CAUSAL GRAPHICAL MODELS FOR SYSTEMS-LEVEL ENGINEERING ASSESSMENT

Victoria Stephenson[1,*], Chris. J. Oates[2], Andrew Finlayson[3], Chris Thomas[4], and Kevin J. Wilson[5]

[1]Department of Architecture and Civil Engineering, University of Bath, UK. Email: v.j.stephenson@bath.ac.uk

[2]School of Mathematics, Statistics and Physics, Newcastle University, UK. Email: chris.oates@ncl.ac.uk

[3]British Geological Survey, UK. Email: afin@bgs.ac.uk

[4]British Geological Survey, UK. Email: cwt@bgs.ac.uk

[5]School of Mathematics, Statistics and Physics, Newcastle University, UK. Email: kevin.wilson@ncl.ac.uk

[*]Corresponding Author

ABSTRACT

Systems-level analysis of an engineered structure demands robust scientific and statistical protocols to assess model-driven conclusions that are often non-traditional and causal in their content. The formal mathematical, statistical, and philosophical foundations of causal inference on which such protocols are based are, nevertheless, not widely understood. The aim of this paper is to communicate the essentials of graph-based causal inference to the civil engineering community, to demonstrate how rigorous causal conclusions – and formal quantification of uncertainty regarding those conclusions – may be obtained in a typical engineered system application and to discuss the value of this approach in the context of engineered system assessment. The concepts are illustrated via a river-weir ecosystem case

study, as an example of decision-making for engineered systems in the built environment. In this setting, we demonstrate how rigorous predictions can be made about the outcome of decisions, that take a lack of prior knowledge about the system into account. The findings highlight to end-users the value in applying this approach, in providing quantitative, probabilistic outputs that counter decision uncertainty at system level.

Keywords: causal inference; directed acyclic graph; river-weir ecosystem; systems engineering

INTRODUCTION

The vast majority of scientific hypotheses are not statistical, but are *causal*. One example of such a causal construct that surrounds a system in the built environment is the changing response of a structure to external loading conditions over time, as a consequence of the natural evolution of internal characteristics such as strength and stiffness, or the intervention on these properties. Protocols to test such hypotheses are well-understood and codified in the modern scientific method, typically a combination of *in silico* model simulations and *in situ* experiments targeted at replicating the causal mechanisms at work. An experimental approach will often seek to produce high quality data to describe only a single causal relationship, through controlling surrounding physical conditions.

A systems-level approach, on the other hand, aims to describe real-world systems by simultaneously assessing *en masse* a collection of causal statements, through employing a protocol of codifying numerous causal hypotheses in the form of a single mathematical or computational model. The model can be produced without experimental data pertinent to every causal statement, able to be constructed from combinations of empirical formulae and first principles, and supplied with elicited quantitative information. The model is then used to produce predictions about the real-world system, either under specified constraints or as the outcome of interventions. The model's performance can further be assessed against a real-world dataset, with strong predictive capability interpreted as evidence in support of the collection of causal statements taken together, which is then used to guide future hypothesis refinement.

The technique aims to establish multiple causal conclusions, using a holistic mixture of mathematical models, statistical techniques and diverse datasets. In doing so it offers the user a route to rigorous prediction about the real-world system, produced via accessible analytical methods and able to function with imperfect and incomplete data. This is useful in the case of engineered systems in the built environment, where data may be limited for structures situated in real environments if such structures are not endemic in the area.

Of central importance in the effort to use systems-level approaches are mathematical and statistical theories of causal inference. These enable the engineer to establish which causal statements are testable from observational data, to adjust for external factors that might confound parameter estimates and model-based predictions, to reason about the transfer of causal conclusions across the engineered system and its physical surroundings, and to provide an honest quantification of the epistemic uncertainty that accompanies all causal conclusions. Our experience is that, while the correlation-causation distinction is appreciated (e.g. Bell *et al.*, 1992; Salvaneschi *et al.*, 1997; Suraji *et al.*, 2001; Cotter, 2015), the useful and powerful logico-deductive theories of graph-based causal inference are not yet well-understood in this trans-disciplinary research field. The aim of this article is to communicate a clear, explicit and practicable introduction to causal inference via a real-world case study from the field of the built environment. The intention is that this presentation will help to accelerate the adoption of formal causal reasoning in the field.

The real-world motivation for this research was to study the Clerkington Weir, an historic river barrier on the river Tyne in south-east Scotland, under the jurisdiction of the Scottish Environmental Protection Agency (SEPA). The weir, which dates from the early 19th century, has been identified as an inhibitor of fish migration and there is an ongoing conversation with a wide variety of stakeholders regarding possible weir removal or modification, such as via the addition of a fish passage structure. Two aspects frustrate this decision landscape; firstly that removal is typically technically complex and costly, and may also be hindered by other factors such as historic weirs being protected (listed) structures. In this case uncertainty regarding the long-term prosperity of the current physical weir-river system is a factor. Secondly, that if weir removal is carried out the impact on the performance of the remaining elements of the system is challenging to predict due to its complexity, and hence there is epistemic uncertainty about the consequences of removal. These could include, for example, changes to river ecology health and river re-routing in the case of removal, or increased flood risk under increased precipitation in the case of non-removal.

To date no formal quantitative probabilistic attempt has been made to predict the consequences at system level of removal or non-removal. The Clerkington Weir is therefore an ideal case study on which to demonstrate the applicability of the concepts of causal inference to a real-world context, as well as highlighting aspects in which these techniques are limited. This presents an opportunity to assess the value of applying causal inference methods to this real world engineered system, where a challenging decision context is being played out, and where addressing uncertainty about system response to intervention is key to moving forwards. Across the system a large number of causal mechanisms are at play. For example, to assess the impact of extreme rainfall events on the structural integrity of a weir it is necessary to posit causal hypotheses for how rainfall affects flow in the river, for how the weir responds to different flow conditions and for how flow induced erosion and scour might act to undermine the integrity of the weir.

The assessment presented here seeks to address three features of the decision landscape; the impetus for the decision (that there is a barrier to natural fish migration), a cause of uncertainty relevant to non-removal (weir condition and design), and an uncertain consequence of removal (alteration of flood risk). Thus we deploy causal techniques to estimate fish passability on the weir, to estimate the unknown weir density and embedment depth, both pertinent to the stability of the weir, and to assess the change in risk of upstream flooding as a result of weir removal. The results deliver distributional and risk based predictions, derived from an explicitly causal model. Using a subset of observed datasets a new set of numeric outputs is presented that describes both currently unknown features and future performance measures of the river-weir system pertinent to the ongoing decision making effort.

The main body of the paper is given to first presenting the elicitation of the causal model used with the case study, then the formal reasoning associated with the questions used to make predictions about the system, and the predictive results derived from them. This is followed by a discussion of the practical and technical challenges faced, the novelty of the approach, and a comparison of the work with other possible methods for obtaining probabilistic predictions of system performance. The conclusions focus on the value-added offered by the causal inference approach over other available methods, especially in the context of generating impact in society, and to highlight the potential gains that more widespread use of these methods would provide. Two appendices are provided with the paper; the first contains an overview of the underpinning frameworks of causal graphical models, the second presents the full extent of the causal model construction.

A CAUSAL GRAPHICAL MODEL OF AN HISTORIC RIVER-WEIR SYSTEM

There are several competing mathematical and philosophical frameworks that attempt to formalise the process of causal deduction, including counterfactuals (Morgan and Winship, 2014), structural equation models (Kline, 2015) and the decision-theoretic approach of Dawid, 2000. This work applies one such framework, due to Pearl, 1995, that is based on a directed acyclic graph (DAG) representation of causal inter-dependencies in the system of interest. The DAG framework has received considerable theoretical attention and is perhaps the approach to causal inference that is most widely-used (Pearl, 2009). Even within the context of DAGs, the term 'cause' has historically received diverse usage. In this paper we adopt a domain-specific (and expert-elicited) notion of causation.

Full details of the mathematical and statistical foundations of the causal inference that leads to the causal DAG presented below can be found in Appendix I. The aim of this section is to illustrate how the mathematical and statistical content of Appendix I can be applied to perform rigorous causal inference in a civil engineering context specifically in relation to an historic river-weir.

The Case Study: Clerkington Weir

Clerkington Weir is a barrier on the river Tyne in south-east Scotland. It is located approximately 1.5 km to the south-west of Haddington and is one of a total of 12 weirs on the river (SEPA, 2018). The River Tyne has a total drainage area of 318.27 km^2; it is sourced in the Moorfoot and Lammermuir Hills and flows in a general north-eastward direction to enter the outer Firth of Forth at Tynemouth. The stream network for the Tyne catchment is shown in Figure 1.

The impact of the weir on fish passage has been highlighted by stakeholders and the possibility of weir modification or removal has been discussed. Conversely, the age of the weir and its perceived cultural and historic significance in the local landscape means it is considered an important feature and, as for other barriers on the Tyne, the added protection of having listed status renders removal a challenging and emotive issue. One impediment to resolution is the absence of quantitative measures of the physical, hydrological and ecological impact, positive or negative, weir modification or removal might result in.

The river-weir ecosystem is complex, containing a large number of components across different domains and multiple inter-dependencies. This hinders the generation of reliable outputs to produce these measures of impact by standard, non-causal statistical methods. For example, it is likely that the Clerkington weir differs in several important respects to other weirs on which data may have been collected and it is therefore unclear how conclusions of a statistical nature, drawn from structures with possibly quite different characteristics, can be meaningfully extracted.

In seeking to determine the best decision for the Clerkington weir, causal links that represent the whole system, must be considered simultaneously in order to prevent inaccurate reasoning about the system. For example, although statistical analysis of fish passage over barriers indicates that the height of a barrier is negatively associated with fish passability (e.g. King and O'Hanley, 2016), it would not be appropriate to reason that removal of Clerkington weir would therefore lead to increased fish stock in the Tyne. This is because such reasoning considers only one causal mechanism in the system, where other causal mechanisms may also exist. It may be the case that weir removal changes the flow in the river in a way that leads to bank erosion and vegetation loss, to the overall detriment of the fish stock. Alternatively, if the face of the weir is supporting denitrifying microbes then removal of the weir may result in increased levels of nitrogen in the river, leading indirectly to a reduction in fish stock.

The Clerkington Weir will be used as a case study, allowing us to demonstrate how causal graphical models can be applied in the context of managing the built environment and its relationship with the surrounding landscape. In particular, we considered three questions in detail:

Q1: To what extent can fish pass over the weir?

Q2: What can be said about the un-observable aspects of masonry structure of the weir?

Q3: To what extent does weir removal reduce the upstream flood risk?

In order to provide clarity in the presentation of our argument regarding the value of applying the causal graphical model to this real world context, we have deliberately limited our attention to a subset of key random variables (RVs) and datasets. This allows for focused discussion of the complex real-world interactions across these variables, that underpins the case for a causal inference approach, and their formal representation in a causal graphical model. It does also determine that the results presented in this paper are illustrative only, and that further development of the model would ideally need to be undertaken if it were to be used as the basis of a real-world decision making tool.

In the remainder of this section, first the main RVs relevant to the river-weir ecosystem are elicited and described. These are then assembled into a causal DAG and the conditional distributions associated with the DAG are described. This is followed by a demonstration of how the causal DAG allows for explicitly causal hypotheses on the river-weir ecosystem to be reasoned about and investigated.

Elicitation of the Causal DAG

The first task in constructing a causal graphical model is to elicit the RVs that will form the vertices in the causal graphical model. These will be generically denoted X_v, for v ranging over an index set V, and include RV's that are physically relevant across the system in relation to the questions being asked, and any associated datasets on which inferences are to be based. In this case study, elicitation was conducted based on discussion with both the stakeholders and various domain experts, including geologists, ecologists and engineers. Note that the set of RVs $\mathbf{X}_V := \{X_v\}_{v \in V}$ presented is a subset of all elicited RVs, to encourage clarity of communication of the analysis and results. These RVs are partitioned into those related to the geometry of the weir, the condition of the weir, the environmental RVs and the available datasets.

Geometry of the Weir

The first RVs elicited were intended to characterise the geometry of the Clerkington Weir, derived from visual observation of the structure on site (Figure 2) and historic documentation regarding the typical design of weirs of a similar age to Clerkington Weir, highlighting features such as the stacking of masonry units on the weir face (Figure 3a) and the use of piled foundations (Figure 3b). To this end, the profile of the weir was assumed to be a non-symmetric trapezoid, characterised by a *length* X_{L} (m), a *weir height* X_{WH} (m), a *slope up* X_{SU} and a *slope down* X_{SD}. All geometric RVs are shown in Figure 4.

No inspection of the below-ground structure was undertaken, hence engineering judgement was relied upon to elicit the foundation design in use at the weir. Documentary evidence suggests that piled foundation solutions were employed for the purposes of ensuring stability in weir structures at the time at which the case study weir was originally constructed, however it was not possible to confirm this directly for the Clerkington weir. Considering the complexity associated with articulating pile behaviour within the causal framework, a simplified approach was taken wherein the foundations were modelled as a rectangular section with an unknown *embedment depth* X_{ED} (m). This reflects the fact that massing below river bed level will almost certainly be present in the weir structure, directly contributing to its stability, without seeking to represent additional frictional aspects of pile performance, which go beyond the scope of the geomorphological and geotechnical information available to the case study. This approach seeks to ensure a worst case stability situation is provided for this first stage assessment.

The masonry construction of a weir of this age would not typically have included the presence of mortar, with inclined bedding of rough, interlocking blocks used to provide shearing resistance across the weir mass. Over time as the weir was continually exposed to the dynamic effects of hydraulic loading and other environmental effects (e.g. bank expansion and contraction), it can be safely assumed likely that the masonry units would have moved relative to each other. As such the presence of voids and other imperfections such as vegetation in the weir structure that lead to a reduction in overall density, from the value that was initially ensured via the masonry laying technique, is anticipated. This is supported by the visible and not insignificant presence of vegetation on the weir face (Figure 2), although the exact location and extent of voiding was not measured. This uncertainty was modelled as an RV, *weir density* X_{WDI} (Nm^{-3}), homogeneous across the weir body.

Condition Variables

Engineering expertise was used to elicit RVs contributory to potential failure modes of the weir, utilising available assessment tools (Pickles *et al.*, 2014; Kennard *et al.*, 1996). Four failure modes were identified; failure due to overturning (EQU1), failure due to sliding (EQU2), failure due to uplift (UPL) and failure due to piping (PIP). These failure modes were each represented by the binary RVs $X^{(i)}_{\mathrm{EQU1}}, X^{(i)}_{\mathrm{EQU2}}, X^{(i)}_{\mathrm{UPL}}, X^{(i)}_{\mathrm{PIP}}$, with 0 representing non-failure and 1 failure occurence, and an index i used to represent the date on which failure is being considered. (Here i runs over an index set that will be denoted $\mathcal{I}$.) Not all failure modes are modelled and in particular internal failure of the weir structure, for example due to fracturing, was not considered. This limits the causal model to that of considering external stability of the weir as a rigid body, whilst enabling the interaction with the river water forces to be fully resolved. A further *condition assessment* binary RV $X_{\mathrm{CA}}^{(i)}$ was taken to equal 1 if, on day i, any of the four failure modes occurred.

Environmental Variables

Several environmental RVs are required to properly characterise the river conditions at the weir. Hydrological considerations motivated the the inclusion of: *bank height* X_{BH} (m), *channel width* X_{CW} (m), *flow* $X_{\mathrm{F}}^{(i)}$ ($\mathrm{m}^3\mathrm{s}^{-1}$) on day i, *upstream water depth* $X_{\mathrm{UWD}}^{(i)}$ (m) on day i and *downstream water depth* $X_{\mathrm{DWD}}^{(i)}$ (m) on day i. A further binary RV $X_{\mathrm{UF}}^{(i)}$ was used to indicate whether an upstream flood had occurred on day i, with 1 representing a flood event. Full designation of the conditions used to classify a flood event are described in Appendix II. Additionally, to represent the failure mode EQU2 it was necessary to include RVs X_{C} and X_{SFA} respectively representing the soil *cohesion* (Nm^{-2}) and the *soil friction angle* (deg).

Ecological considerations led to the inclusion of RVs representing *fish passability* $X_{\mathrm{FP}}^{(i)}$ on day i. Here the passability of the weir for brown trout, one of the species of fish known to populate the Tyne, is considered, such that $X_{\mathrm{FP}}^{(i)}$ takes one of the four categorical values $\{\mathrm{total}, \mathrm{high}, \mathrm{medium}, \mathrm{low}\}$ defined in Baudoin *et al.*, 2014 as indicative of the degree of passability of the weir, according to the weir geometry, flow conditions and fish characteristics (e.g. jumping capacity).

Observed Variables

A limited number of datasets were collated to provide statistical information related to the physical RVs just described. The geometric RVs $X_{\mathrm{L}} = 6.3$ (m), $X_{\mathrm{SU}} = 0.4$, $X_{\mathrm{SD}} = 0.4$, $X_{\mathrm{CW}} = 50$ (m) could be directly observed. The weir height $X_{\mathrm{WH}} = 1.2$ (m) was measured using differential GPS data, shown in Figure 5, obtained on 28th September 2018. In addition, the *bank height* was denoted X_{BH} and was observed as 1.5 (m).

Measurements of flow $X_{\mathrm{F}}^{(i)}$ were obtained from the National River Flow Archive (NRFA, 2019). These consisted of mean daily flow measurements taken from 1981-2000 at three upstream locations, one upstream at Spilmersford on the Tyne and two at intermediate tributaries (Lennoxlove on the Coulston and Saltoun Hall on the Birns) that contribute to the total flow arriving at the weir. The values $X_{\mathrm{F}}^{(i)}$ were calculated as the sum of these three contributors to the total flow at the weir with the index set $\mathcal{I}$ containing approximately 7,300 days in total. The date range used derives from the fully overlapping portion of the three time series that constitute our dataset, in order that additional technical development to handle missing data was not required. Finally, a condition appraisal of the weir indicated that no failure mode has occurred, so that $X_{\mathrm{CA}}^{(i)} = 0$ for all days i in the dataset. In the following we denote by $\mathbf{X}_O$ where $O = \{\mathrm{L}, \mathrm{WH}, \mathrm{SU}, \mathrm{SD}, \mathrm{CW}, \mathrm{BH}, \mathrm{SSD}, \mathrm{F}^{(i)}, \mathrm{CA}^{(i)}\}$, the subset of RVs which together constitute observed nodes in the DAG.

This completes specification of the RV index set V. It remains to specify any causal relationships among the RVs, in a real-world qualitative sense at the level of the DAG and in quantitative terms at the level of conditional and interventional probability distributions. Full details of these relationships, as they derive from physical and empirical functions, are presented in Appendix II. The full DAG model is displayed in Figure 6.

Scientific Reasoning Using the Causal DAG

To illustrate how the causal graphical model enables rigorous and automatic reasoning about scientific hypotheses, the three scientific questions Q1, Q2 and Q3 are considered. Of these, Q1 and Q2 concern the distributional nature of the RVs involved and are not causal in nature; the purpose of these is to demonstrate the type of mathematical calculation involved when using the causal DAG to determine the conditional distribution of a given RV. The third question, Q3, is explicitly causal and relies on the Pearlean interventional structure that we have endowed on the causal DAG to measure the effect of an intervention on the river-weir system.

Q1: Fish Passability

The Clerkington weir is recognised as being as a barrier to fish passage on the Tyne, but to date no quantitative analysis of the river-weir ecosystem has been performed that draws on observed data specific to the physical nature and situation of the weir in the river. As a first example of reasoning based on the articulated graphical model, we consider how the observed data described so far can provide quantitative information concerning the impedence to fish passage posed by the weir. This is formalised as the following question:

Question 1. What is the conditional distributions of fish passability $p(X_{\text{FP}}^{(i)} \mid \mathbf{X}_O)$ on each day i, given the observed datasets $\mathbf{X}_O$?

In what follows we explain how the DAG in Figure 6 enables this question to be precisely answered. First we apply the law of total probability to express the desired conditional distribution as the integral

$$p(X_{\text{FP}}^{(i)} \mid \mathbf{X}_O) \;=\; \int p(X_{\text{FP}}^{(i)}, \mathbf{X}_{V\setminus(O\cup\{\text{FP}^{(i)}\})} \mid \mathbf{X}_O) \, \mathrm{d}\mathbf{X}_{V\setminus(O\cup\{\text{FP}^{(i)}\})}.$$

Then we leverage the definition of the conditional density as

$$\begin{aligned} p(X_{\text{FP}}^{(i)} \mid \mathbf{X}_O) \;&=\; \int \frac{p(X_{\text{FP}}^{(i)}, \mathbf{X}_{V\setminus(O\cup\{\text{FP}^{(i)}\})}, \mathbf{X}_O)}{p(\mathbf{X}_O)} \, \mathrm{d}\mathbf{X}_{V\setminus(O\cup\{\text{FP}^{(i)}\})} \\ &=\; \frac{1}{p(\mathbf{X}_O)} \int p(\mathbf{X}_V) \, \mathrm{d}\mathbf{X}_{V\setminus(O\cup\{\text{FP}^{(i)}\})} \qquad (1) \end{aligned}$$

where in (1) we recognise that the RVs $\mathbf{X}_O$ are not being integrated. At this point we can exploit the conditional independence structure of the DAG using the Markov property in (2) of Appendix I to obtain

$$p(X_{\text{FP}}^{(i)} \mid \mathbf{X}_O) \;=\; \frac{1}{p(\mathbf{X}_O)} \int \prod_{v\in V} p(X_v | \mathbf{X}_{\pi(v)}) \, \mathrm{d}\mathbf{X}_{V\setminus(O\cup\{\text{FP}^{(i)}\})}$$

Each of the terms appearing in the product has been elicited. The term $p(\mathbf{X}_O)$ does not depend on $X_{\text{FP}}^{(i)}$ and can be considered to play the role of a normalisation constant. Numerical techniques, such as implemented in the software discussed in Appendix I, can be used to numerically evaluate these conditional distributions. For the purposes of this paper we implemented a standard Markov chain Monte Carlo method.

Results are displayed in Figure 7. The left panel displays a superposition of the conditional probability distributions $p(X_{\text{DWD}}^{(i)}|\mathbf{X}_O)$ for each of the days i in the dataset. These indicate that, given the geometry of the weir and the observed variation in river flow conditions, the downstream water depth typically does not exceed 0.2 (m) and therefore that the air gap $X_{\text{UWD}}^{(i)} - X_{\text{DWD}}^{(i)}$ is typically at least $X_{\text{WH}} - 0.2 = 1$ (m). It follows that fish passability is rarely better than medium or low in the sense of Baudoin *et al.*, 2014. The right panel displays a superposition of the conditional probability distributions $p(X_{\text{FP}}^{(i)}|\mathbf{X}_O)$ which confirms the barrier effect of the weir on fish passage. The automatic computation of these multiple conditional distributions from the same DAG structure provides for efficient prediction across system variables, and from this greater awareness of the system's state.

It is important to emphasise that these results are driven by *all* of the observed datasets in $\mathbf{X}_O$ and not just a small portion of the available data, and that the correct integration of these multiple and diverse strands of evidence is performed automatically and efficiently through the DAG. This simultaneous conditioning against multiple observed datasets, allows the user to bring all the "knowns" to bear on the posited question and output rigorous and reliable new information from it.

Q2: Density and Embedment Depth

A major source of uncertainty regarding the performance of the river-weir system is the state of the weir itself. The original design of the weir and the extent to which its condition has deteriorated since construction dictates its stability and safety as a structure today, which influences decisions around possible interventions to the system. If the weir is in a state where even minor interventions would instigate weir instability or collapse, then the site works required to modify the weir to install a fish passage, for example, may not be practically possible. Or, if long-term system stability is desired with the weir in-situ, and works to ensure the longevity of the weir against increased flows are extensive, they may not be cost effective.

There is therefore interest in assessing and quantifying the state of the weir. However, a lack of observable features limits the capacity to achieve reliable assessment via survey, as values required to fulfil a majority of the variables needed to determine weir state through consideration of the physical interactions contributing to it are missing. Q2 considers inference across these unobserved variables relating to stability and condition, on the basis of the information that is available, $\mathbf{X}_O$. This is posed in particular as:

Question 2. What is the conditional distribution of the weir density and embedment depth $p(X_{\text{WDI}}, X_{\text{ED}}|\mathbf{X}_O)$ given the observed datasets $\mathbf{X}_O$?

The observed data includes the knowledge of the weir geometry and environmental conditions. It additionally includes the information that failure has not occurred, through the condition assessment RV $X_{\text{CA}}^{(i)}$. This knowledge of the capacity of the weir to withstand the loading conditions to which it has previously been exposed provides important insight into the state of the weir. Being able to condition on this knowledge via the DAG, in combination with the other observed information, allows for a prediction of the unobserved weir density and embedment depth that draws on this knowledge of historic system state. Formal computation of unobserved variables via historic systems level knowledge represents a new offering in the context of decision making around complex engineered systems, especially in relation to historic structures where so many variables are unknown. Proceeding in an analogous manner to Q1, we arrive at the formula

$$p(X_{\text{WDI}}, X_{\text{ED}}|\mathbf{X}_O) \;\propto\; \int \prod_{v \in V} p(X_v|\mathbf{X}_{\pi(v)}) \, \mathrm{d}\mathbf{X}_{V\setminus(O\cup\{\text{WDI,ED}\})},$$

with proportionality up to an implicit normalisation constant.

Results are displayed in Figure 8, indicating the probable upper and lower bounds of embedment depth X_{WD} and weir density X_{WDI} that are consistent with the fact that the weir has not failed under the system conditions contained in the observed nodes in the DAG. It is apparent that X_{WDI} (for which a uniform distribution was elicited) is relatively well-informed by the dataset, with a minimum density of around 10,000 (Nm^{-3}) being plausible under the model. Similarly the model provides a plausible minimum value for X_{ED} of around 0.5 (m). For very small values of either X_{WDI} or X_{ED} the model anticipates a larger value of the other to compensate and to ensure stability of the weir, as would intuitively be expected. The contour plot also provides the joint conditions attributable to the worst case that the weir might plausibly be considered to be in, in terms of its overall stability. Meanwhile the "soft" nature of the plot reflects uncertainty with respect to RVs such as the downstream water depth which play a causal role in failure of the weir.

Again, these results are driven by all of the observed data $\mathbf{X}_O$, with correct integration of these different strands of evidence being performed automatically through the DAG. Such a computation of the jointly probabilistic nature of variables from partial, high level knowledge of a complex real world engineered context is not traditionally available to decision makers, and the ease by which the DAG can compute these represents a significant opportunity to improve the quality of information available in these contexts.

Q3: Weir Removal

Neither Q1 nor Q2 require causal semantics, since they do not countenance an intervention on the system. Intervention is also at the root of the decision context being considered for the weir. A more realistic situation is now considered, where causal semantics are essential, specifically the effect of weir reduction or removal on upstream flood risk. This is an explicitly causal question that can be cast as an intervention on the weir height, X_{WH}, whereby it is set to some other fixed height $h \geq 0$. To make this precise, we now let $X_{\text{WH}}^{(i)}$ be indexed by day i and consider the effect of removal on a future day, denoted $*$, not in the earlier index set $\mathcal{I}$.

Question 3. If an intervention was performed that sets the weir height to h, what is the interventional distribution of an upstream flood $p(X_{\text{UF}}^{(*)} \mid \mathbf{X}_O, \text{do}(X_{\text{WH}}^{(*)} = h))$, given the observed datasets $\mathbf{X}_O$?

To address this question we extend the index set $\mathcal{I}$ to include $*$, leading to a larger causal DAG. Here an intervention is considered on a day $*$ not in the index set $\mathcal{I}$, which can be seen as a degenerate case of Balke and Pearl, 1994. This is a simpler method of approach than other alternately available ones. The intervention could have been posed as a *counterfactual* question where it is asked what *would* have happened on a day $i \in \mathcal{I}$ in the dataset *if* the weir height had been intervened on during that day; such questions are rigorously addressed in the *counterfactual network* approach of Balke and Pearl, 1994.

Now it is required to specify a marginal probability distribution for the newly introduced source node $X_{\mathrm{F}}^{(*)}$, which was taken to be a log-normal distribution fitted to the observed $X_{\mathrm{F}}^{(i)}$. Fits that are consistent with the flow dataset are displayed in the left panel of Figure 9. Then, from the Pearlean structure in (3) of Appendix I:

$$p(X_{\mathrm{UF}}^{(*)} \mid \mathbf{X}_O, \mathrm{do}(X_{\mathrm{WH}}^{(*)} = h)) \;\; \propto \;\; \int \prod_{v \in V} p(X_v | \mathbf{X}_{\pi(v)})\big|_{X_{\mathrm{WH}}^{(*)}=0} \; \mathrm{d}\mathbf{X}_{V \setminus (O \cup \{\mathrm{UF}^{(*)}\})}.$$

Results in the middle panel of Figure 9 indicate that complete removal of the weir ($h = 0$) reduces the per-day risk of an upstream flood event substantially, from 10^{-3} with the weir *in situ* to around 10^{-8} with the weir removed. Utilising the causal DAG to compute this reduction in risk provides the end-user with clarity and confidence regarding the scale of impact associated with undertaking a specific intervention within a larger system of interactions. This is a powerful tool with regards situations where there is a need to make decisions without prior knowledge of their effect. Methods that work to counteract vague and uncertain knowledge contexts explicitly address this real world problem. Additionally, the setting out and structuring of the causal DAG enables multiple causal roots to be explored in the context of interventions, and their impact updated as more data and knowledge is supplied.

For illustration the average causal effect (ACE; see Appendix I) of weir height RV $X_{\mathrm{WH}}^{(*)}$ on the upstream flood RV $X_{\mathrm{UF}}^{(*)}$ is also computed, shown in the right panel of Figure 9. This demonstrates the intuitively sensible fact that there is greater impact achieved on flood risk from reduction in height of a tall weir ($X_{\mathrm{WH}}^{(*)} > 1.3$ (m)) compared to reduction in height of a smaller weir ($X_{\mathrm{WH}}^{(*)} \leq 1.3$ (m)). On the other hand, the ACE is zero for values of $X_{\mathrm{WH}}^{(*)}$

greater than the bank height $X_{\mathrm{BH}} = 1.5$ (m), since a weir higher than the bank guarantees an upstream flood.

DISCUSSION

Causal graphical models constitute a rigorous framework in which deductive causal reasoning can be performed that simultaneously takes all of the identified causal mechanisms into account. The purpose of this study has been to demonstrate the value of applying a causal graphical model framework in an engineered-systems decision appraisal context. The outputs from the DAG-based causal analysis provide explicit insight into system performance, that might otherwise have remained as vague assertions. Without such an approach the answering of the three questions posed (Q1-Q3) would have been reliant on non-causal inference from statistical data (e.g. historic flood occurrence) and fragmented by the use of disparate, localised interaction models within the system (e.g. river flow over a barrier). This represents a valuable change in approach to overall engineered-systems assessment.

It is emphasised that the case study is illustrative only, and does not seek to provide validated proof of the specific case study's system state in the future. For example, the results reported account neither for changes that may have occurred in the flow profile of the Tyne since the flow dataset was obtained, nor for the possibility of more extreme future flow events due to climate change. Detailed justification and criticism of modelling choices would be essential if the conclusions drawn from the causal model are to be used as part of a decision-making tool in the future.

The following section discusses where future developments of the approach could be directed, and the impact of these. This includes refinement of the system assessment to increase the resolution of the causal relationships being articulated; expansion of the approach to situations where the causal structure is itself uncertain; and application of the work to cases where new system knowledge can be uncovered by experimentation.

Physical Model Assertions

Underpinning the validity of the DAG are the physical causal models with which it is constructed. Whilst full causal fidelity in the physical and engineering model structure has been sought for as much as is possible, for reasons of feasibility there remain some approximations and gaps. The list of failure modes used is not exhaustive, and the focus in this first stage assessment was to look at those modes where some degree of observation contributory to the causal structure could be undertaken, such as with the geometry of the weir. Additionally, with some of the failure models less resolute numerical techniques have been applied, such as in the specification of the model for piping failure. These stemmed from a desire to produce a model of the system that was more accessible to end-users than a fully resolute one might be, whilst also seeking to ensure confounding effects were avoided.

Further simplifications come from ignoring certain physical features of the natural system, especially those observed over time periods orders of magnitude greater than the immediate decision context. For example, the possibility of dynamic re-routing of the river, which is known to have historically occurred, was not considered. Changes in the course of the river Tyne through the site have been identified by comparison in a GIS system of: (i) historical Ordnance Survey maps (surveyed in 1855 and 1895); (ii) aerial photographs dating from 1946, 1988 and 2009; and (iii) a GPS survey of the river centreline undertaken in September 2018. An overview of these changes is presented in Figure 10; over the past 150 years the river has clearly migrated across the flood plain at several locations across the site. To properly account for uncertainty with respect to the future route of the river appears to be difficult, yet this has a direct bearing on the possible consequences of weir removal.

Estimation of Causal DAGs

This work presents the situation where all relevant causal mechanisms are elicited from experts (e.g. an engineer) and data is used only to quantify uncertainty with respect to parameters of the mechanisms involved. For engineered systems this situation can be justified, as the causal relationships are by definition designed into reality in the artifact. This provides a strong argument in favour of DAG-based causal deduction, compared to, say, epidemiology where the notion of a "direct cause" may need to be clarified. However, in some applications the edge structure of the causal DAG is itself an unknown object of interest. For example, in this case study this would be relevant to assertions about the system that relate to the down-scaling of very large scale causation into locally observed effects. Such as if climate change were to be explicitly considered, scaling from global temperature rise observations through catchment rainfall accumulation to flow specifically at the weir structure would be a consideration. That scale of model extent is beyond the scope of this assessment however.

Statistical methods have been developed to estimate causal DAGs from so-called "observational" data that arise. These methods require the so-called (causal) Markov and faithfulness conditions to hold (see Appendix I) and are often classified as either "constraint-based" or "score-based". Popular constraint-based methods include the PC algorithm of Spirtes *et al.*, 2000 and Bayesian hybrids of these methods (Claassen and Keskes, 2012), and popular score-based methods include (Meinshausen and Bühlmann, 2006; Bühlmann *et al.*, 2014; Bartlett and Cussens, 2013).

Application to Experimental Design

Once a (causal) DAG has been produced, it can be used to guide the design of future experiments to optimally reduce uncertainty with respect to some (causal) statement(s) of interest related to the (causal) DAG. For instance, if it was desired to reduce uncertainty surrounding the unknown embedment depth X_{ED} but there was no option to undertake a direct measurement then, from the DAG, it is apparent that one could instead seek to obtain information on the weir density W_{WDI} (for example by conducting an ultrasound experiment), which would in turn provide information on the conditionally dependent RV X_{ED}. The statistical literature on experimental design is large and we refer the reader to standard sources (e.g. Chaloner and Verdinelli, 1995) for further detail.

CONCLUSIONS

The presentation of this case study serves to highlight the potential benefits of the causal graphical model framework for systems-level engineering assessment. Without such an approach reliance on observed datasets for prediction and subsequent decision becomes the norm for this context. Whilst empirically robust, these approaches do not in general accommodate the deeper logico-deductive causal inference that is afforded in the causal DAG framework. Furthermore, the general lack of observed data that underpins much characterisation of engineered-systems in the built environment, hinders adoption of empirical approaches. As such, methods such as that presented here, offer a significant opportunity to overcome current epistemic uncertainty that surrounds decision making and intervention strategies in engineering situations, such as weir removal. These methods further represent an opportunity to capture and utilise the knowledge and information that does exist, currently confined largely to human expertise, which cannot assimilate and integrate so explicitly with purely data-derived predictive methods.

The deductive frameworks for causal inference that are presented in this article provide the mathematical, statistical and philosophical tools to address this challenge and to enable the honest quantification of the causal content of a model. New outputs produced by this work quantify the epistemic uncertainty accompanying causal conclusions drawn from the model. The case study of the Clerkington weir demonstrates the potential for these analytical techniques to deliver value in a real-world context, but nevertheless it is clear that further model criticism and refinement would be required for the work to form part of a decision-making tool. It is hoped that this article will help to stimulate further research effort toward adopting and tailoring formal causal models in these engineered-systems contexts.

APPENDIX I. APPENDIX I: CAUSAL GRAPHICAL MODELS: AN OVERVIEW

The aim of this section is to communicate the essentials of causal inference based on a DAG. Before we begin, we note that other excellent introductions to causal inference are available and include Spirtes, 2010; Pearl, 2010; Dawid, 2010. Our article differs in its presentation, being focused toward causal inference in civil engineering applications, but we were nevertheless heavily influenced by these earlier authors, who have each made fundamental contributions to the field.

Non-Mathematical Definitions

Causal inference blends both mathematical and real-world considerations in a unified framework. This means that the definition of certain non-mathematical terms will require context-specific semantics that must be specified. Examples will be provided below, while in the immediate development we follow Dawid, 2010 by indicating non-mathematical terms with `teletype` font.

Denote the collection of all relevant quantities in the engineered system of interest abstractly as $\mathbf{X}_V = \{X_v\}_{v \in V}$, with each quantity X_v being indexed by an element v in some suitable index set V. Our aim below is to build a graphical model that describes causal interdependencies among these quantities. To proceed, we must make precise the following non-mathematical terms:

- a `direct cause` among the $\mathbf{X}_V$
- a `common cause` of the $\mathbf{X}_V$

The semantics that are attached to these non-mathematical terms will be context-dependent. For example, when the X_i represent river level measurements, a `direct cause` between X_i and X_j may be understood to mean that location i is upstream of location j, so that increased river level at i implies more water must also be present at location j, since water flows from upstream to downstream. In this same example a `common cause` may be an external stimulus X^*, such as rainfall across the catchment area, that promotes increased river levels simultaneously at both locations i and j. In the case where X^* is latent (i.e. not included in the set $\mathbf{X}_V$), then variation in X^* can induce a spurious association between X_i and X_j that cannot be explained at the level of the quantities $\mathbf{X}_V$. Such latent `common causes` can be problematic as they require special treatment when performing causal deduction and, in order to simplify our presentation, these will be explicitly ruled out. That is, we will make the strong assumption that all relevant variables have been explicitly included in the set $\mathbf{X}_V$. Finally, it is convenient to call X_i an `indirect cause` of X_j if X_i is not a `direct cause` of X_j but there nevertheless exists a sequence of `direct causes` that connect X_i to X_j.

Graphical Calculus

Once the above non-mathematical terms have been defined for the relevant engineering context, one can formulate a causal graphical model. Recall that a DAG $G = (V, E)$ is comprised of a variable index set V and an edge set $E \subset V \times V$ with the property that there does not exist a directed path starting and ending at the same vertex (e.g. $1 \to 2 \to 3 \to 1$). Such a DAG G is said to be "causal" if (a) an edge $(i, j) \in E$ exists if and only if X_i is a `direct cause` of X_j, and (b) there are no latent `common causes` of the $\mathbf{X}_V$. A causal DAG is distinct from, for example, correlation networks or other types of probabilistic graphical model, though the latter have to some extent been exploited in engineering applications (Fienen *et al.*, 2013; Wu *et al.*, 2015a; Wu *et al.*, 2015b; Tong and Tien, 2017; Bhandari *et al.*, 2017). Rather, we restrict attention to formal causal models in order that rigorous causal conclusions can be derived.

For the moment we assume that the DAG G has been elicited from an expert and is treated as fixed. Practical approaches to elicitation are discussed below, in addition to a discussion of how the assumption of perfect expert elicitation can be relaxed.

In the framework of Pearl, 2009 each X_i holds the status of a random variable (RV), with randomness reflecting either epistemic uncertainty regarding these quantitites within a particular engineered system, or reflecting the fact that many similar engineered systems are being considered, of which the behaviour of a typical, randomly selected member of that population is being studied. The joint probability density function of the RVs is denoted $p(\mathbf{X}_V)$. In order to relate causal DAG models to the RVs we assume in this work the (causal) "Markov" property (Spohn, 1980; Spirtes *et al.*, 2000). This states that, for a (causal) DAG G, the following factorisation of the joint density holds:

$$p(\mathbf{X}_V) = \prod_{v \in V} p(X_v | \mathbf{X}_{\pi(v)}) \tag{2}$$

where $\pi(v)$ denotes the set of parents of vertex v according to the DAG G and $\mathbf{X}_S$ denotes the set of RVs $\{X_v : v \in S\}$. For example, under the Markov property the DAG in Figure 11 implies that the joint density $p(X_1, X_2, X_3)$ can be factorised as $p(X_1)p(X_2|X_1)p(X_3|X_2)$. It further follows from this factorisation that the RV X_1 is conditionally independent of the RV X_3 given X_2, written $X_1 \perp\!\!\!\perp X_3 | X_2$. (In general a "conditional independence relation" is a statement of the form

$$\mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B | \mathbf{X}_C$$

for some index sets $A, B, C \subset V$, meaning that the RVs $\mathbf{X}_A$ and $\mathbf{X}_B$ are *de facto* independent once the value of $\mathbf{X}_C$ is observed.) In order to simplify the presentation in what follows, the converse of the (causal) Markov property, called (causal) "faithfulness", is also assumed. This states that (2) is a maximal factorisation of the joint distribution, meaning that a conditional independence relation $X_i \perp\!\!\!\perp X_j | \mathbf{X}_S$, $i, j \notin S$ for some set $S \subset V$, implies that there does not exist an edge $X_i \to X_j$ in the DAG, and hence X_i cannot be a `direct cause` of X_j (Spirtes *et al.*, 2000).

Note that, although the name "random variable" is used, this framework also includes the possibility that a RV X_v is deterministically related to its parents $\mathbf{X}_{\pi(v)}$ in the DAG, perhaps explicitly through a mathematical formula or implicitly through a computer model. In this case the conditional density $p(X_v|\mathbf{X}_{\pi(v)})$ should be interpreted as probability mass function whose mass is confined to a single point.

The power of the graphical representation G is due to an extensively developed graphical calculus for causal DAGs. That is, there exist algorithmic manipulations of the graph which can be used to determine whether certain probabilistic and causal statements follow as a logical consequence of the elementary causal statements that are encoded in the individual edges of the graph. This can be illustrated with the motif in Figure 12, from which we may conclude that X_i is an `indirect cause` of X_k. Moreover, X_k cannot be an `indirect cause` of X_i, since this would imply that there exists a cycle in G, which is in contradiction to the definition of a DAG. In the case of general G, an important algorithm that we highlight is "d-separation" (Geiger *et al.*, 1990), which allows all implied conditional independence statements among the RVs $\mathbf{X}_V$ to be deduced from the graph G; this provides a convenient data-driven check on the statistical (i.e. non-causal) assumptions that are encoded in a DAG model. The criterion are implemented in software including Daggity (www.dagitty.net). These automatic methods for logical deduction, together with the ease of communication that is afforded by the graphical representation, have helped to contribute to the popularity of DAGs in a variety of research fields, most notably epidemiology (Rothman and Greenland, 2005).

Panel Notation

In applications of graphical models it is common for multiple RVs to appear in parallel in the DAG, as illustrated in the left part of Figure 13. In our case study, for example, each day i in the dataset is associated with a RV representing flow conditions in the river on day i. Such large numbers of RVs can make graphical representations unwieldy and it is therefore common to adopt so-called *panel notation*. An explicit example is given in the right part of Figure 13, wherein the dashed panel is used as a shorthand to indicate that copies of the graphical motif in the panel should be included for each of the indices $i \in \{2, 3, 4\}$.

The Reification Fallacy

At this point the opportunity is taken to emphasise the distinction between DAG models, in the general sense of a probabilistic graphical model, and *causal* DAG models in the specific sense that we have outlined. In particular, while every probability distribution can be factorised as in (2) for some DAG G, it is only causal DAGs for which an edge can be interpreted as a `direct cause` and therein associated with additional context-specific semantics. To assign a causal interpretation to edges in a non-causal DAG is known as the "reification fallacy" and is in general both scientifically and philosophically incorrect (see Section 4.3 of Dawid, 2010).

The reification fallacy is frequently overlooked, both in the over-interpretation of edges in a general (non-causal) graphical model, such as Gaussian graphical models and (non-causal) Bayesian networks, and in the assignment of meaning to higher-order graphical motifs. Further discussion on the mis-understanding of causal inference was provided in Imai *et al.*, 2008.

Expert Elicitation of the DAG

The expert elicitation of a causal DAG can be broken down into three main stages: the elicitation of the variables which form the nodes of the graph, the elicitation of the edges of the DAG, and the elicitation of the conditional probability distributions associated to the DAG, as appearing in (2).

In the engineering context, it is usually most efficient to encourage the expert to work backwards from the relevant failure mode(s) of the engineering system. The initial RVs considered will be called *Level 1* RVs. The expert now considers other features of the problem that might be a `direct cause` of at least one of these failure mode(s). These are *Level 2* RVs. The elicitation process continues to trace back these `direct causes` to their sources. The next layer, called *Level 3* RVs, will contain RVs that are a `direct cause` of a Level 2 RV (and therefore also an `indirect cause` of a failure mode). This process continues until the expert is content that all RVs pertinent to the failure mode(s) have been traced back. The resulting structure is sometimes called a *trace-back graph* (Smith, 2010). It is important at this stage to ensure that each of the RVs have a clear and unambiguous meaning, and could in theory be observed. The vertices of the DAG are taken to be the collection of all RVs just identified, denoted $\mathbf{X}_V$.

For each RV, X_v, the expert identifies a subset $\mathbf{X}_{\pi(v)}$ of the remaining RVs that are considered to be `direct causes` of X_v. The set $\pi(v)$ may be empty, in which case there are no `direct causes` of X_v. The set $\pi(v)$ is interpreted as the index set of the parents of the RV X_v in the DAG. For more information on the elicitation of edges in a DAG, see Chapter 7 of Smith, 2010 and Wilkerson and Smith, 2019.

The third stage, the elicitation of condition distributions for RVs, has been extensively studied in the literature (e.g. Garthwaite *et al*, 2005; O'Hagan *et al*, 2006). The aim is to translate the domain knowledge of an expert regarding a RV X_v (conditional on its parents $\mathbf{X}_{\pi(v)}$ in the DAG), into a probability distribution object. To do so, the expert is usually asked a series of questions about quantities that could, at least in theory, be observed. Questions should also be asked to minimise psychological biases exhibited by individuals when they express probabilistic judgements (O'Hagan *et al*, 2006). If domain knowledge is to be elicited from multiple experts, then an additional step of attempting to resolve multiple judgements into a single probability distribution representing the group is required. There are two main approaches to this: *mathematical aggregation*, which uses a mathematical rule to combine probability distributions, and *behavioural aggregation*, which attempts to bring the experts to a consensus. For more information see Cooke, 1991; O'Hagan and Oakley, 2014; Wilson and Farrow, 2018; Barons *et al.*, 2018.

Pearlean Causal DAGs

One of the main purposes of causal inference is to predict how the engineered system might behave when it is manipulated. To be precise, we introduce the non-mathematical concept of an `intervention`, to which context-specific semantics must be associated. For example, in the context of a weir, an `intervention` might constitute removal of the weir, in effect setting the RV $X_i = 0$ when X_i represents the height of the weir.

Pearl, 2009 popularised a specific class of causal DAG models that behave in a particularly simple way under `intervention`. To make this precise, we consider a subset $S \subset V$ of the RVs on which an intervention may be performed, and denote by $\mathrm{do}(\mathbf{X}_S = \mathbf{x})$ the `intervention` that sets the RVs $\mathbf{X}_S$ to the fixed value $\mathbf{x}_S$. Then we say that a causal DAG G is "Pearlean" if the distribution of the RVs $\mathbf{X}_{V \setminus S}$ under `intervention` satisfies

$$p(\mathbf{X}_{V \setminus S} \mid \mathrm{do}(\mathbf{X}_S = \mathbf{x}_S)) = \prod_{v \in V \setminus S} p(X_v | \mathbf{X}_{\pi(v)})\big|_{\mathbf{X}_S = \mathbf{x}_S} . \tag{3}$$

The notation here means that each instance of a RV in $\mathbf{X}_S$ on the right hand side is held fixed equal to the associated value in $\mathbf{x}_S$; in particular, the behaviour of the joint RV $\mathbf{X}_V$ under an intervention is assumed to be a straight-forward transformation of (and only of) the joint distribution $p(\mathbf{X}_V)$ of $\mathbf{X}_V$ describing $\mathbf{X}_V$ in the non-interventional context. In the Pearlean framework it is only necessary for (3) to hold for the specific subset S of the RVs on which an intervention is actually being considered. For a full discussion of Pearlean causal DAGs relative to more general causal models in which an intervention can change conditional distributions in respects that are not captured by a Pearlean causal DAG, see Section 7 of Dawid, 2010. The effect of `intervention` for a Pearlean causal DAG can also be generalised to interventions that change the distributional nature of the RVs $\mathbf{X}_S$, but details are reserved for standard references (e.g. Eaton and Murphy, 2007; Pearl, 2009).

The additional structure that is encoded in a Pearlean causal DAG is sufficient to allow prediction of the effect of an intervention on the engineered system, as explained next.

Estimation of Causal Effects

An important task in the causal context is to quantify "how much" one RV depends on another. Equivalently, an understanding of the strength of causal dependencies is crucial in the design of a targeted intervention with a causal objective, such as in weir modification or removal, where a minimal, cost-efficient intervention is preferred. Here we demonstrate how this is achieved with the `intervention` semantics that are provided in the Pearlean DAG framework. The "average causal effect" (ACE) of RV X_i on RV X_j is defined as the function

$$\mathrm{ACE}(x) \;=\; \frac{\partial}{\partial x_i} \int X_j \; p(\mathbf{X}_{V \setminus \{i\}} \mid \mathrm{do}(X_i = x_i)) \; \mathrm{d}\mathbf{X}_{V \setminus \{i\}}. \tag{4}$$

The integral in (4) represents the expected value of X_j under the `intervention` $\mathrm{do}(X_i = x_i)$; this is then differentiated with respect to x_i to obtain the sensitivity of this expectation with respect to x_i, which is the ACE. Several alternative measures of causal dependence to the ACE are also widely-used (e.g. Rosenbaum and Rubin, 1983; Pearl, 2001; Hudgens and Halloran, 2012).

Causation in Time

The causal DAG presented in this article does not refer to an explicit time-dependence in the engineered system, yet in many applications the causal semantics are premised on one event being the trigger for another subsequent event. There is therefore a need to distinguish between discrete and continuous time models.

A straight-forward extension to the causal DAG model that captures time-dependence is the "dynamic Bayesian network" (DBN; Ghahramani, 1997). In a DBN, RVs are endowed with a second index $n \in \mathbb{N}$ such that $X_{v,n}$ represents the value of the RV X_v at the nth discrete time point. Often the time points $t_1, t_2, \ldots$ are constrained to be evenly spaced, with increment $\Delta = t_{n+1} - t_n$. A `direct cause` X_u of X_v is represented in the DBN by a collection of edges $X_{u,n} \to X_{v,n+1}$ for each $n \in \mathbb{N}$. The DBN has close connections with vector autoregressive models from econometrics, where the causal framework is related (but not identical) to the Granger causality framework (Granger, 1969). Weir removal at time n_0, for example, in the context of the DBN corresponds to an `intervention` $\mathrm{do}(X_{\mathrm{WH},n} = 0 \;\forall\; n \geq n_0)$ that fixes the height of the weir to zero at all subsequent time points. Estimation of causal effects in DBNs is discussed in Brodersen *et al.*, 2015.

The $\Delta \downarrow 0$ limit of a DBN model is a continuous time model that can, in some cases, be described by a stochastic differential equation (SDE):

$$\mathrm{d}\mathbf{X}_V = \mathbf{f}(\mathbf{X}_V)\mathrm{d}t + \mathbf{g}\mathrm{d}\mathbf{B}$$

Here $\mathbf{f}$, $\mathbf{g}$ are drift and diffusion coefficients and $\mathbf{B}$ is a Brownian motion. The analogous notion of a weir removal `intervention` for the SDE is denoted $\mathrm{do}(X_v(t) = 0 \; \forall \; t \geq t_0)$. In this case, Sokol and Hansen, 2013 argued that a natural definition for the continuous time dynamics under `intervention` is

$$\mathrm{d}\mathbf{X}_V = \mathbf{f}(\mathbf{X}_V \mid \mathrm{do}(X_v(t) = 0 \; \forall \; t \geq t_0))\mathrm{d}t + \mathbf{g}\mathrm{d}\mathbf{B} \tag{5}$$

where

$$\mathbf{f}(\mathbf{X}_V \mid \mathrm{do}(X_v(t) = 0 \; \forall \; t \geq t_0)) = \mathbf{f}(\mathbf{X}_V)|_{X_v=0}.$$

In particular the definition given here can be recovered by applying a fine time discretisation $\Delta = t_j - t_{j-1} << 1$ to the original SDE to obtain a DBN, then using the definition of a Pearlean causal DBN and taking the limit $\Delta \downarrow 0$ to obtain (5). This provides a natural generalisation of Pearlean causal DAGs to model engineering systems that evolve in continuous-time.

Other Causal Graphical Models

The causal DAG is a specific example of a causal graphical model, but other classes of causal graphical model have been developed. In general, a causal model is based on certain non-mathematical definitions and formal axioms for causal reasoning and deduction are stated. Such a model is "graphical" when the causal model can be represented as a graph and the deductive process of drawing conclusions based on the stated axioms can be represented as a sequence of graphical manipulations. Examples of causal graphical models include nested Markov models (Shpitser *et al.*, 2014), chain event graphs (Thwaites *et al.*, 2010; Yu *et al.*, 2020) and graphical models that are induced as the margins of causal DAG models (Evans, 2016); each of these can be used to reason about the presence of unmeasured confounders.

Summary

This completes our brief exposition of causal graphical models in the abstract; the interested reader is directed toward the more technical introductions of Spirtes, 2010; Pearl, 2010; Dawid, 2010 for further detail.

The actual calculation of various probability distributions implied by a DAG can be automated with dedicated software, such as Bayes Fusion (www.bayesfusion.com) and Agena Risk (www.agenarisk.com), along with purpose-built (Perov *et al.*, 2019) and generic probabilistic programming software such as STAN (mc-stan.org). However, most software presumes that all RVs are of the same mathematical type (e.g. discrete, continuous, categorical) and in practice this can impose restrictions on the statistical model in order to fit into such a homogeneous framework. For this reason, as well as to improve the pedagogy, we include explicit probabilistic derivations in the main text.

APPENDIX II. APPENDIX II

This appendix contains full details of the causal DAG model that was used.

Direct Causes and Elicitation of the DAG

Once the RVs $\mathbf{X}_V$ have been specified, the edges of the DAG can be elicited. This is equivalent to specifying the parents of each RV in the DAG. Recall that these represent `direct causes`, as opposed to mere statements about correlation. Certain edges are trivially included; for example an edge $X^{(i)}_{\text{EQU1}} \rightarrow X^{(i)}_{\text{CA}}$ should be included since the weir is defined to have failed the condition assessment whenever one of the failure modes, such as $X^{(i)}_{\text{EQU1}}$, has occurred. In what follows we identify the parents for nodes related to failure modes EQU1, EQU2, UPL and PIP, which draws on traditional techniques from engineering assessment. Description of the remainder of the DAG structure will be deferred to the next section, where the associated conditional distributions are specified.

EQU1: Failure Due to Overturning

The first failure mode we considered was overturning of the weir due to rotation about the toe, as shown in Figure 14a. The assessment here is similar to that used for other engineered retaining structures, with the weight of the structure being resolved into downward forces at the centre of gravity of the structure, resisting the overturning moment instigated by the water pressure behind the back face of the weir.

Two kinds of moment must be resolved; horizontal moments due to water pressure and vertical moments due to weight. The horizontal force exerted by the depth of water on the weir was assumed to be

$$\text{force} = \frac{\rho_{\text{water}} g h^2}{2} X_{\text{CW}} \text{ (N)}$$

where $\rho_{\text{water}} = 9970$ (Nm^{-3}) is the density of water, $g = 9.81$ (Nkg^{-1}) is the gravitational constant and h (m) is the height of the body of water. The force was resolved at one third of the height h of the water, acting at the centroid of the triangular pressure distribution.

The vertical forces due to weight were assumed to be

$$\text{force} = \frac{\rho g a}{2} X_{\text{CW}} \text{ (N)}$$

where $\rho = \rho_{\text{water}}$ (kNm^{-3}) in the case of water or $\rho = X_{\text{WDI}}$ (kNm^{-3}) in the case of weir material and a (m^2) is the cross-sectional area of the body being considered.

The failure mode EQU1 is defined to have occurred when the total clockwise moment about the toe of the weir is > 0. It follows that the parent nodes of $X_{\text{EQU1}}^{(i)}$ in the DAG must include the geometric RVs X_{L}, X_{WH}, X_{SU}, X_{SD}, X_{ED} involved in the moment calculations, in addition to the weir density X_{WDI}, that are needed to determine whether failure mode EQU1 has occurred. Note that, since all moments are proportional to X_{CW}, it is clear that this failure mode occurs independently of the channel width X_{CW} and there is therefore no edge $X_{\text{CW}} \rightarrow X_{\text{EQU1}}^{(i)}$ in the DAG. (This is the case for all four failure modes considered.)

EQU2: Failure Due to Sliding

The second failure mode that we considered was failure due to sliding, which occurs when the friction of the weir and its embedment is overcome by the horizontal force exerted by the water. The friction force was modelled as

$$\text{force} = N \tan(X_{\text{SFA}}) \text{ (N)}$$

where N (N) is the total downward force due to the combined weight of the weir and water, as resolved above in EQU1, and X_{SFA} is the soil friction angle (Novak, 2014). Failure mode EQU2 is defined to have occurred when

$$T > X_{\text{L}} X_{\text{C}} X_{\text{CW}} + N \tan(X_{\text{SFA}})$$

where T is the total horizontal force, as resolved above in EQU1, and X_{C} is the cohesion of the soil. The parents of X_{EQU2} in the DAG therefore include the same geometric RVs required in EQU1, together with X_{SFA} and X_{C}.

UPL: Failure Due to Uplift

The third failure mode considers that the upward water pressure is high enough to vertically displace the weir. This is illustrated in Figure 14b. Uplift pressure is determined through calculation of the hydraulic gradient as

$$P_{\mathrm{uplift}} = \frac{\rho_{\mathrm{water}} g (X_{\mathrm{WH}} - X_{\mathrm{DWD}}) X_{\mathrm{CW}}}{2X_L}.$$

The total pressure downward due to the weight of the floor of the weir is

$$P_{\mathrm{floor}} = \frac{X_{\mathrm{WDI}} g a_{\mathrm{weir}} X_{\mathrm{CW}}}{X_{\mathrm{L}}}$$

where a_{weir} (m^2) is the cross-sectional area of the weir. The density of the floor material and its thickness dictate the resisting pressure. Meanwhile the floor length X_{L} contributes to the hydraulic gradient (Novak, 2014).

The failure mode UPL is defined to have occurred if $P_{\mathrm{floor}} < P_{\mathrm{uplift}}$. The parents of X_{UPL} in the DAG therefore include the geometric RVs X_{L}, X_{WH}, X_{SU}, X_{SD} and X_{ED} required to compute cross-sectional area of the weir, along with $X^{(i)}_{\mathrm{DWD}}$ and X_{WDI}.

PIP: Failure Due to Piping

The final failure mode considered is due to piping, which describes the action of seepage under the floor of the weir. The relationship between the seepage streamline lengths and the hydraulic head in the system defines the exit gradient of the weir system

$$G_e = \frac{X_{\mathrm{WH}} - X_{\mathrm{DWD}}}{X_{\mathrm{L}}},$$

which arises from a simple linear model, more sophisticated methods based on partial differential equations can also be used (Khosla *et al.*, 1954). Different bed soils have different permissible exit gradients and we define failure due to piping to have occurred when $G_e > G_e^*$

where G_e^* is a constant specific to a given soil type. This constant can be determined from literature using the sediment size distribution X_{SSD}. Inspection of sediment samples from Clerkington weir suggested that a value $G_e^* = 0.22$ be used. The parents of X_{PIP} in the DAG therefore are X_{WH}, $X_{\text{DWD}}^{(i)}$, X_{L} and X_{SSD}.

Elicitation of Conditional and Interventional Distributions

To each unobserved RV X_v, $v \in V \setminus O$, we must specify the conditional distribution $p(X_v|\mathbf{X}_{\pi(v)})$ of X_v given its parents $\text{X}_{\pi(v)}$ in the DAG. In the case where there are no parents, this is simply the marginal distribution $p(X_v)$ that must be specified. Several conditional distributions are deterministic and have already been specified when we elicited the edges of the DAG. The remainder of the conditional distributions are now elicited.

Source Nodes

A maximal value for the weir density X_{WDI} was informed by information available for similar material (MacGregor, 1945). In particular, we assumed that X_{WDI} is uniformly distributed between 0 and $0.9 \times 26,000$ (Nm^{-3}) where the factor of 0.9 accounts for visually determined voiding in the weir. The lower bound of 0 allows for the possibility that large sections of the interior of the weir are completely voided. The soil properties X_{SFA}, X_{C} were informed from sediment samples and geology tables. For X_{SFA} an elicited uniform distribution of between 0 and 65 degrees was used, representing a range from pure clay to compact sandy loam. For the embedment depth X_{ED} a uniform distribution between 0 (m) and 3 (m) was elicited.

Intermediate Nodes

For X_{C} we took $p(X_{\text{C}}|X_{\text{SFA}})$ to be Gaussian with mean $(5000/35) \times X_{\text{SFA}}$ (m) and standard deviation 250 (m). For the upstream water depth, seepage under the weir is a possibility, which means that the embedment depth X_{ED} may be relevant. For the present paper we neglect this possibility and simply related $X_{\text{UWD}}^{(i)}$ to the flow $X_{\text{F}}^{(i)}$ on day i as follows:

$$X_{\text{UWD}}^{(i)} = X_{\text{WH}} + \left(\frac{X_{\text{F}}^{(i)}}{c_d g^{1/2} X_{\text{CW}}} \right)^{2/3}$$

where c_d is the discharge coefficient, taken to be 0.9 for our weir. The "crump" model most closely represents the geometry that we studied and we therefore used the associated flow equation from Novak, 2014. An upstream flood is defined to have occurred ($X_{\text{UF}}^{(i)} = 1$) when the upstream water depth $X_{\text{UWD}}^{(i)}$ exceeds the bank height X_{BH}.

The relationship between upstream and downstream water levels is challenging to characterise due to dependence on the downstream flow characteristics of the river (Novak, 2014), and demands hydrological expertise beyond the scope of this project. We proceed with a simple statistical model for $p(X_{\text{DWD}}^{(i)} | X_{\text{UWD}}^{(i)}, X_{\text{WH}})$, namely the approximation

$$X_{\text{DWD}}^{(i)} + X_{\text{WH}} - X_{\text{UWD}}^{(i)} \sim \text{Gamma}(1, 0.1) \tag{6}$$

was used. Here the gamma distribution is in the shape-scale parametrisation and we emphasise that (6) would need to be replaced with a model driven by hydrological considerations in the context of a decision-making tool.

The fish passability RV is determined by two aspects; (i) the overflow *head* at weir $X_{\text{UWD}}^{(i)} - X_{\text{WH}}$ and (ii) the air *gap* $X_{\text{UWD}}^{(i)} - X_{\text{DWD}}^{(i)}$. As discussed in the main text, the RV $X_{\text{FP}}^{(i)}$ is categorical and its value is determined as follows:

$$X_{\text{FP}}^{(i)} = \begin{cases} \text{total} & \text{head} > 0.1, \text{gap} < 0.5 \\ \text{high} & \text{head} > 0.1, 0.5 \leq \text{gap} < 0.9 \\ \text{medium} & \text{head} > 0.1, 0.9 \leq \text{gap} < 1.4 \\ \text{low} & \text{otherwise} \end{cases},$$

based on the detailed analysis of Baudoin *et al.*, 2014.

Interventional Distributions

Beyond eliciting conditional distributions, to address an explicitly causal hypothesis we must specify how these conditional distributions change under an `intervention` on the system. For this purpose we endow our causal graphical model with the Pearlean structure that was previously described. Thus (3) defines the collection of interventional distributions that were used as the basis for causal inferences about the river-weir ecosystem. This crucial final step completes the specification of the causal DAG model.

Data Availability Statement Some or all data, models, or code used during the study were provided by a third party. Direct requests for these materials may be made to the provider as indicated in the Acknowledgements.

Acknowledgements: The authors are grateful to Charles Stevenson for access to the weir at Clerkington, as well as Anna Griffin and colleagues from the Scottish Environmental Protection Agency and Matthew O'Hare from the Centre for Ecology and Hydrology, for useful discussions related to the investigation. The authors wish to thank Professor Jim Smith for detailed feedback on an earlier draft of the manuscript. Data from the UK National River Flow Archive has been used with permission.

REFERENCES

Balke, A., Pearl, J. (1994) Probabilistic evaluation of counterfactual queries. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, (pp. 230-237).

Barons, M.J., Wright, S.N., Smith, J.Q. (2018) Eliciting probabilistic judgments in integration decision support sysems. In *Elicitation: The Science and Art of Structuring Judgements*, Chapter 17 (pp. 445-494), Eds Luis C. Dias, Alec Morton, John Quigley, Springer.

Bartlett, M., Cussens, J. (2013) Advances in Bayesian network learning using integer programming. In proceedings of the 29th Conference on Uncertainty in Artificial Intelligence (UAI 2013), p. 182-191.

Baudoin, J.M., Burgun, V., Chanseau, M., Larinier, M., Ovidio, M., Sremski, W., Steinbach, P. and Voegtle, B., 2014. Assessing the passage of obstacles by fish. Concepts, design and application. Onema, France.

Bell, D., Cox, L., Jackson, S. and Schaefer, P. (1992), January. Using causal reasoning for automated failure modes and effects analysis (FMEA). In Annual Reliability and Maintainability Symposium 1992 Proceedings (pp. 343-353). IEEE.

Bhandari J, Khan F, Abbassi R, Garaniya V, Ojeda R. (2017) Pitting degradation modeling of ocean steel structures using Bayesian network. Journal of Offshore Mechanics and Arctic Engineering. 1;139(5):051402.

Brodersen, K.H., Gallusser, F., Koehler, J., Remy, N., Scott, S.L. (2015) Inferring causal impact using Bayesian structural time-series models. The Annals of Applied Statistics, 9(1), pp.247-274.

Bühlmann, P., Peters, J., Ernest, J. (2014) CAM: Causal additive models, high-dimensional order search and penalized regression. The Annals of Statistics, 42(6), pp.2526-2556.

Chaloner, K. and Verdinelli, I. (1995) Bayesian experimental design: A review. Statistical Science, pp.273-304.

Claassen, T., Heskes, T. (2012) A Bayesian Approach to Constraint Based Causal Inference. In UAI 2012, Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence.

Cooke, R. (1991). Experts in uncertainty: opinion and subjective probability in science. Oxford University Press on Demand.

Cotter, T.S. (2015). Statistical Engineering: A Causal-Stochastic Modeling Research Update. In Proceedings of the International Annual Conference of the American Society for Engineering Management. (p. 1). American Society for Engineering Management (ASEM).

Dawid, A.P. (2000) Causal inference without counterfactuals. Journal of the American Statistical Association, 95(450), pp.407-424.

Dawid, A.P. (2010) Beware of the DAG!. NIPS Causality: Objectives and Assessment, 6, pp.59-86.

Eaton, D., Murphy, K.P. (2007) March. Exact Bayesian structure learning from uncertain interventions. In AISTATS (pp. 107-114).

Evans, R. (2016) Graphs for margins of Bayesian networks. Scandinavian Journal of Statistics, 43 (3), pp 625-648.

Fienen, M.N., Masterson, J.P., Plant, N.G., Gutierrez, B.T. and Thieler, E.R. (2013) Bridging groundwater models and decision support with a Bayesian network. Water Resources Research, 49(10), pp.6459-6473.

Garthwaite, P.H., Kadane, J.B., and O'Hagan, A. (2005) Statistical Methods for Eliciting Probability Distributions, Journal of the American Statistical Association, 100(470), pp.680-701.

Geiger, D., Verma, T. and Pearl, J. (1990). d-separation: From theorems to algorithms. In Machine Intelligence and Pattern Recognition (Vol. 10, pp. 139-148). North-Holland.

Ghahramani, Z. (1997) Learning dynamic Bayesian networks. In *International School on Neural Networks*, Initiated by IIASS and EMFCSC (pp. 168-197). Springer, Berlin, Heidelberg.

Granger, C.W.J. (1969) Investigating Causal Relations by Econometric Models and Cross-spectral Methods. Econometrica. 37 (3): 424-438.

Hudgens, M.G., Halloran, M.E. (2012) Toward causal inference with interference. Journal of

the American Statistical Association.
Imai, K., King, G., Stuart, E.A. (2008) Misunderstandings between experimentalists and observationalists about causal inference. Journal of the Rroyal Statistical Society: Series A, 171(2), pp.481-502.
Kennard, M.F.; Owens, C.L.; Reader, R.A. (1996) Engineering guide to the safety of concrete and masonry dam structures in the UK. CIRIA Report 148, London, UK.
Khosla, A.N.; Bose, N.K.; Taylor, E.M. (1954) Design of Weirs on Permeable Foundation, Publication No.12, Central Board of Irrigation and Power, New Delhi.
King, S. and O'Hanley, J.R., 2016. Optimal fish passage barrier removal—revisited. River Research and Applications, 32(3), pp.418-428.
Kline, R.B. (2015) Principles and practice of structural equation modeling. Guilford Publications.
MacGregor, A.G. (1945) The mineral resources of the Lothians. British Geological Survey Information Services, Internal Report IR/04/017.
Meinshausen, N., Bühlmann, P. (2006) High-dimensional graphs and variable selection with the lasso. The Annals of Statistics, pp.1436-1462.
Morgan, S.L., Winship, C. (2014) Counterfactuals and causal inference. Cambridge University Press.
National River Flow Archive, 2019, `https://nrfa.ceh.ac.uk`, UK Centre for Ecology & Hydrology (UKCEH), Wallingford.
Novak, P., Moffat, A.I.B., Nalluri, C. and Narayanan, R., 2014. Hydraulic structures. CRC Press.
O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., and Rakow, T. (2006). Uncertain judgements: eliciting experts' probabilities. John Wiley & Sons.
O'Hagan, A. and Oakley, J. (2014) SHELF: the Sheffield elicitation framework.
Pearl, J. (1995) Causal diagrams for empirical research. Biometrika, 82(4), pp.669-688.

Pearl, J. (2001) Direct and indirect effects. In Proceedings of the seventeenth conference on uncertainty in artificial intelligence (pp. 411-420). Morgan Kaufmann Publishers Inc.

Pearl, J. (2009) Causality: Models, Reasoning and Inference. Cambridge University Press, Cambridge, second edition, 2009.

Pearl, J. (2010) An introduction to causal inference. The International Journal of Biostatistics, 6(2).

Perov, Y., Graham, L., Gourgoulias, K., Richens, J.G., Lee, C.M., Baker, A. and Johri, S. (2019) Multiverse: Causal reasoning using importance sampling in probabilistic programming. arXiv preprint arXiv:1910.08091.

Pickles, A.; Sandham, R.; Simpson, B.; Bond, A. (2014) Application of Eurocode 7 to the design of flood embankments. C749, CIRIA, London, UK.

Rosenbaum, P.R., Rubin, D.B. (1983) The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), pp.41-55.

Rothman, K.J. and Greenland, S. (2005) Causation and causal inference in epidemiology. American journal of public health, 95(S1), pp.S144-S150.

Salvaneschi, P., Cadei, M. and Lazzari, M. (1997). A causal modelling framework for the simulation and explanation of the behaviour of structures. Artificial Intelligence in Engineering, 11(3), pp.205-216.

SEPA Water Environment Hub `https://www.sepa.org.uk/data-visualisation/water-environment-hub/`

Shpitser, I., Evans, R.J., Richardson, T.S., Robins, J. M. (2014) Introduction to nested Markov models. Behaviormetrika, 41(1), pp.3-39.

Smith, N. (1972) A history of dams. Citadel, UK.

Smith, J.Q. (2010) Bayesian Decision Analysis: Principles and Practice. Cambridge University.

Sokol, A., Hansen, N.R. (2013) Causal interpretation of stochastic differential equations. arXiv:1304.0217.

Spirtes, P., Glymour, C., Scheines, R. (2000) Causation, Prediction and Search. Springer-Verlag, New York, Second edition, 2000.

Spirtes, P. (2010) Introduction to causal inference. Journal of Machine Learning Research, 11(May), pp.1643-1662.

Spohn, W. (1980) Stochastic independence, causal independence, and shieldability. Journal of Philosophical Logic, 9:73–99, 1980.

Strahler, A.N. (1952) Hypsometric (area-altitude) analysis of erosional topography. Geological Society of America Bulletin, 63(11), pp.1117-1142.

Suraji, A., Duff, A.R. and Peckitt, S.J. (2001). Development of causal model of construction accident causation. Journal of construction engineering and management, 127(4), pp.337-344.

Thwaites, P., Smith, J.Q., Riccomagno, E. (2010) Causal analysis with chain event graphs. Artificial Intelligence, 174(12):889-909.

Tong, Y. and Tien, I. (2017). Algorithms for Bayesian network modeling, inference, and reliability assessment for multistate flow networks. Journal of Computing in Civil Engineering, 31(5), p.04017051.

Wilkerson, R.L. and Smith, J.Q. (2019) Customised Structural Elicitation. In *Expert Judgement in Risk and Decision Analysis*, Springer International Series in Operations Research and Management Science, to appear.

Wilson KJ, Farrow M. (2018) Combining judgements from correlated experts. In: Luis C. Dias, Alec Morton, John Quigley, ed. Elicitation: the Science and Art of Structuring Judgement. New York: Springer, pp.211-240.

Wu, W.S., Yang, C.F., Chang, J.C., Château, P.A. and Chang, Y.C. (2015). Risk assessment by integrating interpretive structural modeling and Bayesian network, case of offshore pipeline project. Reliability Engineering & System Safety, 142, pp.515-524.

Wu, X., Liu, H., Zhang, L., Skibniewski, M.J., Deng, Q. and Teng, J. (2015). A dynamic Bayesian network based approach to safety decision support in tunnel construction. Reliability Engineering & System Safety, 134, pp.157-168.

Yu, X., Smith, J.Q., Nichols, L. (2020) Bayesian Learning of Causal Relationships for System Reliability. arXiv:2002.06084.

List of Figures

figures/fig4.pdf

FIG. 1: The stream network for the Tyne catchment. This network was extracted using ArcGIS hydrology tools and an assumed 0.1 km^2 area for stream initiation, which matches well with water courses shown on Ordnance Survey maps. The main branch of the River Tyne can be described as a 6th order stream, using the terminology of Strahler (1952). Clerkington Weir is located approximately one third of the distance down this main branch, at an elevation of 48 m above Ordnance Datum. The weir lies approximately 400 m downriver from where the 5th order Gifford Water joins the Tyne, which creates a sharp rise in its contributing drainage area. At the Clerkington Weir, the river captures a total area of approximately 250 km^2, which represents 79% of the entire Tyne catchment.

figures/fig5.jpg

FIG. 2: Clerkington Weir, as observed in 2018. (Image: Stephenson, 2018.)

figures/fig12a.pdf

(a) Original drawing of the 18th Century Carron Dam in Scotland.

figures/fig12b.pdf

(b) Original drawing of the Saltersford Weir in Cheshire, constructed in the 1820's.

FIG. 3: Historic weir construction in the UK.

FIG. 4: A geometric characterisation of the river-weir system. The random variables annotated on the diagram are defined in the main text.

figures/fig7.pdf

FIG. 5: Data collected on the 28th of September 2018, using a Leica GS08 GPS system. The river bed elevation and water surface elevation was measured with an average interval spacing of 12.5 m.

figures/fig13.pdf

FIG. 6: The elicited DAG for the river-weir ecosystem. [Dark nodes indicate observed random variables and light nodes indicate latent random variables that must be inferred. The index set $\mathcal{I}$ runs over each day from 1981-2000.]

`figures/fig8.pdf`

FIG. 7: Q1: Fish Passability. Left: A superposition of the conditional probability distributions $p(X_{\mathrm{DWD}}^{(i)}|\mathbf{X}_O)$ over downstream water depth $X_{\mathrm{DWD}}^{(i)}$ for each of the days i in the dataset. Right: A superposition of the conditional probability distributions $p(X_{\mathrm{FP}}^{(i)}|\mathbf{X}_O)$ over fish passability $X_{\mathrm{FP}}^{(i)}$ for each of the days i in the dataset.

figures/fig9.pdf

FIG. 8: Q2: Density and Embedment Depth. Contours of constant conditional probability density $p(X_{\mathrm{WDI}}, X_{\mathrm{ED}}|\mathbf{X}_O)$ are displayed.

figures/fig10.pdf

FIG. 9: Q3: Weir Removal. Left: Empirical distribution of upstream flow $X_{\mathrm{F}}^{(i)}$ from the dataset indexed by $\mathcal{I}$ (bars), together with a log-normal distribution (blue) fit to this dataset. Middle: The probability of an upstream flood event ($X_{\mathrm{UF}}^{(*)} = 1$), computed under the interventional probability distribution $p(X_{\mathrm{UF}}^{(*)} \mid \mathrm{do}(X_{\mathrm{WH}}^{(*)} = h), \mathbf{X}_O)$, based on a modified height h for the weir. Right: The average causal effect of weir height h on the probability of an upstream flood event. [In each of panel several blue curves are shown, each based on a different log-normal fit to the dataset and representing the fact that several such distributions could plausibly have given rise to the observed dataset.]

figures/fig11.jpg

FIG. 10: Dynamic re-routing of the Tyne river over the last 150 years, based on historical Ordnance Survey maps (surveyed in 1855 and 1895), aerial photographs dating from 1946, 1988 and 2009, and a GPS survey of the river centreline undertaken in September 2018. The colour of the river centre lines go from white (oldest) to dark blue (most recent). Image from Getmapping plc, courtesy of Google Earth.

figures/fig1.pdf

FIG. 11: Illustration, a 3-variable causal directed acyclic graph (DAG). If the causal Markov property holds, then we may conclude that X_1 is conditionally independent of X_3 given X_2, written $X_1 \perp\!\!\!\perp X_3 | X_2$.

figures/fig2.pdf

FIG. 12: Illustration, part of a causal DAG. From this motif we may conclude that X_k has an indirect causal dependence on X_i, but that X_i does not causally depend on X_k. This is a demonstration of logico-deductive reasoning based on a causal DAG.

figures/fig3.pdf

FIG. 13: Illustration of panel notation. For instances of "parallel" random variables in a DAG, such as X_2, X_3 and X_4 in the left hand DAG, panel notation provides a compact shorthand, as exemplified in the right hand DAG.

(a) Equilibrium failure modes (EQU1, EQU2)

(b) Uplift and piping failure modes (UPL, PIP)

FIG. 14: The four failure modes considered.

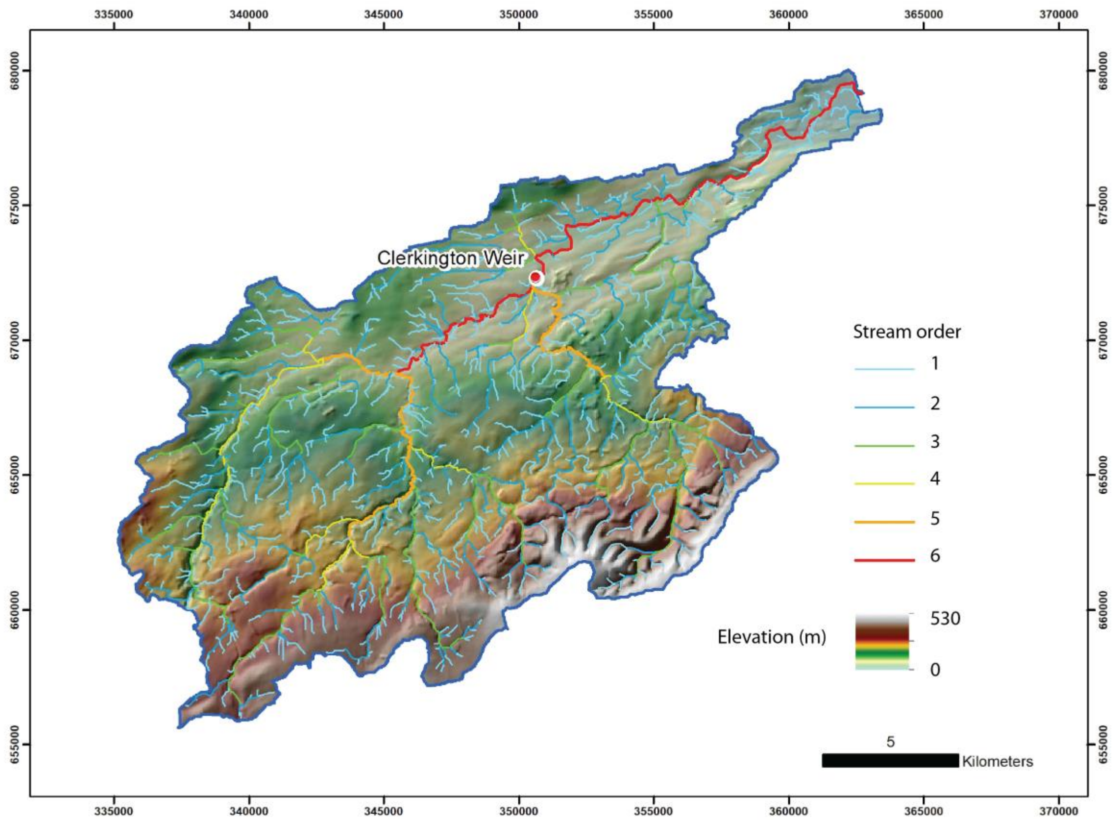

335000
340000
345000
350000
355000
360000
365000
370000
680000
675000
670000
665000
660000
655000
Clerkington Weir
Stream order
1
2
3
4
5
6
Elevation (m)
530
0
5
Kilometers

Fig 1.

Fig. 2

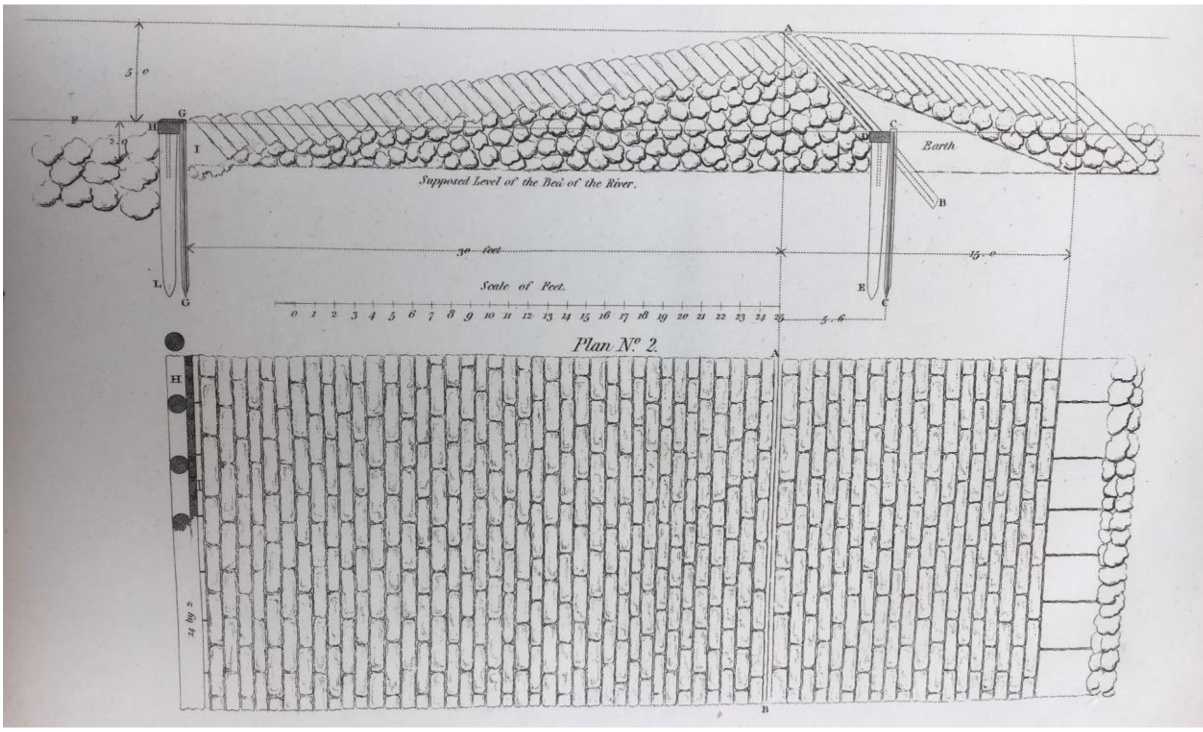

Supposed Level of the Bed of the River.
Earth
30 feet
15 . 0
Scale of Feet.
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
5 . 6
Plan No. 2.

Fig. 3a

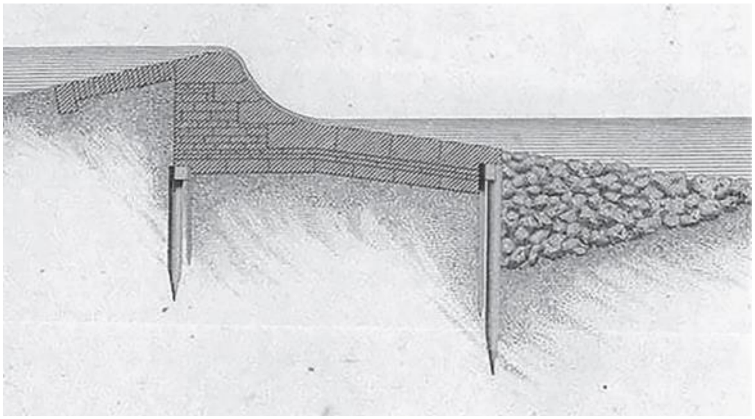


Fig. 3b

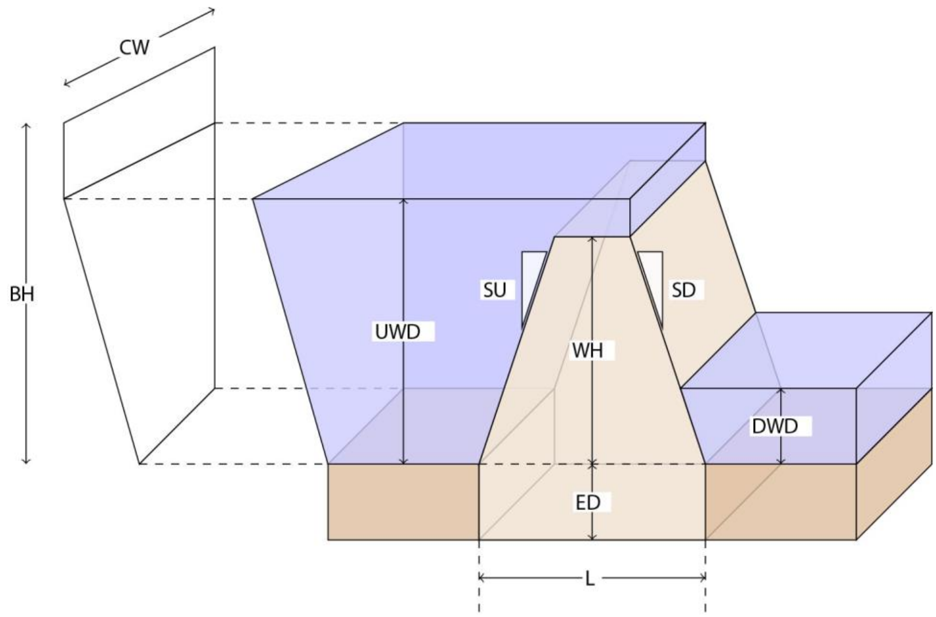

CW
BH
UWD
SU
SD
WH
DWD
ED
L

Fig. 4

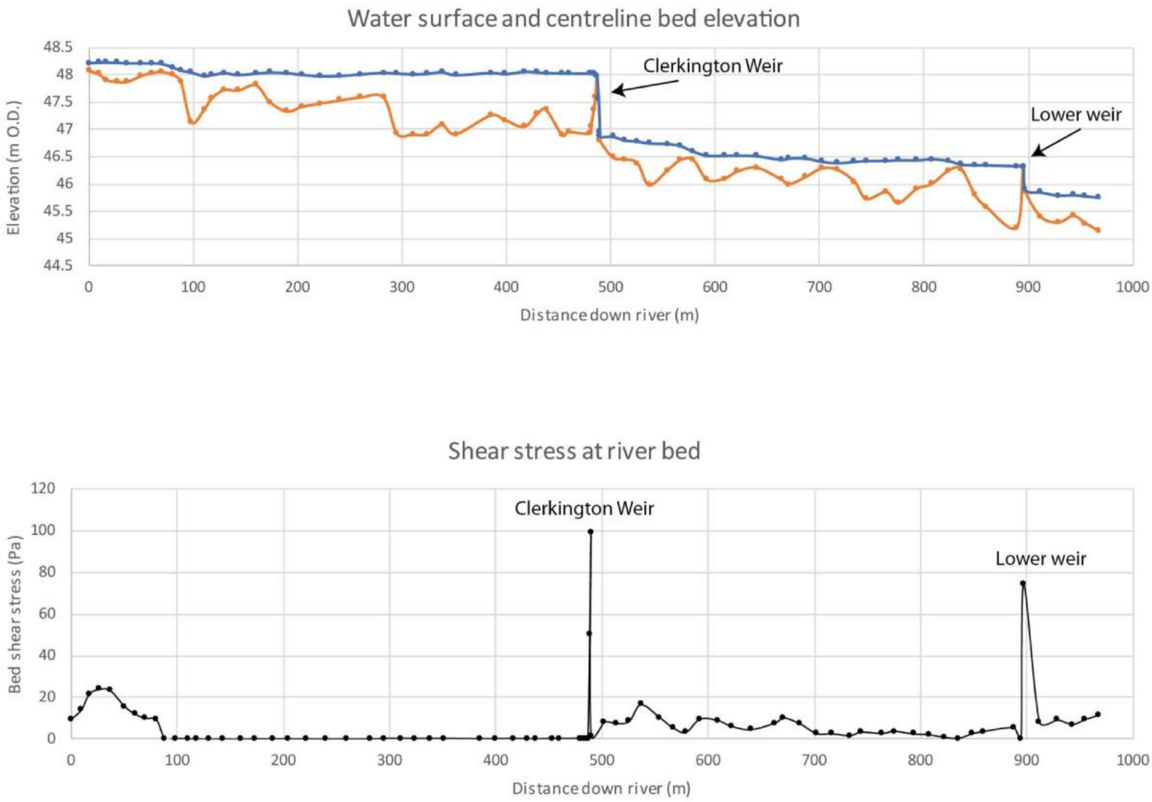

Water surface and centreline bed elevation
Clerkington Weir
Lower weir
Elevation (m O.D.)
Distance down river (m)
Shear stress at river bed
Clerkington Weir
Lower weir
Bed shear stress (Pa)
Distance down river (m)

Fig. 5

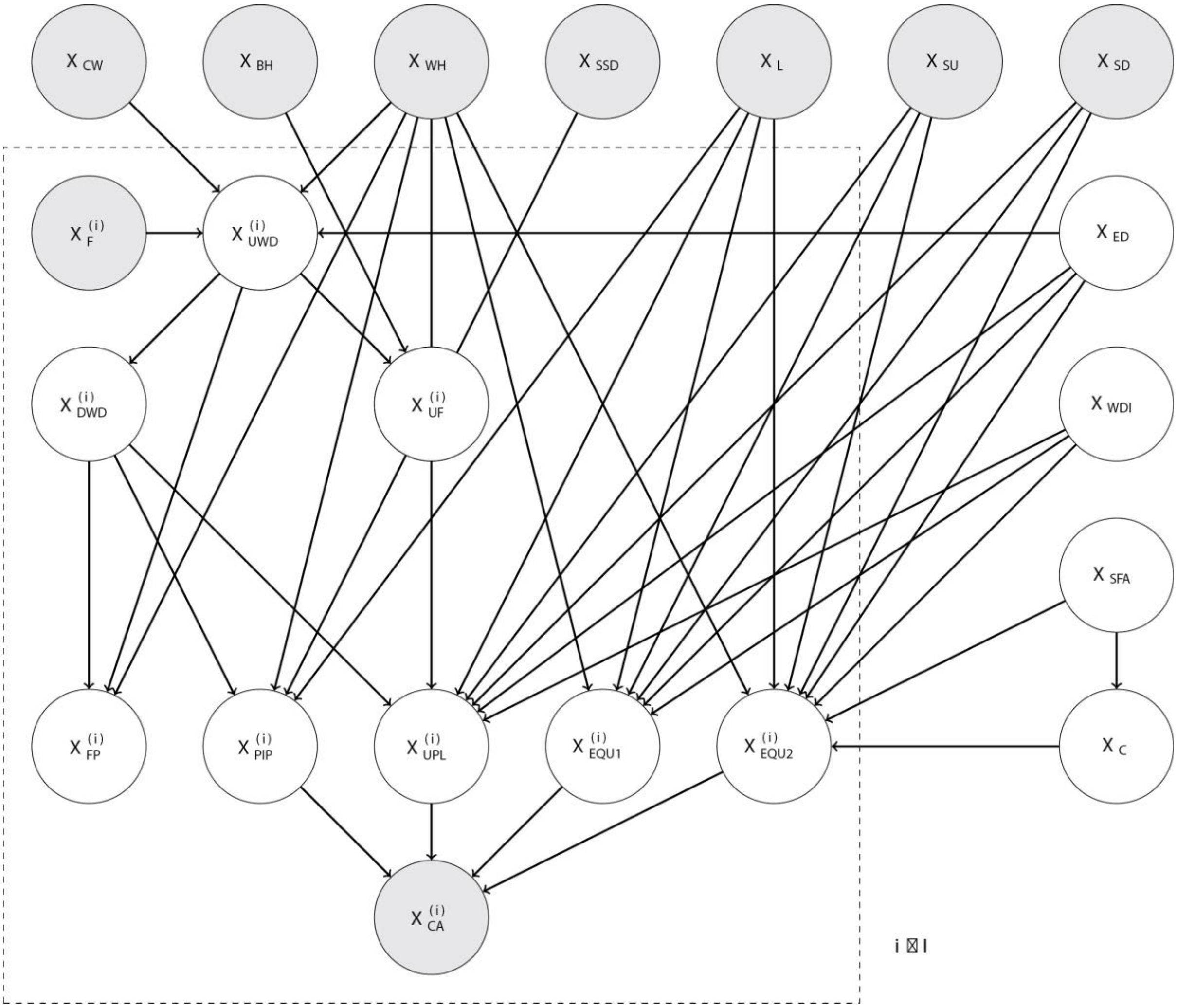

X CW
X BH
X WH
X SSD
X L
X SU
X SD
X F (i)
X UWD (i)
X ED
X DWD (i)
X UF (i)
X WDI
X SFA
X FP (i)
X PIP (i)
X UPL (i)
X EQU1 (i)
X EQU2 (i)
X C
X CA (i)
i ∈ I

Fig. 6

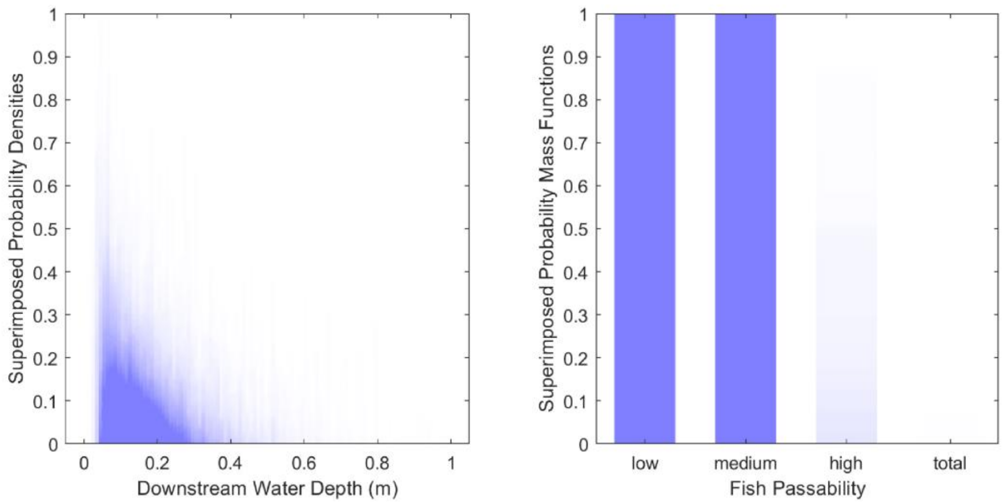
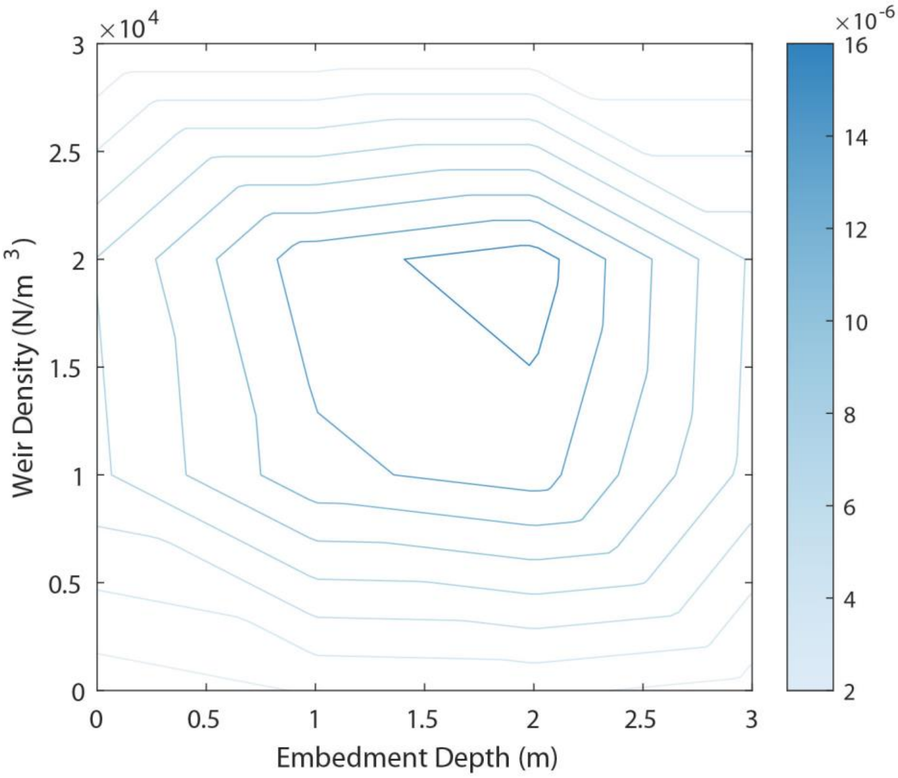
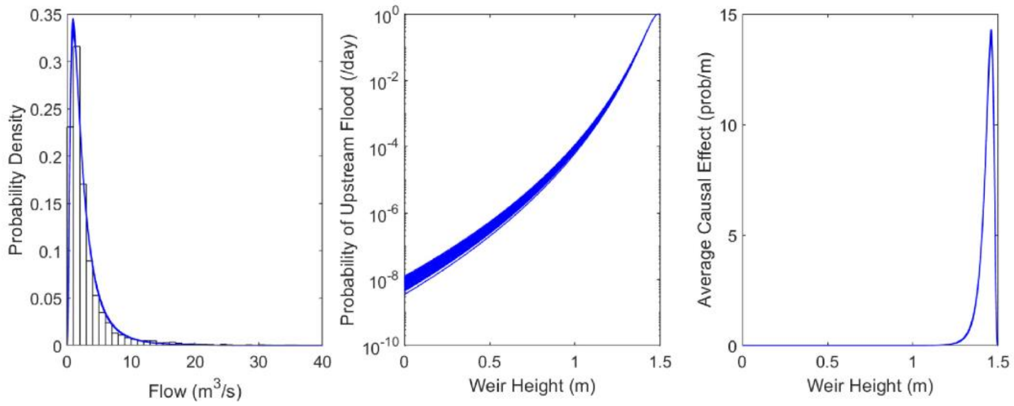

Superimposed Probability Densities
Downstream Water Depth (m)
Superimposed Probability Mass Functions
low
medium
high
total
Fish Passability

Fig. 7



×10^4
×10^-6
Weir Density (N/m 3)
Embedment Depth (m)

Fig. 8



Probability Density
Flow (m3/s)
Probability of Upstream Flood (/day)
Weir Height (m)
Average Causal Effect (prob/m)
Weir Height (m)

Fig. 9

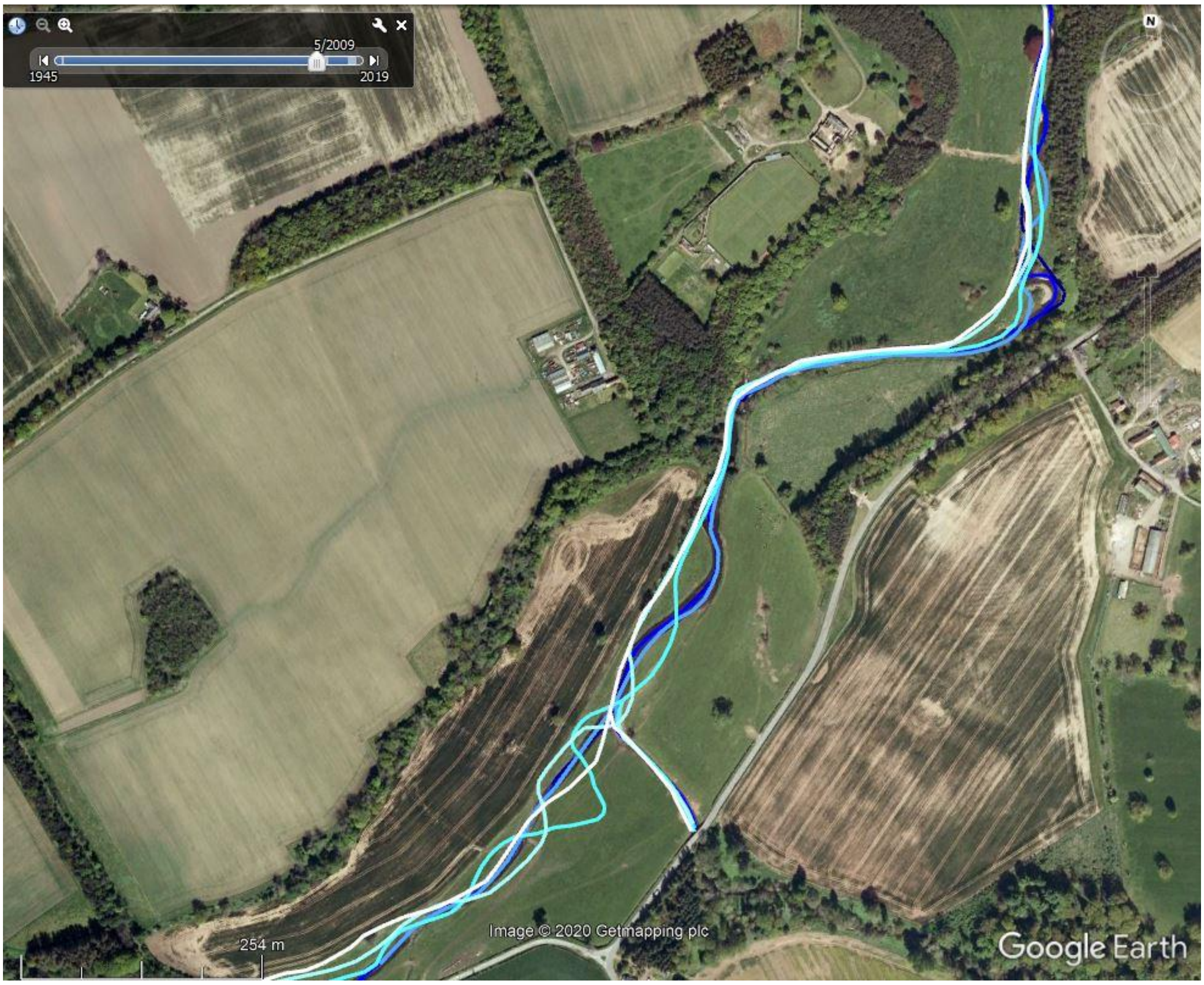


Fig. 10

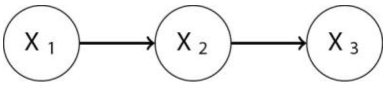


Fig. 11

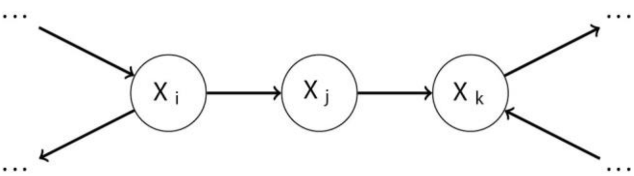


Fig. 12

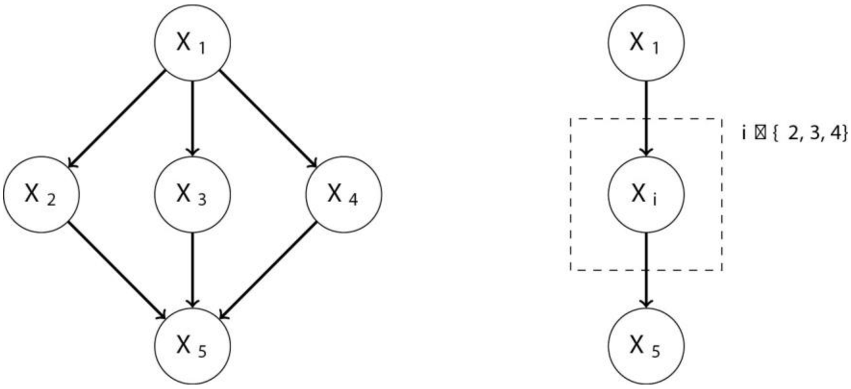


Fig. 13

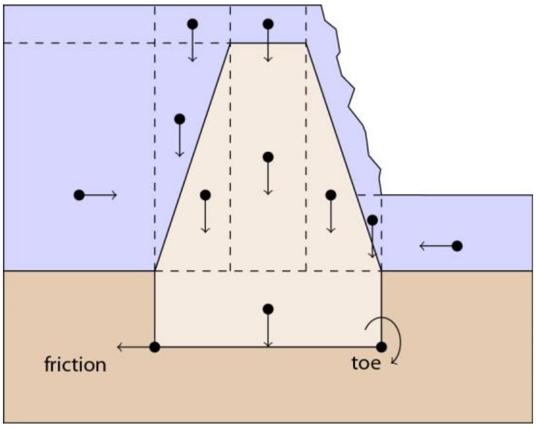


Fig. 14a

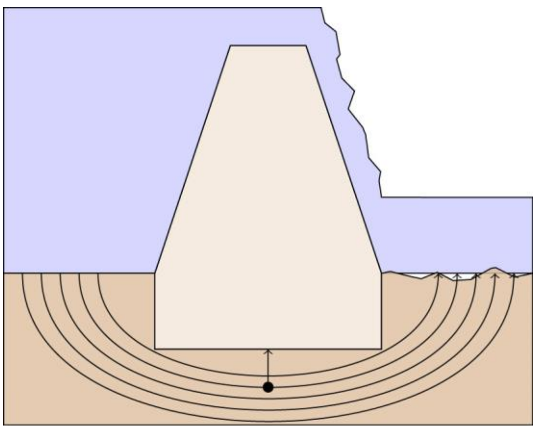


Fig. 14b