Surrogate-based sensitivity analysis for safety assessment of general aviation heavy-fueled engines

Jiaokun CAO*, Shuiting DING, Farong Du
Airworthiness Technologies Research Center, Beihang University, Beijing, 100191, P.R.China

Abstract

This paper has developed a framework for surrogate-based sensitivity analysis in SSA considering the complexity of Heavy-Fueled Engines (HFE) models. Multi-dimensional HFE whole engine models are extreme time consuming and computationally expensive and not practical for engineering applications. To reduce the computational costs, Artificial Neural Networks (ANN) is selected as surrogate model to establish high-fidelity HFE system model. Based on the developed ANN model, global sensitivity analysis is conducted to provide information on parameters importance, which are significant in complex system safety analysis. The results reveal that, by using global sensitivity analysis, the parameters could be ranked with respect to their importance, including first order indices and total sensitivity indices. For the particular study, the importance indices indicate that compression ratio and start angle of injection are more important with respect to the influence on the maximum pressure for HFE. The results also show that ANN-based surrogate model is an efficient way for sensitivity analysis in HFE safety assessment.

1. Introduction

General aviation piston engines have relied nearly solely on gasoline since the World War II. However, as gasoline costs, taxation, safety, and availability apply increasing pressure to the market, the Heavy-Fueled Engines (HFE), which run on diesel, kerosene or other heavy fuels, are gaining popularity. Meanwhile, HFE are imposing new challenges on safety design due to the new features not covered by...
existing regulations\cite{1}. In order to secure the aircraft safety and satisfy the airworthiness requirements, safety assessment progress and system development process should interact with each other, as suggested by System Development Process Model in Aerospace Recommended Practice (ARP) 4754\cite{2}. System Safety Assessment (SSA) verifies whether the safety requirements are met in the implemented design according to ARP 4761\cite{3}.

In the process of SSA, sensitivity analysis is an essential tool. First, at design stage, it is desired to identify the most important factors in order to get supplementary information by questioning the manufacturer on a limited number of elements or by enhancing some aspects of analysis model\cite{4}. And secondly, if the safety requirements were not met, design should be changed until it satisfies the requirements. Hence, it is typically useful to figure out which factors are the most important. Finally, uncertainties may arise from manufacture, operation environment, maintenance, etc. After the safety assessment, some special measures should be taken to ensure that the actual safety level of the most important factors is consistent with those used in SSA, if necessary.

The methodology of integrating sensitivity analysis into SSA has been proposed by literature \cite{5} based on global sensitivity analysis methods. Literature \cite{5} adopted a quasi-dimensional thermodynamics model to represent the engine, while accurate, high-fidelity models are desired in practice. Typically high-fidelity models are time consuming and computationally expensive and poses a serious impediment to the sensitivity analysis. As a result, compact and cheap surrogate models are used to reduce the computation costs. It is worth noting that high-fidelity models still needed during the processes of constructing surrogate models. Indeed, several examples of surrogate-based sensitivity analysis have been provided in the literatures \cite{6, 7}. For these examples, multi-dimensional computationally expensive models are executed due to their small computing scale. However, considering the multi-scale, multi-physics coupling physical properties associated with HFE, even for constructing processes of surrogate models, multi-dimensional HFE whole engine models are still extreme time consuming and computationally expensive and not practical for engineering applications. Hence, reducing the computational costs of high-fidelity models are necessary for HFE sensitivity analysis. In this paper, the whole system is divided into several modules and these modules are modeled in different dimensions to reduce the computational costs. Then the modules are coupled together by using module interfaces which transfer data between different dimensional models. Thus, a high-fidelity HFE system model is established with relatively affordable computational costs.

As for the surrogate model, Artificial Neural Networks (ANN) is selected for its wide use in engine modeling. ANN consists of many interconnected processing elements typically referred to as artificial neurons or nodes. These neurons are able to perform massive parallel computations to establish correlations between input and output data sets\cite{8}. ANN is capable of solving any non-linear problems by acquiring information and recognizing the relationship between input parameters and output responses.

In this study, firstly, a framework for surrogate-based sensitivity analysis in SSA is provided considering the complexity of HFE models. Then, as a specific example, a HFE model constituted by quasi-dimensional thermodynamics modules and multi-dimensional Computational Fluid Dynamics (CFD) modules is established. Finally, a surrogate model is constructed by using ANN and sensitivity analysis is conducted based on the ANN model.

2. Methods

The framework developed in this paper is shown in Fig. 1. First, a deterministic engine model is needed which should include parameters of interest and design outputs to be compared with the defined safety criteria. Second, samples are generated for the factors used in engine model according to factors distribution definition. Last but not the least, safety assessment and sensitivity analysis are conducted
using the data provided by the engine model. If the SSA results meet the safety requirements, the design steps into next stage. Otherwise, design should be improved until the requirements are satisfied. As discussed above, the sensitivity analysis provides helpful information for the design improvements.

A 50 kW two-stroke lightweight heavy-fuel reciprocating engine is considered in this paper. As one of the safety criteria, cylinder pressure should be confirmed to satisfy the power requirements but not to exceed the safety boundary. Then, an engine processes model is necessary. In this whole engine model, in-cylinder process is modeled with three-dimensional CFD methods, including gas flow patterns, fuel spray behavior and combustion performance, while intake and exhaust flows are represented by one-dimensional gas dynamic models to reduce computational costs. During the calculations, the data exchanges between quasi-dimensional and multi-dimensional modules are accomplished by module interfaces. In this case, flows at the connections are assumed to be fully developed. In other words, the patterns of field variables at the interfaces are predefined reasonably. Hence, data from quasi-dimensional modules could be extended to be multi-dimensional, while multi-dimensional data could be compressed to be quasi-dimensional.

![Fig. 1 A framework for surrogate-based sensitivity analysis in SSA](image)

In this study, the variables of interest include rotation speed ($R_s$), compression ratio ($C_r$), start angle of fuel injection ($A_{inj}$), intake temperature ($T_{in}$) and intake pressure ($p_{in}$) and the maximum in-cylinder pressure was chose as the output, considering that the maximum in-cylinder pressure in HFE is much larger than that in traditional gasoline engines. A Back Propagation (BP) ANN model connecting these input variables with the maximum in-cylinder pressure is constructed through the above high-fidelity engine system model. Back-propagation neural networks with Levenberg–Marquardt (LM) as the learning algorithm have successfully been employed to model engine responses$^{[9-13]}$. For these applications, ANN generally includes three types of layers which are input layer, hidden layer and output layer. The ANN configuration adopted in this study is illustrated in Fig. 2. Inputs are transferred to the hidden layer, where activation functions model the complex engine behavior.

![Fig. 2. ANN configuration in this study](image)
Although there is no one network configuration that could be applied universally, in engine modeling applications, typically, the suitable number of neurons for the hidden layer ranges from 10 to 25, and the transfer functions for hidden layer are log-sigmoid and output layer are linear. 340 data points were collected for the ANN surrogate modeling. The ranges of data points were summarized in Table 1. The input and output points were first normalized according to the ranges of the transfer functions. The normalization could be expressed as

\[ X_i = a_{i1} + \left( a_{i2} - a_{i1} \right) \frac{Z_i - Z_{\text{min},i}}{Z_{\text{max},i} - Z_{\text{min},i}}, \]  

(1)

where \( Z_i \) is the original input or output, and \((a_{i1}, a_{i2})\) is selected based on the transfer function. The normalized data points were divided into three parts with the ratio of 70:15:15, including training data, validation data and testing data. These data was used to design ANN through supervised learning approach. A forward pass of ANN generates errors between desired and actual outputs for training data, and weight and bias changes were conducted according to the error and learning rules. The learning process runs iteratively in order to reduce the errors of the training data. However, over-fitting could be encountered during the learning process, where ANN could only produce accurate prediction for the known data but could not give good prediction for new data. This problem is partially addressed by using early-stopping when the errors for the validation data could not decrease.

Table 1 Ranges of engine model parameters

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>Symbols</th>
<th>Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation speed (rev/min)</td>
<td>( R_s )</td>
<td>2300~2600</td>
</tr>
<tr>
<td>Compression ratio</td>
<td>( C_r )</td>
<td>16~18</td>
</tr>
<tr>
<td>Start angle of injection (°)</td>
<td>( A_{\text{inj}} )</td>
<td>-40--20</td>
</tr>
<tr>
<td>Intake temperature (K)</td>
<td>( T_{\text{in}} )</td>
<td>273~313</td>
</tr>
<tr>
<td>Intake pressure (10^5 Pa)</td>
<td>( p_{\text{in}} )</td>
<td>1.3~1.8</td>
</tr>
</tbody>
</table>

Sensitivity measures based on variance decomposition give useful information for parameter priority\textsuperscript{[14]}. These model parameters are modeled by random variables, whose probability distributions are chosen according to engineering experience from former engine models and expert judgments. For this study, Table 2 summarizes the random variables and their distributions. These variables are supposed to be independent.

Assuming that \( Y=f(X) \) is the output of the model and \( X \) is the random vector of model input parameters \( X=(x_1, x_2, \ldots, x_k) \). Then, the first order sensitivity indices are defined as

\[ S_i = \frac{V(E(Y | x_i))}{V_Y}. \]  

(2)

For a system with \( k \) factors there may be interaction terms up to the order \( k \textsuperscript{[15]} \):

\[ \sum_i S_i + \sum_{i<j} S_{ij} + \sum_{i<j<k} S_{ijk} + \cdots + S_{12\ldots k} = 1. \]  

(3)

By substituting \( x_i \) with \( x_{\ldots i} \) in the equation (2), we obtain \( V_x(E(Y | x_{\ldots i})) / V_Y \). By definition this is the first order effect of \( x_{\ldots i} \), which can be demonstrated to equal the sum of all terms in equation (3) that do not
include \( x_i \). Hence \( S_{Ti} = 1 - \frac{V_i(E_i(Y \mid x_i))}{V_Y} \) equals the sum of all terms that do include \( x_i \). The total sensitivity indices are expressed as

\[
S_{Ti} = 1 - \frac{V_i(E_i(Y \mid x_i))}{V_Y} = \frac{E_i(V_i(Y \mid x_i))}{V_Y}
\]  

(4)

where \( V(\bullet) \) and \( E(\bullet) \) denote the variance and the expectation operators, respectively, and \( x_j \) means all variables but \( x_i \).

The parameters associated to the highest values of first order indices are supposed to be the most important factors in terms of their effects on the system output. The total indices give information on the presence of interaction in the model. A significant difference between first order and total indices indicates that interactions with other parameters are very important.

Sobol’s sequence was employed to determine the sensitivity indices by considering the factors separately. Specifically, in this study, formulas proposed by [17] were used.

Table 2 Distributions of engine model parameters for global sensitivity analysis

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>Symbols</th>
<th>Distributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation speed (rev/min)</td>
<td>( R_s )</td>
<td>Uniform, 2300, 2600</td>
</tr>
<tr>
<td>Compression ratio</td>
<td>( C_r )</td>
<td>Normal, 17, 1.83</td>
</tr>
<tr>
<td>Start angle of injection (°)</td>
<td>( A_{inj} )</td>
<td>Uniform, -40, 20</td>
</tr>
<tr>
<td>Intake temperature (°K)</td>
<td>( T_{in} )</td>
<td>Uniform, 273, 313</td>
</tr>
<tr>
<td>Intake pressure (10^5Pa)</td>
<td>( p_{in} )</td>
<td>Uniform, 1.3, 1.8</td>
</tr>
</tbody>
</table>

3. Results and discussions

3.1. ANN modeling

Tolerance value and R-value were adopted in this study to quantitatively evaluate the prediction capability of the ANN. Here, tolerance value is the relative error between predicted and actual data and R-value is the correlation coefficient obtained by performing linear regression between predicted and actual data. Smaller value of tolerance indicates higher accuracy of the predicted data while R-value closer to unity signifies higher accuracy of the predicted data.

Specifically, 10 data points were randomly selected from the test data set to illustrate the prediction capability of the ANN, and these data points are summarized in Table 3. For the selected 10 random data points, their tolerance values and R-values are shown in Fig. 3 and Fig. 4, respectively. The tolerance values are below 0.005 and the R-value is 0.99999. Hence, the tolerance values and the R-value indicate that the developed ANN was able to successfully predict the engine model output, i.e., maximum in-cylinder pressure. Then the developed ANN could be utilized in the following sensitivity analysis.

Table 3 10 random data points from test data set
<table>
<thead>
<tr>
<th>No.</th>
<th>$R_s$(rev/min)</th>
<th>$C_r$</th>
<th>$A_{in}$(°)</th>
<th>$T_{in}$(K)</th>
<th>$p_{in}(10^5$Pa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2405</td>
<td>17.5</td>
<td>-25.7</td>
<td>288.05</td>
<td>1.42</td>
</tr>
<tr>
<td>2</td>
<td>2549</td>
<td>17.3</td>
<td>-21.9</td>
<td>280.64</td>
<td>1.76</td>
</tr>
<tr>
<td>3</td>
<td>2476</td>
<td>16.9</td>
<td>-22.2</td>
<td>290.13</td>
<td>1.43</td>
</tr>
<tr>
<td>4</td>
<td>2465</td>
<td>17.1</td>
<td>-33.3</td>
<td>292.28</td>
<td>1.68</td>
</tr>
<tr>
<td>5</td>
<td>2575</td>
<td>16.6</td>
<td>-26.0</td>
<td>277.82</td>
<td>1.39</td>
</tr>
<tr>
<td>6</td>
<td>2386</td>
<td>17.5</td>
<td>-36.0</td>
<td>296.58</td>
<td>1.44</td>
</tr>
<tr>
<td>7</td>
<td>2527</td>
<td>16.4</td>
<td>-39.4</td>
<td>282.05</td>
<td>1.35</td>
</tr>
<tr>
<td>8</td>
<td>2526</td>
<td>17.4</td>
<td>-25.1</td>
<td>288.38</td>
<td>1.59</td>
</tr>
<tr>
<td>9</td>
<td>2414</td>
<td>16.4</td>
<td>-30.0</td>
<td>296.32</td>
<td>1.64</td>
</tr>
<tr>
<td>10</td>
<td>2470</td>
<td>16.7</td>
<td>-30.4</td>
<td>283.07</td>
<td>1.57</td>
</tr>
</tbody>
</table>

Fig. 3. Tolerance values for 10 test data points

Fig. 4. Linear regression and R-value between predicted and actual data for 10 test data points
3.2. Sensitivity analysis

For this study, $5 \times 10^4$ independent simulations have been performed for each of 5 factors for a total of $3.5 \times 10^5$ model runs.

Fig. 5 and Fig. 6 show the convergence of the mean and variance of normalized maximum pressure, respectively. The maximum pressure is normalized by the mean of maximum pressure. As shown in Fig. 5 and Fig. 6, both of the curves become steady at the end of the simulations.

![Fig. 5. Convergence of the mean for normalized maximum pressure.](image1)

![Fig. 6. Convergence of the variance for normalized maximum pressure.](image2)

Convergence of the first order indices and total indices is displayed in Fig. 7 and Fig. 8 respectively. Table 4 shows the final results of the first order indices and the total indices.

By comparing the convergence histories of the first order indices and the total indices, it is found that the convergence of the total indices is more stable than that of the first order indices. And for the convergence of the first order indices, parameters with higher value have more stable convergence histories than those with lower value.

As the first order indices show, the most important parameters are start angle of injection and compression ratio, followed by intake pressure. The results indicate that start angle of injection and compression ratio should be dedicatedly designed and strictly controlled. It could also be concluded that rotation speed has the least effects on the maximum in-cylinder pressure compared to the other parameters considered here.
The total indices give information on the presence of interaction in the model. The most important factor is compression ratio instead of start angle of injection, which is different from first order indices ranking. This means that compression ratio depends on other parameters more, compared with start angle of injection. However, the ranking is changed due to slight difference. In general, similar to the first order indices, start angle of injection and compression ratio have the largest total indices value.

The results provide information about the importance of the variables of interest. For the particular study, the importance indices indicate that compression ratio and start angle of injection are more important with respect to the influence on the maximum pressure. More efforts should be put on these two variables in the following design and manufacturing procedures or design improvements, if necessary. This conclusion agrees with the general engineering judgment.

Table 4 Importance ranking according to first order indices and total indices

<table>
<thead>
<tr>
<th>Symbols</th>
<th>( S_i )</th>
<th>( S_{Ti} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_s )</td>
<td>0.0009</td>
<td>0.0007</td>
</tr>
<tr>
<td>( C_r )</td>
<td>0.4304</td>
<td>0.4556</td>
</tr>
<tr>
<td>( A_{inj} )</td>
<td>0.4407</td>
<td>0.4492</td>
</tr>
<tr>
<td>( T_{in} )</td>
<td>0.0068</td>
<td>0.0076</td>
</tr>
<tr>
<td>( P_{in} )</td>
<td>0.0911</td>
<td>0.0976</td>
</tr>
</tbody>
</table>
4. Conclusions

This paper has developed a framework for surrogate-based sensitivity analysis in SSA considering the complexity of HFE models. Under the framework, a particular study was applied on general aviation HFE.

The results provide information about the importance of the variables of interest. For the particular study, the importance indices indicate that compression ratio and start angle of injection are more important with respect to the influence on the maximum pressure for HFE.

The results reveal that the framework and corresponding methods are efficient. The results show that ANN-based surrogate model is an efficient way for sensitivity analysis in HFE safety assessment, and the modeling methods provided in this paper could offer high-fidelity data in the process of constructing the surrogate model with relatively low computational costs.

References

