








ARTICLE



<https://doi.org/10.1057/s41599-022-01038-2>

OPEN

Topic modelling exposes disciplinary divergence in research on the nexus between human mobility and the environment

Kerstin K. Zander ^{1,2}✉, Stephen T. Garnett ¹, Harald Sterly ³, Sonja Ayeb-Karlsson ^{4,5,6},
Barbora Šedová⁷, Hermann Lotze-Campen ^{7,8}, Carmen Richerzhagen^{2,9} & Hunter S. Baggen¹

Human mobility is increasingly associated with environmental and climatic factors. One way to explore how mobility and the environment are linked is to review the research on different aspects of the topic. However, so many relevant articles are published that analysis of the literature using conventional techniques is becoming prohibitively arduous. To overcome this constraint, we have applied automated textual analysis. Using unsupervised topic modelling on 3197 peer-reviewed articles on the nexus between mobility and the environment published over the last 30 years, we identify 37 major topics. Based on their language use, the topics were deeply branched into two categories of focus: Impact and Adaptation. The Impact theme is further clustered into sub-themes on vulnerability and residential mobility, while articles within the Adaptation theme are clustered into governance, disaster management and farming. The analysis revealed opportunities for greater collaboration within environmental mobility research, particularly improved integration of adaptation and impact research. The topic analysis also revealed that, in the last 30 years, very little research appears to have been undertaken in migration destinations or on the fate of environmentally influenced migrants during their migration process and after arriving in a new location. There are also research gaps in gender and Indigenous issues within the Impact theme, as well as on adaptive capacity and capacity-building.

¹Charles Darwin University, Northern Institute, Darwin, NT, Australia. ²German Development Institute, Bonn, Germany. ³University of Vienna, Vienna, Austria. ⁴University of Sussex, Brighton, UK. ⁵United Nations University Institute for Environment and Human Security, Bonn, Germany. ⁶University College London, London, UK. ⁷Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany. ⁸Humboldt University of Berlin, Berlin, Germany. ⁹German Aerospace Center, Bonn, Germany. ✉email: kerstin.zander@cdu.edu.au

Introduction

Never have humans been so mobile (IOM, 2020), notwithstanding what is expected to have been a temporary hiatus in 2020/21. Understanding human mobility, which occurs along a continuum from voluntary to involuntary (Boas et al., 2019; McLeman et al., 2021), is fundamental to many aspects of policy-making and planning in human society (Alessandretti et al., 2020). While most movements occur for economic and social reasons (Black et al., 2011; Buckle, 2017; Lekies et al., 2015), others result from conflict, extreme weather events and development-related displacement (Braithwaite et al., 2019; Rogers and Wilmsen, 2020). Climate change and environmental degradation can influence all these types of mobility to different degrees, even where they are not the major drivers (Black et al., 2011, 2013; Hermans and McLeman, 2021). Although research on the environmental-mobility nexus has been expanding (Boas et al., 2019; Hoffmann et al., 2021; Piguet et al., 2018), it has been patchy (Hoffmann et al., 2020), characterised by poor data (Hoffmann et al., 2021; Vinke and Hoffmann, 2020), and often disconnected from the standard migration theories (Black et al., 2011).

To date, reviews of the literature on environmental and climate mobility, including recent meta-analyses, have largely focused on climate as a push factor for the migration decisions of vulnerable people (Cattaneo et al., 2019; Hoffmann et al., 2021; Kaczan and Orgill-Meyer, 2020; Piguet et al., 2018). Hoffmann et al. (2020) undertook a meta-analysis of 30 case studies to quantify the influence of different environmental factors on migration, Beine and Jeusette (2021)—a meta-analysis on 51 empirical studies to investigate methods applied to link climate change to migration and Šedová et al. (2021)—meta-analysis using 116 studies to reveal establish causes for climate migration. Recently, the environmental migration literature has also been accompanied by investigations and acknowledgements of the importance of climate-induced immobility and the ‘trapped’ populations concept (e.g. Ayeb-Karlsson et al., 2018). Other reviews tend to have concentrated on specific regions (Borderon et al., 2019; Piguet et al., 2018; Thalheimer et al., 2021) or on international migration (Obokata et al., 2014). Thompson et al. (2017) reviewed literature specifically on disaster evacuation.

To date no reviews and analyses have included literature focusing on specific environmental factors as attractants, such as climates favourable for agriculture drawing in male labour (Gray and Wise, 2016) or lifestyle migrants seeking mild climates (e.g. Lekies et al., 2015). Nor have they explored the extent to which evacuations following sudden-onset hazards like floods and storms have become permanent given the influence of climate change on event frequency (e.g. Mueller et al., 2020). Often, migration is researched as just one in a portfolio of coping strategies, making this literature less likely to appear in migration reviews. Moreover, the reviews of farmer mobility in poorer countries (e.g. Cattaneo et al., 2019; Piguet et al., 2018) have not been matched by similar studies in urban populations or high-income countries, although environmental and climatic factors are known to play a role in such settings (e.g. Zander et al., 2020).

Reviews on topics as broad as the nexus of environmental change and mobility are becoming more challenging given the volume of research, directly and indirectly, relating to the topic, potentially leading to the separation of different knowledge streams into silos (Tight, 2014). We argue that textual analysis based on machine-based topic modelling can make valuable contributions to overcome this constraint and to avoid a potential separation of different knowledge streams on a given topic into silos (Blei et al., 2003). In the current study, which is the first detailed review of research across the environment–mobility nexus, we applied machine-based unsupervised topic modelling, a text-mining tool used to discover hidden semantic structures in texts. Full texts of nearly 3200 peer-reviewed articles were obtained for the analysis from a systematic literature review on the broad nexus between the environment and human mobility.

Methods

Systematic literature review. Following the Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines, we first used a broad predefined set of selection criteria to retrieve peer-reviewed articles from the Scopus database containing information relevant to environmental mobility, including that influenced by climate change. We kept the key word search (Table 1) very broad to reflect the many dimensions of environmental and climatic push and pull factors influencing human mobility with the aim of capturing even minimally researched issues. The key word search had two key components, a mobility dimension (migrate, relocate, move, displace, evacuate) and an environmental dimension (natural hazard, climate change, climate variability, extreme event or just ‘environment’). Because the scope of the search was so broad, we constrained the sample frame to peer-reviewed articles written in English in the last 30 years and included in the Scopus database. We further excluded book chapters, grey literature and a few subject areas that we deemed as irrelevant (e.g. physics, mathematics, chemistry).

The search, carried out between January and March 2021, found 55,439 articles. The lead author screened all titles, and abstracts if necessary, for relevance. A second reviewer, the last author, checked a random quarter of all articles to confirm the first author’s interpretation but did not find discrepancies. In this step, we excluded most of the articles (90%; Fig. 1) because they were mostly concerned with the movements of animals, chemicals, or pollutants.

Adopting the approach of Shaw et al. (2014), two reviewers then independently screened the full texts of 5314 articles (54 full texts were not available) and excluded 51% because they were not on both human mobility and environmental issues. Many articles, for example, dealt with human mobility, but used the terms ‘environment’ and ‘climate’ in contexts irrelevant to our study, such as ‘social environment’, ‘teaching environment’, ‘legal environment’, ‘work climate’ or ‘political climate’. The same applied to the mobility dimension, with articles extracted that did not deal with humans actively moving. We also excluded articles

Table 1 Scopus search syntax resulting in 55,439 publications (in March 2021).

```
(TITLE-ABS-KEY (migrat* OR relocat* OR move OR moving OR displac* OR evacuat*) AND TITLE-ABS-KEY ("natural hazard*" OR "climat* chang*" OR "extreme event*" OR environment* OR "climate variabilit*")) AND PUBYEAR > 1990 AND (LIMIT-TO (SRCTYPE, "j")) AND (EXCLUDE (SUBJAREA, "ENGI") OR EXCLUDE (SUBJAREA, "COMP") OR EXCLUDE (SUBJAREA, "MEDI") OR EXCLUDE (SUBJAREA, "BIOC") OR EXCLUDE (SUBJAREA, "PHYS") OR EXCLUDE (SUBJAREA, "MATH") OR EXCLUDE (SUBJAREA, "MATE") OR EXCLUDE (SUBJAREA, "ENER") OR EXCLUDE (SUBJAREA, "CHEM") OR EXCLUDE (SUBJAREA, "CENG") OR EXCLUDE (SUBJAREA, "BUSI") OR EXCLUDE (SUBJAREA, "IMMU") OR EXCLUDE (SUBJAREA, "NEUR") OR EXCLUDE (SUBJAREA, "PHAR") OR EXCLUDE (SUBJAREA, "NURS") OR EXCLUDE (SUBJAREA, "HEAL") OR EXCLUDE (SUBJAREA, "VETE") OR EXCLUDE (SUBJAREA, "DENT")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re") OR LIMIT-TO (DOCTYPE, "le")) AND EXCLUDE (PUBYEAR, 2021) AND (LIMIT-TO (LANGUAGE, "English"))
```

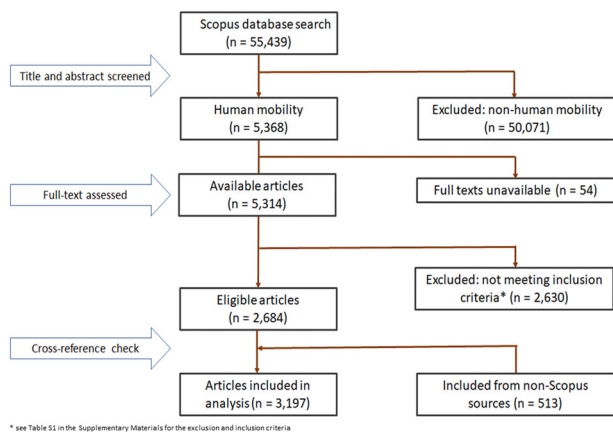


Fig. 1 Flowchart of the systematic literature search. *n* is the number of peer-reviewed articles at each stage of the review process.

that dealt with historical accounts of environmental or climate change mobility and a few other fields of study that were initially found to have both human mobility and environment dimensions but were subsequently determined not to be related to each other (see Table S1 in the Supplementary Materials).

Cross-checking of references with the 10 highly cited articles on the environment–mobility nexus enabled us to identify a further 513 peer-reviewed journal articles from sources other than Scopus (see Part A in the Supplementary Materials). In total, we included the content of 3197 full texts in the analysis. Bibliographic information on all 3197 articles included in the review is provided in a separate Supplementary Material Data file.

Topic modelling and clustering. Based on the keywords applied in the systematic literature review, we assumed that all 3197 peer-reviewed articles selected were broadly about the topic of environmental migration mobility. Topics, which are defined by words that occur frequently together, were revealed by modelling the strength of association between different mobility and hazards search terms used in the literature review and other words. Because there were so many articles, topic modelling was a much more efficient, and probably more objective, means of uncovering hidden patterns of meaning than would have been achieved with manual processing.

A popular topic modelling approach, and one that we applied in our study, is based on latent Dirichlet allocation (LDA) wherein each document is considered a mixture of topics and each word in a document is considered randomly drawn from a document's topics (Blei et al., 2003). LDA is a Bayesian model where each text and word has a probability of belonging to each topic (Blei et al., 2003; Grimmer and Stewart, 2013). The optimal number of topics within the corpus was determined using different metrics (see Part D in the Supplementary Materials).

Because of the large number of articles included in the analysis, we had anticipated identifying many topics. In order to cluster different topics together, we performed clustering on the topic space. Clustering, based on the top 50 stemmed words was undertaken using Jensen–Shannon divergence (Lamberti et al., 2008; Fig. 2), with Ward's hierarchical agglomerative clustering of Euclidean distances measured from the term-topic probability matrix (Murtagh and Legendre, 2014) producing the same hierarchical structure.

Before the full texts could be analysed, they had to be cleaned and processed (see Parts B and C in the Supplementary Materials). This was done using the *quanteda* library (Benoit et al., 2018) in R while the topic modelling was performed using the R library *topicmodels* (Grün and Hornik, 2011).

Results

Topic and cluster extraction. Based on four algorithms (see Part D in Supplementary Materials), we used a LDA model to categorise the texts of the 3197 articles into 37 topics (Fig S2 in Supplementary Materials), which we then summarised using a small set of distinct key words (Table 2) out of the most important 50 words (see Table S2 in Supplementary Materials for a full list of most important stemmed words within each topic). The three most frequent topics, none more than 5% of topics, were *drought adaptation of farmers* (topic 25: 4.4%), *climate change adaptation policy* (topic 37: 4.3%) and *concepts of environmental and climate change migration* (topic 34: 4.2%). Although every article was identified as relevant to human mobility, only 22 of the 37 topics, and 62% of all articles, included a term relating to mobility (i.e. a form of migration, relocation, displacement, evacuation, mobility or movement) within the 50 words identified as being the most important in an article. When an article was not about movements per se, mobility was embedded in the suite of strategies relating to climate change and environmental hazards (e.g. Antwi-Agyei et al., 2018). Ten of the 37 topics, and 34% of all articles, included the term climate change within the 50 most important words.

The subsequent clustering of the 37 topics revealed two clearly distinct branches of scholarship (Fig. 2). Of these, one focussed on the impacts of climate change, disasters and environmental degradation on mobility (Impact theme; 50.2% of articles), the other on how people cope, respond and change their behaviour in order to adapt to climate change, natural hazards and environmental degradation, of which migration, evacuation or relocation are potential strategies (Adaptation theme; 49.8% of articles).

In the Impact theme, the largest of two clusters concerning societal and livelihood vulnerabilities linked to mobility (Vulnerability cluster; 31% of articles), and 19% of articles dealt with different aspects of residential mobility (Residential cluster). Migration that has been the subject of research in this theme could be considered to be largely involuntary and reactive in the continuum of migration agency.

There were three clusters within the Adaptation theme, the largest of which related to governance, concepts, and legal and policy issues of environmental and climate-related mobility (22% of articles; Governance cluster). Twenty-eight per cent of the articles were broadly on climate change adaptation and mobility, either with respect to preparation for or in the aftermath of, disasters (15%; Disaster cluster) or to farming societies (13%; Farming cluster).

The Disaster and Farming clusters were closely related, suggesting that the migration studied was likely to have been planned proactively while the Governance cluster, although distinct, uses a similar language when it refers to adaptive behaviour.

Theme and cluster interpretation

Impact. All topics within the Impact theme include the words 'impact', 'vulnerability' or both as key terms. The Vulnerability cluster covered a diverse array of social vulnerabilities to natural hazards. This branch of literature has mostly treated mobility as a response to impacts that prevents in-situ adaptation and included responses to sea-level rise (topic 1; 3.3%), cyclone-induced displacement (topic 3; 2.7%), farm abandonment (topic 6; 2.4%), responses to wildlife–human conflicts (topic 11; 3.3%) and dam construction where this had displaced populations (topic 12; 2.8%).

Research included within the Residential cluster made almost no reference to climate change. Instead literature in this theme described residential movements as being driven by specific

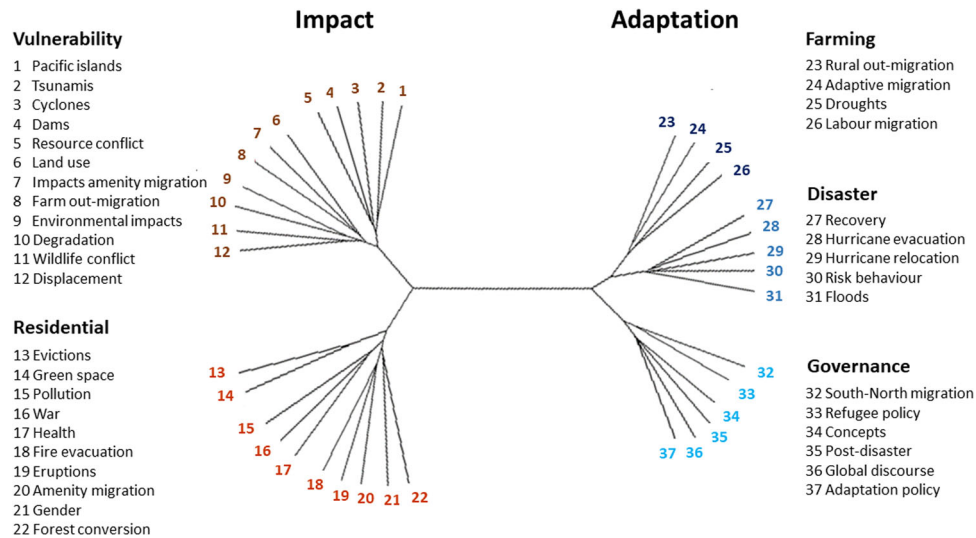


Fig. 2 Phylogram of the topic model. Similarities across the 3177 articles showing two deeply branched themes, five clusters and 37 topics.

environmental hazards such as air pollution (topic 15; 2%), drought and conflict (topic 16; 2.2%), fire risk (topic 18; 2.3%), volcanic eruption (topic 19; 1.6%) and deforestation (topic 22; 2.4%). This category also contained literature that dealt with other environment-related reasons that drive mobility, such as urban displacement due to gentrification (topic 14), eviction of slum dwellers (topic 13), the spread of infectious diseases (topic 17) and lack of or search for natural amenity (topic 20).

Adaptation. Articles within the largest cluster within the Adaptation theme, Governance, aimed at supporting policy-making and decision frameworks for environmental and climate change mobility. Four of the six topics within this theme included ‘climate change’ among the most important words, indicating that these articles dealt with general climate change impacts as a global phenomenon rather than as distinct hazards. Articles on the concepts of environmental and climate change migration (topic 34; 4.2% of all articles) included key literature (e.g. Bates, 2002; Bardsley and Hugo, 2010; Black et al., 2011; Gemenne, 2011; Hunter, 2005; McLeman and Smit, 2006; Warner, 2010), by authors who were among the first scholars to discuss migration as a potential adaptation strategy in response to climate and environmental change. Two of the topics (32 and 33) covered the legal and political aspects of climate change displacement with implications for South to North movements (topic 32) and the risk forecast to be posed by people sometimes characterised as ‘climate refugees’ (topic 33), although the use of the term is disputed (McAdam, 2013).

The Disaster cluster encompassed articles on how people prepare for the impacts of sudden-onset environmental shocks, mainly floods, hurricanes and earthquakes, their immediate evacuation behaviour and how they either return home or leave affected areas for good. This theme had a strong modelling focus (topics 30 and 31). Topics 28 and 29 were very similar, one including literature on shelters and warning systems and temporary movements in response to hurricanes (topic 28) and one dealing with long-term residential movements in the aftermath of hurricanes (topic 29).

The Farming cluster covered literature on climate change adaptation of farming households and poor societies (13% of all articles). Topics within this theme tended to focus on climate change per se or related slow-onset hazards such as drought (topic 25). Two of the four topics in this cluster did not include a mobility term as a keyword, once again highlighting that mobility is just one of several options for adaptation.

Trends in topics. The number of relevant articles published per year has increased, on average, by 16% annually since 1991, although the number has stayed steady or slightly decreased in some years (Fig. 3 and Fig. S1 in the Supplementary Materials). As with climate change research generally (Callaghan et al., 2020), for which there has been a particularly rapid increase in publications since 2007, nearly 80% of articles have been published in the last decade (Fig. 3 and Fig. S3 in the Supplementary Materials). This is three times the annual 5% increase in the peer-reviewed literature on all topics (Walker and Rocha da Silva, 2015) and contrasts with an earlier review of the literature on climate change and migration that identified a peak in 2011 followed by a 5-year decline (1648 publications; Pigué et al., 2018).

While the Impact theme, particularly the Vulnerability cluster, attracted the most research through the 30 years, the proportion of articles relating to Adaptation, particularly Governance, increased relative to other clusters in the last decade, with about 5% more articles published in the Adaptation than the Impact theme since 2016 (Fig. S5 in Supplementary Materials). Some topics within the clusters were relatively important in the first ten years, when fewer than 50 relevant articles were published annually, but attracted little research effort subsequently (Fig. 4). These included ‘War and migration’ (topic 16), ‘Forest to farming conversion’ (topic 22), ‘Degradation driving farmer migration’ (topic 10) and ‘Environmental impacts of human mobility’ (topic 9). Other topics have only recently attracted research attention, such as ‘Climate change adaptation policy’ (topic 37), ‘Global climate change discourse’ (topic 36), ‘Concepts of environmental and climate change migration’ (topic 34) and ‘Climate change refugee policy’ (topic 33), all within the Governance cluster. A spate of articles on hurricane-related relocation (topic 29) and evacuation (topic 28) appeared soon after Hurricane Katrina in the USA (Figs. 4 and S3) and one topic, ‘Drought adaptation of farmers’ (topic 25), attracted little research until 2011 but has since ranked among the most heavily researched topics within the environment–mobility nexus.

Discussion

Breadth of scope. Using unsupervised machine-based topic modelling to extract and classify meaning from documents, which underpins the novelty of the current study within the field of environmental impacts on human mobility, has three major beneficial characteristics—the number of documents that can be analysed, the impartiality of document selection once selection

Table 2 Themes, clusters and topics emerging from textual analysis of peer-reviewed articles on environmental and climate-related migration from 1991 to 2020. Frequent terms within topics (see the full list of terms in Table S1 in SM) and the percentage of articles belonging to each topic are shown. Each topic was classified as being closest linguistically to those to which it is adjacent.

Topic (topic no.)	Most frequent distinct keywords	%
<i>Impact theme</i>		
Cluster 1: Vulnerability		31.1
Relocation of Pacific island societies (1)	Pacific islands, climate change, relocation, migration, sea-level rise, culture, village, impact, vulnerability	3.3
Tsunami impact on coastal population (2)	Coastal, tsunami, sea-level rise, storm surge, scenario, impact, damage, vulnerability	2.5
Cyclone impact on coastal communities (3)	Flood, water, river, coastal erosion, cyclone, fish, village, displacement, livelihood impact, vulnerability, Bangladesh	2.7
Environmental impact of dams (4)	Water, resources, development, impact, environment, ecosystem, impact, dam, energy	2.3
Conflict over resource scarcity (5)	Climate change, security, conflict, violence, food, water, drought, Africa, impact, migration	2.2
Land use changes (6)	China, development, farmland, ecology, water, environment, mountain, growth, settlement abandonment	2.4
Impacts of amenity migration (7)	Migration, population, rural, in-migration, amenity, place, urban, housing, retirement, preference, distance, employment	2.4
Employment driven farm out-migration (8)	Farm labour, rural, employment, village, development, market, industry, capital, skilled migrant	2.2
Environmental impacts of human mobility (9)	Forest, soil, degradation, vegetation change, farming, water, ecological impact, conservation	2.3
Degradation driven migration of farming households (10)	Farming, food, crop production, soil degradation, Africa, migration, season	2.7
Wildlife human conflict (11)	Forest, fish, local resources, wildlife, protection, impact, conservation, access	3.3
Development induced community displacement (12)	Resettlement, relocation, displacement, community, compensation, dam construction, impact, Indigenous	2.8
<i>Cluster 2: Residential</i>		
Eviction of slum dwellers (13)	Urban, settlement, slum dwellers, environment, water, infrastructure, planning, eviction	19.2
Urban green space and displacement (14)	Urban green space, recreation, amenities, housing, development, impact, displacement	1.2
Impact of pollution on location choice (15)	Air pollution, mobility, impact, waste, exposure, health	1.1
War and migration (16)	War, drought, population, movements, border, famine, settlement, forced	2.0
Climate change health impacts (17)	Health, climate change, disease, impact, food, water, heat, mental, vulnerability, displacement, malaria	2.2
Fire evacuation (18)	Fire, evacuation, residential, property, protect, stay, home, Australia, threat	1.7
Volcanic eruptions (19)	Residential, volcanic eruption, community, displacement, green space, gentrification, housing, community	2.3
Amenity migration (20)	Rural, development, residential, nature, lifestyle, amenity, Australia, migrant, impact, home, tourism	1.6
Gender (21)	Women, gender, livestock, men, village, pastoralism, family, traditional, mobility, social	2.3
Forest to farming conversion (22)	Deforestation, farming, Amazon, cattle, coffee, palm oil, settlement, Indigenous, impact	2.4
<i>Adaptation theme</i>		
Cluster 3: Farming		13.4
Rural out-migration (23)	Migration, out-migration, rural, internal, rainfall, farming, climate change, destination, drought	3.5
Models of migration as adaptation (24)	Climate change, farming, scenario, effects, cost, loss, income, migration, adaptation	3.0
Drought adaptation of farmers (25)	Farming, climate change, adaptation, livelihood, strategy, crop, drought, rainfall, food, adopt, diversification, capacity	4.4
Labour migration and remittances (26)	Farming, rural, household, income, labour, migration, remittance, migrant, off-farm	2.5
<i>Cluster 4: Disaster</i>		
Sudden-onset disaster recovery (27)	Disaster, earthquake, flood, typhoon, governance, information, emergency, community, recovery, evacuation, prepared, capacity	14.5
Hurricane evacuation and warning (28)	Evacuation, hurricane, risk, warning, order, shelter, information, emergency, social media	2.5
Hurricane damage and relocation (29)	Disaster, damage, relocation, displacement, return, recovery, displacement, hurricane, Katrina, buyout, property	3.9
Risk behavioural models (30)	Hazard, response, risk, perception, behaviour, preparedness, flood, earthquake, evacuation	2.9
Spatial analysis of flood risk (31)	Flood, spatial, hazard risk, agent-based model, scenario, prediction, pattern	2.3
<i>Cluster 5: Governance</i>		
South-North migration (32)	Countries, global, immigration, internal, Africa, Europe, policy, demographic, poverty, labour, income	22.5
Climate change refugee policy (33)	Refugee, climate change, displacement, internal, protection, policy, legal, governance, convention	3.1
Concepts of environmental and climate change migration (34)	Environment, climate change, migration, displacement, driver, decision, internal, destination, livelihoods, policy	4.0
Post-disaster community resilience (35)	Social, family, community, place, culture, network, migrant, mobility, process, support, response, return	4.2
Global climate change discourse (36)	Climate change, governance, policy, global, discourse, framework, justice, science, society, mobility	3.1
Climate change adaptation policy (37)	Relocation, climate change, community, adaptive capacity, governance, resilience, policy	3.8
		4.3

Each topic was classified as being closest linguistically to those to which it is adjacent.

rules have been established, and, given its automation, the breadth of analysis scope (Dörre et al., 1999). As a consequence, we were able to analyse 3197 peer-reviewed articles selected from Scopus or bibliographically connected to such documents, about twice as many as any attempted previously using narrative approaches (Piguet et al., 2018).

The two most striking results of the analysis are a consequence of the breadth of scope. The first was the diversity of topics—that 37 topics in five clusters should emerge emphasises both the

ubiquity of environmental influences on human mobility behaviours and their complexity. The diversity materialised partly because we did not assume a direction to the relationship between environmental change and mobility; climate change for example can be either a push or a pull factor depending on the context. The variety demonstrates again how inappropriate it is to characterise people moving for reasons related to environmental change as ‘refugees’ (McAdam, 2013)—the term appears as a key word in only one of the topics (33).

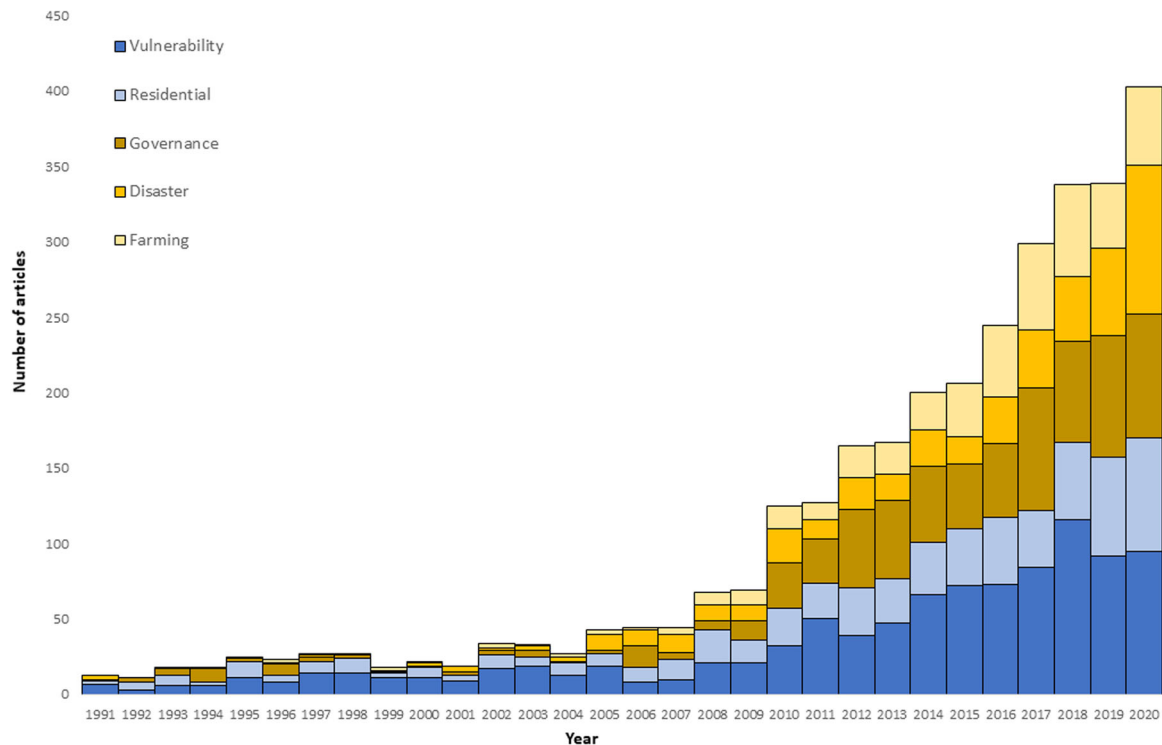


Fig. 3 Number of articles by publication year from 1991–2020. Trends for the five topic clusters in two themes—Impact (blue) and Adaptation (yellow) ($n = 3197$).

The second notable finding was the deep linguistic dichotomy between the Impact and Adaptation research themes. The result suggests a lack of communication between two streams of scholarship that might otherwise have much to share. The Impact theme consists of a rich and diverse body of literature on mobility along the continuum from voluntary to involuntary. The Adaptation theme includes scholarship examining governance, disaster planning and mobility associated with farming, often describing voluntary planned movement, and often to the advantage of those moving. That the topic modelling separated the two so definitively suggests that the two fields of research may draw little on the findings of each other—literature on adaptive planning that builds on the experience of research on involuntary mobility might have been expected to fall between the two research poles.

The distinction between the two themes also reflects an existing distinction between research strands, namely between those studies primarily addressing mobility as a response to climate and environmental change, conceiving mobility as a symptom of failed (in situ) adaptation, and those that address the potential of mobility as an adaptive strategy (e.g. Singh and Basu, 2020). The depth of the dichotomy between the two themes illustrates both the strength and weakness of unsupervised topic modelling. In real life the ontology of the human mobility–environment nexus is complex and impact and adaptation overlap. However, the language adopted in the literature to date has meant such complexity is not recognised by the language-based machine-learning models used here. In future, the dichotomy may be less pronounced either because there is more research that connects the two themes, or there is a greater emphasis on the precision of the language employed so that the continuum between impact and adaptation is more apparent to unsupervised analysis.

As it is, there is some reflection of the overlap in the analyses. Thus, articles within the Impact theme also consider adaptation behaviour and mobility as adaptation strategies in response to

adverse environmental and climatic conditions, just as articles within the Adaptation theme mention the impacts of climate change and environmental hazards. The difference lies in the main focus of the research and the emphasis given to different topics in the interpretation of the results. Within themes, the clusters are more closely related but nevertheless reveal key differences in research emphasis. Urban migration, for example, is all within the Residential cluster and is mostly related to slow-onset hazards such as heat and pollution. Research in these areas emerges as distinct from social vulnerability, which tends to emphasise low-income countries and farming societies (Cattaneo et al., 2019; Piguet et al., 2018). The lack of connection between sea-level rise and urban mobility suggests that research on managed retreat from coastal cities (Mach and Siders, 2021) is in its infancy, although ‘retreat’ is among the 50 key words for ‘Climate change adaptation policy’ (topic 37).

Many topics, and the largest cluster (Vulnerability), do not actually include a mobility-related term among the most important terms even though related to the environment–mobility nexus. This emphasises that human mobility is either just one aspect of behaviour research in the articles or that, for many hazards, mobility is not a viable response. For example, ‘South–North migration’ (topic 32) includes income and labour as key words but not environment or climate-related factors. This means that while research in this area acknowledges that environmental factors might trigger emigration from low-income countries, these have been less important than economic drivers (e.g. Hugo, 2006). Indeed, often the people most vulnerable to environmental change are most likely to be legally, financially and socially prevented from moving (Black et al., 2013).

Climate change is considered patchily across the corpus of literature, featured in only 34% of all articles. The Disaster cluster, for example, makes no mention of climate change. Literature in this cluster is strongly focussed on short-term responses to sudden-onset hazards such as cyclones and floods and appears to make few links to underlying drivers that make such responses

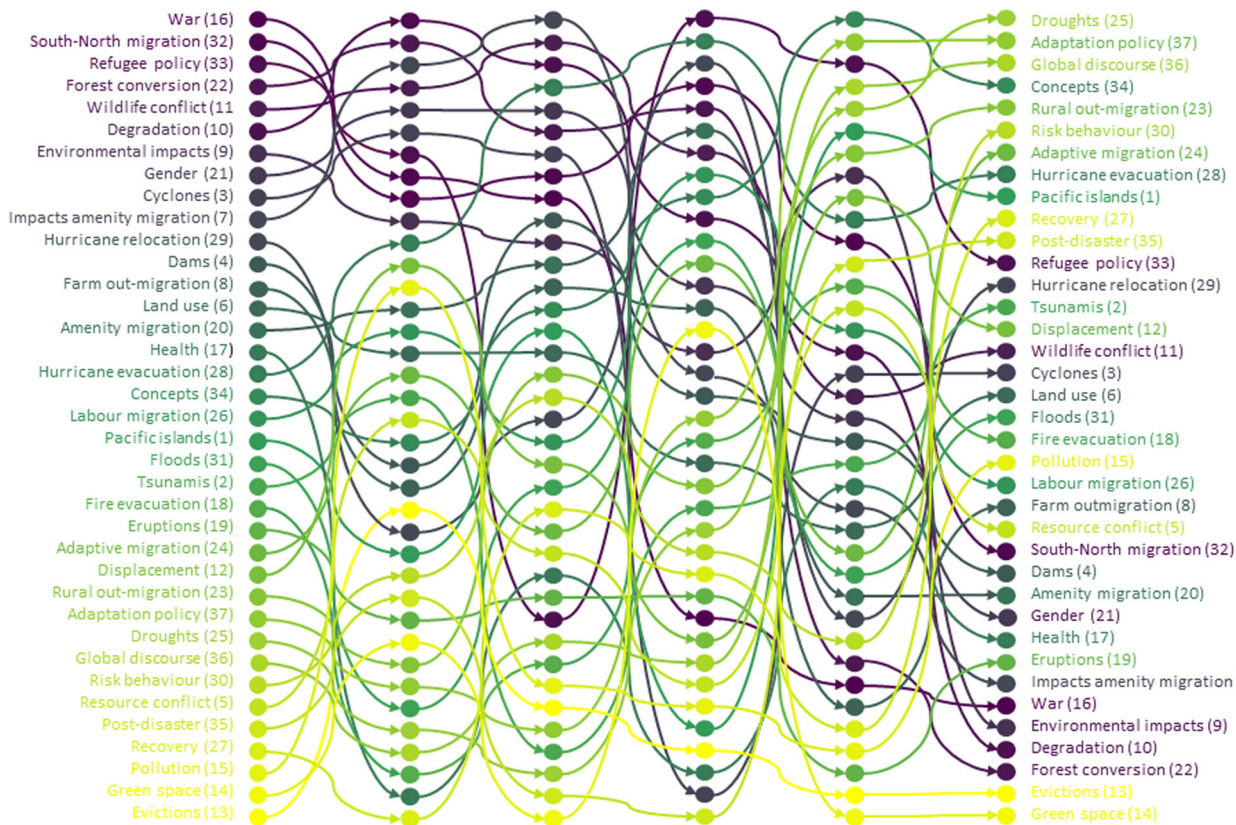


Fig. 4 Topic prevalence ranked over the last 30 years. Articles in each topic are grouped in five-year intervals and ranked from 1 (most prevalent in an interval, top) to 37 (bottom). Numbers in brackets refer to topic number.

increasingly necessary. While migration often implies a long-term and permanent action, evacuation is usually seen as a short-term response, resulting in temporary displacement until people could return after the situation stabilises (e.g. Black et al., 2013). Overall, the topic modelling did not identify research that considered the temporal or spatial scale of mobility. While some topics include countries or regions among their most important words (topics 1, 3, 5, 6, 10, 18, 20, 22, 32), language relating to the origin, direction or destination of mobility is not sufficiently salient to emerge from topic modelling.

Other sets of words appear across multiple topics but with different connotations. For example, ‘Environmental impacts of human mobility’ (topic 9) shares many key words with ‘Environment driven migration’ (topic 10). Although the algorithms underpinning the machine learning classified the research into different topics, further screening is needed to extract the nuanced difference between these topics. A similar example is the differences between ‘Impact of amenity migration’ (topic 7) and ‘Amenity migration’ (topic 20). The former is within the Vulnerability cluster and considers the negative environmental impacts of ‘green migration’ and the impact on rural societies from people in-migrating to the countryside for lifestyle and amenity purposes (e.g. Jones et al., 2003) or the environmental impact of people in-migrating to conservation areas (e.g. Salerno et al., 2014). The latter (topic 20) is clustered with other Residential topics and includes research on people moving to places with higher environmental quality (e.g. Gurran and Blakely, 2007).

Perhaps surprisingly, there were no major trends over the last 30 years in the diversity of topics on which there was peer-reviewed research. This is partly because the proliferation of research in the last half of the period swamps the influence of earlier research and maybe partly because 30 years is too short a

time for some of the trends to become apparent. The one area which has seen a disproportionately large increase in research interest relates to Governance within the Adaptation theme, which is likely to be associated with national legislative responses to the Paris Agreement (Iacobuta et al., 2018).

Research gaps. One means to identify possible research gaps is to look for the mirror-image of topics. Several imbalances stand out. For example, there is a great deal of research on how environmental change has driven human mobility, but only one topic on how mobility is directly affecting the environment (topic 9), whether at the point of origin or the destination. This topic’s prevalence is even declining (Fig. 4) which is in stark contrast to declining habitability of some places which might trigger relocation of many people, planned or unplanned (Horton et al., 2021), with environmental, social and infrastructure consequences for potential destinations. There is mention of impacts under other topics, such as the environmental impacts of large-scale projects such as dams (topic 4) and impacts of amenity migration, which are large of socio-economic nature (topic 7), but the main emphasis of these articles is not on migration.

Similarly, there is research on migration being induced by the health impacts of environmental degradation (topics 15 and 37) but there seems to be none on whether the migration-related health benefits are actually realised, partly for lack of adequate longitudinal studies (Vinke and Hoffmann, 2020). The way that mental health influences people’s migration decisions serve as another area in need of more attention (Kelman et al., 2021). Indeed, there was little research identified in the corpus of literature to 2021 on the trajectories of people who have moved for environmental reasons, whether voluntarily or under duress.

The patchy use of spatial or temporal words also reveals potential research gaps. Disaster risk management, for example,

appears to consider only evacuation not what happens to people who decide never to return. Other terms common in social research, such as Indigenous and gender, appear rarely. The term Indigenous came up in two topics, one on displacement forced by development (topic 12), the other related to Amazon deforestation (topic 22), both within the Impact theme. Indigenous is not mentioned in the top 50 words in topics within the Adaptation theme. Similarly, only one topic (topic 21 within the Impact theme) includes women and gender as keywords, pointing to a lack of gender consideration within the Adaptation literature, whether of women or men, as noted by another review (Obokata et al., 2014). The apparent lack of attention to women extends to relative disinterest in the mobility decisions of families or how families are affected by their members moving. There is one topic on post-disaster resilience (topic 35) that considers, for example, the role of family in wildfire evacuation (Asfaw et al., 2019) or how families cope with the aftermath of disasters by sending out one family member permanently to search for alternative employment opportunities (Mallick, 2014).

Finally, there seems to be limited literature on adaptive capacity with just three topics, one in each cluster of the Adaptation theme: the first investigates people's capacity to cope with sudden-onset (topic 27), the second slow-onset disasters (topic 25), and the third climate change adaptation policy (topic 37).

Study limitation. We identified two main limitations, one related to the applied method, one related to the underlying systematic literature review. While the topic modelling itself was unsupervised and not subjective in terms of the number of topics and the terms, the interpretation of the topics must necessarily involve human judgement. Without expert opinion and context to interpret the results of machine-based topic modelling, they provide little meaning (Philipps, 2018). We used the 50 most frequent words within each topic to assign a label to each topic and a theme to the clusters of similar topics. Sometimes the words themselves did not provide enough context and detail to convey the explicit ideas of the topics (Aggarwal and Zhai, 2012). This might have led to bias and incorrect conclusions (Hillard et al. 2008). When interpreting the topics, we therefore double-checked the articles within each topic and used them to label the topics, particularly when topics had similar keywords.

Overly detailed or complex topics, as well as infrequent topics, may not be detected in topic models (Aggarwal and Zhai, 2012). LDA assumes documents are exchangeable across the corpus, which may not be reasonable if, for example, experts can identify that documents have clear differences (Blei et al., 2003). Similarly, topic modelling must assess similar documents with consistent themes and word use. If this is not the case, then results may be skewed (Hillard et al., 2008). Conversely, it can be difficult to assign words and documents to topics if the texts are too similar to one another (Aggarwal and Zhai, 2012; Hillard et al., 2008). The only way to combat these limitations, as noted above, is for researchers to understand current research and the texts that are being analysed.

Another potential limitation concerns the literature search. There is the possibility that we failed to include relevant articles, for example, those not peer-reviewed and published in Scopus, those not in English and those not using the key search words in the abstract or title. We used cross-reference checks to minimise this limitation and added a substantial amount of literature not initially picked up (513). Also, the large number of articles included in the text analysis gives us great confidence in the validity of the results.

Conclusions

Our review of the peer-reviewed literature on the nexus between environmental change and human mobility contributes to two important fields. First, our analysis of the content of the research on the environment–mobility nexus has identified that, beneath a rapid increase in research interest, there are substantial areas where the language used in the research suggests there is less overlap in interest than might be expected and substantial research gaps. Judging from the machine-based topic analysis, there appears to be potential to include research on opportunities for adaptation with that on community vulnerability, or to extend the interest in evacuations to the fate of those evacuated and the factors that make communities so vulnerable to sudden-onset environmental hazards. The review highlights gaps in research on, for example, the influence of environmental drivers on Indigenous people, women and families, the fate of those who move and the spatial and temporal signature of mobility.

Second, we demonstrate that a (semi-)automated literature review provides important and up-to-date insights into research trends and themes, an approach especially valuable for large and dynamic research fields such as the nexus between human mobility and the environment. This approach has a value not only to researchers but also to policy-makers and funding agencies: automated approaches such as topic modelling can not only help identify emerging themes, gaps and trends but also identify important connecting themes and topics that can serve to build bridges between diverging research communities. Identifying important gaps and possible transdisciplinary bridges allows benchmarks to be set against which future research can be assessed to see whether these deficiencies are being remedied.

Data availability

Supplementary data, the full list of all analysed articles on environmental migration (Excel file), are available to download at the repository: <https://doi.org/10.13140/RG.2.2.29899.92963>.

Received: 20 August 2021; Accepted: 27 December 2021;
Published online: 01 February 2022

References

- Aggarwal CC, Zhai C (Eds.) (2012) Mining text data. Springer, New York
- Alessandretti L, Aslak U, Lehmann S (2020) The scales of human mobility. *Nature* 587:402–407
- Antwi-Agyei P, Dougill AJ, Stringer LC, Codjoe SNA (2018) Adaptation opportunities and maladaptive outcomes in climate vulnerability hotspots of northern Ghana. *Clim Risk Manag* 19:83–93
- Asfaw HW, McGee T, Christianson AC (2019) The role of social support and place attachment during hazard evacuation: the case of Sandy Lake First Nation, Canada. *Environ Hazards* 18:361–381
- Ayeb-Karlsson S et al. (2018) A discursive review of the textual use of 'trapped' in environmental migration studies: the conceptual birth and troubled teenage years of trapped populations. *Ambio* 47:557–573
- Bardsley DK, Hugo GJ (2010) Migration and climate change: examining thresholds of change to guide effective adaptation decision-making. *Popul Environ* 32:238–262
- Beine M, Jeusette L (2021) A meta-analysis of the literature on climate change and migration. *J Demogr Econ* 87:293–344
- Bates DC (2002) Environmental refugees? Classifying human migrations caused by environmental change. *Popul Environ* 23:465–477
- Benoit K et al. (2018) quanteda: an R package for the quantitative analysis of textual data. *J Open Source Softw* 3:774
- Black R et al. (2011) The effect of environmental change on human migration. *Global Environ Change* 21:S3–S11
- Black R et al. (2013) Migration, immobility and displacement outcomes following extreme events. *Environ Sci Policy* 27:S32–S43
- Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. *J Mach Learn Res.* 3:993–1022
- Boas et al. (2019) Climate migration myths. *Nat Clim Change* 9:898–903

- Borderon M et al. (2019) Migration influenced by environmental change in Africa: a systematic review of empirical evidence. *Demogr Res.* 41:491–544
- Braithwaite A, Salehyan I, Savun B (2019) Refugees, forced migration, and conflict: introduction to the special issue. *J Peace Res.* 56:5–11
- Buckle C (2017) Residential mobility and moving home. *Geogr Compass* 11:e12314
- Callaghan MW, Minx JC, Forster PM (2020) A topography of climate change research. *Nat Clim Change* 10:118–123
- Cattaneo C et al. (2019) Human migration in the era of climate change. *Rev Environ Econ Policy* 13:189–206
- Dörre J, Gerstl P, Seiffert R (1999) Text mining: finding nuggets in mountains of textual data. In: *Proceedings of the fifth ACM international conference on knowledge discovery and data mining (KDD-99)*, Fayyad U, Chaudhuri S, Madigan D (eds). ACM Press, New York, pp 398–401
- Gemenne F (2011) Why the numbers don't add up: a review of estimates and predictions of people displaced by environmental changes. *Global Environ Change* 21:S41–S49
- Gray C, Wise E (2016) Country-specific effects of climate variability on human migration. *Clim Change* 135:555–568
- Grimmer J, Stewart BM (2013) Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Political Anal* 21:267–297
- Grün B, Hornik K (2011) topicmodels: an R package for fitting topic models. *J Stat Softw* 40:1–30
- Gurran N, Blakely E (2007) Suffer a sea change? Contrasting perspectives towards urban policy and migration in coastal Australia. *Austn Geogr* 38:113–131
- Hermans K, McLeman R (2021) Climate change, drought, land degradation and migration: exploring the linkages. *Curr Opin Environ Sustain* 50:236–244
- Hillard D, Purpura S, Wilkerson J (2008) Computer-assisted topic classification for mixed methods social science research. *J Inf Technol Politics* 4:31–46
- Hoffmann R, Dimitrova A, Muttarak R, Cuaresma JC, Peisker J (2020) A meta-analysis of country-level studies on environmental change and migration. *Nat Climate Change* 10:904–912
- Hoffmann R, Šedová B, Vinke K (2021) Improving the evidence base: a methodological review of the quantitative climate migration literature. *Global Environ Change* 71:102367
- Horton RM, de Sherbinin A, Wrathall D, Oppenheimer M (2021) Assessing human habitability and migration. *Science* 372:1279–1283
- Hugo GJ (2006) Immigration responses to global change in Asia: a review. *Geogr Res.* 44:155–172
- Hunter LM (2005) Migration and environmental hazards. *Popul Environ* 26:273–302
- Iacobuta G, Dubash NK, Upadhyaya P, Deribe M, Höhne N (2018) National climate change mitigation legislation, strategy and targets: a global update. *Climate Policy* 18:1114–1132
- International Organization for Migration (IOM) (2020) World migration report 2020. IOM, Geneva
- Jones RE, Fly JM, Talley J, Cordell HK (2003) Green migration into rural America: the new frontier of environmentalism? *Soc Nat Resour* 16:221–238
- Kaczan DJ, Orgill-Meyer J (2020) The impact of climate change on migration: a synthesis of recent empirical insights. *Clim Change* 158:281–300
- Kelman I et al. (2021) A review of mental health and wellbeing under climate change in small island developing states (SIDS). *Environ Res. Lett* 16:033007
- Lamberti PW, Majtey AP, Borrás A, Casas M, Plastino A (2008) Metric character of the quantum Jensen–Shannon divergence. *Phys Rev A* 77:052311
- Lekies KS, Matarrita-Cascante D, Schewe R, Winkler R (2015) Amenity migration in the new global economy: current issues and research priorities. *Soc Nat Resour* 28:1144–1151
- Mach KJ, Siders AR (2021) Reframing strategic, managed retreat for transformative climate adaptation. *Science* 372:1294–1295
- Mallick B (2014) Cyclone-induced migration in southwest coastal Bangladesh. *ASIEN* 130:60–68
- McAdam J (2013) Climate change, forced migration, and international law. Oxford University Press, Oxford
- McLeman R, Smit B (2006) Migration as an adaptation to climate change. *Clim Change* 76:31–53
- McLeman R et al. (2021) Conceptual framing to link climate risk assessments and climate-migration scholarship. *Clim Change* 165:24
- Mueller V, Sheriff G, Dou X, Gray C (2020) Temporary migration and climate variation in eastern Africa. *World Dev* 126:104704
- Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J Classif* 31:274–295
- Obokata R, Veronis L, McLeman R (2014) Empirical research on international environmental migration: a systematic review. *Popul Environ* 36:111–135
- Philipps A (2018) How to sort out uncategorisable documents for interpretive social science? On limits of currently employed text mining techniques. In: *Proceedings of the 2nd international conference on advanced research methods and analytics*.
- Piguet E, Kaenzig R, Guélat J (2018) The uneven geography of research on “environmental migration”. *Popul Environ* 39:357–383
- Rogers S, Wilmsen B (2020) Towards a critical geography of resettlement. *Prog Hum Geogr* 44:256–275
- Salerno JD, Mulder MB, Kefauver SC (2014) Human migration, protected areas, and conservation outreach in Tanzania. *Conserv Biol* 28:841–850
- Shaw C, Hales S, Howden-Chapman P et al. (2014) Health co-benefits of climate change mitigation policies in the transport sector. *Nat Clim Change* 4:427–433
- Šedová B, Čizmaziová L, Cook A (2021) A meta-analysis of climate migration literature. *CEPA Discussion Papers* 29, Center for Economic Policy Analysis.
- Singh C, Basu R (2020) Moving in and out of vulnerability: interrogating migration as an adaptation strategy along a rural–urban continuum in India. *Geogr J* 186:87–102
- Thalheimer L, Otto F, Abele S (2021) Deciphering impacts and human responses to a changing climate in East Africa. *Front Clim* 3:84
- Thompson RR, Garfin DR, Silver RC (2017) Evacuation from natural disasters: a systematic review of the literature. *Risk Anal* 37:812–839
- Tight M (2014) Working in separate silos? What citation patterns reveal about higher education research internationally. *High Educ* 68:379395
- Vinke K, Hoffmann R (2020) Data for a difficult subject: climate change and human migration. *Migr Policy Pract* X:16–22
- Walker R, Rocha da Silva P (2015) Emerging trends in peer review—a survey. *Front Neurosci* 9:169
- Warner K (2010) Global environmental change and migration: governance challenges. *Global Environ Change* 20:402–413
- Zander KK, Wilson T, Garnett ST (2020) Understanding the role of natural hazards in internal labour mobility in Australia. *Weather Clim Extremes* 29:100261

Acknowledgements

This project has received funding from the Alexander von Humboldt Foundation under a research fellowships for experienced researchers (recipient: Kerstin Zander).

Competing interests

The authors declare no competing interests.

Ethical approval

Not applicable as this study did not involve human participants.

Informed consent

Not applicable as this study did not involve human participants.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-022-01038-2>.

Correspondence and requests for materials should be addressed to Kerstin K. Zander.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022