University of Reading

The utility of impact data in flood forecast verification for anticipatory actions: case studies from Uganda and Kenya

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Mitheu, F. ORCID: https://orcid.org/0000-0003-1655-449X, Tarnavsky, E. ORCID: https://orcid.org/0000-0003-3403-0411, Ficchì, A., Stephens, E. ORCID: https://orcid.org/0000-0002-5439-7563, Cornforth, R. ORCID: https://orcid.org/0000-0003-4379-9556 and Petty, C. (2024) The utility of impact data in flood forecast verification for anticipatory actions: case studies from Uganda and Kenya. Journal of Flood Risk Management, 18 (1). e12911. ISSN 1753-318X doi: 10.1111/jfr3.12911 Available at https://centaur.reading.ac.uk/111678/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/jfr3.12911

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Received: 30 November 2021 | Revised: 16 February 2023 | Accepted: 23 March 2023

DOI: 10.1111/jfr3.12911

SPECIAL ISSUE

CIWEM Chartered Institution of Water and Environmental Management | Journal of Flood Risk Management | WILEY

The utility of impact data in flood forecast verification for anticipatory actions: Case studies from Uganda and Kenya

Faith Mitheu[1,2] | **Elena Tarnavsky**[2,3] | **Andrea Ficchì**[1,4] | **Elisabeth Stephens**[3,5] | **Rosalind Cornforth**[2] | **Celia Petty**[2,6]

[1]Department of Geography and Environmental Science, University of Reading, Reading, UK

[2]Walker Institute, University of Reading, Reading, UK

[3]Department of Meteorology, University of Reading, Reading, UK

[4]Politecnico di Milano, Milano, Italy

[5]Red Cross Red Crescent Climate Centre, The Hague, The Netherlands

[6]Evidence for Development, University of Reading, Reading, UK

Correspondence
Faith Mitheu, Department of Geography and Environmental Science, University of Reading, Reading, UK.
Email: f.k.mitheu@pgr.reading.ac.uk

Funding information
UK's Natural Environment Research Council (NERC), Grant/Award Number: NE/R007799/1; Forecasts for AnTicipatory HUManitarian action, Grant/Award Number: NE/P000525/1; National-scale IMpact-based forecasting of Flood Risk in Uganda, Grant/Award Number: NE/S00596X/1

Abstract

Skilful flood forecasts have the potential to inform preparedness actions across scales, from smallholder farmers through to humanitarian actors, but require verification first to ensure such early warning information is robust. However, verification efforts in data-scarce regions are limited to only a few sparse locations at pre-existing river gauges. Hence, alternative data sources are urgently needed to enhance flood forecast verification to better guide preparedness actions. In this study, we assess the usefulness of less conventional data such as flood impact data for verifying flood forecasts compared with river-gauge observations in Uganda and Kenya. The flood impact data contains semi-quantitative and qualitative information on the location and number of reported flood events derived from five different data repositories (Dartmouth Flood Observatory, DesInventar, Emergency Events Database, GHB, and local) over the 2007–2018 period. In addition, river-gauge observations from stations located within the affected districts and counties are used as a reference for verification of flood forecasts from the Global Flood Awareness System. Our results reveal both the potential and the challenges of using impact data to improve flood forecast verification in data-scarce regions. From these, we provide a set of recommendations for using impact data to support anticipatory action planning.

KEYWORDS

disaster risk reduction, floods, forecast verification, humanitarian early action, impacts, non-traditional verification data

1 | INTRODUCTION

Climate change, variability, and environmental changes disproportionately affect the agricultural sector in Africa with important implications for anticipatory action as part of humanitarian response. In the agricultural sector, these changes could force smallholder farmers who depend on rain-fed crops or flood-recession agriculture to significantly adjust their farm activities (Ficchì & Stephens, 2019; Ochieng et al., 2016; Salack et al., 2015). In Uganda, farmers need reliable and skilful information on the rainy season onset and amount of rainfall, as well

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Journal of Flood Risk Management* published by Chartered Institution of Water and Environmental Management and John Wiley & Sons Ltd.

as flood occurrence, duration, magnitude, and severity ~1–2 months before the season onset to inform their coping strategies (Mitheu et al., 2022). Decision-makers and humanitarian actors aiming to reduce risks and protect livelihoods are also increasingly considering forecast information to inform the early action mechanisms and operational decisions (Coughlan De Perez et al., 2016; Emerton et al., 2020; Hansen et al., 2019; Lopez et al., 2020; Nidumolu et al., 2020). Given this, the skill of any forecast information provided needs to be transparent and well understood if it is to inform preparedness actions appropriately.

In the context of users' needs, forecasts should be evaluated based on their potential to trigger early actions, which can reduce expected losses if an extreme event occurs (Lopez et al., 2020). The evaluation should also consider the consequences of 'acting in vain', which are particularly important in disaster risk reduction and humanitarian actions (Coughlan De Perez et al., 2015). Indeed, several studies have shown that verified and skilful forecasts have the potential to improve preparedness actions for both the agricultural and humanitarian sectors (Coughlan De Perez et al., 2016; MacLeod et al., 2021; Nidumolu et al., 2020; Nyadzi et al., 2019; Paparrizos et al., 2020). However, this verification is carried out only for regions with long-term historical hydrometeorological observations, typically from in situ stations such as river gauges. In forecast verification, these observations are commonly known as conventional observations (Marsigli et al., 2021).

In data-scarce regions, where conventional observations are limited (Coughlan De Perez et al., 2016; Ogutu et al., 2017), less conventional verification data can be derived from, for example, social media reports, citizen volunteered information, impact/damage reports, and insurance data. The resulting information can be used to bridge the forecast verification gap through non-traditional approaches as they provide a more direct representation of the event (Marsigli et al., 2021). For example, information from insurance databases (Bernet et al., 2017; Cortès et al., 2018), as well as online tools such as Google Trends and Twitter feeds (de Bruijn et al., 2019; Thompson et al., 2022) have been used as reference information to evaluate the occurrence of floods. Impact data have also been used with river-gauge observations to identify the magnitude of discharge that is associated with flooding (Coughlan De Perez et al., 2016). Notably, impact data offer an advantage in the verification of forecast information, because they can be derived from openly accessible data repositories containing quantitative and qualitative information across large spatial areas that enable a better and direct representation of the impacts of the extreme event. The use of impact data in forecast verification can only be possible in areas with exposure and vulnerability for the impact to be reported.

It is worth noting that global data repositories such as the Emergency Events Database (EM-EM-DAT, 2020) and the United Nations Disaster Inventory System (DesInventar [DI]; UNISDR, 2018) are prone to biases due to known limitations (Gall et al., 2009). These limitations include under-reporting/over-reporting of the hazards, aggregated spatial coverage, over-representation of certain locations, and/or focus on the specific type(s) of impacts. Furthermore, differences in the criteria for the inclusion of events in the repositories may result in non-uniformity in the estimates of the impacts reported in each repository. In addition, if unverified, impact data collection methods (e.g., from governments and media) may lead to errors in the resulting information (Guha-Sapir & Below, 2002). Despite these caveats, these data repositories represent a potentially valuable source of less conventional data for monitoring and verifying hazards. For example, impact data can be integrated with other geophysical parameters to sub-categorise flash floods from the primary corresponding disaster type (Kruczkiewicz, Bucherie, et al., 2021). Therefore, if the limitations of impact data are appropriately understood, with guidance on their interpretation and relevant recommendations, impact data can be improved to effectively support anticipatory actions.

In this study, we assess the usefulness of flood impact data to verify flood forecast information across Uganda and Kenya compared with river-gauge observations. We verify the river flood forecast from the Global Flood Awareness System (GloFAS) of the Copernicus Emergency Management Service (Harrigan et al., 2023) using two reference observations. The river-gauge observations and flood impact data were derived from several global and national data repositories.

The study addresses two research questions:

1. How suitable are impact data for verifying flood forecasts compared to river-gauge observations?
2. Where river-gauge observations are limited or unavailable, how best can impact data be used to verify flood forecasts and ensure anticipatory actions are informed?

Through focussed case studies in two East African countries, we investigate the non-traditional approach of forecast verification using impact data relative to the traditional way of verification using river-gauge observations. Consequently, we provide recommendations on how best impact data can be used in areas with no or limited river-gauge observations to increase confidence in the use of forecast products in data-scarce regions.

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2 | CONTEXT

In this section, we describe the case study regions and the datasets used for the analysis, that is, the GloFAS reforecast discharge data, river-gauge observations, and the impact data from several data repositories.

2.1 | Case study regions

The Netherlands-based IKEA Foundation is supporting the Uganda and Kenya Red Cross Societies (URCS and KRCS, respectively) to develop early warning mechanisms to prepare for floods through the Innovative Approaches for Response Preparedness (IARP) project. In Uganda, several high-risk areas were identified using vulnerability and risk layers developed by the National Emergency Operations and Coordination Centre (NECOC), including a total of 15 districts, for the early action protocol (EAP) development. These regions are prone to flooding and waterlogging across the two rainy seasons between May and November (April–May, Long Rains; September–November, Short Rains). In Kenya, flood-prone river basins including Tana, Nzoia, and Athi are considered for the implementation of flood early actions. Examples of early actions include community awareness, distribution of cash and shelter kits, dissemination of early warning information among others (see KRCS, 2021; URCS, 2021).

The case study regions in Uganda and Kenya were selected based on locations with available river-gauge observations. In Uganda, the districts of Katakwi and Amuria on the Akokorio river (hereafter 'Katakwi'), Tororo (Butaleja), and Mbale (Bududa and Manafwa) on Manafwa River (hereafter 'Manafwa'), and Kiboga, Mubende, and Hoima on the Mayanja River (hereafter 'Mayanja') are considered. In Kenya, the county of Tanariver and Garissa on Tana River (hereafter 'Tana'), Busia and Siaya on Nzoia river (hereafter 'Nzoia'), and Taitataveta and Kilifi on Athi river (hereafter 'Athi') have been considered. Figure 1 shows the locations of the river-gauge stations and the affected counties/district in Kenya and Uganda, respectively.

2.2 | GloFAS flood forecasts

GloFAS is an operational global ensemble flood forecasting system developed jointly between the European Commission's Joint Research Centre (JRC), the European Centre for Medium-Range Weather Forecasts (ECMWF), and the University of Reading researchers (Alfieri et al., 2013). The system provides probabilistic extended range discharge forecasts for up to 45 days and seasonal outlooks up to 4 months lead time (Emerton et al., 2018) over the entire globe at a resolution of 0.1°. From GloFAS v3.1 (current operational version), the LISFLOOD hydrological model (van der Knijff et al., 2010) is forced by an

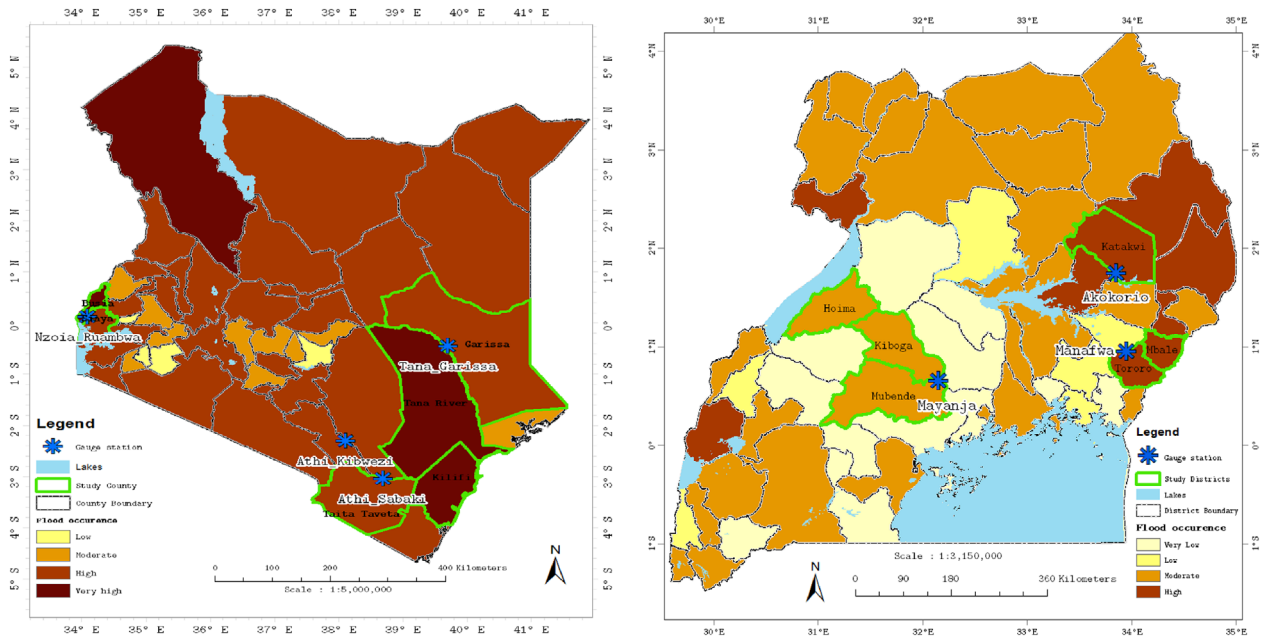

FIGURE 1 Flood occurrence maps for Kenya and Uganda show the study counties/districts and the river gauge locations. The map was created using impact data collated from four different data repositories from 2007 to 2018. The colour scheme represents the number of years out of the 12 years considered when floods occurred ranging from low (1–3 years), moderate (4–6 years), high (7–9 years), and very high (10–12 years).

17533184x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ensemble of medium to extended range meteorological forecasts from the ECMWF Integrated Forecast System to produce 51 ensemble members of daily streamflow at various lead times. LISFLOOD has been calibrated using daily streamflow data at over 1200 river basins worldwide (Hirpa et al., 2018).

GloFAS v3.1 hydrological performance was evaluated for the period 1979–2019 for over 1500 verification stations across the world using various verification metrics (Kling Gupta Efficiency, Bias, variance, etc). Prudhomme and Zsoter (2021) provide details on the hydrological assessment methodology and further discussion on GloFAS performance evaluation. GloFAS provides daily discharge amounts [m^3/s] from which probabilities of flood threshold exceedance can be derived. For flood detection, these forecasts time series are compared against a set of flood thresholds that are derived from the same model climatology (Zsoter et al., 2020) to avoid the impact of systematic biases in the GloFAS climatology on flood forecast probabilities. In this study, we use daily GloFAS v3.1 reforecast discharge data from 2007 to 2018 extracted for the gauge locations in Kenya and Uganda, respectively (Figure 1).

2.3 | Flood thresholds

In the 30-day operational GloFAS forecast interface (https://www.globalfloods.eu/), four different flood return periods (2, 5, 10, and 20 years) are provided and can be used as the thresholds for severe flood events. Zsoter et al. (2020) provide a detailed explanation of how these return periods are computed using GloFAS ensemble reforecasts. Furthermore, thresholds computed as percentiles of the daily river flow time series can also be used to define various hydrological conditions (e.g., high/low river flows) and have been used by several authors to evaluate forecasts from GloFAS or similar forecasting systems (see Alfieri et al., 2013; Arnal et al., 2018; Emerton et al., 2018; MacLeod et al., 2021). For example, high percentiles (90th percentile or greater) have been used to show a high likelihood of floods when the river flow at a gauging station is above that percentile (MacLeod et al., 2021). In the broad hydrological literature, the notation for flow percentiles is not always consistent or clear, so when percentiles are used, the definition needs to be specified clearly.

In this study, we adopt the traditional definition of percentiles used in statistics where a *k*th *percentile* (with *k* in the range of 1–100) for a time series is the level below which (or at which) a *k percentage* of values in its distribution falls (the inclusive definition of percentile is adopted). For example, a 90th percentile is equal to or >90% of the river discharge recorded during the specified period. In flood-related studies, a percentile flow can also be referred to in terms of 'percent exceedance' to indicate the percentage of time that the discharge value is likely to be equalled or exceeded (see; Flow, Excedance and Percentiles, 2023; National River Flow Archive, 2023). Thus, in this study we use the 90th, 95th, and 99th percentile calculated from the re-forecast (all ensemble members) or observed time series of daily discharge, corresponding to high-flow levels exceeded only by a minor portion of the days in the data, that is,10%, 5%, and 1% respectively.

Due to data availability, we followed a percentile-based method to compute flood thresholds for forecast verification similar to previous authors (e.g., Alfieri et al., 2013; Arnal et al., 2018; Emerton et al., 2018; MacLeod et al., 2021). The choice of using these thresholds and not higher Return Periods (e.g., 5-year or 20-year return periods computed from annual maxima) is motivated by the need for robust statistics, given the short data periods available (2007–2018). For the forecasts, these thresholds are lead-time dependent (Zsoter et al., 2020), that is, calculated from the reforecast time series at each given lead time available.

2.4 | River-gauge observations

Observed point-based discharge time series for the river gauges considered here were provided by the Department of Water Resources Management (DWRM) in Uganda and by the Kenya Water Authority (WRA) for Kenya. The time series consists of daily discharge values over long periods with all stations having at least 5 years of daily discharge data over the study period. The river-gauge observations corresponding to the period of the impact data (2007–2018) have been used for the subsequent analysis.

2.5 | Flood impact data

Flood impact data have been used to extend our capability to verify GloFAS flood forecasts beyond conventional observations from sparse gauge networks. The flood impact data contain semi-quantitative and qualitative information on the location and number of reported flood events derived from five different data repositories: (1) Dartmouth Flood Observatory (DFO) Archive (Brakenridge, 2015), (2) DI (UNISDR, 2018), (3) EM-DAT (EM-DAT, 2020), (4) the Global Hazard Weekly Bulletin (PHE, 2015), and (5) local sources (URCS, KRCS, media, etc.) for the 2007–2018 period. These data

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

were collated for Uganda and Kenya for the study regions (districts/counties) for further analysis. The characteristics of these data repositories are summarised in Table 1.

In an ideal situation, an impact would be defined as a combination of the number of people affected and the quantitative estimate of any loss of property and livelihoods. However, the used repositories do not have enough quantitative loss and damage information disaggregated to sub-national administrative units to enable the quantification of impacts and the severity of the flood events. We, therefore, consider the number of flood events reported as a proxy to the impact with an assumption that flood events that result in considerable impacts would be reflected in the data repositories used. The flood events are then classified as either 1 or 0 if the event was reported or not, respectively. The assessment of the number of flood events from the various sources, as well as the overlap (events that are common across the repositories used here), would help understand which data repository is used to identify the highest number of flood events for each study location.

3 | METHODOLOGY

Here, we outline the comparative analysis of river-gauge observations and impact data and the verification of GloFAS flood forecasts using two reference data sets through a set of skill scores. To assess the usefulness of flood impact data in verifying flood forecasts, first, the adequacy of the impact data in supplementing the river-gauge observations is evaluated using Type I and Type II error indices. Second, the flood forecast data are verified using river discharge and impact data as reference, and the verification outcomes based on the probability of detection (POD) and false alarm ratio (FAR) are compared.

3.1 | Comparison of river-gauge observations and impact data

In this part of the analysis, we compare the river-gauge observations and impact data. River-discharge value (Q) that has the potential to cause flooding is defined using the 90th and 95th percentile as the threshold, that is, a flood event (binary) occurs when Q is above the threshold, and it does not occur if Q is below the threshold. The total flood events from impact data are derived from the data repositories while considering the overlaps using the timestamp to avoid duplication in the total events. This means that an event that occurs across all the data repositories for the same timestamp is considered one event. The total flood events from impact data (binary) are then compared with river-gauge observations (binary).

Here, we assess the consistency of impact data false positive and false negative outcomes using a window of 7 days (from the day of the observed event up to 7 days ahead) against the flood events picked from the river-gauge observations. Using a 2×2 contingency table, the false-positive outcome is used to compute the 'Type I error' which represents the ratio of the flood events detected by river-gauge observations with no impacts divided by the total flood events (from gauge observations). Additionally, the false negative outcome is used to compute 'Type II error' which represents the ratio of flood events detected in the impact data and not by river-gauge observations divided by the total number of flood events (from impact data). We first compare the river-gauge data (binary) with the impact data (binary) from the various sources across the locations. Next, we compare the river-gauge observations against impact data from a single data repository to assess if impact data from some repositories are better than others in detecting flood events. Type I and II errors are calculated according to the equations in Table 2.

3.2 | Flood forecast verification using river-gauge observations and impact data

A set of skill scores were used to evaluate the occurrence of forecasted floods from the GloFAS system against river-gauge observations and impact data. The ability of the forecast to discriminate between events and non-events is commonly measured using skill metrics calculated from a 2×2 contingency table. Two skill scores were used to quantify the occurrence of flood events (Wilks, 2006): (1) POD or hit rate, which measures the fraction of observed events that were correctly predicted (perfect score of 1) and (2) FAR, which indicates the fraction of the predicted events that did not occur (perfect score of 0). Table 3 shows the equations used to calculate the skill scores.

In this study, the verification of flood forecast events is based on the need to provide reliable flood forecast information to inform anticipatory actions taken by the communities and humanitarian actors. The preferred verification outcome will therefore depend on the decision-making strategies the actors are willing to take. For example, humanitarian actors might need to decide if actions should be taken based on any forecast probability which might be costly due to the number of events but would ensure reduced losses if the events materialise.

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Characteristics of the data repositories that were used to derive impact data for the study.

Data repository (reference)	Temporal coverage	Criteria for inclusion	Actors/collection methods	Accessibility	Spatial coverage	Parameters
Dartmouth flood observatory (Brakenridge, 2015)	1993 to present	• Large and extreme flood events	News reports, Governments, Flood lists, and remote sensing sources	Available upon request at https://floodobservatory.colorado.edu/	Country-level but provides the centroids for locations affected	Reports the main cause of impacts and categories of loss (see Table S1)
DesInventar (Integrated Research on disaster risk, 2014; UNISDR, 2018)	1994 to present	• One or more human loss and/or • Loss of 1 or more US dollars	National governments and sectoral ministries	Publicly available at www.desinventar.net	Zoning level entry (country, districts, etc.)	Reports as a flood. Reports various categories of loss and damage (see Table S1). Includes qualitative information
Emergency Events Database (EM-DAT, 2020)	1995 to present	• Ten (10) or more people killed • Hundred (100) or more people affected • Declaration of a state of emergency	UN agencies, National Governments, International Federation of Red Cross and Red Crescent Societies (IFRC), and NGOs	Publicly available at https://public.emdat.be/	National and sub-national level	Subtypes of floods and origin, several categories of loss and damage (Table S1)
Global hazards weekly bulletin (PHE, 2015)	2013 to present	• Selected news on floods	Media reports (Flood list)	Publicly available via email bulletins to subscribers and archived independently at http://www.met.reading.ac.uk/sgs02rpa/extreme.html	Provides information specific to the location affected	Reports the location and categories depending on impacts (deaths or displaced)
Local sources (URCS and KRCS)	2000–2018	Depending on the main source	Disaster Relief Emergency Fund (DREF) reports, relief web, flood list, and districts offices	Request sent to National societies to support this research	Disaster prone areas	All the above

17533318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Type I and Type II error equations for the comparative analysis.

Index name	Equation	Score range	Perfect score
Type I error (TI)	$\frac{\text{Number of gauge observed flood events with no impacts reported}}{\text{Total number of flood events from river gauge observations}}$	0–1	0
Type II error (TII)	$\frac{\text{Number of flood events detected by impact data, with no gauge observations}}{\text{Total number of flood events from impact data}}$	0–1	0

TABLE 3 Skill scores used for the verification of forecasts.

Skill score	Equation	Values range	Perfect score
probability of detection (POD)	$POD = \frac{H}{H+M}$	0–1	1
false alarm ratio (FAR)	$FAR = \frac{FA}{H+FA}$	0–1	0

Abbreviations: FA, false alarms; H, Hits; M, Misses.

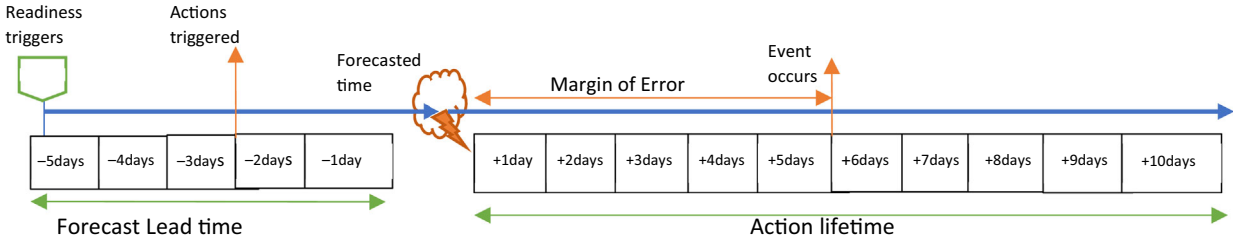

FIGURE 2 Visual representation of the action lifetime and margin of error based on early warning and action triggers.

The second alternative would be if actions should be taken based on a forecast that shows a high likelihood of event occurrence to minimise the expenses that would be incurred if the actions turn out to be in vain (see Lopez et al., 2020).

Various factors identified from the EAPs developed by URCS and KRCS have been adopted in this study. First, a flood forecast with a 60% chance of happening triggers early actions. Hence, we consider forecasts that indicate a forecast probability of 60% and above to correspond to a flood forecast event, and below 60% to correspond to a 'no-flood' event. Second, in the calculations of events correctly forecasted, an action lifetime of 7 days, is considered. 'Action lifetime' is defined as the length of time during which action will remain effective in reducing impacts (Coughlan De Perez et al., 2016). In forecast verification, the action lifetime is commonly known as the 'margin of error' and it is used to give more tolerance to the forecasts such that even if the forecast is late but materialises within the duration of the action lifetime, the actions will still be considered successful. For example, if an action is taken and a flood occurs up to 7 days after the forecasted date, this will still be considered a 'hit' if the action lifetime is >7 days (see Figure 2 for a visual description of the action lifetime and margin of error). Depending on the type of action, the action lifetime can range from 7 to 90 days. This can also vary depending on a specific country's flexibility on the actions to take and the acceptable number of times the stakeholders are willing to 'act in vain'. For Uganda and Kenya, the stakeholders set the probability of 'action in vain' to 50%, indicated using the FAR. From Figure 2, We have considered a margin of error of 5 days and action lifetime of 10 days. However, these parameters can still vary depending on the type of action.

Using distinct flood discharge thresholds (i.e., 90th and 95th percentile) calculated from the GloFAS reforecast data and river-gauge observations, we verify flood forecast using river-gauge observations and impact data as a reference. This study was therefore not meant to evaluate the hydrological performance of GloFAS (calibration and validation of GloFAS time series) but to assess the usefulness of the two reference datasets in forecasts verification. Using a 7 days-action lifetime and 60% probability of flooding, we compute the differences in the skill scores (POD and FAR) for forecast-observed data and forecast-impact data pairs, respectively. Here, if the

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

difference between the 'POD observed' and 'POD impact' is negative and the FAR difference is positive, impact data are more favourable in skill assessment than river-gauge observations and vice versa. Additionally, if the river- gauge observations or impact data (or both) report a flood event for the same days as in the GloFAS flood forecast (within the action lifetime of 7 days from the warning), the reference data (observed or impact) are favourable in skill assessments.

4 | RESULTS

4.1 | Impact data from the data repositories

In Uganda, in two districts (Katakwi and Manafwa) the reported impacts from the data repositories show a higher number of events reported in 2007, 2010, 2011, 2012, and 2018 from DI and DFO as compared with other years. However, the flood events for Mayanja from all the data repositories across the years are low. Table 4 shows the number of events across Uganda and the three locations from 2007 to 2018. The number of flood events from each repository presented in Table 4 is independent, that is, it does not consider any overlap across the repositories.

The analysis of the number of flood events from multiple and single data repositories shows that in Katakwi there are 434 flood events where DI recorded the highest number of events at 36%, followed by DFO at 19% (Table 5). Data collected across Katakwi by URCS also provide a substantial contribution (14%) to the flood events in the area. The overlap from multiple data repositories (EM-DAT, DI, and DFO) contributes to 11% of the total flood events. In Manafwa from a total of 304 events, the highest number of events are from single source DI and overlap between EM-DAT and DFO, at 33% and 28% respectively. EM-DAT alone contributes 14% of the total events. In Mayanja, only two data repositories contribute to the flood events. These are the DI at 23% and EM-DAT at 77% totalling 102 events.

In Kenya, many flood events were reported in 2007, 2010, 2011, 2013, 2015, and 2018 across the country and the three study locations (see Table 4). EM-DAT also records the highest number of flood events across the three locations contrasting with findings in Uganda, whereas DI reported the lowest. For example, in Nzoia EM-DAT represents 69% of the total flood events, local sources contribute 12%, whereas DI covers 6%. The overlaps between the various sources contribute marginally across the locations. For example, EM-DAT and DI together contribute <1% in Tana, 3% in Nzoia, and 1% in Athi (Table 5).

4.2 | How adequate are the impact data in supplementing river-gauge observations in identifying flood events?

The comparative analysis in the three locations in Uganda using combined impact data from the various data repositories and observed gauge data show varied results across locations and thresholds. For example, in Katakwi (Figure 3a) by using the 90th percentile from the river-gauge observations, the impact data capture 60% of all gauged flood events, but 42% of the reported flood events from the impact data do not correspond to flows above the 90th percentile threshold. This could mean that either the threshold is too high, with lower flows still causing impacts or the impacts reported were a result of another form of flooding like flash floods or waterlogging from heavy rainfall. In Manafwa and Mayanja (Figure 3b,c), Type I and Type II errors across the thresholds are high (above 0.5) which could mean that the quality and quantity of available impact data for these locations were not adequate (Type I), and the impacts reported were not as a result of riverine flooding (Type II).

The comparative analysis shows a high Type I error across the 90th and 95th percentile in the Kenyan locations. This means that though the observations indicate flood events, there were no impact data to correspond to these events or the quality of the available impact data was not good enough. On the other hand, the Type II error is also high across the locations, suggesting that impacts reported resulted from different forms of flooding, such as flash floods. For example, in Tana at the 90th percentile, impact data capture only 40% of all gauged flood events, but half of the reported flood events do not correspond to flows above the 90th percentile. Figure 4a–c shows the comparative analysis across the thresholds for Tana, Nzoia, and Athi respectively.

The analysis using a single data repository shows an increase in Type I error in all the locations in Kenya and Uganda (Figure 5a,b). For example, in Katakwi using DI alone results in a Type I error of 0.59 as compared to a Type I error of 0.39 while using four data repositories (DI, EM-DAT, local, and DFO). In Tana, EM-DAT results in a Type I error of 0.79 as compared to 0.61 while using data from all the repositories. Type II error fluctuates across the locations (Figure 5c,d). For example, at the 90th percentile, despite Nzoia having almost the same number of flood events from EM-DAT and local sources, Type II error is higher while using local sources as compared to using EM-DAT (Figure 5d). This shows that at the same (higher) threshold for example at (90th percentile) more events are likely to be missed out (events falling below the threshold) from the local source which takes into consideration more localised events as compared to a high-impact data repository like EM-DAT. In

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

CIWEM Chartered Institution of Water and Environmental Management | Journal of Flood Risk Management—WILEY

TABLE 4 Number of flood events from 2007 to 2018 for Uganda and Kenya locations derived from various data repositories.

Year	Kenya EM-DAT	Tana EM-DAT	Nzoia EM-DAT	Athi EM-DAT	Kenya DI	Tana DI	Nzoia DI	Athi DI	Kenya local	Tana local	Nzoia local	Athi local
2007	110	0	82	32	0	0	0	0	3	0	1	1
2008	57	8	16	20	6	2	15	2	15	3	4	2
2009	15	15	0	15	11	1	2	0	13	1	3	3
2010	97	87	17	87	35	3	1	4	53	8	7	3
2011	50	25	24	0	5	0	1	2	12	1	4	2
2012	27	27	27	27	13	2	3	1	39	5	3	4
2013	60	0	52	1	30	4	9	1	39	4	7	3
2014	0	0	0	0	0	0	0	0	20	1	3	0
2015	45	20	20	0	26	4	5	4	72	6	9	10
2016	11	6	6	0	2	0	0	0	29	2	3	6
2017	9	9	9	9	0	0	0	0	26	4	1	4
2018	79	79	0	0	0	0	0	0	144	61	17	28
Year	**Uganda EM-DAT**	**Katakwi EM-DAT**	**Manafwa EM-DAT**	**Mayanja EM-DAT**	**Uganda DI**	**Katakwi DI**	**Manafwa DI**	**Mayanja DI**	**Uganda DFO**	**Katakwi DFO**	**Manafwa DFO**	**Mayanja DFO**
2007	82	78	78	77	91	62	41	0	82	78	78	0
2008	1	1	0	1	12	1	1	0	0	0	0	0
2009	0	0	0	0	16	0	3	2	0	5	5	0
2010	5	5	5	0	109	48	40	2	11	5	5	0
2011	21	0	20	0	83	45	47	2	41	30	29	0
2012	1	0	1	0	49	37	3	1	29	2	0	0
2013	5	0	0	0	50	28	4	8	2	0	0	0
2014	0	0	0	0	33	1	2	1	5	0	0	0
2015	0	0	0	0	25	0	1	1	0	0	0	0
2016	7	0	0	0	16	0	2	3	7	0	0	0
2017	8	0	0	0	32	1	11	1	8	3	3	0
2018	1	0	1	0	32	28	10	2	17	8	8	0

Abbreviations: DFO, Dartmouth Flood Observatory; DI, DesInventar; EM-DAT, Emergency Events Database.

17533318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 5 Percent of the total number of flood events from multiple (overlaps) and single source data repositories for the study locations in Uganda and Kenya.

Uganda	**Katakwi**	**Manafwa**	**Mayanja**
Number of reports	**434**	**304**	**102**
Sources	**Percent of total events in each location**		
Single source contribution			
DI	**36.41**	**32.57**	**22.55**
EM-DAT	1.38	13.82	**77.45**
DFO	**18.89**	12.5	0
Local sources (URCS)	13.59	0	0
GWHB	0.00	2.30	0
Overlaps			
EM-DAT, DI, and DFO	11.06	4.28	0
EM-DAT and DFO	6.91	**28.29**	0
DI and DFO	8.29	5.59	0
URCS and DI	3.46	0.00	0
EM-DAT and GWHB	0.00	0.66	0
Kenya	**Nzoia**	**Tana**	**Athi**
Number of reports	**316**	**359**	**251**
Sources			
Single source contribution			
EM-DAT	**69.94**	**70.75**	**72.11**
DI	6.33	3.34	3.19
Local sources	**12.03**	**19.22**	**19.92**
Overlaps			
EM-DAT and DI	3.48	0.56	1.20
EM-DAT and local	6.01	5.85	3.19
EM-DAT, DI, and local	2.22	0.28	0.40

Note: The first two sources that represent the highest percentage over each district/county are highlighted in bold.
Abbreviations: DFO, Dartmouth Flood Observatory; DI, DesInventar; EM-DAT, Emergency Events Database.

other words, a data repository that considers a low threshold for inclusion of the event in their database may require a low threshold based on gauge observation to correctly identify the flood events as compared to a data repository that considers high threshold for inclusion.

4.3 | Where river-gauge observations are limited or unavailable, how best can the impact data be used to verify flood forecasts and ensure anticipatory actions are informed?

We plotted the difference between the forecast skill scores (POD and FAR) obtained using the river-gauge observations and impact data (i.e., $POD_{observed} - POD_{impact}$ and $FAR_{observed} - FAR_{impact}$) as a reference for verifying flood forecasts across all the locations and two percentile thresholds to assess their potential in forecast verification (Figure 6). The results show that impact data gives a more favourable assessment of skill as compared to the observed data at the 90th and 95th percentile across lead times in Katakwi (i.e., $POD_{impact} > POD_{observed}$ and $FAR_{impact} < FAR_{observed}$). For other locations at a lead time of up to 15 days, the impact data underestimate the GloFAS skill in terms of POD and FAR. At longer lead times (>15 days), Nzoia shows a good assessment of skill in terms of POD. These outcomes can be associated with the quantity and quality of the impact data that were available for most locations (except Katakwi and partly Nzoia) which also corresponds to the findings in Section 4.2. The highest difference in the POD of up to 0.4 is seen in Mayanja at the 90th percentile while other locations show a difference of below 0.2. The FAR is however spread out

17533318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

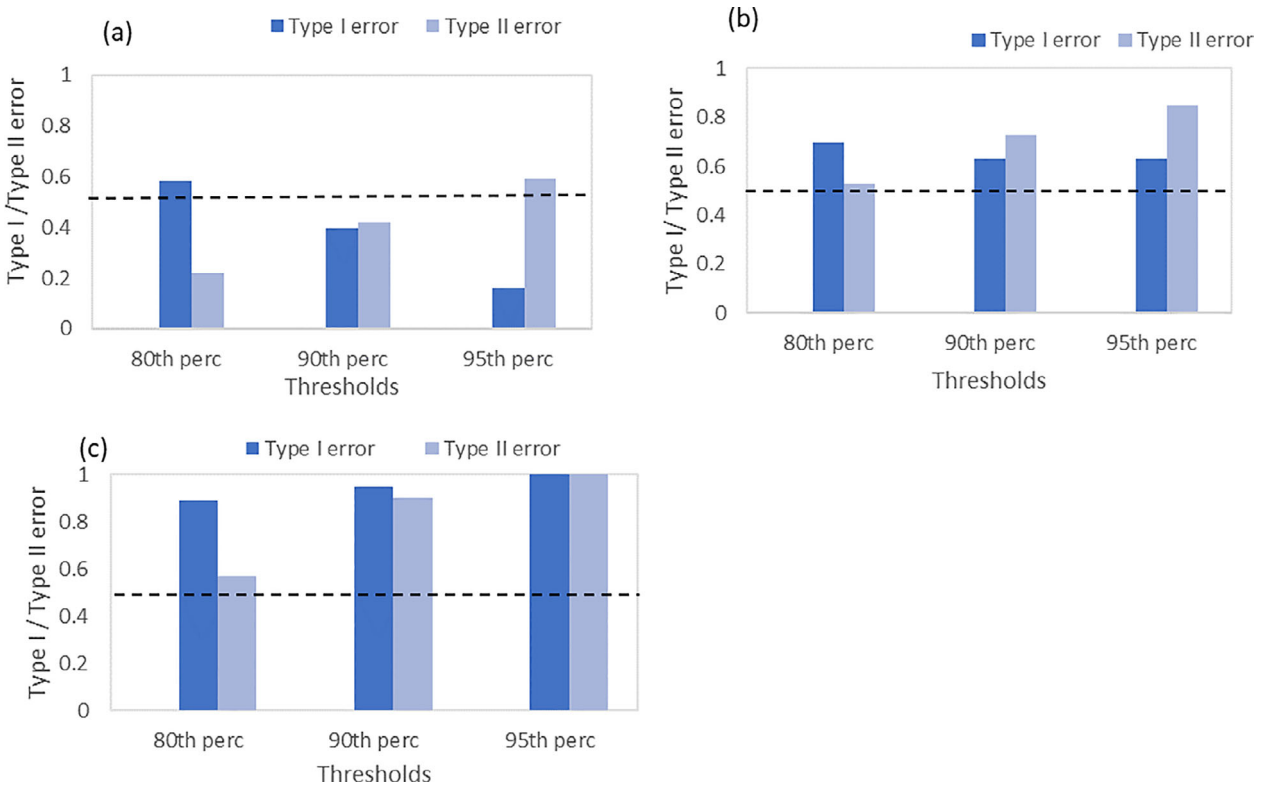

FIGURE 3 Comparative analysis of the impacts (all sources) and observed data at three percentile thresholds (80th, 90th, and 95th) of daily river flows from the gauged stations for (a) Katakwi, (b) Manafwa, and (c) Mayanja in Uganda.

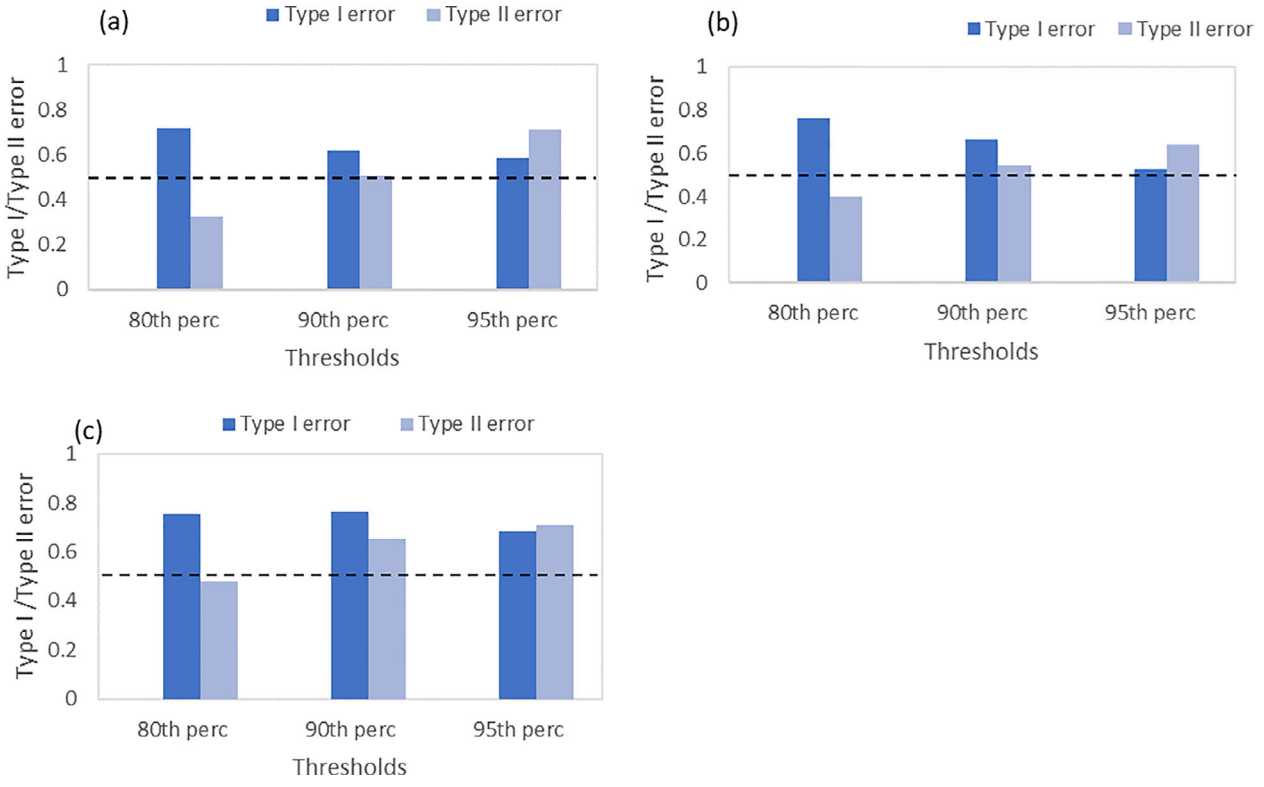

FIGURE 4 Comparative analysis of the impacts and observed data at three percentile thresholds (80th, 90th, and 95th) of daily river flows from the gauged stations for (a) Tana, (b) Nzoia, and (c) Athi in Kenya.

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

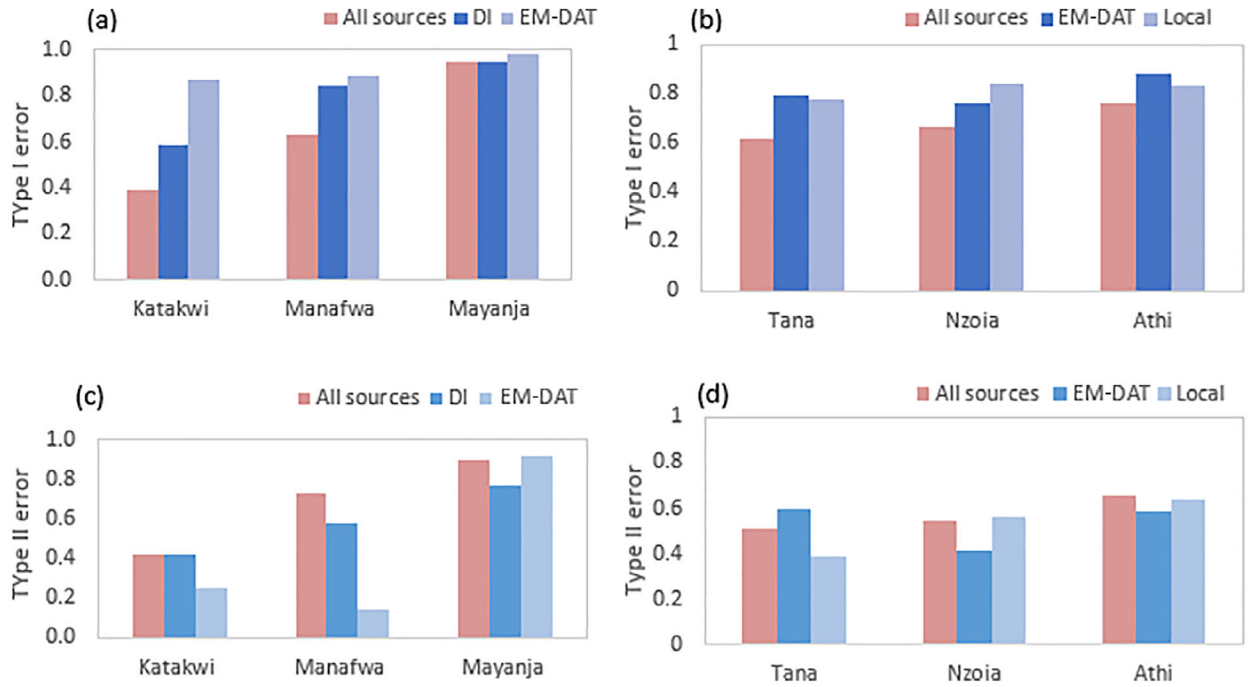

FIGURE 5 Type I and Type II error at 90th percentile for all data repositories (including overlaps) and single source data repositories for (a) Type I in Uganda locations, (b) Type I in Kenya locations, (c) Type II Uganda locations, and (d) Type II Kenya locations. DI, DesInventar; EM-DAT, Emergency Events Database.

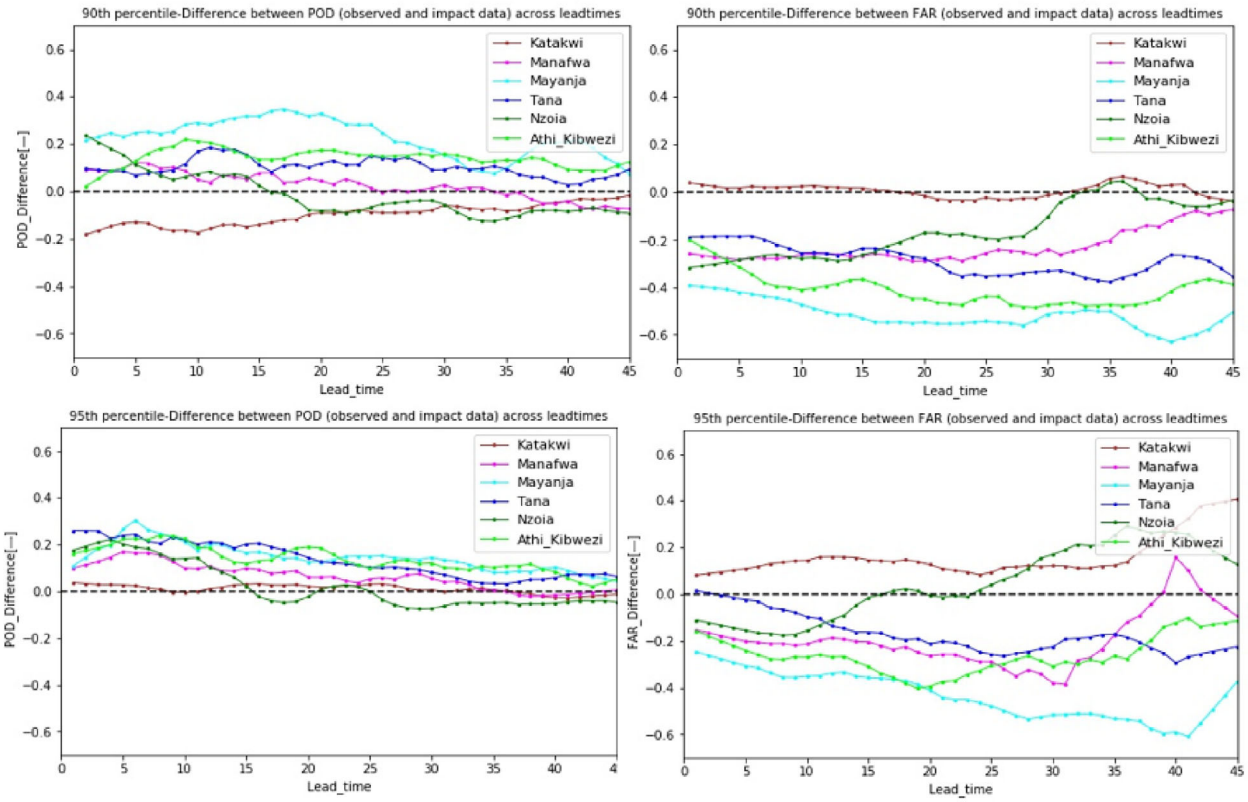

FIGURE 6 Differences in POD and FAR for locations in Uganda (Katakwi, Manafwa, and Mayanja) and Kenya (Tana, Nzoia, and Athi) across lead times at the 90th and 95th percentiles. FAR, false alarm ratio; POD, probability of detection.

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

across locations with a change of about 0.5 in Mayanja and Athi. POD and FAR graphs for the study locations at 90th and 95th percentile using river-gauge observations and impact data are provided in Figure S1.

5 | DISCUSSION

Using less conventional data such as impact data in forecast verification are gaining interest among researchers and practitioners. However, these data sources, just like hydro-meteorological data, are subject to errors and biases (Wilby et al., 2017). Despite these shortcomings, the impact data have the potential to ensure early warning systems are robust. In this section, we discuss the findings and implications of using impact data to verify flood forecasts and the assumptions that have been considered. First, we discuss the available impact data in the East African countries (Uganda and Kenya). Second, we highlight the adequacy of the impact data compared with river-gauge observations and how that may influence forecast verification. Last, we highlight the potential and challenges of using impact data to verify forecast information in data-scarce regions and provide recommendations that can be useful in improving the impact data to ensure effective early actions.

5.1 | What does the available impact data from Uganda and Kenya tell us?

Among the four main data repositories used in this study, DI had the highest number of flood events in Uganda (Katakwi and Manafwa districts), whereas across Kenya and the three counties, EM-DAT reports the highest number of flood events (Table 4). The differences can be associated with the criteria used for the inclusion of impact data in these repositories as well as the country-specific regulations on the collection and systematic reporting of impact data (Osuteye et al., 2017). For example, in Katakwi, if we consider a specific period from 1 August 2007 to 31 October 2007, EM-DAT reported a total of 11 flood events while DI reported 9 flood events (all considering the 7-day window, Section 3.1). Among the events, seven flood events overlap across the repositories while EM-DAT has four distinct events and DI has two distinct events. Therefore, using DI alone will result in fewer (−4) flood events while using EM-DAT alone will result in fewer (−2) flood events. This is just one example and the differences in flood events across data repositories might increase or decrease. Due to such differences, using only one repository can lead to a bias in the outputs generated (e.g., underestimation of event frequency).

Although we disaggregated the impact data into districts and counties, we only used the qualitative information classified as impact/no impact to guide the analysis. This is because there are no direct quantitative loss estimates available for these locations useful in understanding the severity of each flood event. Quantitative estimates are usually reported as aggregated quantities across a region, rather than disaggregated quantities for smaller geographical areas within the region (Gall, 2015). For example, in EM-DAT, the 2007 flooding between August and October that impacted different parts of Uganda are combined as one record (Disaster number 2007-0408; EM-DAT, 2020) with the quantified impact on, for example, the 'number of people affected', also aggregated. The insufficient reporting of quantitative estimates in areas of small spatial coverage can limit the analysis and affect the robustness of any conclusion, especially from a livelihood perspective (Osuteye et al., 2017). In addition, these repositories have differences in the parameters used for reporting. For example, EM-DAT reports only one parameter of 'number of people affected', whereas DI reports the same using two parameters; 'directly affected and indirectly affected'. As also noted in Below et al. (2010), this hinders the direct quantitative comparison between the two data repositories.

5.2 | How adequate are the impact data in identifying thresholds for impactful river flooding and in verifying flood forecasts?

Setting up early warning mechanisms for floods often depends on the thresholds derived from river-gauge data to identify the level at which the river discharge may result in impactful flooding. In data-scarce regions, impact data can help to determine such thresholds (Coughlan De Perez et al., 2016) but this requires a large number of good quality impact data to reduce the chances of over-representation/under-representation of impacts (Ranger et al., 2011). We have found that even within the same country impact data are not consistently available across all locations (Barabadi & Ayele, 2018), which may lead to bias in the outputs. Our analysis shows that using more than one source of impact data reduces the chances of a Type I error or situation where flooding occurs but impact data are not available. For example, although EM-DAT contributes to over 69% of all impact reported in Tana, Nzoia, and Athi respectively, using this repository alone results in an increase in Type I error (flood observed in gauged data but not reported) compared with using all three repositories (EM-DAT, DI, Local; Figure 5b). This can be associated with the inclusion criteria for the various data repositories. For example, for a repository like EM-DAT, only high-impact flood events are represented leaving out low-impact flood events.

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

We have found that the consistency between impact data and river-gauge data varies markedly across the thresholds, but the variability is location-dependent. For example, in Katakwi, there is good correspondence between the river-gauge observations and impact data at the 90th percentile. This suggests impact data can be used to identify river discharge critical thresholds at which impactful flooding occurs. These findings are consistent with scientific literature where impact data have been successfully used to define flood thresholds. For example, Young et al. (2021) used impact reports to determine the rainfall thresholds that resulted in flooding in the urban city of Alexandria, Egypt.

Although we used the percentile-based method to identify flood events, we acknowledge that high-impact events are generally higher than the 99th percentile (MacLeod et al., 2021), but to ensure robustness of the statistical analysis, we adopted the 90th and 95th percentile thresholds as several previous authors did (e.g., Arnal et al., 2018; MacLeod et al., 2021). These percentiles may include low-impact flood events that are likely to affect local limited areas (with relatively high frequency, e.g., 5% of days over a year for the 95th percentile) but are useful in cases where impact data is used in the verification due to the differences in the inclusion criteria of flood events in the various data repositories (see Table 1). In some previous studies, even lower thresholds are used because of data availability limitations, to ensure robustness in the verification. For example, Arnal et al. (2018) used terciles (33rd and 66th percentiles) of the simulated streamflow for the verification of seasonal streamflow forecasts and discussed the need to consider high thresholds such as the 95th percentile if more data were available. We therefore recommend that further studies with possible longer data periods available, should look at the representativeness of results across flood thresholds higher than the 99th percentile.

Other locations in Uganda and Kenya show an increase in Type I (and Type II) error as the river flow threshold decreases (increases). The increase in Type I error can be related to the inadequacy or the low quality of impact data used in this analysis, i.e. for both inadequate impact data (if the repository did not include an observed event) and low-quality data (if the timestamp of the impact data is incorrect) a false positive is produced. Type II error could have resulted if impacts reported were not because of riverine flooding but other subtypes of flooding, and this can also be influenced by the inclusion criteria which are specific to each data repository. Although a repository like EM-DAT differentiates floods using subtypes such as riverine and flash flooding, DI does not include such subtypes. These subtypes would help ensure that flood events are further categorised before analysis to reduce the Type II error. In addition, such differentiation can help in designing appropriate preparedness and response interventions which vary based on the sub-type of flooding (Nauman et al., 2021; Paprotny et al., 2021). To further confirm the source of increase in Type II error, data derived from satellite imagery (e.g., Sentinel-1 and Sentinel-2) could be used to identify if floods occurred as well as their spatial location (with respect to rivers), which can help discriminate riverine floods (Tarpanelli et al., 2022).

The differences in POD and FAR vary across the study locations considered here. Except in Katakwi and partly in Nzoia (>15 days lead time), where we get a more favourable assessment of skill while using impact data, other locations show that using impact data underestimate the GloFAS skill both in terms of POD and FAR. Though the differences are minimal in the majority of the locations, it still means that impact data cannot be adequately used to verify flood forecasts in most locations, as highlighted previously by Gall (2015). However, the available river-gauge observations and impact data could be used to train the hydrological model used in the GloFAS system through calibration and validation in specific locations that show poor detection of flood events. In other words, the available historical impact data and gauge observations can be used to assess the hydrological skill of the GloFAS using scores such as Nash-Sutcliffe efficiency which assesses temporal variability and agreement between the modelled and observed data (see Teule et al., 2020). Overall, being aware of uncertainties that can result in using the available impact data can help ensure the outputs are used appropriately in supporting anticipatory actions.

5.3 | How best can the impact data be used to verify flood forecasts in data-scarce regions?

Our exploratory analysis has highlighted several factors that are affecting the efficacy of impact data for verifying flood forecasts in most of the study locations in Uganda and Kenya. These are inadequacy of events records, poor quality, and spatial resolution/granularity among others. Therefore, using impact data may result in underestimation of forecast skill, leading to reduced confidence in using the forecast to support anticipatory actions. In other words, if we use impact data to verify and it turns out to be unwittingly underestimating the forecast skill, we might discard a forecast that is good enough to support preparedness actions for vulnerable people. Nevertheless, positive results obtained for Katakwi in Uganda and Nzoia in Kenya show that with some improvements, the impact data could be used to determine critical thresholds for flooding and inform the design of early

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

warning mechanisms in data-scarce regions. For such regions, the following improvements would increase the usability of impact data.

5.3.1 | Characterising the gaps/uncertainties

The uncertainties in the impact data should be explicitly stated, as well as the implications for the outputs, especially if the outputs are intended to inform actions. The uncertainty around the estimate can be denoted using standard error, which indicates how far the estimate is from the mean. The standard error can be calculated by dividing the standard deviation by the square root of the sample size (Walker, 2018). From our analysis, the standard error in the FAR calculation varies from 0.02 to 0.05. Therefore, if the recommended forecast FAR to trigger humanitarian action is <0.5, using impact data will require a FAR of <0.4 to minimise actions taken in vain. Continuous operational evaluation of the forecasts is also required in situations where real-time reference data are available.

5.3.2 | Combining databases

A combination of impact data from multiple data repositories should be explored especially if the data is scarce (Barabadi & Ayele, 2018). This can help reduce the biases and possibility of missed events in the reference datasets for forecast verification, because of the differences in the methods and criteria used in the compilation of the various data repositories. For example, comparing river-gauge observations with impact data from all repositories against EM-DAT in Tana resulted in an improvement of the Type I error from 0.8 to 0.6 (Figure 5b). However, the combination should be carefully explored to avoid duplication of entries, especially from repositories fed from the same primary source or if there is a slight difference in the timestamp for the same event. Some of these challenges of replication can be handled by using a tolerance interval such that entries that are within a certain interval are considered one event. In this study, an interval of 7 days was used.

The combination should also consider the differences in the indicators used in each repository. For example, EM-DAT reports the 'number of people affected' as one indicator while DI reports the same in two separate indicators (i.e., 'directly and indirectly affected'). In addition, EM-DAT makes clear differentiations of the disaster type and subtypes, such as riverine flood and flash flooding, whereas DI does not have such differentiation. Such differences make it challenging to combine and compare the data and disaggregate further, for instance, if you want to monitor only a subtype of the disaster. For example, in our analysis, most Type II errors could have resulted from impact data that were not necessarily from riverine flooding.

Harmonising and differentiating these parameters and clarifying their meanings would help minimise these difficulties (Below et al., 2010). This can be done by ensuring that these subtypes are indicated during the data collection process or by applying index-based approaches to differentiate between the various disaster sub-types (see Kruczkiewicz, Bucherie, et al., 2021). In addition, satellite data (e.g., from Sentinel-1 and Sentinel-2) can be used alongside the impacts reports to identify the nature and extent of flooding as well as the spatial location which can help in complementing the impact reports for future applications in forecast verification. The usefulness of satellite images in assessing flood event types and extent has already been demonstrated in several recent studies, although also these datasets have their own current limitations that should be taken into account (see Landuyt et al., 2019; Notti et al., 2018; Tarpanelli et al., 2022).

5.3.3 | Harmonising primary data collection and information management processes

Primary data collection process

primary data collection in most countries is done through normal government procedures. This is mainly done using the damage and needs assessment approach at the local level and the collected data analysed at the national level (see The International Bank for Reconstruction and Development & The World Bank, 2010). If the collected information show that impacts are considerable, the country may decide to seek external support. In this case, the United Nations Office for Coordination of Humanitarian Affairs (UN-OCHA) may coordinate more rapid needs assessments to collect more information using approaches such as the Multi-sector Initial Rapid Assessment (MIRA) framework (Inter-Agency Standing Committee, 2015). Countries can, however, use their own guidelines for collecting the data. In Uganda, the Office of the Prime Minister is tasked with the collection and uploading of impact data in the DI. However, recent interviews in Uganda noted that rapid response assessments and collection of impact data are carried out by various institutions, including the Office of the Prime Minister, the Uganda Red Cross Society, the Humanitarian Open Street mapping team, local NGOs, and the district office, among others (personal communication, October 2020). There is a need to harmonise the data collection process through clear guidelines and dedicated institutions to avoid the probability of competing reports of unknown credibility (Guha-Sapir & Below, 2006).

Furthermore, impact reporting can benefit from improved weather and river-gauge networks. Improving

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

gauge networks can be strategized such that it is done alongside the improvement on impact data collection (Baddour & Douris, 2018). This can ensure improvement in the flood forecasting systems by providing key inputs for hydrological model calibration and forecast verification, as well as for further impact reports verification.

Information management process

impact data collected through primary sources such as in-country institutions are often uploaded to data repositories such as DI. Due to a lack of resources, most countries might not be uploading the collected information regularly. Therefore, the impact data collected are held in internal disaster management systems and managed by the primary institutions. National data repositories could be explored to ensure that all impact data collected in-country is stored in a central in-country repository for ease of accessibility.

5.3.4 | Impact data outside the official public sources

A broader and more accurate collection of temporal and geospatial data on disaster occurrence would ensure improved risk estimations (Bakkensen et al., 2018). An extended search of impact data available at the in-country archives, for example, in private institutions, and insurance companies, but not yet available in the open repositories would therefore help improve the quantity and detail level (spatial–temporal data) of the available impact data. For example, a study by Smith and Katz (2013) shows that a significant under-reporting of disaster loss estimates can occur due to reliance on only public sources because of their ease of accessibility.

5.3.5 | Use of new technologies

New technologies such as artificial intelligence can be used to expand impact data (van den Homberg et al., 2018). Initiatives to expand the impact data, for example, through web scraping, text mining (Margutti & van den Homberg, 2020), and application of earth observation data (Kruczkiewicz, McClain, et al., 2021; Nauman et al., 2021) and social media platforms should be explored. For example, social media platforms like Google Trends and Twitter have shown promising results in the detection and reporting of flood events (de Bruijn et al., 2019; Rossi et al., 2018; Thompson et al., 2022). In addition, an ongoing study by van den Homberg et al. (2022) has shown that flood impact data generated from news articles can complement data from global repositories such as DI both geographically and temporally, improving the usefulness of the data. Ensuring that any new data are interoperable with data from these repositories will require clear technical guidelines and protocols (Wirtz et al., 2014) such as the WMO data standardisation initiative (see Baddour & Douris, 2018).

Overall, impact data represent an important source of less conventional data for monitoring and improving early warning and preparedness actions. There is also great potential for improving these data quantity and quality through strengthening in-country disaster monitoring capabilities and ensuring standardised process of data collection that captures all the relevant data features such as flood extent, gauge level, contact information among others are in place(Integrated Research on disaster risk, 2014).

6 | CONCLUSION

As the world faces an uncertain future due to climate variability, environmental, and climate change, and an increase in extreme hydrometeorological events, investing in early warning early action mechanisms can be an effective way to prepare and adapt to these extreme events. However, such an investment will require understanding how forecast information performs in detecting these extreme events to ensure that anticipatory actions are not taken in vain. While forecast verification has been successful in regions where long-term hydrometeorological observations are available, this is very challenging in data-scarce regions.

Verification of forecasts using non-traditional approaches that use less conventional data would ensure the development of these mechanisms even in locations with scarce/no conventional observations. In this study, we investigated the usefulness of flood impact data to verify flood forecasts. Our findings show that although existing impact data have shortcomings, they also have the potential for flood event analysis and forecast verification and can be used in regions with no long-term hydrometeorological observations. These impact data may, however, require improvement to enhance their utility and make the forecast verification more acceptable and reliable. Among the recommendations outlined above, supporting the national institutions to streamline impact data collection, and expanding impact data using new technologies are of critical importance. Addressing these issues will, however, require a recognition of the role that impact data can play in verifying hydrometeorological forecasts and in identifying trends in extreme events to inform risk management. In addition, a collaborative effort among international humanitarian actors, disaster management institutions, the private sector, and local communities is needed to ensure that quality impact data are collected consistently and made available in near real-time.

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ACKNOWLEDGEMENTS

This research was funded by the UK's Natural Environment Research Council (NERC) Science for Humanitarian Emergencies and Resilience Studentship Cohort (SHEAR SSC) grant number: NE/R007799/1, through the SHEAR FATHUM project (Forecasts for AnTicipatory HUManitarian action, grant number: NE/P000525/1) and the SHEAR catalyst NIMFRU project (National-scale IMpact-based forecasting of Flood Risk in Uganda, grant number: NE/S00596X/1). We appreciate the Department of Water Resources Management (DWRM) in Uganda and the Water Resources Authority (WRA) in Kenya for providing river-gauge data that was used as one of the reference data in our analysis. We are also grateful to the World Meteorological Organisation (WMO) Joint Working Group on Forecast Verification Research (JWGFVR) judging panel for selecting this research as the winner of the second verification challenge for the best new verification metric making use of non-traditional observations.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

DATA AVAILABILITY STATEMENT

The GloFAS v3.1 reforecast are available from Copernicus Climate Change Service- Climate Data Store (https://cds.climate.copernicus.eu/). Impacts reports from DesInventar and EM-DAT are freely available from their respective web pages (https://www.desinventar.net/ and https://www.emdat.be/database). DFO reports are available upon request from the University of Colorado. The R-scripts used in forecast verification are available upon request from the authors.

ORCID

Faith Mitheu https://orcid.org/0000-0003-1655-449X

REFERENCES

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013). GloFAS-global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, *17*(3), 1161–1175. https://doi.org/10.5194/hess-17-1161-2013

Arnal, L., Cloke, H. L., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., Krzeminski, B., & Pappenberger, F. (2018). Skilful seasonal forecasts of streamflow over Europe? *Hydrology and Earth System Sciences*, *22*(4), 2057–2072. https://doi.org/10.5194/hess-22-2057-2018

Baddour, O., & Douris, J. (2018). WMO initiative to standardize data collection of weather, water, climate and space weather extreme events. https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.33/2018/mtg4/S4_5_WMO_initiative_extreme_events.pdf. Accessed August 3, 2022.

Bakkensen, L. A., Shi, X., & Zurita, B. D. (2018). The impact of disaster data on estimating damage determinants and climate costs. *Economics of Disasters and Climate Change*, *2*(1), 49–71. https://doi.org/10.1007/s41885-017-0018-x

Barabadi, A., & Ayele, Y. Z. (2018). Post-disaster infrastructure recovery: Prediction of recovery rate using historical data. *Reliability Engineering and System Safety*, *169*, 209–223. https://doi.org/10.1016/j.ress.2017.08.018

Below, R., Vos, F., & Guha-sapir, D. (2010). Moving towards harmonization of disaster data : a study of six Asian databases, CRED working paper. http://hdl.handle.net/2078.1/113064.

Bernet, D. B., Prasuhn, V., & Weingartner, R. (2017). Surface water floods in Switzerland: What insurance claim records tell us about the damage in space and time. *Natural Hazards and Earth System Sciences*, *17*(9), 1659–1682. https://doi.org/10.5194/nhess-17-1659-2017

Brakenridge, G. R. (2015). *Global active archive of large flood events*. Dartmouth Flood Observatory, University of Colorado. http://floodobservatory.colorado.edu/Archives/index.html.

Cortès, M., Turco, M., Llasat-Botija, M., & Llasat, M. C. (2018). The relationship between precipitation and insurance data for floods in a Mediterranean region (Northeast Spain). *Natural Hazards and Earth System Sciences*, *18*(3), 857–868. https://doi.org/10.5194/nhess-18-857-2018

Coughlan De Perez, E., van den Hurk, B., van Aalst, M. K., Amuron, I., Bamanya, D., Hauser, T., Jongma, B., Lopez, A., Mason, S., Mendler de Suarez, J., Pappenberger, F., Rueth, A., Stephens, E., Suarez, P., Wagemaker, J., & Zsoter, E. (2016). Action-based flood forecasting for triggering humanitarian action. *Hydrology and Earth System Sciences*, *20*(9), 3549–3560. https://doi.org/10.5194/hess-20-3549-2016

Coughlan De Perez, E., van den Hurk, B., van Aalst, M. K., Jongman, B., Klose, T., & Suarez, P. (2015). Forecast-based financing: An approach for catalyzing humanitarian action based on extreme weather and climate forecasts. *Natural Hazards and Earth System Sciences*, *15*(4), 895–904. https://doi.org/10.5194/nhess-15-895-2015

de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., & Aerts, J. C. J. H. (2019). A global database of historic and real-time flood events based on social media. *Scientific Data*, *6*(1), 1–12. https://doi.org/10.1038/s41597-019-0326-9

EM-DAT. (2020). Database | EM-DAT, Centre for Research on the Epidemiology of Disasters. https://www.emdat.be/database.

Emerton, R., Cloke, H., Ficchi, A., Hawker, L., de Wit, S., Speight, L., Prudhomme, C., Rundell, P., West, R., Neal, J., Cuna, J., Harrigan, S., Titley, H., Magnusson, L., Pappenberger, F., Klingaman, N., & Stephens, E. (2020). Emergency flood bulletins for cyclones Idai and Kenneth: A critical evaluation of the use of global flood forecasts for international humanitarian preparedness and response. *International Journal of Disaster Risk Reduction*, *50*, 101811. https://doi.org/10.1016/j.ijdrr.2020.101811

Emerton, R., Zsoter, E., Arnal, L., Cloke, H. L., Muraro, D., Prudhomme, C., Stephens, E. M., Salamon, P., & Pappenberger, F. (2018). Developing a global operational seasonal hydro-meteorological forecasting system: GloFAS-seasonal v1.0. *Geoscientific Model Development*, *11*(8), 3327–3346. https://doi.org/10.5194/gmd-11-3327-2018

Ficchì, A., & Stephens, L. (2019). Climate variability alters flood timing across Africa. *Geophysical Research Letters*, *46*(15), 8809–8819. https://doi.org/10.1029/2019GL081988

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Flow, Excedance and Percentiles (2023). https://thames.me.uk/ThamesDataPercentiles.html.

Gall, M. (2015). The suitability of disaster loss databases to measure loss and damage from climate change. *International Journal of Global Warming*, *8*(2), 170–190. https://doi.org/10.1504/IJGW.2015.071966

Gall, M., Borden, K. A., & Cutter, S. L. (2009). When do losses count? Six fallacies of natural hazards loss data. *Bulletin of the American Meteorological Society*, *90*, 799–809. https://doi.org/10.1175/2008BAMS2721.1

Guha-Sapir, D., & Below, R. (2002). Quality and accuracy of disaster data: A comparative analyse of 3 global data sets. CRED working paper. http://hdl.handle.net/2078.1/179722.

Guha-Sapir, D., & Below, R. (2006). Collecting data on disasters: Easier said than done. *Asian Disaster Management News*, *12*(2), 9–10.

Hansen, J. W., Vaughan, C., Kagabo, D. M., Dinku, T., Carr, E. R., Körner, J., & Zougmoré, R. B. (2019). Climate services can support African farmers' context-specific adaptation needs at scale. *Frontiers in Sustainable Food Systems*, *3*(4), 1–16. https://doi.org/10.3389/fsufs.2019.00021

Harrigan, S., Zoster, E., Cloke, H., Salamon, P., & Prudhomme, C. (2023). Daily ensemble river discharge reforecasts and real-time forecasts from the operational Global Flood Awareness System. *Hydrology and Earth System Sciences Discussions*, *27*(1), 1–19. https://doi.org/10.5194/hess-2020-532

Hirpa, F. A., Salamon, P., Beck, H. E., Lorini, V., Alfieri, L., Zsoter, E., & Dadson, S. J. (2018). Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. *Journal of Hydrology*, *566*(3), 595–606. https://doi.org/10.1016/j.jhydrol.2018.09.052

Integrated Research on disaster risk. (2014). Peril Classification and Hazard Glossary (IRDR DATA Publication No. 1), Integrated Research on Disaster Risk, Beijing. [Preprint]. https://www.irdrinternational.org/knowledge_pool/publications/173.

Inter-Agency Standing Committee. (2015). *Multi-sector initial rapid assessment guidance* (pp. 1–32). *Inter-Agency Standing Committee*. https://interagencystandingcommittee.org/system/files/mira_2015_final.pdf.

KRCS. (2021). Riverine floods early action protocol summary. https://reliefweb.int/report/kenya/kenya-riverine-floods-early-action-protocol-summary-eap2021ke01.

Kruczkiewicz, A., Bucherie, A., Ayala, F., Hultquist, C., Vergara, H., Mason, S., Bazo, J., & de Sherbinin, A. (2021). Development of a flash flood confidence index from disaster reports and geophysical susceptibility. *Remote Sensing*, *13*(14), 2764. https://doi.org/10.3390/rs13142764

Kruczkiewicz, A., McClain, S., Bell, V., Warrick, O., Bazo, J., Mason, J., Vergara, H., & Horna, N. (2021). Chapter 11 - Earth observations for anticipatory action: Case studies in hydrometeorological hazards. In G. J.-P. Schumann (Ed.), *Earth observation for food applications* (pp. 237–253). Elsevier. https://doi.org/10.1016/B978-0-12-819412-6.00011-0

Landuyt, L., van Wesemael, A., Schumann, G. J. P., Hostache, R., Verhoest, N. E. C., & van Coillie, F. M. B. (2019). Flood mapping based on synthetic aperture radar: An assessment of established approaches. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(2), 722–739. https://doi.org/10.1109/TGRS.2018.2860054

Lopez, A., Coughlan de Perez, E., Bazo, J., Suarez, P., van den Hurk, B., & van Aalst, M. (2020). Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings. *Weather and Climate Extremes*, *27*(3), 100167. https://doi.org/10.1016/j.wace.2018.03.006

MacLeod, D., Easton-Calabria, E., de Perez, E. C., & Jaime, C. (2021). Verification of forecasts for extreme rainfall, tropical cyclones, flood and storm surge over Myanmar and The Philippines. *Weather and Climate Extremes*, *33*(4), 100325. https://doi.org/10.1016/j.wace.2021.100325

Margutti, J., & van den Homberg, M. (2020). Text mining of loss data, EGU, p. 20048. https://doi.org/10.5194/egusphere-egu2020-20048

Marsigli, C., Ebert, E., Ashrit, R., Casati, B., Chen, J., Coelho, C. A. S., Dorninger, M., Gilleland, E., Haiden, T., Landman, S., & Mittermaier, M. (2021). Review article: Observations for high-impact weather and their use in verification. *Natural Hazards and Earth System Sciences*, *21*(4), 1297–1312. https://doi.org/10.5194/nhess-21-1297-2021

Mitheu, F., Petty, C., Tarnavsky, E., Stephens, E., Ciampi, L., Butsatsa, J., & Cornforth, R. (2022). Identifying the barriers and opportunities in the provision and use of weather and climate information for flood risk preparedness: The case of Katakwi District, Uganda. *Frontiers in Climate-Climate Services*, *4–22*, 1–35. https://doi.org/10.3389/fclim.2022.908662

National River Flow Archive. (2023). Derived flow statistics. https://nrfa.ceh.ac.uk/derived-flow-statistics.

Nauman, C., Anderson, E., Coughlan de Perez, E., Kruczkiewicz, A., McClain, S., Markert, A., Griffin, R., & Suarez, P. (2021). Perspectives on flood forecast-ba sed early action and opportunities for earth observations. *Journal of Applied Remote Sensing*, *15*(3), 1–17. https://doi.org/10.1117/1.jrs.15.032002

Nidumolu, U., Lim-Camacho, L., Gaillard, E., Hayman, P., & Howden, M. (2020). Linking climate forecasts to rural livelihoods: Mapping decisions, information networks and value chains. *Weather and Climate Extremes*, *27*(6), 100174. https://doi.org/10.1016/j.wace.2018.06.001

Notti, D., Giordan, D., Caló, F., Pepe, A., Zucca, F., & Galve, J. (2018). Potential and limitations of open satellite data for flood mapping. *Remote Sensing*, *10*(11), 1673. https://doi.org/10.3390/rs10111673

Nyadzi, E., Werners, E. S., Biesbroek, R., Long, P. H., Franssen, W., & Ludwig, F. (2019). Verification of seasonal climate forecast toward hydroclimatic information needs of rice farmers in northern Ghana. *Weather, Climate, and Society*, *11*(1), 127–142. https://doi.org/10.1175/WCAS-D-17-0137.1

Ochieng, J., Kirimi, L., & Mathenge, M. (2016). Effects of climate variability and change on agricultural production: The case of small scale farmers in Kenya. *NJAS - Wageningen Journal of Life Sciences*, *77*(1), 71–78. https://doi.org/10.1016/j.njas.2016.03.005

Ogutu, G. E. O., Franssen, W. H. P., Supit, I., Omondi, P., & Hutjes, R. W. A. (2017). Skill of ECMWF system-4 ensemble seasonal climate forecasts for East Africa. *International Journal of Climatology*, *37*(5), 2734–2756. https://doi.org/10.1002/joc.4876

Osuteye, E., Johnson, C., & Brown, D. (2017). The data gap: An analysis of data availability on disaster losses in sub-Saharan African cities. *International Journal of Disaster Risk Reduction*, *26*, 24–33. https://doi.org/10.1016/j.ijdrr.2017.09.026

Paparrizos, S., Smolenaars, W., Gbangou, T., Slobbe, E., & Ludwig, F. (2020). Verification of weather and seasonal forecast

1753318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

information concerning the peri-urban farmers needs in the lower Ganges delta in Bangladesh. *Atmosphere*, *11*(10), 1041. https://doi.org/10.3390/atmos11101041

Paprotny, D., Kreibich, H., & Morales, O. (2021). *A probabilistic approach to estimating residential losses from different flood types, natural hazards*. Springer. https://doi.org/10.1007/s11069-020-04413-x

PHE. (2015). Global hazards weekly bulletin - 14th December to 18th December 2015. http://www.met.reading.ac.uk/~sgs02rpa/extreme15.html#150313.

Prudhomme, C., & Zsoter, E. (2021). GloFAS v3.1 hydrological performance - Copernicus Services - ECMWF Confluence Wiki. https://confluence.ecmwf.int/display/COPSRV/GloFAS+v3.1+hydrological+performance.

Ranger, N., Hallegatte, S., Bhattacharya, S., Bachu, M., Priya, S., Dhore, K., Rafique, F., Mathur, P., Naville, N., Henriet, F., Herweijer, C., Pohit, S., & Corfee-Morlot, J. (2011). An assessment of the potential impact of climate change on flood risk in Mumbai. *Climatic Change*, *104*(1), 139–167. https://doi.org/10.1007/s10584-010-9979-2

Rossi, C., Acerbo, F. S., Ylinen, K., Juga, I., Nurmi, P., Bosca, A., Tarasconi, F., Cristoforetti, M., & Alikadic, A. (2018). Early detection and information extraction for weather-induced floods using social media streams. *International Journal of Disaster Risk Reduction*, *30*(3), 145–157. https://doi.org/10.1016/j.ijdrr.2018.03.002

Salack, S., Sarr, B., Sangare, S. K., Ly, M., Sanda, I. S., & Kunstmann, H. (2015). Crop-climate ensemble scenarios to improve risk assessment and resilience in the semi-arid regions of West Africa. *Climate Research*, *65*, 107–121. https://doi.org/10.3354/cr01282

Smith, A. B., & Katz, R. W. (2013). US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Natural Hazards*, *67*(2), 387–410. https://doi.org/10.1007/s11069-013-0566-5

Tarpanelli, A., Mondini, A. C., & Camici, S. (2022). Effectiveness of Sentinel-1 and Sentinel-2 for flood detection assessment in Europe. *Natural Hazards and Earth System Sciences*, *22*(8), 2473–2489. https://doi.org/10.5194/nhess-22-2473-2022

Teule, T., Couasnon, A., Bischiniotis, K., Blasch, J., & van den Homberg, M. (2020). *Towards improving a national flood early warning system with global ensemble flood predictions and local knowledge*. Institute for Environmental Studies. https://presentations.copernicus.org/EGU2020/EGU2020-507_presentation.pdf

The International Bank for Reconstruction and Development and The World Bank. (2010). Damage, loss and needs assessment. https://reliefweb.int/sites/reliefweb.int/files/resources/TTL Vol1_-WEB.pdf.

Thompson, J. J., Wilby, R. L., Matthews, T., & Murphy, C. (2022). The utility of Google trends as a tool for evaluating flooding in data-scarce places. *Area*, *54*, 203–212. https://doi.org/10.1111/area.12719

UNISDR. (2018). United Nations DesInventar Open Source Initiative. https://www.desinventar.net/.

URCS. (2021). Uganda: Floods early action protocol summary. https://reliefweb.int/report/uganda/uganda-floods-early-action-protocol-summary-eap-number-eap2021ug01.

van den Homberg, M., Monné, R., & Spruit, M. (2018). Bridging the information gap of disaster responders by optimizing data selection using cost and quality. *Computers and Geosciences*, *120*(5), 60–72. https://doi.org/10.1016/j.cageo.2018.06.002

van den Homberg, M., Margutti, J., & Basar, E. (2022). Enriching impact data by mining digital media | PreventionWeb. https://www.preventionweb.net/publication/enriching-impact-data-mining-digital-media.

van der Knijff, J. M., Younis, J., & de Roo, A. P. J. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, *24*(2), 189–212. https://doi.org/10.1080/13658810802549154

Walker, J. A. (2018). Chapter 5 variability and uncertainty (standard deviations, standard errors, confidence intervals) | elements of statistical modeling for experimental biology. https://www.middleprofessor.com/files/applied-biostatistics_bookdown/_book/variability-and-uncertainty-standard-deviations-standard-errors-confidence-intervals.html.

Wilby, R. L., Clifford, N. J., de Luca, P., Harrigan, S., Hillier, J. K., Hodgkins, R., Johnson, M. F., Matthews, T. K. R., Murphy, C., Noone, S. J., Parry, S., Prudhomme, C., Rice, S. P., Slater, L. J., Smith, K. A., & Wood, P. J. (2017). The "dirty dozen" of freshwater science: Detecting then reconciling hydrological data biases and errors. *WIREs Water*, *4*(3), e1209. https://doi.org/10.1002/wat2.1209

Wilks. (2006). Chapter 7-forecast verification. In *Statistical methods in atmospheric sciences* (3rd ed., pp. 255–335). Academic Press.

Wirtz, A., Kron, W., Löw, P., & Steuer, M. (2014). The need for data: Natural disasters and the challenges of database management. *Natural Hazards*, *70*(1), 135–157. https://doi.org/10.1007/s11069-012-0312-4

Young, A., Bhattacharya, B., & Zevenbergen, C. (2021). A rainfall threshold-based approach to early warnings in urban data-scarce regions: A case study of pluvial flooding in Alexandria, Egypt. *Journal of Flood Risk Management*, *14*(2), 1–16. https://doi.org/10.1111/jfr3.12702

Zsoter, E., Prudhomme, C., Stephens, E., Pappenberger, F., & Cloke, H. (2020). Using ensemble reforecasts to generate flood thresholds for improved global flood forecasting. *Journal of Flood Risk Management*, *13*(4), 1–14. https://doi.org/10.1111/jfr3.12658

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mitheu, F., Tarnavsky, E., Ficchì, A., Stephens, E., Cornforth, R., & Petty, C. (2023). The utility of impact data in flood forecast verification for anticipatory actions: Case studies from Uganda and Kenya. *Journal of Flood Risk Management*, e12911. https://doi.org/10.1111/jfr3.12911

17533318x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12911 by University of Reading, Wiley Online Library on [26/04/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License