

## Eliciting social themes of flood mitigation and community engagement studies through text mining.

### Abstract

**Purpose:** Flood preparedness and response from the perspective of community engagement mechanisms have been studied in scholarly articles. However, the differences in flood mitigation may expose social and behavioural challenges to learn from. This study aimed to demonstrate how text mining can be applied in prioritising existing contexts in community-based and government flood mitigation and management strategies.

**Design/Methodology/Approach:** This investigation mined the semantics researchers ascribed to flood disasters and community responses from 2001 to 2022 peer-reviewed publications. Text mining was used to derive frequently used terms from over 15 publications in the Scopus database and Google Scholar search engine after an initial output of 268 peer-reviewed publications. The text mining process applied the topic modelling analyses on the 15 publications using the R studio application.

**Findings:** Topic modelling applied through text mining clustered four (4) themes. The themes that emerged from the topic modelling process were *building adaptation to flooding; climate change and resilient communities; urban infrastructure and community preparedness; and Research output for flood risk and community response*. The themes were supported with geographical flood risk and community mitigation contexts from the US, India and Nigeria to provide a broader perspective.

**Originality/value:** This study exposed the deficiency of “communication, teamwork, responsibility and lessons” as focal themes of flood disaster management and response research. The divergence of flood mitigation in developing nations as compared with developed nations can be bridged through improved government policies, technologies and community engagement.

**Keywords:** Climate change, Community Engagement, Disaster resilience, Flood, Text mining.

### 1. Introduction

Flooding has become major evidence of climate change around the world (Gavin, Leonard-Milsom and Montgomery, 2011; Albright and Crow, 2019, 2019). Emergency Event Database (2022) documented 432 disasters in 2021, higher than the average of 357 from 2001 to 2020. Flooding incidents occurred 223 times in 2021, an upsurge from an annual average of 163 incidents between 2001 and 2020 (Emergency Event Database, 2022). Furthermore, in 2021, 1,282 people died in Indian floods, with economic losses to individuals and the government (Reliefweb, 2022). More recently, flooding claimed over 600 lives in Nigeria and displaced 1.4 million people in 2022 (Reliefweb, 2022). In the United States, flooding through hurricanes and heavy rainfall have led to adaptive measures to counter the unpredictable impacts through collaborative flooding models that engage communities and government stakeholders (Wobus *et al.*, 2019; Sanders *et al.*, 2020). Research and studies into flood mitigation and responses in vulnerable communities such as coastal towns can save lives (Eccles, Zhang and Hamilton, 2019; Avashia and Garg, 2020). Government intervention and plans for community resilience place imperative responsibilities on specific stakeholders, documentation, and plans. Studies into how government and communities respond to flood incidents can foster an improved

1  
2  
3 decision-making (Liu and Li, 2016; Jiménez-Hernández *et al.*, 2021). Geographical  
4 information systems (GIS) and internet-of-things (IoTs) have contributed immensely to  
5 tracking, monitoring and supporting community responses to flooding (Agrawal *et al.*, 2006;  
6 Liu and Li, 2016). The application of technology and research into flood disaster response  
7 required an alignment with community engagements.  
8  
9

10 Community engagement is more challenging in rural areas because of the socioeconomic  
11 situation of residents in such geographical regions (Henderson *et al.*, 2020). In floodplains and  
12 coastal areas around the world, flooding is more prevalent within such communities (Dufty,  
13 2017). Compared with urban regions far from coastal plains, communities, rural areas and  
14 floodplains experience psychological issues and are more affected by the flooding experience.  
15 Dufty (2017) further noted that risk perceptions and self-efficacy are important community  
16 factors that may aid residents of flood-affected communities. Community engagement in flood  
17 incidents must be aimed at saving the lives of flood victims through rescues and providing  
18 temporary shelter, food and basic amenities. Salman (2018) classified the impact of flooding  
19 into economic, non-economic and type of loss as direct and indirect. Direct economic losses  
20 produce Material loss, which includes residential development, industrial development and  
21 agricultural and economic loss affecting the socioeconomic structure of the affected area.  
22 Indirect economic losses are fiscal aid from all stakeholders, including the public and private  
23 sectors (Nkwunonwo, Whitworth and Baily, 2016; Gardiner, Herring and Fox, 2019; Avashia  
24 and Garg, 2020). The non-economic impacts of flooding affect the mental health of the victims,  
25 the influences of which last longer than the responses provide and direct economic losses  
26 (Miller and Hutchins, 2017; Avashia and Garg, 2020). The implications of climate change and  
27 the expected rise in global temperatures necessitates community orientations and preparedness  
28 for eventualities such as flooding. Flooding is a natural disaster which cannot be prevented but  
29 mitigated with an appropriate response. Wobus *et al.* (2019) and Sanders *et al.* (2020) identified  
30 sacrosanct tools such as databases, data from GIS mapping and studies into how communities  
31 and government stakeholders have managed and responded to flooding incidents.  
32  
33  
34  
35

36 Further studies into the plethora of flood and community engagement research will contribute  
37 further knowledge to addressing future incidents. Additionally, previous studies conducted by  
38 Liu and Li (2016); Bevacqua *et al.*, (2020); Yusuf *et al.* (2021); Li and Wang (2022); Avashia  
39 and Garg, (2020); Henderson *et al.*, (2020); and Jiménez-Hernández *et al.*, (2021) on  
40 community engagement have not considered methodological aspects involving machine  
41 learning and social responses. No, community engagement studies have delved into the use of  
42 machine-learning applications to understand how flood victims. Hence, the essence of this  
43 investigation is to create a broader view of how different countries respond to flooding from a  
44 community perspective using a machine-learning approach.  
45  
46  
47

48 This study aims to demonstrate how text mining can be applied in determining existing focal  
49 themes in community flood mitigation and response to juxtapose the results with cases from  
50 the United States, India and Nigeria.  
51  
52  
53

## 54 **2.0 A review of social contexts of community engagement**

55 Community resilience in this study is the ability of neighbourhoods to withstand, adapt and  
56 recover from shocks and stresses in events of disaster (Adebimpe *et al.*, 2018; Mohanty,  
57 Mudgil, and Karmakar, 2020; Yusuf *et al.*, 2021). Community engagement before and after  
58 disasters involves diverse social, economic, environmental and political factors (Campbell-  
59 Arvai and Lindquist, 2021; Puzyreva and de Vries, 2021). Social issues are the underlying  
60

1  
2  
3 typologies that shape community resilience. These themes include social cohesion, trust, social  
4 capital, and community capacity. A growing body of research suggests that social issues play  
5 an important role in community resilience. Babcock and Seebauer (2017) explained that  
6 communities with strong social cohesion are more likely to recover from disasters. Strong  
7 social cohesion aids communities to mobilise resources, share information and support each  
8 other during difficult times. There are different approaches to reinforcing social issues within  
9 the community. Henderson et al. (2020) and Mehring et al. (2018) suggested that building trust  
10 and relationships among community members and encouraging community participation in  
11 decision-making are essential to strengthening social cohesion in flood response. Additional  
12 social support through community understanding, organising members into groups and getting  
13 to know each other is a fundamental resilience typology (Meenar et al., 2018; Pandey, 2019).  
14 Studying social issues in communities can support resilience to shocks resulting from disasters  
15 (Mulvihill, 2023). Knowledge of social themes can protect communities from the negative  
16 effects of disasters. Furthermore, examining the nuances of community resilience to floods,  
17 particularly in the context of social issues, can help ensure that the factors that contribute to  
18 resilience in disasters such as floods are understood to be implemented (Babcock & Seebauer,  
19 2017; Ciampa et al., 2021; Henderson et al., 2020; Mehring et al., 2018; Meenar et al., 2018;  
20 Pandey, 2019). This research contributes to understanding how communities can become more  
21 sustainable and resilient to floods.  
22  
23  
24  
25

26 Additionally, previous studies have used traditional research methods such as qualitative,  
27 quantitative and mixed-method approaches (Babcock & Seebauer, 2017; Ciampa et al., 2021;  
28 Henderson et al., 2020; Mehring et al., 2018; Meenar et al., 2018; Pandey, 2019). These studies  
29 also incorporated research strategies such as focus groups, literature reviews, surveys, and  
30 interviews into the data collection. Secondary data on community resilience to floods are  
31 abundant. However, modern analytical approaches such as machine learning are lacking for a  
32 deeper understanding of social issues that may limit the development of community resilience.  
33 The next section discusses text mining as a machine-learning tool.  
34  
35

### 36 **2.1 Text mining applications for disaster resilience**

37  
38 Text mining, also known as text data analytics, is the process of extracting patterns and useful  
39 textual details in terms of words and topics from written words (Ahadi *et al.*, 2022; Tavana *et*  
40 *al.*, 2022; Thakur and Kumar, 2022). Text mining is a machine learning (ML) approach for  
41 analysing unstructured words through structured functions. Ahadi *et al.* (2022) further  
42 explained that analytical methods such as natural language processing (NLP) could be applied  
43 as a range of computational techniques for evaluating and presenting naturally occurring texts  
44 for high-quality texts. Thakur and Kumar (2022) corroborate Ahadi *et al.* (2022) description of  
45 text mining by noting that it is a statistical analysis for discovering latent information, patterns,  
46 trends, and duplication to foster enhanced conceptual understanding of unstructured words.  
47 Text mining and systematic and bibliometric analysis have been applied to strengthen the  
48 analysis and produce new knowledge (Ahadi *et al.*, 2022, 2022; Tavana *et al.*, 2022). Unlike  
49 systematic and bibliometric reviews where a structure is applied to narrow down the nature of  
50 the context in the study using metadata and topics, text mining uses ML to produce outputs.  
51 Text mining can be applied to word frequency identification, network analysis, co-occurrences,  
52 collocations, topic modelling and visualisation of texts.  
53  
54  
55  
56

57 Text mining has been applied to community information from social media platforms to elicit  
58 the frequency (TF) of flood disasters (Chen and Wang, 2022). Ngamassi *et al.* (2022) used text  
59 data mining from tweets about Hurricane Harvey to understand the needs of people in locations  
60

1  
2  
3 affected by Hurricane Harvey. AlAbdulaali *et al.* (2022) created a dashboard for multimodal  
4 disaster information using social media data and metadata to visualise disasters and support  
5 emergency response. Text mining has been applied in disaster risk reduction, to understand  
6 opinions and requirements for disaster relief and predict responses (Li and Wang, 2022;  
7 Linardos *et al.*, 2022; Ong, Pauzi and Gan, 2022). Accordingly, text mining is a qualitative  
8 data analysis approach in ML. The data sources for text mining may be large, medium, or small,  
9 mostly taken from social media platforms such as Twitter and Facebook posts. The quality of  
10 textual data in most studies conducted by Li and Wang (2022); Linardos *et al.* (2022); Ong,  
11 Pauzi and Gan (2022); Linardos *et al.* (2022); Ong, Pauzi and Gan, (2022); Ngamassi *et al.*  
12 (2022); and Chen and Wang, 2022 are not peer-reviewed publications, but valid because they  
13 are taken directly from people affected by disasters. The issue of fake news and incorrect  
14 information is also difficult to identify in text scraping from social media, and the reliability of  
15 textual data for mining purposes needs to be addressed. This study will only use peer-reviewed  
16 publications and verified government reports in the text-mining process.  
17  
18  
19

20  
21 Text mining is a machine learning tool that is especially useful for studying the resilience of  
22 communities after floods. The strength of text mining as a statistical analysis tool lies in its  
23 ability to quickly and efficiently analyse large amounts of information and potentially identify  
24 themes and patterns that would otherwise go unnoticed (Feldman and Sanger, 2007). This  
25 capability can assist in anticipating flood-related risks and community needs and enable rapid  
26 response to emergencies (Pang and Lee, 2008). Text mining can struggle with context and  
27 nuances (Hotho *et al.*, 2005). Thus, leading to potential misunderstandings in terms of outputs.  
28 Text mining relies heavily on the quality and breadth of source material (Aggarwal and Zhai,  
29 2012), so biased or incomplete data limits its usefulness. Therefore, extracting data of high  
30 quality rather than quantity is essential.  
31  
32  
33

### 34 **3.0 Method and material**

35 The publication search for text mining followed the systematic review guidelines developed  
36 by Higgins *et al.* (2019). The data mining conceptual model designed for this study in Figure  
37 1 commenced with extracting data from Scopus and Google scholar. The keyword search in  
38 Scopus and Google Scholar inputted “*flood AND disaster AND mitigation AND response AND*  
39 *community AND engagement*”. The keywords were derived from the contexts of the study and  
40 aim, as stated in the introduction and critical review in section 2.0. A combined 3,762 results  
41 were obtained. The inclusion criteria from 2001 to 2022 for all publications in books, book  
42 chapters, articles, and conference papers that were all peer-reviewed reduced the list to 268  
43 publications. Upon carefully reviewing the metadata for the results and relevance of the  
44 publications, only 15 peer-reviewed journal publications were downloaded in PDF format for  
45 text mining using R studio version 4.2.3. The criteria for selecting 15 journals were based on  
46 relevance and context to flooding and community engagement. The context of the final text  
47 data extracted from Scopus and Google Scholar focused on the aim of the study, which is to  
48 understand flood disaster mitigation and community engagement, including the response.  
49 Consequently, the scope of the data was from other countries with a focus on cases from the  
50 United States, Nigeria and India to represent countries having recent incidents of global  
51 flooding in recent years. The choice of the USA, Nigeria and India in these studies may limit  
52 the scope of the text mining but is justified by the prevalence of annual flood events and the  
53 existing body of knowledge (Mishra *et al.*, 2018; Brunner *et al.*, 2020; Buba *et al.*, 2021).  
54 Furthermore, text mining differs from bibliometric or scientometric analysis because it applies  
55 ML to analyse words, create clusters, analyse sentiments, create networks and relationships  
56 between terms.  
57  
58  
59  
60

>>>Insert Figure 1<<<

For instance, the years and number of citations should have been looked into when considering text mining. The publication titles, abstracts, and keywords form the basis of the textual data required for the mining activity in R studio. Unsupervised text mining applied the topic modelling to create themes from the identified files in Table 1. Rather the focus was on the most frequently occurring words, the sparsity of the words and network analysis showing the connection of the words to form the themes in this study. In achieving the final stage in Figure 1, topic modelling coding in R studio was conducted. A detailed analysis of text mining is explained in the next section.

>>>Insert Table 1<<<

From Table 1, the documents identified from the systematic review search were evenly distributed across the US, India and Nigeria. The selected publications were about flood risks, government interventions, previous research and databases applied to respond to flooding, GIS, new residential building construction approach and community enlightenment.

#### 4.0 Text mining analyses: Topic modelling

Stage 2 in Figure 1 conceptually highlights this study's broader steps towards text mining. The unsupervised topic modelling analysis used functions already encoded in R studio combined with the variables.

The text frequency is the number of most occurring words in the document, which was essential in the topic modelling using the beta and gamma values to classify the outputs under an unsupervised ML operation (Silge and Robinson, 2017). Beta and gamma distribution of the frequent terms in the 15 documents reveals the generalised uniformity of the terms in terms of intervals of 0 and 1 (beta), and all non-negatives are modelled in gamma (Jambunathan, 1954; Dubey, 1970). The next section details the coding steps taken in Figure 2 to arrive at the beta and gamma topic models.

#### 4.1 Steps in text mining

This section explains how the topic modelling script in Figure 2 is developed in R studio to produce the term model using the beta and gamma distributions for flood disaster risk management and community engagement.

**Step 1.** The library functions “(topic models)” for topic modelling, “(pdf tools)” for reading PDF files, “(tm)” for text mining, “(dplyr)” for data manipulation, “(ggplot2)” for bar charts and visualisation, and “(tidytext)” was made active for data cleaning. The require() or library() attribute can load all required functions.

**Step 2.** Before running loading the PDF data file, the working directory must be set. The “list.files ()” function is used to load the PDF file into R studio. A vector is created for the PDF file at this stage using the “lapply” function.

**Step 3.** Vectors are a collection of strings, logical values and arithmetic expressions. In this regard, the paragraphs are converted into a spreadsheet within R studio. Thus, 26 observations from 15 documents were produced from the paragraphs in the PDF file.

**Step 4.** The corpus is the next category of the script in R studio after creating the vectors. A corpus is a collection of word texts as part of the NLP. The “*Corpus()*” function is used to create the word texts.

**Steps 5.** Data cleaning is conducted using the “*tm\_map()*” to remove all stop words, numbers, whitespaces, and punctuations. This process is important for the tokenization of the word texts in preparation for the Term document matrix.

**Step 6.** The term document matrix is an approach to converting word texts into matrices. Once a table has been created in the Corpus and cleaned by removing all stop words, numbers and punctuations in the file. The “*DocumentTermMatrix()*” (DTM) function is used to filter and inspect the sparsity of specific words in a matrix format.

**Step 7.** To classify the DTM function and words into words into group means, the linear discriminant analysis function known as “*lda()*” is applied to investigate the probability of each term belonging to a group mean.

**Step 8.** The beta distributions are computed using the functions “*tidy()*”, “*group\_by()*”, “*slice\_max()*”, “*ungroup()*”, and “*arrange()*” to create the topic models based on the beta value distributions.

**Step 9.** The beta values are visualised in bar plots using the “*ggplot()*” function. The “*mutate()*” and “*facet\_wrap()*” functions were essential in this visualisation. The “*tidy()*” functions were applied in visualisation and for the gamma distribution analysis.

>>>Insert Figure 2<<<

## 5.0 Results and discussions

The results from the topic modelling ML analysis from Figure 2 are presented as bar plots, line graphs and tables, as expressed in this section.

### 5.1. Topic models: beta and gamma analysis

The purpose of the topic modelling analysis approach in text mining is to categorise into clusters, with the aid of unsupervised ML words with commonalities, to create new themes. The unsupervised ML produces outputs and patterns based on unlabelled training data. In this analysis, four (4) topics were produced as the output of 15 documents, as illustrated in Figure 3. In examining topic 1, flood and risk have a beta value of 0.0225 and 0.060, respectively. These words are more dominant under topic 1 and were combined with the remaining words, such as maps, hazard, building and information, producing the theme “*Building adaptation to flooding*” in Table 2. The dominant word in topic 2 is “*climate*” (beta value = 0.0090), followed by change (beta value = 0.0070). These dominant words combined cost, resilience, management, future, damages, flood, and data to synthesise the “*climate change resilient communities*” theme. Under topic 3, flood is repeated as the term with the highest beta value of 0.0225, risk (beta value = 0.0075), and urban (beta value = 0.007). The theme “*Urban and community preparedness*” was created under topic 3 by compounding the terms above under

1  
2  
3 this topic with flooding, city, water, management, and cities. In topic 4, the “*research output for flood risk and community response*” theme was produced from flood and risk with the same  
4 beta value as the other topics, management with a beta value of 0.007, and “*study*” with a beta  
5 value of 0.065.  
6  
7

8  
9 The topic model terms indicating the beta values from Figure 3 were developed into themes in  
10 Table 2. However, Figure 3 only expressed the direct terms from each document in clusters.  
11 Further distribution of the documents under the topic models is presented using the gamma  
12 distribution in Figure 4.  
13

14 >>>Insert Figure 3<<<

15  
16 >>>Insert Table 2<<<

17  
18  
19  
20  
21 Appendix A signified 60 gamma distributions across 4 topic models for 15 documents inputted  
22 in this study. Document 11 emerged dominant for topic 1 (Building adaptation to flooding)  
23 with a Gamma value of 5.32E-02. Topic 2 (Climate change resilient communities) emerged  
24 from document 13 with a gamma value of 5.55E-06. A gamma value of 9.96E-01 was  
25 associated with document 7 for topic 3 (Urban infrastructure and community preparedness).  
26 Document 13, with a gamma value of 5.55E-06, is dominant in topic 4 (Research output for  
27 flood risk and community response).  
28  
29

30 >>>Insert Figure 4<<<

### 31 32 33 **5.1.1. Topic 1- Building Adaptation to flooding**

34 The gamma chart showing the factors (topics models from Figure 3) illustrates the four charts  
35 and the number of major documents within the topics. For further clarity, in topic 1 under  
36 Figure 4, documents 2, 13 and 14 contributed to topic 1. Table 1, document 2 by Mohanty,  
37 Mudgil and Karmakar (2020) addressed flood management, socio-economic implications and  
38 community engagement. Document 13 was Dhiman *et al.* (2019)’s article on community  
39 engagement for flood risk adaptations. Document 14 concerns Flood management practices by  
40 Kumar *et al.* (2021). The contexts of the documents align with the “*Building adaptation to*  
41 *flooding*” theme, whereby buildings within communities must adapt to the hazard of flooding.  
42 Flooding incidents are increasing globally and will not reduce because of climate change.  
43 Therefore, for communities to adapt effectively to flooding, existing buildings need retrofitting  
44 measures to adapt to disasters like flooding.  
45  
46  
47

### 48 **5.1.2. Topic 2- Climate change resilient communities**

49 The gamma chart of Figure 4 consists of topic 2, which comprises documents 4, 12 and 15.  
50 Islam *et al.* (2022) discussed living with flood to strengthen the communities’ resilience against  
51 flood hazards in Document 4. Gardiner, Herring and Fox (2019) postulated a climate resilience  
52 toolkit for flooding in document 12. Document 15 represents Brunner *et al.* (2020)’s study on  
53 regional flood susceptibilities emerging in river basins with catchments sharing similar  
54 streamflow and climatic regimes. Documents 4, 12 and 15 all contributed to defining the  
55 contexts of the “*climate change resilient communities*” theme. As topic 1 highlighted building  
56 adaptations in communities affected by floods, topic 2 fosters this study by reiterating the  
57 resilience of communities in preparation for the effects of climate change.  
58  
59  
60

### 5.1.3. Topic 3- Urban Infrastructure and community preparedness

The significant components of topic 3 are documents 3, 5, 7 and 11. Document 3 consists of *al. Buba et al. (2021)* article addresses community awareness that can complement traditional flood research for more effective response and mitigation strategies. The study conducted by *Buba et al. (2021)* suggested the application of GIS and probable building and community adaptations to the impacts of natural disasters. *Alam et al. (2018)* investigated flood risk assessment in communities and the application of previous studies in preparing communities for future floods, as contained in Document 5. Document 7 in Table 1 contains *Adebimpe et al. (2018)*'s research on a guide developed to aid government agencies and community stakeholders in achieving resilience integration in building Nigeria and its surroundings. *Nkwunonwo, Whitworth and Baily (2016)* considered the impact of urban risk flood management and its implications on physical, psychological and economic damage to millions of people in Document 11. All 4 documents under topic 3 led to the urban and community preparedness theme that enshrines the contexts of topics 1 and 2. Hence, preparedness for building designs and refurbishments is essential for the built environments of urban areas and local communities to become resilient to the impact of flooding. *Alam et al. (2018)* suggested that community orientation on mitigating and responding to flooding must leverage previous community preparedness research encompassing data and GIS mapping tools.

### 5.1.4. Topic 4- Research output for flood risk and community response

The documents that informed topic 4 (Research output for flood risk and community response) are 1, 8 and 10. Document 1, produced by *Molino, Kenney and Sutton-Grier (2020)*, studied coastal communities and ecosystems facing the impact of climate change. *Sanders et al. (2020)* considered flood hazard classification high-resolution maps for resilient communities in Document 8. *Brisibe and Pepple (2018)*, in Document 10, addressed the application of learning lessons from previous flood events for community resilience. The studies conducted by *Molino, Kenney and Sutton-Grier (2020)*, *Sanders et al. (2020)* and *Brisibe and Pepple (2018)* all advocated the need to delve into previous studies on flood risks and community responses as a tool for effectively supporting community response to flooding. Therefore, topic 4 can only exist with topics, 1,2 and 4.

The next section validates the findings using geographical narratives of the United States, India and Nigeria, as identified in Table 1, and the topic models to bolster the social themes identified above.

## 6.0 Validation of findings

### 6.1 Implications of flood risk and community response from the United States

Coastal communities, such as Hawaii, Caramel-on-the-sea in California, and Miami, Florida, in the US, are prone to flooding because of climate change (*Gardiner, Herring and Fox, 2019*; *Wobus et al., 2019*). More importantly, coastal areas must be more prepared and equipped with the tools to survive flooding. In rural communities prone to flooding, vulnerability can be reduced using non-structural solutions in the form of educational materials, partnerships between non-governmental organisations (NGOs) and local government, and a sourcebook on sustainable flood risk management (*Lumbroso, Ramsbottom and Spaliveiro, 2008*). Further, community preparedness for floods must include awareness of economic and non-economic implications for victims. The response to flood must apply a community-based approach because only communities will understand the needs of households compared with a central response from the government (*Shariff and Hamidi, 2019*).



Further research into the decentralisation and segmentation of communities prone to flooding through simulations can mitigate the impact of flood events on communities. The United States, the UK and developed countries have a better framework for managing and responding to flood risks. This is evident in the studies conducted by (Lumbroso, Ramsbottom and Spaliveiro, 2008; Benson, Lorenzoni and Cook, 2016; Tyler, Sadiq and Noonan, 2019). Decisions to participate in structural and non-structural mitigation efforts and participation in flood risk management programs such as the Federal Emergency Management Agency's Community Rating System program are made at the community level (Sadiq, Tyler and Noonan, 2019; Tyler, Sadiq and Noonan, 2019). Post-flood risk management and resilience hinged on community awareness and level of preparedness based on the previous review of the flood risk management (Munawar *et al.*, 2021). A post-risk review of previous strategies is imperative for effectively improving community flood risks. Mai *et al.* (2020) suggested improvements to flood risk policies, decision-making and capacity development for policymakers and practitioners who share responsibility. Shared responsibility between key stakeholders that are part of flood risk disaster management.

### 6.2 Community engagement and response solutions from India

The Indian National Disaster Response Force (NDRF), Army, Air Force and Navy have always been deployed to rescue and relief operations in all flooding disasters in India. The community's apt response overcame the widespread notion of passive victimhood at the time of disasters. Community resilience, strong social capital, and cohesiveness were collectively employed to respond to the flood (Joseph *et al.*, 2020). The topmost priority of the Sendai Framework is to promote community resilience against disasters (Saja, Sahid and Sutharshanan, 2020; Chisty *et al.*, 2022). Community resilience has been an important tool in disaster response and disaster management. Good emergency communications, on-time disaster response plans, and training exercises strengthen flood-prone communities.

During the flooding events, the fishermen were one of the first responders during the rescue operation. This was made possible through stream-lined communication services and SAR lives imaging to track and rescue in local areas, where the strong community response helps provide an effective response (Mishra *et al.*, 2018; Joy, Kanga and Singh, 2019). The Indian government machinery was grossly inadequate due to the unavailability of many durable canoes and boats and a workforce trained in sailing and diving (Joy, Kanga and Singh, 2019; Hunt and Menon, 2020). Relentless rain prevented proper early warning information from reaching local communities promptly. As a result, people could not respond until floods began nearby (Vishnu *et al.*, 2019; Hunt and Menon, 2020). The fisherman was contacted as a volunteer for a local search party. The issue of community flood management includes mitigation, preparation, response and recovery of all the affected population (Munawar *et al.*, 2021; Li and Wang, 2022). The role of community participation and human involvement in flood-affected areas must be considered. With a growing population in India, in rural and urban areas, human activity is evident and diverse in individual communities (Avashia and Garg, 2020). Community engagement and responses to flooding must be localised, and regional solutions are imperative.

### 6.3 Lessons learned from previous studies and flood mitigation for Nigeria

Nigeria's National Emergency Management Agency (NEMA) is responsible for responding to all national disasters, including flooding within the territory (Yusuf, Salihu, Muhammad and Sule, Ibrahim, 2021). NEMA has offered a limited response and mitigation measures for future occurrences. As early stated in the introductory section of this study, flooding in Nigeria has claimed over 600 lives and displaced millions (Reliefweb, 2022). In 2012, 363 people died,

1  
2  
3 and over 2.1 million people were displaced by flooding in Nigeria (Reliefweb, 2022). The 2022  
4 flooding response by NEMA reflects a lack of community engagement and conscientious effort  
5 to learn from a similar incident recorded in the past. The rapid impact of climate change, as  
6 evidenced by flooding in Nigeria, has aligned with studies conducted by Adebimpe *et al.*  
7 (2018); and Buba *et al.* (2021), whereby improvements of vulnerable buildings and community  
8 engagements will mitigate future flood events in Nigeria. Adebimpe *et al.* (2018); and Buba *et al.*  
9 (2021) studied previous flood events and applied developments from other countries in  
10 North America and South East Asia. Brisibe and Pepple (2018); and Yusuf *et al.* (2021) studied  
11 the idea of learning lessons from previous flood events for community resilience to take  
12 precautionary measures to increase flood resilience in Nigeria. , Nkwunonwo, Whitworth and  
13 Baily (2016) opined that Nigerian stakeholder efforts to address the challenge of flooding have  
14 been limited by a dearth of reliable data and awareness among those affected. This is predicated  
15 on the lack of knowledge about flood risk reduction. Therefore, Nigerian stakeholders,  
16 including government agencies such as NEMA, must invest in research into flood mitigation,  
17 a database for flood incidents, communities and impacts in Nigeria with an inclination to  
18 alleviate the challenges above.  
19  
20  
21  
22  
23

## 24 **7. Contributions to knowledge: Text mining in community engagement and response to** 25 **flooding**

26 Text mining was used in this study to associate research studies in flood disaster mitigation  
27 and response. The findings resulted in social themes drawn from the topic modelling approach,  
28 which resulted in 4 broad themes. From the case of the flooding cases discussed in the US,  
29 India and Nigeria, there is always a lack of communication between the stakeholders (including  
30 the victims) during a response event which generally can create a big gap in providing resources  
31 or relief in time to those in need. The aftermath of flooding can lead to further risks, such as  
32 landslides, economic downturn and public health concerns (Sadiq, Tyler and Noonan, 2019;  
33 Mai *et al.*, 2020; Munawar *et al.*, 2021). Therefore, secondary risks resulting from floods must  
34 be taken into consideration. To curtail secondary flood risks such as landslides, it is vital to  
35 utilise the active members ready to volunteer for help. These volunteers can be where relief  
36 camps were set up near places of worship, schools and community colleges to provide health  
37 care and basic needs to hundreds of people left essentially homeless due to this wave of mass  
38 destruction. Kerala saw a great response from its youth from colleges helping out in local  
39 churches and providing assistance by delivering supplies and recusing people to safer areas  
40 (Young *et al.*, 2022). NGOs set up relief camps and provided aid where needed to cover the  
41 gaps left by the government. Overall the state of Kerala could withstand the severity of the  
42 flood astonishingly well due to community participation and active volunteering.  
43  
44  
45  
46

47 Communication was not a frequent term picked up text mining approach. This is not only  
48 indicative of the limited studies on the influence of communication on flood mitigation and  
49 responses but a wider challenge in practice. The concerns emanating from linkages between  
50 stakeholders at the community, district, state, and national level of Disaster Risk Management  
51 (DRM) authorities, inadequate community-based early warning systems, improper resource  
52 inventories, and absence of training and mock drills were reported in the 2018 Kerala flood  
53 relief activities was evident (Varghese and Yadukrishnan, 2019; Lal *et al.*, 2020). Good  
54 communication with local responders through a group leader can help allocate specific areas  
55 and help the government provide more convenient resources.  
56  
57  
58  
59  
60

## 8. Conclusion and Limitations of the Study

This study investigated the application of text mining in inferring social themes used by researchers in community flood mitigation and responses and compared it with geographical applications of the perspectives. The contributions to knowledge differentiate the findings of this study using text mining to extract relevant concepts that could be fostered in communities in events of flooding. The identified social themes were building adaptation to flooding, climate change in resilient communities, urban infrastructure and preparedness, and research output for flood risk and community response.

The findings position community collaboration, communication and shared responsibility between government, NGOs and local stakeholders as sacrosanct. Teamwork plays an important role in an emergency, and it is important to strengthen communities by promoting public participation. A good way to achieve public participation is to create a platform for communities to participate in village-level disaster management planning, community-based disaster warning systems, simulation drills, first aid, and household-level disaster preparedness. Youths, divers and fishermen may reinforce community social capacities. Preference must be given to appropriate participation through selected leadership roles in the public sector. Every decision counts when lives are at stake, and proper public participation in government activities is just as important as support for relief camps.

Community resilience has emerged as a key factor in determining the outcome of disaster reduction processes. Various examples and past flooding events and threats prove that community involvement and engagement must be used for effective disaster management. To do this, it is important to get government involvement to ensure that all community members can deliver their services in a beneficial way that benefits them in the long run. This will help encourage public participation and engagement in disaster capacity building. It also helps in achieving better awareness.

Quantitative data analysis through ML can be applied to understanding the behaviour of communities towards disasters. The example of flooding in the US, India and Nigeria have been important in determining researchers' responses to disasters by studying the direct implications of previous research outcomes for continuous improvement.

## Appendix

>>>Insert Appendix A<<<

## References

- Adebimpe, O.A. *et al.* (2018) 'Developing Flood Resilient Buildings in Nigeria: A Guide', p. 8.
- Agrawal, S. *et al.* (2006) 'Physical modeling of catchment area by rapid prototyping using GIS data', *Rapid Prototyping Journal*, 12(2), pp. 78–85. Available at: <https://doi.org/10.1108/13552540610652410>.
- Ahadi, A. *et al.* (2022) 'Text Mining in Education—A Bibliometrics-Based Systematic Review', *Education Sciences*, 12(3), p. 210.

- 1  
2  
3 AlAbdulaali, A. *et al.* (2022) 'Designing Multimodal Interactive Dashboard of Disaster  
4 Management Systems', *Sensors*, 22(11), p. 4292.  
5  
6  
7 Alam, A. *et al.* (2018) 'Flood risk assessment of Srinagar city in Jammu and Kashmir, India',  
8 *International Journal of Disaster Resilience in the Built Environment*, 9(2), pp. 114–129.  
9 Available at: <https://doi.org/10.1108/IJDRBE-02-2017-0012>.  
10  
11 Albright, E.A. and Crow, D. (2019) 'Beliefs about climate change in the aftermath of extreme  
12 flooding', *Climatic Change*, 155(1), pp. 1–17.  
13  
14 Avashia, V. and Garg, A. (2020) 'Implications of land use transitions and climate change on  
15 local flooding in urban areas: An assessment of 42 Indian cities', *Land use policy*, 95, p.  
16 104571.  
17  
18 Benson, D., Lorenzoni, I. and Cook, H. (2016) 'Evaluating social learning in England flood  
19 risk management: an 'individual-community interaction' perspective', *Environmental Science  
20 & Policy*, 55, pp. 326–334.  
21  
22 Bevacqua, E. *et al.* (2020) 'More meteorological events that drive compound coastal flooding  
23 are projected under climate change', *Communications earth & environment*, 1(1), pp. 1–11.  
24  
25 Brunner, M.I. *et al.* (2020) 'How Probable Is Widespread Flooding in the United States?',  
26 *Water Resources Research*, 56(10). Available at: <https://doi.org/10.1029/2020WR028096>.  
27  
28 Buba, F.N. *et al.* (2021) 'Assessment of flood vulnerability in some communities in Lokoja,  
29 Kogi State, Nigeria, using Participatory Geographic Information Systems', *International  
30 Journal of Disaster Risk Reduction*, 55, p. 102111. Available at:  
31 <https://doi.org/10.1016/j.ijdr.2021.102111>.  
32  
33  
34  
35 Chen, C.C. and Wang, H.-C. (2022) 'Using community information for natural disaster  
36 alerts', *Journal of Information Science*, 48(5), pp. 718–732.  
37  
38 Chisty, M.A. *et al.* (2022) 'Sendai Framework for Disaster Risk Reduction (SFDRR) and  
39 disaster management policies in Bangladesh: How far we have come to make communities  
40 resilient?', *International Journal of Disaster Risk Reduction*, p. 103039.  
41  
42  
43 Dhiman, R. *et al.* (2019) 'Flood risk and adaptation in Indian coastal cities: recent scenarios',  
44 *Applied Water Science*, 9(1), p. 5. Available at: <https://doi.org/10.1007/s13201-018-0881-9>.  
45  
46  
47 Dubey, S.D. (1970) 'Compound gamma, beta and F distributions', *Metrika*, 16(1), pp. 27–31.  
48  
49 Dufty, N. (2017) 'How can we improve community engagement for floodplain management',  
50 in *Floodplain Management Australia National Conference, "Preparing for the Next Great  
51 Flood," Newcastle, NSW, Australia, May*, pp. 16–19.  
52  
53 Eccles, R., Zhang, H. and Hamilton, D. (2019) 'A review of the effects of climate change on  
54 riverine flooding in subtropical and tropical regions', *Journal of Water and Climate Change*,  
55 10(4), pp. 687–707.  
56  
57  
58 Gabriel Brisibe, W. and Dagogo Pepple, T. (2018) 'Lessons Learnt from the 2012 Flood  
59 Disaster: Implications for Post-flood Building Design and Construction in Yenagoa, Nigeria',  
60

1  
2  
3 *Civil Engineering and Architecture*, 6(3), pp. 171–180. Available at:  
4 <https://doi.org/10.13189/cea.2018.060307>.

5  
6 Gardiner, E.P., Herring, D.D. and Fox, J.F. (2019) ‘The U.S. Climate Resilience Toolkit:  
7 evidence of progress’, *Climatic Change*, 153(4), pp. 477–490. Available at:  
8 <https://doi.org/10.1007/s10584-018-2216-0>.

9  
10  
11 Gavin, N.T., Leonard-Milsom, L. and Montgomery, J. (2011) ‘Climate change, flooding and  
12 the media in Britain’, *Public Understanding of Science*, 20(3), pp. 422–438.

13  
14 Henderson, F. *et al.* (2020) ‘Challenges of community engagement in a rural area: The  
15 impact of flood protection and policy’, *Journal of Rural Studies*, 73, pp. 225–233.

16  
17 Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A  
18 (2019) *Cochrane handbook for systematic reviews*. London: John Wiley & Sons. Available  
19 at:  
20 [https://scholar.google.com/scholar\\_lookup?title=Cochrane+Handbook+for+Systematic+Reviews+of+Interventions&author=Higgins,+J.P.&author=Thomas,+J.&author=Chandler,+J.&author=Cumpston,+M.&author=Li,+T.&author=Page,+M.J.&author=Welch,+V.A.&publication\\_year=2019](https://scholar.google.com/scholar_lookup?title=Cochrane+Handbook+for+Systematic+Reviews+of+Interventions&author=Higgins,+J.P.&author=Thomas,+J.&author=Chandler,+J.&author=Cumpston,+M.&author=Li,+T.&author=Page,+M.J.&author=Welch,+V.A.&publication_year=2019) (Accessed: 24 October 2022).

21  
22  
23 Hunt, K.M. and Menon, A. (2020) ‘The 2018 Kerala floods: a climate change perspective’,  
24 *Climate Dynamics*, 54(3), pp. 2433–2446.

25  
26 Islam, A. *et al.* (2022) ‘Role of in-situ and ex-situ livelihood strategies for flood risk  
27 reduction: Evidence from the Mayurakshi River Basin, India’, *International Journal of*  
28 *Disaster Risk Reduction*, 70, p. 102775. Available at:  
29 <https://doi.org/10.1016/j.ijdr.2021.102775>.

30  
31 Jambunathan, M.V. (1954) ‘Some properties of beta and gamma distributions’, *The annals of*  
32 *mathematical statistics*, pp. 401–405.

33  
34 Jiménez-Hernández, S.B. *et al.* (2021) ‘Coastal Migration Index for Coastal Flooding Events  
35 Increased by Sea Level Rise due to Climate Change: Mexico and Cuba Case Studies’, *Water*,  
36 13(21), p. 3090.

37  
38 Joy, J., Kanga, S. and Singh, S.K. (2019) ‘Kerala flood 2018: flood mapping by participatory  
39 GIS approach, Meloor Panchayat’, *Int J Emerging Techn*, 10(1), pp. 197–205.

40  
41 Kumar, N. *et al.* (2021) ‘A Systematic Review Comparing Urban Flood Management  
42 Practices in India to China’s Sponge City Program’, *Sustainability*, 13(11), p. 6346.  
43 Available at: <https://doi.org/10.3390/su13116346>.

44  
45 Lal, P. *et al.* (2020) ‘Evaluating the 2018 extreme flood hazard events in Kerala, India’,  
46 *Remote Sensing Letters*, 11(5), pp. 436–445.

47  
48 Li, X. and Wang, Y. (2022) ‘Construction of urban flood disaster emergency management  
49 system using scenario construction technology’, *Computational intelligence and*  
50 *neuroscience*, 2022.

1  
2  
3 Linardos, V. *et al.* (2022) 'Machine Learning in Disaster Management: Recent Developments  
4 in Methods and Applications', *Machine Learning and Knowledge Extraction*, 4(2), pp. 446–  
5 473.

6  
7 Liu, C. and Li, Y. (2016) 'Measuring eco-roof mitigation on flash floods via GIS simulation',  
8 *Built Environment Project and Asset Management*, 6(4), pp. 415–427. Available at:  
9 <https://doi.org/10.1108/BEPAM-07-2015-0031>.

10  
11  
12 Lumbroso, D., Ramsbottom, D. and Spaliveiro, M. (2008) 'Sustainable flood risk  
13 management strategies to reduce rural communities' vulnerability to flooding in  
14 Mozambique', *Journal of Flood Risk Management*, 1(1), pp. 34–42.

15  
16  
17 Mai, T. *et al.* (2020) 'Defining flood risk management strategies: A systems approach',  
18 *International journal of disaster risk reduction*, 47, p. 101550.

19  
20  
21 Miller, J.D. and Hutchins, M. (2017) 'The impacts of urbanisation and climate change on  
22 urban flooding and urban water quality: A review of the evidence concerning the United  
23 Kingdom', *Journal of Hydrology: Regional Studies*, 12, pp. 345–362.

24  
25  
26 Mishra, V. *et al.* (2018) 'The Kerala flood of 2018: combined impact of extreme rainfall and  
27 reservoir storage', *Hydrology and Earth System Sciences Discussions*, pp. 1–13.

28  
29  
30  
31 Mohanty, M.P., Mudgil, S. and Karmakar, S. (2020) 'Flood management in India: A focussed  
32 review on the current status and future challenges', *International Journal of Disaster Risk  
33 Reduction*, 49, p. 101660. Available at: <https://doi.org/10.1016/j.ijdr.2020.101660>.

34  
35  
36  
37 Molino, G.D., Kenney, M.A. and Sutton-Grier, A.E. (2020) 'Stakeholder-defined scientific  
38 needs for coastal resilience decisions in the Northeast U.S.', *Marine Policy*, 118, p. 103987.  
39 Available at: <https://doi.org/10.1016/j.marpol.2020.103987>.

40  
41  
42  
43 Munawar, H.S. *et al.* (2021) 'Post-flood risk management and resilience building practices: A  
44 case study', *Applied Sciences*, 11(11), p. 4823.

45  
46  
47  
48 Ngamassi, L. *et al.* (2022) 'Text mining hurricane harvey tweet data: Lessons learned and  
49 policy recommendations', *International Journal of Disaster Risk Reduction*, 70, p. 102753.

50  
51  
52  
53 Nkwunonwo, U.C., Whitworth, M. and Baily, B. (2016) 'Review article: A review and  
54 critical analysis of the efforts towards urban flood risk management in the Lagos region of  
55 Nigeria', *Natural Hazards and Earth System Sciences*, 16(2), pp. 349–369. Available at:  
56 <https://doi.org/10.5194/nhess-16-349-2016>.

57  
58  
59  
60 Omotayo, T.S. *et al.* (2022) 'The System Dynamics Analysis of Cost Overrun Causations in  
UK Rail Projects in a COVID-19 Epidemic Era', *SAGE Open*, 12(2), p. 215824402210979.  
Available at: <https://doi.org/10.1177/21582440221097923>.

Ong, S.-Q., Pauzi, M.B.M. and Gan, K.H. (2022) 'Text Mining in Mosquito-Borne Disease:  
A Systematic Review', *Acta Tropica*, p. 106447.

Sadiq, A.-A., Tyler, J. and Noonan, D.S. (2019) 'A review of community flood risk  
management studies in the United States', *International Journal of Disaster Risk Reduction*,  
41, p. 101327.

- 1  
2  
3 Saja, A.A., Sahid, M.L. and Sutharshanan, M. (2020) 'Implementing Sendai Framework  
4 priorities through risk-sensitive development planning—A case study from Sri Lanka',  
5 *Progress in Disaster Science*, 5, p. 100051.  
6
- 7 Sanders, B.F. *et al.* (2020) 'Collaborative Modeling With Fine-Resolution Data Enhances  
8 Flood Awareness, Minimizes Differences in Flood Perception, and Produces Actionable  
9 Flood Maps', *Earth's Future*, 8(1). Available at: <https://doi.org/10.1029/2019EF001391>.  
10  
11
- 12 Shariff, N.N. and Hamidi, Z.S. (2019) 'Community-based approach for a flood preparedness  
13 plan in Malaysia', *Jambá: Journal of Disaster Risk Studies*, 11(1), pp. 1–6.  
14
- 15 Silge, J. and Robinson, D. (2017) *Text mining with R: a tidy approach*. First edition. Beijing ;  
16 Boston: O'Reilly.  
17
- 18 Tavana, M. *et al.* (2022) 'A Review of Digital Transformation on Supply Chain Process  
19 Management Using Text Mining', *Processes*, 10(5), p. 842.  
20  
21
- 22 Thakur, K. and Kumar, V. (2022) 'Application of text mining techniques on scholarly  
23 research articles: Methods and tools', *New Review of Academic Librarianship*, 28(3), pp.  
24 279–302.  
25
- 26 Tyler, J., Sadiq, A.-A. and Noonan, D.S. (2019) 'A review of the community flood risk  
27 management literature in the USA: lessons for improving community resilience to floods',  
28 *Natural Hazards*, 96(3), pp. 1223–1248.  
29  
30
- 31 Varghese, R.R. and Yadukrishnan, T.A. (2019) 'Role of social media during Kerala floods  
32 2018', *Library Philosophy and Practice*, 2754, pp. 1–15.  
33
- 34 Vishnu, C.L. *et al.* (2019) 'Satellite-based assessment of the August 2018 flood in parts of  
35 Kerala, India', *Geomatics, Natural Hazards and Risk* [Preprint].  
36  
37
- 38 Wobus, C. *et al.* (2019) 'Projecting Changes in Expected Annual Damages From Riverine  
39 Flooding in the United States', *Earth's Future*, p. 2018EF001119. Available at:  
40 <https://doi.org/10.1029/2018EF001119>.  
41
- 42 Young, J.C. *et al.* (2022) 'Social sensing of flood impacts in India: A case study of Kerala  
43 2018', *International Journal of Disaster Risk Reduction*, 74, p. 102908.  
44
- 45 Yusuf, A., Salihu, Muhammad and Sule, Ibrahim (2021) 'Estimating the Cost of Improving  
46 Households' Resilience to Floods through the Construction of a Dam in Some Parts of  
47 Jigawa State, Nigeria', *Gusau International Journal of Management and Social Sciences*,  
48 4(1), pp. 143–167.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Figures

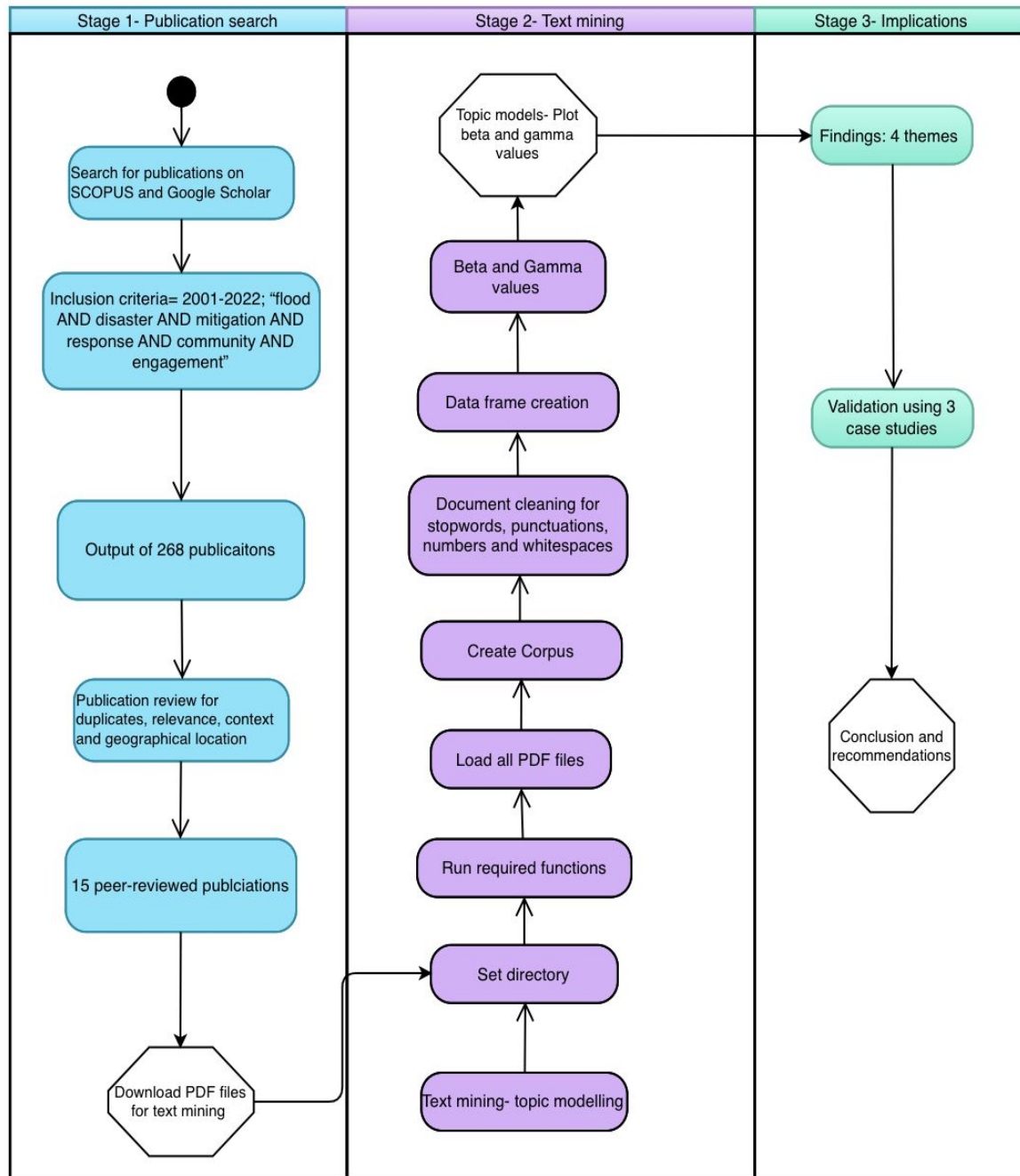


Figure 1. Data analysis conceptual model



```
1 require(topicmodels)
2 require(pdftools)
3 require(tm)
4 require(tidytext)
5 require(ggplot2)
6 require(dplyr)
7
8 # load all PDF files
9 All_Files<- list.files(pattern = "pdf$")
10 All_opinions<- lapply(All_Files, pdf_text)
11 |
12 # Create Corpus
13 document<-Corpus(VectorSource(All_opinions))
14
15 #Document cleaning
16 document<- tm_map(document, content_transformer(tolower))
17 document<-tm_map(document, removeNumbers)
18 document<-tm_map(document, removeWords, stopwords("english"))
19 document<-tm_map(document, removePunctuation, preserve_intra_word_dashes= TRUE)
20 document<-tm_map(document, stripWhitespace)
21
22 # Create DTM
23 DTM<-DocumentTermMatrix(document)
24
25 #Create LDA model
26 Model_lda<- LDA(DTM, k= 4, control =list(seed = 1234))
27 Model_lda
28 |
29 #Beta values showing probability
30 beta_topics<- tidy(Model_lda, matrix = "beta")
31 beta_topics
32
33 #Grouping terms by topic
34 beta_top_terms<- beta_topics %>%
35   group_by(topic) %>%
36   slice_max(beta, n =10) %>%
37   ungroup() %>%
38   arrange(topic, -beta)
39
40 #Display grouped terms on a chart
41 beta_top_terms %>%
42   mutate(term = reorder_within(term, beta, topic)) %>%
43   ggplot(aes(beta, term, fill = factor(topic))) +
44   geom_col(show.legend = FALSE) +
45   facet_wrap(~ topic, scales = "free") +
46   scale_y_reordered()
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```

```

47
48 #Filters terms by topics
49 tidy(DTM) %>%
50   filter(document == 3) %>%
51   arrange(desc(count))
52
53 #Examine per document per topic probability
54 gamma_documents <- tidy(Model_lda, matrix = "gamma")
55 gamma_documents
56
57
58 #Create a data frame of gamma results
59 doc_gamma.df <- data.frame(gamma_documents)
60 doc_gamma.df$chapter <- rep(1:dim(DTM) [1],4)
61
62 #plot gamma results
63 ggplot(data = doc_gamma.df, aes(x = chapter, y= gamma,
64   group = factor(topic), color = factor(topic))) +
65   geom_line()+facet_wrap(~factor(topic), ncol =1)
66

```

Figure 2. R script showing topic modelling involving beta and gamma visualisations

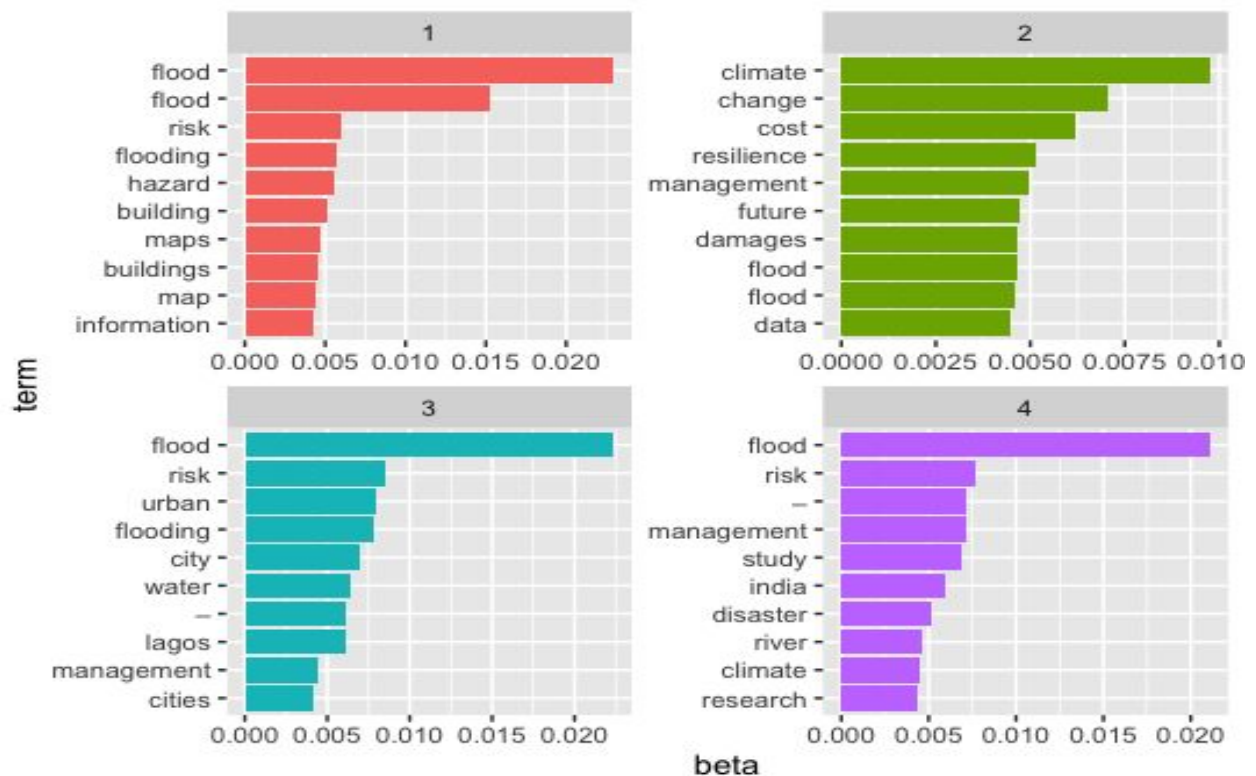


Figure 3. Beta values of topic categories for the terms

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

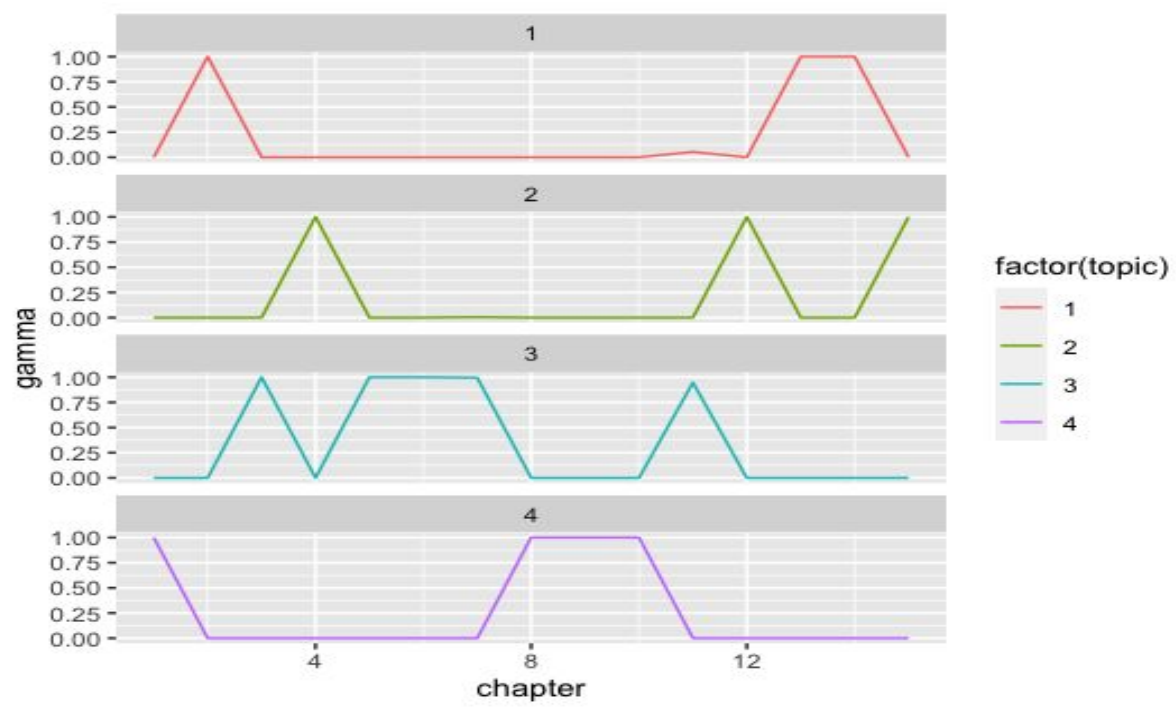


Figure 4. Gamma results for 15 documents and topics

## Tables

Table 1. Selected journal publications for text mining

Doc Nr	Context	Key findings	Study type	Geographical Location	Reference
1	Coastal communities and ecosystems facing the impact of climate change	Government agencies such as the Sea Grant play an important role in supporting stakeholder needs assessments, both in terms of funding and providing the expertise	Document analysis	US	Molino, Kenney and Sutton-Grier (2020)
2	Flood management, socio-economic implications and community engagement	Need to address India's region-specific flood issues and discuss the initiatives of major Indian flood control agencies that focus on current flood management practices	Literature review	India	Mohanty, Mudgil and Karmakar (2020)
3	Aiding flood vulnerable communities using GIS	The study concludes that affected community awareness can complement traditional flood research for more effective response and mitigation strategies.	Participatory Geographic Information System	Nigeria	Buba <i>et al.</i> (2021)
4	Living with flood for strengthening the communities' resilience against flood hazards.	This study proposes two field models (flood-resistant agriculture and wetland fisheries) and one field model of international labour migration.	Questionnaire survey	India	Islam <i>et al.</i> (2022)

1					
2					
3	5	Flood risk assessment in communities and application of previous studies	Attempts to translate the theoretical model of disaster research into practice are very rare. This is primarily due to the complexity associated with the data.	Secondary quantitative study using the risk triangle model	India Alam <i>et al.</i> (2018)
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14	6	Household resilience to floods through dam construction	It is recommended that disaster risk management authorities improve post-flood risk assessment and data collection to support future planning of improved flood control projects	Secondary quantitative study and survey data	Nigeria Yusuf, Muhammad and Sule (2021)
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28	7	Vulnerability of residential buildings to flood	A guide was created to help government agencies and community stakeholders achieve resilience integration in building Nigeria and its surroundings.	Literature review	Nigeria Adebimpe <i>et al.</i> (2018)
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39	8	Flood hazard classification high-resolution maps for resilient communities	Collaborative Flood Modelling supports broad end-user engagement when considering flood risk and provides strong evidence that collaboratively created knowledge can be readily adopted and applied to Flood Risk Management (FRM)	Secondary quantitative study and survey interviews	US Sanders <i>et al.</i> (2020)
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					

9	Projecting Changes in Expected Annual Damages From Riverine Flooding	Adaptation measures that protect against today's 1-in-100-year floods will increasingly pay off in temperate climates by also protecting against more frequent and smaller events.	Secondary quantitative study	US	Wobus <i>et al.</i> (2019)
10	Learning lessons from previous flood event for community resilience	Developers who have taken precautionary measures to increase flood resilience and those who have not. The factors that influenced those decisions and the level of professional involvement.	Survey (questionnaires and interviews)	Nigeria	Brisibe and Pepple (2018)
11	Urban risk flood management	Floods have caused physical, psychological, and economic damage to millions of people in Nigeria. Stakeholder efforts to address the challenge appear to have been limited by a lack of reliable data, and awareness among those affected, including lack of knowledge about flood risk reduction.	Literature review	Nigeria	Nkwunonwo, Whitworth and Baily (2016)
12	Climate resilience toolkit for flooding	Risk management and decision making for climate related hazards	Document and archival analysis	US	Gardiner, Herring and Fox (2019)

1						
2						
3	13	Community engagement for flood risk adaptations	Through proper planning and management with integrated cooperation and engagement of citizens and governments, it is possible to adapt to and mitigate the impacts of natural disasters.	Literature review	India	Dhiman <i>et al.</i> (2019)
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16	14	Flood management practices	Improved intergovernmental and interagency coordination, new avenues of financing, including housing to source local products and public and private companies providing services to boost local economies.	Systematic review	India	Kumar <i>et al.</i> (2021)
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						
31	15	Feasibility of widespread flooding	Regional flood susceptibilities emerge in river basins with catchments sharing similar streamflow and climatic regimes.	Secondary quantitative study	US	Brunner <i>et al.</i> (2020)
32						
33						
34						
35						
36						
37						
38						
39						
40						
41						
42						
43						
44						
45						
46						
47						
48						
49						
50						
51						
52						
53						
54						
55						
56						
57						
58						
59						
60						

Table 2. Theme derivation for flood risk management and community engagement

Topic number	Constituents	Theme
1	Flood, risk, flooding, hazard, building, maps, information	Building adaptation to flooding
2	Climate, change, cost, resilience, management, future, damages, flood, data	Climate change resilient communities
3	Flood, risk, urban, flooding, city, water, management, cities	Urban infrastructure and community preparedness
4	Flooding, risk, management, study, disaster, river, climate, research	Research output for flood risk and community response



## Appendix A. Gamma values for topics clustered within documents

Nr	Document	Topic	Gamma
1	1	1	2.83E-06
2	2	1	1.00E+00
3	3	1	1.66E-06
4	4	1	3.69E-06
5	5	1	1.97E-06
6	6	1	1.47E-06
7	7	1	2.57E-06
8	8	1	1.84E-06
9	9	1	4.87E-06
10	10	1	2.05E-06
11	11	1	5.32E-02
12	12	1	2.81E-06
13	13	1	1.00E+00
14	14	1	1.00E+00
15	15	1	3.69E-06
16	1	2	2.83E-06
17	2	2	4.57E-06
18	3	2	1.66E-06
19	4	2	1.00E+00
20	5	2	1.97E-06
21	6	2	1.47E-06
22	7	2	4.44E-03
23	8	2	1.84E-06
24	9	2	4.87E-06
25	10	2	2.05E-06
26	11	2	3.88E-06
27	12	2	1.00E+00
28	13	2	5.55E-06
29	14	2	1.50E-06
30	15	2	1.00E+00
31	1	3	2.83E-06
32	2	3	4.57E-06
33	3	3	1.00E+00
34	4	3	3.69E-06

---

3	35	5	3	1.00E+00
4	36	6	3	1.00E+00
5	37	7	3	9.96E-01
6	38	8	3	1.84E-06
7	39	9	3	4.87E-06
8	40	10	3	2.05E-06
9	41	11	3	9.47E-01
10	42	12	3	2.81E-06
11	43	13	3	5.55E-06
12	44	14	3	1.50E-06
13	45	15	3	3.69E-06
14	46	1	4	1.00E+00
15	47	2	4	4.57E-06
16	48	3	4	1.66E-06
17	49	4	4	3.69E-06
18	50	5	4	1.97E-06
19	51	6	4	1.47E-06
20	52	7	4	2.57E-06
21	53	8	4	1.00E+00
22	54	9	4	1.00E+00
23	55	10	4	1.00E+00
24	56	11	4	3.88E-06
25	57	12	4	2.81E-06
26	58	13	4	5.55E-06
27	59	14	4	1.50E-06
28	60	15	4	3.69E-06

---