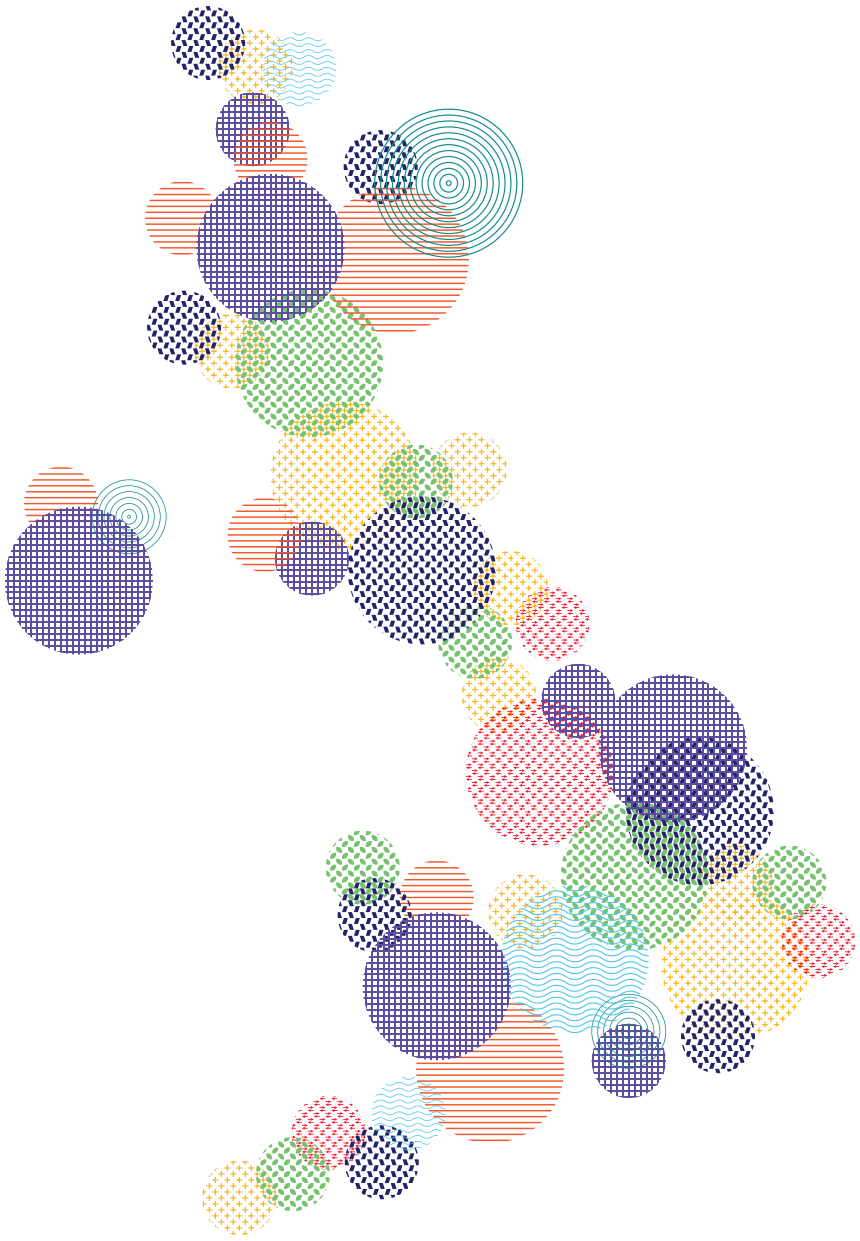
NATIONAL DIGITAL TWIN PROGRAMME

CReDo
Climate Resilience Demonstrator

CReDo Technical Paper 3:
Assessing Asset Failure

March 2022



The Climate Resilience Demonstrator, CReDo, is a climate change adaptation digital twin demonstrator project developed by the National Digital Twin programme to improve resilience across infrastructure networks.

CReDo is a pioneering project to develop, for the first time in the UK, a digital twin across infrastructure networks to provide a practical example of how connected-data and greater access to the right information can improve climate adaptation and resilience. CReDo is the pilot project for the National Digital Twin programme demonstrating how it is possible to connect up datasets across organisations and deliver both private and public good.

Enabled by funding from UKRI, The University of Cambridge and Connected Places Catapult, CReDo looks specifically at the impact of extreme weather, in particular flooding, on energy, water and telecoms networks. CReDo brings together asset datasets, flood datasets, asset failure models and a system impact model to provide insights into infrastructure interdependencies and how they would be impacted under future climate change flooding scenarios. The vision for the CReDo digital twin is to enable asset owners, regulators and policymakers to collaborate using the CReDo digital twin to make decisions which maximise resilience across the infrastructure system rather than from a single sector point of view.

CReDo's purpose is two-fold:

1. To demonstrate the benefits of using connected digital twins to increase resilience and enable climate change adaptation and mitigation.
2. To demonstrate how principled information management enables digital twins and datasets to be connected in a scalable way as part of the development of the information management framework (IMF).[1]

This first phase of CReDo running over the period April 2021 to March 2022 has focused on delivering a minimum viable product to bring the datasets together to offer insight into infrastructure interdependencies and system impact. Separate technical papers have been produced to describe each stage of the project so far:

CReDo Technical Paper 1: Building a cross sector digital twin

CReDo Technical Paper 2: Generating flood data

CReDo Technical Paper 3: Assessing asset failure

CReDo Technical Paper 4: Modelling system impact

CReDo Technical Paper 5: CReDo and the Information Management Framework

The technical papers are nested under the CReDo Overview report, and all CReDo reports and related materials can be found on the Digital Twin Hub, https://digitaltwinhub.co.uk/projects/credo.

[1] IMF - DT Hub Community (digitaltwinhub.co.uk)

Contents

Summary

Climate change is increasing the frequency with which UK infrastructure is threatened by extreme weather events. As part of the CReDo project a digital twin based decision support tool is being designed to help different asset owners evaluate the impact of such extreme events to strategically address these mounting threats.

The novel feature of this tool will be not only to provide assessments concerning the impact on the infrastructure and networks of individual asset owners, but also to communicate information concerning the functionality of assets owned by other companies - where the failure of these assets impinges on the functionality of their own. This technical report outlines one part of this project: how the functionality of individual assets within a network is affected by exposure to water during a flood event. We demonstrate how a Bayesian model of this relationship can be built from the weather data available, the science and the expert judgements of the relevant asset owners. We embody these judgements as a Bayesian network model, which links the operability of each individual asset to a specific kind of flooding incident. We show how these probabilistic models can serve as an interface between the outputs of complex flooding models and the inputs of an operational research model that calculates the knock-on effects of asset failures through the composite network. The methodology and development of this model are illustrated through the application of the model to future flood risk incidents that might happen in a region in East Anglia.

1 Introduction

Climate change is increasing the frequency with which the UK infrastructure is threatened by extreme weather events. To explore the potential impact of future climate conditions, the CReDo project is working to develop a digital twin of key infrastructure networks. This digital twin can be used to help make decisions to better protect the networks in advance of extreme weather events, and ultimately to help inform a real-time response to extreme weather events. The novel feature of this tool is that it will provide the collaborating asset owners - and also crisis management teams - with not only assessments concerning the impact of a weather-induced flooding incident in a future climate on the infrastructure and networks monitored by the *individual* asset owners, but also the *operability of assets owned by other companies* - where the failure of these assets impinges on the functionality of their own. The highly interdependent nature of these infrastructure networks, such as telephone lines relying on power supplies being operational, mean that reliably modelling the impact of an extreme weather event requires accounting for such connections. It is planned that the shared appreciation of the mutual threats described by the digital twin across the different actors will encourage further coordination between the companies in their strategic plans to mitigate these increasing threats.

This report outlines just one component of this development. We demonstrate how it is possible to elicit from asset owners the probabilities that each of their assets might fail, in a particular future flood scenario that makes consideration of the impact of climate changes on extreme weather patterns. Taking these unfolding events, and through working with teams of domain experts drawn from asset owners associated with the local power, water and telecommunication companies, our team demonstrate how it is possible to elicit probability distributions of the failure of each asset and their connections within the network. This information would then be fed to operational researchers who can calculate the knock-on effect on the whole network of each simulated future incident. From a decision-analytic perspective, the digital twin would thus consist of connected digital twins representing hydrology, the failure modes of assets, and the system in which the assets sit, with a decision support layer sitting above this.

One framework for integrating the digital twin's components is the integrating decision support system (IDSS) developed by Smith and others [1–3] where several complex dynamic probability models are knitted together into a single coherent composite. This technology has now been successfully applied in a wide range of settings - see for example [4–7]. In this case, the different groups of experts within CReDo, known as component panels, consist of climate change modellers; weather researchers, especially those specialising in flood modelling; representatives of the three asset owners whose assets would be threatened by the potential flood; and a team of operational researchers who can predict the impact of the unfolding crisis on the whole network of assets given various failure configurations.

The digital twin decision support system would inform asset owners in an intuitive manner of the risk to their own network from both a cascade of failures that a flood would induce within their

own assets, and also from failures within other related asset networks. In the first instance, the system would be designed to help inform the strategic development of asset protection, renewal and replacement by each asset owner that would be robust to the mounting threats posed by climate change. In the medium term the technological development within CReDo is planned to form a component of real time decision support tools that inform crisis teams about the impact of a threatened extreme event, update predictions about the impacts of the unfolding crisis through the timeline of the event, and evaluate the efficacy of various mitigation strategies. This function would be analogous to Bayesian support tools previously designed to evaluate the impact of countermeasure strategies in the presence of crises such as a nuclear accident, food security risks and terrorist threats [4–7] (the term Bayesian in this context means that probabilities are used to represent all uncertainties that are quantified.)

This work centres on constructing a probability model which gives a joint distribution of whether each asset in a network will continue to function at a given time T under current protocols and positioning of assets; the range of extreme flooding incidents considered, for present and future climates, is described in more detail in the hydrology report. Its outputs are then delivered to an operational research module that calculates the consequent impact on the network at that time T, accounting for the likelihood of each individual asset failing.

This flood-to-failure probability model will take inputs from two different sources:

1. **Flood models** within the digital twin provide indication of the extent of flooding in current and future climate conditions of a selected flooding scenario up to and including time T.
2. **Supplementary expert judgements** concerning the finer details of events associated with the extreme flood incidents provide information vital to the evaluation of the actual impact of the flooding. These expert judgements are especially important when mathematical models of climate or flooding do not predict aspects of the development which will be critical to determining whether an asset will be able to operate or not.

For the purposes of this work we only look at the effect of the incident on infrastructure in a particular region, as successfully eliciting accurate probabilities from domain experts requires tightly defined events.

Early in the CReDo project it was realised that within the time frames of the project – accounting for the flooding data available to the team – it would only be possible to demonstrate the digital twin of the networked assets on a single time slice representing the point of maximum disruption in the network. This simple depiction would nevertheless be sufficient to demonstrate the efficacy of the network model, and its power to communicate the potential knock-on effects that a particular flooding incident might have on the combined asset network. On its own, such a tool would already provide the framework for discussing the network-wide implications of policies within various strategic analyses. In future work, we would use this simple demonstrator as a basis to develop a full dynamic model of the failure processes and propagate their effects through the full network, exploring through time how both the flooding incident evolves and the impact of mitigating actions instigated in response.

This document reports our contribution to CReDo. In the next section we provide a formal speci-

fication of the general probabilistic component model that takes as input flooding information and predicts the failure probability of each asset within a given network. It was extremely important that this component probability model – used in the digital twin – was as realistic as it could possibly be. It was therefore necessary to first perform a series of elicitation sessions with the different asset owners to understand the way their assets might fail.

In an exploratory phase, through various video-call interviews with representatives of asset owners, we began to explore the vulnerabilities of the different assets to different kinds of flooding incident and discover some of the key features of these vulnerabilities. Some of the discoveries we made at this stage, and in subsequent elicitation meetings are outlined in Section 3. On the basis of these we were able to build probability models that embedded the generic structural forms of these explanations of failure. We discovered that a Bayesian Network (BN) provided a transparent and feasible framework around which to express this general class. A BN is a probabilistic graphical model that represents a set of variables, in this case the key vulnerabilities of the assets to flooding incidents, and their conditional dependencies via a directed acyclic graph. An example is given in Figure 3. In this report we describe the first prototype model built for asset failures – this was implemented using generic off-the-shelf BN software tools. We provide explanations of the key terms "vulnerability" and "Bayesian Network" below.

Key Terms

Vulnerability: Vulnerability in this report represents the susceptibility of an asset, considered in terms of its probability of failure, to a climate hazard. Within an assessment of the vulnerability we consider factors such as the asset's existing condition, capacity and ability to cope in the presence of the hazard.

We consider the vulnerability in terms of the asset's susceptibility, and not the criticality of the asset in the system, nor the impact of cascading failures within the system, which are considered elsewhere within CReDo.

Bayesian Network (BN): A Bayesian Network is a tool that allows us to represent the variables in a problem, and relationships between them, in a graph. A directed arrow between two variables indicates that a the value of the child ("into") variable depends on the value of the parent ("out of") variable. The absence of an arrow between two variables indicates that the two variables can be thought of as independent, given the values of their parents. By placing arrows carefully in the graph, we can display all of the important relationships between the variables.

The BN structure gives us a model for our uncertainty in all of the variables *together*. When we observe the values of any of the variables, the BN structure enables us to update our uncertainty about all of the variables (using Bayes' Theorem).

Eventually, we will need to build up models of assets as a function of their detailed type and location. Although within the time frame of this project we were not be able to perform this for all assets, we were nevertheless able to demonstrate how this would be performed for one asset facing one type of incident, albeit for a non-stochastic model (that is, for the single time slice

considered within the demonstration of the CReDo digital twin). In Section 4 we describe both how the elicited generic BN structure from Section 3 could be first simplified then embellished into a full probability model using established Bayesian elicitation techniques.

The report ends with a short description of how the techniques explored could be developed further and applied to CReDo, to provide both strategic and real-time decision support in the face of future extreme flooding events, for mitigation of their threats to a network of assets.

The development described in this document is just one part of the CReDo project, designed to investigate the feasibility of implementing a digital twin addressed to inform the progression of such weather-induced flood incidents. We demonstrate how this component could be extended to provide a digital twin – applicable to the different asset owners in helping inform their strategic plans, and eventually to support their real-time counter measures. We would like to thank all the representatives of the different asset owners for their wholehearted engagement in the many elicitation sessions that informed this study. Their contributions have been intrinsic to the positive conclusions about the promise of such methods within a composite model.

2 A probabilistic framework for assessing network operability after a flood

This section provides a formal description of the Integrating Decision Support System (IDSS) used for the digital twin modelling, which will be required by anyone looking to replicate the approach. As such, the IDSS is presented in a general form applicable to assessing the likelihood of asset failure during an extreme flooding event. Readers whose primary interest lies in the specific application, in which this general IDSS has been adapted for the specific flooding scenarios demonstrated within CReDo and detailed above, can omit this section, as the other sections can be read largely independently of it.

2.1 The formal structure of the driving process

Any underlying IDSS would be a composition of various different stochastic models delivered by various different agencies. First of all, complex climate change models would deliver collections of extreme incidents $i \in I$ that might occur because of climate change, as a function of different parameters within the models, and whose outputs would be applied to the geographical area containing the network. On the basis of each generated incident, weather modelers would then apply their modelling tools to predict how the weather consequences associated with how flooding in a future climate incidents might unfold.

Within this IDSS, our team would take these stochastic weather inputs and model their impact on the failure of each of a given set of assets J. These assets would be represented as the nodes of a network and be partitioned across three different asset owners - assets owned by a power company J_1, a water company J_2 and a telecommunications company J_3 - so that $J = J_1 \cup J_2 \cup J_3$. The network would be completed by a set of directed edges. A directed edge would be included from one asset to another - within or across different asset owners - if the failure of the receiving asset might be dependent on whether or not the donating asset at the base of the edge failed. For example, there might be an edge representing pipe-work from one asset to another in a water company or a power connection from a power asset into a telecommunications asset.

Expressed in technical terms, our remit was therefore to develop realistic stochastic transfer functions of a given time series of flooding-induced weather incident $i \in I$ to a failure type, first for each asset, then for each connection within the network. For each given flood-induced weather incident $i \in I$ the failure modelling team would receive:

1. A spatial time series $\{\boldsymbol{d}_t(l:i)\}_{t\in\mathbb{T}}$ over a grid of locations $l \in \mathbb{L}$ within the geography of assets and discrete times $t \in \mathbb{T}$ of the depth of flood - specified until a termination time of the incident t_+. These depths would themselves be stochastic functions provided by flood specialists who would use flood modelling as a tool to generate a spatial grid of flood depths. The granularity defining $\mathbb{T}$ might depend on the type of flooding incident - for example, flash

floods needing finer granularity than river floods because of the relative speeds at which these different types of incident might unfold.

2. For each extreme incident a spatial time series $\{\boldsymbol{r}_t(l:i)\}_{t\in\mathbb{T}}\ l\in\mathbb{L}, i\in I$ of other weather events that would be predicted to occur simultaneously with the flooding event, for example storms or persistent or violent rain, which would threaten the exposed parts of the network and impact on the failure of some of the assets. This could also contain information about potential ground saturation which would impact the effect of the flood.

Within this iteration of CReDo, only observations of type 1. above were available. Note here that both the number of time points $\#(\mathbb{T})$ and the size of the lattice $\#(\mathbb{L})$ are henceforth assumed finite. Below we use the standard time series shorthand for any vector time series $\{X_t\}_{t\geq 0}$ to let $X^{t'} \triangleq \{X_t\}_{0\leq t\leq t'}$.

For each incident $i\in I$ our Bayesian model would take as inputs, features of the probability mass function to time $t,\ 0\leq t\leq t_+$ of the flood depth profile $\boldsymbol{d}_t(l:i)$, and the surrounding weather conditions up to that time $\boldsymbol{r}^t(l:i)$

$$q_t(\boldsymbol{d}_t(l:i),\boldsymbol{r}^t(l:i)) = q_{1t}(\boldsymbol{d}_t(l:i)|\boldsymbol{r}^t(l:i))q_{2t}(\boldsymbol{r}^t(l:i))$$

for each incident $i\in I$ and delivered by JCEEI. Note that the usual rules of conditioning would provide all our team might need: a joint distribution of the multivariate spatial times series of $\{\boldsymbol{D}_t(l:i),\boldsymbol{R}_t(l:i)\}_{t\in\mathbb{T},l\in\mathbb{L}}$ as this applied until the end of each incident $i\in I$.

Our original remit within the context of this project was to build a BN which would take such an input time series and develop a methodology that predicted its effect on the operability of each of the assets in the network in response to the weather-induced flooding in this incident. We learned from the series of elicitation exercises we performed that one of the key features which would determine the operability of the assets as this developed over time was whether or not engineers could obtain access before or during the incident, in order to protect it when working or to repair it if it had failed.

So let $A_t(j:i)$ be an indicator variable denoting whether or not a company's engineers have access to asset $j\in J$. Let $C_t(l:i)$ denote the indicator variable of whether or not permission is granted by any incident control centre for the engineers of the relevant asset owners to have access to the given site located at l. Note that this will be a function of the predicted severity of an incident .

By definition, for a given incident $i\in I$, $A_t(j:i)$ will be a function of $\left\{\boldsymbol{D}^t(l:i),\boldsymbol{R}^t(l:i),\boldsymbol{C}^t(l:i)\right\}_{l\in N(j:i)}$ for asset $j\in J$, where $N(j:i)$ denotes those locations that impinge on access to asset j given the incident i. Notice that $N(j:i)$ will contain the location of the given asset j but also those locations whose flooding might influence the access to this site by engineers - for example the locations of highways into j that might be flooded by incident i or held open for access only for emergency services by a control centre c at a given time t. The density of the time series $\{A_t(j:i)\}_{t\geq 0}$ will be denoted by

$$p_{a(j),t}(a_t(j:i)|\left\{\boldsymbol{d}^t(l:i),\boldsymbol{r}^t(l:i),\boldsymbol{c}^t(l:i)\right\}_{l\in N(j)})$$

Secondly, an asset might fail for indirect causes, if for example a main power source intrinsic to its successfully working fails and there is no back-up to that source – for example the back-up battery fails. Let $\boldsymbol{B}_t(j:i)$ denote the vector of indicators that a back-up is available for the main source delivered by an edge into the node j in the network failing at time t. These failures will typically depend on the same vector of events in the neighbourhood of the asset and so will have conditional mass function given by

$$p_{b(j),t}(b_t(j:i)|a_t(j:i), \left\{\boldsymbol{d}^t(l:i), \boldsymbol{r}^t(l:i), \boldsymbol{c}^t(l:i)\right\}_{l\in N(j:i)})$$

Finally an asset $j \in J$ will not be able to operate at time t - denoted by the indicator variable $F_t(j:i)$ - if either because of the flood the asset itself fails directly - represented by the indicator variable $F_t^{(1)}(j:i)$ or alternatively because the asset fails because an intrinsically important networked source *into* it has failed at that time and any back-up source also fails at that time $F_t^{(2)}(j:i)$. Note that this means that

$$F_t(j:i) = 1 - \left(1 - F_t^{(1)}(j:i)\right)\left(1 - F_t^{(2)}(j:i)\right). \tag{1}$$

We then have the probability of an individual asset $j \in J$ being taken offline because of the direct effect of the flooding incident, given it can receive all the network connections it needs:

$$p_{j,t}(f_t(j:i)|a^t(j:i)\left\{\boldsymbol{d}^t(l(j):i), \boldsymbol{r}^t(l:i), \boldsymbol{c}^t(l:i)\right\}_{l\in N(j)}, \boldsymbol{b}_t(j:i), e_t(j))$$

where $e_t(j)$, $j \in J$ is the event that a source input to an asset $j \in J$ necessary for it to operate - linked by a directed edge into $j \in J$ in the network graph - has failed.

Given these probability assessments at each time step, by marginalising across other variables, we can then derive the conditional probability mass function

$$p^*_{j,t}(f_t(j:i)|\left\{\boldsymbol{d}^t(l:i), \boldsymbol{r}^t(l:i), \boldsymbol{c}^t(l:i)\right\}_{l\in\mathbb{L}}, e_t(j))$$

of failures of assets on the system as a function of the flood depths, other relevant weather events, and emergency protocols in place at the time of the incident. For each $i \in I$ writing $\boldsymbol{F}_t(i) \triangleq \{F_t(j:i) : j \in J\}$ this in turn means that we could deliver the one step ahead joint mass function of asset failures over the network

$$p^*_t(\boldsymbol{f}_t(i)|\left\{\boldsymbol{d}^t(l:i), \boldsymbol{r}^t(l:i), \boldsymbol{c}^t(l:i)\right\}_{l\in\mathbb{L}} \boldsymbol{f}^{t-1}(i))$$

given inputs of the simulated flood consequences and the crisis management protocols in place at the time concerning access permissions.

To calculate the density of the time series of failures across assets we can then simply marginalise across the delivered probability distribution of flood event protocols – the last being deterministic, but variable components of the analysis – to obtain the one step ahead probability mass function $\left\{p^*_t(\boldsymbol{f}_t(i)|\boldsymbol{f}^{t-1}(i))\right\}_{t\in\mathbb{T}}$ for each chosen incident $i \in I$. This then provides a step-by-step unfolding of the incident $i \in I$. This is sampled and fed into the operational research module, which would

calculate the impacts of the failures to the whole of the network, and study various rerouting options to mitigate the impacts of the flood. Explicitly, we can calculate and therefore sample from the full network incident of $\boldsymbol{f}(i) \triangleq (\boldsymbol{f}_0(i), \boldsymbol{f}_i(i), \ldots, \boldsymbol{f}_{t^+}(i))$ whose joint mass function will be

$$p_t^*(\boldsymbol{f}(i)) = \prod_{t=1}^{t^+} p_t^*(\boldsymbol{f}_t(i)|\boldsymbol{f}^{t-1}(i)) \tag{2}$$

In principle the process we needed to go through therefore appeared to be a lengthy but systematic one. By populating the various probability distributions with data and expert judgements associated with an ongoing incident i, and by combining these with meteorological data we could – at least in principle – provide a probabilistic digital twin calibrated to reality which could – given sufficient time – be used for strategic planning, and even crisis management decision support. However for the purposes of CReDo there were several issues that demanded that we simplify this process if we were to be able to demonstrate such a system in a short period of time.

The primary purpose of this analysis was to support strategic discussions about the most vulnerable parts of a network. A collection of flooding incident scenarios is generated from the outputs of flood models, run for current, and a sample of possible future, climate conditions. The conditional probability model we describe here would then be used as a stochastic transfer function. It would take various measures of the advancing threat in time and space, generated from the flooding scenarios investigated in current and future climate conditions above, and project these on to probabilities of various types of failure of each of these individual assets.

From this output we are able to sample the various combinations of failures across different assets owned by the three different companies for each of the flooding scenarios considered. Each instance within the sample then provides a summary description of how asset failures induced by the incident might affect the mutually vulnerable power, sewerage, water and communications assets *as a composite*, as each scenario-induced incident advances in time and space. From this, a final model depicts this network of failures, and calculates its impact on the functions of all these networks.

In this initial phase of the study it was agreed that in the first instance the digital twin would inform the strategic planning of the three companies, in a way that respects their co-dependency within any ongoing incident. The runs used in the scenarios considered by the digital twin would depict the impact of critical flooding incidents used in planning analysis, including consideration of a changing climate. Here we assumed that the current siting of assets, their defences and the protocols for remedial actions of engineers would remain unchanged. These runs would provide a benchmark from which to appraise the vulnerability and criticality of various assets owned by each asset owner to the *composite impact* across all three networks arising from these incidents.

However, the mathematical models could naturally be extended to study the potential benefit to resilience of various different works programmes, and also to provide a template for a real-time digital twin to support crisis managers in coordinating their responses to a flooding incident as it actually unfolds. In the medium term we plan to extend the digital twin so that it could provide such support; further detail on this is provided in the discussion at the end of this report.

2.2 Practical challenges and simplifications

2.2.1 Flooding inputs and models

We learned that the different types of flooding incident could be studied using weather models. However different incidents would use different models. Furthermore, despite being based on extremely sophisticated atmospheric models, standard runs of these simulation models were typically deterministic. Although it would be fairly straightforward to build statistical emulators, which on the basis of different runs of the simulation models could provide stochastic inputs that probabilistically described the development of any unfolding incident, this would be costly in both time and resources and beyond the scope of the project. The data we had available to us were single snapshot summaries of the flood which gave the maximum depth of any simulated flooding incident across a fine grid of locations.

This meant that to demonstrate the methodology we would need to build a probabilistic model only on a single time frame – expressing the failure of each asset as a function of these delivered extremes. In one sense this simplified our task since the inputs of the model were fewer. However, calibrating to a realistic development of a flooding incident was made more difficult because experts would express the effects of the flood in terms of a developing crisis rather than as a single event.

Although this snapshot would provide a very coarse tool for looking at the consequences it would still illustrate the main advantage of the digital twin – i.e. for different asset owners to obtain a transparent picture of how different the failures of the composite of assets induced by a particular type of incident, and owned by different companies, would cascade across the network. Through this a better and more comprehensive perspective of the effect of different types of incidents and the pressure points on the whole network could be explored, within a more comprehensive strategic analysis by each of the asset owners.

2.2.2 Elicitation

Even for a single time slice the task of eliciting failure probabilities across the whole network was a massive one, and not one that could be completed in the space of a few months. We therefore chose first to elicit a general framework that would define failure mechanisms of assets in general, whilst trialling its effect in simulating network failures first by populating it with dummy probabilities. We would then use information from the elicitation sessions to determine the types of information we would need to elicit from domain experts. We would proceed to trial this methodology against just one particular predicted incident, and one of the major assets that would be threatened by the flood. The progress of these elicitations is given in the later sections of this report.

2.2.3 Developing a snapshot class of graphical model

A dynamic model of the multivariate time series we described in the last subsection can be fully expressed as a Dynamic Bayesian Network. However, given the constraints of the project, an

incident that is summarised in one time slice only can be framed around a much simpler class of graphical models called a Bayesian Network (BN). Major advantages (see e.g.[8], [9],[10]) are that BNs have been tried and tested over many years and many analogous applications and thus their efficacy is well understood, and that well-maintained commercial software of various kinds is available. This graphical model enables us to embed the judgements elicited from experts within the asset owner teams. The topology of the BN – embellished by event trees – first describes the broad processes that give rise to various types of failure of various types of asset. Because the BN is fully compatible with a probability model, the framework it provides can then be used as a structure that can quantify probabilistically the likely failure of different equipment as a flood incident progresses. This is achieved by populating these by processes using elicited conditional probability tables in a way we describe.

We note that the BN module we build as part of this support tool enables us to provide a formal and auditable interface with the complex outputs of (weather and) flood models describing the unfolding incident. This is because it explicitly embeds the assertions and assumptions made by informed experts about how the failures in different parts of the system are causally linked together through the mutual exposure of assets and their connections to water. As such it is able to take into account the likely mitigating acts of the asset owners as the incident progresses. An advantage of the BN is transparency: if the relationships represented through the BN are contentious then this inadequacy will become apparent. The topology of the BN can then be adjusted to better represent the knock-on effects of the unfolding processes as these cascade through the network.

In the next section we describe some of the key issues that were elicited from asset owners concerning how assets could fail under various conditions and how they would react to protect assets or restore them to working order when they could obtain access to do this, and then describe how we can produce a generic BN of these. In Section 4, we describe how such methods can be used to generate the types of transfer function we need for our digital twin, and proceed to illustrate our methods. We end the report with a short discussion of how we are embedding these probabilistic models into the composite decision tool being developed within this project.

3 A description of the vulnerability of individual assets to flooding

3.1 The elicitation process

To specify sufficiently precisely an extreme flooding incident for this context, we first needed to understand how such incidents could influence the operability of the individual assets owned by the three different companies contributing to CReDo. We were especially interested in those types of failures that might cascade into other consequent failures – both of that asset owner's own assets and also those of the other two companies. To do this we conducted a number of elicitation exercises. These enabled us to model combinations of features of an incident that, in a generic way, might contribute to a number of assets failing. In this subsection we summarise below some of these risk features. More details are given in the appendices of this report.

The primary goal of each elicitation session was to provide an interface – that could be extended into a probabilistic interface – mapping the flood scenarios provided to their effect on each of the individual assets. To do this, expert judgments needed to be elicited about the probability of failure of an asset conditional on certain broad categories of impactful covariates that might affect this probability.

Our first task was to understand properly how unfolding events associated with precipitation, especially flooding and coincident weather events, threatened different individual assets located within the region; so how such flooding events might cascade to the infrastructure systems. To be realistic, we learned that it was critical to include in this description how these cascading events might routinely induce remedial actions by each of the asset owners, and in what circumstances they would be allowed to perform such remedial acts. Also of critical importance, was to understand similarities and difference in how the different companies perceived these threats – the important drivers and consequences of the unfolding process, which we would then digest into a single snap-shot of the most critical stage of the incident induced crisis expressed by a single BN. Using a network model, these individual component probability models would then be applied to the composite network of critical assets within the demonstration area, providing a composite probability model of the likely consequences of the weather events considered.

It quickly became apparent both through early discussions and the more formal elicitation sessions that the conditioning events relevant to the failure of each critical asset, and the failure scenario across the whole network, would not only depend on the status of assets represented in the network, but also on other events not explicitly represented in the digital twin. An important example of such external events would be blockages along the highways giving access to an asset, which could hamper the access of engineers to a compromised or potentially compromised site. The modelling decision here was whether or not to include these critical external, compromising, secondary events induced by a flood as part of the conditioning descriptions which

might embellish the various instances induced by one of the flooding scenarios or to treat such events as uncertain events averaged out across the overarching model. Given the lack of explicit information on highways the second approach was adopted.

Here, we document how we produced a probability model giving the framework for describing a joint distribution of whether or not each asset considered is contributing to the network at a single *given time* T. We initially focused elicitations on a time T chosen as the time when the network would be maximally disrupted – more specifically when the system was exposed to the highest levels modelled across the grid over the duration of the flood event. This would be most compatible with information about an incident provided by the flood modelling, and would provide a single snapshot of the incident we would use in our prototype model as discussed above. We would then sample this distribution conditional on the flood depths we received up to and including time T, repeating these snapshots for each of the flooding scenarios at collections of different times over the simulated flooding incidents.

Thus, initially we needed to elicit from the relevant asset owners what cascading events leading to the failure of each of their assets might be. We would then use the available flood depth predictions, supplemented by further elicited expert judgements about weather that associates with incidents like the ones in the scenario, to predict the joint failure events across the chosen network. Therefore our elicitation would need to focus on not only the primary, but also the secondary consequences of the critical conditioning states of the assets, which might determine whether or not these assets would continue to function under the incidents being simulated.

3.2 An elicited structural model

3.2.1 Introduction to our methods

In the first instance we elicited a general framework which was sufficiently detailed to apply generically to any asset at risk of failure though a flooding incident. We followed a now standard Bayesian structural elicitation protocol (see e.g. [9], [8], [2],
[11], [12], [13]):

1. We first elicited those *broad features* which we were told could have an impact – directly or indirectly – on the failure of an asset. These provided generic covariates describing the unfolding processes associated with a failure.
2. We then built a *foldback graph* [9] around which to structure a generic BN for a typical asset that might appear in this study.
3. We next used these schemata to first define *random variables that pass the Clarity Test.* These variables would form the building blocks of a probability model built around a BN. Note that this transformed the conceptual processes expressed by the experts into processes which could be defined – at least in principle – through stochastic relationships between features we could measure explicitly.
4. We then *drew a valid BN*. This BN provided a formal framework around which to interrogate the initial processes expressed through the foldback graph, and iteratively and interactively

improve the model of the underlying generic failure processes.

5. Where appropriate we then expressed some of the broad temporal relationships graphically using an *event tree/chain event graph* [14] to provide more details associated with the local relationships populating the variables within this BN. We then specified the typical generic levels of the random variables within this BN that would eventually be used as a template for BNs of specific types of asset in the study.
6. Only then, asset by asset, did we *customise* this probabilistic graphical model to a description of the different processes that might lead *to the failure of each of the specific individual assets* included in the study.
7. For each of these individual assets we then *populated the conditional probability tables* associated with each variable conditioned on all combinations of levels of its parents in the BN. Protocols governing these quantifying elicitations are now very well understood – see e.g. [15]. Here we adopted the SHELF method of elicitation. We note that whenever data is available these methods enable us to embed formally supporting data into these assessments in a principled way. However, the Bayesian method also allows us to embed scientific and engineering knowledge directly when no sample data is available.

Within the timescales of CReDo, we demonstrated steps 6. and 7. one a single asset, as reported in Section 4.

3.2.2 A generic description of causal pathways to flood disruption

The initial traceback graph We depict below (Figure 1) the initial traceback graph [9] we elicited from the first round of elicitations. This summarises the causal pathways traced back from the descriptions of asset failure as described by experts within the different asset owner teams. The idea of this graph is to draw out and feed back the main features and the casual relationships between them, that might take us from an incident into the failure of any asset in the affected area of study.

The directed edges into each node of this graph tell us what features might influence it, enabling us to trace back along a casual pathway to the initiating incident. Although it is wise not to dwell too long on the precise meaning of the nodes of this graph early in an elicitation process, this meaning will be vital as the model becomes more mature. We therefore provide below summary documentation of how these features were extracted from both the elicitation sessions we conducted and later discussions with experts that embellished these.

Different kinds of failure of assets on a given site It appeared from the initial elicitation sessions that there are basically two types of asset failure which can occur, expressed with the formula (1):

1. The asset is taken off-line because of a type of exposure to water that made it *inoperable*. This would include the case when an asset continues to work but does this ineffectually (for example sewage pumps working but not able to do the job intended, or not appropriately because of communication failures). It would also include the case when the asset was taken off-line as a precaution. Note that failures as we classify them here are not always

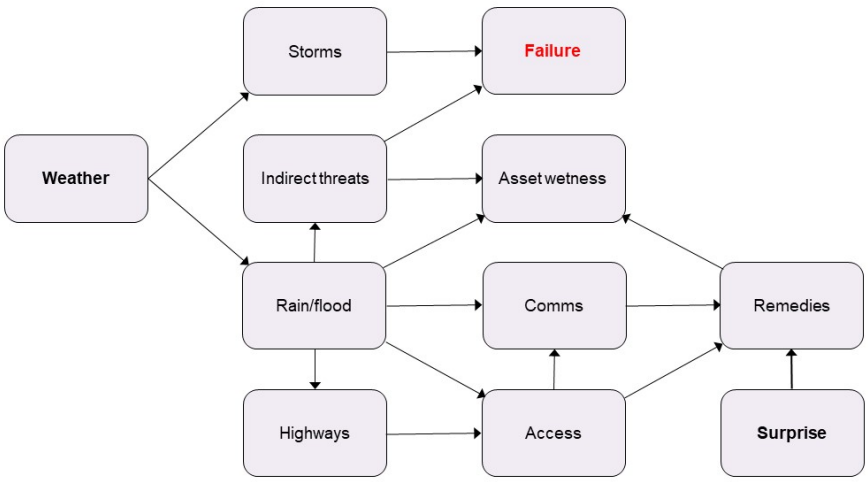


Figure 1: The initial traceback graph.

failures of the equipment itself.

2. The asset was off-line because there was *no mains power feed upstream* (from either the primary or if available secondary source) in a situation where there was no back-up power source. This last depends on the local and global topology of the network, and depends on network information provided by another module. We include in this category the possibility the asset is taken off-line because of exposure to water of *local power cabling* so that no mains power was available and no back-up – for example from a battery – was available either.

Although obviously failures could be partial, for the purposes of this study for simplicity we will always assume that assets either fail or work fully. This greatly simplified the elicitation process, while still enabling us to bound consequences of an incident. Having determined the types of failure to be considered in the study we could then proceed to work backwards from these sequentially to better understand how these failures could have come about.

Direct causes of failure Within the context of the incidents we were considering we discovered that the essential causes of failures of the asset could be classified into three types. The main direct cause of failure was the exposure of the asset to water. We later learned that any such exposure could damage electrical components in a way that cannot quickly be restored. Of course, the precise nature of the exposure – for example the location, extent and duration – that might cause an asset to fail would depend critically on the type of asset under consideration, and the routine protection provided.

Indirect threats such as run-off of chemicals into a water purity plant could cause it to close down (see below). These effects are usually mediated through the increased risk of the exposure to water.

The other major effects are high winds and lightening strikes (usually only relevant to assets owned by the power network and communications network) that could occur during the storms driving a flooding event. These associated events could bring down overhead lines providing power to an asset (which may be remote from it): the outcome of the second point above. We therefore denoted these events as *storms* in the foldback graph. Note that in this latter case – although information about the vulnerability of overhead infrastructure under different wind exposures would be essential information to the power network – information would need to come from different modelling approaches than those taken in this study. We further learned that available climate datasets are typically calibrated for one weather variable (e.g. rain *or* wind), rather than rain *and* wind *together*.

Types and extents of exposure to water/ wetness The extent of exposure to water of the vulnerable parts of an asset – in the graph this is denoted by *wetness* – would depend on the types of extreme conditions arising because of events induced by the rainfall within the incident. But of course, this vulnerability would also depend on the routinely provided protection of the asset (for example the elevation from the ground). We learned that internal components coming in contact with water would tend to be damaged, and resulting electrical fire can quickly cause the asset to fail in ways that are not quick to repair.

We also learned that protection given before and during the incident to these vulnerable parts of the asset could fail through becoming wet. One example here would be the ability to resource power through a battery in order to replace a compromised main power source (our second type of failure listed above). Protection could also consist of activities such as engineers setting up water pumps, moving assets or drying and repairing malfunctioning equipment if they could get to the site in time.

Protection, remedies and surprise The two types of protection are labelled *remedies* in the foldback graph above. Again, the nature of the protection appropriate and available to an asset would be highly asset dependent. However, generically we could state that the greater the exposure to water and the lesser the protection, the higher the probability of failure of any asset would be. As mentioned above, although the types of protection an asset could be given depend heavily on the asset considered, they could nevertheless be usefully categorised generically into two types: protection present during an incident irrespective of whether or not engineers had access to the site from remote locations; and those forms of protection that can only be provided when there is access to the site.

So, the access that engineers might have to an asset before and during the incident could have a significant impact on whether that asset eventually failed. Furthermore, the severe weather conditions envisaged could well determine whether this access was possible. Finally, if electronic communications between engineers tasked to protect the site further and the control centre were disrupted, then even if engineers had physical assess to the site of that asset this may not be enough to successfully protect or repair the asset and return it to functionality. So, in the foldback graph there are directed edges between the remedies node, the *access* node and *comms* (a node representing telephone communications).

Finally, only if the incident were predicted in the hours before it occurred could the asset protection be enhanced. So, if the incident were a surprise, then none of these mobile defences could be put in place. We therefore introduced the vertex *surprise* in the foldback graph to describe another critical component of the story. We later learned that (broadly stated) alerts for fluvial flooding are typically given about 48 hours before an incident, although these would not necessarily be acted upon. The timeline for alerts varies by flooding type. Note that this input would depend on the ability of forecasting tools to predict accurately the incident in the hours before it happened and communicate this to the asset owners in a timely fashion. So the inputs into this part of the picture are rather different from the ones here labelled weather – which predicts in a more abstract way the kinds of weather consequences a particular type of incident might have (within a future climate) as this advances over the studied region: see below.

We note that were engineers and others able to obtain instructions and timely access before or during the incident they could engage in one or more of the following activities:

- Repair or replace any water damaged equipment.
- Install back-up power.
- Relocate any currently operative mobile equipment to higher ground if failures of type 1. are threatened.
- Shield currently operative equipment with e.g. plastic sleeves or protect with sand bags.
- Install pumps to pump water away from the vulnerable parts of the asset or its on-site connections.

Electronic communication disruption This falls into two categories; - the consequences of the loss of phone lines because of the failure of telecommunications assets, and the disruption to mobile and radio signals caused by the incident. Here this is denoted by the node *comms*. It became clear from later elicitations that the comms node tended to apply mostly to smaller assets – often larger assets operated largely autonomously. Furthermore, phone land lines were usually backed up by satellite communications or radio. However, communication transmission failures would always deleteriously affect operations of stretched assets, and would certainly impede or slow down any remedial work.

Access to the site of the asset Experts tended in their general discussions to distinguish two different scenarios:

- The highways leading into the site had been blocked either by infrastructure damage or traffic congestion.
- The site containing the asset had itself has become inaccessible.

The most challenging of these bullets was the first. This was because the covariates that needed to be used to describe these processes involved information about the local highway system and its own exposure to the effects of the incident – an aspect of the problem not explicitly considered within the digital twin, so such information needed to be imported from elsewhere. The relationship to the first is expressed explicitly in the foldback graph as *highways*.

The second was obviously affected directly by the flooding incident here labelled *rain/flood*.

Note that access is useful for two different reasons:

1. Engineers can then inspect equipment and perform remedial action, and/or add new protections like encasing equipment in plastic, moving mobile equipment to higher ground before it stops working, or repair a failing component.
2. It is then possible to transport new equipment in (like batteries) and/or out (like pumped-out water or sewage).

We only later discovered that during a flooding incident engineers might need the permission from the police crisis control centre to access a site. In major incidents this could be unlikely because priority would be given to human safety over the repair of infrastructure damage.

Disruption of highways There were several problems associated with physical communications during a flood. Some of these were directly due to the flooding incident being simulated. We notice here that, despite being critical to whether or not engineers might have access to sites, we did not initially have data relevant to some aspects of infrastructure damage including:

- Loss of bridges.
- Roads made impassable because of surface water or flash flooding.
- Roads impassable because of traffic jams caused by people evacuating themselves from the area or trying to return home to protect their property.

Again, in these circumstances the police control centre may well command that no assess would be given to engineers.

Indirect threats to the functioning of assets caused by the incident Through the elicitation sessions there were certain issues – labelled as *indirect threats* on the foldback graph – that might lead to assets failing or being taken offline. The most critical of these are mentioned below. Notice that each of these is relevant only to a small subset of the assets. It was nevertheless important to include these.

- Run-off (especially of poisonous chemicals into water system from fields because of water saturation or breaching of river/dam) into site [only relevant to water purity plants being closed].
- Roads collapsing onto cables connecting to the site of the asset [only relevant to underground cable junction failures].
- Cable ducts malfunctioning and filling with water exposing cable junctions to wetness [only relevant to underground cable junction failures].

Flooding and rain events Through the design of the incidents considered in the model, the states of the vertices; indirect threats, asset wetness, communication disruption, blockages in the highway system and accessibility of the site, were all determined by the amount of flood water that was invading the area of study - here defined in the foldback graph as *rain/flood*. The major events resulting from extreme rainfall can be usefully identified as:

- Flooding due to breaching of river defences or a dam bursting – especially threatening

flood plains. Notice here that flooding is not only a function of river water level, but also the condition of the defensive walls and other infrastructure to hold back water.
- Flash flooding occurring at various sites.
- Flooding due to sea level rise at high tides and the breaching of sea defences.
- Periods of torrential rain over the asset directly exposing it to water.
- Ground saturation by site. Note here that the season would have a big impact on soil saturation – in the summer there would be much less risk. This affects the impacts of the flash flooding, the speed of transportation of the water to threatened assets and increases the threat of run-off from chemicals.

Weather The root causes comprise both the explicit nature of the weather within the incident, and the extent to which it can be forecast. Features characterising such weather causes would be:

- Storm features which were not about precipitation: these would include high winds – in particular the maximum wind speeds occurring around overhead power lines feeding the assets in question and lightening strikes. Sudden changes in temperature could also cause failure of communication cables.
- Recent history of precipitation before the incident in the month preceding the incident would be important because this would affect the saturation of the ground. This information will also be important in regions upstream of rivers passing through the region because this will affect flooding risks.
- River levels at the time of the incident would be a better surrogate for information about flooding risk – rendering the rainfall history upstream of the river in the area irrelevant to the predictions we need. Note that these river levels will depend on the season in which the incident might occur.
- Torrential rain which might lead to flash flooding near the site of the asset considered and so indirectly threaten assets. We note here that the probability of underground cable junctions failing is greater when there is very heavy rain *after a long dry spell*.

We notice in the above that the probabilities in the model will typically need to depend on the season in which the incident happens and also the time of day (for example to model likely highway congestion).

Comments on this initial phase One point that was made during the elicitation sessions was that when one of the three asset owners, or others – like the highway department – were engaging in renovation or maintenance of infrastructure at the same time as a severe flooding incident then this could have a dramatic negative effect on the resilience of the network. Although it was felt that for any strategic analysis such work could be modelled separately – modelling the structure and interconnections directly onto the network in the process of being repaired – it would nevertheless be critical to bear this in mind in any proper analysis.

The key message here was that any realistic probability model of the consequences of extreme flooding events must consider events happening around an incident, as well as simply the flood

itself. So, flood modelling of a given incident will not alone be sufficient for prediction; for realism in a full model we would need to label each incident with other covariates which capture the threat posed to the different assets. Sometimes the information provided from the additional sources would be greater in volume than the information describing the flood itself. Some of these covariates would be embellishments of the simulated incident, whilst others would be elicited descriptions of the environment of the asset and descriptors associated with the time of the incident (e.g. the season and the time of day).

Note that although we need a considerable amount of expert judgement here to embellish these mechanisms, there will be data from previous flood incidents elsewhere, from the testing of equipment, and from failure data together with flooding exercises, which could be used to benchmark the probabilities elicited from experts and in some cases to combine these formally into the model.

Using the traceback graph as the prototype discrete BN to be implemented in code, we next explored how sampling configurations of the probability distribution of failures could form a component of the digital twin of flood impact.

3.3 Embedding a prototype BN into the digital twin

Taking the discussions described in the previous sections of this report and translating them into automated computer programmes to perform this analysis on-demand is a significant aspect of the innovation of the CReDo project. While the expert discussions can be successful in drawing out a cohesive framework of causal connections described through a graphical structure, translating these discussions into a bespoke software package required careful thought and development.

The first step in this process was for the software development team to be involved in the requirements gathering process. Staff at the Hartree Centre had experience in statistical modelling and software design, and familiarity with both of these concepts was important. The language used to describe the models built through discussions with the asset owners is that of BNs, which are a rich topic in statistical literature. The software team had familiarity with Bayesian statistics and a background in using these models for simple applications. There was extensive experience in writing custom packages for new applications in data science, and implementing these across systems. This gave them sufficient background to engage with the academic leads to translate their ideas into code, and to visualise the constraints and possibilities which came from the computing resources available.

Initial discussions across the expert elicitation team allowed for knowledge and requirements to be shared across differing domains of expertise. It was important for the software team to identify the required inputs and outputs to the model, and what aspects needed to be modular. For example, the knowledge extraction in the first part of the project meant that we needed to be flexible throughout on the exact shape of the BN, so any tools used to build the BN needed to allow for adding and removing connections at a future date. The model needed to be updated through reading the knowledge graph for data on an asset, updating the corresponding nodes and extracting updated probabilities, so it was necessary that the package had this capability.

After an initial search, the `python` package `pgmpy` [16] was identified as well suited to our needs. It allows the user to construct BNs through inputting a list of probabilistic connections, and to define the conditional probability tables for each node of the network. It is predominantly designed for use with tabular CPTs, meaning those with discrete inputs and outputs. As the model consists of many of these, such as the binary presence or absence of a flood defence, this is well suited to the task. The small number of continuous variables, such as the flood depth at a location, which can take any real value, can be allocated to bins before incorporating into the network. This had the benefit of allowing the expert elicitation team to focus on flooding risks for water levels within an arbitrary range. The presence of tabular CPTs did not prohibit the use of continuous functions. If the probability of an event was described by a continuous function, it is possible to map this to discrete bins through averaging outcomes within a given bin and populating a CPT this way.

The process to produce digital representations of the information above began. As the software team had a familiarity with BNs, they could interpret designs and explanations given by the project leads and implement these. For example, if a discussion had led to learning that a flood was likely to damage an asset, but that protection would prevent this, the required structure could be defined and software implementation begun. This meant writing code to define each node and the relationships between them. As only one node type was used for each node, all were defined using the same `TabularCPD` function within the `pgmpy` package. Identifying the outputs of each node, and the other nodes which contain the data to inform the probabilities of these outcomes, is sufficient information. The software team used an iterative process of presenting proposal networks and receiving feedback, to construct networks which progressively got closer to the desired level of representation of the system.

An example BN, for a particular asset class, is shown in Figure 3, the result of several rounds of discussions with asset owners and the elicitation experts. This BN is a particular case of the generic BN structure in Figure 4, which is detailed below. The network contains nine nodes defining some of the complex combinations of factors which define whether this asset is operational. This graph displays the relationships present within the `pgmpy` model and is a direct representation of the code. This information came from discussions between the CReDo expert elicitation team and the asset owners, with no previous constraints on form, demonstrating how complicated structures can be described through careful conversations with domain experts. For example, it became apparent that there was a possibility than an engineer could attend the site and add protection to the asset, given sufficient warning. This is incorporated in the model through the "EngineerAccess" node and the "AddProtection" node, as this event is a combination of the two events.

Each node is defined by a series of probabilities of the outcomes given the input data - "WaterIngress" is defined as the probability of water entering key systems, for a series of water levels at that asset, which are in turn defined in the "SubmergedLevel" node. Table 1 shows a complete table describing this node, defining the CPT. It is necessary to define all possible input values to a node, and all possible output values. Probabilities need to be defined for all possible outputs in the case of all possible inputs, such that all combinations are defined. Probabilities within a column must sum to 1, as for a given observed water level, all possible outcomes are defined

by the node. It is easy to imagine how these tables can grow in size in the presence of a lot of possible input and output categories; this is a common issue with BNs and careful thought is required as to what form of CPT is appropriate for the problem at hand.

If information is observed, we pass this to the network. This is done by assigning observed values to the corresponding nodes in the network, and this has the effect of updating the probabilities in other nodes in the network. To return to the example in Table 1, with no information the probability of water ingress is assessed by considering the probability of flood occurring at each depth. In the presence of information about the flood depth, we can update the likelihood of water ingress. If we observe that there is no flood, we know that ingress is unlikely, but if we observe a high level of water, we can pass this to the network so there is a high probability of water entering the system. This propagates through the network, where information about the depth of water at the "SubmergedLevel" node informs the properties of the "WaterIngress" node, which further updates whether or not a "FloodFailure" is expected to occur. Flood simulations provide expected and maximum water levels at locations within their coordinates, and these can be mapped to assets. The water level would map to the BN by assigning this value to the "SubmergedLevel" node when evaluating the probability of failure for the asset. Through observations made in this fashion, it is possible to extract further information about the functioning of an asset through the combination of data points and probabilistic structure identified by domain experts.

"WaterIngress" node					
	"FloodDepth" node value (metres)				
Outcome	$x < 0.0m$	$0.0 < x \leq 0.2$	$0.2 < x \leq 0.4$	$0.4 < x \leq 0.6$	$0.6 \leq x$
P(Water ingress)	0	0.17	0.5	0.83	1.0
P(No ingress)	1.00	0.83	0.5	0.17	0.0

Table 1: An example of the tabular structure in the nodes in the BNs constructed. The values are initial placeholder values for the "WaterIngress" node in the network shown in Figure 3. The node can have two possible outcomes (either water ingress is present or it is not), and this probability is dependent on the level of water present at the "FloodDepth" node. Tables like this allow us to account for this variability to an arbitrary level of precision, as we can define as many bins as desired. These probabilities are encoded into the network in the way shown in this table. Once these are stored on the node, the model can return expected outcome probabilities given the observations assigned to "FloodDepth".

The final stage is to combine the information in the knowledge graph, incorporate it into the BN and extract failure probabilities based on the information. Data was stored in the knowledge graph describing key information about the assets, and this was accessible to the graphical model. The model is designed to both function as an independent piece of software, capable of evaluating probabilities in a self-contained way, and to run within part of a larger workflow of composite components. The workflow model created by colleagues at DAFNI consists of chaining together many similar components, so each component must adhere to common input and output forms. In the code, this meant that software was written to assume access to a `json` data store containing information updated about all assets, and to save a file in the same format as an output. This output is read by external components to update the central knowledge graph before progressing along the workflow. This single, consistent reference point was important to make the model part of a pipeline of individual components. Storing data in this format meant that this BN is unaffected by the particular approach used to generate the flood data, or any further analysis with the output

of the model. Instead, it is only required that data is stored within this central database in a consistent fashion, allowing a modular software design. It is simple to swap the probabilistic graphical model used; if it is later established that there is an alternative model which captures more of the phenomenology, we can change this with no impact on any other components of the system.

With the information about the assets available, the network models built using `pgmpy` can be used to calculate updated probabilities for each node. In the case of the network in Figure 3, the node of interest is "AssetFunctioning", as we update the knowledge graph with the probability that the observed information leads to the asset failing. However, the network consists of many variables forming a large joint distribution of probabilities, and we need to marginalise over these to get to the distribution of interest. We extract the probability distribution for one node using the `VariableELimination` algorithm. By specifying the node of interest, the algorithm systematically marginalises out variables from the joint probability distribution until only the relevant marginal distribution remains.

```
from pgmpy.inference import VariableElimination
infer = VariableElimination(asset_model)
probabilities = infer.query(["AssetFunctioning"])
print(probabilities)
```

```
Finding Elimination Order: : 0%|          | 0/8 [00:00<?, ?it/s]
Eliminating: BackupSupply: 100%|##########| 8/8 [00:00<00:00, 253.17it/s]
+-------------------------+-------------------------+
| AssetFunctioning        |   phi(AssetFunctioning) |
+=========================+=========================+
| AssetFunctioning(True)  |                  0.8870 |
+-------------------------+-------------------------+
| AssetFunctioning(False) |                  0.1130 |
+-------------------------+-------------------------+
```

Figure 2: A code snippet demonstrating the querying function within the `pgmpy` *package. This shows the commands to extract the probability distribution of the "AssetFunctioning" node in the box at the top, with the outputs shown in the lower part. Running the commands above assumes the absence of any further data, relying on the information present in the node CPTs. Additional information about nodes in the network can be passed to the* `infer.query()` *command to allow evidence to be incorporated into the BN.*

For integration onto the DAFNI platform, the code required to run and evaluate the network was written into a Docker container. Docker containers contain the information which defines the software environment required by the code, and a series of commands to run. When these containers are created, they build a version of the required environment and run the specified commands. A container was defined with the required python environment, location of the code, and a `python` script which queries the knowledge graph and outputs failure probabilities. This was uploaded to the DAFNI platform, able to be run at any point. The complete codebase is available in the BN repository created for this work [17].

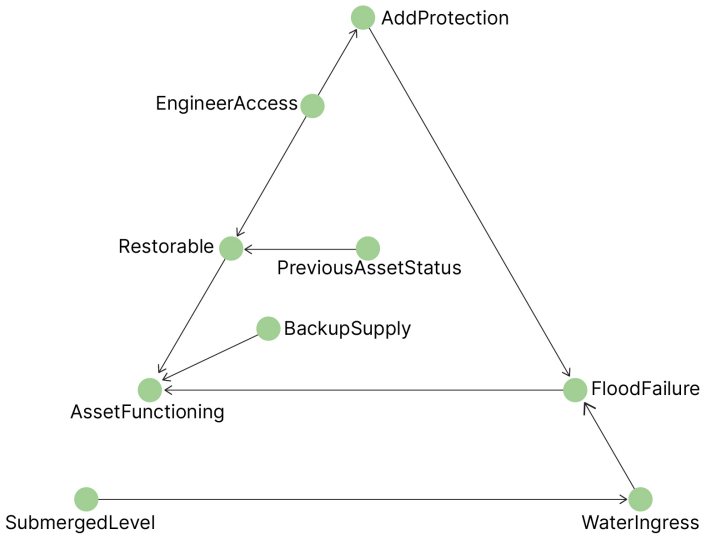


Figure 3: Visualisation of an example BN built using the `pgmpy` *package and visualised with the* `networkx` *package. All causal relationships present in the model are displayed. For an example of the probabilities contained on a node, see Table 1.*

3.4 From a generic traceback graph to specific BNs

3.4.1 Building probability models around the foldback data

Having elicited the coarse structure of the causal mechanisms underpinning these processes, we now examine each of the variables depicted in the graph to define measurement variables (variables which can, in principle at least, be measured/observed). We then draw a BN whose topology is related to the foldback graph but whose vertices are formally defined. This enables us to use the BN to interrogate the heuristic model above and adjust this as necessary.

Here it is necessary to work in the reverse order to the one we used to build the foldback graph: this time working up the casual pathways from the founder nodes until we reach the failure event.

Once this is completed we will be ready to populate the CPTs of the BN, and so specify the numerical form of the probability distribution of the failure events of each asset. Sampling from this distribution will then enable us to emulate the probable consequences on the mutually dependent collection of assets, for the whole area of study, and for each simulated run of a given incident.

From this point, because the causal pathways associated with the failures of different types of asset could be very different from each other, it was most natural to build a BN for each of class of assets in turn. Some of these BNs would have a very simple topology, while others would be much more complicated. However, they mostly share those parts of the description that define the impacts the flood has on the environment around the assets at different locations – those associated with the spread of the water, the geography of the region, and the external infrastructure and highway configuration, will all apply equally well to one asset as to another. There

is however an issue of determining which descriptive measures that might be imported into the system – both concerning the specific incident and also the surrounding environment – are the most appropriate when tracing that incident to its impact on the asset failures in the network.

The critical issue here is to ensure that we can define the variables in such a way that the BN $\mathcal{G}$ is a *valid* description of the domain for each asset considered see e.g. [10], [9]. The graph needs to be interrogated with the experts to check this, and if this is not so its topology must be adjusted until it is.

To be valid, for all nodes in the directed graph not downstream of one of its variables, X must be independent of those nodes given its parents – i.e. those nodes in the graph connected into X by a directed edge. This needs to be true for all nodes X. A careful description and illustrations of how one can work with panels of experts and iteratively modify $\mathcal{G}$ is given in [9]. This process is a particular example of a structural elicitation, see e.g. [11]. We give examples of two such elicited BN's below.

3.5 Later cycles of elicitation embedding asset information and emergency response

Later cycles of elicitation enabled us to interrogate the broad structure of the BNs and to precisely define the random variables in the model. There were three types of elicitation that now need to be conducted.

1. Within the CReDo project, another team had been systematically forming *inventories* in the form of spreadsheets of different types of asset using a common template across the three companies, listing *vulnerabilities of each asset type* to failures associated with flooding. This precious source of information enabled us to critique and embellish the generic BN described in this section, so that the variables within it could be more precisely defined in terms of measurements and unambiguous categories. It also gave us a reliable framework on which to build a BN for an illustrative asset, which would build into this description other features concerning the vulnerability of a specific asset associated with its location and installation
2. It was clear that the functionality of the system would critically depend on *the way that different crisis management teams within each company would routinely respond to a flood alert and a flooding incident*. We therefore interviewed a team of such professionals from one of the companies. We learned about the specific ameliorating actions that would be possible and how these were applied. These have now been embedded in the descriptions of flooding events given in Section 3. One of the key generic points that we were able to elicit was that the implemented preventative countermeasures would typically depend upon the human resources available, and that priority would be given to those assets whose failure would cause the worst immediate impact to customers. We noted that this policy was not one optimising the functionality of the network, but rather minimising the disruption to the company's customers, appropriately prioritised; these two criteria could be very different. One take-home message was that even if there was time to further protect an individual

asset, this might not happen simply because of available resources – so probabilities of extra protection being applied to any non-critical asset were likely to be considerably smaller than might otherwise have been the case.

3. In the final series of elicitations concerning a specific asset, various generic issues were drawn out. The most critical one was that we belatedly learned that once the crisis began *engineers would have to receive permission from a police crisis control centre* to be given access to repair or replace equipment to return their equipment to operability. Because the centre would typically prioritise human safety over the fast return to service of infrastructure, this would inhibit the speed at which engineers could respond even if repairs were technically possible. This point – although not critical to the specification of the present snapshot demonstrator – would be critical to any subsequent dynamic BNs as would other issues they raised (see Appendix B). It could dramatically slow down the return to normal working of each asset. Other remarks pertinent to our snapshot BN have been embedded into the discussion above.

All these issues would need to be addressed within the digital twin for this to provide a high-quality analogue of the effects on a network of given flooding events. In particular, we needed to modify the original BN and to transform this to the version provided in Figure 4.

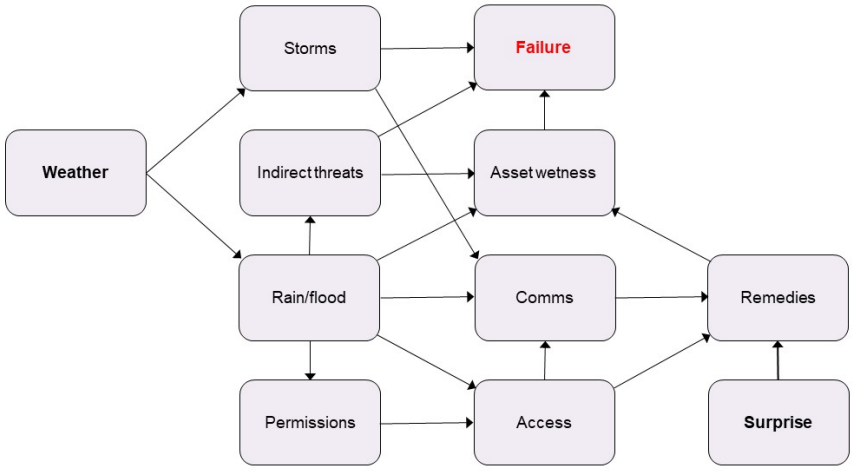


Figure 4: The modified generic BN for the probability of failure of an asset in an extreme flooding event.

4 Eliciting a BN of a specific asset

4.1 Introduction

There were two reasons to dig down during a single well-defined incident at a specific asset. Through this use case we could further check whether we had missed any critical generic features which would challenge the credibility of the prototype digital twin and the generic BN describing how failures might respond to certain extreme flood scenarios. It would also give information about how easy it would be to populate the BN with well-calibrated probabilities (at least in its snapshot form).

The initial model we had built into CReDo was naive. It was deterministic and assumed that an asset would fail if the maximum flood depth exceeded a particular threshold depth at the site of the asset. However, the real failure process was much more nuanced than this. We had demonstrated how a stochastic version could improve on this type of model. Here, by eliciting expert judgements about how specific assets might fail during certain incidents and embedding this information in a bespoke BN, we could demonstrate how the probabilities within the stochastic model could be calibrated to real expert judgements, and so better predict the real impact on the network of selected extreme incidents. In particular, by examining a particular scenario in detail we would be able both to draw out additional features more clearly and also to build in a better understanding of the mechanisms that might lead to failure that would then help frame the probabilities of such events.

To illustrate how the original generic model could be calibrated to each pair of asset and incident, we chose to generate the flooding conditions associated with a single event. We then identified a single asset which was vulnerable to failure when exposed to this event, and forensically investigated the different ways this asset might fail. The chosen asset was a pumping station located north of the town, and the flooding event one which threatened to break through the flood defences of the nearby river. We would then proceed to demonstrate how the BN of this asset could be adapted to its threat, and how standard Bayesian methodologies could then be applied to elicit the probabilities needed for its CPTs to quantify its probabilities of failure under a number of scenarios. The only probabilities we would elicit would be those related to a specialist team drawn from the water company.

We first briefly outline the technology harnessed to elicit the probabilities need to embellish the BN of this problem, and hence deliver a quantification of the uncertainties in the flood as these applied to the given asset.

4.2 A methodology for eliciting probabilities

The probabilities required for the elicitation in this case are those that are required for all of the CPTs in a particular BN. In general, quantitative expert judgements should be elicited in

a structured, transparent and reproducible manner. To ensure that our probability elicitations satisfy these properties in this project, we utilised the Sheffield Elicitation Framework (SHELF) for the elicitation sessions. SHELF is an elicitation protocol which aims to elicit probabilities from a group of experts in such a way that the result represents a consensus of the judgements of the experts; they are owned by a synthetic expert, known as the Rational Impartial Observer (RIO). The session itself comprises two stages: elicitation of individual probabilities from the experts, and elicitation of the group, or consensus, probabilities, arrived at through facilitated discussion amongst the experts.

By probability in this context we mean the subjective beliefs of a particular individual about how likely an event is to happen. Therefore, given the same information and asked for the probability of the same event, two individuals may provide different values based on their own knowledge and expertise. Hence the need to come to consensus via discussion.

Consider the probability that a particular asset fails in a particular future flooding scenario. Interest lies in the value of an indicator variable I which is equal to 1 if the asset fails in a particular scenario and 0 if it doesn't fail (or some other binary event in a CPT) . Then the entire probability distribution for the failure of the asset is captured by a single probability, that is $\Pr(I = 1)$, the probability that the asset fails (and the probability that the asset does not fail follows as $\Pr(I = 0) = 1 - \Pr(I = 1)$). This necessarily captures all of our uncertainty on whether the asset will fail on a particular occasion, both our epistemic uncertainty about the underlying proportion of similar occasions on which the asset would fail, and our aleatory uncertainty about what would happen on this particular occasion.

More comprehensively, we could consider these two aspects separately. That is, we could ask the experts to consider a large number of very similar future scenarios of the same type as that being considered in the same location, and ask for the experts' beliefs about the proportion of occasions on which the asset would fail, which we call p. Questions about the three quartiles of this distribution, for example (the values for which the expert gives probabilities of 0.25, 0.5 and 0.75 to p) would then allow us to fit a suitable continuous probability distribution to the proportion, for example via least squares. We would then obtain $Pr(I = 1)$ by integrating over this distribution, i.e.,

$$Pr(I = 1) = E[p] = \int_0^1 f_0(p)dp$$

where $f_0(p)$ is the prior distribution for p. Given the time constraints in the CReDo project, we chose to elicit $Pr(I = 1)$ directly, although we could explore the more comprehensive strategy via $f_0(p)$ in future iterations.

Another important aspect is that CPTs are structured as categorical variables, such that the probabilities in each column must sum to one. For example, we may consider the probability that a particular flood defence is fully working, partially working or failed under various flooding conditions. This means that we cannot elicit all of the probabilities independently, as it would be unlikely that they will satisfy the sum to one constraint, especially with larger numbers of categories. Instead, we chose to elicit all but one probability in each column of each CPT, and report the resulting final probability (given by one minus the sum of the others) to the experts to check

that this is consistent with their beliefs. If not, then other probabilities in the column would need to be re-elicited.

Given the time constraints in the CReDo project, and the large numbers of probabilities required to populate the CPTs, we used a "simplified" version of SHELF. We used a spreadsheet-style data entry rather than the more interactive elicitation tools available through SHELF, and used the final probability in each column as the sole check of each set of probability statements. A full SHELF elicitation would take place more comprehensively in a future iteration of the CReDo project.

4.3 Modelling of a specific asset and incident

4.3.1 Refining the generic structural BN to a critical asset

Before we elicited any probabilities, we first examined the original BN to check whether and how well it described what might happen to the chosen asset in the chosen incident. The additional generic information this elicitation extracted has already been summarised in an earlier section. Here, we therefore focus on the information relevant to this particular asset and incident. For the next version of the development of the digital twin we would envisage calibrating all the most integral assets, and customising failure probabilities to the specific relationships between their function, the location and siting of their most vulnerable components, and the cascade of events that would eventually lead to failure. The following information about the given asset were critical in assessing its probabilities:

1. Pumping stations can continue to function mechanically but nevertheless be ineffective in the role that they are designed to do. In this sense it is important to decide when the asset can be considered to fail for the purpose of the network for this type of asset. Here, through discussions with experts it was decided to consider both events – it being ineffective and it mechanically failing and the corresponding joint probabilities.
2. Because of the nature of the chosen extreme incident and the location of this asset, there was in fact a high probability that the pump would be overwhelmed. So, in this context the most important probabilities to elicit were the probabilities of mechanical failure given the pumps were overwhelmed.
3. Pumps like these are built to withstand exposure to water. So, when a power supply remained in place the failures of the pump caused by the flood would be mainly due to the electronics inside the station being exposed to water so that the circuitry was fried. One dominant set of probabilities would be for the train of events leading to such events
4. For water to reach the circuitry within the main plant it must first pass through the dry well surrounding the plant.
5. For this to happen the flood water must first break through any barrier protecting the station.
6. Another type of failure would be caused by the electronic junctions on the transformer feeding the asset being exposed to water.
7. Extra protection could increase the effective overtopping height from 900mm to 1200mm - although this extra protection is not completely reliable and might not be available.

8. If power into the plant were disrupted back-up diesel generator and batteries are in place. However, either of these may be unavailable because of theft.
9. If the back-up power is used when the vulnerable parts of the asset are exposed to water downstream of the alternative power source, then the asset will continue to fail.
10. The mechanical failure of the back-up and operation of this plant is unlikely to be affected by a loss of communications because all systems can work on an automatic default.
11. All the vulnerable parts of this asset are located inside it, and so are unlikely to be disrupted by storm damage.
12. Access to perform remedial work during an incident like this is unlikely because access is likely to be needed to be made via other affected areas, and because during such incidents such access is unlikely to be granted.

4.3.2 Events of interest

The company would usually be given 48 hours' notice of an incident. We chose to bin predicted maximum depths of flood water into the categories

$$D = \{\{0\}, (0, 500], (500, 900], \{> 900\}\}$$

because we learned that the permanent flood barriers if fully secure would provide protection up to a depth of 900mm. Flooding on the site which was significant but not overtop the barriers and might seep through was thought to be well expressed as flooding in the range $(0, 500]$.

Here, we denote the probabilities properly delivered by the water company by p, whilst the others denoted by q. Within the context of the type of incident studied, we elicited the following:

1. The probability that the pump was overwhelmed given different levels of consequent flooding at the site $p_o(o)$.
2. The probability water would breach the barriers and trespass into the wet well, and then the wet well would become full and water seep into the inside of the building, given a breach of the barriers, the effective barrier height and the pump being overwhelmed, for different depths of the floodwater at the site $d \in D$, $p_b(b|h, d), d \in D$.
3. The probability that the electronics are fried, given water in the building $p_e(e|d)$, for different depths of flood $d \in D$ when the pumps are overwhelmed.
4. The probability that the main power supply upstream of the station is off, p_u for generically measured flood conditions f.
5. The probability that the transformer into the plant will be compromised $q_t(d)$, $d \in D$ given the different depths of flood, so the main power supply is cut off.
6. Given an upstream power failure (probability $q_m(f)$) the probability that the back-up fails to operate p_c when the asset would otherwise be functioning.

Because many of the possible causal features leading to failure under this incident are not relevant for this asset, the BN can be much simplified such that the CPTs for the failure events are elicited below. Notice here that all the inputs associated with the node defined here as flood depth would, within the IDSS digital twin, be obtained from flood models, while some of the probabilities

associated with the presence of back-up power will be the domain of the power company. So, these probabilities would not be part of the elicitation. The simplified BN used for this asset is provided in Figure 5.

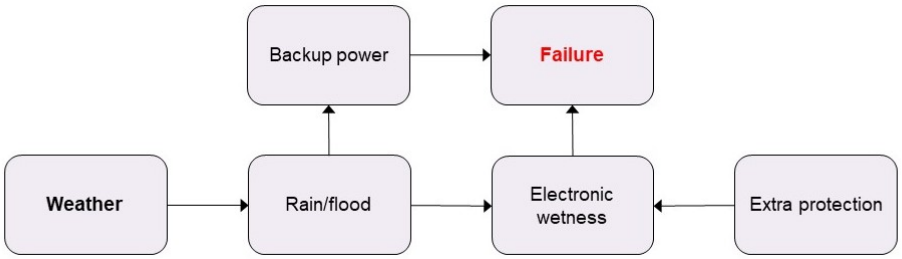


Figure 5: The BN adapted for the probability elicitation at the town's pumping station.

The inputs to the digital twin with its current architecture require us to calculate the probability $p^*(d)$ the asset will mechanically fail and be overwhelmed by the incident given the various depths of flood $d \in D$.

In this case using the definition above, this is the probability of the pump being overwhelmed and the asset having an electrical failure $p_1^*(h,d)$ or when there is a power failure into the plant that cannot be remedied p_2^* when the asset would otherwise be functioning. The usual rules of probability therefore give us that

$$p^*(h,d) = p_1^*(h,d) + p_2^*(h,d)$$

where, by the usual rules of probability and the definitions above

$$p_1^*(h,d) = p_o(o)p_b(b|h,d)p_e(e|d)$$
$$p_2^*(h,d) = p_o(o)(1 - p_1^*(h,d))(q_m(f)p_u + q_t(d))$$

Note here that because in the setting described by the incident the probability of the pumps being overwhelmed was very high, $p^*(h,d)$ is also approximately the probability the asset mechanically fails; i.e. the probability when the flood no longer inhibits the proper working of the pump - the pump nevertheless can no longer do its job.

4.4 The results of the probability elicitation

4.4.1 Overview

There were four experts present for the elicitation, all employees of the water company taking part in the CReDo project. For each probability detailed above, an individual elicitation was conducted, followed by a group discussion and then the elicitation of the consensus probability representing the RIO. The elicitation session lasted for two hours.

A briefing document was provided to the experts prior to the session, and is given in Appendix D. The session began with a summary of the information in the briefing document, followed by a training elicitation question considering the distance between Glasgow and Edinburgh. When the experts were happy with the process, we began the elicitation of the quantities of interest (QoIs) detailed above. A break was provided in the middle of the session.

The time constraints were tight given the number of probabilities to elicit, and so the cut down version of SHELF described earlier was used to elicit the probabilities. The individual probabilities and the consensus probabilities from the elicitation are provided in the top panel of Table 2.

The scenario considered in the elicitation, was defined in the briefing document as

A 48-hour warning of a convective storm surge around the town is given during the summer. This is predicted to last for up to 23 hours with higher peak intensity around the centre of the storm. This will occur when there is an above-average high astronomical tide and sea level rise due to climate change. It is predicted that this occurs concurrently with associated wave overtopping of coastal flood defences and high tide in the river. There is a risk that the level of the river will be sufficiently high to threaten overtopping, although it is unlikely to be overtopped in the 23-hour period. We are interested in the threat to the pumping station.

EA flood information is made available only after the incident has happened.

4.4.2 Rationales for the values chosen

The probability that the pumps were overwhelmed given flooding at the site, $p_o(o)$, was considered first. The experts felt that, within this scenario, the flood depth outside the site would not affect this probability, since when there is standing water of any depth building up around the site, then this indicates a flooding event which is very likely to overwhelm the pumps. The only time this would not occur would be when the flooding did not then infiltrate the network. Thus, the consensus probability chosen was high.

The second probability elicited from the experts was the probability of a complete pump failure in the scenario. In this case, the experts did believe that this probability would change with the changing flood depth. They drew the distinction between a flood depth of below 900mm and above 900m. For a flood depth below 900mm they gave the probabilities in the row p_f in Table 2. In particular, they felt that if the protected barrier was not over-topped then a failure was not inevitable, although this indicates that there would be a lot of water moving around near to the station (e.g. the height of the river and the amount of water moving through the sewers), and water typically finds a way in. For a flood depth of more than 900mm it was felt that, with the water over-topping the barriers to the station, a failure was inevitable, as long as the barriers were over-topped for a reasonable length of time. Thus, only a consensus probability is reported for this situation.

The experts then considered the probability that the back-up generator is unavailable in the event of being called to operate. Call this $p_g(g)$. The generator is automated and so no access is required. The transformer would have failed to call this back-up generator. However, the trans-

former is off site without flood protection and the back-up generator is onsite protected by the flood defences at the site. The experts considered the half a dozen or so reasons a generator wouldn't start, and data on generator starts from within the company. Taken together, this led to the consensus probability chosen.

The elicitation then moved on to the probability that the electronics are fried given that there is water in the building, p_e. The original question was termed in terms of a "fire", although from the discussion it was clear that frying the electronics was what this referred to. Again, the experts felt that this would be similar irrespective of the flood depth outside the station, given that water had somehow made its way into the station. There was relatively strong disagreement in the individual expert probabilities for this question, ranging from 0.05 to 0.7. In the discussion, the expert who gave the highest probability provided information about the relative frequency of electrical fires in similar flooding events previously, and the consensus converged to their probability.

The final two probabilities considered were the probability that, when called for, the batteries have been stolen from the back-up generator, p_s, and the probability that the diesel has been stolen from the back-up generator, p_v. There was a feeling in the group that one of the experts held almost all of the knowledge on these probabilities, and so we moved straight to consensus probabilities for these events. The two types of theft happen separately and so it seemed reasonable to the experts that they could be deemed independent. A combination of the information on the frequency that generators are checked on average (around once a month) and the number of occasions in recent years that diesel and batteries have been found to be stolen, led to the probabilities specified.

Top panel: Probabilities elicited directly					
Probability	Expert 1	Expert 2	Expert 3	Expert 4	Consensus
$p_o(o)$	0.9	0.99	0.99	0.7	0.95
$p_f(f \mid h = 900, d \leq 900)$	0.5	0.35	0.1	0.5	0.425
$p_f(f \mid h = 900, d \geq 900)$					≈ 1
$p_g(g)$	0.8	0.8	0.85	0.85	0.8
$p_e(e)$	0.7	0.25	0.05	0.2	0.7
$p_s(s)$					0.02
$p_v(v)$					0.05
Middle panel: Probabilities calculated from those elicited					
$p_c(c)$					0.26
$p_1^*(h = 900, d)$					0.66
$p_2^*(h = 900, d \leq 900)$					0.10
$p_2^*(h = 900, d \geq 900)$					0.27
Bottom panel: Probability of pumping station failure					
$p^*(h = 900, d \leq 900)$					0.76
$p^*(h = 900, d \geq 900)$					0.92

Table 2: The individual and consensus probabilities elicited during the probability elicitation session (top panel), the consensus probabilities calculated from the elicited probabilities (middle panel) and the estimated probabilities of pumping station failure (bottom panel) for different flood depths.

4.4.3 Estimates resulting from the elicitation

From the consensus probabilities elicited, and provided in the top panel of Table 2, we can calculate some other probabilities of interest. They are provided in the middle panel in Table 2.

The first is the probability that the back-up fails to operate when called on, p_c. This takes as inputs the probability that the generator starts successfully (with all parts present), p_g, the probability that the batteries have been stolen from the generator, p_s and the probability that the diesel has been stolen from the generator, p_v. Assuming independence between these probabilities gives the probability p_c. We see that the back-up generator will work when called approximately three quarters of the time.

We can then find $p_1^*(h, d)$, the probability of the pump being overwhelmed and the asset having an electrical failure. None of the individual probabilities that make up this probability were felt by the experts to change with flood depth (in this scenario for this pumping station), and so this probability does not change with the flood depth. We see that the probability is relatively high, at around two thirds, which is consistent with the view of the experts that a scenario such as this, with standing water outside the station, would represent a severe challenge to this pumping station, in terms of keeping water out.

For the probability $p_2^*(h, d)$, that there is a power failure into the pumping station when the station would otherwise be functioning, we require estimates for the probability that the transformer into the plant will be compromised, q_t and the probability of a power failure prior to the transformer, $q_m(f)p_u$. These were not elicited, and so we choose some illustrative values here to allow calculation of the probability $p_2^*(h, d)$. In practice, we would ideally elicit these values from the power network operators. We choose values of 0.3 and 0.8 for the probability that the transformer is compromised given flood depths of less than 900m and more than 900mm respectively, which are relatively high as a result of the lack of permanent flood defences at the transformer. We suppose that an upstream power failure is unlikely in this scenario, and so give it a probability of below 2%. The resulting values of $p_2^*(h, d)$ for the two different flood depth ranges are provided in the table.

Using the formulas in Section 4.3.2, we can then evaluate the overall probability of failure of the pumping station. This is given, for the two flood depth ranges, in the bottom panel of Table 2. We see that the probability of failure is high, irrespective of flood depth. This is because the failure event which is most likely is an electrical failure, and this would be likely as soon as water entered the pumping station, which in this extreme event, and given the location of the pumping station, would be likely no matter how high the standing water outside the station.

5 Recommendations

We have been able through this study to demonstrate how Bayesian methods can be applied to set up a well-calibrated digital twin. There were a number of lessons learned from this exercise that would frame the next stage of development of a digital twin.

1. Much more realism could be obtained through building a dynamic version of the simple model we have described above, which would have more use for strategic planning as well as real time decision support.
2. One critical omission of network information into the digit twin described here was the network of highways that provide access to the assets. This would be critical to any appraisal of the severity of the consequences of the flood, due to the importance of access for restoration of assets to service once they have failed. This would be especially important for any dynamic extension of the tools we describe here. We would recommend the inclusion of highway networks in any next phase of this development. One issue here will be availability of data on roads – for instance is enough detail on topography available to determine where roads actually flood (e.g. in dips), is there good data available on drainage systems for roads, and how would we model the likelihood of drainage systems not performing as they should?
3. A realistic assessment of the effects of a flooding incident needs to bring into the conversation other events associated with an incident that might be unfolding, because failures in the system might well be caused by other events, such as high winds causing power supply cables to be lost. Although not common in standard flood risk analyses, we would urge that such coincident events be considered. This may require richer data to be associated with the design of storm events beyond precipitation, and climate models to be calibrated for coincidence of rain, wind etc.
4. It might appear that it would be an enormous task to populate this probability model. However, it should be noted that provided the relevant joint probabilities of failure are in the right ballpark – and the extension of the conversation formulae induced by the BN will help ensure they are – then extrapolating across a few cases should ensure a well calibrated model. The stochastic digital twin therefore presents a real promise as an evocative tool to help guide strategic planning. Note that the decompositions associated with the BN enable many what-if analyses of the efficacy of various protective measures designed to mitigate increasing flood risk – like re-siting of vulnerable infrastructure to higher ground.
5. We believe that the methods used here are very important to building meaningful digital twin models, as illustrated by the detail of failure modes formalised based on expert knowledge in the elicitation process. One challenge associated with wide deployment of these methods is that the relevant skills are not widespread even in the statistics and decision analysis research communities – to take advantage of the methods across the industry it would be necessary to provide protocols and software tools to help make them accessible to a wider

range of analysts, and possibly develop the skills base in statistical modelling.

6. There is nothing about the technology defined here that could not be adapted to provide a real time digital twin that could inform a crisis control centre with critical information concerning the unfolding threat posed by an actual extreme incident happened - helping to identify those combinations of assets across different asset owners whose failure might have the most catastrophic effects on the consequences of an incident.

We hope we have demonstrated through this study how probabilistic digital twins can be developed for better understanding the impacts of the increased flooding risks driven by climate changes and the ways these might be used within the strategic planning of asset owners to better ameliorate these risks. In future studies we plan to develop more sophisticated tools to provide exactly this type of support.

References

[1] M. Leonelli and J. Q. Smith, "Bayesian decision support for complex systems with many distributed experts," *Annals of Operations Research*, vol. 235, no. 1, pp. 517–542, 2015. DOI: 10.1007/s10479-015-1957-7.

[2] M. J. Barons, S. K. Wright and J. Q. Smith, "Eliciting probabilistic judgements for integrating decision support systems," in *Elicitation*, Springer, 2017, pp. 445–478. DOI: 10.1007/978-3-319-65052-4_17.

[3] M. Leonelli, E. Riccomagno and J. Q. Smith, "Coherent combination of probabilistic outputs for group decision making: An algebraic approach," *OR Spectrum*, vol. 42, no. 2, pp. 499–528, 2020. DOI: 10.1007/s00291-020-00588-8.

[4] M. Leonelli and J. Q. Smith, "Dynamic uncertainty handling for coherent decision making in nuclear emergency response," in *Proceedings of the winter meeting of the ANS*, 2013.

[5] M. J. Barons, T. C. O. Fonseca, A. Davis and J. Q. Smith, "A decision support system for addressing food security in the united kingdom," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, in press, 2021. DOI: 10.1111/rssa.12771.

[6] A. Shenvi, F. O. Bunnin and J. Q. Smith, *A bayesian decision support system for counteracting activities of terrorist groups*, 2020. arXiv: 2007.04410 [cs.SI].

[7] V. Volodina *et al.*, *Propagating uncertainty in a network of energy models*, 2022. arXiv: 2201.09624 [stat.AP].

[8] K. B. Korb and A. E. Nicholson, *Bayesian Artificial Intelligence*. CRC Press, 2010. DOI: 10.1201/b10391.

[9] J. Q. Smith, *Bayesian Decision Analysis*. Cambridge University Press, 2009. DOI: 10.1017/cbo9780511779237.

[10] J. Pearl, *Causality*. Cambridge University Press, 2000.

[11] R. L. Wilkerson and J. Q. Smith, "Customized Structural Elicitation," in *Expert Judgement in Risk and Decision Analysis*, ser. International Series in Operations Research & Management Science, A. M. Hanea, G. F. Nane, T. Bedford and S. French, Eds., Springer, 2021, pp. 83–113. DOI: 10.1007/978-3-030-46474-5.

[12] J. Smith, R. Procter *et al.*, *Bayesian elicitation behind a firewall: Criminal model building and co-creation*, Turing Research Report (classified), 2021.

[13] X. Yu and J. Q. Smith, "Causal algebras on chain event graphs with informed missingness for system failure," *Entropy*, vol. 23, no. 10, p. 1308, 2021. DOI: 10.3390/e23101308.

[14] R. A. Collazo, C. Gørgen and J. Q. Smith, *Chain Event Graphs: Chapman & Hall/CRC Computer Science and Data Analysis Series*. CRC Press, 2018.

[15] A. O'Hagan *et al.*, "Uncertain judgements: Eliciting experts' probabilities," 2006. DOI: 10.1002/0470033312.

[16] A. Ankan and A. Panda, "Pgmpy: Probabilistic graphical models using python," in *Proceedings of the 14th python in science conference (scipy 2015)*, Citeseer, 2015, pp. 6–11.

[17] *Climate Resiliance Demonstrator collaboration, Bayesian Network Repository*, digitaltwinhub.co.uk/credo/resources/technical, 2021.

A Appendix: From a BN to a factorisation formula and a transfer function

The beauty of the BN once they have been drawn is that enables the calculation of explicit probability distributions as a function of the inputs of numerical descriptions of the contributing parts of the network and its embellishments for each emulated incident. We give below the formulae that enable us to calculate the joint failure of a particular asset given the critical features that define the dangers presented to the asset form a particular unfolding incident.

So suppose that $\mathcal{G}(a)$ is a valid BN of an asset whose m nodes/ random vectors are

$$\left\{Y_{1,T}(a), Y_{2,T}(a), \ldots Y_{m-1,T}(a), F_T(a) \triangleq Y_{m,T}(a)\right\}$$

where $F_T(a)$ denotes the indicator on the failure of an asset of type $a \in A$ at time T. Let the vector of parents of $Y_{k,T}(a)$ in G be denoted by $\boldsymbol{Z}_{k,T}(a), k = 1, 2, \ldots, m$. Each asset a^s located at a site $s \in \mathcal{S}$ will have associated with it a certain set of covariates $\boldsymbol{x}^s$ describing features like its location, height , protection and its connections to other assets in the network - by type.

We now consider a future flooding incident j that has been generated as a possible future event in a potential climate changed environment described through one of the climate scenarios. The scenario faced by this incident will be labelled by two time series. The first $\left\{\boldsymbol{x}_t^i(j) : t = t_0, t_1, \ldots\right\}$ provides the time series provided by the weather/flood modellers from the start of the incident t_0. These will be a part of the routine delivery to the system. These might include such features as the rainfall intensity, whether or not flood water has reached the given site s and its depth at a given time t.

The second, feature, vector time series $\{\boldsymbol{x}_t^e(j) : t = t_0, t_1, \ldots\}$ involves other descriptors of the incident which are exogenous to the particular incident but from the BN can be seen to have implications about whether or not a might fail. These include features like the time of day and season of the incident, whether there are road works locally or maintenance regimes happening at connected sites.

For each incident j Now let $\boldsymbol{x}_t(j) \triangleq \left(\boldsymbol{x}_t^s(j), \boldsymbol{x}_t^i(j), \boldsymbol{x}_t^e(j)\right)$ and let $\boldsymbol{x}^{T^*}(j)$ up to a time T^* denote the vector of inputs that will inform the development of the failure event at time T of a given asset on a site, where T is the time a snapshot of the system is taken. Not that for a real time decision support system we must set $T^* \leq T$ but for planning we can follow a simulated scenario to its endpoint, provided the functionality of the assets do not affect the spread of the flood or the surrounding weather.

Then - directly from the validity of the BN $\mathcal{G}$ - we have that

$$P(F_T(a) = 1|\boldsymbol{x}^{T^*}(j)) = \sum_{y_{1,T}(a),y_{2,T}(a),\ldots y_{m-1,T}(a)} \left\{ \prod_{k=1}^{m} p(y_{k,T}(a)|\boldsymbol{z}_{k,T}(a), \boldsymbol{x}^{T^*}(j)) \right\} \quad (3)$$

where $p(y_{k,T}(a)|\boldsymbol{z}_{k,T}(a), \boldsymbol{x}^{T^*}(j)), k = 1, 2, \ldots, m$ denotes the conditional probability tables elicited from the asset owners. Note here that in practice although $p(y_{k,T}(a)|\boldsymbol{z}_{k,T}(a), \boldsymbol{x}^{T^*}(j))$ could depend very generally on $\boldsymbol{x}^{T^*}(j)$ in practice it will only depend on a few simple functions of this vector and in some cases be independent of it entirely.

So we see here that the probabilities that the elicitation module needs to deliver are simply those in (3), and that these are a simple function of the conditional probability tables

$$\left\{ p(y_{k,T}(a)|\boldsymbol{z}_{k,T}(a), \boldsymbol{x}^{T^*}(j)), k = 1, 2, \ldots, m \right\}$$

constructed by the Bayesian elicitation team as a function of delivered modeling inputs and elicted expert judgements concerning asset failure. So (3) provides us with an explicit transfer function to feed into the model to determine whether or not a particular asset is working at time T.

B Appendix: Briefing document for the probability elicitation

This Appendix has been edited from the original text of the briefing document to remove confidential information.

Introduction

Thank you for agreeing to take part in the probability elicitation. You will be asked to make judgements regarding the likelihood of certain events which could contribute to the failure of the asset in a hypothetical future flooding scenario.

The elicited values will be used as part of the proof of concept for the digital twin developed in the CReDo project. In particular, the values will help us to calibrate the failure probabilities in the scenarios that are being used to demonstrate the digital twin. When reporting the results of this exercise all judgements will be anonymised appropriately (which may be different for internal and external use).

You will be asked to make judgements regarding a number of quantities of interest, provided below.

The scenario we consider

A 48 hour warning of a convective storm surge is given during the summer. This is predicted to last for up to 23 hours with higher peak intensity around the centre of the storm. This will occur when there is an above average high astronomical tide and sea level rise due to climate change. It is predicted that this occurs concurrently with associated wave overtopping of coastal flood defences and high tide in the river. There is a risk that the level of the river will be sufficiently high to threaten overtopping, although it is unlikely to be overtopped in the 23 hour period.

We are interested in the threat to the asset. EA flood information is made available only after the incident has happened.

Quantities of interest

In the elicitation you will be asked about probabilities of certain events of interest, including conditional probabilities such as that of a failure given that water has overtopped the defensive barrier. Where relevant we will consider different flood depths in the vicinity of the pumping station. These are:

(a) between 0mm and 500mm,
(b) between 500mm and 900mm,

(c) above 900mm.

Elicitation questions

For each of the quantities of interest above, you will be asked for your assessment of the probability. This will be your subjective judgement of how likely the event is to happen in the scenario. To help you think about this, you could compare the event to events with well-known probabilities, e.g., do I think this event is more or less likely than getting a head when flipping a coin (probability 1/2) or rolling a six with a dice (probability 1/6).

This means that, given your unique knowledge and experience, your probability is likely to be different to those of your colleagues. This is natural, and does not mean that anyone is wrong. These are one-off events and our knowledge of what would happen during them is imperfect.

Once you have provided your probability we will show you the range of probabilities from all of the experts. This will be used to have a short discussion about the values chosen and the considerations that led to these. Following this we will ask for a probability for the event from the group. This is known as a consensus probability. To come to this probability we ask you to imagine that an impartial observer has been present for the elicitation and discussion, but did not come into the session with their own views. What probability would they assign to the event?

Format of the Elicitation Session

The elicitation session will take place via Teams. It will be facilitated by Kevin Wilson, assisted by Jim Smith, and notes will be taken by Jim Smith and Chris Dent. Sarah Hayes and Benjamin Mawdsley will observe the session, but not participate.

In the session, Kevin will review the information contained in this document and then elicit your probabilities, first for a training question to illustrate the approach, and then for the quantities of interest above.

We are also interested in the rationale for the numbers you provide, and hence the session will be recorded, with your permission.

There will be time to ask any questions you have during the session.

The session will take no longer than two hours, with a break in the middle.

Acknowledgements

Lead Author

Chris Dent
Ben Mawdsley
Jim Q. Smith
Kevin J. Wilson

The Centre for Digital Built Britain at the University of Cambridge's National Digital Twin programme is funded by the Department for Business, Energy and Industrial Strategy via UK Research and Innovation.

Dent, C; Mawdsley, B; Smith, J.Q.; Wilson, K,J (2022). CReDo Technical Report 3: Assessing asset failure. https://doi.org/10.17863/CAM.81780