Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Global Sensitivity Analysis on the Bergman Minimal Model ⋆

Souransu Nandi Tarunraj Singh

Control, Dynamics and Estimation Laboratory, University at Buffalo, Buffalo, NY 14221 USA (e-mail: tsingh@buffalo.edu).

Abstract: In this paper, a novel Global Sensitivity Analysis method is developed and is illustrated on the popular Bergman minimal model for Type 1 Diabetes (T1D). Four parameters are assumed to be uncertain in the model to mimic patient variability. An algorithm is presented to evaluate sensitivity metrics by which the uncertain parameters can be ranked. Results reveal that for a single meal scenario, insulin sensitivity is one of the most important factors after the consumption of meal that influences the blood glucose concentration in people with T1D.

Keywords: Sensitivity Analysis, Polynomial Chaos, Sobol Indices, Insulin Sensitivity

1. INTRODUCTION

Type 1 Diabetes (T1D) is a chronic ailment which without careful regulation results in micro-vascular complications such as retinopathy, neuropathy, nephropathy and macro-vascular complications such as cardiovascular disease and strokes. The American Diabetes Association estimates that the cost associated with the treatment and productivity loss of patients (with diabetes) has risen from $245 billion in 2012 to $327 billion in 2017, a 34 % increase (Association et al. (2018)). Therefore there is a clear motivation from a quality of life and from a health care economics point of view, to develop an Artificial Pancreas (AP) which can emulate the human pancreas as closely as possible. To that end, researchers have developed several mathematical models for predicting the blood glucose levels in T1D patients.

Over the past few decades, models for T1D have increased their fidelity in terms of prediction and performance Palumbo et al. (2013). However, this development takes place at the price of model complexities. Inclusion of systems of differential equations capturing more subsystems of the human physiology, although leads to better model prediction, increases the number of states as well as the parameters. Typically, measurements from real patients are used to estimate these parameters. Errors associated with the experimental process of measurement as well as un-modeled physiological behavior, lead to parameters being identified as random variables with finite support or a probability distribution.

Even with the computational resources available today, the increase in the number of parameters makes it difficult to quantify the state uncertainty in the T1D models as a function of parametric uncertainties. A comprehensive Sensitivity Analysis (SA) would have a profound impact on several avenues of research and development. It would help us better understand and rank which parameters contribute (jointly) to the variation of glucose concentration: thereby motivating researchers to further study those parameters and their interactions. It would also help us figure out the non-significant parameters. This information could be used in a couple of ways. It would give scientists working on understanding the human physiology the added information that mathematically the influence of a certain parameter seems to have little to no effect on the glucose concentration: hence concluding that the particular parameter may not be significant in the biological process. The insignificance of the parameters could also be interpreted by the control and optimization community as an opportunity to reduce the uncertain dimension of the model when solving stochastic optimal control. Other benefits of SA can be found in articles by Saltelli et al. (2004, 2000); Campolongo et al. (2007); Saltelli et al. (2010); Borgonovo and Plischke (2016) and references there in.

This paper presents a novel sensitivity analysis method to rank uncertain parameters of a T1D model. The method presented is a Global Sensitivity Analysis (GSA) technique that is dependent on the probability distribution function (pdf) of the evolving blood glucose concentration. The Bergman model for glucose – insulin dynamics Bergman et al. (1981), is chosen to illustrate the proposed GSA formulation. Parameters in the model are considered to be random variables (including the initial blood glucose level) to represent the inter- and intra-patient variability.

To propagate the uncertainties of the model through time, the popular uncertainty quantification tool: Polynomial Chaos (PC) is used. The time evolving pdfs are then used to rank the parameters of the model. Using a quantitative measure, it is determined which parameter has the most influence on the pdf of the blood glucose level. After comparing the measures, an evaluation is made on the relative importance of the uncertain parameters. Comparisons are also made to the variance based GSA technique using Sobol indices. Final results are seen to be extremely interesting in the context of understanding T1D. It is observed that a couple of parameters which are directly linked to the insulin sensitivity of a patient have significant impact on their blood glucose level.

⋆ Research Sponsored by Juvenile Diabetes Research Foundation (JDRF) under JDRF Grant Key: 1-SRA-2019-823-S-B.

Table 1. Statistical Distances and their mathematical expressions

Statistical Distance ($\mathcal{D}$)	A	Description
Wasserstein	W	$\int_{\Omega_y} \lvert P_Y(y) - Q_Y(y)\rvert dy$
Hellinger	H	$\left[2\left(1 - \int_{\Omega_y} \sqrt{p_Y(y)q_Y(y)}dy\right)\right]^{\frac{1}{2}}$
Total Variation	T	$0.5\int_{\Omega_y} \lvert p_Y(y) - q_Y(y)\rvert dy$
Kolmogorov	K	$\sup_{y\in\Omega_y} \lvert P_Y(y) - Q_Y(y)\rvert$
Bhattacharya	B	$-log(\int_{\Omega_y} \sqrt{p_Y(y)q_Y(y)}dy)$
Cramer-von Mises	C	$\left[\int_{\Omega_y} \lvert P_Y(y) - Q_Y(y)\rvert^2 dy\right]^{\frac{1}{2}}$

The paper has been structured in the following way. Section 2 introduces the moment independent metrics that have been proposed. Section 3 presents an efficient algorithm to determine the afore-mentioned metrics. Section 4 presents the results from the implementation of the GSA metrics on the Bergman model; before ending with concluding remarks in section 5.

2. MOMENT INDEPENDENT METRICS

Similar to Sobol' indices, several sets of metrics are presented (also referred to as classes) which capture the contributions of uncertain variables varied individually (first order effect) and when varied concurrently. The classes are based on observing the disparity between the output pdf of a function and the conditional output pdf evaluated at certain specific locations over the input sub-space.

2.1 Statistical Distance measures

In probability theory, the disparity between two probability measures is often quantitatively determined using metrics called *statistical distances*. Larger the value of the distance, more distinct are the pdfs. Based on the penalty levied on the disparity measures, there exists more than one type of distance. The quantitative value (of distance) changes with the choice of statistical distance used. However, they retain certain basic properties such as non-negativity, symmetry, positive definiteness among others. A list of important or popular Statistical Distances ($\mathcal{D}$) have been presented in Table 1. The first and second columns present the name and a convenient abbreviation of different $\mathcal{D}$s respectively. The final column presents the mathematical expression, where y is used to represent a realization of the random variable Y, $p_Y(y)$ and $q_Y(y)$ are two distinct pdfs over Y, $P_Y(y)$ and $Q_Y(y)$ are the corresponding cdfs and Ω_Y is the support of Y (i.e. $y \in \Omega_Y$). The descriptions for Wasserstein, Hellinger, Total Variation, Kolmogorov distance can be found in Gibbs and Su (2002), Bhattacharya distance in Bhattacharyya (1943) and Cramer-von Mises distance in Baringhaus and Henze (2017). Henceforth in this article, $\mathcal{D}_A(p,q)$ is used to represent the statistical distance with abbreviation A. For example, $\mathcal{D}_W$ would refer to the Wasserstein distance.

We present 6 different measures for GSA (refer to Table 1). The measures are different based on the nature of their penalty with disparity and computational effort required to compute them.

The $\mathcal{D}$s can be used to define certain global sensitivity metrics by comparing disparities between output pdfs and conditional output pdfs. Depending on the nature of conditional pdfs, the metrics are divided into classes. For functional relationships of the form

$$Y = g(\boldsymbol{X}) \tag{1}$$

where $\boldsymbol{X} \in \mathbb{R}^n$ is a vector of n random inputs, $Y \in \mathbb{R}$ is a scalar output and $g : \boldsymbol{X} \to Y$ is a function which maps the inputs to the output; class 1 considers the conditional output pdf $f_{Y|X_i}(y, x_i)$ which is the pdf of the output Y for a given fixed input value of X_i. Class 2 considers the conditional pdf $f_{Y|X_i,X_j}(y, x_i, x_j)$ which is the pdf of the output Y given fixed values of X_i and X_j. Following a similar pattern we get the higher class metrics. Finally, a total effect metric for each individual input X_i is also defined which sums up all the class order metrics which are relevant to X_i. The total effect metric is analogous to the total effect Sobol' indices. Details about the metrics have been elaborated in the following subsections.

2.2 Class 1 metrics

Consider the equation

$$\overline{\mathcal{D}}_i = \int_{\Omega_{X_i}} \mathcal{D}(f_Y(y), f_{Y|X_i}(y, x_i)) f_{X_i}(x_i) dx_i \tag{2}$$

where $\overline{\mathcal{D}}_i$ is the averaged statistical distance over the input subspace Ω_{X_i}, Ω_{X_i} represents the domain of X_i (i.e. $x_i \in \Omega_{X_i}$),

$$f_Y(y) = \int_{\Omega_{\boldsymbol{X}}} f_{\boldsymbol{X},Y}(\boldsymbol{x}, y) d\boldsymbol{x}, \tag{3}$$

$$f_{Y|X_i}(y, x_i) = \int_{\Omega_{\tilde{\boldsymbol{X}}_{\boldsymbol{i}}}} f_{Y,\tilde{\boldsymbol{X}}_{\boldsymbol{i}}|X_i}(\boldsymbol{x}, y) d\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tag{4}$$

$f_{X,Y}(\boldsymbol{x}, y)$ is the joint input output pdf, $\boldsymbol{x}$ is a realization of $\boldsymbol{X}$, $\tilde{\boldsymbol{X}}_{\boldsymbol{i}}$ is the joint variable $\tilde{\boldsymbol{X}}_{\boldsymbol{i}} = [X_1, ..., X_j, ...X_n]^T$ for $j \neq i$, $\tilde{\boldsymbol{x}}_{\boldsymbol{i}}$ is a realization of $\tilde{\boldsymbol{X}}_{\boldsymbol{i}}$, $\Omega_{\tilde{\boldsymbol{X}}_{\boldsymbol{i}}}$ is the domain of $\tilde{\boldsymbol{X}}_{\boldsymbol{i}}$ and $\Omega_{\boldsymbol{X}}$ is the domain of $\boldsymbol{X}$.

$\mathcal{D}(f_Y(y), f_{Y|X_i}(y, x_i))$ in equation (2) quantifies the disparity between the output pdf and the output pdf conditioned on a single input parameter. If the particular input X_i (on its own) is unimportant, it would contribute minimally to the marginalized output pdf: which would mean $f_Y(y)$ and $f_{Y|X_i}(y, x_i)$ are in close proximity: resulting in low values of $\mathcal{D}$. In contrast, if the input parameter (X_i) is indeed influential, it would contribute significantly to the output pdf. This means that the conditioned pdf $f_{Y|X_i}(y, x_i)$ would be far from the marginalized pdf $f_Y(y)$ leading to higher values of $\mathcal{D}$. Hence, on observing what the values of $\mathcal{D}$ are on average, one can estimate the relative importance of inputs (i.e. larger the value of $\overline{\mathcal{D}}_i$, more is the significance of X_i). Note that the metrics $\overline{\mathcal{D}}_i$ only represent the first order effects of the inputs: which means the metric only captures the disparity between the output pdf and the conditioned output pdf when only a single input variable X_i is varied across its domain Ω_{X_i}. It does not account for the effects of the uncertainties when the input variables are varied simultaneously. The effects from varying inputs simultaneously are quantified only via higher order effect metrics (i.e. Class 2 and higher).

2.3 Class 2 metrics

Similar to Class 1 metrics, consider the equation

$$\overline{\mathcal{D}}_{i,j} = \int_{\Omega_{X_i,X_j}} \mathcal{D}(f_Y(y), f_{Y|X_i,X_j}(y,x_i,x_j))\times f_{X_i,X_j}(x_i,x_j)dx_i dx_j \quad (5)$$

where $\overline{\mathcal{D}}_{i,j}$ is the averaged statistical distance over the input subspace Ω_{X_i,X_j} and

$$f_{Y|X_i,X_j}(y,x_i,x_j) = \int_{\Omega_{\tilde{\boldsymbol{X}}_{\boldsymbol{ij}}}} f_{Y,\tilde{\boldsymbol{X}}_{\boldsymbol{ij}}|X_i,X_j}(\boldsymbol{x},y)d\tilde{\boldsymbol{x}}_{\boldsymbol{ij}}. \quad (6)$$

If the combination of X_i and X_j is an important contributor to the uncertainty in the output, then the conditioned pdf, $f_{Y|X_i,X_j}(y,x_i,x_j)$ would be further apart from the marginalized pdf $f_Y(y)$ since X_i and X_j have been fixed and are not being varied. This would yield higher values of the statistical distance $\mathcal{D}$ in equation (5). Similar to the class 1 metric, averaging this distance over the subspace of Ω_{X_i,X_j} would be indicative of the joint contribution of X_i and X_j. Analogous arguments can be made about the joint contribution of m input variables, quantified as class m metic. If there are a total of n inputs, the final metric would be a class n metric which would account for the variation of the output due to the joint variation of all the n inputs.

2.4 Class n Metric

The class n metric is defined as

$$\overline{\mathcal{D}}_{i_1,i_2,\ldots,i_n} = \int_{\Omega_{\boldsymbol{X}}} \mathcal{D}(f_Y(y), f_{Y|\boldsymbol{X}}(y,\boldsymbol{x}))f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x}. \quad (7)$$

Note that $f_{Y|\boldsymbol{X}}(y,\boldsymbol{x})$ is a Dirac delta function because fixing all the inputs would make the output deterministic with a definite value. Therefore we would get

$$f_{Y|\boldsymbol{X}}(y,\boldsymbol{x}) = \delta(y - g(\boldsymbol{X})) \quad (8)$$

and $\mathcal{D}$ would measure the statistical distance between the output pdf and a delta function. Averaging that distance over the entire $\Omega_{\boldsymbol{X}}$ would yield the final class n metric.

2.5 Normalized Metrics

Considering that the values of $\mathcal{D}$ can vary largely depending on the type of statistical distance chosen to be implemented, a normalization of all the metrics are exercised to facilitate comparisons. The normalized metrics are referred to as NS and represent a shorthand for Non-moment based Sensitivity indices. They are defined in the following way:

Class 1 NS:

$$NS_i = \frac{\overline{\mathcal{D}}_i}{\sum \overline{\mathcal{D}}_i + \sum \overline{\mathcal{D}}_{i,j} + \cdots + \sum \overline{\mathcal{D}}_{i_1,\cdots,i_{n-1}}} \quad (9)$$

Class 2 NS:

$$NS_{ij} = \frac{\overline{\mathcal{D}}_{i,j}}{\sum \overline{\mathcal{D}}_i + \sum \overline{\mathcal{D}}_{i,j} + \cdots + \sum \overline{\mathcal{D}}_{i_1,\cdots,i_{n-1}}} \quad (10)$$

$$\vdots$$

and eventually

Class $n-1$ NS:

$$NS_{i_1,\cdots,i_{n-1}} = \frac{\overline{\mathcal{D}}_{i_1,\cdots,i_{n-1}}}{\sum \overline{\mathcal{D}}_i + \sum \overline{\mathcal{D}}_{i,j} + \cdots + \sum \overline{\mathcal{D}}_{i_1,\cdots,i_{n-1}}}. \quad (11)$$

Note that class n metric is not used during normalization since it is a constant that would get added to the denominator of all the NS metrics. As only the relative values of NS are significant, removing a constant from the common denominator does not effect the ranking of the magnitudes of the NS metrics. Evident from the definitions, NS metrics vary between 0 and 1. Closer the value of NS to 1, more is the significance of the joint input uncertainty corresponding to the metric. It should be noted that the Borgonovo metric δ_i is a specific $\overline{\mathcal{D}}_i$ where the chosen $\mathcal{D}$ type is the Total Variation distance and the Gamboa metric is analogous to the case where $\mathcal{D}$ type is the Cramer-von Mises distance. However, they only consider the first order effects and the effect of the joint variation of the uncertain inputs are not investigated. To the authors' best knowledge, a metric based on statistical distances to capture the first order as well as higher order effects of contributing input random variables on the nature of the output uncertainty is investigated for the first time in this work. To quantify the total effect a single input variable has on the uncertainty of the output, (similar to the total effect Sobol' indices), total effect NS_{T_i} metrics are also defined. The metrics are evaluated as

$$NS_{T_i} = \sum_{\mathcal{P}_i} NS_{i_1,\cdots,i_{n-1}} \quad (12)$$

where $\mathcal{P}_j = \{(i_1,\cdots,i_{n-1}) \exists k, 1 \leqslant k \leqslant n, i_k = j\}$. The NS_{T_i} are simply the sum of all the partial sensitivity measures NS where the influence of X_i has been accounted for either in part or in whole.

3. EFFICIENT EVALUATION OF NS INDICES

We often encounter problems in engineering where even a single function evaluation is computationally expensive. For those functions, deriving the aforementioned GSA measures can become impractical especially when output pdfs and conditioned output pdfs need evaluation. Either analytical expressions for these pdfs do not exist, cannot be evaluated or can only be approximated from a large number of sample realizations (eg. Monte Carlo methods): making the calculation of the NS metrics through traditional techniques extremely difficult. Methods to reduce the computational cost and provide tractable alternatives to approximate the NS metrics are presented next.

3.1 Polynomial Chaos based Surrogate Model

Polynomial Chaos (PC) is a probabilistic modeling tool to approximate a stochastic function with a polynomial function of the random variables. First introduced by Wiener (1938), to expand a Gaussian process with the help of an infinite series using Hermite polynomials, PC has been subsequently thoroughly investigated by a number of researchers Ghanem and Spanos (1991); Cameron and Martin (1947); Xiu and Karniadakis (2002). In this paper, we determine a surrogate model $\hat{Y}$ using PC for the true system $Y = g(\boldsymbol{X})$ such that instead of sampling Y (which can be expensive), we can sample $\hat{Y}$ (the surrogate model) instead, relatively cheaply. From PC theory, it is well known that a stochastic function Y can be written as an infinite polynomial series expansion in the form

$$Y = \sum_{i=1}^{\infty} Y_i \Phi_i(\boldsymbol{X}) \quad (13)$$

where Φ_i are certain specific orthogonal set of basis functions in $\boldsymbol{X}$ and $Y_i \in \mathbb{R}$ are their corresponding coefficients. The nature of Φ_i is determined by probability measures of $\boldsymbol{X}$. The orthogonal bases required for some of the popular random variable types can be found in the Wiener-Askey scheme provided in Xiu and Karniadakis (2002). Equation (13) is typically truncated to a finite number of terms as an approximation to yield the surrogate model

$$Y \approx \hat{Y} = \sum_{i=1}^{N} \hat{Y}_i \Phi_i(\boldsymbol{X}). \tag{14}$$

The objective is to determine the coefficients $\hat{Y}_i$ of the surrogate model $\hat{Y}$ so that one can have a simple model of the true stochastic system as a polynomial function of the input stochastic variables.

Traditionally there have been two broad category of methods to find those coefficients: namely Intrusive methods and Non-Intrusive methods (Kim et al. (2013)). In this paper, we briefly review a method from each of those categories.

Intrusive PC In intrusive PC, we look for coefficients which minimize the mean value of the square of the model error, leading to the following closed form expression for the coefficients

$$\hat{Y}_i = \frac{\int_{\Omega_{\boldsymbol{X}}} g \Phi_i f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}}{\int_{\Omega_{\boldsymbol{X}}} \Phi_i^2 f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}}. \tag{15}$$

This method of evaluating the coefficients is also popularly known as the Galerkin projection method.

Although calculating the denominator of equation (15) is trivial, the numerator can be extremely difficult to evaluate for generic non-linear functions and places where g is merely a computer code. Hence in spite of obtaining a clean expression for the coefficients in equation (15) the need for the evaluation of a multivariate integral forms the most profound limitation of this method.

Non-Intrusive PC In order to circumvent multidimensional integrals, the coefficients can also be determined using function evaluations and least squares. Consider the expression: $\boldsymbol{e} =$

$$\underbrace{\begin{bmatrix} \Phi_1(\boldsymbol{x}^{(1)}) & \Phi_2(\boldsymbol{x}^{(1)}) & \cdots & \Phi_N(\boldsymbol{x}^{(1)}) \\ \Phi_1(\boldsymbol{x}^{(2)}) & \Phi_2(\boldsymbol{x}^{(2)}) & \cdots & \Phi_N(\boldsymbol{x}^{(2)}) \\ \vdots & \vdots & \cdots & \vdots \\ \Phi_1(\boldsymbol{x}^{(m)}) & \Phi_2(\boldsymbol{x}^{(m)}) & \cdots & \Phi_N(\boldsymbol{x}^{(m)}) \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_N \end{bmatrix}}_{\boldsymbol{Y_{PC}}} - \underbrace{\begin{bmatrix} y(\boldsymbol{x}^{(1)}) \\ y(\boldsymbol{x}^{(2)}) \\ \vdots \\ y(\boldsymbol{x}^{(m)}) \end{bmatrix}}_{\boldsymbol{y}} \tag{16}$$

where $\boldsymbol{x}^{(i)}$ represents the i^{th} sample point from a total of m samples in the input space, $A \in \mathbb{R}^{m \times N}$ is a matrix whose rows represent the bases evaluated at $\boldsymbol{x}^{(i)}$, $\boldsymbol{y}$ is a vector assimilated using the transformation g over $\boldsymbol{x}^{(i)}$ samples and $\boldsymbol{e}$ is a vector of the approximation errors at each sample point. The coefficients $\boldsymbol{Y_{PC}}$ can be determined by minimizing the 2-norm of $\boldsymbol{e}$ and is given by

$$\boldsymbol{Y_{PC}} = (A^T A)^{-1} A^T \boldsymbol{y}. \tag{17}$$

Equation (17) now provides a simple expression to determine the coefficients of the PC surrogate model from system realizations of the original function.

Since sampling the surrogate is much cheaper than sampling the true system, we now have a tractable way to determine the computationally expensive pdfs $f_Y(y)$, $f_{Y|X_i}(y, x_i)$ and $f_{Y|\tilde{\boldsymbol{X}}_{\boldsymbol{i}}}(y, \tilde{\boldsymbol{x}}_{\boldsymbol{i}})$ using MC methods.

3.2 Efficient Evaluation of the Expectation Integral

Expectation integrals of the form

$$I = \int_{\Omega_{\boldsymbol{X}}} k(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \tag{18}$$

turn up in numerous applications of basic and applied sciences. As a result, many researchers over the years have endeavoured to efficiently evaluate this integral. Some of the most popular methods used have been Monte Carlo (MC) sampling methods (Stroud (1971)), Gauss quadrature (GQ) rules (Stroud and Secrest (1966)), Sparse quadrature rules (Gerstner and Griebel (1998)) and Conjugate Unscented Transform (CUT) rules (Adurthi et al. (2018)). In this paper, the methods are not highlighted (as there is considerable existing literature). However, we do provide commentary on which method to use for the NS metrics.

In all of these methods, the integral I is approximated by a weighted sum of function evaluations

$$I \approx \hat{I} = \sum_{i=1}^{N_{method}} w_i k(\boldsymbol{x}^{(i)}) \tag{19}$$

where $\boldsymbol{x}^{(i)}$ are certain samples from the input space, N_{method} denotes the number of such samples and w_i are specific weights. The aforementioned methods primarily differ in the manner by which the location of the sample points $\boldsymbol{x}^{(i)}$ and their corresponding weights w_i are determined. The benefit of representing an integral with function evaluations lies in the fact that sample realizations are independent of each other and parallel computing techniques can be adopted to evaluate the system realizations simultaneously.

Remarks on the use of MC, GQ and CUT The objective of discussing expectation integrals has been to highlight the fact that in order to determine the values of NS, we first need to evaluate equations (2), (5) and (7) which are expectation integrals.

$\overline{\mathcal{D}}_i$ results from a univariate integral while the other class metrics are outcomes of multidimensional integrals. These integrals are approximated as

$$\overline{\mathcal{D}}_i \approx \sum_{j=1}^{n_{method}} w_j \mathcal{D}(f_{\hat{Y}}(\hat{y}), f_{\hat{Y}|X_i}(\hat{y}, x_i^{(j)})) \tag{20}$$

$$\overline{\mathcal{D}}_{i,j} \approx \sum_{k=1}^{n_{method}} w_k \mathcal{D}(f_{\hat{Y}}(\hat{y}), f_{\hat{Y}|X_i,X_j}(\hat{y}, x_i^{(k)}, x_j^{(k)})) \tag{21}$$

$$\vdots$$

$$\overline{\mathcal{D}}_{i_1,i_2,\cdots,i_n} \approx \sum_{j=1}^{n_{method}} w_j \mathcal{D}(f_{\hat{Y}}(\hat{y}), f_{\hat{Y}|\boldsymbol{X}}(\hat{y}, \boldsymbol{x}^{(j)})) \tag{22}$$

where $x_i^{(j)}$ is the j^{th} sample point out of a total of n_{method} samples in the Ω_{X_i}. Note that y has been replaced by $\hat{y}$ in the equations to represent the surrogate model instead of the true system. n_{method} represents the total number

of sample points in the used sampling algorithm. In the following section, commentary on which sampling method to use for the metrics is provided.

Since the convergence of MC is slow and requires an enormous number of samples to evaluate equations (20) through (22), it is never used to determine $\overline{\mathcal{D}}_i$. GQ is always used to evaluate class 1 metrics $\overline{\mathcal{D}}_i$ since it is an univariate integral and GQ provides the minimal set of points and weights to integrate a polynomial of any order for univariate integrals. For input vectors $\boldsymbol{X}$ which have a uniform distribution or a Gaussian distribution, CUT is preferred to evaluate higher order classes of metrics considering it requires fewer number of points than GQ. However, if the input variables have distributions which are not uniform or Gaussian, GQ should be adopted. Although other sparse quadrature rules have not been discussed, they can also be used instead of GQ to calculate the higher order classes.

3.3 Summarized Review of NS evaluation

This subsection now elaborates the step by step process needed from start to finish to yield the desired NS metrics using results from all the previous sections.

Step 1: Develop the surrogate model $\hat{Y}$ from the model equation $Y = g(\boldsymbol{X})$ using PC.

Step 2: Determine the output pdf: $f_{\hat{Y}}(\hat{y})$. This is done by sampling the input space $\Omega_{\boldsymbol{X}}$, evaluating the surrogate function at each of those samples and plotting the histogram of the outputs $\hat{y}$.

$\overline{\mathcal{D}}_i$ as well as $\overline{\mathcal{D}}_{i_1,\cdots,i_n}$ are approximated by a weighted sum of $\mathcal{D}$s evaluated at strategic points as represented by equations (20) through (22). The j^{th} sample point for $\overline{\mathcal{D}}_i$ is denoted by $x_i^{(j)}$. For each $x_i^{(j)}$ we need $\mathcal{D}(f_{\hat{Y}}(\hat{y}), f_{\hat{Y}|X_i}(\hat{y}, x_i^{(j)}))$ and for each $\mathcal{D}(f_{\hat{Y}}(\hat{y}), f_{\hat{Y}|X_i}(\hat{y}, x_i^{(j)}))$ we need the pdf: $f_{\hat{Y}|X_i}(\hat{y}, x_i^{(j)})$. This pdf is approximated by sampling the $\Omega_{\tilde{\boldsymbol{X}}_{\boldsymbol{i}}}$ space, evaluating the surrogate model at the samples by holding $X_i = x_i^{(j)}$ and finally plotting the histogram of $\hat{y}$. Each value of $\mathcal{D}$ obtained from $f_{\hat{Y}|X_i}(\hat{y}, x_i^{(j)})$ is then stored for assimilation later. In a similar way, for the higher order class of metrics, the surrogate model is randomly sampled while holding the joint input variables at specific values. Subsequently plotting the histogram yields the conditional pdfs.

Step 3: Obtain the weighted sum of all the stored $\mathcal{D}$s to yield the n^{th} class metrics.

Step 4: Determine the normalized metrics NS using equations (9) through (11).

4. GSA OF T1D MODEL

This section presents the global sensitivity analysis study on the Bergman minimal model Bergman et al. (1981), using the NS metrics. Comparisons are also made with total effect Sobol' indices. Finally, based on the relative ranking of the input uncertain parameters, inferences are drawn about the most influential factor that impacts the blood glucose concentration.

Table 2. Type 1 Diabetes Model Parameter

Parameter	Value	Parameter	Value
$p_1^{nominal}$	0.0287 ($\pm$0.0086)	G_b	119.1858
$p_2^{nominal}$	0.0283 ($\pm$0.0085)	I_b	15.3872
$p_3^{nominal}$	$5.035(\pm 1.51)E-5$	d	0.05
p_4	5/54	$G_0^{nominal}$	119.18 ($\pm$35.75)

4.1 Model and Simulation Environment

In this section, we look at the simple Bergman model for our numerical analysis regarding T1D. The following equations for the Bergman model are considered after accounting for a basal insulin input rate:

$$\dot{G}(t) = -(X_b(t) + p_1)G(t) + p_1 G_b + D(t) \tag{23}$$

$$\dot{X}_b(t) = -p_2 X(t) + p_3(I(t) - I_b) \tag{24}$$

$$\dot{I}(t) = -p_4(I(t) - I_b) + U(t) \tag{25}$$

where X_b is an intermediate state to capture the interaction between the blood glucose (G) and insulin (I) concentration. The meal disturbance $D(t)$ to the glucose concentration term is determined using the Fisher model (Fisher (1991)):

$$D(t) = \begin{cases} 0 & t < t_m \\ Be^{-d(t-t_m)} & t \geq t_m. \end{cases} \tag{26}$$

For this study, the meal time t_m is considered to be constant at $t_m = 15$ minutes. B which characterizes the meal quantity is also assumed to be constant with a value of $B = 28.98$ corresponding to a $45gm$ CHO meal (the value of B is derived identical to Nandi et al. (2017)). To compensate for the meal at time t_m, an insulin bolus at the start of the simulation $t = 0$ is administered. This action is simulated by making the insulin input term ($U(t)$) an impulse function lasting for a minute (between $t = 0$ to $t = 1$). The magnitude of the impulse function is determined using the following formula.

$$U(t) = \begin{cases} \dfrac{1000 \times (CHO\ Amount\ in\ g)}{CR \times V_i} & 0 < t \leq 1 \\ 0 & 1 < t \end{cases} \tag{27}$$

where CR is the insulin-to-carb ratio and V_i is the distribution volume of insulin. For a meal comprising $45gm$ of CHO, a patient $CR = 18.477$ and $V_i = 12$, the impulse magnitude turns out to be $202.96\frac{mU}{L}$.

We assume four sources of uncertainties in this sensitivity analysis study. The uncertainties lie with the initial glucose concentration (i.e. $G(t = 0)$) and parameters p_1, p_2 and p_3. Each uncertain parameter is assumed to have a uniform distribution with a nominal mean and a support which varies $\pm 30\%$ about its mean (as an illustrative variation. A more realistic distribution derived from clinical trial data would ideally be more appropriate). Table 2 lists all the parameters used for the simulations. The mean values of the uncertain parameters are marked by the word *nominal*. In terms of the convention used in the paper, the uncertainties are grouped as

$$\boldsymbol{X} = [X_1, X_2, X_3, X_4]^T = [p_1, p_2, p_3, G_0]^T. \tag{28}$$

The objective now is to observe the contribution of each of these uncertainties (as well their joint contribution) towards the uncertainty in the blood glucose concentration over time.

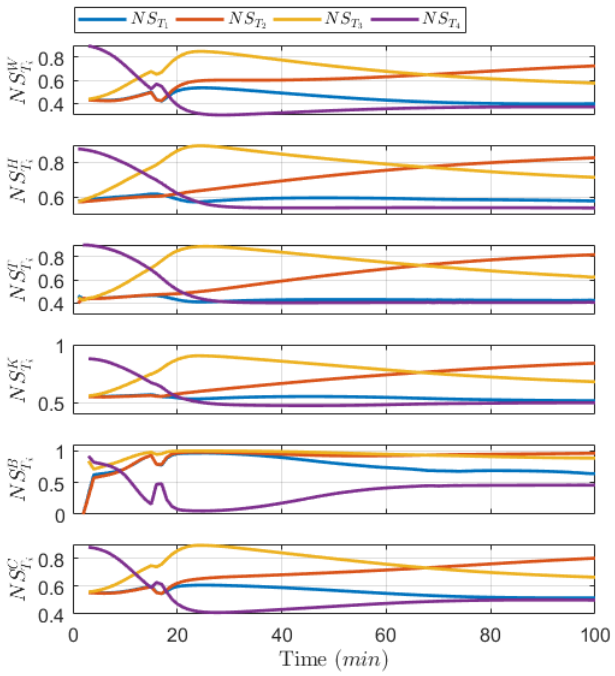

Fig. 1. Variation of NS_{T_i} with time

4.2 Uncertainty Quantification of the Bergman Model

Since there are only four uncertain parameters in our model of study, PC is sufficient in quantifying the propagation of uncertainties through time. Consequently, a PC surrogate model is developed with the idea that the surrogate would be far cheaper to sample as compared to the dynamic system. The basis functions chosen are that of multivariate Legendre polynomials (as recommended by the Wiener-Askey scheme for uniformly distributed inputs) where the multivariate bases are derived from the tensor product of univariate Legendre polynomials. The PC order in each univariate direction is chosen to be $N_{X_i} = 5$. Since the final set of bases is derived from a tensor product of the univariate bases set, the total number of bases become $N = N_{X_1}N_{X_2}N_{X_3}N_{X_4} = 5^4 = 625$.

For this problem, the PC coefficients $Y_i(t)$ of the surrogate model

$$\hat{G}(t) = \hat{Y}(t) = \sum_{i=1}^{625} Y_i(t)\Phi_i(X_1, X_2, X_3, X_4) \qquad (29)$$

are determined using the non-intrusive least squares approach (refer to equation 17). The number of samples used for the method were $m = 100000$. To verify that the surrogate model is of acceptable accuracy, the first couple of moments determined from the PC coefficients and 10000 MC samples of the original stochastic dynamic system were compared and were found to be consistent.

4.3 Computation of NS metrics

Since there are four uncertain input variables, i.e. $n = 4$, we have a total of 15 NS metrics, namely: the class 1 metrics (NS_1, NS_2, NS_3 and NS_4), class 2 metrics ($NS_{1,2}$, $NS_{1,3}$, $NS_{1,4}$, $NS_{2,3}$, $NS_{2,4}$ and $NS_{3,4}$), class 3 metrics ($NS_{1,2,3}$, $NS_{1,2,4}$, $NS_{1,3,4}$ and $NS_{2,3,4}$) and class 4 metric ($NS_{1,2,3,4}$). Each of these metrics are functions of time and hence need evaluation at every instant. The objective is to observe the evolution of the relative importance of each uncertain input with time.

For the class 1 metrics, a 30 point Gauss-quadrature rule is used to average all the statistical distance measures $\mathcal{D}$. As all the input variables are uniformly distributed, we can adhere to CUT for evaluating the average integrals during evaluation of the higher class metrics. There is no need to resort to Gauss quadrature rules. To determine the output pdf and the necessary conditional pdfs, a total of 10000 MC samples are drawn and the histograms of their outputs are recorded.

After each NS metric from every class is derived, the total effect NS metrics are calculated: and the results have been shown (in Figure 1) for all the statistical distance measures. The total time of simulation has been considered to be 100 minutes. The order of ranking estimated from all the statistical distances are seen to be largely consistent. The minimal variations in switching times could be attributed to the fact that the different measures exercise varying penalties on pdf disparities. We see from the figures that initially the most significant parameter is X_4 which is G_0 or the initial blood glucose concentration level. Such an observation is quite intuitive since $G(t)$ is our output of choice and at $t = 0$, G_0 should be most influential. It is comforting to note that the NS metrics are successful at capturing that. However, as time progresses, the influence of G_0 declines as the effect of the insulin bolus kicks in. This is evident from the fact that X_3 and X_2 start dominating. X_3 and X_2 which are essentially parameters p_3 and p_2 are associated with the insulin sensitivity of a person with T1D; where formally, the insulin sensitivity has been defined to be p_3/p_2 (refer to Bergman et al. (1981)). It is interesting to note that for the most part of the simulation (at least until steady state starts setting in), it is the sensitivity to insulin which is most significant. Eventually, for all the statistical measures, the ranking of importance settles to p_2, p_3, p_1 and G_0.

4.4 Computation of Sobol' Indices

To compare the performance of the NS metrics with standard practices of the GSA community, Sobol' indices are derived for the Bergman model under the same simulation setup. It is rather trivial to approximate the Sobol' indices if a PC surrogate model for the system is available. Mere algebraic and polynomial evaluations of the PC coefficients yield the Sobol' indices Sudret (2008). Since the considered model is a dynamic system, we have Sobol' indices which are functions of time which estimate the fraction of the output variance contributed by each input (or their contributions). For a fair comparison with the total effect NS metrics, the total effect Sobol' indices S_{T_i} are presented in Figure 2.

It is interesting to note that a similar pattern emerges from the total effect Sobol' indices when compared to the NS metrics. We observe a decline in the influence of G_0 from an initial position of dominance followed by the rising influence of p_3 and p_2. The settling order of ranking is also similar, given by: p_2, p_3, p_1 and G_0. As a result, we can conclude that quantifying up to the second moment is enough for the Begman model to make inferences on the influence of parameters.

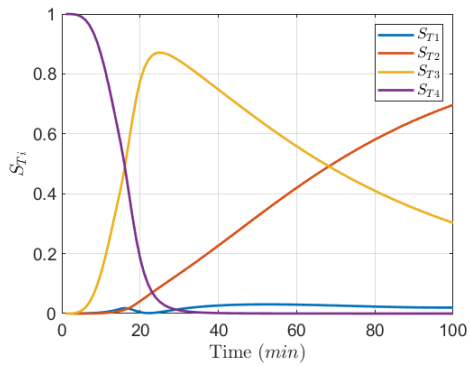

Fig. 2. Variation of the total effect Sobol' indices with time

5. CONCLUSION

This work presents a comprehensive global sensitivity analysis of the Bergman model for type 1 diabetes. The problem is setup for the more traditional mode of treatment where an insulin bolus is administered prior to a meal on top of a basal rate. Certain model parameters have been assumed to be uncertain and the influence of these input uncertainties on the output has been quantified.

As the model is a set of ordinary differential equations, sampling the stochastic system is computationally expensive. As an alternative, polynomial chaos is used to develop a surrogate model to ease the burden of sampling and performing multivariate integrals.

Variance based Sobol' indices as well as non-moment based *NS* metrics are presented to rank the input variables depending on their contributions to the output uncertainty. Since the model is a dynamic system which predicts the blood glucose concentration of a person with type 1 diabetes over time under the effect of a meal and insulin, the GSA measures are evaluated as functions of time. It is seen that the influence each parameter has on the output keeps changing and after a meal, parameters associated with the insulin sensitivity of a person matter the most.

ACKNOWLEDGEMENTS

This material is based upon work supported through the Juvenile Diabetes Research Foundation (JDRF) under JDRF Grant Key: 1-SRA-2019-823-S-B.

REFERENCES

Adurthi, N., Singla, P., and Singh, T. (2018). Conjugate unscented transformation: applications to estimation and control. *Journal of Dynamic Systems, Measurement, and Control*, 140(3), 030907.

Association, A.D. et al. (2018). Economic costs of diabetes in the us in 2017. *Diabetes care*, 41(5), 917–928.

Baringhaus, L. and Henze, N. (2017). Cramér–von mises distance: probabilistic interpretation, confidence intervals, and neighbourhood-of-model validation. *Journal of Nonparametric Statistics*, 29(2), 167–188.

Bergman, R.N., Phillips, L.S., and Cobelli, C. (1981). Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose. *The Journal of clinical investigation*, 68(6), 1456–1467.

Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, 35, 99–109.

Borgonovo, E. and Plischke, E. (2016). Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, 248(3), 869–887.

Cameron, R.H. and Martin, W.T. (1947). The orthogonal development of non-linear functionals in series of fourier-hermite functionals. *Annals of Mathematics*, 48(2), 385–392. URL `http://www.jstor.org/stable/1969178`.

Campolongo, F., Cariboni, J., and Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, 22(10), 1509–1518.

Fisher, M.E. (1991). A semiclosed-loop algorithm for the control of blood glucose levels in diabetics. *IEEE transactions on biomedical engineering*, 38(1), 57–61.

Gerstner, T. and Griebel, M. (1998). Numerical integration using sparse grids. *Numerical algorithms*, 18(3-4), 209.

Ghanem, R.G. and Spanos, P.D. (1991). *Stochastic Finite Elements: A spectral Approach*. Springer New York.

Gibbs, A.L. and Su, F.E. (2002). On choosing and bounding probability metrics. *International statistical review*, 70(3), 419–435.

Kim, K.K.K., Shen, D.E., Nagy, Z.K., and Braatz, R.D. (2013). Wiener's polynomial chaos for the analysis and control of nonlinear dynamical systems with probabilistic uncertainties [historical perspectives]. *IEEE Control Systems*, 33(5), 58–67.

Nandi, S., Singh, T., Mastrandrea, L.D., and Singla, P. (2017). Optimal meal time after bolusing for type 1 diabetes patients under meal uncertainties. In *2017 American Control Conference (ACC)*, 4412–4417. IEEE.

Palumbo, P., Ditlevsen, S., Bertuzzi, A., and De Gaetano, A. (2013). Mathematical modeling of the glucose–insulin system: A review. *Mathematical biosciences*, 244(2), 69–81.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2), 259–270.

Saltelli, A., Chan, K., Scott, E.M., et al. (2000). *Sensitivity analysis*, volume 1. Wiley New York.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley & Sons.

Stroud, A.H. (1971). *Approximate calculation of multiple integrals*. Prentice-Hall.

Stroud, A.H. and Secrest, D. (1966). *Gaussian quadrature formulas*. Prentice-Hall.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety*, 93(7), 964–979.

Wiener, N. (1938). The homogeneous chaos. *American Journal of Mathematics*, 60(4), 897–936. URL `http://www.jstor.org/stable/2371268`.

Xiu, D. and Karniadakis, G.E. (2002). The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2), 619–644.