Unknown Parameter Excitation and Estimation for Complex Systems With Dynamic Performances

Yi-Ping Chen
System Optimization Laboratory,
Department of Mechanical Engineering,
National Taiwan University,
Taipei 10617, Taiwan
e-mail: chenyp@solab.me.ntu.edu.tw

Kuei-Yuan Chan[1]
System Optimization Laboratory,
Department of Mechanical Engineering,
National Taiwan University,
Taipei 10617, Taiwan
e-mail: chanky@ntu.edu.tw

Simulation models play crucial roles in efficient product development cycles, therefore many studies aim to improve the confidence of a model during the validation stage. In this research, we proposed a dynamic model validation to provide accurate parameter settings for minimal output errors between simulation models and real model experiments. The optimal operations for setting parameters are developed to maximize the effects by specific model parameters while minimizing interactions. To manage the excessive costs associated with simulations of complex systems, we propose a procedure with three main features: the optimal excitation based on global sensitivity analysis (GSA) is done via metamodel techniques, for estimating parameters with the polynomial chaos-based Kalman filter, and validating the updated model based on hypothesis testing. An illustrative mathematical model was used to demonstrate the detail processes in our proposed method. We also apply our method on a vehicle dynamic case with a composite maneuver for exciting unknown model parameters such as inertial and coefficients of the tire model; the unknown model parameters were successfully estimated within a 95% credible interval. The contributions of this research are also underscored through multiple cases. [DOI: 10.1115/1.4050107]

Keywords: model validation, excitation, parameter estimation, maneuver design, global sensitivity analysis, Kriging, design and analysis of computer experiment (DACE), Kalman filter

1 Introduction

Technological advances have resulted in products with an increasing number of subsystems and components; this has, in turn, made the modeling of complex systems increasingly central to efficient design and product development. For example, in vehicle design, simulations can aid the design of various subsystems—such as the suspension and chassis, controllers, safety assistance systems, and various ergonomic elements [1]. The reliability and robustness of these simulations are of critical importance when applied to the motion planning and driving strategies of autonomous vehicles [2,3]. Computer models that yield high-confidence results support engineers who must make precise, high-quality decisions.

However, although the model validation process is already mature [4], developing an accurate computer model is still a difficult task. Allen et al. proposed that model validation should account for cases where inconsistencies can occur between computer models and real-world system responses [5]. These cases include the following:

- Model formulation
- Simulation programing
- Parameter estimation/identification
- Numerical accuracy and stability

Of these four cases, parameter estimation remains the toughest and yet most vital problem. Its two sources of difficulties are (1) the coupling/interaction of parameters and (2) measurement noise in real model experiments [6,7]. Because unobservable parameters can only be inferred from the system output, it is almost impossible to identify their real values, and multiple sets of possible solutions may be obtained instead. To address these problems, one must consider how to increase the identifiability for multiple parameters in the validation stage.

In vehicle engineering, although measurement tools, mathematical theories, and simulation environments are well established, few studies have rigorously connected validity analysis of computer models with experimental test data. Typical testing strategies for validating vehicle models follow the vehicle dynamic test maneuvers proposed by the International Organization for Standardization (ISO). For example, Setiawan et al. tried to validate a 14-degrees-of-freedom vehicle model [8] in the context of double lane change maneuvers [9], and Gawade et al. used the average radius of trajectories under steady-state circular driving as a model accuracy index [10]. These testing methods, however, are designed for assessing hardware performance in specific scenarios but not for validating simulation models. Methods for designing operations based on the model validation requirements, such as uncertainty quantification and decoupled parameter effects, are not yet established in the field of vehicle engineering.

Various popular methods for exciting the effects on the parameters have been applied in diverse engineering fields. For example, with robot excitation trajectories which are compositions of fractions of spline curves, by optimizing fitness functions with trajectory-forming matrix and coupling indices, the inertial terms and several parameters of friction models are estimated under the designed operation using the method of least squares [11,12]. Another example in parameter identification of aircraft's control surfaces, parameters are estimated and identified in a single designed and composite flight via Wavelet transform [13]. For autonomous vehicles, Tsai and Chan observed that the effects of sensing errors contribute different trajectory uncertainties; they also performed a procedure on the excitation and quantification of each source of sensing errors [14]. Other excitation techniques, such as natural excitation technique [15] and persistent excitation [16], are mature and widely used methods for analyzing structures

[1]Corresponding author.

Contributed by the Design Automation Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received August 31, 2020; final manuscript received January 27, 2021; published online March 19, 2021. Assoc. Editor: Eun Suk Suh.

and machine tools, respectively. With the reviews from practical applications, it is concluded that when designing excitation operations, operations should be parameterized and uncertainties should be quantified and sorted.

Inspired from the engineering applications and the lack of systematic frameworks on connecting excitation and estimation of model parameters, in this work we propose a general procedure for model parameter excitation and validation. Parameter screening and sensitivity analysis are first conducted to identify importance of parameters [17]. However, not all important parameters can be directly observed and quantified, especially for dynamic systems with varying parameter values. In addition, experiments might not be able to reflect the differences of the parameters of interest. Therefore, in this work, experiments are redesigned to provide valid parameter estimations. Our procedure therefore include a prior screening to identify important parameters, experiment design to maximize the ability to observe the changes in the parameter values, and the estimations of parameter values based on the designed excitation experiments. In what follows, details of the steps will be elaborated with examples.

This paper introduces the proposed method, along with its verification and implementation in an engineering case by assuming that the only difference between a computer model and a real system is the deviation of model parameters. Sections 2 and 3 detail the implementation. Applicability verification and accuracy assessments are expressed in terms of an illustrative mathematical model in Sec. 4. In Sec. 5, details of a three-wheeled x-by-wire (manipulated by control units) vehicle model are described; the proposed method is validated through this dynamic model. Section 6 concludes the research with discussions on the necessity of designing excitation operations.

2 Parameter Excitation Through Global Sensitivity Analysis

Our proposed method, based on our conceptual design, is illustrated in Fig. 1 with two stages: excitation and estimation.

First, in the excitation stage, individual and coupled effects are quantified by an adjusted metamodel-based global sensitivity analysis (GSA) to reduce the computational cost, and an objective function composed of sensitivity indices generates excitation operations through optimization algorithms. Second, in the subsequent parameter estimation stage, the polynomial chaos-based Kalman filter (PC-KF) proposed by Blanchard et al. [17] is implemented and the estimation results are validated per hypothesis tests. The implementation is detailed in the following subsections.

2.1 Definition of System and Operation Parameters. In general, dynamic systems can be generally described by Eq. (1) with generalized output $\mathbf{m}$

$$\mathbf{m} = f(\mathbf{x}, \boldsymbol{\theta}) \tag{1}$$

Each set of operation parameters, denoted as $\mathbf{x} = [x_1, x_2, \ldots, x_l]$, determine how the system can be manipulated and controlled by users. For example, in this research, geometric parameters of planned path, and required velocity, were defined as the operation parameters.

Model parameters, represented as a vector of the form $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_k]$, are the unknown parameters of the model. In this research, measurable/certain/time-invariant specification parameters in the model are considered to be constants, and the immeasurable but fixed parameters are the independent variables of Eq. (1). We assumed that all unknown parameters were uniformly distributed, within their possible ranges. If the system has k model parameters, a k-dimensional multivariate design space is formed as the sampling space for GSA.

The definition of operation parameters and model parameters constitutes a preliminary step in the whole process. With a clear definition on probability density and with a formalized multivariate design space of model parameters, a method to quantify effects by parameters' deviation is detailed in the later sections.

2.2 Global Sensitivity Analysis. GSA is a popular technique for quantifying the influence of model-parameter deviations on system output [18]. GSA (1) entails a comprehensive analysis of the system, (2) covers the whole design space of inputs, and (3) quantifies the interaction effects. Equation (2) shows the decomposition of function variance Var into the main term and a set of interaction using Sobol's GSA method [19]. The sensitivity index of each term can then be obtained by normalizing each term, as represented in Eq. (3).

$$\mathrm{Var} = \left(\sum_{i=1}^{k} V_i\right) + \left(\sum_{i_1=1}^{k}\sum_{i_2=i_1+1}^{k} V_{i_1,i_2}\right) + \left(\sum\sum\sum V_{i_1,i_2,i_3}\right)\cdots + \left(V_{1,\ldots,k}\right) \tag{2}$$

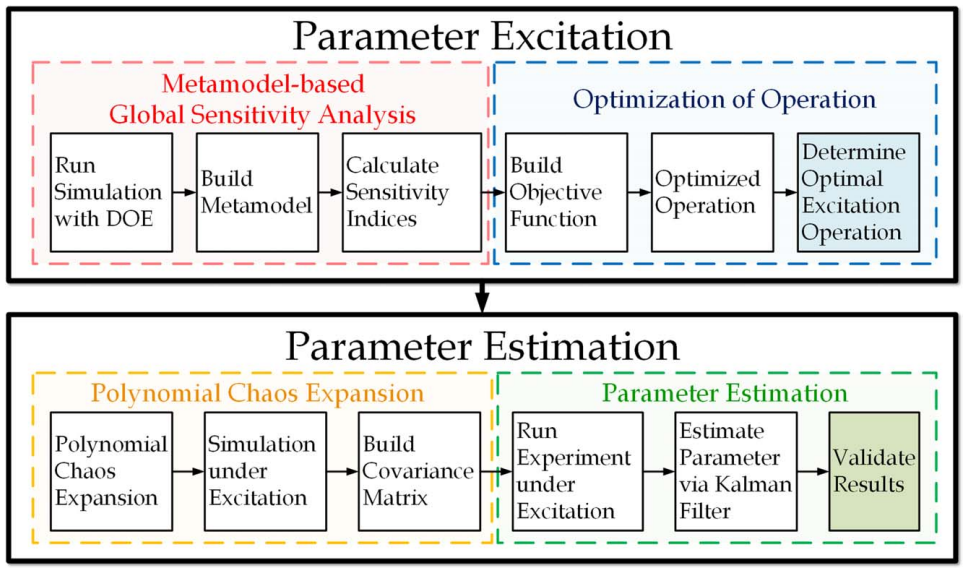

Fig. 1 Flowchart of the proposed method

$$S_{i_1\cdots i_s} = V_{i_1\cdots i_s}/\mathrm{Var} \tag{3}$$

Let us define the *main sensitivity index* (MSI, denoted as S_i) as the partial variance with respect to a single parameter θ_i, and the *total sensitivity index* (TSI, denoted as S_i^t) as the total influence of θ_i considering the main effect and all the interaction effects. Both indices are normalized to 0–1 and expressed as

$$\mathrm{MSI} = S_i = V_i/\mathrm{Var} \tag{4}$$

$$\mathrm{TSI} = S_i^t = S_i + \widehat{S}_{i,\sim i} = 1 - \widehat{S}_{\sim i} \tag{5}$$

where $\widehat{S}_{i,\sim i}$ is the sum of all the $S_{i_1\cdots i_s}$ terms that are involved with the index i and at least one index from $(1, \ldots, i-1, i+1, \ldots, k)$; $\widehat{S}_{\sim i}$ is the sum of all the $S_{i_1\cdots i_s}$ terms that are not involved with the index i.

GSA can be approached using several methods. The following passages introduce procedures for GSA indices.

Obtaining the GSA of the complex system requires some adjustments of the original Sobol method. The steps of the Sobol method, as documented in Ref. [20], include:

(1) *Create low-discrepancy sampling matrices*

This step creates N samples on the k-dimensional model parameter space using a Sobol sequence [21]. A $2k$-dimensional Sobol sequence is formed, and it is split into two independent sampling matrices $\mathbf{P}$ and $\mathbf{Q}$:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \ldots & \mathbf{P}_k \end{bmatrix} = \begin{bmatrix} \theta_{1,1} & \ldots & \theta_{1,k} \\ \theta_{2,1} & \ldots & \theta_{2,k} \\ \vdots & \ddots & \vdots \\ \theta_{N,1} & \ldots & \theta_{N,k} \end{bmatrix},$$
$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \ldots & \mathbf{Q}_k \end{bmatrix} = \begin{bmatrix} \theta_{1,k+1} & \ldots & \theta_{1,2k} \\ \theta_{2,k+1} & \ldots & \theta_{2,2k} \\ \vdots & \ddots & \vdots \\ \theta_{N,k+1} & \ldots & \theta_{N,2k} \end{bmatrix} \tag{6}$$

$\mathbf{P}$ and $\mathbf{Q}$ can generate an $\mathbf{R}^i$ matrix by replacing the ith column of $\mathbf{P}$ with the ith column of $\mathbf{Q}$, where $i = 1, \ldots, k$. Therefore, $\mathbf{R}^i$ can be denoted as

$$\mathbf{R}_i = \begin{bmatrix} \mathbf{P}_1, \ldots, \mathbf{Q}_i, \ldots, \mathbf{P}_k \end{bmatrix} = \begin{bmatrix} \theta_{1,1} & \ldots & \theta_{1,i-1} & \theta_{1,k+i} & \theta_{1,i+1} & \ldots & \theta_{1,k} \\ \theta_{2,1} & \ldots & \theta_{2,i-1} & \theta_{2,k+i} & \theta_{2,i+1} & \ldots & \theta_{2,k} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta_{N,1} & \ldots & \theta_{N,i-1} & \theta_{N,k+i} & \theta_{N,i+1} & \ldots & \theta_{N,k} \end{bmatrix} \tag{7}$$

Thus far, $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}^1, \ldots, \mathbf{R}^k$ are formalized, and each row of these matrices is a sampling point within an experimental design.

(2) *Simulate the model and calculate Sobol indices*

This step simulates the model with the N samples as $g(\mathbf{P})$, $g(\mathbf{Q})$, and $g(\mathbf{R}^i)$ which correspond to input sampling matrices. The GSA indices S_i and S_i^t can be calculated using Eqs. (8) and (9).

$$S_i = \frac{\mathrm{var}(E(g|p_i))}{\mathrm{var}(g)} = \frac{(1/N)\sum_{u=1}^{N} g(\mathbf{Q})_u (g(\mathbf{R}^i)_u - g(\mathbf{P})_u)}{(1/N)\sum_{u=1}^{N} (g(\mathbf{P})_u)^2 - ((1/N)\sum_{u=1}^{N} g(\mathbf{P})_u)^2} \tag{8}$$

$$S_i^t = \frac{\mathrm{var}(E(g|p_{\sim i}))}{\mathrm{var}(g)} = \frac{(1/2N)\sum_{u=1}^{N} (g(\mathbf{P})_u - g(\mathbf{R}^i)_u)^2}{(1/N)\sum_{u=1}^{N} (g(\mathbf{P})_u)^2 - ((1/N)\sum_{u=1}^{N} g(\mathbf{P})_u)^2} \tag{9}$$

By these two steps, MSI and TSI can be obtained. However, this procedure is only suitable for systems with scalar output. The Sobol method must be adjusted for applications to a dynamic output system.

For a complex model with time function outputs, an adjusted Sobol method—where metamodel techniques are added to principal component analysis (PCA)—is introduced in this section. Methods on calculating GSA indices through PCA are adjusted from Refs. [22–24]. To achieve precise calculation on sensitivity indices and also deal with the substantial computational cost of simulation, this research formalized metamodels to substitute for complex systems [19]. This substitution markedly lowered the computational cost with the steps of the operation are listed as follows:

(1) *Create initial sampling*

For a system with k dimensions and N training samples, initial sampling of model parameters is formalized with a Sobol sequence.

(2) *Simulate the model*

Simulate the model with the samples. Assuming that each output of the simulation can be formed by l data points, denoted as $\mathbf{y}_i = [y_i(1), y_i(2), \ldots, y_i(l)]$, the output of N training samples is as follows:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} y_1(1) & y_1(2) & \cdots & y_1(l) \\ y_2(1) & y_2(2) & \cdots & y_2(l) \\ \vdots & \vdots & \ddots & \vdots \\ y_N(1) & y_N(2) & \cdots & y_N(l) \end{bmatrix}_{N\times l} \tag{10}$$

(3) *Principal component analysis*

PCA was executed through the discrete Karlunen–Loéve transform [25], reforming $\mathbf{y}_i$ into a linear combination of principal components and their coefficients, named *PCA coefficients*, which represents $\mathbf{y}_i$ in finite scalar value so that the Sobol method can be applied. Simulation output data are reconstructed in PCA form in Eq. (11) with de, the truncation number to reduce dimensions. The value of de is determined such that the reserved portion of the original data is more than Θ (default as 99.9%). Equation (12) shows that de is the argument of the solution satisfying the desired Θ value. The symbol $\lfloor\rfloor$ indicates the floor of such arguments. $\bar{\mathbf{y}}$ is the average response of $\mathbf{y}$. Note that ϕ_j is the jth principal component and C_j is the corresponding PCA coefficient obtained from Eq. (13); I_j and $\mathbf{B}_j$ are the corresponding eigen values and eigen vectors of covariance matrix in PCA.

$$\mathbf{y}_i = \sum_{j=1}^{\mathrm{de}} \sqrt{I_j}\mathbf{B}_j C_j = \sum_{j=1}^{\mathrm{de}} C_j \phi_j + \bar{\mathbf{y}} \tag{11}$$

$$\mathrm{de} = \left\lfloor \arg\left\{ \frac{\sum_{j=1}^{\mathrm{de}} I_j}{\sum_{j=1}^{l} I_j} > \Theta = 99.9\% \right\} \right\rfloor \tag{12}$$

$$C_{i,k} = \frac{1}{\sqrt{I_i}} \sum_{i=k}^{l} Y_k \mathbf{B}_k \tag{13}$$

(4) *Build metamodels*

Kriging metamodels are created to connect the relationship between sampling points and the model output [26]. During model regression, model enhancement techniques—such as

the infill sampling criteria (ISC)—are used to confirm the goodness of fit. The details of Kriging models, ISC, and accuracy assessment procedures have been discussed in Refs. [26,27]. In this research, a single Kriging model was to predict a single PCA coefficient. Several Kriging models must be formalized, and they are denoted as $\hat{F}_{C_1}(\boldsymbol{\theta}), \ldots, \hat{F}_{C_{de}}(\boldsymbol{\theta})$. By using the initial sampling matrix as the input and $C_1, \ldots, C_{de}$ as the output, the Kriging model can easily predict PCA coefficients as $\hat{\mathbf{C}} = [\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_{de}]$ with respect to any experimental design of model parameters.

(5) *Predict PCA coefficient through metamodel*

After scalar representation of the dynamic output is constructed and the complex simulation model is replaced by several Kriging models, the Sobol method can be applied. In this step, GSA used the sampling matrices **P**, **Q**, and $\mathbf{R}^i$ with k-dimensions, and N_{GSA} samples are formalized by the method executed in step 1. With the formalized Kriging models, PCA coefficients of a complex system with **P**, **Q**, and $\mathbf{R}^i$ can be predicted, denoted as $\hat{g}_1(\mathbf{P}), \ldots, \hat{g}_{de}(\mathbf{P})$, $\hat{g}_1(\mathbf{Q}), \ldots, \hat{g}_{de}(\mathbf{Q})$, and $\hat{g}_1(\mathbf{R}^i), \ldots, \hat{g}_{de}(\mathbf{R}^i)$.

(6) *Global sensitivity analysis*

With total PCA coefficients, GSA of each PCA coefficient can be done by considering PCA coefficients are different and independent model output. By the Sobol method, one may calculate the sets of S_i and S_i^t of PCA coefficients. To fuse all GSA indices into one index, weighted average of each S_i and S_i^t is performed as Eq. (14), where I_j is the corresponding eigenvalue [23].

$$S_i = \frac{\sum_{j=1}^{de}(I_j S_{i,j})}{\sum_{j=1}^{de} I_j}; \quad S_i^t = \frac{\sum_{j=1}^{de}(I_j S_{i,j}^t)}{\sum_{j=1}^{de} I_j}, \quad i = 1, \ldots, k \tag{14}$$

By following these steps, the fused GSA indices of the dynamic system are obtained. Through the use of GSA, where the effects of each unknown parameter are quantified, the identifiability of specific parameters may be compared between different maneuvers.

2.3 Obtaining the Optimal Excitation Using Global Sensitivity Analysis. To effectively estimate the real value of a particular parameter θ_i, the suitable operation $\mathbf{x}^*$ ought to be able to yield the effect contributed from the deviation of θ_i; this effect can be conceptually described as Eq. (15), which shows that under the optimal operation $\mathbf{x}^*$, the system output difference is amplified even if the estimate target θ_i is fine-tuned.

$$\mathbf{x}_i^* = \arg\max \frac{\partial f}{\partial \theta_i} \tag{15}$$

The optimal operation for parameters considers the contribution (GSA indices) of unknown parameters changes under different manipulations. The operation parameters **x**, controlled by users, are the design variables in the optimization. The objective function of a particular model parameter θ_i excitation, following the concept of Eq. (15), can be listed as follows:

$$H_p(S_1 \cdots S_k, S_1^t \cdots S_k^t) = \frac{1}{\mathrm{Var}^{1/n}}\left(\frac{S_i^t - S_i}{S_i} + \frac{S_1 + \cdots + S_k - S_i}{S_i}\right) \tag{16}$$

Equation (16) minimized the interaction effects and main effects of non-observed terms shown as the numerator, while the main effect of θ_i and the total variance of system output under uncertain model parameters are maximized. Note that n in Eq. (16) is a scaling factor to ensure a more balanced optimum between the total variance and the GSA and will be assigned depending on the scale of variance in different cases.

As evident in Fig. 2, as long as the sampling matrix is fixed, the only variable in the GSA process is **x**. In other words, system variance and GSA indices are changed only by adjusting **x**. Therefore, to surrogate the complex process of GSA, the index generator, denoted as $\hat{F}(\mathbf{x})$ in Fig. 2, is formalized to create a continuous regression model. In this research, corresponding single-output index generators of the total variance and each GSA index are formalized by Kriging models, as illustrated in Fig. 3, and denoted as $\hat{F}_{\mathrm{var}}(\mathbf{x})$, $\hat{F}_{m_i}(\mathbf{x})$, $\hat{F}_{t_i}(\mathbf{x})$, and so on. Thus, by substituting the GSA indices into Eq. (16), the resulting optimization problem in

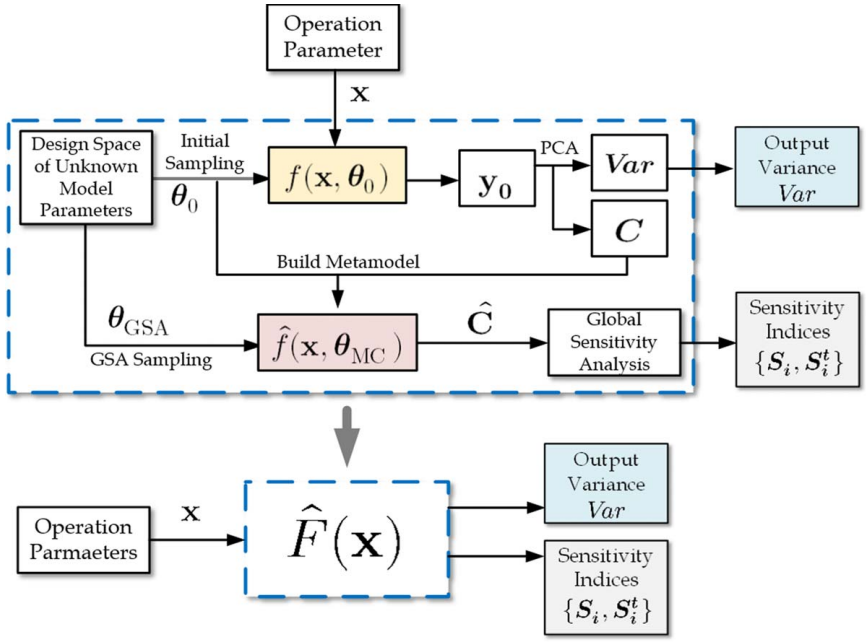

Fig. 2 Index generator

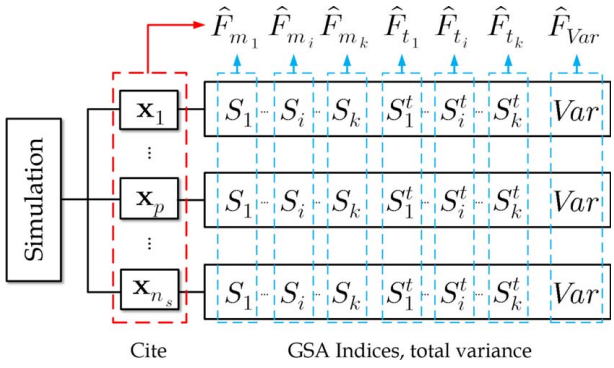

Fig. 3 Structure of index generator

general form is as follows:

$$\begin{aligned} \min \quad & H_p(\mathbf{S}, \mathbf{S^t}) = \frac{1}{\hat{\text{Var}}^{1/n}} \left(\frac{\hat{S}_i^t - \hat{S}_i}{\hat{S}_i} + \frac{\hat{S}_1 + \cdots + \hat{S}_k - \hat{S}_i}{\hat{S}_i} \right) \\ \text{w.r.t} \quad & \mathbf{x} \\ \text{s.t.} \quad & \forall \{\mathbf{x}, \boldsymbol{\theta}\} \in \mathcal{F} \\ \text{where} \quad & \hat{\text{Var}} = \hat{F}_{\text{var}}(\mathbf{x}) \\ & \hat{S}_i = \hat{F}_{m_i}(\mathbf{x}),\ i = 1, \ldots, k \\ & \hat{S}_i^t = \hat{F}_{t_i}(\mathbf{x}),\ i = 1, \ldots, k \end{aligned} \tag{17}$$

Equation (17) shows the intermediate variables on the GSA indices along with the constraints that keep operation in feasible domain. The optimal operation $\mathbf{x}^*$ can be solved through optimization techniques, in this research we used the DIRECT algorithm [28] for its robustness in obtaining results. These procedures constitute a systematic method for designing a dynamic test operation. The deviations of θ_i are well known to be more identifiable; such a property allows for better estimates of unknown model parameters in the validation stage.

3 Parameter Estimation Through Polynomial Chaos-Based Kalman Filter

In the computer model validation stage, techniques for estimating multiple unknown parameters simultaneously under measurement errors are needed. Most engineering applications use Baysian probability-based maximum likelihood estimation or the methods of least squares as in Refs. [29,30]. Blanchard et al. proposed a novel Kalman filter combined with polynomial chaos expansion (PCE) [17] that yielded high accuracy with less samples. This PC-KF is known to be highly applicable, and it was therefore selected as the estimation technique in this research.

3.1 Polynomial Chaos-Based Kalman Filter. This section introduces PC-KF as well as its constituent steps in practical applications. The PCE and PC-KF have been detailed in Refs. [17,31].

A polynomial chaos expanded function can be realized as a regression model, especially when used to describe a random process. A second-order random process $X(\vartheta)$, viewed as a function of random event ϑ, can be expanded as a linear combination of orthogonal polynomial chaos as follows:

$$X(\boldsymbol{\xi}(\vartheta)) = \sum_{j=1}^{\infty} c^j \psi^j(\boldsymbol{\xi}(\vartheta)) \tag{18}$$

where $\psi^j(\boldsymbol{\xi})$ are generalized Wiener–Askey polynomial chaos with respect to n_p-dimensional random variables $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{n_p}) \in \Omega \subseteq \mathcal{R}^{n_p}$. These orthogonal polynomials follow the Galerkin projection [32] as follow:

$$\langle \psi^i, \psi^j \rangle = \delta_{ij}, \quad i, j \in S, \quad \delta_{ij} = \begin{cases} 1 \text{ when } i = j \\ 0 \text{ when } i \neq j \end{cases} \tag{19}$$

The multidimensional basis functions $\psi^j(\boldsymbol{\xi})$ are tensor products of one-dimensional polynomial basis functions $P(\xi_k)$:

$$\begin{aligned} \psi^j(\xi_1, \ldots, \xi_n) &= \prod_{k=1}^{n} P_k^{l_k}(\xi_k), \\ j &= 1, 2, \ldots, S, \quad l_k = 1, 2, \ldots, p_b \end{aligned} \tag{20}$$

where p_b is the maximum order of $P(\xi_k)$, and $S = (n_p + p_b)!/n_p!p_b!$ is the total number of terms.

The update of a system is written as Eq. (21), where the superscripts f and a stand for the predicted (forecast) and estimated (assimilated) values, respectively. $\mathbf{M}$ is a one-step-ahead model that predicts the future state based on the existing information.

$$\begin{bmatrix} y_k^f \\ \theta_k^f \end{bmatrix} = \begin{bmatrix} \mathbf{M}(t_{k-1}, y_{k-1}^a, \theta_{k-1}^a) \\ \theta_{k-1}^a \end{bmatrix} \tag{21}$$

The updating mechanism of PC-KF follows Eq. (22) with considering the covariance of prediction uncertainty.

$$\begin{bmatrix} y_k^a \\ \theta_k^a \end{bmatrix} = \begin{bmatrix} y_k^f \\ \theta_k^f \end{bmatrix} + \mathbf{K}_k \left(z_k - \mathbf{H} \begin{bmatrix} y_k^f \\ \theta_k^f \end{bmatrix} \right) = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \begin{bmatrix} y_k^f \\ \theta_k^f \end{bmatrix} + \mathbf{K}_k z_k \tag{22}$$

$\mathbf{H}$ in Eq. (22) is the observer, $\mathbf{K}$ is the Kalman gain, and $\mathbf{R}_k$ is the noise covariance matrix. Each element in $\mathbf{R}_k$ is the variance of measurement noise corresponding to the state in $\mathbf{H}$:

$$\mathbf{H} = \begin{bmatrix} H_1 & 0 & \cdots & 0 \\ 0 & H_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_{n_s} \end{bmatrix}_{n_s \times n_s}, \quad \mathbf{R}_k = \begin{bmatrix} R_{1k} & 0 & \cdots & 0 \\ 0 & R_{2k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{n_s k} \end{bmatrix}_{n_s \times n_s} \tag{23}$$

$$\begin{aligned} \mathbf{K}_k &= P_k \begin{bmatrix} \mathbf{H}^T \\ 0 \end{bmatrix} (\mathbf{R}_k + \mathbf{H}(P_{yy})_k^f) \mathbf{H}^T)^{-1} \\ &= \begin{bmatrix} (P_{yy})_k^f \mathbf{H}^T \\ (P_{\theta y})_k^f \mathbf{H}^T \end{bmatrix} (\mathbf{R}_k + \mathbf{H}(P_{yy})_k^f \mathbf{H}^T)^{-1} \end{aligned} \tag{24}$$

where P_{yy} and $P_{\theta y}$ are the covariance matrix between state/state and model parameter/state, respectively.

However, for a complex system when linearity assumption is no longer valid, predicting the output states from t_{k-1} to t_k is quite challenging. Thus, in this research, we adopted the method in Ref. [33] to obtain y_k^f and P_k from y_{k-1}^a, θ_{k-1}^a, and P_{k-1}. The computer model only updates the estimated parameters θ_{k-1}^a and then runs simulations from $t = 0$ to $t = t_k$ once again to avoid predicting y_k^f and P_k

Table 1 Operation and unknown model parameters of the illustrative math model

Unknown model parameters	θ_1	θ_2	θ_3
Distribution	$U(0.95, 1.05)$	$U(0.95, 1.05)$	$U(0.95, 1.05)$
Operation parameters	x_1	x_2	x_3
operation space	[5, 10]	[0, 0.5]	[0, 10]

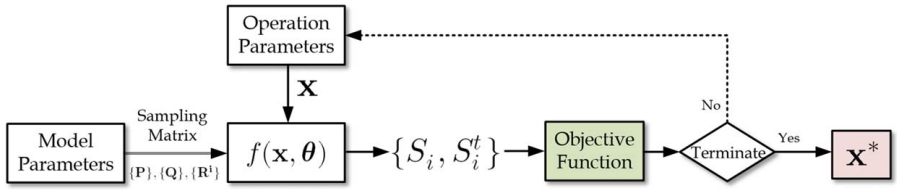

Fig. 4 Direct approach to optimizing exitation operation

directly. The calculation of P_k^f is as follows:

$$P_k^f = \mathrm{Cov}\left(\begin{bmatrix} y_k^f \\ \theta_k^f \end{bmatrix}\right) = \begin{bmatrix} \mathrm{Cov}(y_k^f) & \mathrm{Cov}([y_k^f, \theta_k^f]) \\ \mathrm{Cov}([\theta_k^f, y_k^f]) & \mathrm{Cov}(\theta_k^f) \end{bmatrix} = \begin{bmatrix} (P_{yy})_k^f & (P_{y\theta})_k^f \\ (P_{\theta y})_k^f & (P_{\theta\theta})_k^f \end{bmatrix} \tag{25}$$

To substitute y_k^f and θ_k^f, we assume that N samples of n_p-dimensional $\boldsymbol{\theta}$ are taken and simulated with the n_s output state y at $t = t_k$. Under the same PCE structure, they can be represented as follows:

$$y_{m,k}^f = \sum_{j=1}^{S} (y_{m,k}^f)^j \psi^j(\boldsymbol{\xi}), \quad 1 \le m \le n_s \tag{26}$$

$$\theta_{i,k}^f = \sum_{j=1}^{S} (\theta_{i,k}^f)^j \psi^j(\boldsymbol{\xi}), \quad 1 \le i \le n_p \tag{27}$$

PC-KF can be represented as follows:

$$\begin{bmatrix} \sum_{j=1}^{S} (y_k^a)^j \psi^j(\boldsymbol{\xi}) \\ \sum_{j=1}^{S} (\theta_k^a)^j \psi^j(\boldsymbol{\xi}) \end{bmatrix} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \begin{bmatrix} \sum_{j=1}^{S} (y_k^f)^j \psi^j(\boldsymbol{\xi}) \\ \sum_{j=1}^{S} (\theta_k^f)^j \psi^j(\boldsymbol{\xi}) \end{bmatrix} + \mathbf{K}_k z_k \psi^1(\boldsymbol{\xi}), \quad j = 1, \ldots, S \tag{28}$$

where $\psi^1(\boldsymbol{\xi}) = 1$. By Galerkin projection, each orthogonal term is isolated and left with only the coefficients. Through these processes, the estimated polynomial chaos coefficient of the model parameters can be obtained from the following equation:

$$(\theta_k^a)^j = (\theta_k^f)^j + (P_{\theta y})_k^f \mathbf{H}^T (\mathbf{R}_k + \mathbf{H}(P_{yy})_k^f \mathbf{H}^T)^{-1} (z_k \delta_{1j} - \mathbf{H}(y_k^f)^j), \quad j = 1, \ldots, S \tag{29}$$

By using PC-KF, multiple parameters can be estimated simultaneously by considering measurement errors.

3.2 Results Validations. Given that the only difference between a computer model and a real system is the deviation of model parameters under the designed operation, model parameters should be the argument in Eq. (30).

$$\hat{\boldsymbol{\theta}} = \arg\min \left\| f(\mathbf{x}^*, \boldsymbol{\theta}_{\mathrm{real}}) - f(\mathbf{x}^*, \hat{\boldsymbol{\theta}}) \right\| \tag{30}$$

The true values of many real-world models, however, are unknown and therefore the estimates can only be validated using the system output. The last step of our method is to validate the estimated model parameters through experimental data using hypothesis testing. To overcome the difficulties in dynamic function output, geometric mean of the test statistic of each data point is considered as the equivalent test statistic representing the whole system output. For example, to test the computer model with estimate parameters, first, simulate the model with measurement noise N times and

Table 2 Comparison between optimal excitation operation parameters from two methods

Excite target	$[x_1^*, x_2^*, x_3^*]$	Var
Direct approach		
θ_1	[9.7222, 0.4722, 0.5556]	11,977
θ_2	[9.7222, 0.0278, 9.4444]	9661
θ_3	[5.0001, 0.1140, 9.9997]	3454
Metamodel-based approach		
θ_1	[9.8333, 0.4883, 0.3333]	12,206
θ_2	[9.7222, 0.0278, 9.1667]	9409
θ_3	[5.0333, 0.0250, 9.8444]	3242

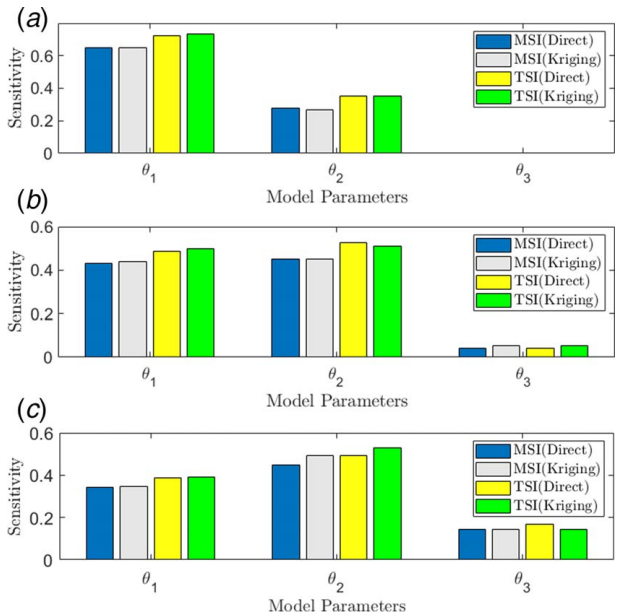

Fig. 5 GSA indices on exciting unknown model parameters: (*a*) GSA indices on exciting θ_1, (*b*) GSA indices on exciting θ_2, and (*c*) GSA indices on exciting θ_3

collect the output data with n equal steps:

$$\mathbf{y}^{\mathrm{sim}} = \begin{bmatrix} y_1^{\mathrm{sim}}(1) & y_1^{\mathrm{sim}}(2) & \cdots & y_1^{\mathrm{sim}}(n) \\ y_2^{\mathrm{sim}}(1) & y_2^{\mathrm{sim}}(2) & \cdots & y_2^{\mathrm{sim}}(n) \\ \vdots & \vdots & \ddots & \vdots \\ y_N^{\mathrm{sim}}(1) & y_N^{\mathrm{sim}}(2) & \cdots & y_N^{\mathrm{sim}}(n) \end{bmatrix} \tag{31}$$

Furthermore, run one experiment and obtain a vector output with the same length as that in Eq. (31):

$$\mathbf{y}^{\mathrm{exp}} = [y^{\mathrm{exp}}(1), y^{\mathrm{exp}}(2), \ldots, y^{\mathrm{exp}}(n)] \tag{32}$$

We then calculate the ith mean and standard deviation of the simulation output, $i = 1, \ldots, n$, and construct its confident interval. Then, calculate the test statistic of $y^{\mathrm{exp}}(i)$ through a comparison with $\mathbf{y}^{\mathrm{sim}}(i) = [y_1^{\mathrm{sim}}(i), y_2^{\mathrm{sim}}(i), \ldots, y_N^{\mathrm{sim}}(i)]^T$, $i = 1, \ldots, n$. The following

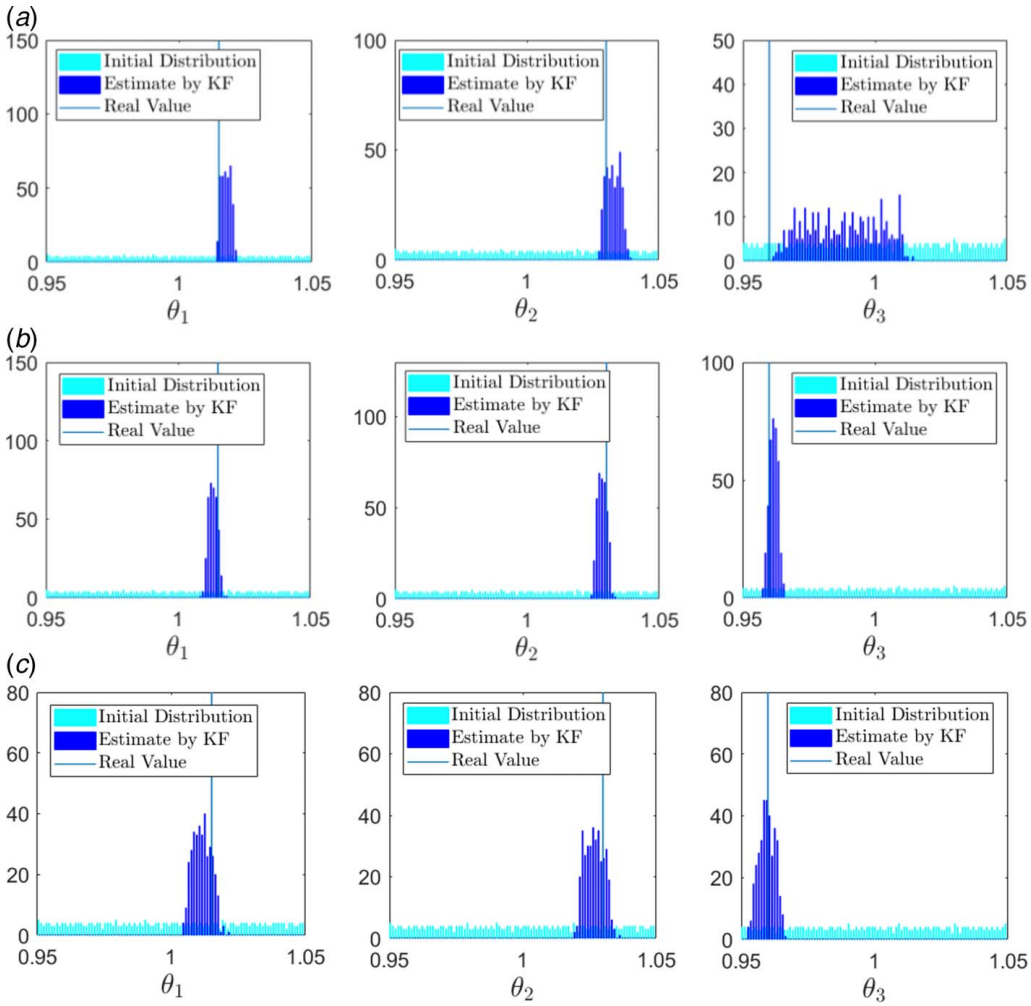

Fig. 6 Parameter estimation when exciting different model parameters: (*a*) parameter estimation with Kalman filter when exciting θ_1, (*b*) parameter estimation with Kalman filter when exciting θ_2, and (c) parameter estimation with Kalman filter when exciting θ_3

Table 3 Estimation result on model parameters; $[\theta_1^{\text{real}}, \theta_2^{\text{real}}, \theta_3^{\text{real}}] = [1.015, 1.03, 0.96]$

	Excite θ_1			Excite θ_2			Excite θ_3		
	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
Estimate	1.017	1.0326	0.9876	1.0125	1.0282	0.9616	0.0112	1.0267	0.9595
Var. of estimate	0.0018	0.0027	0.0137	0.0016	0.0017	0.0017	0.0033	0.0035	0.003
Error value	0.002	0.0026	0.0276	0.0025	0.0018	0.0016	0.0038	0.0033	0.0005
Error percentage	0.20%	0.25%	2.88%	0.25%	0.17%	0.17%	0.37%	0.32%	0.05%

result is obtained:

$$\boldsymbol{\alpha} = [\alpha(1), \alpha(2), \ldots, \alpha(n)] \tag{33}$$

with the equivalent test statistic α_{eq}:

$$\alpha_{\text{eq}} = \sqrt[n]{\alpha(1) \times \alpha(2) \times \cdots \times \alpha(n)} \tag{34}$$

By assuming that only the deviation of model parameters contribute to the output differences of the two models, the null hypothesis H_0 is posited; this hypothesis states that the output data are the same, which means that they are inferred to be from the same parameter sets. The alternative hypothesis H_1 declares the output difference and inaccuracy of estimate parameters. Thus, a test of α_{eq} with 95% confidence entails the following:

$$\begin{cases} \text{if } \alpha_{\text{eq}} \geq 0.975, H_0 \text{ is rejected, } H_1 \text{ stands} \\ \text{if } \alpha_{\text{eq}} < 0.975, H_1 \text{ is rejected, } H_0 \text{ stands} \end{cases} \tag{35}$$

In this proposed method, with excited operation, we expect the identifiability of model parameters to increase and necessarily lead to a more accurate estimation; we validate this expectation through hypothesis testing. In the later sections, the proposed method is demonstrated and verified.

4 An Illustrative Analytical Example

In this section, we apply the proposed method on a math problem to demonstrate the effectiveness and validity of the method.

Fig. 7 X-by-wire tricycle

Suppose a dynamic engineering system can be mathematically described as follows:

$$y(t)=\theta_1 x_1 \cos\left(\frac{\theta_2+x_2}{\theta_3}t\right)+\theta_3 \log(x_2)\left(\sin\left(\frac{1}{x_1}\right)+\sqrt{x_2 t}\theta_3 x_3\right)-3 \tag{36}$$

with operation parameters $\mathbf{x}$ and model parameters $\boldsymbol{\theta}$ that are listed in Table 1.

Set the simulation time $t=0$–15(s) with a 100 Hz sampling rate. $N=600$ low-discrepancy samples are taken within the three-dimensional design space of $\boldsymbol{\theta}$ by a Sobol sequence.

4.1 Metamodel-Based Approach to Global Sensitivity Analysis and Operation Excitation. A three-dimensional low-discrepancy sample set within the design space of operation parameters $\mathbf{x}=\{\mathbf{x}_1,\ldots,\mathbf{x}_p,\ldots,\mathbf{x}_{n_s}\}$ is generated. In this case, $n_s=125$, which is 40 times larger than the number of design variables.

By applying the GSA techniques for dynamic output as presented in Sec. 2.2, each simulation output $\mathbf{y}_i$, $i=1,\ldots,600$, can be transformed into a linear combination of five principal components with PCA coefficients $\mathbf{C}=\{C_1,\ldots,C_5\}$ if 99.9% of information can be reserved. Therefore, five Kriging models are formalized with $\boldsymbol{\theta}$ as the model input, where $\mathbf{C}_1,\ldots,\mathbf{C}_5$ are the respective model outputs. Each Kriging model is assessed by two indices R^2 and *relative average absolute error (RAAE)* [27], and $R^2>0.98$, $RAAE<0.06$ for each ensures their accuracy.

These high fidelity Kriging models replace the original model to obtain the GSA indices of each PCA coefficient as in Sec. 2.2. These GSA indices are then assigned as the model output; index generator, the continuous response surface between the indices and operation sites $\mathbf{x}$, is then generated. Thus, the optimization is formulated as follows:

$$\min H_p(\mathbf{x})=\frac{1}{\hat{\mathrm{Var}}}\left(\frac{\hat{F}_{t_i}(\mathbf{x})-\hat{F}_{m_i}(\mathbf{x})}{\hat{F}_{m_i}(\mathbf{x})}+\frac{\hat{F}_{m_1}(\mathbf{x})+\cdots+\hat{F}_{m_k}(\mathbf{x})-\hat{F}_{m_i}(\mathbf{x})}{\hat{F}_{m_i}(\mathbf{x})}\right)$$
$$\text{w.r.t}\quad \mathbf{x}$$
$$\forall \mathbf{x}\in\mathcal{F} \tag{37}$$

Where the terms with hat notation indicate the index generators and correspond to GSA indices. Operations that excite the unknown parameters θ_1 to θ_3 can be obtained. With metamodels, this method can also be implemented on a computationally expensive system.

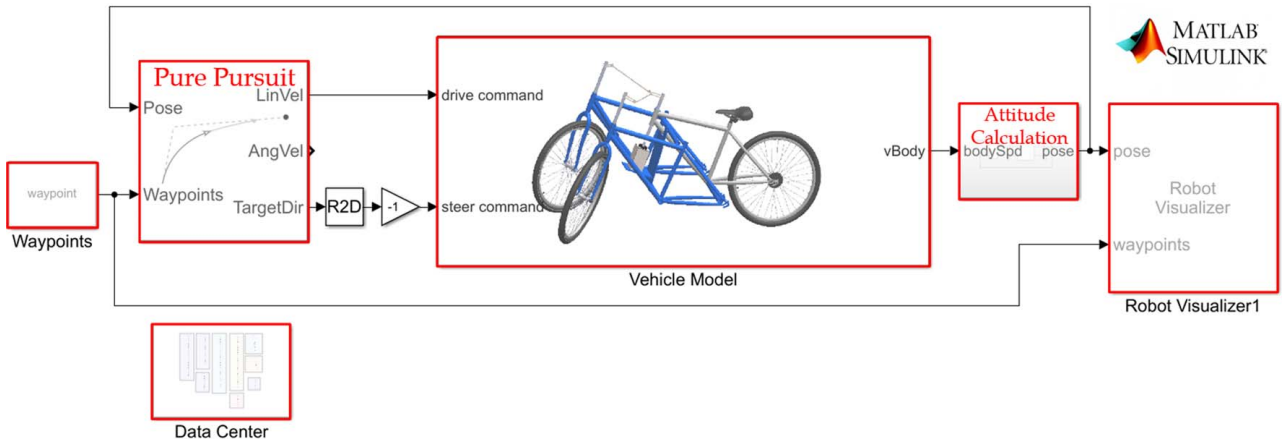

Fig. 8 Vehicle model in Simulink

Table 4 Specifications of tricycle: measurable/temporarily assumed model parameters that must be constant in the vehicle dynamic model

Item	Notation	Value	Unit	Item	Notation	Value	Unit
Front axle to COM	l_1	0.724	m	Gross mass	m	48.3	kg
Rear axle to COM	l_2	0.719	m	Inertial on x axis	I_x	12.1	kg m^2
Left axle to COM	w_1	0.3065	m	Inertial on y axis	I_y	36.7	kg m^2
Right axle to COM	w_2	0.3135	m	Inertial of right wheel	I_{fR}	0.3	kg m^2
Ground to COM	h	0.4896	m	Inertial of left wheel	I_{fL}	0.3	kg m^2
Wheel base	p	1.43	m	Inertial of rear wheel	I_r	0.5	kg m^2
Track	w	0.62	m	Radius of front wheel	r_f^0	0.3302	m
Forward offset	d	0.048	m	Radius of rear wheel	r_f^0	0.3302	m
Angular offset on left wheel	Δ_{fL}	−1	deg	Width of front tire	t_f	0.025	m
Angular offset on right wheel	Δ_{fR}	3.5	deg	Width of rear tire	t_r	0.025	m
Long. max. friction coeff.	μ_{dp}	0.8	–	Tire vertical stiffness	k_z	30,000	N/m
Lat. max. friction coeff.	μ_{dp}	0.8	–	Gravity	g	9.81	m/s^2

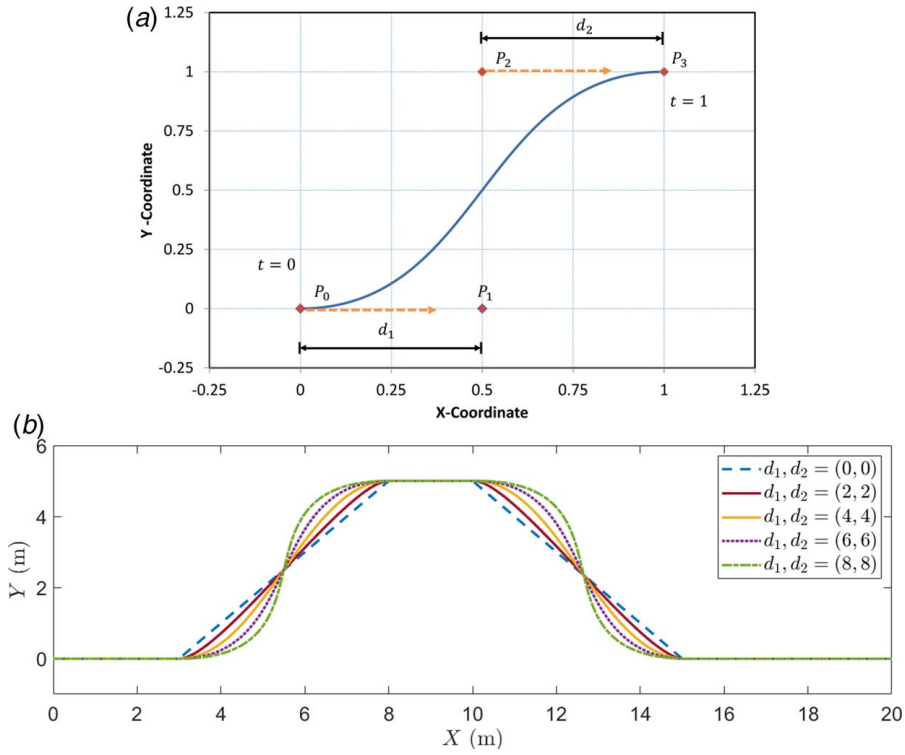

Fig. 9 Double lane change with different parameters: (*a*) Bézier curves and distance parameters and (*b*) Bézier curves with different parameters

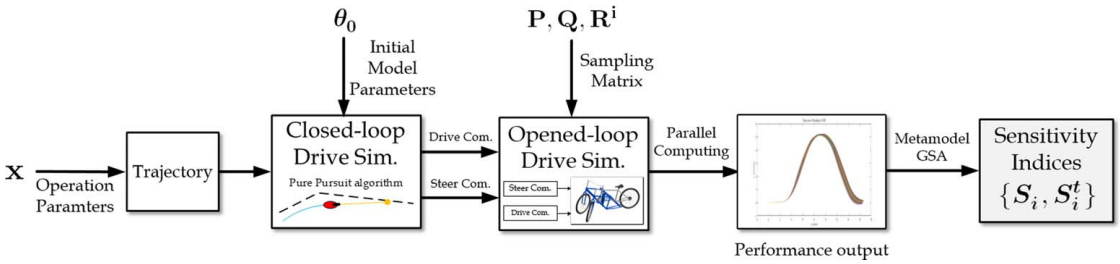

Fig. 10 System operation flowchart

4.2 Direct Approach on Global Sensitivity Analysis for Operation Excitation. The proposed metamodel-based method is verified by comparing with the results from the direct approach, as illustrated in Fig. 4. We also use `DIRECT` algorithm with corresponding GSA indices based on the method in Sec. 2.2.

4.3 Comparisons of Optimal Excitations. The optima of the two methods are compared in Table 2. The results indicate that both methods yield similar results. By substituting $\mathbf{x}^*$ from both methods back into the mathematical model and calculating the GSA indices, the comparison of GSA is illustrated in Fig. 5. If the difference between the two results is small, then the accuracy of the index generator is proven to be sufficiently high and the metamodel-based approach is therefore verified.

4.4 Unknown Parameter Estimation Through Polynomial Chaos-Based Kalman Filter. With verified optimal excitation operations, a PC-KF can be applied to estimate the unknown parameters. In the preliminary setting, n_p was set to 7, and the iteration of PC-KF was $k = 20$. Furthermore, white noise $N(0, 0.1)$ was added as measurement noise to create a simulated real system output $z(t_k)$. We assumed that $\boldsymbol{\theta}$ in the real model were $[\theta_1, \theta_2, \theta_3] = [1.015, 1.03, 0.96]$. The estimation results under three different excitation operations are presented in Fig. 6 and Table 3.

Figure 6 and Table 3 both show that the operation on exciting θ_2 yields the best estimate (with the least error value and variance of estimate) if the mean of each estimated distribution is considered as the estimated value $\hat{\theta}$. Although the total variance is the largest, the variance from θ_3 is small as in Fig. 6(*a*), which results in the wide confident interval. In Fig. 6(*c*), although the fraction from each factor is closer, the confident intervals are still larger than those in Fig. 6(*b*). One may infer from the results that if an unknown parameter can be estimated accurately with high confidence, the variance contributed from its main effect (approximately equal to the product of MSI and total variance) should be greatly increased. This fact proves the importance of designing an objective function to strike a balance between GSA indices and total variance. These tests confirm the conditions, under which an excitation operation is designed, that can improve the model parameters' identifiability values. In this case, the estimated model parameters from the operation that excites θ_2 are $[\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3] = [1.0125, 1.0282, 0.9616]$; these are considered as the final result and will be validated through an output comparison with the real model.

4.5 Results Validations. Validation was executed under an independent operation. In this case, $\mathbf{x}_{val} = [7.5, 0.25, 5]$ was assigned as a validation operation. A real model that carried θ_{real} was required for the conduct of one simulation with added measurement noise $N(0, 0.1)$ as experimental data. A model that carries $\hat{\boldsymbol{\theta}}$ (as a computer model) requires 500 replicates with measurement noise. From 500 outputs with 1500 data points for each, the 95% confident interval of 500 data points in the same period was formalized. Furthermore, 1500 data points in experimental data were tested with the formalized confident interval and 1500 test statistics along with the equivalent test statistic α_{eq} were calculated. In this case, $\alpha_{eq} = 0.7416 < 0.975$, thereby proving the validity of $\hat{\theta}$.

This mathematical model illustrates the implementation and verification of the proposed method. Given appropriate assumptions, this method can be applied to any system that can represent elements in the form of Eq. (1). In the subsequent section, we detail an engineering case.

5 Engineering Case Study: A Tricycle With Dynamic Performance

The proposed method is applied on an x-by-wire tadpole designed tricycle in this section, as shown in Fig. 7. This tricycle is used in our facility as a cheaper test-bed for autonomous vehicles due to its transparent structure and less components. Figure 8 shows the corresponding three-wheeled vehicle dynamic model in Simulink following the ISO-8855:2011 axis definition based on vehicle dynamic and motorcycle dynamic models [34–36]. Table 4 lists the specifications of the tricycle, including some measurable parameters. We can make provisional assumptions regarding immeasurable parameters, which are considered to have a negligible influence.

Six unknown model parameters are selected to be estimated as in Table 5, along with their possible ranges that contain true value θ_{real}. For operation parameters, four testing maneuvers elected as the basis scenarios where each of them are formed with two parameters:

(1) *Double lane change*: Bézier curves [37] in Eq. (38) are used to form a smooth curve as the trajectory, as illustrated in Fig. 9(*a*). Because P_0 and P_3 are fixed, $x_1 = d_1 \in [2, 8]$ and $x_2 = d_2 \in [2, 8]$ are set as the parameters that determine the curvature of the whole curve. Figure 9(*b*) shows the curves with different operation parameters.

$$B(t) = P_0(1-t)^3 + 3P_1(1-t)^2 + 3P_2t^2(1-t) + P_3t^3, \quad t \in [0, 1] \tag{38}$$

(2) *Steady-state cornering*: This maneuver is considered as the "static" test for several lateral variants and is defined by the two following operation parameters in Eq. (39): radius of trajectory $x_1 = r \in [2, 6]$ and required velocity $x_2 = v \in [0.8, 2]$.

$$\begin{cases} x = r\cos\left(2rt - \dfrac{\pi}{2}\right) \\ y = r\sin\left(2rt - \dfrac{\pi}{2}\right) \end{cases} \tag{39}$$

(3) *Chirp sine and inverse chirp sine*: Inspired by the commonly used method in the field of system identification, the chirp sine is the increasing values of its angular frequency with time, where the corresponding inverse chirp sine values are the decreasing values. With these values, the system undergoes a highly dynamic operation. The chirp sine values in

Table 5 Unknown model parameters and their possible range

θ	Item	Notation	l.b.	u.b.	Unit
θ_1	Inertial on z axis	I_z	30	90	kg m^2
θ_2	Cornering stiffness	C_α	500	1500	N/rad
θ_3	Camber stiffness	C_β	25	75	N/rad
θ_4	Self-aligned parameter (SAP)	SAP	0	0.03	m
θ_5	Rolling resistance 1	μ_{r_0}	0.0045	0.0135	–
θ_6	Rolling resistance 2	μ_{r_1}	0.0002	0.00075	v^{-2}

Table 6 Weights of GSA indices and variance fusion in different maneuvers

Observed sate	Double lane change	Steady-state cornering	Chirp sine	Inverse chirp sine
X	0.1	0.4	0.05	0.05
Y	0.7	0.4	0.85	0.85
$\dot{\psi}$	0.2	0.2	0.1	0.1

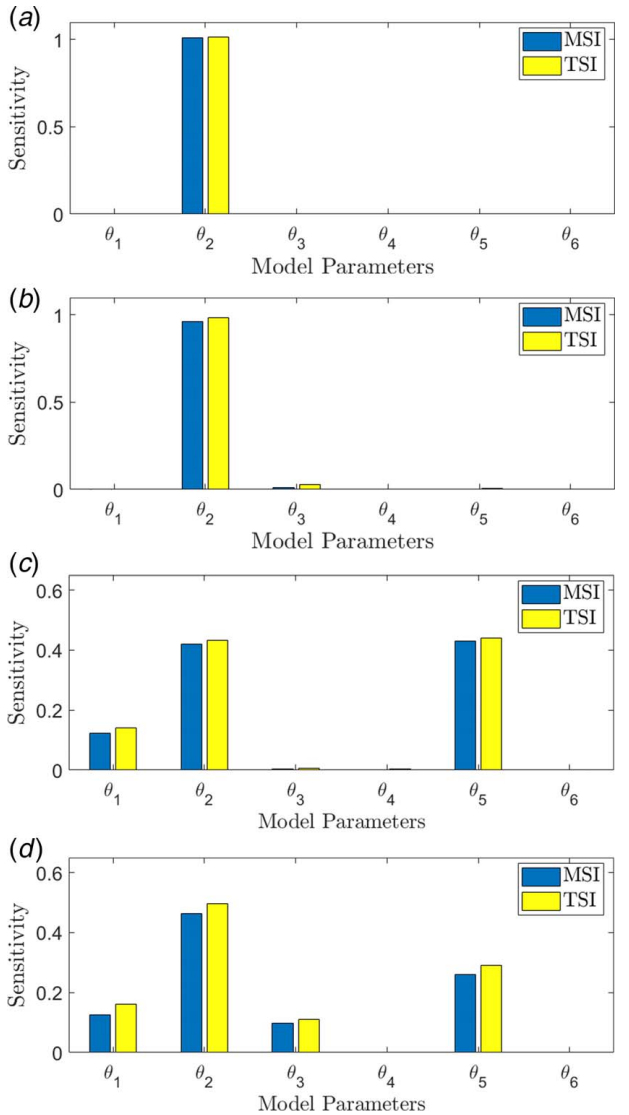

Fig. 11 GSA indices of selected maneuvers: (*a*) GSA indices of selected double lane change, $[x_1^*, x_2^*] = [5.8253, 1.1578]$, (*b*) GSA indices of selected steady-state cornering, $[x_1^*, x_2^*] = [2.0011, 0.90782]$, (c) GSA indices of selected chirp sine, $[x_1^*, x_2^*] = [0.24340.9906]$, and (*d*) GSA indices of selected inverse chirp sine, $[x_1^*, x_2^*] = [0.2, 0.73272]$

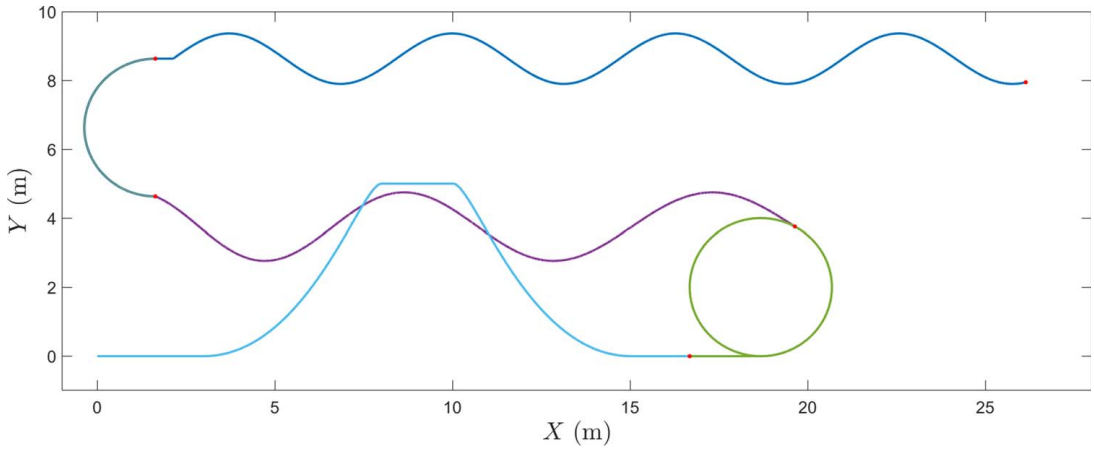

Fig. 12 Composite maneuver for parameter estimation

Table 7 Model parameters in the virtual experiment model

θ_1 I_z	θ_2 C_α	θ_3 C_β	θ_4 SAP	θ_5 μ_{r_0}	θ_6 μ_{r_1}
48.8834	1413.4876	56.618	0.0029	0.007	5.23×10^{-04}

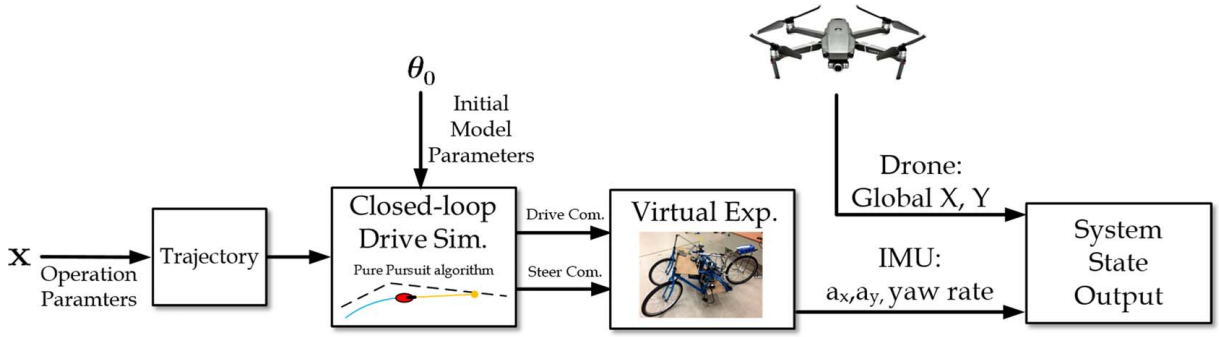

Fig. 13 Structure of the virtual experiment (Inertia Measurement Unit (IMU))

this case are as follows:

$$(x_{\text{chirp}}, y_{\text{chirp}}) = \begin{cases} x = 0.2t \\ y = A \sin((0.5 + 0.05e^{\lambda t})\omega t) \end{cases} \tag{40}$$

where the amplitude $x_1 = A \in [0.5, 1]$ and decay $x_2 = \lambda \in [0.05, 0.25]$ are the controlled parameters, and ω is a constant fixed as 1. The inverse chirp sine values are modeled as follows.

$$(x_{\text{inv.chirp}}, y_{\text{inv.chirp}}) = \begin{cases} x = 0.2t \\ y = A \sin((1 - 10^{-8}e^{\lambda t})\omega t) \end{cases} \tag{41}$$

where $x_1 = A \in [0.5, 1]$, $x_2 = \lambda \in [0.04, 0.2]$, and $\omega = 0.1$.

Figure 10 illustrates the entire process from operation parameters to form trajectories to the system models and then to the final performance output. In the closed-loop drive simulation, the pure pursuit algorithm[2] was used to follow the trajectory formalized by $\mathbf{x}$, and the model parameters in this stage $\boldsymbol{\theta}_0$ were the median of each θ_i called initial model parameters. The closed-loop simulation obtains the corresponding drive and steer commands that are the input to the opened-loop drive for Sobol GSA. In opened-loop drive, sampling matrices $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}^i$ are substituted and performance output can be calculated with parallel computation, and sensitivity indices can be obtained with the same GSA method. With $\mathbf{x}$ and $\boldsymbol{\theta}$ clearly defined and the model operation established, the complete procedure was executed based on Fig. 1.

Table 8 Sensing errors

Objective	Item	Error	Distribution
a_x	IMU acc. error	±3%	$N(0, (0.03a_x)^2)$
a_y	IMU acc. error	±3%	$N(0, (0.03a_y)^2)$
$\dot{\psi}$	IMU angular vel. error	±0.08726 rad/s	$N(0, 0.04363^2)$
X	Drone image error	±0.05 m	$N(0, 0.025^2)$
Y	Drone image error	±0.05 m	$N(0, 0.025^2)$

5.1 The Optimal Excitation Maneuver. The excitation operation is defined as a composite maneuver that can excite most of the unknown model parameters. This section demonstrates how this operation can be built with the four basis maneuvers. Three state outputs—X and Y locations on global coordinates and the yaw rate $\dot{\psi}$—were simultaneously observed and analyzed to perform a

[2]MATLAB *Pure Puresuit*.

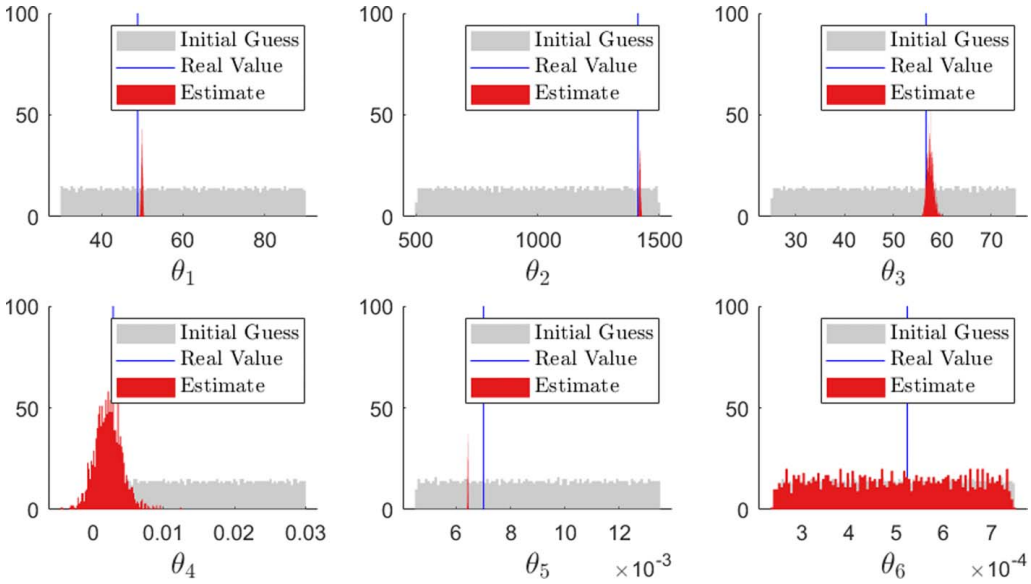

Fig. 14 Estimated results of model parameters through composite maneuver

more comprehensive GSA on the parameters' effects. Each state owns its individual S_i, S_i^t, and Var, with their weighted sum form the fusion of indices of the overall S_i, S_i^t, and Var. The weights in different cases, as listed in Table 6, were assigned to represent the importance of states. For example, in the maneuver of a double lane change, because Y is more crucial than X, Y has the larger weight. A similar situation exists in steady-state cornering; because the importance and performance of X and Y are the same, their assigned weights are identical. At present, there is no systematic method for weight assignment; it can only be done through observing the total variance of each state and through intuition.

The objective function for the optimal excitation operation is formed with the fused GSA indices and total variance as in Eq. (42). As for scale factor n, since most of our studies showed that the total variance are between 10 and 1000 under different operating conditions, therefore we use $n = 3$ to ensure a 0 ~1 scale. For different applications, the values of n should vary.

$$\begin{aligned} \min \quad & H_p(\mathbf{x}) = \frac{1}{\left(\hat{F}_{\text{var}}(\mathbf{x})\right)^{1/3}} \left(\frac{\hat{F}_{t_1}(\mathbf{x}) - \hat{F}_{m_1}(\mathbf{x})}{\hat{F}_{m_1}(\mathbf{x})} + \frac{\hat{F}_{m_1}(\mathbf{x}) + \cdots + \hat{F}_{m_6}(\mathbf{x}) - \hat{F}_{m_1}(\mathbf{x})}{\hat{F}_{m_1}(\mathbf{x})} \right) \\ \text{w.r.t} \quad & \mathbf{x} \\ \text{s.t.} \quad & \forall \{\mathbf{x}, \boldsymbol{\theta}\} \in \mathcal{F} \end{aligned} \tag{42}$$

During the creation of index generators, the system only has two operation parameters in each operation; thus, a five-level full-factorial experimental design for the two-dimensional space of $\mathbf{x}$ was constructed. In each designed experiment, the GSA indices of θ_i were calculated using the metamodel-based GSA method. Note that to attain sufficient accuracy, 1400 initial samples are taken along with ISC applied to confirm $R^2 > 0.99$ and $RAAE < 0.05$ for every Kriging model. Furthermore, the fused accuracy assessment indices were taken to be the weighted sum of the reserved PCA coefficients—an identical approach to that in the fusion method for sensitivity indices.

The composition of optimal operation $\mathbf{x}^*$ from each maneuver can be formed as the final testing maneuver. By re-calculating Eq. (42) for $i = 1, \ldots, 6$ with four maneuvers, respectively, a total of 24 optimal values of $\mathbf{x}^*$ were calculated. In this case, only one optimal set of $\mathbf{x}$ was picked from each maneuver to form the final testing maneuver. The selected set of four operations not only must have a relatively large Var but also the set must also (hopefully) have each θ_i be excited at least once (meaning that its MSI or TSI should be enlarged at least once). The GSA indices of the four selected $\mathbf{x}^*$ —$[x_1^*, x_2^*] = [5.8253, 1.1578]$ in double lane change, $[x_1^*, x_2^*] = [2.0011, 0.90782]$ in steady-state cornering, $[x_1^*, x_2^*] = [0.24340.9906]$ in chirp sine, and $[x_1^*, x_2^*] = [0.2, 0.73272]$ in inverse chirp sine — are illustrated in Fig. 11. With these optimized results, the single composite maneuver is illustrated in Fig. 12. The composite maneuver in this case is mainly used for demonstration and for verifying our methodology; it may differ if the included maneuvers are changed.

In Fig. 11, θ_2 and θ_5 were relatively easier to be excited, θ_1 and θ_3 could be excited in this specific case, and θ_4 and θ_6 were barely excited. These results also show the irrelevance of θ_i as a screening experiment, where the confidence of estimation can be roughly predicted based on the magnitude of indices. Thus, the selection of this set may guarantee acceptable estimations on θ_1, θ_2, θ_3, and θ_5, but this set may also result in coupled effects that may lead to multiple solutions. Thus, the composition of these maneuvers provides multiple test situations, which affords the greatest chance for multiple solutions to be eliminated. With the composite maneuver, the unknown model parameters are expected to be estimated within a single drive.

5.2 Parameter Estimation. In this section, we detail how, with the designed maneuver, PC-KF was applied to estimate $\boldsymbol{\theta}$ using experimental data. However, because we assumed that the computer model and real model differed in only one respect to focus on this issue, experimental data were simulated from a virtual experiment. Virtual experiments and computer simulation models differ in two

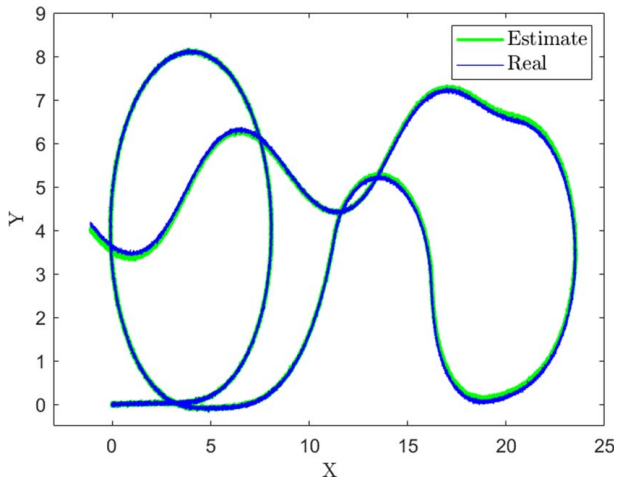

Fig. 15 Trajectory comparison between the updated model and the real model (unit: m)

Table 9 Hypothesis test of each output state and overall system

	X	Y	a_x	a_y	ψ	Overall
Test statistic	0.7659	0.8507	0.7718	0.7425	0.7369	0.7725
Result	Reject H_1	Reject H_1	Reject H_1	Reject H_1	Reject H_1	Reject H_1

respects. First, although the Simulink models are identical, in a virtual experiment, all the model parameters are set as known and fixed, as detailed in Table 7.

Second, simulated sensing errors are added to the observers as white noise elements. In the virtual experiment, five state outputs—which could be directly measured without integration, as noted in Fig. 13—were observed; their error distributions are listed in Table 8 for a_x, a_y, X, Y, and $\dot{\psi}$. With the known sensing errors, PC-KF were ready to be used.

The estimated results with $k = 20$ iterations and updates are illustrated in Fig. 14. As evident in this figure, θ_1, θ_2, θ_3 were successfully estimated; θ_5 was a closed guess with an offset; and θ_4 and θ_6 were, as expected, unable to be estimated with high confidence.

We use the estimated model parameters in the computer model to validate the result under an independent maneuver designed for this stage. Trajectory comparisons between the simulation results (with measurement uncertainty) of the posterior with the real system are shown in Fig. 15. Following the method in Sec. 4.5, with the confident interval contributed from 500 simulation data of the posterior distribution, the equivalent test statistics of the five states were obtained, and the geometric average of the five test statistics of states was taken to be the equivalent test statistic of the whole system output; these findings are listed in Table 9. Following the hypothesis testing in Eq. (34), it is insufficient to only reject H_0 within a 95% credible interval if we wish for the estimated results to be considered as the real system parameters.

One can compare this result with a PC-KF estimation for a randomly picked double lane change maneuver (Fig. 16); such a comparison is the most common method for validating vehicle models, and the two-estimate results are listed in Table 10. In this comparison, the estimated results from a composite maneuver exhibited not

Table 10 Comparison between composite maneuver and randomly picked double lane change

		θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
	Real	48.8834	1413.488	56.618	0.0029	0.007	0.000523
Composite maneuver	Estimate	49.9348	1422.5	57.6706	0.0023	0.0064	0.000495
	Error/span %	1.7523	0.9012	2.1052	2	0.6667	5.0909
	Standard dev.	0.202	2.7908	0.7202	0.0019	1.28×10^{-05}	1.44×10^{-04}
Double lane change	Estimate	50.0684	1516.1	55.2672	0.0209	0.0041	0.000505
	Error/span %	1.975	10.2612	2.7016	163.333	10.888	3.2727
	Standard dev.	0.5183	5.9331	2.0947	0.0077	1.82×10^{-04}	1.44×10^{-04}

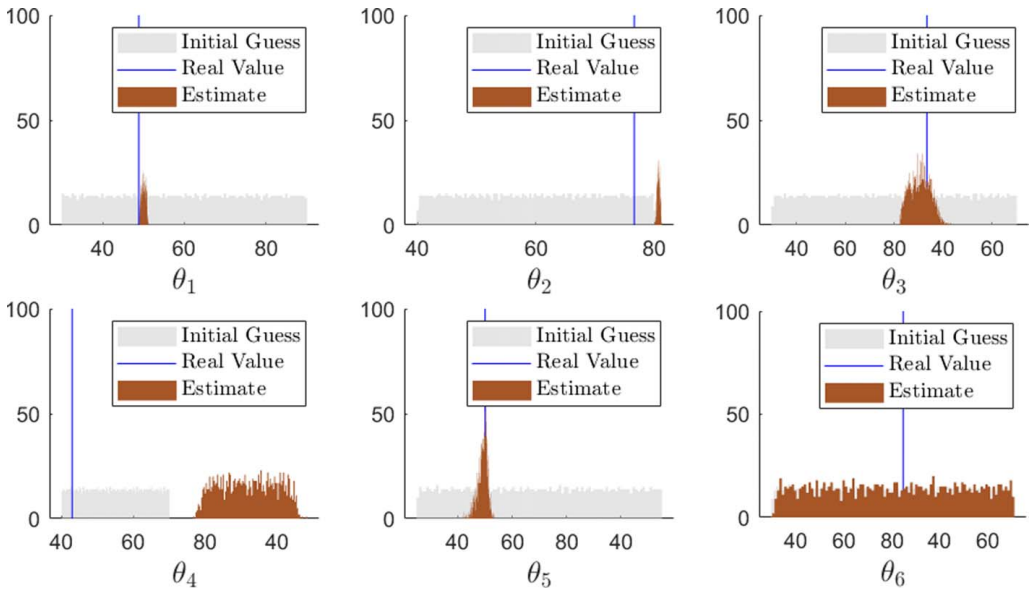

Fig. 16 Estimation result of model parameters through randomly picked double lane change

only more precise estimates with slighter offsets but also greater confidence with a smaller standard deviation. Through this simple comparison, we can see that the designed composite maneuver improved the estimation. Although the improvements were not highly significant, this method still affords us an idea of the validation process from operation design to parameter estimation and model validation.

6 Conclusions

We propose an approach for unknown parameter excitation and validation for dynamic systems. In this approach, operations are designed to ensure maximum emphasis on the effects on model-parameter deviations. Although some parameter estimation techniques, such as Kalman filters, are mature, engineers are still unable to estimate all parameters of a complex system at once due to the existence of multiple-solution cases. Our method provides ideas on how one can magnify the effects of all the parameters of interest through optimal operation design. The method also composes different maneuvers as an optimal operation, reducing the likelihood of multiple solutions. For verification, an implementation of a mathematical model proved that the metamodel-based methods are feasible. We discussed the mechanism underlying the precise estimation of the best conditions for PC-KF. The method's feasibility in engineering applications was demonstrated by applying it to an x-by-wire tricycle. As a result, the estimation of model parameters with the designed operation may lead to more confident and precise estimates. The complete process—including GSA, operation design, parameter estimation, and validation—was not only novel but also demonstrated to result in good performance.

Future studies should address some challenges. First, the relationship between parameter identifiability and the precision/quality of estimation that can be reached remains uncertain. One cannot guarantee a precise estimate under the excitation operation for interaction effects between model parameters. Thus, tuning a single parameter at a time is a difficult task; techniques for estimating all parameters at once continue to be crucial. In other words, the benefit of this proposed method must be quantified in several contexts. Systematic methods must address complex systems with numerous states; thus, future studies must discuss (1) normalizing scale differences on variance and (2) assigning weights for fusing the output. In our future work, we also plan to apply the proposed techniques to identify the parameters of a real mechanical system; we intend to not only test the applicability of this method but also investigate and expand on the basic assumptions on the contributors to model errors. We hope that, in the pursuit of superior computer models, extended methods based on our work will not only manage parameter differences but also quantify and excite model defections.

Acknowledgment

This work is partially supported by the Ministry of Science and Technology in Taiwan under grant MOST 108-2221-E-002-134-MY3. This support is gratefully acknowledged.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The authors attest that all data for this study are included in the paper.

References

[1] Kutluay, E., and Winner, H., 2014, "Validation of Vehicle Dynamics Simulation Models—A Review," Veh. Syst. Dyn., **52**(2), pp. 186–200.
[2] Althoff, M., and Magdici, S., 2016, "Set-Based Prediction of Traffic Participants on Arbitrary Road Networks," IEEE Trans. Intell. Veh., **1**(2), pp. 187–202.
[3] Hirschmann, D., Tissen, D., Schroder, S., and De Doncker, R. W., 2007, "Reliability Prediction for Inverters in Hybrid Electrical Vehicles," IEEE Trans. Power Electron., **22**(6), pp. 2511–2517.
[4] ASME V&V 10 Committee and Others, 2006. "Guide for Verification and Validation in Computational Solid Mechanics." The American Society of Mechanical Engineers, New York, Technical Report No. V&V 10-2006.
[5] Allen, R. W., Rosenthal, T. J., Klyde, D. H., Owens, K. J., and Szostak, H. T., 1992, "Validation of Ground Vehicle Computer Simulations Developed for Dynamics Stability Analysis," Technical Report, SAE Technical Paper.
[6] Aster, R., Borchers, B., and Thurber, C., 2018, *Parameter Estimation and Inverse Problems*, Elsevier, Boston, MA.
[7] Blundell, M., and Harty, D., 2004, *Multibody Systems Approach to Vehicle Dynamics*, Elsevier Butterworth-Heinemann, Oxford, UK.
[8] Setiawan, J. D., Safarudin, M., and Singh, A., 2009, "Modeling, Simulation and Validation of 14 DOF Full Vehicle Model," International Conference on Instrumentation, Communication Information Technology, and Biomedical Engineering, Indonesia.
[9] ISO Central Secretary, 2018, "Passenger Cars—Test Track for a Severe Lane-Change Manoeuvre—Part 1: Double Lane-Change," Standard ISO/TR 3888-1:2018, International Organization for Standardization, Geneva, CH.
[10] Gawade, T., Mukherjee, S., and Mohan, D., 2005, "Six-Degree-of-Freedom Three-Wheeled-Vehicle Model Validation," Proc. Inst. Mech. Eng. Part D J. Autom. Eng., **219**(4), pp. 487–498.
[11] Presse, C., and Gautier, M., 1993, "New Criteria of Exciting Trajectories for Robot Identification," Proceedings IEEE International Conference on Robotics and Automation, Atlanta, GA.
[12] Calafiore, G. C., Indri, M., and Bona, B., 2001, "Robot Dynamic Calibration: Optimal Excitation Trajectories and Experimental Parameter Estimation," J. Rob. Syst., **18**(2), pp. 55–68.
[13] Roeser, M. S., 2018, "Multi-axis Maneuver Design for Aircraft Parameter Estimation," Deutscher Luft- und Raumfahrtkongress 2018, Workshop Item.
[14] Tsai, H.-T., and Chan, K.-Y., 2019, "Investigating the Impact of Component Uncertainty on Autonomous Vehicle Overtaking Maneuvers," ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Anaheim, CA, Aug. 18–21.
[15] James, G., Carne, T. G., and Lauffer, J. P., 1995, "The Natural Excitation Technique (NExT) for Modal Parameter Extraction From Operating Structures," Modal Anal. Int. J. Anal. Exp. Modal Anal., **10**(4), p. 260.
[16] Narendra, K. S., and Annaswamy, A. M., 1987, "Persistent Excitation in Adaptive Systems," Int. J. Control, **45**(1), pp. 127–160.
[17] Blanchard, E. D., Sandu, A., and Sandu, C., 2010, "A Polynomial Chaos-Based Kalman Filter Approach for Parameter Estimation of Mechanical Systems," ASME J. Dyn. Syst. Meas. Control, **132**(6), p. 061404.
[18] Saltelli, A., Tarantola, S., and Chan, K.-S., 1999, "A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output," Technometrics, **41**(1), pp. 39–56.
[19] Chen, W., Jin, R., and Sudjianto, A., 2005, "Analytical Variance-Based Global Sensitivity Analysis in Simulation-Based Design Under Uncertainty," ASME J. Mech. Des., **127**(5), pp. 875–886.
[20] 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," Comput. Phys. Commun., **181**(2), pp. 259–270.
[21] Sobol, I. M., 1967, "On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals," Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, **7**(4), pp. 784–802.
[22] Lamboni, M., Monod, H., and Makowski, D., 2011, "Multivariate Sensitivity Analysis to Measure Global Contribution of Input Factors in Dynamic Models," Reliab. Eng. Syst. Saf., **96**(4), pp. 450–459.
[23] Lamboni, M., Makowski, D., and Monod, H., 2008, "Multivariate Global Sensitivity Analysis for Discrete-Time Models," Technical Report 2008-3, auto-saisine.
[24] Sumner, T., Shephard, E., and Bogle, D., 2012, "A Methodology for Global-Sensitivity Analysis of Time-Dependent Outputs in Systems Biology Modelling," J. R. Soc. Interface, **9**(74), pp. 2156–2166.
[25] Wang, L., 2008, "Karhunen-loeve Expansions and Their Applications," PhD thesis, London School of Economics and Political Science, London.
[26] Huang, Y.-C., and Chan, K.-Y., 2010, "A Modified Efficient Global Optimization Algorithm for Maximal Reliability in a Probabilistic Constrained Space," ASME J. Mech. Des., **132**(6), p. 061002.
[27] Jin, R., 2004, "Enhancements of Metamodeling Techniques in Engineering Design," PhD thesis, University of Illinois at Chicago, Chicago, IL.
[28] Finkel, D., 2003, "DIRECT Optimization Algorithm User Guide," Technical Report, Center for Research in Scientific Computation, North Carolina State University.
[29] Tarantola, A., 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia.
[30] Wesemeier, D., and Isermann, R., 2009, "Identification of Vehicle Parameters Using Stationary Driving Maneuvers," Control Eng. Pract., **17**(12), pp. 1426–1431.

[31] Li, J., and Xiu, D., 2009, "A Generalized Polynomial Chaos Based Ensemble Kalman Filter With High Accuracy," J. Comput. Phys., **228**(15), pp. 5454–5469.
[32] Chan, T. F., and Ng, M. K., 1999, "Galerkin Projection Methods for Solving Multiple Linear Systems," SIAM J. Sci. Comput., **21**(3), pp. 836–850.
[33] Astroza, R., Ebrahimian, H., and Conte, J. P., 2015, "Material Parameter Identification in Distributed Plasticity Fe Models of Frame-Type Structures Using Nonlinear Stochastic Filtering," J. Eng. Mech., **141**(5), p. 04014149.
[34] Jazar, R. N., 2019, *Advanced Vehicle Dynamics*, Springer, New York.
[35] Zandieh, A., 2015, "Dynamics of a Three-Wheel Vehicle With Tadpole Design," Master's thesis, University of Waterloo, Waterloo.
[36] Cossalter, V., 2006, "Motorcycle Dynamics," Lulu.com.
[37] Bellem, H., Schönenberg, T., Krems, J. F., and Schrauf, M., 2016, "Objective Metrics of Comfort: Developing a Driving Style for Highly Automated Vehicles," Trans. Res. Part F: Traffic Psychol. Behav., **41**, pp. 45–54.