Gauging mixed climate extreme value
956        distributions in tropical cyclone regions. *Sci. Rep.* **12**(1). doi:10.1038/s41598-022-08382-y

957    Ramakrishnan, R., Remya, P. G., Mandal, A., Mohanty, P., Arayakandy, P., Mahendra, R. S. & Nair, T.
958        M. B. (2022) Wave induced coastal flooding along the southwest coast of India during tropical
959        cyclone Tauktae. *Sci. Rep.* **12**(1). doi:10.1038/s41598-022-24557-z

960    Shaji, C., Kar, S. K. & Vishal, T. (2014) Storm surge studies in the North Indian Ocean: A review. *Indian
961        J. Mar. Sci.* **43**(2).

962    Siahsarani, A., Karami Khaniki, A., Aliakbari Bidokhti, A. A. & Azadi, M. (2021) Numerical Modeling of
963        Tropical Cyclone-Induced Storm Surge in the Gulf of Oman Using a Storm Surge–Wave–Tide
964        Coupled Model. *Ocean Sci. J.* **56**(3). doi:10.1007/s12601-021-00027-x

965    Slocum, C. J., Razin, M. N., Knaff, J. A. & Stow, J. P. (2022) Does ERA5 mark a new era for resolving
966        the tropical cyclone environment? *J. Clim.* 1–39. doi:10.1175/jcli-d-22-0127.1

967    Sweet, W. V., Hamlington, B. D., Kopp, R. E., Weaver, C. P., Barnard, P. L., Bekaert, D., Brooks, W., et
968        al. (2022) Global and Regional Sea Level Rise Scenarios for the United States. *NOAA Tech. Rep.*
969        *NOS 01* 111 pp. Retrieved from https://oceanservice.noaa.gov/hazards/sealevelrise/noaa-nos-

970    Sweet, W. V, Genz, A. S., Obeysekera, J. & Marra, J. J. (2020) A Regional Frequency Analysis of Tide
971        Gauges to Assess Pacific Coast Flood Risk   . *Front. Mar. Sci.*   . Retrieved from
972        https://www.frontiersin.org/articles/10.3389/fmars.2020.581769

973    Taherkhani, M., Vitousek, S., Barnard, P. L., Frazer, N., Anderson, T. R. & Fletcher, C. H. (2020) Sea-
974        level rise exponentially increases coastal flood frequency. *Sci. Rep.* **10**(1). doi:10.1038/s41598-
975        020-62188-4

976    Tanim, A. H. & Akter, A. (2019) Storm-surge modelling for cyclone Mora in the northern Bay of
977        Bengal. *Proc. Inst. Civ. Eng. Marit. Eng.* **172**(3). doi:10.1680/jmaen.2019.1

978    Timmermans, B. W., Gommenginger, C. P., Dodet, G. & Bidlot, J. R. (2020) Global Wave Height
979        Trends and Variability from New Multimission Satellite Altimeter Products, Reanalyses, and
980        Wave Buoys. *Geophys. Res. Lett.* **47**(9). doi:10.1029/2019GL086880

981    Vanem, E. (2017) A regional extreme value analysis of ocean waves in a changing climate. *Ocean
982        Eng.* **144**. doi:10.1016/j.oceaneng.2017.08.027

983    Vousdoukas, M. I., Voukouvalas, E., Mentaschi, L., Dottori, F., Giardino, A., Bouziotas, D., Bianchi, A.,
984        et al. (2016) Developments in large-scale coastal flood hazard mapping. *Nat. Hazards Earth
985        Syst. Sci.* **16**(8). doi:10.5194/nhess-16-1841-2016

986    Wang, J. & Wang, Y. (2022) Evaluation of the ERA5 Significant Wave Height against NDBC Buoy Data
987        from 1979 to 2019. *Mar. Geod.* **45**(2). doi:10.1080/01490419.2021.2011502

988    Weiss, J. & Bernardara, P. (2013) Comparison of local indices for regional frequency analysis with an
989        application to extreme skew surges. *Water Resour. Res.* **49**(5). doi:10.1002/wrcr.20225

990    Yang, Z., Wang, T., Castrucci, L. & Miller, I. (2020) Modeling assessment of storm surge in the Salish
991        Sea. *Estuar. Coast. Shelf Sci.* **238**. doi:10.1016/j.ecss.2019.106552

# 8. Appendix

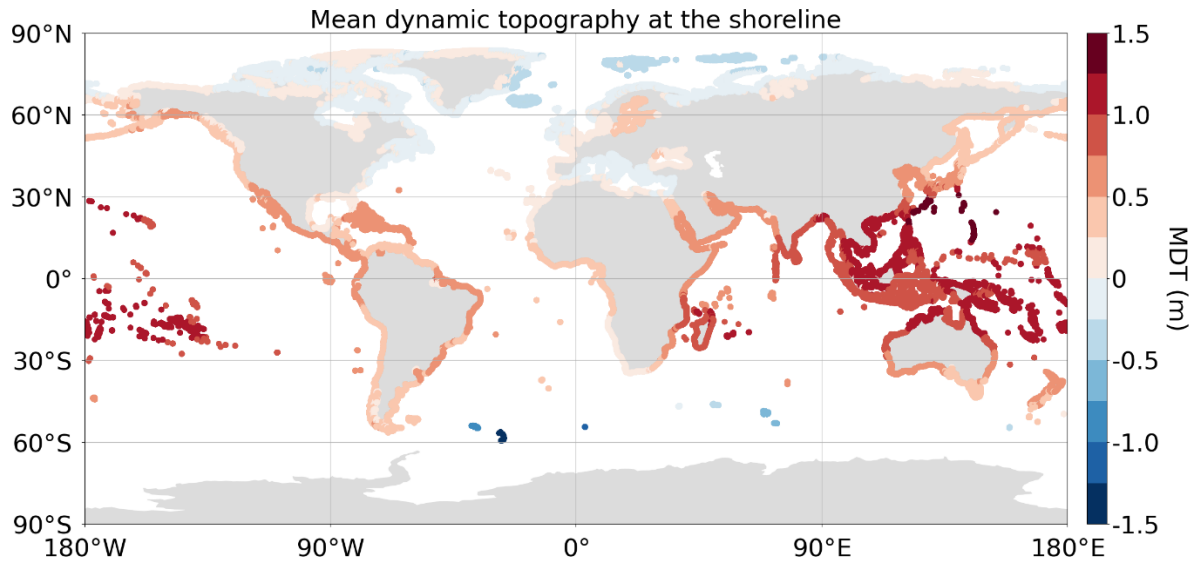
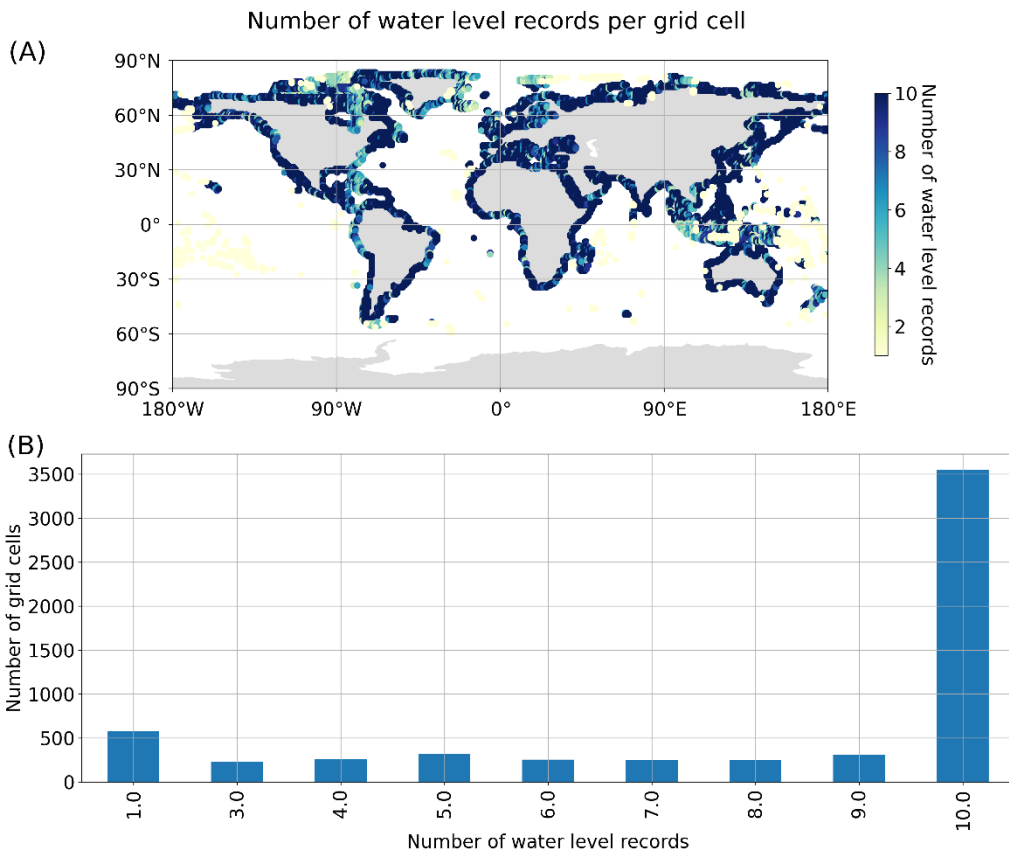



*Figure A1: HYBRID-CNES-CLS18-CMEMS2020 MDT dataset from Mulet et al., (2021), extracted at the shoreline for use in correcting the output from the RFA for future uses such as inundation modelling.*

*Figure A2: (A) Category 4 and 5 IBTrACS hurricane impacting the Queensland coastline between 1980-2022 (Knapp et al., 2010) and (B) equivalent STORM events impacting the same the stretch of coastline (Bloemendaal et al., 2020).*

1002

1003 *Figure A3: The number of water level records used per grid cell (A) as a scatter plot showing the distribution globally, and*
1004 *(B) as a bar plot showing the number of water level records vs the number of grid cells.*

1005

# 9. Code Availability

1007 The Python scripts used for handling the GESLA dataset can be downloaded for:
1008 https://github.com/philiprt/GeslaDataset

1009 The Conda package (Python) used for creating the FES2014 tidal timeseries can found at:
1010 https://anaconda.org/fbriol/pyfes

# 10.    Data availability

1012 GESLA tide gauge data is available at: https://gesla787883612.wordpress.com/downloads/

1013 GTSM data is available at: https://cds.climate.copernicus.eu/cdsapp#!/dataset/sis-water-level-
1014 change-timeseries?tab=overview

1015 ERA5 wave hindcast data is available at:
1016 https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview

1017 FES2014 tidal heights can be downloaded from:

1018 https://www.aviso.altimetry.fr/en/data/products/auxiliary-products/global-tide-fes.html

1019 HYBRID-CNES-CLS18-CMEMS2020 is available at:

1020 https://www.aviso.altimetry.fr/en/data/products/auxiliary-products/mdt/mdt-global-hybrid-cnes-
1021 cls-cmems.html

1022 Copernicus 30m DEM is found at: https://spacedata.copernicus.eu/collections/copernicus-digital-
1023 elevation-model

1024 COAST-RP dataset is downloaded from: https://data.4tu.nl/articles/_/13392314

1025 The data produced in this study is available for academic, non-commercial research only. Please
1026 contact the corresponding author for access.

## 11. Author contributions

1028 T.C. was responsible for coding up the pre-processing the tide gauge and GTSM data, coding up the
1029 RFA and validating the results. N.Q. pre-processed the wave data, including fitting the copula to
1030 predict wave conditions for tide gauge records that extended beyond the hindcast period.  J.G.
1031 created the coastline output points using the Copernicus DEM. I.P. worked on the evaluating the
1032 empirical shape parameter limiter. H.W. assisted in validating the output results from the RFA. S.M.
1033 supplied the GTSM dataset and W.S. provided the RFA methodology which we applied globally. I.H.
1034 and P.B. provided guidance and assistance throughout. T.C. prepared the manuscript with
1035 contributions and editing from all co-authors.

## 12. Competing Interests

1037 The authors declare that they have no conflict of interest.