Grey swan tropical cyclones

Ning Lin[1]* and Kerry Emanuel[2]

We define 'grey swan' tropical cyclones as high-impact storms that would not be predicted based on history but may be foreseeable using physical knowledge together with historical data. Here we apply a climatological–hydrodynamic method to estimate grey swan tropical cyclone storm surge threat for three highly vulnerable coastal regions. We identify a potentially large risk in the Persian Gulf, where tropical cyclones have never been recorded, and larger-than-expected threats in Cairns, Australia, and Tampa, Florida. Grey swan tropical cyclones striking Tampa, Cairns and Dubai can generate storm surges of about 6 m, 5.7 m and 4 m, respectively, with estimated annual exceedance probabilities of about 1/10,000. With climate change, these probabilities can increase significantly over the twenty-first century (to 1/3,100–1/1,100 in the middle and 1/2,500–1/700 towards the end of the century for Tampa). Worse grey swan tropical cyclones, inducing surges exceeding 11 m in Tampa and 7 m in Dubai, are also revealed with non-negligible probabilities, especially towards the end of the century.

The term 'black swan'[1,2] is a metaphor for a high-consequence event that comes as a surprise. Some high-consequence events that are unobserved and unanticipated may nevertheless be predictable (although perhaps with large uncertainty); such events may be referred to as 'grey swans'[3,4] (or, sometimes, 'perfect storms'[5]). Unlike truly unpredicted and unavoidable black swans, which can be dealt with only by fast reaction and recovery, grey swans—although also novel and outside experience—can be better foreseen and systematically prepared for[4,5].

Tropical cyclones (TCs) often produce extreme wind, rainfall and storm surges in coastal areas. Storm surges are especially complex functions of TC characteristics (track, intensity and size) and coastal features (geometry and bathymetry), and they are also the most fatal and destructive aspect of TCs (see ref. 6 for a comprehensive review of global TC surge observations and impacts). Hence, storm surge is an appropriate and practical metric for identifying grey swan TCs. The most infamous TC disasters early this century were attributable to storm surges, but they should not be considered grey swans, as they had been or could have been anticipated based on historical observations and/or experience. Hurricane Katrina (2005), the costliest US natural disaster, generated the highest US recorded surge flooding (~10 m; ref. 7), but its impact on New Orleans, due largely to the levee failure, had been anticipated by various studies[8]. Cyclone Nargis (2008), the worst natural disaster in Myanmar's history and one of the deadliest TCs worldwide, struck Myanmar's Ayeyarwady River Delta at an unusually low latitude (near 16° N) and induced extreme surges (over 5 m); however, the catastrophic fatalities in the hardest-hit areas were largely due to the lack of evacuation plans and cyclone awareness[9], although intense tropical cyclones had been active in the Bay of Bengal and made landfall in Myanmar (for example, in 2006). Hurricane Sandy, which devastated the US Northeast coast in 2012, set the record-high storm tide (3.4 m) at the Battery in New York City (NYC); however, its storm surge (2.8 m) at the Battery was much lower than those of the 1821 NY hurricane (4.0 m; refs 10,11) and more severe grey swan TCs (4.5–5 m) that had been simulated for the region[12,13]. Typhoon Haiyan (2013), the deadliest TC in Philippine history, and probably the most powerful TC to make landfall worldwide,

generated extreme water levels up to 8 m near the most-affected Tacloban area[14], but the water level was comparable to those induced by earlier storms, including a severe typhoon that struck the area in 1897 (7.3 m; refs 6,15).

Prediction of a grey swan TC is meaningful and practically useful only when associated with some likelihood/probabilistic statement; for example, the probability of exceeding the storm surge level induced by the TC in a year is 10^{-3}. The Monte Carlo (MC) method, based on numerous synthetic simulations, is an important way to assess the probability of grey swan TCs. Most current MC methods[16–18] simulate synthetic TCs using (fairly limited) historical TC statistics. In contrast, a statistical–deterministic model[19], which is independent of the TC record, simulates TC environments statistically and generates TCs in the simulated environments deterministically. This statistical–deterministic approach may sometimes be more reliable, as observations of the large-scale TC environment are often better constrained than those of TC characteristics in areas with very limited TC history. It is also more likely to generate unexpected but realistic grey swan TCs, because, rather than extrapolating historical TCs, it applies physical knowledge of TCs and ample observations of the large-scale environment. Moreover, as the synthetic TC environments can be generated for any given climate state, this model can simulate grey swan TCs not only in the current and past climates but also in projected future climates[20]. As TC activity may vary with changing climate[21–24], the model enables quantitative projection of how grey swan TCs will evolve in the future. This statistical–deterministic TC model has been integrated with hydrodynamic surge models[25] into a climatological–hydrodynamic method[13], which has been shown to generate extreme storm surges that are far beyond historical records but are compatible with geologic evidence[26]. The method has been used to study storm surge risk and mitigation strategies for NYC (refs 27,28), and it is applicable to any coastal city. Here we apply the method to another three highly vulnerable regions: Tampa in Florida, Cairns in Australia, and the Persian Gulf; we identify their grey swan TCs as the synthetic TCs that are associated with extremely low annual exceedance probabilities (large mean return periods) of the induced storm surges (see Methods).

[1]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey 08544, USA. [2]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA. *e-mail: nlin@princeton.edu
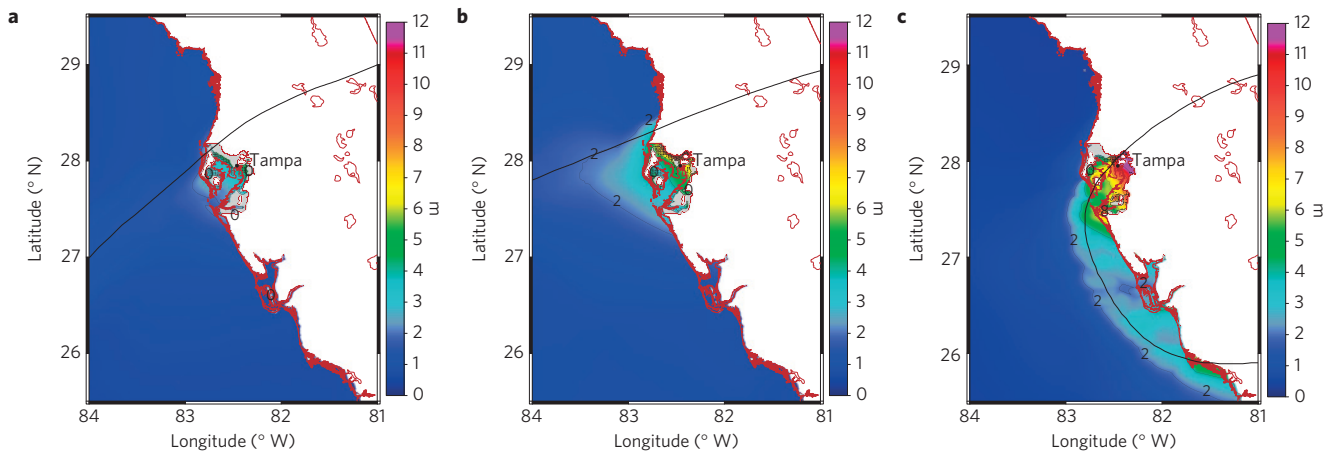
Figure 1 | The 1921 Tampa hurricane compared with two grey swan TCs. a, The 1921 Tampa hurricane simulated based on observed storm characteristics, including 1-min wind intensity (at 10 m) $V_m = 43.1\,\mathrm{m\,s^{-1}}$, minimum sea-level pressure $P_c = 967.8\,\mathrm{mb}$ and radius of maximum wind $R_m = 36.0\,\mathrm{km}$ (when the storm is at its nearest approach point to the site). Simulated surge at Tampa is 4.0 m. **b**, The 'worst' surge (5.9 m) event for Tampa in the NCEP/NCAR reanalysis climate of 1980–2005, with $V_m = 54.7\,\mathrm{m\,s^{-1}}$, $P_c = 953.4\,\mathrm{mb}$ and $R_m = 39.7\,\mathrm{km}$. **c**, The 'worst' surge (11.1 m) event for Tampa in the 2068–2098 climate projected by HADGEM for the IPCC AR5 RCP8.5 emission scenario, with $V_m = 104.3\,\mathrm{m\,s^{-1}}$, $P_c = 829.6\,\mathrm{mb}$ and $R_m = 17.0\,\mathrm{km}$. The shaded contours represent the simulated surge height (m; above MSL) and the black curve shows the storm track.

Tampa

Tampa, located on the central west Florida coast, is highly susceptible to storm surges. Although many fewer storms have made landfall in this area than in regions farther north and west on the Gulf Coast or further south on the Florida Coast, Tampa Bay is surrounded by shallow water and low-lying lands; a 6-m rise of water can inundate much of the Bay's surroundings[29]. Two significant historical events have affected Tampa. The Tampa Bay hurricane of 1848 produced the highest storm tide ever experienced in the Bay, about 4.6 m, destroying many of the few human works and habitations then in the area. The Tampa Bay hurricane of 1921 produced an estimated storm tide of 3–3.5 m, inducing severe damage (10 million in 1921 USD).

To investigate the current TC threat for Tampa we simulate 7,800 Tampa Bay synthetic TC surge events in the observed climate of 1980–2005 (late twentieth century) as estimated from the NCEP/NCAR reanalysis[30]. To study how the threat will evolve from the current to future climates, we apply each of six climate models to simulate 2,100 surge events for the climate of 1980–2005 (control) and 3,100 surge events for each of the three climates–2006–2036 (early twenty-first century), 2037–2067 (middle), and 2068–2098 (late)–under the IPCC AR5 RCP8.5 emission scenario. The six climate models, selected as in ref. 24 from Coupled Model Intercomparison Project Phase 5 (CMIP5), are CCMS4 (denoted as CCMS; NCAR), GFDL-CM3 (GFDL; NOAA), HADGEM2-ES (HADGEM; UK Met Office Hadley Centre), MPI-ESM-MR (MPI; Max Planck Institution), MIROC5 (MIROC; CCSR/NIES/FRCGC, Japan), and MRI-CGCM3 (MRI; Meteorological Research Institute, Japan).

The large synthetic surge database includes many extreme events affecting Tampa. As a comparison, the 1921 Tampa surge event is also simulated (Fig. 1a). The 1921 Tampa hurricane had a track similar to that of the 1848 Tampa hurricane[31], travelling northwestwards over the Gulf of Mexico and making landfall north of Tampa Bay. The 'worst' synthetic case (among 7,800 events) in the reanalysis climate of 1980–2005 has a similar track (Fig. 1b). However, this grey swan TC is more intense (upper Category 3, compared to the lower Category 2 1921 storm), inducing a higher surge at Tampa of over 5.9 m (compared to 4.0 m simulated for the 1921 storm). We have also identified grey swan TCs affecting Tampa that have very different tracks, especially those moving northwards parallel to the west Florida coast before making landfall. For example, Fig. 1c shows an extremely intense storm (104 m s⁻¹;

'worst' case generated under the late twenty-first-century climate projected by HADGEM) that moves northwards parallel to the coast and turns sharply towards Tampa Bay, inducing a storm surge of 11.1 m in Tampa. In such cases, the storm surges are probably amplified by coastally trapped Kelvin Waves. These waves form when the storm travels along the west Florida coast and propagate northwards along the Florida shelf, enhancing the coastal surges, especially when the storm moves parallel to the shelf and at comparable speed to the wave phase speed[32]. This geophysical feature makes Tampa Bay even more susceptible to storm surge.

These grey swan TCs have very low probabilities, which can be quantified only within the full spectrum of events. Figure 2 shows the estimated storm surge level for Tampa as a function of (mean) return period for the reanalysis climate of 1980–2005. The grey swan surge of 5.9 m (Fig. 1b) has a return period of over 10,000 years. In comparison, the 1,000-yr surge is about 4.6 m and the 100-yr surge is about 3.2 m. The observed surge level of the 1921 hurricane (approximately 3.3–3.8 m, as it probably happened at low tide) has an estimated return period of 100–300 years in the 1980–2005 climate. We note here a potentially large uncertainty in the analysis. In the simulations, we take the storm outer radius R_o to be its statistical mean[33] to generate the radius of the maximum wind R_m (see Methods). As shown previously[26], neglecting the statistical variation of storm size may greatly underestimate the surge risk, as the distributions of the size metrics (R_o and R_m) may be positively skewed[33]. Indeed, a sensitivity analysis for Tampa shows that the estimated surge return periods would be significantly reduced if a lognormal distribution of R_o (ref. 33) (with the same mean) was applied; for example, the return period of the 1921 storm surge could be reduced to as little as 60 years (not shown). However, the result is very sensitive to the specific distribution of R_o, which itself is largely uncertain owing to data limitations and lack of fundamental knowledge of what controls the TC size in nature[34,35].

The more severe grey swan surges of above 8 m up to 11 m (Fig. 1c) have extremely low or negligible probabilities in the 1980–2005 climate, but they are projected to happen as 5,000–150,000-yr events in the late twenty-first century. As shown in Fig. 3, the six climate models project that the return period of the storm surges for Tampa will significantly decrease over the twenty-first century, especially for the extremes (grey swans). This increase in storm surge threat is mainly due to the increase in storm frequency and intensity. The magnitude of the surge, especially for the
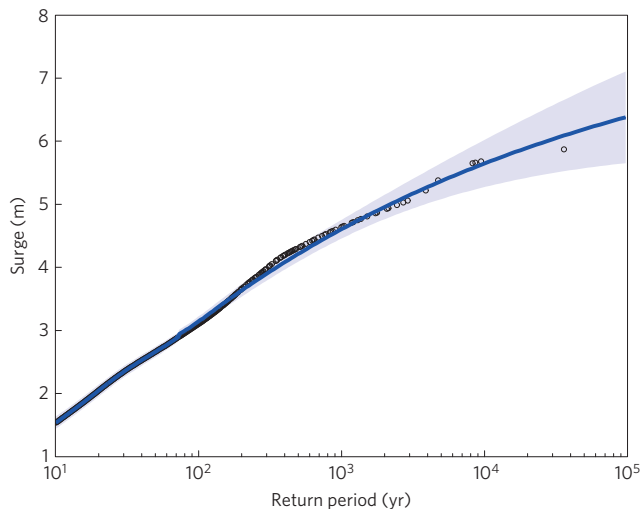
Figure 2 | Estimated storm surge level as a function of return period for Tampa for the NCEP/NCAR reanalysis climate of 1980–2005, based on 7,800 synthetic events. The associated annual frequency of the synthetic events is 0.36. Black dots show the simulated data, and the shading shows the 90% statistical confidence interval.

extremes, is projected to increase by all six models, and the CCSM, HADGEM and MPI models project relatively larger increases (see Supplementary Fig. 1). The overall frequency of the Tampa Bay storms is also projected to increase moderately (<25%) according to the CCSM, HADGEM and MRI models; greatly (<75%) according to MIROC and MPI; or extremely (240%) according to GFDL (noted in Fig. 3). As a result, the CCSM and HADGEM models project the largest increase in the frequency of the grey swans and little change in the normal events, whereas GFDL projects a relatively uniform increase in the frequency of all events, and the other three models

project relatively large (small) increases in the frequency of extremes (normal events). Hence, large uncertainties exist among the climate models in the probable increase of grey swans over the century. For example, a 10,000-yr event in the late twentieth century will become a 1,500–7,000-yr, 1,100–3,100-yr and 700–2,500-yr event in the early, middle and late twenty-first century, respectively, depending on the climate models; and a 1,000-yr event in the late twentieth century will become a 270–1,300-yr, 110–530-yr and 60–450-yr event in the early, middle and late twenty-first century, respectively. (Supplementary Fig. 2 (Supplementary Fig. 3) illustrates, for various levels of events, how the return periods (annual exceedance probabilities) decrease (increase) over the twenty-first century, projected by each of the six climate models.) Here the effect of neglecting the variability of storm size may be relatively small for the projections of the change of the probability. However, this analysis neglects the possible increase of the magnitude of storm size in a warmer climate. Although such an increase in storm size, as suggested by potential intensity theory[36], would further increase the surge risk[13], the effect of climate change on storm size has yet to be investigated observationally and numerically.

Cairns

The TC threat to Cairns, in the far north of Queensland, may not be well recognized. The city is located about 300 km south of Bathurst Bay, which was hit in 1899 by Cyclone Mahina (the most intense TC in the Southern Hemisphere, inducing what may have been the highest surge flooding (13 m) in the historical record[37]). According to the Australian Bureau of Meteorology, at least 53 cyclones have affected Cairns since it was founded in 1876, and several high-intensity storms (for example, Cyclones Larry in 2006 and Yasi in 2011) were near misses. Recent events include Cyclones Justin in 1997, Rona in 1999, and Steve in 2000, all making landfall north of Cairns; although these storms (<Category 2) generated storm surges in Cairns of less than 1 m, they induced major flooding (due also to tide and waves) and significant damage ($100–190 million)
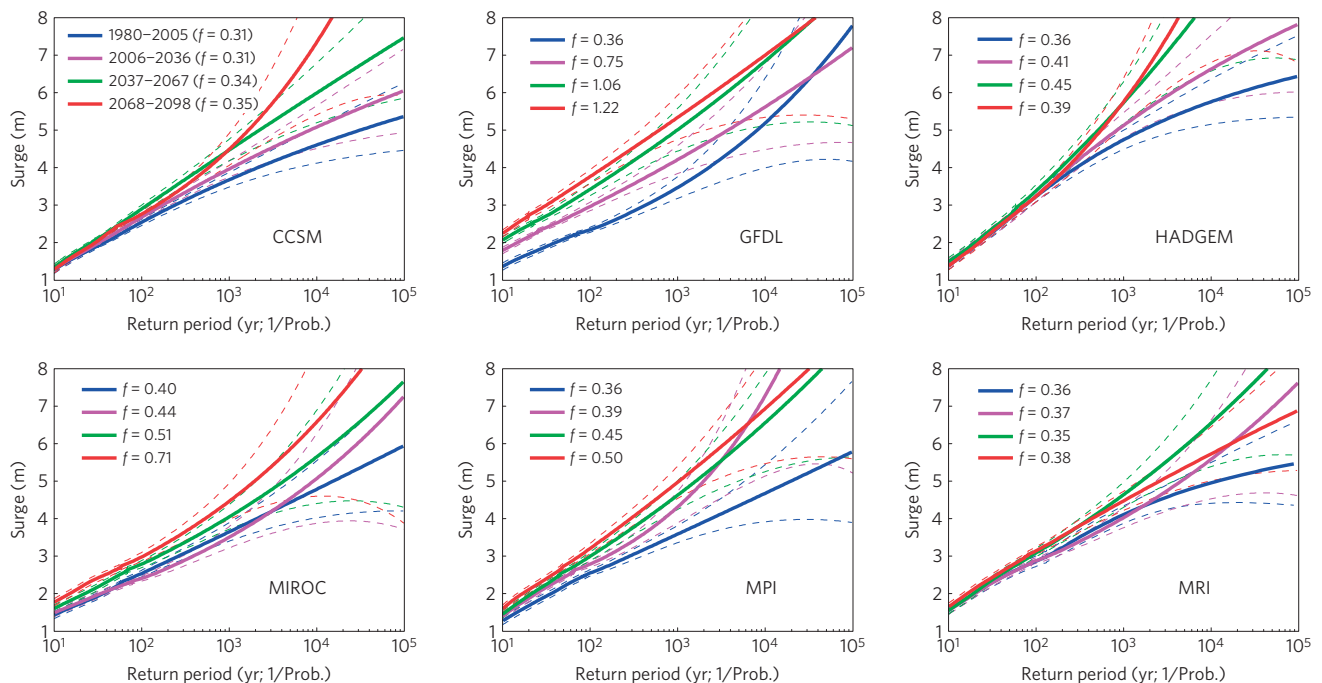

Figure 3 | Estimated storm surge level as a function of return period for Tampa in the climate of 1980–2005 (based on 2,100 events), 2006–2036 (3,100 events), 2037–2067 (3,100 events), and 2068–2098 (3,100 events) projected using each of the six climate models for the IPCC AR5 RCP8.5 emission scenario. The annual frequency (f) is noted for each case. The thin dash curves show the 90% statistical confidence interval. (The data points and goodness of fit for the upper tail are shown in Supplementary Fig. 1.)
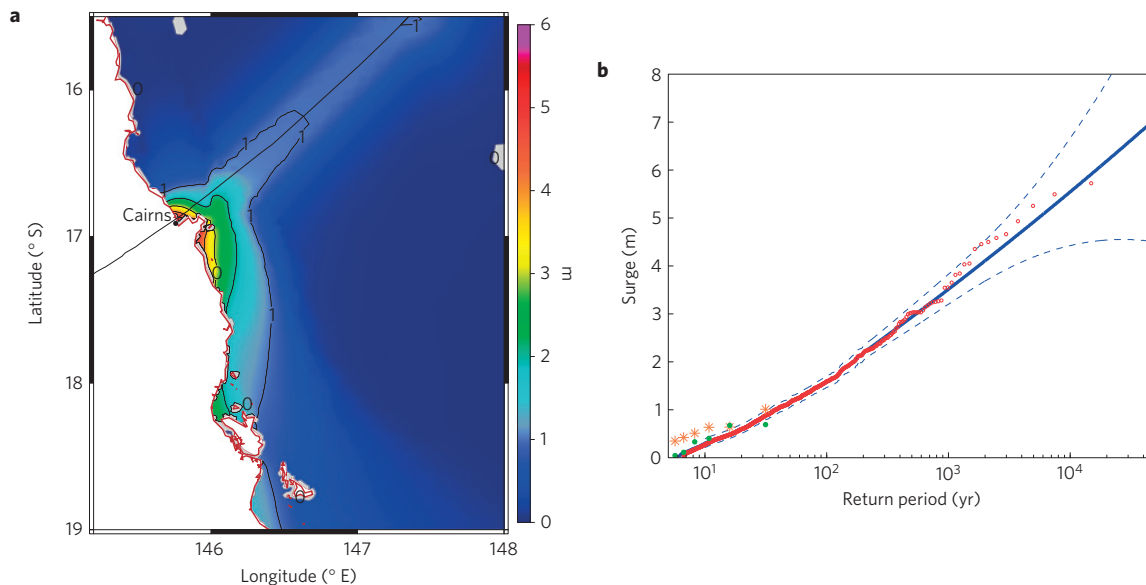
Figure 4 | Storm surge risk analysis for Cairns, Australia, based on 2,400 synthetic events in the NCEP/NCAR reanalysis climate of 1980–2010. The associated annual frequency for the synthetic events is 0.16. **a**, The 'worst' surge (5.7 m) event for Cairns, with $V_m = 79.3 \, \text{m s}^{-1}$, $P_c = 901.1 \, \text{mb}$ and $R_m = 22.3 \, \text{km}$. The shaded contours show the simulated surge height (m; above MSL) and the black curve shows the storm track. **b**, Estimated storm surge level as a function of return period for Cairns. The red dots show the synthetic data, and the dash curves show the 90% statistical confidence interval. Orange dots show the tidal-gauge-observed Cairns storm surges (six in total) between 1980 and 2010; green dots show the modelled surges for these historical TCs (the annual frequency of the historical storms is 0.19).

in the area. (Simulations of these historical cyclones, in comparison with observations, are shown in Supplementary Fig. 4.)

To study the TC threat for Cairns, we simulate 2,400 synthetic Cairns TC surge events in the NCEP/NCAR reanalysis climate of 1980–2010. The 'worst' surge for Cairns is about 5.7 m, induced by an intense storm (80 m s^{-1}) travelling perpendicularly to the coast and making landfall just north of Cairns (Fig. 4a). This grey swan TC is much stronger than Cyclones Justin, Rona and Steve, and makes landfall much closer to Cairns. It resembles a hypothetical Cyclone Yasi that is moderately intensified (by about 10 m s^{-1}) and shifted northwards by about 160 km.

As shown by the estimated surge return curve in Fig. 4b, the grey swan surge of 5.7 m has a return period of over 10,000 years in the 1980–2010 climate. As a reference, the 1,000-yr surge is about 3.5 m, and the 100-yr surge is about 1.6 m. These results are significantly higher than previous estimates based on synthetic storm databases generated by statistically extending the historical storm records. For example, one such study[38] estimated that the 1,000-yr storm surge level for Cairns is about 2.3 m (storm tide of 2.9 m) and the 100-yr surge level is about 1.3 m (storm tide of 2.0 m); another[39] estimated the 10,000-yr storm tide to be 2.6 m, the 1,000-yr storm tide to be 2.2 m, and the 100-yr storm tide to be 1.8 m. The lower estimates in these previous analyses, especially for the most extreme events, were probably deduced by extrapolating the storm record from several decades to thousands or tens of thousands of years. Analyses based on geologic evidence of palaeo coastal inundations also yielded much higher estimates of such extremes for the north of Queensland than these historical-storm-based estimates[40]; our results are more consistent with the geologic evidence (Nott, personal communication). (The estimated return levels based on the synthetic storms also compare well with observed and modelled historical event levels, available for short return periods; Fig. 4b.)

The Persian Gulf

The Persian Gulf is a mediterranean sea of the Indian Ocean, connected to the Arabian Sea through the Strait of Hormuz and Gulf of Oman. The Persian Gulf is comprised of hot, shallow, and highly saline water, which can support the development of intense TCs and storm surges. However, no TC has been observed in the Persian Gulf, and TC development in the Arabian Sea is limited by the region's typically low humidity and high wind shear[41]. Cyclone Gonu (2007), the strongest historical TC in the Arabian Sea (Category 3; 78 fatalities and 4.4 billion in damage), came close to entering the Persian Gulf, making landfall at the mouth of the Gulf on the easternmost tip of Oman and then in southern Iran. It is scientifically interesting and socially important to ask if such a strong TC can travel into the Persian Gulf.

To answer this question, we assess the TC threat for three major cities bordering the Persian Gulf: Dubai, Abu Dhabi and Doha. For each of these cities, we simulate 3,100 synthetic TC surge events in the NCEP/NCAR reanalysis climate of 1980–2010. As the maximum width of the Persian Gulf is only about 340 km, it may be poorly resolved by the NCAR/NCEP reanalysis resolution of 2.5 degrees (about 250 km); thus we also apply a higher-resolution reanalysis data set, the NASA Modern-Era Retrospective Analysis[42] (MERRA; with a resolution of 0.67° × 0.5°), to simulate the TC surge events for Dubai. The obtained surge levels and probabilities, however, are very similar for the two data sets. We here present the result for Dubai from the MERRA reanalysis (whereas the results for Dubai, Abu Dhabi and Doha from the NCEP/NCAR reanalysis are shown in the Supplementary Information). In these simulations, some of the synthetic storms originate in the Arabian Sea and move into the Persian Gulf, but the majority originate, surprisingly, within the Gulf. Moreover, the most extreme surges are all induced by intense storms that originate within the Gulf.

Figure 5a shows the 'worst' surge (among 3,100 events in the climate of 1980–2010) for Dubai. This grey swan TC originates in the northwest region of the Persian Gulf, moves southeastwards in the Gulf, and makes landfall north of Dubai with extremely high intensity (115 m s^{-1}), generating a storm surge of 7.4 m in Dubai. The intensity of this grey swan TC is far beyond the highest observed TC intensity worldwide (Typhoon Haiyan of 87 m s^{-1}). This extremely high wind intensity is due to very large
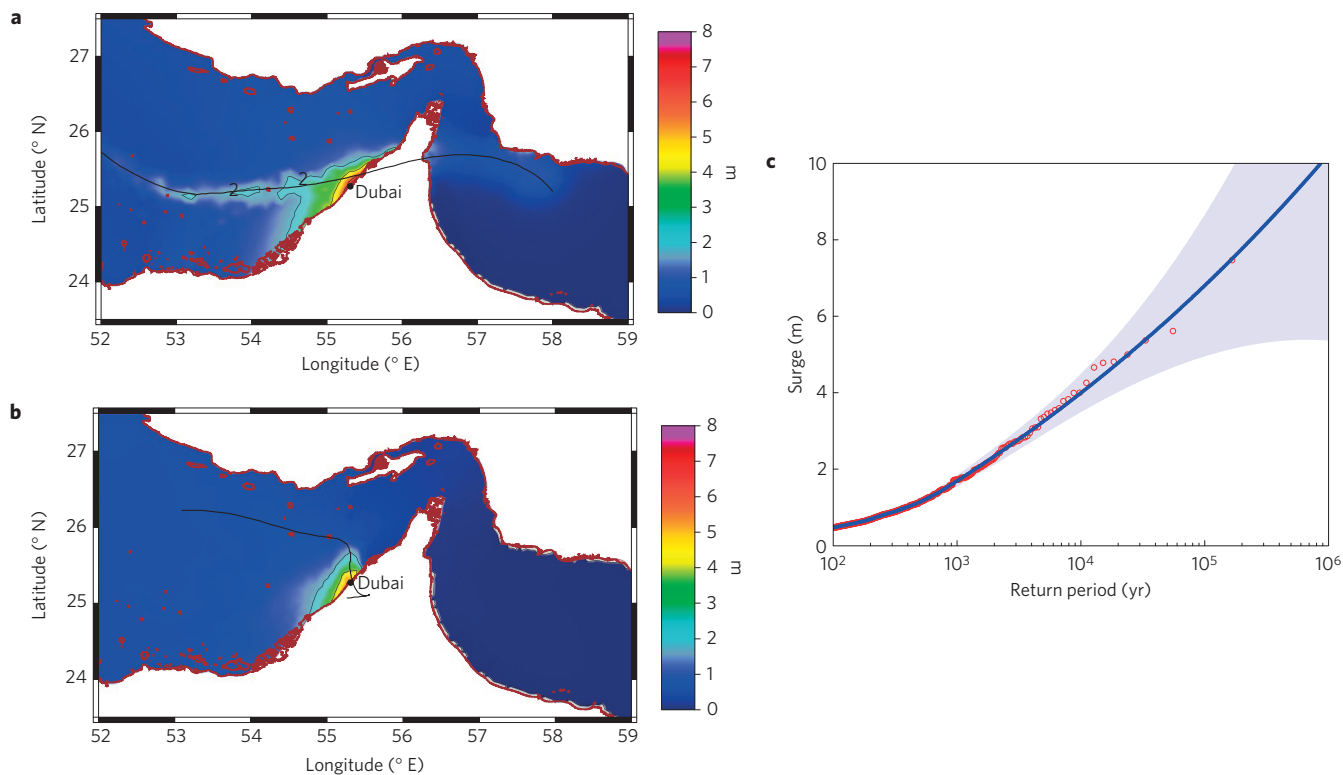
Figure 5 | Storm surge risk analysis for Dubai, based on 3,100 synthetic events in the MERRA reanalysis climate of 1980–2010. The associated annual frequency for the synthetic events is 0.037. **a**, The 'worst' surge (7.5 m) event for Dubai, with $V_m = 114.6$ m s^{-1}, $P_c = 784.2$ mb and $R_m = 13.8$ km. The shaded contours show the simulated surge height (m; above MSL) and the black curve shows the storm track. **b**, The second-'worst' surge (5.6 m) event for Dubai, with $V_m = 65.4$ m s^{-1}, $P_c = 927.3$ mb and $R_m = 21.3$ km. **c**, Estimated storm surge level as a function of return period for Dubai. The dots show the synthetic data, and the shading shows the 90% statistical confidence interval.

potential intensities (PIs), made possible by the area's high sea surface temperature (SST; with summertime peak values in the range of 32–35 °C (ref. 43)) and the deep dry adiabatic temperature profiles characteristic of desert regions. Indeed, the PI calculated (with the method of ref. 44) using the Dammam (Saudi Arabia) atmospheric sounding and an SST of 32–35 °C is between 109 m s^{-1} and 132 m s^{-1}. (The daily PI calculated using the sounding and the Hadley Centre observed SST, shown in Supplementary Fig. 5, confirms this result.) Furthermore, surface cooling from deep-water upwelling is nearly impossible in this highly saline and mixed body of shallow water (with a mean depth of about 36 m and a maximum depth of 90 m), and when, occasionally, the wind shear is small, the storm can fully achieve its potential intensity. (We note, however, that the estimated pressure intensity has not been similarly evaluated, which will be done in the future, but the storm surge is less sensitive to the pressure than to the wind intensity.)

Figure 5b shows the second-highest synthetic surge generated for Dubai. This grey swan TC originates in the southeast region of the Persian Gulf, moves directly towards the coast, and makes landfall almost perpendicularly to the coast and just north of Dubai, generating a storm surge of 5.7 m in Dubai. The storm intensity is moderate (65 m s^{-1}). It is not necessary for the storm to be extremely intense to generate extreme surges; some near 'perfect' combination of track, intensity and size can induce devastating surge inundation in Dubai, given its unusual shallow-water surroundings.

Nevertheless, given the prohibiting atmospheric environment in the region, these extreme grey swan TCs have very low probabilities, with return periods of the order of 30,000–200,000 years (Fig. 5c). Also, the surge level decreases rapidly with decreasing return period. The 10,000-yr surge for Dubai is about 4 m and the 1,000-yr surge is about 1.9 m. The surges for return periods less than 100 years

are very small. Similar and even higher surge levels for Abu Dhabi and Doha are also estimated using the NCAR/NCEP reanalysis (see Supplementary Figs 6 and 7).

We note that these analyses are based on the climate of 1980–2010, during which the Arabian Sea's synthetic TC activity increased, probably owing to a decrease in the wind shear[45]. Thus, although TC development is limited in the Persian Gulf, a large TC threat exists and may be very sensitive to changes of the atmospheric circulation in the region. Moreover, the SST in the Persian Gulf had a significant upward trend during the period of 1950–2010, with an abrupt increase in the 1990–2010 era[43]. Further warming of the ocean may further increase the chance of the Persian Gulf region being struck by an extreme storm.

Final remarks

Assessments of the risks associated with natural hazards such as tropical cyclones have been limited by the comparatively short length of historical records. This limitation is being overcome by the new field of palaeotempestology, which identifies TC events in the geologic record, and by bringing knowledge of storm physics to bear on the problem. Here we have used a physically based climatological–hydrodynamic method to assess the likelihood of highly destructive events for three regions. Uncertainty in storm size induces uncertainty in the estimated probabilities; accounting for the variation of storm size from storm to storm and in different climates, when more information on which becomes available, may yield significantly higher estimated TC threats. In addition to the storm surge that we focus on here, coastal inundation is also affected by the astronomical tide, waves, sea-level rise and future shoreline changes[46], all of which will amplify the impact of grey swan tropical cyclones.

Methods

Methods and any associated references are available in the online version of the paper.

References

1. Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable Fragility* (Random House LLC, 2010).
2. Aven, T. On the meaning of a black swan in a risk context. *Saf. Sci.* **57**, 44–51 (2013).
3. Nafday, A. M. Strategies for managing the consequences of black swan events. *Leadership Manage. Eng.* **9**, 191–197 (2009).
4. Stein, J. L. & Stein, S. Gray swans: Comparison of natural and financial hazard assessment and mitigation. *Nat. Hazards* **72**, 1279–1297 (2014).
5. Paté Cornell, E. On "Black Swans" and "Perfect Storms": Risk analysis and management when statistics are not enough. *Risk Anal.* **32**, 1823–1833 (2012).
6. Needham, H., Keim, B. D. & Sathiaraj, D. A review of tropical cyclone-generated storm surges: Global data sources, observations and impacts. *Rev. Geophys.* **53**, 545–591 (2015).
7. Fritz, H. M. *et al.* Hurricane Katrina storm surge distribution and field observations on the Mississippi Barrier Islands. *Estuar. Coast. Shelf Sci.* **74**, 12–20 (2007).
8. Travis, J. Scientists' fears come true as hurricane floods New Orleans. *Science* **309**, 1656–1659 (2005).
9. Fritz, H. M., Blount, C. D., Thwin, S., Thu, M. K. & Chan, N. Cyclone Nargis storm surge in Myanmar. *Nature Geosci.* **2**, 448–449 (2009).
10. Scileppi, E. & Donnelly, J. P. Sedimentary evidence of hurricane strikes in western Long Island, New York. *Geochem. Geophys. Geosyst.* **8**, Q06011 (2007).
11. Brandon, C. M., Woodruff, J. D., Donnelly, J. P. & Sullivan, R. M. How unique was Hurricane Sandy? Sedimentary reconstructions of extreme flooding from New York harbor. *Sci. Rep.* **4**, 7366 (2014).
12. Lin, N., Emanuel, K. A., Smith, J. A. & Vanmarcke, E. Risk assessment of hurricane storm surge for New York City. *J. Geophys. Res.* **115**, D18121 (2010).
13. Lin, N., Emanuel, K., Oppenheimer, M. & Vanmarcke, E. Physically based assessment of hurricane surge threat under climate change. *Nature Clim. Change* **2**, 1–6 (2012).
14. Mas, E. *et al.* Field survey report and satellite image interpretation of the 2013 Super Typhoon Haiyan in the Philippines. *Nat. Hazards Earth Syst. Sci.* **15**, 817–825 (2015).
15. Bankoff, G. *Cultures of Disaster: Society and Natural Hazard in the Philippines* 232 (Routledge Curzon, 2003).
16. Vickery, P., Skerlj, P. & Twisdale, L. Simulation of hurricane risk in the U.S. using empirical track model. *J. Struct. Eng.* **126**, 1222–1237 (2000).
17. Toro, G. R., Resio, D. T., Divoky, D., Niedoroda, A. W. & Reed, C. Efficient joint-probability methods for hurricane surge frequency analysis. *Ocean Eng.* **37**, 125–134 (2010).
18. Hall, T. M. & Sobel, A. H. On the impact angle of Hurricane Sandy's New Jersey landfall. *Geophys. Res. Lett.* **40**, 2312–2315 (2013).
19. Emanuel, K., Ravela, S., Vivant, E. & Risi, C. A Statistical deterministic approach to hurricane risk assessment. *Bull. Am. Meteor. Soc.* **87**, 299–314 (2006).
20. Emanuel, K., Sundararajan, R. & Williams, J. Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bull. Am. Meteor. Soc.* **89**, 347–367 (2008).
21. Emanuel, K. The dependence of hurricane intensity on climate. *Nature* **326**, 483–485 (1987).
22. Elsner, J. B., Kossin, J. P. & Jagger, T. H. The increasing intensity of the strongest tropical cyclones. *Nature* **455**, 92–95 (2008).
23. Knutson, T. R. *et al.* Tropical cyclones and climate change. *Nature Geosci.* **3**, 157–163 (2010).
24. Emanuel, K. A. Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. *Proc. Natl Acad. Sci. USA* **110**, 12219–12224 (2013).
25. Westerink, J. J. *et al.* A basin- to channel-scale unstructured grid hurricane storm surge model applied to southern Louisiana. *Mon. Weath. Rev.* **136**, 833–864 (2008).
26. Lin, N., Lane, P., Emanuel, K. A., Sullivan, R. M. & Donnelly, J. P. Heightened hurricane surge risk in northwest Florida revealed from climatological–hydrodynamic modeling and paleorecord reconstruction. *J. Geophys. Res. Atmos.* **119**, 8606–8623 (2014).
27. Aerts, J. C. J. H., Lin, N., Botzen, W., Emanuel, K. & de Moel, H. Low-probability flood risk modeling for New York City. *Risk Anal.* **33**, 772–788 (2013).
28. Aerts, J. C. J. H. *et al.* Evaluating Flood resilience strategies for coastal megacities. *Science* **344**, 473–475 (2014).
29. Weisberg, R. H. & Zheng, L. Hurricane storm surge simulations for Tampa Bay. *Estuar. Coast.* **29**, 899–913 (2008).
30. Kalnay, E. *et al.* The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
31. Bossak, B. H. *Early 19th Century US Hurricanes: A GIS Tool and Climate Analysis* PhD dissertation, Florida State Univ. (2003).
32. Morey, S. L., Baig, S., Bourassa, M. A., Dukhovskoy, D. S. & O'Brien, J. J. Remote forcing contribution to storm-induced sea level rise during Hurricane Dennis. *Geophys. Res. Lett.* **33**, L19603 (2006).
33. Chavas, D. R. & Emanuel, K. A. A QuikSCAT climatology of tropical cyclone size. *Geophys. Res. Lett.* **37**, 18 (2010).
34. Rotunno, R. & Emanuel, K. A. 1987: An air-sea interaction theory for tropical cyclones, Part II: Evolutionary study using axisymmetric nonhydrostatic numerical model. *J. Atmos. Sci.* **44**, 542–561.
35. Chavas, D. R. & Emanuel, K. Equilibrium tropical cyclone size in an idealized state of axisymmetric radiative–convective equilibrium*. *J. Atmos. Sci.* **71**, 1663–1680 (2014).
36. Emanuel, K. A. An air-sea interaction theory for tropical cyclones. Part I: Steady-state maintenance. *J. Atmos. Sci.* **43**, 585–605 (1986).
37. Nott, J., Green, C., Townsend, I. & Callaghan, J. The world record storm surge and the most intense southern hemisphere tropical cyclone: New evidence and modeling. *Bull. Am. Meteorol. Soc.* **95**, 757–765 (2014).
38. Hardy, T., Mason, L. & Astorquia, A. *Surge Plus Tide Statistics for Selected Open Coast Locations Along the Queensland East Coast. Queensland Climate Change and Community Vulnerability to Tropical Cyclones. Ocean Hazards Assessment Stage 3* (Queensland Government Report, 2004).
39. Haigh, I. D. *et al.* Estimating present day extreme water level exceedance probabilities around the coastline of Australia: Tropical cyclone-induced storm surges. *Clim. Dynam.* **42**, 139–157 (2014).
40. Nott, J. F. & Jagger, T. H. Deriving robust return periods for tropical cyclone inundations from sediments. *Geophys. Res. Lett.* **40**, 370–373 (2012).
41. Evan, A. T. & Camargo, S. J. A climatology of Arabian Sea cyclonic storms. *J. Clim.* **24**, 140–158 (2011).
42. Rienecker, M. M. *et al.* MERRA: NASA's modern-era retrospective analysis for research and applications. *J. Clim.* **24**, 3624–3648 (2011).
43. Shirvani, A., Nazemosadat, S. M. J. & Kahya, E. Analyses of the Persian Gulf sea surface temperature: Prediction and detection of climate change signals. *Arab. J. Geosci.* **8**, 2121–2130 (2015).
44. Bister, M. & Emanuel, K. A. Low frequency variability of tropical cyclone potential intensity 1. Interannual to interdecadal variability. *J. Geophys. Res.* **107(D24)**, 4801 (2002).
45. Evan, A. T., Kossin, J. P., Eddy' Chung, C. & Ramanathan, V. Arabian Sea tropical cyclones intensified by emissions of black carbon and other aerosols. *Nature* **479**, 94–97 (2011).
46. Woodruff, J. D., Irish, J. L. & Camargo, S. J. Coastal flooding by tropical cyclones and sea-level rise. *Nature* **504**, 44–52 (2013).

Author contributions

K.E. performed numerical modelling of the storms. N.L. carried out storm surge simulations and statistical analysis. N.L. and K.E. co-wrote the paper.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to N.L.

Competing financial interests

The authors declare no competing financial interests.

Methods

Storm generation. The climatological–hydrodynamic method includes three components: storm generation, storm surge simulation and statistical analysis. We use a statistical–deterministic TC model[19] to generate a sufficient number of synthetic TCs in an ocean basin under a given climate to obtain a desired number of TCs that make landfall in a particular coastal area of interest. Weak protostorms are seeded uniformly over the basin within a large-scale environment provided by a reanalysis or climate model data set. Once initialized, the storms move in accordance with the large-scale environmental wind. Along each storm track, the Coupled Hurricane Intensity Prediction System (CHIPS; ref. 47), a dynamic model, is used to simulate the storm intensity according to environmental conditions such as potential intensity, wind shear, humidity and the thermal stratification of the ocean. These wind and environmental conditions are modelled statistically based on the reanalysis or climate model data set. The CHIPS model also predicts the storm radius of maximum wind (R_m), given an externally supplied storm outer radius (R_o). We apply the observed basin mean of R_o based on the historical record[33] and assume it is constant over the lifecycle of a storm (as it is observed to not change much over the storm lifetime[33]). Then we estimate R_m (varying from storm to storm and over the lifecycle of a storm) from CHIPS.

We design specific criteria (a filter) for each study area to select local storms from basin-wide events (and record the corresponding annual frequency of the local storms; the frequency of the basin-wide storms is calibrated to the observation). Various storm tracks can induce significant surges in Tampa Bay, including those that make landfall within or near the Bay as well as those that travel close offshore and parallel to the coast. To capture all these storms, we create a two-line-segment filter encompassing the Bay and surrounding coastal region. One line segment links a point on the coast (82.81° W, 29.17° N) about 180 km north of the Bay's mouth, to a point over the ocean (83.8° W, 27.58° N) about 100 km west of the Bay's mouth. The other line segment links the ocean point (83.8° W, 27.58° N) to a coastal point (82.407° W, 27.0° N) about 70 km south of the Bay's mouth. We select all storms that cross either of these two line segments with intensity greater than 21 m s^{-1}; we call these storms 'Tampa Bay storms.' Simpler, circular filters are created for the other study areas. We create a circle centred in Cairns (145.76° E, 16.91° S) with a radius of 100 km to select all 'Cairns storms' that move into this circle with intensity great than 21 m s^{-1}. Similarly, we create 100-km-radius circular filters centred at Dubai (55.31° E, 25.27° N), Abu Dhabi (54.37° E, 24.47° N) and Doha (51.53° E, 25.28° N).

Given the storm characteristics of selected storms in a study area, we estimate the surface wind and pressure fields using parametric methods fit to the explicitly modelled maximum wind speed, radius of maximum winds, and minimum surface pressure. In particular, the surface wind (1-min wind at 10 m) is estimated by fitting the wind velocity at the gradient height to an analytical hurricane wind profile[48], translating the gradient wind to the surface level with a velocity reduction factor (0.85) and an empirical formula for inflow angles, and adding a fraction (0.55 at 20 degrees cyclonically) of the storm translation velocity to account for the asymmetry of the wind field induced by the surface background wind[49]. The surface pressure is estimated also from a simple parametric model[50].

Surge simulation. With the storm surface wind and pressure fields as input, we apply the Advanced Circulation (ADCIRC) model[25,51] to simulate the storm surge. ADCIRC is a finite element hydrodynamic model that has been validated and applied to simulate storm surges and make forecasts for various coastal regions[25,52]. It allows the use of an unstructured grid with very fine resolution near the coast and much coarser resolution in the deep ocean. The ADCIRC mesh we developed for Tampa covers the entire Gulf of Mexico. The mesh has a peak resolution of about 100 m along the west Florida coast near Tampa and extends on land up to the 10-m height contour in the Tampa Bay area. The meshes developed for other study regions are relatively coarser (given coarser bathymetric data). To capture the effect of storms approaching from various directions, the mesh for Cairns has as its lower boundary the Australian coastline of Queensland, the Northern Territory, and part of Western Australia. The mesh extends over the Indian and South Pacific Oceans (from 114.0° E to 176.0° E) and is bounded above by Indonesia and Indonesian New Guinea. The resolution is about 1 km on the Queensland coast around Cairns. The mesh developed for Dubai covers the entire Persian Gulf and extends over the Arabian Sea (down to 16.0° N). The resolution is about 2 km near Dubai. The same mesh is used for Abu Dhabi and Doha; the resolution around these two locations is about 3–4 km.

To evaluate our surge modelling configuration and ADCIRC meshes, we simulated historical events for Tampa and Cairns (the Persian Gulf has no historical storms) with the storm characteristics obtained from the Best Track databases[53,54]. The simulated storm surge in Tampa for the 1921 hurricane is about 4.0 m (see Fig. 1a in the main article), which is comparable to that observed in this region ($\sim$3.3–3.8 m, considering the storm tide was estimated to be 3.0–3.5 m and happening probably at low tide), given the large uncertainties in both the observed surge level and storm characteristics (especially the size) for this early storm. (Note that in this case, because an observation of R_m is available only at landfall and there is no information about R_o, we estimated R_o from the landfall R_m using an empirical relationship[26] between them and the wind intensity, and then kept the estimated R_o constant to estimate R_m for the time periods before landfall using the empirical relationship.) For Cairns, we simulated storm surges for all six historical Cairns storms between 1980 and 2010 (selected using the same filter as for the synthetic storms) plus Cyclone Yasi in 2011. Simulations are close to the observations for the most significant events, including Cyclones Justin (1997), Rona (1999) and Yasi (see Supplementary Fig. 4), but the simulation underestimates the surge for Cyclone Steve (2000). Not all simulated historical surges match well with the observations individually, mainly owing to the uncertainty in storm size (an empirical estimate[26] of the R_m using the basin mean R_o was applied owing to the lack of observations). However, the simulations compare relatively well with all observations statistically (see Fig. 4b in the main article).

Statistical analysis. Statistical analysis is performed on the synthetic surge data sets. For a specific location and a given climate scenario, we assume the arrival of storms to be a stationary Poisson process, with arrival rate as the storm annual frequency. For each storm arrival, the probability density function (PDF) of its induced storm surge is characterized by a long tail. We apply a peaks-over-threshold (POT) method to model this tail with a generalized Pareto distribution (GPD), using the maximum likelihood method, and the rest of the distribution with non-parametric density estimation. The estimated storm annual frequency and surge PDF are then combined to calculate the (mean) return period (the reciprocal of the annual exceedance probability) for various surge levels[13], with the associated statistical confidence interval calculated using the Delta method[55].

It is noted that the statistical confidence interval is a function of the sample size; it is small for short return periods (where large numbers of samples are generated) and becomes large for long return periods (less samples for extremes). This statistical confidence interval does not account for the epistemic uncertainties in the models that are used to generate the samples. The overall epistemic uncertainty in the estimation may be considered as the combination of the statistical confidence interval and the variation in the estimations using different (and ideally full ranges of) models (for example, different climate model projections presented in the main article).

References

47. Emanuel, K., Des Autels, C., Holloway, C. & Korty, R. Environmental control of tropical cyclone intensity. *J. Atmos. Sci.* **61,** 843–858 (2004).
48. Emanuel, K. & Rotunno, R. Self-stratification of tropical cyclone outflow. Part I: Implications for storm structure. *J. Atmos. Sci.* **68,** 2236–2249 (2011).
49. Lin, N. & Chavas, D. On hurricane parametric wind and applications in storm surge modeling. *J. Geophys. Res.* **117,** D09120 (2012).
50. Holland, G. J. An analytic model of the wind and pressure profiles in hurricanes. *Mon. Weath. Rev.* **108,** 1212–1218 (1980).
51. Luettich, R. A., Westerink, J. J. & Scheffner, N. W. *ADCIRC: An Advanced Three-dimensional Circulation Model for Shelves, Coasts and Estuaries, Report 1: Theory and Methodology of ADCIRC-2DDI and ADCIRC-3DL* DRP Technical Report DRP-92-6 (Department of the Army, US Army Corps of Engineers, Waterways Experiment Station, 1992).
52. Dietrich, J. C. *et al.* Modeling hurricane waves and storm surge using integrally-coupled, scalable computations. *Coast. Eng.* **58,** 45–65 (2011).
53. Landsea, C. W. *et al.* A reanalysis of the 1921–1930 Atlantic hurricane database. *J. Clim.* **25,** 865–885 (2012).
54. Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J. & Neumann, C. J. The International Best Track Archive for Climate Stewardship (IBTrACS) unifying tropical cyclonedata. *Bull. Am. Meteorol. Soc.* **91,** 363–376 (2010).
55. Coles, S. *An Introduction to Statistical Modeling of Extreme Values* (Springer, 2001).