Adaptive hybrid modeling method based on CV-Hypercube

Meimei Xiao[1], Xuelin Xie[1], Jingfang Shen[1,*], Yaohui Li[1,2]

[1] College of Sciences, Huazhong Agricultural University, Wuhan 430070, China,

[2] College of Mechanical and Electrical Engineering, Xuchang University, Xuchang, 461000, Henan, China;

*Corresponding author: sjf_712@mail.hzau.edu.cn

ABSTRACT

As common surrogate models, Kriging and RBF models have been widely used in various fields. Although the Kriging model and the RBF model have their own advantages, when the problems are complex and diverse, a single Kriging or RBF model usually cannot meet the requirements of global approximation. Luckily, the Kriging and the RBF model have good complementarity in performance. In view of this, an adaptive hybrid modeling method (AHM-CVH) based on cross-validation hypercube of Kriging and RBF is proposed in this paper. The CVH adaptive sampling strategy first generates a hypercube centered on the sample point with the largest cross-validation error, then candidate points are randomly sampled in the hypercube, and finally get a new sample point which is farthest from the center and surrounding samples. Eight benchmark functions ranging from 2 to 6 dimensions and an engineering example are validated, and the results show that the AHM-CVH method is superior to the single Kriging or RBF models in performance, and has the characteristics of high accuracy and strong stability.

Keywords: Surrogate model; Ensemble method; Adaptive sampling; Cross validation.

INTRODUCTION

Recently, the exponential growth of computer performance has made simulation experiments widely used in the design of expensive products, such as aircraft, ships, cars, and other complex machinery. However, the simulation of high-precision models requires a lot of computing resources and time (Gu. L., 2001 & Stefan et al., 2010), and such models are often not applicable when time or computing power is limited. In these cases, an approximate surrogate model (or meta-model) built with a small number of sample points can greatly reduce the computation and time (Khuri et al., 2010). Typical surrogate models include kriging (KRG) (Krige. D. G., 1952, Sacks et al., 1989 & Van et al., 2004), radial basis functions (RBF) (Hardy, 1971 & Buhmann. M. D., 2000), support vector regression (SVR) (Gunn. S. R., 1997), artificial neural networks (ANN) (Cheng et al., 1994) and polynomial response surface (PRS) (Shahsavania et al., 2009), etc. Among the above surrogates, Kriging and RBF are the two most popular methods (Li et al., 2017, Echard et al., 2011, Michael et al., 2019 & Shi et al., 2019). The reason is that the Kriging model has excellent approximation accuracy on low-dimensional and high nonlinear problems (Li et al., 2020), and the RBF model has the advantage that it can be applied to high-dimensional nonlinear problems without dimensionality reduction (Jin et al., 2001).

Since the intrinsic characteristics of the function are not a priori, it is challenging to know which

surrogate model best fits the objective function (Zhang et al., 2021). Analogous to the neural network combination principle, Zerpa et al. presented a new surrogate model by weighting PRS model, RBF model and Kriging model, in which the weights were selected by the error of the model on the test set (Zerpa et al., 2005). The modeling accuracy of the model proposed by Zerpa et al. is higher than the individual surrogates. But the above method requires test samples to determine the weights, which is not feasible in practical applications. In view of this, Goel et al. proposed three weight determination methods independent of the test set, one of which is to assign weights based on a global metric named generalized mean square cross-validation error (GMSE) on the current training set (Goel et al., 2007). It is characterized by the same weight for all predicted points in the entire design space. In addition, (Acar et al., 2019) innovatively transformed weight selection into an optimization problem that minimizes GMSE or RMSE, which can avoid the problem of different weights caused by different training sets. By replacing GMSE with pointwise cross-validation error, Acar put forward a method for determining weights using local errors (Acar E., 2010). While the method reduces the number of cross-validations, they also suffer from a loss in performance, since pointwise cross-validation errors are inherently not a substitute for errors at a point. To make use of the advantages of both global and local measures, Chen et al. (Chen et al., 2018) ensembled surrogate models which combining the strengths of global and local measures. However, in their method, the design space should be pre-divided into outer and inner regions for each sample point, and different strategies for evaluating the weight factors are adopted accordingly.

The accuracy of a surrogate model largely depends on its sampling strategy, which can be assorted into one-shot sampling and adaptive sampling. Adaptive sampling does not select all sampling points at once, it uses the existing response and surrogate model to determine the next new sampling point, which can greatly improve the over sampling and under sampling problems caused by one-shot sampling. Thus, this paper developed a novel hybrid model based on adaptive sampling strategy CVH to improve the accuracy of the surrogate models. First, leave-one-out method is used to calculate the cross-validation error (CVE) for each sample point. Usually there is a large uncertainty around the point with large CVE, so a hypercube is generated which is centered on the sample point with the largest CVE. Then a certain number of candidate points are randomly sampled in the hypercube. Finally, we selected the one farthest from the center of the hypercube and neighborhoods as a new sample point, which can avoid the samples from being too dense. The results show that the AHM-CVH model improves the accuracy and robustness of the model.

RELEATED THEORY

2.1 Kriging model

Kriging model consists of a regression model and a stochastic process, its general form can be written as

$$\hat{y} = \sum_{i=1}^{k} \beta_i f_i(x) + z(x) \tag{1}$$

where $f_i(x)$ is the regression polynomial (with zero order, first order and second order), β_i is the unknown coefficient, $z(x)$ is the realization of a stochastic process with zero mean and covariance of

$$Cov[z(x_1), z(x_2)] = \sigma^2 R(\theta, x_1, x_2) \tag{2}$$

where σ^2 is the process variance, and $R(\theta, x_1, x_2)$ is the spatial correlation function with a hyperparameter θ which takes $d = \|x_1 - x_2\|$ as the independent variable.

$$R(\theta, x_1, x_2) = \prod_{i=1}^{n} R_i(\theta_i, x_{1i} - x_{2i}) \tag{3}$$

Given n sample points $x_1, x_2, \cdots, x_n$ and their corresponding responses $Y_1, Y_2, \cdots Y_n$, the predicted value of the Kriging model at any point is

$$\hat{y} = f^T(x)\hat{\beta} + r(x)R^{-1}(Y - F\hat{\beta}) \tag{4}$$

The parameters $\hat{\beta}$ are obtained according to the optimal linear unbiased estimation as

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} Y \tag{5}$$

2.2 Radial basis functions

The RBF model is a linear combination of a set of basis functions, its basic expression is

$$\hat{y} = \sum_{w=1}^{N} \lambda_w \phi(r_w) = \sum_{w=1}^{N} \lambda_w \phi(\|x - x_w\|) \tag{6}$$

where r_w is the Euclidean norm of the matrix, ϕ is a basis function, λ_w is the weight coefficients for the w-th basis function.

Generally, if the selected sampling points $\{x_w\}$ are different in pairs, then $\{\phi(x - x_w)\}$ is linearly independent and can be used as a set of basis for the function space. When $\{x_w\}$ almost fills the space R^d, $\{\phi(x - x_w)\}$ and its linear combination can approximate almost any function. The commonly-used basis functions of RBF model are listed in Table 1.

Table 1. Type of RBF basis functions.

Type	Expression
Cubic	$\phi(r) = (r + c)^3$
Thin plate spline	$\phi(r) = r^2 \ln(cr)$
Gaussian	$\phi(r) = e^{-cr^2}$
Multiquadric	$\phi(r) = \sqrt{r^2 + c^2}$
Inverse-multiquadric	$\phi(r) = \frac{1}{\sqrt{r^2+c^2}}$

Given n samples $\{x_w, y_w\}$ $(w = 1,2,\ldots,n)$, weight coefficients λ_w are obtained by solving the following linear equations.

$$\Phi\lambda = y \tag{7}$$

where $\Phi_{ij} = \phi(\|x_i - x_j\|)$, $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_N\}$, $y = \{y_1, y_2, \ldots, y_N\}$

2.3 Theory of ensemble model

The principle of ensemble model, which combines multiple single surrogate models and weights linearly (Cheng et al., 2020 & Ye et al., 2020), its general form is

$$\hat{y}_{ens}(x) = \sum_{i=1}^{M} w_i(x)\hat{y}_i(x) , \sum_{i=1}^{M} w_i(x) = 1 \tag{8}$$

where $\hat{y}_{ens}(x)$ represents the predicted response value of the hybrid model at x, M is the number of surrogates constituting the hybrid model, $\hat{y}(x_i)$ is the predicted value of the surrogate at x, and $w_i(x)$ represents the weight of the ith surrogate. In general, the weight factors are selected such that the metamodels with high accuracy have large weight factors and vice versa (Zhou et al., 2011).

CONSTRUCTION OF AHM-CVH MODEL

3.1 Cross validation hypercube (CVH) criterion

This paper proposes a novel and versatile adaptive sequence sampling CVH criterion, which selects a new sampling point in the hypercube that is far enough away from the existing sample point to avoid sample clustering.

Assume that $n_{initial}$ initial sample points have been generated by the Latin hypercube sampling (LHS) method. First, the leave-one-out method is used to calculate the cross-validation error of each sample point as formula (9).

$$e_{Loo}(x_i) = \left| y(x_i) - \hat{f}^{-i}(x_i) \right| \tag{9}$$

Then, identify the largest e_{Loo} in the current sample set defined as x_{centre}, the definition of x_{centre} is show in formula (10).

$$x_{centre} = \arg\max\left(e_{Loo}(x_i)\right) \tag{10}$$

The vicinity of x_{centre} is an area with great uncertainty. Adding new sampling points in this area can greatly improve the accuracy of the surrogate model. Therefore, we generate a hypercube with side length 2ε centered on x_{centre}.

Since the LHS sampling is quite uniform, theoretically the $n_{initial}$ initial sample points will divide each dimension of the design space into $n_{initial} - 1$ parts. Thus, the point spacing L_j is calculated in each dimension direction, and the largest L_j is chosen as ε.

$$L_j = \frac{upper_j - lower_j}{n_{initial} - 1}, \ \varepsilon = \max_{j}\{L_j\}, \ j = 1, 2, \cdots D \tag{11}$$

where $upper_j$ and $lower_j$ represent the upper and lower bounds on the jth dimension respectively, D is the size of the dimension. Besides, the m candidate points were randomly generated by LHS method in the hypercube, denoted as P_{can}.

Finally, select a new sample point x_{new} in P_{can} that is farthest from x_{centre} and N, where x_{new} can be written as

$$x_{new} = \arg \max_{x \in P_{can}} \|x - (x_{centre} \cup N)\| \qquad (12)$$

where $\|x - (x_{centre} \cup N)\|$ represents the sum of the distances from x to each point in the set $(x_{centre} \cup N)$, and N represents the neighborhoods of x_{centre}.

Neighborhoods N are found according to the following rules.

$$d_{mean} = mean(\|x_{centre} - S\|)$$

$$N = \{x \in S \backslash \{x_{centre}\} | \ \|x - x_{centre}\| < \alpha \cdot d_{mean}\} \qquad (13)$$

where α is an adjustable parameter ranged from 0 to 1.

To visualize the CVH strategy, the one-dimensional test function is selected to establish an ensemble model of Kriging and RBF (Forrester et al., 2009), and the expression of the one-dimensional function is

$$f(x) = (6x - 2)^2 \sin(12x - 4), \text{where } x \in [0,1]. \qquad (14)$$

The local error of each point after the initial sampling of 5 points is shown in Figure 2(a). It is observed from Figure 1(a) that the 5th point has the largest cross-validation error, and the CVH method will select a new point as shown in Figure 1(b).
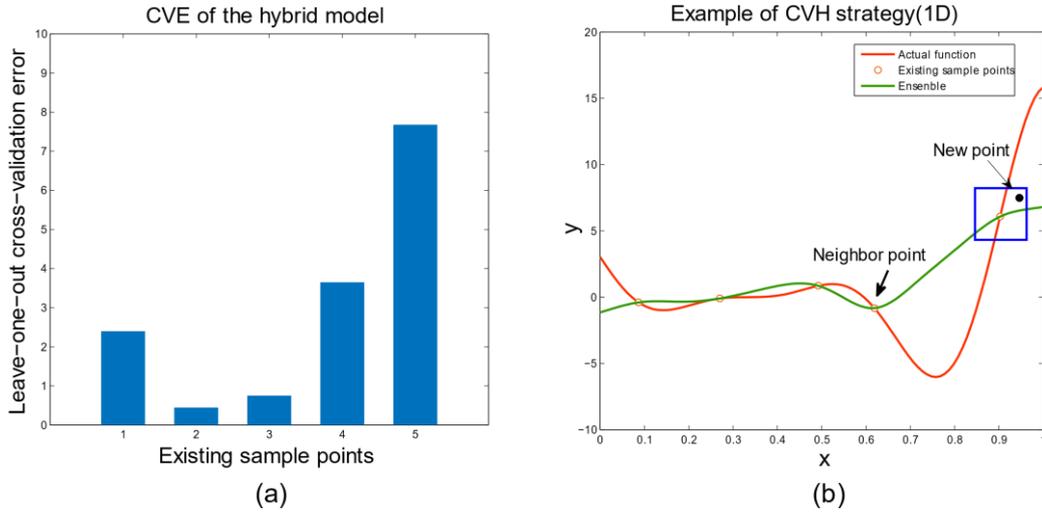

Figure 1. Schematic diagram of CVH strategy**.** (a) CVE of the hybrid model, (b) Example of CVH strategy(1D).

3.2 Modeling process

Based on the adaptive sampling strategy CVH, an adaptive ensemble model of Kriging and RBF can be established, where the weights are selected using the inverse proportional method. The modeling process is shown in Figure 2, and the specific implementation steps of modeling are as follows.
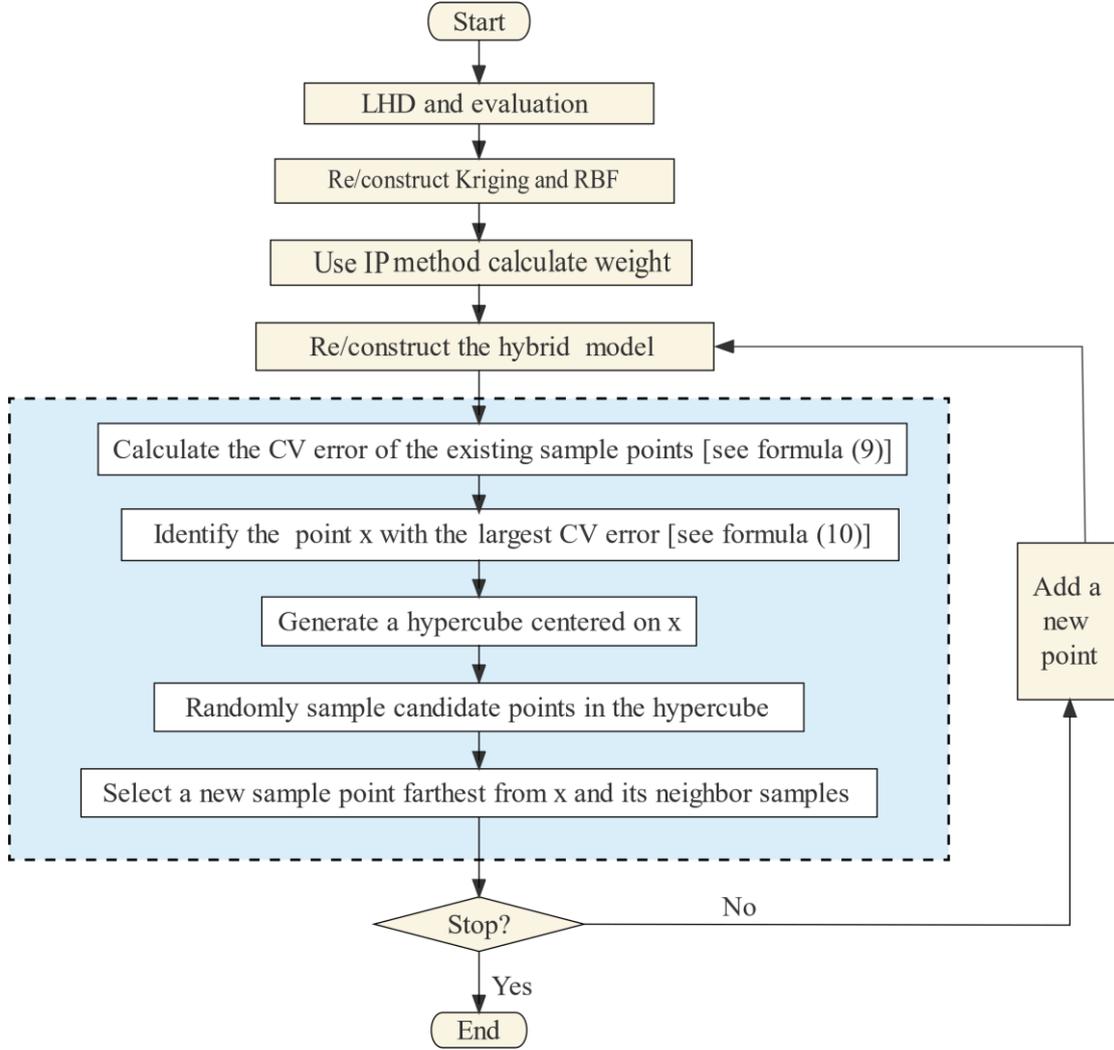
Figure 2. Flowchart of AHM-CVH modeling.

Step1. **Design of initial experimental.** Using Latin Hypercube Design (LHD) to obtain initial sample points, then perform expensive simulation evaluations of the sample points.

Step2. **Re/construct the single surrogate model.** Re/construct the Kriging model and RBF model respectively with sample points.

Step3. **Determine the weights.** The initial weight is set to 0.5, and the weight is calculated using the inverse proportional (IP) method after updating the points (Zerpa et al., 2005). The calculations are show in formula (15)-(17).

$$GMSE_k = \frac{1}{N}\sum_{i=1}^{N}\left(y(x_i) - \hat{f}^{-i}(x_i)\right)^2 \tag{15}$$

$$GMSE_r = \frac{1}{N}\sum_{i=1}^{N}\left(y(x_i) - \hat{g}^{-i}(x_i)\right)^2 \tag{16}$$

$$w_k = \frac{\frac{1}{GMSE_k}}{\frac{1}{GMSE_k}+\frac{1}{GMSE_r}}, w_r = \frac{\frac{1}{GMSE_r}}{\frac{1}{GMSE_k}+\frac{1}{GMSE_r}} \tag{17}$$

Step4. **Re/construct ensemble of models.** A new ensemble model is obtained by linearly combining the updated single model and the updated weights, and the formula is as follows.

$$\hat{y}_{ens}(x) = w_k \hat{y}_k(x) + w_r \hat{y}_r(x) \tag{18}$$

Step5. **Implement CVH criterion.** Implement the CVH criterion to obtain a new point x_{new}, and calculate its true response value y_{new}, see 3.1 for details.

Step6. **Stop criterion.** Since computing resources are limited and using the same number of evaluations facilitates comparisons between different methods, a fixed number of expensive evaluations are used as the stopping criterion. If the model satisfies the stopping condition, the modeling process is terminated. Otherwise, add x_{new} to the sample set and return to *Step2*.

The prediction accuracy of the approximate model directly affects the feasibility and rationality of the optimal solution. Generally, as the accuracy of the model increases, the confidence in the optimal solution based on the approximate model increases. In this paper, a global error metric root mean square error (RMSE) is selected as model evaluation indicators, its mathematical expressions are show in formula (19).

$$RMSE = \sqrt{\frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left(y(x_i) - \hat{y}(x_i)\right)^2} \tag{19}$$

where $y(x_i)$ is the true value of the objective function at x_i, $\hat{y}(x_i)$ is the predicted value of the surrogate model at x_i, and N_{test} is the number of test samples.

NUMERICAL APPLICATION

4.1 Trial setup

To test the modeling accuracy and stability of the AHM-CVH model, the Kriging model, the RBF model, and the AHM-CVH model were used to establish surrogate models for eight benchmark functions respectively. The formula of benchmark functions is shown in Table 2.

Table 2. Formula of benchmark functions.

Function	Dim	Expression
Alpine	2	$f(x) = sin(x_1) sin(x_2) \sqrt{x_1 x_2}$, where $x_{1,2} \in [0,10]$.
Cross in Tray	2	$f(x) = -0.0001 \left(\left\| sin(x_1) sin(x_2) exp\left(\left\| 100 - \frac{\sqrt{x_1^2 + x_2^2}}{\pi} \right\| \right) \right\| + 1 \right)^{0.1}$, where $x_{1,2} \in [-10,10]$.
DropWave	2	$f(x) = -\frac{1 + cos\left(12\sqrt{x_1^2 + x_2^2}\right)}{0.5(x_1^2 + x_2^2) + 2}$, where $x_{1,2} \in [-5.12, 5.12]$.

$$f(x) = \sum_{i=1}^{5} c_i \exp\left(-\frac{1}{\pi}\sum_{j=1}^{2}(x_j - A_{ij})^2\right)\cos\left(\pi\sum_{j=1}^{2}(x_j - A_{ij})^2\right),$$

Lanfermann   2

$$\text{where } x_{1,2} \in [0,10], \ A = \begin{pmatrix} 3 & 5 \\ 5 & 2 \\ 2 & 1 \\ 1 & 4 \\ 7 & 9 \end{pmatrix}, c = \begin{pmatrix} 1 \\ 2 \\ 5 \\ 2 \\ 3 \end{pmatrix}$$

Levy3   3

$$f(x) = \sin^2(\pi\omega_1) + \sum_{i=1}^{d-1}(\omega_1 - 1)^2[1 + \sin^2(\pi\omega_1 + 1)] +$$
$$(\omega_d - 1)^2[1 + \sin^2(2\pi\omega_d + 1)],$$

$$\text{where } x_i \in [-10,10], \omega_i = 1 + \frac{x_i - 1}{4}, for \ all \ i = 1, 2, \dots, D.$$

Shekel4   4

$$f(x) = -\sum_{i=1}^{m}\left(-\sum_{j=1}^{4}(x_j - C_{ji})^2 + \beta_i\right)^{-1},$$

$$\text{where } x_j \in [-2,2], \ j = 1, \cdots, D, \ C = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \end{pmatrix}, \beta = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.4 \\ 0.6 \end{pmatrix}$$

StyblinskiTang5   5

$$f(x) = \frac{1}{2}\sum_{i=1}^{d}(x_i^4 - 16x_i^2 + 5x_i)$$

$$\text{where } x_i \in [-5,5], for \ all \ i = 1, 2, \dots, D$$

Schwefel6   6

$$f(x) = 418.9829d - \sum_{i=1}^{d} x_i \sin\left(\sqrt{|x_i|}\right),$$
$$\text{where } x_i \in [-500,500], for \ all \ i = 1, 2, \dots, D$$

Both training and test sets were sampled by the LHS method, and 10 trials were run to eliminate the effect of random sampling on the results. The specific design of the experimental design is shown in Table 3. This paper selected the linear trend and the exponential correlation function to construct an ordinary Kriging model, the parameter θ is 10. Additionally, the basis function of the RBF model is the cubic function.

Table 3. Test function experiment design.

Dimension	Trial times	Initial sampling	Update sampling	Number of test
2	10	5*dim	8*dim	1000
3-6	10	5*dim	6*dim	5000

4.2 Results of test functions

The results of the four 2D benchmark functions are shown in Figures 3 to 6, and the means of RMSE for 10 trials are shown in Table 4. Each graph contains four subgraphs. The subgraph (a) refers to the true function figure; the subgraph (b) is the notched boxplot; the subgraph (c) is the line chart; the subgraph (d) is the mean bar chart with specific values. The middle line of the notched boxplot represents the median, and the length of the box reflects the degree of RMSE fluctuation to a certain extent. The smaller the RMSE, the higher the prediction accuracy of the model.

Table 4. Results of Means and standard deviation of RMSE.

Functions	Kriging	AHM-CVH	RBF
Alpine	3.8019/ 0.4295	**3.3945/ 0.3199**	4.6165/ 0.6452
Cross in Tray	0.2708/ **0.0237**	**0.2440/** 0.0388	0.2897/ 0.0368
DropWave	0.2365/ 0.0658	**0.1643/ 0.0220**	0.2116/ 0.0585
Lanfermann	1.3296/ **0.3451**	**1.1166/** 0.0970	1.6175/ 0.4438
Levy3	53.0577/ **3.9070**	**47.4220/** 6.7056	49.2320/ 6.8529
Shekel4	0.2600/ **0.0190**	**0.2159/** 0.0237	0.2330/ 0.0632
StyblinskiTang5	172.2294/ 10.9866	**161.3217/** 11.4607	163.9585/ **7.3009**
Schwefel6	**1117.6252/ 49.1597**	1130.6589/ 85.2069	1415.1231/112.1433

(a)

(b)

(c)

(d)
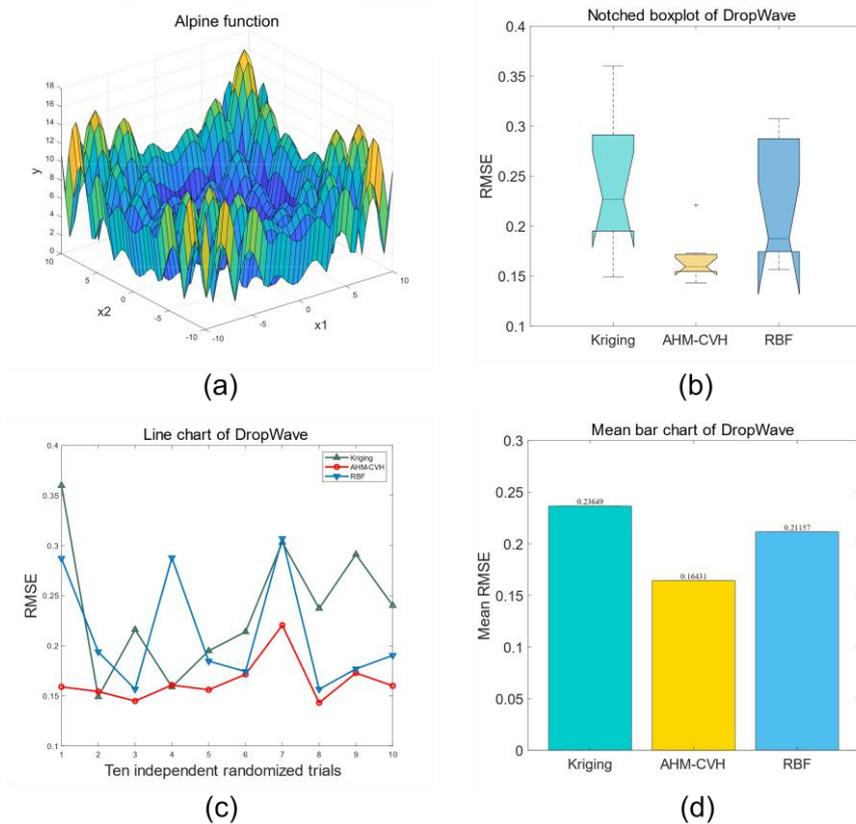
Figure 3. Trial results for the Alpine function. (a) Alpine function, (b) Boxplot, (c) Results of ten replicates of the experiment. (d) Means of RMSE.

Figures 3-6 illustrate the results of the 2D test function, and Figure 7 shows the test results for other dimensions (3 - 6D). These test functions are multimodal and have multiple local extrema. It can be observed that the AHM-CVH model outperforms the single model almost every time for ten replicates (as show in Figure 3 (c), Figure 5 (c) and Figure 6 (c)). Meanwhile, in all 2D test functions, the mean RMSE of the AHM-CVH model outperforms other single functions. In addition, as shown in Figure 7, in the test functions of other dimensions, the performance of the AHM-CVH model is inferior to the single model only under the Schwefel6 function. At this time, as shown in Table 4, the Kriging model has the best performance with an RMSE of 1117.6252, while the RMSE of AHM-CVH is 1130.6589. Nevertheless, the RMSE of the AHM-CVH model is not significantly different from that of the Kriging model, and even better than the single RBF model. In addition, it can be noticed that the range of the boxplot of the AHM-CVH model is

smaller than that of the single Kriging model and the RBF model, which shows that the method has better stability.
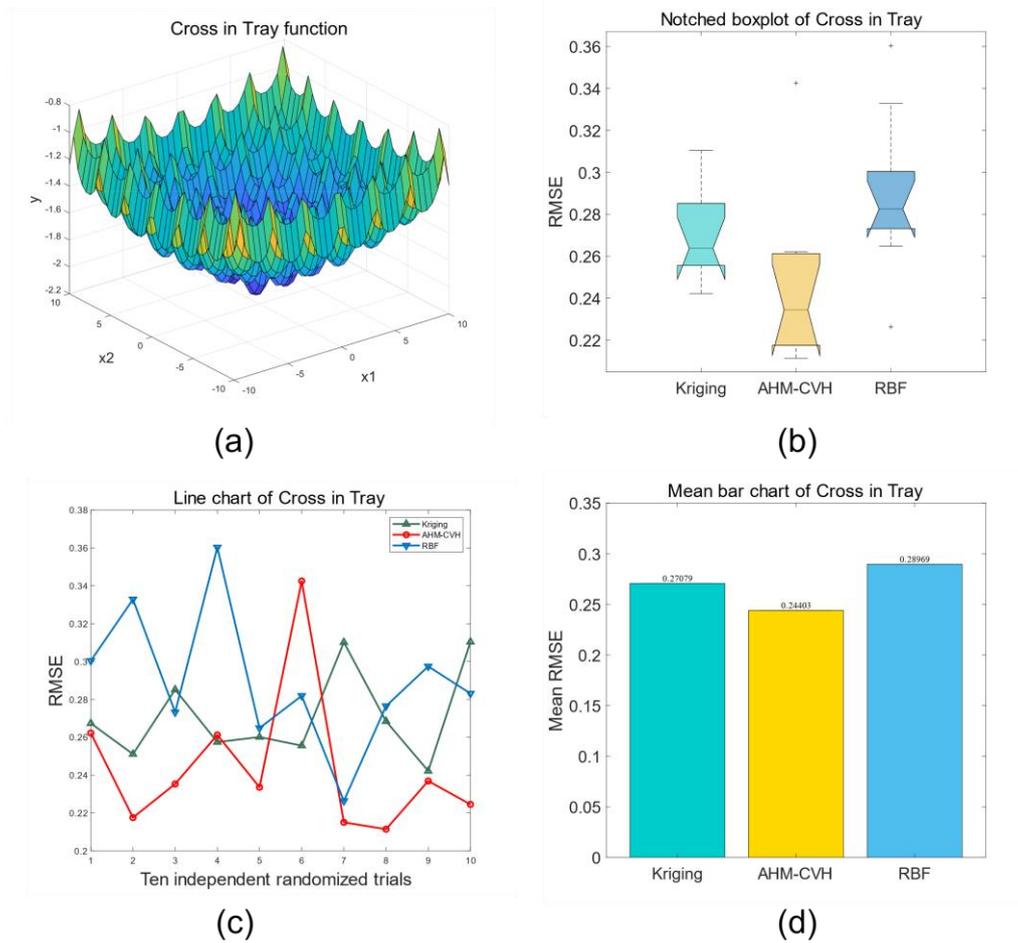

Figure 4. Trial results for the Cross in Tray function. (a) Cross in Tray function, (b) Boxplot, (c) Results of ten replicates of the experiment. (d) Means of RMSE.

Therefore, compared with the single Kriging model and the RBF model, in most cases, the proposed AHM-CVH model has better performance and generalization ability, which is of great significance.
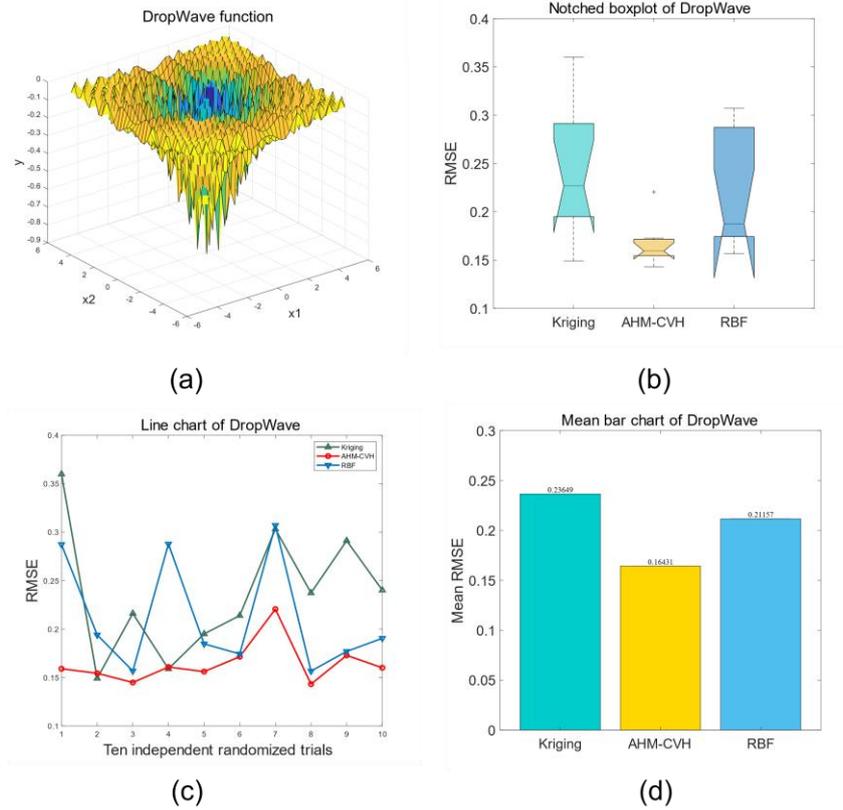
(a)

(b)

(c)

(d)

Figure 5. Trial results for the DropWave function. (a) DropWave function, (b) Boxplot, (c) Results of ten replicates of the experiment. (d) Means of RMSE.
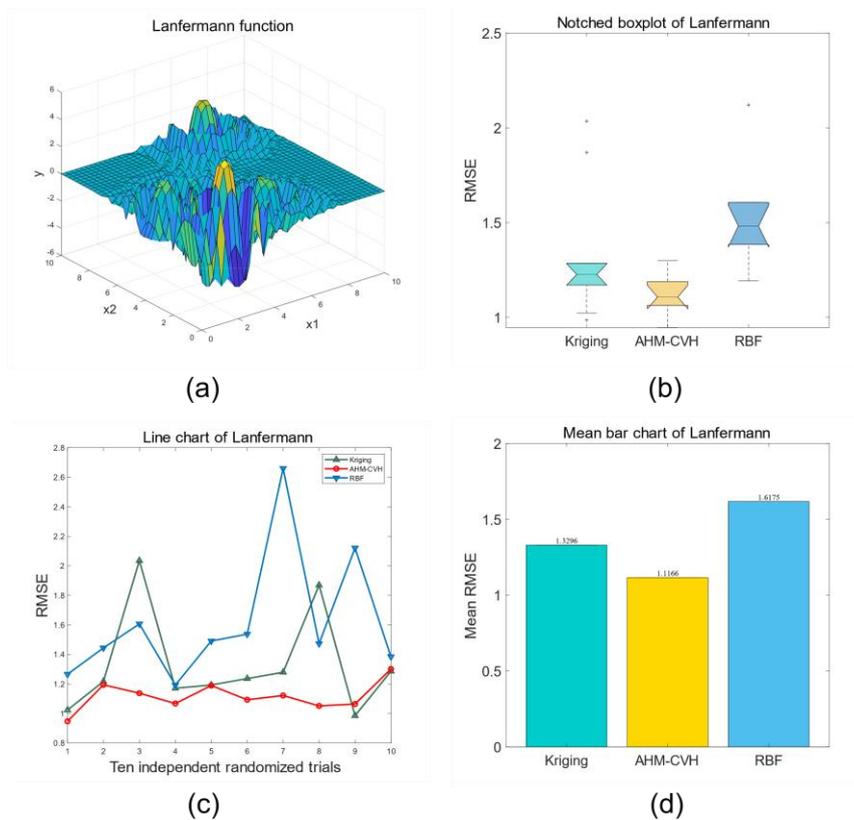

(a)

(b)

(c)

(d)

Figure 6. Trial results for the Lanfermann function. (a) Lanfermann function, (b) Boxplot, (c) Results of ten replicates of the experiment. (d) Means of RMSE.
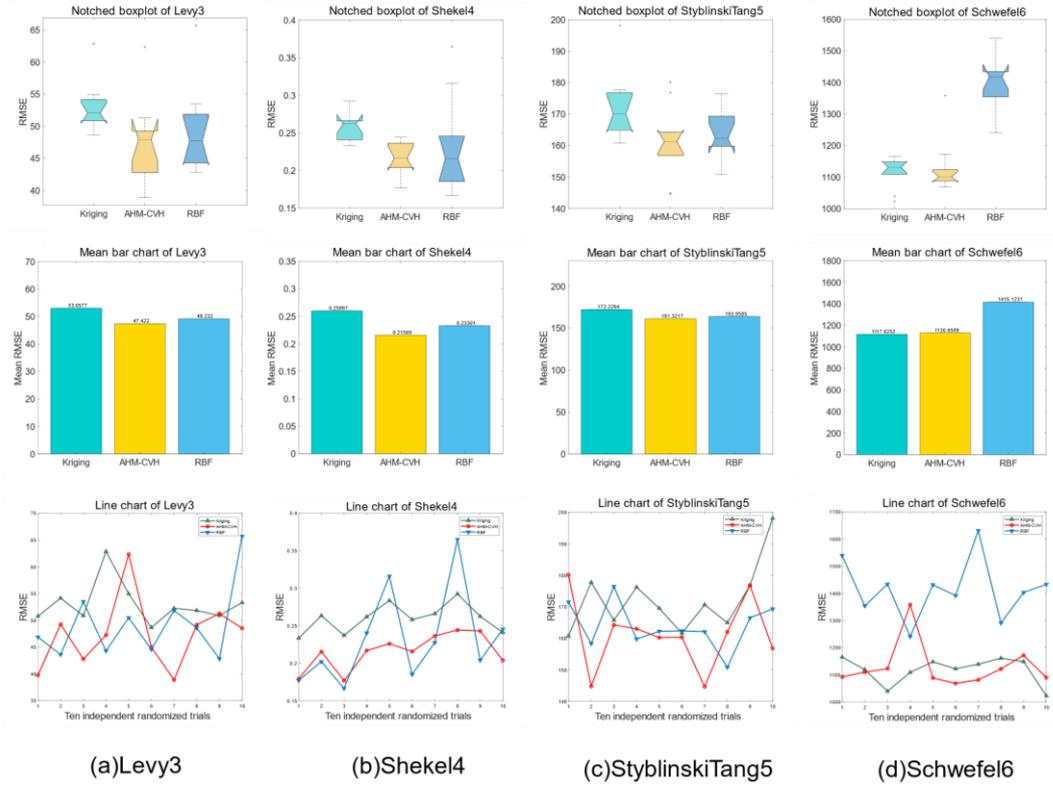
Figure 7. Test results for other dimensions. (a) Levy3 function, (b)Shekel4 function, (c) StyblinskiTang5 function, (d) Schwefel6 function.

4.3 Engineering example

An engineering example is used to verify the validity of the model, this example is a Helical tension cylindrical springs (HTCS) design problem. HTCS are made of wires with circular cross section, which are widely used in automobile, aerospace, construction machinery, elevators and other fields. A schematic diagram of the HTCS design is illustrated in Figure 8.
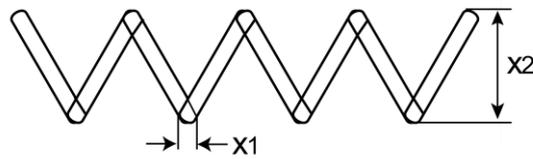

Figure 8. Diagram of the HTCS design

The HTCS design problem has three variables: x_1, x_2, and x_3. Where x_1 is the diameter, x_2 is the mean coil diameter, and x_3 is the number of active coils. Besides, the response function of HTCS is shown in formula (20).

$$f(x) = (2 + x_3)x_1{}^2 x_2. \tag{20}$$

where $x_1 \in [0.05, 2], x_2 \in [0.25, 1.3], x_3 \in [2, 15]$.

The proposed AHM-CVH model is compared with a single Kriging and RBF model, and the RMSE results are shown in Table 5.

Table 5. Results of RMSE.

Kriging	AHM-CVH	RBF
13.1411	**8.3552**	9.5571

As shown in Table 5, the RMSE value of the proposed AHM-CVH model is 8.3552, which outperforms the single Kriging and RBF models. This shows that the AHM-CVH model does have superior performance and is suitable for practical engineering problems.

CONCLUSION

This paper proposed an adaptive update sampling strategy suitable for any surrogate model, and then the Kriging model and the RBF model are ensembled to obtain the AHM-CVH model based on CVH strategy. The effectiveness of AHM-CVH is verified through eight benchmark functions. Finally, the results show that compared with the single kriging model and the RBF model, the proposed AHM-CVH model performs well on most of the test problems, and the generalization ability is significantly better than the single approximate model. At the same time, the model has high accuracy, strong stability and low sensitivity to experimental design. However, the AHM-CVH model does not perform well on all problems. However, it must be mentioned that the proposed AHM-CVH model is based on Kriging and RBF models for ensemble. The existing problems mainly include that with the increase of the function dimension, the modeling time of the Kriging model will increase exponentially. This problem also caused the ensemble model to slow down considerably. In addition, it may affect the accuracy of the model. Consequently, future research can focus on Kriging and RBF ensemble models for high-dimensional problems, especially the time and accuracy of ensemble modeling.

ACKNOWLEDGEMENTS

REFERENCES

Gu, L. 2001. A Comparison of Polynomial Based Regression Models in Vehicle Safety Analysis. ASME Design Engineering Technical Conferences-Design Automation Conference, September.

Stefan, J., Michael, P., Johan, R. & Adam, W. 2010. A method for simulation based optimization using radial basis functions.Optimization and Engineering.11(4) 501-532.

Khuri, A.I. & Mukhopadhyay, S. 2010. Response surface methodology. Wiley Interdisciplinary Reviews: Computational Statistics. 2(2) 128-149.

Krige, D.G. 1952. A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy. 52 201-203.

Sacks, J., Welch, W.J., Mitchell, T. J. & Wynn，H. P. 1989. Design and Analysis of Computer Experiments. Statistical Science. 4 409-423.

van Beers, W.C.M. & Kleijnen, J.P.C. 2004. Kriging interpolation in simulation: a survey. Paper presented at the Simulation Conference. Proceedings of the 2004,   Winter.

Hardy, R.L. 1971. Multiquadratic Equations of Topography and Other Irregular Surfaces. Journal of Geophusical Research.76 1905-1915.

Buhmann, M.D. 2000. Radial basis functions. Acta numerica. 9(1) 1-38.

Gunn, S.R. 1997. Support Vector Machines for Classification and Regression, Technical Report, Image Speech and Intelligent Systems Research Group, University of Southampton, UK.

Cheng, B. & Titterington, D.M. 1994. Neural networks: A review from a statistical perspective. Statistical Science. 9(1) 2-54.

Shahsavania, D. & Grimvall, A. 2009. An adaptive design and interpolation technique for extracting highly nonlinear response surfaces from deterministic models. Reliability Engineering and System Safety. 94 1173-1182.

Li, X., Yang, H., Zhang, C., Zeng, G. M., Liu, Y. G., Xu, W.H., Wu, Y. & Lan, S.M. 2017. Spatial distribution and transport characteristics of heavy metals around an antimony mine area in central China. Chemosphere. 170 17-24.

Echard, B., Gayton, N. & Lemaire, M. 2011. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Structural Safety. 33(2) 145-154.

Michael, R., O'Lenick, C.R., Monaghan, A., Wilhelmi, O., Wiedinmyer, C.,Hayden, M. & Estes, M. 2019. Application of geostatistical approaches to predict the spatio-temporal distribution of summer ozone in Houston, Texas. Journal of exposure science & environmental epidemiolog. 29(6) 806-820.

Shi, L.J., Sun, B.B. & Ibrahim, D.H. 2019. An active learning reliability method with multiple kernel functions based on radial basis function.Structural and Multidisciplinary Optimization. 60(1) 211-229.

Li, Y.H., Shi, J.J., Shen, J.F., Cen, H. & Chao, Y.P. 2020. An adaptive Kriging method with double sampling criteria applied to hydrogen preparation case. International Journal of Hydrogen Energy. 45(56) 31689-31705.

Jin, R., Chen, W. & Simpson, T.W. 2001. Comparative studies of metamodelling techniques under multiple modelling criteria.Structural and Multidisciplinary Optimization. 23(1) 1-13.

Zhang, J., Yue, X.X., Qiu, J.J., Zhang, M.Y. & Wang, X.M. 2021. A unified ensemble of surrogates with global and local measures for global metamodelling. J. Engineering Optimization. 53(3) 1-22.

Zerpa, L.E., Queipo, N.V., Pintos, S. & Salager，J.L. 2005. An optimization methodology of alkaline-surfactant-polymer folding processes using field scale numerical simulation and multiple surrogates. Journal of Petroleum Science and Engineering. 47 197-208．

Goel, T., Hafika, R.T., Shyy, W., Queipo, N.V. 2007. Ensemble of surrogates. Structural and Multidisciplinary Optimization. 33(3) 199-216.

Acar, E. & Rais-Rohani, M. 2009. Ensemble of metamodels with optimized weight factors. Structural and Multidiseiplinary Optimization. 37(3) 279-294.

Acar, E. 2010. Various approaches for constructing an ensemble of metamodels using local measures. Structural and Multidisciplinary Optimization, 42(6) 879-896.

Chen, L.M., Qiu, H.B. Jiang, C., Cai, X.W. & Gao, L. 2018. Ensemble of Surrogates with Hybrid Method Using Global and Local Measures for Engineering Design. Structural and Multidisciplinary Optimization. 57 (4) 1711–1729.

Cheng, K. & Lu, Z.Z. 2020. Structural reliability analysis based on ensemble learning of surrogate models. Structural Safety. 83 101905.

Ye, Y. F., Wang, Z.X. & Zhang, X.B. 2020. An optimal pointwise weighted ensemble of surrogates based on minimization of local mean square error. Structural and Multidisciplinary Optimization. 62(4) 529–542.

Zhou, X.J., Ma, Y.Z. & Li, X.F. 2011. Ensemble of surrogates with recursive arithmetic average. Structural and Multidisciplinary Optimization. 44(5) 651-671.

Forrester, A.I.J. & Keane, A.J. 2009. Recent advances in surrogate-based optimization. Progress in Aerospace Sciences. 45(1–3) 50–79.