Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Probabilistic sensitivity analysis for multivariate model outputs with applications to Li-ion batteries

To cite this article: V. Triantafyllidiis *et al* 2018 *J. Phys.: Conf. Ser.* **1039** 012020

IOP Conf. Series: Journal of Physics: Conf. Series **1039** (2018) 012020 doi:10.1088/1742-6596/1039/1/012020

Probabilistic sensitivity analysis for multivariate model outputs with applications to Li-ion batteries

V. Triantafyllidiis[1], W.W. Xing[1], P.K. Leung[2], A. Rodchanarowan[3], and A.A. Shah[1]

[1]School of Engineering, University of Warwick, Coventry CV4 7AL, UK
[2]Department of Materials, University of Oxford, Oxford, OX1 3PH, UK
[3]Department of Materials Engineering, Faculty of Engineering, Kasetsart University, 50 Ngamwongwan Rd., Ladyao, Chatuchak, Bangkok, 10900, Thailand

Abstract. Full battery models are highly complex, which limits their application to tasks such as optimization and uncertainty quantification. To lower the computational burden, sensitivity analysis (SA) can be used as a precursor to identify the most important parameters in the model, but SA itself relies on a high number of full model evaluations, which has motivated the use of emulators. For high-dimensional output problems, emulators are challenging to construct. In this paper we develop a probabilistic framework for SA of high-dimensional output models using a Gaussian process emulator based on dimensionality reduction. This allows us to perform SA under uncertainty for multi-ouput problems, providing error bounds for the emulator predictions of sensitivity measures. We show how this can be achieved using Monte Carlo sampling or possibly by using semi-analytical expressions with highly efficient sampling. Moreover, we can perform SA for multivariate outputs by ranking the sensitivity measures related to (uncorrelated) coefficients in a basis for the output space.

1. Introduction

The large number of parameters appearing in mathematical models and numerical codes for batteries complicates modelling efforts. Lowering the time cost of simulations by identifying the most influential parameters and studying their effects is an effective precursor to tasks such as design optimization and uncertainty quantification (UQ). This process is referred to as *sensitivity analysis (SA)* [1]. SA methods can be categorized in different ways. In *quantitative* SA the influence of a parameter (usually referred to as "factor" of an *input*) is assigned (reproducibly) a number called a *sensitivity index* or *importance measure*. In *local* SA, the output variability is studied by perturbing an input around a nominal (base) value, while methods that attempt to measure the output variability across the entire input space are termed *global*. For small variations in the inputs local methods may be more computationally efficient. In many cases involving complex nonlinear models, however, local SA methods are inadequate.

In all but the simplest of cases, sensitivity indices must be approximated numerically based on a sample of the inputs (sampling-based SA) together with the corresponding outputs. Output variations are measured by varying one input factor at a time (as in all local methods), or by varying all input factors simultaneously (as in most global methods) using factorial or fractional factorial methods, which takes into account the joint influences of factors due to correlations, but comes at the cost of a higher number of input samples in order to achieve accurate results.

In the context of battery models, very little attention has been paid to SA. In the majority of cases, formal SA methods are not used; the model is simply run multiple times by varying factors individually and inspecting the outputs, using *ad-hoc* measures or by employing visualization tools [2]. In a small number of studies more rigorous approaches have been used, but almost invariably with highly simplified models. Applying formal SA methods to complex battery and fuel cells models is computationally burdensome and often not feasible, particularly with brute force Monte Carlo (MC) approaches. In order to overcome this issue an emulator or meta-model can be used. The emulator itself can be difficult to construct when the input and/or output space is high-dimensional (e.g., input and/or output fields). If the quantity of interest (QoI), which is derived from the output, is a scalar, an alternative is to use an emulator directly between the inputs and QoI. It may be the case, however, that there are multiple QoIs, in which case it would be ideal to emulate the output, especially when other tasks (e.g. UQ) involving different quantities, including perhaps the original output, are to be performed subsequently.

To address these issues, we develop an approach for SA of a nonlinear Li-ion battery model by employing a Gaussian process emulator based on dimensionality reduction to approximate entire charge-discharge curves. QoIs are extracted from the curves in order to perform a SA. We show how it is possible to perform an efficient probabilistic SA in the context of a variance-based approach [4], extending previous results for the scalar case considered by Oakley and O'Hagan [5] to linear functional QoIs derived from the multi-dimensional output. Lastly, using the emulation method we are able to perform a SA of multiple outputs (including in high dimensional spaces) by ranking coefficients in a low-dimensional subspace approximation of the output space.

2. Problem setup and sensitivity analysis

Suppose that the model output is $y = \eta(\mathbf{x};\boldsymbol{\xi}) \in \mathcal{F}$ for some function space $\mathcal{F}$, where $\mathbf{x}$ represents, e.g., space, time or space-time and $\boldsymbol{\xi}^T = (\xi_1, \xi_2, ..., \xi_k) \in \mathcal{X} \subset \mathbb{R}^k$ is a vector of input factors; that is, a spatial, temporal or spatio-temporal field, *parameterized* by inputs $\boldsymbol{\xi}$. The computer model (simulator), on the other hand, provides a finite-dimensional approximation of $\eta(\mathbf{x};\boldsymbol{\xi})$, e.g., at a finite number of points in a spatial and/or temporal grid or in terms of a finite basis. We may write the simulator output as a vector $\mathbf{y} \in \mathbb{R}^d$, in which d is the number of degrees of freedom, e.g., the d components of $\mathbf{y}$ represent values of $\eta(\mathbf{x};\boldsymbol{\xi})$ at d points $\mathbf{x}$. We can therefore consider the simulator as a mapping $\boldsymbol{\eta} : \mathcal{X} \to \mathcal{Y} \subset \mathbb{R}^d$ between a feasible input space $\mathcal{X} \subset \mathbb{R}^k$ and an output space $\mathcal{Y}$, i.e., $\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\xi})$.

Let Q be a scalar QoI that is derived from y *via* a *linear* functional $G : \mathcal{F} \to \mathcal{Q}' \subset \mathbb{R}$. We can instead consider Q as a mapping $F = (G \circ \eta)(\boldsymbol{\xi}) : \boldsymbol{\xi} \mapsto Q$ directly between $\mathcal{X}$ and $\mathcal{Q}$, i.e., $Q = F(\boldsymbol{\xi}) = G(\eta(\mathbf{x};\boldsymbol{\xi}))$. In reality, we have an approximation $q = f(\boldsymbol{\xi})$ of Q, where the mapping $f : \mathcal{X} \to \mathcal{Q}$ is derived from a linear functional $g : \mathbb{R}^d \to \mathcal{Q} \subset \mathbb{R}$ that acts on the simulator outputs $\mathbf{y}$, that is $f(\boldsymbol{\xi}) = (g \circ \boldsymbol{\eta})(\boldsymbol{\xi}) = g(\boldsymbol{\eta}(\boldsymbol{\xi}))$. We develop a SA framework in which the outputs $\mathbf{y}$ are estimated by an emulator. We first describe the SA methods employed and the construction of the emulator. We then show how it is possible to perform SA on QoIs derived from the emulator output, before presenting a probabilistic SA analysis and a method for multi-output SA.

One approach to SA measures the sensitivity of the QoI by estimating the derivative $\partial_{\xi_i} f$ around a nominal (base) point $\overline{\boldsymbol{\xi}}$ using finite differences [6]. The main weakness of such local methods is that they provide no information on how the sensitivity to a given ξ_i depends on the values of the other factors. In order to extend this method to a global analysis we may use multiple base points $\{\overline{\boldsymbol{\xi}}_j\}_{j=1}^r$, which leads to the class of *elementary effect tests* (EETs). Suppose that $\mathcal{X}$ is the unit hypercube, and each direction ξ_i is discretized into p levels (points). The *elementary effect* of ξ_i at $\overline{\boldsymbol{\xi}}_j = (\overline{\xi}_{j,1}, \ldots, \overline{\xi}_{j,k})^T$ is: ‘

$$EE_i^{(j)} \equiv \left[f(\overline{\xi}_{j,1}, \ldots, \overline{\xi}_{j,i} + \delta_i, \ldots, \overline{\xi}_{j,k}) - f(\overline{\xi}_{j,1}, \ldots, \overline{\xi}_{j,i}, \ldots, \overline{\xi}_{j,k}) \right] / \Delta_i, \tag{1}$$

where $\Delta_i \in \{1/(p-1), \ldots, 1-1/(p-1)\}$. Typically, $\Delta_i = \Delta$, $\forall i$. Morris [3] proposed two sensitivities measures for each ξ_i, namely the mean and the standard deviation of the finite distribution F_i over $\{EE_i^{(j)}\}_{j=1}^r$ (or $|EE_i^{(j)}|$). The mean μ_i of EE_i measures the influence of ξ_i, while the standard deviation σ_i measures the degree of interaction with the other factors.

To calculate the statistics for each elementary effect, i.e., μ_i and σ_i, we can randomly choose M base points $\{\bar{\boldsymbol{\xi}}_j\}_{j=1}^M$, then construct so-called *trajectories* in $\mathcal{X}$ of $k+1$ points $\bar{\boldsymbol{\xi}}_j \cup \{\boldsymbol{\xi}_{j,n}\}_{n=1}^k$ for each j. Setting $\boldsymbol{\xi}_{j,0} = \bar{\boldsymbol{\xi}}_j$, the trajectory point $\boldsymbol{\xi}_{j,n}$ is obtained by perturbing a randomly chosen factor of $\boldsymbol{\xi}_{j,n-1}$ by $\pm\Delta$ until all factors have been perturbed, but with the property that $\boldsymbol{\xi}_{j,n}$ and $\boldsymbol{\xi}_{j,n-1}$ differ in only one factor. The model is run at every point in each of the trajectories (a total of $M(k+1)$) to obtain $EE_{i,j}$, $i = 1, \ldots k$, $j = 1, \ldots M$, yielding estimates of μ_i and σ_i.

A more sophisticated approach to SA, embedded in probability, involves treating the inputs as stochastic variables, which leads to a distribution over the QoI [4, 6]. A variance based first-order effect of each input factor is given by $\mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i])$, where $\mathbb{E}[\cdot]$ and $\mathbb{V}\mathrm{ar}(\cdot)$ denote expectation and variance operators with respect to the distribution over a subscripted random variable; or with respect to the distribution $p(\boldsymbol{\xi})$ over $\boldsymbol{\xi}$ if no subscript is present, i.e. $\mathbb{E}[\cdot] \equiv \mathbb{E}_{\boldsymbol{\xi}}[\cdot]$. The quantity $\boldsymbol{\xi}_{\sim i}$ is the vector of all inputs factors *excluding* ξ_i (and similarly for multiple indices). The *first order sensitivity index* (or main effect index) for the input ξ_i is defined as $S_i \equiv \mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i])/\mathbb{V}\mathrm{ar}(q)$, which measures the contribution of the main effect of ξ_i to the total QoI variance. Another measure of sensitivity, defined below, is the *total effect index*, which incorporates interactions between the factors ξ_i:

$$S_{T_i} \equiv \mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[\mathbb{V}\mathrm{ar}_{\xi_i}(q|\boldsymbol{\xi}_{\sim i})]/\mathbb{V}\mathrm{ar}(q) = 1 - \mathbb{V}\mathrm{ar}_{\boldsymbol{\xi}_{\sim i}}(\mathbb{E}_{\xi_i}[q|\boldsymbol{\xi}_{\sim i}])/\mathbb{V}\mathrm{ar}(q). \tag{2}$$

The variance-based SA framework can be couched in terms of the decomposition of the variance of q. Suppose $q = f(\boldsymbol{\xi}) \in L^2(\mathcal{X})$ (square integrable functions defined on $\mathcal{X}$) and $\mathcal{X}$ is (without loss of generality) a unit hypercube $\mathcal{X} = \{\boldsymbol{\xi} | 0 \le \xi_i \le 1; i = 1, \ldots, k\}$. We also assume that the factors are independently and uniformly distributed within $\mathcal{X}$, so that the probability density functions satisfy $p(\xi_{i_1}, \ldots, \xi_{i_l}) = \mathbf{1}_{[0,1]^l}$ for $\{i_1, \ldots, i_l\} \subset \{1, \ldots, k\}$, where $\mathbf{1}_A$ is the indicator function on a set A. The expectation operators $\mathbb{E}_{\boldsymbol{\xi}_{\sim i_1 \ldots i_l}}[\cdot]$ are then unweighted integrals over $\xi_{i_1}, \ldots, \xi_{i_l}$. The function $f(\boldsymbol{\xi})$ can be decomposed in the following way (Hoeffding [7]):

$$f(\boldsymbol{\xi}) = f_0 + \sum_{i=1}^{k} f_i(\xi_i) + \sum_{i=1}^{k} \sum_{j=i+1}^{k} f_{ij}(\xi_i, \xi_j) + \ldots + f_{1\ldots k}(\xi_1, \ldots, \xi_k), \tag{3}$$

where f_0 is a constant, $f_i(\xi_i)$ (the main effect of ξ_i) is a function only of ξ_i, $f_{ij}(\xi_i, \xi_j)$ (the interaction) is a function only of ξ_i and ξ_j, and so on. The condition $\int_0^1 f_{i_1 i_2 \ldots i_s}(\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_s}) d\xi_{i_w} = 0$ is imposed for $1 \le i_1 < i_2 < \ldots < i_s \le k$ and $i_w \in \{i_1, i_2, \ldots, i_s\}$ [7]. Thus, the summands are orthogonal, in the sense that:

$$\int_{\mathcal{X}} f_{i_1 i_2 \ldots i_m}(\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_m}) f_{i'_1 i'_2 \ldots i'_n}(\xi_{i'_1}, \xi_{i'_2}, \ldots, \xi_{i'_n}) d\boldsymbol{\xi} = 0, \tag{4}$$

for $\{i_1, i_2, \ldots, i_m\} \neq \{i'_1, i'_2, \ldots, i'_n\}$. Moreover, $f_0 = \mathbb{E}[q] = \int_{\mathcal{X}} f(\boldsymbol{\xi}) d\boldsymbol{\xi}$, $f_i = \mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i] - f_0$, $f_{ij} = \mathbb{E}_{\boldsymbol{\xi}_{\sim ij}}[q|\xi_i, \xi_j] - f_i - f_j - f_0$, etc.. By squaring and integrating Eq. (3) and using the orthogonality property, we obtain a *decomposition of the total variance* $V = \mathbb{V}\mathrm{ar}(q) = \int_{\mathcal{X}} f^2(\boldsymbol{\xi}) d\boldsymbol{\xi} - f_0^2$, that is, $V = \sum_{i=1}^k V_i + \sum_{i=1}^k \sum_{j=i+1}^k V_{ij} + \ldots + V_{1\ldots k}$, in which $V_i = \mathbb{V}\mathrm{ar}_{\xi_i}(f_i(\xi_i)) = \mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i])$, $V_{ij} = \mathbb{V}\mathrm{ar}_{\xi_i \xi_j}(f_{ij}(\xi_i, \xi_j)) = \mathbb{V}\mathrm{ar}_{\xi_i \xi_j}(\mathbb{E}_{\boldsymbol{\xi}_{\sim ij}}[q|\xi_i, \xi_j]) - \mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]) - \mathbb{V}\mathrm{ar}_{\xi_j}(\mathbb{E}_{\boldsymbol{\xi}_{\sim j}}[q|\xi_j])$, and so on. The terms $V_{i_1 \ldots i_l}$, $l \le k$ are called partial variances and it is obvious that the main effect indices S_i are simply the partial variances normalized by the total variance. We can also

define higher order sensitivity indices by normalizing the $V_{i_1\ldots i_l}$, e.g. the second-order index $S_{ij} = V_{ij}/V$, which measures the effect of interactions between ξ_i and ξ_j on q.

The main and total indices can be computed using a quasi MC method [1] by first generating a matrix $\mathbf{X} = [\xi_{i,j}]$, $i = 1, \ldots, 2k$, $j = 1, \ldots, N$, of N points in the $2k$ hypercube, using a low-discrepancy sequence such as a Latin hypercube. This is done according to the distribution $p(\boldsymbol{\xi})$ over the factors, e.g., $p(\boldsymbol{\xi}) = \mathbf{1}_{[0,1]^k}$ for independent $\xi_i \sim \mathcal{U}[0,1]$. $\mathbf{X}$ is then partitioned into a matrix $\mathbf{A} \in \mathbb{R}^{N\times k}$ consisting of the first k columns and a matrix $\mathbf{B} \in \mathbb{R}^{N\times k}$ consisting of the remaining k columns. This provides two independent sets of N samples in the k hypercube. A third matrix $\mathbf{C}_i$ consists of the columns of matrix $\mathbf{B}$ except the i-th column, which is set to the i-th column of $\mathbf{A}$. The next step is to compute the QoI q by running the model at the selected inputs contained in the sample matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}_i$ to yields vectors $\mathbf{q_A} = f(\mathbf{A})$, $\mathbf{q_B} = f(\mathbf{B})$ and $\mathbf{q_{C_i}} = f(\mathbf{C}_i)$ ($f(\mathbf{A})$ is used to denote vectorized q values from the set consisting of the rows of $\mathbf{A}$). The indices S_i and S_{T_i} are then estimated as follows:

$$S_i = \frac{(1/N)\sum_{j=1}^N q_{\mathbf{A},j} q_{\mathbf{C}_i,j} - f_0^2}{(1/N)\sum_{j=1}^N q_{\mathbf{A},j}^2 - f_0^2}, \quad S_{T_i} = \frac{(1/N)\sum_{j=1}^N q_{\mathbf{B},j} q_{\mathbf{C}_i,j} - f_0^2}{(1/N)\sum_{j=1}^N q_{\mathbf{A},j}^2 - f_0^2}, \tag{5}$$

where $q_{\mathbf{A},j}$ is the j-th coordinate of $\mathbf{q_A}$ (etc.) and $f_0 = (1/N)\sum_{j=1}^N q_{\mathbf{A},j}$ is the sample mean. This procedure is repeated for each $i = 1, \ldots, k$. The first of Eqs. (5) follows from the basic definition $\mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]) = \int_{[0,1]} \mathbb{E}^2_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]d\xi_i - (\int_{[0,1]} \mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]d\xi_i)^2$. The last term is $\mathbb{E}^2[q] = f_0^2$, while the first term can be written as: $\int_{[0,1]} \mathbb{E}^2_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]d\xi_i = \int_{[0,1]^k}\int_{[0,1]^{k-1}} f(\xi_1, \ldots, \xi_k) \times f(x'_1, \ldots, x'_k)d\boldsymbol{\xi}d\boldsymbol{\xi}'_{\sim i}$, i.e., the expectation over $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'_{\sim i}$ of $f(\xi_1, \ldots, \xi_k) \times f(x'_1, \ldots, x'_k)$, which explains the MC estimate in Eq. (5). A similar explanation can be given for the S_{T_i} estimate.

The cost of this procedure is $2N$ runs of the model to generate the matrices $\mathbf{A}$ and $\mathbf{B}$, and an additional Nk runs to obtain the QoIs corresponding to $\mathbf{C}_i$. This give a total of $N(k+2)$, which is much lower than the cost of brute-force MC estimates of $\mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i])$ and $\mathbb{V}\mathrm{ar}_{\boldsymbol{\xi}_{\sim i}}(\mathbb{E}_{\xi_i}[q|\boldsymbol{\xi}_{\sim i}])$. The former, e.g., would require $O(N)$ runs ($N \gg k$) to estimate the inner expectation for a fixed ξ_i, which we would be repeated $O(N)$ times to estimate the outer variance, leading to $O(N^2)$ runs for each i.

3. Gaussian process emulation of the model outputs

Suppose we are given *training points* $\{\mathbf{y}_j\}_{j=1}^m \subset \mathcal{Y}$, which are values of $\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\xi})$ at the *design points* $\{\boldsymbol{\xi}_j\}_{j=1}^m$. Without loss of generality we mean centre the training points: $\mathbf{y}_j \mapsto \mathbf{y}_j - \overline{\mathbf{y}}$, where $\overline{\mathbf{y}} = \sum_{k=1}^m \mathbf{y}_k$. Assume that $\mathcal{Y}$ is a low-dimensional linear subspace of $\mathbb{R}^d$. We derive an approximate basis for $\mathcal{Y}$ using principal component analysis (PCA) [8], i.e., we find a linear transformation $\mathbf{w}(\boldsymbol{\xi}) = \mathbf{V}^T\mathbf{y}$, in which $\mathbf{V} \in \mathbb{R}^{d\times d}$ has orthogonal columns $\mathbf{v}_i$ (a basis for $\mathbb{R}^d$) and the uncorrelated components $w_i(\boldsymbol{\xi})$ of $\mathbf{w}(\boldsymbol{\xi})$ have decreasing variance with i.

Let $\boldsymbol{\Sigma} = \mathbb{E}[\mathbf{y}\mathbf{y}^T]$ be the symmetric and positive definite variance-covariance matrix. The eigenvalue problem $\boldsymbol{\Sigma}\mathbf{v} = \lambda\mathbf{v}$ yields the $\mathbf{v}_i$ and corresponding positive eigenvalues $\lambda_1 > \cdots > \lambda_d$. The components of a point in this basis satisfy $\mathbb{V}\mathrm{ar}[w_i(\boldsymbol{\xi})] = \lambda_i$ and $\mathbb{E}[w_i(\boldsymbol{\xi})w_j(\boldsymbol{\xi})] = 0$ for $i \neq j$. Any point $\mathbf{y} \in \mathcal{Y}$ can be written in the form $\mathbf{y} = \mathbf{V}\mathbf{w}(\boldsymbol{\xi}) = \sum_{i=1}^d w_i(\boldsymbol{\xi})\mathbf{v}_i = \sum_{i=1}^d (\mathbf{v}_i^T\mathbf{y})\mathbf{v}_i$ and an r-dimensional approximation $\mathbf{y}_r \in \mathcal{Y}_r = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ of $\mathbf{y}$ is given by $\mathbf{y}_r = \mathbf{V}_r\mathbf{w}_r(\boldsymbol{\xi}) = \sum_{i=1}^r w_i(\boldsymbol{\xi})\mathbf{v}_i$, where $\mathbf{V}_r = [\mathbf{v}_1 \ldots \mathbf{v}_r]$ and $\mathbf{w}_r(\boldsymbol{\xi}) = (w_1(\boldsymbol{\xi}), \ldots, w_r(\boldsymbol{\xi}))^T$. It can be demonstrated [8] that $\mathbb{E}\left[\|\mathbf{y} - \mathbf{y}_r\|^2\right] = \sum_{i=r+1}^d \lambda_i$, from which a value of r can be selected based on a chosen tolerance. $\boldsymbol{\Sigma}$ is approximated by the sample covariance matrix $\boldsymbol{\Sigma}_S = (1/m)\mathbf{Y}\mathbf{Y}^T$, where $\mathbf{Y} = [\mathbf{y}_1 \ldots \mathbf{y}_m]$. The coefficients $w_i(\boldsymbol{\xi}) = \mathbf{v}_i^T\mathbf{y}$ of a point $\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\xi})$ are assumed to be realizations of uncorrelated (therefore mutually independent) GPs. We emulate these coefficients for test inputs $\boldsymbol{\xi}$ to find low-dimensional approximations $\mathbf{y}_r = \boldsymbol{\eta}_r(\boldsymbol{\xi}) \equiv \mathbf{V}_r\mathbf{w}_r(\boldsymbol{\xi}) \in \mathcal{Y}_r$ [9].

Let us focus on $w_i(\boldsymbol{\xi})$ for some $i \in \{1,\dots,r\}$. We wish to approximate $w_i(\boldsymbol{\xi}) : \mathcal{X} \to \mathbb{R}$ given values of this function (from a PCA) at design points $\{\boldsymbol{\xi}_j\}_{j=1}^m$. In GPR, we place a GP *prior distribution* indexed by $\boldsymbol{\xi} \in \mathcal{X}$ over $w_i(\boldsymbol{\xi})$. The distribution $p(w_i(\boldsymbol{\xi}_1),\dots,w_i(\boldsymbol{\xi}_m))$ for an arbitrary finite collection of indices $\{\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_m\}$ is Gaussian. The GP prior is $w_i(\boldsymbol{\xi})|\boldsymbol{\theta}_i \sim \mathcal{GP}(0, c(\boldsymbol{\xi},\boldsymbol{\xi}';\boldsymbol{\theta}_i))$, which has a zero mean function (we set $w_i(\boldsymbol{\xi}_j) \mapsto w_i(\boldsymbol{\xi}_j) - \overline{w}_i$, where $\overline{w}_i = (1/m)\sum_j w_i(\boldsymbol{\xi}_j)$) and a covariance function $c(\boldsymbol{\xi},\boldsymbol{\xi}';\boldsymbol{\theta}_i)$, dependent upon *hyperparameters* $\boldsymbol{\theta}_i$. We employ a square-exponential function for all i: $c(\boldsymbol{\xi},\boldsymbol{\xi}';\boldsymbol{\theta}_i) = \theta_0 \exp\{-(\boldsymbol{\xi}-\boldsymbol{\xi}')^T \mathrm{diag}(\theta_{i,1},\dots,\theta_{i,k})(\boldsymbol{\xi}-\boldsymbol{\xi}')\}$, where $\boldsymbol{\theta}_i = (\theta_{i,0},\dots,\theta_{i,k})^T$, in which $\theta_{i,1},\dots,\theta_{i,k}$ are the inverse square correlation lengths.

The distribution of $\mathbf{d}_i \equiv (w_i(\boldsymbol{\xi}_1),\dots,w_i(\boldsymbol{\xi}_m))^T$ given $\boldsymbol{\theta}_i$ (i.e., the likelihood) is $p(\mathbf{d}_i|\boldsymbol{\theta}_i) = \mathcal{N}(\mathbf{0},\mathbf{C}_i)$, with covariance matrix $\mathbf{C}_i = [c(\boldsymbol{\xi}_i,\boldsymbol{\xi}_j;\boldsymbol{\theta}_i)]_{i,j=1}^m$. The predictive distribution at new inputs $\boldsymbol{\xi} \in \mathcal{X}$ is obtained from the joint distribution $p(w_i(\boldsymbol{\xi}),\mathbf{d}_i|\boldsymbol{\theta}_i)$ by conditioning on $\mathbf{d}_i$ [10]:

$$\begin{gathered} w_i(\boldsymbol{\xi})|\mathbf{d}_i,\boldsymbol{\theta}_i \sim \mathcal{GP}\left(m_i'(\boldsymbol{\xi}), c_i'(\boldsymbol{\xi},\boldsymbol{\xi}')\right), \\ m_i'(\boldsymbol{\xi}) = \mathbf{c}_i(\boldsymbol{\xi})^T\mathbf{C}_i^{-1}\mathbf{d}_i + \overline{w}_i, \quad c_i'(\boldsymbol{\xi},\boldsymbol{\xi}') = c(\boldsymbol{\xi},\boldsymbol{\xi}';\boldsymbol{\theta}_i) - \mathbf{c}_i(\boldsymbol{\xi})^T\mathbf{C}_i^{-1}\mathbf{c}(\boldsymbol{\xi}'), \end{gathered} \tag{6}$$

in which $\mathbf{c}_i(\boldsymbol{\xi}) = (c(\boldsymbol{\xi}_1,\boldsymbol{\xi};\boldsymbol{\theta}_i),\dots,c(\boldsymbol{\xi}_m,\boldsymbol{\xi};\boldsymbol{\theta}_i))^T$. The hyperparameters can be specified by point estimates [10] such as the maximum log likelihood estimate (MLE): $\boldsymbol{\theta}_{i,MLE} = \arg\max_{\boldsymbol{\theta}_i}(-\ln|\mathbf{C}_i|/2 - \mathbf{d}_i^T\mathbf{C}_i^{-1}\mathbf{d}_i/2)$. This procedure is repeated for each $i = 1,\dots,r$ to obtain $\mathbf{w}_r(\boldsymbol{\xi}) = (w_1(\boldsymbol{\xi}),\dots,w_r(\boldsymbol{\xi}))^T$. Using MLE estimates, we obtain $\mathbb{E}[\mathbf{w}_r(\boldsymbol{\xi})] = \mathbf{m}'(\boldsymbol{\xi}) \equiv (m_1'(\boldsymbol{\xi}),\dots,m_r'(\boldsymbol{\xi}))^T$. The predicted variance of each coefficient is $\mathbb{V}\mathrm{ar}(w_i(\boldsymbol{\xi})) = c_i'(\boldsymbol{\xi},\boldsymbol{\xi})$. The model outputs are therefore distributed as follows (noting that $\mathbb{C}\mathrm{ov}\,(w_i(\boldsymbol{\xi}),w_j(\boldsymbol{\xi})) = 0$ for $i \neq j$):

$$\begin{aligned} \mathbf{y}_r - \overline{\mathbf{y}} = \boldsymbol{\eta}_r(\boldsymbol{\xi}) &= \mathbf{V}_r\mathbf{w}_r(\boldsymbol{\xi}) \sim \mathcal{GP}\left(\mathbf{m}_{\mathbf{y}_r}, \mathbf{c}_{\mathbf{y}_r}(\boldsymbol{\xi},\boldsymbol{\xi}')\right) \\ \mathbf{m}_{\mathbf{y}_r}(\boldsymbol{\xi}) &= \mathbb{E}[\boldsymbol{\eta}_r(\boldsymbol{\xi})] = \mathbf{V}_r\mathbf{m}'(\boldsymbol{\xi}), \\ \mathbf{c}_{\mathbf{y}_r}(\boldsymbol{\xi},\boldsymbol{\xi}') &= \mathbb{C}\mathrm{ov}\left(\boldsymbol{\eta}_r(\boldsymbol{\xi}),\boldsymbol{\eta}_r(\boldsymbol{\xi}')\right) = \mathbf{V}_r\mathrm{diag}(c_1'(\boldsymbol{\xi},\boldsymbol{\xi}'),\dots,c_r'(\boldsymbol{\xi},\boldsymbol{\xi}'))\mathbf{V}_r^T, \end{aligned} \tag{7}$$

3.1. Probabilistic and multivariate sensitivity analysis

The emulation method described above extends Bayesian GP modelling to multiple outputs in a probabilistic manner, furnishing an explicit distribution over the output (Eqs. (7)). We can always extract estimates of the statistics of sensitivity measures using full MC sampling. Take for example the main effect index $S_i = \mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i])/\mathbb{V}\mathrm{ar}(q)$. The expected value and variance of a quantity with respect to the distribution over $\boldsymbol{\eta}_r$ are denoted $\mathbb{E}_{\boldsymbol{\eta}_r}[\cdot]$ and $\mathbb{V}\mathrm{ar}_{\boldsymbol{\eta}_r}(\cdot)$, respectively. Since $q = f(\boldsymbol{\xi}) = g(\boldsymbol{\eta}(\boldsymbol{\xi}))$, a MC estimate of $\mathbb{E}_{\boldsymbol{\eta}_r}[S_i]$ is given by:

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\eta}_r}[S_i] = \mathbb{E}_{\boldsymbol{\eta}_r}\left[\frac{\mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[g(\boldsymbol{\eta}(\boldsymbol{\xi}))|\xi_i])}{\mathbb{V}\mathrm{ar}(g(\boldsymbol{\eta}(\boldsymbol{\xi}))}\right] &= \mathbb{E}_{\boldsymbol{\eta}_r}\left[\frac{\mathbb{E}_{\xi_i}\left[\mathbb{E}^2_{\boldsymbol{\xi}_{\sim i}}[g(\boldsymbol{\eta}(\boldsymbol{\xi}))|\xi_i])\right] - \mathbb{E}^2\left[g(\boldsymbol{\eta}(\boldsymbol{\xi}))\right]}{\mathbb{E}\left[g(\boldsymbol{\eta}(\boldsymbol{\xi}))^2\right] - \mathbb{E}^2\left[g(\boldsymbol{\eta}(\boldsymbol{\xi}))\right]}\right] \\ &\approx \frac{1}{J}\sum_{j=1}^J \frac{N^{-3}\sum_{l=1}^N\left(\sum_{n=1}^N g(\boldsymbol{\eta}^{(j)}(\boldsymbol{\xi}_{\sim i}^{(n)},\xi_i^{(l)}))\right)^2 - N^{-2}\left(\sum_{n=1}^N g(\boldsymbol{\eta}^{(j)}(\boldsymbol{\xi}^{(n)})\right)^2}{N^{-1}\sum_{n=1}^N g(\boldsymbol{\eta}^{(j)}(\boldsymbol{\xi}^{(n)})^2 - N^{-2}\left(\sum_{n=1}^N g(\boldsymbol{\eta}^{(j)}(\boldsymbol{\xi}^{(n)})\right)^2} \end{aligned} \tag{8}$$

where $\boldsymbol{\eta}_r^{(j)}$ is drawn from $p(\boldsymbol{\eta}_r) = \mathcal{GP}\left(\mathbf{m}_{\mathbf{y}_r}(\boldsymbol{\xi}), \mathbf{c}_{\mathbf{y}_r}(\boldsymbol{\xi},\boldsymbol{\xi}')\right)$ and the $\xi_i \sim \mathcal{U}[0,1]$ are independent. The notation $\boldsymbol{\eta}^{(j)}(\boldsymbol{\xi}_{\sim i}^{(n)},\xi_i^{(l)})$ means that $\boldsymbol{\eta}^{(j)}(\boldsymbol{\xi})$ is evaluated at $\xi_i = \xi_i^{(l)}$, $\boldsymbol{\xi}_{\sim i} = \boldsymbol{\xi}_{\sim i}^{(n)}$ for some $i \in \{1,\dots,k\}$. The generic number N in Eq. (8) need not be the same for all MC estimates.

In this MC procedure we interpret $S_i \equiv S_i(\boldsymbol{\eta})$ as a random function of the random vector $\boldsymbol{\eta}$ and obtain samples $\boldsymbol{\eta}^{(j)}(\boldsymbol{\xi})$ (deterministic functions of $\boldsymbol{\xi}$) from $p(\boldsymbol{\eta}_r)$ to approximate the outer integral. We then sample from $\boldsymbol{\xi}_{\sim i} \sim \mathcal{U}[0,1]^{k-1}$ and $\xi_i \sim \mathcal{U}[0,1]$ to approximate the inner integrals of $g(\boldsymbol{\eta}(\boldsymbol{\xi})) = f(\boldsymbol{\xi})$, which is a random function of $\boldsymbol{\xi}$. In practice, we generate the

samples of $\boldsymbol{\xi}_{\sim i}$ and ξ_i, and then sample from the distributions over $w_i(\boldsymbol{\xi})$, $i = 1, \ldots, r$, to obtain partial realizations (at the sampled values of $\boldsymbol{\xi}$) of $\mathbf{w}_r(\boldsymbol{\xi})$, from which we can obtain (partial) realizations of $\boldsymbol{\eta}_r(\boldsymbol{\xi}) = \mathbf{V}_r\mathbf{w}_r(\boldsymbol{\xi})$. In fact, the last step is not necessary since we can work directly with $\mathbf{w}_r(\boldsymbol{\xi})$ to obtain realizations of the QoI $q = g(\boldsymbol{\eta}(\boldsymbol{\xi}))$, i.e., $g(\boldsymbol{\eta}(\boldsymbol{\xi})) = g(\mathbf{V}_r\mathbf{w}_r(\boldsymbol{\xi}))$. $\mathbb{V}\mathrm{ar}_{\boldsymbol{\eta}_r}(S_i)$ is estimated in the same way and the same procedure can be used for S_{T_i} or any other measure.

In the scalar output case (the output being the QoI), Oakley and O'Hagan derived semi-analytical expressions for estimating the expectations $\mathbb{E}_{\boldsymbol{\eta}_r}[\cdot]$ and possibly variances $\mathbb{V}\mathrm{ar}_{\boldsymbol{\eta}_r}(\cdot)$ of several sensitivity measures using only a very small number of MC runs (e.g. $O(1)$ *vs.* $O(N)$ for $\mathbb{E}_{\boldsymbol{\eta}_r}[S_i]$ as required in full MC to estimate $\mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[g(\boldsymbol{\eta}(\boldsymbol{\xi}))|\xi_i]))$ [5]. Equivalent semi-analytical expressions can be established for certain types of scalar QoIs derived from the multivariate output emulator used in this paper, namely QoIs arising from a linear functional of the output. These expressions are derived in the Appendix. Another feature of this method is that we could investigate the *sensitivity of multivariate outputs under uncertainty* to the inputs by *separately* (due to their independence) ranking the sensitivity indices for the coefficients $w_i(\boldsymbol{\xi})$, $i = 1, \ldots, r$, using the procedures described above, i.e., setting $q = w_i(\boldsymbol{\xi})$. Since the $w_i(\boldsymbol{\xi})$ are scalar GPs, the probabilistic analysis of Oakley and O'Hagan [5] is directly applicable.

4. Li-on battery model

We consider a Li-ion battery comprised of a $LiMn_2O_4$ positive electrode and a graphite Li_xC_6 porous negative electrode. The electrolyte consists of a non-aqueous carbonate solvent mixture and a lithium salt $LiPF_6$ in a 1:1 mixture of ethylene carbonate and diethyl carbonate dispersed in an inert polymer matrix. The domain is 1-d (direction x) and the positive and negative current collectors are located at $x = 0$ and $x = L$, respectively. Intercalation of Li is described by a mass balance with diffusion in a pseudo dimension R (into spherical particles) [11]. The solid Li concentration in the positive (negative) electrode c_s^p (c_s^n) is given by:

$$\partial_t c_j^s = R^{-2}\partial_R(R^2 D_j^s \partial_R c_j^s), \tag{9}$$

where D_j^s is the diffusion coefficient of Li in the active material. Here and below, $j = p$ for the positive electrode, $j = n$ for the negative electrode and $j = s$ for the separator. The boundary conditions are $\partial_R c_j^s|_{R=0} = 0$ and $-D_j^s \partial_R c_j^s|_{R=R_p} = (1/aF)\partial_x i_2$, where R_p is the particle radius, a is the specific active surface area and i_2 is the current density in the electrolyte:

$$i_2 = -\kappa_2 \partial_x \phi_2 + \kappa_2 R_U T F^{-1}(1 - t_+^0)(1 + \partial_{\ln c} \ln f_A)\partial_x \ln c, \tag{10}$$

where κ_2 is the effective ionic conductivity, T is the temperature, F is Faraday's constant, R_U is the universal gas constant, ϕ_2 is the electrolyte potential, f_A is the mean molar activity coefficient of the electrolyte, c is the lithium ion (Li^+) concentration and t_+^0 is the transference number of Li^+. The solid phase current density i_1 is governed by Ohm's law: $i_1 = -\kappa_1 \partial_x \phi_1$, where κ_1 is the effective conductivity of the solid and ϕ_1 is the solid-phase potential. Charge conservation demands that $i_1 + i_2 = I$, for a total current density I. The boundary conditions for the potentials (galvanostatic) are $-\kappa_1 \partial_x \phi_1 = I$ at $x = 0, L$ and $-\kappa_2 \partial_x \phi_2 = 0$ at $x = 0, L$, while the electronic charge fluxes are zero and the ionic charge fluxes are continuous at the separator interfaces. The mass balance for Li^+ is:

$$\epsilon_j \partial_t c = \partial_x \left(\epsilon_j D_j \partial_x c\right) - (1 - t_+^0)(\nu_+ F)^{-1}\partial_x i_2, \tag{11}$$

where ϵ_j is the volume fraction of electrolyte, D_j is the effective diffusion coefficient of the Li^+ through the electrolyte, and ν_+ is the number of cations into which a mole of electrolyte dissociates. The flux at both ends of the cell ($-\epsilon_j D_j \partial_x c$ by virtue of the zero ionic charge flux) is set to zero. The current density is given by the Butler-Volmer equation:

$$\partial_x i_j = -aFk_j(c)^{\alpha_{a,j}}(c_t - c_j^s)^{\alpha_{a,j}}(c_j^s)^{\alpha_{c,j}}\left(e^{-\alpha_a F\eta_j/(R_U T)} - e^{\alpha_a F\eta_j/(R_U T)}\right), \tag{12}$$

where α_a and α_c are the charge transfer coefficients, c_t is the total concentration of lithium, k_j is the rate constant for the relevant reaction and $\eta_j = \phi_1 - \phi_2 - U_j$ is the overpotential at the relevant electrode, in which U_j is the corresponding equilibrium potential.

5. Results and discussion

The Li-ion battery model was implemented in COMSOL Multiphysics [12]. A total of 500 simulations were performed by varying the initial state of charge SOC_{in} (initial c_n^s divided by c_t), the particle diameter in the positive electrode R_p and the positive electrode porosity ϵ_p. We set $\boldsymbol{\xi} = (\mathrm{SOC}_{in}, R_p, \epsilon_p)^T \in \mathcal{X} \subset \mathbb{R}^k$ as the input, with the $k = 3$ factors given by the components. The inputs for the 500 simulations were selected using a Sobol sequence. A current pulse $i(t)$ consisting of 12 s of 120 A discharge, followed by 12 s of relaxation (0 A load), and then 12 s of 120 A charge was simulated (galvanostatic operation). The output was taken to be the cell voltage $E_{cell}(t)$[V] at 0.5 s intervals, yielding a total of 73 values at $t = 0, 0.5, 1, \ldots, 35.5, 36$ s. These values were vectorized to form the outputs: $\mathbf{y}^T = \boldsymbol{\eta}(\boldsymbol{\xi}) = (E_{cell}(0), E_{cell}(0.5), \ldots, E_{cell}(35.5), E_{cell}(36)) \in \mathcal{Y} \subset \mathbb{R}^d$, where $\boldsymbol{\eta}$ as before represents the simulator and $d = 73$. The first 100 outputs were reserved for training the emulator and the remaining $m_t = 400$ were used for testing the emulator. The training data set is denoted $\{(\boldsymbol{\xi}_j, \mathbf{y}_j)\}_{j=1}^m$, as before, and the test data set is denoted $\{(\boldsymbol{\xi}_j^*, \mathbf{y}_j^*)\}_{j=1}^{m_t}$.

We consider two QoIs: (i) the energy efficiency $q = f(\boldsymbol{\xi}) = \eta_E = \int_{[d]} i(t)E_{cell}(t)dt \Big/ \int_{[c]} i(t)E_{cell}(t)dt \in \mathcal{Q} = [0,1]$, in which $[d]$ ($[c]$) is the discharge (charge) time interval; .and (ii) the voltage drop during discharge, $q = \Delta V_c \in \mathcal{Q} = \mathbb{R}_+$

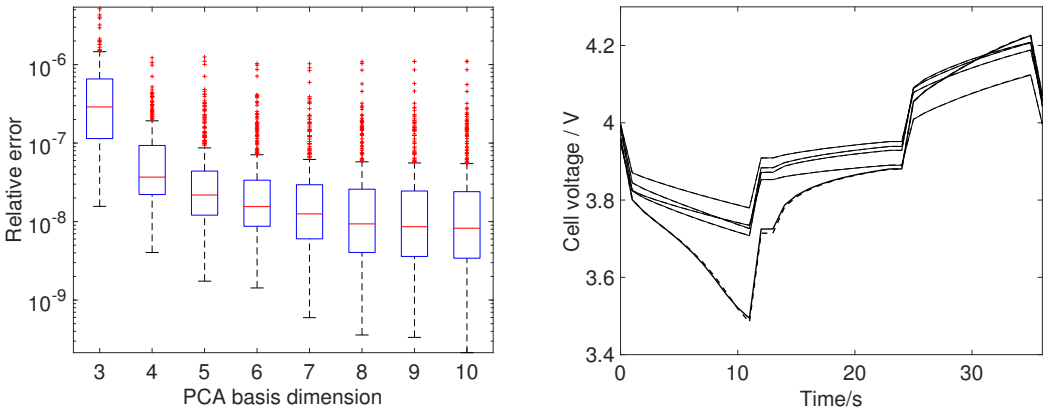

Figure 1. (Left) Boxplots of the emulator error on the test set $\{(\boldsymbol{\xi}_j^*, \mathbf{y}_j^*)\}_{j=1}^{m_t}$ using the training set $\{(\boldsymbol{\xi}_j, \mathbf{y}_j)\}_{j=1}^m$ with $m = 100$. (Right) Example predictions $\overline{\mathbf{y}}_{r,j}^*$ (dashed lines) of E_{cell} during the discharge-charge cycle $\mathbf{y}_j^*$ (solid lines) using $r = 10$. The worst case predictions (highest ϵ^*) are the thick lines and 4 further examples are shown.

Figure 1 shows Tukey boxplots of the relative errors on the test set $\{(\boldsymbol{\xi}_j^*, \mathbf{y}_j^*)\}_{j=1}^{m_t}$ using the training set $\{(\boldsymbol{\xi}_j, \mathbf{y}_j)\}_{j=1}^m$ with $m = 100$ for an increasing r. The errors were defined as $\epsilon^* = ||\overline{\mathbf{y}}_{r,j}^* - \mathbf{y}_j^*||/||\mathbf{y}_j^*||$, in which $\overline{\mathbf{y}}_{r,j}^*$ is the mean GP prediction of $\mathbf{y}_j^*$ using Eq. (7). Example predictions of the discharge-charge cycle are shown in Figure 1, for $r = 10$. The worst case predictions (highest ϵ^*) are shown, alongside 4 further examples. For the SA below we used $m = 100$ and $r = 10$.

5.1. Sensitivity analysis

The SA was performed using the SAFE package developed by Pianosi *et al.* [13]. For the variance-based method we placed a uniform distribution on the factors and sampled points $\mathbf{X} = [\xi_{i,j}]$, $i = 1, \ldots, 2k$, $j = 1, \ldots, N$ ($N = 5000$) in the $2k$ hypercube using a Latin hypercube design. The physical ranges were $0.1 \leq \epsilon_p \leq 0.4$, $0.5 \leq R_p[\mu\text{m}] \leq 2$ and $0.4 \leq \text{SOC}_{in} \leq 0.6$, and the factors were scaled to obtain $\mathcal{X} = [0, 1]^3$. The sampled inputs were used to produce the three input matrices $\mathbf{A} \in \mathbb{R}^{N \times k}$, $\mathbf{B} \in \mathbb{R}^{N \times k}$ and $\mathbf{C}_i \in \mathbb{R}^{N \times k}$, $i = 1, \ldots, k$, from which the QoI values $\mathbf{q_A} = f(\mathbf{A})$, $\mathbf{q_B} = f(\mathbf{B})$ and $\mathbf{q_{C_i}} = f(\mathbf{C}_i)$ were extracted.

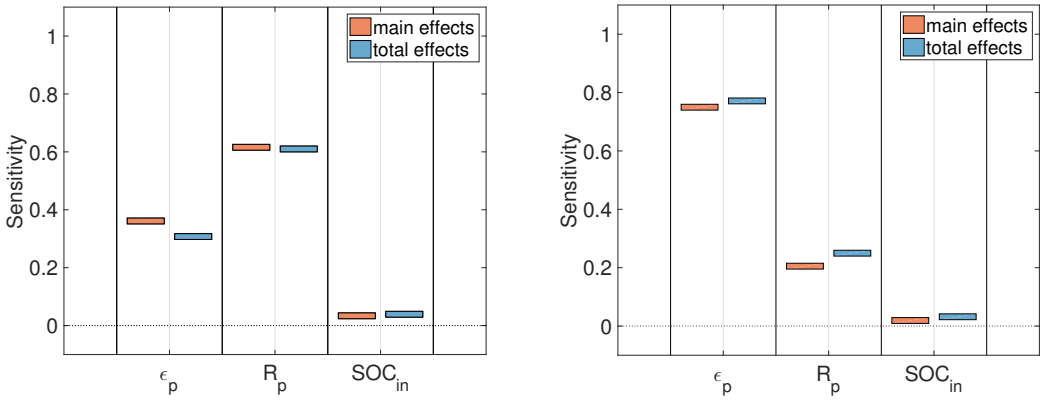

Figure 2. Main and total effects for η_E (left) and ΔV_c (right).

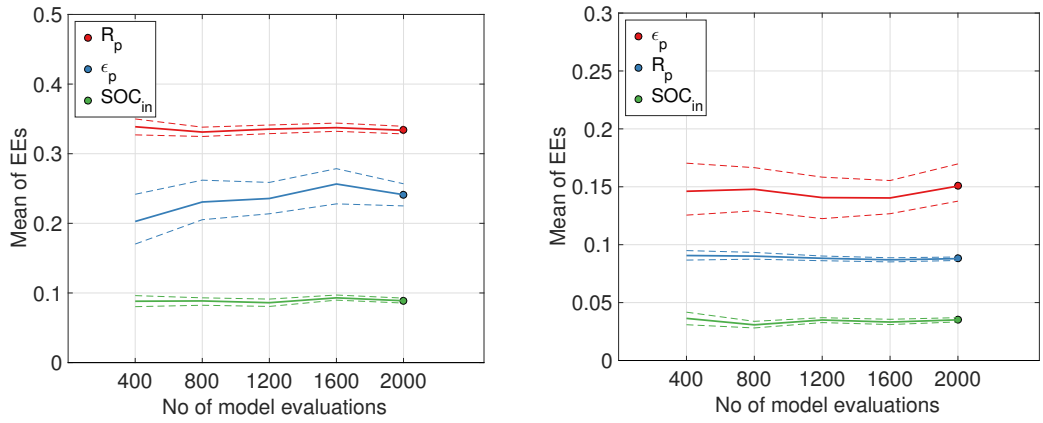

Figure 3. Convergence of the EEs for different numbers of model evaluations with 95% confidence intervals in the case of η_E (left) and ΔV_c (right).

Figure 2 presents both the main and total effects for the three factors. As expected, the particle size and porosity are the most influential, while the initial SOC mainly affects the open-circuit potential so has relatively little influence on q. The porosity determines the effective ionic conductivity (the volume fraction of electrolyte is ϵ_p) and since the ohmic loss is predominantly suffered in the ionic phase, ϵ_p has a major influence on the internal resistance. Moreover, the

reaction rate depends upon the concentration (per unit volume of the electrode) of Li^+ according to the Butler-Volmer law (12), so a restricted supply of Li^+ in the positive electrode will lead to a large concentration overpotential for a fixed current (the overpotential in (12) must increase as c decreases in order to maintain a fixed left hand side, i.e, applied current density). The particle radius determines the level of mass transport resistance for the solid Li (which has to diffuse through the particle to react at $R = R_p$) as well as the specific surface area for reaction (smaller particles lead to higher specific areas). Thus, increasing the particle radius will lead to a higher concentration overpotential and, therefore, a deterioration in performance. For the voltage drop during discharge (ΔV_c), ϵ_p has the greatest influence, followed by R_p and lastly SOC_{in}, as seen in Figure 2. The combination of an increased Ohmic drop and a higher concentration overpotential on the total polarization caused by a lower ϵ_p outweighs the effect of an increased concentration overpotential caused by lowering R_p.

For an EET, uniform distributions were selected for the three factors, which were again scaled to yield $\mathcal{X} = [0, 1]^3$. A major difference between the variance-based method and the EET is the sampling strategy. The EET is highly efficient, requiring only $M(k + 1)$ model evaluations *vs.* $N(k + 2)$ to calculate the main effect indices; in the results above, $N(k + 2) = 5000 \times (3 + 2) = 25000$, which is much higher than typical values of $M(k + 1)$. The trends in the means of the μ_i with confidence intervals (CIs) are depicted in Figure 3 for an increasing number of model evaluations ($M(k + 1)$). The CIs were established using bootstrapping [14], which consists of re-sampling the base points with replacement to produce P copies of the trajectories and for each of the P copies to use the EET to estimate μ_i and σ_i. This provides empirical distributions over μ_i and σ_i from which means and confidence bounds can be estimated.

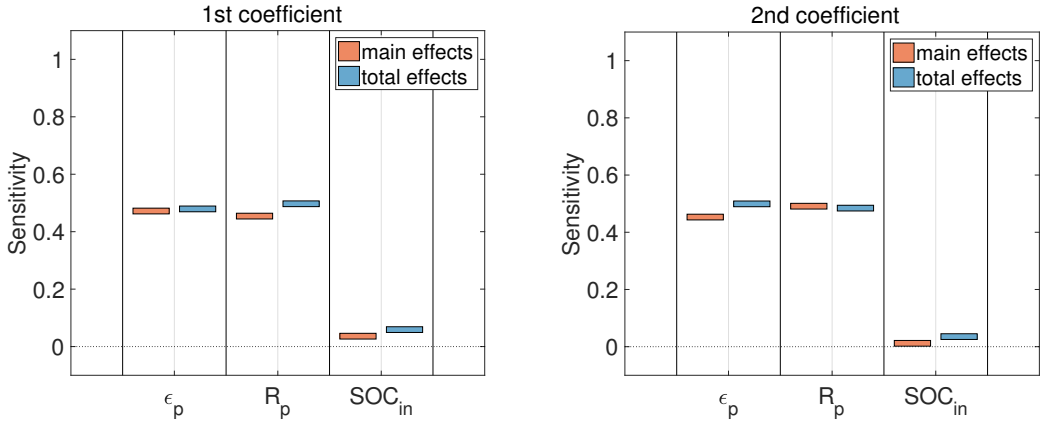

Figure 4. Main and total effects for the first two PCA coefficients (for the cell voltage curve).

The ranking of the inputs is the same as in the variance-based method. The means stabilize at around $M = 500$ (2000 model runs). There are small but noticeable fluctuations in the mean for ϵ_p around the value 0.25 even at much higher values of M but this behaviour is stable. The cause is the broad range of ϵ_p in comparison with the other factors and, therefore, the relatively small number of samples. Moving to a higher value of M the confidence intervals shrink, suggesting greater accuracy in the predictions. The ranking, however, is accurate even for very low numbers of M, which shows that the EET is more efficient than the variance based method. Although time cost is not an issue for the emulator, which provides extremely rapid predictions (on the order of a few seconds for 2000 predictions), in cases where a full simulator

is used the much lower number of model runs for the EET represents an enormous advantage.

To investigate the *sensitivity of the charge-discharge curve*, we examine the main and total effects of the PCA coefficients $w_i(\boldsymbol{\xi})$ using the same Latin hypercube design and $N = 5000$ (we replace, e.g., η_E with $w_i(\boldsymbol{\xi})$, $i \in \{1, \ldots, r\}$). The results are depicted in Figure 4 for w_1 and w_2. Figure 5 shows an example (from the test set) of the contributions from the PCA eigenvectors ($w_i\mathbf{v}_i$, up to $i = 4$) towards the final (mean centred) voltage profile. The first two contributions can be seen to have by far the most influence. The sensitivities of the coefficients w_1 and w_2 are highest for ϵ_p and R_p, with roughly equal contributions from each, while higher-order coefficients (w_3 and w_4) were more heavily influenced by SOC_{in}.

6. Summary and conclusions

SA is often unfeasible with complex computer models. In such cases emulators, of varying degrees of sophistication, can be employed. Quantifying the uncertainty in the emulator predictions is desirable, but this is only achievable for certain approaches. For multivariate outputs (especially in high dimensional spaces), SA under uncertainty is especially challenging, even when the QoI is a scalar. In this paper we propose a GP emulator approach for performing SA under uncertainty when the model output is multivariate (possibly in a high-dimensional space). We present an example for a Li-ion battery, demonstrating that the method can be efficient and accurate. We are able to perform a probabilistic SA on scalar QoIs derived from the output (*via* a linear functional) and also on the output itself by focusing individually on each random principal coefficient. This can be achieved with either full MC sampling or by using semi-analytical expressions that are extensions of those derived by Oakley and O'Hagan [5] for scalar QoIs.

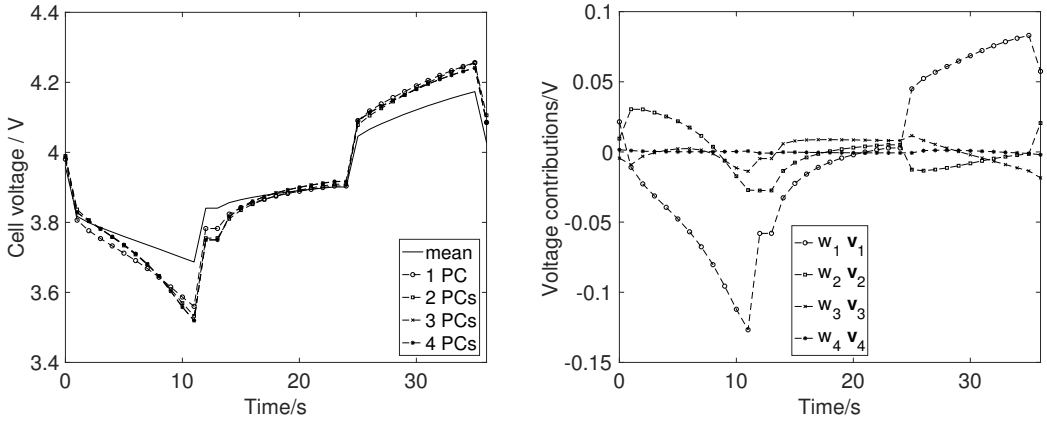

Figure 5. An example of the contributions from the PCA eigenvectors ($w_i\mathbf{v}_i$) towards the final (mean centred) voltage profile. In the left-hand figure, $w_i\mathbf{v}_i$ is successively added to the mean.

References

[1] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

[2] D. Zhou, K. Zhang, A. Ravey, F. Gao, and A. Miraoui, Parameter sensitivity analysis for fractional-order modeling of lithium-ion batteries, *Energies*, 9, 123, 2016.

[3] M. D. Morris, Factorial sampling plans for preliminary computational experiments, *Technometrics*, 33, 161–174, 1991.

[4] I. M. Sobol, Sensitivity estimates for nonlinear mathematical models, *MMCE*, 1, 407–414, 1993.

[5] J. E. Oakley and A. O'Hagan, Probabilistic sensitivity analysis of complex models: a Bayesian approach, *R. Stat. Soc. Ser. B Stat. Methodol.*, 66, 751–769, 2004.

[6] F. Pianosi, K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener, Sensitivity analysis of environmental models: A systematic review with practical workflow, *Environ. Model. Softw.*, 79, 214–232, 2016.

[7] A. Saltelli and I. M. Sobol', Sensitivity analysis for nonlinear mathematical models: numerical experience, *Matematicheskoe Modelirovanie*, 7, 16–28, 1995.
[8] I. Jolliffe, *Principal Component Analysis*. Springer Series in Statistics, Springer, 2002.
[9] D. Higdon, J. Gattiker, B. Williamsa, and M. Rightley, Computer model calibration using high-dimensional output, *J Am. Stat. Assoc.*, 103, 570–583, 2008.
[10] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. MIT Press, Cambridge MA, USA, 2006.
[11] M. Doyle, J. Newman, A. S. Gozdz, C. N. Schmutz, and J.-M. Tarascon, Comparison of modeling predictions with experimental data from plastic lithium ion cells, *J. Electrochem. Soc.*, 143, 1890–1903, 1996.
[12] `https://www.comsol.com/model/1d-lithium-ion-battery-model-for-internal-resistance-and-voltage-loss-determinat-19131`, Last accessed 20 November 2017.
[13] F. Pianosi, F. Sarrazin, and T. Wagener, A matlab toolbox for global sensitivity analysis, *Environ. Model. Softw.*, 70, 80–85, 2015.
[14] B. Efron, Bootstrap methods: Another look at the jackknife, *Ann. Statist.*, 7, 1–26, 01 1979.

Appendix A. Efficient sampling for sensitivity analysis under uncertainty

We consider a scalar linear functional QoI $Q = F(\boldsymbol{\xi}) = (G \circ \eta)(\boldsymbol{\xi})$ as defined in section 2. For example, we may consider $G(y) = \int_{\mathcal{R}} \eta(\mathbf{x};\boldsymbol{\xi}) w(\mathbf{x}) d\mu(\mathbf{x})$, for some measure μ on a compact subset $\mathcal{R}$ of $\mathbb{R}^L$ representing space or time. To keep matters simple, we set $w(\mathbf{x}) \equiv 1/\mu(\mathcal{R})$, use Lebesgue measure and assume that $\eta(\mathbf{x};\boldsymbol{\xi})$ is continuous; then the Riemann and Lebesgue integrals coincide and approximations to $G(y)$ by Riemann sums or Gauss quadratures converge. In reality of course we have a discrete output $\mathbf{y}_r = (y_1, \ldots, y_d)^T = \boldsymbol{\eta}_r(\boldsymbol{\xi})$ that approximates $\eta(\mathbf{x};\boldsymbol{\xi})$ at, say, points $\{\mathbf{x}_l\}_{l=1}^d \subset \mathcal{R}$. Correspondingly, we have a discrete approximation $g(\mathbf{y})$ of $G(y)$ defined by a quadrature $g(\mathbf{y}_r) = \mu(\mathcal{R})^{-1} \sum_{j=1}^{d'} b_j y_{l_j}$, where $\{y_{l_j}\}_{j=1}^{d'} \subset \{y_l\}_{l=1}^d$ (approximating $\eta(\mathbf{x}_{l_j};\boldsymbol{\xi})$, $j = 1, \ldots, d'$) is a subset of the coefficients of $\mathbf{y}_r$ and b_j are quadrature weights. If we are using a Gauss quadrature, the points $\mathbf{x}_{l_j}$ are specified and must be included in the design $\{\mathbf{x}_l\}_{l=1}^d$. For ease of presentation, and without loss of generality, we use a mid-point Riemann sum, so that $f(\boldsymbol{\xi}) = g(\boldsymbol{\eta}_r(\boldsymbol{\xi})) = g(\mathbf{y}_r) = d^{-1} \sum_{l=1}^d y_l$.

Rather than a point estimate of $\mathbf{y}_r$ we have a distribution over functions (7), which leads to distributions over $q = f(\boldsymbol{\xi})$ and therefore over the sensitivity measures, as a consequence of the emulator uncertainty. The typical sensitivity measures employed are $S_i = V_i/V = \mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i])/\mathbb{V}\mathrm{ar}(q)$ and $S_{T_i} = 1 - \mathbb{V}\mathrm{ar}_{\boldsymbol{\xi}_{\sim i}}(\mathbb{E}_{\xi_i}[q|\boldsymbol{\xi}_{\sim i}])/\mathbb{V}\mathrm{ar}(q)$. Oakley and O'Hagan [5] also propose the main effects $f_i = \mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i] - f_0$ as useful graphical summaries of the influences of each variable. We derive approximate estimates of the means and variances of these various quantities, extending the analysis in [5] to multiple output problems. Recalling relationship (7), namely, $\mathbf{y}_r = \sum_{i=1}^r w_i(\boldsymbol{\xi}) \mathbf{v}_i$, and denoting the l-th component of $\mathbf{v}_j$ by v_j^l, we obtain:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]\right] &= \mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}\left[\frac{1}{d}\sum_{l=1}^d y_l \middle| \xi_i\right]\right] = \frac{1}{d}\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}\left[\sum_{l=1}^d\sum_{j=1}^r w_j(\boldsymbol{\xi}) v_j^l \middle| \xi_i\right]\right] \\
&= \frac{1}{d}\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}\left[\sum_{j=1}^r b_j w_j(\boldsymbol{\xi}) \middle| \xi_i\right]\right] = \frac{1}{d}\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}\left[\sum_{j=1}^r b_j \mathbb{E}_{\boldsymbol{\eta}_r}[w_j(\boldsymbol{\xi})] \middle| \xi_i\right] \\
&= \frac{1}{d}\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}\left[\sum_{j=1}^r b_j m_j'(\boldsymbol{\xi}) \middle| \xi_i\right] = \frac{1}{d}\sum_{j=1}^r b_j \int_{[0,1]^{k-1}} m_j'(\boldsymbol{\xi}) d\boldsymbol{\xi}_{\sim i} = \frac{1}{d}\sum_{j=1}^r b_j T_j(\xi_i)
\end{aligned} \tag{A.1}$$

where $b_j = \sum_{l=1}^d v_j^l$ and the functions $T_j(\xi_i)$ are defined by the integrals in the last line. Similarly:

$$\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}[q]\right] = (1/d)\sum_{j=1}^r b_j \int_{[0,1]^k} m_j'(\boldsymbol{\xi}) d\boldsymbol{\xi} = (1/d)\sum_{j=1}^r b_j T_j', \tag{A.2}$$

where T'_j are now constants since the integration is over all input factors. Thus $\mathbb{E}_{\boldsymbol{\eta}_r}[f_i] = \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]] - \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}[q]] = (1/d)\sum_{j=1}^{r} b_j(T_j(\xi_i) - T'_j)$. The integrals defining T_j and T'_j can be approximated numerically at a very low computational cost.

We denote by $\boldsymbol{\xi}_{\sim\mathcal{I}}$ the vector of factors excluding those corresponding to the index set $\mathcal{I} \subset \{1,\ldots,k\}$ and denote by $\boldsymbol{\xi}_{\mathcal{I}}$ the subset of factors corresponding to $\mathcal{I}$. From the definition of covariance we have:

$$
\begin{aligned}
&\mathbb{C}\mathrm{ov}_{\boldsymbol{\eta}_r}\left(\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{I}}}[q|\boldsymbol{\xi}_{\mathcal{I}}], \mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{J}}}[q|\boldsymbol{\xi}_{\mathcal{J}}]\right) \\
&= \mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{I}}}[q|\boldsymbol{\xi}_{\mathcal{I}}]\,\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{J}}}[q|\boldsymbol{\xi}_{\mathcal{J}}]\right] - \mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\boldsymbol{\xi}_{\mathcal{I}}]\right]\,\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{J}}}[q|\boldsymbol{\xi}_{\mathcal{J}}]\right] \\
&= \int_{\boldsymbol{\xi}'_{\sim\mathcal{J}}}\int_{\boldsymbol{\xi}_{\sim\mathcal{I}}} \mathbb{E}_{\boldsymbol{\eta}_r}\left[f(\boldsymbol{\xi})f(\boldsymbol{\xi}')\right] d\boldsymbol{\xi}_{\sim\mathcal{I}} d\boldsymbol{\xi}'_{\sim\mathcal{J}} - \mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{I}}}[q|\boldsymbol{\xi}_{\mathcal{I}}]\right]\,\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{J}}}[q|\boldsymbol{\xi}_{\mathcal{J}}]\right] \\
&= \int_{\boldsymbol{\xi}'_{\sim\mathcal{J}}}\int_{\boldsymbol{\xi}_{\sim\mathcal{I}}} \left\{\mathbb{C}\mathrm{ov}_{\boldsymbol{\eta}_r}\left(f(\boldsymbol{\xi}), f(\boldsymbol{\xi}')\right) - \mathbb{E}_{\boldsymbol{\eta}_r}\left[f(\boldsymbol{\xi})\right]\mathbb{E}_{\boldsymbol{\eta}_r}\left[f(\boldsymbol{\xi}')\right]\right\} d\boldsymbol{\xi}_{\sim\mathcal{I}} d\boldsymbol{\xi}'_{\sim\mathcal{J}} \\
&\qquad\qquad - \mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{I}}}[q|\boldsymbol{\xi}_{\mathcal{I}}]\right]\,\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim\mathcal{J}}}[q|\boldsymbol{\xi}_{\mathcal{J}}]\right] \qquad \text{(A.3)} \\
&= \int_{\boldsymbol{\xi}'_{\sim\mathcal{J}}}\int_{\boldsymbol{\xi}_{\sim\mathcal{I}}} \mathbb{C}\mathrm{ov}_{\boldsymbol{\eta}_r}\left(f(\boldsymbol{\xi}), f(\boldsymbol{\xi}')\right) d\boldsymbol{\xi}_{\sim\mathcal{I}} d\boldsymbol{\xi}'_{\sim\mathcal{J}} \\
&= \int_{\boldsymbol{\xi}'_{\sim\mathcal{J}}}\int_{\boldsymbol{\xi}_{\sim\mathcal{I}}} \sum_{j=1}^{r}\sum_{p=1}^{r} b_j b_p \mathbb{C}\mathrm{ov}_{\boldsymbol{\eta}_r}\left(w_j(\boldsymbol{\xi}), w_p(\boldsymbol{\xi}')\right) d\boldsymbol{\xi}_{\sim\mathcal{I}} d\boldsymbol{\xi}'_{\sim\mathcal{J}} \\
&= \frac{1}{d^2}\int_{\boldsymbol{\xi}'_{\sim\mathcal{J}}}\int_{\boldsymbol{\xi}_{\sim\mathcal{I}}} \sum_{j=1}^{r} b_j^2 c_j(\boldsymbol{\xi},\boldsymbol{\xi}') d\boldsymbol{\xi}_{\sim\mathcal{I}} d\boldsymbol{\xi}'_{\sim\mathcal{J}} = U(\boldsymbol{\xi}_{\mathcal{I}}, \boldsymbol{\xi}_{\mathcal{J}})
\end{aligned}
$$

in which the last step follows from the mutual independence of the w_i, and the posterior covariances $c_j(\boldsymbol{\xi},\boldsymbol{\xi}')$ are given in Eqs. (6). Now, $\mathbb{V}\mathrm{ar}_{\boldsymbol{\eta}_r}(f_i) = \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}^2_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]] - 2\mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]\mathbb{E}[q]] - \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}^2[q]] - \mathbb{E}^2_{\boldsymbol{\eta}_r}[f_i]$ in which the last term on the right hand side is already known. The first to third terms are calculated by using the definition of covariance and Eq. (A.3) with $(\mathcal{I},\mathcal{J}) = (\{i\},\{i\})$, $(\{i\},\emptyset)$ and $(\emptyset,\emptyset)$, together with the previous expressions for $\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]\right]$ and $\mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}[q]]$. The integrals in Eq. (A.3) are again cheap to evaluate. The means of all the f_i, evaluated separately for selected values of ξ_i in $[0,1]$, can be combined on a single plot together with standard deviations to measure the influences of the factors [5] .

We next consider the partial variances V_i. We first note that $\mathbb{E}_{\boldsymbol{\eta}_r}[V_i] = \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{V}\mathrm{ar}_{\xi_i}(\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i])] = \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}_{\xi_i}[\mathbb{E}^2_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]] - \mathbb{E}^2_{\xi_i}[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]]] = \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}_{\xi_i}[\mathbb{E}^2_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]]] - \mathbb{E}_{\boldsymbol{\eta}_r}[\mathbb{E}^2[q]]$, the second term of which is already known. The first term is evaluated as follows:

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\eta}_r}\left[\mathbb{E}_{\xi_i}[\mathbb{E}^2_{\boldsymbol{\xi}_{\sim i}}[q|\xi_i]]\right] = \mathbb{E}_{\boldsymbol{\eta}_r}\left[\int_{[0,1]}\left(\int_{[0,1]^{k-1}} f(\boldsymbol{\xi}^*) d\boldsymbol{\xi}^*_{\sim i}\int_{[0,1]^{k-1}} f(\boldsymbol{\xi}) d\boldsymbol{\xi}_{\sim i}\right) d\xi_i\right] \\
&= \int_{[0,1]}\int_{[0,1]^{k-1}}\int_{[0,1]^{k-1}} \mathbb{E}_{\boldsymbol{\eta}_r}\left[f(\boldsymbol{\xi})f(\boldsymbol{\xi}^*)\right] d\boldsymbol{\xi}^*_{\sim i} d\boldsymbol{\xi}_{\sim i} d\xi_i \\
&= \int_{[0,1]}\int_{[0,1]^{k-1}}\int_{[0,1]^{k-1}} \left\{\mathbb{C}\mathrm{ov}_{\boldsymbol{\eta}_r}\left(f(\boldsymbol{\xi}), f(\boldsymbol{\xi}^*)\right) - \mathbb{E}_{\boldsymbol{\eta}_r}[f(\boldsymbol{\xi})]\mathbb{E}_{\boldsymbol{\eta}_r}[f(\boldsymbol{\xi}^*)]\right\} d\boldsymbol{\xi}^*_{\sim i} d\boldsymbol{\xi}_{\sim i} d\xi_i \qquad \text{(A.4)} \\
&= \frac{1}{d^2}\sum_{j=1}^{r}\int_{[0,1]}\int_{[0,1]^{k-1}}\int_{[0,1]^{k-1}} b_j\left\{b_j c_j(\boldsymbol{\xi},\boldsymbol{\xi}^*) - \sum_{p=1}^{r} b_p w_j(\boldsymbol{\xi}) w_p(\boldsymbol{\xi}^*)\right\} d\boldsymbol{\xi}^*_{\sim i} d\boldsymbol{\xi}_{\sim i} d\xi_i
\end{aligned}
$$

in which $\boldsymbol{\xi}^* = (\xi_1^*,\ldots,\xi_i,\ldots,\xi_k^*)^T$. These integrals are readily and cheaply approximated numerically. A first order Taylor expansion yields $\mathbb{E}_{\boldsymbol{\eta}_r}[S_i] = \mathbb{E}_{\boldsymbol{\eta}_r}[V_i/V] = \mathbb{E}_{\boldsymbol{\eta}_r}[V_i]/\mathbb{E}_{\boldsymbol{\eta}_r}[V]$, from which the main effect indices can be approximated.