Supplementary appendix for

Rainfall events and daily mortality: a multi-country study across 645 locations

1. Methods

1.1 Mortality data

The MCC Network has gathered and regularly updated mortality data from different health authorities across many countries and territories. This data is detailed in our earlier study [1]. In its latest version, the MCC database includes all causes (substituted by non-external cause mortality [classified under ICD-9 codes 0–799 or ICD-10 codes A00-R99] without complete data) from 743 locations spanning 44 countries/regions. cardiovascular mortality encompassing 631 cities across 29 countries/regions, and respiratory mortality data from 639 cities in 28 countries/regions. This dataset provides a crucial foundation for understanding the impact of global environmental health issues, like extreme climate events.

1.3 Definition of rainfall events

We employ the Regional Frequency Analysis (RFA) framework to analyze extreme rainfall events across a variety of climatic and geographical locations. The initial step involved grouping locations based on the second level Köppen climate classification [2], followed by the utilization of latitude, longitude, and annual precipitation sums, as well as the standard deviation of daily precipitation amounts, through K-means clustering [3]. This ensures that location within the same climatic zone and exhibit no significant differences ($p < 0.05$) in precipitation proximities. These precipitation proximities include two measures, both based on 40 years of rainfall data: 1.the annual rainfall amount for each year; 2.the standard deviation of daily precipitation calculated annually for each year. This step is crucial for stabilizing subsequent analyses and accurately reflecting the extreme precipitation conditions of different regions, aligning with established methodologies in climatological studies [4 5].

Following this, the Generalized Extreme Value (GEV) distribution function, a widely used model for extreme events analysis [6 7], is fitted to the precipitation data of each cluster. This

fitting is based on the partial duration series (PDS) of the 40 highest daily precipitation values for each year from 1980 to 2020 for every location within a cluster. The 40-year period was selected in alignment with the availability of mortality data, ensuring the analysis's robustness through a comprehensive temporal scope. This process facilitates the estimation of extreme rainfall thresholds for specified return periods of 1, 2, and 5 years.

Referring to related studies [8 9], the fit of the model for each cluster is rigorously evaluated using the Kolmogorov-Smirnov test [10], with clusters not meeting the significance threshold ($p > 0.05$) subject to refinement. This ensures the reliability of extreme value predictions across different regions and allows for a nuanced understanding of extreme precipitation patterns in varied spatial contexts.

1.2 Climatic, population density, income class, and vegetation coverage

Between 1980 and 2020, we utilized the ERA5-Land dataset[11] to calculate the annual cumulative precipitation across various cities. This dataset provides an extensive analysis of long-term precipitation trends and patterns. Furthermore, we calculated the average annual standard deviation of daily cumulative precipitation, derived from the same dataset, which provides valuable insights into the fluctuation features of precipitation in different urban regions.

For demographic insights, the 2010 population density data were sourced from the Gridded Population of the World (GPW, version 4) dataset for each city, managed by the Socioeconomic Data and Applications Center (SEDAC) [12]. This dataset offers population density estimates with a spatial resolution of 1 km, aligning closely with national censuses and population registers [12]. Using the spatial data on city boundaries, we calculated the average population density for each city, incorporating all grid points within these boundaries. This

dataset effectively captures the diverse population distribution patterns within different urban areas.

The income class information for various locations, as pertaining to their respective countries, was sourced from the World Bank (https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html). Considering the primary coverage period of the mortality data, we opted for the data version from the year 2010 to represent the income class information.

Furthermore, the average percent of vegetation coverage was calculated using the MOD44B Version 6.1 Vegetation Continuous Fields (VCF) yearly product [13], derived from MODIS satellite data. This product presents a continuous, quantitative depiction of land vegetation coverage with a resolution of 250 meters [14]. Based on the 2010 annual data and city boundaries, we computed the average percent of vegetation coverage for each city. This metric establishes the baseline for green cover conditions in each city, which is a critical factor in urban adaptation to many extreme climate events [15 16].

1.3 Sensitivity analyses

We conducted a series of sensitivity analyses to examine the robustness of our analysis results, including:

(1) Different lag settings: we incorporated three additional lag settings of 7, 21, and 28 days to assess the impact of extreme rainfall events.

(2) Df for lag-response: we examined variations in the setting of degrees of freedom of the natural cubic B spline for the extreme rainfall events to evaluate their potential influence.

(3) Df for year for seasonal control: we also examined variations in the setting of degrees of freedom of the natural cubic B spline for the time of the year to evaluate the potential influence of long-term and season trends of death.

(4) Alternative precipitation datasets: as outlined in the main text, we also included different precipitation datasets, applying a consistent method to evaluate the relative risk.

(5) Air pollution: we adjusted for the non-linear and lagged effects of daily 24-h average particulate particles with an aerodynamic diameter of 2.5 μg ($PM_{2.5}$) and maximum 8-h average ozone (O_3) concentrations by adding a distributed lag non-linear model with a duration of lag 0, 1, and 2 d for $PM_{2.5}$ and O_3 into the main model. This sensitivity analysis was conducted in cities where complete meteorological and air pollution data were available during the research period.

(6) Flood events: We discarded the extreme rainfall events that were simultaneously identified as flood exposure days on the same day and within the subsequent 14 days to specifically test the impact of extreme rainfall independent of flood exposure. This approach follows the same identification methods previously done based on the Dartmouth Flood Observatory database in our recent studies [17] [18].

2 Results

Table S1 Statistical overview of 645 locations categorized by country/region

	Income class	No. of city	Study period	No. of 1-in-1	No. of 1-in-2	No. of 1-in-5	All-cause deaths	Respiratory deaths	Cardiovascular deaths
South-East Asia									
Philippines	lower-middle	10	2006-2019	599	79	24	671721	96810	241965
Thailand	upper-middle	50	1999-2008	4201	402	182	1277756	152182	228008
Vietnam	lower-middle	2	2009-2013	74	13	10	108173	8970	24433
South Europe									
Portugal	high	6	1980-2018	623	150	55	1925255	181391	718560
Italy	high	8	2001-2010	163	43	21	378552	37123	100396
Spain	high	49	1990-2014	2015	517	191	2985947	336589	1031607
South America									
Brazil	upper-middle	17	1997-2018	2533	244	86	3836604	458501	1205195
Colombia	upper-middle	5	1998-2013	427	48	20	956539	99816	267900
Costa Rica	upper-middle	1	2000-2017	143	14	3	31117	2651	9320
France Guiana	high	1	2000-2015	649	21	4	7110	NA*	NA
Paraguay	upper-middle	1	2004-2019	72	12	2	48037	4445	15365
Peru	upper-middle	4	2008-2014	72	19	8	111553	NA	NA
South Africa									
South Africa	upper-middle	50	1997-2013	3108	588	224	8434669	1039663	1282363
North Europe									
Estonia	high	2	1997-2018	61	21	5	31381	1394	15634
Finland	high	1	1994-2014	11	3	3	153319	9742	5740
Ireland	high	6	1984-2007	105	30	16	1058215	164139	340321
Netherland	high	5	1995-2016	139	35	18	453395	NA	NA
Norway	high	1	1980-2018	35	14	4	206949	20450	76978

Sweden	high	3	1990-2016	82	26	14	717294	54930	310806
UK**	high	53	1990-2016	1138	370	193	5249116	776863	192123
North America									
Canada	high	22	1986-2015	1918	225	64	2709269	232811	929463
USA***	high	207	1980-2006	17760	3229	1337	30767094	2833989	10913073
East Asia									
Mainland China	upper-middle	4	2003-2015	93	21	11	233476	42538	103738
Japan	high	47	1980-2015	7875	1164	381	34348889	4628076	11246546
South Korea	high	36	1997-2018	4104	590	209	3070414	222314	701638
Taiwan	upper-middle	3	1994-2014	519	58	30	1209868	116518	269444
Central Europe									
Czech Republic	high	4	1994-2015	151	30	16	711910	39853	360040
France	high	13	2000-2015	393	99	36	866419	55951	NA
Germany	high	12	1993-2015	349	85	42	3105865	NA	NA
Moldova	lower-middle	3	2001-2010	65	14	7	57094	NA	NA
Romania	upper-middle	8	1994-2016	358	92	45	951146	NA	NA
Switzerland	high	5	1995-2013	198	50	18	156654	9663	59785
Central America									
Mexico	upper-middle	3	1998-2014	653	20	9	1945994	189906	513720
Australia									
Australia	high	3	1988-2009	227	36	13	1177950	NA	NA

*NA: no mortality data available

** The United Kingdom (UK)

*** The United States of America (USA)

Table S2 I^2 Statistics and Cochran's Q tests results of meta-analyses

	All causes		Cardiovascular		Respiratory	
	I^2	*p*-value*	I^2	*p*-value	I^2	*p*-value
Global	42%	<0.001	34%	<0.001	34%	<0.001
North Europe	21%	0.006	21%	0.012	47%	<0.001
Central Europe	31%	<0.001	24%	0.019	51%	<0.001
South Europe	42%	0.005	32%	<0.001	43%	0.006
East Asia	57%	<0.001	44%	0.004	53%	<0.001
South-East Asia	36%	0.003	41%	<0.001	31%	0.011
North America	61%	<0.001	42%	<0.001	48%	<0.001
Central America	43%	0.106	52%	0.109	51%	<0.001
South America	60%	<0.001	40%	<0.001	37%	<0.001
South Africa	40%	<0.001	25%	0.008	43%	<0.001
Australia	56%	0.453	NA	NA	43%	<0.001
Climate zones						
Tropical	54%	0.005	44%	0.004	42%	0.027
Aric	30%	<0.001	13%	<0.001	51%	<0.001
Temperate	44%	<0.001	25%	<0.001	45%	<0.001
Continental	34%	0.005	17%	0.004	35%	<0.001
Average annual precipitation (mm)						
Q1: < 834.5	49%	<0.001	34%	0.002	44%	0.002
Q2: 834.5 - 1231.5	31%	<0.001	12%	<0.001	54%	<0.001
Q3: 1231.5 - 1487.8	45%	<0.001	22%	0.007	52%	<0.001
Q4: > 1487.8	59%	<0.001	21%	0.019	27%	0.013
Daily precipitation standard deviation						

Q1: < 4.6	61%	<0.001	41%	<0.001	57%	<0.001
Q2: 4.6 - 6.9	35%	<0.001	52%	<0.001	29%	0.002
Q3: 6.9 - 8.3	21%	<0.001	31%	0.003	49%	<0.001
Q4: > 8.3	44%	<0.001	21%	0.021	59%	0.004
Income class						
low	61%	<0.001	47%	<0.001	51%	<0.001
lower-middle	34%	<0.001	31%	0.005	44%	<0.001
upper-middle	45%	<0.001	21%	0.021	38%	0.003
high	51%	<0.001	14%	0.017	34%	<0.001
Average vegetation coverage (%)						
Q1: < 72.8	43%	<0.001	39%	<0.001	45%	<0.001
Q2: 72.8 - 79.0	41%	<0.001	31%	0.003	43%	<0.001
Q3: 79.0 - 83.3	39%	<0.001	11%	0.009	38%	0.002
Q4: > 83.8	43%	<0.001	10%	0.014	24%	0.019
Population density (per km^2)						
Q1: < 116	45%	<0.001	36%	<0.001	54%	<0.001
Q2: 116 - 301	51%	<0.001	27%	<0.001	33%	<0.001
Q3: 301 - 998	49%	<0.001	19%	0.012	22%	<0.001
Q4: > 998	53%	<0.001	21%	0.003	31%	<0.001

[*] *P* values were estimated through Cochran's Q tests

[**] Each city's average annual precipitation, daily precipitation variability, vegetation coverage, and population density were divided into quartiles over values from 738 cities.

Table S3 Pooled cumulative relative risks (95% confidence intervals) of mortality associated with extreme rainfall events with a 5-year return period after changing lag settings, degrees of freedom of the lag-response, and only included non-flood exposure events.

	All causes		Cardiovascular		Respiratory	
	Cumulative RR	P value *	Cumulative RR	P value	Cumulative RR	P value
Primary results	1.08 (1.05, 1.11)	Ref	1.05 (1.02, 1.08)	Ref	1.29 (1.19, 1.39)	Ref
Lag setting						
7	1.06 (1.03, 1.09)	0.88	1.04 (1.00, 1.08)	0.91	1.30 (1.20, 1.40)	0.94
21	1.11 (1.02, 1.20)	0.98	1.06 (0.98, 1.12)	0.89	1.31 (1.10, 1.42)	0.92
28	1.12 (1.01, 1.23)	0.81	1.06 (0.96, 1.16)	0.87	1.32 (1,09, 1.55)	0.89
Df for lag-response:						
5	1.07 (1.04, 1.10)	0.89	1.04 (1.02, 1.06)	0.91	1.30 (1.18, 1.42)	0.93
Df for year for seasonal control						
4	1.10 (1.03, 1.17)	0.92	1.05 (1.01, 1.09)	0.98	1.33 (1.20, 1.51)	0.96
10	1.09 (1.04, 1.14)	0.95	1.04 (1.01, 1.07)	0.99	1.28 (1.20, 1.36)	0.97
Non-flood exposure events**						
	1.10 (1.02, 1.19)	0.89	1.07 (1.02, 1.12)	0.85	1.23 (1.12, 1.34)	0.77

*P value calculated based on the difference between the primary results and the result calculated after modification

** We discarded the extreme rainfall events that were simultaneously identified as flood exposure days on the same day and within the subsequent 14 days to specifically test the impact of extreme rainfall independent of flood exposure.

Table S4 Pooled cumulative relative risks (95% confidence intervals) of mortality associated with extreme rainfall events using alternative precipitation datasets.

		ERA5-Land (Primary)	MSWEP*		ERA5-Land (After 2001)	IMERG**	
	No. of days						
	1-in-1	70,844	71,249		34,139	34,412	
	1-in-2	20,908	23,120		3,640	3,461	
	1-in-5	16,041	16,199		1,507	1,621	
	Estimated RR and 95% confidence interval			**P-value*** **			**P-value**
All causes	1-in-1	0.98 (0.95 - 1.01)	0.99 (0.96 - 1.02)	0.95	0.98 (0.94 - 1.02)	0.98 (0.94 – 1.02)	0.98
	1-in-2	1.03 (0.98 - 1.08)	1.05 (1.01 - 1.09)	0.93	1.04 (1.02 - 1.06)	1.05 (1.02 - 1.08)	0.93
	1-in-5	1.08 (1.05, 1.11)	1.08 (1.03 - 1.13)	0.91	1.10 (1.04 - 1.16)	1.09 (1.04 - 1.14)	0.94
Cardio-vascular	1-in-1	0.98 (0.96 - 1.00)	0.98 (0.95 - 1.01)	0.94	1.03 (0.97 - 1.09)	0.98 (0.95 - 1.01)	0.81
	1-in-2	1.01 (0.98 - 1.04)	0.97 (0.94 - 1.00)	0.80	1.02 (0.96 - 1.06)	1.02 (0.95 - 1.09)	0.92
	1-in-5	1.05 (1.02 - 1.08)	1.06 (1.01 - 1.11)	0.89	1.04 (1.02 - 1.06)	1.06 (1.02 - 1.10)	0.79
Respirat-ory	1-in-1	0.95 (0.89 - 1.02)	0.98 (0.96 - 1.00)	0.94	1.03 (0.96 - 1.10)	1.02 (0.98 - 1.06)	0.93
	1-in-2	1.14 (1.05 - 1.23)	1.16 (1.09 - 1.23)	0.68	1.15 (1.06 - 1.21)	1.18 (1.12 - 1.24)	0.81
	1-in-5	1.29 (1.19 - 1.39)	1.28 (1.20 – 1.36)	0.73	1.23 (1.05 - 1.41)	1.24 (1.10 - 1. 38)	0.78

* The Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset has employed the same methodology as the primary analysis for defining extreme rainfall events.In addition, rainfall days identified as snowfall events in the main analysis have been excluded here as well.

** The Integrated Multi-satelliteR Retrievals for Global Precipitation Measurement (IMERG) has adopted the same methodology as the primary analysis to define extreme rainfall events starting from the year 2001, since the dataset only became available after June 2000. Furthermore, to ensure comparability between different results, the ERA5-Land dataset has also been reanalyzed starting from 2001. In addition, rainfall days identified as snowfall events in the main analysis have been excluded here as well.

*** *P* value calculated based on the difference between the primary results and the result calculated after modification

Table S5 Pooled cumulative relative risks (95% confidence intervals) of mortality associated with the extreme rainfall events in 113 cities*

	Without adjustment	With adjustment**	*P*-value	With adjustment***	*P*-value	With adjustment****	*P*-value
All causes							
1-in-1	1.01 (0.98 - 1.04)	0.99 (0.96 - 1.02)	0.83	0.98 (0.96 - 1.00)	0.89	1.01 (0.97 - 1.05)	0.94
1-in-2	1.02 (0.94 - 1.10)	0.99 (0.95 - 1.03)	0.75	0.99 (0.94 - 1.04)	0.79	0.98 (0.95 - 1.01)	0.84
1-in-5	1.04 (1.01 - 1.07)	1.05 (1.01 - 1.09)	0.85	1.04 (1.01 - 1.07)	0.79	1.03 (1.00 - 1.06)	0.84
Cardiovascular							
1-in-1	0.98 (0.95 - 1.01)	0.98 (0.95 - 1.01)	0.94	0.98 (0.95 - 1.01)	0.93	0.99 (0.97 - 1.01)	0.93
1-in-2	0.99 (0.97 - 1.01)	0.99 (0.96 - 1.02)	0.85	1.00 (0.98 - 1.02)	0.89	1.00 (0.97 - 1.03)	0.83
1-in-5	1.04 (0.99 - 1.09)	1.03 (0.99 - 1.07)	0.86	1.04 (1.01 - 1.07)	0.89	1.02 (0.98 - 1.08)	0.79
Respiratory							
1-in-1	0.98 (0.96 - 1.01)	0.98 (0.95 - 1.01)	0.90	1.00 (0.98 - 1.02)	0.83	0.99 (0.96 – 1.02)	0.89
1-in-2	1.08 (1.03 - 1.13)	1.07 (1.01 - 1.13)	0.82	1.10 (1.04 -1.16)	0.80	1.04 (1.01 - 1.07)	0.78
1-in-5	1.18 (1.06 - 1.30)	1.13 (1.07 – 1.19)	0.69	1.18 (1.03 - 1.32)	0.72	1.17 (1.07 - 1.27)	0.81

* The 113 cities were chosen because they had accessible data on air pollution, including $PM_{2.5}$ and maximum 8-hour ozone levels at the same time. Additionally, these cities had a sufficient number of extreme rainfall days within the study period, meeting the criteria of experiencing more than two extreme rainfall events with a 5-year return period.

** Adjustments were made for controlling the concurrent day (lag 0 day) 24-hour $PM_{2.5}$ concentrations and 8-hour ozone levels in 113 cities.

*** Adjustments were made for controlling the lag 1 d of 24-hour $PM_{2.5}$ concentrations and 8-hour ozone levels in 113 cities.

**** Adjustments were made for controlling the lag 2 d of 24-hour $PM_{2.5}$ concentrations and 8-hour ozone levels in 113 cities.

Table S6 Annual frequency of identified extreme rainfall events by climatic region (per/year)

	No. of 1-in-1 per year	No. of 1-in-2 per year	No. of 1-in-5 per year
Tropical (n=81)	9.13	0.88	0.39
Arid (n=48)	3.38	0.73	0.26
Temperate (n=405)	2.90	0.53	0.21
Continental (n=111)	3.84	0.58	0.20

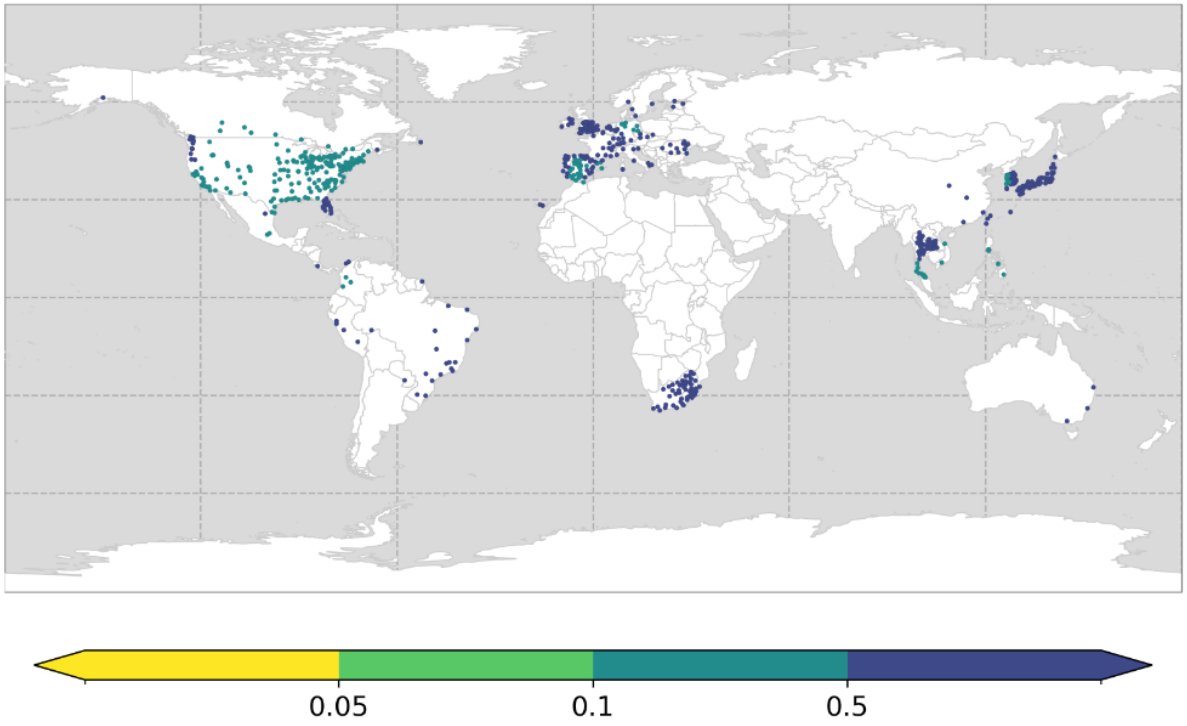

Figure S1 Kolmogorov-Smirnov test results at 5% significance level for GEV distribution fit to annual maximum daily rainfall of each location. *P*-values here represent cluster-based outcomes for included locations.

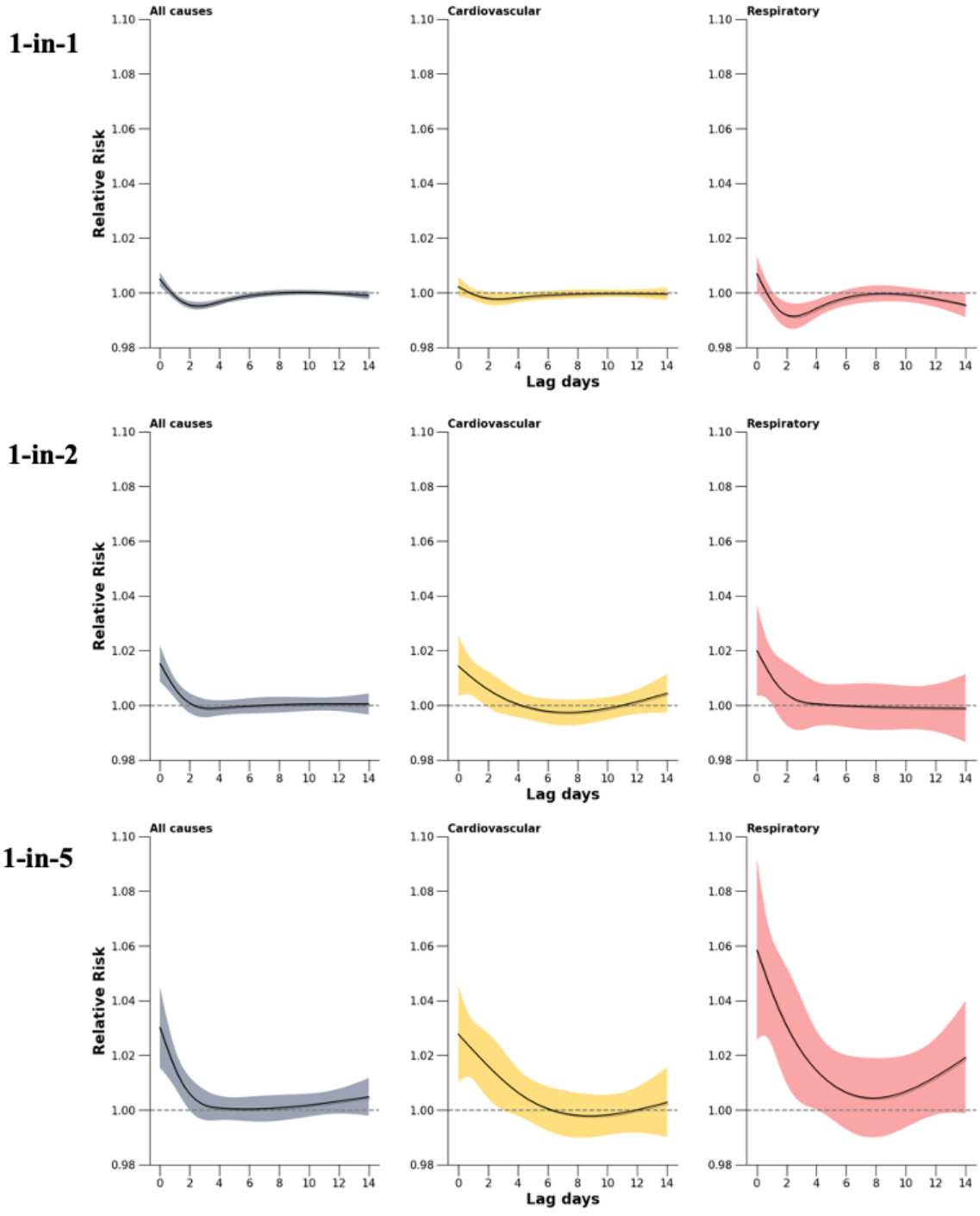

Figure S2 Relative risks of mortality associated with exposure to extreme rainfall events with 3 different return periods during lag 0-14 days in 645 cities from 35 countries/regions

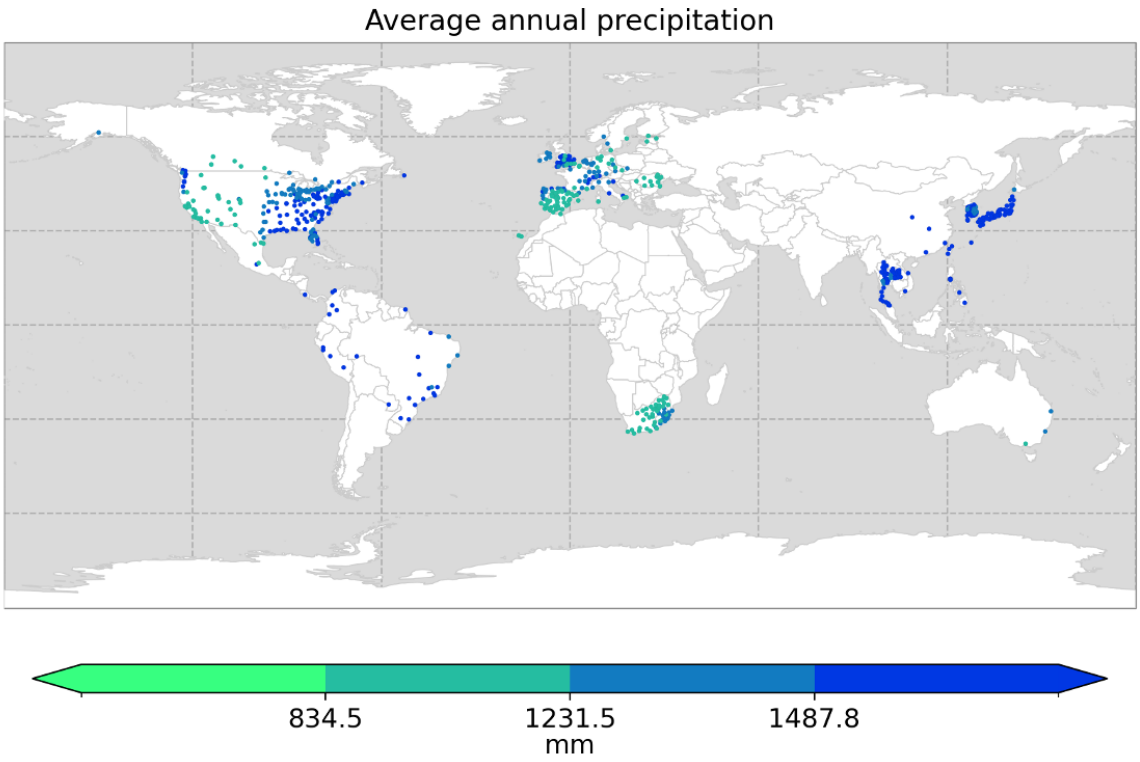

Figure S3 Average annual precipitation over 645 cities from 1980 to 2020.

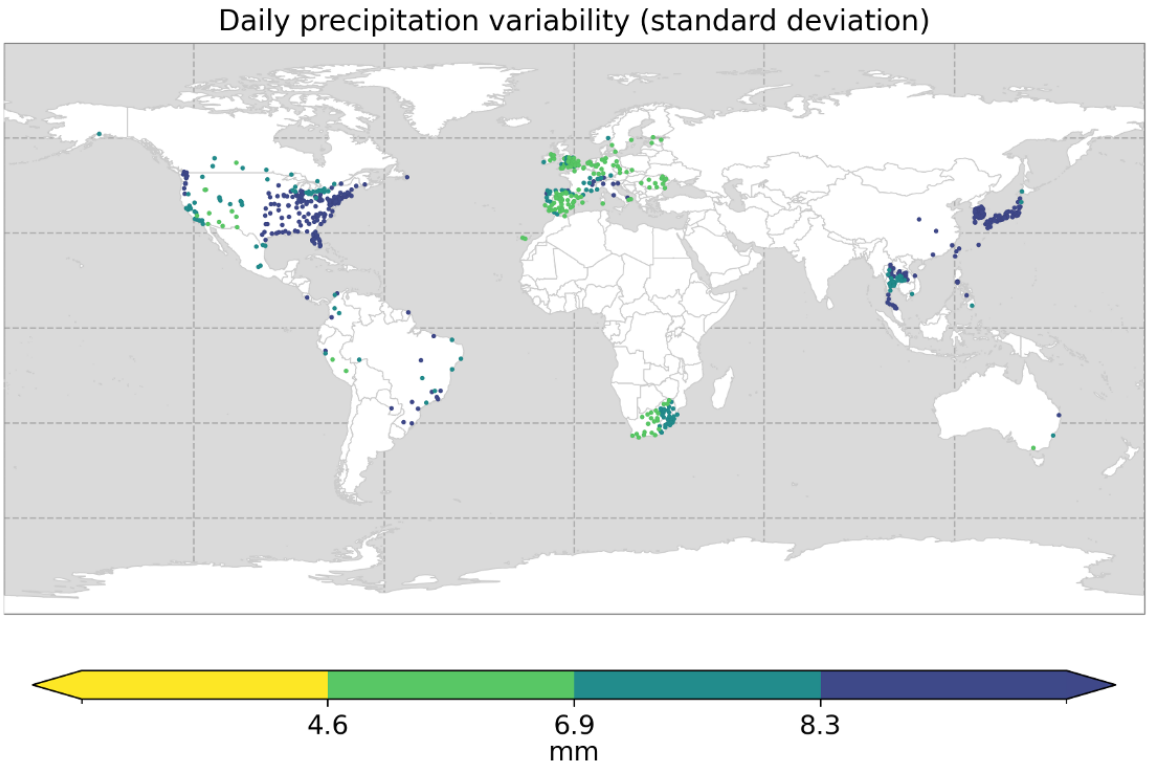

Figure S4 Daily precipitation variability (standard deviation for all rainfall days) over 645 cities from 1980 to 2020.

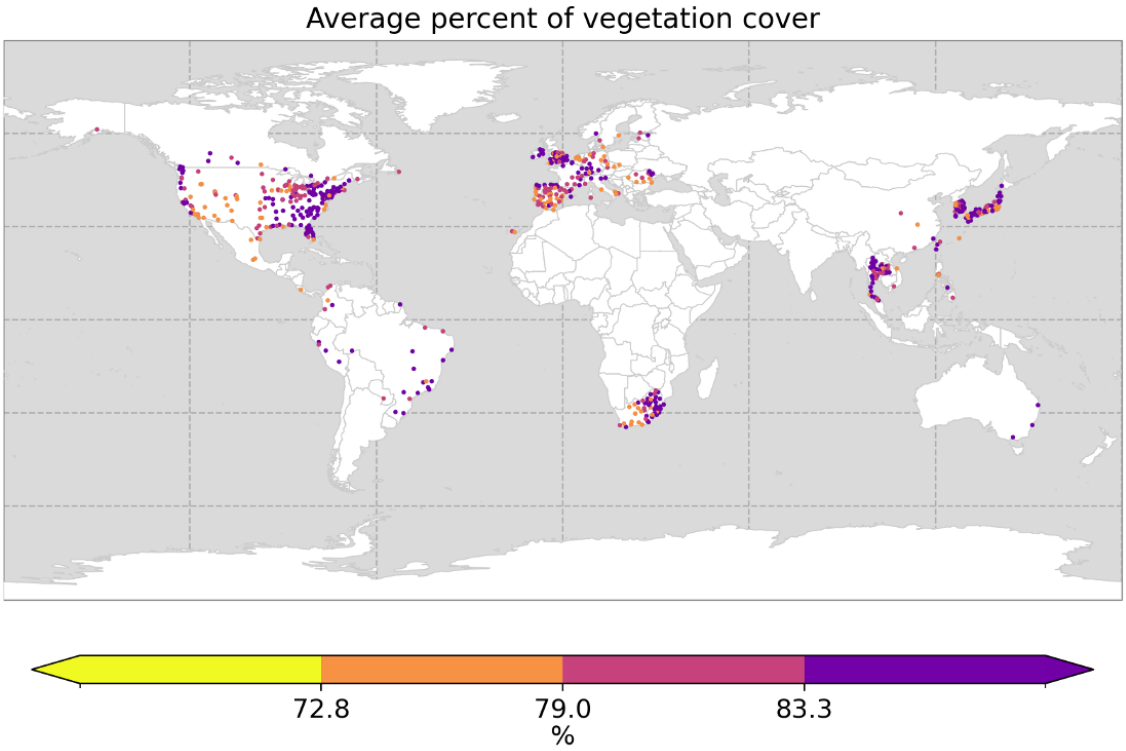

Figure S5 Average percent of vegetation cover over 645 cities.

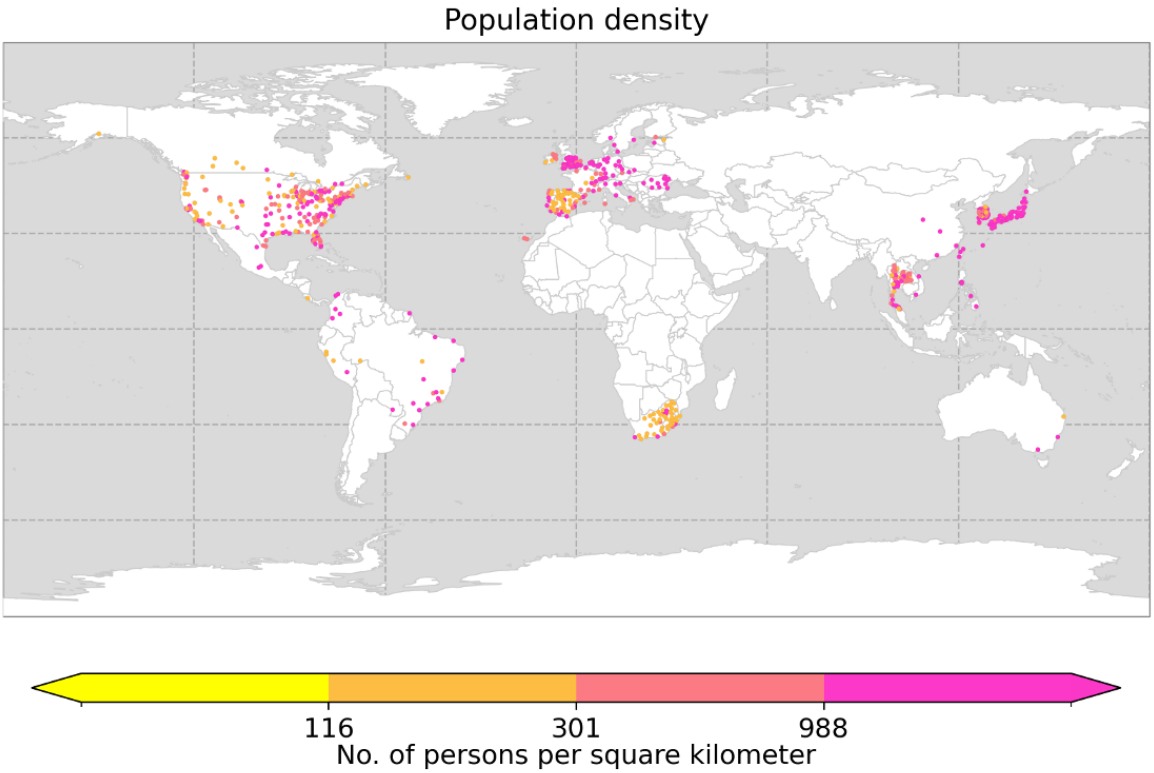

Figure S6 Population density over 645 cities.

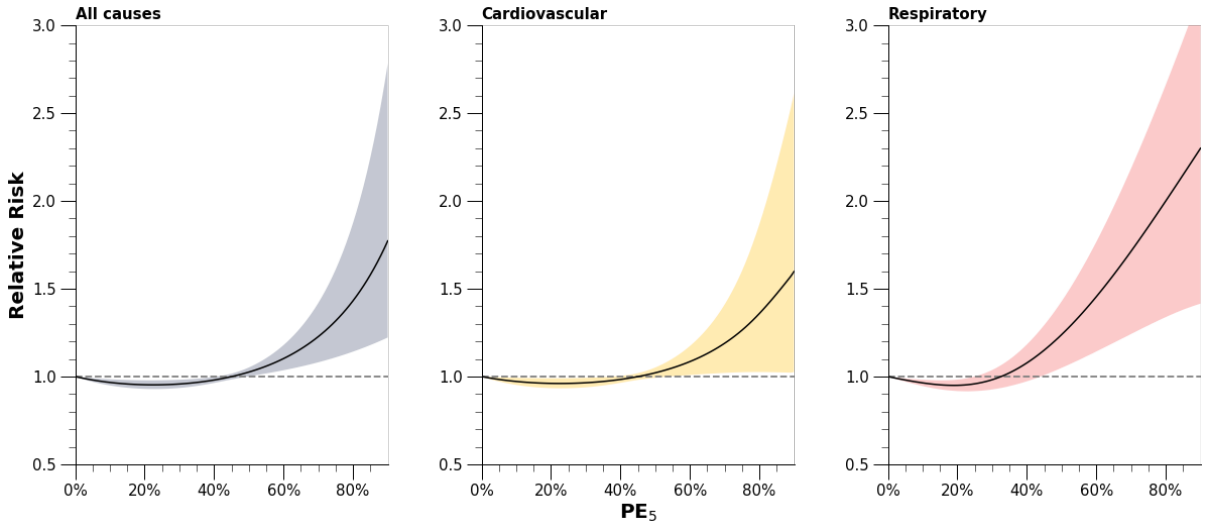

Figure S7 Exposure-response function of the relative risks of all-causes, cardiovascular, and respiratory mortality associated with daily accumulated rainfall exceeding the five-year return period threshold (*PE_5*).

References

1. Gasparrini A, Guo Y, Hashizume M, et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The lancet* 2015;386(9991):369-75.
2. Beck HE, Zimmermann NE, McVicar TR, et al. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific data* 2018;5(1):1-12.
3. Likas A, Vlassis N, Verbeek JJ. The global k-means clustering algorithm. *Pattern recognition* 2003;36(2):451-61.
4. Blanchet J, Ceresetti D, Molinié G, et al. A regional GEV scale-invariant framework for Intensity–Duration–Frequency analysis. *Journal of Hydrology* 2016;540:82-95.
5. Fowler H, Kilsby C. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. *International Journal of Climatology: A Journal of the Royal Meteorological Society* 2003;23(11):1313-34.
6. Zorzetto E, Botter G, Marani M. On the emergence of rainfall extremes from ordinary events. *Geophys Res Lett* 2016;43(15):8076-82.
7. Papalexiou SM, Koutsoyiannis D. Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research* 2013;49(1):187-201.
8. DeGaetano AT. A Smirnov test–based clustering algorithm with application to extreme precipitation data. *Water resources research* 1998;34(2):169-76.
9. de Carvalho JRP, Assad ED, de Oliveira AF, et al. Annual maximum daily rainfall trends in the Midwest, southeast and southern Brazil in the last 71 years. *Weather and Climate Extremes* 2014;5:7-15.
10. Massey Jr FJ. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association* 1951;46(253):68-78.
11. Muñoz-Sabater J, Dutra E, Agustí-Panareda A, et al. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth system science data* 2021;13(9):4349-83.
12. Doxsey-Whitfield E, MacManus K, Adamo SB, et al. Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4. *Papers in Applied Geography* 2015;1(3):226-34.
13. DiMiceli C, Sohlberg R, Townshend J. MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V061. *NASA EOSDIS Land Process DAAC* 2022
14. DiMiceli C, Townshend J, Carroll M, et al. Evolution of the representation of global vegetation by vegetation continuous fields. *Remote Sens Environ* 2021;254:112271.
15. Choi HM, Lee W, Roye D, et al. Effect modification of greenness on the association between heat and mortality: A multi-city multi-country study. *EBioMedicine* 2022;84
16. Urban green spaces: potentials and constraints for urban adaptation to climate change. Resilient Cities: Cities and Adaptation to Climate Change-Proceedings of the Global Forum 2010; 2011. Springer.
17. Yang Z, Huang W, McKenzie JE, et al. Mortality risks associated with floods in 761 communities worldwide: time series study. *bmj* 2023;383
18. He C, Kim H, Hashizume M, et al. The overlooked health impacts of extreme rainfall exposure in 30 East Asian cities. *Nature Sustainability* 2024;7(4):423-31.