



Invited perspectives: How machine learning will change flood risk and impact assessment

Dennis Wagenaar^{1,2}, Alex Curran^{1,3}, Mariano Balbi⁴, Alok Bhardwaj⁵, Robert Soden^{6,7,8}, Emir Hartato⁹, Gizem Mestav Sarica¹⁰, Laddaporn Ruangpan^{11,3}, Giuseppe Molinaro⁷, and David Lallemand^{5,8}

¹Department of flood risk management, Deltares, Delft, the Netherlands

²Institute for environmental studies, VU University, Amsterdam, the Netherlands

³Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands

⁴Structural and Materials Lab, School of Engineering, Universidad de Buenos Aires, Buenos Aires, Argentina

⁵Earth Observatory of Singapore, Nanyang Technological University, Singapore

⁶Columbia University, New York City, New York, USA

⁷GFDRR, World Bank Group, Washington, D.C., USA

⁸Co-Risk Labs, Oakland, California, USA

⁹Planet, San Francisco, USA

¹⁰Institute of Catastrophe Risk Management, Nanyang Technological University, Singapore

¹¹Department of Water Resources and Ecosystems, IHE Delft Institute for Water Education, Delft, the Netherlands

Correspondence: Dennis Wagenaar (dennis.wagenaar@deltares.nl)

Received: 14 October 2019 – Discussion started: 16 October 2019

Revised: 17 February 2020 – Accepted: 15 March 2020 – Published: 29 April 2020

Abstract. Increasing amounts of data, together with more computing power and better machine learning algorithms to analyse the data, are causing changes in almost every aspect of our lives. This trend is expected to continue as more data keep becoming available, computing power keeps improving and machine learning algorithms keep improving as well. Flood risk and impact assessments are also being influenced by this trend, particularly in areas such as the development of mitigation measures, emergency response preparation and flood recovery planning. Machine learning methods have the potential to improve accuracy as well as reduce calculating time and model development cost. It is expected that in the future more applications will become feasible and many process models and traditional observation methods will be replaced by machine learning. Examples of this include the use of machine learning on remote sensing data to estimate exposure and on social media data to improve flood response. Some improvements may require new data collection efforts, such as for the modelling of flood damages or defence failures. In other components, machine learning may not always be suitable or should be applied complementary to process models, for example in hydrodynamic applications. Overall,

machine learning is likely to drastically improve future flood risk and impact assessments, but issues such as applicability, bias and ethics must be considered carefully to avoid misuse. This paper presents some of the current developments on the application of machine learning in this field and highlights some key needs and challenges.

1 Introduction

Exponentially increasing computing power and data, as well as rapidly improving machine learning algorithms to analyse these data, have been changing many aspects of our lives (Manyika et al., 2011). These trends are expected to continue and will undoubtedly keep affecting many scientific, commercial and social sectors (Manyika et al., 2011). Flood risk and impact assessments are no exception to this trend. Flooding yearly affects more people than any other natural hazard type (Jonkman, 2005), and the impact and frequency of flooding events is expected to increase in the future due to urban development and climate change (Kundzewicz et al., 2014). It is therefore an opportunity for researchers and

Table 1. Overview of different types of flood risk and impact assessments.

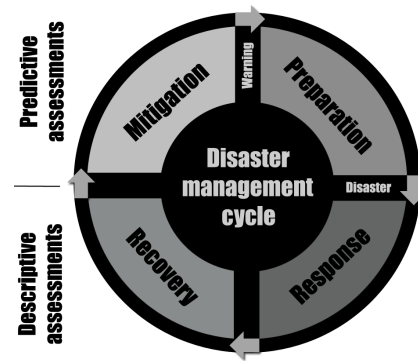
	Predictive	Descriptive
Exposure	Urban growth modelling	Identification of current built-up area
Hazard	Flood modelling	Mapping current and past floods
Impact	Forecasting impact (e.g. damage)	Assessing flood impacts (e.g. damage) after they have occurred

flood managers to tap into the potential of machine learning, taking advantage of their strengths while being cognizant of their limitations. It is also important to anticipate improvements in the capabilities of machine learning methods, so as to plan for forthcoming changes in flood modelling.

When assessing the interaction between floods and society, three different components can be recognized: exposure, hazard and impact (Kron, 2002). Exposure refers to the characteristics of the people and assets that can be affected by flooding. Hazards are the physical characteristics of a flood such as the extent, water depth, duration and flow velocity. Impacts are the effects the hazard has on the exposure. To assess these three components, we make the distinction between flood risk, as the probabilistic analysis of the potential (predictive) impacts of floods, and flood impact assessment, as the post-event assessment of (descriptive) impact from an actual flood event. Table 1 provides examples of predictive and descriptive assessments in relation to the hazard, exposure and impact components. The scope of this paper is limited to the predictive and descriptive assessments shown in Table 1 and does not include potential uses of machine learning in risk awareness or communication strategies.

Flood risk and impact assessments have many different applications. A useful paradigm through which look at these different applications is the “disaster management cycle” (Khan et al., 2008; National Research Council, 2006) (Fig. 1). This cycle delineates the phases between events, i.e. the immediate response to an event, the long-term recovery, the mitigation to prevent future events and the preparation prior to a new forecasted event.

In the response phase, the focus is typically on descriptive hazard, exposure and impact assessments (e.g. Klemas, 2015), sometimes complemented with predictive models if the event descriptive information is not available yet (e.g. a predictive model estimating the number of people affected can be fed by a descriptive hazard model of the flood extent). The challenge in this phase is mostly data reliability. In the recovery phase, descriptive assessments are often used for payouts (e.g. indemnity insurance), and one of the main challenges is ensuring these payouts are timely and reliable. In the mitigation phase, probabilistic predictive models are

**Figure 1.** Disaster management cycle, a common paradigm tool.

used (e.g. Wagenaar et al., 2019), typically for the design of risk-reduction interventions ranging from protective infrastructure to insurance products. The challenge in this phase is model reliability and uncertainties about future developments (e.g. uncertainty in future exposure). In the preparation phase, predictive models are used for emergency planning (e.g. Coughlan de Perez et al., 2016), where the challenge is the reliability, availability and communication of data. Machine learning is capable of generating more reliable and faster models that can help solve some of the current challenges in the disaster management cycle but could also provide new opportunities (GFDRR, 2018).

Machine learning algorithms can find patterns in data and use these patterns to make predictions about new data (Bishop, 2006). For example, when a machine learning algorithm is provided with aerial images of either urban or rural areas and corresponding labels (urban or rural), it can build the capacity to classify new unlabelled aerial images as either urban or rural. Features in the above example would be different components of the aerial images (i.e. pixel tone and locations), and the target variable would be the label (i.e. urban or rural). When a precise value is required as opposed to a label, it is called a “regression task” (e.g. Bishop, 2006). An example of this is in flood damage modelling, where features such as water depth, flow velocity and building materials can be used to predict a target variable such as monetary economic damage based on historical records (e.g. Merz et al., 2013; Wagenaar et al., 2017). Due to the use of labelled training data (e.g. classified images or historic damage examples), regression and classification are called supervised learning tasks. Machine learning method categories also include unsupervised learning and reinforcement learning (see GFDRR, 2018). However, such methods are not discussed in this paper because they are expected to have a smaller short-term impact on the field of flood risk and impact assessments.

The simplest machine learning algorithms have been used for a long time and are often known as basic statistical techniques (e.g. linear regression: Legendre, 1805; Gauss, 1809). More sophisticated machine learning techniques that emerged in the 1980s and 1990s (e.g. decision trees and

neural networks) can find more complex non-linear patterns (Breimann et al., 1984; Rumelhart et al., 1986). Recent advances in machine learning (e.g. convolutional neural networks) make computer vision and other advanced applications possible (Krizhevsky et al., 2012). The more advanced techniques such as decision trees, neural networks and especially convolutional neural networks can find more complex patterns. This is because they allow for more complex non-linear functions to be fitted to the data. Such complex functions require a large number of model coefficients to be set during the training of the model. To set all these coefficients, a lot of training examples are required. In some cases the number of training examples can be reduced with transfer learning techniques (Olivas et al., 2010). These techniques make it possible to re-use knowledge gained from other problems to train a model on a smaller training data set.

From the beginning, machine learning has been used in predictive flood hazard modelling (Solomatine and Ostfield, 2008) mostly as a faster and simpler alternative to process models. A simple example of this is the prediction of river discharge based on upstream rainfall data (e.g. Dibike and Solomatine, 2001). This type of modelling has been practised for a long time but has not displaced the traditional process models. This is probably because the methods are not sufficiently better than traditional methods to offset some disadvantages as discussed in the predictive-hazard section. In recent years, more data have become available through remote sensing, social media (e.g. Fohringer et al., 2015), citizen science (e.g. Annis and Nardi, 2019) and other sources. This impulse of new data combined with machine algorithms could lead to changes in flood risk and impact assessment. Some of these changes have already been highlighted by major international organizations such as the World Bank (GFDRR, 2018).

This invited perspective paper starts with a perspective per risk assessment component as defined in Table 1. These specific perspectives start with a description of the traditional approach for the assessments, followed by a literature review on how machine learning techniques are currently being developed to improve the traditional approach, and then proceed to speculate on potential future improvements. This is followed by a “General perspectives” chapter, in which general trends that come back in the different components are identified and discussed. This includes common challenges (i.e. data limitations, transferability, ethics and bias) and ends with some speculation about the likelihood of future developments.

2 Perspective per component

2.1 Exposure assessment

2.1.1 Descriptive exposure assessments

Descriptive exposure assessments consist of detecting and characterizing (spatial) features such as current buildings, agriculture fields, roads and other infrastructure. Traditionally this has been done by population censuses, building counts and conventional mapping techniques that require ground surveys. Remote sensing is currently changing this. It has become common for aerial and satellite images to be manually digitized and labelled to make building footprints or map roads. This has been done by “crowds” of mappers in “mapathons”, for example using the OpenStreetMap platform. Machine learning is very likely going to drastically change this. Research into automatically labelling remote sensing data has already been going on for some time (e.g. Heermann and Khazenie, 1992; Giacinto and Roli, 2001). It is already being used to label build-up areas based on night-time lights (Goldblatt et al., 2018) or satellite images (Goldblatt et al., 2016). Furthermore, algorithms are already being used to automatically label buildings (Sermanet et al., 2014; Alshehhi et al., 2017; GFDRR, 2018) and map roads (Gao et al., 2019) using aerial/satellite imagery. This will reduce the need for manual detection and will probably provide global availability of such building footprints and road information in the near future.

Part of an exposure assessment is the observation of asset features relevant to risk analysis, for example building materials, building occupancy (e.g. residential or industrial), building height, ground floor elevation, poverty rates in the population etc. This information is typically not available but could be very valuable as input for impact models (e.g. Merz et al., 2013; Wagenaar et al., 2017; Schröter et al., 2014) or, for example, to account for poverty in cost–benefit analyses (e.g. Kind et al., 2016). Similarly, ground floor elevation information could radically improve urban pluvial flood damage modelling as damage from small-scale floods is very sensitive to such variables.

Some work has already been carried out on detecting poverty (Watmough et al., 2019) and building heights (Saadi and Bensaibi, 2014) by satellite imagery. Another source of this building feature information could be 360° street view images combined with computer vision techniques. Such images are available in, for example, the open-source street view data platform Mapillary (Neuhold et al., 2017). Such techniques are already starting to impact earthquake risk assessments, such as in Guatemala, where 360° imagery was fed into Mapillary algorithms in order to automatically detect “soft story” buildings: those most likely to collapse in an earthquake. This was done by having the machine learning algorithm detect features that were indicators of large openings on the ground floor of buildings (large doors, garage

doors, shop windows etc.) (GFDRR, 2018). Computer vision techniques from street level imagery are currently limited to detecting such relatively simple features. However, based on recent advances seen in other computer vision applications (e.g. facial recognition), it is likely that in the future it will be possible to detect more complex building features as well. For computer vision models to detect complex information like ground floor elevation or building materials, it would be necessary to provide labelled examples to the algorithms. Such labelled examples are in some cases already available for some areas, e.g. ground floor elevation (Bouwer et al., 2017) or building materials (Schröter et al., 2018).

2.1.2 Predictive exposure assessments

Predictive exposure mapping consists of estimates of future exposure. This mostly includes modelling to predict urban growth and other changes in land use. It is required for evaluating flood mitigation measures (e.g. Wagenaar et al., 2019) because such measures typically need to function for a long time and should therefore still perform as required after predicted land-use changes. Land-use changes affect the impact of a flood because more damage may occur for the same flood hazard and the flood hazard may become greater because of changes in impervious area and therefore rainfall runoff (Triantakou and Mountrakis, 2013; Mestav Sarica et al., 2019). Predictive exposure assessments for flood risk and impact assessments are currently often not carried out spatially, but rather GDP growth projections are applied to estimate future total exposure values (e.g. van der Most et al., 2014; Wagenaar et al., 2019). This is enough for some studies, but if large changes are expected a land-use change or urban growth model is required.

Urban growth has been modelled with simple machine learning models in the past (e.g. logistic regression) (Samardzic-Petrovic et al., 2017). The use of cellular automata (CA) models has become more common recently (Naghbi et al., 2016). These models assign cells as either urban or non-urban based on specific transition rules. Determining the optimum transition rules is a critical issue for CA modelling (Aarthi and Gnanappazham, 2019). This is sometimes difficult because of human bias, heterogeneity and non-linear relations between driving factors and urban expansion (Naghbi et al., 2016; Xu et al., 2019). To overcome these limitations, machine learning algorithms such as artificial neural networks have been integrated with traditional CA to model urban growth (Aarthi and Gnanappazham, 2019; Naghbi et al., 2016). They then use historical land-use changes (e.g. Song et al., 2015) to learn the transition rules. Complex machine learning models have also been directly applied to urban growth modelling without the CA model structure (Pal and Ghosh, 2017). These improvements, together with more data about past land-use changes and additional computation power, are expected to provide better

future land-use maps and make high-resolution future land-use maps globally available.

2.2 Hazards assessment

2.2.1 Descriptive hazard assessments

Descriptive flood hazard assessment focuses primarily on the response phase, i.e. estimating current inundation extents and depths to assist both emergency responders and those affected directly. This is traditionally achieved using optical remote sensing data, local sensors or manually collected data from observers on the ground. However, the rise of two major data sources, synthetic aperture radar (SAR) and social media, provides a number of opportunities for machine learning to improve upon current flood detection methods.

During a flood event, affected populations frequently produce “user-generated content” or “crowd-sourced” data from social media posts or apps where citizens can report floods (Mazoleni et al., 2017; Assumpção et al., 2018; Annis and Nardi, 2019; UrbanRiskLab, 2019). This is especially the case in urban areas, where internet and social media penetration are higher compared to rural areas. These data are often “tagged” temporally and spatially and can be used by machine learning algorithms for applications such as nowcasting by searching for certain keywords like “flood” (e.g. see Tkachenko et al., 2017; Bischke et al., 2017; Lopez-Fuentes et al., 2017). The method is currently used to map real-time flood extents in several countries (Eilander et al., 2016). Potential future machine learning and computer vision techniques could be extended to estimate water depths and other flood characteristics from posted photos.

Remotely sensed optical data are often used to identify the extents of flooding, but optical sensors are not functional during periods of cloud cover or at night. Furthermore, the temporal resolution often prevents the observation of flash floods. SAR data using the microwave wavelengths of the electromagnetic spectrum can help overcome these problems by providing additional imagery during the night or during cloud cover. Adding night-time and cloud-cover images will provide a higher total temporal resolution. Flood extents are currently determined with statistical methods using thresholds to subsequently identify flood extents, e.g. by using Bayesian method on SAR amplitude time-series data (Lin et al., 2019). Advanced machine learning classification methods are being developed to improve this process, but in order to train them it is necessary to have manually labelled images as training data. Collection of this labelled flood extent information is the main challenge for automatic detection moving forward. Manual methods could harness the power of the crowd, as people are connected through the internet or with mapathons. These approaches could have game-changing implications for the training of machine learning algorithms. Already mapathons are often “trainathons”, where mappers are not only manual digitizers but also labellers and

trainers for automated machine learning methods for the future.

2.2.2 Predictive hazard assessments

Predictive flood hazard assessments consist of predicting future floods and their characteristics such as extents, inundation depths, durations, flow velocities, waves and water levels in rivers or seas. These assessments are applied for short-term forecasting in the preparation phase (preparing for imminent events) and long-term risk analyses for use in flood risk management (mitigation phase).

In flood forecasting, traditional methods of predicting hazard variables can involve a chain of hydrologic and hydraulic models that describe the physical processes. Although such models provide system understanding, they often have high computational and data requirements. Therefore, the use of process models may not always be feasible or necessary in the preparation stage of a disaster. At that moment, accurate and timely outputs become more important than system understanding, and the use of “black-box” machine learning models (e.g. Campolo et al., 2003) is becoming more widespread (Mosavi et al., 2018). The increased speed can create a trade-off with the robustness of forecast models, as changes to the hydraulic system (such as a new structure that could be easily implemented into a hydraulic model) cannot be directly introduced into a trained machine learning model. In addition, machine learning models might not perform well in predicting extremes far outside past observations, since they have not been trained against such extremes.

A review of flood forecasting methods using machine learning by Mosavi et al. (2018) highlights trends such as component and ensemble models (collectively termed “hybrid models”; Corzo and Solomatine, 2014). Hybrid component models assign machine learning a specific task in the modelling process that is either highly complex or not well understood. Examples of this include using machine learning for error correctors (see, for example, studies by Abrahart and See, 2007, and Google Research – Nevo et al., 2019) or flows subject to human influence (Yaseen et al., 2019). Hybrid ensemble methods often use machine learning models to supplement process models, providing robust predictions and uncertainty ranges (Solomatine and Ostfeld, 2008). Such methods benefit from the speed and ability to deal with non-linear multivariable problems of machine learning modelling and the process understanding available in conventional modelling. The review by Mosavi et al. (2018) does not consider gridded/spatial forecasting techniques, but advanced machine learning techniques are starting to be developed for precipitation pattern nowcasting (Xingjian et al., 2015) and flood extents prediction (Chang et al., 2018). Another application of machine learning in the preparation phase is in the real-time control of flood defences and systems (e.g. Lobbrecht and Solomatine, 2002; Castelletti et al., 2010). For example, Lobbrecht and Solomatine (2002) used

machine learning methods to optimize control decisions in the event of communication network breakdowns during extreme storm events.

Another major application for machine learning in long-term risk analysis is “surrogate” modelling (Ong et al., 2003), in which the outputs from process models are used to train computationally less-intensive machine learning models. This can be applied to speed up different types of process models applied in predictive hazard modelling. For example, in flood defence analysis and design, classical reliability techniques such as the first-order reliability method (FORM) and Monte Carlo simulations (Steenbergen et al., 2004), or large-scale risk analyses that utilize them (Curran et al., 2019), can be replicated using a relatively small amount of evaluations as samples (Chojaczyk et al., 2015; Kingston et al., 2011). However, surrogate models may be particularly susceptible to extrapolation problems, where input data outside the range of the training data are introduced (Ghalkhani et al., 2013).

In the mitigation phase a chain of hydrologic and hydraulic models that describe the physical processes is typically applied (e.g. Wagenaar et al., 2019). In general, system understanding is required to assess proposed or potential future changes. In such cases, data-driven approaches are typically not applicable as there are no data about how the system behaves after the changes occur, and hence simulation models are required that describe the physical system.

2.3 Flood impact assessment

2.3.1 Descriptive impact assessments

Descriptive impact assessments consist of making estimates of the flood impact after or during an event. This is traditionally done with manually collected data from observers on the ground. However, such manual ground inspections are slow and require people to enter the disaster area. Remote sensing can be used to get a very quick first impression of the damage to help with disaster response. Such techniques have already been applied, for earthquake and wind damage (e.g. Menderes et al., 2015). For flooding, this is often more difficult because damage inside buildings is difficult to obtain either from aerial or space-based sensors. Only when buildings completely collapse or are removed by strong flows does remote sensing become feasible. This is, for example, the case with flash floods, tsunamis or some storm surges. If 360° street view images are collected after a flood, these could potentially be used for damage assessment. Machine learning techniques could then eventually be used to give a quick first estimate of the damage.

The use of machine learning techniques for automatic detection of damages from remote sensing information (aerial or street view) requires labelled training data from manually collected data from observers on the ground. These data are currently rare. An approach could be to start using re-

mote sensing data to manually label the impact. A way to get around this limitation is to detect changes pre- and post-flood using high-resolution satellite images for urban areas where many buildings are damaged. Pixels with changed information will denote the damage that happened due to the floods. Eventually these data can then be used as training data for cases where only the post-flood images are available within a short time interval after the flood event. This method would however only be relevant to catastrophic floods because it does not address the fact that most damage remains not observable from top view. On top of that, this approach introduces significant new error: (1) error in the change detection signal, (2) error in relating the change to damage and (3) error in training a new model based on those damage labels. Imagery from different angles (e.g. from street view or drones) might be more useful for change detection; however these data would also be more difficult to acquire.

2.3.2 Predictive impact assessments

Predictive flood impact assessments include models that translate hazard and exposure information into socio-economic impacts of the flood. This can include information such as monetary flood damage, casualties, buildings damaged, crop damage, disease outbreak, building materials needed, recovery time, health monitoring of key structures and indirect damage (damages that occur in a different spatial and/or temporal setting than the originating event).

Most predictive flood damage modelling relies on depth-damage functions that describe a relationship between the water depth and monetary flood damage (Merz et al., 2010). They are based either on historical flood damage records (e.g. Thielen et al., 2008; Kreibich et al., 2010) or on expert estimates (e.g. Penning-Rowsell et al., 2005). In practice, many more variables than water depth have an influence on the flood damage (Cammerer et al., 2013; Wagenaar et al., 2016). Therefore, in the scientific literature there has been a move towards multivariable flood damage models that use many variables (e.g. flood duration, velocity, building materials, socio-economic status of inhabitants etc.) instead of just water depth (e.g. Merz et al., 2013; Spekkers et al., 2014; Chinh et al., 2015; Kreibich et al., 2017; Wagenaar et al., 2017; Carisi et al., 2018; Amadio et al., 2019). These models are based on data and machine learning. The problem lies with insufficient data availability to train machine learning models and with the fact that using the models requires a lot of feature data about flood and building characteristics plus socio-economic data about inhabitants (Wagenaar et al., 2017). In the future we expect more data about features to become available from computer vision applied to street view, satellite or drone images (see descriptive exposure section). This would improve the quality of such models, could make it easier to apply them and make the development possible for more areas.

Machine learning could also be applied to predict disease outbreak after floods by combining remote sensing, meteorological and socio-economic data (e.g. Mayfield et al., 2018; Carvajal et al., 2018; Modu et al., 2017; Yomwan et al., 2015; Tiwari et al., 2013; Shively et al., 2015). In a flood event, there is an increased risk of infectious diseases among survivors and displaced persons. For example, measles, diarrhea, acute respiratory infections and malaria can be responsible for many deaths (Lignon, 2006). Predictive modelling of such diseases is rarely carried out, and current approaches mostly focus on simple regression models or process models that simulate the spread of pollutants in the water. One major challenge is that the degree to which such epidemics occur is associated with the regional endemicity of specific diseases, the nature and scope of the disaster, the level of public health infrastructure in place both before and after the event, and the level and efficacy of disaster response (Ivers and Ryan, 2006). Machine learning models could take such complex processes better into account.

Machine learning can be used for structural health monitoring; this has applications in the preparation phase (Pyayt et al., 2014; Jonkman et al., 2018) and in the long-term reliability analysis required in the mitigation phase (Prendergast et al., 2018; Klerk et al., 2019). In the preparation phase for a flood, structural health monitoring is often done by manual inspections of the infrastructure on the ground. For example, in the Netherlands there is a large network of volunteers that can be activated in the event of high river levels to inspect the dikes. In the mitigation phase this is done by geotechnical process models fed by observations from the ground (e.g. De Waal, 2016); this is for example applied to decide on dike strengthening. Machine learning algorithms can help detect damage patterns from sensor data and are currently being used for the monitoring of flood defence structures such as dikes (Pyayt et al., 2011). Similar methods have also been applied to bridges (Neves et al., 2017). The use of both machine learning algorithms and traditional techniques for damage detection during floods is still very scarce (Prendergast et al., 2018; Pyayt et al., 2011); however, integration of structural health monitoring with flood early-warning systems is a very promising field of development for machine learning techniques but would also require training data.

Indirect damages and business interruption are often taken into account simply through a scaling factor of the direct damage (e.g. Wagenaar et al., 2019). More complex models for quantifying such damages include input-output models and general equilibrium models (e.g. Koks et al., 2016). To quantify indirect damages, such as business interruption losses, estimating the time it will take for different assets to be back in full or partial functionality is required. These post-disaster restoration models have started to be formalized in the last few years, primarily focused on earthquake disasters (Kang et al., 2018; Burton et al., 2018). Due to a lack of gathered empirical data on post-disaster recovery, the use of data-intensive machine learning techniques has not yet made an

impact on this discipline. However, the need to probabilistically quantify recovery will require the use of statistical models for calibration or assessments of recovery times, and that might be possible in the near future with the use of new remote sensing and crowd-sourcing technologies to obtain the empirical feature data needed.

3 General perspectives

3.1 Data limitations

Many machine learning applications in flood risk and impact modelling appear to be limited by a lack of data, especially training data needed to build effective machine learning models. This is especially true since the field of flood risk analysis is concerned primarily with extreme events, which are rare, and data-collection during such events is often logistically difficult. The increase in the amount of data around the world does not necessarily imply that this problem will be resolved in the future. Some data are simply not collected, or there are measurement definition or quality issues. To fulfil the potential of machine learning, new data collection efforts will be required, along with data standardization protocols. This will necessitate collaboration between different organizations and stakeholders, setting of data standards and a willingness to share. This problem is common to impact data collected (see Sect. 2.3.1 and 2.3.2), labelled flood extent data (see Sect. 2.2.1), social media hazard data (see Sect. 2.2.1) and first-floor elevation data (see Sect. 2.1.1).

3.2 Transferability of data

A critical assumption behind machine learning techniques is that the data being used to train a model are representative of the situation the model needs to be applied in. For example, a data set on damage to concrete buildings is not fully applicable to modelling the damage to thatched huts. It is therefore important to collect heterogeneous data sets that cover a large spectrum of potential situations (Wagenaar et al., 2018). Data that are not fully applicable can still have some value, for example through domain adaptation or transfer learning (GFDRR, 2018), but applicable data are always required as well. Wagenaar et al. (2020) showed that sample selection bias correction, a form of domain adaptation, helps to improve machine learning impact models in a transfer setting. Furthermore, it is important to work on efficient ways to communicate the applicability of data-driven models.

3.3 Ethics and bias

Significant attention is currently being given to questions of the ethics and bias of machine learning systems across a variety of domains, including facial recognition (Keyes, 2018), automated weaponry (Suchman and Weber, 2016), criminal justice (Eubanks, 2018) and search engines (Noble, 2018).

A number of technology companies and research institutions have developed guidelines for evaluating machine learning systems, but this work is still evolving. Despite similar potential for negative impacts of these tools in flood risk management (Soden et al., 2019), the community has not given these issues as much attention. Such concerns include the potential for reinforcing existing social inequalities and the reduced role of human judgement in modelling processes. These are risks that need to be weighed seriously against the potential benefits of machine learning and explored in greater detail.

Biases in machine learning can occur because of data sets that, for a number of reasons, do not fully represent the phenomena which they are meant to describe (e.g. people are accidentally excluded). For example, we often measure what we have data for, rather than measuring what matters most, or use training data sets that reinforce past problems. For example, if certain settlements are not detected in exposure maps, because they use different construction practices than the settlements used in training data sets, they may not receive emergency aid in the event of a flood. These problems can be mitigated by ensuring modellers understand the context of what they are attempting to model. Other ethical issues raised by machine learning in the flood management context include data ownership, transparency, consent and privacy. For example, some people may object to having their home labelled “vulnerable” on a vulnerability map used by first responders. Privacy concerns may be aggravated by machine learning and other big-data techniques. Ethics problems should be addressed by carefully weighing the benefits of collecting certain data against the related privacy costs, in collaboration with people who may be affected by the outcomes of decisions based on machine learning tools.

An additional ethical concern regarding machine learning in flood risk assessment is misuse of models. In some sectors great advances have been made with machine learning (e.g. facial recognition and self-driving cars). This success for some tasks can lead to an awe-inspiring general attitude towards the techniques (Ames, 2018; Narayanan, 2019). This hype sometimes leads to unwarranted trust in the techniques for tasks machine learning is not (yet) suitable for (Narayanan, 2019). For example, many companies are currently using machine learning for hiring decisions despite well-documented failings of these tools (Narayanan, 2019; Raghavan et al., 2019). In order to avoid such misuse in flood risk assessment, it is important that machine learning implementations are transparent and supervised by independent machine learning and flood risk assessment experts.

Importantly, flood risk assessments are highly data reliant, and the increased attention to questions of ethics and bias in machine learning systems might serve as an opportunity to drive conversations in our field about the limits of disaster data more broadly. Many of the sources of bias or ethical concerns in machine learning systems originate in, or share common roots with, other kinds of data used to understand disaster risks. This includes issues such

Table 2. Future predictions.

	Predictive	Descriptive
Exposure	Likely incremental changes, e.g. improved cellular automata transition rules	Very likely significant changes, e.g. automatic exposure detection including building features
Hazard	Diverse field; changes are more likely to be complementary or specific components of modelling	Likely changes in detection due to remote sensing and social media algorithms.
Impact	Potential for significant changes (i.e. multi-variable data-driven methods)	Significant changes likely for some elements; others will probably remain the same

as (1) property values driving what areas get protected, (2) privacy concerns (which may be aggravated by machine learning and other big-data techniques), (3) how the lack of gender/age/ethnicity-disaggregated data on disaster risk masks differential vulnerabilities and (4) the importance of public participation and the voice of residents of areas portrayed by models as “at risk”. Detailed analyses of specific cases (e.g. Soden and Kauffman, 2019) are urgently needed to make further progress in understanding the consequences of the assessment methods we use to understand disasters.

3.4 Future predictions

In the following section we draw some general conclusions about how machine learning will change flood risk and impact assessments. Table 2 provides an overview of these predictions.

3.4.1 Very likely changes

A few of the trends seem inevitable, primarily in cases where recent technological advances or data that recently became available make next steps obvious. A good example of this is the automatic detection of building footprints and roads from high-resolution remote sensing imagery (see Sect. 2.1.1). This is already possible and will, especially in data-poor areas, drastically improve the quality of the first response and risk calculation. Further advances in the use of machine learning in descriptive hazard assessment through social media are also inevitable (see Sect. 2.2.1), given the amount of data available to social media companies.

3.4.2 Likely and potential changes

This is the category that can be shaped the most by individual innovators, and the majority of the advances discussed in this paper fall under this category. In this case, the innovation still experiences some kind of obstacle that prevents widespread application. It is typically difficult to predict whether such obstacles can be truly removed in the future and how long that will take. Because the field of flood risk and impact assessments is relatively small, the obstacles are often eco-

nomically feasible, which is difficult to assess, combined with conservative users. An example of this is the large-scale collection of impact data which is required for both descriptive and predictive impact modelling (see Sect. 2.3.1 and 2.3.3) or the training data required for descriptive hazard assessments (see Sect. 2.2.1). Sometimes the obstacle is also technical feasibility, for example whether it will really be possible to extract first-floor elevation levels from street view (see Sect. 2.1.2). Innovations are also interdependent; for example, when building feature information can be automatically extracted from street view, impact models will become easier to train and easier to run, and it will make more sense to start collecting the required impact data.

3.4.3 Unlikely changes

For some processes, machine learning may not be the best solution from a theoretical perspective. For example, the processes of how water flows are very well known and can be well approximated with existing equations. It, therefore, does not always make sense to pick a machine learning approach. Another situation in which machine learning is not applicable is when a system is being modelled on which predictions need to be made that cannot have been seen in the data or when we know from an exploratory data analysis that we have no data for it (GFDRR, 2018), for example how a system may behave under never-seen discharges or after new infrastructure has been built (e.g. new dam in the river). In these cases, machine learning may play a role in some components of the model, but process models will very likely remain crucial in simulating the never-before-seen conditions. Especially for predictive hazard models (see Sect. 2.2.2), there are many elements that are unlikely to change with the advance of machine learning.

3.4.4 New practices in flood risk and impact assessments

Most change to flood risk and impact assessments discussed in this paper relate to better models. Such cheaper, faster and more accurate models could possibly yield new practices in

flood risk and impact assessments. Cheaper models would make flood risk and impact assessments feasible to carry out for a larger group of users and are therefore likely to make emergency aid and investments in mitigation measures more efficient. Faster methods may speed up emergency response and recovery, especially when manually collected data from observers on the ground are replaced by remote earth observation. More accurate models may lead to more early actions being feasible (Coughlan de Perez et al., 2016), and hence early actions can be carried out that could not be carried out before, for example more targeted measures during the preparation and response phase of a flood. Such new measures include providing emergency payouts even before the event to the most vulnerable people (e.g. Reuters, 2019), prioritization of emergency measures in buildings, targeted disease outbreak prevention (Coughlan de Perez et al., 2016), early shipping of the right emergency goods (Coughlan de Perez et al., 2016) and prioritization of early harvesting of crops.

Data availability. No data sets were used in this article.

Author contributions. This article was the result of an intensive 2-week-long collaboration during the UR Field Lab in Chiang Mai of June 2019. All authors attended this field lab, where the outline and structure of the paper were agreed upon. DW and AC collected and edited the input from all authors. Specific attention was given to each section by the following authors; Sect. 1 – DW, AC, MB and DL; Sect. 2 – DW, AC, MB, AB, EH, GMS, LR, GM; Sect. 3 – DW, AC, RS, DL. The authors' comments in the discussion were written by DW, AC and DL. All authors reviewed the paper before submission.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This invited perspective paper benefited from the unique intellectual environment created and facilitated through the Understanding Risk Field Lab on urban flooding in Chiang Mai, Thailand, in June 2019 (<https://urfieldlab.com/>, last access: 23 April 2020). We would like to thank the organizers of the event (Robert Soden, David Lallemand, Perrine Hamel, Katherine Barnes and Giuseppe Molinaro) for providing the unique setting to make this paper possible. We would also like to thank the other participants, who provided valuable input on this paper by participating in some of the discussions, namely Gautam Dadhich, Carmen Acosta, Maricar Rabonza, Pamela Cajilig, Rahul Sharma and Wahaj Habib. Furthermore, we would like to thank the editor (Heidi Kreibich) and the reviewers (Fernando Nardi and an anonymous reviewer) for their contributions to the paper.

Financial support. This research has been supported in part by Deltares and the National Research Foundation, Prime Minister's Office, Singapore, under the NRF-NRFF2018-06 award.

Review statement. This paper was edited by Heidi Kreibich and reviewed by Fernando Nardi and one anonymous referee.

References

- Aarathi, A. D. and Gnanappazham, L.: Comparison of Urban Growth Modeling Using Deep Belief and Neural Network Based Cellular Automata Model – A Case Study of Chennai Metropolitan Area, Tamil Nadu, India, *Journal of Geographic Information System*, 11, 1–16, 2019.
- Abrahart, R. J. and See, L. M.: Neural network modelling of non-linear hydrological relationships, *Hydrol. Earth Syst. Sci.*, 11, 1563–1579, <https://doi.org/10.5194/hess-11-1563-2007>, 2007.
- Alshehhi, R., Marpu, P. R., Woon, W., and Dalla Maru, M.: Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks, *ISPRS J. Photogramm.*, 130, 139–149, 2017.
- Amadio, M., Scorzini, A. R., Carisi, F., Essfelder, A. H., Domeneghetti, A., Mysiak, J., and Castellarin, A.: Testing empirical and synthetic flood damage models: the case of Italy, *Nat. Hazards Earth Syst. Sci.*, 19, 661–678, <https://doi.org/10.5194/nhess-19-661-2019>, 2019.
- Ames, M. G.: Deconstructing the algorithmic sublime, *Big Data & Society*, 5, 1–4, <https://doi.org/10.1177/2053951718779194>, 2018.
- Annis, A. and Nardi, F.: Integrating VGI and 2D hydraulic models into a data assimilation framework for real time flood forecasting and mapping, *Geo-spatial Information Science*, 22, 223–236, <https://doi.org/10.1080/10095020.2019.1626135>, 2019.
- Assumpção, T. H., Popescu, I., Jonoski, A., and Solomatine, D. P.: Citizen observations contributing to flood modelling: opportunities and challenges, *Hydrol. Earth Syst. Sci.*, 22, 1473–1489, <https://doi.org/10.5194/hess-22-1473-2018>, 2018.
- Bischke, B., Helber, P., Folz, J., Borth, D., and Dengel, A.: Multi-Task Learning for Segmentation of Building Footprints with Deep Neural Networks, available at: <https://arxiv.org/abs/1709.05932> (last access: 28 April 2020), 2017.
- Bishop, C. M.: *Pattern Recognition and Machine Learning*, Springer, Cambridge, UK, ISBN 978-0-387-31073-2, 2006.
- Bouwer, L. M., Haasnoot, M., Wagenaar, D., and Roscoe, K.: Assessment of alternative flood mitigation strategies for the C-7 Basin in Miami, Florida, Deltares, Delft, the Netherlands, 1230718, 2017.
- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J.: *Classification and regression trees*, Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, USA, ISBN 978-0-412-04841-8, 1984.
- Burton, H. V., Miles, S. B., and Kang, H.: Integrating Performance-Based Engineering and Urban Simulation to Model Post-Earthquake Housing Recovery, *Earthq. Spectra*, 34, 1763–1785, <https://doi.org/10.1193/041017EQS067M>, 2018.
- Cammerer, H., Thielen, A. H., and Lammel, J.: Adaptability and transferability of flood loss functions in residen-

- tial areas, *Nat. Hazards Earth Syst. Sci.*, 13, 3063–3081, <https://doi.org/10.5194/nhess-13-3063-2013>, 2013.
- Campolo, M., Soldati, A., and Andreussi, P.: Artificial neural network approach to flood forecasting in the River Arno, *Hydrolog. Sci. J.*, 48, 381–398, <https://doi.org/10.1623/hysj.48.3.381.45286>, 2003.
- Carisi, F., Schröter, K., Domeneghetti, A., Kreibich, H., and Castellarin, A.: Development and assessment of uni- and multivariable flood loss models for Emilia-Romagna (Italy), *Nat. Hazards Earth Syst. Sci.*, 18, 2057–2079, <https://doi.org/10.5194/nhess-18-2057-2018>, 2018.
- Carvajal, T. M., Viacrusis, K. M., Hernandez, L. F. T., Ho, H. T., Amalin, D. M., and Watanabe, K.: Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines, *BMC Infect. Dis.*, 18, p. 183, 2018.
- Castelletti, A., Galelli, S., Restelli, M., and Soncini-Sessa, R.: Tree-based reinforcement learning for optimal water reservoir operation, *Water Resour. Res.*, 46, W09507, <https://doi.org/10.1029/2009WR008898>, 2010.
- Chang, L., Amin, M. Z., Yang, S. N., and Chang, F.: Building ANN-Based Regional Multi-Step-Ahead Flood Inundation Forecast Models, *Water*, 10, 1283, <https://doi.org/10.3390/w10091283>, 2018.
- Chinh, D., Gain, A., Dung, N., Haase, D., and Kreibich, H.: Multi-Variate Analyses of Flood Loss in Can Tho City, Mekong Delta, *Water*, 8, 6, <https://doi.org/10.3390/w8010006>, 2015.
- Chojaczyk, A., Teixeira, A. P., Neves, L. C., Cardoso, J. B., and Guedes Soares C.: Review and application of Artificial Neural Networks models in reliability analysis of steel structures, *Struct. Saf.*, 52, 78–89, 2015.
- Corzo, P. G. A. and Solomatine, D.: Comparative analysis of conceptual models with error correction, artificial neural networks and committee models, EGU General Assembly 2014, 27 April–2 May 2014, Vienna, Austria, 2014.
- Coughlan de Perez, E., van den Hurk, B., van Aalst, M. K., Amuron, I., Bamanya, D., Hauser, T., Jongma, B., Lopez, A., Mason, S., Mendler de Suarez, J., Pappenberger, F., Rueth, A., Stephens, E., Suarez, P., Wagemaker, J., and Zsoter, E.: Action-based flood forecasting for triggering humanitarian action, *Hydrol. Earth Syst. Sci.*, 20, 3549–3560, <https://doi.org/10.5194/hess-20-3549-2016>, 2016.
- Curran, A., de Bruijn, K. M., Klerk, W. J., and Kok, M.: Large Scale Flood Hazard Analysis by Including Defence Failures on the Dutch River System, *Water*, 11, 1732, <https://doi.org/10.3390/w11081732>, 2019.
- De Waal, J. P.: Basisrapport WBI 2017, Deltares 1230086-002, Delft, the Netherlands, 2016.
- Dibike, Y. B. and Solomatine, D. P.: River flow forecasting using artificial neural networks, *Phys. Chem. Earth Pt. B*, 26, 1–7, 2001.
- Eilander, D., Trambauer, P., Wagemaker, J., and Van Loenen, A.: Harvesting social media for generation of near real-time flood maps, 12th International Conference on Hydroinformatics, HIC, 21 August 2016, Incheon, South Korea, 2016.
- Eubanks, V.: Automating inequality: How high-tech tools profile, police, and punish the poor, St. Martin's Press, New York, USA, 2018.
- Fohringer, J., Dransch, D., Kreibich, H., and Schröter, K.: Social media as an information source for rapid flood inundation mapping, *Nat. Hazards Earth Syst. Sci.*, 15, 2725–2738, <https://doi.org/10.5194/nhess-15-2725-2015>, 2015.
- Gao, X., Klaiber, C., Patel, D., and Underwood, J.: AI is supercharging the creation of maps around the world, Tech@Facebook, available at: <https://tech.fb.com/ai-is-supercharging-the-creation-of-maps-around-the-world/>, last access: 21 August 2019.
- Gauss, C. F.: *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, sumtibus Perthes, F. and Besser, I. H., Hamburg, Germany, <https://doi.org/10.3931/e-rara-522>, 1809.
- GFDRR: Machine Learning for Disaster Risk Management, GFDRR, Washington, D.C., USA, 2018.
- Ghalkhani, H., Golian, S., Saghafian, B., Farokhnia, A., and Shamseldin, A.: Application of surrogate artificial intelligent models for real-time flood routing, *Water Environ. J.*, 27, <https://doi.org/10.1111/j.1747-6593.2012.00344.x>, 2013.
- Giacinto, G. and Roli, F.: Design of effective neural network ensembles for image classification purposes, *Image Vis. Comput.*, 19, 699–707, [https://doi.org/10.1016/S0262-8856\(01\)00045-2](https://doi.org/10.1016/S0262-8856(01)00045-2), 2001.
- Goldblatt, R., You, W., Hanson, G., and Khandelwal, A.: Detecting the boundaries of urban areas in india: A dataset for pixel-based image classification in google earth engine, *Remote Sens.*, 8, 634, <https://doi.org/10.3390/rs8080634>, 2016.
- Goldblatt, R., Stuhlmacher, M. F., Tellman, B., Clinton, N., Hanson, G., Georgescu, M., Wang, C., Serrano-Candela, F., Khandelwal, A. K., Cheng, W., and Balling, R.: Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover, *Remote Sens. Environ.*, 205, 253–275, 2018.
- Heermann, P. D. and Khazenie, N.: Classification of multispectral remote sensing data using a back-propagation neural network, *IEEE T. Geosci. Remote*, 30, 81–88, 1992.
- Ivers, L. C. and Ryan, E. T.: Infectious diseases of severe weather-related and flood-related natural disasters, *Curr. Opin. Infect. Dis.*, 19, 408–414, 2006.
- Jonkman, S. N.: Global Perspectives on Loss of Human Life Caused by Floods, *Nat. Hazards*, 34, 151–175, 2005.
- Jonkman, S. N., Voortman, H. G., Klerk, W. J., and van Vuren, S.: Developments in the management of flood defences and hydraulic infrastructure in the Netherlands, *Struct. Infrastruct. Eng.*, 14, 895–910, 2018.
- Kang, H., Burton, H., and Miao, H.: Replicating the Recovery following the 2014 South Napa Earthquake using Stochastic Process Models, *Earthq. Spectra*, 34, 1247–1266, <https://doi.org/10.1193/012917EQS020M>, 2018.
- Keyes, O.: The misgendering machines: Trans/HCI implications of automatic gender recognition, *Proceedings of the ACM on Human-Computer Interaction*, 2, 88, <https://doi.org/10.1145/3274357>, 2018.
- Khan, A., Khan, H., and Vasilescu, L.: Disaster Management CYCLE – a theoretical approach, *Management and Marketing Journal*, 6, 43–50, 2008.
- Kind, J., Botzen, W. J., and Aerts, C. J. H.: Accounting for risk aversion, income distribution and social welfare in cost-benefit analysis for flood risk management, *WIREs Clim. Change* 2016, 8, e446, <https://doi.org/10.1002/wcc.446>, 2016.
- Kingston, G. B., Rajabalinejad, M., Gouldby, B. P., and Van Gelder, P. H. A. J. M.: Computational intelligence methods for the ef-

- ficient reliability analysis of complex flood defence structures, *Struct. Saf.*, 33, 64–73, 2011.
- Klemas, V.: Remote Sensing of Floods and Flood-Prone Areas: An Overview, *J. Coastal Res.*, 31, 1005–1013, 2015.
- Klerk, W., Schweckendiek, T., Den Heijer, F., and Kok, M.: Value of information of Structural Health Monitoring in Asset Management of Flood Defences, *Infrastructures*, 4, 56, <https://doi.org/10.3390/infrastructures4030056>, 2019.
- Koks, E. E., Carrera, L., Jonkeren, O., Aerts, J. C. J. H., Husby, T. G., Thissen, M., Standardi, G., and Mysiak, J.: Regional disaster impact analysis: comparing input–output and computable general equilibrium models, *Nat. Hazards Earth Syst. Sci.*, 16, 1911–1924, <https://doi.org/10.5194/nhess-16-1911-2016>, 2016.
- Kreibich, H., Seifert, I., Merz, B., and Thielen, A.: Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector, *Hydrolog. Sci. J.*, 55, 1302–1313, 2010.
- Kreibich, H., Botto, A., Merz, B., and Schröter, K.: Probabilistic, Multivariable Flood Loss Modeling on the Mesoscale with BT-FLEMO, *Risk Anal.*, 37, 774–787, <https://doi.org/10.1111/risa.12650>, 2017.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E.: ImageNet Classification with Deep Convolutional Neural Networks, *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 3–6 December 2012, Lake Tahoe, Nevada, USA, 2012.
- Kron, W.: Flood Risk = Hazard × Exposure × Vulnerability, in: *Flood Defence*, edited by: Wu, B. S., Wang, Z. Y., Wang, G. Q., Huang, G. H., Fang, H. W., and Huang, J. C., Science Press, New York, USA, 82–97, 2002.
- Kundzewicz, Z. W., Kanae, S., Seneviratne, S. L., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, R., Kron, W., Benito, G., Honda, Y., Takahashi, K., and Sherstyukov, B.: Flood risk and climate change: global and regional perspectives, *Hydrolog. Sci. J.*, 59, 1–28, <https://doi.org/10.1080/02626667.2013.857411>, 2014.
- Legendre, A. M.: Nouvelles méthodes pour la détermination des orbites des comètes, *Sur la Méthode des moindres carrés*, Firmin Didot, Paris, France, 1805.
- Lignon, B. L.: Infectious Diseases that Pose Specific Challenges After Natural Disasters: A Review, *Seminars in Pediatric Infectious Diseases*, 17, 36–45, <https://doi.org/10.1053/j.spid.2006.01.002>, 2006.
- Lin, Y. N., Yun, S., Bhardwaj, A., and Hill, E. M.: Urban Flood Detection with Sentinel-1 Multi-Temporal Synthetic Aperture Radar (SAR) Observations in a Bayesian Framework: A Case Study for Hurricane Matthew, *Remote Sens.*, 11, 1778, <https://doi.org/10.3390/rs11151778>, 2019.
- Lobrecht A. and Solomatine, D.: Machine Learning in Real-Time Control of Water Systems, *Urban Water* 4, 283–289, 2002.
- Lopez-Fuentes, L., Van de Weijer, J., Bolaños, M., and Skinnemoen, H.: Multi-modal Deep Learning Approach for Flood Detection, *MediaEval'17*, 13–15 September 2017, Dublin, Ireland, 2017.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Hung Byers, A.: Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute, New York City, New York, USA, 2011.
- Mayfield, H. J., Smith, C. S., Lowry, J. H., Watson, C. H., Baker, M. G., Kama, M., Nilles, E. J., and Lau, C. L.: Predictive risk mapping of an environmentally-driven infectious disease using spatial Bayesian networks: A case study of leptospirosis in Fiji, *PLoS Neglect. Trop. D.*, 12, e0006857, <https://doi.org/10.1371/journal.pntd.0006857>, 2018.
- Mazzoleni, M., Verlaan, M., Alfonso, L., Monego, M., Norbiato, D., Ferri, M., and Solomatine, D. P.: Can assimilation of crowdsourced data in hydrological modelling improve flood prediction?, *Hydrol. Earth Syst. Sci.*, 21, 839–861, <https://doi.org/10.5194/hess-21-839-2017>, 2017.
- Menderes, A., Erener, A., and Sarp, G.: Automatic Detection of Damaged Buildings after Earthquake Hazard by Using Remote Sensing and Information Technologies, *Proced. Earth Plan. Sci.*, 15, 257–262, <https://doi.org/10.1016/j.proeps.2015.08.063>, 2015.
- Merz, B., Kreibich, H., Schwarze, R., and Thielen, A.: Review article “Assessment of economic flood damage”, *Nat. Hazards Earth Syst. Sci.*, 10, 1697–1724, <https://doi.org/10.5194/nhess-10-1697-2010>, 2010.
- Merz, B., Kreibich, H., and Lall, U.: Multi-variate flood damage assessment: a tree-based data-mining approach, *Nat. Hazards Earth Syst. Sci.*, 13, 53–64, <https://doi.org/10.5194/nhess-13-53-2013>, 2013.
- Mestav Sarica, G., Zhu, T., and Pan, T.-C.: Flood Exposure of Shenzhen from Past to Future: A Spatio-Temporal Approach using Urban Growth Modeling, *Proceedings of the 7th Annual International Conference on Architecture and Civil Engineering*, 27–28 May 2019, Singapore, 400–405, 2019.
- Modu, B., Polovina, N., Lan, Y., Konur, S., Asyhari, A., and Peng, Y.: Towards a Predictive Analytics-Based Intelligent Malaria Outbreak Warning System, *Appl. Sci.*, 7, 836, <https://doi.org/10.3390/app7080836>, 2017.
- Mosavi, A., Ozturk, P., and Kwok-wing, C.: Review: Flood Prediction Using Machine Learning Models: Literature Review, *Water*, 10, 1536, <https://doi.org/10.3390/w10111536>, 2018.
- Naghbi, F., Delavar, M. R., and Pijanowski, B.: Urban Growth Modeling Using Cellular Automata with Multi-Temporal Remote Sensing Images Calibrated by the Artificial Bee Colony Optimization Algorithm, *Sensors*, 16, 2122, <https://doi.org/10.3390/s16122122>, 2016.
- Narayanan, A.: How to recognize AI snake oil, available at: <https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf> (last access: 27 January 2020), 2019.
- National Research Council: Introduction Facing hazards and disasters: Understanding human dimensions, The National Academies Press, Washington, D.C., USA, <https://doi.org/10.17226/11671>, 2006.
- Neuhold, G., Ollmann, T., Rota Bulò, S., and Kotschieder, P.: The Mapillary Vistas for Semantic Understanding of Street Scenes, *International Conf. on Computer Vision (ICCV)*, 22–29 October 2017, Venice, Italy, 2017.
- Neves, C., González, I., Leander, J., and Karoumi, R.: Structural health monitoring of bridges: a model-free ANN-based approach to damage detection, *J. Civ. Struct. Heal. Monit.*, 7, 689–702, 2017.
- Nevo, S., Anisimov, V., El-Yaniv, R., Giencke, P., Gigi, Y., Hasidim, A., Mushe, Z., Schlesinger, M., Shalev, G., Tirumali, A., Wiesel, A., Zlydenko, O., and Matias, Y.: Machine Learning for Flood Forecasting at Scale, *32nd Conference on Neural Infor-*

- mation Processing Systems (NIPS 2018), 3–8 December 2019, Montréal, Canada, 2019.
- Noble, S. U.: Algorithms of oppression: How search engines reinforce racism, nyu Press, New York City, New York, USA, ISBN 9781479837243, 2018.
- Olivas, E. S., Guerrero, J. D., Martinez-Sober, M., Magdalena-Benedito, J. R., and Serrano López, A. J.: Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, IGI Global, Hershey, PA, USA, 2, 1–852, <https://doi.org/10.4018/978-1-60566-766-9>, 2010.
- Ong, Y. S., Nair, P., and Keane, A. J.: Evolutionary Optimization of Computationally Expensive Problems via Surrogate Modeling, *AIAA Journal*, 41, 4, <https://doi.org/10.2514/2.1999>, 2003.
- Pal, S. and Ghosh, S. K.: Rule based End-to-End Learning Framework for Urban Growth Prediction, *ArXiv*, abs/1711.10801, available at: <https://arxiv.org/abs/1711.10801> (last access: 28 April 2020), 2017.
- Penning-Rowsell, E. C., Johnson, C., and Tunstall, S.: The benefits of Flood and Coastal Risk Management: A Manual of Assessment Techniques, Middlesex University Press, London, UK, 2005.
- Prendergast, L. J., Limongelli, M. P., Ademovic, N., Anžlin, A., Gavin, K., Zanini, M.: Structural Health Monitoring for Performance Assessment of Bridges under Flooding and Seismic Actions, *Struct. Eng. Int.*, 28, 296–307, 2018.
- Pyayt, A., Mokhov, I., Lang, B., Krzhizhanovskaya, V., and Meijer, R.: Machine learning methods for environmental monitoring and flood Protection, *World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 5, <https://doi.org/10.5281/zenodo.1075060>, 2011.
- Pyayt, A. L., Kozionov, A. P., Mokhov, I. I., Lang, B., Meijer, R. J., Krzhizhanovskaya, V. V., and Sloot, P. M. A.: Time-frequency methods for structural health monitoring, *Sensors*, 14, 5147–73, 2014.
- Raghavan, M., Barocas, S., Kleinberg, J., and Levy, K.: Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices, *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 27–30 January 2020, Barcelona, Spain, 469–481, 2019.
- Reuters: Bangladesh tries new way to aid flood-hit families: cash up front, available at: <https://www.preventionweb.net/news/view/66899> (last access: 24 April 2020), 2019.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by back-propagating errors, *Nature*, 323, 533–536, 1986.
- Saadi, S. and Bensaïbi, M.: Detection of Buildings height using satellite monoscopic image, *2nd European Conference On Earthquake Engineering*, 24–29 August 2014, Istanbul, Turkey, <https://doi.org/10.13140/2.1.4985.6005>, 2014.
- Samardzic-Petrovic, M., Kovacevic, M., Bajat, B., and Dragicevic, S.: Machine Learning Techniques for Modelling Short Term Land-Use Change, *ISPRS Int. J. Geo-Inf.*, 6, 387, <https://doi.org/10.3390/ijgi6120387>, 2017.
- Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F., and Merz, B.: How useful are complex flood damage models?, *Water Resour. Res.*, 50, 3378–3395, <https://doi.org/10.1002/2013WR014396>, 2014.
- Schröter, K., Lüdtke, S., Redweik, R., Meier, J., Bochow, M., Ross, L., Nagel, C., and Kreibich, H.: Flood loss estimation using 3D city models and remote sensing data, *Environ. Modell. Softw.*, 105, 118–131, <https://doi.org/10.1016/j.envsoft.2018.03.032>, 2018.
- Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., and LeCun, Y.: Overfeat: Integrated recognition, localization and detection using convolutional networks, *arXiv:1312.6229*, available at: <https://arxiv.org/abs/1312.6229> (last access: 28 April 2020), 2014.
- Shively, G., Sununtnasuk, C., and Brown, M.: Environmental variability and child growth in Nepal, *Health Place*, 35, 37–51, 2015.
- Soden, R. and Kauffman, N.: Infrastructuring the Imaginary: How Sea-level Rise Comes to Matter in The San Francisco Bay Area, in: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 4–9 May 2019, Glasgow, UK, Paper No.: 286, 1–11, <https://doi.org/10.1145/3290605.3300516>, 2019.
- Soden, R., Wagenaar, D., Luo, D., and Tijssen, A.: Taking Ethics, Fairness, and Bias Seriously in Machine Learning for Disaster Risk Management, *Workshop Paper, NeurIPS 2019 Workshop on Machine Learning for the Developing World*, 8–14 December 2019, Vancouver, Canada, 2019.
- Solomatine, D. P. and Ostfield, A.: Data-driven modelling: some past experiences and new approaches, *J. Hydroinform.*, 10, 3–22, 2008.
- Song, X., Sexton, J. O., Huang, C., Channan, S., and Townshend, J. R.: Characterizing the magnitude, timing and duration of urban growth from time series of Landsat-based estimates of impervious cover, *Remote Sens. Environ.*, 175, 1–13, <https://doi.org/10.1016/j.rse.2015.12.027>, 2015.
- Spekkers, M. H., Kok, M., Clemens, F. H. L. R., and ten Veldhuis, J. A. E.: Decision-tree analysis of factors influencing rainfall-related building structure and content damage, *Nat. Hazards Earth Syst. Sci.*, 14, 2531–2547, <https://doi.org/10.5194/nhess-14-2531-2014>, 2014.
- Steenbergen, H. M. G. M., Lassing, B. L., Vrouwenvelder, A. C. W. M., and Waarts, P. H.: Reliability analysis of flood defence systems, *HERON*, vol. 49, 2004.
- Suchman, L. A. and Weber, J.: *Human-machine autonomies. Autonomous Weapons Systems*, Cambridge University Press, Cambridge, UK, 75–102, 2016.
- Thieken, A. H., Olschewski, A., Kreibich, H., Kobsch, S., and Merz, B.: Development and evaluation of FLEMOps – A new flood loss estimation model for the private sector, *WIT Trans. Ecol. Environ.*, 118, 315–324, 2008.
- Tiwari, S., Jacoby, H., and Skoufias, E.: Monsoon Babies: Rainfall Shocks and Child Nutrition in Nepal (March 1, 2013). *World Bank Policy Research Working Paper No. 6395*, available at: <https://ssrn.com/abstract=2241953> (last access: 28 April 2020), 2013.
- Tkachenko, N., Jarvis, S., and Procter, R.: Predicting floods with Flickr tags, *PLoS ONE*, 12, e0172870, <https://doi.org/10.1371/journal.pone.0172870>, 2017.
- Triantakontantis, D. and Mountrakis, G.: Urban Growth Prediction: A Review of Computational Models and Human Perceptions, *Journal of Geographic Information System*, 4, 555–587, 2013.
- UrbanRiskLab: <https://urbanrisklab.org/work#/riskmap/>, last access: 12 August 2019.

- Van der Most, H., Tanczos, I., De Bruijn, K. M., and Wagenaar, D. J.: New, Risk-Based standards for flood protection in the Netherlands, 6th International Conference on Flood Management (ICFM6), 16–18 September 2014, Sao Paulo, Brazil, 2014.
- Wagenaar, D. J., de Bruijn, K. M., Bouwer, L. M., and de Moel, H.: Uncertainty in flood damage estimates and its potential effect on investment decisions, *Nat. Hazards Earth Syst. Sci.*, 16, 1–14, <https://doi.org/10.5194/nhess-16-1-2016>, 2016.
- Wagenaar, D., de Jong, J., and Bouwer, L. M.: Multi-variable flood damage modelling with limited data using supervised learning approaches, *Nat. Hazards Earth Syst. Sci.*, 17, 1683–1696, <https://doi.org/10.5194/nhess-17-1683-2017>, 2017.
- Wagenaar, D., Lüdtke, S., Schröter, K., Bouwer, L. M., and Kreibich, H.: Regional and Temporal Transferability of Multi-variable Flood Damage Models, *Water Resour. Res.*, 54, 3688–3703, <https://doi.org/10.1029/2017WR022233>, 2018.
- Wagenaar, D. J., Dahm, R. J., Diermanse, F. L. M., Dias, W. P. S., Dissanayake, D. M. S. S., Vajja, H. P., Gehrels, J. C., and Bouwer, L. M.: Evaluating adaptation measures for reducing flood risk: A case study in the city of Colombo, Sri Lanka, *Int. J. Disast. Risk Re.*, 37, 101162, <https://doi.org/10.1016/j.ijdr.2019.101162>, 2019.
- Wagenaar, D. J., Hermawan, T., Van den Homberg, M., Aerts, J. C. J. H., Kreibich, H., De Moel, H., and Bouwer, L. M.: Improved transferability of multi-variable damage models through sample selection bias correction, submitted, 2020.
- Watmough, G. R., Marcinko, C. L. J., Sullivan, C., Tschirhart, K., Mutuo, P. K., Palm, C. A., and Svenning, J.: Socioecologically informed use of remote sensing data to predict rural household poverty, *P. Natl. Acad. Sci. USA*, 116, 1213–1218, <https://doi.org/10.1073/pnas.1812969116>, 2019.
- Xingjian, S., Chen, Z., Wang, H., Yeung, D., Wong, W., and Woo, W.: Convolutional LSTM network: A machine learning approach for precipitation nowcasting, in: *Neural Information Processing Systems*, arXiv:1506.04214, available at: <https://arxiv.org/abs/1506.04214> (last access: 28 April 2020), 2015.
- Xu, T., Gao, J., and Coco, G.: Simulation of urban expansion via integrating artificial neural network with Markov chain – cellular automata, *Int. J. Geogr. Inf. Sci.*, 33, 1960–1983, <https://doi.org/10.1080/13658816.2019.1600701>, 2019.
- Yomwan, P., Cao, C., Rakwatin, P., Suphamitmongkol, W., Tian, R., and Saokarn, A.: A study of waterborne diseases during flooding using Radarsat-2 imagery and a back propagation neural network algorithm, *Geomatics, Natural Hazards and Risk*, 6, 289–307, 2015.