

# Satellite imaging reveals increased proportion of population exposed to floods

<https://doi.org/10.1038/s41586-021-03695-w>

Received: 25 August 2020

Accepted: 3 June 2021

Published online: 4 August 2021

 Check for updates

B. Tellman<sup>1,2,3,10</sup>✉, J. A. Sullivan<sup>2,3,4,10</sup>, C. Kuhn<sup>5</sup>, A. J. Kettner<sup>6</sup>, C. S. Doyle<sup>2,7</sup>, G. R. Brakenridge<sup>6</sup>, T. A. Erickson<sup>8</sup> & D. A. Slayback<sup>9</sup>

Flooding affects more people than any other environmental hazard and hinders sustainable development<sup>1,2</sup>. Investing in flood adaptation strategies may reduce the loss of life and livelihood caused by floods<sup>3</sup>. Where and how floods occur and who is exposed are changing as a result of rapid urbanization<sup>4</sup>, flood mitigation infrastructure<sup>5</sup> and increasing settlements in floodplains<sup>6</sup>. Previous estimates of the global flood-exposed population have been limited by a lack of observational data, relying instead on models, which have high uncertainty<sup>3,7–11</sup>. Here we use daily satellite imagery at 250-metre resolution to estimate flood extent and population exposure for 913 large flood events from 2000 to 2018. We determine a total inundation area of 2.23 million square kilometres, with 255–290 million people directly affected by floods. We estimate that the total population in locations with satellite-observed inundation grew by 58–86 million from 2000 to 2015. This represents an increase of 20 to 24 per cent in the proportion of the global population exposed to floods, ten times higher than previous estimates<sup>7</sup>. Climate change projections for 2030 indicate that the proportion of the population exposed to floods will increase further. The high spatial and temporal resolution of the satellite observations will improve our understanding of where floods are changing and how best to adapt. The global flood database generated from these observations will help to improve vulnerability assessments, the accuracy of global and local flood models, the efficacy of adaptation interventions and our understanding of the interactions between landcover change, climate and floods.

Damaging floods are increasing in severity, duration and frequency, owing to changes in climate, land use, infrastructure and population demographics<sup>7,12–14</sup>. An estimated \$651 billion (USD) in flood damages occurred globally from 2000 to 2019<sup>2</sup>. Investments in flood adaptation reduce mortality and asset losses<sup>3,15</sup>. Yet, only 13% of disaster funds are allocated to preparedness, mitigation and adaptation<sup>16</sup>. Fundamental to prioritizing disaster mitigation efforts is quantifying global changes in flood hazard, exposure and vulnerability. We use the IPCC<sup>17</sup> definitions of ‘flood hazard’ as the frequency and magnitude of events and of ‘exposure’ as the people, livelihoods, ecosystems and assets located where a hazard has or could occur. ‘Vulnerability’ is defined as the propensity for loss of lives, livelihoods and property and for other aspects of wellbeing to be adversely affected<sup>18</sup>. Previous global flood exposure and vulnerability studies have relied on modelled flood hazard<sup>7,8,10,19</sup>. One study estimates that population growth in 100-year floodplains (areas with a 1% annual flood probability) outpaced total population by 2.6% from 1970 to 2010 in 22 countries<sup>7</sup>. Vulnerability influences a wide range of adverse outcomes from flood events, including death, disease, psychological trauma, migration, property loss and poverty<sup>20</sup>. Studies enabled by global flood models reveal trends including declines in loss of life and reduced property damage when controlling for hazard

size<sup>3,10,11</sup>. Sub-Saharan Africa is the only region with increasing flood mortality rates since 1990<sup>10</sup>, where urban flooding has been growing and is expected to continue to do so<sup>7,8</sup>.

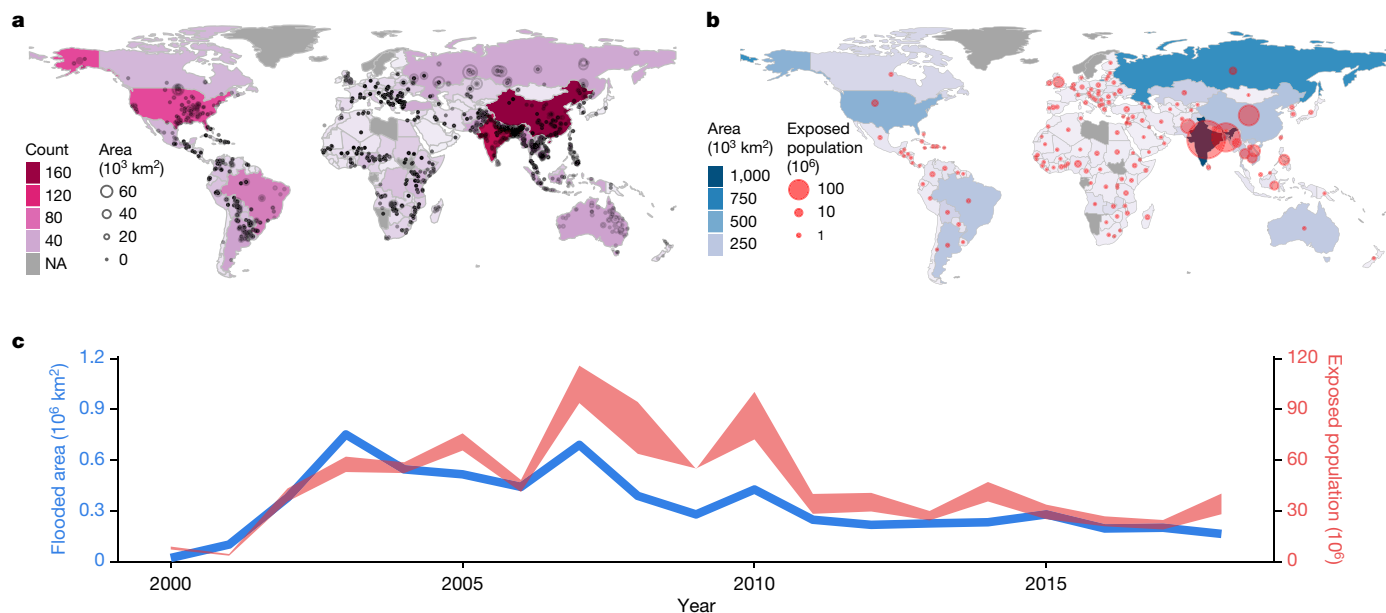
Flood exposure and vulnerability assessments are limited by the uncertainty in hazard models, which is due to the challenges of incorporating rapid anthropogenic change, to inadequate calibration data and to poor quality topographic data<sup>21</sup>. Humans modify land use and rivers, shifting flood water and reshaping exposure<sup>5,22</sup>. Differences in modelling assumptions lead to high disagreement between population and area exposure estimates across models<sup>9,21</sup>. Contrary to models, satellite-based remote sensing can directly observe inundation<sup>23,24</sup>, implicitly accounting for changes in climate, land use and infrastructure that are not reflected in modelled flood extents.

Here we measure global flood exposure from earth-observing satellites. We developed the Global Flood Database to systematically map the maximum observed surface-water extent during 913 large flood events documented by the Dartmouth Flood Observatory (DFO) from 2000 to 2018. The Global Flood Database (<http://global-flood-database.cloudtostreet.info/>) complements existing surface-water products that consist of monthly<sup>25</sup> or daily<sup>26</sup> observations by providing a geospatial event catalogue to aid model calibration and intercomparison<sup>27</sup>. With a

<sup>1</sup>Earth Institute, Columbia University, New York, NY, USA. <sup>2</sup>Cloud to Street, Brooklyn, NY, USA. <sup>3</sup>School of Geography, Development and Environment, University of Arizona, Tucson, AZ, USA.

<sup>4</sup>School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA. <sup>5</sup>School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA. <sup>6</sup>INSTAAR, Dartmouth Flood Observatory, University of Colorado, Boulder, CO, USA. <sup>7</sup>Department of Geography and the Environment, University of Texas, Austin, TX, USA. <sup>8</sup>Google, Mountain View, CA, USA. <sup>9</sup>Science Systems and Applications Inc., Biospheric Sciences Lab, NASA Goddard Space Flight Center, Greenbelt, MD, USA. <sup>10</sup>These authors contributed equally: B. Tellman, J. A. Sullivan.

<sup>✉</sup>e-mail: [beth@cloudtostreet.info](mailto:beth@cloudtostreet.info)



**Fig. 1 | Summary statistics of the Global Flood Database.** **a**, Number of flood events in the Global Flood Database per country (colour scale), along with the centroid locations and area of each flood event (circles). Countries with no observations are shaded grey (NA, not available). **b**, Total (cumulative over 2000–2015) exposed population (circles) and exposed area (colour scale) per country (Supplementary Table 6). **c**, Estimates of annual global population

(right axis, red shading; upper bound, GHSL; lower bound, HRSL) and area inundated (left axis, blue line). The 913 flood events represent those for which high-quality data were available (Methods, ‘Flood map quality control’). Population and area exposure to floods are lower in 2000 and 2001 until a second satellite (MODIS Aqua) was launched, increasing the likelihood of mapping a flood. Base maps: Natural Earth, tmap R package<sup>51</sup>.

spatial resolution of 250 m, the moderate-resolution imaging spectroradiometer (MODIS; a multispectral optical instrument mounted on NASA’s Terra and Aqua satellites that each image the globe daily<sup>28</sup>) resolves large, slow-moving flood events, but has limited ability to resolve urban floods. We estimate exposure trends using methods similar to a previous study<sup>7</sup>, by comparing the proportion of the observed flood-exposed population in 2015 to that in 2000 for each country (equation (6), Methods). Owing to uncertainty in population data<sup>29</sup>, exposure is estimated as a range across two datasets: the global human settlement layer (GHSL)<sup>30</sup> and the high-resolution settlement layer (HRSL)<sup>31</sup>. We then estimate change in flood exposure for the near future (2030), using flood hazard extents from the Global Flood Risk with Image Scenarios (GLOFRIS) model, which is based on present-day emissions scenarios<sup>32</sup> and socio-economic trends<sup>33</sup> for large events (the 100-year return period). We compare observations from the recent past (2000–2015) to modelled estimates in 2030 to identify countries on slowing, continuing or increasing flood exposure trajectories. This analysis may enable prioritization of adaptation measures where flood exposure has been growing or may grow faster than the total population under a changing climate.

### Satellite-observed inundation

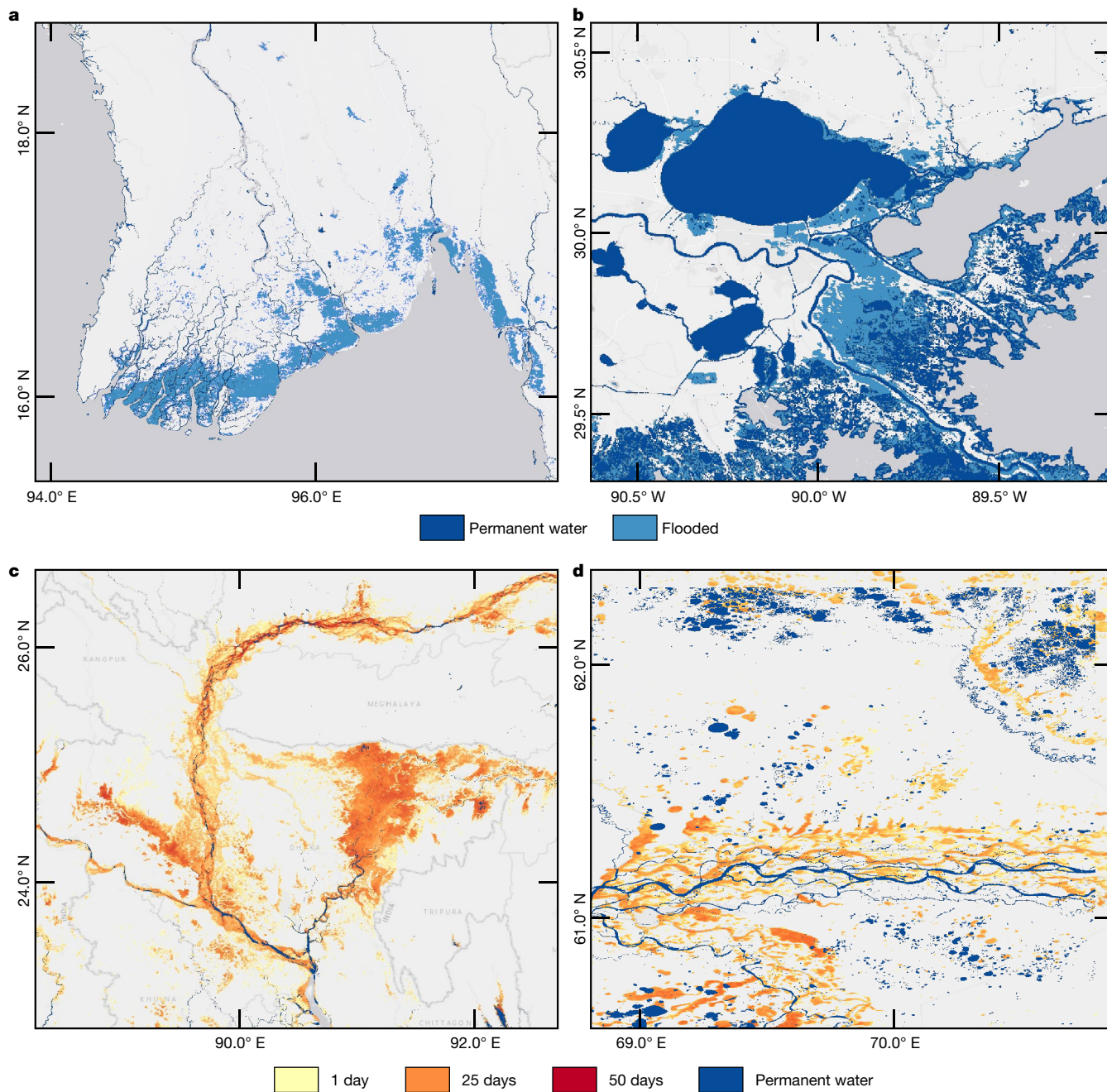
We analyse 12,719 MODIS images from 2000 to 2018 to produce 913 flood maps (Fig. 1a, Extended Data Fig. 1). We detected surface water at 250-m spatial resolution by applying empirically derived and Otsu-optimized thresholds<sup>34</sup> (Extended Data Fig. 2) to the short-wave-infrared, near-infrared and red bands (bands 7, 2 and 1) from MODIS. Results were validated using 30-m-spatial-resolution Landsat scenes coincident with the day of maximum inundation ( $n = 123$  events) for 30,685 points, yielding a mean accuracy of 83% (s.d. = 15%) for empirical thresholds and 80% (s.d. = 12%) for Otsu thresholds (Extended Data Fig. 3). Errors of commission (greater than 65%) were concentrated in northern latitudes, where low sun angle on dark soil causes low reflectance that mimics water<sup>35</sup>. Errors of omission show no geographic pattern (Extended Data Figs. 4, 5).

Of the 3,054 flood events in the DFO catalogue (compiled largely from news reports), we successfully mapped 913 events with mostly cloud-free MODIS observations. We found no temporal bias of events over time due to increased news-media-reporting trends in the DFO catalogue when compared to another database (the Emergency Events Database<sup>36</sup>; Extended Data Fig. 9a). MODIS could not detect floods in 2,141 events because of persistent cloud cover ( $n = 495$  events), small or flash floods ( $n = 300$  events), inaccurate catalogue locations ( $n = 94$ ), complex terrain (for example, dense forest, cities;  $n = 44$  events) or other reasons ( $n = 1,208$ ; Extended Data Fig. 6, Supplementary Table 9). Event maps may underrepresent the maximum flood extent, owing to the aforementioned uncertainty, and damaging floods underrepresented by the media may be absent from the DFO catalogue.

Most events in the Global Flood Database occurred in Asia ( $n = 398$ ; 52 in China and 85 in India), followed by the Americas ( $n = 223$ ; 98 in the US), Africa ( $n = 143$ ), Europe ( $n = 92$ ) and Oceania ( $n = 57$ ; Fig. 1a). Many flood events occurred across multiple countries, giving rise to 2,617 single-country events observed by MODIS. We estimate that 255–290 million people (about 3% of the global population) have been exposed to at least one observed event since 2000 and three flood events on average (735–892 million total exposures; Fig. 1c). Consistent with flood model estimates<sup>37</sup>, 90% of exposure is concentrated in south and southeast Asia. Most flood events were caused by heavy rainfall ( $n = 751$ ), followed by tropical storms or surges ( $n = 97$ ), snow or ice melt ( $n = 52$ ) or dam breaks ( $n = 13$ ). The largest cumulative global inundated area occurred in 2003 and 2007, with highest population exposure in 2007 and 2010. We highlight notable events with high human and socioeconomic losses in Fig. 2a–d.

### Flood-exposed population 2000–2015

The total global population increased by 18.6% from 2000 to 2015, compared with 34.1% in areas of observed inundation. Between 2000 and 2015, 58–86 million people, or 23%–30% of the total population exposed, were newly residing in areas where inundation was observed



**Fig. 2 | Observed inundation and flood duration for selected extreme events.** **a, b**, Observed inundation exceeding permanent water (from the Joint Research Program<sup>25</sup>) for the events in the Global Flood Database with the highest mortality (**a**; cyclone Nargis, Burma, 2008; roughly 100,000 people) and with the most expensive recovery (**b**; hurricane Katrina, USA, 2005;

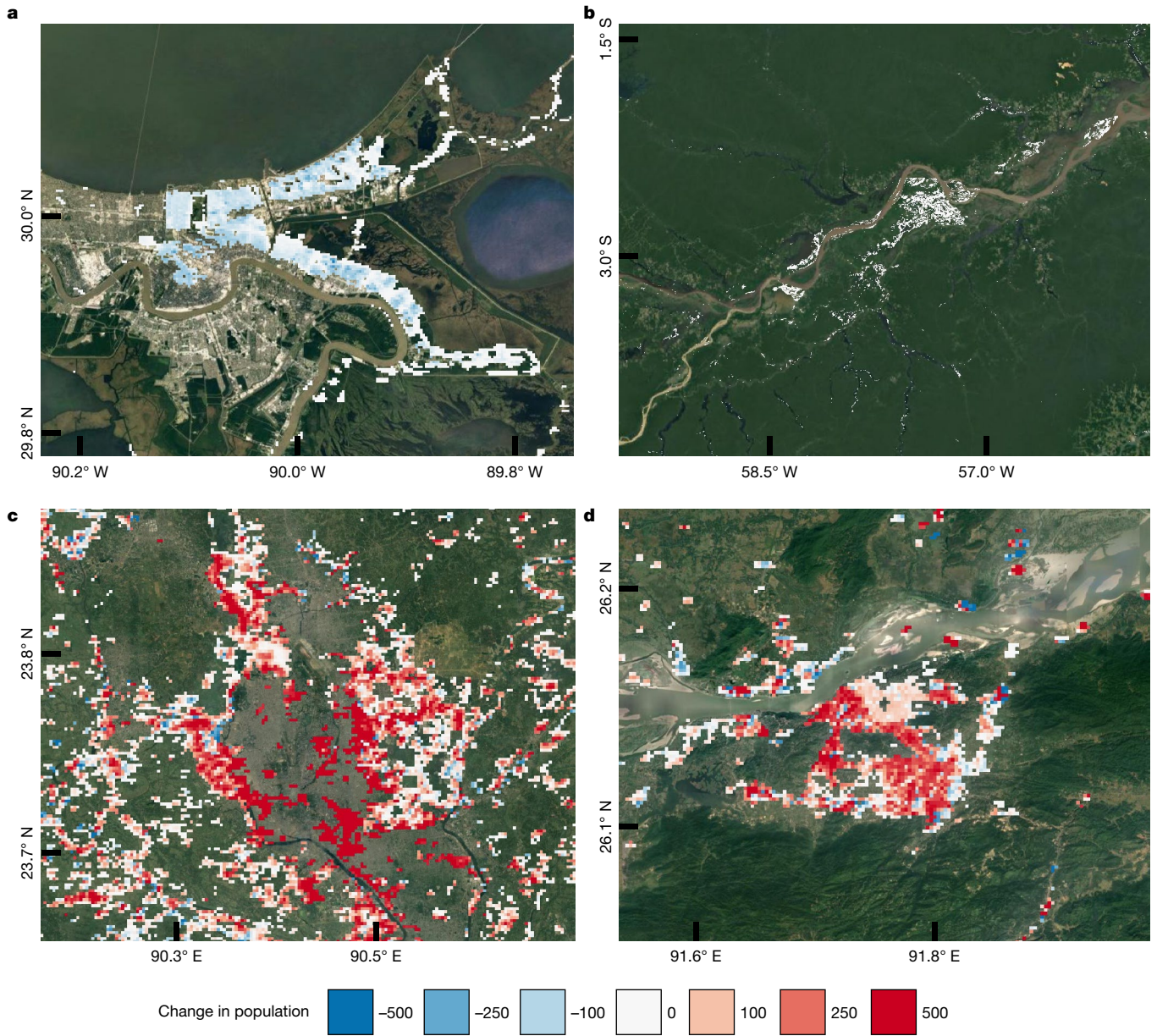
\$60 billion (USD)). **c, d**, Flood duration exceeding permanent water for the events in the Global Flood Database with the highest estimated exposure (**c**; India and Bangladesh, 2004; 27 million people exposed) and with the largest area (**d**; Russia, 2003; 98,000 km<sup>2</sup>). Base maps: Light Gray Canvas, Esri, HERE, Garmin, INCREMENT P, OpenStreetMap contributors and the GIS user community.

at least once. The change in the proportion of the population exposed to large flood events (equation (6), Methods) represents a global-mean increase of 20%–24% (s.d. = 53%) across 119 countries. Increased flood exposure was concentrated in low- and middle-income countries (Fig. 4a). Flood exposure trends are probably underestimated in rapidly urbanizing countries, because urban floods are underrepresented in the Global Flood Database. We excluded 15 countries from the trend analyses because the uncertainty in the population estimates was larger than the estimated trends (Supplementary Discussion). The proportion of the population exposed to floods increased across all flood types, but was highest in regions with floods caused by dam breaks, where it increased by 177% (Supplementary Table 8). Increased exposure

near flood mitigation infrastructure (such as dams) could be due to the levee effect<sup>38</sup>.

The proportion of the population in inundated areas increased by more than 2% in 70 countries and by more than 20% in 40 countries. Example locations with large population growth in observed inundation areas include Guwahati, India and Dhaka, Bangladesh (Fig. 3c, d). Countries with increased flood exposure were concentrated in Asia and sub-Saharan Africa. Large basins in south and southeast Asia (Indus, Ganges-Brahmaputra and Mekong) had the largest absolute numbers of people exposed (17.0–19.9 million, 107.8–134.9 million and 20.2–32.8 million, respectively) and increased proportions of the population exposed to inundation (36%, 26% and 11%, respectively; Supplementary Table 7).





**Fig. 3 | Population dynamics per pixel (250-m resolution) in observed inundated areas, 2000–2018. a**, New Orleans, USA, loses population after hurricane Katrina. **b**, Manaus, Brazil; no population change. **c**, Dhaka, Bangladesh, exhibits increasing population in inundated areas in the

peri-urban zone. **d**, Guwahati, India, an urbanizing town on the Brahmaputra River has repeatedly been exposed to flooding over the past two decades. Base maps: Google, 2015.

In 21 countries, there was little change in the proportion of the population exposed to floods (between  $-3\%$  and  $2\%$  growth), especially where populations have declined in eastern Europe and Russia<sup>39</sup>. Population growth in floodplains was heterogenous across countries. For example, in Brazil, flood exposure increased on average, but little to no population growth was observed in inundated areas recorded in the city of Manaus (Fig. 3b).

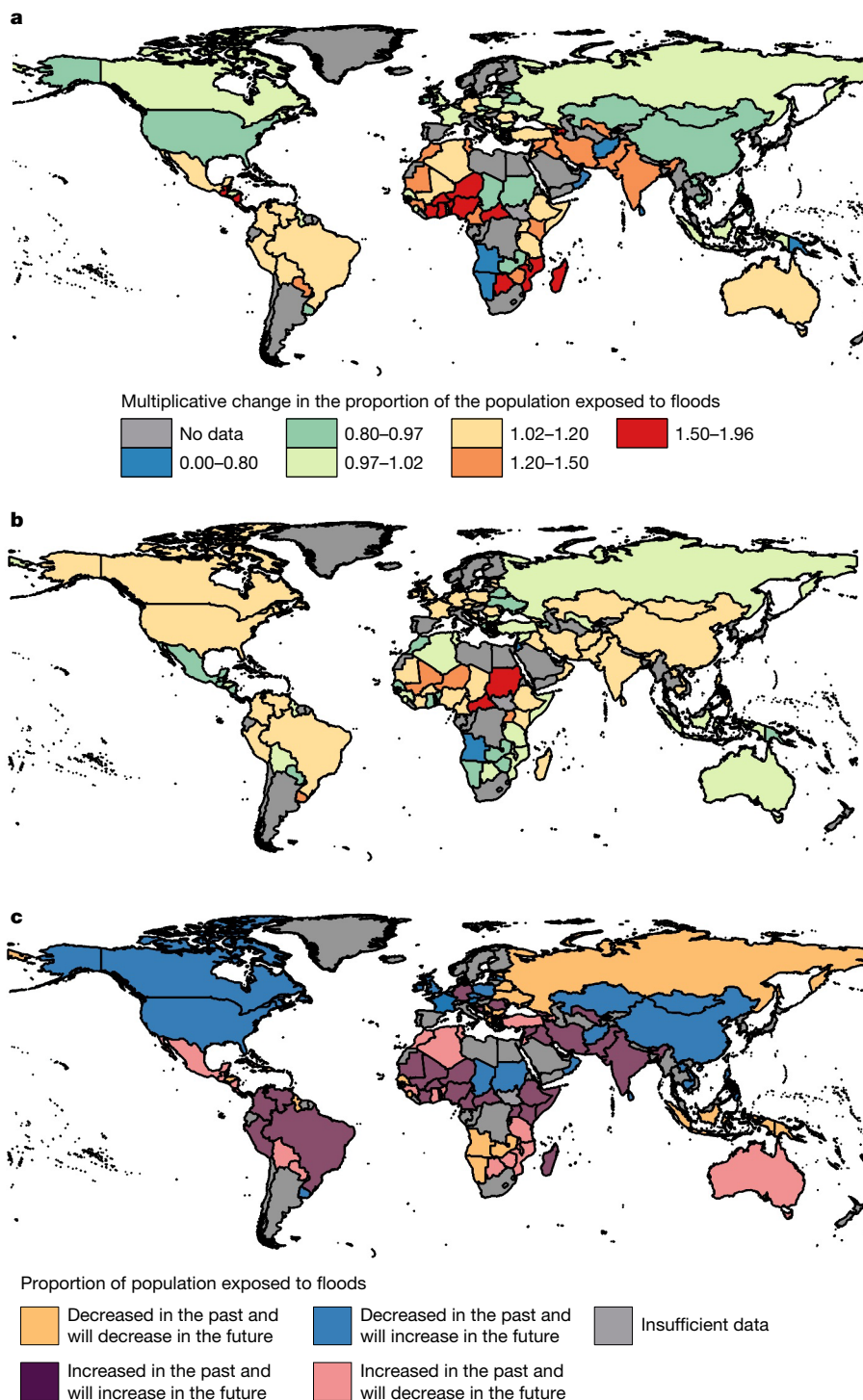
In 28 countries, the proportion of the population exposed to floods decreased by more than  $3\%$ . For example, in the US, the flood-exposed population decreased in New Orleans after hurricane Katrina (Fig. 3a)<sup>40</sup>. Our data indicate that the flood-exposed population decreased in Sri Lanka, potentially because nearly 500,000 people<sup>41</sup> were displaced after the 2004 tsunami, as a result of policies that required residents to relocate 100 m from the shoreline. In the Yangtze basin, the proportion of the population exposed to floods decreased by  $7\%$ . MODIS probably did not capture increases in urban flood exposure in at least eight countries with rapid urbanization (for example, with annual urbanization greater than  $3\%$ ; Angola, Afghanistan, Cambodia, Namibia, Chad,

Senegal, Sierra Leone and Oman<sup>42</sup>, see double asterisks in Supplementary Table 5).

### Estimated flood exposure 2010–2030

We calculated the population that will be exposed to floods in the near future (2010–2030) in countries with sufficient MODIS observations ( $n = 119$  countries), using the World Resources Institute flood-risk analyser Aqueduct<sup>37</sup>. Across these countries, the flood model (GLOFRIS) estimates that 580 million people were exposed to a 100-year-return-period flood in 2010. By 2030, the World Resources Institute<sup>37</sup> estimates that up to 758 million people will be exposed in the 100-year flood zone, with the additional 179.2 million people being exposed as a result of demographic shifts (116.5 million people) or climate change (50.3 million people; assuming representative concentration pathway (RCP) 8.5), and synergistic climate–land use interactions (12.4 million people). The proportion of the population exposed





**Fig. 4 | Change in the proportion of the population exposed to floods observed from 2000 to 2015 and predicted for 2030 per country.**

**a**, Multiplicative change from 2000 to 2015 in the proportion of the population exposed to observed inundation (equation (6), Methods). **b**, Multiplicative change from 2010 to 2030 in the proportion of the population exposed to floods. (equation (7), Methods). The change ranges (coloured shading) in **a** and **b** are from ref.<sup>7</sup> to facilitate comparison. **c**, Countries where the proportion of the population exposed to floods: (1) grew from 2000 to 2015 (multiplicative change >1.02 in **a** and ≤0.97 in **b**; pink; ‘decreasing’ flood exposure); (2) is

expected to grow from 2010 to 2030 (multiplicative change >1.02 in **b** and ≤0.97 in **a**; blue; ‘new’ flood exposure); (3) grew from 2000 to 2015 and is expected to grow from 2010 to 2030 (multiplicative change >1.02 in **a** and **b**; purple; ‘continuously increasing’ flood exposure); and (4) is expected to remain constant or decrease (multiplicative change ≤1.02 in **a** and **b**; orange; ‘never increasing or little change’ in flood exposure). Countries shown in grey had insufficient flood observations or population uncertainty. Base maps: GADM (Global Administrative Areas) 2018, version 3.6.

to floods is expected to increase globally by 2030, but with variation across countries (Fig. 4b; global-mean increase of 4%, s.d. = 90%) and no sensitivity to the return period (Extended Data Fig. 8). In 57 countries,

the increase in flood exposure is expected to outpace future population growth, especially in Asia and Africa<sup>7</sup>. Although we are already halfway towards these projections, they remain uncertain because of

uncertainty in climate<sup>43</sup> and future population models. The difficulty in predicting changes in migration patterns and country-specific urban development means that increases in future flood exposure could be underestimated where urbanization is rapidly increasing.

We compare the change in the proportion of the population exposed to observed large flood events in the recent past (between 2000 and 2015) and predicted to be exposed in the near future (between 2010 and 2030) for the 106 countries with robust population data, using equation (7) (Methods). We identify countries for which a change in flood exposure greater than the population growth is new or continuously increasing, and for which the change in flood exposure relative to population growth is decreasing or the same. For this classification, 'new' signifies flood exposure increasing more rapidly than population only in the future period; 'decreasing' signifies flood exposure increasing more rapidly than population only in the past period; 'continuously increasing' signifies flood exposure increasing more rapidly than population growth in both time periods; and 'never increasing or little change' signifies flood exposure increasing more rapidly than population growth in neither time period.

Nine regions and 32 countries, spread across four continents, have 'continuously increasing' flood exposure (Fig. 4c, Supplementary Tables 4, 5). Five of these countries (four in Africa plus India) exhibit high continuing increases (more than 20%) in the proportion of the population exposed to floods. Five regions and 25 countries will have 'new' flood exposure, concentrated in Europe and North America, with the highest increases (more than 50%) in the flood-exposed proportion in Oman and Sudan. Although 3 regions and 29 countries have 'decreasing' flood exposure, models still estimate that 2.2 million additional people will be exposed to 100-year-return-period floods by 2030 in those countries. Three regions (Melanesia, Central Asia, and Western Asia (the Middle East)) and 20 countries have seen 'never increasing or little change' in flood exposure.

## Discussion

Our results provide evidence from satellite observations that increases in flood exposure are higher (20%–24% from 2000 to 2015) than previously estimated (2.6% from 1970 to 2010)<sup>7</sup>. We find that the proportion of the population exposed to floods increased in 70 countries, across all continents. This finding is in contrast to previous studies that report increases in only 22 or 55 countries, concentrated in sub-Saharan Africa and Asia<sup>7,10</sup>. We identify additional increases in flood exposure in southern Asia, southern Latin America and the Middle East. Our estimates are higher than previous ones probably because our observations capture floods caused by dam breaks, pluvial events and snowmelt, which are not included in global models. In addition to increased flood exposure in the recent past, we identify 57 countries where exposure is predicted to grow, indicating flood-prone development patterns that place lives and livelihoods at risk.

There are four limitations in our analysis: (1) the incomplete event record, which does not include smaller yet impactful flood events<sup>44</sup>; (2) the limited ability of MODIS to map urban floods; (3) the uncertainty in the spatial population distribution; and (4) the uncertainty in predicting climate extremes. The Emergency Events Database<sup>36</sup> estimates that more than 1.1 billion people were exposed to flood events from 2000 to 2018, 159–208 million more people than estimated by our database. Our study probably underestimates flood exposure trends in rapidly urbanizing countries, owing to uncertainty in satellites and population growth models. The population data used in this study tend to overestimate observed flood exposure<sup>29</sup>, with uncertainties too large to reliably estimate a flood trend for 15 countries (Supplementary Discussion).

Future work could improve flood-exposed population estimates by: (1) incorporating more events (for example, through social media<sup>45</sup>) and satellites over longer time periods or at higher resolution; (2) modelling event extents where satellite temporal coverage is insufficient (for

example, flash floods); (3) assigning return periods to compare trends from observations to models; and (4) improving spatial estimates of the past, present and future global population.

The Global Flood Database provides a catalogue of global spatial flood event data at 250-m resolution, available for public download. These data could aid calibration of flood models and comparison to improve modelled flood hazard and exposure estimates. Identifying human settlement growth in areas of observed inundation could inform adaptation strategies such as mitigation and managed retreat<sup>46</sup>. Flood observations may affect the pricing of financial instruments such as municipal bonds<sup>47</sup> and insurance<sup>48</sup>, and may aid planning for a changing (or already changed) tax base. Population growth in the observed inundated areas is largely due to increased economic development and migration to floodplains. Floodplains may be expanding because of increasing impervious surface area<sup>49</sup> and climatic changes<sup>14</sup>. Increasing flood exposure is also rooted in historical and political processes that produce conditions that may make settling in floodplains the only option for vulnerable populations<sup>50</sup>. Vulnerability analyses, together with the improved flood exposure estimates presented here, should drive investment in flood adaptation directed to the people and places that need it most.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-03695-w>.

- Hallegatte, S., Vogt-Schilb, A., Bangalore, M. & Rozenberg, J. *Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters* Ch. 3, 63–77 (The World Bank, 2016).
- CRED, UNDRR. *Human Cost of Disasters. An Overview of the last 20 years: 2000–2019*. <https://reliefweb.int/report/world/human-cost-disasters-overview-last-20-years-2000-2019> (UNDRR, 2020).
- Jongman, B. et al. Declining vulnerability to river floods and the global benefits of adaptation. *Proc. Natl Acad. Sci. USA* **112**, E2271–E2280 (2015).
- Liu, X. et al. High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. *Nat. Sustain.* **3**, 564–570 (2020).
- Grill, G. et al. Mapping the world's free-flowing rivers. *Nature* **569**, 215–221 (2019); author correction **572**, E9 (2019).
- Ceola, S., Laio, F. & Montanari, A. Satellite nighttime lights reveal increasing human exposure to floods worldwide. *Geophys. Res. Lett.* **41**, 7184–7190 (2014).
- Jongman, B., Ward, P. J. & Aerts, J. C. J. H. Global exposure to river and coastal flooding: long term trends and changes. *Glob. Environ. Change* **22**, 823–835 (2012).
- Winsemius, H. C. et al. Global drivers of future river flood risk. *Nat. Clim. Chang.* **6**, 381–385 (2016).
- Trigg, M. A. et al. The credibility challenge for global fluvial flood risk analysis. *Environ. Res. Lett.* **11**, 094014 (2016); corrigendum **13**, 099503 (2018).
- Tanoue, M., Hirabayashi, Y. & Ikeuchi, H. Global-scale river flood vulnerability in the last 50 years. *Sci. Rep.* **6**, 36021 (2016).
- Formetta, G. & Feyen, L. Empirical evidence of declining global vulnerability to climate-related hazards. *Glob. Environ. Change* **57**, 101920 (2019).
- Hoegh-Guldberg, O. et al. in *Global Warming of 1.5 °C. An IPCC Special Report on the Impacts of Global Warming of 1.5 °C Above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (eds Masson-Delmotte, V. et al.) Ch. 3 (IPCC, 2018).
- Alfieri, L. et al. Global projections of river flood risk in a warmer world. *Earths Futur.* **5**, 171–182 (2017).
- Najibi, N. & Devineni, N. Recent trends in the frequency and duration of global floods. *Earth Syst. Dyn.* **9**, 757–783 (2018).
- Ward, P. J. et al. A global framework for future costs and benefits of river-flood protection in urban areas. *Nat. Clim. Chang.* **7**, 642–646 (2017).
- Kellet, J. & Caravani, A. *Financing Disaster Risk Reduction: A 20 Year Story of International Aid* 60 (GFDRR, ODI, 2013).
- IPCC. *Climate Change 2014: Impacts, Adaptation and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Field, C. B. et al.) (Cambridge Univ. Press, 2014).
- Turner, B. L. et al. A framework for vulnerability analysis in sustainability science. *Proc. Natl Acad. Sci. USA* **100**, 8074–8079 (2003).
- Sampson, C. C. et al. A high-resolution global flood hazard model. *Wat. Resour. Res.* **51**, 7358–7381 (2015).
- Rufat, S., Tate, E., Burton, C. G. & Maroof, A. S. Social vulnerability to floods: review of case studies and implications for measurement. *Int. J. Disaster Risk Reduct.* **14**, 470–486 (2015).



21. Ward, P. J. et al. Usefulness and limitations of global flood risk models. *Nat. Clim. Chang.* **5**, 712–715 (2015).
22. Shastry, A. et al. Small-scale anthropogenic changes impact floodplain hydraulics: simulating the effects of fish canals on the Logone floodplain. *J. Hydrol.* **588**, 125035 (2020).
23. Coltin, B., McMichael, S., Smith, T. & Fong, T. Automatic boosted flood mapping from satellite data. *Int. J. Remote Sens.* **37**, 993–1015 (2016).
24. DeVries, B. et al. Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. *Remote Sens. Environ.* **240**, 111664 (2020).
25. Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422 (2016).
26. Ji, L., Gong, P., Wang, J., Shi, J. & Zhu, Z. Construction of the 500-m resolution daily global surface water change database (2001–2016). *Wat. Resour. Res.* **54**, 10270–10292 (2018).
27. Bernhofen, M. V. et al. A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique. *Environ. Res. Lett.* **13**, 104007 (2018).
28. Policelli, F. et al. in *Remote Sensing of Hydrological Extremes* (ed. Lakshmi, V.) 47 (Springer, 2016).
29. Smith, A. et al. New estimates of flood exposure in developing countries using high-resolution population data. *Nat. Commun.* **10**, 1814 (2019).
30. European Commission, Joint Research Centre & Columbia University, Center for International Earth Science Information Network. *GHS Population Grid, Derived from GPW4, Multitemporal (1975, 1990, 2000, 2015)* [https://ghsl.jrc.ec.europa.eu/ghs\\_pop.php](https://ghsl.jrc.ec.europa.eu/ghs_pop.php) (2015).
31. Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN - Columbia University. *High Resolution Settlement Layer (HRSL)* <https://www.ciesin.columbia.edu/data/hrsl/> (2016).
32. Riahi, K. et al. RCP 8.5—a scenario of comparatively high greenhouse gas emissions. *Clim. Change* **109**, 33–57 (2011).
33. Riahi, K. et al. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Glob. Environ. Change* **42**, 153–168 (2017).
34. Otsu, N. A threshold selection method from gray-level histograms. *Automatica* **11**, 23–27 (1975).
35. Nigro, J., Slayback, D., Policelli, F. & Brakenridge, G. R. *NASA/DFO MODIS Near Real-Time (NRT) Global Flood Mapping Product Evaluation of Flood and Permanent Water Detection* [https://floodmap.modaps.eosdis.nasa.gov/documents/NASAGlobalNREvaluationSummary\\_v4.pdf](https://floodmap.modaps.eosdis.nasa.gov/documents/NASAGlobalNREvaluationSummary_v4.pdf) (2014).
36. Centre for Research on the Epidemiology of Disasters. *Em-Dat: the International Disaster Database* <https://public.emdat.be/> (accessed July 2020).
37. World Resources Institute. *Aqueduct Global Flood Risk Maps* <https://www.wri.org/data/aqueduct-floods-hazard-maps> (2015).
38. Collenteur, R. A., de Moel, H., Jongman, B. & Di Baldassarre, G. The failed-levee effect: do societies learn from flood disasters? *Nat. Hazards* **76**, 373–388 (2015).
39. Mykhnenko, V. & Turok, I. East European cities — patterns of growth and decline, 1960–2005. *Int. Plann. Stud.* **13**, 311–342 (2008).
40. Zaninetti, J.-M. & Colten, C. E. Shrinking New Orleans: post-Katrina population adjustments. *Urban Geogr.* **33**, 675–699 (2012).
41. Birkmann, J. & Fernando, N. Measuring revealed and emergent vulnerabilities of coastal communities to tsunami in Sri Lanka. *Disasters* **32**, 82–105 (2008).
42. United Nations, Department of Economic and Social Affairs, & Population Division. *World Urbanization Prospects: the 2018 Revision*. Report No. ST/ESA/SER.A/420 (United Nations, 2019).
43. Fiedler, T. et al. Business risk and the emergence of climate analytics. *Nat. Clim. Change* **11**, 87–94 (2021).
44. Gaillard, J. C. et al. Alternatives for sustained disaster risk reduction. *Human. Geogr.* **3**, 66–88 (2010).
45. de Bruijn, J. A. et al. A global database of historic and real-time flood events based on social media. *Sci. Data* **6**, 311 (2019).
46. Siders, A. R., Hino, M. & Mach, K. J. The case for strategic and managed climate retreat. *Science* **365**, 761–763 (2019).
47. Painter, M. An inconvenient cost: the effects of climate change on municipal bonds. *J. Financ. Econ.* **135**, 468–482 (2020).
48. Surminski, S. & Oramas-Dorta, D. Flood insurance schemes and climate adaptation in developing countries. *Int. J. Disaster Risk Reduct.* **7**, 154–164 (2014).
49. Blum, A. G., Ferraro, P. J., Archfield, S. A. & Ryberg, K. R. Causal effect of impervious cover on annual flood magnitude for the United States. *Geophys. Res. Lett.* **47**, e2019GL086480 (2020).
50. Weichselgartner, J. & Kelman, I. Geographies of resilience: challenges and opportunities of a descriptive concept. *Prog. Hum. Geogr.* **39**, 249–267 (2015).
51. Tennekes, M. tmap: thematic maps in R. *J. Stat. Softw.* **84**, 6 (2018).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2021

## Methods

### Flood event catalogues

We used the DFO flood event catalogue as the source for identifying dates and approximate locations of 4,712 major flood events since 1985 (as of 31 December 2018; Extended Data Fig. 1a). Other publicly available global flood event catalogues, such as the Emergency Events Database (Em-Dat)<sup>36</sup>, have limited location data at the country level. Mapping entire countries when an event occurs in a small area or crosses borders introduced computational challenges and errors. The DFO database provides spatial estimates of flood locations (for example, points and polygons), not available in Em-Dat, that allow us to filter satellite imagery repositories in focused areas for application of flood detection algorithms. DFO also lists the main flood cause, which we simplified into four categories: dams, heavy rain, snow or ice melt, and tropical storms and surges (Supplementary Tables 6, 8).

The inclusion criteria for DFO (primarily large-media-coverage events, including those covered by FloodList (<http://floodlist.com/>)) and Em-Dat (10 or more flood-related deaths or at least 100 people affected) differ and possibly introduce bias. We compared DFO and Em-Dat events temporally and spatially at the country level to assess the differences. We matched the DFO and Em-Dat events over the study period (2000–2018; during the satellite data record) by using country names and overlapping date periods, using the fuzzyjoin R package<sup>52</sup>.

The number of total flood events in the DFO from 2000 to 2018 ( $n = 3,195$ ) is greater than that in Em-Dat before 2009 ( $n = 3,010$ ), but less than that in Em-Dat after 2009 (Extended Data Fig. 9a). The number of flood events per year in DFO and Em-Dat is positively and significantly correlated over time (Pearson correlation  $r = 0.591$ ,  $P < 0.01$ ), consistent with previous results for 1985–2019 ( $r = 0.636$ ,  $P < 0.001$ )<sup>14</sup>. Spatial comparison reveals that DFO reports more floods than Em-Dat in the US (192 more events), Australia (79 more events) and Russia (31 more events), but fewer events in South America (36 fewer events), Central America (30 fewer events), the Caribbean (20 fewer events) and Africa (166 fewer events; 94 fewer in west Africa; Extended Data Fig. 9b). This comparison between the databases suggests that the DFO represents trends in major flood events over time, but may underrepresent floods in Africa and South America.

### Satellite data and inundation detection algorithm

For historical flood observation, we use the MODIS instrument onboard NASA's Terra and Aqua satellites. MODIS is an optical satellite commonly used for inundated area mapping<sup>26,53–57</sup>, is freely available and has had consistent daily coverage since February 2000 and twice-daily coverage since February 2001. The DFO contains 3,127 eligible flood events that co-occurred with MODIS imagery (Extended Data Fig. 1b).

We used the Google Earth Engine platform<sup>58</sup> to preprocess and apply water detection algorithms to the MODIS images. The polygon areas provided by DFO represent approximate areas affected by the events. Therefore, we selected all HydroBASINS Level 4<sup>59,60</sup> watersheds that intersect with the DFO event polygon as our mapping unit (region of interest) for each event. For each event in the database, we collected and analysed every MODIS image acquired over the selected watersheds during the event date range provided in the DFO. In total, we analysed 12,719 individual MODIS tiles across the 3,127 events (Extended Data Fig. 1).

Terra (MOD09GA/GQ) and Aqua (MYD09GA/GQ) MODIS images used in this study were corrected for atmospheric scattering and absorption to provide estimates of surface reflectance at resolutions of 250 m and 500 m<sup>61</sup>. MODIS data provide reflectance values (stored as digital numbers scaled by 10,000) in the visible (457–670 nm) and near-infrared (841–1,250 nm) wavelengths at 250-m resolution; short-wave-infrared (1,628–2,155 nm) wavelengths commonly used to identify surface water are provided at 500-m resolution. We pan-sharpened the short-wave-infrared band to 250 m using an adapted version of the

corrected reflectance algorithm<sup>62</sup> to match the resolution of other bands.

Estimates of inundation extent were produced at 250-m resolution using thresholding approaches based on an existing algorithm<sup>53</sup>. We produced inundation maps for every event using four versions of the algorithm: 3-day standard, 2-day standard, 3-day Otsu and 2-day Otsu.

The 'standard' versions of the algorithm identify water using fixed threshold values on stored reflectance values (digital numbers) of the short-wave-infrared (SWIR) band (band 7; 1,628–1,652 nm) and an index,  $B2B1_{\text{ratio}}$ , defined as

$$B1B2_{\text{ratio}} = \frac{DN_{\text{NIR}} + 13.5}{DN_{\text{red}} + 1,081.1} \quad (1)$$

where  $DN_{\text{NIR}}$  and  $DN_{\text{red}}$  are the digital numbers of the near-infrared (band 2) and red (band 1; 621–670 nm) bands. A pixel is classified as water via the following:

$$\text{pixel}_{\text{water}} = DN_{\text{red}} < C \wedge B1B2_{\text{ratio}} < K_1 \wedge DN_{\text{SWIR}} < K_2 \quad (2)$$

In the standard algorithm,  $K_1 = 0.7$ ,  $K_2 = 675$  and  $C = 2,027$ . The constants in equations (1) and (2) were determined empirically ( $R^2 = 0.91$ ) using regression discharge data from the US Geological Society (USGS)-gauged river reaches<sup>53</sup>.

The 'Otsu' versions of the algorithm adjust the thresholds by estimating  $K_1$  and  $K_2$  (equation (2)) adaptively for each flood event<sup>34,63</sup>. Otsu thresholding requires a bimodal distribution, in our case representing spectral reflectance of water and non-water, to determine a threshold that minimizes interclass variance (that is, misclassification). We extracted a sample of 2,500 water and non-water pixels (1,250 sampled for each class) from a median composite of MODIS images for each flood event, with clouds removed using the internal cloud-state band. Water and non-water pixels for each flood event were differentiated by matching the flood event year to the permanent water classification for that year from the Joint Research Center global surface-water yearly history dataset<sup>25</sup>. From our sample, interclass variance was calculated as the between sum-of-squares (BSS):

$$BSS_T = \sum_k (DN_{T,k} - \overline{DN}_T)^2 \quad (3)$$

where  $DN_{T,k}$  is the average mean surface reflectance (provided as a digital number) in band  $T$  and class  $k$  defined by a selected threshold.  $BSS_T$  was calculated iteratively across each bin of a bimodal histogram, representing candidate threshold values, for  $B2B1_{\text{ratio}}$  and  $DN_{\text{SWIR}}$ . The maximum  $BSS_T$ , and thus the minimized interclass variance, was selected as a threshold for both  $B2B1_{\text{ratio}}$  and  $DN_{\text{SWIR}}$ , and then applied to equation (2) (Extended Data Fig. 2a, b). Using the flood events that passed quality control (see below), the average Otsu thresholds for  $B2B1_{\text{ratio}}$  and  $DN_{\text{SWIR}}$  were  $K_1 = 0.77$  and  $K_2 = 599$ , respectively (Extended Data Fig. 2c, d). Compared to the standard thresholds, the Otsu method provides threshold estimates that represent global water conditions as opposed to USGS gauge data. Although the Otsu method estimates event-optimized thresholds, the fact that the median Otsu thresholds approximate the standard thresholds confirms that the standard thresholds perform consistently on a global basis.

We use equation (2) to classify each MODIS image over a region of interest and period of a flood event. After classifying each MODIS image, using either standard or Otsu versions, we calculate multiday composites to reduce false detections. Using 3-day composites, a pixel maintains a water classification if at least three observations out of a possible six (at least 50%) were classified as water; 2-day composites require two observations out of four (at least 50%). Reducing images to multiday composites removes misclassifications due to cloud shadow, a common misclassification with water<sup>64</sup>. We did not mask clouds



with the MODIS 1-km internal cloud-state band, because it removed large portions of flooded area detectable under thin or cirrus cloud conditions. To prevent confusion between water and terrain shadows, areas with slopes greater than 5° were masked out of the final classification using a digital elevation model<sup>65</sup>, similarly to other water detection studies<sup>57</sup>.

Inundated pixels are defined as those classified as water following the 3- or 2-day compositing and that lie outside of permanent water defined by the Global Surface Water dataset<sup>25</sup>. In the Global Surface Water dataset, pixels are identified as permanent water when the Landsat observations in 1985–1999 and in 2000–2016 have water presence. After post-processing, each flood event has four data products (3-day standard, 2-day standard, 3-day Otsu and 2-day Otsu), each of which contains four bands: (1) the maximum extent of inundation; (2) the number of days inundated; (3) the number of clear observations; and (4) the proportion of clear observations.

## Evaluating the inundation detection algorithm

To assess the accuracy of the Global Flood Database, we identified 123 flood events with coincident Landsat 5, 7 and 8 imagery at 30-m resolution available within 24 h of the day of maximum inundation and less than 20% cloud cover. Maximum inundation dates were estimated by selecting the day (between the start and end dates for each event) with the largest inundated area estimated by the flood detection algorithm. The 123 flood events used for accuracy assessment span 15 biomes, representing diverse landscape conditions<sup>66</sup> (Extended Data Fig. 3).

The number of sampling points selected in remote sensing analysis can affect map accuracy<sup>60</sup>. We conducted sensitivity analysis to determine the number of validation points required to minimize the variance in precision, recall and overall accuracy. We sampled 500 points for 10 floods events, stratified as 25% in permanent water, 50% in flood water and 25% in non-water regions. Points were randomly subsampled, without replacement, to assess accuracy from 0 to 500 points (Extended Data Fig. 4a). We found that the standard deviation in accuracy fell below 0.1 when 250 or more points were sampled, and therefore chose to sample 250 points per flood event for the remainder of the dataset.

Interpretation of validation points was undertaken by a team of analysts who identified each point as water, non-water or no data, totalling 30,685 validation points. These analysts had access to Landsat images visualized in natural colour, false-colour infrared and two indices that highlight water (the normalized difference vegetation index and the modified normalized difference water index)<sup>56,67</sup> to decide whether each pixel was at least 50% dry or wet. Each validation point was assessed by three separate analysts, with the majority vote determining the class of the validation point.

Classification agreement and errors were calculated by comparing per pixel classes from the produced flood maps to the validation data (Supplementary Table 1). Errors of omission ( $\epsilon_{om}$ ) and commission ( $\epsilon_{com}$ )<sup>56</sup> were calculated as follows:

$$\epsilon_{om} = \frac{f_n}{t_p + f_n}, \quad (4)$$

$$\epsilon_{com} = \frac{f_p}{t_p + f_p}, \quad (5)$$

where  $t_p$  is the count of true positives,  $f_n$  is the count of false negatives and  $f_p$  is the count of false positives.

The 3-day standard algorithm performed best, with an overall accuracy of 83%. 43% of floods had accuracies of more than 90% and 65% had accuracies above 75% (Supplementary Table 2). The standard version of the algorithm was more consistent than the Otsu version. Although Otsu thresholds reduced false detections and increased accuracy in

some events, other events resulted in overpredicted flood extent (Extended Data Fig. 4b). Errors of omission had no clear geographic pattern, whereas errors of commission were inflated at higher latitudes (Extended Data Fig. 5).

## Flood map quality control

To create the final library of flood maps, every map underwent a quality control process to eliminate poor-quality maps and choose the best map between the two thresholding methods. Because the 3-day composite versions of the algorithms had higher accuracy than the 2-day composite versions on average, all final maps were chosen from the 3-day composite results (Extended Data Fig. 4c). Each flood map ( $n = 3,195$ ) was visually inspected to assess whether the map was a suitable representation of flooding. We used a quality control procedure similar to that for the NASA flood detection algorithm<sup>35</sup>. Quality control was completed by analysts, using the metrics summarized here (see Supplementary Table 3 for a complete list). Analysts recorded: (1) whether a flood map mapped area additional to permanent water (from water masks<sup>25,68</sup> or Google Earth); (2) whether a flood map was obscured by clouds; and (3) which version of the algorithm (standard or Otsu) best matched visible water from MODIS imagery for the maximum inundation date. They determined the product to be a useful representation of the flood event if it mapped inundation beyond permanent water and was not largely obscured by clouds. To make quality control decisions, analysts viewed the DFO polygon, all original MODIS imagery for the flood event, the standard flood map, the Otsu flood map, underlying high-resolution satellite imagery from Google Earth and a hyetograph of the 95th percentile of precipitation in the region of interest estimated by the PERSIANN data product<sup>69</sup>.

279 flood events were assessed by at least two separate analysts to calculate intercoder reliability. Analysts agreed on classifying the flood event as “a useful representation of the flood event” (Supplementary Table 3, question 3) for 203 events, representing 73% intercoder reliability. Flood events marked as ‘maybe’ or for which analysts disagreed were quality checked by B.T. to make final decisions.

Some floods of low quality may be present in the database that should not have passed quality control, and local knowledge of any area should be leveraged when using these global data. We encourage users to pair our online catalogue of events and flood dates with the MODIS worldview tool (<https://worldview.earthdata.nasa.gov/>) to visually examine whether additional flood extent could be mapped by downloading individual MODIS images where water is present in only 1–2 observations and therefore underestimated in the 3-day composites. Supplementary Table 9 includes the quality control information for each flood event.

Quality control results yield 913 flood maps determined to be useful representations of flooding (29.4% of the all DFO events that were mapped). Maps that used Otsu thresholding (124 flood extents; 13.6%) were shown to better capture flood extent than those that used standard threshold, but most flood maps used the standard threshold (789 flood extents; 86.4%). The Global Flood Database produced by this study therefore includes 789 maps using the standard threshold and 124 maps using event-specific Otsu thresholds.

A large proportion of flood events from the DFO (2,212 events; 43.1% of events mapped) did not reveal areas of widespread flooding and failed quality control (Extended Data Fig. 6). The top three reasons noted for failing quality control are extreme cloud cover ( $n = 495$ ; 16% of events), no standing water beyond existing permanent water ( $n = 300$ ; 10% of events) and unmapped floods in urban areas ( $n = 44$ ; 1.5% of events). MODIS may fail to capture: (1) rapid, flash flood events; (2) small channels of water below 250-m resolution (for example, flooded streets in urban areas<sup>70</sup>); (3) inundation below dense canopy cover (for example, greater than 60%)<sup>71</sup>; and (4) maximum inundation if the event catalogue start and end dates are not inclusive of the peak flood day.

### Estimating observed flood exposure, recent past

To examine flood exposure trends, we use the multiplicative change in the proportion of the population exposed to floods between 2000 and 2015<sup>7</sup>:

$$\text{change}_{\text{fe}}^{2000-2015} = \frac{p_{\text{fe}}^{2015}/p_{\text{tot}}^{2015}}{p_{\text{fe}}^{2000}/p_{\text{tot}}^{2000}} \quad (6)$$

Therefore,  $\text{change}_{\text{fe}}^{2000-2015} = 1.35$  is equivalent to a 35% increase;  $\text{change}_{\text{fe}}^{2000-2015} = 1$  (that is, no change) occurs when the total population and flood-exposed population increase or decrease at the same rate.

Each country's statistic is calculated individually (Supplementary Table 5) and the global mean is an average across countries (weighting each country equally). We also estimate the change in the proportion of the population exposed to floods for distinct flood types (Supplementary Table 8) and for the five basins with the largest total population exposed to floods in our archive (UK, Indus, Ganges-Brahmaputra, Mekong and Yangtze; Supplementary Table 7).

To estimate the global flood-exposed population, we calculated the maximum inundated area across the Global Flood Database between 2000 and 2018. This observed inundated area was intersected with the GHSL for years 2000 and 2015 to calculate the flood-exposed population. GHSL was selected because of its global availability over time with a consistent method, matching resolution of MODIS (250-m pixels) and better accuracy compared to other globally available population data<sup>72-75</sup>. GHSL allocates population from census data (within several years of 2000 and 2015) according to the intensity of built-up-area estimates from Landsat (for approximately 2000 and 2015). Other global gridded population datasets either use inconsistent methods over time (for example, Landsat) or inflate estimates in rural areas because population is not allocated on the basis of built-up area<sup>29,77</sup>.

GHSL population estimates have multiple sources of error, including census estimates, incorrect estimation of built-up area or failing to distribute population in rural areas where forest cover obscures built-up area. Unfortunately, neither the Gridded Population of the World (GPW) nor GHSL datasets provide uncertainty estimates<sup>72</sup>. To understand potential sources of error, we conducted a sensitivity analysis on the flood-exposed population in 2015 estimated by GHSL compared to a higher-resolution dataset, HRSL (Supplementary Discussion). HRSL was selected as a second dataset for its high resolution (30-m) and near global representation ( $n = 183$  countries).

We found that the global-mean bias of HRSL to GHSL was 0.67 (s.d. = 0.40), indicating that GHSL systematically predicts higher flood-exposed populations. The bias of GHSL to predict more exposed population was not constant by region; bias was three times as high in Africa compared to Europe (Supplementary Discussion, Extended Data Fig. 7). We estimate all absolute numbers of exposed population in a range, using upper and lower bounds estimated from the two population datasets. Countries ( $n = 15$ ) for which the potential population error spread was higher than the flood trend are not included in trend analyses (Supplementary Table 5, single asterisks).

Owing to the potential noise of scattered singular flood pixels, especially along coastlines, which contain mixed pixels at the ocean-land interface, we removed isolated pixels (not connected to at least two other pixels) for area and population calculations. This reduced the population exposure count globally by approximately 20 million people, but did not change the results of the comparison to global flood models or trends.

The flood-exposed population was estimated per country by summing populations residing in the observed floodplain for years 2000 and 2015. Countries with a ratio of flood maps to total known flood events from the DFO of less than 0.13 (the 50th percentile across all countries) were marked as having insufficient data ( $n = 86$  countries;

Extended Data Fig. 6c), leaving 119 countries for this trend analysis. Country estimates of the population inundated in at least one observed flood event from 2000 to 2018 are significantly correlated with flood exposure estimates from GLOFRIS<sup>78</sup> for the 100-year return period ( $r = 0.89$ ,  $P < 0.001$ ; Extended Data Fig. 10). These results suggest that the distribution of the flood-exposed population recorded in the Global Flood Database is consistent with results from a flood model, and that the data may be used to compare past and future trends.

### Estimating modelled flood exposure, near-term future

Estimates of the population exposed to future floods is taken from Aqueduct<sup>37</sup>, for each country available in the Global Flood Database and for which robust population data are available based on uncertainty analyses ( $n = 106$ ). These data are from the output of GLOFRIS, which uses an average of five climate model outputs coupled to a hydrologic and hydraulic model. We used the RCP 8.5 climate model results and the 100-year flood zone, which is consistent with other flood exposure studies<sup>7,10,11,79</sup>, and the SSP2 socioeconomic pathways scenario (which predicts future growth will follow historical patterns)<sup>33</sup>. Flood exposure estimates for 2030 remain uncertain because climate models exhibit high uncertainty for extreme events and often disagree on precipitation trends<sup>43</sup>. The multiplicative change in the proportion of the population at risk of flood exposure is

$$\text{change}_{\text{fe}}^{2010-2030} = \frac{p_{\text{fe}}^{2030}/p_{\text{tot}}^{2030}}{p_{\text{fe}}^{2010}/p_{\text{tot}}^{2010}} \quad (7)$$

where  $p_{\text{fe}}^{2030}$  assumes RCP 8.5 and SSP2, and  $p_{\text{tot}}^{2030}$  assumes SSP2. We assessed the sensitivity of our choice of the 100-year flood zone and found little variation in trends in the population at risk of flood exposure across return periods (Extended Data Fig. 8a-d, Supplementary Discussion).

We use estimates of flood-exposure population from Aqueduct<sup>80</sup> and summarize the methods here. The 2010 flood-exposure population data in this product were estimated by intersecting the GLOFRIS inundated area with the Landsat 2010 gridded population<sup>81</sup> corrected by the SSP2 2010 population cell estimates. 2030 population estimates in rural areas downscale SSP2 country estimates proportional to the 2010 Landsat distribution. 2030 urban population estimates downscale SSP2 country projections, using projected urban land use from the Netherlands Environmental Assessment Agency 2UP model<sup>82</sup> and local suitability for population growth. This 2030 projection does not take into account urban-urban or urban-rural migration patterns, or differentiate population growth suitability per country in protected areas or flood zones (which could be high in the Global South)<sup>83</sup>. Future flood-risk estimates may be overestimated in rural areas and underestimated in urban areas, which would mean that flood exposure trends reported here are probably underestimated in rapidly urbanizing regions.

### Data availability

The MODIS Collection 6 datasets analysed here are available in the NASA LP DAAC at the USGS EROS Center (<https://lpdaac.usgs.gov/products/mod09gav006/>, <https://lpdaac.usgs.gov/products/mod09gqv006/>) and are mirrored in the Google Earth Engine data catalogue ([https://developers.google.com/earth-engine/datasets/catalog/MODIS\\_006\\_MOD09GA](https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MOD09GA), [https://developers.google.com/earth-engine/datasets/catalog/MODIS\\_006\\_MYD09GQ](https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MYD09GQ)). The MODIS NRT Global Flood Product is available in the NASA LANCE Near Real-Time Data and Imagery service (<https://earthdata.nasa.gov/earth-observation-data/near-real-time/mcdwd-nrt>). The Landsat 5 TM, 7 ETM and 8 OLI surface reflectance products used for the accuracy assessment are available from USGS (<https://earthexplorer.usgs.gov/>) and are mirrored in the Google Earth Engine data catalogue (<https://developers.google.com/>



# Article

earth-engine/datasets/catalog/LANDSAT\_LT05\_C01\_T1\_SR, [https://developers.google.com/earth-engine/datasets/catalog/LANDSAT\\_LE07\\_C01\\_T1\\_SR](https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LE07_C01_T1_SR), [https://developers.google.com/earth-engine/datasets/catalog/LANDSAT\\_LC08\\_C01\\_T1\\_SR](https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_SR)) The datasets generated for this study from the Global Flood Database are available on the Cloud to Street website (<http://global-flood-database.cloudtostreet.ai>) and are mirrored in Google Earth Engine ([https://developers.google.com/earth-engine/datasets/catalog/GLOBAL\\_FLOOD\\_DB\\_MODIS\\_EVENTS\\_V1](https://developers.google.com/earth-engine/datasets/catalog/GLOBAL_FLOOD_DB_MODIS_EVENTS_V1)). Supplementary Tables provide summary estimates for each event, and all data may be downloaded from <http://global-flood-database.cloudtostreet.ai/>. Source data are provided with this paper.

## Code availability

Google Earth Engine's web interface allows the flood mapping algorithm defined in equations (1) and (2) to be applied on any MODIS images. Code to make all figures and flood maps are publicly available at [https://github.com/cloudtostreet/MODIS\\_GlobalFloodDatabase](https://github.com/cloudtostreet/MODIS_GlobalFloodDatabase).

52. Robinson, D., Bryan, J. & Elias, J. *fuzzyjoin* version 0.1.6, <https://cran.r-project.org/web/packages/fuzzyjoin/fuzzyjoin.pdf> (2020).
53. Brakenridge, R. & Anderson, E. in *Transboundary Floods: Reducing Risks Through Flood Management* (eds Marsalek, J. et al.) 1–12 (Springer, 2006).
54. Sakamoto, T. et al. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sens. Environ.* **109**, 295–313 (2007).
55. Islam, A. S., Bala, S. K. & Haque, M. A. Flood inundation map of Bangladesh using MODIS time-series images: flood inundation map of Bangladesh. *J. Flood Risk Managem.* **3**, 210–222 (2010).
56. Boschetti, M., Nutini, F., Manfron, G., Brivio, P. A. & Nelson, A. Comparative analysis of normalised difference spectral indices derived from modis for detecting surface water in flooded rice cropping systems. *PLoS One* **9**, e88741 (2014).
57. Klein, I., Dietz, A., Gessner, U., Dech, S. & Kuenzer, C. Results of the global WaterPack: a novel product to assess inland water body dynamics on a daily basis. *Remote Sens. Lett.* **6**, 78–87 (2015).
58. Gorelick, N. et al. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
59. Lehner, B. & Grill, G. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **27**, 2171–2186 (2013).
60. Congalton, R. G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **37**, 35–46 (1991).
61. Vermote, E. F., El Saleous, N. Z. & Justice, C. O. Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sens. Environ.* **83**, 97–111 (2002).
62. Gumley, L., Desclotres, J. & Schmaltz, J. *Creating Reprojected True Color MODIS Images: A Tutorial* [https://ftp.ssec.wisc.edu/pub/willemm/Creating\\_Reprojected\\_True\\_Color\\_MODIS\\_Images\\_A\\_Tutorial\\_process.pdf](https://ftp.ssec.wisc.edu/pub/willemm/Creating_Reprojected_True_Color_MODIS_Images_A_Tutorial_process.pdf) (Univ. Wisconsin-Madison, 2010).
63. Donchyts, G., Schellekens, J., Winsemius, H., Eisemann, E. & van de Giesen, N. A. 30 m resolution surface water mask including estimation of positional and thematic differences using Landsat 8, SRTM and OpenStreetMap: a case study in the Murray-Darling Basin, Australia. *Remote Sens.* **8**, 386 (2016).
64. Verpoorter, C., Kutser, T. & Tranvik, L. Automated mapping of water bodies using Landsat multispectral data: automated mapping of water bodies. *Limnol. Oceanogr. Methods* **10**, 1037–1050 (2012).
65. Danielson, J. J. & Gesch, D. B. *Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010)*. Report No. OFR 2011-1073 (US Geological Survey, 2011).
66. Dinerstein, E. et al. An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* **67**, 534–545 (2017).
67. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **27**, 3025–3033 (2006).
68. Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P. & Sohlberg, R. A. A new global raster water mask at 250 m resolution. *Int. J. Digit. Earth* **2**, 291–308 (2009).
69. Ashouri, H. et al. PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* **96**, 69–83 (2015).
70. Chini, M. et al. Sentinel-1 InSAR coherence to detect floodwater in urban areas: Houston and hurricane Harvey as a test case. *Remote Sens.* **11**, 107 (2019).
71. Hawker, L. et al. Comparing earth observation and inundation models to map flood hazards. *Environ. Res. Lett.* **15**, 124032 (2020).
72. Leyk, S. et al. The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth Syst. Sci. Data* **11**, 1385–1409 (2019).
73. Archila Bustos, M. F., Hall, O., Niedomysl, T. & Ernstson, U. A pixel level evaluation of five multitemporal global gridded population datasets: a case study in Sweden, 1990–2015. *Popul. Environ.* **42**, 255–277 (2020).
74. Chen, R., Yan, H., Liu, F., Du, W. & Yang, Y. Multiple global population datasets: differences and spatial distribution characteristics. *ISPRS Int. J. Geoinf.* **9**, 637 (2020).
75. Pesaresi, M. et al. *Operating Procedure for the Production of the Global Human Settlement Layer from Landsat Data of the Epochs 1975, 1990, 2000, and 2014*. Report No. JRC97705 (Publications Office of the European Union, 2016).
76. Bhaduri, B., Bright, E., Coleman, P. & Urban, M. L. LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal* **69**, 103–117 (2007).
77. Center for International Earth Science Information Network (CIESIN), Columbia University. *Gridded Population of the World, Version 4 (GPWv4): Population Density* <https://doi.org/10.7927/H4NP22DQ> (SEDAC, 2016; accessed August 2019).
78. Ward, P. J. et al. Assessing flood risk at the global scale: model setup, results, and sensitivity. *Environ. Res. Lett.* **8**, 044019 (2013).
79. Irvine, P. J., Srivier, R. L. & Keller, K. Tension between reducing sea-level rise and global warming through solar-radiation management. *Nat. Clim. Chang.* **2**, 97–100 (2012).
80. Ward, P. J. et al. *Aqueduct Floods Methodology* (World Resources Institute, 2020).
81. Bright, E. A., Coleman, P. R., Rose, A. N. & Urban, M. L. *LandScan 2010* <https://landscan.ornl.gov/landscan-datasets> (2011).
82. van Huijstee, J., van Bommel, B., Bouwman, A. & van Rijn, F. *Towards an Urban Preview: Modelling Future Urban Growth With 2UP* <https://www.pbl.nl/en/publications/towards-an-urban-preview> (PBL Netherlands Environmental Assessment Agency, 2018).
83. UN Habitat. *Informal Settlements* <https://unhabitat.org/habitat-iii-issue-papers-22-informal-settlements> (2016).

**Acknowledgements** We thank students at the University of Washington for quality checking and validating the Global Flood Database: S. Lau, C. C. Wu, P. K. Kamaluddin, W. Pietsch, S. Lau, G. J. Kim, H. Tran and J. Crinkey. We thank students from the University of Texas at Austin, including C. Norton. We thank the Cloud to Street staff for quality checking additional portions of the database: S. Weber and E. Glinskis. We thank B. L. Turner II, U. Lall, V. Lakshmi, P. Kareiva, C. Tuholkse, J. Doss-Gollin, D. Farnham and A. Massmann for providing comments on the manuscript. Funding for this project was provided by a Google Earth Engine Research Award and NASA award 80NSSC18K0426, 16-GE016-0068, Integrating Global Remote Sensing and Modeling Systems for Local Flood Prediction and Impact Assessment. This research made use of the LandScan 2010 High Resolution global Population Data Set, copyrighted by UT-Battelle, LLC, operator of Oak Ridge National Laboratory under contract no. DE-AC05-00OR22725 with the United States Department of Energy. The United States Government has certain rights in this dataset. Neither UT-Battelle, LLC nor the United States Department of Energy, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the dataset.

**Author contributions** B.T. conceived the study and led the population exposure analysis, writing and project administration. J.A.S. conceived the study and algorithm, designed the validation, led the data curation and contributed to writing and editing. C.K. led algorithm accuracy and validation analysis and quality control. C.S.D. contributed to algorithm development and data visualization and curation. A.J.K. contributed to quality control, writing and editing. G.R.B. contributed to editing and data curation of the event database. T.A.E. contributed to data curation. D.A.S. contributed to quality control and data visualization.

**Competing interests** Two of the authors (B.T. and C.D.) are employed by and hold stock in a company, Cloud to Street, that sells flood observation and satellite monitoring technology. The data in this paper are free and open, and the company expects no direct financial benefit from this study.

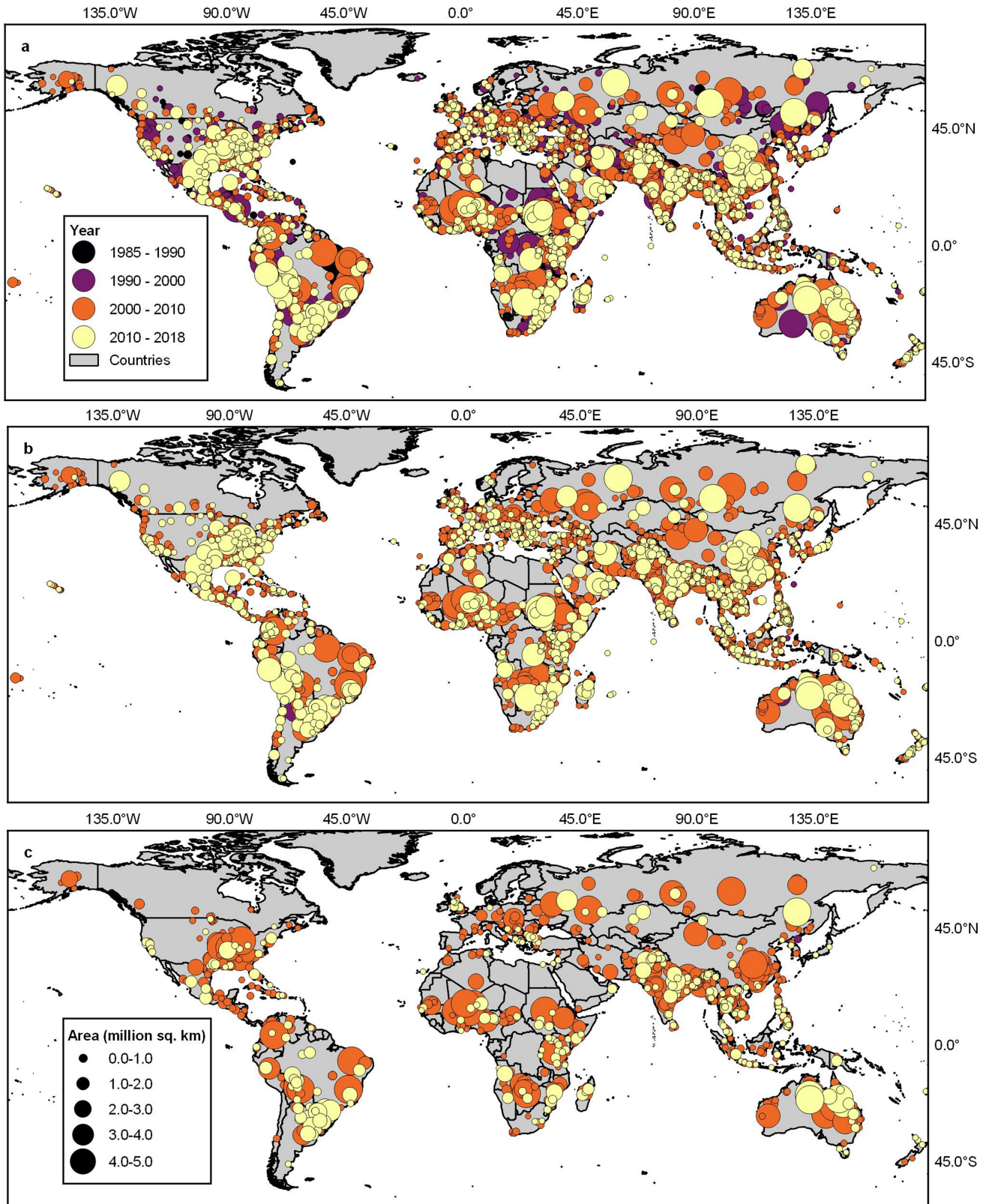
## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-03695-w>.

**Correspondence and requests for materials** should be addressed to B.T.

**Peer review information** Nature thanks Brenden Jongman, Albert van Dijk and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

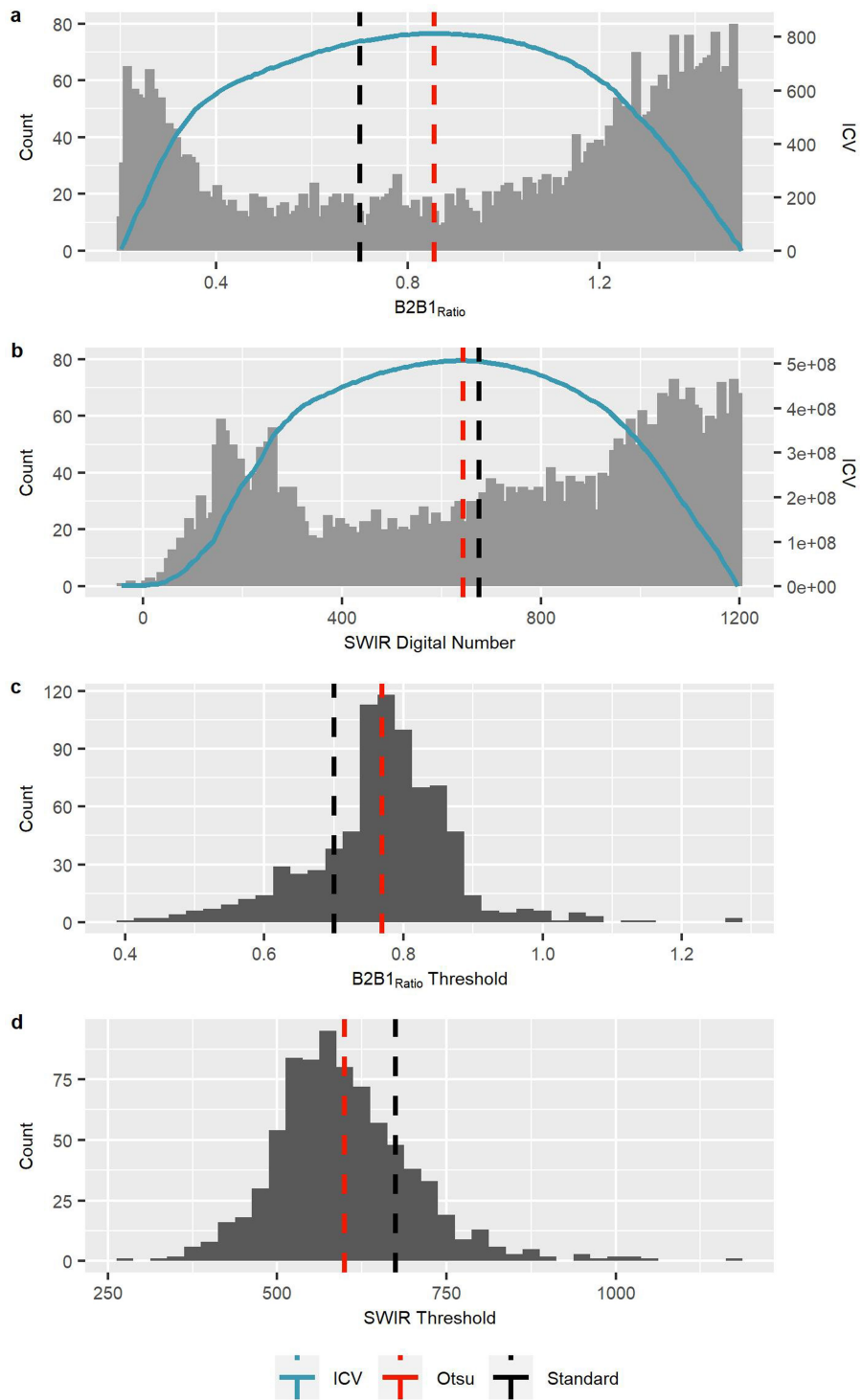
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Global distribution of flood events catalogued by the DFO and the Global Flood Database. a-c,** Flood events in the DFO from 1 January 1985 to 31 December 2018 ( $n = 4,712$ ; **a**), in the DFO and coincident with

MODIS imagery ( $n = 3,127$ ; **b**) and that passed the quality control evaluation for the Global Flood Database ( $n = 913$ ; **c**). Base map: US government Large Scale International Boundaries (LSIB) Polygons (2017).

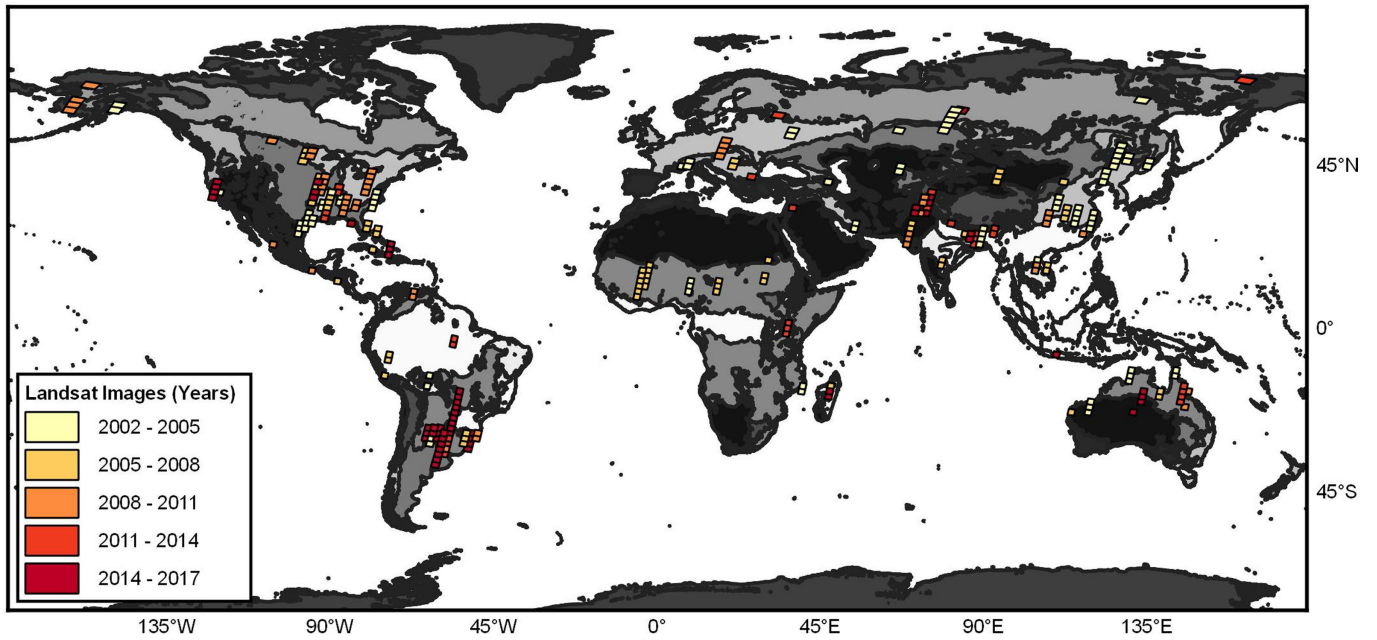




**Extended Data Fig. 2 | Example of bimodal histograms used to calculate adaptive thresholds for water classifications that approximate, on average, the standard versions of the water classification thresholds.**

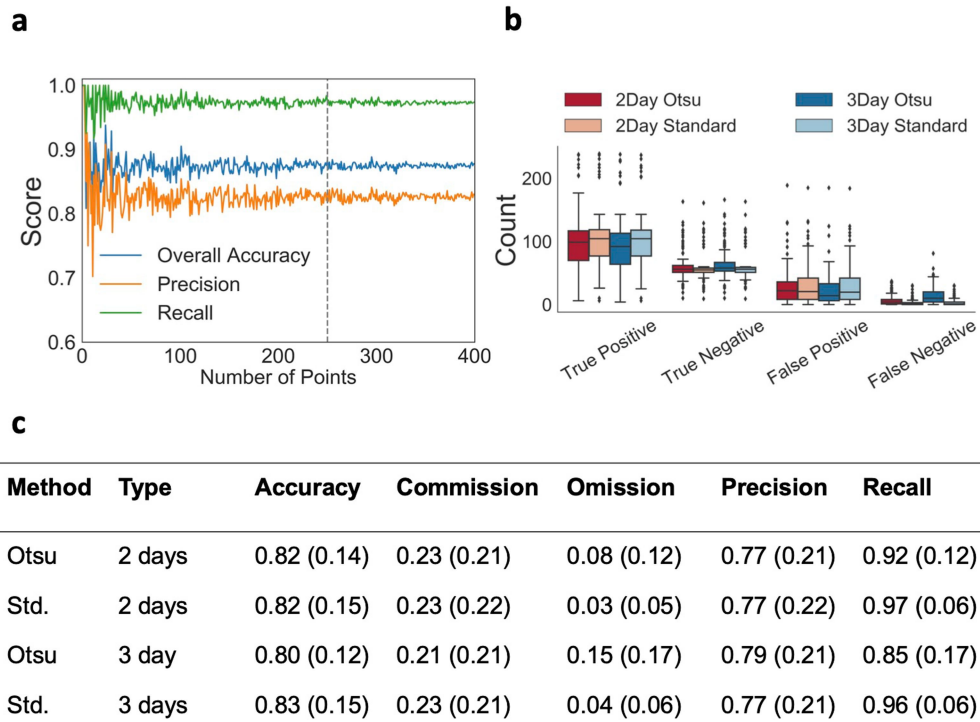
**a, b,** Example bimodal histograms (left axes) with interclass variance (ICV; blue lines, right axes) extracted from MODIS imagery used to determine optimal thresholds for  $B2B1_{ratio}$  ( $K_1$ ; **a**) and  $DN_{SWIR}$  ( $K_2$ ; **b**). The dashed red and black lines reflect the estimated Otsu and standard thresholds, respectively.

**c, d,** Distribution of estimated Otsu thresholds calculated for each flood event across the Global Flood Database ( $n = 913$ ), for  $B2B1_{ratio}$  ( $K_1$ ; **c**) and  $DN_{SWIR}$  ( $K_2$ ; **d**). The average Otsu threshold across the Global Flood Database for  $B2B1_{ratio}$  ( $K_1 = 0.77$ ; dashed red line in **c**) and  $DN_{SWIR}$  ( $K_2 = 599$ ; dashed red line in **d**) are comparable to the standard thresholds ( $K_1 = 0.70$ , dashed black line in **a**;  $K_2 = 675$ , dashed black line in **b**).



**Extended Data Fig. 3 | Distribution of Landsat 5, 7 and 8 imagery.** The Landsat images used to evaluate 123 flood events in the Global Flood Database are globally distributed, cover 15 biomes<sup>66</sup> and span 15 years. Flood events for validation were selected on the basis of availability of Landsat imagery. Eligibility of imagery included conditions that Landsat imagery occurred

within 24 h of the maximum day of inundation, intersected the flood area and had less than 20% cloud cover. Biome shapefile for the base map (grey shading; <https://ecoregions.appspot.com/>) indicates varying vegetation patterns that may influence the accuracy of our remote sensing algorithm.

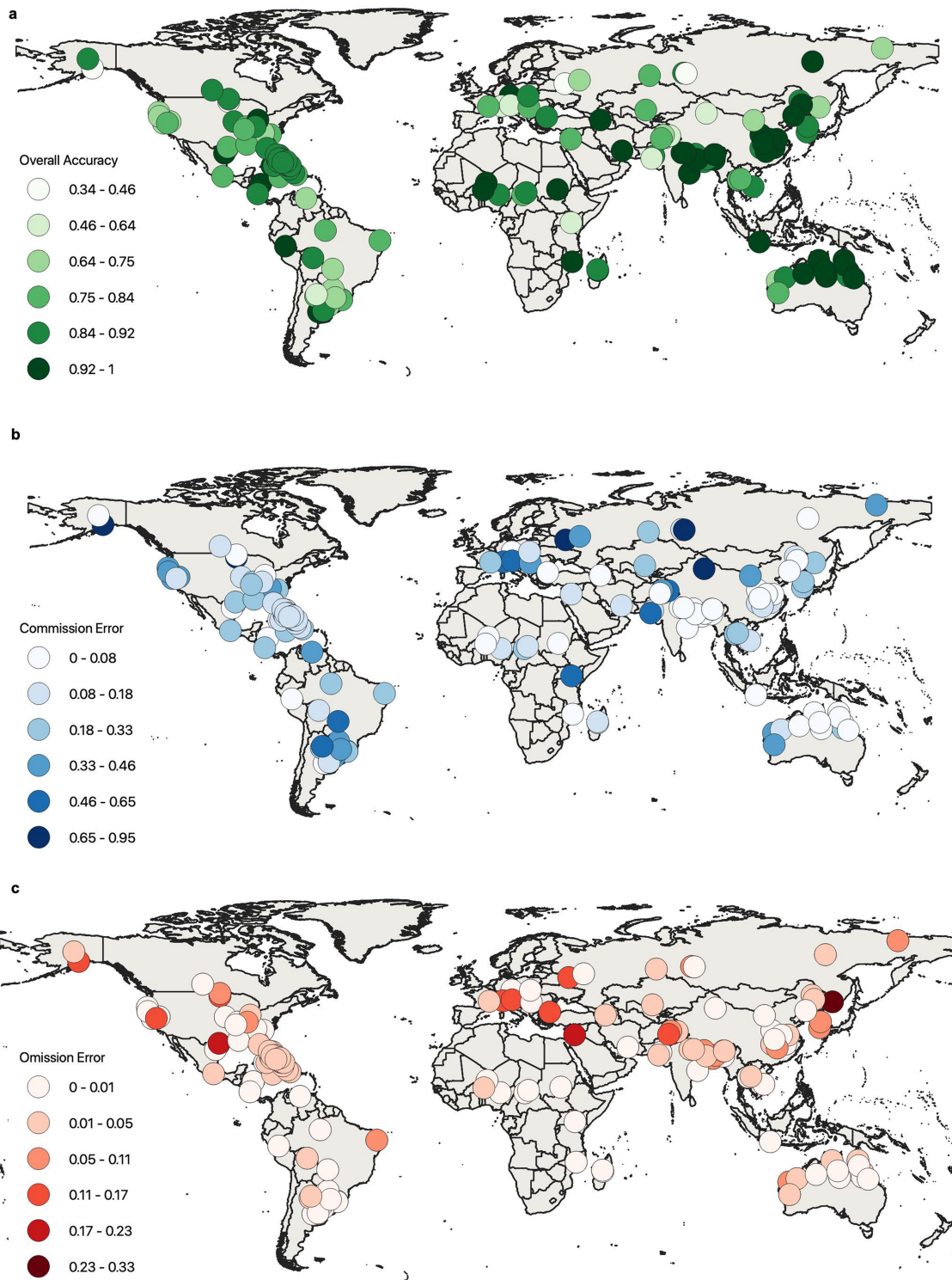


**Extended Data Fig. 4 | Global Flood Database accuracy metrics.**

**a**, Sensitivity plot of accuracy metrics with random sampling of 500 points for 10 flood events. **b**, Error analysis showing the distribution of true positive ( $t_p$ ), true negative ( $t_n$ ), false positive ( $f_p$ ) and false negative ( $f_n$ ) rates for each of the four water detection methods, where the centre line represents the median, the hinges represent the 25th and 75th percentiles, the upper whisker extends

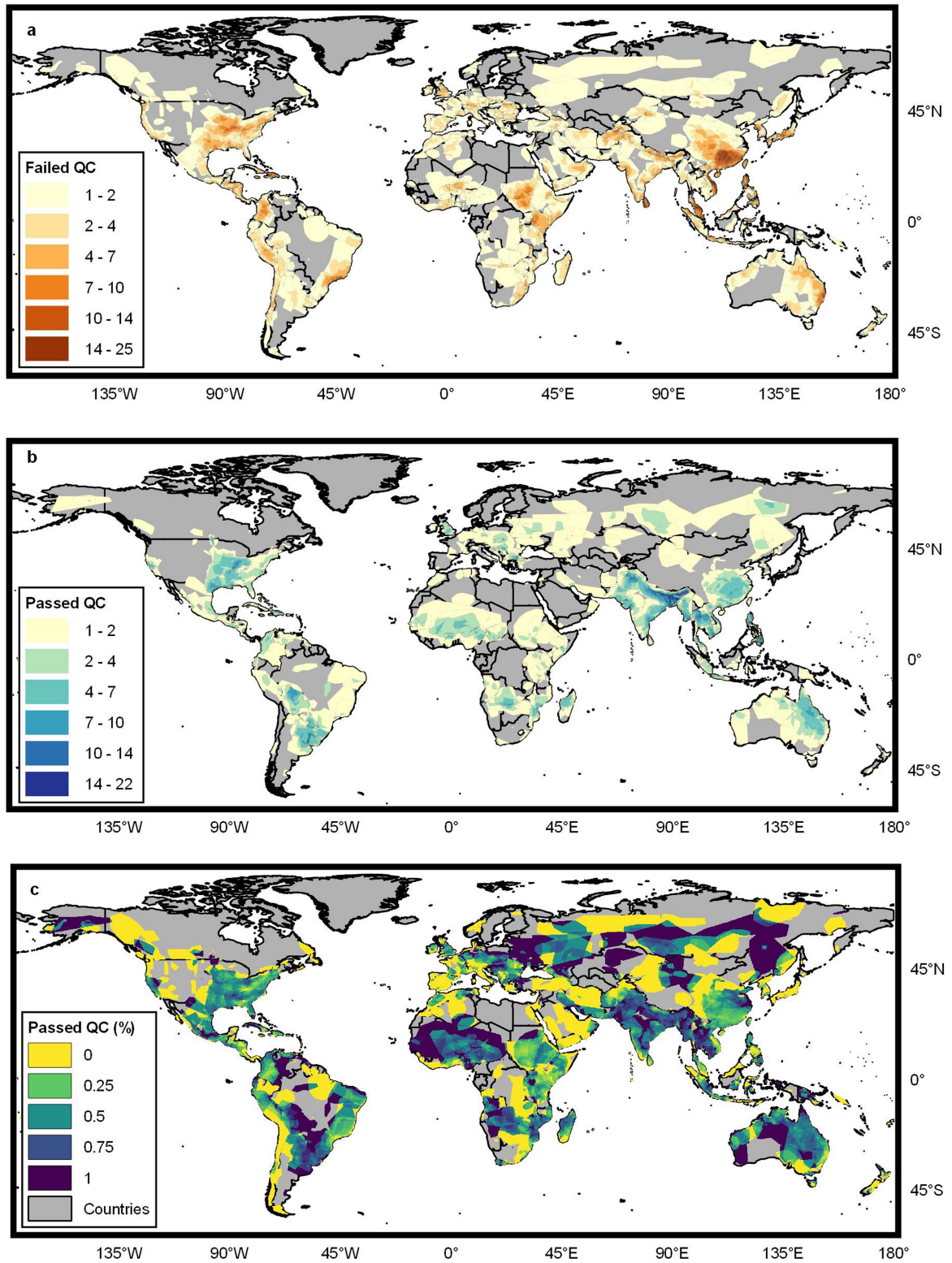
from the hinge to the largest value no further than  $1.5 \times \text{IQR}$  (interquartile range) and the lower whiskers extends from the hinge to the smallest value no further than  $1.5 \times \text{IQR}$ , and the dots show outlier points outside the whisker range. **c**, Accuracy statistics, given as the mean and standard deviation (s.d), summarized for each thresholding method and image composite choice (metrics per event in Supplementary Table 2).





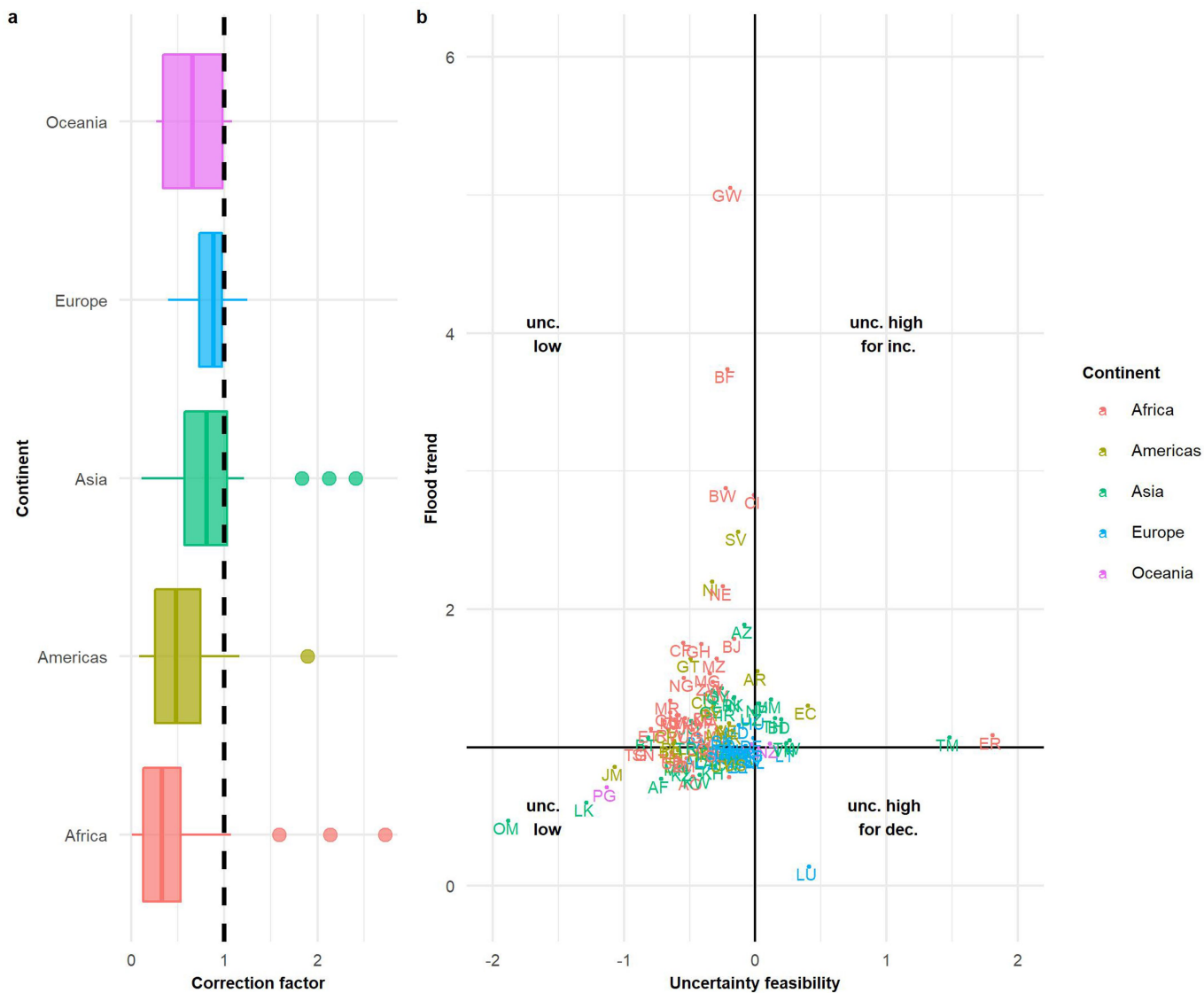
**Extended Data Fig. 5 | Global distribution of accuracy metrics based on 123 flood events from the Global Flood Database. a, b,** Overall accuracy (a) is consistently distributed across the globe, whereas errors of commission (b)

are inflated at higher latitudes and errors of omission (c) are lower than errors of commission and have no clear spatial pattern. Base maps: GADM (Global Administrative Areas) 2018, version 3.6.



**Extended Data Fig. 6 | Results of the quality control assessment.** Coverage of the Global Flood Database is well represented in southeast USA, Central America, South America, southeast Asia, Australia, west Africa and east Africa.

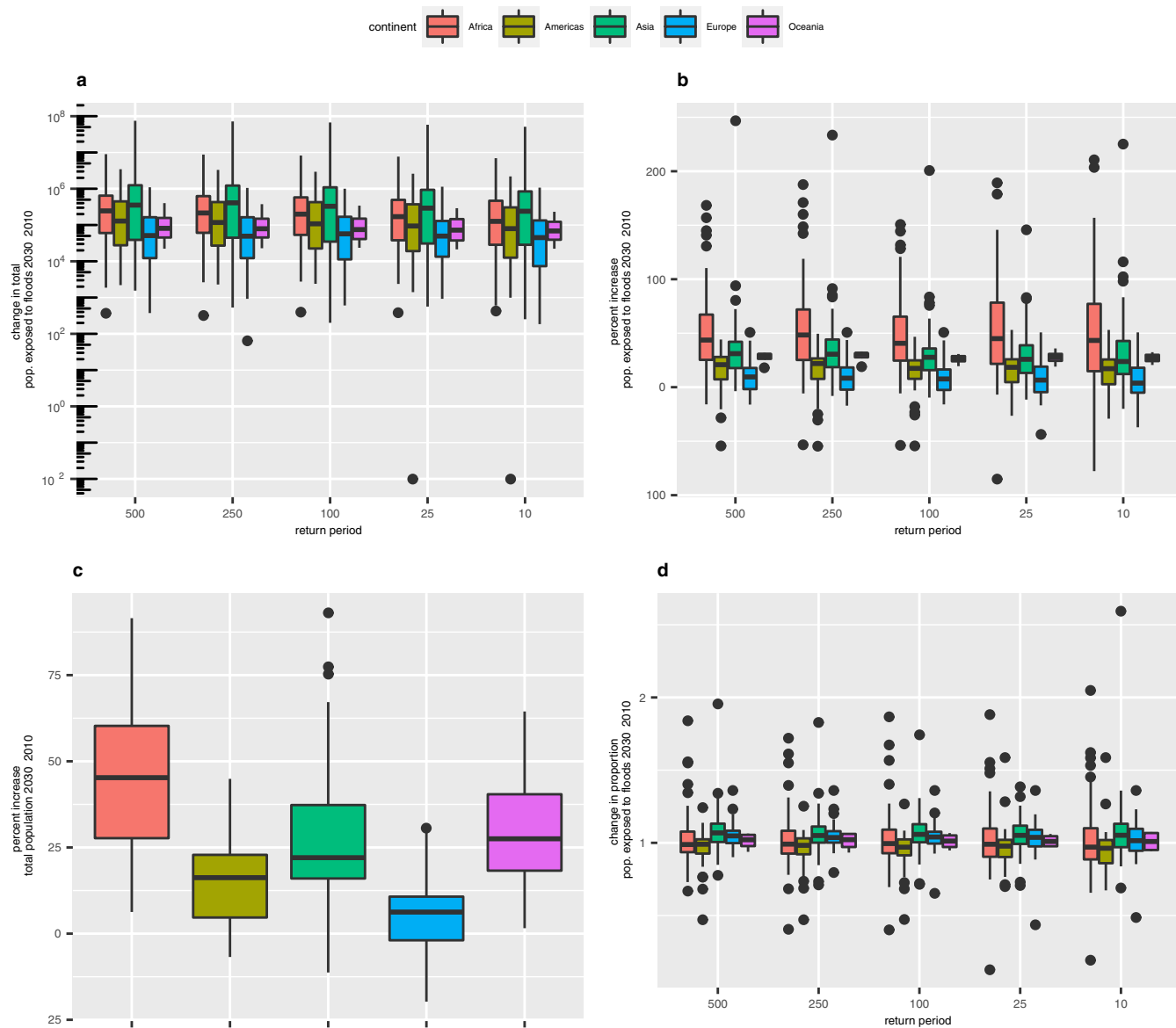
**a-c,** Counts of flood events that passed (**b**) or failed (**a**) quality control (QC), and the proportion of events that passed as a ratio of total flood events (**c**). Base map: US government LSIB Polygons (2017).



**Extended Data Fig. 7 | Population uncertainty analysis. a,** Correction factors determined as the ratio of HRSL flood exposure to GHSL flood exposure per continent, where the centre line represents the median, the hinges represent the 25th and 75th percentiles, the upper whisker extends from the hinge to the largest value no further than  $1.5 \times \text{IQR}$  (interquartile range) and the lower whiskers extends from the hinge to the smallest value no further than  $1.5 \times \text{IQR}$ , and the dots show points outside the whisker range. **b,** Uncertainty

analysis for each country plotted against flood exposure trends (2000–2015). Countries in the upper right quadrant represent regions with plausibly high uncertainty in population estimates, such that we are not confident in their flood exposure trends (unc., uncertainty; inc., increasing trends; dec., decreasing trends). Countries are labelled by their ISO 1366 standard Alpha-2 two-letter country codes (from LSIB) and coloured by continent.

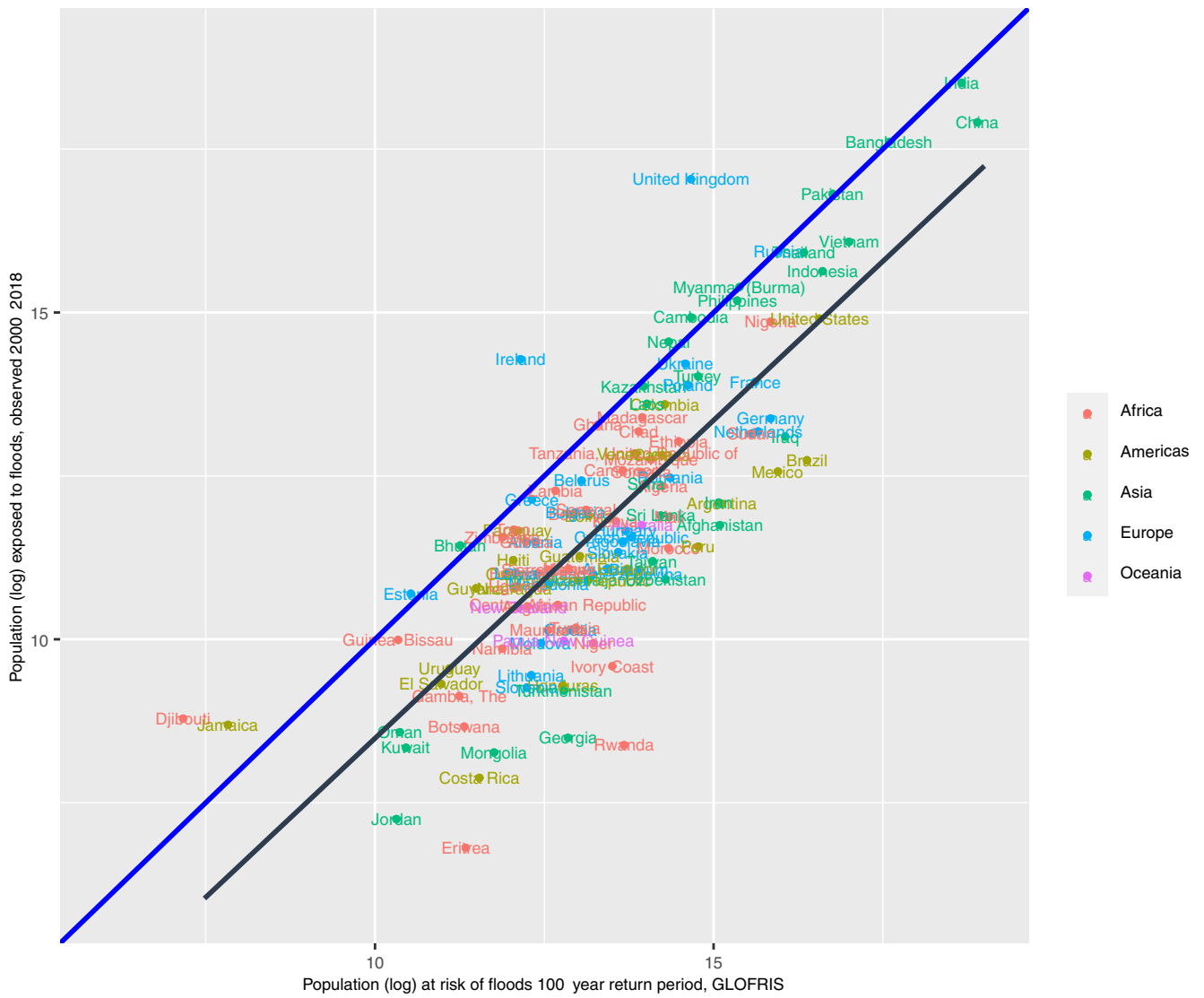




**Extended Data Fig. 8 | Sensitivity analysis of return periods to changes in flood exposure, 2010–2030. a,** Change in total population exposed to floods (logarithmic scale; 2030 minus 2010), for each return period, summarized by continent. **b,** Percentage increase in the population exposed to floods from 2010 to 2030, for each return period. **c,** Percentage increase in total population from 2010 to 2030 for each continent. **d,** Multiplicative change in the

proportion of the population exposed to floods (equation (7), Methods). For all boxplots (a–d), the centre line represents the median, the whiskers represent the 25th and 75th percentile, the upper whisker extends from the hinge to the largest value no further than 1.5\*IQR (interquartile range) and the lower whiskers extends from the hinge to the smallest value no further than 1.5\*IQR, and the dots show points outside the whisker range.





**Extended Data Fig. 10 | Comparison for each country of the population exposed to floods in a 100-year event (using 2010 climate and population) determined from GLOFRIS and an estimate of the total population exposed from the Global Flood Database, 2000–2018.  $r = 0.89, P < 0.001$ . The blue line**

represents parity between the population determined from GLOFRIS and from the Global Flood Database line; the black line is the linear regression line. Country names are coloured by continent.