

RESEARCH ARTICLE

# Decision-theoretic inspection planning using imperfect and incomplete data

Domenic Di Francesco<sup>1,\*</sup> , Marios Chryssanthopoulos<sup>2</sup>, Michael Havbro Faber<sup>3</sup> and Ujjwal Bharadwaj<sup>4</sup>

<sup>1</sup>National Structural Integrity Research Centre, TWI Ltd., Granta Park, Great Abington, United Kingdom

<sup>2</sup>Department of Civil and Environmental Engineering, University of Surrey, Guildford, United Kingdom

<sup>3</sup>Department of the Built Environment, Aalborg Universitet, Aalborg Øst, Denmark

<sup>4</sup>TWI Ltd., Great Abington, United Kingdom

\*Corresponding author. E-mail: [domenicdifrancesco@gmail.com](mailto:domenicdifrancesco@gmail.com)

**Received:** 26 February 2021; **Revised:** 05 October 2021; **Accepted:** 14 October 2021

**Keywords:** Bayesian data analysis; multilevel modeling; partial pooling of information; risk based inspection; value of information

## Abstract

Attempts to formalize inspection and monitoring strategies in industry have struggled to combine evidence from multiple sources (including subject matter expertise) in a mathematically coherent way. The perceived requirement for large amounts of data are often cited as the reason that quantitative risk-based inspection is incompatible with the sparse and imperfect information that is typically available to structural integrity engineers. Current industrial guidance is also limited in its methods of distinguishing quality of inspections, as this is typically based on simplified (qualitative) heuristics. In this paper, Bayesian multi-level (partial pooling) models are proposed as a flexible and transparent method of combining imperfect and incomplete information, to support decision-making regarding the integrity management of in-service structures. This work builds on the established theoretical framework for computing the expected value of information, by allowing for partial pooling between inspection measurements (or groups of measurements). This method is demonstrated for a simulated example of a structure with active corrosion in multiple locations, which acknowledges that the data will be associated with some precision, bias, and reliability. Quantifying the extent to which an inspection of one location can reduce uncertainty in damage models at remote locations has been shown to influence many aspects of the expected value of an inspection. These results are considered in the context of the current challenges in risk based structural integrity management.

## Impact Statement

The primary novel component of this work is that a value of information calculation has been integrated with a partial pooling degradation model, relating information from inspection data across multiple locations. In addition, a Bayesian imputation model (of missing data) was used to estimate the prior damage model. Inspections of engineering structures typically consist of indirect measurements of some complex degradation process, sometimes in challenging environmental conditions. Even when state of the art technologies are employed, the information that they provide is imperfect. Improved probabilistic quantification of interdependencies may contribute to the development of risk-optimal inspection and maintenance planning. This is expected to be of particular benefit to ageing structures, associated with limited, incomplete, and lost information.

## 1. Introduction

### 1.1. Structural integrity management

Structural integrity management is defined as a structure's ability to function "effectively" and "efficiently" (Health and Safety Executive, 2009). This joint optimization of safety and economy is a basis for selecting appropriate investments for in-service structures. Since all structures degrade or become damaged to some extent, there may be points in time when it is worthwhile investing in activities that reduce the risk of failure. Identifying where, when and how to inspect (or more generally, to collect information) is an on-going challenge to engineers.

Fundamentally, this requires a mechanism for integrating information from an inspection with existing models of structural condition. There is also the requirement for a means of determining whether or not the expected utility of an inspection exceeds its expected cost, that is whether or not the inspection is expected to be worthwhile. As outlined in this paper, Bayesian inference provides the mathematical basis for achieving this.

### 1.2. Epistemic uncertainty

In Kiureghian and Ditlevsen (2009) various uncertainties that should be considered in structural engineering are outlined. The authors identify that contributions from *epistemic* (statistical uncertainty due to quality and availability of information) and *aleatory* (associated with inherent variability) uncertainty can be treated similarly, and this can be achieved using Bayesian methods, where both are represented in the posterior distribution. Probabilistic methods in engineering have historically focused on aleatory uncertainty and various Bayesian methods have been proposed to also account for epistemic uncertainty (Sankararaman and Mahadevan, 2013; Nannapaneni and Mahadevan, 2016). The influence of epistemic uncertainty will be largest when only small amounts of data are available. In such cases, when neglecting epistemic uncertainty, engineers will overestimate the precision of their models, which can lead to ineffective risk management. Engineering structures are often associated with the absence of large amounts of data and the presence of subject matter knowledge DNVGL (2017c). This is often cited as the reason that quantitative risk-based inspection (RBI) is infeasible in practice (API, 2016; DNVGL, 2017b). Alternatively, this can be considered a reason that structural engineering is well suited to Bayesian methods (Ang and Tang, 2007). Consequently Bayesian methods have been proposed for managing uncertainty in this context for many decades (Tang, 1973). Incorporating epistemic uncertainty is more challenging using other methods. For instance, maximum likelihood estimation is primarily concerned with obtaining point estimates of parameters, although with sufficient data the variation can also be estimated (Faber, 2012).

### 1.3. System effects

Inspection planning of systems of components within a structure (or systems of multiple structures) should consider possible dependencies between individual components (Straub and Faber, 2004). Information (such as detection of the presence or absence of damage) at one location should inform estimates of the condition of other locations (Straub and Faber, 2005). This concept is intuitive, but incorporating this effect in a quantitative inspection plan requires models capable of characterizing this dependency. In this paper, partial pooling of information between locations using multilevel Bayesian models, as detailed in Section 3, is proposed as a solution to this challenge. Note that for the purposes of this paper, a multilevel model is considered to be a more general form of a hierarchical model and is used interchangeably with *partial pooling model*. In Section 4, it is demonstrated how the results of such a model can be integrated with a decision analysis for evaluating prospective inspection options.

Sources of variation in degradation rates are expected to differ between equipment that is manufactured at scale through more established or controlled processes, and structural systems, which may be comprised of less common arrangements. In this paper, it is proposed that a binary classification of a structure or structural component as being either unique, or nominally identical to others in a population is an simplification that can be overcome. This is discussed in more detail in Section 2.

Note that there are also system effects in consequence modeling (Straub and Faber, 2004). For instance, the expected cost of failure of a structural element may be less for structures designed with redundancy. Additionally, the cost of multiple failures may not increase linearly. The work presented in this paper focuses on modeling failure probability and not failure consequences. However, Bayesian methods of quantifying uncertainty are expected to be equally valid for estimating failure consequences, although the details will be conditional on the preferences of the asset operator (and other stakeholders), which are generally not publicly available. Conversely, certain failure mechanisms are common across welded steel structures and components, and can therefore be discussed more generally.

## 2. Risk Based Inspection

### 2.1. Historical developments

RBI procedures have been developed for many industries, including petrochemical plant (Wintle et al., 2001; API, 2008; API, 2016), ship hulls (Lloyd's Register, 2017), offshore structures (Health and Safety Executive, 2009; DNVGL, 2015; Bureau Veritas, 2017; Guédé, 2018), and offshore topsides equipment (DNVGL, 2017b). Despite the variation in application, these procedures share many common features. Specifically, they are all concerned with a risk optimal allocation of resources based on some evaluation of the probability and consequence of failure. However, these high-level principles allow for significant variation in their application, and catastrophic failures have been recorded at sites with RBI systems in place. A review of two such incidents Clarke (2016) highlighted (among other issues) that procedures which prioritize inspections to be completed within a specified time-frame are “doomed to failure.” Rather, inspections that are shown to be required should be completed. In addition, this report also recommended that any decisions which results from an RBI assessment should be auditable and accountable.

Some industries make direct use of large datasets to infer expected probabilities of failure. In API (2008), the idea of a *generic failure frequency* (that summarizes failures from large numbers of components, which have been assumed to belong to a single population) is central to the recommended RBI practices. In the case of structures, this is generally not considered suitable as the variation between different structures is considered too great (Thoft-Christensen and Baker, 1982; Straub, 2004). While structures may vary more than mass manufactured process equipment, there will be some degree of commonality that they share. Conversely, no two mass manufactured components degrade in a completely identical manner and so they will also exhibit some variation. The modeling approach detailed in Section 3 is capable of accounting for this on a continuous scale.

A method originally proposed in Madsen (1987) and still used in conjunction with structural reliability analysis (Ditlevsen and Madsen, 2007) is presented in Equations (1) and (2). It shows scheduling of inspections before a probability of failure,  $\text{Pr}(\text{Fail})$ , (due to some active degradation mechanism) exceeds a threshold. Updating of  $\text{Pr}(\text{Fail})$  following inspections assumes that no damage is detected. Here  $d$  represents the extent of damage,  $d_C$  is the critical dimensions beyond which failure is predicted and  $H$  is the extent of the damage reported by an inspection. This calculation can be completed using methods of structural reliability, or sampling procedures.

$$\text{Pr}(\text{Fail}) = \text{Pr}(d_C \leq d), \quad (1)$$

$$\text{Pr}(d_C \leq d | H \leq 0) = \frac{\text{Pr}(d_C \leq d \cap H \leq 0)}{\text{Pr}(H \leq 0)}. \quad (2)$$

This approach does account for inspection reliability, as fewer inspections are shown to be required when a method with a higher probability of detection (PoD) is used. However, it needs to be combined with some decision rule (such as any detected damage will be immediately repaired). A more flexible approach may consider the prior likelihood of whether damage will be detected, and use this to help quantify the value of the data that each inspection will provide.

**2.2. Quantitative methods (Bayesian decision analysis)**

Inspection data,  $z$ , of various type and quality may be available to engineers, and can be related to uncertain parameters of interest,  $\theta$ , using a probabilistic model,  $f(z|\theta)$ . This can be used to update a prior model,  $\pi(\theta)$ , to obtain a posterior model,  $\pi(\theta|z)$  using Bayes’ theorem, as shown in Equation (3). Understanding the requirements for information collection requires these probabilistic models to be combined with a decision analysis (Jaynes, 2003) and this principle has been acknowledged in the context of statistical experimental design for decades (Berger, 1985). Quantification of the expected value of information (VoI) is a special application of Bayesian decision analysis. It is sometimes referred to as preposterior decision analysis (Jordaan, 2005), and sometimes as multistage decision-making under uncertainty (Gelman et al., 2014). Its application to structural integrity management has been investigated extensively in recent years (Pozzi and Der Kiureghian, 2011; Straub, 2014; Di Francesco et al., 2021).

$$\pi(\theta|z) \propto \pi(\theta) \cdot f(z|\theta). \tag{3}$$

A decision tree representation of the question of how, and to what extent, an inspection is expected to benefit risk management is shown in Figure 1. Here, available inspection options,  $E$ , can provide additional data,  $Z$  (unless the option not to inspect is selected). Following this, risk mitigation options,  $A$  (again, including the option to do nothing) are available. These decisions will influence the parameters describing structural condition,  $\Theta$ . The utility of this outcome is a function of each of these outcomes,  $u[e, a, \pi(\theta|z)]$ .

The joint optimization over the two decision nodes (shown as squares in Figure 1) is presented in Equation (4). It simply states that the optimal inspection,  $e^*$ , and risk mitigation,  $a^*$ , are those associated with the expected maximum utility (or minimum cost).

$$e^*, a^* = \arg \max_{e \in E, a \in A} u[e, a, \pi(\theta|z)]. \tag{4}$$

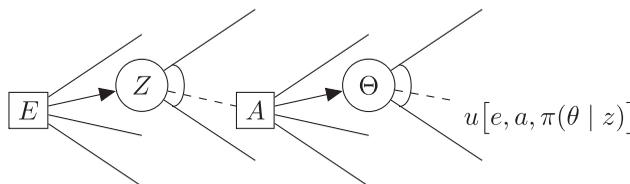
The expected value of any given inspection,  $\text{VoI}_{e_i}$ , is the difference in expected utility with,  $u[e_i, a_{e_i}^*, \pi(\theta|z)]$ , and without,  $u[a^*, \pi(\theta)]$ , the information it provides, see Equation (5). As indicated by the multiple branches in Figure 1, this calculation requires integration over various possible outcomes. For instance, a continuum of possible inspection results  $z \in Z$  can be accounted for by sampling from a prior probabilistic model and averaging the subsequent expected costs. This method has been used in Section 4.

$$\text{VoI}_{e_i} = u[e_i, a_{e_i}^*, \pi(\theta|z)] - u[a^*, \pi(\theta)]. \tag{5}$$

VoI analysis is capable of identifying risk-optimal inspection strategy for nonintuitive decision problems. In other instances, the risk-optimal solution may be self-evident to engineers, however, even in such cases a formal, reproducible analysis will make it easier to revise or audit.

**2.3. Inspection quality**

The American Petroleum Institute (API) has developed RBI guidance for petrochemical industry (API, 2008; , 2016). These documents acknowledge that the quality of an inspection determines the extent to which uncertainty can be reduced, and that this can directly influence an estimated probability of failure,



**Figure 1.** Decision tree representation of inspection evaluation.

obtained by a structural reliability analysis. Bayesian analysis is listed as a method of accounting for *Inspection Effectiveness*, but the API procedures provide a qualitative ranking method. An inspection is assigned a rank on a scale of A (representing a highly effective inspection) to E (representing an ineffective inspection). A damage factor, which is used to obtain a semi-quantitative estimate of the probability of failure, is reduced based on the inspection effectiveness rank. A similar qualitative categorization is provided in an RBI recommended practice for offshore topsides equipment (DNVGL, 2017b). Such heuristics allow for the intuitive concept of inspection quality to be accounted for, albeit in a possibly imprecise way.

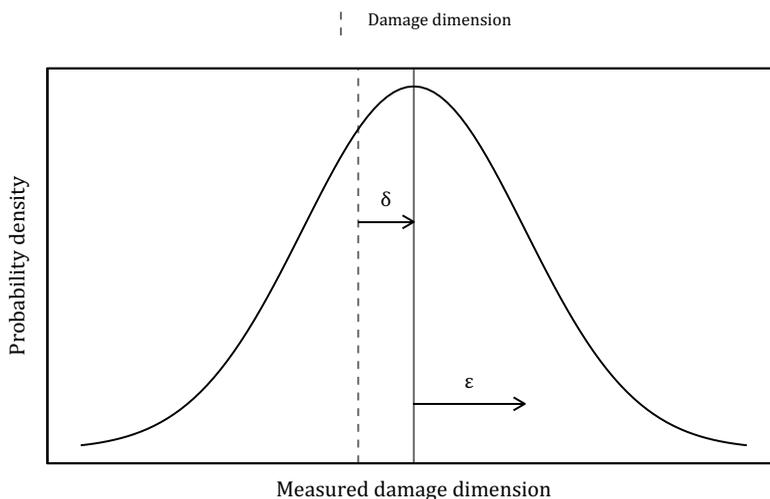
Quantifying the expected value of an inspection requires mathematical characterization of the information provided and a model relating it to the parameters of interest (Faber, 2000; Straub, 2004; Straub et al., 2006; Di Francesco et al., 2021). Assuming perfect information results in overestimation of the VoI of inspection activities. The mechanism for describing how probabilistic models are updated based on new information (such as inspection) is Bayes' theorem, as shown in Equation (3). Here, the *prior* models of parameters,  $\pi(\theta)$  are combined with a *likelihood* function,  $f(z|\theta)$  (which describes the imperfect features of the inspection data,  $z$ ), to generate an updated, *posterior* model,  $\pi(\theta|z)$ . Reducing uncertainty in probabilistic models of damage allows for better informed decisions and improved risk management.

### 3. Modeling Imperfect Inspection Data

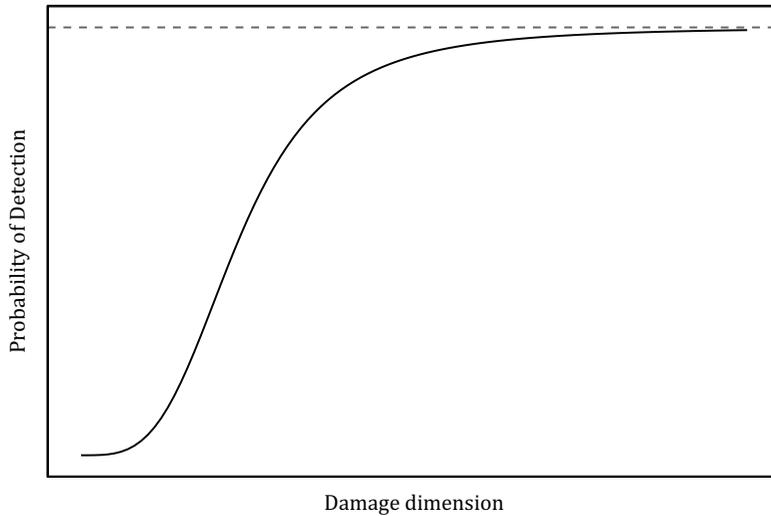
#### 3.1. Imperfect data

The value of an inspection is dependent on (among other factors) the *quality* of the data that will be collected. Data from a more precise and reliable inspection will reduce uncertainty to a greater extent. These qualities should be reflected in its expected value. A detailed review of approaches to mathematically characterize such features of an imperfect inspection is provided in Di Francesco et al. (2021). In Ali et al. (2020), an alternative categorization of imperfect information (based on that proposed in Nielsen et al. (2019)) is also presented in the context of VoI.

As identified in Di Francesco et al. (2021), the precision, bias and reliability of inspection data can be modeled as shown in Equations (6) and (7). Here, a measurement of the extent of damage,  $d_{\text{meas}}$ , is normally distributed with a mean equal to the true extent,  $d$  plus any bias,  $\delta$ , with some sizing variance,  $\epsilon$ . This model is shown graphically in Figure 2. The probability of this damage being detected by the inspection can be modeled using a logistic regression, where  $\alpha$  and  $\gamma$  represent a linear regression



**Figure 2.** Gaussian sizing accuracy model.



**Figure 3.** Logistic regression probability of detection (PoD) model.

co-efficient and intercept, respectively, on the log-odds scale (Health and Safety Executive, 2006). Such a model is presented in Figure 3.

$$d_{\text{meas}} \sim N(d + \delta, \varepsilon), \quad (6)$$

$$\text{PoD} = \frac{\exp - (\alpha + \gamma \cdot \ln(d))}{1 + \exp - (\alpha + \gamma \cdot \ln(d))}. \quad (7)$$

### 3.2. Incomplete (missing) data

Engineers may often be required to work with incomplete datasets, which is less frequently accounted for in existing VoI scientific literature. In Qin et al. (2015), cases of an inspection not detecting damage are considered to be missing data. In Di Francesco et al. (2021), such instances are treated as information in a Bayesian updating procedure. Rather, missing data represents the absence of information and some examples of this are listed below:

- **Operational constraints:** Some part of an inspection not completed, perhaps due to time constraints, or unsafe weather conditions.
- **Inspection technology not operating:** Excursions from normal operating conditions, or mechanical damage, could cause electronic components to stop functioning.
- **Lost data:** An analysis may be considering historical inspection and maintenance data. If these data were recorded before a data management system, or when it was operated by a previous asset owner, then it may be at least partially lost.

The standard approach for analysis of incomplete datasets is to disregard cases with one or more missing data points. This is known as a *complete case analysis*, and is only appropriate when the reduced dataset has not been biased by the removal of incomplete cases. An alternative approach is *imputation of missing data*, which uses available information to estimate the missing values.

As outlined in Gelman et al. (2020) and McElreath (2020), whether the data are missing at random will determine how it should be imputed. For instance, if there is a causal link between the value of the data and its missingness status then this should be accounted for in a model.

The Bayesian framework is well suited to probabilistic imputation of missing data. Missing values can be inferred in the same way as other (unobserved) model parameters such that probabilistic estimates are identified, which are consistent with the other data, in the context of the specified model.

### 3.3. Multilevel modeling approach to quantifying system effects

In Straub and Faber (2004), the authors suggest that the condition of multiple, so-called *hot spots*,<sup>1</sup> on a structure should somehow be related. In Straub and Faber (2005), it is proposed that this can be achieved through a co-variance matrix that describes common factors between locations. In practice, some of the underlying factors for these dependencies may not be intuitive, and even for those that are identified it is unlikely that the specific data required will be available for estimating correlation co-efficients or copula parameters. The solution proposed in this paper is to use multilevel Bayesian models Gelman (2006) to partially pool information between locations.

When fitting a conventional Bayesian model, a set of parameters,  $\theta$ , are estimated based on some measurements,  $z$  (described using a likelihood function), and prior models,  $\pi(\theta)$ . Often, engineers must incorporate data from different sources. In this context, information from an inspection at one location can help inform models of structural condition at another location. When differing locations are modeled as a discrete heterogeneity in inspection data, Bayesian multilevel models can be used to estimate degradation rates, as detailed in this Section.

When combining data from multiple sources, engineers may choose to create a single (fully pooled) population and use this to fit a model. This approach neglects any variation between the datasets, and failing to account for this will introduce bias to any predictions. For instance, if the condition of two hot spots differs significantly, then a model which fully pools data from inspections at these locations will result in some average, which may not accurately predict the condition of either. This can be addressed by fitting separate models to each dataset, but could introduce excessive variance to the predictions, since the models are now being estimated based on less evidence. The optimal solution will be somewhere in between these two extremes of dependency and can be realized by fitting a Bayesian multilevel model (Gelman, 2006; Gelman and Hill, 2007).

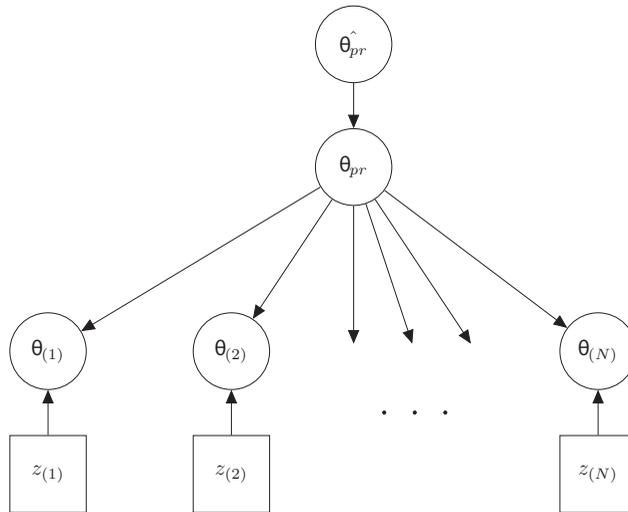
A graphical representation of a partial pooling Bayesian model is shown in Figure 4. Here, a separate set of parameters are estimated for each group, but population level priors (which control the extent of pooling for each parameter) are also estimated from the data (and hyperpriors). Further discussion on the extent of pooling from a multilevel model is presented in Section 4.

As demonstrated in Di Francesco et al. (2020), Bayesian multilevel models can be used to improve estimates of structural condition at locations remote to where an inspection has been completed. If there is some commonality between the condition of a location being inspected and the condition of another location of interest, then engineers can learn indirectly something about the latter, by inspecting the former. This has the potential to increase the value of the inspection activity, since the same data is now being used to update models at additional locations.

In a different setting, this approach has also been demonstrated for combining evidence from different test datasets to estimate the parameters of a fatigue crack growth model (Di Francesco et al., 2020). The benefit of allowing for possible variability between tests (which were completed in different labs, on different steels) was quantified using out of sample predictive performance criteria.

The Bayesian models presented in this work have been fit using Markov Chain Monte Carlo (MCMC) sampling. The suitability of Hamiltonian Monte Carlo (a modern MCMC algorithm) as a procedure for sampling from multilevel models is demonstrated in Betancourt and Girolami (2015). The calculations presented in this paper make use of the no u-turn implementation of Hamiltonian Monte Carlo (Hoffman and Gelman, 2014) in the probabilistic programming language, Stan (Carpenter et al., 2017).

<sup>1</sup> Note that in the cited work a hot spot is considered to be high risk locations of active degradation.



**Figure 4.** Structure of a multilevel (partial pooling) Bayesian model for estimating parameters,  $\theta$ , from data,  $z$ , priors,  $\theta_{pr}$ , and hyperpriors,  $\hat{\theta}_{pr}$ .

#### 4. Example Application: Inspecting for Corrosion Damage

##### 4.1. Problem overview

Corrosion is considered as the degradation mechanism in the example in this section, primarily because new data from inspections are typically used to update estimates of degradation rate (Health and Safety Executive, 2002). Estimation of corrosion rates is a statistical challenge. Unlike traditional alternatives, Bayesian methods can be used to obtain estimates from small datasets, accounting for imperfect features of the data available. A simple model of corrosion rate,  $\beta$ , using the inspection data,  $d_i$  is proposed in Equations (8) and (9).

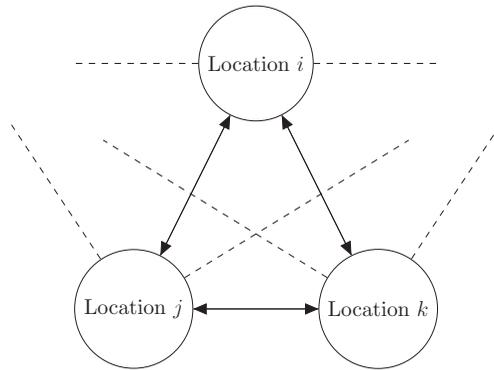
$$\beta_i = \frac{d_{meas}^{t=T_2} - d_{meas}^{t=T_1}}{T_2 - T_1}, \tag{8}$$

$$\beta \sim N(\mu_\beta, \sigma_\beta). \tag{9}$$

The true corrosion rate (in mm year<sup>-1</sup>) has been assigned as being normally distributed with a mean value of 1 and a standard deviation of 0.2. In the case of the first inspection measurements are unbiased ( $\delta = 0$ ) with sizing accuracy parameter,  $\varepsilon = 0.5$  (see Equations (6) and (7)). The second inspection is more precise, with  $\delta = 0$  and  $\varepsilon = 0.3$ .

The challenge of corrosion rate estimation is extended to consider the additional challenge multiple locations. Thirteen separate sites of corrosion are considered, with identifiers 1–13. Sites 1–10 were measured at location A and sites 11–13 were measured at location B. There may be reasons to expect differing corrosion rates between remote locations, for example, due to variation in the local environment and the performance of any corrosion protection systems. Consequently, combining these data into a single population will result in biased predictions and using separate models (and separate data) will result in additional variance. This will be especially problematic if at least one location has only very limited evidence. A partial pooling model will allow for information to be shared so that measurements can inform predictions at different locations. This principle this could be extended to many more locations, as proposed in Di Francesco et al. (2020) and illustrated in Figure 5.

The corrosion site with identifier 11 (at location B) was assumed not to have been measured during the second inspection. Therefore, a probabilistic estimate of the extent of the damage at this time has been imputed, as described in Section 3.



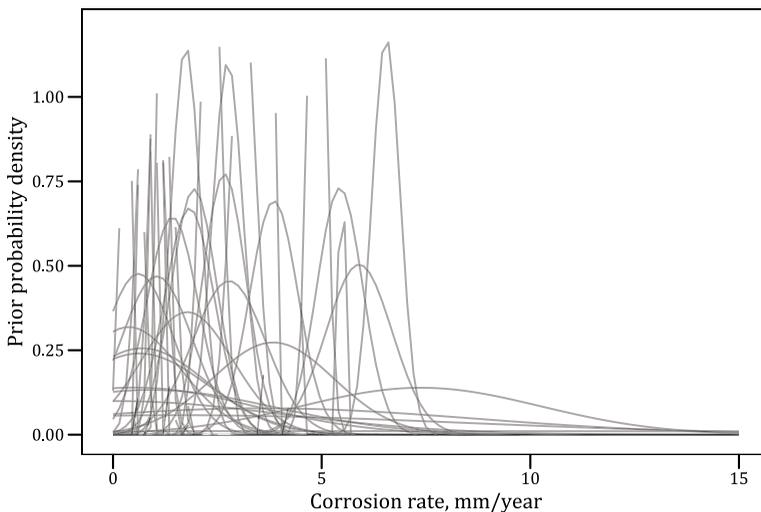
**Figure 5.** Representation of dependencies between inspected locations (*i*) and non-inspected locations ( $\bar{i}$ ) in a partial pooling Bayesian model.

**4.2. Prior predictive simulation**

Sampling from the prior distributions of a statistical model allows for a graphical representation of the information they contain on a meaningful outcome scale. This approach (known as prior predictive simulation) is advocated for in Gelman et al. (2017), Gabry et al. (2019), and McElreath (2020). This is especially important when parameters undergo some transformation, such as in generalized linear models, whereas the normal likelihood model for corrosion rate is comparatively simple. Nevertheless, prior predictive simulation for the model described by Equations (8)–(11) is shown in Figure 6. This demonstrates that a large range of corrosion rates are considered credible (in advance of conditioning on any inspection data). They are broadly constrained to be of the order of a few mm year<sup>-1</sup>, and corrosion specialists may be able to provide additional information, when the priors are presented visually in this way.

$$\mu_\beta \sim N\left(\frac{1}{2}, 3\right), \tag{10}$$

$$\sigma_\beta \sim \exp\left(\frac{1}{3}\right). \tag{11}$$



**Figure 6.** Prior predictive simulation of corrosion rate.

### 4.3. Value of information calculation procedure

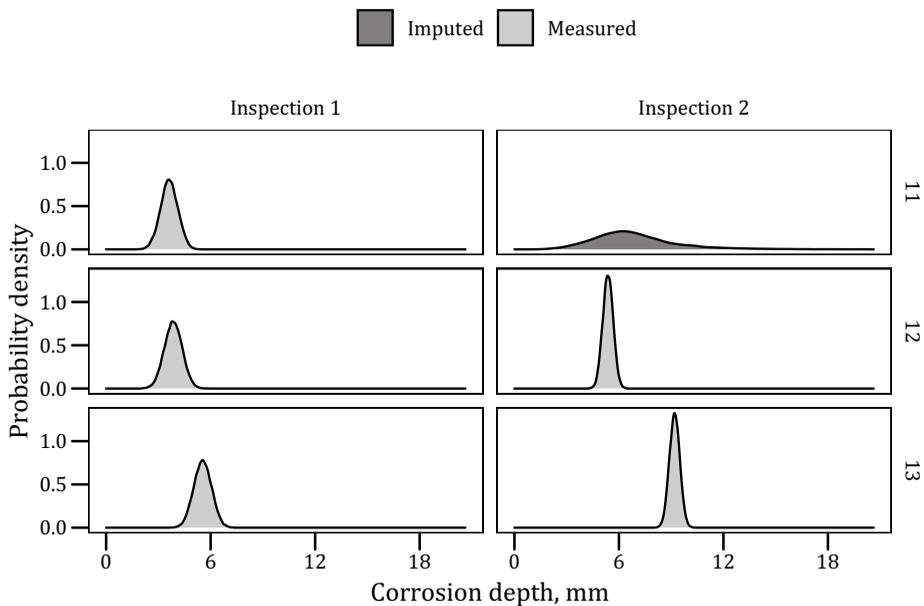
Figure 7 shows the measurements from location *B*, including the missing data which has been imputed using an independent (no pooling) model. Figure 8 shows the data from all sites, including the missing data, which has been imputed using a multilevel (partial pooling) model.

Figure 9 shows the 1,000 samples from the imputed depths at site 11 that have been selected for the VoI analysis. The samples (shown as histograms) are compared with the MCMC samples from the posterior distributions (shown as density plots). These samples represent possible measurements from the proposed inspection of site 11.

In Figure 10, the predictions of corrosion at site 11 from both the independent and partial pooling model are compared. On this scale, it is evident that the uncertainty in this parameter has been reduced by leveraging the information from location *A*. However, in both models, it has not been quantified as precisely as the sites which were measured (as shown in Figures 7 and 8). A VoI analysis can be used to quantify the expected value of completing the inspection at the location of the missing data.

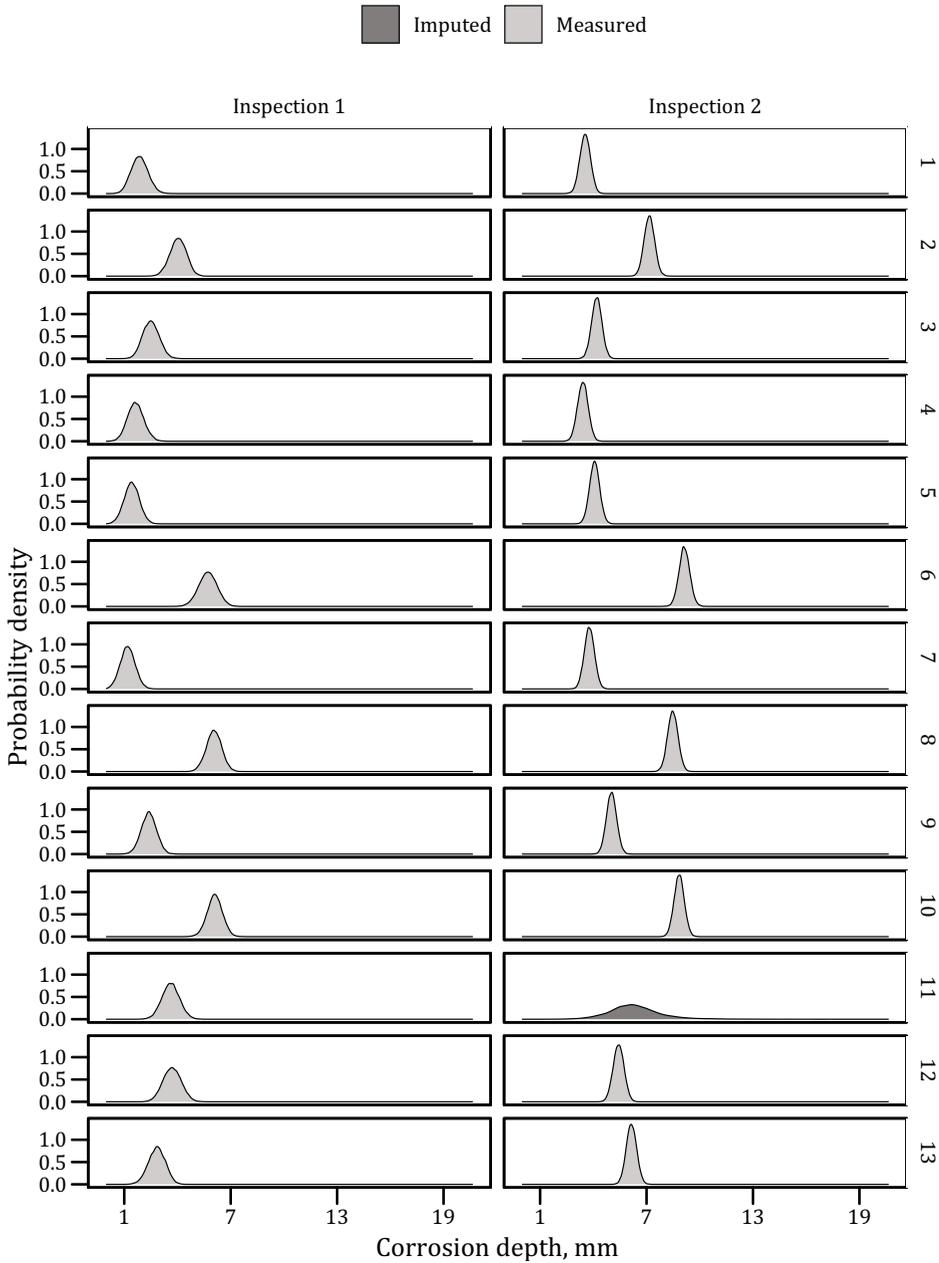
As shown in Equation (5), the VoI is the difference between the utility with (posterior expected cost) and without (prior expected cost) the inspection. The expected costs are calculated based on the below decision problem.

All sites of corrosion damage in this simulated example are considered to be axially oriented in a 20-inch, schedule 60-Grade *B* (ASME International and American Petroleum Institute, 2012) linepipe. The failure pressure has been taken to be 38.6 MPa. This results in a hoop stress just below the specified minimum yield strength and could represent an overpressure event during normal operation. Considering only loads from internal pressure (and assuming a constant aspect ratio of  $10^2$ ), the modified ASME B31G assessment method (ASME International, 2012) has been used to predict failure. Similar procedures are provided in BSI (2019) and DNVGL (2017a). The failure stress,  $S_{\text{fail}}$ , is related to the corrosion depth,  $d$ ,



**Figure 7.** Bayesian estimate of missing data using independent models.

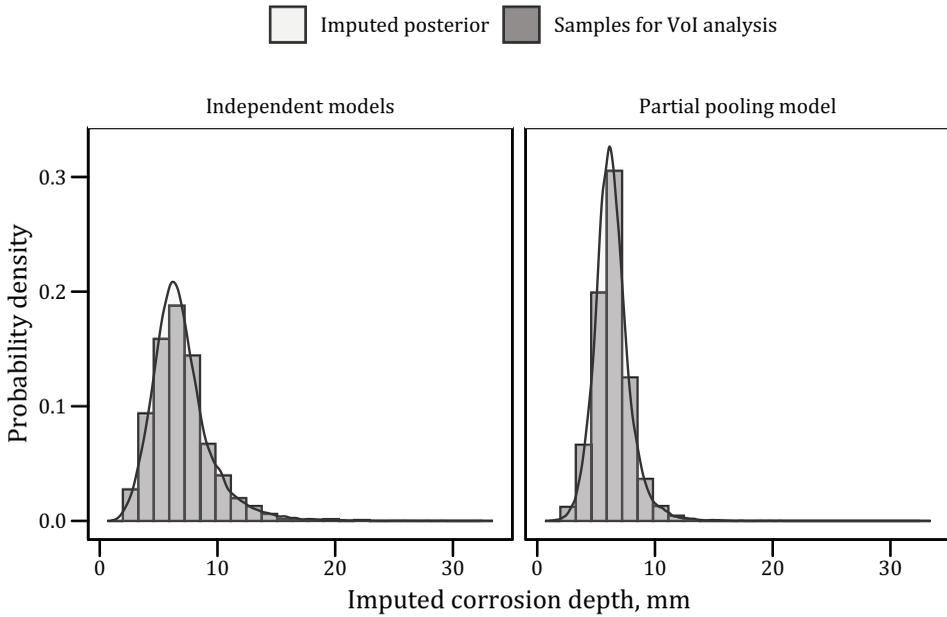
<sup>2</sup> In practice the length of such corrosion could also be measured by inspection and an analogous model could even be used to model the corrosion rate in this dimension. However, to avoid repetition, a constant aspect ratio (corrosion length divided by corrosion depth) has been assumed in this example.



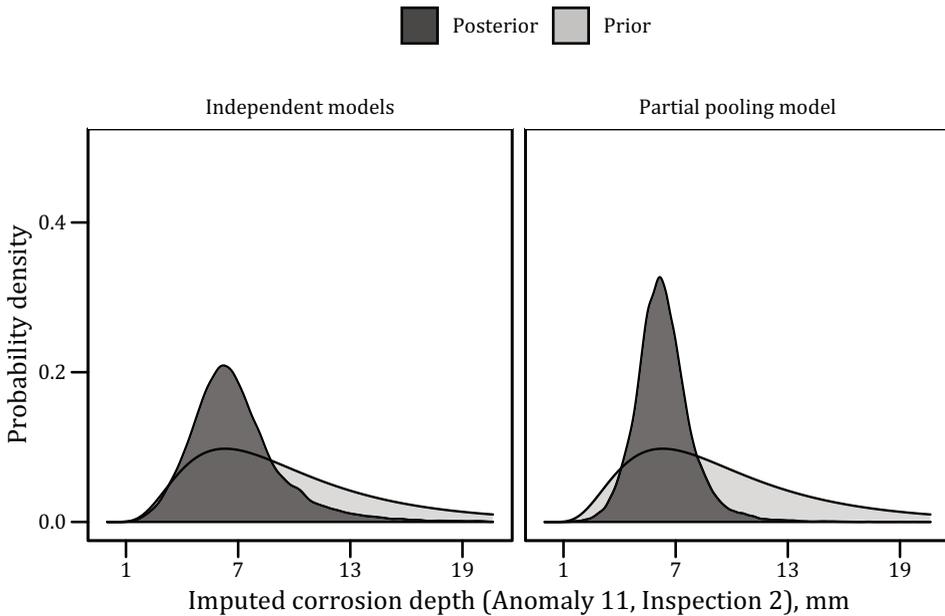
**Figure 8.** Bayesian estimate of missing data using multilevel (partial pooling) model.

axial extent,  $l$ , normalized axial extent,  $l_z$ , and pipeline (outer) diameter,  $D$ , as shown in Equation (12). Here, the flow strength,  $S_{\text{flow}}$  is equal to the mean of the yield and tensile strength and  $M$  is defined in Equations (13) and (14).

$$S_{\text{fail}} = S_{\text{flow}} \cdot \left[ \frac{1 - 0.85 \cdot \frac{d}{t}}{1 - 0.85 \cdot \frac{d}{t} \cdot M^{-1}} \right], \tag{12}$$



**Figure 9.** Samples from the Bayesian imputation models for the missing data site.



**Figure 10.** Comparison of Bayesian estimates of missing data between two model structures.

$$M = \begin{cases} \sqrt{1 + 0.6275 \cdot l_z - .003375 \cdot l_z^2}, & \text{for } l_z \leq 50, \\ 0.032 \cdot l_z + 3.3, & \text{for } l_z > 50, \end{cases} \tag{13}$$

$$l_z = \frac{l^2}{D \cdot t}. \tag{14}$$

Following the second inspection (and the identification of the apparently high corrosion rates, a repair plan is required for the next year. Prior probabilities of failure for each site are presented in Figure 11. The calculation outlined in this section details how the expected value of completing the inspection at site 11 can be calculated for both independent and partial pooling Bayesian models. As detailed in Section 1, this work does not include detailed cost modeling. For the purposes of this example, normalized values have been assumed, where the cost of each repair is 1.0 and the cost of each failure is 10.0.

Differences in the probabilities of failure between the two model structures are primarily due to the differing corrosion rates that have been estimated. As shown in Figure 12, the corrosion rate can be estimated more precisely (with less variance in the probabilistic estimate) at location A, and this is because more inspection data are available than at location B. However, both locations can be estimated more precisely by the partial pooling model. The reduced probability mass at the tails in the predictions from the partial pooling model results in extreme corrosion rates being considered less likely. This includes very low corrosion rates, which would have reduced the estimated probability of failure of large damage, and very high corrosion rates, which would have increased the estimated probability of failure of small damage. This effect is evident in Figure 11, and is larger at the sites at location B, where there is a greater improvement in the corrosion rate predicted by the two models.

As discussed in Section 3, the extent of pooling between locations is not arbitrary. Rather it is determined from the variance parameters of the (population level) priors for the parameters that are to be partially pooled. An example of the posterior distribution of the mean corrosion rate is presented in

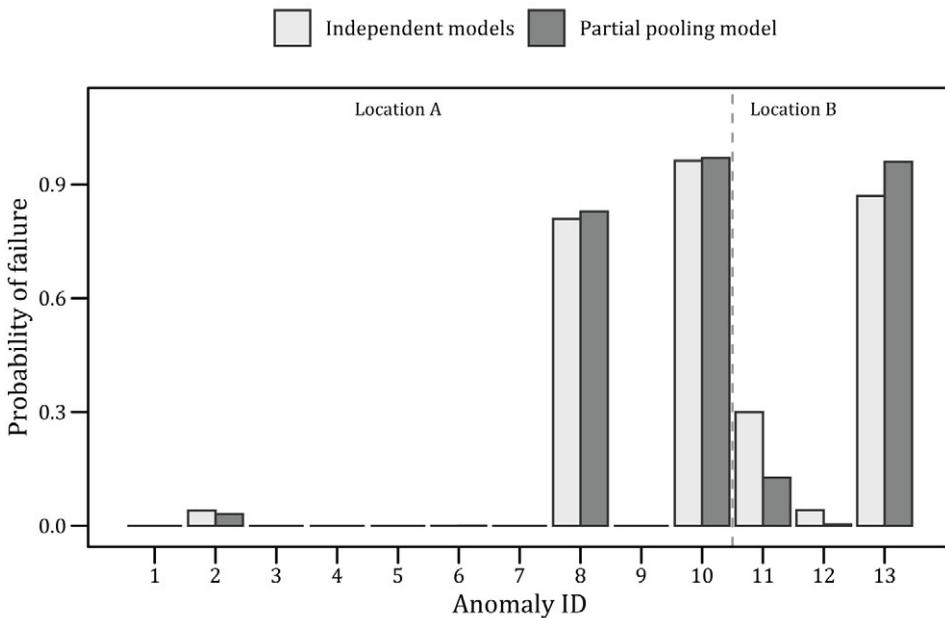
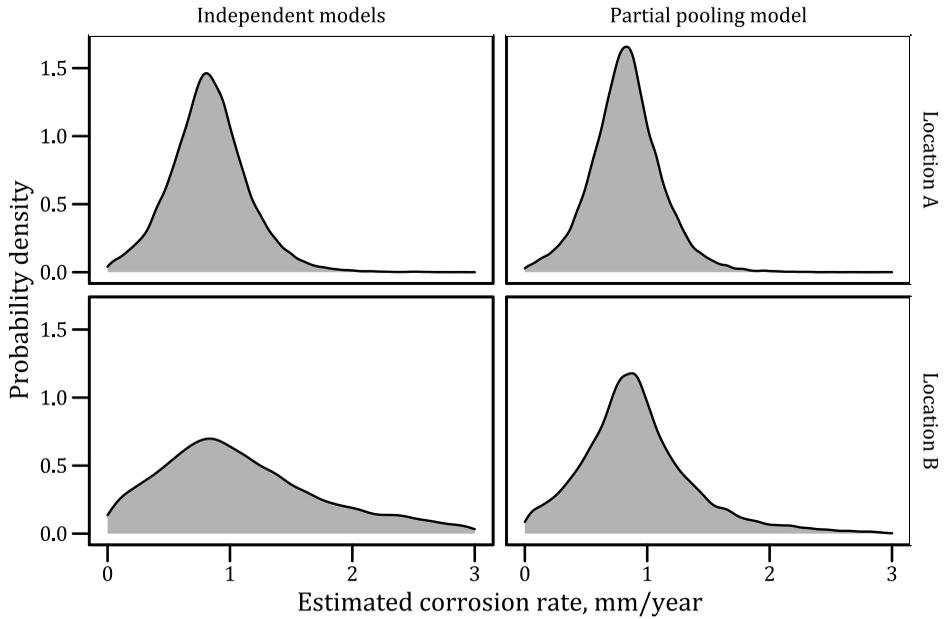
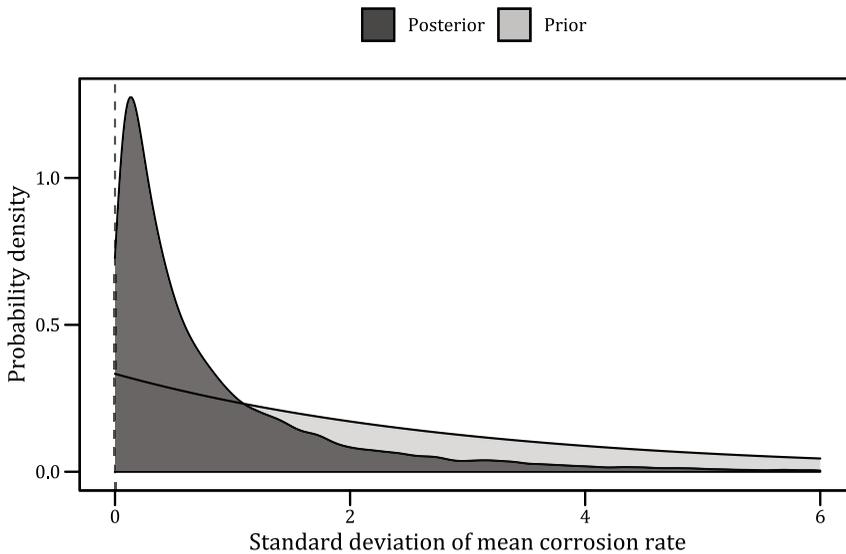


Figure 11. Probability of failure for each corrosion site from both model structures.



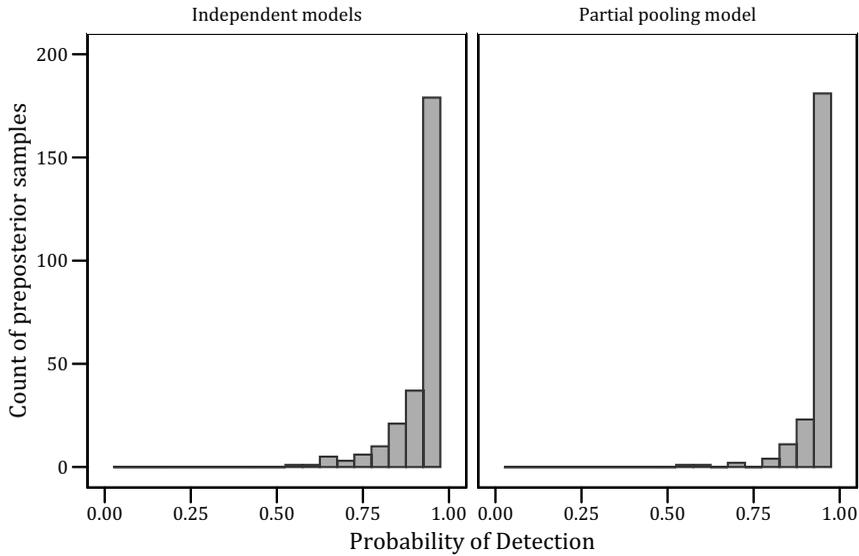
**Figure 12.** Posterior distribution of corrosion rates from each model structure.



**Figure 13.** Posterior distribution of parameter controlling pooling of the mean corrosion rate.

Figure 13. Here, values approaching 0 would indicate very little variation between the locations and would therefore justify a greater extent of pooling of information between them. Conversely, very large values would be associated with models that exhibit very little pooling.

In this example, the two locations have been assumed to share the same corrosion rate, and this have been identified by the model as shown by the posterior distribution in Figure 13. A consequence of this



**Figure 14.** Histogram of simulated probability of detection (PoD) for each of the samples.

assumption is that the model is expected to identify evidence of commonality, and partially pool estimates of corrosion rates. This model structure is equally valid for examples when “true” corrosion rates differ between locations, although this will result in less pooling.

## 5. Results

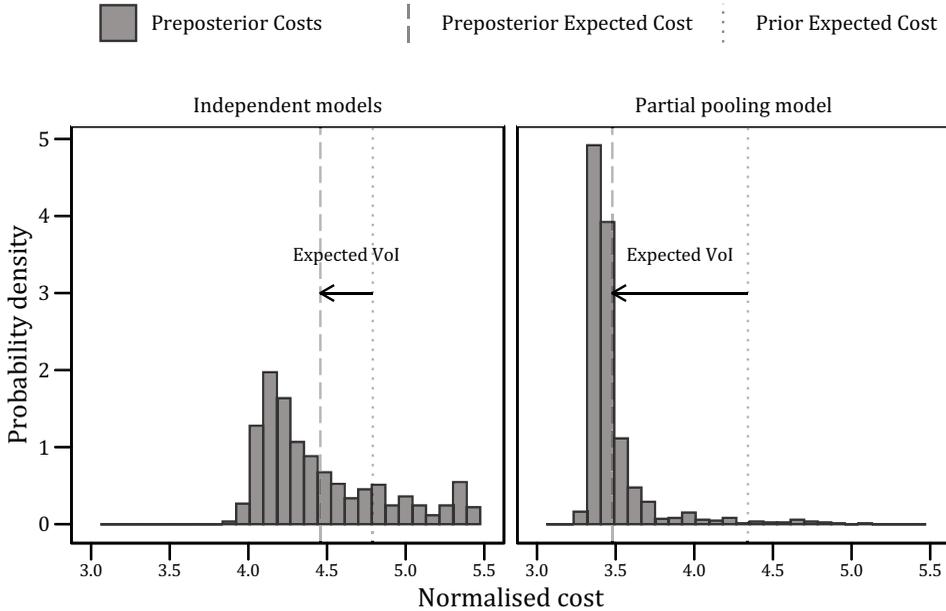
For each sample in Figure 9, the Bayesian models are re-evaluated and a decision analysis is completed. A PoD was estimated, based on the PoD curve defined in Equation (7), with parameter values  $\alpha = -3$  and  $\gamma = 5$ . As shown in Figure 14, these probabilities may appear to be relatively high, but recall that earlier, reduced damage was previously identified at these locations. In cases where the damage was not expected to be detected, reduced corrosion rates were estimated and propagated through the decision analysis.

The subsequent *preposterior costs* are shown in Figure 15, which also includes the difference between the expected value of the prior and preposterior costs (labeled as the expected VoI). These results are also shown in Table 1 as proportions of the cost of a repair. As stated in Section 4.1, the cost of each failure is assumed to be 10 times the cost of each repair, with the repair given a unit cost. Thus, the expected VoI of the additional inspection is estimated to be equal to 33% of the repair cost (for independent models) and circa 86% of the same cost (for the multilevel model).

## 6. Discussion

The additional expected value of partially pooled inspection data (shown in Table 1 and Figure 15) is attributed to the wider use of the data (in the context of the multilevel corrosion rate model). However, it is also acknowledged that precise values cannot be generalized since VoI analysis is case-dependent and will vary with new utility functions and structural reliability assessments.

An additional novel feature of the VoI analysis, is that the prior damage model (which is used to generate hypothetical inspection data) has been obtained using a Bayesian imputation model. That is to say, it is a probabilistic model, consistent with available inspection data (of some known precision and reliability) at all other locations.



**Figure 15.** Comparison of expected value of inspection between two model structures.

**Table 1.** Results from the value of information (VoI) analysis.

Decision analysis	Independent models	Multilevel model
Expected prior cost	4.791	4.341
Expected preposterior cost	4.458	3.479
Expected VoI	0.333	0.862

The simulated corrosion rate did not vary between locations, so the model was expected to identify evidence of similarity and therefore pool estimates, as was shown to be the case in Figures 12 and 13. Further work could be completed to investigate the extent of pooling when there is more variation between locations. When the variation extends to the point that a single underlying model is no longer considered appropriate for all locations, multilevel models are no longer suitable. A full probabilistic analysis of such data could provide quantitative justification for separate models, but this finding could be hidden if outliers are removed, or any other *cherry picking* of data are completed.

## 7. Computational Considerations

### 7.1. Fitting multilevel Bayesian models

Transitioning from conventional Bayesian models to a multilevel structure increases the number of parameters to be estimated significantly. Each parameter being partially pooled is now estimated for each exchangeable group, and a prior is also now inferred from the data. Navigating a higher dimensional probability space will be more challenging for any algorithm and so the efficiency of these models suffer. When using partial pooling models to account for system effects in VoI analysis, these issues are

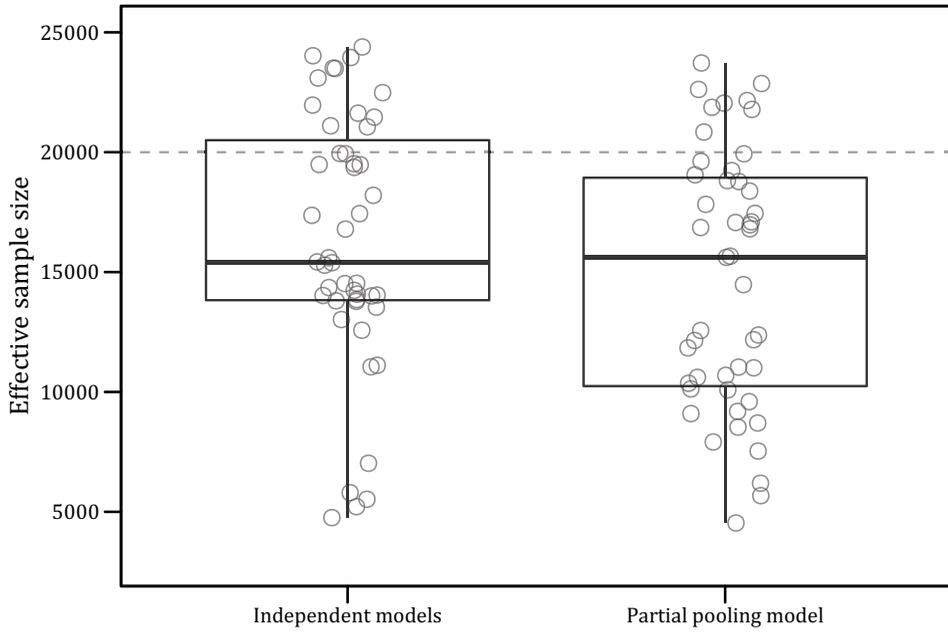


Figure 16. Effective sample size of parameters from each model structure.

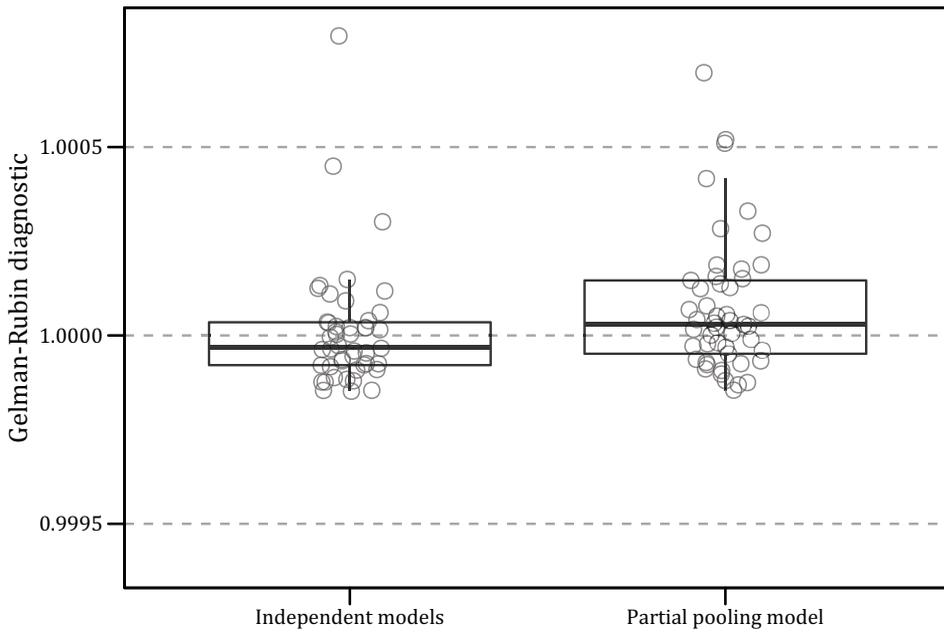


Figure 17.  $\hat{R}$  for parameters from each model structure.

compounded with the existing challenges of computational preposterior decision analysis, described in Straub (2014), Pozzi and Der Kiureghian (2011), and Di Francesco et al. (2021).

## 7.2. Model diagnostics

A comparison of quantitative MCMC diagnostics are shown for each of the models in the example calculation in Section 3. The effective sample size,  $n_{\text{eff}}$  and the Gelman–Rubin auto-correlation measure,  $\hat{R}$  (Gelman et al., 2014) are shown in Figures 16 and 17, respectively for each model structure. In Figures 16 and 17, each data point represents an unobserved parameter that the model has estimated. In Figure 18, each data point represents a separate Markov chain.

Although there is no substantial evidence of a lack of convergence, excessive auto-correlation or insufficient length of any of the Markov chains, it is noteworthy that the parameters with the lowest value of  $n_{\text{eff}}$  and the value of  $\hat{R}$  that deviates most from 1.0 are both from the partial pooling model.

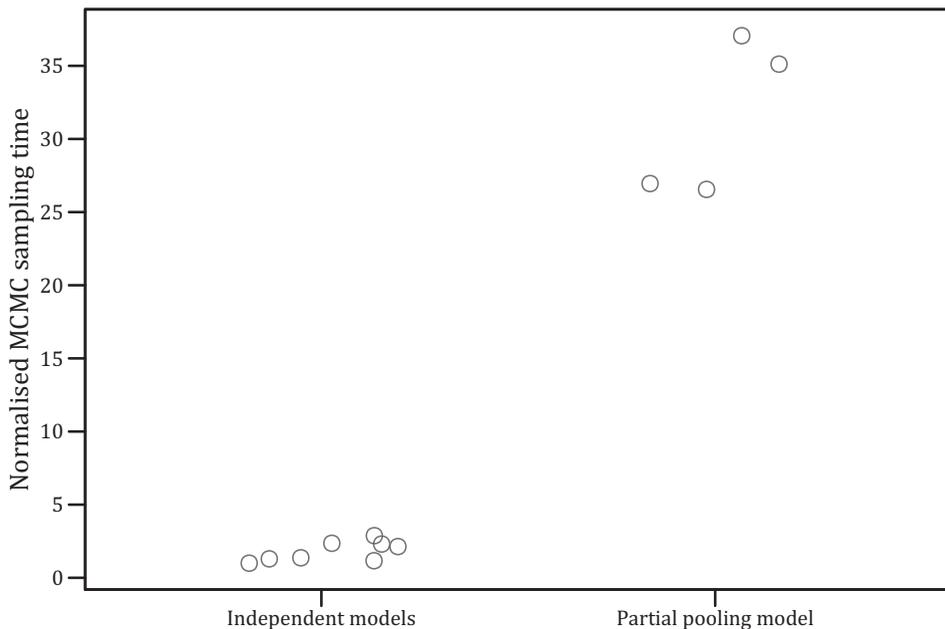
In Figure 18, the time taken to obtain samples from the partial pooling model are shown to be (approximately) between 10 and 40 times longer in this analysis. Additional time and/or computational resources will be required when working with multilevel models.

## 8. Conclusions

Bayesian multilevel models are proposed as a solution for probabilistic modeling when only limited data are available. The Bayesian framework is sufficiently flexible to interchange missing data with model parameters, and multilevel (partial pooling) models introduce dependencies between discrete groups.

These models are also compatible with quantitative inspection planning (VoI analysis), although they are associated with greater computational requirements.

Bayesian imputation of missing data has been used to generate prior models for the VoI calculation, which are consistent with all other sources of information (locations where inspection data is available) in the context of the model. An additional possible application of this method is in estimation of the expected



**Figure 18.** Run-times for each model structure.

value of inspections that were not completed. Although it is not possible to act upon these calculations, since the inspection opportunity has passed, it could be used to inform policy and further demonstrate the utility of statistical decision analysis in industry. In such cases, the model could either make use of current information or only use what would have been available at the time. The selected approach would have implications for the meaning of the result.

An example calculation was presented that demonstrated a procedure for partial pooling inspection data, to estimate corrosion rates at two distinct locations. The expected VoI was found to be higher when estimated using a partial pooling than when using independent (no pooling) models. In the partial pooling model, the proposed measurement would also have informed predictions at the other location.

Whilst it is difficult to extrapolate specific conclusions about these effects (since outcomes from the decision analysis are heavily context dependent), this work has shown how multilevel Bayesian models can be applied in the field of VoI.

## Nomenclature

Parameter	Meaning	Units
$a$	risk mitigation action	—
$a^*$	expected optimal risk mitigation action	—
$A$	available risk mitigation actions	—
$e$	prospective inspection	—
$e^*$	expected optimal inspection	—
$E$	available inspection options	—
$d$	corrosion depth	mm
$d_C$	critical corrosion depth	mm
$d_{\text{meas}}$	measured corrosion depth	mm
$D$	outer diameter	mm
$H$	extent of damage as identified from an inspection	mm
$l$	axial extent of corrosion	mm
$l_z$	normalized axial extent of corrosion	mm
$M$	empirical bulging factor	—
$n_{\text{eff}}$	effective sample size from Markov chain	—
$\text{Pr}(\text{Fail})$	probability of failure	—
$\hat{R}$	Gelman–Rubin diagnostic parameter	—
$S_{\text{fail}}$	failure stress	MPa
$S_{\text{flow}}$	flow strength	MPa
$u$	normalized utility	—
$t$	thickness	mm
$T$	time	years
$z$	inspection data	—
$Z$	domain of possible inspection outcomes	—
$\alpha$	probability of detection model parameter (log-odds intercept)	—
$\beta$	corrosion rate	mm year <sup>-1</sup>
$\gamma$	probability of detection model parameter (log-odds gradient)	—
$\delta$	inspection bias	mm
$\varepsilon$	inspection sizing uncertainty	mm
$\theta$	generic uncertain parameter(s)	—
$\theta_{\text{pr}}$	prior model parameter(s)	—
$\hat{\theta}_{\text{pr}}$	hyperprior model parameter(s)	—
$f(z \theta)$	likelihood function for data, $z$	—
$\mu_{\beta}$	mean corrosion rate	mm year <sup>-1</sup>
$\pi(\theta)$	prior distribution of parameter(s), $\theta$	—
$\pi(\theta z)$	posterior distribution of parameter(s), $\theta$	—
$\sigma_{\beta}$	standard deviation of corrosion rate	mm year <sup>-1</sup>

**Acknowledgments.** The work was enabled through, and undertaken at, the National Structural Integrity Research Centre (NSIRC), a postgraduate engineering facility for industry-led research into structural integrity established and managed by TWI through a network of both national and international Universities.

**Data Availability Statement.** The Bayesian partial pooling and imputation model has been made available as a Stan program: <http://dx.doi.org/10.17632/93kknv5dz5.1>.

**Author Contributions.** Conceptualization: D.D.F.; Software: D.D.F.; Methodology: D.D.F., M.C., and M.H.F.; Formal analysis: D.D.F.; Validation: D.D.F.; Visualization: D.D.F.; Supervision: M.C., M.H.F., and U.B.; Project administration: U.B.; Writing – original draft: D.D.F.; Writing – review & editing: D.D.F., M.C., and M.H.F.

**Funding Statement.** This publication was made possible by the sponsorship and support of Lloyd’s Register Foundation and the Engineering and Physical Sciences Research Council (EPSRC). Lloyd’s Register Foundation helps to protect life and property by supporting engineering-related education, public engagement and the application of research.

**Competing Interests.** The author declares no competing interests exist.

## References

- Ali K, Qin J and Faber M (2020) On information modeling in structural integrity management. *Structural Health Monitoring* <https://doi.org/10.1177/1475921720968292>
- American Petroleum Institute (2008) *Risk based inspection technology. API RP 581.*
- American Petroleum Institute (2016) *Risk-based inspection. API RP 580 third edition.*
- Ang A and Tang W (2007) *Probability Concepts in Engineering*, 3rd ed. Hoboken, NJ: Wiley.
- ASME International (2012) *Manual for determining the remaining strength of corroded pipelines. ASME B31G forty-fifth edition.*
- ASME International, American Petroleum Institute (2012) *Specification for line pipe. API Specification 5L forty-fifth edition.*
- Berger JO (1985) *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Berlin: Springer.
- Betancourt M and Girolami M (2015) Hamiltonian Monte Carlo for hierarchical models. *Current Trends in Bayesian Methodology with Applications Chapter 4*, 79–103. <https://doi.org/10.1201/b18502>
- BSI (2019) *Guide to methods for assessing the acceptability of flaws in metallic structures. BS 7910 fourth edition.*
- Bureau Veritas (2017) *Risk-based structural integrity management of offshore jacket structures. Guidance Note NI 624 DT R00 E.*
- Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker MA, Guo J, Li P, and Riddell A (2017) Stan: a probabilistic programming language. *Statistical Software* 76 1–32. <https://doi.org/10.18637/jss.v076.i01>
- Clarke I (2016) The pitfalls of risk based inspection: an insurer’s view of the Chevron Richmond and Tesoro Anacortes incidents. *Chemical Engineering Transactions* 48, 757–762. <https://doi.org/10.3303/CET1648127>
- Di Francesco D, Chryssanthopoulos M, Havbro Faber M, and Bharadwaj U (2020) Bayesian multi-level modelling for improved prediction of corrosion growth rate. *Proceedings of International Conference on Offshore Mechanics and Arctic Engineering Volume 2B: Structures, Safety, and Reliability*. <https://doi.org/10.1115/OMAE2020-18744>
- Di Francesco D, Chryssanthopoulos M, Havbro Faber M, and Bharadwaj U (2020) Consistent and coherent treatment of uncertainties and dependencies in fatigue crack growth calculations using multi-level Bayesian models. *Reliability Engineering and System Safety* 204C. <https://doi.org/10.1016/j.res.2020.107117>
- Di Francesco D, Chryssanthopoulos M, Havbro Faber M, and Bharadwaj U (2021) Evaluation of inspection features including exposure risk using a value of information analysis. *Civil Engineering and Environmental Systems* 38:1, 36–58. <http://dx.doi.org/10.1080/10286608.2021.1887154>
- Ditlevsen O and Madsen HO (2007) *Structural Reliability Methods*. Hoboken, NJ: John Wiley & Sons Ltd. <http://od-website.dk/books/OD-HOM-StrucRelMeth-Ed2.3.7.pdf>. (accessed 02 July 2018).
- DNVGL (2015) *Probabilistic methods for planning of inspection for fatigue cracks in offshore structures. DNVGL-RP-C210* <https://www.dnvgl.com/rules-standards/>. (accessed 20 September 2018).
- DNVGL (2017a) *Integrity management of submarine pipeline systems. DNVGL-RP-F116* <https://www.dnvgl.com/rules-standards/>. (accessed 35 May 2019).
- DNVGL (2017b) *Risk based inspection of offshore topsides static mechanical equipment. DNVGL-RP-F116* <https://www.dnvgl.com/rules-standards/>. (accessed 05 September 2018).
- DNVGL (2017c) *DNVGL-RP-F101: corroded pipelines. DNVGL-RP-F101* <https://www.dnvgl.com/rules-standards/>. (accessed 13 November 2018).
- Faber MH (2000) Reliability based assessment of existing structures *Progress in Structural Engineering and Materials* 2, 247–253.
- Faber MH (2012) *Statistics and Probability Theory in Pursuit of Engineering Decision Support*, vol. 18. Berlin: Springer. <https://doi.org/10.1007/978-94-007-4056-3>

- Gabry J, Simpson D, Vehtari A, Betancourt M and Gelman A** (2019) Visualization in Bayesian workflow. *Royal Statistical Society. Series A: Statistics in Society* 182, 389–402. <https://doi.org/10.1111/rssa.12378>
- Gelman A** (2006) Multilevel (hierarchical) modeling: what it can and cannot do. *Technometrics* 48, 432–435. <https://doi.org/10.1198/004017005000000661>
- Gelman A, Carlin JBB, Stern HSS and Rubin DBB** (2014) *Bayesian Data Analysis*, 3rd ed. London: Chapman & Hall.
- Gelman A and Hill J** (2007) *Data Analysis Using Regression and Multi-level/Hierarchical Models*, 1st ed. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>
- Gelman A, Hill J and Vehtari A** (2020) *Regression and Other Stories*, 1st ed. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781139161879>
- Gelman A, Simpson D and Betancourt M** (2017) The prior can often only be understood in the context of the likelihood. *Entropy* 19, 1–13. <https://doi.org/10.3390/e19100555>
- Guédé F** (2018) Risk-based structural integrity management for offshore jacket platforms. *Marine Structures* 58, 223–241. <https://doi.org/10.1016/j.marstruc.2018.04.004>
- Health and Safety Executive** (2002) Guidelines for use of statistics for analysis of sample inspection of corrosion. <https://www.hse.gov.uk/research/rpddf/rr016.pdf>. (accessed 27 December 2019).
- Health and Safety Executive** (2006) Research Report 454 Probability of Detection (PoD) curves Derivation, applications and limitations. <http://www.hse.gov.uk/research/rpddf/rr454.pdf>. (accessed 19 November 2019).
- Health and Safety Executive** (2009) Structural Integrity Management Framework for Fixed Jacket Structures. <http://www.hse.gov.uk/research/rpddf/rr684.pdf>. (accessed 26 March 2019).
- Hoffman MD and Gelman A** (2014) The no-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Machine Learning Research* 15, 1593–1623. <https://dl.acm.org/doi/10.5555/2627435.2638586>
- Jaynes ET** (2003) *Probability Theory: The Logic of Science*, 1st ed. Cambridge: Cambridge University Press.
- Jordaán I** (2005) *Decisions under Uncertainty*, 1st ed. Cambridge: Cambridge University Press.
- Kiureghian AD and Ditlevsen O** (2009) Aleatory or epistemic? Does it matter? *Structural Safety* 31, 105–112. <https://doi.org/10.1016/j.strusafe.2008.06.020>
- Lloyd's Register** (2017) *Guidance notes for the risk based inspection of hull structures. Version 2.0* [info.lr.org/en-gb/guidance-notes-for-risk-based-inspection-for-hull-structures/](http://info.lr.org/en-gb/guidance-notes-for-risk-based-inspection-for-hull-structures/). (accessed 05 September 2018).
- Madsen HO** (1987) Model updating in reliability analysis. *Proceedings of International Conference on Applications of Statistics and Probability in Soil and Structural Engineering* 5, 564–577.
- McElreath R** (2020) *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, 2nd ed. London: Chapman & Hall. <https://doi.org/10.1201/9780429029608>
- Nannapaneni S and Mahadevan S** (2016) Reliability analysis under epistemic uncertainty. *Reliability Engineering and System Safety* 155, 9–20. <https://doi.org/10.1016/j.res.2016.06.005>
- Nielsen L, Glavind ST, Qin J and Faber M** (2019) Faith and fakes – dealing with critical information in decision analysis. *Civil Engineering and Environmental Systems* 36 32–54. <https://doi.org/10.1080/10286608.2019.1615476>
- Pozzi M and Der Kiureghian A** (2011) Assessing the value of information for long-term structural health monitoring. *Proceedings of Health Monitoring of Structural and Biological Systems 2011* 510, 812–825.
- Qin HW, Zhou W and Zhang S** (2015) Bayesian inferences of generation and growth of corrosion defects on energy pipelines based on imperfect inspection data. *Reliability Engineering and System Safety* 144, 334–342. <http://dx.doi.org/10.1016/j.res.2015.08.007>
- Sankararaman S and Mahadevan S** (2013) Distribution type uncertainty due to sparse and imprecise data. *Mechanical Systems and Signal Processing* 37, 182–198. <https://doi.org/10.1016/j.ymsp.2012.07.008>
- Straub D** (2004) *Generic approaches to risk based inspection planning for steel structures. Thesis*. <https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/148213/eth-1550-01.pdf>. (accessed 08 June 2018).
- Straub D** (2014) Value of information analysis with structural reliability methods. *Structural Safety* 49, 75–85. <http://dx.doi.org/10.1016/j.strusafe.2013.08.006>
- Straub D and Faber MH** (2004) System effects in generic risk based inspection planning. *Journal of Offshore Mechanics and Arctic Engineering* 126, 265–271.
- Straub D and Faber MH** (2005) Risk based inspection planning for structural systems. *Structural Safety* 27, 335–355. <https://doi.org/10.1016/j.strusafe.2005.04.001>
- Straub D, Sørensen JD, Goyet J and Faber M** (2006) Benefits of risk based inspection planning for offshore structures. *Proceedings of International Conference on Offshore Mechanics and Arctic Engineering*.
- Tang W** (1973) Probabilistic updating of flaw information. *Journal of Testing and Evaluation* 1, 459–467.
- Thoft-Christensen P and Baker MJ** (1982) *Structural Reliability Theory and its Applications*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-68697-9>.
- Wintle JB, Kenzie BW, Amphlett GJ and Smalley S** (2001) Best practice for risk based inspection as part of plant integrity management.

**Cite this article:** Di Francesco D, Chryssanthopoulos M, Faber MH and Bharadwaj U (2021). Decision-theoretic inspection planning using imperfect and incomplete data. *Data-Centric Engineering*, 2: e18. doi:10.1017/dce.2021.18