arXiv:2301.04315v1 [stat.AP] 11 Jan 2023

基于多项式混沌的 Shapley效应估计

Adrian Stein[a], Tarunraj Singh[a]

[a]*Department of Mechanical and Aerospace Engineering, University at Buffalo (SUNY), Buffalo, NY 14260-4400, USA*

摘要

本文提出一种估计 Shapley效应的方法，将其用作全局敏感性度量，以量化不确定模型参数的相对重要性。本文提出采用多项式混沌展开（一种成熟的代理模型构建方法）来估计 Shapley效应。多项式混沌能够将随机过程转化为确定性模型，进而高效地计算关注量的统计矩。这些统计矩包括条件方差，可通过代数运算映射为 Shapley效应。本文通过蒙特卡罗模拟验证基于多项式混沌的 Shapley效应估计，并在基准 Ishigami函数、动态SEIR流行病模型和 Bergman 1型糖尿病模型上进行测试。结果表明，与 Sobol 指数相比，该方法能够正确地对 Ishigami函数中的不确定变量进行排序；同时也展示了动态模型中模型参数随时间变化的排序。

关键词：不确定性量化、敏感性分析、Sobol、Shapley、多项式混沌展开

1. 引言

不确定性无处不在，通常分为偶然不确定性（不可约）和认知不确定性（可约）。不确定性量化（UQ）涵盖不确定性表征，包括对不确定性进行参数化或非参数化表示；还涵盖不确定性传播，指综合模型输入中的不确定性所导致的输出不确定性的过程。UQ 还包括全局敏感性分析，旨在根据不确定输入对输出不确定性的贡献对其进行排序。

受模型降阶、不确定性下的决策、可靠性工程和可解释机器学习等需求的推动，对系统或过程不确定性进行量化受到了广泛关注 [1]。局部敏感性分析是一种已有数十年历史的方法，用于评估系统输出相对于标称参数发生微小扰动时的变化。同时，全局敏感性分析（GSA）致力于对所有输入变量在其作用域内的重要性进行排序。需要注意的是，GSA需要预先了解不确定输入变量的分布。由于上世纪末计算能力显著提升，GSA开始受到广泛关注；当时最常见的两种GSA方法是：（1）Morris方法 [2] 和（2）Sobol分析 [3]。后者即Sobol分析，是一种基于方差的分析，要求输入变量彼此独立。Sobol和Gershman [4] 提出了一种基于导数的方法，该方法在不确定域上进行积分，并将积分结果作为全局敏感性度量。Kucherenko对其进行了修改 [5] ，并将其称为基于导数的全局敏感性度量（DGSM）。与此同时，Sobol和Kucherenko揭示了Sobol 指数与DGSM之间的关系[6], ，这对于高维系统十分有用 [7]。Nandi等人 [8] 提出采用一种称为共轭无迹变换的基于抽样的方法进行数值积分，以更高效地评估Kucherenko的 [5] DGSM度量。

*通讯作者*Email address:* tsingh@buffalo.edu (Tarunraj Singh)

全局敏感性指数通常通过蒙特卡洛（MC）抽样或准蒙特卡罗抽样进行评估，最近则多采用多项式混沌展开（PCE）。蒙特卡洛方法采用“抽取并冻结”方法 [3][9] ，而PCE代理模型可以通过侵入式方法或非侵入式方法构建。在侵入式方法中，通过变量分离展开控制方程并应用伽辽金投影，这被称为侵入式方法。非侵入式方法是一种基于抽样的方法，同时由于其属于黑箱方法，编程相对容易 [10] [11]。然而，对于包含许多输入变量的系统，众所周知，PCE可能会带来较大的计算负担 [12]。Sudret进一步揭示了GSA与PCE之间更深层次的联系 [12], ，他指出可以将Sobol 指数通过代数方式映射到多项式混沌（PC）系数上。受通过蒙特卡洛或准蒙特卡洛计算DGSM度量所带来的计算负担启发，Sudret和Mai [13] 也找到了利用多项式混沌系数计算DGSM度量的方法。

自认识到全局敏感性分析应成为数学建模不可或缺的一部分以来，它已应用于诸多不同领域 [1]。Harman和Johnston [14] 详细介绍了如何将侵入式多项式混沌展开应用于动力系统。他们选用了易感者-感染者-康复者（SIR）疾病模型，将其结果与蒙特卡洛结果进行比较，同时开展了Sobol分析。更复杂的疾病模型是易感者-暴露者-感染者-康复者（SEIR）疾病模型，Jensen等人 [15]采用了该模型。他们基于新冠肺炎传播的丹麦数据，通过多项式混沌展开开展了Sobol分析。为更快地计算Sobol 指数，Blatman和Sudret [16] 提出了稀疏PCE，Lin等人 [17] 将其用于锂离子电池的三维多物理场模型。在控制工程领域，Singh等人 [18] 使用多项式混沌展开设计鲁棒输入整形器。Sobol分析也被用于分析纯电动汽车的能耗 [19], 、贝叶斯网络 [20], 、1型糖尿病建模 [21], 、荷电状态估计 [22]、锂离子电池的热失控 [23] [24], 、水下滑翔机 [25] ，以及预混燃烧器的氮氧化物浓度敏感性分析 [26]。

Borgonovo [27] 提出了一种非矩类全局敏感性度量，通过比较两个概率密度函数（pdf）之间的距离，对不确定输入变量的相对重要性进行排序。Nandi和Singh [28] 还提出了多种基于统计距离的非矩类全局敏感性度量，包括Wasserstein距离、Hellinger距离和Kolmogorov距离。他们提出使用多项式混沌代理模型，以减轻估计统计距离所带来的计算负担。这两篇论文以及Chun等人 [29] 的一篇论文均证实，Sobol 指数对Ishigami函数中三个不确定变量给出的排序是不正确的。基于方差的Sobol指数按重要性由高到低将变量排序为“1,2,3”，而非矩类指标则将其排序为“2,1,3”。这种差异可归因于Ishigami函数具有双峰概率密度函数这一特征，基于方差的Sobol度量无法对此进行正确排序。

Shapley值是 Lloyd Shapley 于1953 [30] 提出的一个概念。在合作博弈中，它根据参与者对联盟的价值，为参与者公平地分配收益。Shapley效应通过考虑参与者可能加入的所有联盟，量化每位参与者的平均边际贡献。将 Shapley效应用作全局敏感性度量的代理指标，其中一个优势是可以处理相关的输入变量 [31]。对于大量不确定变量，使用 Shapley效应的缺点是：评估参与者在大量联盟中的边际成本会带来较高的计算成本 [9] [32] [33] [34] [35]。

Owen [36] 指出了 Shapley效应的一个独特性质：Shapley效应介于一阶 Sobol 指数与总Sobol指数之间。Iooss 和 Prieur [35] 在研究相关输入变量时发现，“夹层”效应并非总是成立。Owen 指出，Sobol 指数易于计算，可用于界定独立变量的 Shapley效应。

近年来，Shapley效应已用于机器学习中的特征选择、模型可解释性和模型降阶，例如通过剪枝 [37]。可解释人工智能领域正利用 Shapley效应帮助理解机器学习模型的输出决策，因为机器学习模型的一大缺点是其黑箱架构导致模型不透明。一个经典的应用案例是为拟建跑道分配着陆费：根据飞机尺寸确定费用，而飞机尺寸与所需跑道长度相关 [38]。Shapley效应已应用于水力模型 [39], 生物制造过程风险 [40], 孤岛微电网的潮流模型 [41],

药物敏感性 [42], 地裂缝 [43], 以及疲劳叶片载荷估算 [44]。一些研究人员比较了 Sobol 指数和 Shapley效应，例如: [45], [46], [47], [35]。大多数研究人员使用蒙特卡洛抽样计算 Shapley效应，例如： [48] [49] [50] [33] [34] [25], 仅举几例。

近年来，全局敏感性分析领域涌现出了一些新思路。其中，Gayrard 等人 [51],提出将具有独立增量的随机过程作为输入进行随机处理。Song 等人 [34],则借助半值的概念，建立了一阶效应、总效应与 Shapley效应之间的联系。Benoumechiara 等人 [52] 使用高斯过程（克里金模型）计算 Shapley效应，速度比蒙特卡洛方法更快，并通过小提琴图展示了研究结果。Benoumechiara 等人 [52] 和 Plischke 等人 [32] 采用了与 [29] [27], 略有不同的 Ishigami 函数参数设置，根据 Shapley 效应得出了 “2,1,3” 的排序。Benoumechiara 等人 [52] 指出了 Sobol 指数与 Shapley效应在线性高斯模型中的关系，并强调了 Sudret 关于根据多项式混沌系数计算 Sobol 指数的研究 [12] 。他们最后指出：“如果能为 Shapley效应建立这样一种分解，将会很有意思。”这是因为，即使使用克里金模型估算 Shapley效应，计算成本仍然很高，尤其是在高维情形下。本文同样致力于建立多项式混沌系数与 Shapley效应之间的关系。我们将介绍两个静态模型：[52] [32] 框架下的 Ishigami函数，以及一个包含四个不确定输入变量的用户定义函数。为展示所提方法在动力学模型中的应用，我们还将研究 SEIR疾病模型和 Bergman 1型糖尿病模型。需要强调的是，本文仅关注独立输入变量，并以超立方体表示不确定域。

本文结构如下：第2 节介绍基于方差的敏感性分析，包括 Sobol 分析、Shapley 值定理、PCE、多项式混沌系数与 Sobol 指数之间的映射，以及 Borgonovo 度量。第3 节介绍静态问题，包括 Ishigami 函数和用户定义函数。第4 节展示了全局敏感性分析在 SEIR 疾病模型和 Bergman 1型糖尿病模型中的应用。结论见第5节。

2. 全局敏感性度量概述

本节回顾了成熟的全局敏感性度量，包括作为一种基于方差的度量的 Sobol 指数，以及作为一种非矩度量的 Borgonovo delta度量。本节还提出将 Shapley效应用作一种度量，并说明如何通过代数运算将多项式混沌系数映射为 Shapley 效应。

2.1. Sobol 指数

假设给定一个函数:

$$y = f(x), \quad x \in \mathcal{R}^n \tag{1}$$

其中， y 是一个标量输出， x 表示输入参数，假定这些参数均为实数。ANOVA[53] 分解可写为 [12] ：

$$f(x_1, ..., x_n) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1 \le i < j \le n} f_{ij}(x_i, x_j) + ... + f_{1,2,...,n}(x_1, ..., x_n) \tag{2}$$

其中 f_0 是一个 c 常数， f_i 完全取决于 x_i 和$f_{i,\ j}$ 完全取决于 x_i 和 x_j。此外，以下条件成立 :

$$\int_0^1 f_{1,2,...,s}(x_1, x_2, ..., x_s) dx_k = 0, \ \ 1 \le k \le s \tag{3}$$

这意味着输入变量彼此独立。可以表述为:

$$f_o = \int_{\mathcal{R}} f(x)dx = \mathbb{E}[y] \tag{4a}$$

$$f_i(x_i) = \int_{\mathcal{R}} f(x)dx_{\sim i} - f_0 = \mathbb{E}[y|x_i] - f_0 \tag{4b}$$

$$f_{ij}(x_i, x_j) = \int_{\mathcal{R}} f(x)dx_{\sim i,j} = \mathbb{E}[y|x_i, x_j] - f_i(x_i) - f_j(x_j) - f_0 \tag{4c}$$

其中， $\sim i$ 表示对除 i 以外的所有变量进行积分，并意味着 x_i 是一个固定的输入变量。显然， $f(x)$的总方差可按如下方式计算:

$$\mathbb{V} = Var(f(x)) = \int_{\mathcal{R}} f^2(x)dx - f_0^2 \tag{5}$$

其中，总方差可以分解为:

$$\mathbb{V} = \sum_{i=1}^{n} \mathbb{V}_i + \sum_{1 \le i < j \le n} \mathbb{V}_{ij} + ... + \mathbb{V}_{1,2,...,n} \tag{6}$$

其中，我们定义集合 s 为 $s = 1$，2，.....，n。与式（4a）-（4c）类似，我们可以将方差表示为:

$$\mathbb{V}_i = \mathbb{V}(\mathbb{E}[y|x_i]) \tag{7}$$

$$\mathbb{V}_{ij} = \mathbb{V}(\mathbb{E}[y|x_i, x_j]) - \mathbb{V}_i - \mathbb{V}_j. \tag{8}$$

现在可以计算 Sobol 指数:

$$S_{i_1,...,i_s} = \frac{\mathbb{V}_{i_1,...i_s}}{\mathbb{V}} \tag{9}$$

其中，已知

$$\sum_{i}^{n} S_i + \sum_{1 \le i < j \le n} S_{ij} + ... + S_{1,2,...n} = 1. \tag{10}$$

总Sobol指数可计算为:

$$S_{T_i} = \sum_{\kappa_i} S_{1,...,s} \tag{11}$$

其中 $\kappa_i = \{(i_1, ..., i_s): \ \exists k, 1 \le k \le s, i_k = i\}$。

2.2. Shapley 值概念

每个变量的 Shapley效应定义为 [35]:

$$Sh_i = \frac{1}{d!} \sum_{u \subseteq N \setminus \{i\}} |u|!(d - 1 - |u|)! \left(c(u \cup \{i\}) - c(u)\right) \tag{12}$$

$$\leftrightarrow Sh_i = \frac{1}{d} \sum_{u \subseteq N \setminus \{i\}} \binom{d-1}{|u|}^{-1} \left(c(u \cup \{i\}) - c(u)\right) \tag{13}$$

其中， d 是变量的数量， $c(...)$ 是联盟的价值。 $c(u \cup \{i\})$ 是包含变量 i 的联盟， $c(u)$ 则是不包含变量 i的联盟。Shapley效应是一种重要性度量，表示变量 i 对联盟 u的边际贡献。Shapley 值概念具有以下特征 [35]:

- 适用于各参与者贡献不相等的情况
- 参与者之间的合作比各自独立行动更有益
- Shapley效应不会为负
- Shapley效应之和为 1

这些特性得出以下结论：

- 联盟 u 至少是空集 $\varnothing$
- 考察方程 (13)，显然，联盟 $|u|$ 不能包含超过 $d-1$ 个参与者，因为它已经是一个子集，并且从集合中排除了参与者 i，因此这类联盟最多有 $d-1$ 个。

Iooss 和 Prieur [35] 提出了一种用于衡量联盟价值的基于方差的度量：

$$c(u) = \frac{Var\left(E[Y|X_u]\right)}{Var(Y)}. \tag{14}$$

通过将不 确定变量视为合作博弈中的参与者，可以将 Shapley效应用作全局敏感 性度量。

2.3. 多项式混沌展开

PCE 的形式如下： [10] [18]

$$y(\zeta) = \sum_{i=0}^{P} a_i \Phi_i(\zeta) \tag{15}$$

其中， y 是输出变量， a_i 是多项式混沌系数， Φ_i 是基函数。 P 是 PCE 的阶数。基函数的正交条件如下： [10]

$$\langle \Phi_i(\zeta), \Phi_j(\zeta) \rangle = \int_\Omega \Phi_i(\zeta)\Phi_j(x)w(\zeta)d\zeta = g_i^2 \delta_{ij} \tag{16}$$

其中 g_i 是由内积得到的系数， δ_{ij} 是克罗内克δ。 $w(x)$ 是 ζ的概率密度函数。维纳–阿斯基方案为标准随机变量提供了合适的多项式基函数，如表 1所示。

表 1：用于多项式混沌展开的维纳–阿斯基方案。 k_l 和 k_u 分别是用户定义的下限和上限。

ζ的分布	多项式族	支撑集
高斯分布	埃尔米特	$(-\infty, \infty)$
伽马分布	拉盖尔	$[0, \infty)$
beta	雅可比	$[k_l, k_u]$
均匀	勒让德	$[k_l, k_u]$

多项式混沌系数可以通过侵入式方法或非侵入式方法确定。侵入式方法需要计算内积，并且对于某些非线性函数，可能无法得到闭式表达式。非侵入式方法可以被视为一种黑箱方法，需要求解最小二乘问题来确定系数。本文采用侵入式方法

其中，系数通过利用伽辽金投影确定。考虑方程（15），其可近似为：

$$y(\zeta) = f(\zeta) \approx f_{PC} = \sum_{i=0}^{P} a_i \Phi_i(\zeta). \tag{17}$$

方程 (17) 左侧表示已知且需要用一个 P <style id='8'>阶多项式进行近似的方程，其系数为 a_i，如方程 (17) 右侧所示。假设 ζ 服从均匀分布，并以选择 $P = 2$ 为例，可得多项式混沌展开：

$$y(\zeta) = f(\zeta) \approx f_{PC} = a_0\Phi_0(\zeta) + a_1\Phi_1(\zeta) + a_2\Phi_2(\zeta). \tag{18}$$

对方程 (18) 进行伽辽金投影，得到：

$$\int f(\zeta)\Phi_0(\zeta)w(\zeta)d\zeta = g_0^2 a_0 \tag{19a}$$

$$\int f(\zeta)\Phi_1(\zeta)w(\zeta)d\zeta = g_1^2 a_1 \tag{19b}$$

$$\int f(\zeta)\Phi_2(\zeta)w(\zeta)d\zeta = g_2^2 a_2. \tag{19c}$$

根据方程 (19a)-(19c)，多项式混沌系数 a_0 至 a_2 可通过除以常数因子 g_0 至g_2求得。对于含有多个随机变量的问题，基函数由一维基函数的张量积确定，如表 2 所示，其中两个随机变量均服从均匀分布。

表 2：系统中两个不确定变量（均服从均匀分布）对应的多项式，次数为 $P =$ 2.

$\Phi_i(\zeta_1)$ / $\Phi_i(\zeta_2)$	1	ζ_1	$(3\zeta_1^2-1)/2$
1	1	ζ_1	$(3\zeta_1^2-1)/2$
ζ_2	ζ_2	$\zeta_1\zeta_2$	
$(3\zeta_2^2-1)/2$	$(3\zeta_2^2-1)/2$		

多项式 Φ_{ij} 表示多项式基函数，其中 i 和 j 分别是 ζ_1 和ζ_2 变量的次数。因此，系统的多项式混沌展开可写为：

$$\begin{aligned} y(\zeta_1,\zeta_2) = f(\zeta_1,\zeta_2) \approx f_{PC} = a_{00}\Phi_{00}(\zeta_1,\zeta_2) + a_{10}\Phi_{10}(\zeta_1,\zeta_2) + a_{01}\Phi_{01}(\zeta_1,\zeta_2)... \\ ... + a_{11}\Phi_{11}(\zeta_1,\zeta_2) + a_{20}\Phi_{20}(\zeta_1,\zeta_2) + a_{02}\Phi_{02}(\zeta_1,\zeta_2). \end{aligned} \tag{20}$$

将所有最高阶为 $P = 2$ 的 Φ_{ij} 组合相乘，并对 ζ_1 和 ζ_2 积分，可得到包含 6 个未知数的 6 个方程，从而可以确定多项式混沌系数 a_{ij}。

2.4. 多项式混沌系数与 Sobol 指数之间的映射

随机输出的均值和方差通过多项式混沌展开表示，可按如下方式计算 [12]:

$$y_{\bar{PC}} = \mathbb{E}\left[f_{PC}(\zeta)\right] = \int_{\mathcal{R}} \sum_{i=0}^{P} a_i\Phi_i(\zeta) \underbrace{\Phi_0(\zeta)}_{=1} w(\zeta)d\zeta = a_0 \tag{21}$$

$$\mathbb{V}_{PC} = Var\left(\sum_{i=0}^{P} a_i\Phi_i(\zeta)\right) = \int_{\mathcal{R}} \left(\sum_{i=0}^{P} a_i\Phi_i(\zeta) \sum_{i=0}^{P} a_i\Phi_i(\zeta) - a_0^2\right) w(\zeta)d\zeta = \sum_{i=1}^{P} a_i^2\, \mathbb{E}\left[\Phi_i^2(\zeta)\right] \tag{22}$$

考虑一个单变量多项式族，并将其记为 $C_P(x)$ $P \in \mathbb{N}$，其中 P为多项式的次数。C_P 例如可以表示勒让德多项式族。当系统中有多个变量存在不确定性时，这一思想可以推广为一个阶数为 M （变量个数）、次数为 P的多变量多项式，其中单变量多项式相乘可构成次数为 P_m的多变量多项式。然而，所得次数 P_m 必须小于或等于 P。因此，我们定义：

$$\alpha = \{\alpha_i : i = 1, ..., M\}, \quad \alpha_i \geq 0, \quad \sum_{i=1}^{M} \alpha_i \leq P \tag{23}$$

其中引入了一个集合， α_i 是自然数，对应于关于变量 x_i的第 i 个一元多项式的次数。随后，多变量多项式 Φ_i 可以表示为 M 个一元多项式的乘积：

$$\Phi_i = \Phi_\alpha(x_1, ..., x_M) = \prod_{i=1}^{M} \Phi_{\alpha_i}(x_i). \tag{24}$$

满足方程 n_C （23）的多项式数量为 [12]:

$$n_C = \frac{(M+P)!}{M!P!}. \tag{25}$$

定义$\psi_{i_1, \ ..., \ i_s}$，它由 n_C α元组组成，其中 i_1、...、 i_s 对应于所选取的变量集合。因此，某些元组元素为零。

$$\psi_{i_1,...,i_s} = \forall \alpha : \begin{cases} \alpha_k > 0 \ \ \text{,when} \ \ k \in \{i_1, ..., i_s\} \\ \alpha_k = 0 \ \ \text{,when} \ \ k \notin \{i_1, ..., i_s\} \end{cases} \tag{26}$$

其中 $k = 1$、...、 M。方程（26）表明，任意 k若属于 $\{i_1, ..., i_s\}$则为非零。现在，的 ANOVA 分解可表示为：

$$\begin{aligned} f_{PC} = a_0 + \sum_{1 \leq i_1 \leq M} \sum_{\alpha \in \psi_{i_1}} a_\alpha \Phi_\alpha(x_{i_1}) + \sum_{1 \leq i_1 < i_2 \leq M} \sum_{\alpha \in \psi_{i_1, i_2}} a_\alpha \Phi_\alpha(x_{i_1}, x_{i_2}) + ... \\ ... + \sum_{1 \leq i_1 < ... < i_s \leq M} \sum_{\alpha \in \psi_{i_1, ..., i_s}} a_\alpha \Phi_\alpha(x_{i_1}, ..., x_{i_s}) + ... + \sum_{\alpha \in \psi_{1,2,...,M}} a_\alpha \Phi_\alpha(x_1, ..., x_M). \end{aligned} \tag{27}$$

请注意，每一项$\sum_{\alpha \in \psi_{i_1, \ ..., \ i_s}} a_\alpha \Phi_\alpha$（$x_{i_1}$、...、 x_{i_s}）可理解为：仅当元组 α完全取决于变量 x_{i_1}、...、 x_{i_s}时，才会考虑项 $a_\alpha \Phi_\alpha$（x_{i_1}、...、 x_{i_s}）。相应地，基于多项式混沌的 Sobol 指数可以计算为 [12]:

$$S_{i_1,...,i_s} = \sum_{\alpha \in \psi_{i_1,...,i_s}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}}. \tag{28}$$

既然现在可以将多项式混沌系数映射到 Sobol 指数，那么是否也可以根据多项式混沌系数计算 Shapley效应呢？可以说，我们只需计算所有联盟的价值$c(\{...\})$。对于具有 n 个未知量的系统，Sobol 指数可以按如下方式映射为所有联盟的价值 $c(\{...\})$

所有联盟如下：

$$c(\{i_1\}) = \sum_{\alpha \in \psi_{i_1}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}} \tag{29a}$$

$$c(\{i_1, i_2\}) = \sum_{\alpha \subset \psi_{i_1,i_2}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}} + \sum_{\alpha \in \psi_{i_1,i_2}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}} \tag{29b}$$

$$\vdots$$

$$c(\{i_1, ..., i_s\}) = \sum_{\alpha \subset \psi_{i_1,...,i_s}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}} + \sum_{\alpha \in \psi_{i_1,...,i_s}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}} \tag{29c}$$

$$\vdots$$

$$c(\{1, 2, .., M\}) = \sum_{\alpha \subset \psi_{1,2,...,M}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}} + \sum_{\alpha \in \psi_{1,2,...,M}} \frac{a_\alpha^2 \, \mathbb{E}\left[\Phi_\alpha^2\right]}{\mathbb{V}_{PC}} \approx 1. \tag{29d}$$

注意，$\alpha \subset \psi_{i_1,...,i_s}$ 表示 $\{i_1, ..., i_s\}$ 的所有真子集的可能组合，而 $\alpha \in \psi_{i_1,...,i_s}$表示 $\{i_1, ..., i_s\}$的交集。对于集合 $\{i_1, ..., i_s\}$，我们可以推导出 $2^s - 1$ 个真子集，因为空集是除自身以外任何集合的真子集。方程（29a）-（29d）中给出的价值可以直接代入方程（12）中。

如果将多项式混沌系数映射到 Sobol 指数，就可以计算 Shapley效应：

$$Sh_i = S_i + \frac{1}{2} \sum_{1 \le i < j \le M} S_{ij} + \frac{1}{3} \sum_{1 \le i < j < k \le M} S_{ijk} + ... + \frac{1}{M-1} \sum_{1 \le i < ... < s \le M} S_{i,...,s} + \frac{1}{M} S_{1,2,...,M}. \tag{30}$$

图1 展示了使用方程（14）和多项式混沌系数计算 Shapley效应的差异。这揭示了 Sobol 指数与 Shapley 效应之间的关系。此外，给定 $c(u)$ 的所有值后，可以映射到相应的 S_i、S_{ij}，.....反之亦然。

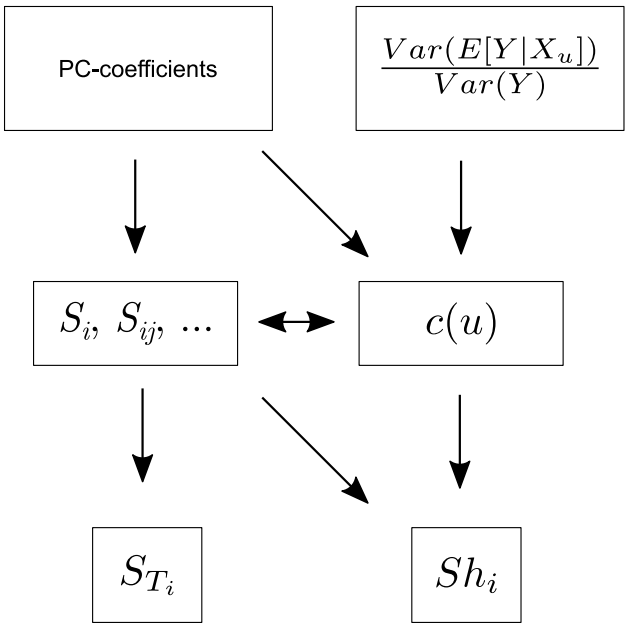


图 1：Sobol 指数与 Shapley效应之间的联系。

2.5. Borgonovo度量

Borgonovometric [27] 是一种用于衡量无条件概率密度函数与条件概率密度函数之间偏移的方法，如图 2所示。

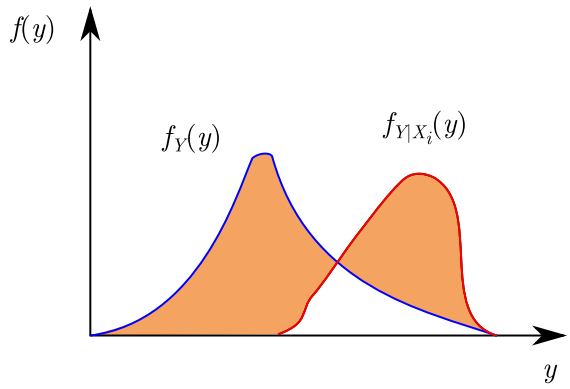


图2：偏移量是两个概率密度函数之间差异的面积（阴影区域 ）.

假设函数 Y 有 i 个不确定变量。随后可利用偏移量对所有不确定变量的重要性进行排序。其基本思路是，对输出 Y 影响较小的变量产生的偏移量较小，而影响较大的变量产生的偏移量较大。两个概率密度函数之间偏移量的定义记作 Z，可写为 [27]:

$$Z(X_i) = \int |f_Y(y) - f_{Y|X_i}(y)| dy \tag{31}$$

其中， $f_Y(y)$ 是输出 Y的概率密度函数，而 $f_{Y|X_i}(y)$ 是给定变量 x_i时的条件概率密度函数。此外， $f_Y(y)$ 与 $f_{Y|X_i}(y)$ 之间的期望偏移量可写为:

$$E_{X_i}\left[Z(X_i)\right] = \int f_{X_i}(x_i) \left(\int |f_Y(y) - f_{Y|X_i}(y)| dy \right) dx_i. \tag{32}$$

最后，可由期望偏移量推导出 Borgonovo指数 [27] ，即:

$$\delta_i = \frac{1}{2} E_{X_i}\left[Z(X_i)\right] \tag{33}$$

其中， δ_i 是针对输出 Y的不确定变量 x_i 的矩无关敏感度度量。

3. 代数示例

本节将介绍两个代数示例，包括 Ishigami函数和一个用户定义函数。

3.1. Ishigami函数

一个基准mark题是 Ishigami函数，尤其是因为其双峰概率密度函数。函数定义如下:

$$Y = \sin(x_1) + a \sin(x_2)^2 + b x_3^4 \sin(x_1) \tag{34}$$

其中 $a = 7$ 和 $b = 0.1$ 是与文献中所用参数一致的参数。 x_1、 x_2 和 x_3 是不确定变量，假设所有变量均在 $-\pi$ 和 π之间均匀分布。Ishigami函数的概率密度函数如图 3所示。

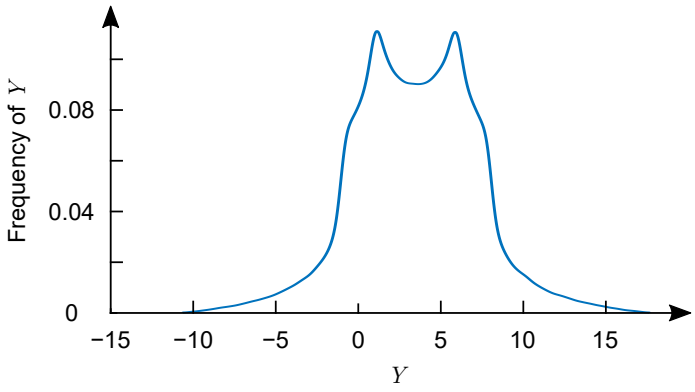


图3：当 $a = 7$ 和 $b = 0.1$时 Ishigami函数的概率密度函数。

期望值和总方差可以通过对所有三个不确定变量进行积分来解析求得：

$$E[Y] = \mu = \left(\frac{1}{2\pi}\right)^3 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \sin(x_1) + 7\sin(x_2)^2 + 0.1x_3^4 \sin(x_1) dx_1 dx_2 dx_3 \tag{35}$$

$$Var(Y) = \left(\frac{1}{2\pi}\right)^3 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \left(\sin(x_1) + 7\sin(x_2)^2 + 0.1x_3^4 \sin(x_1) - \mu\right)^2 dx_1 dx_2 dx_3. \tag{36}$$

为确定 Sobol 指数，需要计算三个一阶、三个二阶和一个三阶指数 。对于 一阶Sobol指数，条件方差可计算如下：

$$E[Y|x_i] = \mu_i = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \sin(x_1) + 7\sin(x_2)^2 + 0.1x_3^4 \sin(x_1) d\mathcal{X}_{\sim i} \tag{37}$$

$$Var(E[Y|x_i]) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\mu_i - \frac{1}{2\pi}\int_{-\pi}^{\pi} \mu_i dx_i\right)^2 dx_i \tag{38}$$

$$\rightarrow \mathbb{V}_i = Var(E[Y|x_i]) \tag{39}$$

其中， $d\mathcal{X}_{\sim i}$ 表示对除 x_i以外的所有变量进行积分。对于二阶 Sobol 指数，条件方差可具体计算如下：

$$E[Y|x_{ij}] = \mu_{ij} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sin(x_1) + 7\sin(x_2)^2 + 0.1x_3^4 \sin(x_1) d\mathcal{X}_{\sim i,j} \tag{40}$$

$$Var(E[Y|x_{ij}]) = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \left(\mu_{ij} - \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mu_{ij} dx_i dx_j\right)^2 dx_i dx_j \tag{41}$$

$$\rightarrow \mathbb{V}_{ij} = Var(E[Y|x_{ij}]) - \mathbb{V}_i - \mathbb{V}_j. \tag{42}$$

对于三阶Sobol指数，条件方差可以具体计算如下：

$$E[Y|x_{123}] = \mu_{123} = \sin(x_1) + 7\sin(x_2)^2 + 0.1x_3^4 \sin(x_1) \tag{43}$$

$$Var(E[Y|x_{123}]) = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \left(\mu_{123} - \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mu_{123} dx_1 dx_2 dx_3\right)^2 dx_1 dx_2 dx_3 \tag{44}$$

$$\rightarrow \mathbb{V}_{123} = Var(E[Y|x_{123}]) - \mathbb{V}_{12} - \mathbb{V}_{13} - \mathbb{V}_{23} - \mathbb{V}_1 - \mathbb{V}_2 - \mathbb{V}_3. \tag{45}$$

解析计算得到的 Sobol 指数列于表3 ， 以便与使用多项式混沌展开代理模型确定的结果进行比较。

表 3：Ishigami函数的 Sobol 指数。

指数	解析	PCE阶数			
		$P=3$	$P=5$	$P=7$	$P=9$
S_1	0.3139	0.6582	0.3699	0.3166	0.3140
S_2	0.4424	0.0540	0.3545	0.4377	0.4423
S_3	0	0	0	0	0
S_{12}	0	0	0	0	0
S_{13}	0.2437	0.2878	0.2756	0.2456	0.2437
S_{23}	0	0	0	0	0
S_{123}	0	0	0	0	0
$S_{T,1}$	0.5576	0.9460	0.6455	0.5623	0.5577
$S_{T,2}$	0.4424	0.0540	0.3545	0.4377	0.4423
$S_{T,3}$	0.2437	0.2878	0.2756	0.2456	0.2437

表3 显示，随着 PC 阶数的增加，Sobol 指数逐渐收敛于解析解。为确定 Shapley效应，也可以使用方程（14）解析计算联盟的价值，结果列于表4。

表 4：Ishigami函数的 Shapley效应。

价值	解析	PCE阶数			
		$P=3$	$P=5$	$P=7$	$P=9$
$c(\{1\})$	0.3139	0.6582	0.3699	0.3166	0.3140
$c(\{2\})$	0.4424	0.0540	0.3545	0.4377	0.4423
$c(\{3\})$	0	0	0	0	0
$c(\{1,2\})$	0.7563	0.7122	0.7244	0.7544	0.7563
$c(\{1,3\})$	0.5576	0.9460	0.6455	0.5623	0.5577
$c(\{2,3\})$	0.4424	0.0540	0.3545	0.4377	0.4423
$c(\{1,2,3\})$	1	1	1	1	1
Sh_1	0.4357	0.8021	0.5077	0.4395	0.4358
Sh_2	0.4424	0.0540	0.3545	0.4377	0.4423
Sh_3	0.1218	0.1439	0.1378	0.1228	0.1219

请注意，空联盟的价值为 $c(\{\varnothing\}) = 0$。与表3中的 Sobol 指数类似，表4 展示了随着 PC 阶数 P 提高，联盟价值逐渐收敛至其解析值，因此也可以观察到 Shapley效应逐渐收敛至解析解。剩下的问题是：计算 Shapley效应与 Sobol分析究竟有何不同？Sobol 和 Shapley 可用于对系统中的不确定变量的重要性进行排序。这意味着，排名较高的不确定变量对输出变量扰动的贡献更大。通过图1中的 Sobol 指数来计算 Shapley效应，我们可以尝试考察这些指数的加权。表5 专门针对 Ishigami函数编制，有助于深入理解这种加权。

表 5：计算 Ishigami函数的总Sobol指数和 Shapley效应时，Sobol 指数加权方式的差异。

Sobol 指数	S_T的权重	Sh的权重	解析
S_1	1	1	0.3139
S_2	1	1	0.4424
S_3	1	1	0
S_{12}	1	1/2	0
S_{13}	1	1/2	0.2437
S_{23}	1	1/2	0
S_{123}	1	1/3	0

如果使用表5 中的解析值，就可以根据 Sobol 指数按如下方式计算 Shapley效应：

$$Sh_1 = S_1 + \frac{1}{2}S_{12} + \frac{1}{2}S_{13} + \frac{1}{3}S_{123} = 0.4357 \tag{46a}$$
$$Sh_2 = S_2 + \frac{1}{2}S_{12} + \frac{1}{2}S_{23} + \frac{1}{3}S_{123} = 0.4424 \tag{46b}$$
$$Sh_3 = S_3 + \frac{1}{2}S_{13} + \frac{1}{2}S_{23} + \frac{1}{3}S_{123} = 0.1218. \tag{46c}$$

相反，总Sobol指数可按如下方式计算：

$$S_{T,1} = S_1 + S_{12} + S_{13} + S_{123} = 0.5576 \tag{47a}$$
$$S_{T,2} = S_2 + S_{12} + S_{23} + S_{123} = 0.4424 \tag{47b}$$
$$S_{T,3} = S_3 + S_{13} + S_{23} + S_{123} = 0.2437. \tag{47c}$$

将方程（46a）—（46c）与方程（47a）—（47c的结果进行比较可以看出，与每个指数权重相同的 Sobol 方法相比，Shapley 对高阶Sobol指数赋予了较低的权重。从 Sobol 的角度评估排序，建议将不确定变量排列为 x_1、 x_2 和 x_3；而 Shapley 建议的顺序是x_2、 x_1 和 x_3。事实上，如果仅存在一阶Sobol指数，而高阶指数均为零，那么总Sobol指数将与 Shapley效应相同。

正如 [32] 和 [52], 所指出的，不确定变量 x_2 比 x_1更重要。评估不确定变量排序最明确的方法，是按照小节 2.5 中所述计算 Borgonovo指数，或计算其他统计距离 [28] ，因为这些方法基于概率密度函数，计算成本可能较高。进行 300,000 次蒙特卡罗评估以计算 Borgonovo指数后发现，不确定变量的排序为 x_2、x_1 和 x_3，这与 Shapley 建议的排序相同。三种方法的结果均列于表 6。

表 6：Ishigami函数的总Sobol指数、Shapley效应和Borgonovo指数。

变量	总Sobol指数	Shapley效应	Borgonovo指数
x_1	0.5576	0.4357	0.2392
x_2	0.4424	0.4424	0.4222
x_3	0.2437	0.1218	0.1957

3.2. 用户定义函数

我们考虑一个多项式函数，以说明 Shapley效应的评估，其中 ：

$$Y = 1.9x_1^2 + 2x_2^2 + 1.05x_3x_1^3 + 0.35x_4 \tag{48}$$

假设 x_1 至 x_4 服从均匀分布，且每个变量的取值介于 -1 和 1 之间。图 4展示了用户定义函数的概率密度函数。

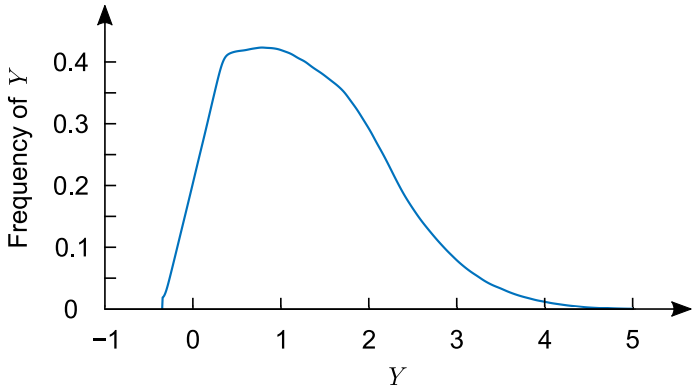


图4：用户定义函数的概率密度函数。

Sobol 指数列于表7 ，显然，当 PCE阶数为 $P = 4$时，我们可得到精确的解析解。由于用户定义函数的最高阶数为 4，阶数为 4 的多项式混沌展开可以精确复现原函数。请注意，除 S_{13}外，所有高阶Sobol指数均为 0。

表7：用户定义函数的 Sobol 指数。

指数	解析	PCE阶数			
		$P = 1$	$P = 2$	$P = 3$	$P = 4$
S_1	0.4169	0	0.4215	0.4215	0.4169
S_2	0.4619	0	0.4670	0.4670	0.4169
S_3	0	0	0	0	0
S_4	0.0530	1	0.0536	0.0536	0.0530
S_{12}	0	0	0	0	0
S_{13}	0.0682	0	0.0579	0	0.0682
S_{14}	0	0	0	0	0
S_{23}	0	0	0	0	0
S_{24}	0	0	0	0	0
S_{34}	0	0	0	0	0
S_{123}	0	0	0	0	0
S_{124}	0	0	0	0	0
S_{134}	0	0	0	0	0
S_{234}	0	0	0	0	0
S_{1234}	0	0	0	0	0
$S_{T,1}$	0.4851	0	0.4794	0.4794	0.4851
$S_{T,2}$	0.4619	0	0.4670	0.4670	0.4619
$S_{T,3}$	0.0682	0	0.0579	0.0579	0.0682
$S_{T,4}$	0.0530	1	0.0536	0.0536	0.0530

表8 展示了使用更高阶PCE时联盟价值和Shapley效应的收敛情况。同样，当PCE阶数为 $P = 4$时，可以观察到多项式混沌展开近似与解析结果完全相同。

表8：用户定义函数的 Shapley效应。

价值	解析	PCE阶数			
		$P=1$	$P=2$	$P=3$	$P=4$
$c(\{1\})$	0.4169	0	0.4215	0.4215	0.4169
$c(\{2\})$	0.4619	0	0.4670	0.4670	0.4619
$c(\{3\})$	0	0	0	0	0
$c(\{4\})$	0.0530	1	0.0536	0.0536	0.0530
$c(\{1,2\})$	0.8788	0	0.8884	0.8884	0.8788
$c(\{1,3\})$	0.4851	0	0.4794	0.4794	0.4851
$c(\{1,4\})$	0.4699	1	0.4751	0.4751	0.4699
$c(\{2,3\})$	0.4619	0	0.4670	0.4670	0.4619
$c(\{2,4\})$	0.5149	1	0.5206	0.5206	0.5149
$c(\{3,4\})$	0.0530	1	0.0536	0.0536	0.0530
$c(\{1,2,3\})$	0.9470	0	0.9464	0.9464	0.9470
$c(\{1,2,4\})$	0.9318	1	0.9421	0.9421	0.9318
$c(\{1,3,4\})$	0.5381	1	0.5330	0.5330	0.5381
$c(\{2,3,4\})$	0.5149	1	0.5206	0.5206	0.5149
$c(\{1,2,3,4\})$	1	1	1	1	1
Sh_1	0.4510	0	0.4504	0.4504	0.4510
Sh_2	0.4619	0	0.4670	0.4670	0.4619
Sh_3	0.0341	0	0.0290	0.0290	0.0341
Sh_4	0.0530	1	0.0536	0.0536	0.0530

从Sobol的角度评估排序，建议将不确定变量依次排列为 x_1、 x_2、 x_3和 x_4；而Shapley效应建议的排序为 x_2、 x_1、 x_4 和 x_3。表9 给出了Borgonovo指数，其排序与基于Shapley效应的排序相同。

表9：用户定义函数的总Sobol指数、Shapley效应和Borgonovo指数。

变量	总Sobol	Shapley	Borgonovo
x_1	0.4851	0.4510	0.2734
x_2	0.4619	0.4619	0.3309
x_3	0.0682	0.0341	0.0178
x_4	0.0530	0.0530	0.0800

4. 动态系统

动态模型的全局敏感性分析对于控制和预测问题至关重要。因此，为了便于说明，本节将介绍两个模型：SEIR疾病模型 [15] 以及用于模拟1型糖尿病的Bergman模型。

4.1. SEIR模型

非线性SEIR模型由以下方程表示:

$$\frac{\partial S}{\partial t} = -\beta \frac{I(t)S(t)}{N} \tag{49a}$$
$$\frac{\partial E}{\partial t} = \beta \frac{I(t)S(t)}{N} - \sigma E(t) \tag{49b}$$
$$\frac{\partial I}{\partial t} = \sigma E(t) - \gamma I(t) \tag{49c}$$
$$\frac{\partial R}{\partial t} = \gamma I(t) \tag{49d}$$

其中，假设 β、 σ 和 γ 为不确定参数。 S、 E、 I 和 R 分别表示易感者、暴露者、感染者和康复组。变量 N 表示总人口，为简化起见，设为 1。不确定参数可解释如下:

- $\gamma = \frac{1}{\tau_{inf}}$，其中 τ_{inf} 是传染状态的时间常数
- $\sigma = \frac{1}{\tau_{inc}}$，其中 τ_{inc} 是暴露状态的时间常数
- $\beta = R_o\gamma$，其中 R_o 是再生数

我们假设它们服从如下均匀分布 [14]:

$$\beta \sim U(2.5, 5.5) \tag{50a}$$
$$\sigma \sim U(0.5, 1.5) \tag{50b}$$
$$\gamma \sim U(0.5, 1.5) \tag{50c}$$

这还可以进一步改写为:

$$\beta = \beta_0 + \beta_1\zeta_1 \tag{51a}$$
$$\sigma = \sigma_0 + \sigma_1\zeta_2 \tag{51b}$$
$$\gamma = \gamma_0 + \gamma_1\zeta_3 \tag{51c}$$

其中，下标 $(.)_0$ 表示均值，$(.)_1$ 用于将具有任意边界的均匀分布映射到区间 $[-1,\ 1]$。随机SEIR变量的多项式混沌展开为:

$$S(t,\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty} S_i(t)\Phi(\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} S_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) \tag{52a}$$
$$E(t,\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty} E_i(t)\Phi(\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} E_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) \tag{52b}$$
$$I(t,\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty} I_i(t)\Phi(\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} I_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) \tag{52c}$$
$$R(t,\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty} R_i(t)\Phi(\zeta_1,\zeta_2,\zeta_3) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} R_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3). \tag{52d}$$

多项式 $\Phi(\zeta_1,\zeta_2,\zeta_3)$由 ζ_1、 ζ_2 和 ζ_3的勒让德多项式的所有张量积组成。现在，常微分方程可表示为:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{dS_{ijk}(t)}{dt}\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) = -\left(\beta_0+\beta_1\zeta_1\right)\ldots$$

$$\ldots\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{q=0}^{\infty}S_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3)\times I_{mnq}(t)\Phi_m(\zeta_1)\Phi_n(\zeta_2)\Phi_q(\zeta_3) \tag{53a}$$

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{dE_{ijk}(t)}{dt}\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) = \left(\beta_0+\beta_1\zeta_1\right)\ldots$$

$$\ldots\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{q=0}^{\infty}S_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3)\times I_{mnq}(t)\Phi_m(\zeta_1)\Phi_n(\zeta_2)\Phi_q(\zeta_3)\ldots$$

$$\ldots-\left(\sigma_0+\sigma_1\zeta_2\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}E_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) \tag{53b}$$

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{dI_{ijk}(t)}{dt}\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) = \left(\sigma_0+\sigma_1\zeta_2\right)\ldots$$

$$\ldots\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}E_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3)-\left(\gamma_0+\gamma_1\zeta_3\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}I_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) \tag{53c}$$

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{dR_{ijk}(t)}{dt}\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3) = \left(\gamma_0+\gamma_1\zeta_3\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}I_{ijk}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3). \tag{53d}$$

为近似式（49a）-（49d）中所示的原始系统，将每个随机维度上的维纳-阿斯基基函数进行张量积所构成的正交基函数子空间投影到该系统上进行伽辽金投影，可得以下方程:

$$\frac{dS_{uvw}}{dt} = -\frac{\beta_0}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}\sum_{m=0}^{P}\sum_{n=0}^{P-m}\sum_{q=0}^{P-m-n}S_{ijk}(t)I_{mnq}(t)\Psi_2(1)\ldots$$

$$\ldots-\frac{\beta_1}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}\sum_{m=0}^{P}\sum_{n=0}^{P-m}\sum_{q=0}^{P-m-n}S_{ijk}(t)I_{mnq}(t)\Psi_2(\zeta_1) \tag{54a}$$

$$\frac{dE_{uvw}}{dt} = \frac{\beta_0}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}\sum_{m=0}^{P}\sum_{n=0}^{P-m}\sum_{q=0}^{P-m-n}S_{ijk}(t)I_{mnq}(t)\Psi_2(1)\ldots$$

$$\ldots+\frac{\beta_1}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}\sum_{m=0}^{P}\sum_{n=0}^{P-m}\sum_{q=0}^{P-m-n}S_{ijk}(t)I_{mnq}(t)\Psi_2(\zeta_1)-\sigma_0E_{uvw}(t)-\frac{\sigma_1}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}E_{ijk}(t)\Psi_1\left(\zeta_2\right) \tag{54b}$$

$$\frac{dI_{uvw}}{dt} = \sigma_0E_{uvw}(t)+\frac{\sigma_1}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}E_{ijk}(t)\Psi_1\left(\zeta_2\right)-\gamma_0I_{uvw}(t)-\frac{\gamma_1}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}I_{ijk}(t)\Psi_1\left(\zeta_3\right) \tag{54c}$$

$$\frac{dR_{uvw}}{dt} = \gamma_0I_{uvw}(t)+\frac{\gamma_1}{\Psi_3}\sum_{i=0}^{P}\sum_{j=0}^{P-i}\sum_{k=0}^{P-i-j}I_{ijk}(t)\Psi_1\left(\zeta_3\right). \tag{54d}$$

Ψ_1、 Ψ_2 和 Ψ_3 的计算方式如下：

$$\Psi_1\left(F\left(\zeta_1,\zeta_2,\zeta_3\right)\right)=\int_{\Omega_3}\int_{\Omega_2}\int_{\Omega_1}F\left(\zeta_1,\zeta_2,\zeta_3\right)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3)\Phi_u(\zeta_1)\Phi_v(\zeta_2)...$$
$$...\times\Phi_w(\zeta_3)p_1\left(\zeta_1\right)p_2\left(\zeta_2\right)p_3\left(\zeta_3\right)d\zeta_1d\zeta_2d\zeta_3 \tag{55}$$

$$\Psi_2\left(F\left(\zeta_1,\zeta_2,\zeta_3\right)\right)=\int_{\Omega_3}\int_{\Omega_2}\int_{\Omega_1}F\left(\zeta_1,\zeta_2,\zeta_3\right)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3)\Phi_m(\zeta_1)\Phi_n(\zeta_2)\Phi_q(\zeta_3)...$$
$$...\times\Phi_u(\zeta_1)\Phi_v(\zeta_2)\Phi_w(\zeta_3)p_1\left(\zeta_1\right)p_2\left(\zeta_2\right)p_3\left(\zeta_3\right)d\zeta_1d\zeta_2d\zeta_3 \tag{56}$$

$$\Psi_3=\int_{\Omega_3}\int_{\Omega_2}\int_{\Omega_1}\left(\Phi_u\left(\zeta_1\right)\right)^2\left(\Phi_v\left(\zeta_2\right)\right)^2\left(\Phi_w\left(\zeta_3\right)\right)^2p_1\left(\zeta_1\right)p_2\left(\zeta_2\right)p_3\left(\zeta_3\right)d\zeta_1d\zeta_2d\zeta_3 \tag{57}$$

和 p_1、 p_2 和 p_3 分别是 ζ_1、 ζ_2 和 ζ_3 的概率密度函数。请注意， Ψ_3 只是一个系数，它是将方程（53a）-（53d）左侧投影到正交基函数上后得到的。

4.1.1. 均值和方差

为便于说明，考虑一组感染者。均值（期望值）可按如下方式计算：

$$\begin{aligned}\mathbb{E}\left[I(t,\zeta_1,\zeta_2,\zeta_3)\right]&=\mathbb{E}\left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}I_{ijk}(t)\Phi_1\left(\zeta_1\right)\left(\zeta_2\right)\left(\zeta_3\right)\right]\\&=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}I_{ijk}(t)\int_{\Omega_3}\int_{\Omega_2}\int_{\Omega_1}\Phi_i\left(\zeta_1\right)\Phi_j\left(\zeta_2\right)\Phi_k\left(\zeta_3\right)p_1\left(\zeta_1\right)p_2\left(\zeta_2\right)p_3\left(\zeta_3\right)d\zeta_1d\zeta_2d\zeta_3\\&=I_{000}(t)\end{aligned} \tag{58}$$

方差则为：

$$\mathbb{V}\left[I(t,\zeta_1,\zeta_2,\zeta_3)\right]=\mathbb{E}\left[I(t,\zeta_1,\zeta_2,\zeta_3)^2\right]-\left(\mathbb{E}\left[I(t,\zeta_1,\zeta_2,\zeta_3)\right]\right)^2$$
$$=\mathbb{E}\left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{q=0}^{\infty}I_{ijk}(t)I_{mnq}(t)\Phi_i(\zeta_1)\Phi_j(\zeta_2)\Phi_k(\zeta_3)\Phi_m(\zeta_1)\Phi_n(\zeta_2)\Phi_q(\zeta_3)\right]-I_{000}(t)^2$$
$$=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\left(I_{ijk}\right)^2\int_{\Omega_3}\int_{\Omega_2}\int_{\Omega_1}\left(\Phi_i\left(\zeta_1\right)\right)^2\left(\Phi_j\left(\zeta_2\right)\right)^2\left(\Phi_k\left(\zeta_3\right)\right)^2...$$
$$....p_1\left(\zeta_1\right)p_2\left(\zeta_2\right)p_3\left(\zeta_3\right)d\zeta_1d\zeta_2d\zeta_3-\left(I_{000}(t)\right)^2. \tag{59}$$

此外，我们可以按照方程（7）和（8）所示，直接计算条件方差，从而求得 Sobol 指数和 Shapley效应。

首先，我们比较多项式混沌展开（PCE）与蒙特卡洛估计，以评估使用 PCE 计算全局敏感性分析指标的近似精度。采用基于方差的方法计算敏感度时，只需知道期望值和方差，因此我们从期望值和总方差两方面比较蒙特卡洛方法与 PCE。蒙特卡洛模拟采用了 5,000 个样本；我们认为这些样本足以捕捉系统的不确定性并反映各状态的行为。PCE 的计算阶数达到$P=4$。对于初始条件，假设 $S_{000}=0.99$, $E_{000}=0$, $I_{000}=0.01$ 和 $R_{000}=0$。所有其他多项式混沌状态均为 0。随后，我们任意选取 15 天作为仿真时间。图 5 展示了期望值随时间的变化。红色虚线表示采用蒙特卡洛方法计算的易感者、暴露者、感染者和康复组的结果。蓝线表示 PCE 的结果；PCE 展开阶数越高，实线的蓝色就越深。从图 5左侧可以看出，PCE 近似结果

正逐渐收敛至蒙特卡洛模拟的结果。图 5 右侧展示了蒙特卡洛方法与 PCE 之间的绝对误差随时间的变化。总体而言，绝对误差相当小；不过，平均来看，PCE 阶数越高，随时间变化的误差越小。

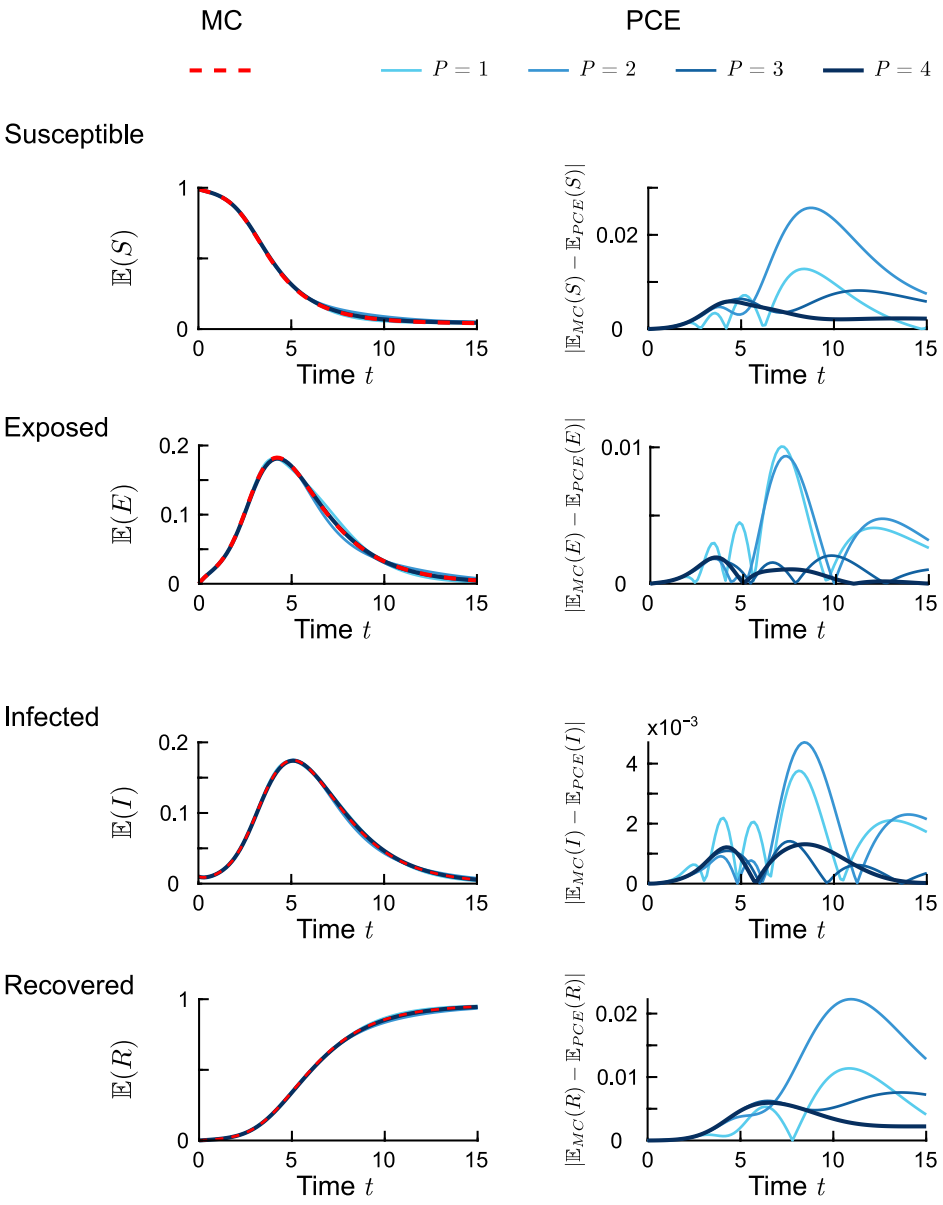


图5：易感者、暴露者、感染者和康复组的期望值：蒙特卡罗模拟（5000 个样本）与多项式混沌展开的比较（左图）。蒙特卡罗模拟（5000 个样本）与多项式混沌展开所得期望值之间的绝对误差（右图）。

图6 通过红色虚线展示了易感组、暴露组、感染组和康复组的蒙特卡洛模拟所产生的总方差随时间的变化。蓝色实线表示多项式混沌展开的模拟结果，随着多项式混沌展开阶数提高，蓝色逐渐加深。与期望值相比，总方差对多项式混沌展开阶数的依赖性更强。图6 左侧清楚地表明，较低的 PC 阶数会导致对蒙特卡洛结果的近似效果不佳；然而，阶数为 $P=4$ 时，结果与蒙特卡洛结果非常接近。图6 展示了

所有阶数和状态下，蒙特卡洛与多项式混沌展开之间的绝对误差。

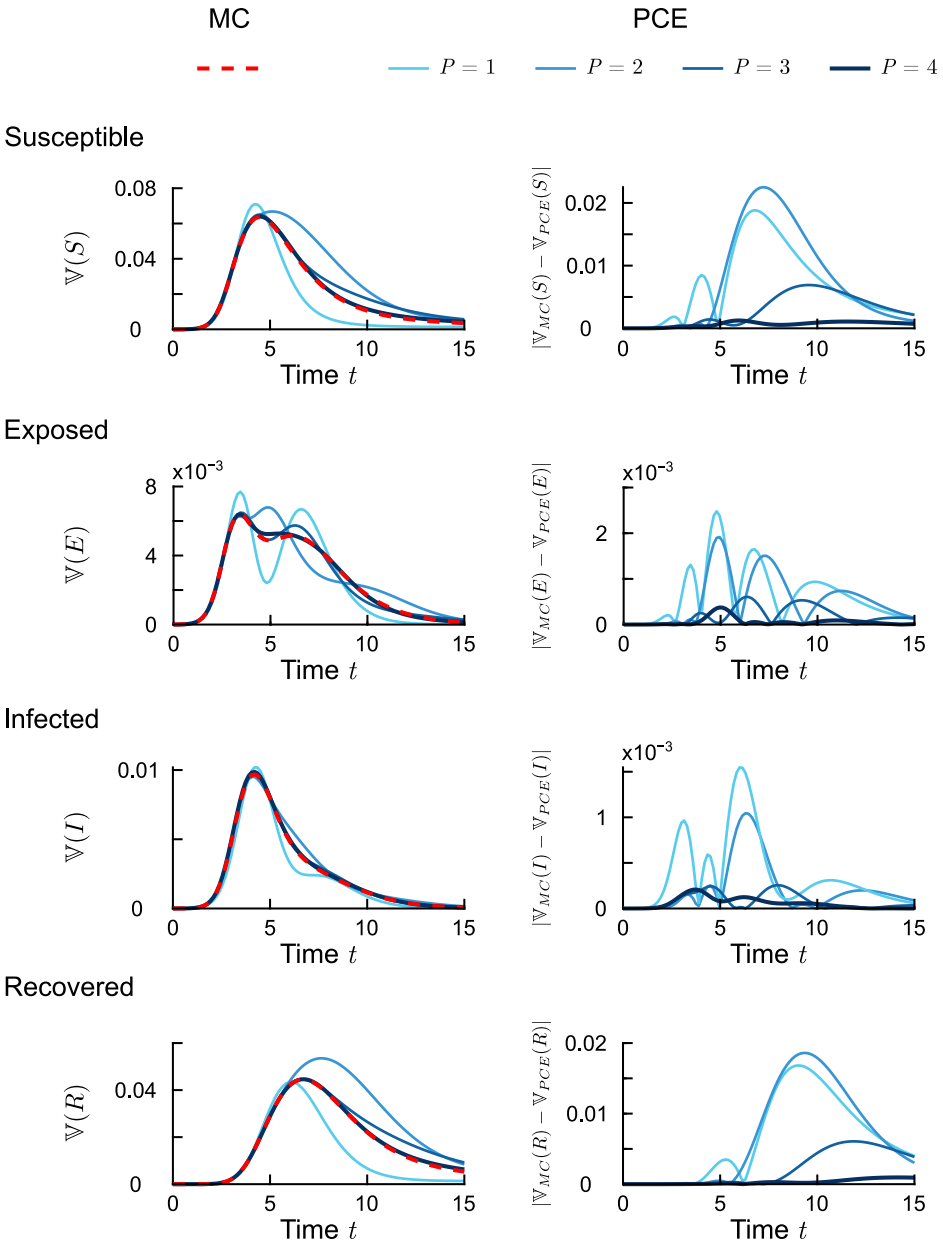


图6：易感组、暴露者、感染者和康复组的蒙特卡罗模拟（5000个样本）与多项式混沌展开的方差对比（左）。蒙特卡罗模拟（5000个样本）与多项式混沌展开之间的方差绝对误差（右）。

在疾病预测中，了解感染组何时达到峰值以及峰值大小十分重要。给定多项式混沌系数 I_{ijk} 和基函数 $\Phi(\zeta_1, \zeta_2, \zeta_3)$，我们可以通过对变量 ζ_1、ζ_2 和 ζ_3 进行抽样，轻松计算 $I(t)$的统计量。这些变量在 $[-1, 1]$上均匀分布。多项式混沌系数只需计算一次；通过对随机变量抽样，即可合成 $I(t)$随时间变化的不同轨迹。根据这些模拟结果，可以计算并绘制 $I(t)$的最大值和峰值时间，并以二维直方图呈现。图7 展示了感染组峰值时间和峰值大小的二维热力图。值得注意的是，多项式混沌展开计算此类热力图的速度要快得多，而无需进行.....

样本方法，例如经典的蒙特卡洛方法。在图7 中，峰值时间最可能出现在时间单位 5 左右，幅度约为 0。18，这与图5中期望值所反映的情况一致。

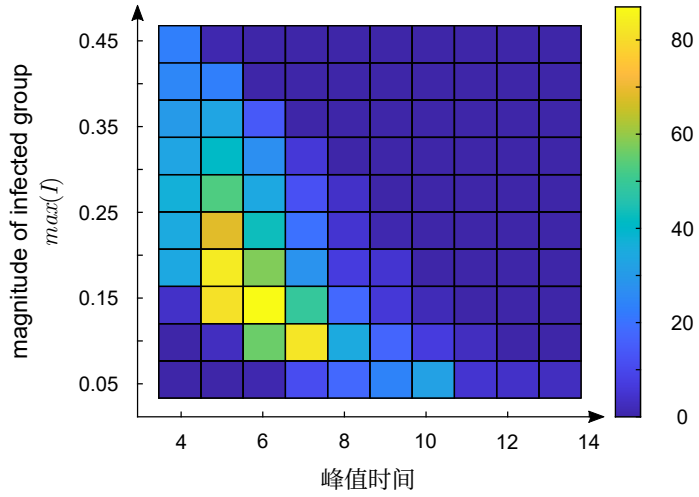


图7：感染组幅度和峰值时间的热图。

剩下的问题是：对于动力系统，Shapley效应是否像第3.1小节和第3.2小节所介绍的静态方程那样，与 Sobol 指数有所不同？表 5 表明，与计算总Sobol指数相比，Shapley 对高阶Sobol指数赋予的权重较低。图8展示了三个不确定变量 β、 σ 和 γ 对感染组随时间变化的重要性，其中实线表示 Shapley效应，虚线表示总Sobol指数。在子图 a) 中，变量 γ 起初对感染组具有很高的重要性。根据给定的初始条件并查看方程 (49a)-(49d)，可以看出，γ 起初的重要性很高是合理的，因为感染组会通过 γ 将其数值减少到康复组。在初始时刻，变量 β 和 σ 对感染组没有影响，因此其排名很低，这是因为暴露组的初始值为 0。随着时间推移， γ 的重要性下降，而 β 和σ 的重要性上升，同时 β 的取值高于 σ。这是因为 β 推动暴露组增长，随后暴露组转变为感染组。回看图5，期望值在时刻 6 达到峰值；由方程 (49c) 可明显看出，如果 I 变大，变量 γ 就会产生较大影响。图8 a) 中也体现了这一点。当感染组减少时， γ 的重要性降低，而 β和 σ 再次变得更为重要，因为它们推动了 I的增长。

图 8 展示了以下指标随时间的变化：a) 将总Sobol指数和Shapley效应绘制在同一尺度上。可以看到，Shapley效应在所有时刻的总和均为1，而总Sobol指数则并非如此。不过，图 8 a) 也展示了两种方法下所有不确定变量的相对重要性。要解释总Sobol指数与Shapley效应为何不同，我们需要查看图 8 b) 和 c) 所示的高阶Sobol指数。由于加权方式不同，Shapley效应与总Sobol指数出现了差异。可以看出，尤其是在这些高阶Sobol指数达到峰值时，总Sobol指数与Shapley效应并不一致（例如时间约为 6 时）；而当它们接近 0 时（时间约为 5），Shapley效应与总Sobol指数几乎相同。不过，要从图 8 a) 得出“排序”顺序相当困难。

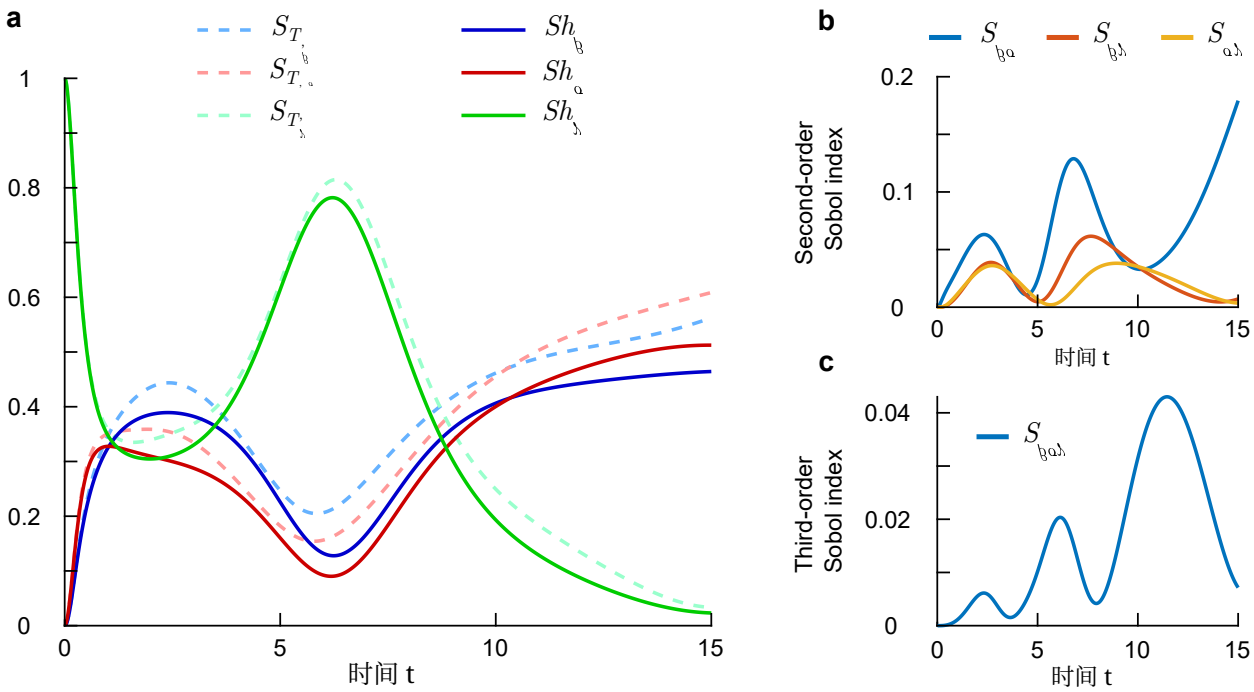


图8：通过四阶多项式混沌展开得到的不确定变量对感染组的重要性。a）总Sobol指数（虚线）与Shapley效应（实线）。b）二阶 Sobol 指数。c）三阶 Sobol 指数。

因此，我们引入图 9，其中蓝色、红色和绿色分别表示不确定变量β、 σ 和 γ。在初始时刻与最终时刻之间，按照直观的第 1、第 2 和第 3 名排序方案，给出了各不确定变量的排名。在时间 1 到 4 之间可以看到不确定变量排名的显著差异，而这段时间恰好早于感染组期望值达到峰值。这些信息可能对试图采取措施减缓大流行病中感染增长的机构非常有帮助。总Sobol指数将变量 σ 排在第 2 位，直至时间接近 3；而 Shapley效应则表明， σ 与 γ 之间的排名更早发生互换（时间约为 2。5）。感染组的期望值达到峰值后，总Sobol指数与Shapley效应相近。

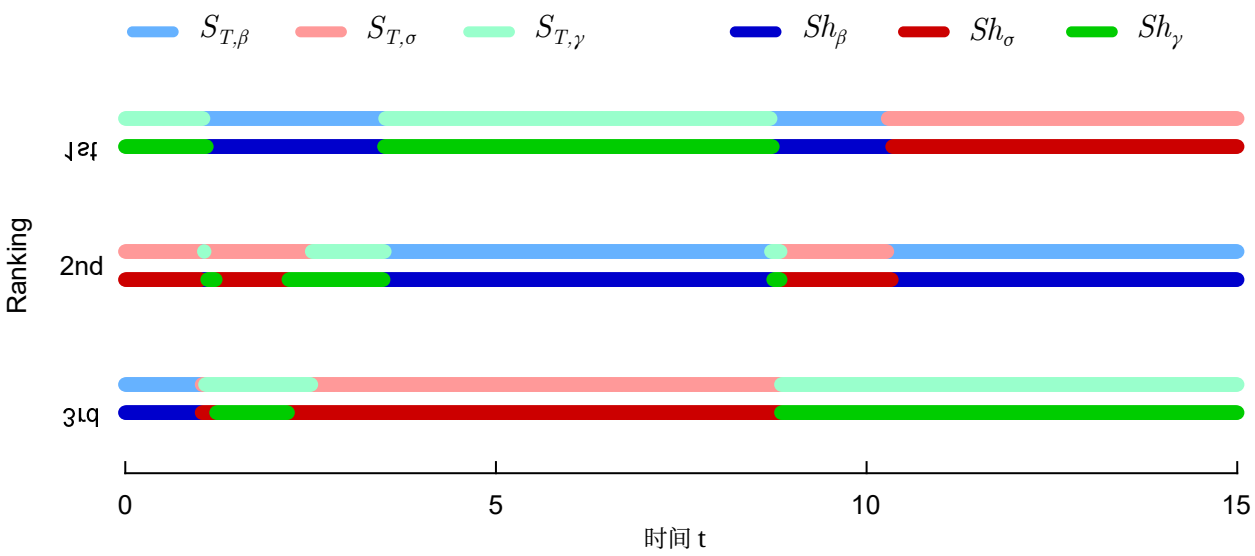


图9：根据总Sobol指数和Shapley效应对 SEIR模型中的不确定参数进行排序。

Owen [36] 指出，Shapley效应被“夹逼”在一阶Sobol指数和总阶Sobol指数之间。图10 展示了不确定变量 β、 σ 和 γ对应的情况，其中虚线、实线和点线分别表示总Sobol指数、Shapley效应和一阶Sobol指数。对于 β 和 σ ，在时间点 6 附近，夹逼幅度尤其大，此时高阶Sobol指数达到峰值，如图8 b) 和 c) 所示。相比之下，在时间点 5 附近，夹逼宽度可以忽略不计，此时高阶Sobol指数几乎为 0。

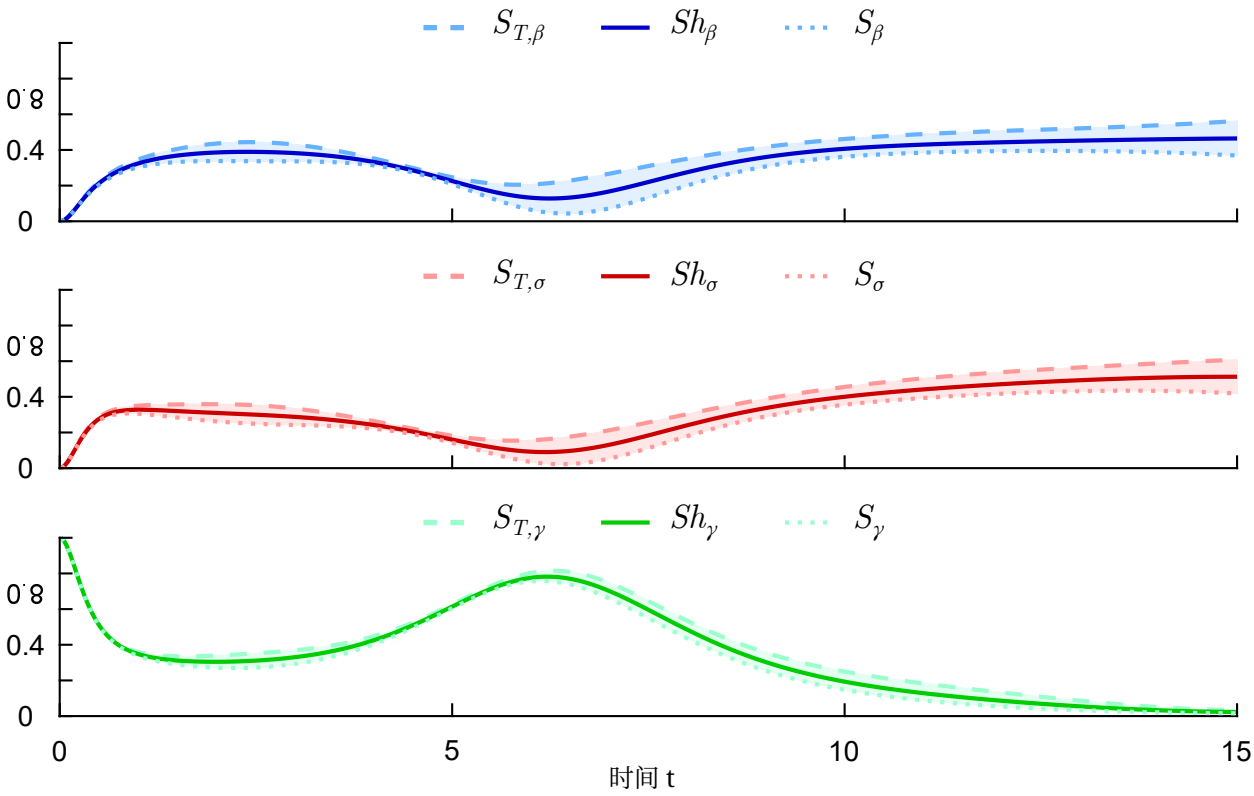


图10：沙　　夹逼：Shapley效应位于总Sobol指数（上界）和一阶Sobol指数（下界）之间界）），用于易　感者-暴露者-感染者-康复者疾病模型。

4.2. 1型糖尿病 Bergman模型

研究1型糖尿病（T1D）的一种常用模型是最简Bergman模型 [21]:

$$\dot{G}(t) = -\left(X(t) + p_1\right) G(t) + p_1 G_b + D(t) \tag{60a}$$

$$\dot{X}(t) = -p_2 X(t) + p_3 \left(I(t) - I_b\right) \tag{60b}$$

$$\dot{I}(t) = -p_4 \left(I(t) - I_b\right) + U(t) \tag{60c}$$

其中 G是　　血糖水平，X 是中间状态，I 是胰岛素浓度。
变量 $D(t)$ 是　　进餐扰动项，假设其符合 Fisher模型 [54]:

$$D(t) = \begin{cases} 0 & t < t_m \\ Be^{-d(t-t_m)} & t \geq t_m \end{cases} \tag{61}$$

其中 B 表示进餐量，假设为 $B = 28.98$，根据含 45 gm 碳水化合物（CHO）的一餐计算，进餐时间为 $t_m = 15$ 分钟 [21]。该模拟考虑在 $t = 0$进行胰岛素推注。模拟中，胰岛素推注由时变项 $U(t)$ 表示，见方程 (60c)。输入 $U(t)$ 定义为:

$$U(t) = \begin{cases} \frac{1000 \cdot CHO(\text{in g})}{CR \cdot V_i} & 0 < t < 1 \\ 0 & 1 < t \end{cases} \tag{62}$$

其中，CR 是胰岛素与碳水化合物比值，V_i 是胰岛素分布容积。对于一顿含 45 克碳水化合物的餐食，$CR = 18.477$ 和 $V_i = 12$，因此 $U(0 < t < 1) = 202.96$ mU/L。不确定性存在于

参数 p_1， p_2， p_3 和初始葡萄糖水平 G_0。我们假设它们均匀分布如下 [21]:

$$p_1 \sim U(0.0201, 0.0373) \tag{63a}$$
$$p_2 \sim U(0.0198, 0.0368) \tag{63b}$$
$$p_3 \sim U(3.525e^{-5}, 6.545e^{-5}) \tag{63c}$$
$$G_0 \sim U(83.43, 154.93). \tag{63d}$$

应用多项式混沌展开时，应注意初始血糖是通过伽辽金投影计算的。请注意，初始血糖水平需除以状态方程经伽辽金投影得到的系数。为比较蒙特卡洛与多项式混沌展开，选取了 10，000 个蒙特卡洛样本。系统还使用 $X_{0000} = 0$ 和 $I_{0000} = 15.3872$进行初始化，而 X 和 I 的其余展开状态则初始化为 0。图 11 展示了血糖水平的期望值和总方差。多项式混沌展开阶数越高，蒙特卡洛方法与多项式混沌展开方法之间的误差越小。图 11右侧显示，即使提高多项式混沌展开的阶数，仍存在较小误差。

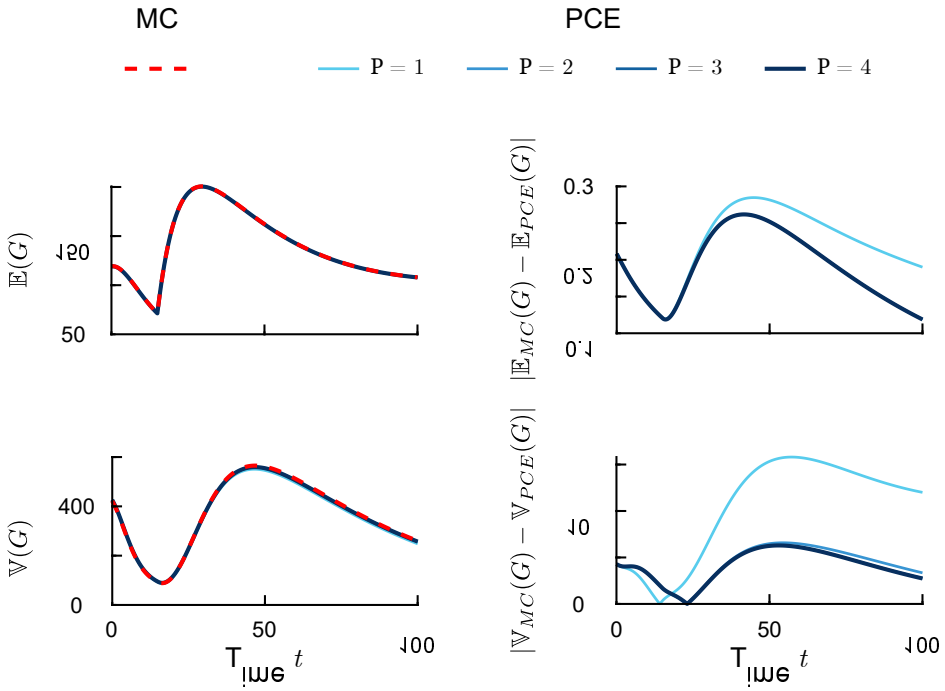


图 11：血糖水平的期望值和方差：左侧为蒙特卡罗模拟（10000 个样本）与多项式混沌展开的比较；右侧为蒙特卡罗模拟（10000 个样本）与多项式混沌展开所得期望值和方差的绝对误差。

图 12 a) 显示了相对于血糖水平， p_1、 p_2、 p_3和 G_0 随时间变化的总Sobol指数和Shapley效应。正如预期，初始血糖水平最为重要；与此同时，随着时间推移， p_3 的重要性在约25分钟时达到峰值，随后持续下降。这是由于方程（60c）中的胰岛素输入使胰岛素水平升高，因此方程（60b）中的不确定参数 p_3 变得更加重要。可以观察到，变量 p_2的重要性随时间增加，而变量 p_1 的重要性在整个时间段内名义上都是最低的。然而，图12 a) 还显示，最终Shapley效应与总Sobol指数出现分歧，这是由于图12 b)、c)和d)所示的高阶Sobol指数所致；这些指数除在仿真时间接近结束时外都相对较小。图12 a)中的插图展示了这种分歧；不过，这并不会改变系统的排序。

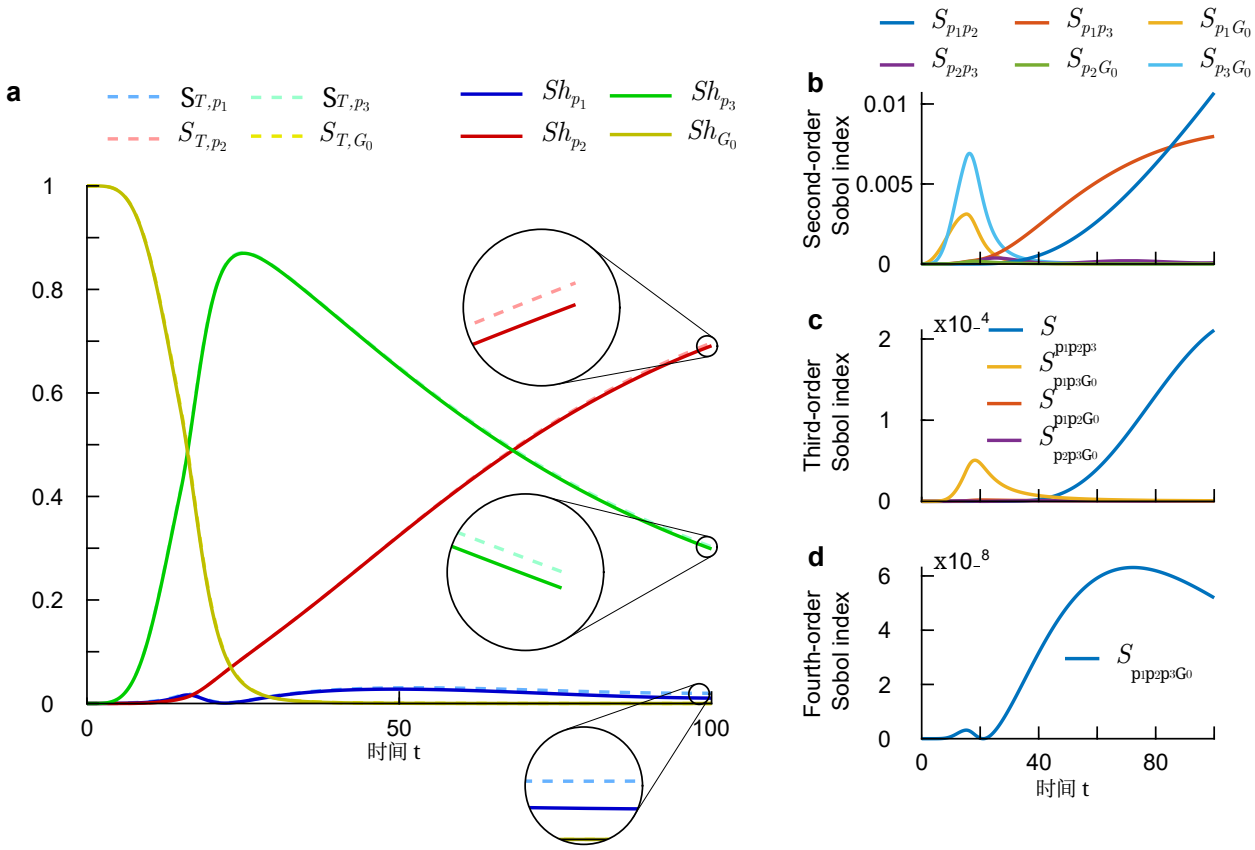


图12：基于 4 阶多项式混沌展开得出的不确定变量对血糖水平的重要性。a）总Sobol指数（虚线）与 Shapley效应（实线）。b）二阶 Sobol 指数。c）三阶 Sobol 指数。d）四阶 Sobol 指数。

图 13 展示了相对于血糖水平，不确定变量随时间变化的排序。Shapley效应给出的结果与总Sobol指数完全相同。这是由于高阶Sobol指数的数值较小，对总Sobol指数和Shapley效应计算的影响也较小。因此，两种方法给出的排序相近。如果将SEIR模型中的高阶Sobol指数（如图8 所示）与图12中的结果进行比较，可以看到，SEIR模型中的二阶 Sobol 指数比Bergman 1型糖尿病模型中的高一个数量级。三阶 Sobol 指数甚至高两个数量级。

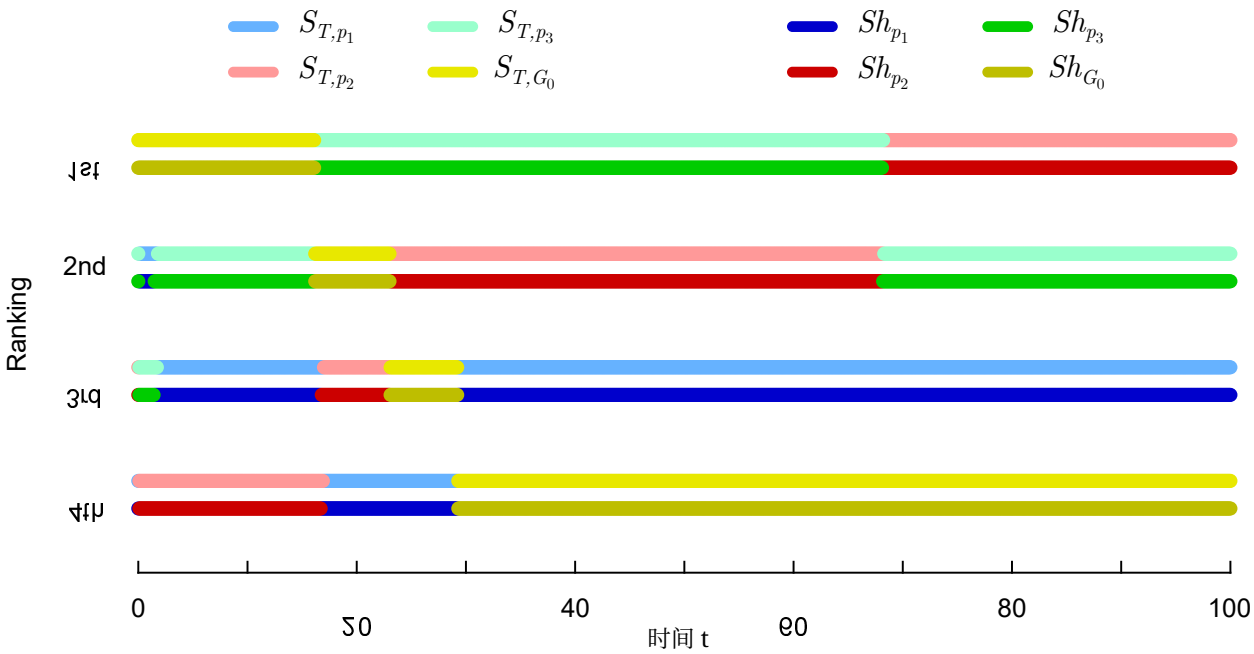


图13：Rankin　　根据总Sobol指数和 Shapley效应得出的 Bergman模型中不确定参数的 g　　.

图14 展示了第一阶和总Sobol指数对不确定变量的夹逼效应p_1， p_2， p_3 和 G_0。如图12所示，除仿真时间末期的 p_1 外，总Sobol指数与Shapley效应的差异并不大。因此，在仿真后期的大部分时间里， p_1都可以观察到较宽的夹逼效应。

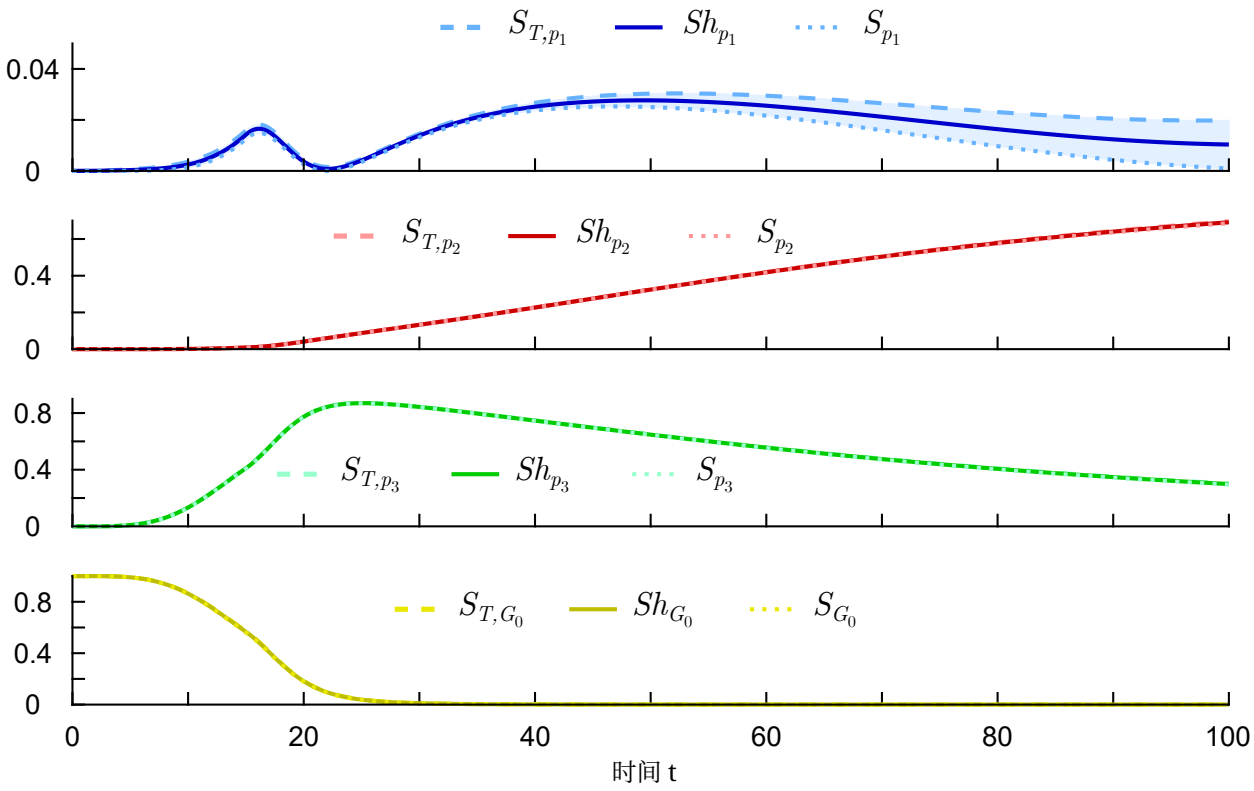


图14: Shapley效应在总Sobol指数（上界）与一阶Sobol指数（下界）之间的夹持模型。

5. 结论

本文提出将Shapley值概念用作基于方差的全局敏感性度量，并将其与成熟的基于方差的Sobol指数进行比较。除了多项式模型外，本文还针对Ishigami基准问题评估了Shapley效应。值得注意的是，对于Ishigami函数，Shapley效应给出的排序与总Sobol指数相冲突；这一事实已通过Borgonovo delta函数及其他统计距离加以说明。为解决计算Shapley效应时计算负担较重的问题，本文提出采用多项式混沌展开，高效评估计算Sobol指数和Shapley效应所需的总方差和条件方差。实际上，一旦确定了多项式混沌模型，就可以通过代数方法计算Sobol指数和Shapley效应。本文采用多项式混沌方法，说明了如何确定SEIR模型和Bergman 1型糖尿病模型中不确定变量的Shapley效应。

对总Sobol指数、Shapley效应和Borgonovo度量的比较，展示了使用Shapley效应对不确定变量进行排序的潜力。除糖尿病研究外，我们还计划研究控制器设计技术，在这些技术中，可利用Shapley值概念降低控制器对不确定性的敏感度。为此，建立从多项式混沌系数到Shapley效应的映射，将为鲁棒控制器设计提供一种计算成本低的方法。

致谢

作者感谢美国国家科学基金会通过CMMI奖项编号2021710对本项工作的支持。

参考文献

[1] S. Razavi, A. Jakeman, A. Saltelli, C. Prieur, B. Iooss, E. Borgonovo, E. Plischke, S. Lo Piano, T. Iwanaga, W. Becker, S. Tarantola, J. H. Guillaume, J. Jakeman, H. Gupta, N. Melillo, G. Rabitti, V. Chabridon, Q. Duan, X. Sun, S. Smith, R. Sheikholeslami, N. Hosseini, M. Asadzadeh, A. Puy, S. Kucherenko, H. R. Maier, 敏感性分析的未来：系统建模与政策支持不可或缺的学科，Environmental Modelling & Software 137 (2021) 104954。doi:10.1016/j.envsoft.2020.104954。[2] M. D. Morris，用于初步计算实验的析因抽样方案，Technometrics 33 (2) (1991) 161。doi:10.2307/1269043。[3] I. M. Sobol，非线性数学模型的敏感性分析，Mathematical Modelling and Computational Experiment 1 (1993) 407–414。URL https://cir.nii.ac.jp/crid/1570009751284338304

[4] I. M. Sobol'，A. Gresham，关于一种替代性全局敏感性估计量，载于：1995年SAMO会议论文集，1995年，第40–42页。[5] S. Kucherenko, M. Rodriguez-Fernandez, C. Pantelides, N. Shah，基于蒙特卡洛的基于导数的全局敏感性度量评估，Reliability Engineering & System Safety 94 (7) (2009) 1135–1148。doi:10.1016/j.ress.2008.05.006。[6] I. M. Sobol'，S. Kucherenko，基于导数的全局敏感性度量及其与全局敏感性指数的联系，Mathematics and Computers in Simulation 79 (10) (2009) 3009–3017。doi:10.1016/j.matcom.2009.01.023。[7] B. Iooss，P. Lemaˆıtre，全局敏感性分析方法综述（2014）。URL http://arxiv.org/pdf/1404.2405v1 [8] S. Nandi, T. Singh, P. Singla，使用共轭无迹变换进行基于导数的全局敏感性分析，载于：2019年美国控制会议（ACC），IEEE，2019年，第2458–2463页。doi:10.23919/ACC.2019.8815110。[9] T. A. Mara, W. E. Becker，利用多项式混沌展开对具有相关输入的模型输出进行敏感性分析，Reliability Engineering & System Safety 214 (6) (2021) 107795。doi:10.1016/j.ress.2021.107795。[10] K.-K. K. Kim, D. E. Shen, Z. K. Nagy, R. D. Braatz，Wiener多项式混沌在具有概率不确定性的非线性动力系统分析与控制中的应用 [历史视角], IEEE Control Systems 33 (5) (2013) 58–67。doi:10.1109/MCS.2013.2270410。[11] D. Xiu, G. E. Karniadakis，用于随机微分方程的Wiener–Askey多项式混沌，SIAM Journal on Scientific Computing 24 (2) (2002) 619–644。doi:10.1137/S1064827501387826。[12] B. Sudret，利用多项式混沌展开进行全局敏感性分析，Reliability Engineering & System Safety 93 (7) (2008) 964–979。doi:10.1016/j.ress.2007.04.002。[13] B. Sudret, C. V. Mai，利用多项式混沌展开计算基于导数的全局敏感性度量，Reliability Engineering & System Safety 134 (3) (2015) 241–250。doi:10.1016/j.ress.2014.07.009。[14] D. B. Harman, P. R. Johnston，将随机Galerkin方法应用于参数存在不确定性的流行病模型，Mathematical Biosciences 277 (2016) 25–37。doi:10.1016/j.mbs.2016.03.012。[15] B. C. Jensen, A. P. Engsig-Karup, K. Knudsen, E. Augeraud, M. Banerjee, J.-S. Dhersin, A. d'Onofrio, T. Lipniacki, S. Petrovskii, C. Tran, A. Veber-Delattre, E. Vergu, V. Volpert，利用多项式混沌对流行病建模进行高效不确定性量化和基于方差的敏感性分析，Mathematical Modelling of Natural Phenomena 17 (2022) 8。doi:10.1051/mmnp/2022014。[16] G. Blatman, B. Sudret，利用稀疏多项式混沌展开高效计算全局敏感性指数，Reliability Engineering & System Safety 95 (11) (2010) 1216–1229。doi:10.1016/j.ress.2010.06.015。[17] N. Lin, X. Xie, R. Schenkendorf, U. Krewer，锂离子电池三维多物理场模型的高效全局敏感性分析，Journal of The Electrochemical Society 165 (7) (2018) A1169–A1183。doi:10.1149/2.1301805jes。[18] T. Singh, P. Singla, U. Konda，基于多项式混沌的鲁棒输入整形器设计，Journal of Dynamic Systems, Measurement, and Control 132 (5) (2010) 76。doi:10.1115/1.4001793。[19] M. Braband, M. Scherer, H. Voos，纯电动汽车经济型模型预测纵向运动控制的全局敏感性分析，Electronics 11 (10) (2022) 1574。doi:10.3390/electronics11101574。[20] P. Ballester-Ripoll, M. Leonelli，概率图模型中的全局敏感性分析（2021）。URL http://arxiv.org/pdf/2110.03745v1 [21] S. Nandi, T. Singh，Bergman最小模型的全局敏感性分析，IFAC-PapersOnLine 53 (2) (2020) 16112–16118。doi:10.1016/j.ifacol.2020.12.431。[22] G. Fan, X. Li, R. Zhang，降阶电化学模型温度相关参数的全局敏感性分析及不同温度下的鲁棒荷电状态估计，Energy 223 (5) (2021) 120024。doi:10.1016/j.energy.2021.120024。[23] A. S. Yeardley, P. J. Bugryniec, R. A. Milton, S. F. Brown，锂离子电池热失控研究：基于高斯过程的全局敏感性分析，Journal of Power Sources 456 (2020) 228001。doi:10.1016/j.jpowsour.2020.228001。[24] X. Lai, Z. Meng, S. Wang, X. Han, L. Zhou, T. Sun, X. Li, X. Wang, Y. Ma, Y. Zheng，基于Sobol'方法的电动汽车锂离子电池等效电路模型全局参数敏感性分析，Journal of Cleaner Production 294 (3) (2021) 126246。doi:10.1016/j.jclepro.2021.126246。[25] H. Wu, W. Niu, S. Wang, S. Yan，使用Sobol'方法分析控制参数误差和电流参数对水下滑翔机运动精度的敏感性，Applied Ocean Research 110 (1) (2021) 102625。doi:10.1016/j.apor.2021.102625。[26] S. Yousefian, G. Bourque, R. F. D. Monaghan，预混燃烧器运行条件和化学动力学参数导致的氮氧化物排放不确定性量化，Journal of Engineering for Gas Turbines and Power 140 (12) (2018) 2721。doi:10.1115/1.4040897。

[27] E. Borgonovo，一种新的不确定性重要性度量，Reliability Engineering & System Safety 92 (6) (2007) 771–784。doi:10.1016/j.ress.2006.04.015。[28] S. Nandi、T. Singh，基于统计距离的全局敏感性分析度量，International Journal for Uncertainty Quantification 11 (6) (2021) 1–30。doi:10.1615/Int.J.UncertaintyQuantification.2021035424。[29] M.-H. Chun、S.-J. Han、N.-I. Tak，一种基于累积分布函数变化的距离度量的不确定性重要性度量，Reliability Engineering & System Safety 70 (3) (2000) 313–321。doi:10.1016/S0951-8320(00)00068-5。[30] L. S. Shapley，17. n人博弈中的一个值，载于：H. W. Kuhn、A. W. Tucker（编），《博弈论研究贡献》（AM-28），第II卷，普林斯顿大学出版社，1953年，第307–318页。doi:10.1515/9781400881970-018。[31] E. Algaba、V. Fragnelli、J. S´anchez-Soriano，《夏普利值手册》，第1版，Chapman & Hall/CRC运筹学丛书，Chapman & Hall/CRC，博卡拉顿，2019年。doi:10.1201/9781351241410。[32] E. Plischke、G. Rabitti、E. Borgonovo，计算用于敏感性分析的夏普利效应，《SIAM/ASA不确定性量化期刊》9 (4) (2021) 1411–1437。doi:10.1137/19M1304738。[33] M. I. Radaideh、S. Surani、D. O'Grady、T. Kozlowski，夏普利效应用于少群截面基于方差的敏感性分析，《核能年刊》129 (8) (2019) 264–279。doi:10.1016/j.anucene.2019.02.002。[34] E. Song、B. L. Nelson、J. Staum，用于全局敏感性分析的夏普利效应：理论与计算，《SIAM/ASA不确定性量化期刊》4 (1) (2016) 1060–1083。doi:10.1137/15M1048070。[35] B. Iooss、C. Prieur，具有相关输入的敏感性分析中的夏普利效应：与Sobol'指数的比较、数值估计及应用（2019）。URL http://arxiv.org/pdf/1707.01334v7 [36] A. B. Owen，Sobol'指数与夏普利值，《SIAM/ASA不确定性量化期刊》2 (1) (2014) 245–251。doi:10.1137/130936233。[37] B. Rozemberczki、L. Watson、P. Bayer、H.-T. Yang、O. Kiss、S. Nilsson、R. Sarkar，夏普利值在机器学习中的应用（2022）。URL http://arxiv.org/pdf/2202.05594v2 [38] S. C. Littlechild、G. F. Thompson，飞机着陆费：一种博弈论方法，《贝尔经济学期刊》8 (1) (1977) 186。doi:10.2307/3003493。[39] L. Pheulpin、N. Bertrand、V. Bacchi，相关输入参数下的不确定性量化与全局敏感性分析：在基础二维水力模型中的应用，LHB 108 (1) (2022) 221。doi:10.1080/27678490.2021.2015265。[40] W. Xie、B. Wang、C. Li、D. Xie、J. Auclair，面向质量源于设计和稳定性控制的可解释生物制造过程风险与敏感性分析，《海军研究后勤学》（NRL）69 (3) (2022) 461–483。doi:10.1002/nav.22019。[41] H. Wang、Z. Yan、X. Xu、K. He，评估可变可再生能源发电对孤岛微电网潮流的影响，《IEEE Access》6 (2018) 71339–71349。doi:10.1109/ACCESS.2018.2881189。[42] Y. Chen、L. Juan、X. Lv、L. Shi，药物敏感性预测的生物信息学研究，《药理学前沿》12 (2021) 799712。doi:10.3389/fphar.2021.799712。[43] Y. Li、N. Friedman、P. Teatini、A. Benczur、S. Ye、L. Zhu、C. Zoccarato，过度抽取地下水导致地裂缝的影响因素敏感性分析，《随机环境研究与风险评估》36 (1) (2022) 201。doi:10.1007/s00477-022-02237-8。[44] L. Schr¨oder、N. K. Dimitrov、J. Aasted Sørensen，将人工神经网络用作风力涡轮机载荷代理模型时的不确定性传播与敏感性分析，《物理学杂志：会议系列》1618 (4) (2020) 042040。doi:10.1088/1742-6596/1618/4/042040。[45] J. A. Carta、S. D´ıaz、A. Casta˜neda，一种应用于风电场功率输出估算模型的全局敏感性分析方法，《应用能源》280 (2020) 115968。doi:10.1016/j.apenergy.2020.115968。[46] B. Broto、F. Bachoc、M. Depecker、J.-M. Martinez，针对相互独立的变量组的敏感性指数，Mathematics and Computers in Simulation 163 (6) (2019) 19–31。doi:10.1016/j.matcom.2019.03.008。[47] A. B. Owen、C. Prieur，用夏普利值衡量相关输入的重要性（2017）。URL http://arxiv.org/pdf/1610.02080v3 [48] M. Il Idrissi、V. Chabridon、B. Iooss，将夏普利效应应用于具有相关输入的可靠性导向敏感性分析的进展与应用，Environmental Modelling & Software 143 (2021) 105115。doi:10.1016/j.envsoft.2021.105115。[49] T. Goda，一种基于夏普利效应的全局敏感性分析简易算法，Reliability Engineering & System Safety 213 (2021) 107702。doi:10.1016/j.ress.2021.107702。[50] M. B. Heredia 、 C. Prieur 、 N. Eckert，基于聚合夏普利效应的全局敏感性分析：在雪崩灾害评估中的应用，Reliability Engineering & System Safety 222 (4) (2022) 108420。doi:10.1016/j.ress.2022.108420。[51] E. Gayrard、C. Chauvi`ere、H. Djellout、P. Bonnet、D.-P. Zappa，具有独立增量的随机过程全局敏感性分析，《概率工程力学》62 (1) (2020) 103098。doi:10.1016/j.probengmech.2020.103098。[52] N. Benoumechiara、K. Elie-Dit-Cosaque、B. Bouchard、J.-F. Chassagneux、F. Delarue、E. Gobet、J. Lelong，具有相关输入的敏感性分析中的夏普利效应：自助法与基于克里金法的算法，《ESAIM：论文集与综述》65 (2019) 266–293。doi:10.1051/proc/201965266。[53] S. S. Kucherenko等，全局敏感性指数在非线性数学模型中的应用：综述，Wilmott Mag 1 (2005) 56–61。[54] M. E. Fisher，一种用于控制糖尿病患者血糖水平的半闭环算法，《IEEE生物医学工程学报》38 (1) (1991) 57–61。doi:10.1109/10.68209。