Probabilistic Sensitivity Analysis of Optimised Preventive Maintenance Strategies for Deteriorating Infrastructure Assets

A. Daneshkhah[a,*], N. G. Stocks[a], P. Jeffrey[b]

[a] *Warwick Centre for Predictive Modelling, School of Engineering, University of Warwick, Covetry CV4 7AL, UK*
[b] *Cranfeld Water Science Institute, School of Water, Energy, and Environment, Cranfeld University, Cranfeld, MK43 0AL*

Abstract

Efficient life-cycle management of civil infrastructure systems under continuous deterioration can be improved by studying the sensitivity of optimised preventive maintenance decisions with respect to changes in model parameters. Sensitivity analysis in maintenance optimisation problems is important because if the calculation of the cost of preventive maintenance strategies is not sufficiently robust, the use of the maintenance model can generate optimised maintenances strategies that are not cost-effective. Probabilistic sensitivity analysis methods (particularly variance based ones), only partially respond to this issue and their use is limited to evaluating the extent to which uncertainty in each input contributes to the overall output's variance. These methods do not take account of the decision-making problem in a straightforward manner. To address this issue, we use the concept of the Expected Value of Perfect Information (EVPI) to perform decision-informed sensitivity analysis: to identify the key parameters of the problem and quantify the value of learning about certain aspects of the life-cycle management of civil infrastructure system. This approach allows us to quantify the benefits of the maintenance strategies in terms of expected costs and in the light of accumulated information about the model parameters and aspects of the system, such as the ageing process. We use a Gamma process model to represent the uncertainty associated with asset deterioration, illustrating the use of EVPI to perform sensitivity analysis on the optimisation problem for age-based and condition-based preventive maintenance strategies. The evaluation of EVPI indices is computationally demanding and Markov Chain Monte Carlo techniques would not be helpful. To overcome this computational difficulty, we approximate the EVPI indices using

*Corresponding author
Email addresses: a.daneshkhah@warwick.ac.uk (A. Daneshkhah), N.G.Stocks@warwick.ac.uk (N. G. Stocks), p.j.jeffrey@cranfield.ac.uk (P. Jeffrey)

Gaussian process emulators. The implications of the worked numerical examples discussed in the context of analytical efficiency and organisational learning.
Keywords: Cost-benefit analysis, Deterioration models, Expected Value of Partial Perfect Information, Gaussian process, optimised maintenance, Time Input emulator, Uncertainty quantification

1. Introduction

The cost effective life-cycle management of civil infrastructure systems is highly dependent on the determination of optimal maintenance and rehabilitation strategies. The determination of optimal maintenance decisions is widely recommended [6] as an effective way of minimising system downtime and corresponding maintenance costs. For instance, Dobbs et al. [1] report that maintenance costs for infrastructure systems such as water energy, rail, etc. are rapidly rising and current estimates suggest that optimized maintenance strategies could save \$100bn p.a. on global infrastructure costs. Infrastructure maintenance practices have traditionally been premised on one of two strategies; Corrective Maintenance (CM) which involves repairing failed components and systems, or Preventative Maintenance (PM) which involves the systematic inspection and correction of incipient failures before they develop into major defects. Recent years have seen increasing dominance of PM approaches with overall costs demonstrated to (perhaps counter-intuitively) be lower than for a CM strategy. PM is widely used to mitigate asset deterioration and reduce the risk of unexpected failure and as a strategy can be sub-classified into two approaches; time-based maintenance (TBM), where maintenance activities take place at predetermined time intervals, and condition-based maintenance (CBM) where interventions are prompted by information collected through condition sensing and monitoring processes (either manual or automated). Ahmad and Kamaruddin [6] provide an extensive review comparing TBM against CBM (see also [2, 3, 4, 5]).

Preventive maintenance strategies (both time and condition based) are widely used for infrastructure life-cycle management decision making. These strategies can be planned and scheduled and their costs are typically lower than those for CM strategies. However, early preventive maintenance intervention adds little to the reliability of the system and can lead to unnecessary costs, hence maintenance strategies often comprise a combination of preventative and corrective approaches. The challenge is then to identify the optimal PM decision that achieves the best balance between these types of maintenance and minimise overall maintenance costs, controlled over an appropriate time period. The central challenge for those who wish to make informed PM decisions is that determining the time to first inspection, maintenance intervention, or replacement is confounded by model parameter uncertainties associated with the adopted failure, deterioration, repair, or maintenance model. Consequently, SA of the model output (to identify an optimal maintenance strategy) with respect to the changes in the model parameters is of great interest. In this paper we investigate the issue of SA for maintenance optimisation models. To achieve this, we consider time based and condition based preventive maintenance strategies for infrastructure systems under continuous deterioration. Both strategies are discussed in detail in [6, 11] and references therein. Under TBM, a component is replaced (or perfectly repaired) either at failure (CM) or when it has reached age T - whichever occurs first. The central objective of a TBM decision problem is to determine the replacement time which minimizes expected total cost. The CBM strategy involves the periodic inspection of a component/structure at a fixed time interval T_i and cost C_i. At the i^{th} inspection, one of the following actions might be taken: (i) the system is operating satisfactorily and no action is required to be taken; (ii) immediate preventative maintenance is required to avoid component or system failure; (iii) a failure is identified and corrective maintenance (or a perfect repair) is required to restore the system's functionality (see Subsection 5.2 and [11] for further details). The optimal maintenance decision under the CBM strategy is taken as the inspection time and the PM ratio which are similarly determined by minimising the cost function of interest. The decision under a CBM policy for a deteriorating component constitutes a two-dimensional optimisation problem, whilst for the TBM case the aim is to find the critical age as a single variable. It has been argued that the types of PM strategy discussed above is more useful in practice (particularly for larger and more complex systems) since it removes the need to record component ages ([6, 7]).

As inferred above, the preventive maintenance policy cost function is influenced by both the deterioration model and repair model's parameters. Thus, the calculation of a mean cost rate for a particular preventive maintenance policy is not sufficiently robust because of the uncertainty around parameter values, and the corresponding maintenance model can generate inefficient outcomes. In other words, the identification of an optimal maintenance intervention becomes sensitive to the model parameters creating uncertainty as to the optimal strategy. Variance based approaches [14] offer a partial answer to this problem and can be used to assess the degree to which uncertainty in each variable contributes to the overall variance in model output. However, these approaches do not take account of the decision-making context properly. In order to address this issue, we make use of the concept of the Expected Value of Partially Perfect Information (EVPPI). The EVPPI provides a decision-informed SA framework which enables researchers to determine the key parameters of the problem and quantify the value of learning about certain aspects of the system ([8, 7]). In maintenance studies ([9, 10]), this information can play an important role, particularly where we are interested in not only identifying an optimal maintenance decision but in also gathering additional information about the system characteristics including the deterioration process to improve the robustness of decisions.

The determination of EVPPI involves the calculation of multi-dimensional integrals that are often computationally demanding, making conventional numerical integration or Monte Carlo simulation techniques infeasible in practice. To partially overcome this computational difficulty, we follow the work of [7, 8], and execute SA through the use of Gaussian process emulators. The following section presents a well-known probabilistic model of deterioration; the Gamma process model, and discusses how this relates to TBM and CBM maintenance optimisation problems. We go on to describe how Gaussian Process (GP) emulators can be used to compute EVPPIs within the context of decision-theoretic SA. Robust optimised maintenance decisions are then derived for two forms of PM policy using several illustrative settings of varying complexity. We conclude by discussing the implications of our approach and identify opportunities for future work.

2. Deterioration models

Infrastructure asset deterioration processes are uncertain and can best be regarded as stochastic. Two previous studies have demonstrated the values of using Gamma process models to analysis the deterioration of physical assets. Pandey et al. [11] compared the use of random variable and gamma process models in the life-cycle management of infrastructure systems. They demonstrated that the random variable model cannot capture the temporal variability associated with the evolution of asset degradation. As a consequence, this model tends to underestimate the life-cycle cost due to the lack of consideration of temporal uncertainty. Van Noortwijk [12] extensively reviewed the application of stochastic deterioration processes, and particularly the use of the Gamma process model in maintenance. He concluded that gamma processes are well suited for modelling the temporal variability of deterioration, and of particular value when determining optimal inspection and maintenance decisions.

We now briefly introduce the *Gamma process* for deterioration modelling of an ageing asset. In mathematical terms, a gamma process is a stochastic process with independent non-negative increments having a gamma distribution ([11, 12]). The Gamma process with a shape function $\nu(t) > 0$ and scale parameter $\xi > 0$ is a continuous-time stochastic process $\{X(t), t \geq 0\}$ with the following properties:

1. $Pr(X(0) = 0) = 1$
2. $X(\iota) - X(t) \sim Ga(\nu(\iota) - \nu(t), \xi), \forall\ \iota > t \geq 0$
3. $X(t)$ has independent increments

and where $\nu(t)$ is a non-decreasing, right-continuous, real-valued function of $t \geq 0$ with $\nu(0) \equiv 0$.

Let $X(t)$ denote the deterioration at time $t \geq 0$, and let $X(t)$ follows a gamma process with the shape function $\nu(t) > 0$ and scale parameter $\xi > 0$, then the probability density function of $X(t)$ is given by

$$f_{X(t)}(x) = Ga(x \mid \nu(t), \xi) = \frac{(x/\xi)^{v(t)-1}}{\xi\Gamma(v(t))} \exp\{-x/\xi\}, \quad \text{for} \quad x \geq 0 \tag{1}$$

The structural failure for a deteriorating structure or component is defined as an event when its deteriorating resistance, denoted by $R(t) = r_0 - X(t)$, falls short of the applied stress s. The initial resistance r_0 and s are assumed to be fixed and known. We denote $\rho = (r_0 - s) > 0$ as the available design margin or a failure threshold. We let the time at which failure occurs be denoted by the lifetime T (also called the first hitting time of level ρ). Since the deterioration of a component at time t is given by Eq. (1), the cumulative lifetime distribution of this is then given by

$$F_T^G(t) = Pr(T \leq t) = Pr(X(t) \geq \rho) = 1 - \mathcal{G}(\rho; \nu(t)t, \xi) \tag{2}$$

where $\mathcal{G}(\rho; \nu(t)t, \xi)$ denote the cumulative distribution function of the deterioration model at ρ.

Expression (2) features outstanding duality between a component's deterioration and its lifetime that makes the Gamma process model tractable for cycle-life management analysis. It should be noted that the lifetime probability density function, denoted by $f_T^G = \frac{\partial}{\partial t} F_T^G(t)$, has no closed form expression, and the corresponding maintenance optimisation problem requires a computationally fast and powerful numerical method.

3. Optimal Preventive Maintenance Policy

The central objective of a preventive maintenance (TBM or CBM) optimisation model is to determine the value of the decision variable T (replacement time or inspection time) that optimizes a given objective function amongst the available alternative maintenance decisions. For instance in a TBM policy, the optimisation problem is usually defined over a finite time horizon $[0, t]$, and the objective function, denoted by $C(t)$, represents costs over the interval $[0, t]$. For infinite horizon models, we seek to minimise the long-term average costs [13]. If a life cycle of an asset is defined over the period between two consecutive maintenance/replacements, then the expected cost per unit of time under decision T (which could be either optimised maintenance time or inspection interval) is given by

$$\mathcal{C}(T|\boldsymbol{\theta}) = \frac{C(T|\boldsymbol{\theta})}{L(T|\boldsymbol{\theta})} \tag{3}$$

where $C(T|\boldsymbol{\theta})$ is the expected cost during the system's life cycle, $L(T|\boldsymbol{\theta})$ is the expected length of the life cycle or length of time between two consecutive replacements/repairs, and $\boldsymbol{\theta}$ is the vector of deterioration/failure and time to repair/replacement. We assume that system/component failure and time to repair or replacement is a random variable characterized by a distribution as discussed in Section 5.

The following formula is an example of the expected cost per unit of a component under a general TBM policy

$$\mathcal{C}(T) = \frac{c_1 F(T) + c_2 R(T)}{T \cdot R(T) + \int_0^T t f(t) \mathrm{d}t + \tau} \tag{4}$$

where $F(T)$ is the failure distribution function of a system at time T (or probability of unplanned replacement due to an unexpected failure), $R(T) = 1 - F(T)$ is the probability of planned replacement at time T, c_1 is the cost of a corrective maintenance, c_2 is the cost of planned replacement and τ is the expected duration of replacement. The objective is then to identify the optimal strategy T^* that corresponds to the minimum cost rate (cost per unit of time), that is;

$$T^* = \arg\min_{T>0} \{\mathcal{C}(T)\}. \tag{5}$$

A similar method is used to determine the optimised CBM strategy. The cost function in this policy is the mean cost rate which is defined as

$$\mathcal{K}(t_I, \upsilon) = \frac{E[C(t_I, \upsilon)]}{E[L(t_I, \upsilon)]} \tag{6}$$

where $E[C(t_I, \upsilon)]$ represents the renewal cycle cost, $E[L(t_I, \upsilon)]$ is the renewal cycle length, t_I is the inspection time interval and υ is the PM ratio. The details of numerator and denominator of the mean cost rate will be given in Section 5. The objective is then to find t_I^* and υ^* so that $\mathcal{K}(t_I^*, \upsilon^*)$ becomes the minimal cost solution.

3.1. Uncertainty quantification via decision-informed sensitivity analysis

The optimal maintenance strategies derived by minimizing the expected cost rate is influenced by characteristics such as the deterioration process or failure behaviour of the system and the characteristics of maintenance tasks (including repair/replacement policy, maintenance crew and spare part availability etc.). These characteristics are subject to uncertainty, prompting study of the sensitivity of an optimal maintenance strategy with respect to changes in the model parameters and other uncertain inputs. Such an analysis improves understanding of the 'robustness' of the derived inferences or predictions of the model, and, offers a tool for determining the critical influences on model predictions ([14]). Zitrou et al. [7] summarise the main sensitivity measures and discuss their values and applications in an extensive SA. They conclude that a simple yet effective method of implementing SA is to vary one or more parameter inputs over some plausible range, whilst keeping the other parameters fixed, and then examine the effects of these changes on the model output. Although this method is straightforward to implement and interpret, it becomes inconvenient where there are large numbers of model parameters or when the model is computationally intensive.

In order to resolve this difficulty, we use a variance-based method for SA ([14]). This approach can capture the fractions of the model output variance which are explained by the model inputs. In addition, it can also provide the total contribution to the output variance of a given input - i.e. its marginal contribution and its cooperative contribution. The contribution of each model's input to the model output variance serves as an indicator of how strong an influence a certain input or parameter has on model output variability. However, within a decision-making context like the maintenance optimisation problem, we are primarily interested in the effect of parameter uncertainty on corresponding utility or loss. To achieve this objective, we use the concept of EVPPI as a measure of parameter importance ([7, 8]). Incorporating the value of information (or EVPPI) in a sensitivity analysis allows the decision-maker (or model user) to relate the importance of each uncertain input parameter directly to the decision problem at hand, something that is lacking in a traditional variance-based sensitivity analysis method. The EVPPI approach thus allows the application of SA to the maintenance optimisation model and identifies the model parameters for which collecting additional information (learning) prior to the maintenance decision would have a significant impact on total cost.

Monte Carlo sampling can be used to estimate partial EVPIs [18], but again, in the case of computationally expensive models this may not be practical due to the numbers of model runs typically required. Oakley [8] shows how Gaussian process emulators can be used to obtain estimates more efficiently in this case.

4. Decision-informed sensitivity analysis

4.1. Sensitivity analysis

The mean cost rate induced by a specific maintenance strategy (chosen value for T or t_I) is effected by features like the deterioration process of individual structure/system and the aspects of the replacement/repair task. As these aspects are part of a real-world system, they are then subject to uncertainty. It is thus of key importance to investigate sensitivity of the maintenance model with respect to these uncertain aspects.

Sensitivity analysis is widely used in modelling to examine whether alternative assumptions or modelling choices lead to different predictions or inference. In general, there are two types of approach: 'local and 'global sensitivity analysis. The aim of the former is to evaluate the change in output, $f(\boldsymbol{\theta})$ due to small perturbations in the input from some baseline value/choice, and typically involves the consideration of partial derivatives of the function under study with respect to the variables, $\partial f(\boldsymbol{\theta})/\partial\theta_i$ ([34, 35]). When $f(\cdot)$ is non-linear in its inputs, $\mathbf{x}$ and small perturbations of the inputs do not adequately reflect the input uncertainty, a local sensitivity analysis is unlikely to be a plausible approach. In this situation, a global sensitivity analysis can be used to examine how the output varies as the inputs vary over some range. Where we are interested in reducing uncertainty about model inputs by collecting more data, a global sensitivity approach may identify how to prioritize data collection by identifying the most important uncertain inputs.

There are two approaches to global sensitivity analysis: variance-based methods, and decision-theoretic approaches based on the expected value of perfect information. The variance-based global sensitivity analysis method is extensively reviewed in [36], and its applications can be found in [37]. The two most useful measures of input importance within the variance-based approach are the main effect index ($z_i(\theta_i) = E(f(\boldsymbol{\theta})|\theta_i) - E(f(\boldsymbol{\theta})|\theta_i)$) and the total sensitivity index. A third concept, related to the main effect index, is the main effect plot, which can be used to display graphically the relationship between an input and the output ([16, 14]).

There are various computational methods for estimating these sensitivity measures. One of the earliest proposed approaches was the FAST (Fourier amplitude sensitivity test, [36, 14]) which involves evaluating simulator outputs at inputs along a curve which explores the input space, oscillating at different frequencies in each input dimension. Other approaches relate enhancements on simple Monte Carlo sampling [38]. The computation of the sensitivity indices for the complex functions (e.g., consists of non-linear terms or expressed based on a complicated mathematical formulae) would be very challenging. In these situations, the emulators can be then used for computationally expensive simulators. In [17], the GP emulator was used to compute sensitivity indices and produce main effects plots (see also [25, 39]).

Variance-based measures are more concerned with the individual elements within vector outputs (or simply scalar outputs) and express what fraction of the variance of $f(\boldsymbol{\theta})$ can be attached to an uncertain input variable θ_i, or any subset of $\boldsymbol{\theta}$. However, these approaches do not take account of the decision-making context properly. In order to tackle this drawback, a sensitivity analysis method based on the concept of value of information which allows the decision-maker to relate the importance of each uncertain input parameter directly to the decision problem at hand was developed in [8].

In the field of life-cycle management of civil infrastructure, the value of information concept is widely used to determine the optimum preventive maintenance policy or condition monitoring strategy. For instance, a methodology based on partially observable Markov decision process was proposed in [40] to calculate the value provided by condition monitoring systems for infrastructure assets. This was achieved by combining "value of information" concepts with Markov sequential decision process.

The determination of the benefits offered by the two condition monitoring technologies can then be ascertained and the decision maker can choose the most appropriate one in an informed manner. In order to understand the factors that influence the information value, sensitivity analysis on the specific model parameters are carried out. In order to understand the impact of accuracy, the parameter can be varied, keeping other parameters constant, and the resulting total expected costs can be calculated for each technology. In a similar study [41] Markov chains and simulation techniques were used to quantify the benefits of condition monitoring for wind turbines by conducting sensitivity analysis to operational parameters.

A comprehensive overview of the mathematical framework for estimating the value of information adapted to life-cycle analysis of structural systems was provided in [42, 43]. It was shown the computation of the expected value of information relating to decisions on maintenance of the civil infrastructure systems requires a large number of life-cycle analyses, and the computational cost can be very high when decisions concern the systems that are modelled with complex computational models[44]. In order to tackle this computational burden, it was suggested to use the Kriging meta-models.

In this paper, we provide a holistic approach for guiding making optimised decisions in the presence of uncertainty using value of information analysis. We show how global sensitivity analysis can be conducted within the framework of preventive maintenance decision making, based on the concept of the expected value of perfect information. It should be noted that the variance based sensitivity analysis method is considered as a special case of this approach. The computational challenges are tackled using computationally efficient meta-models known as Gaussian process emulators which enable us to compute the value of information indices (including EVPI and EVPPI) of complex scenarios. In this section, we describe how GP emulators can be used to compute EVPPIs within the context of decision-theoretic sensitivity analysis.

A comprehensive overview of the mathematical framework for estimating the value of information adapted to life-cycle analysis of structural systems is provided in [42, 43]. It was shown the computation of the expected value of information relating to decisions on maintenance of the civil infrastructure systems requires a large number of life-cycle analyses, and the computational cost can be very high when decisions concern the systems that are modelled with complex computational models [44]. In order to tackle this computational burden, it was suggested to use the Kriging meta-models.

In this paper, we provide a holistic approach for guiding making optimised decisions in the presence of uncertainty using value of information analysis. We show how global SA can be conducted within the framework of preventive maintenance decision making, based on the concept of the expected value of perfect information. It should be noted that the variance based SA method is considered as a special case of this approach. The computational challenges are tackled using computationally efficient meta-models known as Gaussian process emulators which enable us to compute the value of information indices (including EVPI and EVPPI) of complex scenarios. In this section, we describe how GP emulators can be used to compute EVPPIs within the context of decision-theoretic SA.

4.2. Expected Value of Perfect Information

To briefly recap, the objective function of interest to us is the expected cost function (e.g., the cost rate function given in Equation (4) for TBM or the mean cost rate given in (6) for CBM). These cost functions take reliability and maintenance parameters as uncertain inputs (denoted by $\boldsymbol{\theta}$) and a decision parameter, T (which could be critical age or periodic inspection interval). A strategy parameter (which is fixed) needs to be selected in the presence of unknown reliability and maintenance parameters. These unknown parameters can be modelled by a joint density function, $\pi(\boldsymbol{\theta})$. In the maintenance optimisation setting, the decision maker can choose the strategy parameter T (from a range or set of positive numbers) where each value of T corresponds to a maintenance decision. The decision T is selected so that the following utility function is maximised

$$U(T, \boldsymbol{\theta}) = -\mathcal{C}(T; \boldsymbol{\theta}) \tag{7}$$

where $\mathcal{C}(T; \boldsymbol{\theta})$ is a generic cost function per unit of time given the unknown parameters $\boldsymbol{\theta}$.

Suppose that we need to make a decision now, on the basis of the information in $\pi(\boldsymbol{\theta})$ only. The optimal maintenance decision (known as *baseline* decision), given no additional information, has expected utility

$$U_0 = \arg\max_{T>0} E_{\boldsymbol{\theta}}\left[U(T, \boldsymbol{\theta})\right] \tag{8}$$

where

$$E_{\boldsymbol{\theta}}\left[U(T, \boldsymbol{\theta})\right] = -\int_{\boldsymbol{\theta}} \mathcal{C}(T; \boldsymbol{\theta})\pi(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \tag{9}$$

Now suppose that we wish to learn the precise value of a parameter θ_i in $\boldsymbol{\theta}$ before making a decision (e.g., through exhaustive testing; new evidence elicited from the domain expert). Given θ_i, we are still uncertain about the remaining input parameters, $\boldsymbol{\theta}_{\underline{i}} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n)$, and so we would choose the maintenance strategy to maximise

$$E_{\boldsymbol{\theta}|\theta_i}\left[U(T, \boldsymbol{\theta})\right] = -\int_{\boldsymbol{\theta}_{\underline{i}}} \mathcal{C}(T; \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \theta_i)\mathrm{d}\boldsymbol{\theta}_{\underline{i}} \tag{10}$$

The expected utility of learning θ_i is then given by

$$U_{\theta_i} = E_{\theta_i}\left[\arg\max_{T>0} E_{\boldsymbol{\theta}|\theta_i}\left\{U(T,\boldsymbol{\theta})\right\}\right] \tag{11}$$

Now, learning about parameter θ_i before making a maintenance decision will benefit the decision-maker by

$$\text{EVPI}_{\theta_i} = E_{\theta_i}[U_{\theta_i}] - U_0. \tag{12}$$

Therefore, the quantity EVPI_{θ_i}, known as the partial Expected Value of Perfect Information (partial EVPI or EVPPI), is a measure of the importance of parameter θ_i in terms of the cost savings that further learning (data collection) will achieve.

EVPIs allow for SA to be performed in a decision-theoretic context. However, the computation of partial EVPIs as in (12) requires the evaluation of expectations of utilities over many dimensions. Whereas the one-dimensional integral $E_{\theta_i}[U_{\theta_i}]$ can be evaluated efficiently using Simpson's rule, averaging over the values of multiple parameters is computationally intensive. One way to approximate these expectations is to use a Monte Carlo numerical method. However, the Monte Carlo based integration methods require a large number of simulations which make the computation of the EVPPIs impractical. Zitrou et al. [7] propose an alternative method for resolving this problem by utilizing a Gaussian Process emulator based SA to the objective function of interest. This method enables computation of the multi-dimensional expectations at a limited number of model evaluations with relative computational ease. We develop this method further for the purposes mentioned above.

4.3. Gaussian Process Emulators

An emulator is an approximation of a computationally demanding model, referred to as the *code*. An emulator is typically used in place of the code, to speed up calculations. Let $\mathcal{C}(\cdot)$ be a code that takes as input a vector of parameters $\boldsymbol{\theta} \in \mathcal{Q} \subset \mathbb{R}^q$, for some $q \in \mathbb{Z}_+$, and has output $y = \mathcal{C}(\boldsymbol{\theta})$, where $y \in \mathbb{R}$. As we will see later on, this is not a restrictive assumption, and we will let $y \in \mathbb{R}^s$, for some $s \in \mathbb{Z}_+$. For the time being, let $\mathcal{C}(\cdot)$ be a deterministic code, that is for fixed inputs, the code produces the same output each time it 'runs'.

An emulator is constructed on the basis of a sample of code runs, called the *training set*. In a Gaussian Process emulation context, we regard $\mathcal{C}(\cdot)$ as an unknown function, and use a $q-$

dimensional Gaussian Process (GP) to represent prior knowledge on $\mathcal{C}(\cdot)$, i.e.

$$\mathcal{C}(\cdot) \sim N_q(m(\cdot), v(\cdot, \cdot)) \tag{13}$$

We subsequently update our knowledge about $\mathcal{C}(\cdot)$ in the light of the training set, to arrive at a posterior distribution of the same form.

Expression (13) implies that for every $\{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n\}$ output $\{\mathcal{C}(\boldsymbol{\theta}_1), \ldots \mathcal{C}(\boldsymbol{\theta}_n)\}$ has a prior multivariate normal distribution with mean function $m(\cdot)$ and covariance function $v(\cdot, \cdot)$. There are many alternative models for the mean and covariate functions $m(\cdot)$. Here, we use the formulation in line with [15], and assume

$$m(\boldsymbol{\theta}) = h(\boldsymbol{\theta})^\intercal \boldsymbol{\beta} \tag{14}$$

for the mean function, and

$$v(\boldsymbol{\theta}, \boldsymbol{\theta'}) = \sigma^2 c(\boldsymbol{\theta}, \boldsymbol{\theta'}). \tag{15}$$

for the covariance function. In (14), $h(\cdot)$ is a vector of q known regression functions of $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ is a vector of coefficients. In (15), $c(\cdot, \cdot)$ is a monotone correlation function on $\mathbb{R}^+$ with $c(\boldsymbol{\theta}, \boldsymbol{\theta}) = 1$ that decreases as $|\boldsymbol{\theta} - \boldsymbol{\theta'}|$ increases. Furthermore, the function $c(\cdot, \cdot)$ must ensure that the covariance matrix of any set of outputs is positive semi-definite. Throughout this paper, we use the following correlation function which satisfies the aforementioned conditions and is widely used in the Bayesian Analysis of Computer Code Outputs (BACCO) emulator ([8, 16]) for its computational convenience:

$$c(\boldsymbol{\theta}, \boldsymbol{\theta'}) = \exp\{-(\boldsymbol{\theta} - \boldsymbol{\theta'})^\intercal \boldsymbol{R}(\boldsymbol{\theta} - \boldsymbol{\theta'})\} \tag{16}$$

where $\boldsymbol{R}$ is a diagonal matrix of positive smoothness parameters (also known as length scales). $\boldsymbol{R}$ determines how close two inputs $\boldsymbol{\theta}$ and $\boldsymbol{\theta'}$ need to be such that the correlation between $\mathcal{C}(\boldsymbol{\theta})$ and $\mathcal{C}(\boldsymbol{\theta'})$ takes a particular value. For mathematical tractability, the conjugate prior form for β and σ^2, the normal inverse gamma distribution, is assumed ([17]):

$$p(\beta, \sigma^2) \propto (\sigma^2)^{-\frac{1}{2}(\kappa+q+2)} \exp\{-\{(\beta - \mathbf{z})^T V^{-1} (\beta - \mathbf{z}) + a\}/(2\sigma^2)\}$$

where the hyperparameters $\mathbf{z}, V, a$ and κ (the number of regressors in the mean function) are known.

The cost function of interest $\mathcal{C}(\cdot)$ is evaluated at N design points $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$ to generate the outputs $\mathbf{y}^T = (\mathcal{C}(\boldsymbol{\theta}_1), \ldots \mathcal{C}(\boldsymbol{\theta}_N))$. The following set $\mathcal{D} = \{(\boldsymbol{\theta}_i, \mathcal{C}(\boldsymbol{\theta}_i)), i = 1, \ldots, N\}$ is then considered as the data required to train the standard GP. These design points are chosen based on a suitable space filling design, such as Max-Min Latin Hypercube scheme which is designed to ensure the multi-dimensional parameters pace is fully covered without having to use a very large sample size which is required in the Monte Carlo based methods. As a result, we only need to evaluate $\mathcal{C}(\boldsymbol{\theta})$ at limited input points. Since $\boldsymbol{\theta}$ is unknown, the beliefs about $\boldsymbol{\theta}$ is represented by the probability distribution $\pi(\boldsymbol{\theta})$. Therefore, the choice of the design points will also depend on $\pi(.)$ (the choice of design points is discussed in [19]). The standardised posterior distribution of $\mathcal{C}(\cdot)$ given $\mathcal{D} = \{(\boldsymbol{\theta}_i, \mathcal{C}(\boldsymbol{\theta}_i)), i = 1, \ldots, N\}$ is

$$\frac{\mathcal{C}(\boldsymbol{\theta}) - m^*(\boldsymbol{\theta})}{\hat{\sigma}\sqrt{c^*(\boldsymbol{\theta}, \boldsymbol{\theta}')}} \mid \mathcal{D}, \boldsymbol{R} \sim t_{q+N} \tag{17}$$

where t_{q+n} is a student t random variable with $n + q$ degrees of freedom,

$$\begin{aligned}
m^*(\boldsymbol{\theta}) &= \mathbf{h}(\boldsymbol{\theta})^T \hat{\beta} + \mathbf{t}(\boldsymbol{\theta})^T A^{-1}(\mathbf{y} - H\hat{\beta}) \\
c^*(\boldsymbol{\theta}, \boldsymbol{\theta}') &= c(\boldsymbol{\theta}, \boldsymbol{\theta}') - \mathbf{t}(\boldsymbol{\theta})^T A^{-1} \mathbf{t}(\boldsymbol{\theta}') + \\
&(\mathbf{h}(\boldsymbol{\theta})^T - \mathbf{t}(\boldsymbol{\theta})^T A^{-1} H)(H^T A^{-1} H)^{-1}(h(\boldsymbol{\theta}')^T - \mathbf{t}(\boldsymbol{\theta}')^T A^{-1} H)^T \\
&\mathbf{t}(\boldsymbol{\theta})^T = (c(\boldsymbol{\theta}, \boldsymbol{\theta}_1), \ldots, c(\boldsymbol{\theta}, \boldsymbol{\theta}_n)) \\
&H^T = (\mathbf{h}(\boldsymbol{\theta}_1), \ldots, \mathbf{h}(\boldsymbol{\theta}_n))
\end{aligned}$$

and

$$A = \begin{pmatrix} 1 & c(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) & \ldots & c(\boldsymbol{\theta}_1, \boldsymbol{\theta}_n) \\ c(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1) & 1 & & \vdots \\ \vdots & & \ddots & \\ c(\boldsymbol{\theta}_n, \boldsymbol{\theta}_1) & \ldots & & 1 \end{pmatrix}$$

$$\begin{aligned}
\hat{\beta} &= V^*(V^{-1}\mathbf{z} + H^T A^{-1}\mathbf{y}) \\
\hat{\sigma}^2 &= \frac{\{a + \mathbf{z}^T V^{-1}\mathbf{z} + \mathbf{y}^T A^{-1}\mathbf{y} - \hat{\beta}^T (V^*)^{-1}\hat{\beta}\}}{(N + q - 2)} \\
V^* &= (V^{-1} + H^T A^{-1} H)^{-1}
\end{aligned}$$

The outputs corresponding to any set of inputs will now have a multivariate t-student distribution as presented in (17). The resulting t-distribution is obtained as a marginal distribution for $\mathcal{C}(\boldsymbol{\theta})$ after integrating out the hyperparameters β and σ^2. It is not tractable to remove analytically the smoothness parameters $\mathbf{R}$, and we deal with uncertainty in $\mathbf{R}$ by sampling from the posterior distribution of $\mathbf{R}|\mathbf{y}$ using MCMC methods (see [20]). These estimates can be obtained by using the posterior mode approach, and cross validation.

The GP emulators developed above are useful tools for uncertainty and SA ([8, 17]) and it has been shown that they perform better than standard Monte-Carlo methods in terms of both accuracy of model output and computational effort. This is mainly due to their analytical efficiency which can be used to evaluate $E[\mathcal{C}(\boldsymbol{\theta})]$ and $Var[\mathcal{C}(\boldsymbol{\theta})]$ relatively fast. Thus, it is trivial to show that if $\mathcal{C}(\boldsymbol{\theta}) \sim GP(\cdot,\cdot)$, then

$$E[\mathcal{C}(\boldsymbol{\theta})] = \int\limits_{\boldsymbol{\theta}} \mathcal{C}(\boldsymbol{\theta})\pi(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \qquad (18)$$

follows a GP distribution.

In order to perform the decision-theoretic sensitivity approach, we need to compute the partial EVPIs given in (12). By using an emulator, the expected value of the utility function $U(T, \boldsymbol{\theta})$ for each decision variable T, including the first and second moments can be rapidly computed with relatively low much computational effort. In recent years, GP emulators have been extensively used for a wide range of applications including sensitivity/uncertainty analysis [25, 16, 17], calibration [20], forecasting [22, 49, 50], optimisation [50, 7], etc. A detailed comparison of the use of Monte-Carlo and emulator methods to deal with uncertainty and sensitivity analyses and relevant examples is provided in [16] showing that both methods can provide an estimate for the model/quantity of interest, with an error term to represent model uncertainty. Model uncertainty can be reduced by executing multiple model runs which, in the case of Monte-Carlo methods can run in to the tens or hundreds of thousands. In the case of the GP emulator, the set of model runs is used to construct the emulator and achieving acceptable accuracy would require only a handful of runs for a model with just one or two inputs, or up to a few hundred for a complex function of many inputs. Therefore, achieving the desired precision can be a cumbersome business for a complex model even with a handful of input variables when using the Monte-Carlo methods. [16] draws a similar conclusion in computing the sensitivity measures for an application in the field of health economics. He shows that achieving negligible bias may require a very large number of simulations. This can lead to evaluate $\mathcal{C}(\boldsymbol{\theta})$ numerous times (of the order of 10000) at different values of $\boldsymbol{\theta}$ to achieve a sufficiently small bias using the Monte Carlo sampling method for a simple case study.

In another study presented in [25], an emulator-based sensitivity analysis was used to examine the changes in system availability and reliability with respect to changes in time-to-failure and time-to-repair distribution parameters. It was shown that only tens to hundreds of model runs are required to construct an emulator (depending on the complexity of the system under study), and subsequently compute the variance-based sensitivity measures while the computation of the same sensitivity indexes would require millions of model runs using the Monte-Carlo method.

In this study, we are interested in identifying the robust optimised PM strategy T^* which minimizes the cost rate function given in (7). This optimization problem can be addressed using two sub approaches. In the first approach, the PM strategy, T belongs to a finite set $\mathcal{T}_m = \{T_1, T_2, \ldots, T_m\}$, and the main objective is to identify the optimal decision among this finite set of decision options. Oakley [8] addressed this issue for a limited number of the available decisions in a health economics context. Zitrou et al. [21, 7] use the same method to find the robust optimized maintenance action. In this framework, a separate GP model is first fitted to approximate the mapping between $\boldsymbol{\theta}$ and $\mathcal{C}(\boldsymbol{\theta}, T_i)$ for each T_i. The computed partial EVPI for each T_i is then used to select the optimized PM strategy over a subset of the parameter space.

In the case that the decision space is not finite or consists of many decision options, the methodology addressed in these works is not useful and practical. In addition, regardless of the model complexity and the model runs required to compute the EVPPI for each decision option, the conventional Monte Carlo based methods will also not be useful when the decision space is not finite [21, 7]. Consequently, we adopt the multi-output Gaussian process models proposed in [22]. They propose various methods to deal with the modelling of multivariate computer code model outputs including *Multi-Output emulator (MO)*, *Many Single Outputs (MS)* and *Time Input (TI)* emulators. The MO emulator is a multivariate version of the single output emulator, where the dimension of the output space is υ. This process allows for the representation of any correlations existing among the multiple outputs. The MS emulator procedure treats the multi-outputs of the function of interest, $\{Y_1, \ldots, Y_s\}$ as independent random variables, and emulates each output Y_j separately. This means that s separate GP emulators are built, each describing the utility for each decision $T \in \mathcal{T}_m$. This is the model that is used in [8, 21]. Finally, the TI emulator is a single-output emulator that considers decision variable T as an additional input parameter. In this paper, we show that this model can be used to find the robust optimised PM when $\mathcal{T}_m$ does not have to be a finite space, and cost rate function $\mathcal{C}(\boldsymbol{\theta}, T)$ can be determined for any value of T over any interval, as (T_l, T_u). In other words, the optimised maintenance strategy T identified using the TI emulator can be a continuous variable, and the expectations of any order of $\mathcal{C}(\boldsymbol{\theta}, T)$ are continuous functions of T, and the utilities of the optimal strategies are calculated without restricting the decision-maker to choose amongst a pre-determined, finite number of options. This feature of the TI emulator outweighs the general correlation structure provided by the MO emulator (see [22]). In the next section, we briefly introduce the TI emulator and demonstrate how it can be used to identify the optimized PM strategy.

4.4. The TI Emulator

Suppose that the optimal decision T in a maintenance optimisation problem (critical age or periodic interval) belongs to an infinite set $\mathfrak{T} = (T_l, T_u)$. We consider T as an additional code input and we are interested in building a single-output emulator to approximate the utility function, $U(T, \boldsymbol{\theta}) = -\mathcal{C}(T; \boldsymbol{\theta})$. As mentioned above and shown in the related literature, the computation of $E_{\boldsymbol{\theta}}[U(T, \boldsymbol{\theta})]$ and $E_{\boldsymbol{\theta}|\theta_i}[U(T, \boldsymbol{\theta})]$ for any $T \in \mathfrak{T}$, required to calculate EVPI and the partial EVPI, using the TI emulator would be very fast and efficient.

The main challenge is to estimate the hyper-parameters of the TI emulator, based on the generated training dataset consisting of code outputs $\mathbf{y} = (y_1 = \mathcal{C}(\boldsymbol{x_1}), \dots, y_N = \mathcal{C}(\boldsymbol{x_N}))$, where $(\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_N})^\intercal$ are design points defined as follows:

$$x_l = (T_i, \boldsymbol{\theta_j}), \quad l = 1, 2, \dots, N = s \times n$$

where T_i is a maintenance decision ($i = 1, \dots, s$) and $\boldsymbol{\theta_j}$ are (reliability, maintainability) parameter values ($j = 1, \dots, n$).

The choice of design points affects how well the emulator is estimated. Here, we choose equally spaced points $\{T_1, \dots, T_s\}$ so that interval $\mathfrak{T}$ is properly covered. Points $(\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \dots, \boldsymbol{\theta_n})^\intercal$ are generated using Latin hypercube sampling (see [23]), which ensures that the multidimensional parameter space is sufficiently covered.

As mentioned earlier, building a TI emulator requires the inversion of an $N \times N$ matrix. Given the size of the training set, this can be computationally challenging. Essentially, there are two ways to build the TI emulator: (1) fit a Gaussian process directly to the whole training set $\mathbf{y}$ obtained as described above; (2) separate $\mathbf{y}$ and fit two Gaussian Processes: one on the set of design points $(\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \dots, \boldsymbol{\theta_n})$ and one on the time input data points $\{T_1, \dots, T_s\}$. Multiple authors ([24, 22, 7]) have concluded that the first approach based on fitting a single GP to the whole training set $\mathbf{y}$

takes longer, but that it produces more accurate results. In addition, they have shown that the relative mean squared error of the posterior predictive mean of the first model (based on fitting a single Gaussian process) is much smaller than when fitting two Gaussian process. We therefore follow their suggestion and fit a single GP to the full training set.

The baseline maintenance strategy is the choice of T that maximises utility, and the baseline expected utility in (8) is now calculated as

$$U_0 = \max_{T \in \mathfrak{T}} E_{\mathcal{C}}\{E_{\boldsymbol{\theta}}\left[\mathcal{C}(T, \boldsymbol{\theta})\right]\} \tag{19}$$

and the utility of the optimal strategy in (10), after learning the value of Θ_i, becomes

$$U_{\theta_i} = \max_{T \in \mathfrak{T}} E_{\mathcal{C}}\{E_{T,\boldsymbol{\theta}|\theta_i}\left[\mathcal{C}(T, \boldsymbol{\theta})\right]\}. \tag{20}$$

Bayesian quadrature ([15]) allows us to compute the expectations given in (8) and (10) relatively fast based on the fitted GP to $\mathbf{y}$. The details of the approximation of this type of integral (expectation) in terms of the fitted GP can be found in [25]. The computation steps of computing EVPI and partial EVPI are given in Algorithm 1.

In Equation (21), $\mathcal{R}_i$ and $\mathbf{W}_i$ are given by

$$\mathcal{R}_i = \int \mathbf{h}(\boldsymbol{\theta}, T_i)^T \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}, \quad \mathbf{W}_i^T = \int \mathbf{t}(\boldsymbol{\theta}, T_i)^T \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

The computation of these integrals is trivial. For example, if inputs are normally distributed, and the correlation and mean functions are respectively given in (14) and (16),), these integrals can be evaluated analytically. If the inputs are not normally or uniformly distributed, then numerical or Monte Carlo integration can be used without significant computational effort [8].

As we are interested in conducting a global sensitivity analysis (how the output varies as the inputs vary over some range), then the following prior distribution defined over the input parameters would be plausible:

$$\pi(\boldsymbol{\theta}) = \prod_{i=1}^{q} U(a_i, b_i)$$

where the hyper-parameters a_i and b_i are determined based on information elicited from experts or published studies (e.g., see [11]).

Algorithm 1 The computation of EVPI and Partial EVPI of the given cost function.

1: **Require:** The cost function of interest: $\mathcal{C}(T,\boldsymbol{\theta})$; the prior distribution over $\boldsymbol{\theta}$; and the set of the possible strategy options: $\{T_1,\dots,T_s\}$.

2: Using Max-min Latin hypercube, generate the design points of size n over $\mathcal{Q}$, as $(\boldsymbol{\theta_1},\boldsymbol{\theta_2},\dots,\boldsymbol{\theta_n})$.

3: By including the set of the strategy options, expand the design points to

$$\{x_l=(T_i,\boldsymbol{\theta_j}),\quad l=1,2,\dots,N=s\times n,\quad i=1,\dots,s,\quad j=1,\dots,n\}$$

4: Evaluate, $N=n\times s$ values of $\{y_l=\mathcal{C}(x_l),\ l=1,\dots,N\}$

5: Fit a TI Emulator to $\{(x_l,y_l),\ l=1,\dots,N\}$

6: Estimate: A^{-1}; $\hat{\boldsymbol{\beta}}=(\hat{\beta}_0,\hat{\beta}_1,\dots,\hat{\beta}_q,\hat{\beta}_T)$; and $\hat{\sigma^2}$;

7: For each strategy, T_i, compute

$$E_{\mathcal{C}(T_i,.)}[E_{\boldsymbol{\theta}}[\mathcal{C}(T_i,\boldsymbol{\theta})]]=\mathcal{R}_i\hat{\boldsymbol{\beta}}+\mathbf{W}_iA^{-1}(\mathbf{y}_i-H\hat{\boldsymbol{\beta}}) \tag{21}$$

8: $U_0=\max_i E_{\mathcal{C}(T_i,.)}[E_{\boldsymbol{\theta}}[\mathcal{C}(T_i,\boldsymbol{\theta})]]$

9: For each strategy, T_i, compute

$$E_{\mathcal{C}(T_i,.)}[E_{T_i,\boldsymbol{\theta}|\theta_j}[\mathcal{C}(T_i,\boldsymbol{\theta})]]=\mathcal{R}_{ij}\hat{\boldsymbol{\beta}}+\mathbf{W}_{ij}A^{-1}(\mathbf{y}_i-H\hat{\boldsymbol{\beta}}) \tag{22}$$

10: $U_{\theta_j}=\max_i E_{\mathcal{C}(T_i,.)}[E_{T_i,\boldsymbol{\theta}|\theta_j}[\mathcal{C}(T_i,\boldsymbol{\theta})]]$

11: $EVPI_{\theta_j}=E_{\theta_j}[U_{\theta_j}]-U_0$

By choosing this prior distribution, $\mathcal{R}_i$ and $\mathbf{W}_i^T$ can be analytically evaluated as follows:

$$\mathcal{R}_i=\int \mathbf{h}(\boldsymbol{\theta},T_i)^T\pi(\boldsymbol{\theta})d\boldsymbol{\theta}=\hat{\beta}_0+\sum_{l=1}^{p}\hat{\beta}_lE_\pi(\theta_l)+\hat{\beta}_TT_i,$$
$$\mathbf{W}_i^T=\int \mathbf{t}(\boldsymbol{\theta},T_i)^T\pi(\boldsymbol{\theta})d\boldsymbol{\theta},$$

where the j^{th} element of $\mathbf{W}_i^T$, associated with the j^{th} design point, is given by

$$\int \exp\{-((\boldsymbol{\theta},T_i)-(\boldsymbol{\theta}_j,T_i))^T\hat{\mathbf{R}}((\boldsymbol{\theta},T_i)-(\boldsymbol{\theta}_j,T_i))\}\pi(\boldsymbol{\theta})d\boldsymbol{\theta},$$

which can also be analytically evaluated,

$$\mathcal{R}_{ij} = \int \mathbf{h}(\boldsymbol{\theta}, T_i)^T d\pi(\boldsymbol{\theta}|\theta_j) = \hat{\beta}_0 + \hat{\beta}_j\theta_j + \sum_{l=1, l\neq j}^{p} \hat{\beta}_l E_\pi(\theta_l) + \hat{\beta}_T T_i,$$

$$\mathbf{W}_{ij}^T = \int \mathbf{t}(\boldsymbol{\theta}, T_i)^T d\pi(\boldsymbol{\theta} \mid \theta_j),$$

where the l^{th} element of $\mathbf{W}_{ij}^T$, associated with the l^{th} design point, is given by

$$\int \exp\{-((\boldsymbol{\theta}, T_i) - (\boldsymbol{\theta}_l, T_i))^T \hat{\mathbf{R}}((\boldsymbol{\theta}, T_i) - (\boldsymbol{\theta}_l, T_i))\} d\pi(\boldsymbol{\theta}|\theta_j),$$

which can be analytically evaluated, $\hat{\mathbf{R}} = diag\{\hat{r}_1, \ldots, \hat{r}_q, \hat{r}_T\}$ and $\mathbf{h}(\boldsymbol{\theta}, T)^T = (1, \boldsymbol{\theta}, T)$.

We use R and GEM-SA packages to fit the GP to the training points and then approximate the expected utilities and their corresponding uncertainty bounds. To calculate the aforementioned expected utilities, the calculations are carried out based on the discretisation of the interval $\mathcal{S}$ (maintenance decision) and the support of the joint prior distribution of the parameters $\pi(\boldsymbol{\theta})$. It is apparent that the computation of these expectation can become quite expensive by choosing a finer discretisation. The following section presents two illustrative examples. The focus here is on the way emulators can be used to perform SA based on EVPI, providing a resource efficient method for maintenance strategy identification and identifying targets for institutional learning (uncertainty reduction). In the first example we build an emulator for a TBM optimisation problem and in the second example find a robust CBM strategy for a civil structure using emulator-based SA.

5. Numerical examples

5.1. Time-based maintenance decisions model

Under the TBM policy (also known as age-based replacement), the system or component under study is in one out of two operating conditions; working or failed. System failure is identified immediately and corrective maintenance (CM) actions are undertaken to restore the system to its original condition. Regardless of the system condition, the system is renewed when it reaches a predetermined time (or age) T^*. In the TBM optimisation problem, the main challenge is to identify the optimal time to maintenance to minimise overall maintenance costs. This optimisation problem is usually defined over a finite horizon $[0, t]$, and we seek to minimise the objective cost function $\mathcal{C}(t)$ over this time interval.

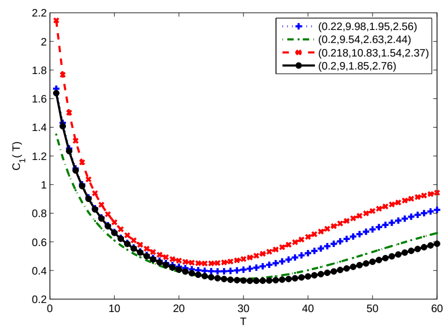

Figure 1: Total long-run average costs per unit time function for different values of $\boldsymbol{\theta} = (\gamma, \xi, \alpha, \beta)$ for Gamma-process

It can be illustrated [26, 28] that the cost per unit of time, as defined in (3) for a deteriorating component under the TBM strategy is equivalent to

$$\mathcal{C}_G(t;\boldsymbol{\theta}) = \frac{C(t|\boldsymbol{\theta})}{L(t|\boldsymbol{\theta})} = \frac{C_F[1-\mathcal{G}(\rho;\gamma t,\xi)] + C_P\mathcal{G}(\rho;\gamma t,\xi)}{\int_0^T \mathcal{G}(\rho;\gamma t,\xi)dt + \tau(\boldsymbol{\theta}_2)} \tag{23}$$

where the cumulative distribution function of system failure (due to deterioration) is represented by $\mathcal{G}(\rho;\gamma t,\xi)$ as defined in Section 2, the unexpected replacement of the component cost is denoted by C_F, each preventive maintenance action costs C_P $(0 < C_P \leq C_F)$, and $\tau(\boldsymbol{\theta}_2)$ is the expected duration of the maintenance action, and is defined by

$$\tau(\boldsymbol{\theta}_2) = \int_0^\infty t g_T(t;\boldsymbol{\theta}_2)dt \tag{24}$$

where $g_T(t;\boldsymbol{\theta}_2)$ is the time to repair (or replacement) distribution, and $\boldsymbol{\theta}_2$ is the set of repair distribution parameters. The repair distribution, $g_T(t;\boldsymbol{\theta}_2)$ is assumed to follow a Gamma distribution with α and β as shape and scale parameters respectively. A more general age-based replacement (and inspection) policy can be found in [33].

For numerical illustration, we follow [11] and set $C_F = 50$ and $C_P = 10$. Figure 1 illustrates how the expected cost rates change over the decision variable T for specific values of parameters, $\boldsymbol{\theta} = (\gamma, \xi, \alpha, \beta)$ and the given costs. It can be clearly concluded that the optimal replacement time would change by varying the parameter values. As a result, the sensitivity of the optimal maintenance strategy should be examined with respect to the changes of the input parameters to achieve robust optimised TBM decisions.

The decision-maker proposes the following prior distribution on $\boldsymbol{\theta}$

$$\pi(\boldsymbol{\theta}) = \pi_1(\gamma)\pi_2(\xi)\pi_3(\alpha)\pi_4(\beta) \tag{25}$$

where each of these parameters individually is uniformly distributed as follows

$$\gamma \sim U(0.18, 0.22),\ \xi \sim U(9, 11),\ \alpha \sim U(1, 3),\ \beta \sim U(2, 3)$$

where $U(a_1, b_1)$ denote a uniform density function defined over (a_1, b_1).

It can be shown that the cost function in (23) has a unique optimal solution (according to Theorem 1 given in [26]). When the uncertainty in input parameters $\boldsymbol{\theta}$ are included, the optimal maintenance decision will lie in the interval, $I = [25, 35]$ (see [7] for the technical details of the existence of such an interval for the considered cost rate function).

In order to lower the computational load of computing the value of information measures (EVPPIs) as the SA index, a TI emulator is fitted to the cost rate function $\mathcal{C}_R(t; \boldsymbol{\theta})$. The total training data-points to build this emulator is 1260 and selected as follows. We first generate 60 design points from $\pi(\boldsymbol{\theta})$, using the Latin hypercube design (see [19]). We then calculate the cost rate function (as a computer code) at each design point for 21 values of T (i.e., $T = 25, 25.5, 26, \ldots, 35$).

Using the fitted Gaussian process, the baseline optimal decision is derived at $T = 28.2$ where the corresponding maximum utility is $U_0 = 0.369$. So, if there is no additional information available on individual input parameters, apart from the prior information, the optimal time to maintenance is at 28.2 time units. The maximum expected net benefit (or cost saving) that a decision maker can gain by selecting the optimal maintenance time at $T = 28.2$, given no information, will be $U_0 = 0.369$ monetary unit. Further benefit can be achieved if additional information about the values of the parameters can be provided before making any decision. For example, suppose that ξ is known before making a decision. Table 1 provides the detailed information about the optimal decisions for the different values of ξ, γ, α and β. For instance, when the scale parameter, ξ, of the lifetime distribution of a component under study takes values in (9.05, 9.25), then the cost rate is minimum for $T = 32.5$, but if $\xi \in (10.25, 10.75)$, then the optimal maintenance decision is $T = 26.75$.

Range	T	Range	T
Parameter γ			
(9, 9.05)	35	(9.75, 10.25)	29.5
(9.05, 9.25)	32.5	(10.25, 10.75)	26.75
(9.25, 9.75)	29	(10.75, 11)	25
Parameter ξ			
(0.18,0.1890)	28.4	(0.1890,0.22)	28.2
Parameter α			
(1,1.35)	29.25	(2.05,2.45)	29.25
(1.35,1.55)	28.25	(2.85.2.95)	28.25
(1.55,1.95)	27.25	(2.65,2.85)	27.25
(1.95,2.05)	28	(2.45,2.55)	28
(2.95,3)	29.5		
Parameter β			
(2,2.07)	28.7	(2.69,2.81)	28.2
(2.07,2.21)	28.2	(2.81, 2.93)	28.5
(2.21,2.69)	27.9	(2.93,3)	29

Table 1: Optimal TBM decisions when a parameter of interest is known prior the maintenance decision.

The values of the EVPPIs along with the uncertainty intervals for this case are given in Table 2. By learning the values of input parameters, the decision-maker could select the maintenance time that maximises the expected utility for a particular value of the parameter of interest. For instance, if the decision maker learns about the value of "α" with the details given in Table 1, before making any decision, the expected increase in utility of learning α will be 0.3361 (in monetary unit) which is gained on the top of the situation when a decision was made based on no information (or the prior information only). The benefits that can be gained by learning α and β (the shape and scale parameters of the repair distribution) are much higher than γ and ξ. In addition, knowing α and β prior to the decision shows the most substantial differentiation between optimal strategies. Thus, these parameters are 'important' in the sense that reducing uncertainty about their values is likely to trigger selection of a different strategy.

Figure 2 summaries the SA of the cost rate function with respect to the changes of the model input parameters at $T = 28.2$. In this figure, the variance contribution of each parameter to the total variance of the cost rate at $T = 28.2$ is shown. The variance contribution of ξ, γ and α are 46%, 26% and 24% respectively based on only 60 data-points at $T = 28.2$, while β covers only 4% of total variance. In other words, this analysis exposes the behaviour of the expected cost at a specific time for different values of the parameters. Figure 3 illustrates how expected cost $E_{\boldsymbol{\theta}|\theta_i}\left[-\mathcal{C}_R(t;\boldsymbol{\theta})\right]$ when $T = 28.2$ changes with different values of the parameters (i.e., $(\eta, \delta, \alpha, \beta)$), along a 95% uncertainty bound (the thickness of the band).

θ_i	$EVPPI_i$	$C.I$
γ	0.0049	(0.0047, 0.0051)
ξ	0.0075	(0.0077, 0.0079)
α	0.3361	(0.3359, 0.3363)
β	0.3359	(0.3357, 0.3361)

Table 2: Estimated EVPPIs based on the fitted GP emulator for the parameters of the GP deterioration model for the TBM policy.

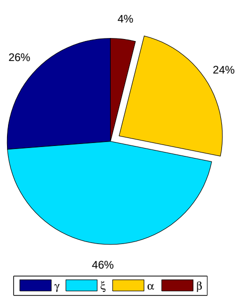

Figure 2: The variance contribution of each input parameters to the mean cost rate of the TBM policy at $T = 28.2$ for the GP deterioration model.

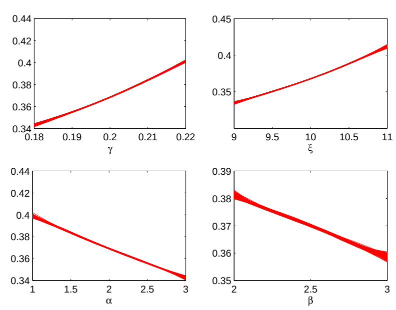

Figure 3: Expected utilities and 95% uncertainty bounds for $T = 28.2$ when the parameters are completely known before the maintenance decision.

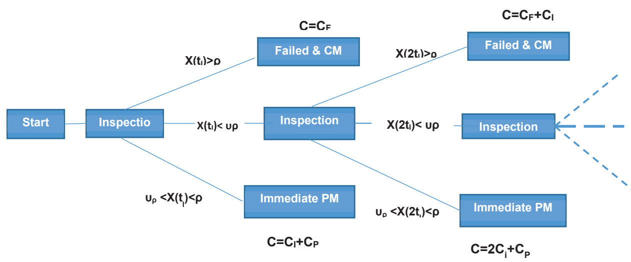

Figure 4: CBM decision tree for the GP deterioration model.

5.2. *The CBM policy under the GP deterioration model*

The inspection and replacement scenarios under the Gamma process deterioration model are more convoluted and complicated due to the temporal uncertainty (see [11]). The CBM policy under the GP deterioration model is illustrated in Figure 4 and explained as follows:

1. The system at the i^{th} inspection is at sound operating state (if $X(it_I) < \upsilon t_I$), and no action is required to be taken at this stage.
2. Immediate PM should be done (when $\upsilon\rho < X(it_I) < \rho$) to prevent any unexpected failure
3. A failure is identified at the i^{th} inspection (if $X(it_I) > \rho$), and subsequent CM is required to restore the system.

where $0 < \upsilon < 1$ is called PM ratio, and $\upsilon\rho$ is the threshold for the PM which is a fraction of the failure threshold.

According to renewal theory ([29, 26]), the mean cost rate for the CBM policy under the GP deterioration model is given by

$$\mathcal{K}_G(t_I, \upsilon; \boldsymbol{\theta}) = \frac{E[\mathcal{C}_{UG}(t_I, \upsilon; \boldsymbol{\theta})]}{E[\mathcal{L}_{DG}(t_I, \upsilon; \boldsymbol{\theta})] + \tau_r} \tag{26}$$

where the expected cost associated with a renewal cycle is given by

$$E[\mathcal{C}_{UG}(t_I, \upsilon; \boldsymbol{\theta})] = C_P + (C_F - C_P)[1 + \sum_{n=1}^{\infty} \mathcal{G}(\upsilon\rho; n\gamma t_I; \xi)] -$$

$$(C_F - C_I - C_P)[\mathcal{G}(\rho; \gamma t_I; \xi) + \sum_{n=1}^{\infty} \int_0^{\upsilon\rho} \mathfrak{g}(z; n\gamma t_I; \xi)\mathcal{G}(\rho - z; \gamma t_I; \xi)dz]$$

and the mean cycle length is as follows

$$E[\mathcal{L}_{DG}(t_I,\upsilon;\boldsymbol{\theta})] = \int_0^{t_I} \mathcal{G}(\rho;\gamma t;\xi)dt + \sum_{n=1}^{\infty}\int_0^{\upsilon\rho}\int_0^{t_I} \mathfrak{g}(z;n\gamma t_I;\xi)\mathcal{G}(\rho - z;\gamma t;\xi)dtdz]$$

where $\mathfrak{g}(z;n\gamma t_I;\xi)$ denote to gamma density function with $n\gamma t_I$ as shape and ξ as scale parameter, and $\boldsymbol{\theta} = (\gamma,\xi,\alpha,\beta)$.

The objective in the CBM policy is to find the optimal inspection time and PM ratio so that the corresponding mean cost rate becomes minimum, that is,

$$(t_I^*,\upsilon^*) = \arg\min_{(t_I,\upsilon)}\{\mathcal{K}_G(t_I,\upsilon;\boldsymbol{\theta})\}$$

As discussed in [29, 11], one can conclude that the optimal inspection time (t_I) is unique and will lie in an interval derived from the system information, failure and the characteristics of the inspection and replacement tasks. These decision variables would clearly change by varying the parameter value of $\boldsymbol{\theta}$. As a result, the sensitivity of the determined inspection time and PM ratio parameters should be examined with respect to the changes of the input parameters to achieve robust optimised CBM decisions.

The PM ratio, υ is considered as an extra parameter and included into the uncertain parameters input, that is, $\boldsymbol{\psi} = (\gamma,\xi,\alpha,\beta,\upsilon)$, where γ,ξ are respectively the shape and scale parameters of the GP deterioration model given in (2), α,β are respectively the shape and scale parameters of the maintenance distribution. The corresponding joint prior distribution is given by

$$\pi(\boldsymbol{\psi}) = \pi_1(\gamma)\pi_2(\xi)\pi_3(\alpha)\pi_4(\beta)\pi_5(\upsilon) \tag{27}$$

where

$$\gamma \sim U(0.2, 0.4),\ \xi \sim U(9, 12),\ \alpha \sim U(1, 3),\ \beta \sim U(2, 3),\ \upsilon \sim U(0.2, 0.8)$$

We first generate 80 design points generated from the joint distribution of $\boldsymbol{\psi}$ (using the Latin hypercube desin) and then evaluate the mean cost rate, $\mathcal{K}_G(t_I,\upsilon;\boldsymbol{\psi})$. An emulator based SA is implemented using this data, $\mathcal{D} = \{(\boldsymbol{\psi}^{(i)},\mathcal{K}_G(t_I^0;\boldsymbol{\psi}^i)), i = 1,\ldots,80\}$ at a fixed inspection time, $t_I^0 = 24.5$. From the variance contribution fractions of these parameters shown in Figure 5, it is evident that the PM ratio (covers 83% of the total variance) has a substantial role on determining the optimal inspection interval and minimising the maintenance costs.

Due to the importance of the PM ratio in determining the optimal inspection interval, the robustness of t_I with respect to the changes in $\boldsymbol{\theta}$ at some fixed values of υ is examined. We first let

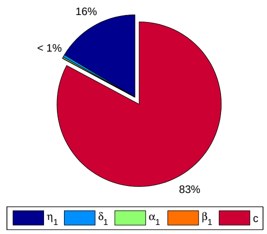

Figure 5: The variance contribution of each input parameters to the mean cost rate of the CBM policy at $t_I = 26.5$ for the GP deterioration model.

$\upsilon = 0.55$. To train the TI emulator, we generate 60 design points from the prior joint distribution of $\boldsymbol{\theta}$ over the range of the parameters given above. For each of these design points we then calculate the mean cost rate, $\mathcal{K}_G(t_I, \upsilon^{(1)}; \boldsymbol{\theta})$ for 35 values of t_I, in particular for $t_I = 18, 18.5, \ldots, 35$. The final training set is comprised of 2100 points.

The baseline optimal inspection time is derived at $t_I = 33.64$ where the corresponding maximum utility is $U_0 = 0.669$. The maximum expected net benefit (U_0) shows the decision maker's gain (in monetary unit) corresponds to the optimal inspection time at $t_I = 33.64$ which derived without any knowledge of the input parameters' values. Table 3 shows the optimal inspection interval decisions when the values of γ, ξ, α and β are learned prior to making any decision about the inspection time. For example, if the decision maker learns that $\gamma \in (0.2, 0.226)$, the baseline decision for the inspection time will not be changed. But, if it was learned that $\gamma \in (0.226, 0.234)$, the optimal inspection time should be $t_I = 31.44$.

The values of the estimated EVPPIs along with the uncertainty intervals are given in Table 4. These values illustrate the expected increase in utility of learning each input parameter before making any decision regarding the optimal inspection time. For example, if the decision maker learns about the value of γ in advance, the expected net benefit will increase to 0.145 (in monetary unit) more than the maximum expected net benefit, U_0. A similar interpretation can be made about the benefits of learning ξ, α and β based on their estimated EVPPIs given in Table 4. From these results, it can be concluded that γ (the shape parameters of the lifetime distribution) is the most important factor in the sense that knowing its value prior to making any decision would result in substantial cost savings and reduced uncertainty about the optimal inspection strategy. A similar conclusion can be derived from Figure 6 which summaries the variance fractions of each parameter to the total variance of the cost rate at $t_I = 33.64$. It also confirms that γ which covers about 92% of total variance of the mean cost rate is the most important factor affecting the maintenance cost.

As demonstrated above, the optimal inspection decision is very sensitive to υ's changes (see Figure 5). As a result, by changing υ value from 0.55 to 0.75, the derived results would be changed dramatically and this extreme behaviour at these two points is the the main reason behind selecting $\upsilon = 0.55$ and $\upsilon = 0.75$ for the SA of the cost function (and the optimal inspection strategy) with respect to the changes in parameter values. We list the possible changes of the SA when $\upsilon = 0.75$ as follows

- The optimal inspection interval, $t_I \in [25, 39]$
- The baseline optimal inspection interval is $t_I = 29.76$ (corresponding to the maximum benefit of $U_0 = 0.951$).
- Based on the computed EVPPIs of the parameters, β and γ are in order the most important factors in reducing the uncertainty about the optimal inspection interval (see Table 5).
- At the baseline decision ($t_I = 29.76$), ξ, γ and α are the most important factors affecting the maintenance costs (see Figure 7).

Table 6 shows the optimal inspection interval decisions when $\upsilon = 0.75$ and the values of γ, ξ, α and β are learned prior to making any decision about the inspection time.

Range	t_I	Range	t_I
Parameter γ			
(0.2,0.226)	33.64	(0.278, 0.29)	25.65
(0.226,0.234)	31.44	(0.290, 0.294)	25.14
(0.274,0.278)	26.16	(0.294, 0.318)	24.50
(0.234,0.254)	29.56	(0.318,0.338)	23.44
(0.254,0.27)	27.6	(0.338, 0.398)	22.06
(0.270,0.274)	26.50	(0.398,0.4)	19.70
Parameter ξ			
(9,9.03)	28.20	(9.51,9.57)	24.80
(9.03, 9.09)	27.86	(9.57, 9.63)	24.46
(9.09, 9.15)	27.52	(9.63, 9.69)	24.12
(9.15, 9.33)	26.50	(9.69, 11.31)	23.44
(9.33,9.51)	25.40	(11.31, 12)	20.44
Parameter α			
$(1, 1.10) \cup (1.30, 1.42)$	24.46	$(1.42, 1.58) \cup (1.78, 3)$	23.44
(1.10, 1.30)	25.14	(1.58, 1.78)	21.54
Parameter β			
(2,2.07)	21.54	(2.27,2.41)	25.24
$(2.07, 2.21) \cup (2.51, 2.69) \cup (2.85, 2.95)$	23.5	(2.69, 2.85)	22.76
$(2.21, 2.27) \cup (2.45, 2.51) \cup (2.95, 3)$	24.60		

Table 3: The optimal inspection interval, t_I when a parameter is known prior the maintenance decision for the CBM policy and under the GP deterioration model for $\upsilon = 0.55$.

θ_i	$EVPPI_i$	$C.I$
γ	0.145	(0.142, 0.148)
ξ	0.14	(0.137, 0.143)
α	0.1375	(0.134, 0.141)
β	0.1378	(0.1344, 0.1412)

Table 4: The estimated EVPPIs for the parameters of the GP deterioration model for the CBM policy when $\upsilon = 0.55$.

θ_i	$EVPPI_i$	$C.I$
γ	0.0128	(0.0122, 0.0133)
ξ	0.0095	(0.0089, 0.0099)
α	0.0092	(0.0087, 0.097)
β	0.0149	(0.0144, 0.0154)

Table 5: The estimated EVPPIs for the parameters of the GP deterioration model for the CBM policy when $\upsilon = 0.75$.

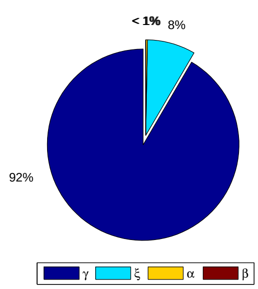

Figure 6: The variance contribution of each input parameters to the mean cost rate of the CBM policy at $t_I = 33.64$ for the GP deterioration model when $\upsilon = 0.55$.

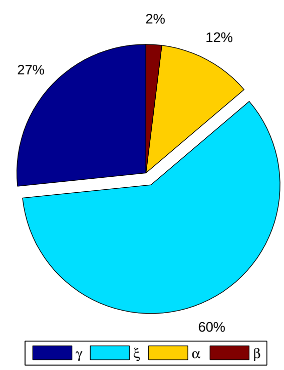

Figure 7: The variance contribution of each input parameters to the mean cost rate of the CBM policy at $t_I = 29.76$ for the GP deterioration model when $\upsilon = 0.75$.

Range	t_I	Range	t_I
Parameter γ			
(0.2, 0.3615)	30.88	(0.3615, 0.3855)	30.66
(0.3855, 0.4)	30.32		
Parameter ξ			
(9, 9.10)	36.76	(9.74, 9.90)	35.64
(9.10, 9.38)	36.48	(9.90, 10.10)	35.36
(9.38, 9.58)	36.20	(10.10, 10.98)	35.08
(9.58, 9.74)	35.92	(10.98, 12)	35.36
Parameter α			
1,1.38)	30.18	(2.14,2.26)	29.48
(2.34, 2.78)	30.32	(2.78, 2.92)	30.88
(1.38, 1.66)	39	(2.92, 3)	31.44
(1.66, 2.14)	28.64		
Parameter β			
(2, 2.07)	29.22	(2.39, 2.51)	31.44
(2.07, 2.11)	30.20	(2.51, 2.59)	30.32
(2.07, 2.15)	30.54	(2.59, 2.65)	29.48
(2.15, 2.39)	31.72	(2.65, 3)	28.36

Table 6: The optimal inspection interval, t_I when a parameter is known prior the maintenance decision for the CBM policy and under the GP deterioration model for $\upsilon = 0.75$.

6. Discussion and conclusions

In this paper we have investigated the robustness of preventive maintenance policies (TBM and CBM) as they relate to a deteriorating infrastructure system with respect to the changes of the lifetime and repair distributions' parameters using a decision-informed SA approach. The concept and application of Expected Value of Perfect Information (EVPI) have been furthered to help the decision-maker in choosing an optimised maintenance decision (critical age or inspection interval) out of the infinite set of decisions. Using this sensitivity method, analysts can examine the effect of parameter uncertainty on cost calculations, resulting in more robust maintenance decisions with respect to changes in parameter values. When planning inspections or predicting the remaining useful life of an asset, engineers must assess the benefits of the additional information that can be obtained and weigh them against the cost of these measures. The methodology developed in this paper provides an efficient framework to quantify these benefits, and possibly revise decisions based on the aggregation of the information including the system deterioration process, maintenance aspects, etc. The computation of the EVPPI requires the evaluation of multi-dimensional integrals which are often computationally exhausting. We have demonstrated how the Gaussian process emulator can be used to reduce the computational burden associated with the EVPI-based SA. In particular, we have used a Time-Input GP emulator to obtain expected utilities as continuous functions of the decision parameter (critical age or inspection interval). One of the main practical benefits of using such an emulator is that it does not restrict the decision-maker/engineer to choosing a maintenance decision from a limited number of decision options. This flexibility enables the decision maker to take maintenance decisions which are as precise as possible in the presence of parameter uncertainty which in turn would have a considerable effect on the overall cost of the maintenance strategy.

We have applied this sensitivity approach in the life-cycle management of infrastructure systems under continuous deterioration through two illustrative examples comprise both time-based (or age replacement policy) and condition-based maintenance strategies. The sensitivity results have identified the most 'important' parameters in terms of the benefit to be achieved by 'learning'. It is shown that the optimal strategy may change if a parameter becomes known prior to a maintenance decision, and this may have significant effect on the resulting cost. For instance, under the time-based maintenance example, the shape and scale parameters of the repair distribution were found to be the main influencing factors affecting the cost calculations and consequently the optimal maintenance decision. In contrast, under CBM and when $\upsilon = 0.55$, the shape and scale parameters of the lifetime distribution play the primary role in determining the cost-effective inspection strategy. Identifying important parameters in this way can provide guidance on reliability testing, monitoring or inspection. The EVPI-based SA presented here can be used for other maintenance optimisation problems including problems with imperfect maintenance ([30]), or delay-time maintenance ([31]), considered as one of the more effective preventive maintenance policies for optimising inspection planning. An efficient condition-based maintenance strategy which allows us to prevent system/-component failure by detecting the defects via an optimised inspection might be identified using the SA proposed in this paper to determine a robust optimal solution for delay-time maintenance problems and the expected related cost when the cost function parameters are either unknown or partially known.

Finally, the method articulated in this contribution might usefully be extended to calculate the EVPI measures associated with decisions at multiple points in time. In many contexts, maintenance decisions can be made at multiple points in time, at which different amounts of information from the monitoring system are available. A classic example is the monitoring and inspection of a deteriorating structure. In this situation, the EVPI measures should be computed so that the maintenance decisions could be optimized sequentially. Gramacy and Polson [32] proposed a sequential design and optimization approach for a complex system using particle learning of Gaussian process which could be very useful in computing the corresponding EVPPIs. We would encourage further developments in this field to enhance engineers' ability to make informed decisions about infrastructure maintenance and rehabilitation.

In this work, we have been concerned with computing the value of information indices and determining optimised CBM or TBM based on available information and a given cost function. In [45], it was discussed that whilst most existing autonomous condition monitoring systems provide functions for data collection they lack decision support functionality. It thus becomes crucial to understand the link between the information we have to hand and our ability to make informed decisions about asset management. The quality of information provided by the condition monitoring system is another important factor which influences the effectiveness of maintenance decisions and thus the performance of the asset [46]. For instance, the accuracy of information regarding the rate of asset degradation is critical to improving civil infrastructure life-cycle management. In order to better evaluate the accuracy of information (and the quality of the corresponding maintenance decisions) provided by condition monitoring, the value of information methodology has been used to compare the benefits offered by these techniques and the factors that affect the value delivered by them.

Autonomous condition monitoring systems (e.g., sensors) provide higher quality information in comparison to more traditional approaches such as visual inspection [40]. However, sensor location, sensitivity, and parameter recording frequency across multiple components and assets become important determinants of robust decision making. The value of information approach proposed in this paper can be used to guide more efficient information collection by identifying high information value locations for sensors and sensor arrays [47]. Furthermore, the approach might be extended to determine the timing of condition-based maintenance interventions using data from multiple sensors or time sequenced measurements from a single sensor.

It would be also interesting to extend the methodology proposed in this paper to determine the condition-based maintenance when data comes from multiple sensors or time sequenced measurements from a single sensor are combined. In this situation a data fusion should be first employed for improving condition monitoring, quality of information and system health assessment and then integrated with the condition-based maintenance system [48]. The EVPPI methodology presented in this paper can play a key role in making a decision of fusing data/features from multiple sensors which could result in improving the information quality and decision accuracy.

ACKNOWLEDGMENTS

This research was supported by funding from the UK Engineering & Physical Sciences Research Council (ICIF Project - Grant EP/K012347/1) and the European Commission (TRUST Project - Grant 265122). Additional support from the UK Engineering & Physical Sciences Research Council (Strategic Package: Centre for Predictive Modelling in Science and Engineering - Grant No. EP/L027682/1) is acknowledged.

References

[1] Dobbs,R., Pohl,H., Lin,D., Mischke, J., Garemo, N., Hexter, J., Matzinger, S., Palter, R., and Nanavatty, R. (2013). *Infrastructure productivity: How to save $1 trillion a year*, McKinsey Global Institute.

[2] Dekker, R. (1996). Applications of maintenance optimization models: a review and analysis, *Reliability Engineering & System Safety*, **51**, 229-240.

[3] Nicolai, R.P. and Dekker, R. (2007). A review of multi-component maintenance models. In: Aven, T and Vinnem, JE (Eds.), *Risk, Reliability and Societal Policy, VOLS 1-3 - VOL 1:*

Specialisation Topics; VOL 2: Thematic Topics; VOL 3: Application Topics, Taylor & Francis LTD, London, England, 289-296.

[4] Wang, H. Z. (2002). A survey of maintenance policies of deteriorating systems, *European Journal of Operational Research*, **139(3)**, 469-489.

[5] Zhang, X, and Gao, H. (2011). Determining an Optimal Maintenance Period for Infrastructure Systems, *Computer-Aided Civil and Infrastructure Engineering*, **27**, 543-554.

[6] Ahmad, R., and Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application, *Computers & Industrial Engineering*, **63**, 135149.

[7] Zitrou, A. and Bedford, T. and Daneshkhah, A. (2013). Robustness of maintenance decisions: Uncertainty modelling and value of information. *Reliability Engineering & System Safety*, **120**, 6071.

[8] Oakley, J. E. (2009).Decision-Theoretic Sensitivity Analysis for Complex Computer Models, *Technometrics*, **51(2)**, 121-129.

[9] Ellis, H., Jiang, M., and Corotis, R. (1995). Inspection, maintenance and repair with partial observability. *Journal of Infrastructure Systems*, **1(2)**, 9299.

[10] De Jonge, B., Dijkstra, A. S., and Romeijnders, W. (2015). Cost benefits of postponing time-based maintenance under lifetime distribution uncertainty, *Reliability Engineering & System Safety*, **140**, 1521.

[11] Pandey, M.D., Yuan, X.X., and Van Noortwijk, J.M. (2009). The influence of temporal uncertainty of deterioration on life-cycle management of structures, *Structure and Infrastructure Engineering*, **5(2)**, 145156.

[12] Van Noortwijk, J.M. (2009). A survey of the application of gamma processes in maintenance, *Reliability Engineering and System Safety*, **94**, 221.

[13] Dekker, R. and Plasmeijer, R. P. (2001). Multi-parameter maintenance optimisation via the marginal cost approach, *Journal of the Operational Research Society*, **52**, 188-197.

[14] Saltelli, A. (2002). Sensitivity Analysis for Importance Assessment, *Risk Analysis*, **22(3)**, 579-590.

[15] O'Hagan, A (1991). Bayes-Hermit Quadrature, *Journal of Statistical Planning & Inference*, **91**,245-260.

[16] O'Hagan, A. (2006).Bayesian analysis of computer code outputs: A tutorial, *Reliability Engineering & System Safety*, **91**, 1290-1300.

[17] Oakley, J. and O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: a Bayesian approach, *Journal of the Royal Statistical Society Series B*, **66**, 751-769.

[18] Oakley, J.E., Brennan, A., Tappenden, P., and Chilcott, J.B. (2006) Sample Sizes for Monte Carlo Partial EVPI Calculations, *Journal of Health Economics*, **29(3)**, 468-477.

[19] Sacks J, Welch, W.J., Mitchell,T.J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical Science*, **4(4)**, 40935.

[20] Kennedy, M. C. and O'Hagan, A. (2001). Bayesian calibration of Computer models (with discussion), *Journal of the Royal Statistical Society: Series B*, **63**, 425-464.

[21] Zitrou, A. and Bedford, T. and Daneshkhah, A. (2011). Robustness of maintenance decisions: uncertainty modelling and value of information. In: Grall A,Soares GD, editors. Advances in safety, reliability and risk management: ESREL2011, pp: 940-948.

[22] Conti, S. and O'Hagan, A. (2010). Bayesian emulation of complex multi-output and dynamic computer models, *Journal of Statistical Planning and Inference*, **140**, 640-651.

[23] McKay, M.D., Beckman, R.J., and Conover, W.J. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics*, **21(2)**, 239-245.

[24] Rougier, J. (2008). Efficient Emulators for Multivariate Deterministic Functions,*Journal of Computational and Graphical Statistics*, **14(4)**, 827-843.

[25] Daneshkhah, A., Bedford, T. (2013). Probabilistic sensitivity analysis of system availability using Gaussian processes, *Reliability Engineering and System Safety*, **112**, 8293.

[26] Aven, T., and Dekker, R. (1997). A useful framework for optimal replacement models, *Reliability Engineering & System Safety*, **58**, 61-67.

[27] Barlow, R., and Proschan, F. (1967). *Mathematical Theory of Reliability*, SIAM series in applied mathematics. New York: John Wiley & Sons.

[28] Barlow, R., and Proschan, F. (1965). *Mathematical Theory of Reliability*, John Wiley & Sons, New York.

[29] Park, K.S. (1988). Optimal continuous-wear limit replacement under periodic inspections. *IEEE Transactions on Reliability*, **37(1)**, 97102.

[30] Pham, H. and Wang, H. (1996). Imperfect maintenance, *European Journal of Operational Research*, **94(3)**, 425-438.

[31] Christer, A. H. and Wang, W. (1995). A delay-time-based maintenance model of a multi-component system, *IMA Journal of Mathematics Applied in Business & Industry*, **6**, 205-222.

[32] Gramacy, R.B. and Polson, N. (2011). Particle learning of Gaussian process models for sequential design and optimization. *Journal of Computational and Graphical Statistics*, **20(1)**, 102-118.

[33] Scarf, P.A., Cavalcante, C.A.V., Dwight, R.A., and Gordon, P. (2009). An age-based inspection and replacement policy for hererogenous components. *IEEE Transactions on Reliability*, **58(4)**, 641-648.

[34] Turanyi, T. and Rabitz, H. (2000) Local methods. In A. Saltelli, K. Chan and M. Scott (eds), Sensitivity Analysis. Chichester: John Wiley & Sons, Ltd.

[35] Smith, J.Q. and Daneshkhah, A. (2010). On the Robustness of Bayesian Networks to Learning from Non-conjugate Sampling. *International Journal of Approximate Reasoning*, **51**, 558 -572.

[36] Chan, K., Tarantola, S., Saltelli, A. and Sobol', I.M. (2000) Variance-based methods. In A. Saltelli, K. Chan and M. Scott (eds), *Sensitivity Analysis*. Chichester: John Wiley & Sons, Ltd.

[37] Saltelli, A. and Tarantola, S. (2002) On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal. *Journal of the American Statistical Association*, **97**, 702-709.

[38] Morris, M.D., Moore, L.M. and McKay, M.D. (2008) Using orthogonal arrays in the sensitivity analysis of computer models. *Technometrics*, **50**, 205-215.

[39] Marrel, A., Iooss, B., Laurent, B. and Roustant, O. (2009) Calculations of Sobol indices for the Gaussian process metamodel. *Reliability Engineering and System Safety*, **94**, 742-751.

[40] Srinivasan, R., and Parlikad, A. K. (2013). Value of condition monitoring in infrastructure maintenance, *Computers & Industrial Engineering*, **66**, 233-241.

[41] McMillan, D., and Ault, G. (2008). Condition monitoring benefit for onshore wind turbines: Sensitivity to operational parameters. *IET Renewable Power Generation*, **2**, 60-72.

[42] Konakli, K., and Faber, M.H. (2014). Value of information analysis in structural safety. *Proc., 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014)*, Liverpool, U.K., 1605-1614.

[43] Thns, S., and Faber, M. (2013). Assessing the value of structural health monitoring. *Proc., 11th Int. Conf. on Structural Safety and Reliability, New York.*

[44] Konakli, K., Sudret, B., and Faber, M.H. (2015). Numerical Investigations into the Value of Information in Lifecycle Analysis of Structural Systems, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, B4015007-1-13.

[45] Ouertani, M.Z and Parlikad, A.K.N., and McFarlane, D.C. (2008) Towards an approach to select an asset information management strategy. *International Journal of Computer Science and Applications*, **5**, 25-44.

[46] Borek, A., Parlikad, A.K, Woodall, P., and Tomasella, M. (2014).A risk based model for quantifying the impact of information quality, *Computers in Industry*, **65**, 354-366.

[47] Malings, C., Pozzi, M. (2016). Value of information for spatially distributed systems: Application to sensor placement, *Reliability Engineering and System Safety*, **154**, 219-233.

[48] Niu, G., Yang, B.S., Pecht, M. (2010). Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance, *Reliability Engineering and System Safety*, **95**, 786-796.

[49] Roberts, S., Osborne, M. A., Ebden, M., Reece, S., Gibson, N., and Aigrain, S. (2013). Gaussian processes for time series modelling. *Philosophical Transactions of the Royal Society (Part A)*, **371**, 1-25.

[50] Osborne, M. (2010). *Bayesian Gaussian Processes for Sequential Prediction, Optimisation and Quadrature*. A PhD thesis, The University of Oxford.