Supporting Information

Ward et al. 10.1073/pnas.1409822111

SI Materials and Methods

The impacts of flooding were calculated using the global flood impact module of GLOFRIS, described in ref. 1. For clarity, we also briefly describe the data and methods used in this impacts module in the following paragraphs. Note that in this paper we used updated versions of the exposure datasets.

Affected Population and Affected GDP. In GLOFRIS, affected population and GDP are estimated using maps showing downscaled population and GDP. For this study, we used maps derived from the IMAGE model (Integrated Model to Assess the Global Environment) (2, 3), which were further downscaled to a horizontal resolution of 30 × 30 arcseconds using LandScan population maps (4). The affected population and GDP were calculated as the sum of the population or GDP located in areas that are shown as flooded in the inundation maps.

Urban Damage. This risk indicator provides an estimate of urban damage, and the calculation uses a map of asset values in urban areas to represent economic exposure and a stage-damage function to represent vulnerability. The asset value map is based on a land-use map, together with an estimate of urban asset values per square kilometer. The land-use data are taken from the HYDE database (5). For each grid cell, HYDE shows the fractional area of land with urban land cover; the horizontal resolution is 5 × 5 arcminutes. Using these data, we calculated the urban area per grid cell at a horizontal resolution of 30 × 30 arcseconds. The next step was to assign an economic value to the urban area per square kilometer data, which was carried out using the method described in ref. 6. Vulnerability was represented by applying a stage-damage function. The function used here is the average of the high and low urban density land-class functions used in the Damagescanner model (7, 8) and can be found in ref. 1. In this paper we specifically estimated relative anomalies in risk during the different phases of ENSO, and hence the choice of the function is not so important. For future analyses looking at absolute risk estimates it is important to account for spatial variations in vulnerability (9).

SI Discussion

Validation of Hydrological and Hydraulic Models. The discharge results of PCR-GLOBWB have been validated in previous studies against observed data on mean monthly discharge (10) and peak annual discharge (1), and the flood volume results of DynRout have been validated against GRACE satellite data of terrestrial water storage (11). In general, the models have been found to show fair to good performance.

In terms of simulating peak annual discharge, the model was previously validated against observed daily discharge data from the Global Runoff Data Centre (GRDC; www.bafg.de/cln_007/nn_266918/GRDC), using gauging stations with daily data availability for more than 25 y (during 1958–2000) and an upstream area above 125,000 km^2. For these stations, the relative error was calculated for discharge with a return period of 10 y, based on extrapolations using the Gumbel distribution fit to the observed and modeled time series. Generally, the relative error was found to be reasonable (between −25 and +25% for 37 of the 53 stations).

However, these validations did not assess the performance of the model in terms of its ability to simulate relative differences in peak discharges between the different phases of ENSO, which is the aspect of most crucial importance for the current paper. Therefore, in this study we carried out extra validation to specifically examine this aspect. We used daily discharge data from the GRDC database, using only those stations with upstream areas greater than 10,000 km^2 for which daily data are available for every day of the hydrological year in at least 15 hydrological years between 1959 and 2000 (i.e., 722 stations).

In Fig. S7 we show the correlation (Spearman's rank) between the modeled and observed values of the maximum daily discharge (Q_{max}) per hydrological year. Generally, the agreement is good. For half of the stations, the correlation coefficient is greater than 0.6, and it is greater than 0.4 for 80% of stations. The correlation is relatively low in northern high-latitude regions. This may be due to earlier noticed biases in the precipitation data in northern latitudes owing to snow undercatch of the used rain gauges (1). Low correlation is also found in several of the gauging stations assessed in Central America.

We also examined the agreement between the modeled and observed data in terms of the relative change in Q_{max} between El Niño and non-El Niño years (Fig. S8*A*) and between La Niña and non-La Niña years (Fig. S8*B*). The figure shows that for the vast majority of stations modeled and observed median Q_{max} show either no significant difference between El Niño (La Niña) and non-El Niño (non-La Niña) years or significant differences of the same sign. For the other stations there is a statistically significant difference in modeled median Q_{max} between El Niño (La Niña) and non-El Niño (non-La Niña) years but none for observed data (or vice versa). There are no stations at which modeled and observed median Q_{max} show significant changes between El Niño/non-El Niño years or La Niña/non-La Niña years with different signs. Hence, the model seems to simulate well the relative differences in peak annual discharge between the different phases.

Validation of Impact Assessment Results. The globally aggregated impact results of the modeling cascade have been validated in ref. 1. For the current paper, we carried out further validation at regionally disaggregated levels, first by the large regions defined in ref. 12 and shown in Table S3 and, second, by country (Table S4). We assessed the correlation (Spearman's rank) between annual reported losses and annual modeled urban damages and between annual reported fatalities and annual modeled exposed population over the period 1990–2000. The reported fatalities and losses were taken from Munich Re's NatCatSERVICE database (13). For this analysis, we only included events for which the main cause of the event is river flooding (i.e., "floods"), and not other forms of floods, for example flash floods, dam breaks, and tsunamis. Both reported and modeled annual impacts were extracted for hydrological years, rather than calendar years. The validation comparisons are intended to give a broad picture of whether the modeled impacts pick up interannual differences in impacts similar to those recorded in the loss database. The modeled and reported impacts are not the same variables (e.g., affected population is compared with reported fatalities). Moreover, loss databases themselves are inherently limited and suffer from many problems; for a detailed discussion see ref. 14. We used the period 1990–2000 only because this is the latest 10-y period for which we have simulated impacts, and the loss database is generally considered to be more reliable for the most recent period compared with early time periods (14).

In Table S3 we show the correlation between annual reported losses and annual modeled urban damages and between annual reported fatalities and annual modeled exposed population for

the world regions. Generally, we see reasonable to good agreement. A major exception is for North Africa, where both pairs of variables show negative correlation. This is not a surprising result for two reasons. First, simulating flood events in this arid region is particularly difficult (1), and second, there are very few reported events in this region ($n = 32$). Hence, we cannot make any statistically robust statements on whether the discrepancy results from the quality of the modeled data or from the reported data, or both. In Eastern Asia, the correlation is almost zero between reported losses and urban damage (rho = 0.52 for reported fatalities and modeled exposed population), and for Eastern Europe and Central Asia the correlation is slightly negative between reported fatalities and modeled exposed population (and positive, although weak, between reported losses and urban damage). For Eastern Europe and Asia, this may be related to the hydrological modeling results: The validation results for modeled vs. observed peak annual discharge (Fig. S7) show low to negative correlation for many gauging stations. The same cannot be said for Eastern Asia, where the discharge validation results are good. The fact that the correlation between reported fatalities and modeled exposed population is reasonably strong in this region, yet the correlation is negative for reported losses vs. urban damage, may be indicative of lower quality of reported loss information. Until recently, many states in the region felt they should not share information of this kind with the rest of the world (14). However, the poor performance for exposed population vs. reported fatalities may simply be due to there being no clear (linear) relationship between these two variables in this region. In our calculations of urban damage we do include a vulnerability component, by means of using a depth-damage function. However, we did not use such a component for exposed population.

The same analysis was carried out using reported and modeled impacts for the 10 countries for which the largest number of reported events are available (Table S4). For most countries, the correlation is again reasonable to strong. An exception is the Russian Federation, and to a lesser extent India. As above, the poor correlation for the Russian Federation may be a result of the fact that our hydrological model does not seem to simulate interannual variations in peak annual discharge very well in this region (compared with observed data), especially in the eastern part. In addition, the Russian Federation and India (as well as many other countries) may have suffered from relatively large numbers of events that remain undocumented, owing to low media penetration and insufficient data collection and sharing. Finally, the reported loss data include both damages in urban and rural areas, whereas the damage model only reflects urban damages. The latter fact may lead to discrepancies in developing countries, such as India, where large parts of the country consist of rural areas and small settlements that are not represented in the global urban density databases.

Comparison with Past Results. ENSO's influence on observed floods has only been examined in a few regions, mainly in the western United States and Australia. An analysis of observed daily discharge data for the Santa Cruz River at Tucson, Arizona (15) showed the magnitude of a 100-y flood to be significantly larger during El Niño years. For the cell in which this river is found, we simulated a positive anomaly in 100-y flood volume (+25%), and for the FPU in which this river is found (Colorado) we simulated a positive anomaly in annual expected damages (+18%). A study of ENSO relationships with peak observed discharges at 303 locations in the western United States (16) found that in El Niño years days with high daily discharge occur more frequently than average over the southwestern United States and less frequently than average over the northwestern United States, with an almost opposite pattern for La Niña years.

We also see this general pattern reflected in our flood-risk results. We found higher exposed population and GDP during El Niño years in the southwestern United States and lower values in the northwest. For urban damage, however, the El Niño signal is weak. During La Niña years, we found lower urban damages, exposed GDP, and exposed population in the southwest and higher urban damages in the northwest. However, the pattern is not clear in basins located in the central part of the western United States. For example, for the Sacramento–San Joaquin basin we simulated negative damage anomalies during La Niña conditions and weak (not reliable) anomalies during El Niño conditions. This is related to hydroclimatological differences between the northern (Sacramento) and southern (San Joaquin) parts of this basin. During La Niña conditions, the simulated anomalies in flood volumes are lower across almost the entire Sacramento–San Joaquin basin, whereas during El Niño conditions negative anomalies are more dominant in most parts of the Sacramento basin, with positive anomalies in most parts of the San Joaquin basin.

For Australia, studies have been carried out to condition flood return periods on indices of ENSO and the Interdecadal Pacific Oscillation (IPO) (17, 18). These studies show graphs of log-normal values of a regional flood index value for New South Wales for return periods between 1 and 100 y, conditioned on El Niño and La Niña years only. The results clearly show higher values of the flood index during La Niña years compared with El Niño years (the climatological mean is not shown). This is in agreement with our maps of flood volume anomalies during La Niña years (positive anomaly) and El Niño years (negative anomalies) in New South Wales. Ishak et al. (19) analyzed a database of annual maximum stream flows in Australia and found that indices of ENSO, IPO, and the Southern Annular Mode can account for most of the observed trend in annual maximum stream flows.

In the northern coastal region of Peru, analyses of observed time series of annual floods for 13 rivers show strong positive peak flow anomalies during El Niño years (20). For this region, we also simulated high positive anomalies in flood volume and damages (+43% for the latter).

Main Differences in Results Between Indicators. We simulated anomalies in risk between the different phases of ENSO for the different risk indicators, namely, annual expected urban damage, exposed GDP, and exposed population. We found that the sign of anomalies is very similar across the different indicators. However, there are clear differences in the strength of the anomalies, and hence also in the number and distribution of reliable anomalies.

In Table S1 we show the percentage of the Earth's land area for which basins showed a reliable positive or negative anomaly during El Niño and/or La Niña years. Note that Antarctica and Greenland are excluded from the analyses. The anomalies per basin are shown in Fig. 2 (annual expected urban damage), Fig. S4 (annual exposed GDP), and Fig. S5 (annual exposed population).

From Table S1 we see large differences in the percentage of land area affected by the different anomalies. For example, during La Niña years similar percentages of land surface experienced positive and negative anomalies (10% and 13% of total land area, respectively). However, for exposed GDP and exposed population, negative anomalies were simulated across much larger areas than is the case for positive anomalies. Figs. S4 and S5 reveal a simple explanation for these differences. For example, during La Niña years, negative anomalies in exposed GDP and population are simulated in the large FPU representing northeastern Russia. However, this anomaly is not simulated for urban damage. For a large part, this explains the difference noted earlier in this paragraph.

A comparison of the maps of anomalies in annual expected urban damage (Fig. 2) with those for annual affected GDP shows, in general, a similar pattern in terms of the signal of the

anomalies. However, there are also some clear differences in terms of the differences in magnitude of the anomalies. Notable differences exist in South America, particularly around the region of the Paraná and San Francisco basins, and in the western basins. In North America, there are several basins in the western and south-central parts of the United States showing reliable anomalies in annual expected urban damage during different ENSO phases, but no anomalies in annual exposed GDP. The strength of the anomalies in southern Africa also tends to be stronger for urban damage than for affected GDP. However, there are regions that show reliable anomalies in affected GDP but no anomalies in urban damage. Further research is required to understand the reasons for these differences, which could be related to several factors. One factor is the location of the exposed elements at risk. So-called urban damages can only occur in regions where urban area is located in flood-prone areas (i.e., where it is potentially exposed to a flood hazard). The same is the case for affected GDP and population. Because the geographical locations of these elements differ, so do the absolute impacts of flooding, and by extension the relative differences between ENSO phases. Another factor of large importance is that the urban damage indicator is dependent on inundation depth, whereas the other indicators are only dependent on the binary presence or absence of inundation. In our approach, the spatial distribution of GDP is directly related to the population density per cell. Hence, the results in ENSO anomalies for these two indicators are very similar.

1. Ward PJ, et al. (2013) Assessing flood risk at the global scale: Model setup, results, and sensitivity. *Environ Res Lett* 8:044019.
2. Van Vuuren DP, Lucas PL, Hilderink H (2007) Downscaling drivers of global environmental change: Enabling use of global SRES scenarios at the national and grid levels. *Glob Environ Change* 17(1):114130.
3. Bouwman AF, Kram T, Klein Goldewijk K (2006) *Integrated Modelling of Global Environmental Change: An Overview of IMAGE 2.4* (Netherlands Environmental Assessment Agency, Bilthoven, The Netherlands).
4. Bright EA, Coleman PR, Rose AN, Urban ML (2010) *LandScan 2010 High Resolution Global Population Data Set* (Oak Ridge National Laboratory, Oak Ridge, TN).
5. Klein Goldewijk K, Beusen A, Van Drecht G, De Vos M (2011) The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Glob Ecol Biogeogr* 20(1):73–86.
6. Jongman B, Ward PJ, Aerts JCJH (2012) Global exposure to river and coastal flooding: Long term trends and changes. *Glob Environ Change* 22(4):823–835.
7. Klijn F, Baan PJ, De Bruijn KM, Kwadijk JC, Van Buren R (2007) Overstromingsrisico's in Nederland in een veranderend klimaat: Verwachtingen, schattingen en berekeningen voor het project Nederland Later (WL | Delft Hydraulics, Delft, The Netherlands).
8. Ward PJ, De Moel H, Aerts JCJH (2011) How are flood risk estimates affected by the choice of return-periods? *Nat Hazard Earth Sys* 11:3181–3195.
9. Jongman B, et al. (2012) Comparative flood damage model assessment: Towards a European approach. *Nat Hazard Earth Sys* 12:3733–3752.
10. Van Beek LPH, Wada Y, Bierkens MFP (2011) Global monthly water stress: I. Water balance and water availability. *Water Resources Res* 47(7):W07517.
11. Wada Y, Van Beek LPH, Bierkens MFP (2012) Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resources Res* 48(6):W00L06.
12. Kummu M, Ward PJ, De Moel H, Varis O (2010) Is physical water scarcity a new phenomenon? Global assessment of water shortage over the last two millennia. *Environ Res Lett* 5(3):034006.
13. Munich Re (2013) NatCatSERVICE Database (Münchener Rückversicherungs-Gesellschaft, Munich).
14. Kron W, Steuer M, Löw P, Wirtz A (2012) How to deal properly with a natural catastrophe database – analysis of flood losses. *Nat Hazard Earth Sys* 12:535–550.
15. Cayan DR, Webb RH (1992) El Niño / Southern Oscillation and streamflow in the western United States. *El Niño: Historical and Paleoclimatic Aspects of the Southern Oscillation*, eds Diaz HF, Markgraf V (Cambridge Univ Press, Cambridge, UK), pp 29–68.
16. Cayan DR, Redmond KT, Riddle LG (1999) ENSO and hydrologic extremes in the western United States. *J Clim* 12:2881–2893.
17. Kiem AS, Franks SW, Kuczera G (2003) Multi-decadal variability of flood risk. *Geophys Res Lett* 30(2):1035.
18. Kiem AS, Verdon-Kidd DC (2013) The importance of understanding drivers of hydroclimatic variability for robust flood risk planning in the coastal zone. *Australian J Water Resources* 17(2):126–134.
19. Ishak EH, Rahman A, Westra S, Sharma A, Kuczera G (2013) Evaluating the non-stationarity of Australian annual maximum floods. *J Hydrol (Amst)* 494:134–145.
20. Waylen PR, Caviedes CN (1986) El Niño and annual floods on the north Peruvian littoral. *J Hydrol (Amst)* 89:141–156.
21. Livezey RE, Chen WY (1983) Statistical field significance and its determination by Monte Carlo techniques. *Mon Weather Rev* 111(1):46–59.

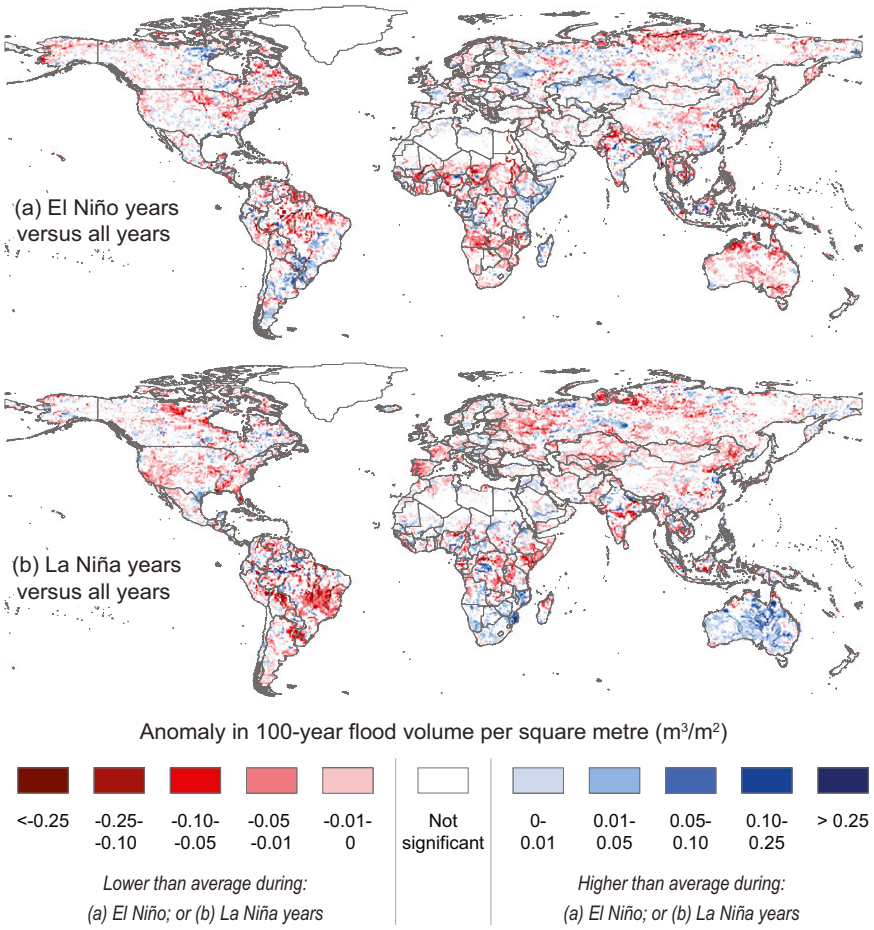

Fig. S1. Absolute anomaly in flood volumes normalized to surface area per cell (cubic meter per square meter) with return periods of 100 y during (*A*) El Niño years and (*B*) La Niña years (compared with all years). Statistical significance was assessed by bootstrapping ($\alpha = 0.05$), using 1,000 repetitions. Field significance of the gridded results was assessed using the binomial distribution (21) and found to be highly significant ($P < 0.001$).

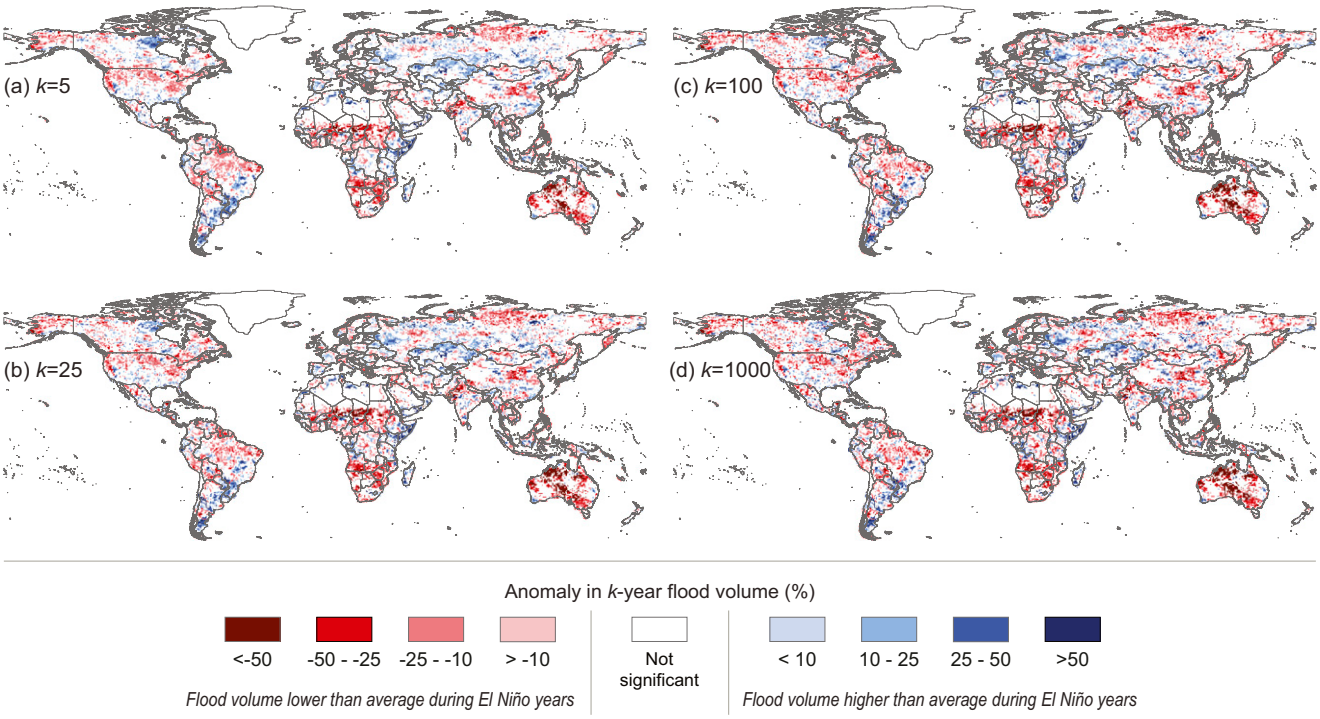

Fig. S2. Percentage anomaly in flood volumes during El Niño years (compared with all years) with a return period of (*A*) 5 y, (*B*) 25 y, (*C*) 100 y, and (*D*) 1,000 y. Statistical significance per cell was assessed by bootstrapping ($\alpha = 0.05$), using 1,000 repetitions. Field significance of the gridded results was assessed using the binomial distribution (21) and found to be highly significant ($P < 0.001$). Compared with the percentage anomalies in flood volumes during El Niño years with a return period of 100 y (Fig. 1*A*), the strengths of the anomalies generally show a small increase for floods with shorter return periods. In most regions this change is very small, and hence there is little change between the categories shown in the figure. The patterns of the anomalies therefore remain similar across different return periods.

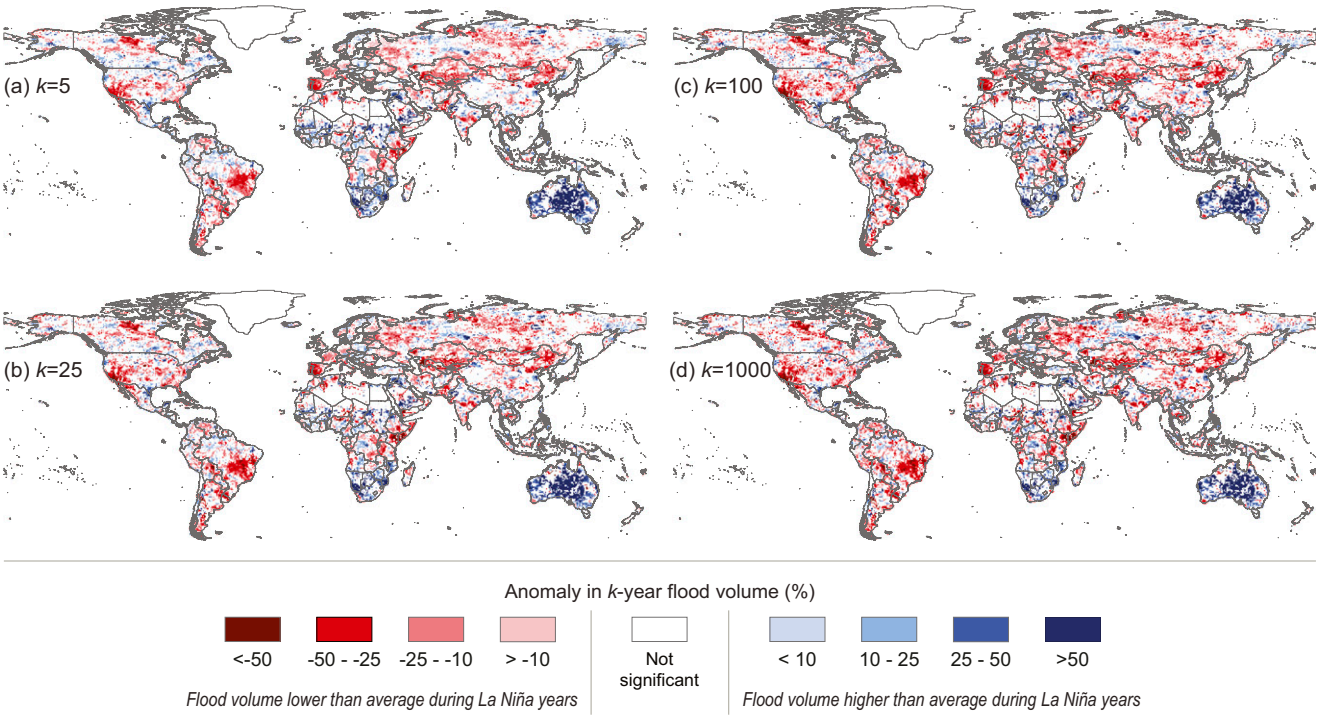

Fig. S3. Percentage anomaly in flood volumes during La Niña years (compared with all years) with a return period of (*A*) 5 y, (*B*) 25 y, (*C*) 100 y, and (*D*) 1,000 y. Statistical significance per cell was assessed by bootstrapping ($\alpha = 0.05$), using 1,000 repetitions. Field significance of the gridded results was assessed using the binomial distribution (21) and found to be highly significant ($P < 0.001$). Compared with the percentage anomalies in flood volumes during La Niña years with a return period of 100 y (Fig. 1*B*), the strengths of the anomalies generally show a small increase for floods with shorter return periods. In most regions this change is very small, and hence there is little change between the categories shown in the figure. The patterns of the anomalies therefore remain similar across different return periods.

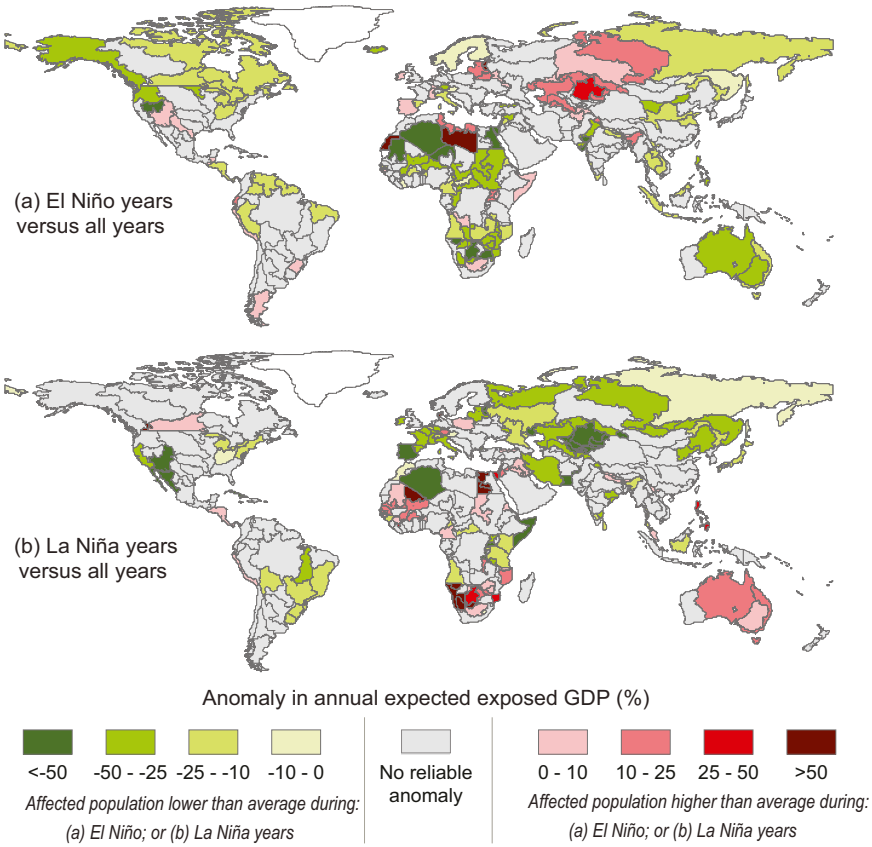

Fig. S4. Percentage anomaly per FPU in annual expected exposed GDP for (*A*) El Niño years and (*B*) La Niña years (compared with all years).

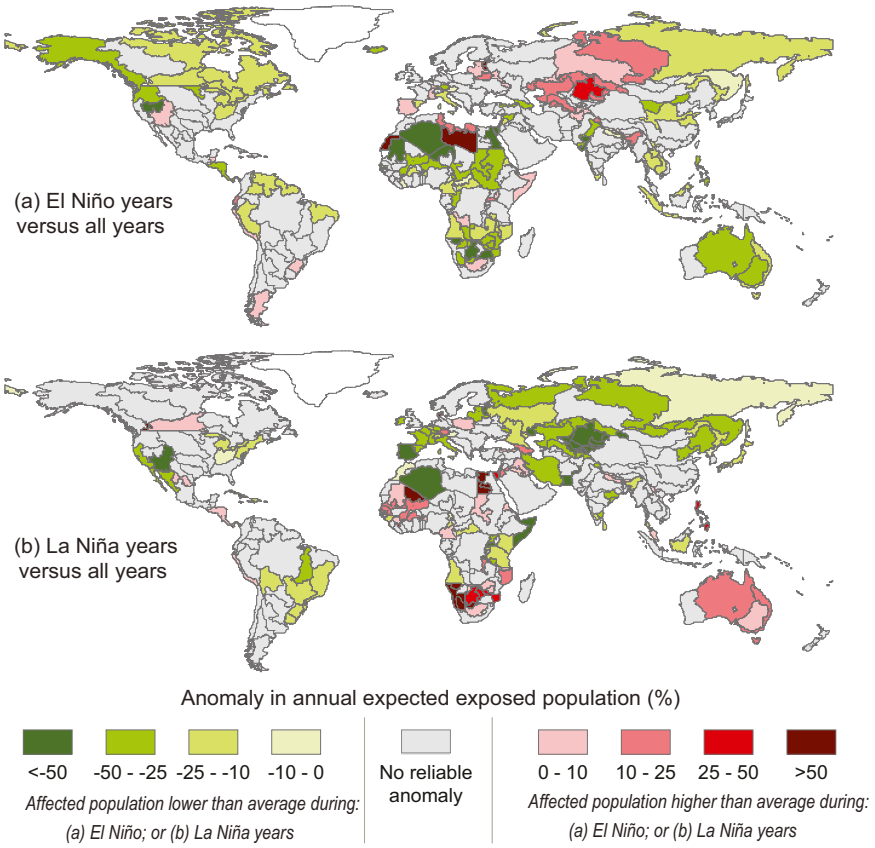

Fig. S5. Percentage anomaly per FPU in annual expected exposed population for (*A*) El Niño years and (*B*) La Niña years (compared with all years).

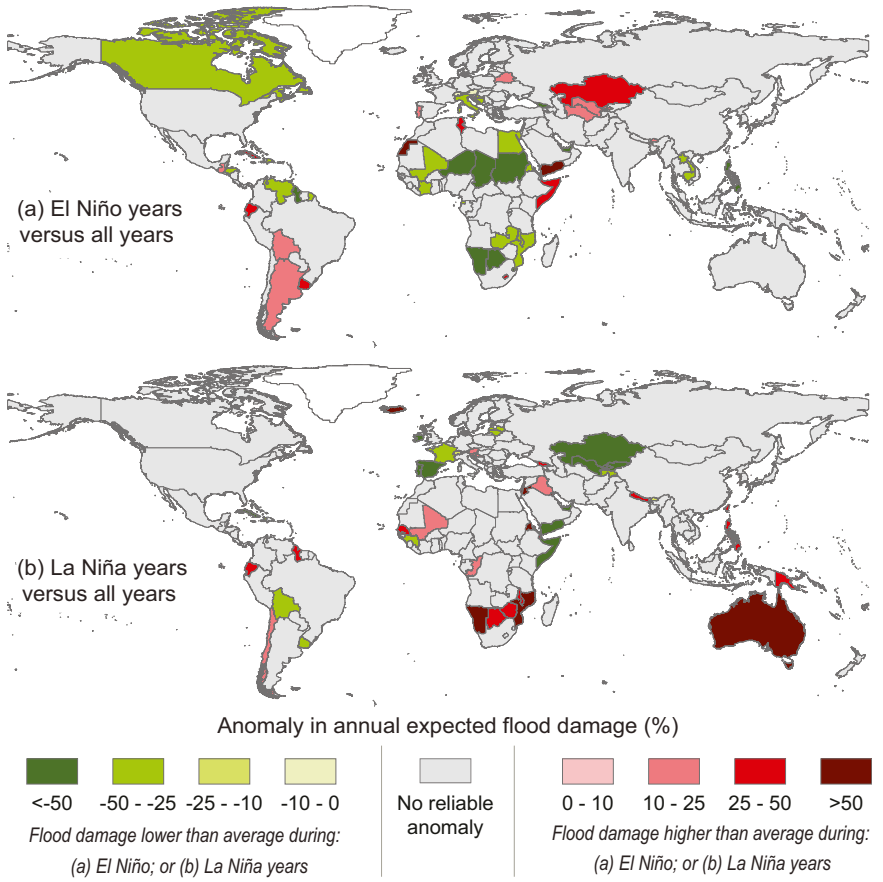

Fig. S6. Percentage anomaly per country in annual expected damage in urban areas for (*A*) El Niño years and (*B*) La Niña years (compared with all years). The results show the importance of the aggregation scale on the overall results in terms of anomalies in risk. Comparing the results at the country scale with those at the FPU scale shown in Fig. 2 shows that country-scale assessments mask important regional ENSO influences on risk. This is especially the case in larger countries where there are regionally opposite influences of ENSO on hydroclimate. A clear example is the United States: Based on the country-aggregated results, there is no reliable anomaly in this country during El Niño or La Niña years, yet the regionally disaggregated influence is high.

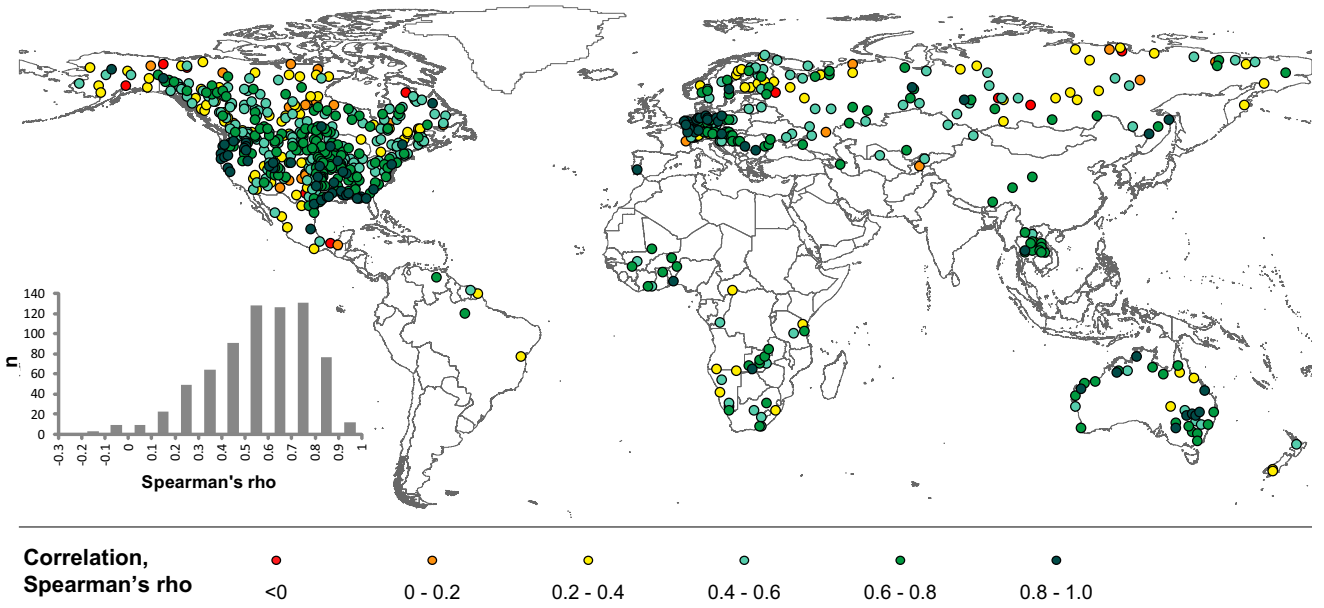

Fig. S7. Correlation (Spearman's rank, rho) between modeled and observed Q_{max} time series over the period 1958–2000. (*Inset*) Histogram of values for individual gauging stations, $n = 722$. For a discussion of these validation results, see *SI Discussion*, *Validation of Hydrological and Hydraulic Models*).

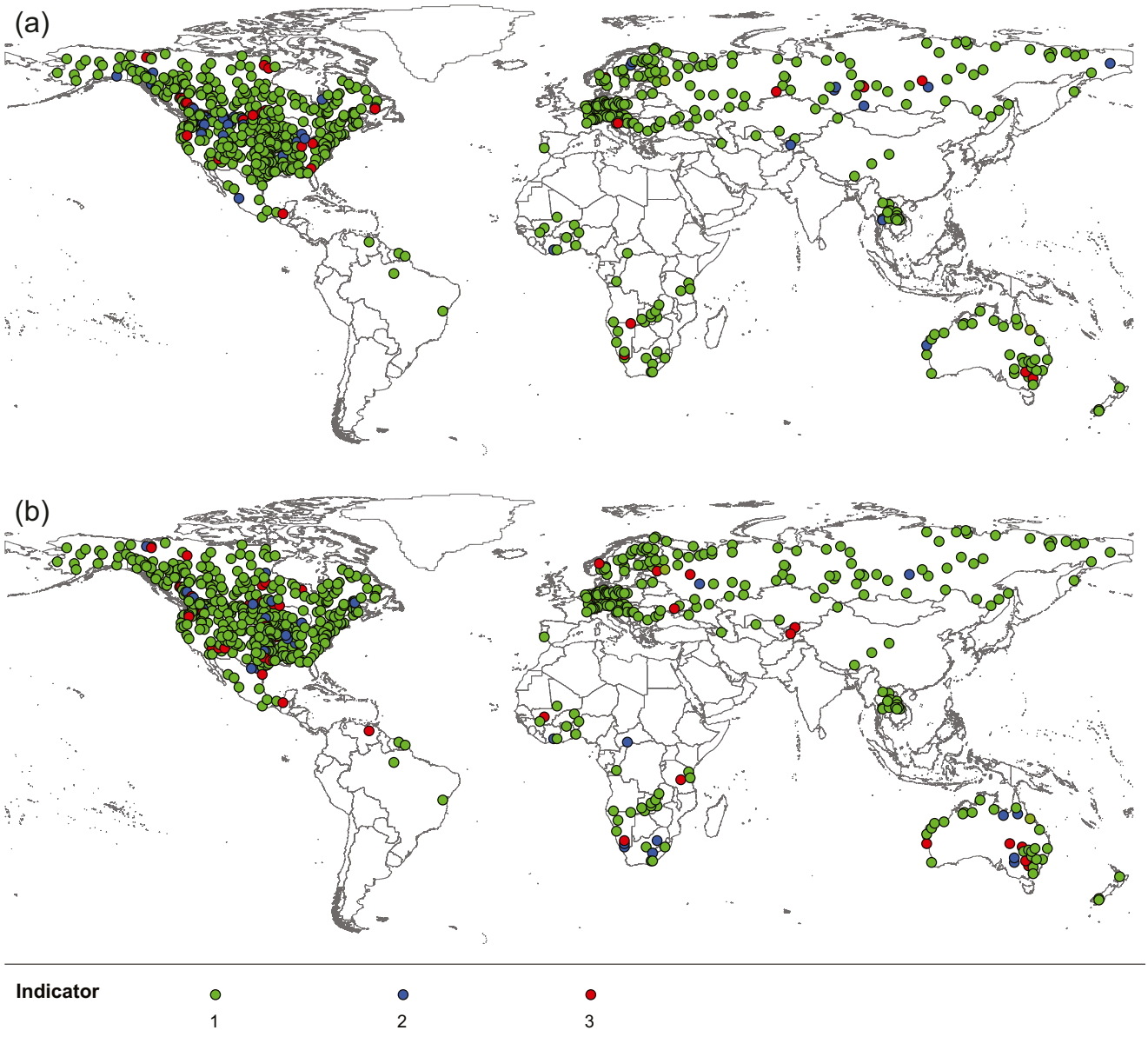

Fig. S8. Indicator of the agreement between modeled and observed median Q_{max} for (*A*) El Niño years compared with non-El Niño years and (*B*) La Niña years compared with non-La Niña years. Indicator 1 means that both modeled and observed median Q_{max} show either no significant difference or significant differences with the same sign. Indicator 2 means that modeled median Q_{max} shows no significant difference, whereas the observed Q_{max} does show a significant difference. Indicator 3 means that observed median Q_{max} shows no significant difference, whereas modeled Q_{max} does show a significant difference. There are no stations at which modeled and observed median Q_{max} show significant differences with the opposite sign. Statistical significance was assessed using a two-tailed Mann–Whitney *U* test, $\alpha = 0.05$. For a discussion of these validation results, see *SI Discussion*, *Validation of Hydrological and Hydraulic Models*.

Table S1. Percentage of land area (excluding Antarctica and Greenland) with reliable positive and negative anomalies in flood risk during El Niño and La Niña years

	El Niño		La Niña	
Flood risk indicator	Positive anomaly	Negative anomaly	Positive anomaly	Negative anomaly
Urban damage	10	19	10	13
Exposed GDP	12	39	13	30
Exposed population	12	38	14	30

Table S2. Percentage anomalies in globally aggregated values of modeled annual expected urban damage based on El Niño and La Niña years only (compared with results based on all years)

	Anomaly per ENSO phase, %	
Protection standard, y	El Niño	La Niña
2	−8.9	−14.2
5	−6.8	−10.9
10	−6.2	−9.8
25	−6.0	−8.9
50	−5.8	−8.7
100	−5.9	−8.8
250	−5.8	−8.9
500	−5.9	−9.0

Results are shown for different assumptions of a nominal protection standard (expressed as a return period) against flooding. None of the anomalies is reliable. The simulations carried out for this paper assume that no infrastructural flood protection measures are in place, such as dikes and retention areas. A past study has shown that the absolute risk estimates are strongly influenced by the assumed flood protection standard (1). Here, we assessed how anomalies in risk between ENSO phases are affected by assuming different nominal flood protection standards. At the global scale, we found that the flood protection standard has a fairly small influence on the simulated risk anomalies.

Table S3. Correlation (Spearman's rank, rho) between reported impacts (13) and modeled impacts over the period 1990–2000

Country	*n*	Spearman's rank, rho	
		Reported losses vs. urban damage	Reported fatalities vs. exposed population
Australia and Oceania	69	0.71	0.48
Central America	69	0.42	0.21
Eastern Asia	102	−0.03	0.52
Eastern Europe and Central Asia	123	0.24	−0.08
Indian subcontinent	131	0.15	0.32
Latin America	108	0.60	0.61
Middle East	71	0.50	0.81
Middle and South Africa	123	0.65	0.49
North Africa	32	−0.10	−0.18
North America	126	0.77	0.59
Southeastern Asia	127	0.62	0.77
Western Europe	140	0.70	0.68

Values are shown for the geographical regions shown in ref. 12. *n* is the number of reported events in the period used in the analysis. For a discussion of these results, see *SI Discussion*, *Validation of Hydrological and Hydraulic Models*.

Table S4. Correlation (Spearman's rank, rho) between reported impacts (13) and modeled impacts over the period 1990–2000 per country

Country	*n*	Spearman's rank, rho	
		Reported losses vs. urban damage	Reported fatalities vs. exposed population
United States	87	0.71	0.53
China	54	0.70	0.58
India	47	0.19	0.22
Russian Federation	45	0.15	−0.07
Canada	39	0.72	0.37
Indonesia	38	0.27	0.78
Australia	34	0.68	0.37
Brazil	32	0.42	0.50
Bangladesh	28	0.54	0.52
South Africa	27	0.63	0.51

Values are shown for the 10 countries with the largest number of observed flood events (*n*). For a discussion of these results, see *SI Discussion*, *Validation of Hydrological and Hydraulic Models*.