Received: 8 December 2017 | Revised: 11 June 2018 | Accepted: 13 June 2018

DOI: 10.1002/wcc.543

WILEY WIREs CLIMATE CHANGE

ADVANCED REVIEW

Big Data Approaches for coastal flood risk assessment and emergency response

James A Pollard[1] | Tom Spencer[1] | Simon Jude[2]

[1]Department of Geography, University of Cambridge, Downing Place, Cambridge, UK

[2]School of Water, Energy and Environment, Cranfield University, College Road, Cranfield, UK

Correspondence
Tom Spencer, Department of Geography, University of Cambridge, Downing Place, Cambridge, UK CB23EN.
Email: ts111@cam.ac.uk

Funding information
NERC/ESRC Data, Risk and Environmental Analytical Methods (DREAM) Centre, Grant/Award Number: NE/M009009/1

Edited by Anita Engels, Domain Editor, and Mike Hulme, Editor-in-Chief

Big Data Approaches (BDAs) refers to the combined use of historic datasets, incoming data streams, and the array of related technologies designed to shed new light on societal and environmental complexities through novel organizational, storage, and analytical capabilities. Despite widespread recognition of the commercial benefits of BDAs, application in the environmental domain is less well articulated. This represents a missed opportunity given that the dimensions used to characterize BDAs (volume, variety, velocity, and veracity) appear apt in describing the intractable challenges posed by global climate change. This paper employs coastal flood risk management as an illustrative case study to explore the potential applications in the environmental domain. Trends in global change including accelerating sea level rise, concentration of people and assets in low-lying areas and deterioration of protective coastal ecosystems are expected to manifest locally as increased future flood risk. Two branches of coastal flood risk management are considered. First, coastal flood risk assessment, focusing on better characterization of hazard sources, facilitative pathways, and vulnerable receptors. Second, flood emergency response procedures, focusing on forecasting of flooding events, dissemination of warnings, and response monitoring. Critical commentary regarding technical, contextual, institutional, and behavioral barriers to the implementation of BDAs is offered throughout including a discussion of two fundamental difficulties associated with applying BDAs to coastal flood risk management: the role of BDAs in the broader flood system and the skill requirements for a generation of data scientists capable of implementing Big Data Approaches.

This article is categorized under:
Social Status of Climate Change Knowledge > Knowledge and Practice
Climate, History, Society, Culture > Technological Aspects and Ideas

KEYWORDS

Big Data, coastal flood risk management, flood emergency response, flood risk assessment, source-pathway-receptor

1 | INTRODUCTION

Virtually all writing that offers an opinion on the buzzword "Big Data" begins by grappling with a definition of the term itself (Chen, Mao, & Liu, 2014; Kitchin, 2014; Vitolo, Elkhatib, Reusser, & Macleod, 2015). Various technology companies, consultancies, and government bodies have sought to capture the essence of Big Data as a springboard for discussions of its defining features, potential value, and associated risks (Bughin, Chui, & Manyika, 2012; Chen et al., 2014). This article refers to "Big Data Approaches" (BDAs; Jagadish, 2015) to encompass both the datasets themselves and the array of related technologies, including: the internet of things, parallel processing, cloud computing, machine learning, geotagging, data mining, and natural language processing. BDAs collate, store, and analyze both legacy datasets and incoming data streams enabling the critical conversion from raw data to value. Despite a plethora of commercial applications (Bughin et al., 2012), articulation of the benefits of applying BDAs in environmental domains is in its infancy (Gabrys, 2016; Salmond, Tadaki, & Dickson, 2017). This has prompted a number of questions: does the environment *need* Big Data (Hsu, 2013)? can Big Data *save* the environment (Shubert, 2015)? and what are the implications of Big Data for the broader question of environmental sustainability (Keeso, 2014)?

BDAs have been popularly described according to three "Vs": volume, velocity, and variety (Mattmann, 2013; Miller & Goodchild, 2015). This characterization has since expanded to consider a number of additional "Vs" (veracity being the most common fourth) and motivated studies dedicated solely to characterizing exactly what constitutes Big Data (Jagadish, 2015; Kitchin & McArdle, 2016). Notwithstanding the value of more detailed characterizations, the four "Vs"provide a useful framework that emphasizes BDAs as distinguished by more than sheer volume (Jagadish, 2015). Appropriately, the four "Vs" appear apt in characterizing the challenges of climate change (Ford et al., 2016). The global scale of climate change attests to the volume of impacts to be expected. The impacts themselves varying from rising sea levels (Hay, Morrow, Kopp, & Mitrovica, 2015; Nicholls et al., 2014) to shifting circulation patterns (Bader et al., 2011) with unprecedented rates of change in extinctions (Barnosky et al., 2011), atmospheric, and oceanic compositions (Doney, Fabry, Feely, & Kleypas, 2008). Furthermore, climate scientists are dogged by the veracious question of attribution: the ability to extract the signal from noise despite the complex dynamic and thermodynamic interactions between the Earth's many spheres (Otto, 2016; Otto et al., 2016; Stott, 2016).

Coastal systems embody the complexities of climate change, requiring an appreciation of oceanic, atmospheric, terrestrial, and anthropogenic processes. Global trends such as accelerating sea level rise, concentration of people and assets in low-lying areas, and deterioration of protective coastal ecosystems are likely to be manifested locally through increased flood risk (Nicholls et al., 2014). In the first global-scale analysis of the drivers of coastal flooding, Rueda et al. (2017) found that 75% of coastal regions have the potential for "very large" flooding events. Extreme water level events responsible for flooding (such as surges and high tides) are imposed atop mean sea level, leading to suggestions that future sea level rise could generate a doubling of flood frequencies within decades in certain regions (Vitousek et al., 2017). Alongside studies that quantify the source of future flood risk, it is necessary to consider landward propagation of these hazards via pathways toward receptors (Sayers, Hall, & Meadowcroft, 2002). Focusing on coastal megacities, (Hallegatte, Green, Nicholls, & Corfee-Morlot, 2013) suggest that in the absence of defense maintenance and upgrading, global coastal flood losses could exceed US$ 1 trillion annually. Furthermore, socioeconomic scenario modeling suggests that up to 50 million additional people could be flooded by 2080 compared with 1990 figures (Nicholls, 2004). As reported in the latest Intergovernmental Panel on Climate Change Report (IPCC, 2014) carbon-dependent societal processes have resulted in a degree of "lock-in" to this higher risk future (due to lagged responses to observed atmospheric temperature rise), regardless of stabilization in global mean temperature. Consequently, coastal flood risk managers face a substantial challenge, both in the assessment of coastal flood risk and in executing effective emergency management procedures when flooding occurs.

This paper explores potential applications of BDAs to address challenges associated with both these branches of coastal flood risk management. First, application to coastal flood risk assessment, focusing on improved understanding of the flood system to better characterize the interaction between hazard sources, facilitative pathways, and vulnerable receptors (BDAs 1 and 2). Second, application to flood emergency response procedures, focusing on forecasting of flooding events themselves, dissemination of warnings and monitoring of responses (BDAs 3, 4, and 5). This is achieved through a systematic literature review which identifies five BDAs and considers their application to coastal flood risk management. Critical commentary is offered throughout including a discussion of two fundamental difficulties associated with applying BDAs to coastal flood risk management: (a) the role of BDAs in tackling incertitude (uncertainty, ambiguity, and ignorance) introduced by the broader flood system and (b) the skill requirements for a generation of data scientists capable of implementing BDAs.

2 | BDAS FOR COASTAL FLOOD RISK ASSESSMENT

A persistent challenge in assessing coastal flood risk is incorporating the full complexity of interactions between marine, terrestrial, atmospheric, and anthropogenic variables. In a management context, this challenge is most acute at a spatial scale of tens to hundreds kilometers and a temporal scale from years to decades (Figure 1). At shorter spatiotemporal scales, coastal change is determined by time varying phenomena (e.g., waves and tides) and risk assessment focuses on the contribution of individual components of the flood system. At geological scales, these phenomena average out, lending greater importance to

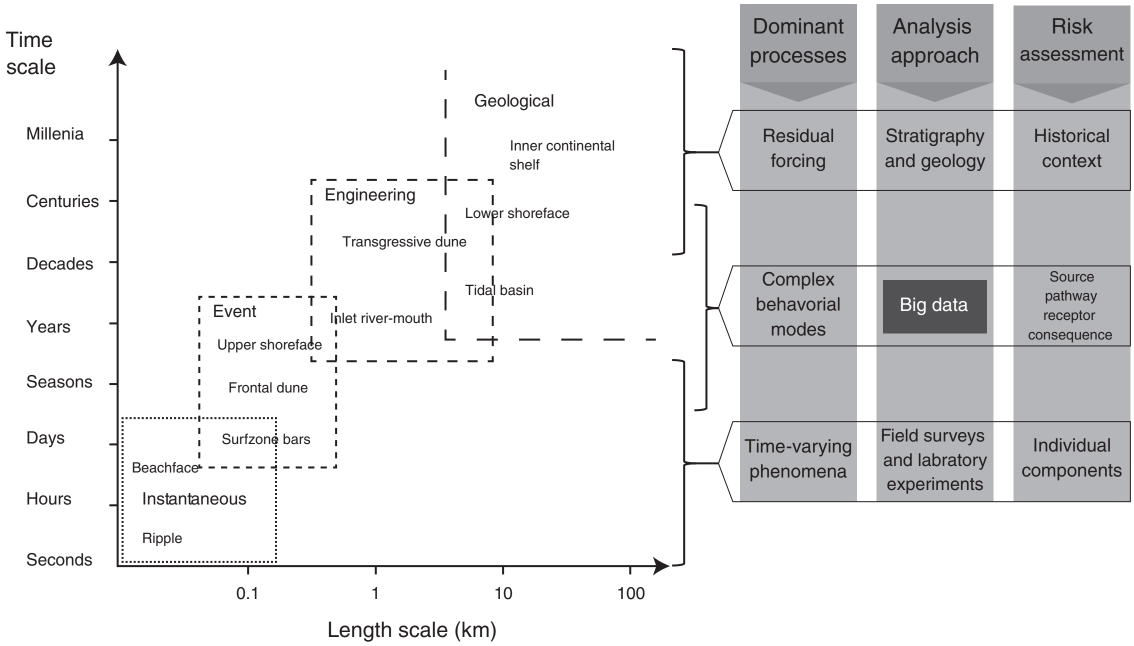

FIGURE 1 Big Data Approaches are especially well suited to address problems at the event and engineering scale (Reprinted with permission from Cowell and Thom (1994). Copyright The Authors, 2018)

residual forcing (e.g., sea level trends). Risk assessment over this timescale provides historical context to events taking place in the present day. It is when both time-varying and residual forcing must be considered that coastal systems exhibit complex "behavioral modes" (self-regulation, hysteresis, and nonlinearity among others) necessitating risk assessments that consider the whole system (Cowell & Thom, 1994; Narayan et al., 2014; Sayers et al., 2002). One systems approach to coastal flood risk assessment is known as the Source-Pathway-Receptor-Consequence (SPRC) framework. The SPRC approach takes place in advance of the flood event itself and seeks to characterize the interaction between hazard sources (surge, waves, rain), facilitative pathways (defenses, coastal bathymetry, ecosystems) and vulnerable receptors (residential and commercial property, critical infrastructure) to establish the potential consequences (flooding, increased insurance premiums, psychological impacts) of a given risk "event." It is at these "event" and "engineering" scales that BDAs can be applied to greatest effect (Figure 1).

2.1 | Big Data Approach 1: Synthesis and Harmonization of Coastal Datasets

This first BDA concerns characterization of the SPRC framework in a given coastal context. This requires understanding of the behavior of coastal systems during flooding events, which relies in the first instance on empirical observation of coastal system responses to a range of forcing conditions. In the United Kingdom, and elsewhere, there has been a relatively sudden and dramatic expansion in the availability of high-quality datasets which can be used to catalogue and analyze how hazard sources interact with facilitative pathways to impact on vulnerable receptors, with varied consequences. The UK Government Department of Environment, Food and Rural Affairs (Defra) have undertaken an "Open Data Initiative" to publish geospatial data in the public domain (Defra, 2013; Rumson, Hallett, & Brewer, 2017). It is now possible, for example, to download coastal aerial photography, Light Detection and Ranging (LiDAR) and bathymetric data from data archives held at http://environment.data.gov.uk in formats that can be easily manipulated using Geographic Information System (GIS) software.

Extracting value from such varied datasets necessitates a BDA. These morphological datasets can be combined with relevant hydrological forcing variables using data reduction techniques capable of harmonizing data structures of varying temporal and spatial frequency. In the United Kingdom, water level and wave conditions are recorded at 15- and 30-min intervals, respectively, while cross-shore profiles and aerial photography are gathered at monthly to annual frequency. Harmonization can be achieved through the construction of wave probability density functions to represent wave properties for the period prior to the cross-shore profile survey (Walstra, Hoekstra, Tonnon, & Ruessink, 2013). One application of this approach is to isolate extreme flooding events and establish how hazard interaction with the coastal morphology influenced the ultimate flood impacts.

Applying synthesis and harmonization techniques to extreme flooding event characterization has been achieved in the United Kingdom where openly available datasets have been systemically collated to produce a database of coastal flooding named "SurgeWatch" (www.surgewatch.org/). The current version (V2.0, 2017) of the database includes 329 coastal flood events between 1915 and 2016, based on both quantitative hydrodynamic data and "soft" data sources (Haigh et al., 2017). An "enhanced systematic commentary" based on the SPRC framework accompanies the 53 most severe events in the hope that this information will prove instructive to future coastal flood risk management in the United Kingdom (Haigh & Nicholls, 2017). Analysis of this database is already delivering important insights. Thus, for example, 86% of extreme sea level events were found to be only moderate surges but imposed on high tides, indicating a dominance of the tidal component (Haigh et al., 2016). Analysis has also revealed examples of simultaneous flooding on unconnected stretches of coastline (Santos, Haigh, & Wahl, 2017; Wyncoll et al., 2016). This has important implications for risk assessment given the potential for stretching emergency services beyond capacity (Wahl et al., 2017; Wyncoll et al., 2016). Such a database may also be of value in coastal consultations. A persistent finding in flood risk perception studies is the importance of prior flooding experience, with the exact influence on risk perception and subsequent action dependent on the experience of the individual (Baan and Klijn, 2004; Brilly and Polic, 2005; Burn, 1999; Burningham, Fielding, & Thrush, 2008; Kellens et al., 2011; Keller, Siegrist, & Gutscher, 2006; Siegrist and Gutscher, 2006). When individuals have experienced an event first hand, the ability to draw on a database such as SurgeWatch could provide valuable stimulus material for discussion around future flood risk management options. This is strengthened further if information from the database can be adapted for three-dimensional visualization, shown to be an effective tool for participatory coastal management (Jude, 2008; Jude, Jones, Andrews, & Bateman, 2006).

Extending such efforts to the regional scale, an event-based database has been established for 10 European case study sites as part of the Resilience Increasing Strategies for Coasts Tool-Kit (RISC-KIT) project (Ciavola, Harley, & den Heijer, 2017). Furthermore, there is potential for remote sensing to expand these local and regional databases to the global scale, with the aim of achieving continuous and complete cataloguing of surge hazard and associated coastal vulnerability (Brakenridge et al., 2013). It is, however, important to note that in global context, the United Kingdom maintains a high-quality coastal monitoring program, supported by a strong institutional network led by the UK Met Office. They make use of an estimated 20 million observations each day, less than 10% of the total 335 million archived daily observations (Met Office, 2016). Elsewhere, monitoring networks are not so comprehensive, raising more fundamental socioeconomic and political issues that may have deep-seated origins. A recent study into water challenges in northern India identified numerous interdependent barriers to effective adaptation in the face of climate driven water risks (Azhoni, Holman, & Jude, 2017a, 2017b). Inaccessibility to information and data was found to result from barriers including negative work culture and attitudes, resource limitations, and weak intra-institutional networks. Inaccessibility to information and data was in turn found to propagate governance challenges and result in low implementation of adaptation options (Azhoni et al., 2017a). This study draws attention to the context in which data is collected, stored, and made publicly accessible (or not) and highlights a wider set of challenges for attention before BDAs can realistically provide input to coastal flood risk assessment.

2.2 | Big Data Approach 2: Handling and Validating Satellite Imagery

Alongside airborne aerial photography and LiDAR surveys, space-based satellites provide high resolution, multispectral Earth monitoring systems. Although not exhaustive, the satellites listed in Table 1 demonstrate the range of available resolutions, revisit times, and sensors. Google Earth Engine provides a platform for analysis of satellite imagery with a catalogue of >200 remotely sensed and supporting datasets (Google, 2017). This catalogue amounts to >5 petabytes (5 million gigabytes) of data and is continuously updated with approximately 6,000 scenes added daily, typically uploaded within 24 hr of capture (Gorelick et al., 2017). In this case, it is not only the data but also an application programming interface (API) that is provided

TABLE 1 Selected earth observation satellites launched since 2000. Where satellite series perform similar functions, the most recently launched satellite is listed

Satellite	Launch year	Resolution (m)	Revisit time (days)	Outputs
QuickBird	2001	0.6–2.4	1–3.5	Panchromatic and multispectral
RapidEye	2008	5	5.5	Multispectral
Cryosat-2	2010	250	30	Altimetry
RISAT-1	2012	3–50	25	SAR
SARAL	2013	n/a	35	Altimetry
Landsat-8	2013	15–30	16	Panchromatic, multispectral, infrared
Sentinel-1	2014	5–40	12	SAR
Spot-7	2014	1.5–6.0	1	Panchromatic and multi-spectral
Sentinel-2	2015	10–60	5	Multispectral
Sentinel-3	2016	n/a	27	Multispectral and altimetry
Jason-3	2016	n/a	10	Altimetry
WorldView-4	2016	0.25	4.5	Multispectral and altimetry

to allow data exploration and analysis. Through harnessing cloud-based computing, Google Earth Engine facilitates global-scale geospatial analysis without the financial and technological burden of supercomputing hardware. Performing a similar service, but specifically for environmental applications is "EarthServer," an EU-funded data access and analytical engine for "Big Earth" data services aimed at addressing some of the challenges of data handling that have this far limited the application of Big Data analytics environmental science (Baumann et al., 2016).

One example of an integrated data collection, storage, and access service is that provided by the European Space Agency. The present European Copernicus system consists of a suite of satellites, each mounted with sensors tailored to a specific aspect of earth observation, with a global coverage between 84° North and 56° South (Copernicus Program, 2017; Nadim, 2016). Supporting this data collection, the Copernicus governance system consists of a number of national or European "Entrusted Entities," that are tasked with providing core terrestrial, atmospheric and marine environment monitoring, and emergency management functions (European Commission, 2013).

In addition to the core services provided centrally, through granting free, open access to their data, Copernicus encourages downstream applications. Such applications are delivered by public and private operators with the intention of adding value through integrating other datasets with Copernicus satellite data. Thus, for example, the EU FP7 funded Foreshore Assessment using Space Technology (FAST) project aims to deliver a tool capable of deriving foreshore characteristics from satellite imagery to inform nature based flood defense (FAST, 2017). This intention is predicated on empirical research that has demonstrated the flood protection capabilities of foreshore ecosystems under extreme water level and wave conditions (Möller et al., 2014). Yet the ability of these ecosystems to provide protection depends critically on several foreshore characteristics including ecosystem structural properties, habitat position within the tidal range, and erosional susceptibility (Morris, Gomez-Enri, & van der Wal, 2015). While investigations into these characteristics could be achieved in the field on a site by site basis, satellite technology offers the potential to conduct rapid, extensive foreshore assessment at considerably reduced cost. Having established the ecosystem characteristics responsible for flood protection services, standard generic parameters are derived to allow for the prediction of the flood defense capabilities of unmeasured areas (Morris et al., 2015). Next, the relationship between foreshore characteristics and their representation in the visible and nonvisible spectrum as collected by satellite sensors is established. This requires field-based datasets to validate the inferences made based on satellite imagery analysis. Finally, the influence of predicted foreshore characteristics on flood severity is quantified through flood modeling. This is packaged into the "MI-SAFE viewer" (http://fast.openearth.eu/), an interactive online viewing platform for end users. The BDA developed through the FAST project demonstrates the potential for national- (even global) scale flood risk assessment that is cost-effective, sensitive to dynamic changes in the foreshore and openly accessible.

The above FAST project example emphasizes that field-based research remains essential. In fact, the Copernicus initiative to which it belongs emphasizes three equally important pillars (in-situ measurement, data harmonization and standardization and product delivery) alongside satellite monitoring. Yet maintaining field-based observation alongside satellite monitoring is set to become an increasingly difficult task. Sentinel-2 has a revisit time of 5 days, and while not every image will be useable (principally owing to weather conditions) this temporal resolution simply cannot be matched in the field. There is a danger that "data fictions" will emerge whereby satellite-derived earth surface characteristics are taken to reflect natural processes without critical appraisal of the proxy assumptions involved (Nadim, 2016). These assumptions do not mean that the satellite data is unusable, but they do draw attention to the need for corroborating evidence. Given that field datasets appear so essential to corroborate remotely sensed big data, it is equally necessary to cultivate effective data management and sharing practices that have been noted as lacking by scientists across varied environmental domains (Hampton et al., 2013; Hsu, Martin, McElroy, Litwin-Miller, & Kim, 2015).

3 | BDAs FOR COASTAL FLOOD EMERGENCY RESPONSE

BDAs are well suited to offer insights into flood risk assessment in advance of the event itself. This is crucial in informing long-term mitigation and adaptation planning. However, the present is increasingly characterized by unprecedented global climate change and resultant nonanalogue future environmental conditions (Naylor et al., 2017; Stott, 2016). The challenge of characterizing the near-present and future conditions that coastal environments will encounter is further complicated by imperfect knowledge of the many ways in which risk sources, pathways, and receptors may interact. Such a combination of uncertainty and potentially unprecedented events suggests that coastal flooding can be expected to continue in the future. As such, coastal decision makers must be capable of appropriate and informed flood emergency response. Emergency response procedures encompass monitoring, forecasting, and detection (through international, national and regional networks) to warning (assembling an effective risk communication message), dissemination (through television, radio, internet, press, and person) to response (the emergency services and postevent recovery efforts).

3.1 | Big Data Approach 3: Integrating Process-based modeling and bayesian networks

Storm monitoring, forecasting, and detection form the first step in the emergency response chain. The stochastic and case-specific nature of storms lends itself to a process-based modeling approach (Lorenzo-Trueba & Ashton, 2014; Roelvink et al., 2009). The requirement for these processes to be resolved over large geographical extents and to incorporate strong spatial gradients both alongshore and crossshore means this approach is highly computer intensive (Bertin, 2016). Bertin et al. (2014) made progress in overcoming this limitation when hindcasting the flooding associated with the mid-latitude Atlantic storm Xynthia (2010). The Bertin et al. (2014) model achieved a resolution of just a few meters, this detail being necessary to resolve the dykes responsible for protecting the coastlines of the Bay of Biscay, France. Parallel processing was implemented to run the model across 144 CPU cores resulting in a runtime of 8 days. Resolving small-scale processes explicitly in this way would be impossibly computer intensive without parallel processing techniques. While important to longer term mitigative actions such as land-use planning and infrastructure design standards, runtimes spanning over several days still excludes this kind of modeling from usefully contributing to the emergency response chain. First, a storm might make landfall before the model has finished running, and second, changes to the storm characteristics during the model run might render the results obsolete.

One way around these difficulties is to perform numerical modeling in advance and use the results to train a Bayesian Network (BN). BN are one branch of Bayesian modeling which apply joint probability to update user-defined prior beliefs about a future event using data from past events. Once a BN has been trained on historical flood events, it can act as a surrogate for process-based modeling, removing the long computer processing times that would otherwise be required (Poelhekke et al., 2016). Inclusion of BN as a BDA derives from an ability to handle missing data, impressive prediction accuracy even with small sample sizes (Kontkanen, Myllymaki, Silander, Tirri, & Grunwald, 1997), an ability to incorporate domain knowledge alongside quantitative data (Heckerman, Geiger, & Chickering, 1995) and explicit treatment of uncertainty in a probabilistic sense (Aguilera, Fernández, Fernández, Rumí, & Salmerón, 2011; Uusitalo, 2007). The probabilistic set-up of BNs provides a quantitative approach to dealing with the veracity associated with bringing together such a diverse range of data sources, each of which are accompanied with different types and magnitudes of uncertainty (Skinner, Rocks, & Pollard, 2016). For example, Bertin et al. (2014) report that for storm Xynthia flooding extents were overestimated in some locations and underestimated in others, in part because of wave run-up and infragravity wave effects. This uncertainty can be propagated through a Bayesian model through the specification of vague prior distributions for these specific components (Jensen & Nielsen, 2007). Model outputs (in this case, likelihood of a flood event) are accompanied by percentile-based credibility intervals indicating whether a certain event will occur within user-specified uncertainty bounds. Moreover, the ability to integrate new data as it becomes available enables predictions to be updated as storms near landfall (Poelhekke et al., 2016). This has been demonstrated for the Praia de Faro, Portugal, where 20 years of data were used to train a BN capable of probabilistically predicting onshore erosion, overwash depth, and flow velocity (Poelhekke et al., 2016). Application of BN in this way could help to alleviate purported side lining of highly varied datasets (Jagadish, 2015) ensuring that emergency responders have access to the full suite of data relevant to a given flooding event.

Despite the clear potential for BN to be deployed for emergency flood response, no case studies of successful applications yet exist in the literature. One reason for this is that the BN must be trained to a specific coastal locality limiting applicability elsewhere, even within the same coastal management zone. Validation datasets are another critical limitation. For example, the above Praia de Faro study lacked a real-time storm event for validation of the model and so instead had to rely on partitioning past observations to produce independent training and validation datasets (Poelhekke et al., 2016). In addition to these technical difficulties, certain contextual challenges must be overcome. Responsible institutions tend to already have in place established emergency response procedures. Integrating new techniques such as BN will necessitate acquiring appropriate hardware and software, training of employees, integration and trialing periods, and appraisal and assessment. This is complicated by the fact that trialing would have to occur during the extremely high-pressured scenario of a genuine flood event. These are all reasons why the application of BN in emergency response is primarily limited to speculation by coastal researchers as opposed to practical implementation by decision-making actors.

3.2 | Big Data Approach 4: Ensemble Forecasting

Through ensemble forecasting, BDAs contribute to real time, and near-real time, emergency response. The latest models adopt a probabilistic approach. Thus, for example, the Global Ensemble Forecast System combines 21 separate forecasts to provide scenario-based forecasts on a global scale with associated uncertainty bounds (NOAA, 2017). The premise of ensemble forecasting is to generate multiple forecasts based on slightly adjusted initial conditions. The uncertainty associated with alternative model outcomes can be probabilistically attributed based up the prevalence of each outcome. For example, if high onshore winds are simulated in every model there is a higher probability that this will occur than in a scenario which is only reproduced in half of the total possible outcomes. This means that management decisions can be weighted toward the most likely scenario while contingency measures may be implemented to address the possibility that the most likely scenario will not materialize.

The issue of data processing time is particularly relevant in the context of ensemble forecasting for emergency flood response. There is a trade-off between the amount of time spent cleaning and preparing datasets for forecasting and the need for timely information to be disseminated to those likely to be impacted. In the context of modeling storm surges, and other natural phenomena, this is referred to as the Forecasters' Dilemma (McCallum & Young, 1989). The critical question is how close to the event impact can we wait until a warning must be issued? Increasing the amount of time spent understanding the sources of coastal risk will typically reduce the uncertainty associated with the final decision (for example to issue an evacuation order), but it also increases the likelihood that the decision will be made too late to be effectively executed. A further complication is introduced by the possibility of false warnings which can contribute to a deterioration of trust between those receiving and those issuing warnings, leading to an erosion of legitimacy and culture in which emergency advice is not heeded and acted upon. Lacking trust in the institutions responsible for issuing warnings is just one of several behavioral barriers that can inhibit effective emergency response. Flood risk perception studies have identified diverse influences on risk perception from socioeconomic (Botzen, Aerts, & Van Den Bergh, 2009; Burningham et al., 2008; Green et al., 1991) and psychometric (Keller et al., 2006; Siegrist & Gutscher, 2006; Siegrist & Gutscher, 2008; Terpstra, Gutteling, Geldof, & Kappe, 2006) to cultural (Demeritt, Nobert, Cloke, & Pappenberger, 2013) and environmental viewpoints (Botzen et al., 2009; Kellens et al., 2011; O'Neill, Brereton, Shahumyan, & Clinch, 2016). The exact combinations of influences, and their relative importance is specific to the individual, meaning that a single risk communication will be interpreted and acted upon in a multitude of ways (Burn, 1999). This illustrates that even if BDAs are effectively executed, resulting in well-informed and justified emergency decision-making, they cannot guarantee that the decisions taken will elicit the desired response. Applying BDAs in this context should not be accompanied with the expectation of resolving these important behavioral influences on individual reasoning.

3.3 | Big data Approach 5: Natural Language Processing of Social Media

Alongside physical variable datasets, social media provides real-time information about flooding events (Saravanou, Valkanas, Gunopulos, & Andrienko, 2015; Smith, Liang, James, Lin, & Liang, 2015). Analysis of Twitter during flood events in the United Kingdom (Saravanou et al., 2015), United States (Vieweg, Hughes, Starbird, & Palen, 2010), Pakistan (Murthy & Longwell, 2013), Thailand (Kongthon, Haruechaiyasak, Pailai, & Kongyoung, 2012), and the Philippines (Takahashi, Tandoc Jr, & Carmichael, 2015) has established the potential for social media as a first response data source, as well as a tool for dissemination of emergency warnings. The value of Twitter data derives from its temporal (immediate) and spatial (geotagging) relation to the event in question (Leetaru, Wang, Padmanabhan, & Shook, 2013). One organization translating this potential into practice is FloodTags®. They analyze Twitter data alongside other social and traditional media sources to inform management decisions during water crises (FloodTags, 2017). Currently based in the Philippines, the wider application of this tool is obvious. FloodTags report that during the 2015 flooding in Jakarta, related tweets reached 900 per minute, and included data on water depth as well as location and source of flooding (Rowling, 2015). Elsewhere, Twitter-derived water depths have been applied to inform real-time flood modeling of extreme pluvial flooding in Newcastle, United Kingdom (Smith et al., 2015), demonstrating the potential of this information to identify areas of highest risk during an event. Moreover, combining social media data with physical modeling insights addresses the uneven spatial distribution of active Tweeters.

Twitter data does not conform to a standardized format, requires significant preprocessing, and undergoes censorship before posting. All these characteristics complicate the extraction of useful information, especially given the time pressures experienced by emergency response services. Employing human operators to extract pertinent information from Twitter for real-time data assimilation to the emergency response operations is likely to be prohibitively costly, intensive work. To achieve data extraction at a rate sufficient to usefully input to emergency response in real-time would require a large team of analysts with technical flood risk management backgrounds that are available “on call” for the moment when emergency strikes. Instead, partial or fully automated techniques using “natural language processing” algorithms can extract key words and (ideally) the context in which they are written to provide quantitative summations of millions of data points (Figure 2). A further characteristic relating specifically to Tweets is the 280 character limit. Vieweg et al. (2010) identify “High Yield Twitterers” as those who Tweet close to the character limit and consequently tend to think carefully about the content of tweets, aiming to convey as much pertinent information as possible. Perhaps an intelligent language processing algorithm could highlight such users and prioritize the information they post.

Integrating social media data into the emergency response systems requires careful consideration about the quality of data, the way in which it has been produced and the useful information that can be extracted. In the Newcastle, United Kingdom, example mentioned above, 1,800 “related tweets” were identified. Yet, after filtering for (a) appropriate timestamp (during

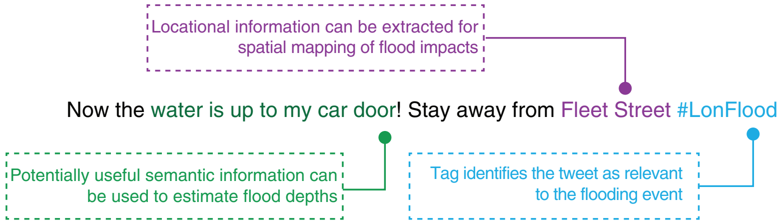

FIGURE 2 Extracting pertinent flood risk information form Twitter (Reprinted with permission from Smith et al. (2015). Copyright The Authors, 2018)

first 4 hr of the event) (b) sufficiently specific locational information, and (c) semantically "relevant" terms (e.g., "knee-height water"), just 43 tweets containing hazard relevant information remained (Smith et al., 2015). One analysis of flooding along the Red River, USA (2009) found that just 18% tweets were geotagged, further demonstrating that large quantities of "relevant" tweets may quickly be reduced to a sample of dubious statistical significance (Vieweg et al., 2010).

The subject of timing raises further challenges. Twitter data possesses a temporal aspect (with users more likely to post at certain times of day) while hazards such as surges are temporally indiscriminate. In fact, a contributory factor to the devastating impacts of the 1953 North Sea surge was its landfall during the night when most people were asleep (Pollard & Pollard, 2017). Furthermore, those active on social media reflect only a certain segment of society—and one that is arguably less vulnerable to extreme weather events and their aftermath. It is also unclear how the reliability or relevance of a given tweet could be assessed before including it in analysis that will be input to response agencies. Here, more than anywhere, ease of access should not be a mandate for use (Table 2).

4 | DISCUSSION

A review of relevant literature has identified five BDAs that hold the potential to improve risk assessment prior to flooding events and enhance emergency response procedures during the events themselves. Additionally, technical, contextual, institutional, and behavioral challenges associated with the approaches described have been discussed. Given the effort required to overcome these challenges, it is crucial that BDAs are applied critically to ensure that they genuinely inform coastal flood risk management rather than being employed as justification for decisions that would have been taken regardless. In addition to the challenges associated with specific approaches, two more intractable difficulties with applying BDAs to coastal flood risk management deserve attention: the place of BDAs in the broader flood system and the skill requirements for a generation of data scientists capable of implementing the BDAs.

4.1 | BDAs in a Broader Coastal Management Context

This paper has established numerous opportunities and challenges involved in applying BDAs to enhance coastal flood risk assessment and emergency response procedures. It is important to recognize that the coastal flood risk assessment and emergency response chains exist within a broader coastal risk management context, and beyond that, within an institutional, political cultural, and behavioral context (Evans et al., 2006; Narayan et al., 2012; Sayers et al., 2002; Thorne, Evans, & Penning-Rowsell, 2007). Explicit attention toward this broader context moves away from the notion of flood risk assessment and emergency response as linear frameworks (Narayan et al., 2014) and, crucially, draws attention to the broader context within which coastal flood risk management decisions are taken (Schanze, 2006).

The BDAs described in this paper deal exclusively with flood risk and yet in many coastal locations flooding is one of numerous potential risks. For example, a growing literature seeks to investigate the way in which erosion and flooding risks interact in specific localities to improve joint erosion-flooding risk assessment (Dawson et al., 2009; Grilli, Spaulding, Oakley, & Damon, 2017; Van Dongeren et al., 2014). Beyond this, a critical influence on whether decisions are ultimately financed and implemented is political climate (Kates, Colten, Laska, & Leatherman, 2006; Olasky, 2006). While resource constraints often disappear during emergency response procedures, they tend to rematerialize when it comes to financing longer term investments. Consequently, coastal risk reducing measures must compete against an array of alternative expenditure streams. Settling these kinds of decisions involves moral and political considerations that cannot be scientifically resolved. In this sense, the decision to undertake specific coastal risk management projects might be considered as trans-scientific (Weinberg, 1974) and an example of a "wicked problem" (Rittel & Webber, 1973), postnormal science (Funtowicz & Ravetz, 1993) or mode two science (Gibbons et al., 1994; Table 3). Drawn from the science and technology studies literature, these classifications emphasize that scientifically founded data analysis is limited in its usefulness for answering contentious policy

TABLE 2 Summary of BDAs for coastal flood risk assessment and emergency response procedures

Big Data Approach	Datasets	Tools (skill requirements)	Coastal flood risk applications	Decision contexts
1 Synthesis and harmonization of coastal datasets	Historical maps Aerial photographs LiDAR Coastal topography surveys Wave buoys Tide gauges Receptor databases	Statistical techniques (e.g., Probability density functions) Open source databases GIS Domain-specific knowledge	National flood event meta-analysis Flood risk communication Historical context for contemporary flood events Community-based decision-making	Risk assessment (SPRC characterization) Mitigation (through identifying areas in need of protection) Adaptation (through identifying persistently flooded areas where future investment may be less cost-effective)
2 Handling and validating satellite data	Multispectral satellite imagery Field study Numerical model outputs	Programming (e.g., Python or Java for Google Earth Engine) Machine learning algorithms Cloud-based computing (Google Earth Engine, EarthServer) Numerical modeling Fieldwork Domain-specific knowledge	Ecosystem flood protection services Global-scale flood risk analyses Flood risk communication	Risk assessment (SPRC characterization) Mitigation (through justifying habitat protection) Adaptation (identifying areas where habitat creation could occur) Resilience (through identifying synergistic ecosystem benefits)
3 Integrating process-based modeling and Bayesian Networks	Storm validation datasets Coastal zone features Receptor databases	Parallel processing Bayesian statistics Domain-specific knowledge	Probabilistic flood forecasting	Emergency response (monitoring, forecasting and detection) Preparation (most likely scenarios could be used as a basis for emergency response planning)
4 Ensemble forecasting	Global and regional climate models Weather stations Tide gauges Wave buoys	Ensemble forecasting system Supercomputers Domain-specific knowledge	Probabilistic flood forecasting Informing appropriate warning messages for flood risk communication	Emergency response (monitoring, forecasting and detection) Preparation (most likely scenarios could be used as a basis for emergency response planning)
5 Natural language processing of social media	Social media platform entries Geolocation	Natural language processing GIS Domain-specific knowledge	Real-time flood event information Resource allocation during flooding events.	Emergency response dissemination (a valuable channel for warning dissemination) Response (through helping to prioritize when numerous areas are simultaneously flooded)

Note. BDA = Big Data Approaches; GIS = Geographic Information System.

TABLE 3 Characterizing science given the expectation of application to public policy problems

Term	Meaning	Reference
Trans-science	When public policy asks questions of science that cannot be answered purely by scientific enquiry.	Weinberg (1974)
Wicked problems	When a scientific basis is sought to address policy problems that cannot be definitively described, nor can their solutions be universally agreed.	Rittel & Webber (1973)
Postnormal science	Science where the scale of system uncertainties and the decision stakes are large. Quantifiable uncertainty expands into indeterminacy and ignorance.	Funtowicz & Ravetz (1993)
Mode two science	Multiple competing knowledge claims; science driven by application to public policy problems; and assessed on these successes rather than scientific autonomy.	Gibbons et al. (1994)

questions (Sarewitz, 2004). The role of BDAs in informing coastal risk management policy will therefore be tempered by moral and political opinions (Kingdon, 1995; Rose, 2014).

Expanding the scope of BDAs to consider broader coastal flood risk management applications necessitates a discussion of the intractable forms of uncertainty that are introduced as a result. There is aleatory (inherent variability in the system) and epistemic (our limited knowledge about the system) uncertainty already contained within the risk assessment and emergency response chains, but when broader contexts are considered, questions of ambiguity, uncertainty, and ignorance bare even greater consequence (Stirling, 2007). Figure 3 establishes that risk, strictly defined, should refer to instances where the probabilities and outcomes of an event are known (Stirling, 2007). But this is only one of a number of possible knowledge states. Figure 3 also shows that uncertainty occurs when outcomes are known, but probabilities are undetermined; ambiguity characterizes situations when the outcomes themselves are under question; and ignorance reigns when both probabilities and outcomes are contested.

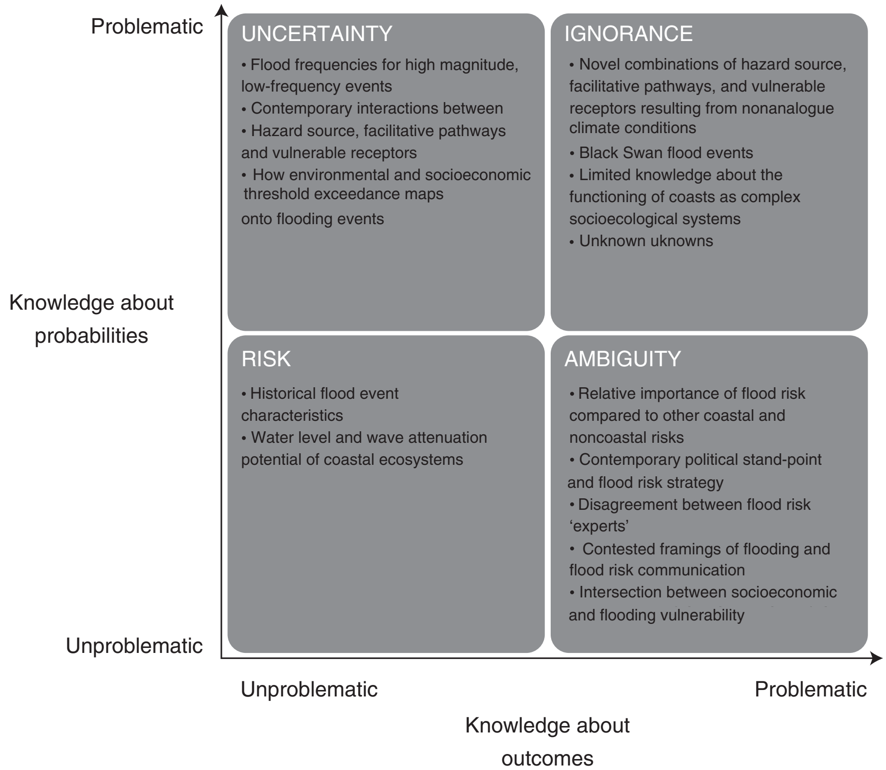

FIGURE 3 Flood events in the context of incertitude. This figure applies Stirling's (2007) categorization of incomplete knowledge states to flooding

BDAs can help to address some of the challenges presented in Figure 3, but not all. For example, BDA 1 (synthesis and harmonization of coastal datasets) may assist decision makers in situations of ambiguity by systematically cataloguing past flooding events. This can be used to determine the importance of flooding relative to other risks. It may also aid flood risk communication by drawing on historical events that communities themselves experienced (Siegrist & Gutscher, 2008). Indeed the construction of the database itself might represent an important collaborative (or coproductive) exercise between flood experts and impacted communities to assist in the selection of desired outcomes from flood risk management (Lane et al., 2011). Alternatively, the probabilistic treatment of uncertainty in both BDA 3 (integrating process-based modeling and BNs) and BDA 4 (ensemble forecasting) provides a valuable approach for dealing with uncertainty in the face of nonanalogue climate conditions and novel flood characteristics. Through access to the full range of possible outcomes and associated probability estimates, coastal decision makers can prioritize the most likely scenario while also preparing contingency measures to address other possible outcomes (Cowell, Thom, Jones, Everts, & Simanovic, 2006).

Although BDAs prove informative in situations of risk, uncertainty and ambiguity, they appear limited in their ability to inform decisions under conditions of ignorance. This derives largely from the characteristics of ignorance itself. It is difficult to apply BDAs in situations where appropriate data to construct probabilities is nonexistent and where there is no agreed outcome to inform appropriate tool selection. Here, it may be necessary to defer to more general coastal management approaches which invoke discussions around coastal flood resilience (Taussik, 2010). Given an inability to know future flood outcomes or their probability of occurrence, resilience thinking suggests that resources should be invested in ensuring systems are capable of withstanding event impacts and recovering to (at least) pre-event states (Southwick, Bonanno, Masten, Panter-Brick, & Yehuda, 2014). One manifestation of this decision-making under uncertainty is "no or low regret" options. These are defined as outcomes which "generate net social benefits with no or low regrets irrespective of the future outcome of climate change" (Cheong et al., 2013, p. 787). Such decisions should seek synergistic effects and multiple benefits. For example, saltmarsh restoration not only offers flood protection services but also habitat creation, which may be in line with certain political commitments, (such as the EU habitat Directive, 2008) in addition to offering reduced coastal erosion, carbon sequestration, and aesthetic and recreational benefits (Cheong et al., 2013; Friess et al., 2012). Identifying and pursing outcomes such as these facilitates effective decision-making even under conditions of ignorance (Environment Agency, 2015).

4.2 | Skill requirements for Coastal Data Scientists

In addition to negotiating various manifestations of uncertainty associated with the broader contextual influences, appropriate and effective implementation of BDAs relies on research and policy communities possessing the skills necessary to do so

(Table 2). It is those who have a detailed knowledge of coastal systems and experience managing them that are in the best position to ask appropriate questions of the data available and to be critical of the results that BDAs generate. This chimes with well-versed calls for a new generation of "data scientists" (Mattmann, 2013) that possess both domain-specific knowledge and the analytical skills necessary to apply big data techniques in novel contexts. Domain-specific knowledge is essential since carefully crafted research questions must remain the starting point to ensure that research continues to be data-driven but not data-determined. Similarly, among policy communities, decision makers require an understanding of the technical nature of BDAs used, their applicability to a specific management scenario, and the questions that they are unable to answer. Additionally, those responsible for data collection and provision must endeavor to record and publish appropriate metadata, again requiring domain-specific knowledge and an awareness of the potential end uses.

In the broader coastal risk management, it is these *coastal data scientists* that will be best equipped to make sense of the "datafication of the environment" (Nadim, 2016) and unpick the complexities of the "datascapes" (Salmond et al., 2017) that they are charged with cataloguing, analyzing, and interpreting. Even then, technical skills are not sufficient. Throughout, this paper has highlighted institutional, contextual, and behavioral barriers which may inhibit effective application of BDAs in the coastal flood risk management context. Coastal data scientists must also engage effectively in these aspects of the challenge. Simply applying more data to a pre-existing problem is unlikely to deliver the full potential of BDAs in coastal risk management. Communicating the validity of Big Data techniques and the political salience of the insights generated is a vital accompaniment to necessary technical analytical skills (see Table 2).

5 | CONCLUSION

Coastal flood risk is a global phenomenon (Kron, 2013; Rueda et al., 2017). Cumulative losses from flooding already dominate relative to other "natural" hazards (Kron, 2009). Furthermore, they are set to increase in the future due to intensified hazard sources, deterioration of protective defenses and ecosystems, and increased value and vulnerability of risk receptors (Hallegatte et al., 2013; Stott, 2016; Zhang, Douglas, & Leatherman, 2004). To some extent, the future scene is already set as lags in the interaction between socioeconomic and environmental spheres commit the global community to a higher flood risk future, demanding appropriate and effective coastal flood risk management.

BDAs are a suite of technological innovations designed to address the challenge of organizing, storing, and analyzing legacy datasets and incoming data streams. The ultimate aim is to apply BDAs to shed new light on societal and environmental complexities through novel organizational, storage and analytical capabilities. This paper emphasizes the codependence of datasets and the selection of appropriate tools to extract value from them. Five BDAs are identified and their potential to enhance coastal flood risk assessment and emergency response procedures is critically discussed. While these BDAs offer opportunities for improved decision-making in varied aspects of both decision chains, they are also accompanied by specific technical contextual, institutional, and behavioral barriers. These barriers must be overcome if the BDAs outlined here are to practically and genuinely inform coastal flood risk management.

In addition to specific implementation barriers, two more fundamental challenges have been identified. Establishing whether BDAs can continue to provide useful insights under conditions of incertitude (uncertainty, ambiguity, and ignorance) is critical given that coastal flood risk managers of the future must respond to nonanalogue environmental conditions. Through assimilating varied datasets, and probabilistic representations of uncertainty, the BDAs described here display capabilities under conditions of uncertainty and ambiguity. It is also necessary to recognize the limited ability of such approaches to cope with ignorance. In these instances, it may also be necessary to draw on the rhetoric of coastal resilience and associated no and low regret strategies. Another fundamental constraint on the application of BDAs is the skill sets of coastal researchers, policy and decision makers. This extends beyond the technical skills necessary to execute BDAs. In the future, there will be the need to interpret the insights generated from initial applications of BDAs and integrate them into existing management structures.

The coastal system is just one among many environmental systems that can expect to be directed toward nonanalogue conditions by global climate trends. Having established the potential theoretical and practical gains offered by BDAs in coastal research, the priority now must be to assess the suitability of BDAs in other environmental domains. Perhaps then the opportunities and challenges of applying BDAs to the full breadth of 21st century climate challenges can be established.

ACKNOWLEDGMENTS

This research is funded by a NERC/ESRC Data, Risk and Environmental Analytical Methods (DREAM) Centre for Doctorial Training PhD studentship [NE/M009009/1].

CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

RELATED WIREs ARTICLES

Using adaptation tipping points to prepare for climate change and sea level rise: a case study in the Netherlands
Responsibility for meeting the costs of adaptation

REFERENCES

Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., & Salmerón, A. (2011). Bayesian networks in environmental modelling. *Environmental Modelling and Software*, *26*, 1376–1388.

Azhoni, A., Holman, I., & Jude, S. (2017a). Contextual and interdependent causes of climate change adaptation barriers: Insights from water management institutions in Himachal Pradesh, India. *Science of the Total Environment*, *576*, 817–828.

Azhoni, A., Holman, I., & Jude, S. (2017b). Adapting water management to climate change: Institutional involvement, inter-institutional networks and barriers in India. *Global Environmental Change*, *44*, 144–157.

Baan, P., & Klijn, F. (2004). Flood risk perception and implications for flood risk management in the Netherlands. *International Journal of River Basin Management*, *2*, 113–122.

Bader, J., Mesquita, M. D. S., Hodges, K. I., Keenlyside, N., Østerhus, S., & Miles, M. (2011). A review on Northern Hemisphere sea-ice, storminess and the North Atlantic Oscillation: Observations and projected changes. *Atmospheric Research*, *101*, 809–834.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., … Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, *471*, 51–57.

Baumann, P., Mazzetti, P., Ungar, J., Barbera, R., Barboni, D., Beccati, A., … Wagner, S. (2016). Big data analytics for Earth Sciences: The EarthServer approach. *International Journal of Digital Earth*, *9*, 3–29.

Bertin, X. (2016). Storm surges and coastal flooding: Status and challenges. *La Houille Blanche*, *2*, 64–70.

Bertin, X., Li, K., Roland, A., Zhang, Y. J., Breilh, J. F., & Chaumillon, E. (2014). A modeling-based analysis of the flooding associated with Xynthia, central Bay of Biscay. *Coastal Engineering*, *94*, 80–89.

Botzen, W., Aerts, J., & Van Den Bergh, J. (2009). Dependence of flood risk perceptions on socioeconomic and objective risk factors. *Water Resources Research*, *45*, 1–15.

Brakenridge, G. R., Syvitski, J. P. M., Overeem, I., Higgins, S. A., Kettner, A. J., Stewart-Moore, J. A., & Westerhoff, R. (2013). Global mapping of storm surges and the assessment of coastal vulnerability. *Natural Hazards*, *66*, 1295–1312.

Brilly, M., & Polic, M. (2005). Public perception of flood risks, flood forecasting and mitigation. *Natural Hazards and Earth System Science*, *5*, 345–355.

Bughin, J., Chui, M., & Manyika, J. (2012). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly*. Retrieved from https://www.mckinsey.com/industries/high-tech/our-insights/clouds-big-data-and-smart-assets-ten-tech-enabled-business-trends-to-watch

Burn, D. (1999). Perceptions of flood risk: A case study of the Red River flood of 1997. *Water Resources Research*, *35*, 3451–3458.

Burningham, K., Fielding, J., & Thrush, D. (2008). 'It'll never happen to me': Understanding public awareness of local flood risk. *Disasters*, *32*, 216–238.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*, 171–209.

Cheong, S.-M., Silliman, B., Wong, P. P., van Wesenbeeck, B., Kim, C.-K., & Guannel, G. (2013). Coastal adaptation with ecological engineering. *Nature Climate Change*, *3*, 787–791.

Ciavola, P., Harley, M. D., & den Heijer, C. (2017). The RISC-KIT storm impact database: A new tool in support of DRR. *Coastal Engineering*, *134*, 24–32.

Copernicus Program. (2017). *Copernicus.eu* [WWW Document]. Retrieved from http://www.copernicus.eu/

Cowell, P. J., & Thom, B. G. (1994). Morphodynamics of coastal evolution. In R. W. G. Carter & C. D. Woodroffe (Eds.), *Coastal evolution: Late quaternary shoreline morphodynamics* (pp. 33–86). Cambridge, England: Cambridge University Press.

Cowell, P. J., Thom, B. G., Jones, R. A., Everts, C. H., & Simanovic, D. (2006). Management of Uncertainty in predicting climate-change impacts on beaches. *Journal of Coastal Research*, *221*, 232–245.

Dawson, R. J., Dickson, M. E., Nicholls, R. J., Hall, J. W., Walkden, M. J. A., Stansby, P. K., … Watkinson, A. R. (2009). Integrated analysis of risks of coastal flooding and cliff erosion under scenarios of long term change. *Climatic Change*, *95*, 249–288.

Defra. (2013). *Defra open data strategy* (pp. 29). London, England: Defra.

Demeritt, D., Nobert, S., Cloke, H., & Pappenberger, F. (2013). The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management. *Hydrological Processes*, *27*, 147–157.

Doney, S. C., Fabry, V. J., Feely, R. A., & Kleypas, J. A. (2008). Ocean acidification: The other CO 2 problem. *Annual Review of Marine Science*, *1*, 169–192.

Environment Agency. (2015). *No regrets planning for sea level rise and climate change and investing in adaptation*. London, England: Environment Agency.

European Commission. (2013). *Satellites reveal how sea level rise changes or coastlines*, European Space Agency.

Evans, E., Hall, J., Penning-Rowsell, E., Sayers, P., Thomas, C., & Watkinson, A. (2006). Future flood risk management in the UK. *Water Management*, *159*, 53–61.

FAST. (2017). *Foreshore assessment using space technology* [WWW Document]. Retrieved from http://www.fast-space-project.eu

FloodTags. (2017). *FloodTags* [WWW Document]. Retrieved from https://www.floodtags.com/

Ford, J. D., Tilleard, S. E., Berrang-Ford, L., Araos, M., Biesbroek, R., Lesnikowski, A. C., … Bizikova, L. (2016). Opinion: Big data has big potential for applications to climate change adaptation. *Proceedings of the National Academy of Sciences*, *113*, 10729–10732.

Friess, D. A., Spencer, T., Smith, G. M., Möller, I., Brooks, S. M., & Thomson, A. G. (2012). Remote sensing of geomorphological and ecological change in response to saltmarsh managed realignment, The Wash, UK. *International Journal of Applied Earth Observation and Geoinformation*, *18*, 57–68.

Functowicz, S., & Ravetz, J. (1993). Science for the post-normal age. *Futures*, *25*, 739–755.

Gabrys, J. (2016). Practicing, materialising and contesting environmental data. *Big Data & Society*, *3*, 1–7.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies* (p. 192). London, England: SAGE.

Google. (2017). *Google Earth engine* [WWW Document]. Retrieved from https://earthengine.google.com/

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Green, C. H., Tunstall, S. M., & Fordham, M. H. (1991). The risks from flooding: Which risks and whose perception? *Disasters*, *15*(3), 227–236.

Grilli, A., Spaulding, M. L., Oakley, B. A., & Damon, C. (2017). Mapping the coastal risk for the next century, including sea level rise and changes in the coastline: Application to Charlestown RI, USA. *Natural Hazards*, *88*, 389–414.

Haigh, I. D., & Nicholls, R. J. (2017). *Coastal flooding*. MCCIP Science Review, 108–114. Retrieved from http://www.mccip.org.uk/impacts-report-cards/full-report-cards/2017-10-year-report-card/

Haigh, I. D., Ozsoy, O., Wadey, M. P., Nicholls, R. J., Gallop, S. L., Wahl, T., & Brown, J. M. (2017). An improved database of coastal flooding in the United Kingdom from 1915 to 2016. *Scientific Data*, *4*, 1–10.

Haigh, I. D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., … Gouldby, B. (2016). Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK. *Scientific Data*, *3*, 1–14.

Hallegatte, S., Green, C., Nicholls, R. J., & Corfee-Morlot, J. (2013). Future flood losses in major coastal cities. *Nature Climate Change*, *3*, 802–806.

Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., … Porter, J. H. (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment*, *11*, 156–162.

Hay, C. C., Morrow, E., Kopp, R. E., & Mitrovica, J. X. (2015). Probabilistic reanalysis of twentieth-century sea-level rise. *Nature*, *517*, 481–484.

Heckerman, D., Geiger, D., & Chickering, D. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, *20*, 197–243.

Hsu, A. (2013). Does the environment need big data [WWW Document]. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/angel-hsu/does-the-environment-need_b_3568529.html

Hsu, L., Martin, R. L., McElroy, B., Litwin-Miller, K., & Kim, W. (2015). Data management, sharing, and reuse in experimental geomorphology: Challenges, strategies, and scientific opportunities. *Geomorphology*, *244*, 180–189.

IPCC. (2014). Climate change 2014: Synthesis report. In R. K. Pachauri & L. A. Meyer (Eds.) [Core Writing Team]. *Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (151 pp.). Geneva, Switzerland: Author.

Jagadish, H. V. (2015). Big data and science: Myths and reality. *Big Data Research*, *2*, 49–52.

Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian network and decision graph*. New York, NY: Springer-Verlag.

Jude, S. (2008). Investigating the potential role of visualization techniques in participatory coastal management. *Coastal Management*, *36*, 331–349.

Jude, S., Jones, A. P., Andrews, J. E., & Bateman, I. J. (2006). Visualisation for participatory coastal zone management: A case study of the Norfolk Coast, England. *Journal of Coastal Research*, *226*, 1527–1538.

Kates, R. W., Colten, C. E., Laska, S., & Leatherman, S. P. (2006). Reconstruction of New Orleans after Hurricane Katrina: A research perspective. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 14653–14660.

Keeso, A. (2014). *Big data and environmental sustainability: A conversation starter* (Working paper series 14-04, pp. 34). Oxford, England: Smith School.

Kellens, W., Zaalberg, R., Neutens, T., Vanneuville, W., & De Maeyer, P. (2011). An analysis of the public perception of flood risk on the Belgian Coast. *Risk Analysis*, *31*, 1055–1068.

Keller, C., Siegrist, M., & Gutscher, H. (2006). The role of the affect and availability heuristics in risk communication. *Risk Analysis*, *26*, 631–639.

Kingdon, J. W. (1995). *Agendas, alternatives, and public policies* (2nd ed.). New York, NY: Longman.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, *79*, 1–14.

Kitchin, R., & McArdle, G. (2016). What makes big data, big data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, *3*, 1–10.

Kongthon, A., Haruechaiyasak, C., Pailai, J., & Kongyoung, S. (2012). The role of Twitter during a natural disaster: Case study of 2011 Thai Flood. In *Proceedings of PICMET '12: Technology Management for Emerging Technologies* (pp. 2227–2232). Vancouver.

Kontkanen, P., Myllymaki, P., Silander, T., Tirri, H., & Grunwald, P. (1997). Comparing predictive inference methods for discrete domains. In *Proceedings of Sixth International Work on Artificial Intelligence and Statistics* (pp. 311–318). Florida.

Kron, W. (2009). Flood insurance: From clients to global financial markets. *Journal of Flood Risk Management*, *2*, 68–75.

Kron, W. (2013). Coasts: The high-risk areas of the world. *Natural Hazards*, *66*, 1363–1382.

Lane, S. N., Odoni, N., Landström, C., Whatmore, S. J., Ward, N., & Bradley, S. (2011). Doing flood risk science differently: An experiment in radical scientific method. *Transactions of the Institute of British Geographers*, *36*, 15–36.

Leetaru, K., Wang, S., Padmanabhan, A., & Shook, E. (2013). Mapping the global Twitter heartbeat: The geography of Twitter. *First Monday, 18*. Retrieved from http://firstmonday.org/article/view/4366/3654

Lorenzo-Trueba, J., & Ashton, A. D. (2014). Rollover, drowning, and discontinuous retreat: Distinct modes of barrier response to sea-level rise arising from a simple morphodynamic model. *Journal of Geophysical Research - Earth Surface*, *119*, 779–801.

Mattmann, C. A. (2013). Computing: A vision for data science. *Nature*, *493*, 473–475.

McCallum, E., & Young, M. V. (1988). *Spot the Great Storm. The forecasters' dilemma* (Meteorological office). Bracknell, England: Royal Meteorological Society.

Met Office. (2016). *Big Data drive* [WWW Document]. Retrieved from http://www.metoffice.gov.uk/services/data-provision/big-data-drive

Miller, H. J., & Goodchild, M. F. (2015). Data-driven geography. *GeoJournal*, *80*, 449–461.

Möller, I., Kudella, M., Rupprecht, F., Spencer, T., Paul, M., van Wesenbeeck, B. K., … Schimmels, S. (2014). Wave attenuation over coastal salt marshes under storm surge conditions. *Nature Geoscience*, *7*, 727–731.

Morris, E. P., Gomez-Enri, J., & van der Wal, D. (2015). Copernicus downstream service supports nature-based flood defence. *Sea Technology*, *56*, 23–26.

Murthy, D., & Longwell, S. A. (2013). Twitter and disasters. *Information, Communication & Society*, *16*, 837–855.

Nadim, T. (2016). Blind regards: Troubling data and their sentinels. *Big Data & Society*, *3*, 205395171666630.

Narayan, S., Hanson, S., Nicholls, R., Clarke, D., Willems, P., Ntegeka, V., & Monbaliu, J. (2012). A holistic model for coastal flooding using system diagrams and the Source-Pathway-Receptor (SPR) concept. *Natural Hazards and Earth System Science*, *12*, 1431–1439.

Narayan, S., Nicholls, R. J., Clarke, D., Hanson, S., Reeve, D., Horrillo-Caraballo, J., … Vanderlinden, J. P. (2014). The SPR systems model as a conceptual foundation for rapid integrated risk appraisals: Lessons from Europe. *Coastal Engineering*, *87*, 15–31.

Naylor, L. A., Spencer, T., Lane, S. N., Darby, S. E., Magilligan, F. J., Macklin, M. G., & Möller, I. (2017). Stormy geomorphology: Geomorphic contributions in an age of climate extremes. *Earth Surface Processes and Landforms*, *42*, 166–190.

Nicholls, R. J. (2004). Coastal flooding and wetland loss in the 21st century: Changes under the SRES climate and socio-economic scenarios. *Global Environmental Change*, *14*, 69–86.

Nicholls, R. J., Hanson, S. E., Lowe, J. A., Warrick, R. A., Lu, X., & Long, A. J. (2014). Sea-level scenarios for evaluating coastal impacts. *WIREs Climate Change*, 5, 129–150.

NOAA. (2017). *National oceanic and atmospheric administration* [WWW Document]. Retrieved from https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-ensemble-forecast-system-gefs

Olasky, M. (2006). *The politics of disaster: Katrina, big government, and a new strategy for future crises*. Nashville, TN: Thomas Nelson.

O'Neill, E., Brereton, F., Shahumyan, H., & Clinch, J. (2016). The impact of perceived flood exposure on flood-risk perception: The role of distance. *Risk Analysis*, *36*, 2158–2186.

Otto, F. E. L. (2016). Extreme events: The art of attribution. *Nature Climate Change*, *6*, 342–343.

Otto, F. E. L., van Oldenborgh, G. J., Eden, J., Stott, P. A., Karoly, D. J., & Allen, M. R. (2016). The attribution question. *Nature Climate Change*, *6*, 813–816.

Poelhekke, L., Jager, W. S., van Dongeren, A., Plomaritis, T. A., McCall, R., & Ferreira, O. (2016). Predicting coastal hazards for sandy coasts with a Bayesian network. *Coastal Engineering*, *118*, 21–34.

Pollard, M. & Pollard, J. A. (2017). *North Sea surge: The story of the East Coast floods of 1953* (2nd ed.). Suffolk, England: Terence Dalton Ltd.

Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*, 155–169.

Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., & Lescinski, J. (2009). Modelling storm impacts on beaches, dunes and barrier islands. *Coastal Engineering*, *56*, 1133–1152.

Rose, D. C. (2014). Five ways to enhance the impact of climate science. *Nature Climate Change*, *4*, 522–524.

Rowling, M. (2015). Tweets turned into flood maps that could help save lives [WWW Document]. *Reuters*. Retrieved from http://www.reuters.com/article/us-disaster-flood-socialmedia/tweets-turned-into-flood-maps-that-could-help-save-lives-idUSKBN0N51W020150414

Rueda, A., Vitousek, S., Camus, P., Tomás, A., Es, A., Inigo, J., … Reguero, B. G. (2017). A global classification of coastal flood hazard climates associated with large-scale oceanographic forcing. *Natural Science Reports*, *7*, 1–8.

Rumson, A. G., Hallett, S. H., & Brewer, T. R. (2017). Coastal risk adaptation: The potential role of accessible geospatial Big Data. *Marine Policy*, *83*, 100–110.

Salmond, J. A., Tadaki, M., & Dickson, M. (2017). Can big data tame a “naughty” world? *The Canadian Geographer/Le Géographe Canadian*, *61*, 52–63.

Santos, V. M., Haigh, I., & Wahl, T. (2017). Spatial and temporal clustering analysis of extreme wave events around the UK coastline. *Journal of Marine Science and Engineering*, *5*, 28.

Saravanou, A., Valkanas, G., Gunopulos, D., & Andrienko, G. (2015). Twitter floods when it rains: a case study of the UK floods in early 2014. In *WWW 2015 Companion* (pp. 1–14).

Sarewitz, D. (2004). How science makes environmental controversies worse. *Environmental Science & Policy*, *7*, 385–403.

Sayers, P. B., Hall, J. W., & Meadowcroft, I. C. (2002). Towards risk-based flood hazard management in the UK. *Proceedings of the Institution of Civil Engineers—Civil Engineering*, *150*, 36–42.

Schanze, J. (2006). Flood risk management: A basic framework. In J. Schanze, E. Zeman, & J. Marsalek (Eds.), *Flood risk management: Hazards, vulnerability and mitigation measures* (p. 319). Netherlands: Springer.

Shubert, W. (2015). Dust in the wind: How data visualization can help the environment [WWW Document]. *Scientific American*. Retrieved from https://blogs.scientificamerican.com/guest-blog/dust-in-the-wind-how-data-visualization-can-help-the-environment/

Siegrist, M., & Gutscher, H. (2006). Flooding risks: A comparison of lay people's perceptions and expert's assessments in Switzerland. *Risk Analysis*, *26*, 971–979.

Siegrist, M., & Gutscher, H. (2008). Natural hazards and motivation for mitigation behavior: People cannot predict the affect evoked by a severe flood. *Risk Analysis*, *28*, 771–778.

Skinner, D. J. C., Rocks, S. A., & Pollard, S. J. T. (2016). Science of the total environment where do uncertainties reside within environmental risk assessments? Expert opinion on uncertainty distributions for pesticide risks to surface water organisms. *Science of the Total Environment*, *572*, 23–33.

Smith, L., Liang, Q., James, P., Lin, W., & Liang, Q. (2015). Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. *Journal of Flood Risk Management*, *10*, 370–380.

Southwick, S. M., Bonanno, G. A., Masten, A. S., Panter-Brick, C., & Yehuda, R. (2014). Resilience definitions, theory, and challenges: Interdisciplinary perspectives. *European Journal of Psychotraumatology*, *5*, 1–14.

Stirling, A. (2007). Risk, precaution and science: Towards a more constructive policy debate. Talking point on the precautionary principle. *EMBO Reports*, *8*, 309–315.

Stott, P. (2016). How climate change affects extreme weather events. *Science*, *352*, 1517–1518.

Takahashi, B., Tandoc, E. C., Jr., & Carmichael, C. (2015). Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines. *Computers in Human Behavior*, *50*, 392–398.

Taussik, J. (2010). Planning for coastal resilience: Best practice for calamitous times. *Journal of Environmental Policy and Planning*, *12*, 331–333.

Terpstra, T., Gutteling, J., Geldof, G., & Kappe, L. (2006). The perception of flood risk and water nuisance. *Water Science and Technology*, *54*, 431–439.

Thorne, C., Evans, E., & Penning-Rowsell, E. (2007). *Future flooding and coastal erosion risks* (p. 505). London, England: Thomas Telford.

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, *203*, 312–318.

Van Dongeren, A., Ciavola, P., Viavattene, C., de Kleermaeker, S., Martinez, G., Ferreira, O., … McCall, R. (2014). RISC-KIT: Resilience-increasing strategies for coasts—toolKIT. *Journal of Coastal Research*, *70*, 366–371.

Vieweg, S., Hughes, A.L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: What twitter may contribute to situational awareness. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10* (pp. 1079–1088). New York, NY: ACM Press.

Vitolo, C., Elkhatib, Y., Reusser, D., & Macleod, C. J. A. (2015). Environmental Modelling & Software Web technologies for environmental Big Data. *Environmental Modelling and Software*, *63*, 185–198.

Vitousek, S., Barnard, P. L., Fletcher, C. H., Frazer, N., Erikson, L., & Storlazzi, C. D. (2017). Doubling of coastal flooding frequency within decades due to sea-level rise. *Scientific Reports*, *7*, 1399.

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., & Slangen, A. B. A. (2017). Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nature Communications*, *8*, 1–12.

Walstra, D. J. R., Hoekstra, R., Tonnon, P. K., & Ruessink, B. G. (2013). Input reduction for long-term morphodynamic simulations in wave-dominated coastal settings. *Coastal Engineering*, *77*, 57–70.

Weinberg, A. (1974). Science and trans-science. *Minerva*, *10*, 290–222.

Wyncoll, D., Haigh, I., Gouldby, B., Hames, D., Laeger, S., Wall, A., . . ., Hammond, A. (2016). Spatial analysis and simulation of extreme coastal flooding scenarios for national-scale emergency planning. In *E3S Web of Conferences 7*, 1001.

Zhang, K., Douglas, B. C., & Leatherman, S. P. (2004). Global warming and coastal erosion. *Climatic Change*, *64*, 41–58.

How to cite this article: Pollard JA, Spencer T, Jude S. Big Data Approaches for coastal flood risk assessment and emergency response. *WIREs Clim Change*. 2018;e543. https://doi.org/10.1002/wcc.543