ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Can citizen scientists provide a reliable geo-hydrological hazard inventory? An analysis of biases, sensitivity and precision for the Rwenzori Mountains, Uganda

To cite this article: John Sekajugo *et al* 2022 *Environ. Res. Lett.* **17** 045011

IOP Publishing Environ. Res. Lett. 17 (2022) 045011 https://doi.org/10.1088/1748-9326/ac5bb5

ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
8 November 2021

REVISED
3 March 2022

ACCEPTED FOR PUBLICATION
8 March 2022

PUBLISHED
21 March 2022

LETTER

Can citizen scientists provide a reliable geo-hydrological hazard inventory? An analysis of biases, sensitivity and precision for the Rwenzori Mountains, Uganda

John Sekajugo[1,2,3,*], **Grace Kagoro-Rugunda**[3], **Rodgers Mutyebere**[1,5], **Clovis Kabaseke**[1], **Esther Namara**[1], **Olivier Dewitte**[4], **Matthieu Kervyn**[2] and **Liesbet Jacobs**[5,6]

[1] School of Agriculture and Environmental sciences, Mountains of the Moon University, Fort Portal, Uganda
[2] Department of Geography, Vrije Universiteit, Brussels, Belgium
[3] Department of Biology, Mbarara University of Science and Technology, Mbarara, Uganda
[4] Department of Earth Sciences, Royal Museum for Central Africa, Tervuren, Belgium
[5] Department of Earth and Environmental Sciences, KU Leuven, Leuven, Belgium
[6] Ecosystem & Landscape Dynamics, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
[*] Author to whom any correspondence should be addressed.

E-mail: john.sekajugo@vub.be

Keywords: citizen science, geo-observer, landslides, floods, reliability, tropical Africa, disaster response

Abstract

Spatio-temporal inventory of natural hazards is a challenging task especially in rural or remote areas in the Global South where data collection at regional scale is difficult. Citizen science, i.e. involvement of no-experts in collecting information and co-creation of knowledge with experts to solve societal and environmental problems, has been suggested as a viable approach to tackle this bottleneck, although the reliability of the resulting data is often questioned. Here we analyse an inventory of geo-hydrological hazards (landslides and floods) reported by a network of citizen scientists in the Rwenzori Mountains, Uganda, established since 2017. We assess the precision, sensitivity and potential biases affecting this citizen science-based hazard inventory. We compare the citizen science-based records with two independent inventories, one collected through systematic fieldwork and another by PlanetScope satellite imagery mapping for the period between May 2019 and May 2020. The precision of the geo-observer data is higher (99% and 100% for landslides and floods, respectively) than that of satellite-based data (44% and 84%, respectively) indicative of fewer false positives in the former inventory. Also, citizen scientists have a higher sensitivity in reporting landslides (51%) compared to satellite imagery (39%) in addition to being able to report the events a few days after the occurrence. In contrast, the sensitivity of satellite-based flood detection is higher than that of citizen scientists. The probability of landslide events being reported by citizen scientists depends both on citizen scientists and hazard specific features (impact, landslide-citizen scientist home distance, landslide-road access distance and altitude). Although satellite imagery mapping could result in a spatially less biased inventory, small landslides are often missed while shallow ones can easily be confused with freshly cleared vegetation. Also, in a dominantly cloudy environment, it can take several days to weeks before a cloud-free satellite image can be obtained. In summary, the typically rapid response time of citizen scientists can result in faster information with high reliability at the risk of missing out almost half of the occurrences. Citizen scientists also provide more data on impact and type of land use, something difficult to achieve using satellite imagery. Working with farmers at village level as citizen scientists can facilitate covering a wider geographical area while reducing the area monitored by each citizen scientist at the same time.

1. Introduction

Climate-related disasters are on the rise (IFRC 2016, UNISDR 2015a, Papathoma-Köhle *et al* 2016). They continue to claim lives of many, destroy infrastructure, damage property, and disrupt livelihood activities that eventually undermine the welfare of affected societies. According to the Sendai framework for disaster risk reduction (2015–2030), increasing exposure and vulnerabilities have steadily increased the risk to natural hazards and the associated disasters (UNISDR 2015b). To be able to foster national disaster risk reduction strategies, it is necessary to promote local risk assessment studies based on continuous data collection on disaster occurrences and the impacts they cause. Disaster risk is categorized into intensive and extensive risk. Intensive risk is associated with large impact events of low frequency (Hamdan 2015). Extensive risk on the other hand involves small, frequent and widely spread hazard events such as landslides and floods (UNDRR 2015). Such geo-hydrological hazard events rarely make it to national media, making their assessment challenging despite their cumulative impact being large (Hamdan 2015). This is particularly the case in developing nations especially the tropical mountainous regions of Africa where spatially dispersed small hazards are common but their impact remain under-documented (UNISDR 2015a, UNDRR 2019a). In these mountainous regions, high population densities are often found, frequently on the rise and combined with high societal vulnerabilities (UNISDR 2015a, CRED 2017, Zhou *et al* 2019).

Bottlenecks in regional and subnational Disaster Risk Reduction (DRR) include insufficient institutional capacities, lack of data and inadequate budget allocations (UNISDR 2015b, AUC 2018, UNDRR 2019a). Comprehensive multi-hazard analysis can contribute to better risk management through: (a) a clear understanding of geographical concentrations of the disaster risk, (b) quantification of potential impacts, and (c) identification of key risk drivers (Pondard and Daly 2011). This requires spatially and temporally explicit data records including impact caused. The increasing availability of very high spatial resolution satellite imagery (such as IKONOS, QuickBird, PlanetScope, and Pléiades) presents an opportunity for rapid and accurate visual geo-hydrological hazard mapping at different spatial and temporal scales (Rabonza *et al* 2016, Monsieurs *et al* 2018, Talisay *et al* 2019, Thirumurugan and Krishnaveni 2019, Jain *et al* 2021). However, in tropical environments, persistent cloud cover, often limits use of optical satellite remote sensing for hazard detection (Robinson *et al* 2019). Visual satellite image interpretation can also be affected by mis-classifications as human-induced seasonal land use changes can be perceived as potential geo-hydrological hazards (Dewitte *et al* 2022, Jain *et al* 2021).

In the recent past, citizen science (CS) has gained tremendous recognition (Danielsen *et al* 2014, Aceves-Bueno *et al* 2017, McKinley *et al* 2017, Hicks *et al* 2019) although it has existed for long time (Silvertown 2009, Eitzel *et al* 2017). This is partly due to the increasing realization among scientists of the benefits of this practice wherein non-experts and experts engage and collaborate to solve real-world problems (Cohn 2008). Citizen scientists (CSs) can engage in various stages of the research cycle ranging from collecting data to disseminating study findings (Dickinson *et al* 2012, Aceves-Bueno *et al* 2017, Cunha *et al* 2017, Phillips *et al* 2018). CS therefore has been interpreted and defined differently depending on context and field in which it is being applied. According to Kullenberg and Kasperowski (2016), CS is understood to involve participation of citizens in the collection and analysis of data for research in different fields including ecology, biology and conservation (Kullenberg and Kasperowski 2016). It can also involve environmental monitoring and inventory projects (Haklay *et al* 2021). In the context of DRR, the involvement of citizens in recording or mapping hazards can be an example of CS projects (Hicks *et al* 2019). In this research, the project involves the participation of non-expert citizens in mapping geo-hydrological hazards and their impacts with goal of scientifically understanding the spatio-temporal distributions, vulnerabilities and risks imposed by the disasters. Participatory approaches enable data collection on a larger geographic scale and over a longer time period than is possible in more conventional scientific research (Cohn 2008). Furthermore, collaborating with community members in disaster assessment facilitates complete understanding of disaster risk in a local context (Bwambale *et al* 2022).

In 2017, a CS network, coined the *geo-observer network*, was established in the Rwenzori Mountains, Uganda (Jacobs *et al* 2019) mainly to (a) record and report environmental hazards (including dates of occurrence and impact caused), (b) disseminate scientific research findings and participate in developing and implementing DRR strategies. This was based on an understanding that people in the affected communities are better placed to monitor their own places and participate in developing DRR strategies. The network currently consists of 30 CSs, trained to detect, and report information on eight different hazards (landslides, floods (here referring to overflow of large volumes of water beyond the usual normal levels mainly during or after rainfall), droughts, crop and livestock pests and diseases, earthquakes, lightnings, windstorms, and hailstorms) by filling a simple structured questionnaire programmed in a

KoboCollect application using smartphone technology (Jacobs *et al* 2019). With a mobile internet connection, the information recorded in a standard form (including location, type of hazard, size, picture of the affected elements and impact on housing, population, crops and physical infrastructures) is uploaded to a central *KoboCollect* account managed by a research assistant based at Mountains of the Moon University. This online database contains all information recorded for each event, including GPS coordinate, pictures, and the ID of the reporting citizen scientist. Each geo-observer of the network is responsible to report all events for a specific territory, typically one or two parish (es). A parish is a second smallest administrative unit in Uganda next to a village. The parish sizes are not homogeneous as the boundaries are mostly defined by physical features such as rivers, streams, wetlands, roads and gorges. Although the network has demonstrated potential for CS-based hazard mapping Jacobs *et al* (2019), the growing inventory has to date not been used for hazard or risk assessment. Prior to valorisation of the data, i.e. using the CS-based hazard location and impact on society to constrain risk assessment and inform risk reduction strategies, in hazard and risk assessments, the reliability of the citizen-based dataset needs to be assessed (Jacobs *et al* 2019) as citizen science based data might be affected by biases or issues of reliability or completeness (Bird *et al* 2014, Lewandowski and Specht 2015, Haworth 2016, Aceves-Bueno *et al* 2017, Haworth *et al* 2018, Jacobs *et al* 2019). This study therefore is aimed at evaluating the data quality of the citizen-based inventory in the Rwenzori Mountains. Focusing on landslides and floods, we validate the CSs reports through systematic fieldwork, analyse the precision and sensitivity of the inventory, and compare these to an inventory constructed using visual interpretation of very high-resolution satellite imagery. Finally, we identify potential underlying factors determining the detectability of landslides by CSs.

2. Materials and methods

2.1. Study area

Rwenzori Mountains are located in tropical Africa (Peel *et al* 2007) across the southwestern border of Uganda and eastern D.R Congo. This mountainous region suffers from frequent landslide (Jacobs *et al* 2016a) and floods (Jacobs *et al* 2016b). The Ugandan side of the mountains is characterized mainly by protected zones (national parks, forest, and wildlife reserves) and inhabited zones (figure 1). This study focuses on landslide and flood events on the inhabited Ugandan side of the Rwenzori Mountains, specifically those occurring in the two districts of Kasese and Bundibugyo.

2.2. Landslide and flood inventory

2.2.1. Citizen science-based inventory

The CS based inventory was established from a selection of 18 CSs active in 30 selected parishes out of the 41 covered by CSs in the study districts (figure 1). In selecting the parishes, in addition to ensuring a good spatial spread over the two study districts, considerations were made to ensure representation of both parishes from which many reports had been received and those with little or no reports at all. The methodology for data-collection on landslides and floods is described in (Jacobs *et al* 2019). This dataset contains among others, details on the location (GPS coordinates), timing and damage inflicted by a hazard event (individual landslide or flood occurrence) and is typically collected within a timespan of a few days after occurrence. Although the data collection started in February 2017, we here narrow the dataset to landslides and floods reported between 1 May 2019 and 30 May 2020 because landslide scars, particularly for shallow and smaller landslides, are not always preserved in the landscape due to quick vegetation regrowth or reclamation by agriculture. As soon as the hazard reports are submitted by the CSs, they go through a manual validation process by a research assistant to remove reports with obvious errors like presence of incorrect hazard pictures (pictures not related to the hazard being reported) and suspicious impact (Jacobs *et al* 2019). This was followed with automated screening using a Python script to identify potential inconsistences (including GPS position precision, date of recording compared with date of occurrence). In the 1 year period considered in this study, 292 and 75 landslide and flood reports had been submitted by the CSs. After the manual and automated screening, the consolidated dataset consisted of 255 landslide and 59 flood reports.

2.2.2. Satellite imagery-based inventory

The remote sensing (RS)-based inventory was built for the same period (1 May 2019 to 30 May 2020) as the retained CS-based inventory from orthorectified 3 m resolution visual optical images of PlanetScope (Planet 2016). PlanetScope offers among other products visual geometrically and atmospherically corrected 3-band; Red Green Blue (RGB) and analytic 4-band (RGB and Near Infra-Red) images for different places across the globe with a one-day revisit time (Planet 2021). The imagery is collected as a series of framed overlapping scenes from a single satellite in a single pass (Planet 2019). To ensure the images acquired were suitable for visualization application, a 15% cloud threshold was used before download. However, due to persistent cloud cover, it was not possible to get cloud free images covering the two study districts at the same time at all times. For each clear image of one district, the nearest in

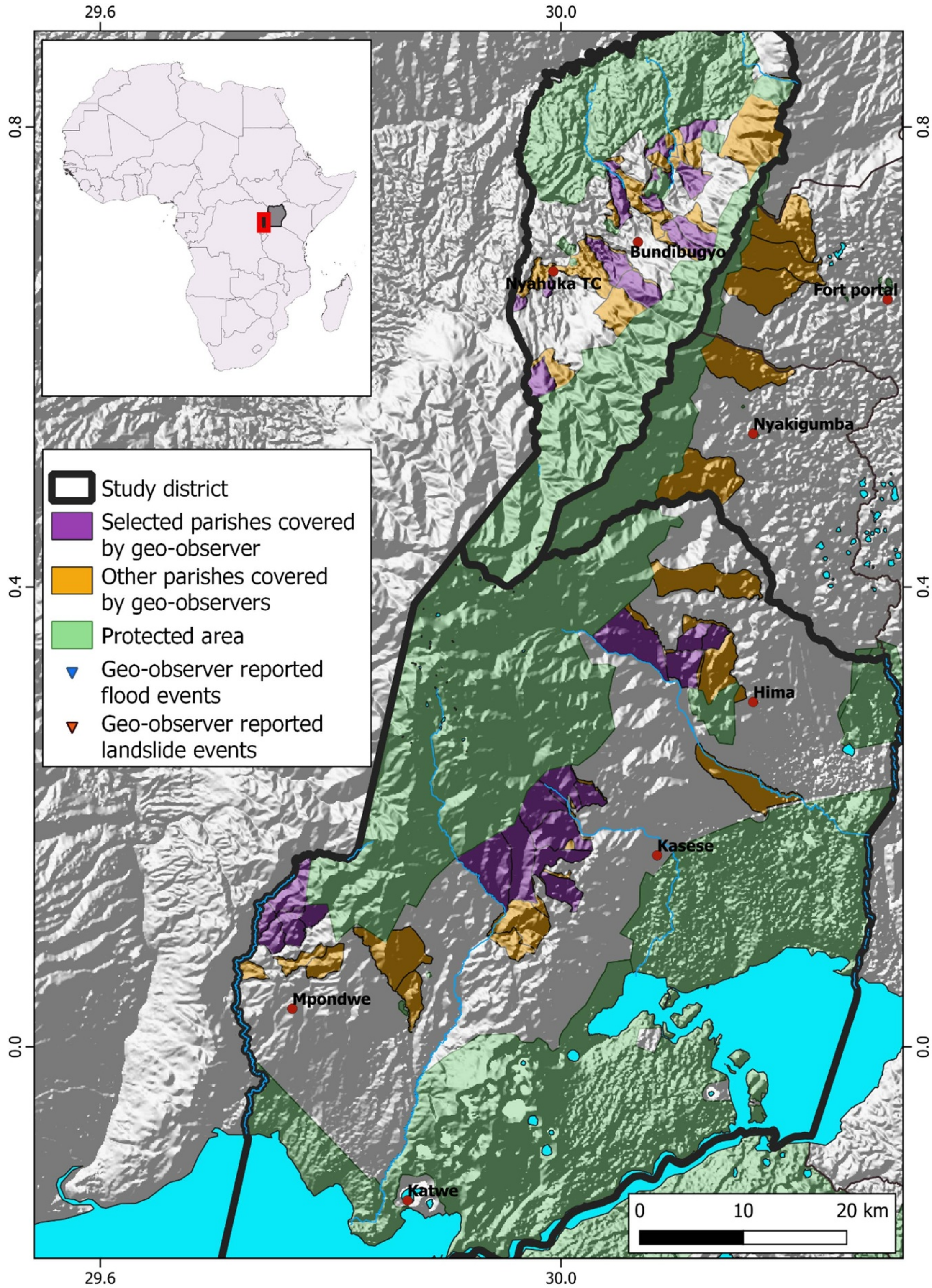

Figure 1. Map showing the study districts of Bundibugyo and Kasese, and landslide and flood events reported by geo-observers for the period 1 May 2019 to 30 May 2020 in the parishes selected for this study.

time image of the other district was used. From each image (appendix A), potential landslides and flood events were digitized in a GIS environment. By overlaying the already constructed shapefile of digitized events on a successive satellite image, double digitizing was avoided. A total of 458 landslide and 59 flood potential events respectively were digitized from the satellite images.

2.2.3. Expert-based field assessment and validation

From mid-September to mid-December 2020, a field assessment and validation was conducted by the first

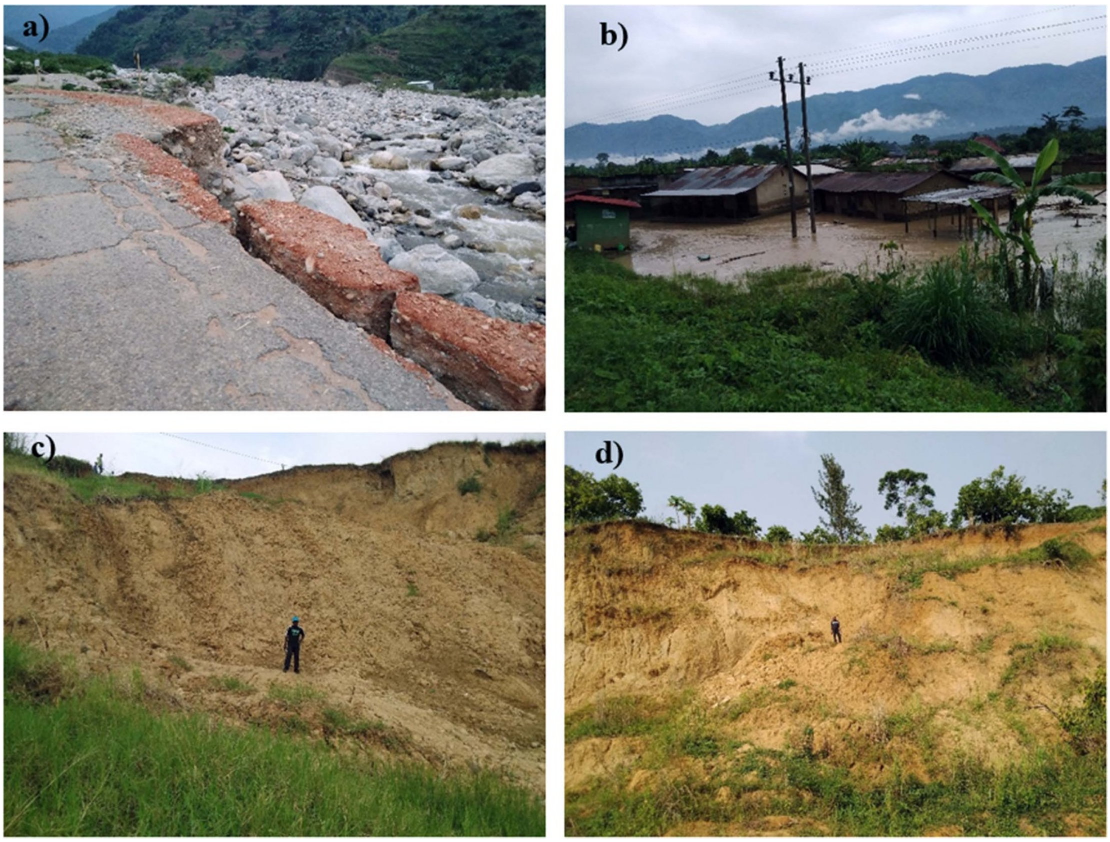

Figure 2. Examples of flood events along river Nyamwamba (a), Kasese district (5 August 2020) and river Kirumya (b), Bundibugyo district (16 November 2020) as well as landslides in Bundikeki (c) and Kirindi (d) parishes in Bundibugyo district (10 January 2021 and 6 February 2021, respectively).

author to systematically map landslides and flood occurrences (figure 2) in the 30 selected parishes and verify the events reported in the CS database and those events identified through visual interpretation of satellite imagery. For each parish, a local guide was hired to ensure all the possible areas of the administrative unit were reached. This facilitated mapping as many landslide and flood events that occurred in the period of May 2019 and May 2020 as possible. The CSs reports, and the RS-based inventory were also verified by systematically visiting these locations.

2.3. Data validation

Each event that would be found to exist in the field and correctly identified by the CSs and/or mapped from the satellite imagery, was labelled True Positive (TP). Those erroneously mapped as landslide or flood events or found not to exist in the field, were labelled False Positives (FP). Events that were found to exist in the field but had been missed by either the CSs or the RS inventory process were labelled False Negatives (FN). The process is illustrated using a hypothetical field space in figure 3. These classifications were then used in subsequent analyses outlined below and summarized in the schematic figure 4.

2.3.1. Sensitivity, precision, F1

Sensitivity (also called true positive rate or recall rate) is used to refer to the proportion of existing positives that are correctly identified or detected as positives (Powers 2008, Tharwat 2018, Chicco and Jurman 2020). In previous studies, sensitivity has been used to measure the performance of diagnostic tests (Genders *et al* 2012) and classification performance in computational linguistics (Powers 2008). In this study, sensitivity is used to refer to the proportion of landslides/floods that occurred and were correctly reported by CSs and/or mapped from satellite imagery. It is defined as given in equation (1).

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})}. \tag{1}$$

Precision (confidence) has been applied in information retrieval and data mining studies to measure the predictive performance of models. It refers to the proportion of predicted positives that are truly positives (Powers 2008, Boughorbel *et al* 2017). In this study it denotes the proportion of detected landslides or floods that were real landslides or floods in the field. It is calculated as given in equation (2).

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}. \tag{2}$$

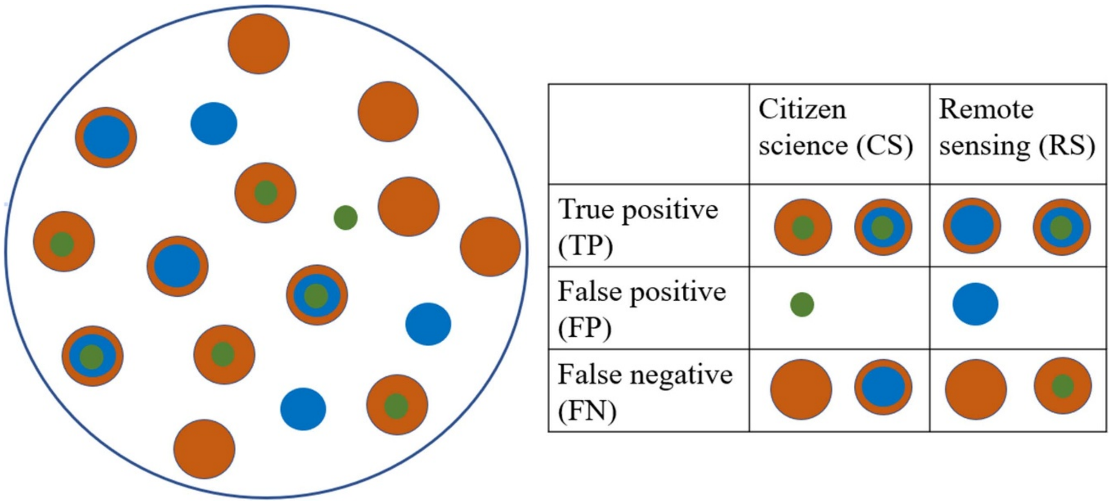

Figure 3. Illustration of a hypothetical field space where brown dots represent events (landslide/floods) that occurred, green dots represent CSs mapped hazards while blue dots represent RS-based inventory. TP represent events found to exist in the field and correctly identified by the CSs or mapped from the satellite imagery. FP represent features erroneously mapped as landslide or flood events or found not to exist in the field while FN represent events that were detected in the field but had been missed by either the CSs or the RS inventory process.

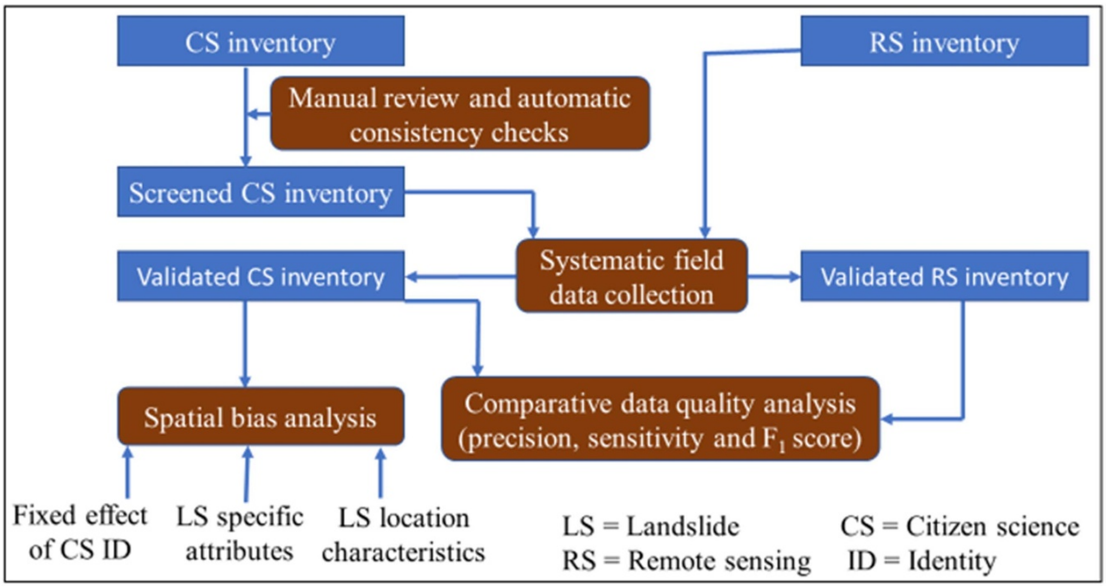

Figure 4. Schematic flow of the entire study with data inventories (blue fill), inventory verification, quality, and bias analysis (dark brown fill).

The F_1 score (F measure) is an integration of both sensitivity and precision. It gives the harmonic mean of sensitivity and precision and ranges between 0 and 1 (Tharwat 2018). Although the metric has received some criticism (Flach and Kull 2015), it is still a widely used measure of classification or predictive performance studies (Genders *et al* 2012, Chicco and Jurman 2020, Chicco *et al* 2021). Here it was applied to measure the overall performance of CSs and satellite in detecting landslide and flood occurrences. The higher the F_1 score, the better the detective performance of the hazard inventory approach. It is calculated as in equation (3).

$$F_1\,\mathrm{score} = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}. \tag{3}$$

The comparison of the CS-based- and satellite-based inventories with the field data and subsequent calculation of sensitivity, precision and F_1 statistics allowed a direct contrasting of the data quality of the two former inventories (figure 4).

2.3.2. Spatial bias analysis

CS-based inventories can be affected by spatial biases (Zhang and Zhu 2018). In our study, the landslide inventory in particular could be affected by biases: landslide events are spatially more dispersed than floods and are more frequent but smaller in size, which could render them less noticeable. CSs might have more knowledge or be more likely to visit areas closer to their homestead (Pernat *et al* 2021). In addition, landslide inventories by CSs but also by experts (Dickinson *et al* 2010) can be affected by a bias towards those areas closer to the road network (Wang *et al* 2020). This creates incompleteness issues in such citizen-based databases. Here we examine the potential sources of spatial biases by analysing factors that determine whether a citizen scientist detects and records a hazard event. To do this, a binary logistic regression model was applied. Logistic regression is a widely used modelling approach in determining relationships between a binary dummy dependent variable and explanatory variables (Coughlin *et al* 1992, Genders *et al* 2012, Nishadi 2019). In our case, the binary dependent variable is whether or not a landslide that was confirmed in the field was reported by the citizen scientist (1) or not (0). If landslide reporting ability is represented as $\mathrm{Ls}_{\mathrm{report}}$ and $X_1, X_2, X_3, \ldots, X_n$ represent a set of potential explanatory variables, the probability of a landslide being reported by CSs is given as shown in equation (4).

$$P\left(\mathrm{Ls}_{\mathrm{report}}=1\right) = \frac{1}{1+e^{-(\alpha+\beta_1X_1+\beta_2X_2+\beta_3X_3+\ldots\ldots\ldots\ldots\beta_nX_n)}}. \quad (4)$$

where $P\left(\mathrm{Ls}_{\mathrm{report}}\right)$ is the probability of a landslide being detected and β_1 to β_n are regression coefficients. The first set of explanatory variables used in equation (4) model were landslide related factors. These included landslide size, slope, altitude, landslide-citizen scientist home distance, landslide-road access distance and impact (weighted sum of effects on humans, infrastructures and crops, appendix B). Although even a low impact landslide event can be detected and reported by a citizen scientist, the more disastrous an event is, the higher is the chance that such an event will be discussed on different platforms and the easier it might be for a citizen scientist to learn about it and detect it. The regression analysis was repeated while accounting for a fixed effect of CSs (citizen scientist specific unique attributes). To achieve this, unique identity (citizen scientist ID) numbers were introduced in the model. To gain an insight on how much variance is explained by each of the factors included in the regression analysis, a univariate logistic regression analysis was performed for each of the variables. The Tjur's coefficient of discrimination, D, was then used as an indicator of the individual contribution for a specific variable in explaining variance in landslide detectability by the CSs (Tjur 2009, Allison 2014). We then investigated which factors could explain variations in the observed individual odds ratios of CSs. This explorative analysis was done through a Spearman rank correlation matrix involving individual odds ratios of CSs, age, years spent reporting, size of the area monitored and number of reports per citizen scientist.

Table 1. Results from validation of CS and RS inventory through systematic fieldwork.

	Citizen scientists		Satellite imagery	
Outcome	Landslides	Floods	Landslides	Floods
True positive	218	46	165	48
False positive	3	0	208	9
False negative	207	20	260	18
Total	428	66	633	75
Sensitivity (%)	51	70	39	73
Precision (%)	99	100	44	84
F_1 score	0.7	0.8	0.4	0.8

3. Results

3.1. Landslide and flood detection

Out of the 255 and 59 landslides and floods retained in the CS-based inventory after manual review and automated screening of incoming reports, 221 landslides and 46 floods were verified in the field. Of the 459 landslides and 59 floods identified on satellite images, 373 and 57 respectively could be verified in the field (table 1). The remaining instances could not be verified due to accessibility constraints in the field and were not considered in further analysis. Results show that CSs identify and report landslide and flood hazards in their localities with 99% and 100% precision respectively (table 1). Table 1 however indicates also a relatively low sensitivity for CS-based landslide and flood detection (51% and 70% respectively). RS-based landslide and flood detection is less precise particularly for landslides (39%) in addition to having even lower sensitivity. RS-based flood detection has a higher sensitivity compared with CS. Overall, floods are more easily detected by both CSs and RS as implied by F_1 scores of 0.8 and 0.8 respectively.

3.2. Citizen scientists versus RS-based landslide hazard detection

Table 2 shows a direct comparison between CS and RS inventories in the form of a confusion matrix. Only 16% of the 425 landslides mapped in the field were detected by both satellite and CS and 26% were detected by neither inventories. Remarkably, 35% of the landslides was detected by the CS but not by the satellite interpretation and 23% of satellite detected landslides were not identified by CSs. The high sum of this diagonal demonstrates that both inventories are to a large extent complementary and typically do not detect the same events. While the average size of the landslides detected and not detected by the CSs is not statistically significant (Prob $> z = 0.088$), satellite imagery detects larger events and tends to miss out on the small ones (Prob $> z = 0.000$). Through logistic regression (appendix C; table 8), it was found out that landslide detection by RS is significantly influenced by

Table 2. Citizen scientists versus RS-based landslide detection.

	CS-based inventory			
RS-based inventory	Detected	Not detected	Total	Average size (m^2)
Detected	69 (16%)	96 (23%)	165 (39%)	3713 ± 6196
Not detected	149 (35%)	111 (26%)	260 (61%)	792 ± 1558
Total	218 (51%)	207 (49%)	425 (100%)	—
Average size (m^2)	1489 ± 3413	2385 ± 5010	—	—

Table 3. Logistic regression results for landslide related factors. z statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Variable	Odds ratio	Univariate Tjur's D
Landslide impact	1.099^{***}	0.028
Landslide geo-observer home distance	0.771^{***}	0.055
Landslide road distance	0.873^{**}	0.005
Slope	1.023	0.024
Land slide area	1	0.012
Altitude	1.001^{**}	0.009
Constant	0.293^{**}	—
Tjur's D	0.138	0.133
	Prob > $chi^2 = 0.000$	

Table 4. Correlation matric for citizen scientists' individual odds ratios and potential determinants.

Variables	Individual odds ratios	Age	Period of operation	Area covered	Number of reports submitted
Individual odds ratios	1.000	—	—	—	—
Age	0.778[a]	1.000	—	—	—
Period of operation	0.234	0.541	1.000	—	—
Area covered	0.371	0.283	−0.353	1.000	—
Number of reports submitted	0.238	0.287	−0.187	0.563	1.000

[a] Significance at 0.05 confidence level.

landslide size, slope and altitude (Tjur's D = 0.226). A marginal effect post estimation showed that the chances of missing out some landslides on a satellite image increase as the landslide area reduces.

3.3. Factors that influence citizen science-based landslide detectability

Results of the logistic regression show that the variation in landslide detectability by CSs is significantly influenced by the landslide's impact, the distance between citizen scientist's residence and the landslide, the distance from the access roads to landslide and altitude (table 3). The odds ratios confirm the hypothesis that larger impact and shorter distance increase the likelihood of landslide events being detected. However, based on the Tjur's D, these factors only explain a small proportion (Tjur's D = 0.138) of the variations in landslide detectability. Through univariate analysis, a Tjur's D for each factor was generated to gain an understanding of the maximum contribution of each individual factor to the variance in landslide detectability by CSs (table 3 column 3). Based on the univariate analysis results, the landslide distance from the citizen scientist's home plays the most important role, followed by impact and slope in influencing landslide detectability.

To check potentially controlling factors related to the individual CSs included a fixed effect for each individual in the logistic regression. This increased the explanatory power of the model (Tjur's D = 0.435 from 0.138). In a final step, we explored which personal characteristics could influence the resulting odds ratios by simple correlation with the variables as presented in table 4.

4. Discussion

4.1. Sensitivity and precision in CS- and RS-based hazard detection

CSs can correctly identify and report georeferenced hazard information comparable to research experts. Although no control trial was done to quantify the added value of the training workshops, this high precision rate after training workshops and in-field practice concurs with what Lewandowsli and Spencht (2015) report on the importance of adjusting protocols and volunteers training in improving data quality (Lewandowski and Specht 2015). According to Paul *et al* (2014), CS-based data collection programs result in datasets with quality attributes similar to those established systematically by experts (Paul *et al* 2014). However, their inventories (CSs) suffer incompleteness challenges especially those that involve mapping small, frequent but spatially spread events like landslides (Rohan *et al* 2021).

Despite the relatively low sensitivity, the CS database outperformed the RS-based inventory in terms of sensitivity, precision and F_1 scores when considering landslides. The low sensitivity and precision of RS-based inventory compared to CSs and other studies are due to several factors. We used satellite images at a 3 m resolution while other studies used Google Earth images that offer a 0.3–0.6 m resolution (Hao *et al* 2020, Ubaidulloev *et al* 2021). The images available in Google Earth are usually acquired with a lower temporal resolution compared with the PlanetScope images, and therefore may capture landslide scars that are already (partially) concealed by the regeneration of the vegetation (Dewitte *et al* 2022). This reduction of the spectral signature of the landslides identified via Google Earth is however counterbalanced by the higher spatial resolution of the images. In addition, CSs visit the field after a few days, allowing them to identify very small landslides that cannot be seen on a satellite image. Furthermore, the inventory was limited to areas monitored by CSs. These areas are characterized by cultivated landscapes made of parcels that present shape that could be mistaken for landslides on the satellite image. The confusion matrix (table 2) demonstrates that the CS inventory and the RS inventory are largely complementary. However, the CS inventory in turn is affected by biases predominantly relating to a fixed individual citizen scientist effect and to lesser extent relating to the landslide characteristics, impact, and location. The further the event from the residence of the citizen scientist or the community access roads, the higher are the chances of such an event being missed. The magnitude of impact and nature of the element impacted are also important. This supports our hypothesis that more destructive events are more readily communicated or discussed on different platforms and so easier to detect by CSs. People are more motivated to report events that affect their houses or cause injuries/death than those that affect crops and pastures. This could partially be due to the perception that CSs reports reach humanitarian agencies who eventually distribute relief items mainly to those whose houses get damaged. Most landslide events that affects houses in the study area are usually small resulting from (a) steep slope cuts while establishing platforms for house construction and (b) small scale cropping farming on the steep slopes. This partially explains why CSs report more of small events. The observed improvement in the explanatory power of the model after introducing the citizen scientist ID implies that CSs specific characteristics (here bundled into one fixed effect) play an important role in influencing the likelihood of landslides being detected. Thus, while CS-based inventories might depend on socio-cultural, infrastructural, motivational factors of the CSs and nature of the hazard, satellite imagery largely depends on nature of the target object in the feature space (size of the object) and the prevailing atmospheric conditions (Agapiou *et al* 2011, Selva 2021). In general, it should be noted that the ability of CSs to detect landslides that have occurred largely depend on whether they get informed about their occurrence. The positive correlation between number of reports and age of CSs (table 4) can be explained by the fact that the older the CSs are, the higher the chance of them having a stable residence in their village and the wider the social network built from whom news about landslide occurrences can be received. In addition, older CSs tend to stay longer in the reporting network as indicated by a strong positive correlation ($r = 0.55$) between years of operation and CSs age (table 4).

CS-based inventories and RS-based inventories not only differ in which landslides are detected, but both inventory methods also fundamentally differ in the type of information that can be collected: CSs can collect data on the impact associated with the hazards that is not visible through remotely sensed information. RS-based mapping on the other hand provides an opportunity to record geo-hydrological hazards even in inaccessible areas (Ji *et al* 2020). On the other hand, while revisit periods of very high-resolution satellites are shortening, image interpretation is oftentimes still hampered by cloud coverage, while CSs tend to visit, and map affected sites within few days after occurrence (Jacobs *et al* 2019). The implication here is that CS-based landslide detection and very high-resolution ($\leqslant$3 m) satellite imagery are complementary in spatial extents, types of landslides and types of information collected, but not substitutes. Because of this complementarity, CS data represent an added value to hazard assessment and risk analysis for example in determining thresholds for landslide triggering factors (Monsieurs *et al* 2019) and supplementing incomplete landslide inventories (Samodra *et al* 2018). CSs generate reliable georeferenced hazard information including information on timing and damage inflicted, particularly in data-poor countries like Uganda where extensive disaster risk is high. Such events do not easily make it to social or formal media platforms in addition to not meeting the impact thresholds of the existing national (DesInventar) data repositories (UNDRR 2019b) and global (EM-DAT) databases (IFRC). To reduce incompleteness challenges, it is relevant for projects and programs aimed at spatio-temporal explicit hazard assessment to complement high spatial resolution satellite image-based hazard inventories with CS hazard sensing. By visiting the affected scenes within a few days and documenting the impact caused, CS-based inventory can be useful in temporal landslide assessment and vulnerability analysis which are prerequisites for risk analysis.

4.2. Limitations of the data used in the analysis

Although we find the quality of the CS-based inventory of good quality, maintaining a big network of CSs to cover a large geographical area over time can be challenging and require investment to equip, train and continuously monitor the activities of CSs. Although CSs report geo-hydrological hazards with high precision, the inventory should be carefully used for regional susceptibility and risk assessments as it is not free from spatial biases. In addition, the satellite imagery digitization was entirely made by the main author of this manuscript. Although efforts were made to do it systematically, it does not completely erase possibility of omitting or including certain features that would respectively be or not be included by other users.

5. Conclusion

This study demonstrated the reliability of CS-based hazard detection and its complementarity with RS-based inventory. It also showed that by integrating CS-based hazard inventories with existing high resolution satellite imagery in spatio-temporal geo-hydrological hazard studies can help to overcome the incompleteness challenges. The Sendai framework for DRR advocates for inclusive national and local hazard assessment and risk management (UNISDR 2015b) and so CS provides a viable approach not only for rural and remote mountainous areas such as the Rwenzori ranges, but for all at-risk regions where traditional observation networks and infrastructures are weak or absent and local communities directly impacted. The network presents a desired multidimensional approach in cogenerating knowledge and disseminating research findings and could enhance participatory disaster risk management at local level.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This research was supported by the Flemish Inter University Council (VLIR-UOS) through the TEAM project (UG2019TEA492A105) titled 'Digital citizen science for community-based resilient environmental management (D-SiRe)' implemented in a collaboration between Vrije Universiteit Brussels and KU Leuven on one hand, Mbarara University of Science and Technology and Mountains of the Moon University on the other hand. We thank all the geo-observers in the Rwenzori network for their commitment. This work would not have been possible without them. We also thank the local government and community people for supporting the CS network.

Ethical statement

The Digital citizen science for community-based resilient environmental management (D-SiRe) project under which this research has been conducted was registered and approved by Uganda National Council for Science and Technology (UNCST) with registration number NS126ES.

Appendix A. List of dates and months (table 5) in which images were acquired to build the RS-based inventory

Table 5. List of dates and months in which images were acquired to build the RS-based inventory.

Bundibugyo district	Kasese
3 May 2019	3 May 2019
5 July 2019	14 and 18 July 2019
13 September 2019	22 September 2019
27 December 2019	30 December 2019 and 1 January 2020
16 February 2020	12, 13 (area near D.R Congo border only) and 16 February 2020
8 April 2020	7 and 8 April 2020
13 May 2020	13 May 2020
26 May 2020	26 and 30 May 2020

Appendix B. Landslide combined impact (tables 6 and 7) value development through multicriteria additive weighing

Landslide impact is an important variable to consider while analysing factors that influence whether an event can be detected and reported or not in citizen science (CS)-based hazard assessment. Although event a low impact landslide event can be detected and reported by a citizen scientist, we hypothesize that the more disastrous an event is, the higher is the chance that such an event will be discussed on different platforms and the easier it will be for a citizen scientist to detect it.

The nature of impact and type of element affected also matters. For instance, an event that involves a fatality attracts more public attention than the one involving only crop damage. Similarly, an event that blocks an access road can be easier to detect than the one that only cause it to crack. Therefore, using type of element affected, nature of impact and magnitude of impact caused criteria, different weights were assigned based on experience and expert judgement as shown in tables 6 and 7. For each type of element affected, the weights assigned were converted to relative weights by dividing the biggest wight assigned to the smallest so that all values range from 0 to 1. Since one landslide event can impact on multiple elements, the relative weights of different affected elements for each of the respective event were added together to come up with the total impact value for that particular landslide event.

Table 6. Criteria for weighing of impacts based on the type of element and nature of impact on a scale of 1–5 (1—attract low public attention, 5—attract high public attention).

Road damage ranking	Weight	House damage ranking	Weight	Direct impact on people	Weight	Crop damage	Weight
No damage	1	No house	1	No impact	1	No crop damage (pasture or forest)	1
Cracked	2	Cracked	2	Displaced	3	Partially damaged	2
Covered by debris/soil	3	Wall/roof damage	4	Hospitalized	4	Completely destroyed	3
Destroyed	4	Destroyed	5	Some people killed	5	—	—

Table 7. Criteria for weighing impact based on the magnitude of impact on different types of elements at a scale of 1–10 (1—attract low public attention, 10—attract high public attention).

Number people displaced	Weight	Number of people hospitalized	Weight	Number people killed	Weight	Number of houses affected	Weight	Length of road affected (m)	Weight	Area of crop damage (Sq.m)	Weight
0	0	0	0	1	5	0	0	0	0	⩽100	1
⩽5	3	⩽5	4	2	8	⩽5	2	⩽5	1	⩽500	2
⩽10	5	⩽10	6	>2	10	⩽10	4	⩽10	2	⩽1000	3
⩽15	7	⩽15	8	—	—	⩽15	6	⩽15	4	⩽2000	4
⩽20	8	>15	10	—	—	⩽20	8	⩽20	6	⩽3000	5
⩽25	9	—	—	—	—	⩽25	9	⩽25	8	⩽4000	6
>25	10	—	—	—	—	>25	10	⩽30	9	⩽6000	7
—	—	—	—	—	—	—	—	>30	10	⩽8000	8
—	—	—	—	—	—	—	—	—	—	⩽10 000	9
—	—	—	—	—	—	—	—	—	—	>10 000	10

Appendix C. Logistic regression results (table 8) and marginal effect post estimation (figure 5) to determine threshold for landslide area effect on landslide detectability using satellite imagery

A logistic regression was performed for landslide detectability through satellite imagery as a dependent variable and slope, landslide area, altitude and NDVI as independent variables. Slope, landslide area and altitude influence positively on the detectability. The marginal effect post estimation results (figure 5) show that as the landslide area reduces, the chances of missing out some landslides on a satellite image increase.

Table 8. Logistic regression results for factors that influence landslide detection using satellite imagery. z statistics in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Variable	Odds ratio
Slope	.963***
Land slide area	1***
Altitude	.999***
Normalized difference vegetation index (NDVI)	30.565
Constant	.959
Tjur's D	0.226
Prob > chi^2	0.000

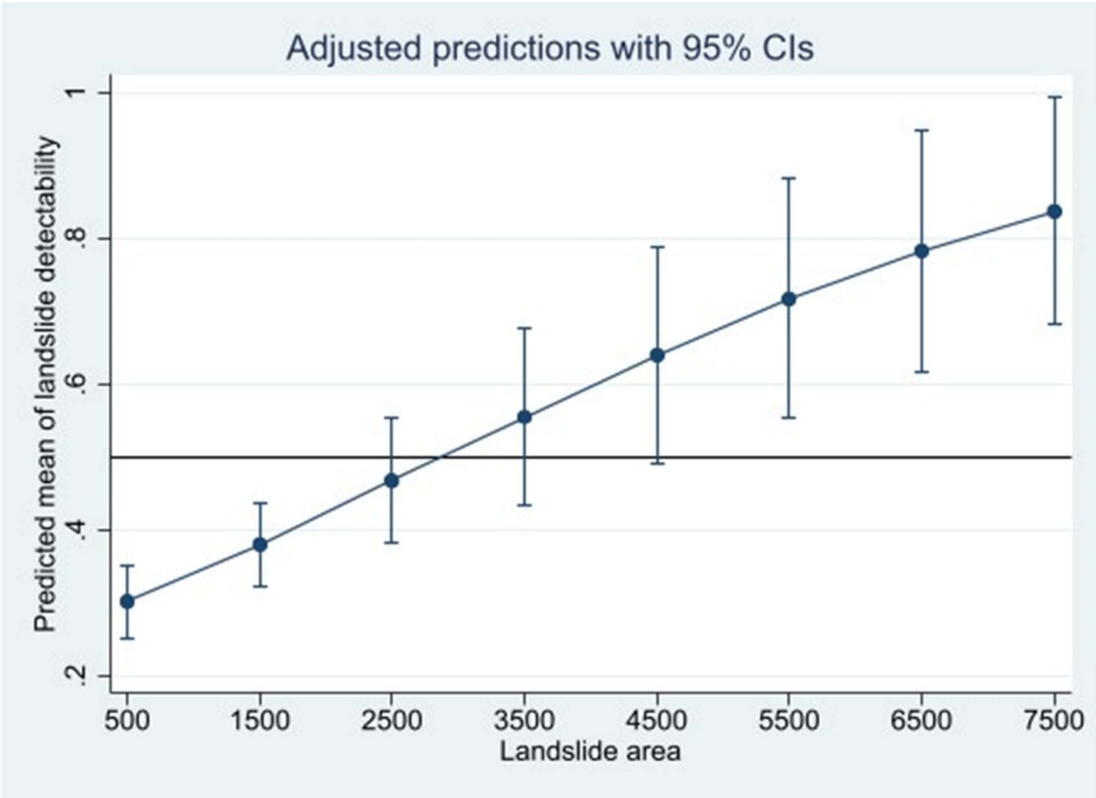

Figure 5. Marginal effect plot of landslide detectability using satellite imagery against landslide area.

ORCID iD

John Sekajugo https://orcid.org/0000-0002-8904-0898

References

Aceves-Bueno E, Adeleye A S, Feraud M, Huang Y, Tao M, Yang Y and Anderson S E 2017 The accuracy of citizen science data: a quantitative review *Bull. Ecol. Soc. Am.* **98** 278–90

Agapiou A, Hadjimitsis D G, Papoutsa C, Alexakis D D and Papadavid G 2011 The importance of accounting for atmospheric effects in the application of ndvi and interpretation of satellite imagery supporting archaeological research: the case studies of Palaepaphos and Nea Paphos sites in Cyprus *Remote Sens.* **3** 2605–29

Allison P D 2014 Measures of fit for logistic regression *SAS Glob. Forum 2014* vol 2 pp 1–12

AUC 2018 Building disaster resilience to natural hazards in Sub-Saharan African regions, countries and communities annual report annual report

Bird T J *et al* 2014 Statistical solutions for error and bias in global citizen science datasets *Biol. Conserv.* **173** 144–54

Boughorbel S, Jarray F and El-Anbari M 2017 Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric *PLoS One* **12** 1–17

Bwambale B, Nyeko M, Sekajugo J and Kervyn M 2022 The essential contribution of indigenous knowledge to understanding natural hazards and disaster risk: historical evidence from the Rwenzori (Uganda) *Nat. Hazards* **110** 1847–67

Chicco D and Jurman G 2020 The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation *BMC Genomics* **21** 1–13

Chicco D, Tötsch N and Jurman G 2021 The matthews correlation coefficient (Mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation *BioData Min.* **14** 1–22

Cohn J P 2008 Citizen science can volunteers do real research? *Bioscience* **58** 192–7

Coughlin S S, Trock B, Criqui M H, Pickle L W, Browner D and Tefft M C 1992 The logistic modeling of sensitivity, specificity, and predictive value of a diagnostic test *J. Clin. Epidemiol.* **45** 1–7

CRED 2017 Economic losses, poverty & disasters, 1998-2017. 01 pp 1–7

Cunha D G F, Marques J F, de Resende J C, de Falco P B, de Souza C M and Loiselle S A 2017 Citizen science participation in research in the environmental sciences: key factors related to projects' success and longevity *An. Acad. Bras. Cienc.* **89** 2229–45

Danielsen F *et al* 2014 Linking public participation in scientific research to the indicators and needs of international environmental agreements *Conserv. Lett.* **7** 12–24

Dewitte O, Depicker A, Moeyersons J and Dille A 2022 Mass movements in tropical climates *Ref. Modul. Earth Syst. Environ. Sci.* 338–49

Dickinson J L, Shirk J, Bonter D, Bonney R, Crain R L, Martin J, Phillips T and Purcell K 2012 The current state of citizen science as a tool for ecological research and public engagement *Front. Ecol. Environ.* **10** 291–7

Dickinson J L, Zuckerberg B and Bonter D N 2010 Citizen science as an ecological research tool: challenges and benefits *Annu. Rev. Ecol. Evol. Syst.* **41** 149–72

Eitzel M V *et al* 2017 Citizen science terminology matters: exploring key terms *Citizen Sci. Theory Pract.* **2** 1

Flach P A and Kull M 2015 Precision-Recall-Gain curves: PR analysis done right *Adv. Neural Inf. Process. Syst.* **1** 838–46

Genders T S S, Spronk S, Stijnen T, Steyerberg E W, Lesaffre E and Hunink M G M 2012 Methods for calculating sensitivity and specificity of clustered data: a tutorial *Radiology* **265** 910–6

Haklay M *et al* 2021 Contours of citizen science: a vignette study *R. Soc. Open Sci.* **8** 0–2

Hamdan F 2015 Intensive and extensive disaster risk drivers and interactions with recent trends in the global political economy, with special emphasis on rentier states *Int. J. Disaster Risk Reduct.* **14** 273–89

Hao L, Rajaneesh A, Van Westen C, Sajinkumar K S, Martha T R, Jaiswal P and McAdoo B G 2020 Constructing a complete landslide inventory dataset for the 2018 monsoon disaster in Kerala, India, for land use change analysis *Earth Syst. Sci. Data* **12** 2899–918

Haworth B T 2016 Assessing the potential, application, and implications of volunteered geographic information in disaster risk reduction *Geography* The University of Sydney

Haworth B T, Bruce E, Whittaker J and Read R 2018 The good, the bad, and the uncertain: contributions of volunteered geographic information to community disaster resilience *Front. Earth Sci.* **6** 1–15

Hicks A, Barclay J, Chilvers J, Armijos M T, Oven K, Simmons P and Haklay M 2019 Global mapping of citizen science projects for disaster risk reduction *Front. Earth Sci.* **7** 1–18

IFRC 2016 *Resilience: saving lives today, investing for tomorrow World disasters report* International Federation of Red Cross

Jacobs L, Dewitte O, Poesen J, Maes J, Mertens K, Sekajugo J and Kervyn M 2016a Journal of African Earth Sciences Landslide characteristics and spatial distribution in the Rwenzori *J. Afr. Earth Sci.* **134** 917–30

Jacobs L, Kabaseke C, Bwambale B, Katutu R, Dewitte O, Mertens K, Maes J and Kervyn M 2019 The geo-observer network: a proof of concept on participatory sensing of disasters in a remote setting *Sci. Total Environ.* **670** 245–61

Jacobs L, Maes J, Mertens K, Sekajugo J, Thiery W, van Lipzig N, Poesen J, Kervyn M and Dewitte O 2016b Reconstruction of a flash flood event through a multi-hazard approach: focus on the Rwenzori Mountains, Uganda *Nat. Hazards* **84** 851–76

Jain S, Khosa R and Gosain A K 2021 Landslides hazard mapping using high-resolution satellite data *Lect. Notes Civ. Eng.* **86** 95–110

Ji S, Yu D, Shen C, Li W and Xu Q 2020 Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks *Landslides* **17** 1337–52

Kullenberg C and Kasperowski D 2016 What is citizen science?—A scientometric meta-analysis *PLoS One* **11** 1–16

Lewandowski E and Specht H 2015 Influence of volunteer and project characteristics on data quality of biological surveys *Conserv. Biol.* **29** 713–23

McKinley D C *et al* 2017 Citizen science can improve conservation science, natural resource management, and environmental protection *Biol. Conserv.* **208** 15–28

Monsieurs E *et al* 2018 Landslide inventory for hazard assessment in a data-poor context: a regional-scale approach in a tropical African environment *Landslides* **15** 2195–209

Monsieurs E, Dewitte O and Demoulin A 2019 A susceptibility-based rainfall threshold approach for landslide occurrence *Nat. Hazards Earth Syst. Sci.* **19** 775–89

Nishadi A S T 2019 Predicting heart diseases in logistic regression of machine learning algorithms by Python Jupyterlab *Int. J. Adv. Res. Publ.* **3** 69–74

Papathoma-Köhle M, Promper C and Glade T 2016 A common methodology for risk assessment and mapping of climate change related hazards-implications for climate change adaptation policies *Climate* **4** 8

Paul K, Quinn M S, Huijser M P, Graham J and Broberg L 2014 An evaluation of a citizen science data collection program forrecording wildlife observations along a highway *J. Environ. Manage.* **139** 180–7

Peel M C, Finlayson B L and McMahon T A 2007 Updated world map of the Köppen-Geiger climate classification *Hydrol. Earth Syst. Sci.* **11** 1633–44

Pernat N, Kampen H, Jeschke J M and Werner D 2021 Citizen science versus professional data collection: comparison of approaches to mosquito monitoring in Germany *J. Appl. Ecol.* **58** 214–23

Phillips T, Porticella N, Constas M and Bonney R 2018 A framework for articulating and measuring individual learning outcomes from participation in citizen science *Citizen Sci. Theory Pract.* **3** 3

Planet 2016 1610.06_Spec Sheet_Combined_Imagery_Product_Letter_ENGv1. Planet Labs

Planet 2019 *Planet Imagery Product Specification* (Santa Marta: Planet Labs Inc) p 56

Planet 2021 *Planet Imagery Product Specifications* (Papua Province: Planet Labs Inc) pp 1–56

Pondard N and Daly M 2011 *Natural hazards risk modelling: an approach providing risk management solutions for local government* GNS Miscellaneous Series 38 Institute of Geological and Nuclear Sciences

Powers D M W 2008 Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation

Rabonza M L, Felix R P, Lagmay A M F A, Eco R N C, Ortiz I J G and Aquino D T 2016 Shallow landslide susceptibility mapping using high-resolution topography for areas devastated by super typhoon Haiyan *Landslides* **13** 201–10

Robinson T R, Rosser N and Walters R J 2019 The spatial and temporal influence of cloud cover on satellite-based emergency mapping of earthquake disasters *Sci. Rep.* **9** 1–9

Rohan T J, Wondolowski N and Shelef E 2021 Landslide susceptibility analysis based on citizen reports *Earth Surf. Process. Landf.* **46** 791–803

Samodra G, Chen G, Sartohadi J and Kasama K 2018 Generating landslide inventory by participatory mapping: an example in Purwosari Area, Yogyakarta, Java *Geomorphology* **306** 306–13

Selva M 2021 The quality of remote sensing optical images from acquisition to users *Remote Sens.* **13** 10–13

Silvertown J 2009 A new dawn for citizen science *Trends Ecol. Evol.* **24** 467–71

Talisay B A M, Puno G R and Amper R A L 2019 Flood hazard mapping in an urban area using combined hydrologic-hydraulic models and geospatial technologies *Glob. J. Environ. Sci. Manage.* **5** 139–54

Tharwat A 2018 Classification assessment methods *Appl. Comput. Inform.* **17** 168–92

Thirumurugan P and Krishnaveni M 2019 Flood hazard mapping using geospatial techniques and satellite images—a case study of coastal district of Tamil Nadu *Environ. Monit. Assess.* **191** 31

Tjur T 2009 Coefficients of determination in logistic regression models—a new proposal: the coefficient of discrimination *Am. Stat.* **63** 366–72

Ubaidulloev A, Kaiheng H, Rustamov M and Kurbanova M 2021 Landslide inventory along a national highway corridor in the Hissar-Allay Mountains, Central Tajikistan *GeoHazards* **2** 212–27

UNDRR 2015 Global assessment report: loss data and extensive risk analysis pp 1–10

UNDRR 2019a *Global Assessment Report on Disaster Risk Reduction* (Geneva: United Nations Office for Disaster Risk Reduction (UNDRR))

UNDRR 2019b DesInventar Sendai User Manual Data Management pp 1–57

UNISDR 2015a Making development sustainable: the future of disaster risk management Global Assessment Report on Disas- ter Risk Reduction

UNISDR 2015b Sendai framework for disaster risk reduction 2015-2030

Wang S, Zhou Y, Musili P M, Mwachala G, Hu G and Wang Q 2020 Inventory incompleteness and collecting priority on the plant diversity in tropical East Africa *Biol. Conserv.* **241** 108313

Zhang G and Zhu A X 2018 The representativeness and spatial bias of volunteered geographic information: a review *Ann. GIS* **24** 151–62

Zhou Y, Varquez A C G and Kanda M 2019 High-resolution global urban growth projection based on multiple applications of the SLEUTH urban growth model *Sci. Data* **6** 1–10