Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Missed Reports of Extreme Heat Events Exacerbate Vulnerability and Losses

Haozhi Pan

panhaozhi@sjtu.edu.cn

Shanghai Jiao Tong University https://orcid.org/0000-0002-0709-632X

Shiyu Liang

Shanghai Jiao Tong University

Yiming Yang

Shanghai Jiao Tong University https://orcid.org/0000-0002-4283-2804

Ye Zheng

Shanghai Jiao Tong University

Chaofan Sun

Shanghai Jiao Tong University

Shenghao Yan

Shanghai Jiao Tong University

Hanbo Huang

Shanghai Jiao Tong University

Hao Zheng

Shanghai Jiao Tong University

Temitope Sogbanmu

University of Lagos

Alecia Bennett-Bryan

Caribbean Biodiversity Fund

Joe Mulligan

KTH Royal Institute of Technology

Zahra Kalantari

KTH Royal Institute of Technology

Fengping Wang

Shanghai Jiao Tong University https://orcid.org/0000-0002-3429-8410

Zhifu Mi

University College London https://orcid.org/0000-0001-8106-0694

Enhui Liao

Shanghai Jiao Tong University https://orcid.org/0000-0002-5516-7831

Article

Keywords: Extreme heat, Media reporting bias, Climate vulnerability, Multilingual large-language models, Climate Data Justice

Posted Date: April 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9186502/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: There is **NO** Competing Interest.

Missed Reports of Extreme Heat Events Exacerbate Vulnerability and Losses

Shiyu Liang[1†], Yiming Yang[2†], Ye Zheng[1†], Haozhi Pan[2*†], Chaofan Sun[3*†], Shenghao Yan[2], Hanbo Huang[1], Hao Zheng[1], Temitope Sogbanmu[4], Alecia Bennett-Bryan[5], Joe Mulligan[6], Zahra Kalantari[7*], Fengping Wang[3], Zhifu Mi[8], Enhui Liao[3*]

[1]John Hopcroft Center for Computer Science, Shanghai Jiao Tong University, Shanghai, China.
[2]School of International and Public Affairs, Shanghai Jiao Tong University, Shanghai, China.
[3]School of Oceanography, Shanghai Jiao Tong University, Shanghai, China.
[4]Ecotoxicology and Conservation Unit, Department of Zoology, University of Lagos, Akoka, Lagos, Nigeria.
[5]Caribbean Biodiversity Fund, Saint Andrew, Jamaica.
[6]Division of Strategic Sustainability Studies, KTH Royal Institute of Technology, Stockholm, Sweden.
[7]Department of Sustainable Development, Environmental Science and Engineering, KTH Royal Institute of Technology, Stockholm, Sweden.
[8]The Bartlett School of Sustainable Construction, University College London, London, UK.

*Corresponding author(s). E-mail(s): panhaozhi@sjtu.edu.cn; scf024@sjtu.edu.cn; zahrak@kth.se; ehliao@sjtu.edu.cn;
†These authors contributed equally to this work.

Abstract

Climate risks are systematically under-observed in the world's most vulnerable regions, compounding exposure by delaying preparedness and adaptation. Using 2.1 million news articles across 184 countries and 40 languages, combined with AI-assisted geolocation and human-in-the-loop validation, we expose global under-reporting of extreme heat events. We find that only a tiny fraction of identified extreme heat events enters widely used global disaster databases, while 63.38% remain unreported even across the most comprehensive global and local news coverage. Under-reporting is concentrated in places like equatorial Africa, the Amazon basin, small island states, such as the Caribbean, and polar adjacent regions, where limited early warning capacity, high vulnerability, and linguistic distance from dominant languages intersect. Under the high-emissions scenario SSP5-8.5, by 2050, the largest future heat burdens are concentrated in highly vulnerable countries that coincide with the most severe under-reporting of extreme heat. These countries are projected to incur economic losses 60.09% higher and mortality-related losses 46.40% higher than better-reported countries. Yet pervasive under-reporting suggests that even these severe loss projections remain conservative for regions with the highest vulnerability, reinforcing the urgency of loss-and-damage action.

Keywords: Extreme heat, Media reporting bias, Climate vulnerability, Multilingual large-language models, Climate Data Justice

1 Introduction

Climate change is intensifying the frequency, severity, and reach of extreme weather and climate events globally [1, 2]. However, these crises remain chronically under-reported in many parts of the world, especially in regions with limited monitoring infrastructure, geopolitical isolation, and linguistic exclusion from the English-dominated global information sphere [3, 4], while disaster records tend to be biased toward large, insured losses in well-connected countries. The result is a skewed global understanding of climate risk, resulting in fragmented disaster datasets and reporting blind spots that distort both perception and policy [5, 6].

Of critical importance for climate resilience, many of the most under-reported regions are also the most exposed and least equipped to cope with climate hazards. Small island states, equatorial lowlands, mountainous highlands, and polar regions face disproportionate threats from compound extremes, such as sea-level rise, glacial retreat, drought, and heatwaves [7, 8]. The very factors that make these regions vulnerable to heatwaves also suppress their visibility. Sparse weather monitoring networks, limited internet and news media penetration, and linguistic isolation mean that many extreme heat events never enter the global news stream [9].

For example, the most widely used global disaster database, EM-DAT, lists very sparse significant heatwave impacts in Sub-Saharan Africa since 1900 [10], despite the fact that the region is one of the most evident sufferers of heatwaves. On the other hand, more than half of EM-DAT's recorded heatwave events come from just 9 countries (mostly in Europe and Asia). Failure to bridge reporting voids for the four-decade intensification of Caribbean heat stress threatens to leave the 60% of the population projected for severe water stress under SSP2 scenarios to suffer within a monitoring and adaptation vacuum [11]. Heat and drought in rural Africa or the Indian subcontinent, for instance, may be reported only in local dialects or social media, if at all, and thus escape international notice [12, 13]. As a result, large swathes of the Global South and highly exposed regions remain effectively invisible in global climate discourse and datasets [5, 14]. As a result, large swathes of the Global South and exposed regions remain effectively invisible to global climate discourse and datasets, reinforcing a form of epistemic and policy marginalization [5, 14].

This invisibility is not merely informational; it has structural consequences. Under-reporting reduces the visibility of needs, leading to the misallocation of adaptation funds, insurance mechanisms, and humanitarian aid [7, 14]. Vulnerable populations, who have contributed the least to global emissions, are systematically under-represented in both datasets and decision-making, compounding climate injustice and leaving communities unprepared for emerging extremes [1, 5, 11, 14].

To systematically quantify these global blind spots, we conducted a comprehensive survey of extreme heat events reported in news media and compared them against physically simulated heatwave occurrences. Here we introduce a novel AI-driven framework to identify and expose these hidden events (Fig. 1). The overall pipeline integrates large-scale multilingual news aggregation, AI-assisted screening and structuring, and spatial validation within a unified workflow. We leverage state-of-the-art large language models trained on an exhaustive search on global and local multilingual news to detect and geo-locate extreme heat events [15]. In parallel, we incorporate high-resolution gridded climate projections, bias-corrected reanalyses, and socioeconomic vulnerability datasets to map regions where recorded impacts fall short of any media reports [16, 17]. Then, we identified and highlighted potential socioeconomic and health losses that would be subject to systemic under-reporting by 2050 under increased climate extreme scenarios.

Our results reveal large and systematic reporting gaps in extreme heat impacts, with 63.38% of identifiable extreme heat events lacking any corresponding media or disaster database record globally. Exhaustive AI driven analysis of multilingual news show that reported heat events capture only a small fraction of actual occurrences, with only a extremely fraction of events recorded in EM-DAT and 2.68% covered by influential global media, with pronounced under-reporting concentrated in places such as equatorial Africa, South Asia, and polar regions. These reporting gaps are associated with structural conditions. Countries with limited climate monitoring capacity, higher environmental and socioeconomic vulnerability, and lower climate readiness tend to exhibit larger reporting deficits. Under SSP5.8.5, although the most under-reported countries do not currently show significantly higher

observed economic or health losses, by 2050, countries in the bottom 10% of news coverage are projected to experience economic damages that are 60.09% higher and mortality related losses 46.40% higher than in well-reported countries, with particularly pronounced impacts across Africa.

2 Mainstream Media Severely Under-report while Local News Fills Partial Gap

Our analysis indicates that EM-DAT captures fewer than ten extreme heat events globally per year, despite physical simulations identifying on the order of 30,000 heat events annually (Fig. 2f). This stark imbalance indicates that the vast majority of extreme heat events remain absent from the datasets commonly used for global climate risk assessment. Mainstream global news coverage is similarly sparse and geographically uneven. Only 2.68% of extreme heat events receive attention from editorially influential global and regional media outlets, defined by large audiences, substantial reach, and frequent citation or professional recognition. In contrast, wealthier regions in North America, Northern and Western Europe, and Australia exhibit comparatively comprehensive coverage, with EM-DAT records disproportionately concentrated in a small number of countries, even as comparable or larger heat burdens occur elsewhere.

At the global scale, extreme heat impacts are systematically under-observed. Using an LLM-based survey covering more than 180 countries and 40 languages, we find that 63.38% of identifiable extreme heat events have no corresponding media report or disaster database entry, indicating a pervasive visibility gap even under exhaustive global and local news searches.

This visibility gap persists despite recent growth in media coverage (Fig. 2). Although reporting of extreme heat has increased gradually since around 2015, these improvements are highly uneven across regions. Gains are concentrated in well-resourced countries, while poorer, remote, and linguistically marginalized regions remain persistently under-covered in both global and local news.

Our results reveal profound and systematic geographic asymmetries in the global observation of extreme heat impacts (Fig. 3). Under-reporting is acutely concentrated in climate-vulnerable regions such as polar areas, equatorial Africa, the Amazon basin, and parts of Asia. Conversely, several high-income regions exhibit substantial media coverage of heat episodes that do not meet standard scientific thresholds for extreme events.

Pronounced reporting deficits emerge in regions where media infrastructure and monitoring capacity are limited. When adjusted for population exposure (Fig. 3C), large populations in Africa, Eastern Europe, parts of Asia, and Pacific Island states are shown to experience extreme heat events without corresponding documentation. The disparity is most severe in mainstream global media and disaster databases, which are disproportionately anchored in English-dominated media ecosystems and Global North institutional networks. As a result, Panels B and D show that under-reporting by influential international outlets is especially acute in Western Africa, Sub-Saharan Africa, southeastern Asia, and South America, while local multilingual news only partially fills this gap. Furthermore, the distribution of events in the primary disaster database EM-DAT (Fig. 3E) reveals reporting "bright spots" almost exclusively in Western Europe, North America, and East Asia, demonstrating a significant spatial bias in the foundational records used for risk assessment.

3 Under-reporting Deepens Vulnerability, Under-preparedness, Colonial Legacy and Language Barriers

Our analysis further reveals that the most severe under-reporting of extreme heat events is not randomly distributed but spatially concentrated in regions already identified as climate hotspots, areas where high physical exposure intersects with low adaptive capacity (Fig. 4 A-B), encompassing climate-critical zones such as the Arctic, Siberia, the Congo Basin, West Africa, Amazonia, and numerous Small Island Developing States (SIDS).

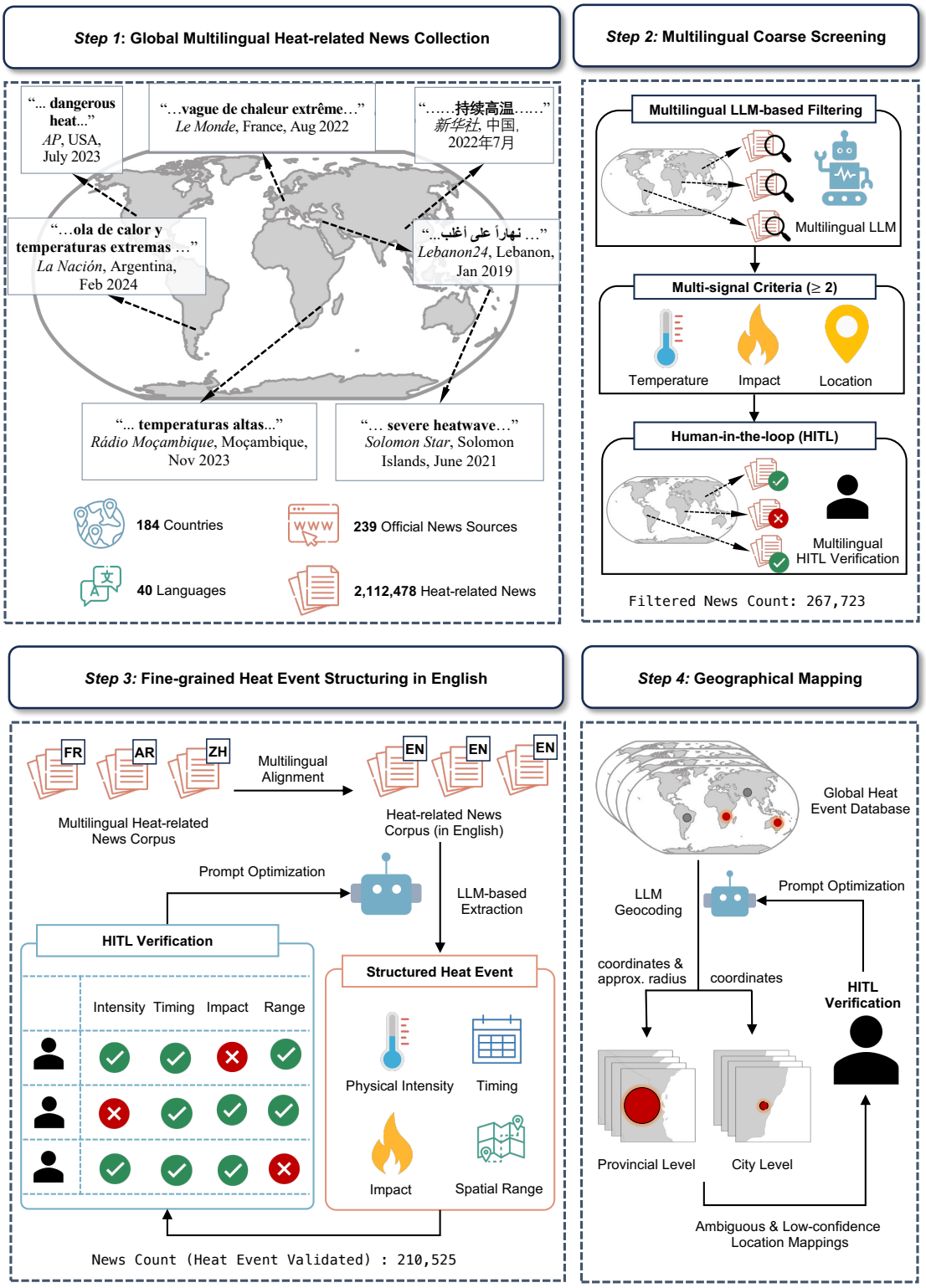


Fig. 1 AI–human collaborative identification and structuring of global extreme heat event reporting. a, Large-scale multilingual news aggregation spanning 184 countries and 40 languages (2015–2024). **b**, LLM-based screening coupled with human-in-the-loop verification for candidate extreme heat events. **c**, Cross-lingual alignment and AI-assisted structuring of heat events, with iterative human validation. **d**, LLM-enabled geocoding with manual verification at provincial and city scales, establishing the global heat event news database.

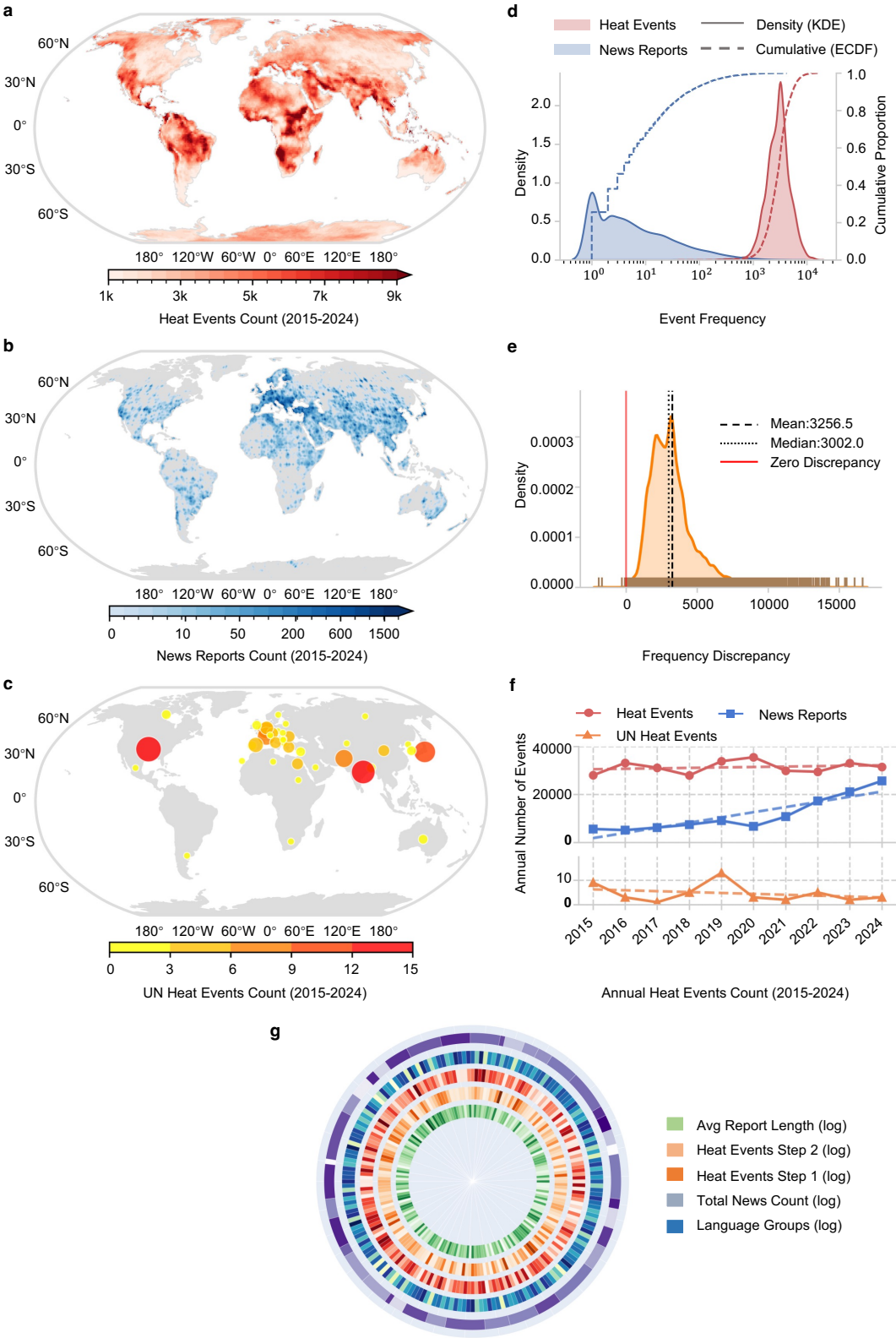


Fig. 2 Global and local extreme heat event news visualization. a–c, Spatial distribution of extreme heat events factually identified via meteorological data **(a)**, compared to the count of corresponding news reports **(b)** and heat events recorded in the UN EM-DAT database **(c)** from 2015 to 2024. **d–f**, Statistical distribution of event frequencies **(d)** and frequency discrepancies **(e)**, alongside annual trends **(f)** showing the divergence between actual occurrences and media visibility over time. **g**, Multidimensional visualization of reporting characteristics, including report length, total news count, and language group distribution.

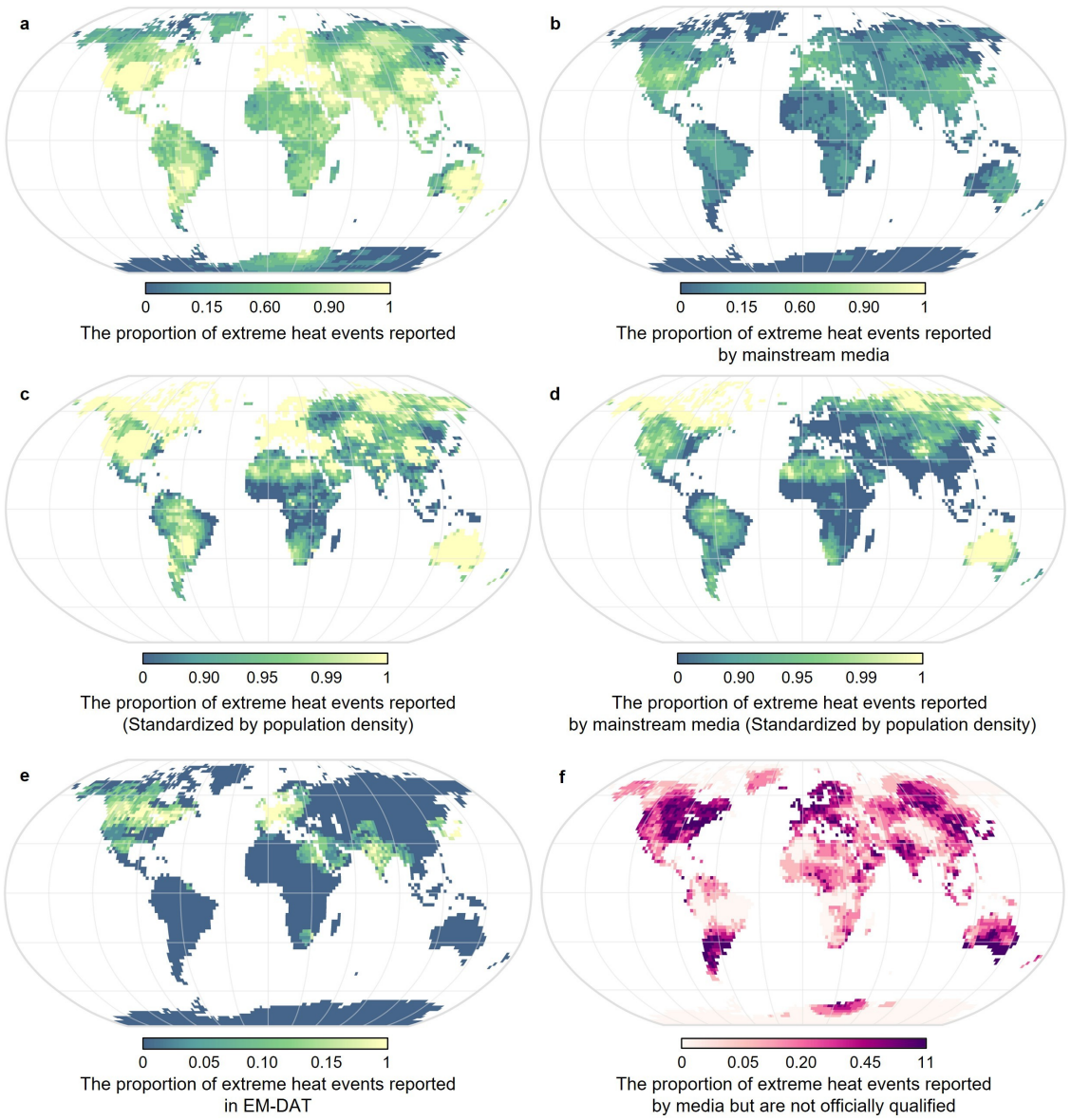


Fig. 3 Comparison of actual extreme heat events, news reports, and disaster database records. a, Global map of the proportion of extreme heat events not reported, with dark shading highlighting regions of high under observation. **b**, Proportion of events missed by mainstream global media. **c–d**, Population-weighted reporting deficits, identifying areas where significant human exposure occurs without documentation. **e**, Proportion of meteorological heat events absent from the EM-DAT database. **f**, Instances of "over-reporting," showing heat events documented in the media that do not meet standard scientific thresholds for extreme temperature or duration.

The relationship between latitude and news coverage follows a distinct M-shaped pattern (Fig. 4C), with systematic troughs in equatorial and high-latitude regions. This pattern suggests that the combination of climatic extremity, sparse monitoring infrastructure, and distance from global media hubs creates a double burden of invisibility for communities in these zones.

Critically, these under-reported zones align closely with regions of elevated institutional and socioeconomic vulnerability. Fig. 4D demonstrates that systemic under-reporting is strongly associated with structural disadvantage. Small Island Developing States (SIDS) and countries with low climate preparedness (as measured by the ND-GAIN Index) exhibit substantially higher shares of extreme heat events without any media or database records. While Least Developed Countries (LDCs) show elevated under-reporting, developing countries outside the LDC category, often characterized by rapid urbanization, strained infrastructure, and intermediate institutional capacity, record the largest absolute shares of unreported events. This pattern underscores how reporting gaps are embedded within

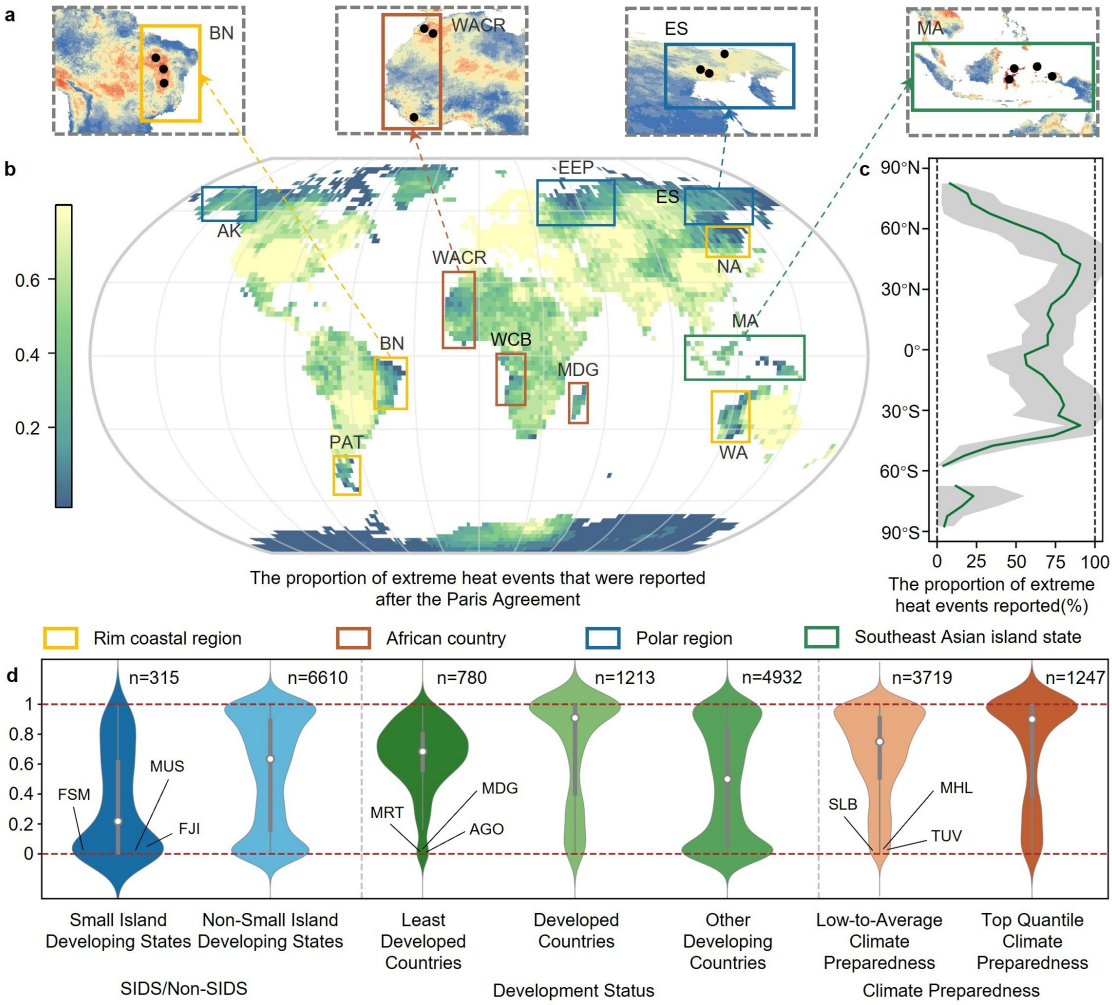


Fig. 4 Characteristics of regions with severe extreme heat under-reporting. **a**–**b**, Global maps identifying reporting gaps across 184 countries, with boxes highlighting equatorial and high-latitude regions. **c**, Plot showing the ratio of reported to detected extreme heat events as a function of latitude. **d**, Violin plots of the proportion of unreported events, stratified by UN development categories and ND-GAIN climate preparedness levels. White markers indicate the median; vertical lines represent the interquartile range. Regional abbreviations: AK (Alaska), EEP (East European Plain), ES (Eastern Siberia), NA (Northeast Asia), WACR (West Africa's Coastal Region), BN (Brasil Nordeste), WCB (Western Congo Basin), MDG (Madagascar), MA (Malay Archipelago), WA (Western Australia), and PAT (Patagonia). Country abbreviations: FSM (Federated States of Micronesia), MUS (Mauritius), FJI (Fiji), MRT (Mauritania), MDG (Madagascar), AGO (Angola), SLB (Solomon Islands), MHL (Marshall Islands), and TUV (Tuvalu).

broader matrices of vulnerability, where limited monitoring capacity, constrained media ecosystems, and lower climate readiness converge to render extreme heat impacts invisible to the global record.

The systemic under-reporting of extreme heat is fundamentally anchored in structural deficits in monitoring capacity and information access (Fig. 5A-B). Exemplified by many African nations where internet penetration remains below 30 percentage. Countries with lower climate preparedness scores and limited ICT infrastructure exhibit markedly larger shares of undetected heat events. This underscores a critical feedback loop. Regions with the weakest capacity to observe and communicate climate hazards are precisely those where such hazards are most poorly documented, reinforcing a cycle of invisibility and unmet adaptation needs.

Beyond infrastructure, reporting patterns are also shaped by linguistic geography and information connectivity (Fig. 5C–D). At the global scale, proximity to mainstream reporting languages is significantly and positively associated with higher heat event visibility, consistent with the centralization of global media flows. However, this relationship is not uniform. Despite formal use of colonial languages such as English, French, and Portuguese, many Francophone and Anglophone African countries remain

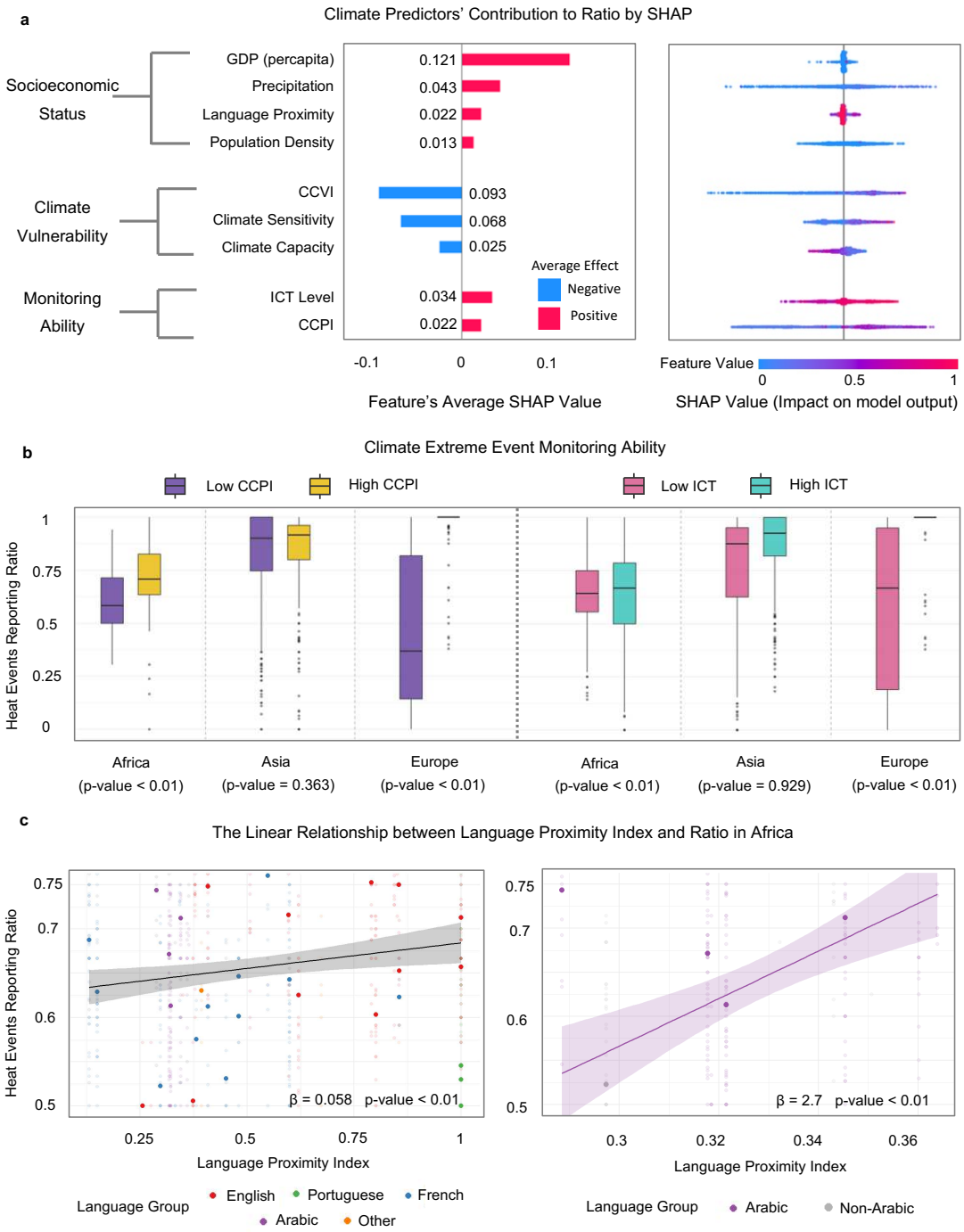


Fig. 5 Structural drivers of extreme heat under-reporting. **a**, Variable importance and marginal effects of socioeconomic, vulnerability, and monitoring predictors on the reporting ratio. **b**, Regional comparison of heat event reporting ratios stratified by climate policy performance (CCPI) and information technology (ICT) levels across Africa, Asia, and Europe. **c**, Linear regressions illustrating the relationship between the Language Proximity Index and reporting ratios in Africa, separated by major language groups. Statistical significance is indicated by p-values and confidence.

severely under-covered. In contrast, several Arabic-speaking nations, supported by contiguous media ecosystems and shared regional news networks, demonstrate comparatively higher reporting levels. Together, these patterns indicate that linguistic proximity alone does not guarantee visibility; rather, it operates through the vitality, regional embeddedness, and connectivity of media environments.

Crucially, in many under-reported regions, heat impacts are primarily discussed in vernacular and Indigenous languages, such as the over 3,000 languages spoken across Africa [18], with Nigeria alone home to 520 [19]. These languages are largely absent from global news aggregation and disaster databases, leaving rural, less-educated, and information-isolated populations particularly invisible. Thus, reporting disparities are not merely a function of technical capacity, but also of deeply embedded linguistic and colonial legacies that determine whose experiences are amplified and whose are erased from the global climate record.

4 Future Losses Will be Highest Where Heat Impacts Are Least Reported

Building on existing projections of future heatwave impacts [20], we reanalyze projected economic and health losses by stratifying countries according to the degree of extreme heat under-reporting and vulnerability identified in this study. While these forward-looking models provide crucial estimates of future risk, they remain constrained by the same biased historical data that underpin them—relying on disaster databases and economic records that systematically under-count impacts in precisely the most vulnerable regions. By overlaying our empirically derived under-reporting metrics onto these projections, we can assess how much the true burden of future heat may be underestimated for invisibilized populations. This comparison reveals systematic differences between under-reported and better-reported countries. Across countries in the top 10% of under-reporting, projected heat-related economic losses and mortality-related losses are 69.50% and 39.78% higher, respectively, than in well-reported countries (Fig. 6).

These disparities are also evident at the regional scale (Fig. 6B–C). In Asia, the gap between under-reported and other countries is the largest, with Indonesia exhibiting the highest projected economic losses and Malaysia the highest projected mortality-related losses. Both are island nations highly exposed to heat-related climate impacts yet characterized by limited visibility of extreme heat events in global reporting. African countries show the highest average projected losses overall, although the gap between under-reported and other countries is comparatively smaller. Within Africa, Rwanda stands out with high projected economic and mortality-related losses combined with substantial reporting deficits. Europe exhibits the lowest projected losses among regions, but marked disparities persist between well-reported and under-reported countries, with losses and reporting gaps concentrated primarily in Eastern Europe.

The divergence between under-reported and better-reported countries widens under higher-emissions scenarios. Under SSP5.8.5, projected losses increase more rapidly in under-reported countries than elsewhere, producing the largest relative gaps across scenarios (Fig. 6D).

Finally, we identify a compounded vulnerability pattern that intensifies over time under high-emissions scenarios. Countries that are simultaneously under-reported, classified as LDCs, and situated in the highest climate-vulnerability quantile experience the most severe projected outcomes. While these countries do not currently exhibit significantly higher observed economic or health losses, projections indicate that by 2050, these countries will experience economic losses 60.09% higher and mortality-related losses 46.40% higher than well-reported countries, and by 2060 economic losses 44.31% higher and mortality-related losses 41.57% higher. This divergence suggests that present-day reporting deficits mask emerging risks that become increasingly visible only as climate impacts intensify, highlighting a growing disconnect between present-day loss records and future vulnerability as under-reported countries account for an increasing share of projected global losses.

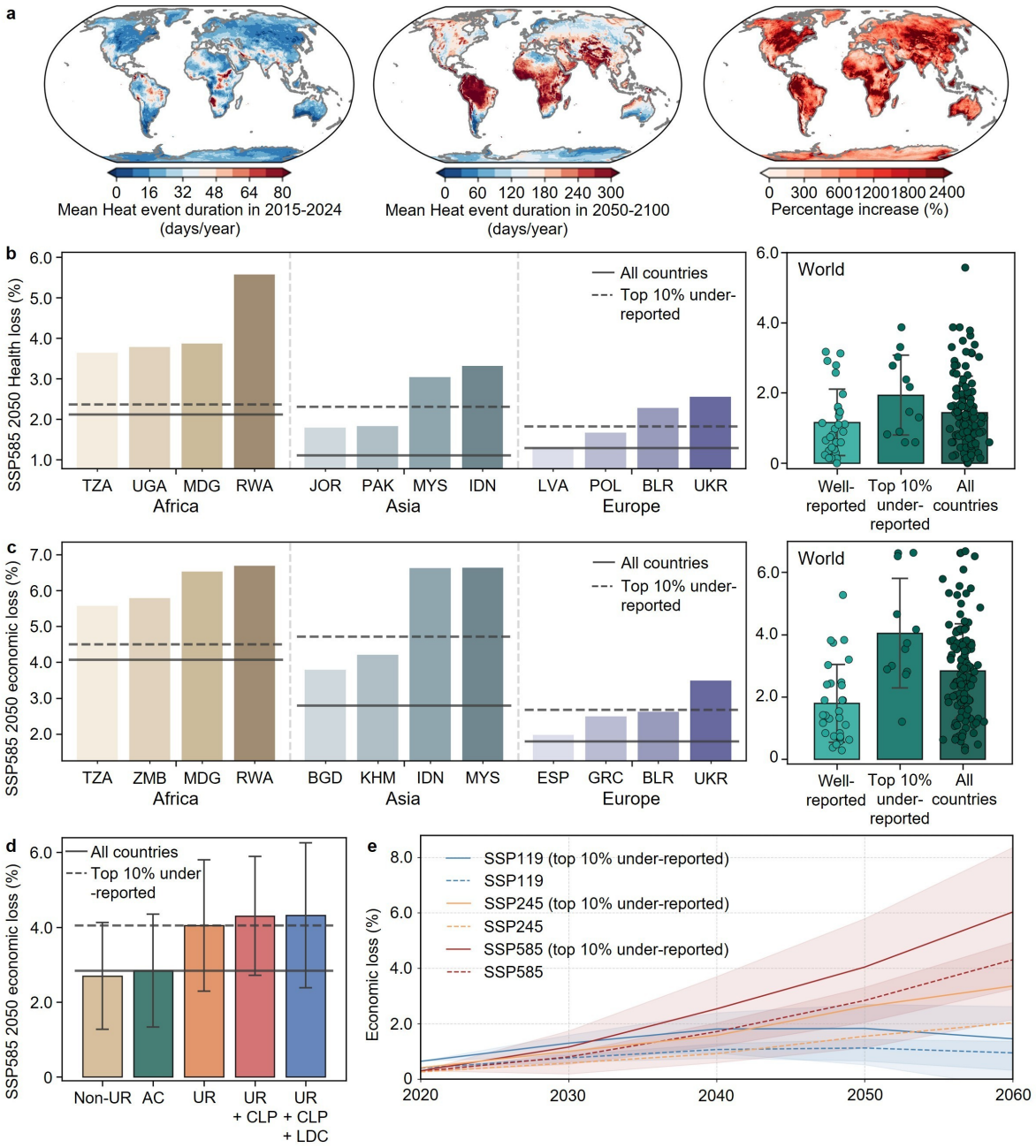


Fig. 6 Future projections of heat-related losses for countries with severe under-reporting. **a**, Global maps showing mean heat event duration for the 2015–2024 baseline and 2050–2100 projections, with the percentage increase illustrated under the SSP5.8.5 scenario. **b–c**, Projected 2050 health **(b)** and total economic **(c)** losses as a percentage of GDP, stratified by region (Africa, Asia, Europe) and compared across well-reported and under-reported country cohorts. Scatter plots indicate the distribution of individual country values. **d**, Comparison of 2050 economic losses across different vulnerability classifications, including under-reported (UR), climate least prepared (CLP), and Least Developed Countries (LDC). **e**, Projected increase in economic loss shares from 2020 to 2060 for the most under reported countries compared to the global average across three emissions scenarios (SSP1.19, SSP2.45, SSP5.8.5), with shaded areas representing 95% confidence intervals. Country abbreviations: TZA (Tanzania), UGA (Uganda), MDG (Madagascar), RWA (Rwanda), JOR (Jordan), PAK (Pakistan), MYS (Malaysia), IDN (Indonesia), LVA (Latvia), POL (Poland), BLR (Belarus), UKR (Ukraine), ZMB (Zambia), BGD (Bangladesh), KHM (Cambodia), ESP (Spain), and GRC (Greece). Group classifications: UR (Top 10% under-reported), AC (All countries), CLP (Climate least prepared), and LDC (Least developed countries).

5 Discussion and Implications

Our findings highlight the strong spatial co-occurrence of vulnerability, low preparedness, and under-reporting of extreme heat impacts. Regions characterized by weak infrastructure and limited early-warning systems consistently exhibit both higher exposure to heat extremes and lower visibility in reporting records [21]. Climate change has already severely exposed vulnerable tropical regions, sub-Saharan Africa, and South Asia [22] and will suffer disproportionally from future warming [23].

This pattern is most acute in Africa, where increasing heatwave intensity (often exacerbated by El Niño variability) intersects with rapid urbanization and land-use change. The challenge of documentation is further compounded by extreme linguistic fragmentation; the continent hosts over 3,000 Indigenous languages, with Nigeria alone accounting for 520 [18]. This linguistic diversity, while culturally rich, presents a structural barrier to centralized climate news reporting, contributing to the widening deficit between physical heat events and their public record. Populations particularly in rural areas across sub-Saharan Africa, and vulnerable populations with little or no formal education in addition to lack of or limited access to information gateways are often at a loss. Several SIDS similarly lack the fiscal space to implement observation and reporting structures necessary for adaptive capacity, leaving governments unable to fulfill their central role in helping populations cope with shocks where insurance markets remain underdeveloped and financial sectors lack resilience, limiting post-shock risk sharing and resource mobilization.

Weakened climate monitoring infrastructure has high political costs. Recent policy decisions, notably during the Trump administration, illustrate how rolling back observation networks can worsen inequalities and erode resilience. In effect, dismantling climate science domestically amplifies global inequities. When key agencies are gutted or paralyzed by politics, vulnerable communities suffer most, as they lose early warnings and information needed for adaptation [24]. In this context, reliance on historical disaster databases for loss modeling may further obscure risk in regions where vulnerability and under-reporting coincide spatially, reinforcing uncertainty in global assessments rather than reflecting a complete accounting of climate impacts [25, 26].

To address these gaps, we recommend policies to strengthen reporting and monitoring through coordinated national and global action. First, invest in and empower climate journalism in local languages. Training programs should equip reporters in vulnerable regions with scientific literacy and tools for disaster reporting. Local NGOs, health workers and citizen networks often see disasters first; formalizing their reports could capture events bypassing the media. International bodies could fund platforms where local agencies submit damage assessments to a central database.

Second, national meteorological and hydrological services in low- and middle-income countries should be strengthened. Funding should flow into the global basic observing networks, weather stations, radiosonde launches, and satellite receiving stations. Building on emerging continental and UN-led multi-hazard early-warning initiatives, sustained investment in surface observations, upper-air monitoring, and satellite data reception will be critical for both forecasting and impact validation. This technical capacity should support the integration of heat-health alerts into existing multi-hazard early warning systems (MHEWS), which is vital to ensure warnings move beyond simple air temperature to provide act

Finally, we urge rapid operationalization of loss-and-damage finance and adopt broad criteria that recognize vulnerability. Financing disbursement could use objective climate exposure indices rather than requiring formal event records. Adaptation funds should similarly consider at-risk status when allocating resources, insulating the poorest countries from measurement bias. Similarly, we suggest that insurance and relief instruments to be triggered by meteorological thresholds or by index measures, rather than reported claims.

We also acknowledge limitations in our study. Our study focuses on heat-related mortality, but many other climate impacts may also be missed, such as droughts and floods that often cause major crop failures or displacements without immediate deaths. Future work should extend these techniques to identify under-reported floods, storms, droughts, and other multi-hazard events.

6 Method

6.1 Data

Meteorological Data for Heatwave Detection and projection: Daily near-surface air temperature data were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis product. We used hourly 2-m air temperature fields spanning 1985–2024, aggregated to daily maximum temperature ($T_{\max}$). ERA5 provides global coverage at $0.25^\circ \times 0.25^\circ$ horizontal resolution and has been extensively validated for representing temperature variability and extremes. The 2015–2024 (post-Paris Agreement) period is used as the analysis window. ERA5 data were accessed through the Copernicus Climate Data Store and processed consistently across all grid cells, with missing values retained as NaN and excluded from climatological computations. Detailed preprocessing procedures, climatological baseline construction, and raster generation steps are provided in Supplementary Information, Section S1.

To assess future heatwave dynamics, we employed simulations from the Coupled Model Intercomparison Project Phase 6 (CMIP6), using output from the CESM2 (Community Earth System Model version 2) [27, 28]. Daily maximum temperature fields for 2015–2100 were extracted under the SSP5-8.5 high-emissions scenario. Model outputs were regridded to $1^\circ \times 1^\circ$ spatial resolution via bilinear interpolation to facilitate global-scale analysis and integration with socioeconomic datasets. The regridding procedure and severe heat-day projection framework are described in Supplementary Information, Section S10.1.

Multilingual Global Extreme Heat Event News Dataset: To capture reported social and ecological impacts of extreme heat, we constructed a high-resolution, multilingual global news database covering 2015–2024 (post-Paris Agreement). News articles were collected using a custom web-crawling framework sourcing at least 2 independent outlets per country across 184 countries, explicitly extending beyond English-language and Global North–centric media systems.

The pipeline processed content in 40 languages, covering main languages used in the identified from the public media platforms of the 184 countries, yielding 2,673,892 raw articles, which were subsequently deduplicated to remove syndicated or repeated content. This resulted in 2,112,478 unique articles, forming the foundation for event identification and extraction. Example corpus samples are provided in Supplementary Information, Section S3, while the complete deduplicated corpus metadata are documented in Supplementary Information, Sections S2–S3.

Future Heat-Related Socioeconomic Loss Data: Country-level projections of future heat-related economic and mortality losses are obtained from the global dataset published by Sun et al. (2024) . The dataset provides decadal-average national estimates of heat-related health losses, labour productivity losses, and indirect economic losses transmitted through global supply chains, expressed as percentages of gross domestic product (GDP), for 141 regions worldwide under multiple Shared Socioeconomic Pathways (SSPs). These data are used as externally sourced outcome variables and are merged at the country level with heatwave detection and reporting metrics derived independently in this study. No recalculation, reparameterization, or modification of the original loss estimates is performed.

Socioeconomic, Cultural Correlates and Vulnerability Indicators: To evaluate the structural context of under-reporting, we integrated several global indicators. National Information and Communication Technology (ICT) levels were sourced from the International Telecommunication Union (ITU). Gross Domestic Product (GDP) per capita and population density were obtained from the World Bank and SEDAC. Institutional preparedness and vulnerability were quantified using the ND-GAIN Country Index and the Climate Change Vulnerability Index (CCVI). Finally, a Language Proximity Index was constructed using the Ethnologue database to measure the distance between local languages and globally dominant media languages (English, French, Spanish).

6.2 Heatwaves Detection

In this study, heatwaves were identified following the definition used in the reference [29]. For each grid cell, daily maximum temperature during 2015-2024 for the historical and present-day and during

2015-2100 for future projections were compared with a 90th percentile threshold derived from the climatological baseline period 1985-2014 using an 11-day moving window centered on each calendar day. A heatwave event was defined as a period of at least five consecutive days when daily maximum temperature exceeded this threshold. Consecutive events separated by two days or fewer were merged into a single continuous event. The total number of heatwave events and the cumulative duration for each grid cell were subsequently calculated. We further conducted sensitivity analyses using alternative definitions of heat events, including periods exceeding the 95th percentile for at least two consecutive days and the 97.5th percentile for at least four consecutive days, which showed minimal impact on our results. Detailed formulations of the threshold construction, seasonal aggregation, event-merging rules, and sensitivity analyses are provided in Supplementary Information, Section S1.

6.3 Multilingual Global and Local Extreme Heat Event News Dataset

To construct a comprehensive inventory of social–ecological heat impacts, we developed a high-resolution multilingual news database covering 2015–2024 (post-Paris Agreement). The complete source selection protocol, keyword retrieval strategy, deduplication procedures, and corpus construction workflow are detailed in Supplementary Information, Sections S2–S6.

As illustrated in Fig. 1, the analytical pipeline comprises four sequential stages: (1) global multilingual news collection, (2) multilingual coarse screening using LLM-based filtering and human-in-the-loop verification, (3) fine-grained heat event structuring through cross-lingual alignment and AI-assisted extraction, and (4) geospatial mapping with hierarchical coordinate assignment and expert validation.

The process commenced with Raw News Data Collection. Specifically, we employed a custom web crawling framework to harvest news articles from at least three distinct sources across 184 countries, ensuring global coverage beyond the Anglosphere. The pipeline processed content in 40 languages to capture localized reporting of heat events that might otherwise be overlooked in global English-language media. Following the initial acquisition of approximately 2.67 million raw articles, a rigorous deduplication protocol was applied to eliminate syndicated content and redundant reprints. This yielded a refined corpus of 2,112,478 unique articles, serving as the foundational dataset for subsequent event extraction.

6.4 LLM-Assisted News Validation, Classification and Geo-locating

Subsequently, we implemented a multi-stage analytic pipeline utilizing Large Language Models (LLMs) to identify, validate, and structure heat events from unstructured text, designed to balance computational efficiency with semantic reasoning capabilities. Detailed prompt examples, model hyperparameter settings, and representative output samples are illustrated in Fig. 1. The multilingual screening logic, structured extraction prompts, and geocoding workflow are described in Supplementary Information, Sections S4–S6.

Step 1: Heat Event Detection (Coarse Screening). To manage the volume of the collected multilingual corpus, the initial filtration was conducted using Llama 3.1-8B-Instruct [30] to determine data relevance. As detailed in Step 2 of Fig. 1, we implemented a criteria-driven Chain-of-Thought (CoT) prompting protocol. The model synthesized an evaluative rationale based on three multidimensional indicators: explicit temperature references, documented impacts (e.g., health or economic ramifications), and specific geospatial markers. Sample retention was conditional upon the satisfaction of at least two criteria. Integral to this workflow was a semantic context disambiguation mechanism, operationalized via negative constraint prompting. By scrutinizing the semantic domain that differentiates meteorological contexts (e.g., "humidity," "residents") from abstract sociopolitical usages (e.g., "parliamentary debate," "financial market"), the model systematically filtered metaphorical homonyms (e.g., "heated discussion"), thereby enforcing strict physical relevance prior to final classification.

Step2: Validation and Extraction (Fine-Grained). Candidate articles subsequently underwent high-precision processing using DeepSeek V3 [31]. As shown in Step 3 of Fig. 1, this stage utilized a sequential translation-validation workflow. Non-English texts were first translated into English to ensure semantic standardization. Following translation, a secondary validation step was applied to minimize false

positives using the same criteria as the previous stage. Validated articles were then parsed to extract standardized occurrence times and specific location entities.

Step 3: Geographical Mapping. The final module, powered by DeepSeek V3, converted textual location descriptions into geospatial coordinates (Step 4 of Fig. 1). The system adopted a hierarchical resolution strategy: precise centroids were generated for city-level reports, while broader regions (provincial level and above) were mapped using central coordinates with an estimated radius to represent the spatial extent of the event.

6.5 Human-in-the-Loop Validation

To ensure data integrity and mitigate generative hallucinations, we integrated a rigorous Human-in-the-Loop (HITL) protocol. Expert reviewers annotated random data subsets at three critical junctures: (1) *Detection Calibration*, where human feedback was used to refine the Llama 3.1 decision boundaries, specifically training the model to distinguish physical heat events from semantic ambiguities (e.g., rejecting metaphorical "heated debates") to minimize false positives; (2) *Metadata Verification*, which audited DeepSeek V3 outputs to ensure high-fidelity translation and the accurate standardization of complex relative temporal references into ISO formats; and (3) *Geospatial Auditing*, where reviewers confirmed the precision of coordinate assignments, validating the model's hierarchical logic in distinguishing between precise city-level point coordinates and broader provincial-level regional estimates. This iterative feedback loop facilitated the dynamic optimization of prompt engineering, establishing a high-confidence alignment between automated classification and expert consensus. Calibration thresholds, correction taxonomy, and post-refinement stability analyses are documented in Supplementary Information, Section S5.3.

Socioeconomic, Institutional, and Vulnerability Covariate Data: To contextualize spatial and national patterns of heatwave under-reporting, we compiled a harmonized set of socioeconomic, institutional, linguistic, and environmental covariates from authoritative global datasets. Economic development indicators, including GDP per capita (constant 2015 USD) and internet users (% of population), were obtained from the World Bank World Development Indicators (WDI, 2023 edition). Climate vulnerability and adaptive capacity were represented using the ND-GAIN Country Index (2022 release) produced by the University of Notre Dame Global Adaptation Initiative. Climate policy and governance performance was captured using national scores from the Climate Change Performance Index (CCPI 2024; data year 2023) published by Germanwatch. Country classifications for least developed countries (LDCs) and small island developing states (SIDS) follow United Nations designations. To capture structural communication and reporting asymmetries, we incorporated a language proximity index derived from the global linguistic proximity dataset [32]. Linguistic proximity measures pairwise distances between languages based on lexical and phylogenetic similarity. For each country, the primary national language was defined as any language spoken by more than 100 million people globally, and proximity was calculated relative to this set of dominant global reporting languages. All covariates were standardized to a country-level reference frame and temporally aligned with the 2015–2024 study period prior to integration with under-reporting metrics. Covariate harmonization procedures, normalization methods, and statistical modeling specifications are detailed in Supplementary Information, Section S9.

6.6 High Resolution Global Extreme Heat Event Simulation and Backcasting

In this study, heatwaves were defined as a period of at least five consecutive days during which the daily maximum temperature exceeded the 90th percentile of the climatological baseline following the framework of [29]. Sensitivity analyses using alternative definitions ($\geq$ 2 days above the 95th percentile or $\geq$ 4 days above the 97.5th percentile) showed minimal impact on our results. Detection was implemented in MATLAB for spatially gridded datasets, with the detailed computation described below.

Step 1: Climatology and Threshold Calculation. For each grid cell, a daily climatology and percentile threshold were calculated for the baseline period (1985-2014). To account for the seasonal cycle, an

11-day sliding window centered on each calendar day (366 days) was used to pool daily maximum temperatures across all baseline years. The 90th percentile of this pooled distribution defined the threshold for heatwave detection, while the mean defined the climatology. Both climatology and percentile thresholds were further smoothed using a 31-day running mean to reduce abrupt fluctuations. It should be noted that February 29 was treated by interpolating the threshold and climatology between February 28 and March 1. This ensured a consistent day-of-year sequence across all years, allowing the sliding window and percentile calculation to function correctly without bias from leap years. Grid cells with no valid temperature data were represented as NaN, and these values were excluded from climatology and threshold calculations.

$$T_{\text{clim}}(d) = \text{mean}[T(t) \mid |\text{doy}(t) - d| \leq 5]$$
$$T_{90}(d) = \text{P90}[T(t) \mid |\text{doy}(t) - d| \leq 5]$$

Where T_{clim}, T_{90} denote the climatology and the 90th percentile threshold, $T(t)$ is the daily maximum temperature, $\text{doy}(t)$ is the day-of-year (ranging from 1 to 366), and $P90$ denotes the 90th percentile.

Step 2: Identification of Heatwave Events. Heatwave events were identified at each grid cell by comparing daily maximum temperatures against the corresponding daily threshold T90 for the baseline climatology. A heatwave was defined as a period of consecutive days where the temperature exceeded the 90th percentile threshold for at least 5 days. This duration threshold follows the canonical definition ensuring that events represent persistent, rather than transient, warm anomalies. If the interval between two events was 2 days or less, the events were combined into a single heatwave event. For each detected event, metrics were calculated: onset date, end date, duration, maximum intensity, mean intensity, cumulative intensity, and spatial location (longitude, latitude) of the event. The intensity of each event was defined as the difference between the daily maximum temperature and the climatological mean temperature (T_{clim}). Event counts were then aggregated by season, based on the start date of each heatwave.

The full simulation, backcasting, and severe heat-day derivation procedures are described in Supplementary Information, Section S10.

6.7 Sensitivity Analysis Under Different Heatwave Definitions

To evaluate the robustness of our results to the choice of heatwave definition, we conducted sensitivity experiments using alternative percentile and duration thresholds. In addition to the baseline definition ($\geq$ 5 consecutive days exceeding the 90th percentile of the climatological baseline), two additional definitions were tested: (1) events exceeding the 95th percentile for at least two consecutive days, and (2) events exceeding the 97.5th percentile for at least four consecutive days.

Across all three definitions, including both the contemporary period (2015–2024) and future projections (2050–2100), we found consistent spatial patterns in the distribution of heatwave days and event frequency. As expected, the 95th-percentile threshold with a minimum duration of two days produced the largest number and duration of heatwave events, whereas the 97.5th-percentile threshold with a minimum duration of four days yielded the fewest. Importantly, despite these quantitative differences, the spatial distribution of heatwave occurrence remained nearly identical across definitions. Since our study focuses primarily on the proportion of under-reported events across countries, these variations in threshold selection have negligible influence on the overall conclusions. Full sensitivity results, including comparative spatial maps and quantitative deviation metrics, are presented in Supplementary Information, Section S1.5 and Section S11.

6.8 Identification of Under-Reported Heatwave Regions

Under-reported heatwave regions were identified by integrating independently derived heatwave detection outputs with multilingual heat-event news records for the period 2015–2024. News data were preprocessed to construct a geocoded and temporally standardized event dataset by merging sources, extracting geographic coordinates, assigning events to seasonal periods, and removing duplicates and

invalid records. This yielded structured reported-event records with explicit locations, spatial extents, and seasonal labels. In parallel, heatwave detection results were represented as global raster datasets that record the presence of detected heatwave events for each grid cell across 40 seasonal periods (2015–2024), with consistent spatial resolution and temporal delineation.

Reported events were spatially matched to the detection rasters on a season-by-season basis, to avoid any double-counting and measurement differences between news reporting counts and heat event counts. Each reported event was mapped to its corresponding raster grid using geographic coordinate transformation, and a spatial buffer was applied to represent its reported geographic extent, with buffer width adjusted by latitude to account for meridional convergence. For each grid cell and season, the presence or absence of detected heatwaves and reported heatwave events was binarized (1 = present, 0 = absent). Under-reporting was quantified by aggregating information across all seasons and computing a reported-to-detected event ratio for each grid cell. Ratio is defined as the total number of reported events divided by the total number of recorded events, where reported events are binarized as 1 if a grid cell is covered by any reported heatwave event in a given season and 0 otherwise, and recorded events correspond to binarized heatwave detection raster values. Summation is performed across all 40 seasonal datasets. Grid cells with no recorded heatwave events across all seasons were masked as NaN to avoid spurious ratios. This procedure identifies locations where heatwaves are consistently detected but weakly or not reported. We additionally replicated the above analysis using news data from mainstream media sources (as detailed in Supplementary Information, Section S7) and disaster records from EM-DAT, and incorporated population density data from WorldPop to adjust selected analytical components.

The resulting grid-level reporting-ratio dataset, containing longitude, latitude, and aggregated under-reporting ratios for all valid grid cells, constitutes an intermediate analytical product, with access details provided in the Data Availability section. This spatially explicit dataset underpins subsequent visualization, regional aggregation, and country-level analyses of reporting disparities.

6.9 Latitudinal and Structural Disparities in Heatwave Under-Reporting

To examine the global latitudinal pattern of under-reporting of extreme heat events, we used gridded raster dataset of country-level ratios of reported to observed extreme heat events. Grid-cell centroids of the raster file are converted to geographic coordinates based on the native affine transform. All valid grid cells were then aggregated into 5° latitude bands spanning 90°S to 90°N (36 intervals in total). For each band, we calculated the mean under-reporting ratio and its standard deviation, which were subsequently expressed as percentages. Importantly, the shaded area was constructed explicitly between upper and lower bounds of the distribution, avoiding closure to the zero baseline, thereby preserving the true dispersion structure along latitude. These enable clearer visual comparison of under-reporting intensity in the Arctic, Antarctic, and equatorial zones relative to other regions. Full aggregation procedures, bin definitions, and statistical testing criteria are provided in Supplementary Information, Section S8.

To evaluate structural disparities in global under-reporting of extreme heat events, we classified countries into three sets of comparative groups. First, countries were divided into Small Island Developing States (SIDS) and Non-SIDS, following the classification adopted by Vousdoukas, M. et al[33]. Second, according to country classification provided by the United Nations Conference on Trade and Development, countries were categorized as Least Developed Countries (LDC), Developed Countries (DC), and Other Developing Countries (Other). Third, climate preparedness was defined using ND-GAIN index scores from the Notre Dame Global Adaptation Initiative: countries in the top quartile ($\geq$75th percentile) were labeled Top Quantile Climate Preparedness, while the remaining countries were classified as Low-to-Average Climate Preparedness.

For each grouping scheme, the distribution of the under-reporting ratio was compared between categories. Prior to inferential testing, we assessed normality of each group using the Kolmogorov–Smirnov test against a normal distribution and evaluated homogeneity of variances using Levene's test. If both normality and variance homogeneity assumptions were satisfied, independent-samples t-tests were applied (two-group comparisons) or pairwise t-tests (for the three development-status categories).

When assumptions were violated, we employed the non-parametric Mann–Whitney U test for two-group contrasts. For the three-category development-status comparison, all pairwise contrasts (LDC–DC, LDC–Other, DC–Other) were evaluated. Statistical significance was defined at $P < 0.05$. This hierarchical testing framework ensured methodological consistency while accommodating non-normality and heteroscedasticity in the global distribution of under-reporting ratios. Lists of least developed and climate-unprepared regions with $\leq 50\%$ extreme heat media coverage are provided in the Supplementary Information, Section S7.

6.10 Statistical Analysis of Socioeconomic and Cultural Factors Associated with Heatwave Under-reporting

To examine the socioeconomic, institutional, and environmental factors associated with cross-national variation in heatwave under-reporting, we employed statistical analysis combining interpretable machine learning with stratified descriptive analyses. The analytical sample comprises grid-level observations, with predictor variables spatially matched to each observation. The variable include the national-level reporting ratio, aggregated from grid-level under-reporting metrics, as well as climate vulnerability (ND-GAIN), economic development (GDP per capita), Internet users rate (ICT), climate policy performance (CCPI), linguistic proximity to mainstream languages, and background climatic conditions. We used an extreme gradient boosting (XGBoost) model as a flexible mapping between covariates and reporting ratios. The model was trained on a random subset of samples and evaluated on held-out data to ensure stability of associations. To interpret the relative contribution of each covariate to reporting disparities, we computed Shapley Additive Explanation (SHAP) values, which quantify the marginal contribution of each variable to deviations from the global mean reporting ratio without imposing linearity assumptions. Global feature importance was summarized using mean absolute SHAP values across all observations, providing a ranked assessment of which structural factors are most consistently associated with under-reporting. Complete model specification, covariate harmonization, and SHAP attribution procedures are detailed in Supplementary Information, Section S9.

To complement the machine learning findings and examine specific associations in greater detail, we conducted two targeted statistical approaches: grouped comparisons and linear regression analyses. These methods provide transparent, interpretable insights into how key structural factors relate to reporting ratios. For grouped comparisons, countries were first grouped by continent, and then we selected two variables—CCPI and ICT—and classified countries within each continent into high and low categories using the upper and lower terciles of the variable distribution. Boxplots were generated to compare reporting ratios between groups, with differences tested using either a t-test or Mann–Whitney U test based on normality assessments (K–S test). Continent-level analyses were restricted to Africa, Asia, and Europe to avoid spurious inference from small samples. For linear regression analyses, we examined the association between language proximity and reporting ratios specifically for African countries, given the continent's linguistic diversity. Nations were categorized into five linguistic groups based on their mainstream language: English, French, Portuguese, Arabic, and Others, and a simple linear regression was fitted to quantify the relationship across the entire African subset. A separate analysis was conducted exclusively for countries within the Arabic language group to test whether the relationship held for a non-globally dominant language community. Group-specific mean reporting ratios were calculated and highlighted in scatter plots.

6.11 Stratified and Regional Association Analyses for 2050 global extreme heat event loss

To assess how under-reporting of extreme heat may bias projections of future heatwave-related losses, we draw on global impact projections to 2050 from Sun et al., which estimate health losses, labour productivity losses and indirect losses under multiple future scenarios, with total economic losses defined as their aggregate. Rather than generating new impact estimates, we reorganize countries according to the degree of extreme heat under-reporting identified in this study and re-express the projected damages within this reporting-based framework, with data access details provided in the Data Availability section. Although the heatwave definition in Sun et al. (2025) differs slightly from ours,

sensitivity analyses indicate highly consistent spatial patterns of heatwave duration across definitions, suggesting that projected loss patterns are robust to metric choice. We then compare projected losses between severely under-reporting and well-reporting countries to quantify how systematic reporting gaps may distort the magnitude and geographic distribution of future heatwave-related economic and health risks. Detailed percentile stratification rules, country grouping logic, and scenario comparison procedures are provided in Supplementary Information, Section S10.

Future losses were organized and categorized under the SSP585 scenario for 2050 using two impact indicators: health loss (%) and economic loss (%). Country-level projections were first linked to the mean reporting ratio of extreme heat events. Countries were then ranked by reporting performance and classified into percentile-based groups. The bottom decile of the reporting ratio ($\leq$10th percentile) was defined as the Top 10% under-reported, representing countries with the most severe reporting deficits, whereas All countries includes the full country sample irrespective of reporting level. For continental comparison, we additionally highlight selected countries within each continent that exhibit particularly high projected losses. And we report mean projected losses for both All countries and the Top 10% under-reported subset, allowing assessment of how severe reporting gaps reshape regional loss profiles. At the global scale (World), three groups are shown: Well-reported ($\geq$90th percentile of reporting ratio), Top 10% under-reported ($\leq$10th percentile), and All countries (entire sample). This percentile-based stratification isolates the upper and lower tails of reporting performance and quantifies how systematic under-reporting may bias projected health and economic losses.

Total economic losses (%) under SSP585 in 2050 were extracted at specially defined country-group levels. Countries were ranked by the mean reporting ratio of extreme heat events, and the 10th percentile threshold was used to define the Top 10% under-reported (UR) group. Five analytical groups were constructed: Non-UR (reporting ratio >10th percentile), All countries (AC), UR ($\leq$10th percentile), UR + CLP (climate least prepared), and UR + CLP + LDC (least developed countries). Group means and standard deviations were calculated for economic loss (%) and visualized using bar charts with error bars, horizontal reference lines denote AC and UR averages to facilitate comparison. Temporal dynamics were further examined from 2020 to 2060 under SSP119, SSP245 and SSP585. For each scenario, economic loss (%) was plotted for All countries and the Top 10% under-reported countries, enabling assessment of how reporting deficits modify loss trajectories across scenarios.

Data availability

All meteorological reanalysis data used in this study are publicly available. ERA5 hourly 2-m air temperature data (1985–2024) were obtained from the Copernicus Climate Data Store (CDS; https://cds.climate.copernicus.eu/). CMIP6 CESM2 model outputs under the SSP5-8.5 scenario were downloaded from the Earth System Grid Federation (ESGF; https://esgf-node.llnl.gov/).

Country-level socioeconomic indicators, including GDP per capita and Internet usage rates, were obtained from the World Bank World Development Indicators (https://data.worldbank.org/). Climate vulnerability and preparedness indices were sourced from the ND-GAIN Country Index (https://gain.nd.edu/). Climate policy performance scores were derived from the Climate Change Performance Index (https://ccpi.org/). Country classifications for Least Developed Countries (LDCs) and Developed Countries were obtained from the United Nations Conference on Trade and Development (https://unctadstat.unctad.org/). Small Island Developing States (SIDS) classifications follow the United Nations designation framework. Population projections were obtained from the United Nations World Population Prospects (https://population.un.org/wpp/). Baseline mortality statistics were obtained from the World Health Organization (https://www.who.int/). Population density data were sourced from WorldPop (https://www.worldpop.org/). Disaster database records were obtained from EM-DAT (https://www.emdat.be/). Country-level projections of future heat-related economic and health losses were obtained from the publicly released dataset of Sun et al. (2024).

The multilingual news corpus was constructed from 239 publicly accessible national news websites spanning 184 countries and 40 languages during 2015–2024. Due to copyright and licensing restrictions, full-text news articles cannot be redistributed. However, structured event-level metadata generated in

this study—including publication date, outlet domain, language label, geocoded coordinates, administrative resolution level, seasonal assignment, and reporting-ratio metrics—are available in the project repository (https://github.com/OTTO-OTO/ExtremeHeat) and will be archived in a permanent DOI-based repository upon publication.

The final validated geocoded extreme heat event dataset comprises 210,525 structured heat-event records. Grid-level reporting-ratio rasters, latitude-band aggregates, country-level reporting ratios, and model-ready analytical tables underlying all figures are available in the project repository.

Code availability

All code used for data preprocessing, heatwave detection, climatology construction, raster aggregation, spatial matching, reporting-ratio computation, statistical modeling, robustness testing, SHAP attribution analysis, and visualization is publicly available at: https://github.com/OTTO-OTO/ExtremeHeat. The repository includes MATLAB (R2023a), Python (version 3.10), and R(version 4.5.2) scripts required to reproduce the analytical workflow. Key Python packages include NumPy, SciPy, pandas, rasterio, scikit-learn, statsmodels, XGBoost, and SHAP.

Multilingual screening and structured extraction were conducted using publicly available large language models (Meta Llama 3.1-8B-Instruct and DeepSeek-V3) under fixed deterministic inference settings, as described in the Supplementary Information (Sections S4–S6). No fine-tuning or parameter adaptation of the underlying language models was performed. All stochastic processes were executed using fixed random seeds where applicable to ensure computational reproducibility.

All intermediate processing stages—including screened article identifiers, structured extraction outputs, geocoding logs, reporting-ratio calculations, model training artifacts, and robustness-testing outputs—are archived within the repository and can be independently audited.

AI-assisted writing statement

Portions of the manuscript text were refined for clarity and language using large language models. All scientific content, analyses, interpretations, and conclusions were conceived, conducted, and verified by the authors. The authors take full responsibility for the integrity and originality of the work.

Author contributions

S.L., H.P., and E.L. jointly conceived the study, designed the interdisciplinary framework, and supervised the overall research direction. S.L., Y.Z., H.H., and H.Z. developed the artificial intelligence pipeline, including multilingual data aggregation, large-language-model-assisted screening, and geospatial extraction. E.L. and C.S. designed and implemented the physical climate simulation and backcasting analyses. H.P., Y.Y., and S.Y. conducted the loss modeling, indicator construction, and socioeconomic vulnerability analyses. T.S., J.M., Z.K., F.W., and Z.M. provided regional expertise, country-level contextual validation, and cross-national interpretation of findings. S.L., H.P., and E.L. drafted the manuscript. All authors contributed to interpretation, critical revision, and approved the final manuscript.

Competing Interests

The authors declare no competing interests.

Supplementary information

Supplementary Methods, Extended Data Tables, robustness analyses, and additional figures accompany this paper.

References

[1] Zhang, B. *et al.* Anthropogenic climate change doubled the frequency of compound drought and heatwaves in low-income regions. *Commun. Earth Environ.* **5**, 715 (2024).

[2] Bellprat, O. *et al.* Towards reliable extreme weather and climate event attribution. *Nat. Commun.* **10**, 1732 (2019).

[3] Hase, V. Climate injustice through unequal news. *Nat. Clim. Change* **15**, 129–130 (2025).

[4] Hase, V. *et al.* Climate change in news media across the globe: An automated analysis of issue attention and themes in climate change coverage in 10 countries (2006–2018). *Glob. Environ. Change* **70**, 102353 (2021).

[5] Dolšak, N. & Prakash, A. Three faces of climate justice. *Annu. Rev. Polit. Sci.* **25**, 283–301 (2022).

[6] Vu, H. T. *et al.* Nationalizing a global phenomenon: A study of how the press in 45 countries and territories portrays climate change. *Glob. Environ. Change* **58**, 101942 (2019).

[7] UNDRR & WMO. Global status of multi-hazard early warning systems: Target G. Tech. Rep., United Nations Office for Disaster Risk Reduction and World Meteorological Organization (2022).

[8] Rentschler, J. *et al.* Stronger resilience through better risk data. Policy Research Working Paper, World Bank (2023). Available at: https://openknowledge.worldbank.org/handle/10986/39470.

[9] Kelder, T. *et al.* How to stop being surprised by unprecedented weather. *Nat. Commun.* **16**, 2382 (2025).

[10] Harrington, L. J. & Otto, F. E. Reconciling theory with the reality of African heatwaves. *Nat. Clim. Change* **10**, 796–798 (2020).

[11] Di Napoli, C., Allen, T., Méndez-Lázaro, P. A. & Pappenberger, F. Heat stress in the caribbean: climatology, drivers, and trends of human biometeorology indices. *International Journal of Climatology* **43**, 405 (2022).

[12] Thompson, V. *et al.* The most at-risk regions in the world for high-impact heatwaves. *Nat. Commun.* **14**, 2152 (2023).

[13] Lamptey, B. *et al.* Challenges and ways forward for sustainable weather and climate services in africa. *Nat. Commun.* **15**, 2664 (2024).

[14] Thaker, J. *et al.* English and regional media coverage of the 2022 heatwave in india. *Environ. Commun.* (2024).

[15] Rahman, S. M. *et al.* Generative Artificial Intelligence for Environmental Assessment: A New Paradigm for Sustainability Analysis. *Environ. Manage.* **76**, 93 (2026).

[16] Camps-Valls, G. *et al.* Artificial intelligence for modeling and understanding extreme weather and climate events. *Nat. Commun.* **16**, 1919 (2025).

[17] Schiller, J. *et al.* Artificial intelligence in environmental and Earth system sciences: explainability and trustworthiness. *Artif. Intell. Rev.* **58**, 316 (2025).

[18] Epstein *et al.* *The Language of African Literature* (Africa World Press, Trenton, 1998).

[19] Eberhard *et al.* Ethnologue: Languages of the world (2025). Online version: https://www.ethnologue.com/country/NG/ (retrieved February 3, 2026).

[20] Sun, Y. *et al.* Global supply chains amplify economic costs of future extreme heat risk. *Nature* **627**, 797–804 (2024).

[21] World Meteorological Organization (WMO). State of the climate in Africa 2022. Tech. Rep. WMO-No. 1330, World Meteorological Organization (2023). Available at: https://library.wmo.int/idurl/4/57694.

[22] Nangombe, S. *et al.* Record-breaking climate extremes in Africa under stabilized 1.5 °C and 2 °C global warming scenarios. *Nat. Clim. Change* **8**, 375–380 (2018).

[23] Lange, S. *et al.* Projecting exposure to extreme climate impact events across six event categories and three spatial scales. *Earth's Future* **8**, e2020EF001616 (2020).

[24] Lavelle, M. Congressional democrats assail Trump's plans for deep cuts to government science. *Inside Climate News.* (2025, April 14).

[25] Schewe, J. *et al.* State-of-the-art global models underestimate impacts from climate extremes. *Nat. Commun.* **10**, 1005 (2019).

[26] Newman, R. & Noy, I. The global costs of extreme weather that are attributable to climate change. *Nat. Commun.* **14**, 6103 (2023).

[27] Danabasoglu, G. *et al.* The community earth system model version 2 (CESM2). *J. Adv. Model. Earth Syst.* **12**, e2019MS001916 (2020).

[28] Simpson, I. R. *et al.* An Evaluation of the Large-Scale Atmospheric Circulation and Its Variability in CESM2 and Other CMIP Models. *J. Geophys. Res.:Atmos.* **125**, e2020JD032835 (2020).

[29] Hobday, A. J. *et al.* A hierarchical approach to defining marine heatwaves. *Prog. Oceanogr.* **141**, 227–238 (2016).

[30] Grattafiori *et al.* The llama 3 herd of models (2024). Preprint at https://arxiv.org/abs/2407.21783.

[31] Liu *et al.* Deepseek-v3 technical report (2024). Preprint at https://arxiv.org/abs/2412.19437.

[32] Gurevich, T. *et al.* A dataset on linguistic connectivity across and within countries. *Sci. Data* **12**, 542 (2025).

[33] Vousdoukas, M. *et al.* Small island developing states under threat by rising seas even in a 1.5 °C warming world. *Nat. Sustain.* **6**, 1552–1564 (2023).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ReportingSummary.pdf
- ExtremeLLMSM260321.pdf