Multi-Response Weighted Adaptive Sampling Approach Based on Hybrid Surrogate Model

JIANGQI LONG, YAOQING LIAO, AND PING YU
School of Mechanical and Electrical Engineering, Wenzhou University, Wenzhou 325035, China
Corresponding author: Ping Yu (yuping55@zju.edu.cn)

ABSTRACT In order to improve the fitting accuracy and optimization efficiency of the surrogate model, a multi-response weighted adaptive sampling (MWAS) approach based on the hybrid surrogate model was proposed and implemented to a multi-objective lightweight design of car seats. In this approach, the sample discreteness index in the input design space was calculated by the maximum and minimum distance approach (MDA), the fitting uncertainty index of output response was calculated by a strategy based on the weighted prediction variance (WPV), and the two indices are combined by the weight coefficients. In the iterative process, the weight coefficients of the two indices were determined according to the accuracy of the hybrid surrogate model. The balance of global and local accuracy was realized by considering the sample dispersion and the fitting uncertainty of the surrogate model comprehensively. Numerical examples of single-response and multi-response systems showed that the proposed approach has excellent sampling efficiency and robustness. Moreover, the results of actual engineering application showed that the hybrid surrogate model constructed through MWAS could significantly improve the efficiency of model optimization. Hence, a high-precision optimization solution to the multi-objective lightweight design of passenger car rear seat was obtained.

INDEX TERMS Adaptive sampling, hybrid surrogate model, multi-objective lightweight, multi-response systems.

I. INTRODUCTION

Safety, energy conservation and environmental friendliness are the three major themes of the development of the automobile industry. While lightweight is one of the most important means to achieve these goals. Researchers all over the world have made tremendous efforts on this aspect [1]–[4]. However, in the field of automobile lightweight design, there is little research on passenger car rear seats. Therefore, it is of great significance to reduce the weight of the passenger car rear seat for the development of the automobile industry while satisfying the safety performance and riding comfort [5].

With the rapid development of computer hardware and software, high-fidelity simulations are often used to replace real-life experiments to reduce the overall time and cost. However, the simulation for a complex system with high dimensional and multi-output is computationally expensive. A widely used strategy is to utilize surrogate models and replace the expensive simulation model during the process.

Various types of optimization designs (multi-objective optimization, reliability-based design optimization, and multi-disciplinary design optimization) can be carried out quickly and conveniently, based on surrogate models and optimization algorithms [6]. Currently, mainstream surrogate models such as Polynomial Response Surface (PRS) models, Kriging (KRG) models, Radial Basis Function (RBF) models, and Elliptic Base Function (EBF) models have been widely applied to engineering optimization and achieved greater successes [7]–[12]. However, when solving multi-objective optimization problems with high dimensions and high degree of nonlinearity, it is difficult to obtain reasonably optimized results while relying only on a single surrogate model [13]. To solve this problem, Zetpa et al. [14] have first proposed the concept of a hybrid surrogate model and then applied it to alkaline surfactant polymer flooding to improve oil recovery. Chen et al. [15] and Zhang et al. [16] have proposed a new hybrid surrogate model which combined the advantages of global and local measures, and the appropriate trade-off between the two measures was made through the weight coefficients. Similarly, Yin et al. [17] have proposed a multi-region optimization hybrid surrogate model, which was...
applied to two engineering examples of thin-walled column crushing and airbag buffering. The results showed that the prediction accuracy of the hybrid surrogate model proposed in the above literatures was higher than other single surrogate models.

In the field of multi-objective engineering optimization research, the accuracy of the surrogate model largely depends on the number of sample points and their location distribution on the design space [18]. Therefore, the distribution of the sample points in a reasonable location has become the key factor for engineering optimization. At present, there are mainly two types of sample techniques. The one-time sampling approach generates all sample points at one time through experimental design which is quick and easy. However, it requires a large number of samples to explore the design space, and re-sampling is required when the sample data cannot meet the expected requirements. On the other hand, the adaptive sampling approach uses the experimental design to obtain a certain number of initial sample points and followed by the determination of the next sample point according to the information obtained from the initial sample points and relevant evaluation criteria [19]. Compared with the one-time sampling approach, the adaptive sampling approach has better flexibility and high efficiency. In addition, the latter approach can effectively avoid over-sampling. Therefore, scholars over the world have conducted a large number of studies on how to obtain the next sampling point of the information about the last iteration. Xiong et al. [20] have formulated the adaptive sampling as an optimization problem. This approach combined the maximum and minimum distance with the projection distance, and the next sample point was obtained by algorithm optimization. Liu et al. [19] have used Monte Carlo approach and space reduction strategy combined with local boundary search algorithm to achieve adaptive sampling. However, the determination of sample points is independent of the output characteristics in these approaches. To improve the overall model accuracy efficiently, an adaptive sampling approach should utilize the characteristics of both inputs and outputs for choosing sample points. Jiang et al. [21] have proposed an adaptive sampling approach based on technique for order preference by similarity to an ideal solution (TOPSIS) and Delaunay triangulation, which achieved a balance between global exploration and local exploitation during the sampling process. Yang and Xue [22] have proposed a weighted adaptive sampling approach that comprehensively considered the influence of sample quality of the input and output parameter space. Jin et al. [23] have introduced the mean squared error (MSE) method for global meta-modeling. In this approach, the new sample point \( x_C \) with the largest mean squared error is selected form the existing surrogate model (created based on the existing sample set \( P_C \)) to complete the adaptive sampling. However, this technique is no longer applicable when meta-modeling techniques other than Kriging are used. Most of the current adaptive sampling techniques are only applicable to single-response problems. In addition, most of the engineering optimization problems are multi-response system problems. So, it has certain limitations. At the same time, there are limited studies on the optimization efficiency and application value of the combination of hybrid surrogate model and adaptive sampling approach.

In this article, a multi-response weighted adaptive sampling (MWAS) approach based on the hybrid surrogate model is proposed. This approach comprehensively considers the dispersion of sample points in the input design space and the fitting uncertainty of output response. Hence, the balance of global and local accuracy optimization can be achieved. The WPV of the hybrid surrogate model is used to identify areas with large local errors and the MDA approach is used to acquire the dispersion of sample points in the design space. In the iterative process, the weight coefficients are adaptively selected according to the accuracy of the hybrid surrogate model. Using this approach for sampling the design space, the sampling efficiency and the fitting accuracy of the surrogate model can be improved effectively.

II. THEORY OF MODEL OPTIMIZATION
A. HYBRID SURROGATE MODEL

Hybrid surrogate model was constructed by using weighted linear combinations of different surrogate models, and it is one of the best ways to improve the model prediction accuracy [24]. Its basic form is defined as:

\[
\begin{align*}
\hat{y}_{HS} &= \sum_{i=1}^{N} w_i \hat{y}_i(x) \\
\sum_{i=1}^{N} w_i &= 1
\end{align*}
\]

where \( \hat{y}_{HS} \) is the predicted response by the hybrid surrogate model, \( w_i \), \( \hat{y}_i \) are the weight and predicted response of \( i \)th sub-model respectively, \( N \) is the number of sub-models in the hybrid surrogate model.

In the hybrid surrogate model, the size of the weight coefficient reflects the predictive ability of the sub-model, and the more accurate sub-model should be assigned a larger weight coefficient. The root mean square error (RMSE) heuristic weighting scheme based on cross-validation was adopted to construct a hybrid surrogate model [25]. The formula is shown in Equation (2):

\[
\begin{align*}
\hat{E} &= \frac{1}{N} \sum_{i=1}^{N} E_i, \quad E_i = RMSE_i = \sqrt{\frac{1}{M} \sum_{j=1}^{M} (y_j^i - \hat{y}_j^i)^2} \\
\end{align*}
\]

where \( y_j^i \) is the response of \( j \)th sample, \( \hat{y}_j^i \) is the predicted response obtained for the surrogate model constructed using all sample points except the \( j \)th sample, \( E_i \) is the RMSE of \( i \)th sub-model, \( M \) is the number of sample, \( \alpha \) and \( \beta \)
is respectively controlling the importance of averaging and sub-models, with values of 0.05 and -1.

The fitting accuracy of the surrogate model directly affects the feasibility and rationality of the optimal solution. Only if the surrogate model meets the accuracy requirements then the optimized solution has credibility. In order to assess the fitting accuracy of the surrogate model, two frequently used evaluation indicators were adopted: relative root mean square error (RRMSE) and relative maximum absolute error (RMAE) [26], [27]. The mathematical definitions of the two indicators are shown in Equation (3):

\[
\text{RRMSE} = \sqrt{\frac{1}{n_t} \sum_{i=1}^{n_t} (y_i - \hat{y}_i)^2} \\
\text{RMAE} = \frac{\text{Max} \left \{ \frac{\text{std} \left[ y_i - \hat{y}_i \right]}{\text{std} \left[ y \right]} \right \}}{\text{std} \left[ y \right]}
\]

where \(y_i\) and \(\hat{y}_i\) are the actual response and the predicted response of \(i\)th test sample points respectively, \(n_t\) is the number of test sample points, and \(\text{std}\) is the standard deviation of \(y_i\).

RRMSE represents the global accuracy of the surrogate model. The closer the value of RRMSE is to 0, the higher the global accuracy of the model. On the flip side, RMAE reflects the local accuracy of the surrogate model. The smaller the value of RMAE indicates the higher regional accuracy of the surrogate model.

According to research by Pan [28], when the hybrid surrogate model contains 3 to 5 sub-models, the prediction accuracy was the highest. Therefore, this paper constructs 9 single surrogate models (as shown in Table 1) based on different parameter settings, including 3 PRS models, 4 KRG models, 1 RBF model, and 1 EBF model. Then they were sorted according to the GMSE, and the three single surrogate models with the highest fitting accuracy were weighted to construct a hybrid surrogate model.

### B. EXPERIMENTAL DESIGN

The selection of the preliminary sample points is the premise and foundation of the surrogate model construction. Mainstream experimental designs include optimal Latin hypercube design (OLHD), full factorial design, central composite design, and orthogonal array [29]–[33]. The OLHD is an experimental design approach proposed by Mckay et al. [29] to decrease the number of actual engineering simulation analysis experiments. This approach maximizes the stratification of each edge distribution to ensure full coverage of each variable range and then performs efficient sampling from the patchy distribution interval, which has very good space-filling and balance. Therefore, this paper adopts the OLHD to obtain the initial sample points. In the OLHD approach, the number of sampling points is usually selected according to the design variables and design space, and the number should not be less than the expected minimum sample point of the fitting surrogate model, as shown in Equation (4) [34].

\[ N_C \geq 5n_s \]  

where \(N_C\) is the number of sample points, \(n_s\) is the dimension of the design space.

### III. MULTI-RESPONSE WEIGHTED ADAPTIVE SAMPLING (MWAS) APPROACH

In the proposed approach, the accuracy of the surrogate model is improved by considering the sample discreteness index in the input design space and the fitting uncertainty index of output response. In the iterative process, according to the influence of new sample points on the accuracy of the surrogate model, the weight coefficients of the two indices were determined to achieve the balance between global and local accuracy.

### A. CALCULATION OF THE FITTING UNCERTAINTY INDEX OF OUTPUT RESPONSE BASED ON WEIGHTED PREDICTION VARIANCE

A strategy based on the prediction variance was used to calculate the deviation of the sub-model and the hybrid surrogate model at an unknown point in the input design space. This approach needs several different surrogate models to predict the response of candidate point \(x\), and select the point with the largest divergence as the update point. The degree of divergence at point \(x\) is usually identified as the prediction variance of hybrid surrogate model [35]. The expression is shown in Equation (5):

\[
\sigma_{\hat{y}_i}^2(x) = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i(x) - \bar{y}(x))^2 \\
\bar{y}(x) = \frac{\sum_{i=1}^{N} y_i(x)}{N}
\]

where \(y_i\) is the predicted value of \(i\)th sub-model, \(\bar{y}\) is the average of the predicted value of \(N\) sub-models.

As can be recognized from Equation (5), the greater the predictive variance, the greater the difference between a sub-model and the hybrid surrogate model at that point. It indicated a great local error in the region. However, when the difference of the predicted value between a sub-model and the hybrid surrogate model is too obvious, the weight coefficient of the sub-model is too small. It is unreasonable for the next sample point completely determined by the sub-model. Meanwhile, the engineering optimization design problem is mostly a multi-response system, but the above-mentioned research is only applicable to the single response output. Combining the prediction variance strategy with the weight coefficient of the sub-model, a weighted prediction variance is proposed and used as a fitting uncertainty index.

The weighted prediction variance firstly uses a hybrid surrogate model constructed by \(N\) different sub-models to fit the multi-response system problem, and predicts all the
responses of the candidate point \( x \). The expression is shown in Equation (6):

\[
\begin{align*}
\sigma^2_{WPV_j}(x) &= \sum_{i=1}^{N} \omega_{ji} \left[ y_i(x) - y_{HS_i}(x) \right]^2 \\
1 &\leq j \leq t; \quad 1 \leq i \leq N
\end{align*}
\]

(7)

where \( \omega_{ji} \) is the weight coefficient of the \( ji \)th sub-model, \( t \) is the number of responses.

On the basis of Equation (5), the weighted predictive variance (WPV) expression was established by combining the weighted coefficients of the sub-model:

\[
\begin{align*}
\sigma^2_{WPV}(x) &= \sum_{i=1}^{N} \omega_{i} \left[ y_i(x) - y_{HS_i}(x) \right]^2 \\
1 &\leq j \leq t; \quad 1 \leq i \leq N
\end{align*}
\]

(6)

where \( \omega_{i} \) is the weight coefficient of \( j \)th sub-model in \( it \)h response.

For multi-response system problems, different responses may produce an order of magnitude difference in weighted prediction variance, so the weighted prediction variance of each response was normalized:

\[
\hat{\sigma}^2_{WPV_j}(x) = \frac{\sigma^2_{WPV_j}(x) - \min \sigma^2_{WPV_j}(x)}{\max \sigma^2_{WPV_j}(x) - \min \sigma^2_{WPV_j}(x)}
\]

(8)

where \( \sigma^2_{WPV_j} \) is the weighted prediction variance of \( j \)th response, \( \hat{\sigma}^2_{WPV_j} \) is the normalized value of \( \sigma^2_{WPV_j} \), \( \max \sigma^2_{WPV_j} \) and \( \min \sigma^2_{WPV_j} \) are the maximum and minimum WPV of \( j \)th response, respectively.

Finally, a multi-response weighted predictive variance expression was established by combining the correlation degrees of each response:

\[
\sigma^2_{WPV}(x) = \sum_{i=1}^{t} \omega_{j} \hat{\sigma}^2_{WPV_j}(x)
\]

(9)

where \( \omega_{j} \) is the weight coefficient of \( j \)th response which is obtained by the entropy weight approach based on the initial sample data [36]. When \( t = 1 \), it is the single response version of the weighted prediction variance.

A one-dimensional test function was used to illustrate that the weighed prediction variance can efficiently identify the areas with large prediction deviations.

\[
y = 2x^2 - 5x - 2(2 \cos \pi x - 3 \sin \pi x) + 10
\]

(10)

The 9 surrogate models in Table 1 were constructed from 7 sample points uniformly distributed in the design space. Figure 1(a) show the difference between the predicted curve of the hybrid surrogate model and the actual function curve. It can be seen from the figure that the weighted prediction variance has a good correlation with the actual error. The location of the new sample point is the region with the largest actual deviation. Figure 1(b) show the comparison between the weighted prediction variance and the actual deviation square. The results showed that weighted prediction variance and the actual deviation have the same increase and decrease trend. The positions of the maximum points of the weighted prediction variance curve and the actual deviation curve are consistent. It indicated that this approach can effectively identify the area with large local error.

From the above research, it can be concluded that fitting uncertainty index can effectively identify areas with large local error. However, this approach only considers the output response. Furthermore, in the absence of constraints, update points tend to cluster in a small area. When points are clustered together, the matrix used to fit the surrogate model will have ill-conditioned mutations, which will lead to a poor fit.

### B. Calculation of the Sample Discreteness Index in the Input Design Space Based on Maximum and Minimum Distance Approach

To solve the problem that poor fitting of the surrogate model caused by sample point aggregation, and to make the sample distribution sparse in areas with insensitive fitting accuracy, a distance term \( d \) was added to make the trade-off between uniformity and sparseness of samples in different regions. However, it is difficult to determine a suitable distance term \( d \) in practical applications. When \( d \) is too large, the sample points in the area with large local error will be filtered out, and the area with heavy local error cannot be effectively explored.

### Table 1. Basic information of several sub-models.

| ID | Sub-model          | Details                                                                 |
|----|-------------------|-------------------------------------------------------------------------|
| 1  | PRS2              | Full model of degree 2, 3 and 4 for Polynomial Response Surface          |
| 2  | PRS3              |                                                                         |
| 3  | PRS4              |                                                                         |
| 4  | KRG-Gaussian      | Four different Kriging models are selected by correlation models        |
| 5  | KRG-Exponential   |                                                                         |
| 6  | KRG-Matern Linear |                                                                         |
| 7  | KRG-Matern Cubic  |                                                                         |
| 8  | RBF               | A type of neural network employing a hidden layer of radial units and an output layer of linear units |
| 9  | EBF               | Similar to RBF but use elliptical units in place of radial units        |

---

J. Long et al.: Multi-Response Weighted Adaptive Sampling Approach Based on Hybrid Surrogate Model

VOLUME 9, 2021


\[ \max f(x_{n+1}) = \max (\omega_1 f_d(x_{n+1}) + \omega_2 f_i(x_{n+1})) \]

\[ \omega_1 + \omega_2 = 1 \]

where \( f_d(x_{n+1}) \) and \( f_i(x_{n+1}) \) are sample discreteness index and fitting uncertainty index respectively, \( \omega_1 \) and \( \omega_2 \) are the weight coefficients of \( f_d(x_{n+1}) \) and \( f_i(x_{n+1}) \) respectively.

The sample discreteness index and fitting uncertainty index were normalized to avoid misleading the acquisition of new sampling points caused by the difference in orders of magnitude.

\[ f_d(x_{n+1}) = \frac{d_{\min}(x_{n+1}, x_{A_j}) - \min d_{\min}(x_{n+1}, x_{A_j})}{\max d_{\min}(x_{n+1}, x_{A_j}) - \min d_{\min}(x_{n+1}, x_{A_j})} \]

\[ f_i(x_{n+1}) = \frac{\sigma^2_{WPV}(x_{n+1}) - \min \sigma^2_{WPV}(x_{n+1})}{\max \sigma^2_{WPV}(x_{n+1}) - \min \sigma^2_{WPV}(x_{n+1})} \]

In the adaptive sampling process, the global accuracy was gradually improved with the increasing of sample points, and it was more reliable when looking for the next sample point. Therefore, \( \omega_1 \) is a large number in the early stage of adaptive sampling, and \( \omega_2 \) takes a larger value in the later stage. The values of \( \omega_1 \) and \( \omega_2 \) are determined by Equation (14) and Equation (15):

\[ R^2 = 1 - \frac{\sum_{i=1}^{n_t} (y_i - \tilde{y}_i)^2}{\sum_{i=1}^{n_t} (y_i - \bar{y}_i)^2} \]

where \( \bar{y}_i \) is the average value of the actual response value, \( n_t \) is the number of test sample points, \( R^2 \) reflects the global accuracy of the surrogate model. The closer the value of \( R^2 \) is to 1, the higher the global accuracy of the model [31].

D. PROCESS OF MULTI-RESPONSE WEIGHTED ADAPTIVE SAMPLING APPROACH

Figure 2 displays a flowchart of the MWAS approach. Firstly, the initial sample set \( P_C \) with \( n \) sample points was generated by OLHD, and the responses of all sample points were achieved by simulation analysis or experiment. Secondly, the initial sample set \( P_C \) was used to construct the 9 surrogate models in Table 1, and the hybrid surrogate model was
constructed by using three sub-models with the highest accuracy. Then, the two indices (fitting uncertainty and sample discreteness) of the candidate sample points were calculated, and the next update point was determined by solving the optimization problem in Equation (12). Import the update point into the simulation model to obtain the sample response value of the kth iteration. Finally, the hybrid surrogate model was reconstructed by using the total sample set PA, and the fitting accuracy evaluation index (RRMSE, RMAE) was calculated depending on Equation (3). If surrogate model accuracy meets the convergence criteria or computed resources (budget or time) are exhausted, no additional samples are needed and the final sample set of PA is the output. Otherwise, continue to cycle sampling to obtain a new sample point.

IV. NUMERICAL EXAMPLE

A. TEST FUNCTION

In order to evaluate the practical feasibility of the MWAS. Three test functions were applied to form three single-response and four multi-response systems for test analysis. The three test functions are shown in Equation (16-18). Figure 3 shows the basic shapes of the three test functions.

\[ F_1: f_1(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_2^2 + x_1 - 7)^2, \quad x_1, x_2 \in [-3, 3] \]  
(16)

\[ F_2: f_2(x_1, x_2) = 2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} + x_1x_2 + x_2^2, \quad x_1, x_2 \in [-3, 3] \]  
(17)

\[ F_3: f_3(x_1, x_2) = 3(1 - x_1)^2 \exp(-x_1^2 - (x_2 + 1)^2) - 10(\frac{x_1}{5} - x_1^3 - x_2^5) \exp(-x_1^2 - x_2^2) - \frac{1}{3} \exp(-(x_1 + 1)^2 - x_2^2), \quad x_1, x_2 \in [-3, 3] \]  
(18)

B. TEST SCHEME

For comparison, the other four approaches were adopted to test the examples, including the MSE [23], MDA, WPV and OLHD. For the four adaptive sampling approaches, 10 initial sample points were generated by OLHD, and 25 sample points were gradually collected through their respective adaptive sampling principles. While OLHD collected 35 sample points at one-time. Accuracy evaluation indices RRMSE and RMAE of the surrogate model were calculated based on 1000 test sample points obtained by OLHD. Global and local accuracy (RRMSE and RMAE) of the surrogate model constructed with the same total sample size was used to measure the sampling efficiency. The smaller the RRMSE and RMAE under the equal sample size, the better the sampling efficiency.
The proposed single response version of MWAS was applied to three single response functions for analysis. Accuracy results of different sampling methods are listed in Table 2, and the best results are indicated in bold.

It can be seen from Table 2 that when the sample size was the same, the global and local accuracy obtained by the MWAS approach was higher than the other four sampling methods. Compared with the WPV and MDA approaches, the MWAS approach (which comprehensively considers the fitting uncertainty of output response and the sample dispersion) has higher sampling efficiency. At the same time, the local accuracy of the surrogate model constructed by WPV is better than that of MDA, MSE and OLHD approaches.

Comparing MWAS with the other four sampling methods, improvements of RRMSE and RMAE are shown in Figure 4. MWAS method significantly improved the accuracy of the surrogate model compared to the other four sampling approaches, and was more prominent in the test function F3, which changed dramatically in response.

To compare the performance of the proposed MWAS on multi-response system problems, five sampling methods were applied to four multi-response systems. Table 3 shows the model accuracy of five multi-response system problems. The bottom of the table shows the average and standard deviation of the three test functions for the accuracy of the surrogate model in the different multi-response system. In Table 3, the best results are indicated in bold.

Table 3 shows that under the same number of samples, MWAS method achieved the best model accuracy, followed by WPV. From the results of the standard deviation, it can be concluded that MWAS approach has better robustness in the F1 and F3 functions, while it performs worse than the other approach in the F2 function. Meanwhile, the proposed MWAS approach can provide a more accurate surrogate...
TABLE 3. The accuracy results of multi-response systems by different sampling methods.

| Multi-response systems | Test functions | Accuracy criteria | MWAS | WPV  | MDA  | MSE  | OLHD |
|------------------------|----------------|------------------|------|------|------|------|------|
|                        | F1             | RRMSE            | 0.088| 0.187| 0.079| 0.116| 0.092|
|                        |                | RMAE             | 0.234| 0.845| 0.541| 0.629| 0.723|
|                        | F2             | RRMSE            | 0.074| 0.182| 0.235| 0.155| 0.170|
|                        |                | RMAE             | 0.319| 0.644| 1.071| 0.596| 1.082|
|                        | F3             | RRMSE            | 0.071| 0.042| 0.054| 0.079| 0.068|
|                        |                | RMAE             | 0.324| 0.200| 0.308| 0.400| 0.707|
|                        | F1+F2          | RRMSE            | 0.301| 0.309| 0.426| 0.371| 0.390|
|                        |                | RMAE             | 1.030| 1.049| 1.986| 1.269| 1.539|
|                        | F3             | RRMSE            | 0.141| 0.099| 0.154| 0.212| 0.141|
|                        |                | RMAE             | 0.699| 0.562| 1.134| 0.734| 1.071|
|                        | F2+F3          | RRMSE            | 0.280| 0.335| 0.446| 0.326| 0.445|
|                        |                | RMAE             | 0.894| 1.198| 2.034| 1.194| 2.029|
|                        | F1+F3+4         | RRMSE            | 0.066| 0.099| 0.074| 0.088| 0.085|
|                        |                | RMAE             | 0.204| 0.446| 0.301| 0.339| 0.503|
|                        | F3             | RRMSE            | 0.223| 0.181| 0.176| 0.272| 0.154|
|                        |                | RMAE             | 0.815| 0.672| 0.776| 0.838| 1.308|
|                        | F3             | RRMSE            | 0.284| 0.408| 0.500| 0.458| 0.442|
|                        |                | RMAE             | 1.269| 1.506| 2.280| 1.696| 2.040|
|                        | F1             | RRMSE            | 0.075| 0.109| 0.069| 0.094| 0.082|
|                        |                | RMAE             | 0.254| 0.497| 0.383| 0.456| 0.644|
|                        | AVE            | RRMSE            | 0.146| 0.154| 0.188| 0.213| 0.155|
|                        |                | RMAE             | 0.611| 0.626| 0.994| 0.723| 1.154|
|                        | F3             | RRMSE            | 0.288| 0.351| 0.457| 0.385| 0.426|
|                        |                | RMAE             | 1.064| 1.251| 2.100| 1.386| 1.869|
|                        | F1             | RRMSE            | 0.009| 0.06 | 0.011| 0.016| 0.010|
|                        |                | RMAE             | 0.051| 0.266| 0.112| 0.125| 0.100|
|                        | SD             | RRMSE            | 0.061| 0.039| 0.034| 0.048| 0.012|
|                        |                | RMAE             | 0.212| 0.047| 0.156| 0.099| 0.109|
|                        | F3             | RRMSE            | 0.009| 0.042| 0.031| 0.055| 0.025|
|                        |                | RMAE             | 0.155| 0.190| 0.129| 0.221| 0.234|

model for multi-response system problems with different characteristics. At the same sample size, the model accuracy of the test function in the multi-response system was lower than that in the single-response system because as the size of the system increased, the effect of output responses with more different characteristics on sample update point location selection needs to be considered.

V. LIGHTWEIGHT DESIGN OF PASSENGER CAR REAR SEAT

A. LUGGAGE COLLISION TEST OF PASSENGER CAR REAR SEAT

Figure 5 shows the principle of a luggage collision test of passenger car rear seat. According to the requirements of GB15083-2006, a passenger car rear seat assembly and two test specimens with a size of 300 mm × 300 mm × 300 mm, edge chamfer of 20 mm and mass of 18 kg were placed in the trolley test bench. Then, the acceleration-time curve illustrated in Figure 5(a) was applied to the trolley test bench to simulate a passenger car collision. Figure 5(b) shows the state of the passenger car’s rear seat at a certain moment during the collision [38]. The finite element model of the luggage collision test was established by HyperMesh software and solved by using Ls_Dyna software (as shown in Figure 5(c)).

B. SELECTION OF DESIGN VARIABLES AND RESPONSES

During the luggage collision test, the main force and energy-absorbing component was the backrest skeleton. Therefore, according to the symmetry and functionality of the backrest skeleton structure, the tube and plate parts of the backrest skeleton were simplified into 5 sets of optimized parts (as showed in Figure 6) marked as $P_1$ to $P_5$. The thickness
and material of the optimized parts are invoked as design variables, which were marked as $t_1$-$t_5$ and $m_1$-$m_5$ respectively.

Table 4 shows the main performance parameters of candidate materials for optimized parts, including yield strength (SIGY), ultimate tensile strength (UTS), percentage elongation (PE), and relative cost. The candidate materials are divided into three groups concerning a gradual increasing value of SIGY. Table 5 exhibited the value ranges of the design variables.

According to regulation, the front profile of the headrests is not allowed to be moved out of the 150 mm horizontal-vertical plane in front of the seat design reference point during the test. The backrest skeleton and its fasteners are permitted to have a certain degree of deformation, provided that they can’t fail. The maximum strain criterion was adopted as the failure criterion, and introduced the strain index (as shown in Equation (19)) to judge whether the component fails. When the strain index is greater than 1, it means that the component is in a failure state; otherwise, it is in a safe state [39]. Therefore, the maximum displacement of the headrest ($L_i$), the strain index of each optimized parts ($Q_i$), the total mass ($M$) and total price ($C$) of the optimized parts are used as evaluation indicators for safety performance and lightweight.

$Q_i = \frac{S_i}{E_i}$  \hspace{1cm} (19)

where $Q_i$ represents the strain index of $i$th component, $S_i$ is the maximum plastic strain of $i$th component, $E_i$ is the elongation of the material used in $i$th component.
In summary, this 10-dimensional multi-response system problem with 8 output responses can be expressed as:

$$\begin{align*}
M(\mathbf{t}, \mathbf{m}), \ C(\mathbf{t}, \mathbf{m}), \ L(\mathbf{t}, \mathbf{m}), \\
Q_i(\mathbf{t}, \mathbf{m}), \ i = [1, 2, \ldots , 5], \\
\mathbf{t} = [1.0, 1.1, \ldots , 2.2], \\
\mathbf{m} = [101, 102, 103, 104, 201, 202, 203, 204, 301, 302, 303, 304]
\end{align*}$$

where $M(\mathbf{t}, \mathbf{m})$ and $C(\mathbf{t}, \mathbf{m})$ are the total mass and total price of optimized parts respectively, $L(\mathbf{t}, \mathbf{m})$ is the maximum displacement of the headrest, $Q_i(\mathbf{t}, \mathbf{m})$ is the strain index of $i$th optimized part, $\mathbf{t}$ is the thickness design variable, $\mathbf{m}$ is the material design variable.

### C. ADAPTIVE SAMPLING SCHEME

Considering that the high-fidelity collision simulation is a very time consuming work. In this section, only four adaptive sampling approaches were used for comparison. Firstly, the OLHD method was used to obtain 100 initial sample points and 2000 candidate sample points respectively. Then, four adaptive sampling approaches were used to iteratively add 60 new sample points from the candidate sample points on the basis of the surrogate model constructed by the initial sample points. Finally, 50 random sample points were used to estimate the accuracy of each surrogate model to judge the sampling efficiency of the four adaptive sampling approaches.

Table 6 shows the model accuracy of each response in the lightweight design of passenger car rear seats, and the best results are indicated in bold. It can be seen from the table that, under the same number of sample points, the surrogate model accuracy obtained by MWAS method was the best except for response $C$. Among them, the three lower nonlinear responses of $M$, $C$, and $L$ have lower requirements on the number of samples. With a large number of sample points, the four methods all achieved high accuracy, among which MWAS and MSE performed better. Compared with the other three adaptive sampling methods, MWAS has a great improvement in global and local accuracy for five high nonlinear responses from $Q_1$ to $Q_5$. It shows that the MWAS method has useful application value for complex engineering optimization problems.

### D. MULTI-OBJECTIVE OPTIMIZATION

The thickness and material types of the optimized parts were simultaneously taken as design variables and the strain index of the optimized parts were taken as constraints. Then, the total mass and price of the optimized parts, together with the maximum displacement of the headrest, were selected as three conflicting optimization objectives. The constraint value was set to be 0.8, after the error of simulation results and the safety factor of the product were taken into account. Therefore, the final multi-objective optimization design mathematical model of passenger car rear seat is shown in Equation (21):

$$\begin{align*}
\text{Minimize} \quad & \{ M(\mathbf{t}, \mathbf{m}), C(\mathbf{t}, \mathbf{m}), L(\mathbf{t}, \mathbf{m}) \}, \\
\text{S.t.} \quad & Q_i(\mathbf{t}, \mathbf{m}) \leq 0.8, \ i = [1, 2, \ldots , 5], \\
& \mathbf{t} = [1.0, 1.1, \ldots , 2.2], \\
& \mathbf{m} = [101, 102, 103, 104, 201, 202, 203, 204, 301, 302, 303, 304]
\end{align*}$$

NSGA-II algorithm was utilized to solve the multi-objective optimization problem of Equation (21) based on the hybrid surrogate model [40]. The parameter setting of NSGA-II algorithm was presented in Table 7. Due to the reason that the optimization results obtained by NSGA-II algorithm were stochastic, 10 runs were performed and then the optimal solution set with the best Pareto frontier distribution were taken as the final choice. Finally, 106 Pareto optimal solutions were obtained after 12,000 evaluations. Figure 7 displays the distribution of Pareto solution set in the objective space. It could be seen from the figure that the distribution of Pareto solution set in the objective space is a continuous and narrow spatial surface. It shows that the

| Responses | Accuracy criteria | MWAS | WPV | MDA | MSE |
|-----------|------------------|------|-----|-----|-----|
| $M$       | RRMSE            | 0.017| 0.023| 0.033| 0.034|
|           | RMAE             | 0.045| 0.058| 0.114| 0.070|
| $C$       | RRMSE            | 0.105| 0.106| 0.153| 0.104|
|           | RMAE             | 0.280| 0.228| 0.403| 0.213|
| $L$       | RRMSE            | 0.116| 0.122| 0.117| 0.121|
|           | RMAE             | 0.258| 0.258| 0.286| 0.261|
| $Q_1$     | RRMSE            | 0.438| 0.514| 0.584| 0.481|
|           | RMAE             | 1.253| 1.314| 1.388| 1.367|
| $Q_2$     | RRMSE            | 0.329| 0.386| 0.411| 0.361|
|           | RMAE             | 0.649| 0.895| 1.184| 0.930|
| $Q_3$     | RRMSE            | 0.332| 0.366| 0.348| 0.412|
|           | RMAE             | 0.829| 0.970| 0.925| 0.864|
| $Q_4$     | RRMSE            | 0.349| 0.536| 0.492| 0.508|
|           | RMAE             | 1.126| 1.703| 1.387| 1.288|
| $Q_5$     | RRMSE            | 0.235| 0.292| 0.315| 0.312|
|           | RMAE             | 0.720| 0.753| 0.819| 0.729|
| AVE       | RRMSE            | 0.240| 0.293| 0.307| 0.292|
|           | RMAE             | 0.645| 0.772| 0.813| 0.715|

### D. MULTI-OBJECTIVE OPTIMIZATION

The thickness and material types of the optimized parts were simultaneously taken as design variables and the strain index of the optimized parts were taken as constraints. Then, the total mass and price of the optimized parts, together with the maximum displacement of the headrest, were selected as three conflicting optimization objectives. The constraint value was set to be 0.8, after the error of simulation results and the safety factor of the product were taken into account. Therefore, the final multi-objective optimization design mathematical model of passenger car rear seat is shown in Equation (21):

$$\begin{align*}
\text{Minimize} \quad & \{ M(\mathbf{t}, \mathbf{m}), C(\mathbf{t}, \mathbf{m}), L(\mathbf{t}, \mathbf{m}) \}, \\
\text{S.t.} \quad & Q_i(\mathbf{t}, \mathbf{m}) \leq 0.8, \ i = [1, 2, \ldots , 5], \\
& \mathbf{t} = [1.0, 1.1, \ldots , 2.2], \\
& \mathbf{m} = [101, 102, 103, 104, 201, 202, 203, 204, 301, 302, 303, 304]
\end{align*}$$

NSGA-II algorithm was utilized to solve the multi-objective optimization problem of Equation (21) based on the hybrid surrogate model [40]. The parameter setting of NSGA-II algorithm was presented in Table 7. Due to the reason that the optimization results obtained by NSGA-II algorithm were stochastic, 10 runs were performed and then the optimal solution set with the best Pareto frontier distribution were taken as the final choice. Finally, 106 Pareto optimal solutions were obtained after 12,000 evaluations. Figure 7 displays the distribution of Pareto solution set in the objective space. It could be seen from the figure that the distribution of Pareto solution set in the objective space is a continuous and narrow spatial surface. It shows that the
TABLE 8. Comparison of the original and optimum responses.

| Type               | M     | L     | C     | Q1    | Q2    | Q3    | Q4    | Q5  |
|--------------------|-------|-------|-------|-------|-------|-------|-------|-----|
| Purposes           | Min   | Min   | Min   | ≤0.8  | ≤0.8  | ≤0.8  | ≤0.8  | ≤0.8|
| Original           | 10.53 | 84.88 | 11.76 | 0.284 | 0.371 | 0.609 | 1.295 | 0.351|
| Post-Lightweight   | 7.42  | 86.70 | 8.29  | 0.433 | 0.762 | 0.594 | 0.657 | 0.269|
| Simulation         | 7.35  | 85.26 | 8.04  | 0.425 | 0.784 | 0.614 | 0.685 | 0.268|
| Error%             | 2.97  | 1.66  | 0.90  | 1.90  | 2.91  | 3.45  | 4.20  | 0.43 |
| Change%            | 31.63 | -0.45 | 30.20 | -49.65| -111.32| -0.82 | 47.10 | 23.65|

FIGURE 7. Distribution of Pareto solutions in the three-objective space.

The comprehensive performance of the optimized result is fine, and there won’t be a large number of optimal solutions with superior performance for a single objective. However, selecting a specific solution based on subjective consciousness and engineering experience has a certain degree of randomness. Therefore, this paper uses the TOPSIS method [41] to select the optimal compromise solution from the multi-objective optimization problem.

E. VERIFICATION OF LIGHTWEIGHT DESIGN

After obtaining the best compromise solution, it is very important to verify whether the optimized result is appropriate. Table 8 exhibits the comparison before and after the lightweight of the rear seat of a passenger car. It can be seen from the table that the mass and relative cost of the optimized parts in the optimized solution are reduced from the original value of 10.53 kg to 7.42 kg, and 11.76 to 8.04 (i.e. the weight is reduced by 31.63% and the relative cost is reduced by 30.20%). At the same time, the displacement of the headrest has increased slightly, from the original value of 84.88 to 85.26 mm (i.e. an increased by 0.45%), but there is always a large margin from the 150 mm required by regulations. Compared with the original value, the strain index of the optimized parts has increased or decreased, but they are all within the range of the safety factor and meet the regulation requirements. Therefore, in lightweight design, it is more feasible to use the TOPSIS method to select the optimal compromise solution. The error between the optimized solution of each response and its simulation value were below 4.5%, indicated that the hybrid surrogate model based on the MWAS method has a better accuracy guarantee. One completed luggage crash simulation for an passenger car rear seat needs 6 to 8 hours. The computational cost of lightweight design of passenger car rear seats will be greatly reduced by adopting the proposed MWAS method based on the hybrid surrogate model.

VI. CONCLUSION

With the intention to improve the accuracy of the surrogate model and optimize efficiency, in this article a WPV approach based on the hybrid surrogate model was proposed to identify areas with large predicted deviations, and a MWAS approach was established by combining it with MDA.

(a) Through a one-dimensional test function, it is shown that the weighted prediction variance and the actual error have the same increase and decrease trend, and the maximum value was located in the same area, indicated that the approach can effectively identify areas with large local errors.

(b) Based on numerical examples and engineering cases, the sampling efficiency of MWAS, WPV, MDA, MSE and OLHD sampling methods were compared. All results showed that under the same number of sample points, MWAS has the highest accuracy of the surrogate model, that is, MWAS has the highest sampling efficiency. It indicated that MWAS can efficiently guide adaptive sampling and is useful for complex engineering optimization problems with long solution time and small maximum allowable sampling points.

(c) The MWAS method was applied to the rear seats of passenger cars for structure-material integration multi-objective lightweight design. The optimization solution obtained by NSGA-II algorithm optimization was compared with the pre-optimization scheme. Under the premise of ensuring that multiple performance indicators of the seat meet the requirements of regulations. The mass and cost of the optimized parts were reduced by 31.63% and 30.20%, respectively. Meanwhile, the headrest displacement increased by 0.45%. The error between the optimized solution of each response and its simulation value was below 4.5%. It can be seen that the hybrid surrogate model has an ideal fitting accuracy in the luggage collision simulation of passenger car rear seats.
JIANGQI LONG is currently a Professor with the College of Mechanical and Electrical Engineering, Wenzhou University. His main research interests include performance analysis, lightweight design theory of car body structure, and structural design and multidisciplinary optimization of auto parts.

YAOQING LIAO was born in 1996. He is currently pursuing the degree with the School of Mechanical and Electrical Engineering, Wenzhou University. His main research interest includes lightweight design of auto parts.

PING YU is currently an Assistant Professor with College of Mechanical and Electrical Engineering, Wenzhou University. Her main research interests include structure optimization design of vehicles and force sensors design for vehicles and robots.

* * *