ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Unknown risk: assessing refugee camp flood risk in Ethiopia

To cite this article: Mark V Bernhofen *et al* 2023 *Environ. Res. Lett.* **18** 064041

ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
12 December 2022

REVISED
24 April 2023

ACCEPTED FOR PUBLICATION
25 May 2023

PUBLISHED
6 June 2023

LETTER

Unknown risk: assessing refugee camp flood risk in Ethiopia

Mark V Bernhofen[1,2,*], Faye Blenkin[2] and Mark A Trigg[2]

[1] Smith School of Enterprise and the Environment, University of Oxford, Oxford, United Kingdom
[2] School of Civil Engineering, University of Leeds, Leeds, United Kingdom
[*] Author to whom any correspondence should be addressed.

E-mail: mark.bernhofen@smithschool.ox.ac.uk

Keywords: flood risk, refugees, camps, global flood models

Supplementary material for this article is available online

Abstract

The number of global refugees has been rising annually for the last decade. Many of these refugees are housed within camps, in temporary structures, vulnerable to the impacts of flooding. The flood risk of refugees is not well understood. Flood risk guidance available for camp planners and managers is vague, and existing flood risk data is often lacking in the remote areas where camps are typically located. We show how global data should, and should not, be used to assess refugee flood risk in Ethiopia; a country hosting 725 000 refugees, primarily from four neighboring countries, in 24 camps. We find that global population (GP) datasets, typically used in national flood risk assessments, do not accurately capture camp populations (CPs). Even the most accurate GP datasets are missing three fifths of camp flood exposure. We propose, and test, alternative approaches for representing exposure that combine reported estimates of CP with data on camp area, building footprints, and population density. Applying these approaches in our national flood risk assessment, we find that 95.8% of camps in Ethiopia are exposed to flooding of some degree and between 143 208 (19.8%) and 182 125 refugees (25.2%) are exposed to a 1% annual exceedance probability flood (100 year return period). South Sudanese refugees are the nationality most exposed to flooding, but Eritrean refugees are the nationality most exposed to flooding with a high risk to life. Promisingly, we find that many camps may be set up in such a way that reduces the exposure of refugees to flooding. Our study demonstrates that global data, augmented with local data, can be useful for understanding the flood risk of refugee camps. The consistent scalable approach can be used as a first-order analysis of risk, identifying risk hotspots, and help to prioritize further detailed analyses to inform within-camp adaptation.

1. Introduction

At the end of 2021, there were over 27 million refugees globally, the majority escaping violent conflict or persecution in their home countries [1]. Climate change is expected to further exacerbate future migration, as climate impacts make regions less hospitable [2] and increase the risk of violent conflict [3–5]. Refugee camps house over 24% (6.6 million) of refugees globally [6]. These camps are susceptible to extreme climate events such as flooding due to their remote location, temporary shelters, and densely populated inhabitants. There is growing awareness that refugees need to be incorporated into national flood risk planning and response [7].

Published guidance and standards provide little direction to camp planners and managers about how to assess flood risk [8–10]. To compound this, there is often a lack of existing flood risk data because public land made available for refugee camps by the government is typically far away from urban centers—where flood risk information is more prevalent. This lack of guidance and data, coupled with the costs of commissioning a local flood risk assessment, means that often such assessments are not carried out and the risk to camps and inhabitants is unknown.

Global flood risk data, such as global flood model (GFM) hazard maps and global population (GP) maps, could represent a low-cost and easy-to-implement solution for understanding refugee camp flood risk [11]. These datasets are becoming increasingly relevant at the national scale, especially in regions lacking existing flood risk data [12]. However, there is still significant disagreement between these global datasets in rural areas [13, 14], where most refugee camps are located. Previous studies have shown that GP maps underrepresent populations in refugee camps in Uganda [15] and informal settlements in Kenya [16]. However, the utility of global data in the context of refugee flood risk has not yet been examined.

We explore different approaches for using global data in the assessment of refugee flood risk in Ethiopia, a country with a well-established history of hosting refugees as a result of its progressive open-door policy and working rights for refugees [17]. Today, Ethiopia hosts over 870 000 refugees, over 98% of which originate from four neighboring countries: Sudan, South Sudan, Eritrea, and Somalia [18]. Over 90% (725 000) of these refugees are housed in 24 camps throughout the country. Camp populations (CPs) are dynamic—compounded by the ongoing civil war in northern Ethiopia, which has already resulted in the destruction of two refugee camps in the region and subsequent relocation of refugees [17]. Camps in Ethiopia have a history of flooding. In 2014, heavy rainfall led to floods in the Gambela region, which inundated two camps hosting South Sudanese refugees and led to the eventual abandonment of the camps [19]. This problem is not going away. Refugee numbers have risen every year over the past decade [1] and are projected to continue increasing in the near future [20]. Concurrently, a changing climate will bring more frequent and intense flood events for most of Africa [21]. Coupled increases in exposure and hazard frequency have the potential to exacerbate the future flood impacts felt by refugees [22]. However, to date, the data available for understanding refugee flood risk has been limited, hampering adaptation actions by camp planners and managers. Further, the degree to which camps within a country may be impacted represents an unknown risk as they are typically missed from top-down national assessments [12] or assessed on a site-by-site basis [23, 24].

In this paper, we map the location and detail of all current refugee camps in Ethiopia and carry out a building-level flood risk assessment in each camp using a high-resolution GFM. In doing so, we critically analyze the usefulness and limitations of global data used at this scale and propose a methodology combining global and local data for refugee camp flood risk assessments. The results of our analysis are intended to inform camp planners, managers, and decisions makers and provide a consistent view of the flood risk faced by refugees at the camp, regional, and national level.

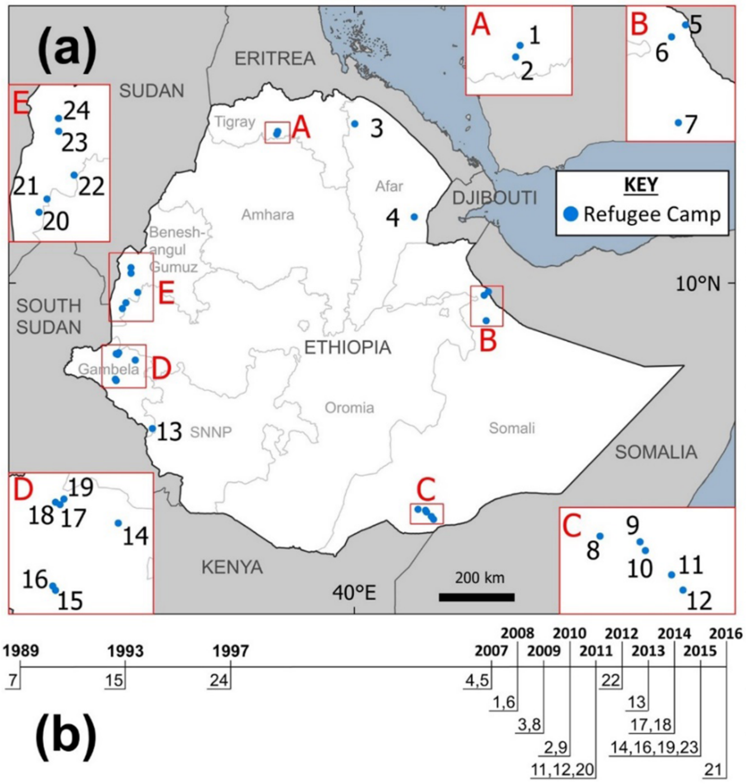
2. Methods

2.1. Mapping refugee camps

Ethiopian refugee camps were mapped using a combination of country-level information from UNHCR, remote sensing data, OpenStreetMap data, and building footprint data. We identified 24 active refugee camps in Ethiopia (see figure 1). Using data on camp locations provided by UNHCR [25], we manually mapped camp boundaries using remote sensing data and populated the camps with building footprints from OpenStreetMap (accessed June 2022) [26] and Google Open Buildings (v1 2021 vintage) [27, 28]. Building footprint data was available in 84% of camps. CPs were taken from the latest UNHCR camp reports. A detailed description of the camp mapping methodology, as well as maps of the 24 camps, can be found in the supplementary material.

2.2. GP data assessment

We assessed the accuracy of six widely used GP datasets at capturing refugee CPs in Ethiopia. The GP datasets assessed include: the Global Human Settlement Population 2019 revision (GHS-POP R2019, 2015 population, 270 m resolution) [29, 30] and 2022 revision (GHS-POP R2022, 2020 population, 100 m resolution) [31, 32] the Facebook High Resolution Settlement Layer (HRSL, 2020 population, 30 m resolution) [33], LandScan (2019 population, 900 m resolution), WorldPop unconstrained (2020 UN adjusted population, 90 m resolution) [34], and WorldPop constrained (2020 UN adjusted population, 90 m resolution) [35]. The GP datasets differ in their temporal resolution by up to five years. For each dataset, we choose its most recent vintage for our analysis, as would typically be done in assessments of present-day flood risk. Global dataset based CP estimates were calculated by summing the population cells that fell within the camp boundaries. These global dataset population estimates were then compared with UNHCR reported CPs. The within camp accuracy of the population datasets was calculated using the mean absolute percentage error (MAPE) metric and the cross-camp accuracy was calculated using the mean error (ME) and mean absolute error (MAE) metrics. The relationship between the within-camp accuracy of GP data and camp characteristics was calculated using two bivariate correlation statistics: Pearson's correlation (r) and Spearman's rank correlation (ρ). Camp characteristics considered include: camp area, camp age, CP, and camp distance from nearest urban center [36]. For both correlation statistics, the direction of the relationship is indicated by the sign of the correlation statistic and the strength of the correlation indicated by the absolute value of the statistic, where 1 is perfect correlation and 0 is no correlation. To test the statistical significance of



ID	NAME	REGION	ORIGIN
1	Mai Aini	Tigray	Eritrea
2	Adi Harush	Tigray	Eritrea
3	Barahle	Afar	Eritrea
4	Aysaita	Afar	Eritrea
5	Aw-Barre	Somali	Somalia
6	Sheder	Somali	Somalia
7	Kebribeyah	Somali	Somalia
8	Bokolmanyo	Somali	Somalia
9	Melkadida	Somali	Somalia
10	Kobe	Somali	Somalia
11	Helaweyn	Somali	Somalia
12	Buramino	Somali	Somalia
13	Okugo	Gambela	South Sudan
14	Jewi	Gambela	South Sudan
15	Pugnido	Gambela	South Sudan
16	Pugnido 2	Gambela	South Sudan
17	Tierkidi	Gambela	South Sudan
18	Kule	Gambela	South Sudan
19	Nguenyiel	Gambela	South Sudan
20	Tongo	Benishangul-Gumuz	Sudan and South Sudan
21	Gure-Shombola	Benishangul-Gumuz	South Sudan
22	Bambasi	Benishangul-Gumuz	Sudan
23	Tsore	Benishangul-Gumuz	Sudan and South Sudan
24	Sherkole	Benishangul-Gumuz	Sudan and South Sudan

Figure 1. (a) Location of refugee camps in Ethiopia. (b) Timeline of when each camp was established.

the correlation, we calculate each metrics' respective p values.

2.3. Flood exposure assessment

We used global flood hazard maps from Fathom (www.fathom.global) because of their representation of small rivers and streams. The Fathom model explicitly models flooding on rivers with upstream drainage areas as small as 50 km^2 using the 2D inertial wave formulation of the shallow water equations [37, 38]. The inundation of even smaller streams is captured implicitly in Fathom's pluvial rain-on-grid model, as flooding in headwater catchments is typically flashy [39]. The representation of small rivers was important, as a preliminary analysis (detailed in the supplementary material) found that the majority of refugee camps in Ethiopia are exposed to flooding from small rivers with upstream drainage areas less than 500 km^2 [14, 40]. We used Fathom Global 2.0 maps, which improve on the Fathom Global 1.0 maps [38] through the incorporation of higher accuracy terrain [41] and hydrography [42] data. In our analysis, we combined the Fathom Global 2.0 fluvial and pluvial flood maps at 3 arc second resolution (~90 m at the equator) to ensure that flooding on all sizes of river channels is captured.

GFMs enable flood risk to be assessed in data-scarce regions, such as Ethiopia. However, using these models in such regions also poses a validation problem, as there is typically limited data to test the reliability of the models. Previous model intercomparison work found large disagreement between available GFMs in Africa [43]. Despite this, comparative GFM evaluation has shown that the Fathom model shows minimal bias and performs well when validated against historical flood events in Nigeria and Mozambique [44]. We perform an additional validation (outlined in detail in the supplementary material) of the Fathom model using observational data from two recent flood events in Ethiopia and find that the model's performance is in line with previous evaluations of large-scale flood models [45].

Flood exposure was calculated for each camp in four ways (illustrated in figure 2) to explore what data inputs are necessary to accurately capture refugee camp flood exposure. In the GP approach (figure 2(a)), exposure was calculated by intersecting a GP map with a global flood map. This top-down approach is representative of how exposure is typically calculated in global [46, 47] and national scale [12, 13] assessments. In the CP approach (figure 2(b)), the reported UNHCR CP was distributed evenly across the camp area. Exposure was calculated by multiplying the proportional area of inundation within the camp boundary by the reported UNHCR CP. In the CP and global population density (CP + GPD) approach, the UNHCR reported

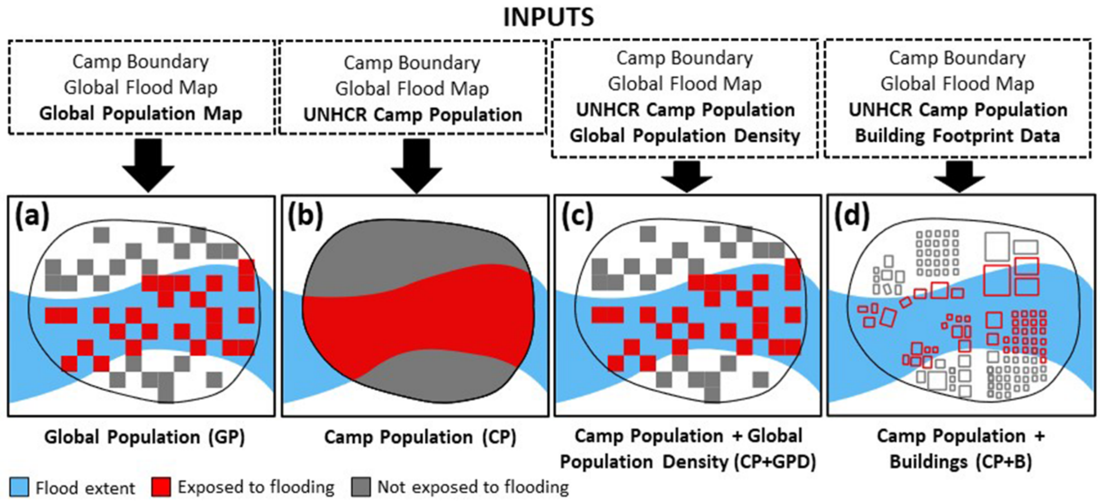


Figure 2. Illustration of three methods used for calculating camp flood exposure. (a) Global Population (GP) approach, which intersects a global flood map with a GP map and calculates the exposure within the camp boundary. (b) Camp Population (CP) approach, where UNHCR reported CP is distributed evenly within the camp boundary and flood exposure is calculated as the proportion of inundation within the camp boundary. (c) Camp Population + Global Population Density (CP + GPD) approach, where UNHCR reported CP is distributed across the camp area weighted by the relative population density of the GP map. (d) Camp Population + Buildings (CP + B) approach, where UNHCR reported CP is distributed evenly across building footprints within the camp boundary and flood exposure is calculated as the proportional number of buildings that are inundated. Datasets in bold are the datasets that change in each approach.

CP was distributed across the camp area, weighted by the relative population density of the GP map. In the final camp population + buildings (CP + B) approach (figure 2(c)), reported UNHCR CPs were evenly distributed across building footprint data (OpenStreetMap and Google Open Buildings, where available), irrespective of the size of the building footprint. Exposure was then calculated by intersecting the global flood maps with the building footprint data. We chose to classify a building as exposed if the flood map touched any part of the building footprint. A preliminary comparison between this approach and an approach which considers exposure only to building centroids found that the footprint approach overpredicted exposure to a 100 year flood relative to the centroid approach by 2.5% and 3.2% for the OpenStreetMap and Google Open Buildings data, respectively. Although the building footprint approach overestimates exposure relative to the building centroid approach, we argue that even partial inundation of a building can be significant and thus use the footprint approach in our analysis. Refugee camp flood exposure was calculated for ten flood return periods (from the 5 year to the 1000 year flood). Average annual exposure (AAE) was calculated by solving the Riemann integral of the exposure probability curve, assuming no exposure occurs outside the return period bounds.

2.4. Flood risk assessment

To calculate risk, we introduced simple hazard vulnerability thresholds into our analysis. Hazard vulnerability thresholds are commonly used to communicate the associated risks of flooding. Flood hazard intensities (typically depths and velocities) are assigned different thresholds corresponding to different levels of impact. These thresholds will typically be location and context specific. For example, in the UK the Environment Agency assigns flood depth thresholds for different levels of property damage [48] and in Australia different combined depth and velocity thresholds are reported for risk to children, adults, and motor vehicles [49, 50]. To our knowledge, no specific hazard vulnerability thresholds exist for refugee camps. As such, we introduced four simple flood depth thresholds based off existing approaches but tailored to the context of refugee camps (see table 1). Depth ranges were assigned four risk categories (low, medium, high, and very high) according to the immediate risk to life.

While we recognize that flood flow velocity is an important indicator of vulnerability, particularly when assessing risk to human life [50], we note that global flood modeling products do not currently make extreme flow velocities available. Due to the coarse nature of the global flood modeling computations, depth averaged velocities are not considered indicative of local flood hazard velocities [11]. This limits risk assessments with global model outputs to flood extent and depth.

We carried out the flood risk assessment for three different return periods of flooding, each representing a different likelihood of occurrence. The 10 year flood represents a frequent flood event, the 100 year flood represents an infrequent flood event, and the 1000 year flood represents an extremely infrequent event. In the flood risk analysis, we used only one exposure dataset/approach per camp. In camps where

Table 1. Flood hazard vulnerability thresholds.

Risk category	Flood depth (m)	Description
Low	<0.15	Low immediate risk to life. Stagnant water at this depth could pose longer term health risks. Low-cost flood adaptation strategies could be effective.
Medium	0.15–0.5	Some risk to life for vulnerable groups (e.g. children, elderly, and disabled). Structural flood adaptation strategies could be effective.
High	0.5–1.5	Large immediate risk to life for all groups. Significant damage to structures within camp.
Very High	>1.5	Substantial immediate risk to life. Structures within camp destroyed.

building footprint data from multiple sources is available, we used the dataset with the most footprints (under the crude assumption it would be the most complete); if only one footprint dataset was available, we used that; if no footprint data was available, we used the CP + GPD approach (outlined in figure 2(c)).

3. Results and discussion

3.1. Accuracy of GP data

GP datasets do a poor job at estimating refugee CPs in Ethiopia. In the 24 camps examined, GP datasets estimated, on average, only 27% of the refugee CP reported by UNHCR. The total refugee CP in Ethiopia (725 033) was underestimated by between 407 979 (56.2%) and 712 050 (98.2%), depending on the GP map used. There were only six instances (out of 144 camp and GP data combinations) where a GP dataset estimated a larger CP than UNHCR (see figure 3(a)). In general, error within the camps was high (figure 3(c)); 13 out of 24 camps had MAPE scores greater than 0.8. In exploring the relationships between camp characteristics and camp accuracy, we found no statistically significant relationship between the accuracy of the population data within the camps and the age of the camp, the area of the camp, or the distance of the camp to the nearest urban area. The only statistically significant relationship was a moderate positive relationship between camp MAPE and the population of the camp ($r = 0.46$; $\rho = 0.47$), suggesting that the population maps are less able to capture highly populated camps accurately.

There were significant differences in the accuracy of GP datasets. Four datasets performed particularly poorly: GHS-POP R2019, LandScan, WorldPop Unconstrained, and WorldPop Constrained. The 2015 vintage of GHS-POP R2019 makes it particularly unsuitable for studies such as this. Recently established camps such as Gure-Shombola (established in 2016) will not be captured in datasets produced before this time. Temporal accuracy is not the only issue affecting the GHS-POP R2019 dataset, which predicts no population in 19 out of 24 camps. This is because GHS-POP R2019 was primarily developed to map built-up areas [30] and its approach, which relies on distributing census data across satellite derived settlements, has been shown in previous studies to underestimate rural populations [51, 52]. WorldPop Unconstrained, on the other hand, applies a population distribution approach that is intended to better capture rural populations. The approach involves distributing 'unassigned' census data across vast areas of 'unpopulated' land under the rationale that communities not captured in the underlying satellite data will be implicitly represented in the population map. In the context of refugee camps, this approach is of limited value as camps tend to be small and densely populated and are thus underrepresented in the unconstrained approach. This is evidenced by the fact WorldPop Unconstrained predicted the lowest total CP of all the datasets. WorldPop Constrained applies a population distribution approach more in line with GHS-POP R2019, where census data is distributed across satellite identified settlements. Despite being slightly more accurate than WorldPop Unconstrained (figure 3(b)), error for this dataset remains high, which points to issues in the underlying census or settlement data. LandScan is the second worst performing population dataset. Its coarse resolution (~1 km) may contribute to its high error scores, as the area of individual refugee camps in Ethiopia is generally small (median camp area is 2.9 km^2) and LandScan population data will be insufficiently granular to capture camps of this size accurately.

The two most accurate GP maps were HRSL and GHS-POP R2022. Previous work found that HRSL performed favorably when compared with other human settlement data in Ugandan refugee camps [15]. The authors suggest that higher resolution satellite imagery, from which these datasets are derived, leads to greater accuracy. Our results support this hypothesis. Both HRSL and GHS-POP R2022 are derived from high-resolution satellite imagery: 0.5 m and 10 m resolution, respectively. Shelters within the camps, which can be as small as 4 m across [53], are better captured in the population maps derived from high resolution satellite data. The importance of the resolution of the underlying satellite data is highlighted when comparing the accuracy of GHS-POP R2022 and GHS-POP R2019. One of the improvements of GHS-POP R2022 from GHS-POP R2019 is the incorporation of higher resolution Sentinel-2 imagery (10 m) into the settlement classification; previously, only Landsat imagery (30 m) was used [54].

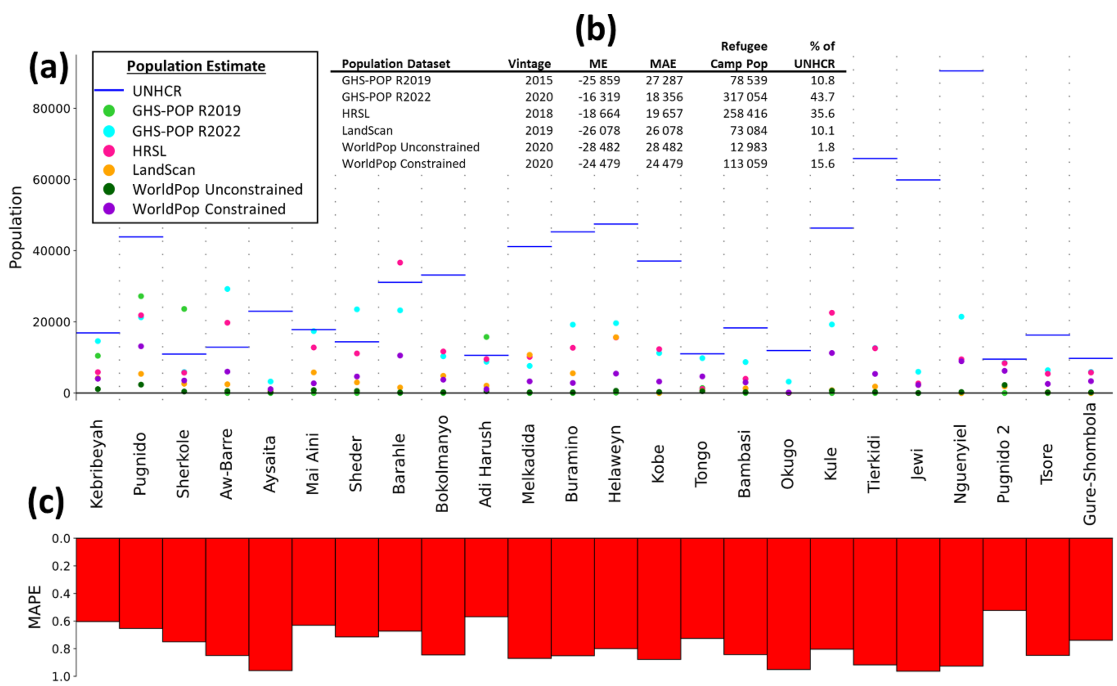


Figure 3. Accuracy of global population datasets at replicating UNHCR reported camp populations. (a) Dot plot of refugee camp population estimates. Coloured dots represent global population dataset population estimates. Horizontal blue lines represent the latest UNHCR reported camp populations. (b) Mean error (ME) and mean absolute error (MAE) of each global population dataset at replicating UNHCR reported camp populations across all 24 camps. Also reported are each global population dataset's total national camp refugee estimates expressed as an absolute value and as a percentage of UNHCR total national camp population (725 033). (c) Bar chart of the mean absolute percentage error (MAPE) of the global population data at replicating UNHCR camp population total for each camp. NOTE: Camps in figure arranged based on when they were established, where left is older, and right is more recent.

As a result, MAE was reduced by 8931 (figure 3(b)). Additional dataset improvements will also have contributed to this reduction in error, such as the inclusion of 2020 population estimates and updates to the population disaggregation approach.

Our results indicate a chronic underestimation of refugee CPs in Ethiopia by existing GP datasets. We have identified a number of factors that contribute to this underestimation, including: the temporal accuracy of the datasets, the resolution of the datasets and their inputs, and their approach to population disaggregation. The key limitation of these datasets, however, is the underlying census data they use. Refugees are typically not included in national censuses [55]. Rapidly changing refugee flows during crises also means any national population survey that included refugee estimates would quickly become outdated. Even the most accurate GP map estimates less than half of the actual Ethiopian refugee CP total. Given these findings, we would caution against using GP maps 'as is' in any study of refugee flood risk.

3.2. Refugee camp flood exposure

In Ethiopia, 23 out of 24 refugee camps (95.8%) are exposed to flooding of some degree. The five most exposed camps in Ethiopia (calculated using the CP approach) are Tierkidi, Nguenyiel, Barahle, Jewi, and Pugnido (see figure 4(a)). These camps have estimates of average annual population exposed to flooding of 4736 (7.2% of CP), 2313 (2.6% of CP), 1784 (5.7% of CP), 1441 (2.4% of CP), and 1280 (2.9% CP), respectively. Four of these five most exposed camps are located in the Gambela region, which is one of the most flood-prone regions in the country [56]. Refugee flood exposure in Gambela is more than five times greater than any of the other four regions in Ethiopia hosting refugees. The camps in Gambela are home to South Sudanese refugees, who are disproportionately more exposed to flooding than other refugees in Ethiopia.

The approach we used to calculate flood exposure had a significant impact on our flood exposure estimates. For the 100 year return period flood, the GP approach (calculated using HRSL) estimated 57 330 refugees exposed (7.9%). Compared to approaches that incorporate UNHCR CP estimates, the GP approach significantly underestimates exposure. The CP approach (which assumes UNHCR population is spread evenly across camp area) estimated 182 125 refugees exposed (25.1%) to the 100 year flood and the CP + GPD approach (which uses HRSL population density to weight population distribution across the camp) estimated 143 208 refugees exposed (19.8%) to the 100 year flood. These results show that even the most accurate GP maps underpredict between three-fifths and two-thirds of flood

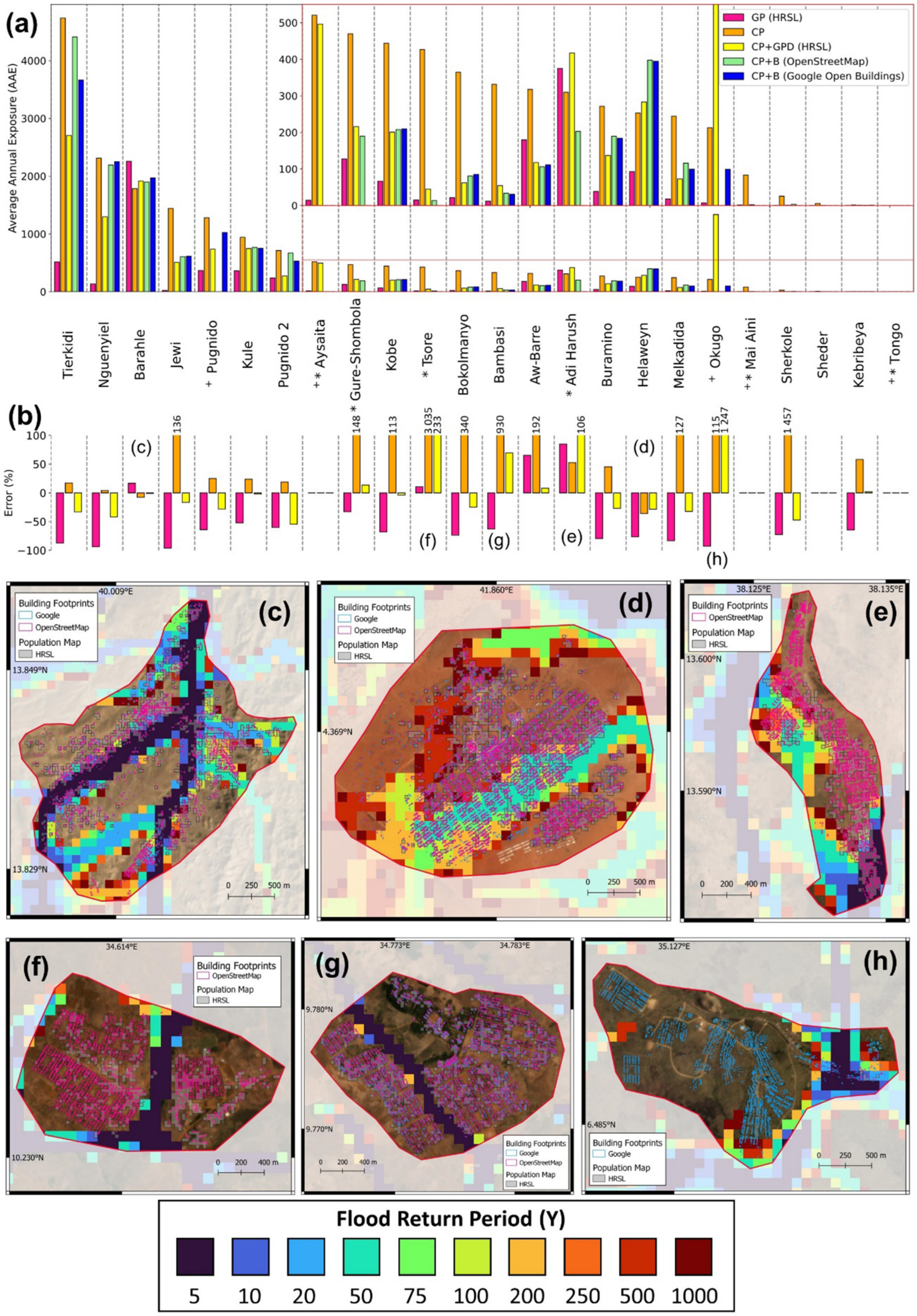


Figure 4. (a) Average annual population exposure to flooding in each refugee camp. Coloured bars represent the dataset/approach used to calculate exposure. Where, GP is Global Population, CP is Camp Population, CP + GPD is Camp Population and Global Population Density, and CP + B is Camp Population and Building Footprints. (b) Percentage error of GP, CP and CP + GPD approaches relative to the CP+B approach. (c)–(h) Maps of selected camps with corresponding flood extent and distribution of building footprints (c) Barahle refugee camp (d) Helaweyn refugee camp (e) Adi Harush refugee camp (f) Tsore refugee camp. (g) Bambasi refugee camp. (h) Okugo refugee camp. Note: Symbols next to camp names indicate which building footprint map was not available in each camp: + Google * OpenStreetMap.

exposed refugees, relative to approaches which incorporate UNHCR CP data. It also illustrates the deficiencies of the GP dataset's underlying census data. Incorporating better estimates of CP into the HRSL dataset increases national exposure estimates by a factor of 2.5. Worryingly, the GP approach is substantially underrepresenting flood exposure in four out of the five of the most exposed camps (see figure 4(a)), suggesting that the most exposed camps might be overlooked in traditional flood risk assessments if reliant only on global data. We performed a similar comparison using GHS-POP R2022 and found the results to be similar to those presented above (100 year GP exposure of 56 702 and GP + GPD exposure of 146 370). We replicate figure 4 using the GHS-POP R2022 data in the supplementary material.

To further enhance our flood exposure estimates, we incorporated building footprint data (where it was available) into our analysis of refugee flood exposure. In figure 4(b), we treat the CP + B approach as 'truth' and calculate the percentage error of the other three approaches (GP, CP, and CP + GPD) relative to the CP + B approach. Unsurprisingly, the GP approach significantly underestimates flood exposure relative to CP + B (−46.8% average percentage error in camps with building footprint data). Total AAE in camps where footprint data was available was 4857 for the GP approach, 16 890 for the CP approach, and 11 127 for the CP + GPD approach; compared to 12 823 for the CP + B approach. The CP approach generally overpredicted exposure relative to the CP + B and CP + GPD approaches, which emphasizes the importance of considering the location of structures within the camp, as these may be organized in a way to reduce exposure to flooding.

In figures 4(c)–(h) we highlight interesting camps that illustrate some of the characteristics discussed above and some additional issues that global data used in this context may present. Barahle (figure 4(c)) and Helaweyn (figure 4(d)) are the only two camps in Ethiopia where the CP approach predicts less exposure than the CP + B approach, indicating that structures within these camps have been constructed in areas of higher flood exposure relative to a scenario where population is distributed evenly throughout the camp. Interestingly, a large proportion of the exposure in these camps is to higher return period (>50 year RP) flooding, suggesting the risk recognition of rarer floods in these camps may be lacking. Figures 4(e)–(g) all show camps with structures that, for the most part, are set up in such a way as to reduce exposure to flooding. The majority of structures in Adi Harush (figure 4(e)), Tsore (figure 4(f)), and Bambasi (figure 4(g)) camps are located outside of the floodplain, something that can only be observed through the incorporation of building footprint data (and to some degree, population density data). Our findings corroborate previous work that found that humans tend to make rational decisions about flood risk, but the observation of this rational behavior is limited by the granularity of the data used in risk assessments [13]. Figure 4(h) illustrates a potential drawback to using the GP + GPD approach. Despite the fact that most of the structures within the Okugo camp are located outside the flood extent (according to the Google footprint data), the HRSL population density data indicates that the population (represented by only six cells) within the camp is predominantly located towards the west of the camp, within the flood extent. This leads to the large CP + GPD overestimation of AAE for the camp (1247% error) and emphasizes the importance of performing sense checks when using global data in refugee flood exposure assessments.

3.3. Refugee camp flood risk

By incorporating simple, flood hazard vulnerability thresholds (table 1) into our analysis we were able to calculate the flood risk faced by camp inhabitants. Promisingly, most flood exposed refugees (between 80%–85%, depending on flood return period) were exposed to flooding with low risk to life (depths below 0.15 m) (see figure 5(a)). Despite the low risk to life, flooding at these depths can pose significant health risks as latrines within camps can flood, contaminating the flood water and increasing the risk of disease [57]. The tents and temporary structures within camps are also vulnerable to floods of this depth as they may not have a foundation slab and many refugees sleep on the floor with no higher floor to move their possessions to during a flood event. Camps exposed to this category of flood risk can introduce flood adaptation measures to increase the resilience of structures and inhabitants of the camp. Measures such as elevating pit latrines, retrofitting existing shelters, and improving drainage can be applied with minimal need to restructure or relocate the camp. Such adaptation measures would be of relevance to camps such as Tierkidi, Nguenyiel, Kule, and Pugnido, where over 90% of the exposure is low risk to life. In camps where the immediate risk to life is more acute, the internal relocation of tents and facilities or the relocation of the camp may be necessary, as adaptation measures would be less effective for flooding at these depths. In the Okugo camp, where relative exposure is low (ranked 19th for flood exposure) but risk to life is high (316 people exposed to high-risk flooding for a frequently occurring 10 year event), structures with high risk exposure would benefit from internal relocation within the camp. In camps with a large proportion of inhabitants exposed to risky flooding (for example the Barahle camp, where 26% of inhabitants are exposed to medium-risk flooding for a 100 year event) it may be necessary to relocate the camp, as internal relocation of over 8000 refugees may not be feasible. Distinguishing between the levels

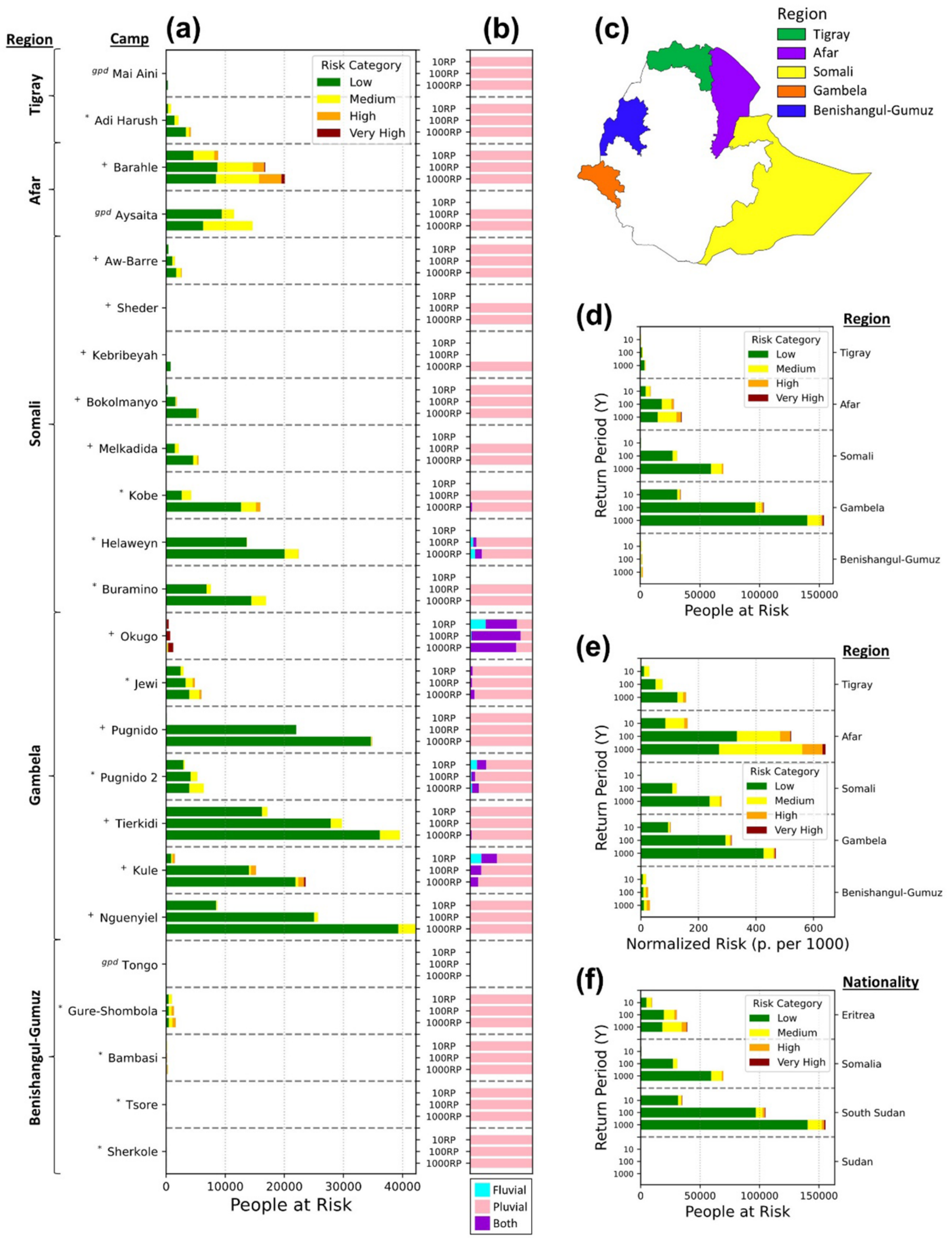


Figure 5. (a) Refugee camp flood risk classifications for frequent (10 year Return Period (RP)) infrequent (100 year RP) and extremely infrequent (1000 year RP) flooding. (b) Proportion of flooded area within each camp caused by fluvial flooding, pluvial flooding, or both types of flooding. No bar indicates no flood extent in the camp for the given return period. (c) Regions in Ethiopia housing refugee camps. (d) Absolute flood risk by region. (e) Relative flood risk by region. (f) Flood risk by nationality (in camps housing both Sudanese and South Sudanese refugees, we assume an even split). *Note:* To calculate risk to the camps we choose the building footprint dataset with the most buildings in each camp, where no building footprint dataset is available, we use the camp population and global population density approach. The symbols next to each camp name indicates which approach was used to calculate risk in each camp: + Google * OpenStreetMap ^{gpd} Camp Population and Global Population Density (CP + GPD).

of risk exposure within a camp can aid in determining within-camp adaptation actions.

We also distinguish between the type of flooding within each camp in figure 5(b). Refugee camps in Ethiopia are mostly at risk from pluvial flooding because the camps are predominantly located in smaller catchments. The type of flooding can influence the type of adaptation measure taken within a camp [58]. Pluvial flood risk reduction measures may include improved camp drainage, retrofitting of structures, nature-based solutions, or early warning systems [59]. In camps where the risk from fluvial flooding (such as in Okugo or Kule) is significant, this information could be used to relocate refugees from high-risk areas (see for example Ho, Vu [60]) or inform the construction of structural adaptation measures.

We can draw some regional conclusions about the flood risk faced by refugees in Ethiopia. The Benishangul-Gumuz region is of little concern in a flood risk context; despite housing five refugee camps, both the absolute (figure 5(d)) and relative (figure 5(e)) flood risk in the region is minimal. The two regions most at-risk from flooding are Afar and Gambela. For flooding with at least a medium risk-to-life, Afar has the most refugees at risk, both in an absolute and relative sense. Despite there being only two refugee camps in Afar: Barahle and Aysaita, both camps have a significant proportion of refugees exposed to medium-risk flooding or greater. Both these camps also house Eritrean refugees, which is the nationality most exposed to medium-risk flooding or greater (see figure 5(f)), and thus at greater risk of further relocation. Understanding the regional disparities of refugee flood risk within Ethiopia can be useful high-level information for decision makers in prioritizing funding for camp flood adaptation.

3.4. The usefulness and limitations of our approach

Our refugee camp flood risk assessment combines global flood hazard maps, building footprint data, local CP estimates, and flood hazard vulnerability thresholds to estimate refugee flood risk at the camp, regional, and national level. Results of an assessment such as this can be useful for understanding the severity and distribution of flood risk to refugees within a country and can help prioritize action and funds to reduce this flood risk. Understanding the risk exposure within camps can help ensure the appropriate adaptation measures are taken. However, there are also several limitations to our approach that should be addressed. We group these limitations under the three components of risk: hazard, exposure, and vulnerability.

Despite being amongst the most resolved ($\sim$90 m) global flood maps, the granularity of our camp risk assessment means that the hazard maps we use are still at the coarse end of the scale for this application. These GFMs cannot capture the intricate dynamics of flood flow within camps and should not be used for detailed design and planning of within-camp adaptation. Instead, results should be used to prioritize the camps in which detailed local flood modeling should be carried out. These detailed assessments can be expensive, but have been shown to be invaluable in understanding risk within camps in Syria [24] and Bangladesh [23].

Although the incorporation of building footprint data significantly enhanced our risk analysis, these datasets still cannot capture the dynamic population changes within the camps. The Google building footprint data we use is static, representative of the situation in June 2021 when the footprints were extracted. Similarly, OpenStreetMap data is updated in a fragmented manner, given its approach that relies on community uploads. As such, the OpenStreetMap data in certain camps may be more up-to-date than others. The internal organization of camps is also much more complex than simply assuming population is evenly distributed across camp buildings. Structures within the camps have different uses; in addition to shelters, camps will typically contain latrines, showers, refuse pits, health facilities, schools, markets, feeding centers, and administrative buildings [10]. Understanding which facilities within a camp are at risk is important for prioritizing adaptation.

Our measure of vulnerability is highly generalized, based on categorizations developed in the global north, which will not be entirely applicable in the context of Ethiopian refugee camps. It is also based on flood depth and extent with no consideration of flood flow velocity, which will be an important component of camp vulnerability, given both the temporary nature of structures within the camp and the large proportion of pluvial risk identified in the camps. Further, vulnerability profiles will vary from camp to camp and depend on factors such as the construction material and design of shelters, the effectiveness of existing adaptation measures, and any pre-existing vulnerabilities within the camps. We have also not considered the social vulnerability of camp inhabitants. Factors such as economic status, age, gender, marital status, experience of flooding, and household size can all affect disaster resilience [61–63]. Disaggregating this information from global datasets (such as [64] and [65]) will pose similar problems to the issues we identified with the census derived population estimates in this study. For certain camps, the UNHCR has begun reporting population demographics. This information could be incorporated into studies such as this to begin to unpack the social vulnerabilities of camp inhabitants. To properly assess the vulnerability of camps, it is essential to engage with camp planners and managers, who will have a better understanding of the unique vulnerabilities of the camps they design and operate.

4. Conclusion/next steps

In our critical evaluation of global flood exposure estimates, we found significant differences in the accuracy of GP datasets at capturing the population of refugee camps in Ethiopia. The most accurate GP datasets were GHS-POP R2022 and HRSL which, across the 24 camps, captured only 43.7% and 35.6% of the UNHCR reported CPs, respectively. We conclude that GP datasets do a poor job at estimating refugee CPs and should not be used in studies of refugee flood risk. Instead, we propose alternative approaches for representing exposure that incorporate locally reported estimates of CP. Using these approaches, we found that 95.8% of refugee camps were exposed to flooding of some degree and between 143 208 (19.8%) and 182 125 refugees (25.2%) were exposed to a 100 year return period flood. By incorporating information on the spatial distribution of camp inhabitants into our exposure analysis (using either building footprints or GPD maps), we found that estimates of AAE decreased significantly, suggesting that camp planners are making rational decisions about the internal organization of camps with respect to potential flood exposure. To understand the different risks facing these camps, we introduced hazard vulnerability thresholds into our analysis. We found that most refugees in Ethiopia are exposed to pluvial flooding, and the majority of this (80%–85%) exposure is a low risk to life. Many of the exposed camps would benefit from within-camp adaptation measures. At a regional level, we found that Afar and Gambela were the regions most at risk. South Sudanese refugees were by far the most exposed nationality to flooding, but Eritrean refugees were the nationality most exposed to flooding with a high risk to life.

Our analysis has demonstrated the utility of combining local and global data for assessing refugee camp flood risk in Ethiopia. Results from such an analysis can be useful for understanding the degree of risk faced by refugees in a country and for prioritizing further detailed risk analyses. We have highlighted the key limitations of our approach: the coarse flood hazard data, the simplified assumption of camp anatomy, and the generalized hazard vulnerability thresholds. Future work should integrate the knowledge of camp planners and managers into the risk analysis, as they have an intimate understanding of the internal organization and vulnerabilities of camps. Further research could also incorporate and evaluate additional global datasets such as night-time light data, which could be useful (due to its high temporal resolution) for capturing changes in camp boundaries and populations, something the GP and building footprint datasets are not able to do. This study only considered refugee camps, but internally displaced people represent another vulnerable group often housed in temporary camps, whose risk is unknown. The consistency afforded by global datasets means such an analysis can be easily reproduced or scaled. Our paper has shown how such an analysis can be useful for identifying risk hotspots in Ethiopia. Scaling this work globally to include all camps housing forcibly displaced people could help international agencies understand their global risk exposure and provide a consistent first-order assessment of risk for managers and planners at the level of every camp.

Data availability statement

Fathom 2.0 global flood model data are available for academic purposes and were provided by Fathom.

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.7962039.

Acknowledgments

This work was supported by the Water Security and Sustainable Development Hub funded by the UK Research and Innovation's Global Challenges Research Fund (GCRF) [Grant No. ES/S008179/1] and by the National Environment Research Council (NERC) [Grant No. NE/V017756/1]. The authors would like to thank the members of the Global Flood Partnership (GFP), who have helped to shape this research through discussions and feedback at annual GFP meetings. Thanks to Zoe Hagan and Niall McNamara, MEng students from the School of Civil Engineering at the University of Leeds for their dissertation work which helped contributed to this research topic.

ORCID iDs

Mark V Bernhofen https://orcid.org/0000-0002-4919-0111
Mark A Trigg https://orcid.org/0000-0002-8412-9332

References

[1] UNHCR 2022 *Global Trends—Forced Displacement in 2021* (Copenhagen: Statistics and Demographics Section, UNHCR Global Data Service) (available at: https://www.unhcr.org/media/40152)
[2] Clement V *et al* 2021 Groundswell part 2: *Acting on Internal Climate Migration* (Washington, DC: World Bank) (available at: http://hdl.handle.net/10986/36248) License: CC BY 3.0
[3] Schleussner C-F, Donges J F, Donner R V and Schellnhuber H J 2016 Armed-conflict risks enhanced by climate-related disasters in ethnically fractionalized countries *Proc. Natl Acad. Sci. USA* **113** 9216–21
[4] Mach K J *et al* 2019 Climate as a risk factor for armed conflict *Nature* **571** 193–7
[5] von Uexkull N and Buhaug H 2021 Security implications of climate change: a decade of scientific progress *J. Peace Res.* **58** 3–17
[6] UNHCR Refugee camps explained 2021 (available at: www.unrefugees.org/news/refugee-camps-explained/#What%20is%20a%20refugee%20camp?) (Accessed: 15 August 2022)
[7] UNHCR 2015 UNHCR, displacement and disaster risk reduction (UNHCR) (available at: https://www.unhcr.org/media/33134)
[8] Sphere Project 2011 Sphere handbook: humanitarian charter and minimum standards in disaster response, 2011
[9] Davis J, Lambert R and R R 2002 *Engineering in Emergencies: A Practical Guide for Relief Workers* (Rugby: Practical Action Publishing)
[10] UNHCR 2018 *Camp Site Planning Minimum Standards, in UNHCR Emergency Handbook* (UNHCR) (available at:

https://emergency.unhcr.org/emergency-assistance/shelter-camp-and-settlement/camps/camp-site-planning-minimum-standards)
[11] Trigg M A *et al* 2020 *Global Flood Models, in American Geophysical Union Monograph: Global Drought and Flood* (New York: Wiley) (https://doi.org/10.1002/9781119427339.ch10)
[12] Bernhofen M V *et al* 2022 The role of global data sets for riverine flood risk management at national scales *Water Resour. Res.* **58** e2021WR031555
[13] Smith A, Bates P D, Wing O, Sampson C, Quinn N and Neal J 2019 New estimates of flood exposure in developing countries using high-resolution population data *Nat. Commun.* **10** 1814
[14] Bernhofen M V, Trigg M A, Sleigh P A, Sampson C C and Smith A M 2021 Global flood exposure from different sized rivers *Nat. Hazards Earth Syst. Sci.* **21** 2829–47
[15] Van Den Hoek J and Friedrich H K 2021 Satellite-based human settlement datasets inadequately detect refugee settlements: a critical assessment at thirty refugee settlements in Uganda *Remote Sens.* **13** 3574
[16] Thomson D R, Gaughan A E, Stevens F R, Yetman G, Elias P and Chen R 2021 Evaluating the accuracy of gridded population estimates in slums: a case study in Nigeria and Kenya *Urban Sci.* **5** 48
[17] UNHCR 2021 Ethiopia country refugee response plan January 2022—December 2022 (UNHCR)
[18] UNHCR 2022 *Operational Data Portal Country View: Ethiopia* (available at: https://data.unhcr.org/en/country/eth) (Accessed 24 August 2022)
[19] OCHA 2014 Ethiopia: Areas of Humanitarian Concerns Snapshot (as of September 2014) (OCHA)
[20] DRC 2022 Global displacement forecast 2022—July update (Danish Refugee Council)
[21] Seneviratne S I *et al* 2021 Weather and climate extreme events in a changing climate *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*
[22] Kam P M, Aznar-Siguan G, Schewe J, Milano L, Ginnetti J, Willner S, McCaughey J W and Bresch D N 2021 Global warming and population change both heighten future risk of human displacement due to river floods *Environ. Res. Lett.* **16** 044026
[23] ISCG 2019 Summary report: natural hazard risk analysis taskforce, Cox's bazar (Inter Sector Coordination Group (ISCG))
[24] REACH *et al* 2021 Syria—West Dana sub-district | idleb: IDP camps and informal sites flood susceptibility and flood hazard assessment (available at: https://reliefweb.int/report/syrian-arab-republic/syria-west-dana-sub-district-idleb-idp-camps-and-informal-sites-flood)
[25] UNHCR Ethiopia 2020 Ethiopia—refugee camp locations (Humanitarian Data Exchange)
[26] OpenStreetMap contributors 2022 *OpenStreetMap Data Extracts* (available at: https://download.geofabrik.de)
[27] Sirko W *et al* 2021 Continental-scale building detection from high resolution satellite imagery (arXiv:2107.12283)
[28] Bernhofen M V, Blenkin F and Trigg M A 2023 Data to support the publication 'Unknown risk: assessing refugee camp flood risk in Ethiopia' (1.0.1) [Data set] *Zenodo* https://doi.org/10.5281/zenodo.7962039
[29] Schiavina M F S and MacManus K 2019 GHS population grid multitemporal (1975, 1990, 2000, 2015) (European Commission, Joint Research Centre (JRC))
[30] Freire S *et al* 2020 Enhanced data and methods for improving open and free global population grids: putting 'leaving no one behind' into practice *Int. J. Digit. Earth* **13** 61–77
[31] Freire S *et al* 2016 Development of new open and free multi-temporal global population grids at 250 m resolution *Geospatial Data in a Changing World; Association of Geographic Information Laboratories in Europe (AGILE) (Organiser)* (AGILE) JRC100523 (available at: https://publications.jrc.ec.europa.eu/repository/handle/JRC100523)
[32] Schiavina M, Freire S and MacManus K 2022 GHS-POP R2022A—GHS Population Grid Multitemporal (1975–2030) (European Commission, Joint Research Centre)
[33] Tiecke T G L *et al* 2017 Mapping the world population one building at a time (arXiv:1712.05839)
[34] WorldPop and CIESIN 2018 Global high resolution population denominators project
[35] Bondarenko M *et al* 2020 Census/projection-disaggregated gridded population datasets, adjusted to match the corresponding UNPD 2020 estimates, for 51 countries across sub-Saharan Africa using building footprints WorldPop (University of Southampton) (https://doi.org/10.1016/j.compenvurbsys.2019.101444)
[36] Florczyk A *et al* 2019 GHS urban centre database 2015, multitemporal and multidimensional attributes, R2019A (European Commission & Joint Research Centre)
[37] Bates P D, Horritt M S and Fewtrell T J 2010 A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling *J. Hydrol.* **387** 33–45
[38] Sampson C C, Smith A M, Bates P D, Neal J C, Alfieri L and Freer J E 2015 A high-resolution global flood hazard model *Water Resour. Res.* **51** 7358–81
[39] Borga M, Stoffel M, Marchi L, Marra F and Jakob M 2014 Hydrogeomorphic response to extreme rainfall in headwater systems: flash floods and debris flows *J. Hydrol.* **518** 194–205
[40] Bernhofen M V *et al* 2021 River flood susceptibility map (available at: http://archive.researchdata.leeds.ac.uk/id/eprint/828)
[41] Yamazaki D, Ikeshima D, Tawatari R, Yamaguchi T, O'Loughlin F, Neal J C, Sampson C C, Kanae S and Bates P D 2017 A high-accuracy map of global terrain elevations *Geophys. Res. Lett.* **44** 5844–53
[42] Yamazaki D, Ikeshima D, Sosa J, Bates P D, Allen G H and Pavelsky T M 2019 MERIT hydro: a high-resolution global hydrography map based on latest topography dataset *Water Resour. Res.* **55** 5053–73
[43] Trigg M A *et al* 2016 The credibility challenge for global fluvial flood risk analysis *Environ. Res. Lett.* **11** 10
[44] Bernhofen M V *et al* 2018 A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique *Environ. Res. Lett.* **13** 104007
[45] Mester B, Willner S N, Frieler K and Schewe J 2021 Evaluation of river flood extent simulated with multiple global hydrological models and climate forcings *Environ. Res. Lett.* **16** 094010
[46] Alfieri L, Bisselink B, Dottori F, Naumann G, de Roo A, Salamon P, Wyser K and Feyen L 2017 Global projections of river flood risk in a warmer world *Earths Future* **5** 171–82
[47] Hirabayashi Y, Tanoue M, Sasaki O, Zhou X and Yamazaki D 2021 Global exposure to flooding from the new CMIP6 climate model projections *Sci. Rep.* **11** 3740
[48] Environment Agency 2019 What is the risk of flooding from surface water map? (Environment Agency)
[49] Smith G and Cox R 2019 Safety design criteria, in Australia rainfall runoff-a guide to flood estimation. Book 6: flood hydraulics
[50] Cox R, Shand T and Blacka M 2010 Revision project 10: appropriate safety criteria for people. Stage 1 report. Australian rainfall and runoff (Water Research Laboratory, The University of New South Wales)
[51] Pesaresi M, Corbane C, Julea A, Florczyk A, Syrris V and Soille P 2016 Assessment of the added-value of sentinel-2 for detecting built-up areas *Remote Sens.* **8** 299
[52] Liu F, Wang S, Xu Y, Ying Q, Yang F and Qin Y 2020 Accuracy assessment of global human settlement layer (GHSL) built-up products over China *PLoS One* **15** e0233164
[53] UNHCR 2016 *Shelter Design Catalogue*

[54] Schiavina M *et al* 2022 GHSL data package 2022, in JRC scientific information systems and databases report (European Commission, Joint Research Centre)

[55] IOM 2018 Guidelines for the harmonization of migration data management in the ECOWAS region (Global Migration Data Analysis Centre)

[56] Haile A T, Kusters K and Wagesho N 2013 Loss and damage from flooding in the Gambela region Ethiopia *Int. J. Glob. Warm.* **5** 483–97

[57] Borges Pedro J P, Oliveira C A D S, de Lima S C R B and von Sperling M 2020 A review of sanitation technologies for flood-prone areas *J. Water Sanit. Hyg. Dev.* **10** 397–412

[58] Tiepolo M, Belcore E, Braccio S, Issa S, Massazza G, Rosso M and Tarchiani V 2021 Method for fluvial and pluvial flood risk assessment in rural settlements *MethodsX* **8** 101463

[59] Yin J, Gao Y, Chen R, Yu D, Wilby R, Wright N, Ge Y, Bricker J, Gong H and Guan M 2023 Flash floods: why are more of them devastating the world's driest regions? *Nature* **615** 212–5

[60] Ho J C *et al* 2021 Chapter 6—from cloud to refugee camp: a satellite-based flood analytics case-study in Congo-Brazzaville *Earth Observation for Flood Applications* ed G J P Schumann (Amsterdam: Elsevier) pp 131–46

[61] Black R, Adger W N, Arnell N W, Dercon S, Geddes A and Thomas D 2011 The effect of environmental change on human migration *Glob. Environ. Change* **21** S3–S11

[62] Oppenheimer M *et al* 2014 Emergent risks and key vulnerabilities *Climate Change 2014 Impacts, Adaptation, and Vulnerability: Part A: Global and Sectoral Aspects, Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press) pp 1039–100

[63] Ayanlade A, Smucker T A, Nyasimi M, Sterly H, Weldemariam L F and Simpson N P 2023 Complex climate change risk and emerging directions for vulnerability research in Africa *Clim. Risk Manage.* **40** 100497

[64] Kummu M, Taka M and Guillaume J H A 2018 Gridded global datasets for gross domestic product and human development index over 1990–2015 *Sci. Data* **5** 180004

[65] Nirandjan S, Koks E E, Ward P J and Aerts J C J H 2022 A spatially-explicit harmonized global dataset of critical infrastructure *Sci. Data* **9** 150