AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2015WR017464

Special Section:
The 50th Anniversary of Water Resources Research

Key Points:

- Surprise is a neglected element in flood risk assessment and management
- Limits of predictability and cognitive biases need closer attention
- Approaches for reducing surprise and its malicious consequences are discussed

Correspondence to:
B. Merz,
bmerz@gfz-potsdam.de

Citation:
Merz, B., S. Vorogushyn, U. Lall, A. Viglione, and G. Blöschl (2015), Charting unknown waters—On the role of surprise in flood risk assessment and management, *Water Resour. Res.*, *51*, doi:10.1002/2015WR017464.

Received 4 MAY 2015
Accepted 16 JUL 2015
Accepted article online 21 JUL 2015

Charting unknown waters—On the role of surprise in flood risk assessment and management

B. Merz[1,2], S. Vorogushyn[1], U. Lall[3,4], A. Viglione[5], and G. Blöschl[5]

[1]GFZ German Research Center for Geosciences, Section 5.4 - Hydrology, Potsdam, Germany, [2]Institute of Earth and Environmental Science, University of Potsdam, Potsdam, Germany, [3]Columbia Water Center, Columbia University, New York, New York, USA, [4]Department of Earth and Environmental Engineering, Columbia University, New York, New York, USA, [5]Institute of Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Vienna, Austria

Abstract Unexpected incidents, failures, and disasters are abundant in the history of flooding events. In this paper, we introduce the metaphors of terra incognita and terra maligna to illustrate unknown and wicked flood situations, respectively. We argue that surprise is a neglected element in flood risk assessment and management. Two sources of surprise are identified: (1) the complexity of flood risk systems, represented by nonlinearities, interdependencies, and nonstationarities and (2) cognitive biases in human perception and decision making. Flood risk assessment and management are particularly prone to cognitive biases due to the rarity and uniqueness of extremes, and the nature of human risk perception. We reflect on possible approaches to better understanding and reducing the potential for surprise and its adverse consequences which may be supported by conceptually charting maps that separate terra incognita from terra cognita, and terra maligna from terra benigna. We conclude that flood risk assessment and management should account for the potential for surprise and devastating consequences which will require a shift in thinking.

1. Terra Incognita and Terra Maligna

Old maps often labeled unexplored regions of the world as *terra incognita*, or unknown land. Some mapmakers left these areas blank, others let their imagination run wild. Paolo Forlani, for example, drew a gigantic terra incognita with mountain chains and fabulous animals in the very south of his 1565 world map (Figure 1). Fierce beasts such as dragons or large serpents were often depicted to populate these areas, associating terra incognita with terra maligna, i.e., a country where wicked and dreadful encounters could happen. The filling-in of maps projected the society's myths, fantasies, and fears, brought a sense of security, and reduced the "*aesthetic and epistemological anxiety*" [*Cosgrove*, 1999, p. 10] generated by empty spaces in maps [*Murray*, 2005]. Interestingly, the Venetian cartographer Fra Mauro challenged this tradition as early as 1460 when he asked for humility in the face of the unknown: "*If anyone considers incredible the unheard-of things I have set down here, let him do homage to the secrets of nature, rather than consult his own intellect. For nature conceives of innumerable things, of which those known to us are fewer than those not known, and this is so because nature exceeds understanding.*" (quoted in *Whitfield* [1994, p. 2] and *Murray* [2005, p. 105]).

In this paper, we argue that today's actors, involved in flood risk assessment and management, are in a similar position as cartographers in the Middle Ages and Renaissance when charting maps. Although much of the spectrum of flood risks has been explored and documented, there remain blank spaces that may lead to devastating consequences. For illustration, we introduce the metaphors of terra incognita and terra cognita. Phenomena in terra cognita are those we know and have thought about. Terra incognita encompasses what we do not know or perceive wrongly. Our perception and knowledge of phenomena are always based on some kind of model for the situation at hand. Models range from mental models that represent or symbolize external realities, to mathematical models that quantify reality numerically. All of these models are based on assumptions, available data, and experience. We consider phenomena to be located in terra incognita when our mental and/or mathematical models do not include them or when they include them wrongly.

10.1002/2015WR017464

Figure 1. Map by Paolo Forlani, Venice, 1565, showing a gigantic terra incognita in the southern hemisphere, populated by fabulous animals [*Library of Congress*, 2015].

In case an event from terra cognita happens, society is in a good position of coping with it since, in a favorable economic and political setting, adaptation to known risks is always a governance objective. This does not mean that there are no adverse consequences, but there is planning to mitigate the risk. If an event from terra incognita occurs, however, society is caught by surprise. Following *Taleb* [2007], we introduce the notions of terra benigna where the consequences of surprise are benign, and terra maligna, where our limited knowledge may lead to devastating consequences.

Surprise plays the key role in this paper. Surprise has been variably referred to as indeterminacy, deep uncertainty, wild uncertainty, known unknowns, and unknown unknowns. *Macgill and Siu* [2004] differentiate between uncertainties for which distributions can be assigned and indeterminacy where phenomena cannot be known, measured, controlled, or predicted. In his book on Black Swans, *Taleb* [2007] identifies mild uncertainty where processes and risks are predictable (if uncertain), and wild uncertainty where they are just not predictable. *Cox* [2012] defines deep uncertainty as situations where trustworthy models are not available, experts disagree about the probable consequences of alternative decisions, and where there is not enough objective information to support rational decision making. Former United States Secretary of Defense Donald Rumsfeld noted in a press conference in February 2002: "*. . .there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say, we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know*." We collectively refer to these types of uncertainty by the term *surprise*. A central notion of this paper is that surprise is inherently subjective and depends on the respective actor, be it an expert or a lay person. What would be a known unknown at a certain time to one person, could be an unknown unknown to another.

Examples for terra incognita and terra maligna in flood risk assessment and management abound. One example is the summer flood in Central Europe in 2002 which incurred losses of 15 billion Euros. The flood was caused by intensive and widespread rainfall associated with a storm track that advected moisture from the Mediterranean to Central Europe. Although this meteorological situation was well known for its flood potential, the consequences such as more than 130 dike breaches in the German part of the Elbe catchment took authorities and residents by surprise. The authorities were aware that the dikes would probably fail when overtopped but such scenarios had neither been considered nor communicated to the public prior to the event. The survey of *Kuhlicke* [2010] in the township of Eilenburg at an Elbe tributary revealed that the majority of the citizens could simply not imagine that such a situation could happen. Society had developed a blind spot where knowledge about the environment had been forgotten, neglected, or suppressed, and where major events were outside societal imagination, i.e., in terra incognita from a citizen perspective. Massive damage due to leaky oil tanks was another surprise from terra maligna as citizens had been unaware of the potential consequences of the link between flooding and non-flood-proofed oil tanks.

The 2011 Thailand flood is a second example. A sequence of five typhoon landfalls led to floods that persisted over 90 days, with over 2.5 million people displaced, an estimated $46 billion in damages, and an over $32 billion impact on the manufacturing sector that reverberated through global automobile and electronics supply chains for a period of nearly 2 years subsequent to the event [*Haraguchi and Lall*, 2015]. Although it was known to the insurance industry that there existed significant flood hazard in Thailand, the actual accumulation of assets and the global implications to supply chains were a surprise for them. For the multinational manufacturing sector that had developed economically efficient supply chains with just-in-time delivery, the event was an extreme surprise in terms of the scale of its impact.

The third example is more complex. In 1963, a massive landslide into the Vajont reservoir, Italy, caused a flood wave that overtopped the dam and resulted in nearly 2000 fatalities when the downstream valley was inundated [*Müller-Salzburg*, 1987; *Delle Rose*, 2012; *Di Baldassarre et al*., 2014]. Based on preliminary geological explorations, deep-seated landslides were considered to be unlikely by engineers because, among others, boreholes showed no evidence of weak layers at depth and the rock walls were judged as firm according to seismic surveys [*Kilburn and Pretley*, 2003]. During the first few years of the reservoir filling, slow downhill creep was observed [*Semenza and Ghirotti*, 2000]. It became evident that a big landslide was unavoidable, and the engineers started to build a bypass tunnel [*Datei*, 2005], which would have connected the two branches of the basin if the expected landslide would have divided it into two parts. Even though several other signals of instability and fragility of the slope were reported [*Merlin*, 1983], and new geological surveys had suggested the existence of a deep ancient landslide [*Müller*, 1964; *Bianchizza and Frigerio*, 2013], the dam owner downplayed the risk and continued the operations. Lowering the water level seemed to control the slope movement. Over the following few years, filling and lowering of the reservoir was alternated as the dam engineers attempted to keep the creep velocity under control. This was based on the wrong hypothesis that the creep was caused by the initial saturation of the slope material and could be controlled by further lowering and subsequent filling of the reservoir [*Semenza and Ghirotti*, 2000, and references therein]. In the end, the dam engineers were not taken by surprise of the occurrence of the event, but by its actual speed and magnitude because of wrong perceptions, overconfidence, and wishful thinking. Moreover, this terra incognita led to malign consequences also because no warning or evacuation orders were issued in time to the population.

Despite recent progress in assessing and managing flood risks, the mental and mathematical models involved do not usually consider the potential for surprise as the above examples demonstrate. In this paper, we argue that terra incognita and terra maligna are presently not fully acknowledged in flood risk management. To better cope with surprise and its malicious consequences, we explore approaches for understanding and reducing the potential for surprise in a given situation.

2. Sources of Surprise in Flood Risk Assessment and Management

Where does surprise come from? We suggest that there are two sources of surprise in flood risk assessment and management: the complexity of flood risk systems that confounds the predictability of failure scenarios, and biases in human perception and thinking.

2.1. Lack of Predictability: System Complexity as a Source of Surprise

Flood risk results from the interaction of hazard, exposure, and vulnerability [*UNISDR*, 2013; *IPCC*, 2012]. Hazard represents the probability and intensity of flooding. Exposure describes the elements-at-risk such as people and their assets that may be affected by flooding. Vulnerability describes the susceptibility of elements-at-risk to be adversely affected. These components of risk are controlled by a variety of factors. Flood risk systems are, therefore, complex systems which are not easily understood nor managed. Although there have been various attempts at defining system complexity [e.g., *Casti*, 1986; *Pincus*, 1991; *Allison and*

Table 1. Factors That Increase the Potential for Surprise and Contribute to Terra Incognita

Factor	Explanation
Nonlinearity	Nonlinearity is present when a small change in one variable leads to a nonproportional change in the dependent variable, or if threshold effects occur. For example, a small increase in one variable leads to a sudden transition from near accident to severe accident.
Interdependence, feedbacks	Interdependence in space, and in time, between elements of the system studied, or combinations thereof, may lead to surprise. Temporal interdependence means that a process is related in time, e.g., today's state has an effect on the future state. Spatial interdependence relates to process at one location affecting the process at other locations. Interdependence between elements means that a certain process or element is related to another element or process and may involve feedback processes.
Nonstationarity	We define nonstationarity here as temporal change in the structure or/and the boundary conditions of the system studied, rather than as a change in the statistics of a state variable. Such changes imply that our experience and knowledge are less valuable and harder to extrapolate into the future than for a stationary situation.

Hobbs, 2006], no universal definition has emerged. We suggest that the complexity of flood risk systems may stem from their nonlinearity, interdependencies, and nonstationarity (Table 1). Not accounting for these characteristics in mental or mathematical models of a flood risk system will reduce predictability and lead to surprise.

Nonlinear systems tend to be less predictable than linear systems as illustrated by *Blöschl and Zehe* [2005] for the case of the nonlinear dependence of runoff generation on the moisture state of the soil. Predictability is even more limited in flood risk systems where socioeconomic processes with nonlinear relationships are common. The response of society to a flood neither scales linearly with the hydrological nor with the economic severity of that flood. From the late 1980s, a number of severe floods occurred in Germany which only triggered limited response, such as the construction of additional flood retention basins in the affected catchments. In contrast, the 2002 flood led to comprehensive policy changes at the national scale, such as mandatory inclusion of residual risk scenarios in dam safety analyses and the harmonization of flood protection regulations [*Petrow et al*., 2006]. On the other hand, the 2013 flood with comparable economic losses and much higher hydrological severity than that in 2002 [*Blöschl et al*., 2013b; *Schröter et al*., 2015] did not lead to commensurate societal responses. Obviously, the 2002 event was perceived as a signal that flood risk management had to be improved while the others were not.

Other causes of complexity are interdependencies and feedbacks. Flood risk systems usually contain many elements that interact with each other. A case in point is the influence of dike breaches and flood retention on downstream locations. Simulations of *Vorogushyn et al*. [2012] along the Elbe River showed that increasing flood retention upstream leads to smaller probabilities of downstream dike overtopping as would be expected. Less obviously, they also showed that retention may increase the probabilities of downstream dike failure mechanisms such as piping and microinstability due to an increase in the flood duration. Such spatial interdependencies are rarely taken into account in practice, and yet may lead to counterintuitive and unexpected risk patterns. Flood risk systems may also contain more complex feedback loops between the hydrological and societal system components [*Sivapalan et al*., 2012]. One example is the levee effect which relates to the observation that flood protection may increase vulnerability as the protection reduces the frequency of flooding and people tend to move into protected areas due to a false sense of safety. Such dynamics have been explored by sociohydrology models that represent the feedbacks between economy (in terms of wealth), technology (in terms of level of flood protection), hydrology (in terms of flood magnitudes and damage), politics (in terms of urban planning), and society (in terms of risk awareness) based on coupled nonlinear differential equations [*Di Baldassarre et al*., 2013; *Viglione et al*., 2014].

A further cause of complexity is the nonstationarity of the flood risk system itself. Flood risk analysts and managers usually draw (for good reasons) narrow boundaries as they define their system, without considering changes in the boundary conditions and/or the system structure. For example, flood behavior often exhibits decadal variability with flood-rich and flood-poor periods [*Hall et al*., 2014], but flood design is widely based on the stationarity assumption which may underestimate the probability of extreme events [*Jain and Lall*, 2001] and thus increase the potential for surprise. Another example is the 2011 Thailand flood

Table 2. Cognitive Biases That Contribute to Flawed Risk Estimates and Management[a]

Type of Bias	Description
Wishful thinking	Being overly optimistic and letting own preferences affect expectations of the future. Overestimating the influence on events over which one has little or no control.
Erring on the side of least drama	Asking for greater levels of evidence in support of dramatic, alarming conclusions.
Selective perception	Analyzing problems in terms of own experience and background.
Availability bias	Associating higher occurrence probability to events that can be more easily remembered or imagined.
Recency bias	Putting stronger weight on more recent events.
Illusory correlation bias	Perceiving patterns, trends, correlations when they do not exist.
Illusion of certainty, overconfidence	Emotional need for certainty when none exists. Overrating one's own knowledge.
Hindsight bias	Overestimating the likelihood with which one would have correctly predicted known outcomes.
Confirmation bias	Favoring information that confirms existing preconceptions or hypotheses.

[a]The table is based on *Makridakis* [1988] with additional information from *Evans* [1982], *Kahneman et al*. [1982], *Sjöberg* [1982], *March et al*. [1991], *Bunn and Salo* [1993], *Nickerson* [1998], *Pidgeon and O'Leary* [2000], *Gigerenzer et al*. [2008], *Kahneman* [2011], and *Brysse et al*. [2013].

where reinsurance risk analysts had excluded accelerating interconnectedness of global supply chains over past decades from their system boundaries.

2.2. Psychological Sources of Surprise—Cognitive Biases

Cognitive biases are another source of surprise. Human perception and decision making are heavily plagued by biases in intuitive judgments [*Tversky and Kahneman*, 1974; *Kahneman*, 2011]. Table 2 provides a summary of such biases. Some of these biases lead to overly optimistic risk estimates and management decisions, tending to neglect dramatic events or consequences. Other biases lead to overconfidence of one's system understanding, where people are unaware of how far off their perception is from the actual risk.

In the case of the Vajont flood, the overconfidence of the project engineers in their ability to control the creep velocity and system behavior through the reservoir water level may have contributed to the disaster. Overconfidence is a widespread phenomenon that results in a tendency to overrate one's knowledge. Psychological experiments consistently show that the distribution representing the true error has a larger variance than the subjective error. According to *Hammitt and Shlyakhter* [1999], 20–45% of true values are typically outside the subjective 98% confidence region. There is also a strong psychological need for certainty even when none exists, termed the illusion of certainty bias [*Gigerenzer et al*., 2008]. This prompts, for example, the belief that the future can be forecasted even in cases when the empirical evidence suggests otherwise [*Makridakis*, 1988].

Other examples for biases in flood risk assessment and management are the availability and recency biases where people rate those events as more probable that are more readily available to them, i.e., events they can recall from memory or imagine easily. Recent events tend to dominate over earlier events. An extreme flood scenario is often judged as more likely, when it is presented vividly and in a concrete way, when it tells a good story, or when a similar incident has occurred recently. This process of recency has been modeled as collective memory in the sociohydrology model of *Viglione et al*. [2014]. A case in point is the 2002 Elbe flood. Before the flood, lay people and experts in Germany would have assigned a smaller probability of occurrence (or even a probability of zero) to a flood scenario with more than 100 dike breaches compared to their assessment of today, after such a scenario has actually occurred. We can safely assume that the perceived probability of this scenario has increased between 2002 and today, whereas the actual probability has decreased because numerous polders were built and dikes were strengthened in the meantime.

An example of selective perception is the media attention that surrounds extreme events and creates an impression that they are unprecedented. One such case is the "unprecedented polar vortex" of winter 2013/2014 that kept the Northeastern U.S. cold, and Northern Europe warm. This phenomenon was first described in 1853 [*Littell*, 1853], and there has been recent study of such phenomena, including the possible role of changing climate [*Rasmussen and Turner*, 2003; *Reichler et al*., 2012]. In combination with a recency bias that lets decision makers and society "forget" about these events given their infrequent occurrence, there is either over-recognition (e.g., Hurricane Sandy), or only short-term attention that does not translate into coherent action (e.g., polar vortex). Hurricane Sandy's landfall in New York City galvanized a significant amount of action across the U.S. in the belief that such events may become routine under a future climate. It is unclear whether the history or the science can be used to make such an assertion with confidence, yet it has had a significant impact on public perception.

Another cognitive bias (manifest as hindsight bias or illusory correlation bias) is related to the clustering of extreme events, which is possible given the quasi-oscillatory nature of climate dynamics. As illustrated in *Jain and Lall* [2001], temporal structure in climate can lead to persistent increase or decrease in regional flooding, which, depending on the length of record relative to the time scale of the underlying oscillation, may be misinterpreted as a structural change, leading to an inappropriate extrapolation into the future, or ignored, leading to a perception of overdesign or underdesign relative to the actual outcomes. For instance, a design based on a 100 year flood level estimated from the historical record may be called into question if the region experiences multiple 100 year floods soon after project completion. Based on media reports, such occurrences are not infrequent. However, the reverse (the overdesign case) is rarely reported, leading to an asymmetry in perception. Finally, the setting of a new record, which is often presented as evidence of change, and hence contributes to cognitive bias, can be an issue, even if there is no real change in the system. *Vogel et al*. [2001] demonstrate that many record floods can simply be ascribed to the fact that as the duration of the data grows, new records will appear, purely from the behavior of a random process.

Overall, cognitive biases are particularly strong (1) in complex situations with many factors involved, (2) when uncertainty is high and events are probabilistic, (3) when the outcome of decisions is not closely reported back, (4) when people are under stress, and (5) when strong beliefs are held [*Sjöberg*, 1982; *Bunn and Salo*, 1993]. Points 1–3 clearly apply to flood risk assessment and management. Particularly problematic is the lack of direct and meaningful response on the outcome of decisions. Learning requires feedback about the outcome, however, in flood risk assessment and management there is either no feedback due to the rarity of events or the feedback is ambiguous (for example because relationships are probabilistic). When outcome feedback is slow or unclear, individuals and organizations tend to repeat decisions simply because they have made them before [*March et al*., 1991]. Stress levels (point 4) are also common in flood situations, where people tend to get a more narrow perspective, regress to more primitive levels of decision making, and deal successfully only with small amounts of simultaneous information [*Sjöberg*, 1982; *Fahy and Proulx*, 1997]. All of these factors may aggravate the crisis.

3. Charting Terra Incognita: What Is the Potential for Surprise?

The current gold standard of charting terra incognita, i.e., for understanding the limits of knowledge, is uncertainty analysis. Methods include uncertainty propagation, ensemble methods, sensitivity analysis, and regional error analysis [e.g., *Apel et al*., 2004, 2006; *Pappenberger et al*., 2006; *Hall and Solomatine*, 2008; *Merz and Thieken*, 2009; *De Moel and Aerts*, 2011; *Gupta et al*., 2013]. Although these methods are very valuable, they are not designed to capture surprise. They estimate uncertainty by mathematical models developed from past observations and experience of the problem at hand. When flood risk systems are afflicted by deep uncertainty, trustworthy models do not exist. "Observations of risk" are rarely available, and risk estimates cannot be validated in the traditional sense by comparing them against observations [*Hall and Anderson*, 2002]. The important events are extremes, outliers, or unrepeatable processes, such as the landslide in the Vajont case, which makes the accumulation of knowledge and the learning through experience difficult. Some insight into possible disaster scenarios can be gained from similar settings elsewhere (e.g., a flood similar to Vajont may occur in other river valleys downstream of reservoirs), but there remains a plethora of unique possibilities.

Some authors have attempted to capture the potential for surprise by formal methods. *Shlyakhter* [1994] and *Hammit and Shlyakhter* [1999] proposed to inflate the estimated uncertainties by safety factors determined from relevant historical data to account for unsuspected errors. However, "*ignorance cannot be aware of itself*" [*Macgill and Siu*, 2004, p. 327] and the truly unthinkable cannot be imagined upfront. This is related to Rumsfeld's notion of unknown unknowns. Yet it is possible to reflect about the boundaries of what is known. For example, we may not be able to identify a priori human errors that cause failure of a flood

Table 3. Reasons for Surprise of Some Flood Disasters

Disaster	Who Was Surprised?	Complexity as a Source of Surprise	Psychological Sources of Surprise
Elbe 2002 flood: massive dike breaches and devastating losses	Engineers and authorities	*Nonlinearity*: Dike breach is threshold process.	*Availability and recency biases*: Dike failures of this order of magnitude were outside imagination. Missing experience of dike failures (last significant fluvial dike breaches in Germany occurred in 1882/1883 at the Rhine). *Wishful thinking*: Consequences of load exceeding dike resistance were not considered in safety concepts.
Elbe 2002 flood: massive damage due to flooded oil tanks in private households	Citizens	*Nonlinearity*: Oil contamination is threshold process. *Nonstationarity*: Oil tanks were not in existence during nineteenth century floods	*Availability and recency biases*: Missing awareness of oil spill hazard. Missing experience with floods and emergency measures. *Wishful thinking*: Flooding hazard was not linked to possibility of household flooding and associated consequences.
Thailand 2011 flood: massive direct and indirect losses	Insurance industry	*Interdependence*: Significant interdependences between flooding in Thailand and global manufacturing industry via global supply chains. *Nonstationarity*: Rapid accumulation of industrial assets and increasing role of Thailand in global manufacturing industry.	*Availability and recency biases*: Rapid rate of asset accumulation and changes in vulnerability of the global manufacturing industry not considered.
Vajont dam disaster 1963	Dam engineers	*Nonlinearity*: Slope movement highly nonlinear. Landslide and overtopping of dam structure are threshold processes. *Feedbacks*: Fast draw down of the lake accelerated the speed of the landslide.	*Illusory correlation bias*: Wrong perception of relationship between reservoir water level and mass movement. *Overconfidence*: Engineers overrated their knowledge about the system and their ability to control the mass movement by water level regulation. *Wishful thinking*: Possibility of such a massive landslide and disastrous consequences not realistically perceived.

defense system, but we may be able to assess whether this system shows characteristics where human errors are possible or even probable. We suggest that flood risk assessment and management should invest more efforts on reflecting about what may be hidden by specifically considering the two sources of surprise, system complexity, and psychological biases. Table 3 illustrates how the two sources of surprise may have contributed to terra incognita for the flood disasters introduced in section 1. In order to assist in reflecting on possible biases, we propose to map surprising events into the space of potential for surprise defined by system complexity and psychological biases. When doing this we are in a similar position as cartographers in the Middle Ages and Renaissance when charting maps. We are charting terra incognita and terra cognita of floods, to separate the known from the unknown. We do this from the perspective of the stakeholders involved at the time the event happened, mimicking a process Paolo Forlani must have gone through when compiling the map shown in Figure 1.

Figure 2 shows an attempt of such a mapping. The axes indicate the potential for surprise arising from the two sources. Events in terra incognita are those the stakeholders had not included in their mental or mathematical models, or had included them wrongly. Events in terra cognita are those that had been anticipated

10.1002/2015WR017464

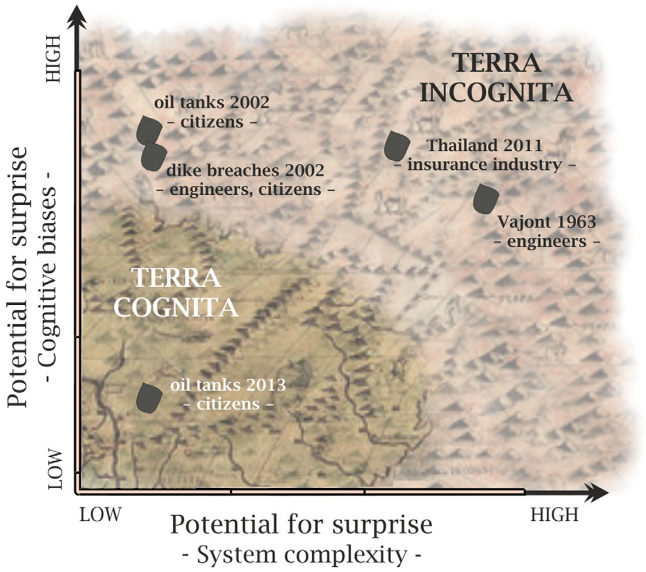

Figure 2. Mapping terra incognita: an attempt to understand the potential for surprise. Each event is linked to the perspective of a specific group of stakeholders such as citizens or engineers.

by the stakeholders. The events are plotted with different degrees of potential for surprise, similar to different degrees of geographical knowledge represented in old maps. The positions on the map were chosen by qualitative reasoning on the basis of Table 3. This is clearly a subjective process. However, this process could be formalized by methods such as choice experiments [e.g., *Adamowicz et al.*, 1998; *Morrison et al.*, 1999; *Balcombe et al.*, 2014] that explore the judgment of a set of experts or stakeholders by questionnaires, eliciting responses that allow a ranking of the potential for surprise. The scales of the axes are relative scales, in line with the ranking outcome of a choice experiment.

In the case of the 2002 Elbe flood, the hydrometeorological characteristics were extreme but not exceptional in terms of system complexity. The major surprises resulted from cognitive biases of engineers, authorities, and affected citizens. Hence, these two scenarios (dike breaches and oil contamination) are plotted in the top left corner of Figure 2. Massive dike failures and oil contamination were not considered as realistic scenarios, although any engineer would have confirmed that these dikes would fail when overtopped, and the link between flood hazard and contamination by oil tanks is straightforward in hindsight. Availability and recency biases and wishful thinking were likely present in all stakeholder groups. Interestingly, for the 2013 flood, the damage due to oil contamination in Germany was much smaller than in 2002. The disastrous experiences in 2002 resulted in legislative amendments and made citizens aware of this problem. The lesson had been learned and the biases were strongly reduced by the 2002 experience.

The Vajont and Thailand events are plotted in the top right corner of Figure 2 as both sources of surprise contributed to these disasters. In the Vajont case, the entire system comprising the reservoir, the surrounding hillslopes, and the downstream areas showed strong nonlinear behavior, and the feedbacks between the reservoir operation and land movement contributed to triggering the landslide. This complexity of the physical system was combined with cognitive biases. The danger of land sliding into the reservoir and a resulting flood wave was clearly known. However, there was the wrong perception of the governing mechanism leading to initial creep and increasing velocities [*Semenza and Ghirotti*, 2000]. This mental model was combined with overconfidence and wishful thinking. In the Thailand case, the changes in the global manufacturing industry and the strongly increased interdependence had not been considered in the mental and mathematical models of the insurance industry. In situations, where the system itself undergoes rapid changes, actors are particularly prone to availability and recency biases.

These disasters illustrate that the sources of surprise can vary vastly from case to case. Although we cannot see what is hidden in terra incognita, charting the unknown waters in terms of the potential for surprise may be extremely useful in flood risk assessment and management.

4. Charting Terra Maligna: What Are the Consequences of Surprise?

Events from terra cognita are those that stakeholders have included in their mental or mathematical models. Given time and opportunity for adaptation, these events will not be disastrous. Events from terra incognita are different as they come as a surprise. Some of these events may lead to small consequences as they are from terra benigna, but others may be devastating as they are from terra maligna.

The surprising events have now been charted in Figure 3 in a similar fashion as in Figure 2 but in terms of terra benigna and terra maligna. The horizontal axis in Figure 3 represents the potential for surprise,

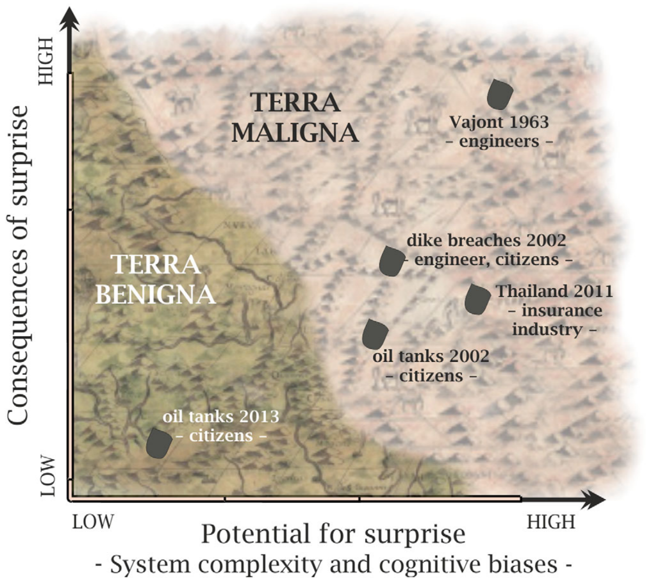

Figure 3. Mapping terra maligna: an attempt to understand the potential of malicious consequences if a surprising event happens.

encompassing both sources of surprise (system complexity and psychological biases) from Figure 2. The vertical axis represents the maliciousness or viciousness of the consequences in case a surprise happens. It is important to note that the notion of maliciousness adopted here is different from the notion of damage which is the usual measure of negative consequences of floods. It is a more personal notion of how people experience a flood involving aspects of hostility and meanness. Whether the consequences of a surprising event are considered malign, therefore, is subjective and depends on the individual. In the Vajont case with 2000 fatalities, the maliciousness is obvious. Oil contamination of a single residential building as a consequence of inundated oil tanks may also be perceived as malign by the owner of the house, who may have toiled for years to afford it, as oil spills usually require the entire house to be demolished.

Similar to separating terrae cognita and incognita, the map in Figure 3 has been drawn based on qualitative reasoning. Choice experiments could again be used to specify the degree of maliciousness from the perspective of a given stakeholder. Clearly, the widespread practice of limiting flood risk assessments to direct economic losses is too narrow because monetary units often do not express whether an event is really vicious to a stakeholder. Economic aspects cover only one part of how the consequences of an event are perceived. This broader perspective is in line with the bottom-up or vulnerability-based approach [*Blöschl et al.*, 2013a] that is motivated by a social paradigm where risk is not defined in monetary but in more qualitative terms. In this approach, the main goal is not to find the most economic management strategy but to ensure the well-being of people by reducing vulnerability and enhancing resilience.

In a practical project context, it would therefore be most useful to understand in advance whether, for a given flood risk system and stakeholder perspective, surprises could be malicious. If taken by surprise, could stakeholders end up in terra maligna? If our mental and/or mathematical model is wrong, if the system does not behave as we assume, or if the future does not evolve according to our scenarios, could the consequences be malicious? Models and assumptions then need to be evaluated by the harmful consequences of their errors rather than by their plausibility. This is the model robustness question of *Ben-Haim* [2012, p. 1644]: How wrong can the models be and still guarantee the outcome is acceptable? Similarly, *Taleb* [2007] warns against depending on models that may have large-scale harmful consequences if being wrong.

The location of an event in terra incognita does not necessarily imply it is in terra maligna. Consider the uncertainty in estimating the 10,000 year flood for two cases. In one case, the estimate is used for designing the spillway of a dam, and in the other case, it is used for assessing the safety of a nuclear power plant located at a river. The dam may be constructed in a way that a flood much larger than the estimate will not lead to dam failure and the resulting flood wave may not lead to devastating consequences downstream if in a remote place. For the case of the nuclear power plant, however, a flood much larger than the estimate may overwhelm the safety measures and lead to nuclear contamination. The dam example is in terra benigna, while the power plant example is in terra maligna.

The consequences of surprise may vary vastly from case to case. It is therefore not only important to evaluate our mental and mathematical models according to their plausibility but also according to the harm they cause when they are wrong. Beyond the traditional question of how close the models are to reality, we

Table 4. Approaches for Reducing the Potential for Surprise by Better Understanding and Modeling Flood Risk Systems

Aspect	Approaches
System boundaries of mental and mathematical models	Draw the boundaries of the system studied deliberately. Avoid narrowly defined perspectives. Increase dimensionality of risk assessments.
Mathematical models	Understand sensitivity, uncertainty, and assumptions of the model. Generate multiple plausible models, given available knowledge and data.
Data	Assemble evidence as inclusive as possible. Use temporal, spatial, and causal information expansion.
Failure and future development scenarios	Generate a large and diverse range of scenarios. Explore the possible (and not only the probable). Use backward inference in scenario generation. Use risk assessment and scenario generation methods from other disciplines.

need to address the model robustness question of whether there are malicious consequences in case the models are wrong. We need to chart the unknown waters in terms of terra benigna and terra maligna.

5. Moving Out of Terra Incognita

What can we do to reduce the potential for surprise and move future events out of terra incognita? Addressing the sources of surprise, we can attempt to better understand the complexity of the flood risk system and we can attempt to reduce cognitive biases. The first moves the future event to the left in Figure 2, the second moves it down.

5.1. Increasing Predictability by Better Understanding Flood Risk Systems

The better we understand and the better we are able to model the flood risk system under study, the smaller the potential for surprise. The aim is to enlarge our mental models, seek process understanding, and incorporate it in the risk models. As *Blöschl and Montanari* [2010] noted, offering insightful explanations may be more helpful than perfecting the estimates of what are inherently uncertain predictions. Models of flood risk systems should therefore not be black boxes, in particular, when the data samples are small or when the event under consideration is unique, such as in the Vajont case. Interdependencies, nonlinearities, and nonstationary effects in flood risk systems that are prone to cause surprise need to be considered in detail. Table 4 proposes a number of approaches for better understanding flood risk systems that address the system boundaries, the mathematical models, the data to parameterize and validate the models, and the failure and future development scenarios.

The mental or mathematical models used in risk assessment cannot represent every detail of reality, and the modeled system is typically much narrower than reality. Drawing the system boundaries therefore requires careful attention as they determine what is included in the assessment and what is not. For example, flood hazard analysis traditionally focuses on flood peaks alone, as in U.S. Bulletin 17B [*Interagency Advisory Committee on Water Data*, 1982], but other event characteristics may be more relevant for society. For example, the most disastrous flood in the Mekong Delta in recent history was the 2000 flood. Its peak discharge was not extraordinarily high, but the flood volume was the largest recorded in the 88 years of observation, causing extended inundation, prolonged water logging, and huge damage [*Nguyen et al*., 2014]. Contamination of flood water may be of high importance for flood damage to private households, yet it is rarely considered [*Thieken et al*., 2005]. In investigating vulnerability, it is important to go beyond the assessment of inundation areas and direct economic damages and address, for example, the implications of inundations for the needs of particular population groups during evacuation and emergency response. An overly narrow perspective may also miss the interplay of a number of events or risk reduction measures. For example, the 2013/2014 winter floods in the UK were caused by the combined effect of persistent rainfall, strong winds, and coastal surges. Double events may be the most relevant in practice [e.g., *Blöschl et al*., 2013b] but are rarely considered in hydrological design. Similarly, downstream effects of flood defense measures may be missed by a narrow system perspective [*Vorogushyn et al*., 2012].

Another example which calls for broader system boundaries is flood forecasting. Hydrologists tend to focus on forecasting from a mechanistic rainfall-runoff event perspective. Consequently, a large or persistent rainfall event that causes floods is inevitably treated with an element of surprise. Expanding the hydrologists' "control volume" to include the global ocean-atmosphere circulation and mechanisms of atmospheric moisture transport associated with large floods has the potential of reducing some of this surprise. This could come from increasing the understanding of the mechanisms that lead to major floods, or from seasonal climate forecasts that could be linked probabilistically to a shift in conditions that may change the likelihood of floods. *Nakamura et al*. [2013] and *Lu et al*. [2013] provide evidence of persistent organization in atmospheric states that can bring in repeated and persistent waves of moisture into a region leading to large-scale flooding, in the Ohio River basin, in the U.S., and for France, respectively. *Steinschneider and Lall* [2015] demonstrate the predictability of such mechanisms in space and time, while *Kwon et al*. [2008] and *Lima et al*. [2015] provide examples of preseason flood prediction using climate precursors. The importance of such work is that often there is short-term predictability of such events from atmospheric precursors, which can be used to reduce surprise. *Collier and Skees* [2012] and *Surminski* [2014] report on the application of these ideas to develop index insurance products, that can have payouts even before the flood season, based on the forecast index, to facilitate preparation for a potential extreme flood. This translates surprise from a purely hydrologic analysis to financial instruments that aid disaster preparation and risk mitigation. A thorough and explicit reflection on the system boundaries is therefore needed, evaluating any neglected factors and the consequences of any flaws in the assumptions made. Expanding the perspective and increasing the dimensionality of risk assessments are necessary for reducing the potential for surprise.

Once the system boundaries are set, mathematical models usually provide the basis for decision making, e.g., for designing flood protection. Whereas uncertainties are increasingly considered, analyses often focus on the model inputs and model parameters as uncertainty sources. Reducing the potential for surprise requires that, additionally, model structure choice and other assumptions are scrutinized for their effects on the management decisions. *Jain and Lall* [2001] offer an example which shows that the traditional assumption of independent and identically distributed (iid) flood peaks may lead to surprise. Considering ENSO-related climate fluctuations led to a substantially larger number of exceedances of the design flood compared to the iid assumption. Another example is the flood maps provided by official flood protection programs. These maps, typically, are pieced together using local estimates and do not attempt to represent the correlation between risks at different locations. Spatial independence, however, may not be a valid assumption when assessing the likelihood of large-scale flood events, their required emergency resources, and the associated losses.

Model development is based on experience and data. *Hall and Anderson* [2002] argue that, because evidence about causal phenomena and system responses is scarce, we should be as inclusive as possible in assembling evidence when we are dealing with extreme or unrepeatable hydrological processes. The notion of flood frequency hydrology suggested by *Merz and Blöschl* [2008a,b] aims at expanding the information horizon in flood estimation. Rather than limiting the problem to a statistical estimation problem with incomplete discharge data, it aims at assembling a range of hydrological information including evidence from similar events in the region, events from back in history including chronicles and flood marks in the landscape, and runoff generation mechanisms. This concept will require a shift in thinking from recipe-based estimation to a broader exploration of flood processes involving hydrological reasoning and formal methods [*Viglione et al*., 2013]. Some countries, such as Germany, have already adopted these concepts in their national guidelines [*Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall*, 2012].

The models are finally used to derive scenarios, either failure scenarios (what can go wrong?) or future development scenarios (how could the future look like?). Traditionally, hydrological scenarios are sought that are as close as possible to the perceived reality. However, reducing surprise from a complex flood risk system requires a scenario space with as diverse a range of alternatives as possible. In his discussion on hydrologic predictability, *Kumar* [2011] suggests that for the anticipation of surprise it is imperative that hydrologic science explores the "possible" with just as much vigor as the "probable." He argues that the exploration of novel phenomena differs from the prediction of future events. Typically, our mindset focuses on the latter case and we attempt to constrain the model to the probable trajectory to obtain the least uncertainty when using the model in the forecast mode. However, the reduction of surprise requires exploring the space of all possible trajectories. In fact, the original concept of the scenario approach developed by the corporate world from the late 1960s was very much about surprises. An example is Shell Oil's scenario of a sharp increase in oil prices which was unthinkable at the time but did materialize in 1973 [*Van Der*

Heijden, 1996]. It would be good to go back to the original concepts and use scenarios to explore the unpredictable rather than to use them as pseudoforecasts [*Hall et al*., 2014]. The call to explore the "possible" also applies to the case of short-term flood forecasting. For understandable reasons, the development of flood forecasting models typically focuses on simple solutions which are able to represent past events. Although such models serve well in the case of nonsurprising events, their forecasts might fail when processes come into play that have not occurred in the observational period. An example is the large forecast error at some locations along the Elbe River in 2002 where the forecast model had not taken into account the effects of dike breaches on the downstream flood wave. Process-based models which attempt to include interactions, nonlinearities, and nonstationarities might help to explore the limits of predictability also in the case of operational forecasting systems.

We need to let imagination play. *Pidgeon and O'Leary* [2000] argue that we should attempt to step temporarily beyond (or even suspend) assumptions about the likely hazard and its consequences. The flood risk management study of *Wardekker et al*. [2010] for the city of Rotterdam, Netherlands, widens the scenario space by exploring imaginable surprises, something they term "wildcards." These wildcards include the collapse of the thermohaline circulation, port freezing events, port malaria incidents, a modified German water safety policy, enduring heat waves and droughts, extreme storms, and failure of the storm surge barrier during an extreme storm. *Henrion and Fischoff* [1986] propose that we focus on extreme possibilities, for example, asking "imagine that, ten years hence, today's best estimate proves to be off by four standard deviations; how could you explain it?" Exploratory modeling can be an important part of such an exercise of asking "if-then" questions [*Blöschl et al*., 2013a]. Another approach for generating a diverse scenario space is backward inference [*Bunn and Salo*, 1993]: Assume a future state or a certain failure and search for possible causes and developments which could have led to this situation. This backward approach tends to be better suited to anticipate surprise than forward inference, where the present is taken as a starting point to simulate realistic but nonsurprising events. In this respect, flood risk assessment and management may learn from research and practices in other safety relevant fields, such as nuclear safety and air traffic safety. A range of methods have been developed in these fields to detect and quantify scenarios with adverse impacts. One of the earliest methods is the "Failure mode and effects analysis" where as many components and subsystems as possible are screened to identify failure modes, and their causes and effects [*Rausand and Hoylan*, 2004].

The conclusion to broaden the perspective resonates with *Kumar* [2015] and *Sivapalan* [2015]. The latter argues in his comment on sociohydrology that offering sustainable solutions requires hydrologists to broaden their scope, including especially the human dimensions. Similarly, *Kumar* [2015] calls for an integrated approach and a broad contextual view of water in all its complexities to address water security and emergent environmental risks.

5.2. Addressing Psychological Sources of Surprise

Cognitive biases are hardwired in the human brain but there are ways of reducing their effectiveness. Table 5 lists approaches for debiasing which could be used in flood risk assessment and management. In the following, we illustrate a few of these debiasing approaches by examples.

In the 2002 Elbe flood (see Table 3), availability and recency biases could have been reduced by adopting a broader perspective beyond the easily available and recent experiences. Presenting information in a balanced way including all aspects of the risk at hand and considering the fundamental factors that affect the process of interest (Table 5) would have involved the assessment and communication of the consequences of extreme scenarios and flood defense failures. Remembering that "no evidence of disaster" does not mean "evidence of no disaster," and remembering that cycles, change points, etc. exist and that not all ups or downs are permanent (Table 5) would have created an awareness that the absence of recent extremes does not guarantee safety in the future. Finally, spending extra time on undesirable and threatening outcomes and including third, unaffected parties (Table 5) would have reduced the effects of wishful thinking.

A second, more generic example concerns the lessons learned from past disasters or avoided disasters. The analysis of such events is frequently plagued by availability and recency biases, selective perception bias, hindsight bias which may all lead to the overgeneralization that the current risk reduction measures are appropriate. One approach of reducing these biases is to include people with different perspectives in the lessons learned study. Different people with different experience and expertise will help learn different

Table 5. Approaches for Debiasing the Flawed Human Perception and Decision Making of Table2[a]

Type of Bias	Approaches for Debiasing
Wishful thinking	Include third, unaffected party.
	Have different persons make the assessment independently.
	Spend extra time on undesirable and threatening outcomes.
Erring on the side of least drama	Allow no alarming scenarios/conclusions to go unmentioned.
Selective perception	Include people with different background and experience.
	Explain and justify reasoning and subjective assessments.
	Avoid tendency to overgeneralize from great failures or successes; failure or success may hinge on luck, timing or slight changes in the boundary conditions.
Availability bias	Present information in a balanced way including all aspects of the risk at hand.
	Remember that "no evidence of disaster" does not mean "evidence of no disaster."
Recency bias	Consider the fundamental factors that affect the process of interest.
	Remember that cycles, change points, etc. exist and that not all ups or downs are permanent.
Illusory correlation bias	Verify statistical significance and causal relationships of patterns.
	Assess the diagnosticity of evidence.
Illusion of certainty, overconfidence	Disclose all assumptions and choices in the risk assessment.
	Justify reasoning and subjective assessments.
	Search for reasons or circumstances why choice might be wrong.
Hindsight bias	Experience history more richly by attending to multiple interpretations of historical events.
	Attempt to learn different lessons from the same experience.
	Check claims about forecasting accuracy against reality.
Confirmation bias	Lay open all assumptions and choices in the risk assessment and justify reasoning and subjective assessments.
	Search for disconfirming information and attempt to falsify your hypotheses.

[a]The table is based on *Makridakis* [1988] with additional information from *Evans* [1982], *Kahneman et al.* [1982], *Sjöberg* [1982], *March et al.* [1991], *Bunn and Salo* [1993], *Nickerson* [1998], *Pidgeon and O'Leary* [2000], *Gigerenzer et al.* [2008], *Kahneman* [2011], and *Brysse et al.* [2013].

lessons from history. Disparate perspectives will help reduce selective perception, availability, and recency biases. Further, the reasoning and strength of evidence should be scrutinized. The members of a lessons learned study should be required to justify their assumptions and reasoning and document the strength (or absence) of evidence. Finally, the analysis of near-misses that under slightly different circumstances could have developed into a disaster may help augment the sparse history of actual disasters [*March et al.*, 1991; *Pidgeon and O'Leary*, 2000]. Forcing oneself to imagine how near-misses could have developed into disasters has a large potential for reducing cognitive biases in risk assessment and management.

6. Moving Out of Terra Maligna

The possibilities of reducing the potential for surprise by the approaches discussed in section 5 are limited. No matter how comprehensive the efforts, how advanced the models and data analyses, the future may still have surprise in stock for us, and the consequences may be malicious. To move out of terra maligna, we need to design risk management measures in such a way that they are able to cope with surprising events and future developments. In Figure 3, one would like to shift the events from terra maligna to terra benigna.

The key factors of designing flood risk systems in a way they can cope with surprise are robustness and adaptivity. A risk reduction measure is robust if it performs acceptably despite large deviations from the assumed development and/or despite large errors in the system understanding. In other words, it is robust if it is immune to surprise [*Ben-Haim*, 2012]. In contrast to traditional risk reduction strategies that aim at identifying some optimum option, robust strategies aim at meeting an acceptable critical requirement. A risk reduction measure is adaptable or flexible if it can be easily adjusted to changing (and possibly unforeseen) future conditions [*Walker et al.*, 2013]. Extended discussions about the concepts of robustness and adaptivity and their relationship to the concept of resilience can be found in *Lempert et al.* [2003], *Ben-Haim* [2012], *Cox* [2012], or *Walker et al.* [2013].

Concepts of robustness and adaptivity from other fields are now being adopted in flood risk management [*de Bruijn*, 2004; *Merz et al.*, 2010; *Klijn et al.*, 2012; *Lempert et al.*, 2013; *Mens et al.*, 2011, 2014]. *Mens et al.* [2014] propose quantifying robustness by three criteria of system response (e.g., economic losses) to a disturbance (e.g., flood peak): (1) the resistance threshold beyond which the impact becomes greater than 0, (2) the proportionality with which the response increases with increasing disturbance, and (3) the manageability which describes the ability to keep the response level below a point beyond which recovery becomes difficult. A defense system with a high-resistance threshold, a nonlinear disturbance-response relationship, and a low manageability is less robust than a "soft-failure" system which does not fail catastrophically but incrementally. For example, spillover segments of dikes that withstand overtopping allow controlled flooding of predefined areas that may be subdivided into segments depending on the potential consequences.

Decentralization, diversification, and redundancy can further enhance system robustness and adaptivity. *Aerts et al*. [2008] illustrate, for an area in Netherlands, that more diverse risk reduction measures are favorable when planning for an uncertain future. Nonstructural measures, such as enhancing the self-protective behavior of residents, may also contribute to robustness and adaptivity in risk management and reduce flood losses [*Grothmann and Reusswig*, 2006]. If residents are aware of their risk and of their possibilities to undertake effective precautionary measures, flood damage will be reduced, regardless of the future development of the flood hazard. While the benefits of robustness and adaptability are clear, the discussion has yet to converge and deliver indicators for evaluating the robustness and adaptivity of different system designs.

The notions of robustness and adaptivity have a lot to do with reducing the vulnerability of the system to unexpected flood disasters. Earlier in the paper, we noted that it is essential to draw the system boundaries in a way that the important processes are included. This also applies to vulnerability. In fact, the vulnerability of a society is determined by a host of factors including variations in wealth, social equality, food availability, health and education status, physical and institutional infrastructure, and access to natural resources, in addition to the technological capabilities [*Wilby and Dessai*, 2010]. There is ample evidence that, for a society as a whole, the former factors are just as or even more important for reducing vulnerability than technological approaches alone, and this may even be the case at the scale of individuals. For example, based on a comparison of hurricane effects in Cuba, Haiti, and the Dominican Republic, *Pichler and Striessnig* [2013] found that better education had clear effects on reducing vulnerability through awareness about crucial information, faster and more efficient responses to alerts, and better postdisaster recuperation, as well as longer-term effects. *Muttarak and Lutz* [2014] summarize the findings from 11 studies around the world which all provide evidence of the positive impact of formal education on vulnerability reduction. Highly educated individuals and societies were better prepared, suffered lower impact, and recovered faster from disasters than their less educated counterparts. Similarly, there are close links between institutional failure, underdevelopment, and vulnerability to disasters [e.g., *Ahrens and Rudolph*, 2006].

Making individuals and societies more robust against surprises therefore goes beyond the design of spillways and flood management plans. We need to adopt a broader view and bring in social dimensions such as culture and governance through a bottom-up or vulnerability based approach [*Blöschl et al*., 2013a] motivated by a social paradigm. Sociohydrology can play an important part in this by helping define viable pathways leading to robust and flexible strategies of flood risk management that move surprising flood events from terra maligna to terra benigna.

7. Conclusions

In this paper, it is argued that surprise is a neglected element in flood risk assessment and management and should be accounted for by risk analysts and managers. This paper develops concepts to address surprise. Figure 4 provides a checklist of the individual steps.

A request for reflecting on the potential of surprise may seem paradoxical at first sight as mapping known unknowns, or even worse, unknown unknowns may be difficult. We argue that reflecting on the sources of surprise may help understand if a situation is "surprise-prone," if things may be hidden. We may not be able to identify a priori all the failure modes of a complex system, but it may be possible to anticipate whether predictions will be realistic or surprises should be expected. Similarly, reflecting on potentially malicious consequences of surprises is important. In some instances, events with large uncertainties may not be malign, for example, when designing dikes along a low-land river where the residents in the hinterland can

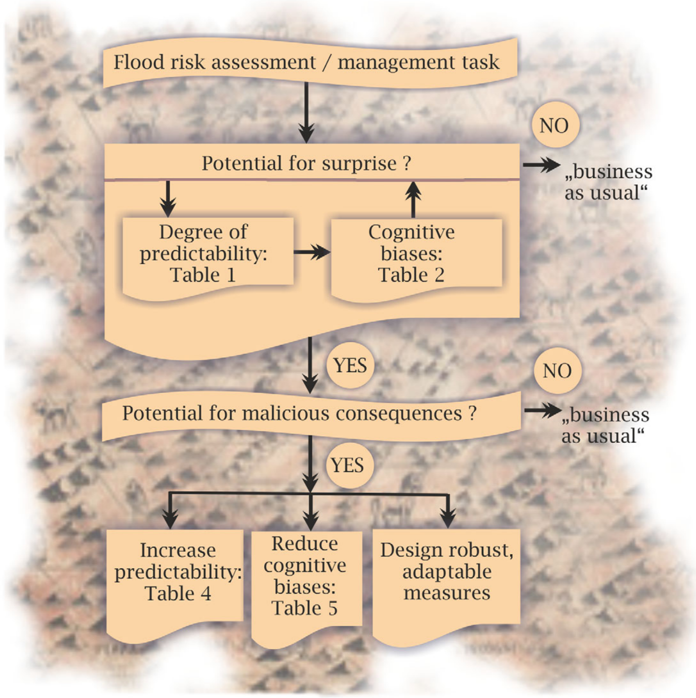

Figure 4. Checklist for accounting for surprise in flood risk assessment and management.

be warned and evacuated. In contrast, the flood safety of a nuclear power plant can pose dramatically different threats, and this difference in maliciousness should be considered.

The following conclusions can be drawn from this paper:

1. Flood risk systems often show significant complexity which limits the predictability of extreme events or future developments. The potential for surprise can be quantitatively understood by reflecting on the factors that limit predictability. This requires the analysis of nonlinearities, interdependencies, and nonstationarities in flood risk systems.
2. Formal uncertainty analyses are helpful in understanding the limits of knowledge but cannot capture surprise. They should be complemented by approaches that are specifically targeted at reducing or anticipating surprise. A number of approaches are proposed here; additional research is needed to substantiate and complete this list.
3. A second source of surprise is cognitive biases in perception and decision making. Flood risk assessment and management are particularly prone to these biases due to the rarity and uniqueness of extremes, and the nature of human risk perception. We propose that risk analysts and risk managers reflect explicitly on the potential for falling into such mental traps.
4. Debiasing approaches should be actively developed and applied to reduce flaws in perception and decision making. The cooperation between hydrologists and psychologists will be useful in these activities.
5. The question of whether unexpected events or developments could lead to malicious consequences should be an integral part of any flood risk assessment and management.
6. Surprise is highly subjective, so quantitative indicators of the potential for surprise and their malicious consequences may be difficult to define. Choice experiments could be used to rank events with respect to their degree of surprise.
7. In case surprises and malicious consequences cannot be ruled out, risk reduction strategies should attempt to move events from terra maligna to terra benigna. Robustness and adaptivity are important traits of such strategies. More research is needed on adopting a broader view and bringing in social dimensions such as culture and governance through a bottom-up or vulnerability-based approach.
8. Future generations of hydrologists and engineers should be able to understand complex systems, to embrace perspectives from natural and social sciences, and they should be well aware of unintended

effects of human decisions. Education in hydrology and water resources management needs to provide these holistic perspectives by extending current curricula with systems thinking, interdisciplinarity, and the human dimensions of hydrology.

Exploring terra incognita and terra maligna will be a clear shift in perspective from the status quo that focuses on quantifying flood risk and future failure scenarios. For example, an official flood map could mark flood-protected areas as "...not inundated unless the flood defense fails...." Such statements would acknowledge the underlying potential for surprise. Charting maps, metaphorically speaking, of the potential for surprise, to separate terra incognita from terra cognita, and terra maligna from terra benigna would go a long way toward making flood risk assessment and management less prone to surprise.

Acknowledgments
We thank the ERC Advanced Grant, "FloodChange," project 291152, for financial support. U. Lall acknowledges the support of the U.S. Army Corps of Engineers through an IPA for his contribution to this work. No data were used in producing this manuscript.

References

Adamowicz, W., P. Boxall, M. Williams, and J. Louviere (1998), Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation, *Am. J. Agric. Econ.*, *80*(1), 64–75.

Aerts, J. C. J. H., W. Botzen, A. van der Veen, J. Krywkow, and S. Werners (2008), Dealing with uncertainty in flood management through diversification, *Ecol. Soc.*, *13*(1), 41.

Ahrens, J., and P. M. Rudolph (2006), The importance of governance in risk reduction and disaster management, *J. Contingencies Crisis Manage.*, *14*(4), 207–220.

Allison, H. E., and R. J. Hobbs (2006), *Science and Policy in Natural Resource Management: Understanding System Complexity*, Cambridge Univ. Press, Cambridge, U. K.

Apel, H., A. Thieken, B. Merz, and G. Blöschl (2004), Flood risk assessment and associated uncertainty, *Nat. Hazards Earth Syst. Sci.*, *4*, 295–308.

Apel, H., A. Thieken, B. Merz, and G. Blöschl (2006), A probabilistic modeling system for assessing flood risks, *Nat. Hazards*, *38*(1–2), 79–100, doi:10.1007/s11069-005-8603-7.

Balcombe, K., M. Bitzios, I. Fraser, and J. Haddock-Fraser (2014), Using attribute importance rankings within discrete choice experiments: An application to valuing bread attributes, *J. Agric. Econ.*, *65*(2), 446–462.

Ben-Haim, Y. (2012), Why risk analysis is difficult, and some thoughts on how to proceed, *Risk Anal.*, *32*(10), 1638–1646.

Bianchizza, C., and S. Frigerio (2013), Domination of or adaptation to nature? A lesson we can still learn from the Vajont, *Ital. J. Eng. Geol. Environ.*, *6*, 523–531.

Blöschl, G., and A. Montanari (2010), Climate change impacts-throwing the dice?, *Hydrol. Processes*, *24*(3), 374–381.

Blöschl, G., and E. Zehe (2005), On hydrological predictability, *Hydrol. Processes*, *19*(19), 3923–3929.

Blöschl, G., A. Viglione, and A. Montanari (2013a), Emerging approaches to hydrological risk management in a changing world, in *Climate Vulnerability: Understanding and Addressing Threats to Essential Resources*, pp. 3–10, Elsevier, Oxford, U. K.

Blöschl, G., T. Nester, J. Komma, J. Parajka, and R. A. P. Perdigão (2013b), The June 2013 flood in the Upper Danube basin, and comparisons with the 2002, 1954 and 1899 floods, *Hydrol. Earth Syst. Sci.*, *17*, 5197–5212.

Brysse, K., N. Oreskes, J. O'Reilly, and M. Oppenheimer (2013), Climate change prediction: Erring on the side of least drama?, *Global Environ. Change*, *23*, 327–337, doi:10.1016/j.gloenvcha.2012.10.008.

Bunn, D. W., and A. A. Salo (1993), Forecasting with scenarios, *Eur. J. Oper. Res.*, *68*, 291–303.

Casti, J. L. (1986), On system complexity: Identification, measurement, and management, in *Complexity, Language, and Life: Mathematical Approaches*, pp. 146–173, Springer, Berlin.

Collier, B., and J. Skees (2012), Increasing the resilience of financial intermediaries through portfolio-level insurance against natural disasters, *Nat. Hazards*, *64*(1), 55–72.

Cosgrove, G. R. (Ed.) (1999), *Mappings*, London: Reaktion Books.

Cox, L. A. T., Jr. (2012), Confronting deep uncertainties in risk analysis, *Risk Anal.*, *32*(10), 1607–1629.

Datei, C. (2005), *Vajont. La storia idraulica* [in Italian], 137 pp., Libreria Int. Cortina, Padova.

De Bruijn, K. (2004), Resilience indicators for flood risk management systems of lowland rivers, *Int. J. River Basin Manage.*, *2*(3), 199–210.

De Moel, H., and J. C. J. H. Aerts (2011), Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates, *Nat. Hazards*, *58*(1), 407–425.

Di Baldassarre, G., A. Viglione, G. Carr, L. Kuil, J. L. Salinas, and G. Blöschl (2013), Socio-hydrology: Conceptualising human-flood interactions, *Hydrol. Earth Syst. Sci.*, *17*(8), 3295–3303, doi:10.5194/hess-17-3295-2013.

Di Baldassarre, G., J. S. Kemerink, M. Kooy, and L. Brandimarte (2014), Floods and societies: The spatial distribution of water-related disaster risk and its dynamics, *WIREs Water*, *1*(2), 133–139, doi:10.1002/wat2.1015.

Delle Rose, M. (2012), Decision-making errors and socio-political disputes over the Vajont dam disaster, *Disaster Adv.*, *5*(3), 144–152.

Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall, (2012), Ermittlung von Hochwasserwahrscheinlichkeiten [Estimating flood probabilities], *Tech. Rep. DWA-M. 552*, Hennef, Germany.

Evans, J. St. B. T. (1982), Psychological pitfalls in forecasting, *Futures*, *14*(4), 258–265.

Fahy, R. F., and G. Proulx (1997), Human behavior in the World Trade Center evacuation, in *Proceedings of the 5th International Symposium on International Association for Fire Safety Science*, pp. 713–724, International Association for Fire Safety Science, Boston, Mass.

Gigerenzer, G., W. Gaissmaier, E. Kurz-Milcke, L. M. Schwartz, and S. Woloshin (2008), Helping doctors and patients make sense of health statistics, *Psychol. Stud. Public Interest*, *8*(2), 53–96.

Grothmann, T., and F. Reusswig (2006), People at risk of flooding: Why some residents take precautionary action while others do not, *Nat. Hazards*, *38*, 101–120.

Gupta, H. V., G. Blöschl, J. J. McDonnell, H. H. G. Savenije, M. Sivapalan, A. Viglione, and T. Wagener (2013), Synthesis, in *Runoff Prediction in Ungauged Basins—Synthesis across Processes, Places and Scales*, edited by G. Blöschl et al., chap. 12, pp. 361–383, Cambridge Univ. Press, Cambridge, U. K.

Hall, J., and M. Anderson (2002), Handling uncertainty in extreme and unrepeatable hydrologic processes—The need for an alternative paradigm, *Hydrol. Processes*, *16*, 1867–1870.

Hall, J., et al. (2014), Understanding flood regime changes in Europe: A state of the art assessment, *Hydrol. Earth Syst. Sci.*, *18*, 2735–2772, doi:10.5194/hess-18-2735-2014.

Hall, J. W., and D. Solomatine (2008), A framework for uncertainty analysis in flood risk management decisions, *J. River Basin Manage.*, *6*(2), 85–98.

Hammitt, J. K., and A. I. Shlyakhter (1999), The expected value of information and the probability of surprise, *Risk Anal.*, *19*(1), 135–152.

Haraguchi, M., and U. Lall (2015), Flood risks and impacts: A case study of Thailand's floods in 2011 and research questions for supply chain decision making, *Int. J. Disaster Risk Reduct.*, doi:10.1016/j.ijdrr.2014.09.005, in press.

Henrion, M., and B. Fischoff (1986), Assessing uncertainty in physical constants, *Am. J. Phys.*, *54*, 791–797.

Interagency Advisory Committee on Water Data (1982), *Guidelines for Determining Flood Flow Frequency: Bulletin 17B of the Hydrology Subcommittee*, 183 pp., Off. of Water Data Coord., U.S. Geol. Surv., Reston, Va.

IPCC (2012), *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*, Cambridge Univ. Press, Cambridge, N. Y.

Jain, S., and U. Lall (2001), Floods in a changing climate: Does the past represent the future?, *Water Resour. Res.*, *37*(12), 3193–3205.

Kahneman, D. (2011), *Thinking, Fast and Slow*, 499 pp., Farrar, Straus and Giroux, N. Y.

Kahneman, D., P. Slovic, and A. Tversky (Eds.) (1982), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge Univ. Press, Cambridge, U. K.

Kilburn, C. R. J., and D. N. Pretley (2003), Forecasting giant, catastrophic slope collapse: Lessons from Vajont, Northern Italy, *Geomorphology*, *54*, 21–32.

Klijn, F., J. M. Knoop, W. Ligtvoet, and M. J. P. Mens (2012), In search of robust flood risk management alternatives for the Netherlands, *Nat. Hazards Earth Syst. Sci.*, *12*(5), 1469–1479.

Kuhlicke, C. (2010), The dynamics of vulnerability: Some preliminary thoughts about the occurrence of 'radical surprises' and a case study on the 2002 flood (Germany), *Nat. Hazards*, *55*, 671–688, doi:10.1007/s11069-010-9645-z.

Kumar, P. (2011), Typology of hydrologic predictability, *Water Resour. Res.*, *47*, W00H05, doi:10.1029/2010WR009769.

Kumar, P. (2015), Hydrocomplexity: Addressing water security and emergent environmental risks, *Water Resour. Res.*, doi:10.1002/2015WR017342, in press.

Kwon, H. H., C. Brown, and U. Lall (2008), Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling, *Geophys. Res. Lett.*, *35*, L05404, doi:10.1029/2007GL032220.

Lempert, R., N. Kalra, S. Peyraud, Z. Mao, S. B. Tan, D. Cira, and A. Lotsch (2013), *Ensuring Robust Flood Risk Management in Ho Chi Minh City*, The World Bank, Washington, D. C.

Lempert, R. J., S. W. Popper, and S. C. Bankes (2003), *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*, 187 pp., The Rand Pardee Cent.

Library of Congress (2015), Vniversale descrittione di tvtta la terra conoscivta fin qvi, Map by Paolo Forlani, published 1565 in Venezia by F. Berteli. [Available at www.loc.gov/item/85690919.]

Lima, C. H., Lall, U., T. J. Troy, and N. Devineni (2015), A climate informed model for nonstationary flood risk prediction: Application to Negro River at Manaus, Amazonia, *J. Hydrol.*, *522*, 594–602.

Littell, E. (1853), *Air Maps, Littell's Living Age No. 495*, 430 pp., Littell Son and Co., N. Y.

Lu, M., U. Lall, A. Schwartz, and H. Kwon (2013), Precipitation predictability associated with tropical moisture exports and circulation patterns for a major flood in France in 1995, *Water Resour. Res.*, *49*, 6381–6392, doi:10.1002/wrcr.20512.

Macgill, S. M., and Y. L. Siu (2004), The nature of risk, *J. Risk Res.*, *7*(3), 315–352.

Makridakis, S. (1988), Metaforecasting. Ways of improving forecasting accuracy and usefulness, *Int. J. Forecasting*, *4*, 467–491.

March, J. G., L. S. Sproull, and M. Tamuz (1991), Learning from samples of one or fewer, *Org. Sci.*, *2*(1), 1–13.

Mens, M. J. P., F. Klijn, K. M. De Bruijn, and E. Van Beek (2011), The meaning of system robustness for flood risk management, *Environ. Sci. Policy*, *14*(8), 1121–1131.

Mens, M. J. P., F. Klijn, and R. M. J. Schielen (2014), Enhancing flood risk system robustness in practice: Insights from two river valleys, *Int. J. River Basin Manage.*, *13*, 297–304.

Merlin, T. (1983), *Sulla pelle viva* (in Italian), Cierre edizioni, Come si costruisce una catastrofe, Il caso del Vajont, Verona.

Merz, B., and A. H. Thieken (2009), Flood risk curves and uncertainty bounds, *Nat. Hazards*, *51*, 437–458, doi:10.1007/s11069-009-9452-6.

Merz, B., J. Hall, M. Disse, and A. Schumann (2010), Fluvial flood risk management in a changing world, *Nat. Hazards Earth Syst. Sci.*, *10*, 509–527.

Merz, R., and G. Blöschl (2008a), Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information, *Water Resour. Res.*, *44*, W08432, doi:10.1029/2007WR006744.

Merz, R., and G. Blöschl (2008b) Flood frequency hydrology: 2. Combining data evidence, *Water Resour. Res.*, *44*, W08433, doi:10.1029/2007WR006745.

Morrison, M., J. Bennett, and R. Blamey (1999), Valuing improved wetland quality using choice modeling, *Water Resour. Res.*, *35*(9), 2805–2814, doi:10.1029/1999WR900020.

Müller, L. (1964), The rock slide in the Vajont Valley, *Rock Mech. Eng. Geol.*, *2*, 148–212.

Müller-Salzburg, L. (1987), The Vajont catastrophe—A personal review, *Eng. Geol.*, *24*, 423–444.

Murray, C. (2005), Mapping terra incognita, *Polar Rec.*, *41*(217), 103–112.

Muttarak, R., and W. Lutz (2014), Is education a key to reducing vulnerability to natural disasters and hence unavoidable climate change?, *Ecol. Soc.*, *19*(1), 42, doi:10.5751/ES-06476-190142.

Nakamura, J., U. Lall, Y. Kushnir, A. W. Robertson, and R. Seager (2013), Dynamical structure of extreme floods in the US Midwest and the United Kingdom, *J. Hydrometeorol.*, *14*(2), 485–504.

Nickerson, R. (1998), Confirmation bias: A ubiquitous phenomenon in many guises, *Rev. Gen. Psychol.*, *2*, 175–220.

Nguyen, M. V., D. Nguyen, N. N. Hung, B. Merz, and H. Apel (2014), Large-scale suspended sediment transport and sediment deposition in the Mekong Delta, *Hydrol. Earth Syst. Sci.*, *18*, 3033–3053, doi:10.5194/hess-18-3033-2014.

Pappenberger, F., H. Harvey, K. Beven, J. Hall, and I. Meadowcroft (2006), Decision tree for choosing an uncertainty analysis methodology: A wiki experiment, *Hydrol. Processes*, *20*, 3793–3798, doi:10.1002/hyp.6541.

Petrow, T., A. H. Thieken, H. Kreibich, B. Merz, and C. H. Bahlburg (2006), Improvements on flood alleviation in Germany: Lessons learned from the Elbe flood in August 2002, *Environ. Manage.*, *38*(5), 717–732.

Pichler, A., and E. Striessnig (2013), Differential vulnerability to hurricanes in Cuba, Haiti, and the Dominican Republic: The contribution of education, *Ecol. Soc.*, *18*(3), 31.

Pidgeon, N., and M. O'Leary (2000), Man-made disasters: why technology and organizations (sometimes) fail, *Safety Sci.*, *34*, 15–30.

Pincus, S. M. (1991), Approximate entropy as a measure of system complexity, *Proc. Natl. Acad. Sci. U. S. A.*, *88*(6), 2297–2301.

Rasmussen, E. A., and J. Turner (2003), *Polar Lows: Mesoscale Weather Systems in the Polar Regions*, 174 pp., Cambridge Univ. Press, Cambridge, U. K.

Rausand, M., and A. Hoylan (2004), *System Reliability Theory: Models, Statistical Methods, and Applications, Wiley Ser. Probab. Stat.*, 2dn ed., John Wiley & Sons, Inc., Hoboken, N. J.

Reichler, T., J. Kim, E. Manzini, and J. Kroger (2012), A stratospheric connection to Atlantic climate variability, *Nat. Geosci.*, *5*, 783–787, doi:10.1038/ngeo1586.

Schröter, K., M. Kunz, F. Elmer, B. Mühr, and B. Merz (2015), What made the June 2013 flood in Germany an exceptional event? A hydro-meteorological evaluation, *Hydrol. Earth Syst. Sci.*, *19*, 309–327, doi:10.5194/hess-19-309-2015.

Semenza, E., and M. Ghirotti (2000), History of the 1963 Vaiont slide: The importance of geological factors, *Bull. Eng. Geol. Environ.*, *59*, 87–97.

Shlyakhter, A. I. (1994), An improved framework for uncertainty analysis: Accounting for unsuspected errors, *Risk Anal.*, *14*(4), 441–447.

Sivapalan, M. (2015), Debates—Perspectives on socio-hydrology: Changing water systems and the "tyranny of small problems"—Socio-hydrology, *Water Resour. Res.*, *51*, 4795–4805, doi:10.1002/2015WR017080.

Sivapalan, M., H. H. G. Savenije, and G. Blöschl (2012), Socio-hydrology: A new science of people and water, *Hydrol. Processes*, *26*, 1270–1276.

Sjöberg, L. (1982), Aided and unaided decision making: Improving intuitive judgement, *J. Forecasting*, *1*, 349–363.

Steinschneider, S., and U. Lall (2015), A hierarchical Bayesian regional model for nonstationary precipitation extremes in Northern California conditioned on tropical moisture exports, *Water Resour. Res.*, *51*, 1472–1492, doi:10.1002/2014WR016664.

Surminski, S. (2014), The role of insurance in reducing direct risk: The case of flood insurance, *Int. Rev. Environ. Resour. Econ.*, *7*(3–4), 241–278.

Taleb, N. N. (2007), *The Black Swan: The Impact of the Highly Improbable*, 444 pp., Random House, N. Y.

Thieken, A. H., M. Müller, H. Kreibich, and B. Merz (2005), Flood damage and influencing factors: New insights from the August 2002 flood in Germany, *Water Resour. Res.*, *41*, W12430, doi:10.1029/2005WR004177.

Tversky, A., and D. Kahneman 1974), Judgement under uncertainty: Heuristics and biases, *Science*, *185*, 1124–1131.

UNISDR (2013), *Global Assessment Report on Disaster Risk Reduction 2013, From Shared Risk to Shared Value: The Business Case for Disaster Risk Reduction*, U. N. Int. Strategy for Disaster Reduct. Secr., Geneva.

Van Der Heijden, K. (1996), *Scenarios: The Art of Strategic Conversation*, 321 pp., John Wiley, Chichester, U. K.

Viglione, A., R. Merz, J. S. Salinas, and G. Blöschl (2013), Flood frequency hydrology: 3. A Bayesian analysis, *Water Resour. Res.*, *49*, 675–692, doi:10.1029/2011WR010782.

Viglione, A., G. Di Baldassarre, L. Brandimarte, L. Kuil, G. Carr, J. L. Salinas, A. Scolobig, and G. Blöschl (2014), Insights from socio-hydrology modelling on dealing with flood risk—Roles of collective memory, risk-taking attitude and trust, *J. Hydrol.*, *518*, Part A(0), 71–82, doi:10.1016/j.jhydrol.2014.01.018.

Vogel, R. M., A. Zafirakou-Koulouris, and N. C. Matalas (2001), The frequency of record breaking floods in the United States, *Water Resour. Res.*, *37*(6), 1723–1731, doi:10.1029/2001WR900019.

Vorogushyn, S., K.-E. Lindenschmidt, H. Kreibich, H. Apel, and B. Merz (2012), Analysis of the detention basin impact on dike failure probabilities and flood risk for a complex channel-dike-floodplain system, *J. Hydrol.*, *436/437*, 120–131, doi:10.1016/j.jhydrol.2012.03.006.

Walker, W. E., M. Haasnoot, and J. H. Kwakkel (2013), Adapt or perish: A review of planning approaches for adaptation under deep uncertainty, *Sustainability*, *5*, 955–979, doi:10.3390/su5030955.

Wardekker, J. A., A. de Jong, J. M. Knoop, and J. P. van der Sluijs (2010), Operationalising a resilience approach to adapting an urban delta to uncertain climate changes, *Technol. Forecasting Soc. Change*, *77*(6), 987–998, doi:10.1016/j.techfore.2009.11.005.

Whitfield, P. (1994), *The Image of the World: 20 Centuries of World Maps*, Pomegranate Artbooks, San Francisco, Calif.

Wilby, R. L., and S. Dessai (2010), Robust adaptation to climate change, *Weather*, *65*(7), 180–185.