



# Innovations in the use of data facilitating insurance as a resilience mechanism for coastal flood risk

Alexander G. Rumson\*, Stephen H. Hallett

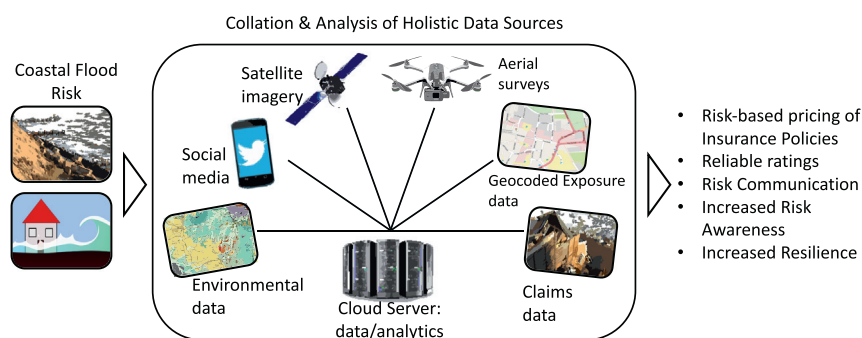
School of Water, Energy and Environment, Cranfield University, Bedford, UK



## HIGHLIGHTS

- Data-driven insurance is capable of redistributing and mitigating coastal flood risk.
- Insurance sector interviews reveal scope for adoption of emerging data technologies.
- Data innovations allow insurance to operate an evidence-based risk advisory service.
- Satellite-derived data holds untapped potential to enhance insurance risk analyses.
- Big Data allows fusion of vast empirical datasets with claims data to generate insight.

## GRAPHICAL ABSTRACT



## ARTICLE INFO

### Article history:

Received 27 September 2018

Received in revised form 20 December 2018

Accepted 10 January 2019

Available online 14 January 2019

Editor: Ralf Ludwig

### Keywords:

Risk analytics

Adaptation

Remote sensing

Big Data

## ABSTRACT

Insurance plays a crucial role in human efforts to adapt to environmental hazards. Effective insurance can serve as both a measure to distribute, and a method to communicate risk. In order for insurance to fulfil these roles successfully, policy pricing and cover choices must be risk-based and founded on accurate information. This is reliant on a robust evidence base forming the foundation of policy choices. This paper focuses on the evidence available to insurers and emergent innovation in the use of data. The main risk considered is coastal flooding, for which the insurance sector offers an option for potential adaptation, capable of increasing resilience. However, inadequate supply and analysis of data have been highlighted as factors preventing insurance from fulfilling this role. Research was undertaken to evaluate how data are currently, and could potentially, be used within risk evaluations for the insurance industry. This comprised of 50 interviews with those working and associated with the London insurance market. The research reveals new opportunities, which could facilitate improvements in risk-reflective pricing of policies. These relate to a new generation of data collection techniques and analytics, such as those associated with satellite-derived data, IoT (Internet of Things) sensors, cloud computing, and Big Data solutions. Such technologies present opportunities to reduce moral hazard through basing predictions and pricing of risk on large empirical datasets. The value of insurers' claims data is also revealed, and is shown to have the potential to refine, calibrate, and validate models and methods. The adoption of such data-driven techniques could enable insurers to re-evaluate risk ratings, and in some instances, extend coverage to locations and developments, previously rated as too high a risk to insure. Conversely, other areas may be revealed more vulnerable, which could generate negative impacts for residents in these regions, such as increased premiums. However, the enhanced

*Abbreviations:* ABM, Agent Based Model; ANN, Artificial Neural Network; API, Application Programming Interface; CAT, Catastrophe; GIS, Geographical Information Systems; GUI, Graphical User Interface; EA, Environment Agency (UK); EO, Earth Observation; HPCC, High Performance Computing Cluster; LIDAR, Light Detection and Ranging; IoT, Internet of Things; LM TOM, London Market Target Operating Model; NLP, Natural Language Processing; QA, Quality Assurance; RDBMS, Relational Database Management System; SaaS, Software as a Service; SAR, Synthetic Aperture Radar; SQL, Structured Query Language.

\* Corresponding author at: School of Water, Energy and Environment, Building 53, Cranfield University, Bedford MK43 0AL, UK.

E-mail address: [al.rumson@cranfield.ac.uk](mailto:al.rumson@cranfield.ac.uk) (A.G. Rumson).

risk awareness generated, by new technology, data and data analytics, could positively alter future planning, development and investment decisions.

© 2019 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Insurance permits the issue of information asymmetry, between the insurer and the insured, to be addressed. In correctly rating risk, insurance can thus enable risk-transfer between clients (policyholders) and the global insurance and capital markets (Cervantes-Godoy et al., 2013; von Dahlen and von Peter, 2012). As a result of highly developed, globalised, reinsurance markets, risk from catastrophic (CAT) losses occurring in a single locality can be transferred across the world (Abdullayev, 2014; von Dahlen and von Peter, 2012). In a perfectly competitive market, the market would be price setting. However, this is rarely the case in practice, so for markets to generate risk-reflective pricing (or an actuarial fair rate (Kunreuther et al., 2016)), underwriters and actuaries require access to accurate, up to date information detailing the nature of the risks associated with each class of business (Actuaries Institute, 2016). Insurance pricing also provides a mechanism for risk signalling, which can act to raise awareness and encourage risk-averse behaviour (Bin et al., 2006; Hudson et al., 2016). If market distortions occur though, this message can become diluted, resulting in adverse societal consequences (Hudson et al., 2016). In such cases, if the risk reducing element of insurance is lost, a moral hazard could be created (Surminski and Oramas-Dorta, 2014). Governments can play a crucial role in relation to insurance. For example, for flood insurance, this can take the form of land planning, investments in adaptations, and provision of cover to some of the most vulnerable (OECD, 2016; Surminski, 2014) (which may be increasingly necessary given climate change predictions (Lamond and Penning-Rowsell, 2014; Thistlethwaite et al., 2018)). However, some government interventions, such as last resort insurance coverage, have been reported to create market distortions, preventing insurance from fulfilling its full socio-economic potential (Crick et al., 2018; Kunreuther et al., 2016), or reducing the incentive for households to take preventative adaptation measures (Surminski, 2014). Yet finding an insurance arrangement that optimises risk reduction is not simple and has been acknowledged as an international challenge (Surminski and Oramas-Dorta, 2014).

Since the 1970s losses have been growing (especially from weather-related incidents), with non-insured losses growing the fastest (von Dahlen and von Peter, 2012; Kunreuther et al., 2016; OECD, 2016). For insurance and reinsurance markets to function effectively, it is essential for risks to be both priced appropriately, and for coverage to be extended to those in need. Recent advances in realm of data and analytics are reported to have increased the supply of reinsurance for flood risk (The American Academy of Actuaries, 2017). For both insurance and reinsurance, it is essential for analysts to supply the information required to allow exposure management, so aggregation of risks and exposure to natural perils, can be established (Andrews et al., 2008). The opportunities presented by the vast stores of data which are continually becoming available (Actuaries Institute, 2016; Choi and Lambert, 2017; Rumson and Hallett, 2018), open up possibilities for risk to be priced more accurately (Stoekli et al., 2018). Inevitably more accurate risk evaluations (and potentially the use of 'Big Data' (Actuaries Institute, 2016)) will create losers as well as beneficiaries, for example some geographical areas reassessed as being higher risk, may currently benefit from unrealistically priced insurance premiums. In such cases, current policy holders may be priced out of the market (Collinson, 2017). On aggregate though, this kind of outcome is socially optimal, and can result in insurers lowering their risk ratings, and premiums, for other, less vulnerable locations. This can address the pressing problem of asset underinsurance (Kunreuther, 1984; Kunreuther et al., 2016; Lloyd's, 2018a), and potentially result in increased investment and a rise in

sustainable developments in more resilient areas. Positive outcomes may also be generated, such as areas previously being regarded as off limits to investors becoming an attractive option and potentially, as a consequence, regional economic regeneration occurring. This paper reveals how emergent innovation in the use of data, can improve the ability of coastal flood insurance, to facilitate adaptation and increase resilience. Literature cited within this paper reveals how the potential for insurance to increase resilience to coastal flooding has been acknowledged. However, the role of data, in ensuring the effective functioning of insurance, has been widely overlooked. This work seeks to address this issue.

## 2. Methods

In addressing the issue of how to increase the capacity of insurance to act as a resilience increasing mechanism, our research considers how data is consumed within the insurance industry and the potential role of innovations in the use of data and analytics. This has entailed researching data sources, data analytics, and methods of communicating information outputs. There are abundant suppliers of data and analytics in this field, however there is currently a lack of rigorous academic evaluation addressing the associated range of data-related challenges and opportunities. The first part of the research comprises a literature review, considering the role of insurance in relation to flood risk adaptation in coastal areas. The literature review drew on a wide range of sources including academic papers, grey literature and industry related websites. Multiple combinations of key words and phrases were used within literature searches, these included: coast\*, flood\*, insurance, reinsurance, adaptation, resilience, 'risk mitigation', data, 'data source\*', 'data analytics', 'geospatial data', and 'flood model\*'. Over 30 relevant academic papers were identified, however emphasis was placed on using more recent literature, as such, the majority of academic sources cited were published within the last 10 years.

In the sections, following the literature review, the role of data and analytics is addressed, drawing on feedback obtained from 50 semi-structured interviews with a broad range of practitioners, working in and associated with, the London Insurance market (including risk engineers, brokers, actuaries, underwriters, analysts and managers), and representatives of firms who supply data and analytics (such as CAT modellers, specialist insurance analytics firms, flood modellers, and suppliers of geospatial data). In many instances single interviews were conducted with two or more representatives of an organisation. In terms of the backgrounds of those interviewed, this can be loosely categorised as follows: 20 were from the insurance sector, 6 from data providers, 8 from insurance specific analytics firms (such as CAT modellers), 3 from more general data analytics organisations, 10 from satellite data analytics suppliers, and 3 from the field of Big Data solutions. The use of Earth Observation (EO) data, emerged as a prominent theme, as such feedback on advances in the use of EO data was provided in interviews with representatives of multiple organisations who work in this field.

All interviews were completed within a 3-month period (November 2017 to January 2018). Interviews commenced with a briefing on the nature of the research being conducted and the neutral position of interviewer, who was not connected or sponsored by any organisation linked to the insurance industry. A standardised set of questions were covered, which addressed topics outlined in Fig. 1; the questions differed depending on the category of organisation the interviewee belonged to. Interviews were not recorded, however extensive notes were made from which transcripts were produced. Interview

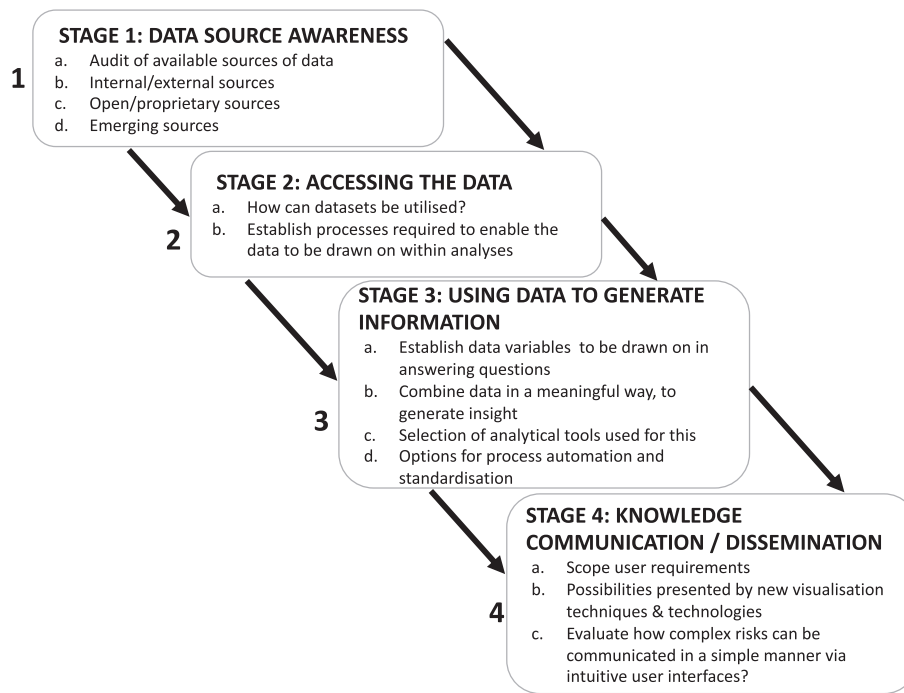


Fig. 1. Data utilisation for insurance.

transcripts were analysed systematically, and responses given grouped into themes. The themes evolved through a process of manually comparison, and following this generation of word counts for specific terms. The most prominent themes identified have been elaborated on and led to development of the subsequent sections of this paper. The London market was selected as a case study as it is one of the oldest and most comprehensive insurance markets in the world, covering international risks, and a wide range of multinational firms are represented within it. Furthermore, it is quoted to be ‘the largest global hub for commercial and specialty risk’ (London Market Group, 2018). Tidal flooding is also a significant issue for London, which has led to implementation of innovative adaptations, such as the Thames Barrier (Linham and Nicholls, 2010). Lloyd’s provided a focal point for this research, and assistance was provided by Lloyd’s Data Lab, in selecting and securing appropriate interviewees. Details of specific contributors are omitted to protect their identities, also none of the companies mentioned have been associated with opinions expressed by their employees (who have been anonymised).

### 3. Flood risk adaptation in coastal areas: the role of insurance

Those settling, using and working in coastal locations must contend with numerous hazards. Of these, flooding is one of the most prominent and can be severe and extensive. Flood impacts are compounded by the presence of critical infrastructure in coastal areas, e.g. due to requirements for ocean access (i.e. for oil, gas, and renewables) or the need for water cooling (nuclear power plants). Additionally, most major global cities are sited on or near the coast due to needs for port and shore access. Recent events in the Caribbean and United States reveal how extreme weather-related hazards cause devastating effects in coastal areas, and how losses are transferred to insurers (Lloyd’s, 2018a). Flooding is one of the major perils which generated losses from these events, resulting in calls for improvement of the flood modelling process (Lloyd’s, 2018a). In many locations impacted, a primary form of defence is artificial protection, such as engineered coastal defences. However, there are a number of problems with such structural defences (Crichton, 2008). Where such measures prove inadequate in ensuring resilience in the resident populations (Kunreuther et al.,

2016), devastation reaped by events such as hurricane winds and storm surges, can reach beyond what individuals are capable of covering financially. Therefore, insurance against natural perils, such as flooding, is considered a significant element within coastal management (Clark, 1998), which can facilitate recovery (Viavattene et al., 2018), and has been termed a ‘catalyst for resilience’ (Kunreuther et al., 2016). The type of risk covered by insurance is fortuitous risk, which is a risk related to accidental or chance events. In this sense the risk of flooding is more suited to insurance than erosion, which in many locations is inevitable.

Insurance is acknowledged as having a crucial role in redistributing risk (von Dahlen and von Peter, 2012). Insurability, or lack of this, can also serve as a tool to raise awareness of the real risks associated with settling in coastal areas, deterring investment in high risk, hazard-prone locations. In this sense, insurance has a role as a planning instrument in relation to controlling impacts on flood plain geography (Crichton, 2008). It also has a clear role within the housing market, in that market value of houses are seen to reflect perceived risk (Jongman et al., 2014; Pilla et al., 2018). In influencing asset values, insurance can also affect developers’ decisions to build in coastal areas. This is noted by Botzen & van den Bergh (Botzen and van den Bergh, 2008) who highlight how varying premiums can serve to reduce risk indirectly, by reducing the desirability to settle in high-risk areas. The application of larger more granular datasets, is highlighted to contribute to such improvements (Actuaries Institute, 2016). In fact, flood insurance premiums are said to account for up to 80% of reductions in real-estate prices in flood plains (Filatova et al., 2011). This form of house price discounting, although unpopular with real-estate owners, can lead to less overall damage arising from flooding. Furthermore, Filatova et al. (Filatova et al., 2011) conclude that when combined with building on higher ground, insurance can offer the best means of communicating risk.

Kron (Kron, 2013) describes insurance as covering a range of activities on the coast other than just real estate, including: fish farms, bio-fouling of hydraulic structures and vessels by toxic algae, and indirect impacts on hotels and resorts. Flood hazards can generate physical damage to households, businesses and infrastructure, but can also result in pollution and impacts to human health and welfare, as well as creating

widespread business disruption and supply chain shocks (which can exceed direct damages) (Jongman, 2018). These risks are exacerbated by urbanisation and increasing population densities in coastal areas (Kron, 2013; Kunreuther et al., 2016). Insurance has a clear role in creation of incentives to reduce such risk (Kunreuther et al., 2016). Yet standard, static, insurance risk assessments based on a limited number of data variables can underestimate risk (Haer et al., 2017). Noting this, a wide range of data is required to enable comprehensive analysis of flood risk. This extends beyond peril data and projected hazard propagation, to data relating to human behaviour and use of areas at risk of flooding (Yang et al., 2018). Data also needs to be provided covering wider consequences of infrastructure failures, for example that related to roads, power stations, water supply, and port facilities (Kunreuther et al., 2016). One method which has been applied to understanding relationships between such data is Agent Based Models (ABM). For example ABMs have been applied to understanding disruptions generated by environmental hazards to critical infrastructure, resulting in power outages (Walsh et al., 2018). Increasing numbers of coastal flood models are also becoming available which draw on such holistic data sources, including the CFFlood model (Mokrech et al., 2014), which allows impact analyses, under varying socio-economic and climate scenarios (combining environmental and human datasets).

### 3.1. Adaptation through insurance

Insurance's role as an adaptation mechanism extends beyond influencing real-estate prices and development decisions; in its ability to communicate risk, insurance can also spur and encourage investment in adaptive capacity at a household level (Filatova et al., 2011; Hudson et al., 2016). Yet, the ability for insurance to function as an adaptation mechanism depends on its availability, the resulting coverage achieved, and is based on the premise that insurers are able to operate in an open market place. However, this is not the case in many countries, as approaches to state and private flood insurance provision vary across the world (Crichton, 2008, 2004; Lamond and Penning-Rowsell, 2014), and in some instances this can create distortions. Throughout Europe a number of different approaches have been adopted in relation to insurance of coastal flood risk. In the UK a private system operates for flood risk, yet no insurance is available against erosion (Dávila et al., 2014). In France a public/private partnership exists, in which flood insurance is mandatory as a part of buildings insurance (Hudson et al., 2016). For the past 60 years in the Netherlands there has been a public flood compensation scheme, yet this has been considered inefficient, and now private schemes are being looked to (Botzen and van den Bergh, 2008). Germany has combined flood insurance within private insurance packages, resulting in only 10% coverage for flood risk (Botzen and van den Bergh, 2008). The French system can be seen as effective in that it secures close to 100% coverage, yet the method of implementation creates distortions, and the French national insurance system is regarded as not supporting reduction of individual risk (Botzen and van den Bergh, 2008). By contrast, in the UK 'insurance companies differentiate premiums based on geographical risk characteristics' which reward settlement in low-risk areas; an ethical problem exists though due to low coverage (30%) in poor households (Botzen and van den Bergh, 2008). Similar ethical challenges have been reported in other parts of the world also, such as in Australia (Actuaries Institute, 2016).

Within the UK a unique situation has prevailed, taking the shape of a 'Statement of principles' (Association of British Insurers, 2005) between the government and the insurance industry, whereby the government commits to build defences whilst insurance companies continue to provide cover for flooding (Jongman et al., 2014; Surminski and Eldridge, 2015). However, this agreement failed to take account of affordability of the insurance cover provided to those living in areas vulnerable to flooding. To address this issue, in 2015 the British government introduced a scheme labelled Flood RE ("Flood RE," n.d.). This aims to enable those who live in properties at the highest risk of flooding, to gain

affordable home insurance. The current scheme, although aiming to address a serious concern for those living in areas prone to flooding (i.e. the inability to insure their real estate assets), has been found likely to generate moral hazard, due to it overlooking the risk signalling aspect of insurance, which can encourage more risk-averse behaviour (Surminski and Eldridge, 2015). This results from policy pricing being decoupled from true levels of risk, for policies underwritten through Flood RE. Nevertheless, it does discourage future developments in high risk locations, as the scheme only applies to houses constructed prior to 2009. As such, new homes built in flood plains, wouldn't be covered by the scheme, and therefore may be uninsurable.

Similarities can be drawn between Flood RE and the French insurance system, which as a result of imposing uniform premiums (or a universal surcharge) fails to account for varying levels of risk, and has been branded an inefficient risk communication mechanism (Dávila et al., 2014). Elsewhere in Europe other challenges are faced in relation to the application of insurance to coastal flood risk. In Spain there are issues hampering effective utilisation of insurance, these relate to perceived weaknesses of the law courts and low confidence levels, due to many claims not being paid (Dávila et al., 2014). Italy also suffers from a confidence problem due to low levels of trust in national institutions and insurance companies (Dávila et al., 2014). Nevertheless in some locations beyond Europe, such as the USA, insurance is seen to be a 'primary tool of improving location choice in flood prone areas' (Filatova et al., 2011).

It is important for insurance schemes to incorporate risk reduction elements, widening their focus beyond risk transfer alone (Surminski and Oramas-Dorta, 2014). Where this aspect has been neglected, such as in an example of state backed insurance for unsustainable developments on barrier islands, the availability of insurance is reported to exacerbate problems (McNamara and Werner, 2008). As such, insurance policies which encourage rebuilding in high risk locations, as opposed to resettlement, can negatively impact future resilience. Roberts (Roberts, 2012) outlines how policies addressing coastal risks can generate an unintended outcome, whereby the burden of compensation for developments in high-risk areas can fall on society. Additionally, many countries operate cross-subsidy insurance coverage (or a bundle system, where other perils are combined with flooding (Crichton, 2008; Lamond and Penning-Rowsell, 2014)), this offers the benefit of reducing insurance premiums through further spreading risk. However, it can also prove an excuse for inaction by those settling in high risk areas (Dávila et al., 2014), as it can dilute the apparent risk posed by specific perils. In some extreme circumstances, increasing property values in areas prone to flooding (potentially resulting from increased levels of protection), can render insurance in these areas impossible (Jongman et al., 2014). In fact, a high proportion of properties in flood plains remain uninsured (von Dahlen and von Peter, 2012; Landry and Jahan-Parvar, 2011; Lloyd's, 2018a; OECD, 2016). This can pose a direct barrier to insurance achieving the resilience increasing function described. In order to tackle issues such as this, Roberts (Roberts, 2012) proposes a form of compulsory insurance, arguing that in practice only obligatory insurance schemes appear capable of establishing a fully functioning community of insureds. Filatova et al. ((Filatova et al., 2011), p. 169) concur, stating how compulsory insurance can force homeowners 'to face the social cost of locating in a flood plain'.

### 4. The role of data in provision of insurance as an adaptation mechanism

Relevant data and information play a significant role in ensuring the desired outcomes are achieved through insurance cover. The success of flood insurance schemes is said to be reliant on how sophisticated a country's insurers are in mapping flood risks (Dávila et al., 2014). In order for insurance to function as an adaptation mechanism (through communicating and redistributing risk effectively), those who provide and underwrite insurance policies are required to use representative

risk matrices for rating specific locations. For these assessments to characterise accurately the vulnerability and resilience of assets, comprehensive datasets are required, representing environmental factors, adaptation measures, past and projected impacts and consequences. Through combining insurers' internal data, such as damage reports, with household level information, can reveal drivers for implementation of household adaptation measures (Osberghaus, 2017). Additionally, human behaviour needs to be accounted for. ABMs can be well suited to this task. For example, ABMs have revealed how risk averse behaviour, in response to increased risk awareness can have serious implications, altering projected risk ratings (Crick et al., 2018; Dawson et al., 2011; Dubbelboer et al., 2017; Haer et al., 2017; Jenkins et al., 2017; Yang et al., 2018). They have also been used to reveal which insurance arrangements can prove most conducive to risk reduction (Crick et al., 2018). Furthermore, through incorporating behavioural responses of humans, to hazard events such as storm surge flooding, ABMs have allowed areas vulnerable to disruption to be identified (Dawson et al., 2011).

It is deemed insufficient for insurers to merely have access to data, they require the capacity to process, analyse and communicate outputs from the data. In line with this a framework has been devised, involving four stages (represented below in Fig. 1). This framework has been used as the basis to structure feedback, received from interviews undertaken. The following sections address various aspects of these four stages, from outlining data sources, to methods of accessing these, and discussions around internal data, Open Source data, the application of Earth Observation (EO) data, Big Data, data analytics, how data is drawn on by underwriters, and finally challenges.

## 5. Data utilisation: London insurance market interview feedback

### 5.1. Data sources

Through conversations and interviews conducted for this study, a number of datasets were highlighted as being of particular importance and interest for flood risk evaluations for the insurance industry. These datasets have been split into themes, which for simplicity have been grouped by the approximate category they belong to (Table 1). For all themes listed, direct reference was made to associated data, within at least one interview.

### 5.2. Data access

There are many ways of obtaining data. Open Source data can be obtained free of charge, whilst proprietary data can be downloaded, if a subscription is bought or a one-off payment made. Further to this, data can be 'scraped' from web sources (one insurers and one representative of an analytics company, reported obtaining data this way). This process can draw on hash tags and geotags and even involve scraping social media feeds for information related to hazards such as floods. Data obtained through social media sites (such as Twitter) are increasingly recognised as offering a valuable input to flood risk analytics and recently a number of studies have been undertaken focusing on this area (de Bruijn et al., 2018; Jongman et al., 2015; Smith et al., 2017; Wang et al., 2018). Another method of obtaining and delivering data, is web-feeds (such as Application Programming Interfaces (API)). It is possible to embed data feeds from external websites within a user interface or webpage. This has been reported an increasingly common method for many analytics firms to provide underwriters with data.

No value is derived from holding data which cannot be drawn on when needed though, and this can result in the data providing 'DRIP' support - namely Data Rich Information Poor (Wilding et al., 2017). To avoid this, careful consideration needs to be given to how the data can be utilised effectively. It is regarded essential to focus initially on data inputs, and to transform the data into a format which can be readily worked with. In some cases, this can necessitate seeking out those

**Table 1**  
Data themes commonly utilised in Insurance Flood Risk Analysis.

Category	Data themes
Environmental	<ul style="list-style-type: none"> <li>• Environmental risks (extreme weather events, natural disasters, climate change, loss of natural capital, air/soil/-water pollution)</li> <li>• General information on local environments</li> <li>• Historical records of contamination and pollution events</li> <li>• Threats to natural resources (e.g. salinization of aquifers from flooding)</li> <li>• Land use change (urbanisation, industrialisation, changes in exposure)</li> <li>• Oceanographic records and projections (wave climates, sea temperature, water quality, marine fauna/flora, coastal processes)</li> <li>• Tidal data</li> <li>• Past storm surge events and impacts</li> <li>• River and estuarine data (river levels, flow rates)</li> <li>• Natural capital/habitats/ecosystem services (quantification, loss/gain)</li> <li>• Records and predictions of beach/loss creation (change calculations, modelling outputs)</li> <li>• Contaminant and pollution sources in flood plains</li> <li>• Location of landfill and sewage sites</li> </ul>
Flooding	<ul style="list-style-type: none"> <li>• Flooding records, predictions (extents taken from aerial imagery, EO data, water level gauges)</li> <li>• Flood risk exposure (publicly available modelling outputs)</li> <li>• Flood defences/adaptations (location and condition)</li> <li>• Flood protection offered by natural habitats</li> <li>• Flood damage costs (records of financial impacts to people, property, business and infrastructure)</li> <li>• Flood specific geotagged social media data (text, images, videos -revealing extents of flooding and impacts)</li> <li>• Inundation modelling outputs</li> </ul>
Geological	<ul style="list-style-type: none"> <li>• Earthquakes, subsidence, landslides -monitoring data and projections</li> <li>• Geological stability of urban areas</li> <li>• Geomorphological changes in coastal areas (derived from LIDAR, EO data analysis, Terrestrial Laser scanning)</li> </ul>
Weather	<ul style="list-style-type: none"> <li>• Archive climate data (used in claims assessment)</li> <li>• Records of CAT events</li> <li>• Predictions -short- and long-range projections of weather and climate patterns</li> </ul>
Satellite Earth Observation	<ul style="list-style-type: none"> <li>• Satellite feeds for claims (drawn on by loss adjusters to reveal extents of damage)</li> <li>• Derived products - change detection (revealing erosion subsidence, land use/land cover change)</li> <li>• Asset identification (drawing on automated processes or manual analysis)</li> <li>• Archive data (can form inputs to machine learning processes)</li> </ul>
Cadastral/location Data/Topographic Data	<ul style="list-style-type: none"> <li>• Natural capital monitoring -loss/gain/condition</li> <li>• Accurate and up to date digital maps</li> <li>• Geocoding data - Boundary datasets, area codes, wards</li> <li>• Building footprints and other relevant BIM (building information management) data</li> <li>• Terrain data (Digital Terrain Models (DTMs), Digital Surface Models (DSM))</li> <li>• Roads, rail, and other infrastructure</li> <li>• Identification of critical infrastructure (through looking at traffic data and human movements, supply chains)</li> </ul>
Corporate	<ul style="list-style-type: none"> <li>• Business activity</li> <li>• Audit data from clients (for companies seeking insurance)</li> <li>• History of companies <ul style="list-style-type: none"> <li>- Distribution of company assets</li> <li>- Value of business</li> <li>- Legal proceedings filed against company</li> </ul> </li> <li>• Commercial properties mix in area</li> <li>• Lines of business (for companies seeking insurance)</li> <li>• Supply Chains -revealing complex risks</li> </ul>
Insurance Specific	<ul style="list-style-type: none"> <li>• Flood-related claims</li> <li>• Exposure data -identification of assets in flood zones</li> <li>• Policy insight - premiums, cancellations and gaps in cover</li> <li>• CAT models</li> <li>• Modelling inputs from clients: <ul style="list-style-type: none"> <li>- descriptors</li> <li>- location</li> </ul> </li> </ul>

Table 1 (continued)

Category	Data themes
	<ul style="list-style-type: none"> <li>- type of asset</li> <li>- policy considered</li> <li>- 3rd party data</li> </ul>
Social/Economic	<ul style="list-style-type: none"> <li>• Human movements -footfall, use of coastal areas</li> <li>• Human Health -revealing health related impacts from flooding</li> <li>• Costs of rebuilding houses/structures</li> <li>• Integration of supply chains -vulnerability to disruption of business (past impacts, claims can be used as an indicator)</li> <li>• Demographics</li> <li>• Population distribution</li> <li>• Property/land values</li> <li>• Urbanisation -population concentration in urban areas, loss of natural habitat</li> <li>• Economic activity -identification of core industries, how flood events have impacted these.</li> <li>• Road use -traffic flow data</li> <li>• Recreation and tourism data</li> <li>• Spatialized indices of deprivation</li> <li>• Human behavioural data</li> <li>• ABM outputs -giving indications of cascading risks, adaptive behaviour, insurance policy take-up</li> </ul>
Risk/Hazard	<ul style="list-style-type: none"> <li>• Threat data -relating to flood hazards</li> <li>• Impacts and damage levels</li> <li>• Indicators of how buildings react to peril intensity               <ul style="list-style-type: none"> <li>- Vulnerability characteristics</li> <li>- Vulnerability classifiers</li> </ul> </li> <li>• Key infrastructure at risk: roads, rail, ports, water, energy, telecoms, undersea structures</li> </ul>

with specific scientific or technical background to interpret the data. Scrutiny of data sources is also important. When drawing on multiple data sources the work required to process these extends beyond shifting data into the correct format, to more comprehensive reviewing of the source and in some cases calibration and QA of the data used.

### 5.3. Internal data

Insurers hold valuable internal data which may relate to past impacts, claims, and performance. The main type of internal data referred to by parties interviewed, was claims data. This can include simple information such as loss coordinates, which prove valuable when assessing property underwriting rates for example. Claims data can further be used for predicting claims frequency and loss, however it is restricted, not always shared across a market and can vary in resolution (as reported by analysts questioned). Exposure data can also be classed as internal data, yet it is typically derived from clients. For property cover, exposure data can include location, building type, construction type, occupancy, and year built. Exposure data can prove problematic (as reported by insurers and analysts), in that it can differ widely depending on source and be hard to obtain. It can also prove difficult to determine if exposure data is accurate, as it can be vague, incomplete or not presented in a usable form. Aside from claims and exposure, another type of valuable internal data highlighted is that termed policy insights, this can detail factors such as premiums, cancellations and gaps in cover.

Suggestions have been advanced by data analysts and actuaries questioned that, in dealing with internal data, an initial step may be to focus on structuring and standardising data capture, data cleaning and archiving. Data can be fragmented, with claims data restricted to a few lines of text, frequently including slang and poor spelling. This prevents machines from being able to process it in an intelligent manner. As such, simple analytical strategies have been adopted based around typologies and manual approaches, such as counting key word frequencies. Another challenge is that industry records often take the form of narratives. This presents a particular hurdle when trying to analyse

systematically large numbers of records. In order to undertake statistical analysis, such qualitative information needs to be converted into quantitative data, to enable like-for-like comparisons. This can be summed up in the requirement for structured data capture. In some instances, advanced techniques can be applied such as Natural Language Processing (NLP), for location or argument extraction, for example (de Bruijn et al., 2018; Gritta et al., 2018; Roth et al., 2018; Wang et al., 2018). Changes in the methods of exposure data capture is also regarded (by a CAT modeller questioned) as one of many factors necessitating creation of associated standards.

Representatives from data analytics firms have reported problems in sourcing data from insurance companies, particularly around claims, where varying levels of information are made available. Given this, in some instances, loss adjusters have been used for supplying claims data, including companies such as Crawford and Cunningham Lyndsey. Supply of internal data has been highlighted by multiple practitioners as an issue restraining progress, especially given that it can embellish and validate analysis using external data sources.

### 5.4. Open source data

A variety of opinions on the use and value of Open Source data were encountered among those parties interviewed. Open Source data can be viewed in terms of the broad possibilities it presents for data reuse, that is data collected for one purpose, yet made freely available to use for another. Many interviewees regard Open Source data as beneficial to their work, this can be due to funding limitations, which necessitate the use of free data wherever possible. Others believe that not only data should be made available Open Source, but also methods, as the available data may not always be the specific type needed. One true evangelist (a senior-level market figure) stated that Open Data 'should be at the core of the data types available around the world, as it enables further innovation, with it being not just a public good, but a public requirement'.

In contrast to some of these positive sentiments, many challenges have been highlighted. When considering data obtained from a wide range of countries, there can be issues relating to data source and reliability. The level of data that is provided Open Source, and associated standards can vary significantly depending on levels of respective government funding. In some instances, the use of Open Source data is reported to actually involve higher costs internally, than drawing on well calibrated and regulated sources. This can be due to the data being incomplete, inconsistent and error-bound. In many instances, the user is said to have no concept of these issues until the data has been downloaded. Therefore, to obtain something of adequate quality, and completeness, much time may need to be devoted to searching. In many parts of Europe, freely available data is limited. Conversely, in the USA there are a plethora of sources of open data, but reservations have been expressed (by one data analyst) over the quality of this data and there being a lack of associated metadata provided. Insurers are said to require a comprehensive appreciation of what they are using, so in this sense (where metadata is lacking), many open sources are deemed unsuitable. Additionally, many insurers do not have the resources to address the inherent complexity of some of the open data outputs. Yet analytics firms have overcome this hurdle by outsourcing data processing tasks to lower income countries. To add to these challenges, non-public sector organisations are reported to be slow to open up their data or just fail to make any data available for free. This can be due to issues such as the need to recoup the costs of data collection.

Despite reservations on the use of Open Source data, there are a large number of Open Data sources now available, and this is increasing daily. For example, in the USA there are considerable data now available which can be used by insurers, such as that relating to wind, hail, fire, and crime. The UK is seen to be improving, especially in relation to datasets made available by the Environment Agency (EA), who have

for example a detailed mapping program in place involving the collection of airborne flood-plain and coastal LIDAR data, which is freely disseminated. This can be used to consider building footprints and floods. The EA dataset, the Risk of Flooding from Rivers and Seas (RFRS), has been named as a valuable input by flood modellers, as have data on flood defences, such as, 1 in 5 year event defences, recently released in 2017. Yet some modellers report this defence data to be problematic, in that it can be incomplete. Aside from the UK, a wide range of global Open Data portals exist, a selection of which can be viewed on OpenDataSoft's website (<https://opendatainception.io>), in which over 2600 sources are listed (Mercier, 2015). Also, an Open Source data search portal (<https://toolbox.google.com/datasetsearch>) has recently been released by Google (Castelvecchi, 2018).

##### 5.5. Application of advances in satellite Earth Observation (EO) data

In the interviews conducted with representatives of organisations working with EO data, emphasis was placed on gaining an understanding of how EO data can be applied to the insurance of flood risks. Over the last 10 years the cost of satellite technology has reduced significantly. As such, many opportunities are arising to draw on the data products created by a new generation of satellites (Bowler, 2018), and through web-mapping interfaces, and API feeds, it is possible to leverage vast volumes of EO data. EO missions now have the ability to generate repeat coverage of the globe daily. This is even being achieved using low cost miniaturised satellites, by companies such as Planet, using its Dove, medium resolution platforms. Terrain mapping derived from higher resolution EO data (such as through stereo imagery techniques (DeWitt et al., 2017)) is now approaching that obtained from airborne LIDAR missions, with the added benefit over LIDAR of regular repeat imagery, for vast spatial extents. In the past EO data exhibited problems such as imprecision, yet now this has been largely overcome and there are a wide range of options that can be drawn on, such as multispectral imagery (visible, infrared, and thermal), radar, and microwave. Infrared EO data was reported useful by one flood modeller questioned as it deals with problems related to cloud cover. Near-infrared is well suited to delineating water bodies (Adam et al., 2014). Microwave and SAR (such as NovaSAR-S) also bypass cloud cover issues (Lavender et al., 2016), and have been applied successfully to flood mapping. Furthermore, microwave data is being used for near-real-time flood detection and mapping in the Global Flood Detection System (GFDS) (Jongman et al., 2015).

There are many low-cost providers of this data, with the European Space Agency's (ESA) Sentinel 1 providing free and consistent SAR data for the whole of Europe. To keep costs down one company reported that their insurance related baseline product is primarily based on free-to-access Copernicus Sentinel-1 data. One example of how this data is being applied to insurance is provided in the work completed by Hénaff et al. (Hénaff et al., 2018) who combined Copernicus elevation data with historical claims data to make predictions on global insured values in flood risk areas. For flood risk analysis, there are numerous opportunities presented by Open Access remotely sensed data (from a wide range of sources), such as relating to altimetry, Digital Elevation Models (DEM), optical, and radar images, (as demonstrated by Ekeu-Wei and Blackburn (Ekeu-wei and Blackburn, 2018)). The UK Space Agency is championing many innovative projects involving application of EO data to flood risk analysis (UK Space Agency, 2017a). One company questioned, Pixalytics, has been working with the UK Space Agency, in developing Virtual Water Gauge software which uses satellite altimetry to determine water heights in estuaries, rivers and lakes, and has been used for analysis and detection of flood events (UK Space Agency, 2017b). Besides drawing on free to use EO data, many companies are also drawing on Open Source software where possible (Albano et al., 2017; Joseph and Kakade, 2014) (so lowering costs further). A prominent Open Source software drawn on for modelling flood inundation is LISFLOOD-FP (The University of Bristol, n.d.), this has frequently

been combined with other free to use software in completion of coastal flood risk assessments using EO data (De Angeli et al., 2018).

Research and development in this field is continually generating innovations increasing options for application of the technology. Easily accessible user interfaces, such as Google Earth Engine, can enable a vast archive of EO data from different sources to be mixed as required (without cost). Firms also have been supplying (Software as a Service) SaaS platforms enabling clients to test the effects of new and existing insurance policies. One field advancing rapidly is interferometry. The technology can be used, to assess the risk of subsidence and can monitor millimetre changes in land height (Ramieri et al., 2011). Other emerging areas such as the use of Stereo imagery techniques, are being applied to flood risk assessment (Mashaly and Ghoneim, 2018), these techniques are being used by DigitalGlobe and Terrabotics, to generate 3D images from satellite data, and can be used to look at steep slopes and build terrain datasets with sub-metre accuracy. If insurers have high-risk areas which need monitoring, they can also commission satellite missions in advance, to capture detailed, high-resolution data. Furthermore, a recent innovation implemented by the company Earth-i allows colour video to be captured from space (Werner, 2018). This may prove useful to loss adjusters, for example, in analysing disruptions generated by flood events.

Automation in the processing of EO data is resulting in huge reductions in the time spent working with the raw data. Automatic change detection is possible, for assessment of flood risk and extents (Geller, 2017), and Artificial Intelligence (AI) techniques are being used to identify flooded houses, blocked roads and bridges, and estimate depth of flood waters (David Grason, 2018). Examples provided by DigitalGlobe reveal how precise damage to properties can be quantified. And in relation to their WorldView satellites, automated processes have been deployed for tasking satellites to acquire images of areas impacted by flooding, based on interpretation of social media data (Cervone et al., 2016). This is important as flood impact mapping needs to be reactive, which makes manually tasking satellites in advance difficult. Companies are moving away from manual processing of EO data, through automation possible in the cloud (Tsarouchi, 2018). Geospatial service frameworks have been developed which allow parallel processing, expandable on multiple instances within the cloud; this was reported by one firm to 'cut processing times by an order of magnitude'. One EO data analytics firm report to have developed a cloud-based parallel processing platform, drawing on a wide range of sources, including those of Airbus and DigitalGlobe. This platform is reported to cut EO data processing times significantly. Cloud-based platforms can also facilitate implementation of both traditional algorithms and Deep Learning techniques (Chen et al., 2018; DigitalGlobe, 2017).

For the insurance industry satellites can be used in disaster response, exposure management and for underwriting solutions. Satellites have been used as a remote validation tool, to contribute to audit trails and to take the place of site visits. One benefit of this is that EO data is impartial and unbiased, and just reveals what is on the ground. Furthermore, it is being combined with on the ground intelligence and Internet of things (IoT) monitoring outputs. One company, Sensonomic, is drawing together such data sources within ABMs to reveal behavioural interactions between individuals and organisations, which can generate answers as to what drives risk exposure. Such modelling processes also benefit from fusing EO data with insurers' internal data such as claims, to refine, calibrate and validate outcomes.

Within Lloyd's, claims teams have been drawing on EO data following a spate of recent hurricanes in the Caribbean and the USA, to assess damage (Lloyd's, 2018a). Analysis completed, post Hurricane Matthew (which first made landfall over Haiti on 4th October 2016), highlighted damaged properties in the wake of its path. Satellite imagery was also utilised to monitor the devastation reaped by hurricanes Harvey and Irma (Lloyd's Market Association, 2017) (which made landfall on 18th August 2018 and 5th September 2018, respectively). Similarly, firms supplied loss adjusters with imagery of the situation on the ground

between hurricanes Irma and Maria, so they could check damage, and place losses appropriately (Lloyd's, 2018a). Obtaining such detailed satellite imagery, representing specific flooded locations, has proven possible, yet can be costly, due to requirements for tasking satellites to focus on desired locations. One firm (McKenzie Intelligence Service) is reported to have combined satellite imagery with CCTV footage, so street level damage could be viewed. Feedback received from the market indicated that this was useful. Combining EO data with other ground-based sources, is important and can be essential for ensuring its validity and usability. Sources such as CCTV footage, river level gauges, and images uploaded to social media platforms, have all been used to compliment and validate EO data. Many insurers have expressed interest in using EO data in the future, and satellite data has frequently been drawn on as court specific evidence given its high validity. There can be issues in application of the data though; one EO data analyst reported encountering licencing problems when providing EO data for insurance use. Yet, despite such issues, both image and radar data (in particular) are shown to provide insurers with a wide range of possibilities, involving using solely the raw data or thematic data derived through the application of automatic classification. However, to-date, the insurance industry has proved slow in adopting EO data and derived products.

### 5.6. Big Data opportunities

A recent report by the American Academy of Actuaries stated that 'The combination of powerful computers and "Big Data" has transformed understanding of hazards such as flood' (The American Academy of Actuaries, 2017). 'Increased computing power and availability of higher detailed harmonised datasets' has also been acknowledged as enabling detailed flood analysis at various spatial scales ((de Moel et al., 2015), p.882). One CAT modeller interviewed, concurred this, and stated that Big Data has improved their modelling work, particularly for floods, and as a result higher levels of precision are possible. The term Big Data does not only apply to large data Volumes but also large Varieties and Velocities of data (termed the 3Vs of Big Data) (Jagadish, 2015). This involves the ability to store, process and analyse structured and unstructured data, combining archive and real-time streaming data (Jagadish, 2015; Marr, 2015). The data is generated from a wide range of sources and is increasing in availability (much of it being Open Source). For example, these can take the form of more conventional data such as database entries, stored in Relational Database Management Systems (RDBMS) or real-time streaming data being generated by IoT sensor networks, satellites (Maier et al., 2012; Moszynski et al., 2015; SmartBay, 2017), and websites (Singhal et al., 2018). Big Data is reported as being widely applied within risk analyses (Choi and Lambert, 2017). In attempting to understand how the various fields associated with 'Big Data' can relate to an industry such as insurance, a framework is provided below in Table 2. This framework provides a chronological listing of stages associated with aspects of Big Data. These phases align with the data utilisation stages outlined in Fig. 1.

To enable external and internal data related to insurance to be analysed and knowledge extraction to take place (Stage 3) it is necessary for data to be stored and processed in an effective way (Stage 2). Technology firms provide infrastructure and software tools to enable this, many such as Hortonworks and Cloudera base their solutions primarily on software developed by the Apache Software Foundation (The Apache Software Foundation, 2018). This software is Open Source and is the product of the interactions of over 30,000 contributors who commit code to Apache projects. The software tools and technologies include Hadoop, MapReduce, Apache Spark, Nifi, HBase, Hive, MongoDB and many others. The software forms an 'ecosystem' (Marz and Warren, 2015) in which different functions are performed by individual software elements, relating to distributed storage and processing, data mining, analysis and ultimately query and knowledge extraction. The analytics

firm LexisNexis provide an Open Source alternative to some of the Apache software, in their HPCC Systems (LexisNexis, 2018).

In relation to knowledge extraction, a wide range of analytical tools are drawing on 'Big Data' in attempting to better understand risk. Techniques such as machine learning are increasingly being looked to (Peters, 2017). This is an area in which vast stores of data, now available, such as that generated by satellites, can be combined effectively with insurer's internal data. Geocalibrated claims data, for example, have been drawn on to verify and calibrate machine learning algorithms developed to make flood predictions, using EO data (Hénaff et al., 2018). There are also examples revealing how ANNs could be adopted for spotting patterns, and understanding relationships between data variables, such as those related to environmental hazards (Bezuglov et al., 2016; Chang et al., 2018; Joseph and Kakade, 2014). It is becoming increasingly possible to draw on alternative data sources in analysis of flood events. This can take the form of mining social media data, such as Tweets, using geoparsing to extract location data (de Bruijn et al., 2018). Making sense of large quantities of unstructured data is a huge challenge and techniques such as NLP, geocoding and Computer Vision, have been employed to extract flood-related data from social media (Twitter) and crowdsourced data (from Mycoast (<https://mycoast.org>)) (Wang et al., 2018). This field is in its infancy though, and the study by Wang using Computer Vision for urban flood modelling is reported to be the first of its kind (Wang, 2018). Other examples exist revealing how microwave EO data, has been combined with social media data to map flood impacts (Jongman et al., 2015). In fact, Twitter data is emerging as a useful source to combine with EO data, and other data inputs, to reveal extents of flooding in near real time (Li et al., 2018; Panteras and Cervone, 2018).

The emergence of modelling processes focusing on human behaviour was introduced earlier, this is an area which social media data is also forming a valuable input. Du et al. (Du et al., 2017) demonstrate this in their model of individual flood evacuation behaviour, in which they also focus on transport networks. Outputs of such analysis could prove useful for revealing flood-related infrastructure stresses and disruptions. For example, this can relate to a single flood event, generating a multitude of secondary impacts, such as disruptions to business, supply chains, and utilities failures. Examples, such as that provided by Papadopoulos et al. (Papadopoulos et al., 2017) demonstrate how large quantities of unstructured social media data, can be drawn on effectively to improve resilience of supply chains and critical infrastructure. Having access to large stores of data, covering a wide range of themes could prove instrumental in understanding the factors involved in systemic risk scenarios, such as those provided within the simulated catastrophe stress tests performed by Lloyd's (the Realistic Disaster Scenarios (Lloyd's, 2018b)).

Internal data such as claims information, detailing past losses, play a vital role in the validation and calibration of new analytical techniques and models (Christie et al., 2018; OECD, 2016), and as such a lack of data relating to past insured losses can prove a factor limiting their development. Claims data has been drawn on successfully to validate flood and hydrological models such as the 2D BASEMENT simulation, by Zischg et al. (Zischg et al., 2018). EO data now available, revealing impacts, are also being drawn on to validate flood extents, and in predictive modelling of flooding (Ekeu-wei and Blackburn, 2018; OECD, 2016). This data is particularly useful for more remote and developing parts of the world where traditional datasets are lacking (Ekeu-wei and Blackburn, 2018). Reports indicate that data mining methods have been utilised to obtain the required EO data, this can be essential given the data volumes involved (Lavender et al., 2016).

In addition to the 'Big Data' sources already mentioned, the Internet of things (IoT) is a rapidly emerging field which holds promise, to create 'Smart Insurance', in which policies can be based on detailed historical datasets generated by networks of automatic sensors embedded in homes, businesses, machinery and infrastructure. The sensors or 'things' are uniquely identifiable and connected to the Internet, with 'sensing/



**Table 2**  
Big Data Framework.

Stage	Processes	Considerations
1. Data source awareness	Data collection	<ul style="list-style-type: none"> <li>• Inclusion of holistic data sources</li> <li>• Availability of Open Source data</li> <li>• Internal datasets -claims data (unstructured/semi-structured)</li> <li>• Archive/real-time streaming data</li> <li>• Utilisation of emerging data sources               <ul style="list-style-type: none"> <li>- IoT</li> <li>- Social media</li> <li>- Satellite EO</li> <li>- Mobile telematics</li> <li>- Free text (emails, web logs, transcriptions, notes)</li> </ul> </li> </ul>
2. Accessing the data	Data Ingestions and Storage	<ul style="list-style-type: none"> <li>• Data source veracity</li> <li>• Database choice: RDBMS or distributed database (SQL or NoSQL)</li> <li>• Storage solutions: on premise/cloud/hybrid cloud</li> <li>• Requirements for permanent/on demand (elastic) processing capacity</li> <li>• Cloud vendor selection</li> <li>• Data Warehouse/Data Lake</li> <li>• Data security</li> </ul>
	Selection of software infrastructure; data processing requirements	<ul style="list-style-type: none"> <li>• Database software selection based on data types (structured/unstructured/semi-structured)</li> <li>• Parallel processing options and requirements (availability of compute power)</li> <li>• Open Source/proprietary software?</li> <li>• Automated processes for data ingestion and collation</li> <li>• Streaming data processing</li> <li>• Processing and analysis requirements for different data formats: free text, graph, audio, point cloud, imagery, video</li> <li>• Geocoding: by address, postal delivery code, boundary, geotagged data, geoparsing</li> </ul>
3. Using the data to generate information	Analytics and knowledge extraction	<ul style="list-style-type: none"> <li>• Possibilities for advanced geospatial analytics</li> <li>• Options for drawing on graph, text and time series analytics</li> <li>• Ability and requirements to run distributed batch processing tasks for compute intensive workloads (e.g. for actuarial calculations)</li> <li>• Artificial Intelligence and Machine learning's role – discovering patterns (claims), feature detection, classification of land use/land cover, change detection (buildings/infrastructure)</li> <li>• Vast quantities of EO data stored in the cloud, used for training machine learning algorithms for flood and impact detection</li> <li>• Cloud based parallel processing facilitating development of Deep Learning techniques [78]</li> <li>• Computer Vision applied to video/image analysis -to detect flood extents and damage post event</li> <li>• The ability to derive meaning from unstructured messy data through NLP and other techniques</li> <li>• The ability to combine real-time streaming data with archive data</li> <li>• Deployment of Artificial Neural Networks (ANN) for real-time flood inundation modelling [87]</li> <li>• Development of automated workflows for targeted collection, processing and analysis of data (i.e. satellites tasked to collect data for flood sites based on analysis of social media data [75])</li> <li>• Application of text analytics (e.g. NLP) to claims data</li> <li>• Application of predictive modelling functionality, for example involving: Linear Regression, Logistic Regression, Decision Trees, and Random Forests.</li> <li>• Developmental opportunities available using programming notebook interfaces such as Apache Zeppelin (<a href="https://zeppelin.apache.org">https://zeppelin.apache.org</a>) and Jupyter (<a href="https://jupyter.org/index.html">https://jupyter.org/index.html</a>)</li> </ul>
4. Knowledge communication/ dissemination	User interfaces and data visualisation	<ul style="list-style-type: none"> <li>• Means to validate analytical outputs</li> <li>• Web based user interfaces</li> <li>• Graphical User Interfaces (GUI)</li> <li>• Live data feeds incorporated into interfaces</li> <li>• Outputs of on-the-fly analysis available to users (e.g. for analysis of impacts and claims data)</li> <li>• Advanced intuitive dashboards</li> <li>• Advances in 3D visualisations of geospatial data</li> <li>• Virtual/Augmented Reality</li> <li>• SaaS</li> </ul>

actuation and potential programmability capabilities', and data generated by these 'things' can theoretically be collected 'anywhere, anytime by anything' (Hassan et al., 2018). For example, the real-time data made available through IoT devices connected to cloud services, can be used to give updates on the severity of disaster events in real time. Cases, such as that provided by Koduru et al. (Koduru et al., 2018) reveal how such IoT networks could be applied to insurance of flooding and other disasters. Feedback received during interviews with insurance practitioners highlighted how insurers are currently engaging in Proof of Concepts (PoCs) with analytics firms, which involve use of IoT data and sensor deployment. These PoCs concern multiple lines of business, not just flooding, and there appears an appetite to fund future use and deployment of IoT sensors, if insight generated through their use proves effective. For example, if data feeds obtained from these sensors, proved reliable enough to be used in policy pricing or loss assessments, this could justify their utilisation.

In respect to flooding, one area of IoT application is monitoring of storm surges and water levels. In the USA this has been demonstrated through 'StormSense', which has been deployed as part of a smart cities initiative, for real-time monitoring of flood events, and has provided data inputs to subsequent inundation modelling (Loftis et al., 2018). A benefit of IoT is that the sensors can prove cheap and reliable and their outputs can be effectively combined with, social media, crowd sourced, and remote sensing data, for evaluating flood risk in densely populated locations such as 'mega-cities' (Ogie et al., 2018). However, such diverse, and dense data streams are associated with a range of uncertainties and can contain spurious and incomplete data. Given this robust methods are required to fill the gaps and to interpolate and infer values where data is missing or unreliable (Koivumäki et al., 2010). Monrat et al. (Monrat et al., 2011) set out one way of dealing with such uncertainties, for data relating to flooding, using a Belief Rule Based Expert System (BRBES) with Apache Spark, generating real time

flood maps. Their Big Data platform enables replicated storage of vast quantities of data in separate nodes to ensure data integrity and fault tolerance. This allows the data to be analysed by the BRBES, which tackles inherent uncertainties. Examples, such as those detailed above, indicate how advances in collection and analysis of Big Data can potentially be drawn on to enhance flood risk analyses processes.

### 5.7. The use of data analytics for insurance

Of the companies involved in modelling and insurance specific data analytics interviewed, their areas of focus were: CAT risk modelling, general insurance analytics, geospatial threat analysis, flood risk modelling, and property analytics. Feedback supplied related to sources of data draw on, methods and technologies used, and data innovations being implemented.

Analytics firms interviewed acknowledge the requirement to draw upon and fuse data from many different sources and typically state they are data agnostic. A necessary consideration is that data from disparate sources come with varying standards. In dealing with this issue, many such as one geospatial analytics firm questioned undertake extensive data cleaning. Furthermore, a CAT modeller highlighted how new standards for data capture can be required to enable so many sources to be combined: 'due to changes in exposure data capture, having standards becomes necessary'. Given the wide range of data being drawn upon, data aggregation becomes increasingly important. This was reported a goal of multiple insurance analytics firms, who build databases from insurers' internal data in addition to leveraging Open Source data. There are many further issues which require consideration, such as compatibility of the data being used; it is also necessary to focus on data granularity. One CAT modeller, who engages in flood modelling, specified how they require data at an individual property level for their analysis. For example, adaptations implemented at both a regional and household level need to be accounted for in insurer's flood risk calculations (Garvin et al., 2016; Osberghaus, 2017; Thistlethwaite et al., 2018; Yang et al., 2018). This can necessitate incorporation of more granular data, enabling variations in risk exposure, over smaller distances, to be realised (Schwartz, 2018) (as illustrated by an example taken from the Netherlands (Jongman et al., 2014)). Scale was noted as a significant issue in relation to insurance risk assessments. Risk has been aggregated at the level of postcodes in England (Dávila et al., 2014), yet now flood analytics firms are producing assessments at a household level (Garvin et al., 2016). This is benefitted by advances in satellite radar and LIDAR data collection techniques, resulting in detailed terrain data now being made available by commercial suppliers, with quoted resolutions of up to 1 m, globally (Intermap, 2018).

Many geospatial analytics firms report to be actively engaged in seeking out new sources of publicly available data including satellite imagery, LIDAR data, and private drone footage. Yet satellite EO data has not been widely drawn upon to-date, by the insurance specific modelling firms interviewed. Feedback indicated that this can be due to scepticism on its application and reliability. Insurers questioned, highlighted a requirement for line of business and peril specific use cases, demonstrating proven suitability of the technology. However, one property analytics firm stated, that in their analysis they draw on multiple kinds of EO data, such as multispectral and satellite radar imagery, to allow assessment of the impacts of flood and fire events. For the UK, publicly available data such as that made available by British Geological Survey (BGS), and the EA, are drawn on by many modelling firms. Several of these sources provide real-time data feeds (e.g. web services), these are increasingly being incorporated in models, such as those provided by one insurance analytics firm questioned. Many of the general insurance analytics firms are developing solutions for bringing together both insurer's internal data and external sources, and in doing this, harmonising data standards.

There are many similarities between the types of analytical methods employed by firms. Their methods are seen to have common goals such

as enabling underwriters to screen and price risk, through provision of common loss metrics. From the responses received, location data stands out as being especially important. Many firms specialise in dealing with location data such as one property analytics firm, who host location, building, environmental and financial data. Location-based analysis also provides a prime focus of a geospatial analytics firm, who provide exposure management for underwriters. Determining the accurate geographic locations of risk is regarded as crucial by insurers, this hinges on the ability to geocode correctly, especially for accumulation calculations, for which an accurately geocoded source of data, such that relating to buildings, is deemed essential (Garvin et al., 2016). The process of geocoding aids geospatial analysis of risk and is particularly useful for flood risk analytics (as reported by a flood risk analyst). Understanding the geographic distribution of risk was reported a particularly important aspect of analysis carried out by CAT modellers, with one CAT modeller highlighting the importance of geospatial risk analysis in enabling individual locations to be focused on in calculation of risk premiums. As such, Geographical Information Systems (GIS) are a software tool widely used by many insurance analytics firms and are utilised to generate property risk profiles based on geospatial attributes.

Probabilistic modelling techniques drawing on statistical and mathematical analysis, form a central component of the methods used by most firms. These have commonly been coupled with depth damage curves (André et al., 2013; Dávila et al., 2014; de Moel et al., 2015; Hsu et al., 2011; Penning-Rowsell et al., 2013), in making predictions of financial impacts of flooding. Companies also increasingly draw on emerging techniques, for example a CAT modeller reported using Computer Vision to detect damage to an area post event, whilst one general insurance analytics firm reported making attempts to improve the underwriting process by applying machine learning to claims prediction, renewals and accumulation reporting. Open Source tools are being made available by a number of firms. This includes a CAT modelling software platform (provided by a CAT modeller questioned), which draws on probabilistic methods such as Monto Carlo simulations, reporting hazard intensities, exposure and probabilities of loss at specific locations. More general Open Source software was also reported as being used by many firms, such as MongoDB, a NoSQL document-oriented database. Some firms have developed extensive in-house software capabilities. For example, one insurance analytics firm has created their own Open Source Big Data analytics platform. They report drawing primarily on their own technology, using many 'scalable automated linking technologies', which can be statistical based, incorporating probabilistic functions.

An increasing number of firms are adopting the cloud to host and deliver their solutions, and multiple analytics firms have reported migrating data currently stored in local servers to cloud environments. Many of the GIS solutions one property analytics firm operate draw on the automation and scalability possible in a cloud environment, to enable high-resolution geospatial data to be accessed in real-time via web mapping interfaces. One CAT modeller stated that all their 'future development will be completed in a cloud environment', as 'clients want to be able to access big data at scale from across the enterprise, and the cloud allows this'. Additionally, machine learning is deemed much more suited to the cloud due to the possibility for on-demand scaling of compute power. The cloud has not been adopted by all though and many such as a flood risk analytics firm, use their internal data centres to host modelling data, whilst another Insurance analytics firm reported using a conventional data warehouse. In relation to distribution and visualisation of analytical outputs, companies commonly provide their outputs as web-feeds or in the form of GUIs. Numerous analytics firms provide insurers with API feeds so that analytical outputs can be incorporated within existing dashboards. SaaS provision was also reported to be increasing in popularity as a delivery mechanism for risk analytics. This can allow firms to run their analytics solutions in a web browser. SaaS options can also allow providers to implement updates remotely, and to bypass compatibility requirements for integrating their solutions with an insurer's internal IT systems.

### 5.8. Communication of information outputs: underwriting

There are a range of core roles within the insurance industry who are heavily reliant on data in analysis of risk. Among these actuaries and underwriters stand out as the most prominent. Within the feedback received during interviews, the role of the underwriter was focused on, as such, the following discussion covers some data specific aspects related to underwriting. The process of underwriting risk is a fundamental function within the insurance industry and involves risk selection and fast decision making. However, the process is not always transparent, with underwriters regarded as utilising their own internal intelligence and idea of cover price. Some have even claimed that ‘underwriters innately know risky places’ (as reported by a senior insurance practitioner). Insurance cover results from interactions between underwriters and brokers. Moreover, both broker and underwriter need to have a firm grasp on how technology can be drawn on to generate answers.

The underwriter decides on the cover a client is happy doing business with. To enable them to do this they require access to tools, such as an electronic dashboard which can generate answers based on entering simple identifiers. Information supplied to the underwriter can be taken into consideration in pricing models. For example, CAT models enable underwriters to distinguish what and where to insure, geographical spreads, transfer of risk, and financial strength. Their challenge is added to by the practicalities of the underwriting process, resulting in individual underwriters not always being aware of the wider risk picture, such as that associated with cascading and systemic risks.

A challenge for underwriters and those providing them with information (which has been continually repeated by those interviewed for this report) is the lack of time underwriters have to make important decisions and to review information. One particular challenge is how data is served at the point of decision making, given that underwriters may have only minutes to price risk and decide on cover. Such quick decisions do not allow time for underwriters to review the data in great detail. Underwriters consulted report the need to set up in excess of 500 deals per year. Given this, they are unable to devote time to navigating complex user interfaces to retrieve information. Therefore simple, intuitive dashboards are required, presenting a clear view of loss. This requirement has prevented organisations, such as one analytics firm questioned, from implementing GIS tools for underwriting, which they deem are better suited to be used by modelling teams. GIS applications have proved overly complex to be utilised for underwriting and can place constraints on teams. Abstracting this complexity is deemed a requirement, as underwriters require distilled metrics at their disposal.

One underwriter noted ‘A few pieces of choice information can change an underwriters mind’. As such, more general data is required at an actuarial level, than is needed by underwriters who require more granular specific information on facilities and clients. A core requirement is for the potentially huge amounts of data available, to be turned into something useful. Provenance of data sources is also important. Underwriters questioned have stated how they draw on information obtained from internet search engines and geospatial information obtained from Google Earth. Mapping platforms such as Google Earth, are being used by underwriters are to gain an understanding of properties, building materials, roof types, among other features. Yet, information in web-mapping applications aimed at the general public, can be out of date or poorly presented. Google Earth is undoubtedly a useful resource, but images can be many years out of date.

### 5.9. Challenges

Many challenges to the effective utilisation of data have been established from feedback provided from those working with and using data types such as those listed in Table 1. A number of challenges have been detailed in the previous sections, some of the more prominent of these are expanded on here.

Core inputs to CAT models have been reported as difficult to obtain, especially those with the appropriate level of detail and in a usable format. Such inputs include information on the built environment (for certain countries) and calibrated loss data. Many UK insurers and analysts are said to struggle with local authorities not providing them with the information they need. Builders have also been highlighted as not wanting to share information, with those such as flood modelling companies. Yet information relating to new housing developments (for example) is important, especially when used in response to CAT events, where environmental data needs to be merged with information about buildings, and other factors, to produce loss estimates.

Many challenges have been reported when trying to obtain datasets for a wider range of countries. Whilst the UK, and parts of the USA, have 5 m resolution flood data, from which depth of water can be estimated, attempting to source data for Africa and Eastern European states was reported, by multiple flood modellers questioned, to be difficult. Insurance cover is increasingly being provided in geographical areas where policies were not written previously. These new markets can pose fresh challenges, especially in relation to data standards and availability. As such, it has been reported as difficult to obtain the required datasets for modelling risk in some lower income countries. Nevertheless, a range of opportunities are presented by emerging sources such as EO data, to obtain global datasets (Ekeu-wei and Blackburn, 2018), many of which are available Open Source.

States, such as the USA, who provide an extensive variety of Open Source and proprietary datasets, may fall down in certain areas such as provision of geological data, where experts in geology who compile the datasets may not have adequately considered how clients want to use the data. This is a common problem reported by analysts, for scientific datasets in many countries, where some government sources are said to release maps that are not usable, due to problems with complexity. Another factor reported as presenting a barrier to utilising international datasets is language, this can necessitate diverse translation requirements and additional time and resources being devoted to processing data inputs. A common problem encountered, when obtaining data for different regions, is with data existing in various formats and levels of completeness. Data needs to be transformed to regular formats, which can be a time consuming and burdensome process, although many tools are available to facilitate this. Analysts have stated that if these tasks could be pooled by a central body, it would result in time and cost savings, avoiding duplication of efforts. The London Market Target Operating Model (LM TOM) (<https://tomsupports.london>) is one example of such and industry wide initiative. This relates to data capture and access, involving creation of a central data repository. This was greeted with enthusiasm from those spoken with from across the market. Private initiatives have also emerged which are seeking to address these challenges, such as Oasis Hub (<https://oasishub.co>).

One specific problem highlighted is that many who make decisions based on data can be unaware of the limitations of the data they are using. As a result, too much confidence can be placed on the data, resulting in skewed scenario creation. Scepticism was voiced by many well-established insurance practitioners, about the use of data and reliance on models ‘bought, but not understood’. In line with this, many have stated that there are ingrained attitudes held within the industry that may act as a barrier to changes being implemented. This has been cited as a factor contributing to slower take up of technological developments in the insurance industry compared with other areas of financial services, such as investment. Furthermore, one supplier of technology stated that they can use ‘most of their time educating insurance syndicates, and more than actually supplying products’.

From a data-driven perspective, insurance is seen to be behind the times (Miller, 2018), in its reliance on generalised linear models, and expert opinion, such as that of warranty surveyors. This can be especially so for risk engineers, whose main tools are qualitative, with expert judgements, and surveys with clients, determining if engineers should be sent to a site. For many it may appear simpler and more reliable to

resort to drawing on expert opinion instead of using unknown data analytics methods. Unfortunately, expert opinion has proven an inadequate method for capturing the dynamic nature of many risks. Drawing on larger empirical datasets in evaluations, can allow fairer pricing of risk and data signals can act to allow filtering of portfolios. Yet, a lack of knowledge sharing, across the industry, is said to pose a barrier to this.

Adoption of the most appropriate technologies by insurers has also been flagged as an issue (Libarikian et al., 2017). Data analytics firms report that many clients in the insurance industry are currently using outdated IT, and that they (the analytics firms) are not in a position to enforce change. Many firms admit to being in their infancy in the use of advanced data techniques, especially in relation to their own data (Heale, 2014). Yet consideration of advanced methods, such as those detailed in Section 5.6, can prove essential in understanding how complex hazards translate to loss, generating the resulting financial impacts. In line with adoption of such advanced forms of analytics, many believe that in the future data scientists and actuaries should work more closely. For example, benefits could be gained through data scientists drawing on actuarial understanding of the data, and mixing this with modern techniques, ensuring risk selection is closely aligned with risk profiles.

Innovations, such as the use of flood maps, have been reported as altering the fortunes of those covering this peril. Furthermore, underinvestment in flood modelling has been reported, by a senior figure in a flood modelling firm, as contributing to some of the highest profile losses sustained by insurers, over the last twenty years. As such the OECD have cited the lack of high-quality flood maps, in some countries, as an impediment to effective financial management of flood risk (OECD, 2016). Yet, the adoption of geospatial data analytics has altered risk ratings for many areas. High resolution geospatial data sources are now capable of supplying insurers with datasets at an individual building scale. This can enable differentiation of risk premiums at an individual policy level. Effective use of geospatial data can potentially allow more stable areas to be identified within high-risk zones, which can allow companies to be more aggressive in pricing policies covering these areas. The opposite can also be the case with higher risk areas being identified (Collinson, 2017) and potentially avoided.

## 6. Discussion

The previous sections have presented a critique of how data is currently being used within risk analytics covering environmental perils (primarily flooding). This was primarily based on feedback received from those working in and associated with the London Insurance market. Within the London market (and Lloyd's) there is a shift to digitalisation and adoption of modern practices (Carnegie-Brown, 2017; Tischhauser, 2017). Initiatives to maximise the potential offered by data, such as the ongoing LM TOM are evidence of this. Adoption of new methods and techniques have been witnessed at various levels within organisations, data entry teams are building and adopting new tools, and it has been acknowledged that data capture and processing tasks need standardising. Many insurers are actively engaging with analytics firms seeking to apply new technologies to insurance use cases. Furthermore, the development of 'Insurtech' is evidence of the widespread impact which digitalisation is having on most aspects of the insurance industry (Stoeckli et al., 2018).

In terms of deriving value from data, four key aspects were highlighted by the authors (detailed in Fig. 1): 1. knowing data is there; 2. having access to it; 3. making sense of it, and 4. using it. In the following discussion themes relating to these areas are covered in detail.

### 6.1. Data (knowing it is there and having access to it)

In evaluating insurance-related risk, the value derived from internal industry data can be maximised when it is combined with external

feeds (Deshpande, 2018; Zischg et al., 2018). Furthermore, data is becoming available that can remove ambiguity in the pricing of risk, this can relate to new methods of data capture, such as that from IoT devices, satellite-based sensors, the internet (e.g. social media), and initiatives resulting in data sharing. However, an understanding of data veracity (i.e. data quality, source and validity) is essential before a decision is made to use the data. Also, the complexity of some data sources can necessitate specialist interpretation before they can be used. The increasing availability of Open Source data presents an opportunity to enrich analyses of risk. This open data can spur innovation, acting as a raw material to enable development of new forms of analytics. As such, many firms report to be actively engaged in seeking out new sources of publicly available data. Yet there can be limits imposed on the use of Open Source data, including: lack of data for some regions and countries, poor data quality, and a lack of accompanying metadata. Specific requirements which have been repeatedly highlighted are for standardisation and structured data capture. Furthermore, in some cases application of advanced methods such as NLP could enable narratives and qualitative data to be systematically analysed.

### 6.2. Analysis (making sense of the data)

Insurance data analytics should involve fusing data from many sources, providing a holistic view of risk. Techniques becoming available, can enable datasets collected for one purpose to be reused and combined and offer potential for higher-level insights to be derived. Examples have been provided illustrating how EO, IoT, and social media data have been utilised in such a way. Location data has been highlighted as important and can reveal the geographic distribution of risk. In line with this, GIS is regarded a suitable software tool and is widely used by many insurance analytics firms. For flood risk analysis it is especially important to consider the granularity of data drawn on, and if this is adequate to reveal household level risk, and to account for localised or individual adaptation measures. Furthermore, flood risk maps used by insurers need to include data on adaptations (updated regularly) (Beck et al., 2018; de Moel et al., 2015). In relation to property level adaptations, this can necessitate development and consideration of associated standards (Bonfield, 2016). Data related to human behaviour also represents an important factor, which should be included within analysis, and can alter predicted risk ratings for areas. ABMs have been highlighted as one tool which can be applied to this area.

With the volume and variety of available data sources rapidly expanding, an overview of the potential storage options, software infrastructure, and processing techniques, is required so that data can be handled and retrieved in an efficient manner. This work generated some limited findings related to the use and suitability of Big Data and cloud technologies. Open Source modelling software, such as that provided by the Apache Software Foundation, is both provided and being drawn on increasingly. There is also a rise in Open Source flood modelling software being developed, an example of such is *FloodRisk* (Albano et al., 2017). A growing number of firms are now looking to the cloud to host and deliver their solutions, due to on demand compute power, automation, real-time data access, and options to undertake data mining and machine learning (some of which utilises vast global archives of satellite data). Automation possible in cloud environments can also reduce requirements for manual intervention, potentially lowering costs. Yet many in the industry are still wary of shifting data to the cloud, for example, due to security concerns.

### 6.3. Communication (using the outputs)

In considering how data is consumed, the work has focused particularly on the requirements of underwriters. Yet the wider issues raised also consider the needs and requirements of other key actors, such as actuaries, brokers and loss adjusters. Irrespective of user's role, a

thorough appreciation of user requirements and level of domain expertise, is required, so value derived through the previous analytical steps is not squandered. A requirement has been identified for those with knowledge of how data is consumed within the industry, to act as an interface between insurers and more specialist data analysts. Also, it is important to communicate, to those making decisions based on data, the range of available data sources and their limitations. There is a heavy focus, by many analytical firms on methods of delivering their outputs; outputs are being provided as web-feeds or in the form of GUIs. Additionally, SaaS is increasing in popularity.

#### 6.4. Challenges

This research highlights how progress is being made in adoption of new data sources and methods within the London market, however there are numerous challenges related to the use of data and analytics which need to be addressed. Insurers have yet to fully embrace the wide range of opportunities presented by data innovations. The industry (particularly the London market) is deemed by many to operate in an old-fashioned manner and the way data is consumed can be outdated. In many instances, a heavy reliance on expert opinion, qualitative evidence and subjective judgements, has been revealed. Nevertheless, emerging data sources have been identified, which can augment or displace some traditional methods and expert opinions. Furthermore, it is now possible to draw on large (real-time) datasets, linked to actual events, which can displace some in-use analytical methods reliant on statistical sampling. Through the application of advanced analytical processes information outputs can be generated from this data, which can replace gross assumptions, inherent in previous and current assessments of risk, and in doing so reduce uncertainty.

#### 7. Conclusion

Effective insurance can act as both as a measure to distribute, and a method to communicate risk. In relation to coastal flooding hazards, insurance has been clearly identified as one potential resilience increasing mechanism. In addition to insurance providing a safety net, if premiums are risk-based, it can also serve as a signalling mechanism, communicating levels of risk. However, insurance markets need to be freely functioning in order for them to fulfil this role. Examples have been provided within this paper of how market distortions are common in many countries, which can preclude risk-based pricing of flood insurance. This can act to reduce insurance's ability to incentivise risk-averse behaviour, as can a lack of insurance coverage, and flooding being bundled with other perils. However, for insurance to operate effectively and mitigate risk, it is reliant on the provision of accurate data. Such data can also reduce information asymmetries and has a central role in revealing exposure and ensuring policies are appropriately priced. This topic formed the main focus of this paper, and extensive interview-based research was undertaken, centring on the use of data within the London insurance market. In discussing feedback received, the process of data utilisation was split into a number of stages. These were: a. data sources, b. data access, c. data analytics, and d. communication of information outputs. Each stage was considered in turn, and associated challenges and opportunities highlighted.

Through focussing in detail on how data is utilised in insuring risks, it is deemed possible, by the authors, to optimise insurance's role as an instrument to mitigate risks associated with environmental hazards and other perils. A range of opportunities are presented by the increasing availability of 'Big Data' sources, advanced data mining and analytical techniques. Social media, EO, IoT, and crowd sourced data can be drawn on to provide more granular, higher resolution, up-to-date intelligence about environmental risks and their consequences. More traditional sources of information, such as claims data, still prove invaluable, and new techniques can be drawn on to improve how these are utilised. Advances in the field of 'Big Data' management and

analytics, can allow vast bodies of archive and streaming data (in a variety of formats), to form an evidence base for insurers to draw on. This can result in empirical data forming the basis of risk pricing, which can displace more subjective methods previously relied on, and in doing so, reduce moral hazard. Moreover, findings generated through this work have revealed how the extensive range of data sources, and analytical techniques on offer, can be effectively incorporated within insurance risk analyses. This can facilitate a process of evidence-based decision making, increasing the probability for insurance to generate socially optimal outcomes.

#### Acknowledgements

The authors acknowledge the staff of Lloyd's who contributed and provided guidance to this work, in particular Craig Civil. Also acknowledged are Coastal Partnership East, Great Yarmouth Borough, North Norfolk, Suffolk Coastal and Waveney District Councils and, in particular, Bill Parker and Paul Mackie, for input and guidance to this work. British Geological Survey, in particular Anna Harrison and Katherine Lee, are acknowledged for their guidance. Timothy Brewer of Cranfield University is acknowledged for guidance and support provided. This work was supported by the UK Natural Environment Research Council [NERC Ref: NE/M009009/1], Suffolk Coastal District Council [VGP/00043027], and the British Geological Survey [GA/16S/010].

#### Data access statement

Due to confidentiality agreements with research collaborators, supporting data (in the form of interview transcripts) cannot be made openly available.

#### References

- Abdullayev, K., 2014. *Global Reinsurance Market*. University of Vienna.
- Actuaries Institute, 2016. *The Impact of Big Data on the Future of Insurance*. pp. 1–36.
- Adam, E., Mutanga, O., Odindi, J., Abdel-Rahman, E.M., 2014. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *Int. J. Remote Sens.* 35, 3440–3458. <https://doi.org/10.1080/01431161.2014.903435>.
- Albano, R., Mancusi, L., Sole, A., Adamowski, J., 2017. FloodRisk: a collaborative, free and open-source software for flood risk analysis. *Geomatics, Nat. Hazards Risk* 8, 1812–1832. <https://doi.org/10.1080/19475705.2017.1388854>.
- André, C., Monfort, D., Bouzit, M., Vinchon, C., 2013. Contribution of insurance data to cost assessment of coastal flood damage to residential buildings: insights gained from Johanna (2008) and Xynthia (2010) storm events. *Nat. Hazards Earth Syst. Sci.* 13, 2003–2012. <https://doi.org/10.5194/nhess-13-2003-2013>.
- Andrews, T., Billen, G., Kim, H., Lui, A., Pollack, C., Simpson, L., Smith, D., Underhill, S., Whittle, D., 2008. *The insurance of flood risks*. Institute of Actuaries of Australia 16th General Insurance Seminar (Coolumb, Australia).
- Association of British Insurers, 2005. ABI statement of principles on the provision of flood insurance [WWW Document]. URL: <https://www.salford.gov.uk/media/384971/abistatementofprinciples2005.pdf>, Accessed date: 22 September 2018.
- Beck, M.W., Losada, I.J., Menéndez, P., Reguero, B.G., Díaz-Simal, P., Fernández, F., 2018. The global flood protection savings provided by coral reefs. *Nat. Commun.* 9. <https://doi.org/10.1038/s41467-018-04568-z>.
- Bezuglov, A., Blanton, B., Santiago, R., 2016. Multi-output Artificial Neural Network for Storm Surge Prediction in North Carolina.
- Bin, O., Kruse, J.B., Landry, C.E., 2006. *Flood Hazards, Insurance Rates, and Amenities: Evidence From the Coastal Housing Market*.
- Bonfield, P., 2016. *The Property Flood Resilience Action Plan*.
- Botzen, W.J.W., van den Bergh, J.C.J.M., 2008. Insurance against climate change and flooding in the Netherlands: present, future, and comparison with other countries. *Risk Anal.* 28, 413–426. <https://doi.org/10.1111/j.1539-6924.2008.01035.x>.
- Bowler, T.R., 2018. A new space race. In: James, T. (Ed.), *Deep Space Commodities: Exploration, Production and Trading*. Springer International Publishing, Cham, pp. 13–19. [https://doi.org/10.1007/978-3-319-90303-3\\_2](https://doi.org/10.1007/978-3-319-90303-3_2).
- de Bruijn, J.A., de Moel, H., Jongman, B., Wagemaker, J., Aerts, J.C.J.H., 2018. TAGGS: grouping tweets to improve global geoparsing for disaster response. *J. Geovisualization Spat. Anal.* 2, 2. <https://doi.org/10.1007/s41651-017-0010-6>.
- Carnegie-Brown, B., 2017. Chairman outlines vision for insurance market of the future [WWW Document]. URL: <https://www.lloyds.com/news-and-risk-insight/speeches/2017/london-market-conference-2017>, Accessed date: 22 September 2018.
- Castelvecchi, D., 2018. Google enters data ecosystem. *Nat. [News]* 561, 161–162. <https://doi.org/10.1038/d41586-018-06201-x>.

- Cervantes-Godoy, D., Kimura, S., Antón, J., 2013. Smallholder Risk Management in Developing Countries, OECD Food, Agriculture and Fisheries Papers, No. 61, OECD Food, Agriculture and Fisheries Papers. 61. <https://doi.org/10.1787/5k452k28wjlj-en>.
- Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J., Waters, N., 2016. Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *Int. J. Remote Sens.* 37, 100–124. <https://doi.org/10.1080/01431161.2015.1117684>.
- Chang, L.C., Amin, M., Yang, S.N., Chang, F.J., 2018. Building ANN-based regional multi-step-ahead flood inundation forecast models. *Water* 10, 1283. <https://doi.org/10.3390/w10091283>.
- Chen, Y., Fan, R., Yang, X., Wang, J., Latif, A., 2018. Extraction of Urban Water Bodies from High-Resolution Remote-Sensing Imagery Using Deep Learning. *Water* 10, 585. <https://doi.org/10.3390/w10050585>.
- Choi, T.-M., Lambert, J.H., 2017. Advances in Risk Analysis with Big Data. *Risk Anal.* 37, 1435–1442. <https://doi.org/10.1111/risa.12859>.
- Christie, E.K., Spencer, T., Owen, D., McIvor, A.L., Möller, I., Viavattene, C., 2018. Regional coastal flood risk assessment for a tidally dominant, natural coastal setting: North Norfolk, southern North Sea. *Coast. Eng.* 134, 177–190. <https://doi.org/10.1016/j.coastaleng.2017.05.003>.
- Clark, M.J., 1998. Flood insurance as a management strategy for UK coastal resilience. *Geogr. J.* 164, 333–343.
- Collinson, P., 2017. So direct line, why has my home insurance gone up by 442%? [WWW Document]. URL. [Guardhttps://www.theguardian.com/money/2017/sep/23/direct-line-home-insurance-up-assessing-flood-risk](https://www.theguardian.com/money/2017/sep/23/direct-line-home-insurance-up-assessing-flood-risk), Accessed date: 22 September 2018.
- Crichton, D., 2004. The role of private insurance companies in managing flood risks in the UK and Europe. *Urban Flood Risk Management: Introduction-1st International Expert Meeting on Urban Flood Management*. CRC Press, p. 83.
- Crichton, D., 2008. Role of insurance in reducing flood risk. *Geneva Pap. Risk Insur. Issues Pract.* 33, 117–132. <https://doi.org/10.1057/palgrave.gpp.2510151>.
- Crick, F., Jenkins, K., Surminski, S., 2018. Strengthening insurance partnerships in the face of climate change – insights from an agent-based model of flood insurance in the UK. *Sci. Total Environ.* 636, 192–204. <https://doi.org/10.1016/j.scitotenv.2018.04.239>.
- Dávila, O.G., Stithou, M., Pescaroli, G., Pietrantonio, L., Koundouri, P., Díaz-Simal, P., Rulleau, B., Touili, N., Hissel, F., Penning-Rowell, E., 2014. Promoting resilient economies by exploring insurance potential for facing coastal flooding and erosion: evidence from Italy, Spain, France and United Kingdom. *Coast. Eng.* 87, 183–192. <https://doi.org/10.1016/j.coastaleng.2013.12.007>.
- Dawson, R.J., Peppe, R., Wang, M., 2011. An agent based model for risk-based flood incident management. *Nat. Hazards* 59, 167–189. <https://doi.org/10.1007/s11069-011-9745-4>.
- De Angeli, S., D'Andrea, M., Cazzola, G., Dolia, D., Duo, E., Reboria, N., 2018. Coastal Risk Assessment Framework: comparison of modelled fluvial and marine inundation impacts, Bocca di Magra, Ligurian coast, Italy. *Coast. Eng.* 134, 229–240. <https://doi.org/10.1016/j.coastaleng.2017.09.011>.
- Deshpande, K., 2018. Collection and optimisation of insurance data: challenges now and in the future [WWW Document]. URL. [LexisNexis Insur. Insightshttp://blogs.lexisnexis.com/insurance-insights/2018/06/collection-and-optimisation-of-insurance-data-challenges-now-and-in-the-future/](https://blogs.lexisnexis.com/insurance-insights/2018/06/collection-and-optimisation-of-insurance-data-challenges-now-and-in-the-future/), Accessed date: 22 September 2018.
- DeWitt, J.D., Warner, T.A., Chirico, P.G., Bergstresser, S.E., 2017. Creating high-resolution bare-earth digital elevation models (DEMs) from stereo imagery in an area of densely vegetated deciduous forest using combinations of procedures designed for lidar point cloud filtering. *GIScience Remote Sens.* 54, 552–572. <https://doi.org/10.1080/15481603.2017.1295514>.
- DigitalGlobe, 2017. GBDX stories: building detection on mosaics using deep learning [WWW Document]. URL. <http://gbdxstories.digitalglobe.com/building-detection-large-scale/>, Accessed date: 22 September 2012.
- Du, E., Cai, X., Sun, Z., Minsker, B., 2017. Exploring the role of social media and individual behaviors in flood evacuation processes: an agent-based modeling approach. *Water Resour. Res.* 53, 9164–9180. <https://doi.org/10.1002/2017WR021192>.
- Dubbelboer, J., Nikolic, I., Jenkins, K., Hall, J., 2017. An Agent-Based Model of Flood Risk and Insurance. *J. Artif. Soc. Soc. Simul.* 20. <https://doi.org/10.18564/jasss.3135>.
- Ekeu-wei, I.T., Blackburn, G.A., 2018. Applications of Open-Access Remotely Sensed Data for Flood Modelling and Mapping in Developing Regions. *Hydrology* 5, 39. <https://doi.org/10.3390/hydrology5030039>.
- Filatova, T., Mulder, J.P.M., Veen, A., 2011. Coastal risk management: how to motivate individual economic decisions to lower flood risk? *Ocean Coast. Manag.* 54, 164–172. <https://doi.org/10.1016/j.ocecoaman.2010.10.028>.
- Flood RE [WWW Document], n.d. URL <https://www.floodre.co.uk/about-us/> (accessed 9.22.18).
- Garvin, S., Hunter, K., McNally, D., Barnett, D., Dakin, R., Nexis, L., Street, D.O., 2016. Property Flood Resilience Database: An Innovative Response for the Insurance Market. , p. 22002 <https://doi.org/10.1051/e3sconf/20160722002>.
- Geller, C., 2017. Automated flood extent identification using WorldView imagery for the insurance industry. *Remote Sens. Agric. Ecosyst. Hydrol.* Xix 10421. <https://doi.org/10.1117/1.2278075>.
- Grason, David, 2018. How we use machine learning for targeted location monitoring [WWW Document]. URL. [DigitalGlobehttp://blog.digitalglobe.com/developers/how-we-use-machine-learning-for-targeted-location-monitoring/](http://blog.digitalglobe.com/developers/how-we-use-machine-learning-for-targeted-location-monitoring/), Accessed date: 22 September 2018.
- Gritta, M., Pilehvar, M.T., Limsopatham, N., Collier, N., 2018. What's missing in geographical parsing? *Lang. Resour. Eval.* 52, 603–623. <https://doi.org/10.1007/s10579-017-9385-8>.
- Haer, T., Botzen, W.J.W., de Moel, H., Aerts, J.C.J.H., 2017. Integrating household risk mitigation behavior in flood risk analysis: an agent-based model approach. *Risk Anal.* 37, 1977–1992. <https://doi.org/10.1111/risa.12740>.
- Hassan, Q.F., Khan, A. ur R., Madani, S.A. (Eds.), 2018. *Internet of Things*. Chapman and Hall/CRC, New York.
- Heale, B., 2014. Data quality is the biggest challenge [WWW Document]. *Moody's Anal. Risk Perspect.* 3, 30–37 URL. <https://www.moodyanalytics.com/risk-perspectives-magazine/managing-insurance-risk/insurance-regulatory-spotlight/data-quality-is-the-biggest-challenge>, Accessed date: 22 September 2018.
- Hénaff, Q., Jerrari, O., Ali, H., Brettschneider, M., Blaquière, S., 2018. On the use of insurance geodata to map flood-risk areas. *EGU General Assembly Conference Abstracts*, p. 14924.
- Hsu, W.-K., Huang, P.-C., Chang, C.-C., Chen, C.-W., Hung, D.-M., Chiang, W.-L., 2011. An integrated flood risk assessment model for property insurance industry in Taiwan. *Nat. Hazards* 58, 1295–1309. <https://doi.org/10.1007/s11069-011-9732-9>.
- Hudson, P., Botzen, W.J.W., Feyen, L., Aerts, J.C.J.H., 2016. Incentivising flood risk adaptation through risk based insurance premiums: trade-offs between affordability and risk reduction. *Ecol. Econ.* 125, 1–13. <https://doi.org/10.1016/j.ecolecon.2016.01.015>.
- Intermap, 2018. Nextmap [WWW Document]. URL. <https://www.intermap.com/products/nextmap/data>, Accessed date: 22 September 2018.
- Jagadish, H.V., 2015. Big Data and Science: Myths and Reality. *Big Data Res.* 2, 49–52. <https://doi.org/10.1016/j.bdr.2015.01.005>.
- Jenkins, K., Surminski, S., Hall, J., Crick, F., 2017. Assessing surface water flood risk and management strategies under future climate change: insights from an Agent-Based Model. *Sci. Total Environ.* 595, 159–168. <https://doi.org/10.1016/j.scitotenv.2017.03.242>.
- Jongman, B., 2018. Effective adaptation to rising flood risk. *Nat. Commun.* 9, 9–11. <https://doi.org/10.1038/s41467-018-04396-1>.
- Jongman, B., Koks, E.E., Husby, T.G., Ward, P.J., 2014. Increasing flood exposure in the Netherlands: implications for risk financing. *Nat. Hazards Earth Syst. Sci.* 14, 1245–1255. <https://doi.org/10.5194/nhess-14-1245-2014>.
- Jongman, B., Wagemaker, J., Romero, B., de Perez, E., 2015. Early Flood Detection for Rapid Humanitarian Response: Harnessing Near Real-Time Satellite and Twitter Signals. *ISPRS Int. J. Geo-Information* 4, 2246–2266. <https://doi.org/10.3390/ijgi4042246>.
- Joseph, C.K., Kakade, S., 2014. Predicting impact of natural calamities in era of Big Data and data science. *International Environmental Modelling and Software Society (IEMSS) 7th Intl. Congress on Env. Modelling and Software*. San Diego, CA, USA.
- Koduru, S., Prasad Reddy, P., Padala, P., 2018. Integrated disaster management and smart insurance using cloud and internet of things. *Int. J. Eng. Technol.* 7, 341. <https://doi.org/10.14419/ijet.v7i2.6.10777>.
- Koivumäki, L., Alho, P., Lotsari, E., Käyhkö, J., Saari, A., Hyyppä, H., 2010. Uncertainties in flood risk mapping: a case study on estimating building damages for a river flood in Finland. *J. Flood Risk Manag.* 3, 166–183. <https://doi.org/10.1111/j.1753-318X.2010.01064.x>.
- Kron, W., 2013. Coasts: The high-risk areas of the world. *Nat. Hazards* 66, 1363–1382. <https://doi.org/10.1007/s11069-012-0215-4>.
- Kunreuther, H., 1984. Causes of Underinsurance against Natural Disasters. *Geneva Pap. Risk Insur.* 9, 206–220.
- Kunreuther, H., Michel-Kerjan, E., Tonn, G., 2016. Insurance, Economic Incentives and Other Policy Tools for Strengthening Critical Infrastructure Resilience: 20 Proposals for Action 56.
- Lamond, J., Penning-Rowell, E., 2014. The robustness of flood insurance regimes given changing risk resulting from climate change. *Clim. Risk Manag.* 2, 1–10. <https://doi.org/10.1016/j.crm.2014.03.001>.
- Landry, C.E., Jahan-Parvar, M.R., 2011. Flood insurance coverage in the coastal zone. *J. Risk Insur.* 78, 361–388. <https://doi.org/10.1111/j.1539-6975.2010.01380.x>.
- Lavender, S., Haria, K., Cooksley, G., Farman, A., Beaton, T., 2016. Understanding the Future Market for NOVASAR-S flood Mapping. *Living Planet Symposium*, p. 380.
- LexisNexis, 2018. HPC systems [WWW Document]. URL. <https://hpcsystems.com>, Accessed date: 22 September 2018.
- Li, Z., Wang, C., Emrich, C.T., Guo, D., 2018. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. *Cartogr. Geogr. Inf. Sci.* 45, 97–110. <https://doi.org/10.1080/15230406.2016.1271356>.
- Libarikian, A., Javanmardian, K., McElhane, D., Majumder, A., 2017. Harnessing the potential of data in insurance [WWW Document]. URL. [McKinsey Co.https://www.mckinsey.com/industries/financial-services/our-insights/harnessing-the-potential-of-data-in-insurance](https://www.mckinsey.com/industries/financial-services/our-insights/harnessing-the-potential-of-data-in-insurance), Accessed date: 22 September 2018.
- Linham, M.M., Nicholls, R.J., 2010. Technologies for Climate Change Adaptation: Coastal Erosion and Flooding (... Adaptation: Coastal Erosion and Flooding).
- Lloyd's, 2018a. After the Storms Harvey, Maria, Irma and.
- Lloyd's, 2018b. Realistic disaster scenarios (RDS) [WWW Document]. URL. <https://www.lloyds.com/market-resources/underwriting/realistic-disaster-scenarios-rds>, Accessed date: 25 September 2018.
- Lloyd's Market Association, 2017. Market briefing: claims imagery & intelligence service [WWW Document]. URL. [http://www.lmalloyds.com/LMA/News/Blog/Claims\\_Imagery\\_Intelligence\\_Service\\_-\\_4\\_September\\_2017.aspx](http://www.lmalloyds.com/LMA/News/Blog/Claims_Imagery_Intelligence_Service_-_4_September_2017.aspx), Accessed date: 22 September 2018.
- Loftis, J.D., Forrest, D., Katragadda, S., Spencer, K., Organski, T., Rhee, S., Nguyen, C., 2018. StormSense: a New Integrated Network of IoT Water Level Sensors in the Smart Cities of Hampton Roads, VA. *Mar. Technol. Soc. J.* 52, 1–30. <https://doi.org/10.4031/MTSJ.52.2.7>.
- London Market Group, 2018. Why London? Telling the London market story [WWW Document]. URL. <https://www.londonmarketgroup.co.uk/why-london-overview>, Accessed date: 22 September 2018.
- Maier, D., Megler, V.M., Baptista, A.M., Jaramillo, A., Seaton, C., Turner, P.J., 2012. Navigating oceans of data. *Scientific and Statistical Database Management Conference*, pp. 1–19 [https://doi.org/10.1007/978-3-642-31235-9\\_1](https://doi.org/10.1007/978-3-642-31235-9_1).
- Marr, B., 2015. *Big Data: Using SMART Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance*. Wiley, Cichester.

- Marz, N., Warren, J., 2015. *Big Data: Principles and Best Practices of Scalable Realtime Data Systems*.
- Mashaly, J., Ghoneim, E., 2018. Flash flood hazard using optical, radar, and stereo-pair derived DEM: Eastern Desert, Egypt. *Remote Sens.* 10, 1204. <https://doi.org/10.3390/rs10081204>.
- McNamara, D.E., Werner, B.T., 2008. Coupled barrier island-resort model: 1. Emergent instabilities induced by strong human-landscape interactions. *J. Geophys. Res. Earth Surf.* 113, 1–10. <https://doi.org/10.1029/2007JF000840>.
- Mercier, R., 2015. How we put together a list of 2600+ open data portals around the world for the open data community [WWW Document]. URL: [www.opendatasoft.com/2015/11/02/how-we-put-together-a-list-of-1600-open-data-portals-around-the-world-to-help-open-data-community/](http://www.opendatasoft.com/2015/11/02/how-we-put-together-a-list-of-1600-open-data-portals-around-the-world-to-help-open-data-community/), Accessed date: 22 September 2018.
- Miller, P., 2018. Insurance data and technology: challenges and opportunities for our industry [WWW Document]. Inst. Risk Insur. Knowl. Gr. URL: <https://www.theinstitutes.org/guide/insurance-data-and-technology-challenges-and-opportunities-our-industry>, Accessed date: 22 September 2018.
- de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E., Ward, P.J., 2015. Flood risk assessments at different spatial scales. *Mitig. Adapt. Strateg. Glob. Chang.* 20, 865–890. <https://doi.org/10.1007/s11027-015-9654-z>.
- Mokrech, M., Kebede, A.S., Nicholls, R.J., Wimmer, F., Feyen, L., 2014. An integrated approach for assessing flood impacts due to future climate and socio-economic conditions and the scope of adaptation in Europe. *Clim. Chang.* 128, 245–260. <https://doi.org/10.1007/s10584-014-1298-6>.
- Monrat, A.A., Islam, R.U., Hossain, M.S., Andersson, K., 2011. A belief rule based flood risk assessment expert system using real time sensor data streaming. *Proceedings of the 43rd IEEE Conference on Local Computer Networks Workshops (LCN Workshops)*. IEEE Computer Society, p. 2018.
- Moszynski, M., Kulawiak, M., Chybicki, A., Bruniecki, K., Bieliński, T., Łubniewski, Z., Stepnowski, A., 2015. Innovative Web-Based Geographic Information System for Municipal Areas and Coastal Zone Security and Threat Monitoring Using EO Satellite Data. *Mar. Geod.* 38, 203–224. <https://doi.org/10.1080/01490419.2014.969459>.
- OECD, 2016. *Financial Management of Flood Risk*.
- Ogie, R.L., Perez, P., Win, K.T., Michael, K., 2018. Managing hydrological infrastructure assets for improved flood control in coastal mega-cities of developing nations. *Urban Clim.* 24, 763–777. <https://doi.org/10.1016/j.uclim.2017.09.002>.
- Osberghaus, D., 2017. The effect of flood experience on household mitigation—evidence from longitudinal and insurance data. *Glob. Environ. Chang.* 43, 126–136. <https://doi.org/10.1016/j.gloenvcha.2017.02.003>.
- Panteras, G., Cervone, G., 2018. Enhancing the temporal resolution of satellite-based flood extent generation using crowdsourced data for disaster monitoring. *Int. J. Remote Sens.* 39, 1459–1474. <https://doi.org/10.1080/01431161.2017.1400193>.
- Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S.J., Fosso-Wamba, S., 2017. The role of Big Data in explaining disaster resilience in supply chains for sustainability. *J. Clean. Prod.* 142, 1108–1118. <https://doi.org/10.1016/j.jclepro.2016.03.059>.
- Penning-Rowsell, E., Priest, S., Parker, D., Morris, J., Tunstall, S., Viavattene, C., Chatterton, J., Owen, D., 2013. *Flood and Coastal Erosion Risk Management*. Routledge, London.
- Peters, G., 2017. *Statistical Machine Learning and Data Analytic Methods for Risk and Insurance*. <https://doi.org/10.2139/ssrn.3050592>.
- Pilla, F., Gharbia, S.S., Lyons, R., 2018. How do households perceive flood-risk? The impact of flooding on the cost of accommodation in Dublin, Ireland. *Sci. Total Environ.* 650, 144–154. <https://doi.org/10.1016/j.scitotenv.2018.08.439>.
- Ramieri, E., Hartley, A., Barbanti, A., Santos, F.D., Gomes, A., Hilden, M., Laihonon, P., Marinova, N., Santini, M., 2011. *Methods for Assessing Coastal Vulnerability to Climate Change*. ETC CCA Technical Paper (Bologna (IT)).
- Roberts, R.S., 2012. *Economic Strategies for Coastal Disaster Risk-reduction: A Case Study of Exmouth, Western Australia*. PhD. Murdoch University.
- Roth, B., Conforti, C., Poerner, N., Karn, S., Schütze, H., 2018. *Neural Architectures for Open-type Relation Argument Extraction*.
- Rumson, A.G., Hallett, S.H., 2018. Opening up the coast. *Ocean Coast. Manag.* 160. <https://doi.org/10.1016/j.ocecoaman.2018.04.015>.
- Schwartz, J., 2018. National flood insurance is underwater because of outdated science [WWW Document]. *Sci. Am.* 318 (3) URL: <https://www.scientificamerican.com/article/national-flood-insurance-is-underwater-because-of-outdated-science/>, Accessed date: 22 August 2018.
- Singhal, A., Pant, R., Sinha, P., 2018. *AlertMix: A Big Data Platform for Multi-source Streaming Data*.
- SmartBay, 2017. Ireland [WWW Document]. URL: <http://www.smartbay.ie/>, Accessed date: 15 May 2017.
- Smith, L., Liang, Q., James, P., Lin, W., 2017. Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. *J. Flood Risk Manag.* 10, 370–380. <https://doi.org/10.1111/jfr3.12154>.
- Stoekli, E., Dremel, C., Uebernickel, F., 2018. Exploring characteristics and transformational capabilities of InsurTech innovations to understand insurance value creation in a digital world. *Electron. Mark.* 1–19. <https://doi.org/10.1007/s12525-018-0304-7>.
- Surminski, S., 2014. The role of insurance in reducing direct risk—the case of flood insurance. *Int. Rev. Environ. Resour. Econ.* 7, 241–278. <https://doi.org/10.1561/101.00000062>.
- Surminski, S., Eldridge, J., 2015. Flood insurance in England – an assessment of the current and newly proposed insurance scheme in the context of rising flood risk. *J. Flood Risk Manag.* 1–21. <https://doi.org/10.1111/jfr3.12127>.
- Surminski, S., Oramas-Dorta, D., 2014. Flood insurance schemes and climate adaptation in developing countries. *Int. J. Disaster Risk Reduct.* 7, 154–164. <https://doi.org/10.1016/j.jidrr.2013.10.005>.
- The American Academy of Actuaries, 2017. *The National Flood Insurance Program: Challenges and Solutions* (Washington, D.C.).
- The Apache Software Foundation, 2018. No title [WWW Document]. URL: <https://www.apache.org>, Accessed date: 22 September 2018.
- The University of Bristol, n.d. LISFLOOD-FP [WWW Document]. URL: <http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/> (accessed 9.22.18).
- Thislethwaite, J., Minano, A., Blake, J.A., Henstra, D., Scott, D., 2018. Application of re/insurance models to estimate increases in flood risk due to climate change. *Geoenvironmental Disasters* 5, 8. <https://doi.org/10.1186/s40677-018-0101-9>.
- Tischhauser, P., 2017. Leading the insurance market to a new destination [WWW Document]. Bost. Consult. Gr. URL: <https://www.bcg.com/publications/2017/insurance-market-new-destinations-inga-beale.aspx>, Accessed date: 22 September 2018.
- Tsarouchi, G., 2018. *Drought and flood mitigation service for Uganda*. EGU General Assembly 2018 (Vienna Austria).
- UK Space Agency, 2017a. *Space for smarter government program: natural hazards risk management* [WWW Document]. URL: <https://www.spaceforsmartergovernment.uk/natural-hazards/>, Accessed date: 22 September 2018.
- UK Space Agency, 2017b. *Case study: Pixalytics – optical and microwave extension for floodwater mapping (OMEF)* [WWW Document]. URL: <https://spaceforsmartergovernment.uk/case-study/pixalytics-optical-and-microwave-extension-for-floodwater-mapping-omef/>, Accessed date: 22 September 2018.
- Viavattene, C., Jiménez, J.A., Ferreira, O., Priest, S., Owen, D., McCall, R., 2018. Selecting coastal hotspots to storm impacts at the regional scale: a Coastal Risk Assessment Framework. *Coast. Eng.* 134, 33–47. <https://doi.org/10.1016/j.coastaleng.2017.09.002>.
- von Dahlen, S., von Peter, G., 2012. *Natural Catastrophes and Global Reinsurance – Exploring the Linkages*. BIS Q. Rev. Available at SSRN: <https://ssrn.com/abstract=2206335>.
- Walsh, T., Layton, T., Wanik, D., Mellor, J., 2018. *Agent Based Model to Estimate Time to Restoration of Storm-Induced Power Outages*. *Infrastructures* 3, 33. <https://doi.org/10.3390/infrastructures3030033>.
- Wang, R., 2018. *Big Data of Urban Flooding: Dance with Social Media, Citizen Science, and Artificial Intelligence*. 20th EGU General Assembly, EGU2018, Proceedings from the Conference, p. 404.
- Wang, R.Q., Mao, H., Wang, Y., Rae, C., Shaw, W., 2018. Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. *Comput. Geosci.* 111, 139–147. <https://doi.org/10.1016/j.cageo.2017.11.008>.
- Werner, D., 2018. *British startup Earth-i shares first color video from its VividX2 satellite, a prototype for imagery constellation* [WWW Document]. URL: <http://spacenews.com/earth-i-unveils-first-color-video-captured-by-its-vividx2-satellite/>, Accessed date: 24 September 2018.
- Wilding, T.A., Gill, A.B., Boon, A., Sheehan, E., Dauvin, J.C., Pezy, J.P., O'Beim, F., Janas, U., Rostin, L., De Mesel, I., 2017. Turning off the DRIP ('Data-rich, information-poor') – rationalising monitoring with a focus on marine renewable energy developments and the benthos. *Renew. Sust. Energ. Rev.* 74, 848–859. <https://doi.org/10.1016/j.rser.2017.03.013>.
- Yang, L.E., Scheffran, J., Süsser, D., Dawson, R., Chen, Y.D., 2018. Assessment of Flood Losses with Household Responses: Agent-Based Simulation in an Urban Catchment Area. *Environ. Model. Assess.* 1–20. <https://doi.org/10.1007/s10666-018-9597-3>.
- Zischg, A.P., Mosimann, M., Bernet, D.B., Röthlisberger, V., 2018. Validation of 2D flood models with insurance claims. *J. Hydrol.* 557, 350–361. <https://doi.org/10.1016/j.jhydrol.2017.12.042>.

# Innovations in the use of data facilitating insurance as a resilience mechanism for coastal flood risk

Rumson, Alexander G.

2019-01-14

Attribution 4.0 International

---

Rumson AG, Hallett SH. (2019) Innovations in the use of data facilitating insurance as a resilience mechanism for coastal flood risk. *Science of the Total Environment*, Volume 661, April 2019, pp. 598-612

<https://doi.org/10.1016/j.scitotenv.2019.01.114>

*Downloaded from CERES Research Repository, Cranfield University*