Design optimization with variable screening by interval-based sensitivity analysis

Qi Chang [1], Changcong Zhou [2], Matthias G.R. Faes [3], Marcos A. Valdebenito [4]

Abstract: Design optimization problems are very common in engineering practice. Determining their solution may be challenging when many design variables are involved. A means to cope with such large number of design variables consists of first screening influential variables which drive the objective function the most. Then the optimization is carried out with respect to the influential variables while the other noninfluential variables are fixed at specific values. There is no doubt that an accurate identification of influential variables is crucial for high-dimensional optimization problems. In this paper, an interval-based sensitivity index is introduced to identify the influential variables and is theoretically compared with other two types of existing indices. The performance of these indices for dimensionality reduction in optimization is examined by means of a test function. Then, the proposed procedure for high-dimensional design optimization with variable screening is analyzed considering two illustrative examples. Then, the proposed strategy is applied to a practical engineering problem involving an aeronautical hydraulic pipeline. The results show that the interval sensitivity index is an effective tool and is superior to other two existing sensitivity indices for variable screening in design optimization.

Keywords: Optimization; Variable screening; Sensitivity; Interval; Hydraulic pipeline

---

[1] Ph.D. Candidate, Department of Engineering Mechanics, Northwestern Polytechnical University, Youyi West Road 127, 710072 Xi’an, China. Email: qichang@mail.nwpu.edu.cn

[2] Professor, Department of Engineering Mechanics, Northwestern Polytechnical University, Youyi West Road 127, 710072 Xi’an, China (corresponding author). Email: changcongzhou@nwpu.edu.cn

[3] Professor, Chair for Reliability Engineering, TU Dortmund University, Leonhard-Euler-Strasse 5, 44227 Dortmund, Germany. Email: matthias.faes@tu-dortmund.de

[4] Chief Engineer, Chair for Reliability Engineering, TU Dortmund University, Leonhard-Euler-Strasse 5, 44227 Dortmund, Germany. Email: marcos.valdebenito@tu-dortmund.de

Introduction

In engineering practice, a design optimization problem involves improving a predefined performance measure by selecting input variables of the considered system subject to certain constraints (Jamian et al. 2014; Martins and Lambe 2013; Valdebenito and Schuëller 2010; Wu et al. 2021). An optimization problem can be formulated as:

$$\begin{aligned} &\text{Find } \boldsymbol{X}^* = \{X_1^*, X_2^*, \cdots, X_n^*\} \\ &\text{to minimize } f(\boldsymbol{X}) \\ &\text{subject to } G_i(\boldsymbol{X}) \leqslant 0,\ (i = 1, 2, \cdots, NC) \\ &\qquad\qquad \boldsymbol{X}^l \leq \boldsymbol{X} \leq \boldsymbol{X}^u, \end{aligned} \tag{1}$$

where $\boldsymbol{X}$, n, f, G_i, NC, $\boldsymbol{X}^l$ and $\boldsymbol{X}^u$ are the design variable vector, number of design variables, objective function, the *i*-th constraint function, number of constraints, lower bound vector of design variables, and upper bound vector of design variables, respectively. Note that the design variables are constrained in a range of upper and lower bounds, i.e., $\boldsymbol{X} \in \boldsymbol{X}^I = [\boldsymbol{X}^l, \boldsymbol{X}^u]$ where the superscript "I" denotes that they are represented as an interval. The complexity of the design optimization problem is, to some extent, determined by the number of design variables n and the size of the design space $\boldsymbol{X}^I$ (Jamian et al. 2014). In general, design variables are chosen as much as possible based on the experience of designers and engineers in the initial design stage. This may result in a high-dimensional complex optimization model (Jamian et al. 2014; Martins and Lambe 2013).

High-dimensional complex optimization problems often encounter obstacles that mainly come from two aspects. First, the evaluation of the numerical model associated with a system often requires a considerable computational cost that may take several hours for a single simulation, and the design optimization usually requires hundreds or even thousands of iterative evaluations (Cho et al. 2014). Therefore, the computational cost of high-dimensional optimization problems is usually expensive. Second, optimization problems in engineering practice are often multidisciplinary, which will increase the complexity of the problem and can lead to the failure of convergence to an optimal solution (Cho et al. 2014; Spagnol et al. 2019; Wang et al. 2021, 2018). It is worth noting that engineers need to repeatedly utilize the established optimization model many times during the product design iteration process. Thus, high-dimensional complex optimization models are often unsuitable for subsequent use by engineers. In this situation, an effective approach, that is, design optimization based on pre-treatment of variables or so-called optimization with variable screening, is to perform design optimization in two stages (Spagnol et al. 2019; Wang et al. 2018). First, a sensitivity analysis is performed to rank the effect of all input variables and identify the most influential ones. Then, the optimization problem is remodeled by considering only these selected influential variables, while the other noninfluential variables are fixed at specific values in its space. It should be noted that variable screening is not the same as variable reformulation or feature projection, such as principal component analysis (Jolliffe 2002) or active subspace (Constantine et al. 2013) where the design space is changed.

Sensitivity analysis (also known as importance analysis) is a crucial step in many applications (Borgonovo and Plischke 2016; Wei et al. 2015). Through sensitivity analysis we gain essential insights into model behaviour, their structure and their response to changes on the model inputs (Borgonovo and Plischke 2016). In the past few decades, several sensitivity analysis methods have been developed, such as difference-based methods (Sobol' and Kucherenko 2009), parametric regression techniques (Härdle and Simar 2003), random forest techniques (Breiman 2001), variance-based methods (Homma and Saltelli 1996; Sobol' 1993) and moment-independent methods (Borgonovo 2007). Among them, variance-based methods have been particularly popular since the works of Sobol' (Sobol' 1993; Sobol' and Kucherenko 2009) were published (Homma and Saltelli 1996), and have been effectively applied to many engineering problems (Zhang et al. 2020).

Sensitivity analysis has emerged as a variable screening tool for high-dimensional complex optimization problems in various disciplines (Cho et al. 2014; Fesanghary et al. 2009; Li et al. 2020; Liu et al. 2020; Lu et al. 2015; Spagnol et al. 2019). In statistics, important variables among all candidate variables were accurately identified by a data-based sensitivity analysis (Duarte Silva 2001). Especially in statistical learning theory, different kinds of sensitivity analysis methods have been used to choose a subset of all input variables to model the system effectively (Guyon and Elisseeff 2003; Marra and Wood 2011; Marrel et al. 2008). In hydrology and earth systems (Li et al. 2013b), qualitative sensitivity analysis methods were employed to find 2-8 important variables from a total of 40 variables, and the total effect of Sobol's sensitivity method was used to quantify the contribution of each variable to the total variance of the model output (Li et al. 2013b). In addition, sensitivity analysis was used as a screening tool to reduce the computational cost associated with multi-objective design and rehabilitation of water distribution systems (Fu et al. 2012). In power and energy systems, sensitivity analysis was adopted to quantitatively compare the impacts of 24 variables in three major performances of a net-zero energy building (a promising solution to the worsening energy and environmental problems) (Zhang et al. 2020). Also, sensitivity analysis-based multi-objective optimization was performed to select the important operating parameters of a sinter cooler, thus achieving the optimal indicator parameters and obtaining the corresponding operating conditions (Tian et al. 2018). In the domain of structural design, sensitivity analysis methods were introduced or modified to carry out the design optimization of aeronautical hydraulic pipelines, which demonstrated that sensitivity-based variable screening methods can to some extent improve the efficiency of the optimization (Li et al. 2020; Wang et al. 2018; Zhang et al. 2019). Moreover, sensitivity-based variable screening methods have been used for automotive crashworthiness design (Craig et al. 2005), composite fuselage frame design (Gao et al. 2019), occupant restraint system design (Liu et al. 2020), thin-walled structure design, and dynamic design optimization of multi-motor driving transmission systems (Shu et al. 2018).

Regarding the sensitivity indices that the above-mentioned works used, most of them are variance-based measures, such as Sobol' indices. However, for the variable screening of optimization problems, variance-based indices may have some limitations because they only focus on the effect of a variable or multiple variables on the total variance of the model output rather than the model output itself, which may fail to provide a reasonable evaluation of the importance of each input (Fort et al. 2016; Spagnol et al. 2019). Moreover, variance-based indices are appropriately implemented in a probabilistic framework where inputs are represented by probabilistic information (e.g., a normal distribution or uniform distribution). However, the design variables of an optimization problem are usually described by interval models where no probabilistic information can be used (Martins and Lambe 2013). A compromise treatment is to directly consider the interval design variables as normal distribution or uniform distribution (Arwade et al. 2010; Fesanghary et al. 2009; Zhang et al. 2020), which is a clear violation of the interval paradigms as artificial information is added to the analysis (Chang et al. 2021; Zhou et al. 2021). There are relatively few works related to sensitivity analysis of interval variables (Faes and Moens 2020; Wang et al. 2018). The available approach seems to be the work of Wang et al. (Wang et al. 2018), which is also used as screening tools in the design optimization problem. Unfortunately, Wang's approach struggles with some drawbacks: the considered output has no clear physical meaning, and the indices may become inadequate in specific circumstances, which will be elaborated and compared in this work. Recently, Chang et al. (Chang et al. 2022) proposed a new interval sensitivity index, denoted as interval-based sensitivity index in the following, which can provide an intuitive interpretation on the behaviour of model response with respect to interval inputs. The model response can be regarded as the objective function of a design optimization problem, and the interval inputs are exactly consistent with the interval constraints of design variables. In this work, the interval-based sensitivity index is employed to screen the influential design variables for high-dimensional complex optimization problems. Meanwhile, the effectiveness of the interval-based sensitivity index will be compared with the indices of Sobol' and Wang, by its ability to find the influential variables.

The remainder of this work is organized as follows. Firstly, the interval-based sensitivity index is reviewed and is compared with other indices by a test function. In accordance with the design optimization strategy based on the variable screening by sensitivity analysis, which is explained by two illustrative optimization examples. Subsequently, the design optimization strategy is applied to an aeronautical hydraulic pipeline system. Finally, the conclusions are drawn.

Review and interpretation of the interval-based sensitivity index

In this section, the general concept of the interval field and its normalized form are introduced to discuss the connection between the interval field and variable sensitivity. Then, the interval-based sensitivity index, with a corresponding computational strategy, is reviewed, and its related characteristics and advantages are explained with the help of a test function. The index will be used to rank all input variables and screen the influential variables in next sections.

Interval field

A model $Y = f(\boldsymbol{X})$, with all classes of functions $f: \mathbb{R}^n \mapsto \mathbb{R}, \boldsymbol{X} \to Y$, is considered, where $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ is the vector of *n*-dimensional input variables and Y is the output. As reported in some optimization studies (Faes et al. 2017; Jiang et al. 2014; Wang et al. 2019, 2021), the design variables $\boldsymbol{X}$ are assumed to be independent and expressed by the interval model as

$$X_i \in X_i^I = [X_i^l, X_i^u], \tag{2}$$

where the superscript "I" denotes that the parameter is interval valued, $i = 1, 2, \cdots, n$, X_i^l and X_i^u with $X_i^l \leqslant X_i^u$ being respectively the lower and upper bounds of X_i. In addition, an interval can be represented by its centre value $X_i^c = (X_i^u + X_i^l)/2$ and its radius $X_i^r = (X_i^u - X_i^l)/2$. In case $\boldsymbol{X}$ belongs to an interval vector $\boldsymbol{X}^I$, the output is also an interval, which is defined as

$$Y^I = [Y^l, Y^u] = [f^l, f^u] = [\min_{\boldsymbol{X} \in \boldsymbol{X}^I} f(\boldsymbol{X}), \max_{\boldsymbol{X} \in \boldsymbol{X}^I} f(\boldsymbol{X})]. \tag{3}$$

When X_i is fixed to a specific value in the interval $[X_i^l, X_i^u]$ and all other inputs are intervals, the corresponding output $Y_{X_i}^I$ is a subinterval of Y^I, i.e., $Y_{X_i}^I \subseteq Y^I$. The bounds of $Y_{X_i}^I$ are calculated as

$$\begin{aligned} Y_{X_i}^I &= [Y_{X_i}^l, Y_{X_i}^u] = [f^l(X_i), f^u(X_i)] \\ &= [\min_{\boldsymbol{X}_{\sim i} \in \boldsymbol{X}_{\sim i}^I} f(X_1, \cdots X_{i-1}, X_i, X_{i+1}, \cdots X_n), \max_{\boldsymbol{X}_{\sim i} \in \boldsymbol{X}_{\sim i}^I} f(X_1, \cdots X_{i-1}, X_i, X_{i+1}, \cdots X_n)], \end{aligned} \tag{4}$$

where $\boldsymbol{X}_{\sim i}$ denotes the vector of all input variables except X_i and $\boldsymbol{X}_{\sim i}^I$ denotes the interval vector associated with all variables except for X_i^I. Furthermore, when X_i loops through every element within $[X_i^l, X_i^u]$, the output will be an interval field $f^I(X_i)$, which is shown in Fig. 1 (A). The whole interval field $f^I(X_i)$ is enveloped by the lower bound $f^l(X_i)$ and upper bound $f^u(X_i)$, with $f^l(X_i) \leqslant f^u(X_i)$. The centre value and radius of the interval field can be defined as (Jiang et al.

2014)

$$f^c(X_i) = \frac{f^u(X_i) + f^l(X_i)}{2} \quad (5)$$

and

$$f^r(X_i) = \frac{f^u(X_i) - f^l(X_i)}{2}. \quad (6)$$

According to interval theory, the centre value $g^c(X_i)$ reflects the general degree of the interval field, and the radius reflects the deviation degree of the interval field.

To facilitate the comparison of interval quantities and have a clearer and more intuitive interpretation of the relationships between interval variables and the model output, the interval field can be normalized as (Jiang et al. 2014)

$$f^I(X_i) = f^I(X_i^c + X_i^r \delta_i) = g^I(\delta_i), \quad (7)$$

where $\delta_i \in [-1, 1]$ is the normalized interval variable. Accordingly, the interval field $f^I(X_i)$ is transformed into the normalized interval field $g^I(\delta_i)$ (shown in Fig. 1 (B)) which can be described by $g^c(\delta_i)$ and $g^r(\delta_i)$:

$$g^c(\delta_i) = \frac{g^u(\delta_i) + g^l(\delta_i)}{2} \quad (8)$$

$$g^r(\delta_i) = \frac{g^u(\delta_i) - g^l(\delta_i)}{2}. \quad (9)$$

A recently proposed interval-based sensitivity index

For all of the input variables $X_1, X_2, \cdots, X_n$, the corresponding normalized interval field $g^I(\delta_i)$ $(i = 1, 2, \cdots, n)$ is usually different, and the difference between them is mainly determined by the influences of model inputs $\boldsymbol{X}$ on the model output Y (Chang et al. 2022). This is reflected in the radius of the interval field $g^I(\delta_i)$. Two special examples are shown in Fig. 1 (C) and (D):

(i) In Fig. 1 (C), the lower bound $g^l(\delta_i)$ is always equal to Y^l (i.e., $g^l(\delta_i) \equiv Y^l$) and the upper bound $g^u(\delta_i)$ is always equal to Y^u (i.e., $g^u(\delta_i) \equiv Y^u$), which indicates that no matter what value X_i takes in the interval $[X_i^l, X_i^u]$, the model output $Y_{X_i}^I$ remains unchanged and equals Y^I (i.e., $Y_{X_i}^I \equiv Y^I$). In this case, the input variable X_i has no influence on the model output, making X_i a noninfluential variable.

(ii) In Fig. 1 (D), the lower bound $g^l(\delta_i)$ and the upper bound $g^u(\delta_i)$ coincide with each other, i.e., $g^l(\delta_i) \equiv g^u(\delta_i)$, which indicates that the uncertainty of the model output Y is eliminated when X_i takes a specific value in the interval $[X_i^l, X_i^u]$. In this case, the model output is dominated by the input variable X_i. which makes X_i an important variable.

For the most common situation of the normalized interval field $g^I(\delta_i)$ shown in Fig. 1 (E), the whole area (surrounded by four lines $Y = Y^l$, $Y = Y^u$, $\delta_i = -1$ and $\delta_i = 1$) is divided into three parts: A_u enclosed by $g^u(\delta_i)$ and $Y = Y^u$, A_l enclosed by $g^l(\delta_i)$ and $Y = Y^l$, and A_{field} enclosed by $g^l(\delta_i)$ and $g^u(\delta_i)$. According to the description, the larger A_{field} is, the less important the variable is (e.g., the case of Fig. 1 (C)). In contrast, the smaller A_{field} is, the more important the variable is (e.g., the case of Fig. 1 (D)). Therefore, an interval sensitivity index C_i was defined by Chang et al. (Chang et al. 2022) as

$$C_i = \frac{A_u(i) + A_l(i)}{A_{total}(i)}, \tag{10}$$

where $A_{total}(i) = A_u(i) + A_l(i) + A_{field}(i)$. Furthermore, C_i can be derived by the following expression (Chang et al. 2022):

$$\begin{aligned} C_i &= \frac{A_u(i) + A_l(i)}{A_{total}(i)} \\ &= \frac{A_{total}(i) - A_{field}(i)}{A_{total}(i)} = 1 - \frac{A_{field}(i)}{A_{total}(i)} \\ &= 1 - \frac{1}{2(Y^u - Y^l)} \int_{-1}^{1} (g^u(\delta_i) - g^l(\delta_i)) \mathrm{d}\delta_i. \end{aligned} \tag{11}$$

In addition, the interval sensitivity index can be extended when we wish to know the joint contribution of two or more variables for the model output, which are derived in Appendix A. More information about the interval sensitivity index can be found in (Chang et al. 2022).

After introducing the interval sensitivity index, a test function will be studied next to compare the sensitivity analysis results of three types of indices including the interval-based sensitivity index, Sobol' indices (reviewed in Appendix B), and Wang's indices (reviewed in Appendix C).

An illustrative function to interpret the interval sensitivity index

A four-variable function is considered as (Chang et al. 2021)

$$f(\boldsymbol{X}) = X_1X_2 - 2X_2X_3 + 3X_3X_4 - 4X_4X_1 + 4X_1X_2X_3X_4, \tag{12}$$

where $\boldsymbol{X} = (X_1, X_2, X_3, X_4)$ are the input variables. Two cases are considered for this function, which are listed in Table 1. It should be noted that the Sobol' indices in the following are calculated on the premise that all input variables are assumed to be uniformly distributed in its interval. Sobol's indices are included for comparative reasons, and stem from the case where an analyst would invoke the maximum entropy principle to first estimate a distribution based on the available information (i.e., bounds), followed by the sensitivity analysis. Further, it should be noted that both sets of indices have a fundamentally different meaning behind them. Nonetheless, in view of selecting the most appropriate parameters to perform optimization, the usage of these indices is similar, which warrants comparison.

For Case 1, the results of C_i are shown in Fig. 2 (A), and its corresponding normalized interval fields are shown in Fig. 2 (D), where the order of dashed area is $A_3 > A_2 \approx A_4 > A_1$. This indicates that the order of the variable importance is $X_3 > X_2 \approx X_4 > X_1$. In addition, the results of Wang's indices (ζ_i and ς_i) and Sobol' indices (S_i and $S_{\mathrm{T}i}$) are shown in Fig. 2 (B) and (C), respectively. The variable importance results of the three types of indices are the same (i.e., $X_3 > X_2 \approx X_4 > X_1$) except that the result of radius sensitivity ς_i is $X_3 > X_2 \approx X_4 \approx X_1$.

For Case 2, the results of C_i are shown in Fig. 3 (A), and its corresponding normalized interval fields are shown in Fig. 3 (D), where the order of dashed area is $A_4 > A_1 \approx A_3 > A_2$. According to Fig. 3 (A) and (D), $C_4 > C_1 \approx C_3 > C_2$ indicates that the order of the variable importance is $X_4 > X_1 \approx X_3 > X_2$. In Fig. 3 (B), the orders of Wang's indices are $\zeta_1 > \zeta_2 > \zeta_4 > \zeta_3$ and $\varsigma_1 > \varsigma_3 > \varsigma_4 > \varsigma_2$, respectively. In Fig. 3 (C), the order of Sobol' indices $S_i \approx 0$ $(i = 1, 2, 3, 4)$ and $S_{\mathrm{T}4} > S_{\mathrm{T}1} > S_{\mathrm{T}3} > S_{\mathrm{T}2}$, respectively. The variable importance result of the total effect index $S_{\mathrm{T}i}$ is almost the same as the interval-based sensitivity index C_i. Nevertheless, the variable importance order cannot be obtained by Wang's indices because the ζ_i and ς_i results are inconsistent.

Discussion on the interval sensitivity index

According to the introduction of the interval-based sensitivity index and the analysis results of the illustrative function, four aspects of discussions need to be emphasized to offer deeper insights into the index.

First, three characteristics of the interval-based sensitivity index need to be emphasized:

(i) As shown in previous sections, the proposed index C_i is the ratio of the left dashed to the total rectangular area; therefore, the lower bound of C_i is 0, while the upper bound of C_i is 1, i.e., $0 \leqslant C_i \leqslant 1$. The larger C_i is, the more important X_i is. Conversely, the smaller C_i is, the less important X_i is.

(ii) If $C_i = 0$ (as shown in Fig. 1 (C)), the uncertainty of X_i has no influence on the model output Y. Therefore, X_i is a noninfluential variable in this case.

(iii) If $C_i = 1$ (as shown in Fig. 1 (D)), the uncertainty of X_i has dominant influence on the model output Y. Thus, X_i is the most important variable in this case.

Second, similar to conventional sensitivity indices, such as Sobol' indices (Sobol' 1993; Sobol' and Kucherenko 2009; Wei et al. 2015a) and Wang's indices (Wang et al. 2018), the interval-based sensitivity index C_i can be used for the sensitivity analysis of the input variables, thus providing a deeper understanding of the considered system and more comprehensive analysis results for engineering designers. At the same time, the index can be used as a variable screening tool for optimization problems, which will be specifically implemented in next sections. In particular, those variables with relatively large index C_i are selected as influential variables, and then the optimization problem is solved by considering only these selected influential variables, while the other noninfluential variables are evaluated at specific values in its space. Moreover, as shown in Fig. 2 (D) and Fig. 3 (D), the interval-based sensitivity analysis can not only rank the importance of input variables, but also present the behaviour of model response with respect to each interval input, thus providing the guidance for the determination/realization of value-fixation of noninfluential variables. That is to say, fixing these noninfluential variables can be determined with the help of the interval-based sensitivity analysis results, rather than crudely fix all noninfluential variables at its initial values (e.g., lower bounds, upper bounds, or centre values). The specific implementation will be elaborated in next section.

Third, the classical Sobol' indices (reviewed in Appendix B) are variance-based indices (Sobol' 1993; Sobol' and Kucherenko 2009; Wei et al. 2015). For the design optimization with variable screening, selecting Sobol' indices as a screening tool may have three limitations, which are analyzed as follows:

(i) Sobol' indices reflect the contribution of input variables to the model output variance (rather than the model output itself), which is just a variance-based criterion, rather than a comprehensive characteristic description of the model output (Cho et al. 2014).

(ii) In design optimization problems, the design variables are often confined by the lower and upper bounds of intervals rather than characterized by probabilistic models. It may be inappropriate to directly consider the design variable as a specific probabilistic model (e.g., normal distribution or uniform distribution).

(iii) Sobol' indices can only offer the importance ranking results, but not the behaviour of model response with respect to each interval input. Hence, Sobol' indices may not provide sufficient information for fixing the value of noninfluential variables.

Fourth, Wang's indices (Wang et al. 2018) (reviewed in Appendix C) can be used to identify influential variables in design optimization problems, though three drawbacks need to be emphasized:

(i) Wang's indices reflect the contribution of input variables to $\eta = Y^c/Y^r$, which is the ratio of the centre value to the radius of the model output, rather than the model output itself. Moreover, η has no clear physical meaning.

(ii) According to the introduction of Appendix C, $\eta = Y^c/Y^r$ is equal to 0 when $Y^c = 0$. In this case, the definition of Eqs. (C6) and (C7) (i.e., $\zeta_i = \left|\eta - [\eta|X_i]^c\right|/\eta$ and $\varsigma_i = [\eta|X_i]^r/\eta$) are invalid because η is the denominator in Eqs. (C6) and (C7). Therefore, Wang's indices can only be used in the case of $Y^c \neq 0$. Meanwhile, when η is approximately equal to 0 (i.e., $\eta \approx 0$), the results of ζ_i and ς_i are likely to be influenced by calculation errors, which can be clearly realized by the comparation results of Case 1 and Case 2 in the illustrative function of the previous section.

(iii) Same as Sobol' indices, Wang's indices cannot offer insight on the behaviour of model response with respect to each interval input. As a result, Wang's indices may not be informative for fixing values of noninfluential variables. The details will be elaborated and examined in next sections.

In addition, the effectiveness and superiorities of the interval-based sensitivity index will be further compared with Sobol' indices and Wang's indices by its ability to find the influential variables of design optimization problems in the illustration examples of next section.

Computational strategy for calculating the interval sensitivity index

As shown in Eq. (11), the interval-based sensitivity index C_i can be calculated using Y^l, Y^u, $g^l(\delta_i)$, and $g^u(\delta_i)$. The calculation of them is addressed in detail in the following.

Calculation of Y^l and Y^u

The calculation of Y^l and Y^u is a classical interval analysis problem, which has been discussed in depth by Faes et al. (Faes et al. 2017; Faes and Moens 2020). In this section, Y^l and Y^u are obtained by solving the following two optimisation problems:

$$\begin{aligned}&\text{Find } \boldsymbol{\delta}^* = \{\delta_1^*, \delta_2^*, \cdots, \delta_n^*\}\\&\text{to minimize } g(\boldsymbol{\delta})\\&\text{subject to } -1 \le \delta_i \le 1,\ (i = 1, 2, \cdots, n)\end{aligned} \tag{13}$$

and

$$\begin{aligned}&\text{Find } \boldsymbol{\delta}^{**} = \{\delta_1^{**}, \delta_2^{**}, \cdots, \delta_n^{**}\}\\&\text{to maximize } g(\boldsymbol{\delta})\\&\text{subject to } -1 \le \delta_i \le 1,\ (i = 1, 2, \cdots, n).\end{aligned} \tag{14}$$

We can observe that $Y^l = g(\boldsymbol{\delta}^*)$ and $Y^u = g(\boldsymbol{\delta}^{**})$. In this work, the surrogate optimisation (SO) algorithm (Wang and Shoemaker 2014), integrated as "surrogateopt" function in MATLAB software, is adopted to implement the two optimisation problems (i.e., Eqs. (13) and (14)) to obtain Y^l and Y^u.

Calculation of $g^{l}(\delta_i)$ and $g^{u}(\delta_i)$

For the calculation of $g^{l}(\delta_i)$ and $g^{u}(\delta_i)$, a double-loop strategy can be adopted (Chang et al. 2022). In the inner layer, the minimum and maximum values $g^{l}(\delta_i)$ and $g^{u}(\delta_i)$ of the output response for the given value δ_i are obtained by solving the following two optimization problems:

$$\begin{aligned}&\text{Find } \boldsymbol{\delta}'_{\sim i}=\{\delta'_1,\delta'_2,\cdots,\delta'_{i-1},\delta'_{i+1}\cdots,\delta'_n\}\\&\text{to minimize } g(\boldsymbol{\delta})\\&\text{subject to } -1\le\boldsymbol{\delta}_{\sim i}\le 1\\&\qquad\qquad\delta_i=\delta_{i(h)}\end{aligned}\tag{15}$$

and

$$\begin{aligned}&\text{Find } \boldsymbol{\delta}''_{\sim i}=\{\delta''_1,\delta''_2,\cdots,\delta''_{i-1},\delta''_{i+1}\cdots,\delta''_n\}\\&\text{to maximize } g(\boldsymbol{\delta})\\&\text{subject to } -1\le\boldsymbol{\delta}_{\sim i}\le 1\\&\qquad\qquad\delta_i=\delta_{i(h)}\ ,\end{aligned}\tag{16}$$

where $h=1,2,\cdots,n_{\text{sample}}$, n_{sample} denotes the number of generated grid points of δ_i. We can observe that $g^{l}(\delta_{i(h)})=g(\boldsymbol{\delta}'_{\sim i},\delta_{i(h)})$ and $g^{u}(\delta_{i(h)})=g_h(\boldsymbol{\delta}''_{\sim i},\delta_{i(h)})$. Also, SO algorithm as described in previous section is adopted to solve Eqs. (15) and (16). In the outer layer of the double-loop strategy, the lower bound curve $g^{l}(\delta_i)$ and upper bound curve $g^{u}(\delta_i)$ of the interval field $g^{I}(\delta_i)$ can be approximately fitted by global metamodeling with adaptive sampling. The details of global metamodeling method are described in (Chang et al. 2022) and the practical implementation of the strategy is summarized in Algorithm 1.

The optimisation algorithm SO is called to calculate the lower/upper bounds $g^{l}(\delta_{i(h)})/g^{u}(\delta_{i(h)})$ at a series of grid points $\delta_{i(h)}$. The initial points $\{\delta_{i(1)},\cdots,\delta_{i(n_{ini})}\}$ and their corresponding lower/upper bounds $\{g^{l}(\delta_{i(1)}),\cdots,g^{l}(\delta_{i(n_{ini})})\}/\{g^{u}(\delta_{i(1)}),\cdots,g^{u}(\delta_{i(n_{ini})})\}$ are firstly used to construct a metamodel, then the new point $\delta_{i(\text{new})}$ is selected, by maximizing the mean square error (Liu et al. 2018), to update the model until the stopping criterion is met. Finally, $g^{l}(\delta_i)$ and $g^{u}(\delta_i)$ are obtained to calculate the interval-based sensitivity index C_i.

Design optimization with variable screening by the interval-based sensitivity

index

General process

For the design optimization model expressed in Eq. (1), all input variables $\boldsymbol{X}=\{X_1,X_2,\cdots,X_n\}$ can be divided into two categories, i.e., top influential variables $\boldsymbol{X}_{\mathbf{Top}}=\{X_{\mathrm{Top}(1)},X_{\mathrm{Top}(2)},\cdots,X_{\mathrm{Top}(m)}\}$ and noninfluential variables $\boldsymbol{X}_{\mathbf{\sim Top}}=\{X_{\sim\mathrm{Top}(1)},X_{\sim\mathrm{Top}(2)},\cdots,X_{\sim\mathrm{Top}(n-m)}\}$, by sensitivity ranking results. The subscript "Top" refers to variables that are significantly more influential than other variables. As can be seen, $\boldsymbol{X}=\{\boldsymbol{X}_{\mathbf{Top}},\boldsymbol{X}_{\mathbf{\sim Top}}\}$ and $m\leqslant n$. Then, the design optimization problem can be reformulated as

$$\begin{aligned}
&\text{Find } \boldsymbol{X}^*_{\mathbf{Top}}=\{X^*_{\mathrm{Top}(1)},X^*_{\mathrm{Top}(2)},\cdots,X^*_{\mathrm{Top}(m)}\}\\
&\text{to minimize } f(\boldsymbol{X}_{\mathbf{Top}},\boldsymbol{X}_{\mathbf{\sim Top}})\\
&\text{subject to } G_i(\boldsymbol{X}_{\mathbf{Top}},\boldsymbol{X}_{\mathbf{\sim Top}})\leqslant 0,\ (i=1,2,\cdots,NC)\\
&\qquad\qquad \boldsymbol{X}^l_{\mathbf{Top}}\le \boldsymbol{X}_{\mathbf{Top}}\le \boldsymbol{X}^u_{\mathbf{Top}}\\
&\qquad\qquad \boldsymbol{X}_{\mathbf{\sim Top}}=\boldsymbol{X}^{sv}_{\mathbf{\sim Top}},
\end{aligned}\tag{17}$$

where $\boldsymbol{X}^l_{\mathbf{Top}}$ and $\boldsymbol{X}^u_{\mathbf{Top}}$ are the lower and upper bound vectors of influential design variables $\boldsymbol{X}_{\mathbf{Top}}$, respectively. $\boldsymbol{X}^{sv}_{\mathbf{\sim Top}}$ is the selected specific value vector of the noninfluential variables $\boldsymbol{X}_{\mathbf{\sim Top}}$. As illustrated in the discussion on the interval sensitivity index, Sobol' and Wang's indices cannot offer insight on the behaviour of model response with respect to each interval input. Therefore, all noninfluential variables $\boldsymbol{X}_{\mathbf{\sim Top}}$ are crudely fixed at its centre values (i.e., $\boldsymbol{X}_{\mathbf{\sim Top}}=\boldsymbol{X}^c_{\mathbf{\sim Top}}$) (or lower bounds (i.e., $\boldsymbol{X}_{\mathbf{\sim Top}}=\boldsymbol{X}^l_{\mathbf{\sim Top}}$) and upper bounds (i.e., $\boldsymbol{X}_{\mathbf{\sim Top}}=\boldsymbol{X}^u_{\mathbf{\sim Top}}$)). Meanwhile, for the interval-based sensitivity index, the *k*-th noninfluential variable $X_{\sim\mathrm{Top}(k)}$ $(k=1,2,\cdots,n-m)$ is fixed as a specific value $X^{sv}_{\sim\mathrm{Top}(k)}$, where $X^{sv}_{\sim\mathrm{Top}(k)}\in[X^l_{\sim\mathrm{Top}(k)},X^u_{\sim\mathrm{Top}(k)}]$ is the value that corresponds to the minimum of the whole interval field $f^I(X_{\sim\mathrm{Top}(k)})$, i.e., $f^I(X^{sv}_{\sim\mathrm{Top}(k)})=Y^l$. Correspondingly, when solving a maximization problem, the specific value $X^{sv}_{\sim\mathrm{Top}(k)}$ will be the value that corresponds to the maximum of the whole interval field $f^I(X_{\sim\mathrm{Top}(k)})$, i.e., $f^I(X^{sv}_{\sim\mathrm{Top}(k)})=Y^u$. The general process of the design optimization with variable screening can be summarized in Fig. 4.

It can be seen that the objective function f and constraint function G_i change from the original *n*-dimensional functions to *m*-dimensional functions, and the number of the design variables is reduced from the original n to m, which will effectively simplify the optimization problem. It should be noted that the specific value of m is determined by the designers based on sensitivity analysis results, model accuracy requirements and acceptable calculating costs, which will be specifically demonstrated by optimization examples in the next section. In addition, for a complex optimization problem in engineering practice (e.g., the aeronautical hydraulic pipeline) where the model response is obtained by a time-consuming black box (e.g., finite element analysis), the sensitivity analysis may require a lot of computational cost that designers cannot afford. In this case, metamodels can be adopted to reduce intensive computational costs of the sensitivity analysis (Chang et al. 2022).

Illustrative examples of the design optimization with variable screening

Two illustrative optimization examples are conducted to examine the design optimization strategy and to compare the ability of different sensitivity indices to find the optimal results (including the maximum result and the minimum result) and their effect on variable screening. It should be noted that the Sobol' indices are calculated on the premise that all input variables are assumed to be uniformly distributed in its interval. Besides, the sequential quadratic programming (SQP) optimizer is employed for the two illustrative optimization examples.

A ten-dimensional optimization example

A ten-dimensional optimization example (Wang et al. 2018) is expressed as

$$
\begin{aligned}
&\text{Find } \boldsymbol{X}^* = \{X_1^*, X_2^*, \cdots, X_{10}^*\} \\
&\text{to maximize/minimize } f(\boldsymbol{X}) = X_1^2 + X_2^2 - X_3^2 - X_{10}^2 + X_1 X_2 - 14X_1 - 16X_2 \\
&\qquad + (X_3 - 10)^2 + 4(X_4 - 5)^2 + (X_5 - 3)^2 + 2(X_6 - 1)^2 + 5X_7^2 \\
&\qquad + 7(X_8 - 11)^2 + 2(X_9 - 10)^2 + (X_{10} - 7)^7 + 20X_3 + 14X_{10} - 50 \\
&\text{subject to } X_i^l \le X_i \le X_i^u,\ (i = 1, 2, \cdots, 10),
\end{aligned}
\tag{18}
$$

where the lower and upper bounds of all interval variables are assumed to be 5 and 6, respectively, i.e., $X_i^l = 5$ and $X_i^u = 6$ for $i = 1, 2, \cdots, 10$. The actual lower and upper bounds of the model output are $f^l(\boldsymbol{X}) = 391.75$ and $f^u(\boldsymbol{X}) = 572$, and the corresponding centre and radius are

$f^c(\boldsymbol{X}) = 481.875$ and $f^r(\boldsymbol{X}) = 90.125$.

As shown in Fig. 5 (A), the importance ranking obtained by the interval-based sensitivity index C_i is $X_8 > X_7 > X_9 \approx X_6 > X_5 \approx X_4 > X_1 \approx X_2 > X_3 \approx X_{10}$. Fig. 5 (D) shows the normalized interval field used to obtain C_i $(i = 1, 2, \cdots, 10)$. Fig. 5 (E) is the normalized interval field used to obtain C_{78}, which indicates the joint contribution of X_7 and X_8 to the model output $f(\boldsymbol{X})$. The importance ranking obtained by Wang's indices (Fig. 5 (B)) and Sobol' indices (Fig. 5 (C)) are the same as that of the index C_i. It should be noted that the three types of sensitivity indices associated with X_3 and X_{10} are equal to zero, (i.e., $C_3 = \zeta_3 = \varsigma_3 = S_3 = S_{\mathrm{T}3} = 0$ and $C_{10} = \zeta_{10} = \varsigma_{10} = S_{10} = S_{\mathrm{T}10} = 0$), and the three types of indices associated with X_1 and X_2 are very small and close to zero. The value of the model output is basically unchanged when these four variables (X_3, X_{10}, X_1 and X_2) change within its interval range. In contrast, the three types of indices associated with X_8, X_7, X_6 and X_9 are obviously larger than other variables, which means that X_8, X_7, X_6 and X_9 have a considerable influence on the model output.

In Fig. 5 (F), the optimization results (including the maximum value $f_{\max}$ and minimum value $f_{\min}$) as well as the computational cost (measured by N_{call}, i.e., the number of function calls for SQP optimizer during the optimization process) corresponding to different sets of design variables are given. In the horizontal abscissa, Top m $(m = 1, 2, \cdots, 10)$ represents the first m most influential design variables according to the sensitivity analysis. For example, Top 4 means that X_8, X_7, X_6 and X_9 are considered in the optimization, and the rest are fixed at a specific value (such as centre values, i.e., $\boldsymbol{X}_{\sim\mathbf{Top}} = \boldsymbol{X}^c_{\sim\mathbf{Top}}$, or a set of selected values, i.e., $\boldsymbol{X}_{\sim\mathbf{Top}} = \boldsymbol{X}^{sv}_{\sim\mathbf{Top}}$) in its intervals. When all the noninfluential variables are fixed at their centre values (i.e., $\boldsymbol{X}_{\sim\mathbf{Top}} = \boldsymbol{X}^c_{\sim\mathbf{Top}}$), the maximum and minimum values are 544.99 and 413.00 in the case of Top 2, 564.00 and 396.00 in case of Top 4, 569.75 and 393.75 in the case of Top 6, and 572.00 and 391.75 in the case of Top 10 (all variables). Meanwhile, when all the noninfluential variables are fixed at a set of selected values $\boldsymbol{X}^{sv}_{\sim\mathbf{Top}}$ based on the information of the interval-based sensitivity analysis result, the maximum and minimum values are 572.00 and 391.75 for all the case of Top *m*. Take the minimum optimization of Top 2 case as an example, the Top 2 influential variables (i.e., X_8 and X_7) are selected as design variables while the noninfluential variables (i.e., $X_1 \sim X_6$, X_9 and X_{10}) are fixed at $\boldsymbol{X}_{\sim\mathbf{Top}}^{sv}$. According to Fig. 5 (D), $\boldsymbol{X}_{\sim\mathbf{Top}}^{sv} = \{X_1^l,\ X_2^c,\ X_3^\forall,\ X_4^l,\ X_5^l,\ X_6^l,\ X_9^u,\ X_{10}^\forall\}$ where the superscript $\forall$ of X_3 and X_{10} means any value in its intervals because the change of X_3 and X_{10} have no effect on the model output, and the superscripts *c*, *l* and *u* are the centre value, lower bound and upper bound, respectively. At the same time, when performing maximization of Top 2 case, the fixed values for the noninfluential variables can be selected as $\boldsymbol{X}_{\sim\mathbf{Top}}^{sv} = \{X_1^u,\ X_2^c,\ X_3^\forall,\ X_4^u,\ X_5^u,\ X_6^u,\ X_9^l,\ X_{10}^\forall\}$ based on the normalized interval field $g^I(\delta_i)$ in Fig. 5 (D).

In addition, the number of function calls during the optimization process (including N_{call} of $f_{\max}$ and $f_{\min}$) increases with the increase of variables. Satisfactory optimization results can be obtained by considering the influential design variables, e.g., Top 4, and the computational cost can be reduced by more than half in the case of Top 4. At the same time, when the noninfluential variables $\boldsymbol{X}_{\sim\mathbf{Top}}$ are fixed at $\boldsymbol{X}_{\sim\mathbf{Top}}^{sv}$, the optimal results (i.e., $f_{\max} = 572.00$ and $f_{\min} = 391.75$) of Top 10 (all variables) optimization can also be obtained by Top *m* ($m < 10$) optimization case, which can be seen in Fig. 5 (F). In other words, it is effective and practical to perform design optimization with these influential variables selected by the interval-based sensitivity index C_i. Moreover, the interval-based sensitivity analysis can provide the behaviour of model response with respect to each interval inputs, so as to give the guidance for the determination/realization of value-fixation of noninfluential variables. As can be seen in Fig. 5 (F), this strategy is obviously useful for finding the minimum/maximum value of the problem.

A twenty-dimensional optimization example

A twenty-dimensional optimization example (Craig et al. 2005; Welch et al. 1992) is expressed as

$$
\begin{aligned}
&\text{Find } \boldsymbol{X}^* = \{X_1^*, X_2^*, \cdots, X_{20}^*\} \\
&\text{to maximize/minimize } f(\boldsymbol{X}) = 5X_{12}/(1+X_1) + 5(X_4 - X_{20})^2 + X_5 + 40X_{19}^3 - 5X_{19} \\
&\qquad + 0.05X_2 + 0.08X_3 - 0.03X_6 + 0.03X_7 - 0.09X_9 - 0.01X_{10} \\
&\qquad - 0.07X_{11} + 0.25X_{13}^2 - 0.04X_{14} + 0.06X_{15} - 0.01X_{17} - 0.03X_{18} \\
&\text{subject to } X_i^l \le X_i \le X_i^u,\ (i = 1, 2, \cdots, 20)\ .
\end{aligned} \tag{19}
$$

The lower and upper bounds of all interval variables are assumed to be -0.5 and 0.5, respectively, i.e., $X_i^l = -0.5$ and $X_i^u = 0.5$ for $i = 1, 2, \cdots, 20$. The actual lower and upper bounds of the model output are $f^l(\boldsymbol{X}) = -8.152$ and $f^u(\boldsymbol{X}) = 13.3125$ (Craig et al. 2005; Welch et al. 1992), and the corresponding centre and radius are $f^c(\boldsymbol{X}) = 2.5803$ and $f^r(\boldsymbol{X}) = 10.7323$. The function is evidently nonlinear, and that there are six influential variables (X_1, X_4, X_5, X_{12}, X_{19} and X_{20}) as discussed in (Craig et al. 2005; Welch et al. 1992).

The results of three types of indices (i.e., the interval-based sensitivity index, Wang's indices and Sobol' indices) are compared in Fig. 6, where (A) is the interval-based sensitivity index C_i, (B) is Wang's indices including the centre value sensitivity ζ_i and radius sensitivity ς_i and (C) is the Sobol' indices including the main effect S_i and the total effect $S_{\mathrm{T}i}$. The six influential variables are successfully identified by C_i, ζ_i, ς_i and $S_{\mathrm{T}i}$. It can be observed that the main effect S_i corresponding to X_1 is approximately equal to 0, primarily due to the interaction influence between X_1 and the most influential variable X_{12} in the response function. In addition, the variable importance orders obtained by the five indices are different from each other, where the top influential variables are listed in Fig. 6 (A)-(C).

To compare the ability of these three types of indices to find the influential variables, the optimization results (including the maximum value $f_{\max}$ and minimum value $f_{\min}$) as well as the computational cost (measured by N_{call}, i.e., the number of function calls for SQP optimizer during the optimization process) corresponding to different sets of design variables are given in Fig. 7. Similar to the previous section, in the horizontal abscissa, Top m ($m = 1, 2, \cdots, 20$) represents the first m most influential design variables according to the sensitivity analysis. Therefore, the optimization results gradually tend to the optimal solution (i.e., $f_{\max} = 13.3125$ and $f_{\min} = -8.152$) with the increase of design variables. Additionally, the number of function calls (including N_{call} of $f_{\max}$ and $f_{\min}$)

increases with the increase of variables. For the case of Top 1, the maximum and minimum values associated with five indices (C_i, ζ_i, ς_i, S_i and $S_{\mathrm{T}i}$) are the same, i.e., $f_{\max} = 2.5$ and $f_{\min} = -2.5$. In the case of Top 2, the maximum and minimum values associated with four indices (C_i, ζ_i, ς_i and S_i) are the same ($f_{\max} = 5$ and $f_{\min} = -5$), while those associated with $S_{\mathrm{T}i}$ are $f_{\max} = 3.75$ and $f_{\min} = -2.5$, so the ability of $S_{\mathrm{T}i}$ is inferior to other four indices (C_i, ζ_i, ς_i and S_i) in the case of Top 2. The situations in the Top 3-6 and the Top 20 can be concluded in the same manner, where ς_i, $S_{\mathrm{T}i}$ and S_i are inferior to C_i and ζ_i in the case of Top 3, ς_i and S_i are inferior to C_i, ζ_i and $S_{\mathrm{T}i}$ in the case of Top 5, and S_i is inferior to the other four indices (i.e., C_i, ζ_i, ς_i and $S_{\mathrm{T}i}$) in the case of Top 6. The number of inferiors (i.e., the number of times that the Top *m* extrema associated with an index is inferior to the Top *m* extrema associated with one of other four indices) of the five indices are shown in the top left corner of Fig. 7, where the interval-based sensitivity index C_i and the centre value sensitivity ζ_i of Wang's indices are obviously superior to the other three indices in this example.

In addition, it can be seen that the minimum result (-7.5) of the Top 3 associated with C_i is very close to the minimum result (-8.152) of the Top 20, and the corresponding numbers of function calls of the Top 3 and Top 20 are 40 and 820, respectively, where more than 95% of computational cost is saved. At the same time, the maximum result (12.5) of the Top 5 associated with C_i is very close to the maximum result (13.3125) of the Top 20, and the corresponding number of function calls of the Top 5 and the Top 20 are 55 and 451, respectively, where more than 87% of the computational cost is saved.

Moreover, the above-discussed Top *m* ($m = 1, 2, \cdots, 6$) optimization is performed in the case of the noninfluential variables are fixed at its centre values (i.e., $\boldsymbol{X}_{\sim\mathbf{Top}} = \boldsymbol{X}^{c}_{\sim\mathbf{Top}}$). When all the noninfluential variables are fixed at a set of selected values $\boldsymbol{X}^{sv}_{\sim\mathbf{Top}}$ based on the information of the interval-based sensitivity analysis result, the maximum and minimum values are 13.3125 and -8.152 for all the case of Top *m*. Take the minimum optimization of Top 4 case as an example, the Top 4 influential variables (i.e., X_{12}, X_1, X_{19} and X_{20}) are selected as design variables while the noninfluential variables (i.e., $X_2 \sim X_{11}$ and $X_{13} \sim X_{18}$) are fixed at

$\boldsymbol{X}_{\sim\mathbf{Top}}^{sv} = \{X_2^l, X_3^l, X_4^\forall, X_5^l, X_6^u, X_7^l, X_8^\forall, X_9^u, X_{10}^u, X_{11}^u, X_{13}^c, X_{14}^u, X_{15}^l, X_{16}^\forall, X_{17}^u, X_{18}^u\}$ which are selected based on the normalized interval field $g^I(\delta_i)$ of the interval-based sensitivity analysis. As can be seen from above discussion, performing the design optimization with these few influential variables (selected by the index C_i), while the rest of noninfluential variables are fixed at the selected values $\boldsymbol{X}_{\sim\mathbf{Top}}^{sv}$, is effective and practical.

Application for the design of an aeronautical hydraulic pipeline

In this section, the interval-based sensitivity index C_i will be applied to a practical engineering problem of aeronautical hydraulic pipelines, which was first presented in Wang's work (Wang et al. 2018) and studied in (Li et al. 2020; Zhang et al. 2019; ZHOU et al. 2019). First, we introduce the problem of aeronautical hydraulic pipelines, then perform the optimization design with variable screening based on the interval-based sensitivity analysis result, before finally discussing the findings.

Introduction of the aeronautical hydraulic pipelines

For most modern aircrafts, hydraulic pipelines are functional units that transmit hydraulic oil to drive a series of mechanisms on board (Ouyang et al. 2012; Tang et al. 2011), such as flaps, landing gears, and folding mechanisms. At the same time, hydraulic pipelines undertake complex working conditions, such as external shock, vibration, or other dynamic loads. In particular, aeronautical hydraulic pipelines are quite different from those in other engineering areas, which is mainly reflected in three aspects (Zhou et al. 2019):

(i) Low stiffness. The stiffness of aeronautical hydraulic pipelines is relatively low due to space and weight constraints.

(ii) Long pipelines and a considerable number of curves. Aeronautical hydraulic pipelines are relatively long because of the long distances between actuators and pumps. Meanwhile, hydraulic pipelines are curved frequently, because of the arrangement of structural parts and the confined space on board. This may have impacts on the stability of the pipelines.

(iii) Complicated vibration environment. There are many vibration sources, such as engines, pumps and aerodynamics that can cause excitations of hydraulic pipelines.

In engineering practice, the stress and displacement are two quantities of concern, and the failure of aeronautical hydraulic pipeline is mainly due to fatigue load caused by vibration. To be specific, the stress values of all nodes of the hydraulic pipeline change with the excitation. In this study, the excitation of hydraulic pipelines is considered as stochastic due to random external vibration, the complexity of structures, and variations in aircraft loads (Wang et al. 2018). For a specific node, standard deviation of stress reflects the degree of stress change in the whole vibration cycle. The greater the standard deviation of stress, the easier the fatigue failure will occur. Therefore, the maximum standard deviation of stress is chosen as the objective of optimisation design. Meanwhile, the maximum standard deviation of displacement is treated as a constraint in this work. In general, a series of hoops (which are usually dozens or even hundreds) are used to fix hydraulic pipelines, thus preventing instability caused by long spans and excitation, which can obviously lower the stress and displacement in the vibration environment (Wang et al. 2018; Zhou et al. 2019). Therefore, the constraint location plays an important role on the safety of hydraulic pipelines. So, the coordinates of the hoop locations are selected as design variables.

As shown in Fig. 8 (A), the finite element model of an aeronautical hydraulic pipeline is built in ANSYS 17.1, and its related parameters are listed in the lower part of Fig. 8 (A) (Zhou et al. 2019). ρ_P, D, t, ρ_f, T, P, E and μ represent the material density, outer radius, thickness, fluid density, ambient temperature, internal pressure (which is caused by the flow of oil in pipeline), elastic modulus and Poisson's ratio, respectively. Twenty-eight design variables (coordinates of hoops) are selected from 21 hoop locations in Fig. 8 (B), which are listed in Table 2. It should be reminded that more than one coordinate is considered at some hoop locations including 45, 62, 69, 72, 75 and 85. The stochastic excitation is given in the form of an acceleration power spectral density (PSD) function, which is shown in Fig. 8 (C). In this case, the stress results of the hydraulic pipeline are obtained in Fig. 8 (D), where the maximum standard deviation of stress is $2.69 \times 10^7 (\mathrm{Pa})$.

The goal of the optimization is to find the optimal locations of the pipeline hoops to minimize the maximum standard deviation of stress. Therefore, the objection function is $\sigma_{\sigma\max}$. In addition, the maximum standard deviation of displacement $D_{\max}$ of the hydraulic pipeline is the constraint. The optimization problem of the hydraulic pipeline can be expressed as

$$\begin{aligned}
&\text{Find } \boldsymbol{X}^* = \{X_1^*, X_2^*, \cdots, X_{28}^*\} \\
&\text{to minimize } \sigma_{\sigma\max} \\
&\text{subject to } D_{\max} - D' < 0 \\
&\qquad\qquad X_i^l \leqslant X_i \leqslant X_i^u \ \ (i = 1, 2, \cdots, 28),
\end{aligned} \tag{20}$$

where D' is the maximum standard deviation threshold value of displacement, which is set as 5.5×10^{-4}. In other words, $D_{\max} < 5.5\times10^{-4}$ (m) must be satisfied in the whole optimization process. X_i^l and X_i^u are the lower and upper bounds of the design variable X_i, and are shown in Table 2. More detailed information on the maximum standard deviation of displacement $D_{\max}$ of the aeronautical hydraulic pipeline can be found in (Wang et al. 2018).

Sensitivity analysis and the optimization with variable screening

The variable sensitivity results are obtained by the interval-based sensitivity index C_i, which are shown in Fig. 8 (E). It should be noted that the interval-based sensitivity analysis is performed based on a Kriging metamodel, which is built by 500 calls of the aeronautical hydraulic pipeline model in total. The specific procedure of the metamodel construction can be found in (Liu et al. 2018). In the application, the optimisation problem was required, by engineering designers, to have no more than 12 design variables. According to Fig. 8 (E), ten most influential design variables can be selected, which are X_{24}, X_{14}, X_{13}, X_{12}, X_{19}, X_{15}, X_{25}, X_6, X_{27} and X_3, and other design variables have no or little effect on the maximum standard deviation of stress. Therefore, the 10 influential design variables are selected for the design optimization with variable screening (i.e., Top 10 optimization), which can be expressed as

$$\begin{aligned} &\text{Find } \ \boldsymbol{X}'_{\mathbf{Top}} = \{X'_3, X'_6, X'_{12}, X'_{13}, X'_{14}, X'_{15}, X'_{19}, X'_{24}, X'_{25}, X'_{27}\} \\ &\text{to minimize } \ \sigma_{\sigma\max}(\boldsymbol{X}_{\mathbf{Top}},\ \boldsymbol{X}_{\sim\mathbf{Top}}) \\ &\text{subject to } \ D_{\max} - D^t < 0 \\ &\qquad\qquad \boldsymbol{X}^l_{\mathbf{Top}} \leqslant \boldsymbol{X}_{\mathbf{Top}} \leqslant \boldsymbol{X}^u_{\mathbf{Top}} \\ &\qquad\qquad \boldsymbol{X}_{\sim\mathbf{Top}} = \boldsymbol{X}^{sv}_{\sim\mathbf{Top}}, \end{aligned} \tag{21}$$

where $\boldsymbol{X}^{sv}_{\sim\mathbf{Top}}$ is the selected specific value vector of the noninfluential variables $\boldsymbol{X}_{\sim\mathbf{Top}}$, $\sigma_{\sigma\max}$ and $D_{\max}$ are directly obtained by the finite element model of the aeronautical hydraulic pipeline. According to the interval-based sensitivity analysis, the noninfluential variables X_j ($j = 1, 2, 4, 5, 7\sim11,\ 16\sim18, 20\sim23, 26, 28$) are fixed as 34.84, 38.29, 45.51, 60.46, 16.03, 9.82, 10.72, 28.83, 12.10, 3.74, 2.63, 5.67, 37.45, 2.63, 5.68, 19.03, 1.88, 13.33, 1.23, 5.98, and 4.48, respectively. The optimization results of the traditional optimization method (i.e., all-variable optimization or the Top 28 optimization) and the optimization results with variable screening (i.e., the Top 10 optimization) are shown in Fig. 8 (F). The SQP optimizer is employed for the aeronautical hydraulic pipeline optimization. The optimized maximum standard deviation stress of the Top 28 is $2.42 \times 10^7(\text{Pa})$ (shown in the right contour results of Fig. 8 (F)), while the optimized maximum standard deviation stress of the Top 10 is $2.43 \times 10^7(\text{Pa})$ (shown in the left contour results of Fig. 8 (F)). It should be noted that the Top 10 optimization converges after 181 iterations, while the Top 28 optimization does not converge until the maximum allowable iterations (i.e., 550).

Discussions about the results

By comparing the optimization history of the Top 28 optimization and the Top 10 optimization (shown in Fig. 8 (F)), it is evident that the optimal result ($2.42 \times 10^7(\text{Pa})$) of the Top 28 optimization is almost equal to the optimal result ($2.43 \times 10^7(\text{Pa})$) of the Top 10 optimization, indicating that the eliminated 18 design variables have little effect on the maximum standard deviation of stress of the hydraulic pipeline. Therefore, the interval-based sensitivity index C_i is effective to identify these influential design variables and eliminate the remaining noninfluential design variables. The calculation cost for SQP optimizer has been reduced from 550 iterations to 181 iterations during the optimization, which is a decrease of 68%. It is worth noting that, although variable selection by sensitivity analysis incurs additional computational cost, engineers do not use the optimization model only once, but rather use it repeatedly throughout subsequent analyses of the product. Furthermore, for some complex engineering problems, the result of Top *m* optimization may be better than the result of the optimization with all variables (Wang et al. 2018). In addition, it should be noted that when the concerned model is a complex engineering problem without explicit performance function, such as the aeronautical hydraulic pipelines in this work, the calculation of sensitivity indices (including Sobol' indices, Wang's indices, and the interval-based sensitivity index) is time-consuming. In this case, the multi-fidelity surrogate model (Alemazkoor et al. 2022; Meng et al. 2021) may be a suitable choice. First, the sensitivity indices are obtained by a low-fidelity surrogate model. Then, the high-fidelity finite element model is used for the optimization.

As can be seen, in the high-dimensional optimization problem, eliminating those noninfluential design variables can greatly simplify the optimisation problem, and at the same time, relatively adequate results can also be obtained. Ultimately, engineers can use the simplified models (rather than the original high-dimensional complex models) in their subsequent work.

Conclusion

This paper introduces the interval-based sensitivity index and shows how the sensitivity analysis can be used for solving high-dimensional optimization problems. Based on the introduction of the interval field, the interval-based sensitivity index is reviewed and compared with other existing indices by an illustrative function. The screening procedure based on this sensitivity index is applied to high-dimensional design optimization problems involving two illustrative optimization examples and an aeronautical hydraulic pipeline. The results show that the interval-based sensitivity index can provide comprehensive effect information of input variables to the model output, thus effectively screening out influential variables to remodel the high-dimensional optimization problems and providing the guidance for the determination/realization of value-fixation of noninfluential variables.

Moreover, compared with Sobol' and Wang's indices, the interval-based sensitivity index is easy to use and has a superior performance.

It should be noted that the interval-based sensitivity analysis remains largely unexplored, and related studies are still rare. This paper explores a theoretical foundation for design optimization with variable screening by interval-based sensitivity analysis. It is admitted that sensitivity analysis is time-consuming, especially when the concerned models is a complex engineering problem. The interval-based sensitivity index is heuristic and more efficient computing strategies for evaluating this index will be further explored.

In this work, we focus on the single-objective optimization problem where we only need to evaluate the variable importance to a single model output. An important and practical perspective is to investigate the application of the proposed approach to multi-objective optimization problems by extending the interval sensitivity index. At the same time, the interval sensitivity index can also be extended to time-dependent reliability-based design optimization problems.

Appendix A: Extension of the interval sensitivity index to consider the joint effect of two or more input variables

The case of two variables (X_i and X_j) is derived as (Chang et al. 2022)

$$
\begin{aligned}
C_{ij} &= \frac{D_u(ij) + D_l(ij)}{D_{total}(ij)} \\
&= \frac{D_{total}(ij) - D_{\mathrm{NIP}}(ij)}{D_{total}(ij)} = 1 - \frac{D_{\mathrm{NIP}}(ij)}{D_{total}(ij)} \\
&= 1 - \frac{1}{4(Y^u - Y^l)} \int_{-1}^{1}\int_{-1}^{1} \left(g^u(\delta_i, \delta_j) - g^l(\delta_i, \delta_j)\right) \mathrm{d}\delta_i \mathrm{d}\delta_j \\
&= 1 - \frac{1}{2(Y^u - Y^l)} \int_{-1}^{1}\int_{-1}^{1} g^r(\delta_i, \delta_j) \mathrm{d}\delta_i \mathrm{d}\delta_j
\end{aligned}
\tag{A22}
$$

and the case of m variables $(X_i, X_j, \cdots, X_k)$ is derived by the following expression (Chang et al. 2022):

$$
\begin{aligned}
C_{\underbrace{ij\cdots k}_{m}} &= \frac{D_u(ij\cdots k)+D_l(ij\cdots k)}{D_{total}(ij\cdots k)} \\
&= \frac{D_{total}-D_{\mathrm{NIP}}(ij\cdots k)}{D_{total}(ij\cdots k)} = 1-\frac{D_{\mathrm{NIP}}(ij\cdots k)}{D_{total}(ij\cdots k)} \\
&= 1-\frac{1}{2^m(Y^u-Y^l)}\int_{-1}^{1}\cdots\int_{-1}^{1}\left(g^u(\delta_i,\delta_j,\cdots,\delta_k)-g^l(\delta_i,\delta_j,\cdots,\delta_k)\right)\mathrm{d}\delta_i\mathrm{d}\delta_j\cdots\mathrm{d}\delta_k \\
&= 1-\frac{1}{2^{m-1}(Y^u-Y^l)}\int_{-1}^{1}\cdots\int_{-1}^{1} g^r(\delta_i,\delta_j,\cdots,\delta_k)\mathrm{d}\delta_i\mathrm{d}\delta_j\cdots\mathrm{d}\delta_k,
\end{aligned}
\tag{A23}
$$

where $D(ij)$ indicates a two-dimensional domain and $D(ij\cdots k)$ indicates a *m*-dimensional domain. It should be noted that g indicates interval fields in Eqs. (A1) and (A2).

Appendix B: Review of Sobol' indices

For a model $Y=f(\boldsymbol{X})$ where all elements of $\boldsymbol{X}$ are independent, the variance decomposition of the model response Y can be expressed as (Homma and Saltelli 1996; Sobol' 1993; Sobol' 2001)

$$V_Y = \sum\nolimits_i V_i + \sum\nolimits_i \sum\nolimits_{j>i} V_{ij} + \cdots + V_{12\cdots n}, \tag{B1}$$

where $V_i = \mathrm{Var}\left[\mathrm{E}(Y|X_i)\right]$ and $V_{ij} = \mathrm{Var}\left[\mathrm{E}(Y|X_i, X_j)\right] - V_i - V_j$. Dividing both sides of Eq. (B1) by V_Y yields:

$$1 = \sum\nolimits_i S_i + \sum\nolimits_i \sum\nolimits_{j>i} S_{ij} + \cdots + S_{12\cdots n}, \tag{B2}$$

where $S_i = V_i / V_Y$ is the main effect index of X_i and $S_{ij} = V_{ij}/V_Y$ is the second order effect between X_i and X_j. Finally, the total effect index $S_{\mathrm{T}i}$ is defined as (Homma and Saltelli 1996)

$$S_{\mathrm{T}i} = S_i + \sum\nolimits_{i\neq j} S_{ij} + \cdots + S_{12\cdots n} = \frac{\mathrm{E}\left[\mathrm{Var}(Y|X_{-i})\right]}{V_Y} = \frac{V_{\mathrm{T}i}}{V_Y}. \tag{B3}$$

The main index S_i and total index $S_{\mathrm{T}i}$ are often used to measure the importance of each random input. $S_i > S_j$ indicates that X_i is more important than X_j in the sense of individual contribution to the model output variance. Similarly, $S_{\mathrm{T}i} > S_{\mathrm{T}i}$ indicates that X_i contributes more to the model output variance than X_j. It should be noted that the total effect index $S_{\mathrm{T}i}$ is usually used when screening the influential variables (Cho et al. 2014; Wei et al. 2015). A more detailed description can be found in (Homma and Saltelli 1996; Sobol' 1993; Sobol' 2001; Wei et al. 2015).

Appendix C: Review of Wang's indices

For a model $Y = f(\boldsymbol{X})$, the input variables $\boldsymbol{X}$ are intervals and the output Y is also an interval, i.e., $Y \in Y^I = [Y^l, Y^u]$ where Y^l and Y^u are the lower and upper bounds of Y^I. Its centre value and radius can be defined as $Y^c = (Y^u + Y^l)/2$ and $Y^r = (Y^u - Y^l)/2$, respectively. Then, the centre value Y^c and radius Y^r can be combined into the following form (Guo and Lu 2015; Guo and Lu 2001):

$$\eta = \frac{Y^c}{Y^r}. \tag{C1}$$

When X_i is fixed at a specific value in its interval, the impact of the uncertainty of X_i on η will be eliminated. If X_i takes different values in its interval, η will also be an interval that can be expressed as $\eta|X_i$. The lower and upper bounds of $\eta|X_i$ can be expressed as

$$[\eta|X_i]^l = \min_{X_i \in [X_i^l, X_i^u]} \eta|X_i \tag{C2}$$

and

$$[\eta|X_i]^u = \max_{X_i \in [X_i^l, X_i^u]} \eta|X_i. \tag{C3}$$

The centre value and radius of $\eta|X_i$ are given as

$$[\eta|X_i]^c = \frac{[\eta|X_i]^u + [\eta|X_i]^l}{2} \tag{C4}$$

and

$$[\eta|X_i]^r = \frac{[\eta|X_i]^u - [\eta|X_i]^l}{2}. \tag{C5}$$

Finally, two non-probabilistic sensitivity indices (centre value sensitivity ζ and radius sensitivity ς) can be expressed as (Li et al. 2013; Wang et al. 2018)

$$\zeta_i = \frac{|\eta - [\eta|X_i]^c|}{\eta} \tag{C6}$$

and

$$\varsigma_i = \frac{[\eta|X_i]^r}{\eta}. \tag{C7}$$

Wang's two sensitivity indices quantify the effect of the interval variables on η (which is combined by centre value Y^c and radius Y^r of model output Y) from the perspective of the centre value and radius, respectively. A detailed description of Wang's indices can be found in (Wang et al. 2018). In addition, it should be pointed out that Wang's two sensitivity indices described in Eqs.

(C2-C7) have the same form and definition as Li's importance measures (Li et al. 2013), although Li's measures are concerned with a non-probabilistic reliability index rather than system output response.

Data Availability Statement

All data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No. NSFC51975476). The first author is sponsored by Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University (Grant No. CX2022012).

References

Alemazkoor, N., A. Louhghalam, and M. Tootkaboni. 2022. "A multi-fidelity polynomial chaos-greedy Kaczmarz approach for resource-efficient uncertainty quantification on limited budget." *Computer Methods in Applied Mechanics and Engineering*, 389: 114290.

Arwade, S. R., M. Moradi, and A. Louhghalam. 2010. "Variance decomposition and global sensitivity for structural systems." *Engineering Structures*, 32 (1): 1–10.

Borgonovo, E. 2007. "A new uncertainty importance measure." *Reliability Engineering and System Safety*, 92 (6): 771–784.

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

Breiman, L. 2001. "Random forests." *Machine Learning*, 45 (1): 5–32.

Chang, Q., C. Zhou, P. Wei, Y. Zhang, and Z. Yue. 2021. "A new non-probabilistic time-dependent reliability model for mechanisms with interval uncertainties." *Reliability Engineering and System Safety*, 215: 107771.

Chang, Q., C. Zhou, M. A. Valdebenito, H. Liu, and Z. Yue. 2022. “A novel sensitivity index for analyzing the response of numerical models with interval inputs.” *Computer Methods in Applied Mechanics and Engineering*, 400: 115509.

Cho, H., S. Bae, K. K. Choi, D. Lamb, and R. J. Yang. 2014. “An efficient variable screening method for effective surrogate models for reliability-based design optimization.” *Structural and Multidisciplinary Optimization*, 50 (5): 717–738.

Constantine, P. G., E. Dow, and Q. Wang. 2013. “Active subspace methods in theory and practice: applications to kriging surfaces.” *SIAM Journal on Scientific Computing*, 36(4), 1500–1524.

Craig, K. J., N. Stander, D. A. Dooge, and S. Varadappa. 2005. “Automotive crashworthiness design using response surface-based variable screening and optimization.” *Engineering Computations*, 22 (1): 38–61.

Duarte Silva, A. P. 2001. “Efficient variable screening for multivariate analysis.” *Journal of Multivariate Analysis*, 76 (1): 35–62.

Faes, M., J. Cerneels, D. Vandepitte, and D. Moens. 2017. “Identification and quantification of multivariate interval uncertainty in finite element models.” *Computer Methods in Applied Mechanics and Engineering*, 315: 896–920.

Faes, M., and D. Moens. 2020. “Recent trends in the modeling and quantification of non-probabilistic uncertainty.” *Archives of Computational Methods in Engineering*, 27, 633–671.

Fesanghary, M., E. Damangir, and I. Soleimani. 2009. “Design optimization of shell and tube heat exchangers using global sensitivity analysis and harmony search algorithm.” *Applied Thermal Engineering*, 29 (5–6): 1026–1031.

Fort, J. C., T. Klein, and N. Rachdi. 2016. “New sensitivity analysis subordinated to a contrast.” *Communications in Statistics-Theory and Methods*, 45 (15): 4349–4364.

Fu, G., Z. Kapelan, and P. Reed. 2012. “Reducing the complexity of multiobjective water distribution system optimization through global sensitivity analysis.” *Journal of Water Resources Planning and Management*, 138 (3): 196–207.

Gao, B., Y. Ren, H. Jiang, and J. Xiang. 2019. “Sensitivity analysis-based variable screening and reliability optimisation for composite fuselage frame crashworthiness design.” *International Journal of Crashworthiness*, 24 (4): 380–388.

Guo, S. X., and Z. Z. Lu. 2015. “A non-probabilistic robust reliability method for analysis and design optimization of structures with uncertain-but-bounded parameters.” *Applied Mathematical Modelling*, 39(7), 1985–2002.

Guyon, I., and A. Elisseeff. 2003. “An introduction to variable and feature selection.” *Journal of Machine Learning Research*, 3: 1157–1182.

Härdle, W., and L. Simar. 2003. “Applied multivariate statistical analysis.” *Springer Nature*.

Homma, T., and A. Saltelli. 1996. “Importance measures in global sensitivity analysis of nonlinear models.” *Reliability Engineering and System Safety*, 52 (1): 1–17.

Jamian, J. J., M. N. Abdullah, H. Mokhlis, M. W. Mustafa, and A. H. A. Bakar. 2014. “Global particle swarm optimization for high dimension numerical functions analysis.” *Journal of Applied Mathematics*, 36(2): 329193

Jiang, C., B. Y. Ni, X. Han, and Y. R. Tao. 2014. “Non-probabilistic convex model process: A new method of time-variant uncertainty analysis and its application to structural dynamic reliability problems.” *Computer Methods in Applied Mechanics and Engineering*, 268: 656–676.

Jolliffe, I. T. 2002. “Principal component analysis.” *Journal of Marketing Research*, 87 (4): 513.

Li, G., Z. Lu, L. Tian, and J. Xu. 2013. “The importance measure on the non-probabilistic reliability index of uncertain structures.” *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 227 (6): 651–661.

Li, J., Q. Y. Duan, W. Gong, A. Ye, Y. Dai, C. Miao, Z. Di, C. Tong, and Y. Sun. 2013b. “Assessing parameter importance of the common land model based on qualitative and quantitative sensitivity analysis.” *Hydrology and Earth System Sciences*, 17 (8): 3279–3293.

Li, L., Q. Chang, C. Zhou, W. Wang, and Z. Zhang. 2020. “Sensitivity analysis-based optimization: a case study with the MTBF of an aeronautical hydraulic pipeline system.” *International Journal of*

Computational Methods, 17 (5): 1940019.

Liu, H., Y. S. Ong, and J. Cai. 2018. "A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design." *Structural and Multidisciplinary Optimization*, 57 (1): 393–416.

Liu, Q., X. Wu, X. Han, J. Liu, Z. Zhang, and S. Guo. 2020. "Sensitivity analysis and interval multi-objective optimization for an occupant restraint system considering craniocerebral injury." *Journal of Mechanical Design, Transactions of the ASME*, 142 (2): 024502.

Lu, Z., J. Liu, Y. Liu, R. Ding, and F. Yang. 2015. "The interval sensitivity analysis and optimization of the distribution network parameters considering the load uncertainty." *International Journal of Electrical Power and Energy Systems*, 64: 931–936.

Marra, G., and S. N. Wood. 2011. "Practical variable selection for generalized additive models." *Computational Statistics and Data Analysis*, 55 (7): 2372–2387.

Marrel, A., B. Iooss, F. Van Dorpe, and E. Volkova. 2008. "An efficient methodology for modeling complex computer codes with Gaussian processes." *Computational Statistics and Data Analysis*, 52 (10): 4731–4744.

Martins, J. R., and A. B. Lambe. 2013. "Multidisciplinary design optimization: a survey of architectures." *AIAA journal*, 51 (9): 2049–2075.

Meng, X., Z. Wang, D. Fan, M. S. Triantafyllou, and G. E. Karniadakis. 2021. "A fast multi-fidelity method with uncertainty quantification for complex data correlations: Application to vortex-induced vibrations of marine risers." *Computer Methods in Applied Mechanics and Engineering*, 386: 114212.

Ouyang, X., F. Gao, H. Yang, and H. Wang. 2012. "Modal analysis of the aircraft hydraulic-system pipeline." *Journal of aircraft*, 49 (4): 1168–1174.

Guo, S. X., and Z. Z. Lu. 2001. "Interval arithmetic and static interval finite element method." *Applied Mathematics and Mechanics*, 22 (12): 1390–1396.

Shu, R., J. Wei, D. Qin, T. C. Lim, and A. Zhang. 2018. "Global sensitivity analysis and dynamic optimization of multi-motor driving transmission system." *Structural and Multidisciplinary Optimization*, 58 (2): 797–816.

Sobol', I. M. 1993. "Sensitivity analysis for non-linear mathematical models." *Mathematical Modeling and Computational Experiment*, 1: 407-414.

Sobol', I. M., and S. Kucherenko. 2009. "Derivative based global sensitivity measures and their link with global sensitivity indices." *Mathematics and Computers in Simulation*, 79 (10): 3009–3017.

Sobol', I. M. 2001. "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates." *Mathematics and Computers in Simulation*, 55 (1): 271–280.

Spagnol, A., R. Le Riche, and S. Da Veiga. 2019. "Global sensitivity analysis for optimization with variable selection." *SIAM-ASA Journal on Uncertainty Quantification*, 7 (2): 417–443.

Tang, Z., Z. Lu, D. Li, and F. Zhang. 2011. "Optimal design of the positions of the hoops for a hydraulic pipelines system." *Nuclear Engineering and Design*, 241 (12): 4840–4855.

Tian, W., C. Jiang, B. Ni, Z. Wu, Q. Wang, and L. Yang. 2018. "Global sensitivity analysis and multi-objective optimization design of temperature field of sinter cooler based on energy value." *Applied Thermal Engineering*, 143: 759–766.

Valdebenito, M. A., and G. I. Schuëller. 2010. "A survey on approaches for reliability-based optimization." *Structural and Multidisciplinary Optimization*, 42 (5): 645–663.

Wang, L., D. Liu, Y. Yang, and J. Hu. 2019. "Novel methodology of non-probabilistic reliability-based topology optimization (NRBTO) for multi-material layout design via interval and convex mixed uncertainties." *Computer Methods in Applied Mechanics and Engineering*, 346: 550–573.

Wang, L., Y. Liu, D. Liu, and Z. Wu. 2021. "A novel dynamic reliability-based topology optimization (DRBTO) framework for continuum structures via interval-process collocation and the first-passage theories." *Computer Methods in Applied Mechanics and Engineering*, 386: 114107.

Wang, W., C. Zhou, H. Gao, and Z. Zhang. 2018. "Application of non-probabilistic sensitivity analysis in the optimization of aeronautical hydraulic pipelines." *Structural and Multidisciplinary Optimization*, 57 (6): 2177–2191.

Wang, Y., and C. A. Shoemaker. 2014. "A general stochastic algorithmic framework for minimizing expensive black box objective functions based on surrogate models and sensitivity analysis." *arXiv preprint*, arXiv:1410.6271.

Wei, P., Z. Lu, and J. Song. 2015a. "Variable importance analysis: a comprehensive review." *Reliability Engineering and System Safety*, 142: 399–432.

Wei, P., Z. Lu, and J. Song. 2015. "A comprehensive comparison of two variable importance analysis techniques in high dimensions: application to an environmental multi-indicators system." *Environmental Modelling and Software*, 70: 178–190.

Welch, W. J., R. J. Buck, J. Sacks, H. P. Wynn, T. J. Mitchell, and M. D. Morris. 1992. "Screening, predicting, and computer experiments." *Technometrics*, 34 (1): 15–25.

Wu, J., O. Sigmund, and J. P. Groen. 2021. "Topology optimization of multi-scale structures: a review." *Structural and Multidisciplinary Optimization*, 63 (3): 1455–1480.

Zhang, Y., X. Zhang, P. Huang, and Y. Sun. 2020. "Global sensitivity analysis for key parameters identification of net-zero energy buildings for grid interaction optimization." *Applied Energy*, 279: 115820.

Zhang, Z., C. Zhou, W. Wang, and Z. Yue. 2019. "Optimization design of aeronautical hydraulic pipeline system based on non-probabilistic sensitivity analysis." *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 233 (5): 815–825.

Zhou, C., Q. Chang, H. Zhao, M. Ji, and Z. Shi. 2021. "Fault tree analysis with interval uncertainty: a case study of the aircraft flap mechanism." *IEEE Transactions on Reliability*, 70 (3): 944–956.

Zhou, C., Z. Zhang, F. Liu, and W. Wang. 2019. "Sensitivity analysis for probabilistic anti-resonance design of aeronautical hydraulic pipelines." *Chinese Journal of Aeronautics*, 32 (4): 948–953.

Table 1 Two cases for the four-variable function with corresponding inputs and outputs

	$\boldsymbol{X}^l$	$\boldsymbol{X}^u$	$f^l(\boldsymbol{X})$	$f^u(\boldsymbol{X})$	$f^c(\boldsymbol{X})$	$f^r(\boldsymbol{X})$
Case 1	(5, 5, 5, 5,)	(6, 6, 6, 6)	2450	5112	3781	1331
Case 2	(-0.5, -0.5, -0.5, -0.5)	(0.5, 0.5, 0.5, 0.5)	-2.25	2.75	0.25	2.5

Algorithm 1 Pseudocode of the strategy to calculate the interval-based sensitivity index C_i

```
Generate initial points {δ_i(1), ⋯, δ_i(n_ini)}
Call surrogate optimisation algorithm
Obtain the corresponding lower/upper bounds {g^l(δ_i(1)), ⋯, g^l(δ_i(n_ini))}/{g^u(δ_i(1)), ⋯, g^u(δ_i(n_ini))}
Use the training data set to construct a metamodel of g^l(δ_i)/g^u(δ_i)
While n_ini + n_new < n_ε or max σ̂²(δ_i) > σ_ε²
    % n_ini + n_new is the total number of model calls
    % max σ̂²(δ_i) is the maximum of mean square error
    Use an adaptive sampling criterion to sample a new point δ_i(new)
    Call surrogate optimisation algorithm
    Obtain its corresponding lower/upper bound g^l(δ_i(new))/g^u(δ_i(new))
    Update training data set by involving δ_i(new) and g^l(δ_i(new))/g^u(δ_i(new)) respectively
End While
Export final metamodels of the lower/upper bound g^l(δ_i)/g^u(δ_i)
The final metamodels g^l(δ_i) and g^u(δ_i) are used to obtain A_NIP(i)
C_i = 1 − A_NIP(i)/A_total
```

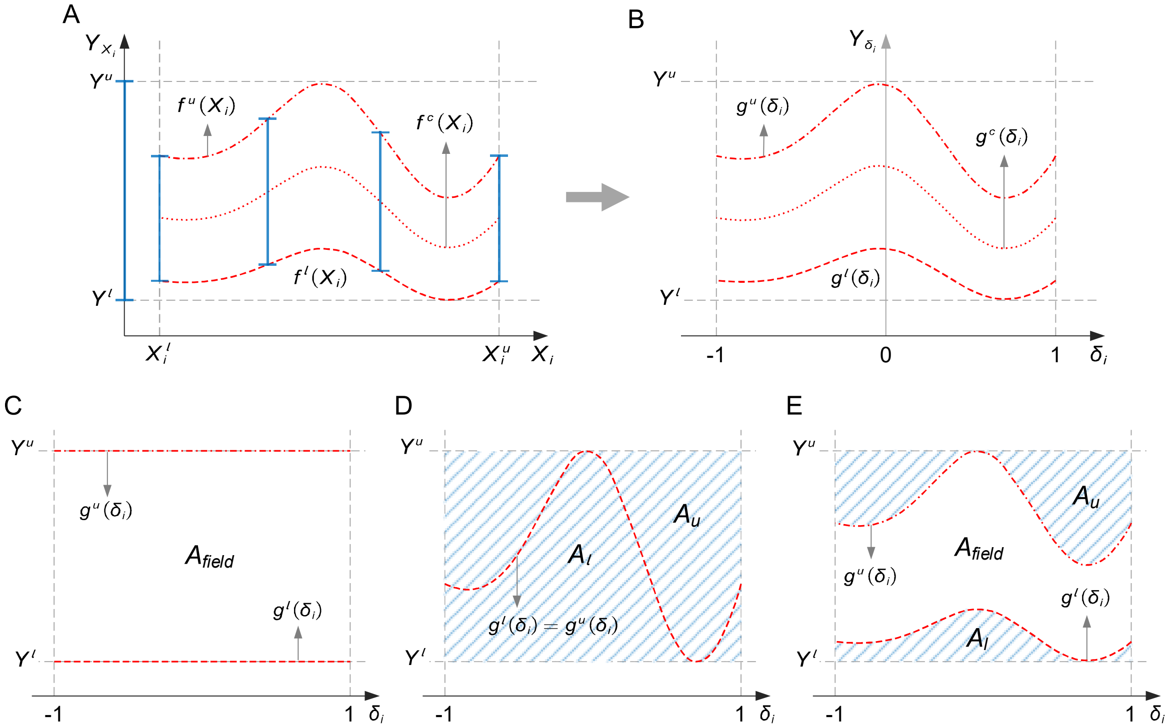
Table 2 Section of input variables and their initial value, lower bounds, and upper bounds.

Variables	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}
Position	14^X	17^X	20^X	23^X	26^X	30^X	33^X	36^X	39^X	40^X	42^X	45^X	45^Z	51^X
$\boldsymbol{X}^0$ (mm)	43	35	93	55	56	110	158	88	93	290	117	30	30	67
$\boldsymbol{X}^l$ (mm)	34	28	72	43	47	97	147	68	78	287	111	25	25	53
$\boldsymbol{X}^u$ (mm)	52	42	114	67	65	123	169	108	108	293	123	35	35	81

Variables	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	X_{26}	X_{27}	X_{28}
Position	54^X	59^X	62^X	62^Y	65^X	69^X	69^Y	72^X	72^Y	75^X	75^Y	82^X	82^Y	82^Z
$\boldsymbol{X}^0$ (mm)	63	36	29	59	56	187	20	111	13	78	7	54	28	50
$\boldsymbol{X}^l$ (mm)	53	29	26	48	50	183	18	88	12	61	5	43	25	40
$\boldsymbol{X}^u$ (mm)	73	43	32	70	62	191	22	134	14	95	9	65	31	60

* The superscript X, Y and Z of the position number are the coordinate of the hoops.

Fig 1



A
Y_{X_i}
Y^u
Y^l
$f^u(X_i)$
$f^c(X_i)$
$f^l(X_i)$
X_i^l
X_i^u
X_i
B
Y_{δ_i}
Y^u
Y^l
$g^u(\delta_i)$
$g^c(\delta_i)$
$g^l(\delta_i)$
-1
0
1
δ_i
C
Y^u
Y^l
$g^u(\delta_i)$
A_{field}
$g^l(\delta_i)$
-1
1
δ_i
D
Y^u
Y^l
A_u
A_l
$g^l(\delta_i) = g^u(\delta_i)$
-1
1
δ_i
E
Y^u
Y^l
A_u
A_{field}
$g^u(\delta_i)$
$g^l(\delta_i)$
A_l
-1
1
δ_i

Fig 2

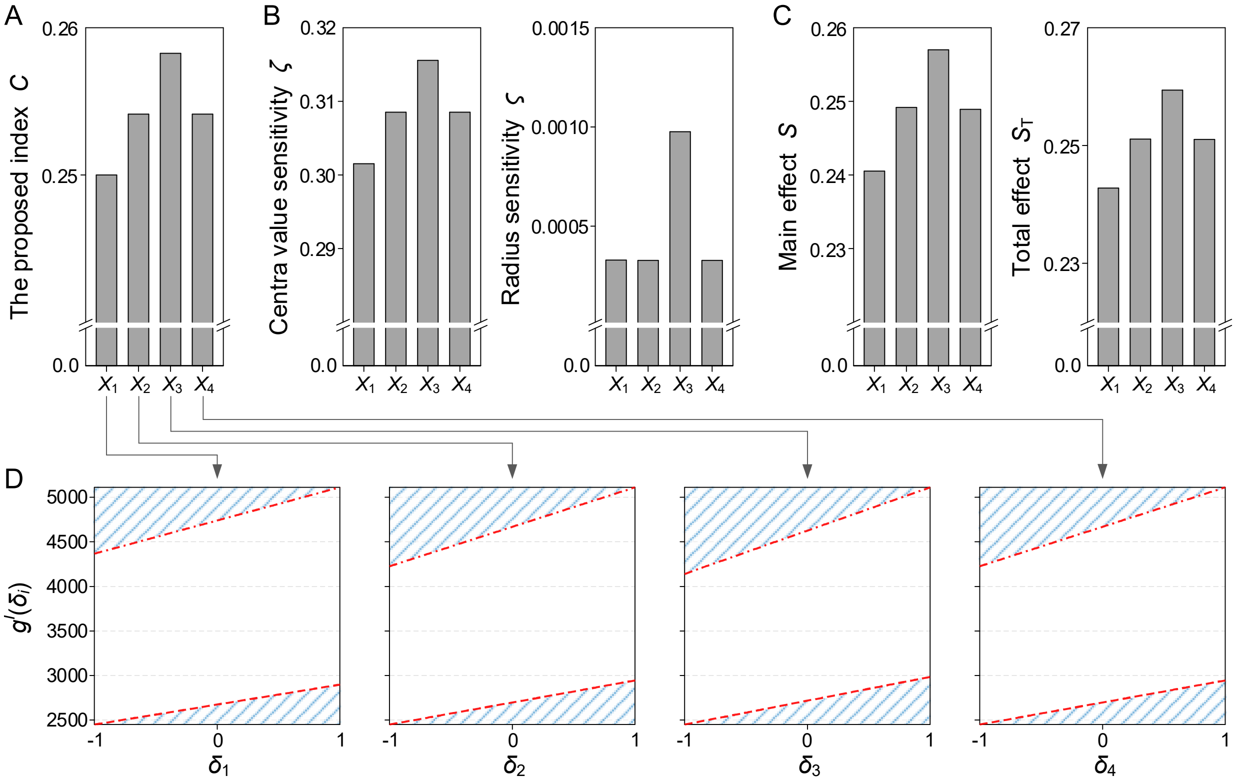
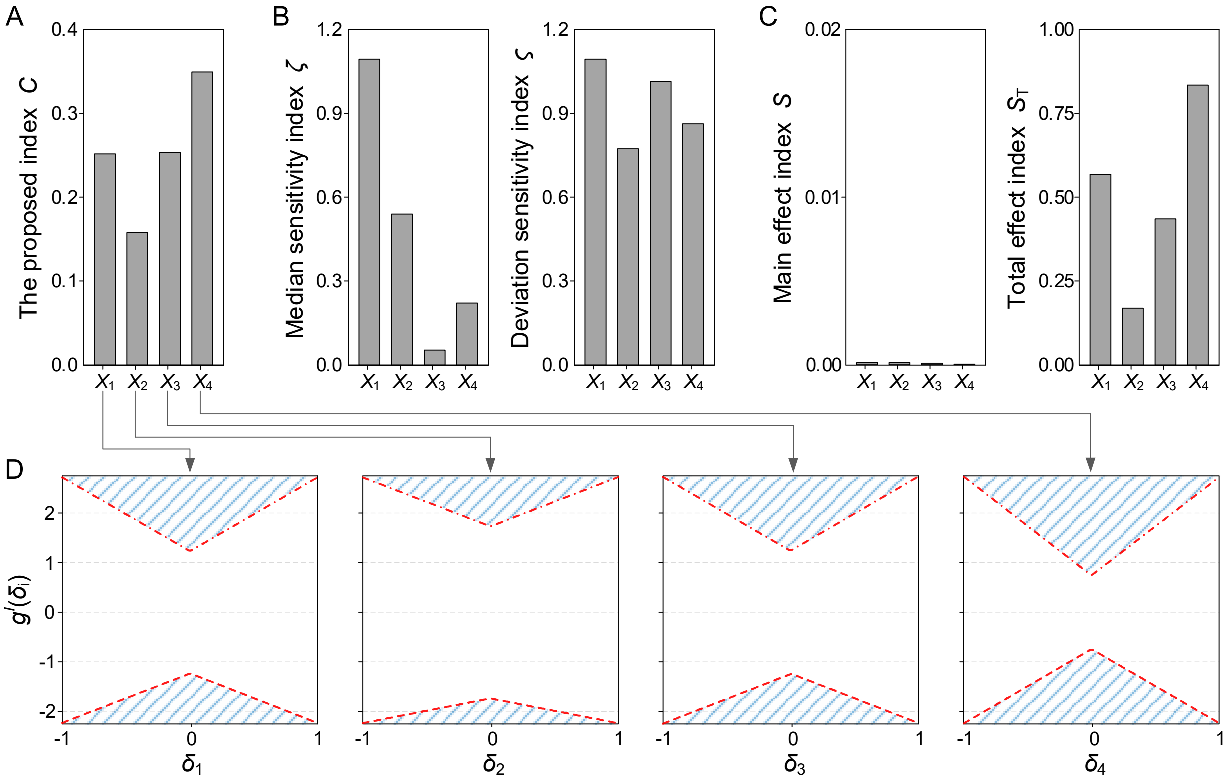

A
B
C
D
The proposed index C
Centra value sensitivity ζ
Radius sensitivity ς
Main effect S
Total effect S_T
0.26
0.25
0.0
0.32
0.31
0.30
0.29
0.0015
0.0010
0.0005
0.24
0.23
0.27
X_1 X_2 X_3 X_4
$g'(\delta_i)$
5000
4500
4000
3500
3000
2500
-1
0
1
δ_1
δ_2
δ_3
δ_4

Fig 3



A
B
C
D
The proposed index C
Median sensitivity index ζ
Deviation sensitivity index ς
Main effect index S
Total effect index S_T
X_1 X_2 X_3 X_4
$g^I(\delta_i)$
δ_1
δ_2
δ_3
δ_4

Fig 4

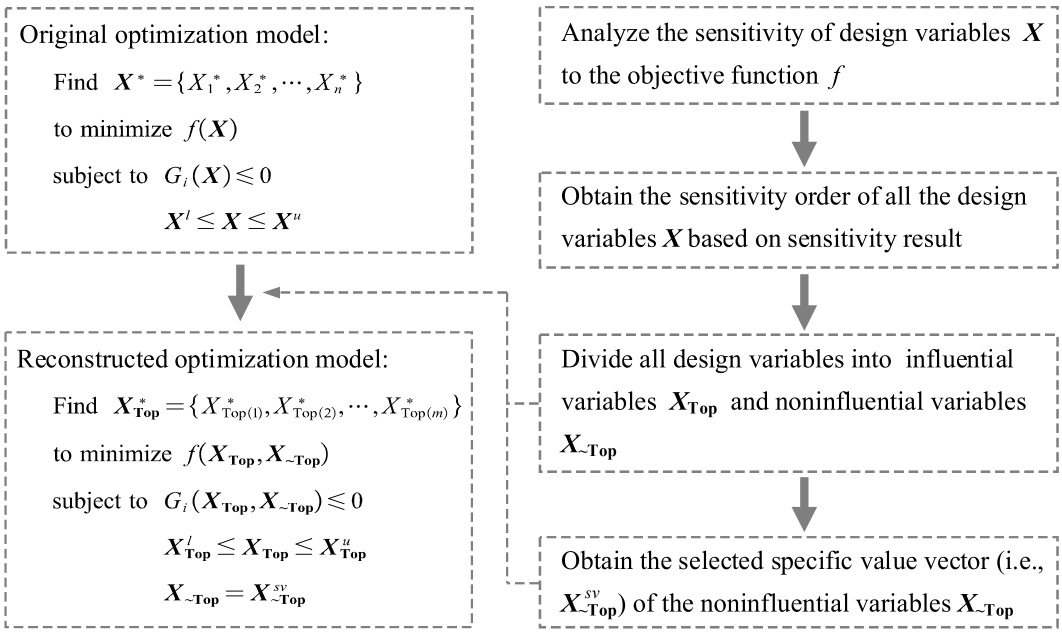

Original optimization model:
Find $\boldsymbol{X}^* = \{X_1^*, X_2^*, \cdots, X_n^*\}$
to minimize $f(\boldsymbol{X})$
subject to $G_i(\boldsymbol{X}) \leqslant 0$
$\boldsymbol{X}^l \leq \boldsymbol{X} \leq \boldsymbol{X}^u$
Reconstructed optimization model:
Find $\boldsymbol{X}^*_{\mathbf{Top}} = \{X^*_{\mathrm{Top}(1)}, X^*_{\mathrm{Top}(2)}, \cdots, X^*_{\mathrm{Top}(m)}\}$
to minimize $f(\boldsymbol{X}_{\mathbf{Top}}, \boldsymbol{X}_{\mathbf{\sim Top}})$
subject to $G_i(\boldsymbol{X}_{\mathbf{Top}}, \boldsymbol{X}_{\mathbf{\sim Top}}) \leqslant 0$
$\boldsymbol{X}^l_{\mathbf{Top}} \leq \boldsymbol{X}_{\mathbf{Top}} \leq \boldsymbol{X}^u_{\mathbf{Top}}$
$\boldsymbol{X}_{\mathbf{\sim Top}} = \boldsymbol{X}^{sv}_{\mathbf{\sim Top}}$
Analyze the sensitivity of design variables $\boldsymbol{X}$ to the objective function f
Obtain the sensitivity order of all the design variables $\boldsymbol{X}$ based on sensitivity result
Divide all design variables into influential variables $\boldsymbol{X}_{\mathbf{Top}}$ and noninfluential variables $\boldsymbol{X}_{\mathbf{\sim Top}}$
Obtain the selected specific value vector (i.e., $\boldsymbol{X}^{sv}_{\mathbf{\sim Top}}$) of the noninfluential variables $\boldsymbol{X}_{\mathbf{\sim Top}}$

Fig 5

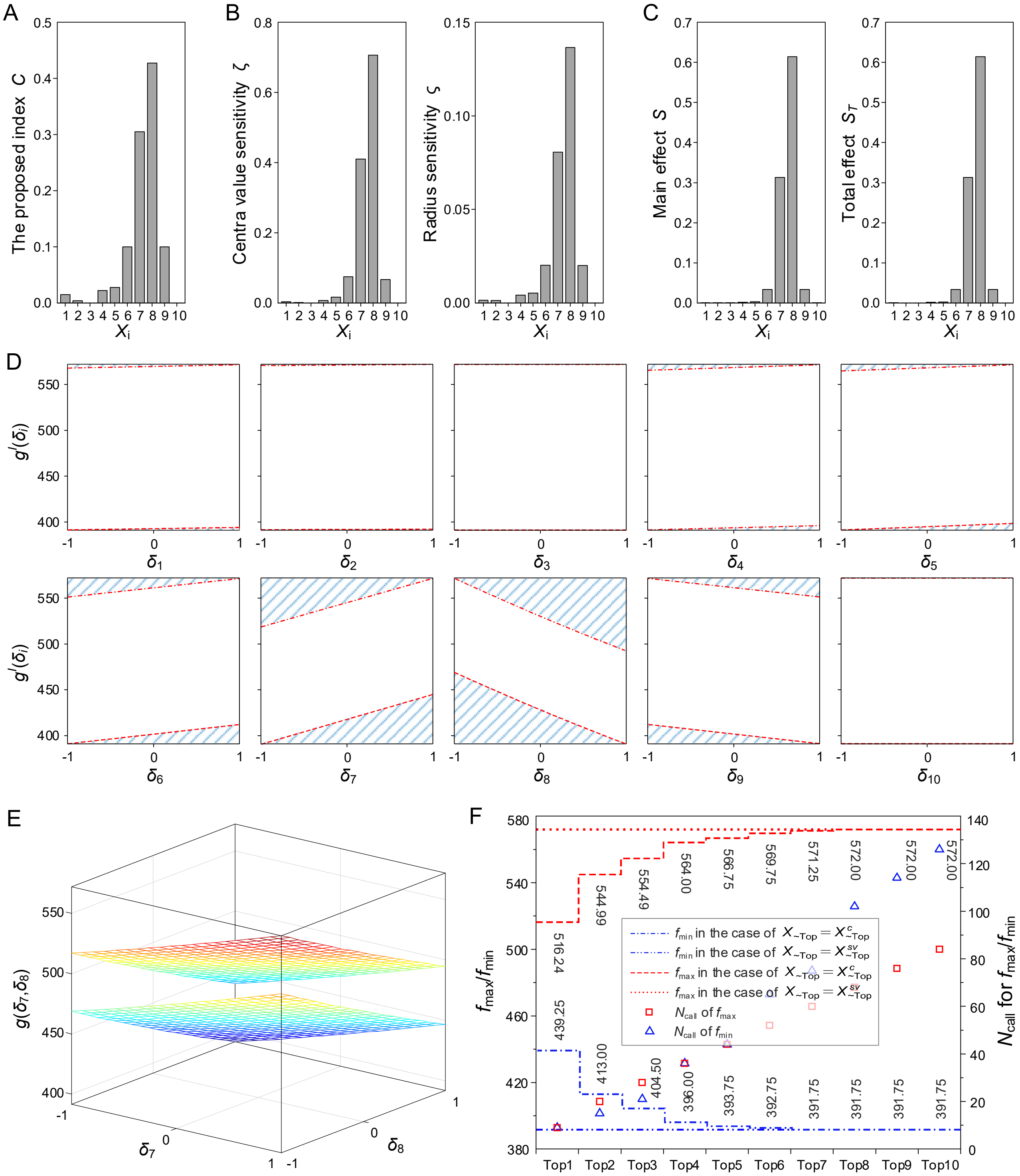

A
The proposed index C
B
Centra value sensitivity ζ
Radius sensitivity ς
C
Main effect S
Total effect S_T
X_i
D
$g'(\bar{\delta}_i)$
E
$g(\delta_7,\delta_8)$
F
f_{max}/f_{min}
N_{call} for f_{max}/f_{min}
f_{min} in the case of $X_{\sim Top} = X^{c}_{\sim Top}$
f_{min} in the case of $X_{\sim Top} = X^{sv}_{\sim Top}$
f_{max} in the case of $X_{\sim Top} = X^{c}_{\sim Top}$
f_{max} in the case of $X_{\sim Top} = X^{sv}_{\sim Top}$
N_{call} of f_{max}
N_{call} of f_{min}
Top1 Top2 Top3 Top4 Top5 Top6 Top7 Top8 Top9 Top10

Fig 6

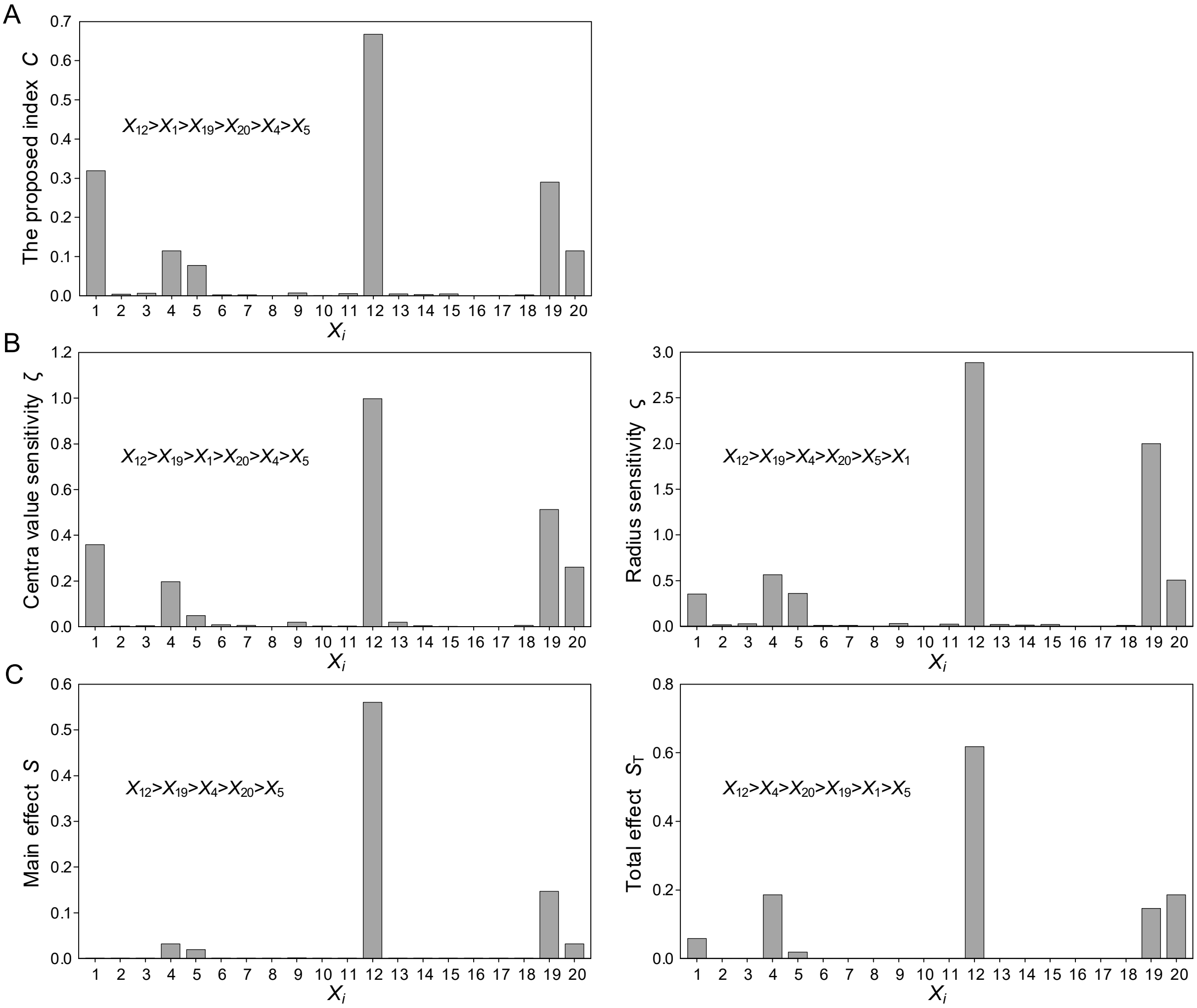
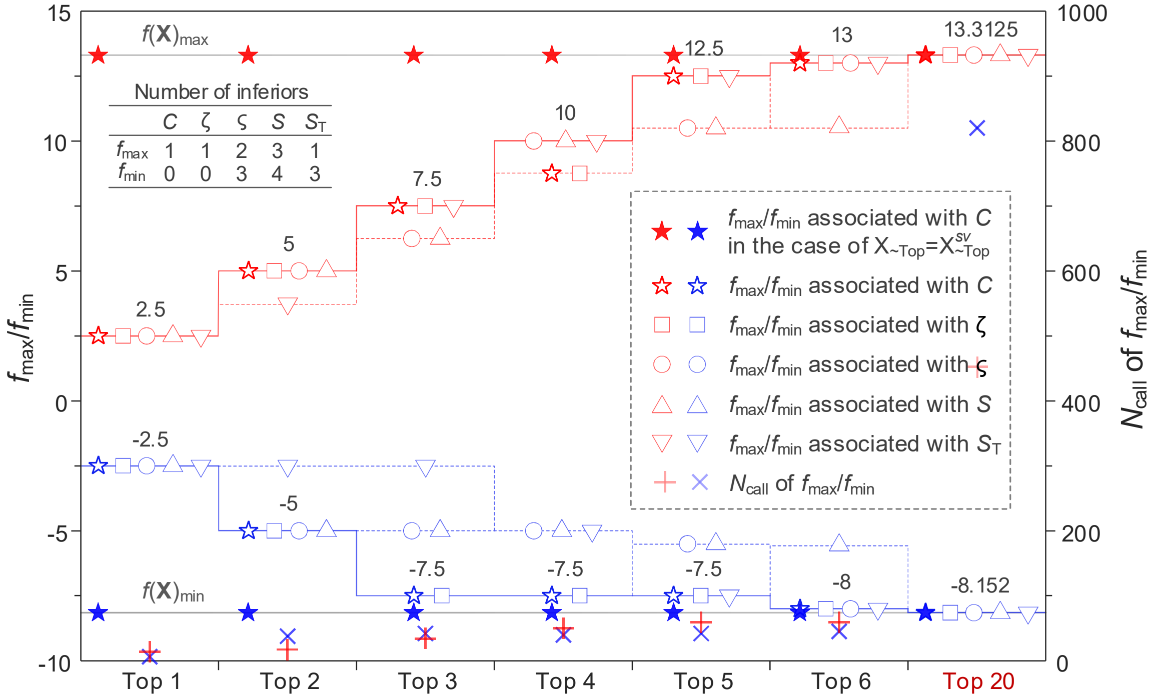
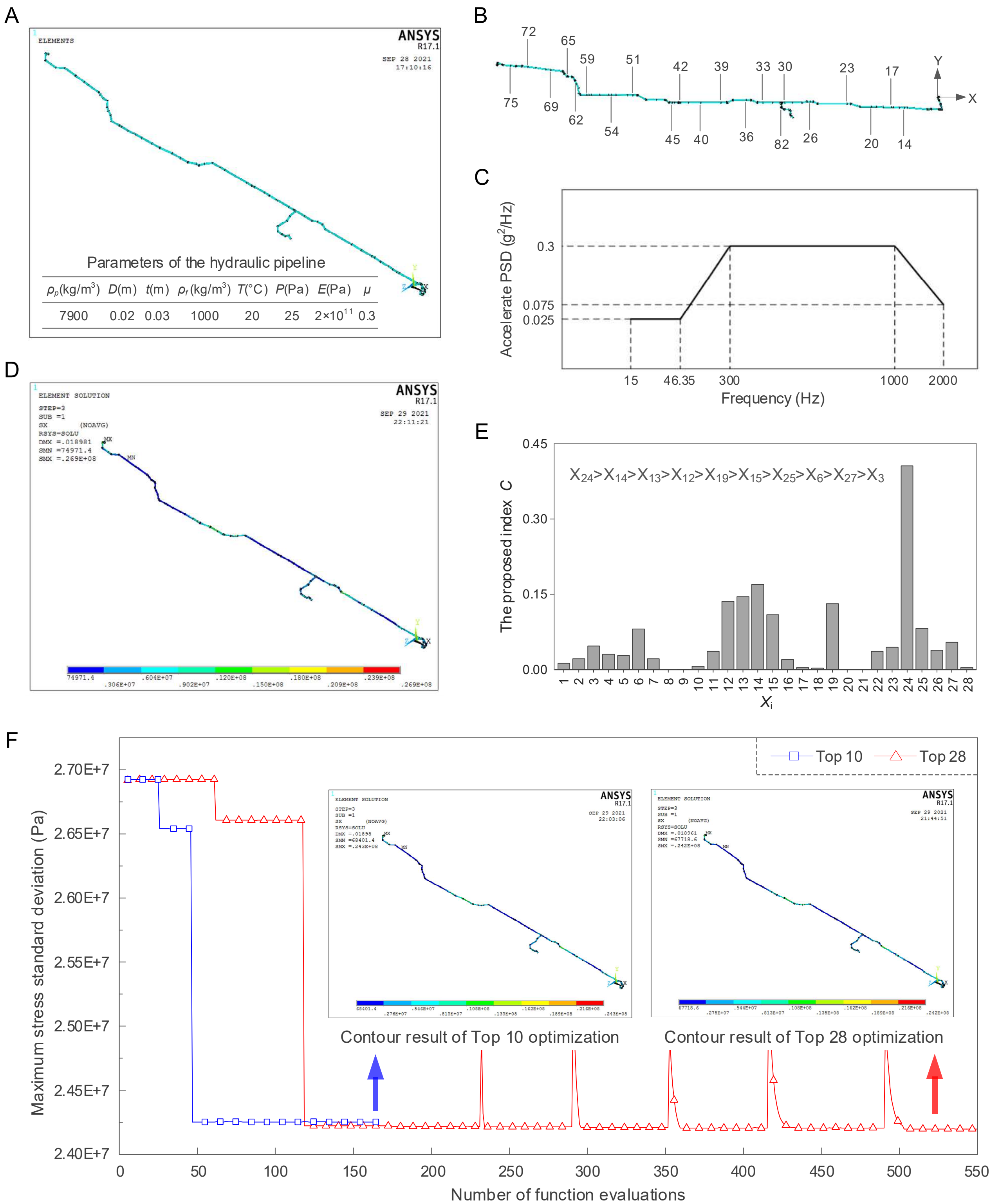

A
The proposed index C
X12>X1>X19>X20>X4>X5
Xi
B
Centra value sensitivity ζ
X12>X19>X1>X20>X4>X5
Xi
Radius sensitivity ς
X12>X19>X4>X20>X5>X1
Xi
C
Main effect S
X12>X19>X4>X20>X5
Xi
Total effect ST
X12>X4>X20>X19>X1>X5
Xi

Fig 7



$f(\mathbf{X})_{max}$
$f(\mathbf{X})_{min}$
Number of inferiors
C ζ ς S S_T
f_{max} 1 1 2 3 1
f_{min} 0 0 3 4 3
2.5
5
7.5
10
12.5
13
13.3125
-2.5
-5
-7.5
-7.5
-7.5
-8
-8.152
f_{max}/f_{min} associated with C in the case of $X_{\sim Top}=X^{sv}_{\sim Top}$
f_{max}/f_{min} associated with C
f_{max}/f_{min} associated with ζ
f_{max}/f_{min} associated with ς
f_{max}/f_{min} associated with S
f_{max}/f_{min} associated with S_T
N_{call} of f_{max}/f_{min}
f_{max}/f_{min}
N_{call} of f_{max}/f_{min}
Top 1
Top 2
Top 3
Top 4
Top 5
Top 6
Top 20
15
10
5
0
-5
-10
1000
800
600
400
200
0

Fig 8



A
ANSYS
R17.1
ELEMENTS
SEP 28 2021
17:10:16
Parameters of the hydraulic pipeline
ρp(kg/m³) D(m) t(m) ρf(kg/m³) T(°C) P(Pa) E(Pa) μ
7900 0.02 0.03 1000 20 25 2×10¹¹ 0.3
B
72 65 59 51 42 39 33 30 23 17
75 69 62 54 45 40 36 82 26 20 14
Y X
C
Accelerate PSD (g²/Hz)
0.3
0.075
0.025
15 46.35 300 1000 2000
Frequency (Hz)
D
ANSYS
R17.1
ELEMENT SOLUTION
STEP=3
SUB =1
SX (NOAVG)
RSYS=SOLU
DMX =.018981
SMN =74971.4
SMX =.269E+08
SEP 29 2021
22:11:21
74971.4 .306E+07 .604E+07 .902E+07 .120E+08 .150E+08 .180E+08 .209E+08 .239E+08 .269E+08
E
The proposed index C
0.45
0.30
0.15
0.00
X24>X14>X13>X12>X19>X15>X25>X6>X27>X3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
Xi
F
Maximum stress standard deviation (Pa)
2.70E+7
2.65E+7
2.60E+7
2.55E+7
2.50E+7
2.45E+7
2.40E+7
0 50 100 150 200 250 300 350 400 450 500 550
Number of function evaluations
Top 10
Top 28
Contour result of Top 10 optimization
Contour result of Top 28 optimization

Fig. 1 Illustration of the interval field for variables: (A) is an interval field $f^I(X_i)$; (B) is a normalized interval field $g^I(\delta_i)$; (C) and (D) are two special interval field cases where the interval-based sensitivity index $C_i = 0$ and $C_i = 1$, respectively; (E) is a normalized interval field used to define the interval-based sensitivity index C_i.

Fig. 2 The sensitivity results in Case 1: (A) the interval-based sensitivity index C_i; (B) Wang's indices; (C) Sobol' indices; (D) normalized interval fields to obtain C_1-C_4.

Fig. 3 The sensitivity results in Case 2: (A) the interval-based sensitivity index C_i; (B) Wang's indices; (C) Sobol' indices; (D) normalized interval fields to obtain C_1-C_4.

Fig. 4 The general process of the design optimization with variable screening

Fig. 5 The sensitivity results and optimization results: (A) the interval-based sensitivity index C_i; (B) Wang's indices; (C) Sobol' indices; (D) normalized interval field to obtain C_i; (E) normalized interval field to obtain C_{78}; (F) the maximum and minimum value (with its corresponding function calls) in the case of Top 1 to Top 10.

Fig. 6 The sensitivity results where (A) the interval-based sensitivity index C_i, (B) Wang's indices and (C) Sobol' indices

Fig. 7 The maximum and minimum values (with its corresponding function calls) obtained by three types of sensitivity indices (five indices in total) in the case of Top 1-6 and Top 20.

Fig. 8 Results of the aeronautical hydraulic pipeline where (A) shows the finite element model and its parameters; (B) node indices of hoops; (C) acceleration PSD function of stochastic excitation; (D) contour result of stress standard deviation before optimization; (E) the results of the interval-based sensitivity index; (F) the optimization history of maximum standard deviation of stress.