Improved Kriging model based on weighted self-adaptive differential evolution and application for structural monitoring data

Junlong Wang¹ and Hongyue Liu¹, *

¹School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China
*Corresponding author: lhy320208@shu.edu.cn

Abstract. The integrity of structural health monitoring data plays an important role in extracting data features. An improved Kriging model based on weighted self-adaptive differential evolution algorithm (KWSDE) is proposed to repair the spatial missing data. In the KWSDE model, the control parameters of variogram in Kriging are optimized by a novel differential evolution (DE) with both scale factor and crossover rate adaptation. Besides, a weighted least-square method is presented for the optimization process to restrain the variable scale of the fitness function, which makes it more suitable for the lack of monitoring data. Finally, a distributed optical fiber monitoring experiment is designed to verify the reliability and functionality of the proposed model. Compared with the Kriging based on DE model (KDE), the interpolation precision of the KWSDE model can be improved by more than 10%, which has great significance for its engineering application.

1. Introduction

In the field of structural health monitoring, there may appear loss of spatial continuity and autocorrelation data due to external random disturbances, input signals saltation and other harsh environment [1]. It is the major direction of this research to timely discover and repair those missing data. Numerous researches have been carried out to figure out this problem. A multi-stage approach was established for structural damage identification by using modal strain energy and evolutionary optimization technique [2]. 1-D convolutional neural networks (CNNs) had achieved state-of-the-art performance in vibration-based structural damage detection [3]. The noise associated with Fiber Bragg Grating (FBG) sensor signal could be reduced by adopting the block-level-thresholding wavelet denoising method [4]. An anomaly data detection method of computer vision and deep learning-based was proposed to identify and clear abnormal data [5]. The deep neural networks were applied into a 33 km optical fiber sensing system to recognize and classify the signals of the external intrusion events [6]. The shrinking effect of fiber at high temperature could be avoided by annealing the fiber [7]. The monitoring data is reconstructed together with the spatial position information on the monitoring object so that it indeed belongs to the category of spatial data. As a typical spatial data interpolation method, Kriging has great performance in processing such spatial data.

Many scholars have studied improvement measures to optimize the accuracy of Kriging interpolation. Recent years have witnessed a spurt of progress in artificial intelligence, the chaotic ant-lion algorithm (CALO) was employed to seek suitable parameters of the variogram to improve spatiotemporal Kriging [8]. An adaptive Kriging by using a refined U learning function was utilized to
enhance the reliability of Kriging meta-model [9]. Furthermore, the optimization efficiency of the system was realized by using L-SHADE algorithm, in which Kriging was as the adaptive surrogate model [10]. In addition, scholars have proposed other related fitting methods of variogram such as the maximum likelihood method, linear programming method, genetic algorithm and so on [11]. Although all the above algorithms can maintain the diversity of solutions to the beginning of search progress, and accelerate the convergence speed at the end of the execution, they have limitations of so many initial setting parameters.

To effectively process the missing data of structural health monitoring due to artificial factors or harsh external conditions, an improved Kriging repair model based on a weighted self-adaptive DE algorithm (KWSDE) is brought forward. The new adaptive parameters of scale factor and crossover rate are utilized to avoid premature iteration and speed up convergence. Besides, the proposed weighted least squares method is mainly to constrain the range of variables in objective function. Moreover, the KWSDE model is applied to repair structural monitoring data based on distributed optical fiber. The experimental results demonstrate that the precision of the KWSDE model is improved by more than 10% compared with the KDE model.

2. The KWSDE model

Ordinary Kriging is a spatial interpolation method based on the spatial analysis of variogram to estimate the regional variation without bias. In geostatistics, spatial correlation is represented by variogram, which assumes that the mean value of localized second-order stationary random vector is constant and the covariance between random variables depends on the distance between sampling points. The variation function \( \gamma(h) \) of regional variation \( Z(x) \) can be expressed as follows:

\[
\gamma(h) = \frac{1}{2} \text{var}[Z(x + h) - Z(x)].
\]

(1)

The existed nugget effect in Eq.(1) helps maintain the conditions of covariance matrix. Therefore, as shown in Figure 1, the parameters of nugget value \( C_0 \), abutment value \( C \) and variation value \( a \) will determine the fitting precision of variogram.

![Figure 1. Nugget effect](image)

As a powerful and effective stochastic global optimization algorithm, the DE algorithm has been proven the quality to optimize these nugget effect parameters. The basic operations of this algorithm over the generations mainly include the initialization, mutation, crossover and selection.

There should be a certain range about the parameters \( C_0^*, C^* \) and \( a^* \). The values of parameters are initialized within the range of lower and upper bounds as below:

\[
\begin{align*}
C_0^*_{ij} &= \text{rand}_{ij}(0,1) \cdot (C_{0\text{min}} - C_{0\text{max}}) \\
C^*_{ij} &= \text{rand}_{ij}(0,1) \cdot (C_{\text{min}} - C_{\text{max}}) \\
a^*_{ij} &= \text{rand}_{ij}(0,1) \cdot (a_{\text{min}} - a_{\text{max}})
\end{align*}
\]

(2)

where \( i \) belongs to [1, NP], and \( j \) is [0, D]. The \( \text{rand}_{ij}(0,1) \) represents random numbers evenly distributed on the interval (0, 1).

In the DE algorithm, the control parameters \( F \) and \( CR \) constrain the evolution direction and iteration speed. The scale factor \( F \) is used to control amplification of two different individuals typically
in the range of $[0, 2]$. A larger $F$ increases the probability of escaping from the local optimum. Obviously, an appropriate value of $F$ can achieve a global optimum in a short time. The self-adaptive $F$ can be expressed as follows:

$$F = F_0 * 2^{\cos \left( \frac{H_m}{H_m + \pi/2} \right)}$$  \hspace{1cm} (3)

where $F_0$ is the initial scale factor, and the new parameter $\alpha$ ensures that the value of scale factor gradually decreases from $2F$ to $F$. $H_m$ denotes the maximum generation, and $H_i$ represents current evolutionary of generation. The cosine function guarantees the value of $\alpha$ increases monotonically in the interval $[0, 1]$.

Similar to the $F$, the constant $CR$ in crossover operation can be dynamically adjusted as shown in Eq.(9) according to the population diversity

$$CR = CR_{min} * 2^{\sin \left( \frac{H_i}{H + \pi/2} \right)}$$  \hspace{1cm} (4)

where $CR_{min}$ is the initial crossover probability. The sine function guarantees that the power of the exponential increases monotonically in the interval $[0, 1]$. The CR in iteration process will monotonously increase in range of $[CR_0, 1]$. By this way, we attempt to maintain both exploitation (with small values) and exploration (with large values) power throughout the entire evolution process.

The evaluation of fitness function is mainly aimed at multivariate optimization based on variogram, which satisfies the minimum principle. One better minimized fitness value of the individuals will be favored for next generation. The specific selected operation can be expressed as follows:

$$s_i(g + 1) = \begin{cases} u_i(g), & \text{if } f(u_i(g)) \leq f(s_i(g)) \\ s_i(g), & \text{otherwise} \end{cases} \hspace{1cm} i = 1, 2, ..., D, \hspace{1cm} (5)$$

where $f(x)$ denotes the fitness function. The multi-parameter optimization problem of variogram can be established as the issue of discovering the optimal design variables to optimize $F(x) = [f_1(x), f_2(x), ..., f_k(x)]$, where $F(x) \in \mathbb{R}^k$ is target function. Traditional objective moderation function of variogram is the sum of the squares of the difference between the test value $\gamma(h)$ and the theoretical value $\gamma(h)$ of the variogram. An improved coefficient based on weighted least square method is established to optimize the objective function. The new fitness function $f_{new}$ is as follows:

$$f_{new} = \sum_{i}^{n} \lambda_i \left[ \gamma(h_i) - \gamma^*(h_i) \right]^2, \hspace{1cm} (6)$$

where $\mu_i$ represents the weight coefficient essentially affected by the lag distance $h_i$, the logarithm $N_i$ of the sample points and the absolute value of variogram of sample points. The weighted factor $\lambda_i$ is as follow:

$$\lambda_i = \frac{h_i}{h \gamma^*(h_i)} \hspace{1cm} (7)$$

where $\bar{\gamma}(h)$ denotes $\frac{1}{n} \sum_{i=1}^{n} \gamma^*(h_i)$. $h$ is the average lag distance. $N$ is the sum of sample logarithms.

The cross-validation method is utilized to verify the accuracy of KWSDE interpolation. The test standard is to select the nearest sample points according to the calculation of the root-mean-square error (RMSE) and mean-absolute error (MAE) between Kriging estimation value $\hat{y}_i$ and actual monitoring value $y_i$. The smaller the values of RMSE and MAE are, the better the prediction effect is.

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2}, \hspace{1cm} (8)$$

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^{m} |y_i - \hat{y}_i| \hspace{1cm} (9)$$

The updated $f_{new}$ is inputted into the selection operation (Eq.(6)) until termination criterion is satisfied. Once the optimal solution of the parameters is generated, we can obtain the theoretical function of variogram. Then, the accurate estimated value can be calculated by Kriging interpolation. When the above work is done, an improved Kriging model based on the weighted self-adaptive DE algorithm is formed. In conclusion, the flowchart of KWSDE model is shown in Figure 2.
3. Experimental results and discussion

In order to monitor the damage caused by non-metallic materials of automobile spoilers through thermal expansion, the real-time monitoring of high temperature thermal cycling was designed as the experimental condition. The strain cloud diagram of finite element thermal analysis was shown in Figure 3-(a). The phenomenon of strain concentration occurred around the edge and groove of the test piece. As reported by this finite element analysis result and structural dimensions of test piece, the distributed optical fiber layout was designed to monitor the deformation of the strain concentration position (Figure 3-(b)). Optical fiber was attached to the upper surface of structure to be tested. The effective data monitoring part of the sensor was from start point A to terminal point W.

Figure 3. The test piece: (a) the finite element thermal analysis diagram; (b) distributed fiber sensor layout

Figure 4. Thermal deformation system monitoring platform
Figure 4 exhibits the structure of thermal deformation monitoring system. The distributed optical fiber demodulator based on OFDR technology was LUNA ODiSI-B. The non-metallic experimental material of bonded optical fiber was placed in an electric constant temperature drying oven. It’s easier to collect enough data with more loss when spatial resolution is larger. The drying oven was used to simulate the environment of the test piece, whose cycle temperature was set from -40 °C to 120 °C. Once the fiber demodulator started collecting data, the whole system kept running continuously.

Due to artificial interference factors, the collected initial data can generate enough data loss, which facilitates the verification of the KWSDE model. Here, the spatial position information of the sampling point is the input independent variable of Kriging interpolation. The experimental data of one random sample moment is selected as test set. Figure 5 is the visual repair performance of monitoring data from a random sampling point, which retains vibration data characteristics at the same time. According to Table 1, the MAE and RMSE of the KWSDE model are decreased by 12.54% and 28.61% respectively, while the fitting accuracy is increased by 10.71% in the thermal deformation experiment.

| MAE  | RMSE  | $R^2$ |
|------|-------|-------|
| KDE  | 11.5124 | 17.4810 | 0.8396 |
| KWSDE | 10.0687 | 14.4912 | 0.9467 |
| (KWSDE – KDE)/KDE | 12.54% | 28.61% | 10.71% |

Figure 5. Processing effect of random sampling point

Figure 6 exhibits intuitive performance of the repaired results, which are respectively repaired by the KDE model and the KWSDE model. The position 4646–4664.1 mm of fiber length is the reference temperature of the environment corresponds to point G to point H in Figure 3-(b), and the KWSDE model retains this characteristics of the monitoring data well. Compared with the repair effect of KDE model (Figure 6-(b)), the interpolation effect of the proposed KWSDE model corresponding to Figure 6-(c) is significant better, which indicates the high accuracy and feasibility. In summary, according to the monitoring data verification of thermal-cycling experiment based on the optical fiber technology, the results fully prove the feasibility and accuracy of the proposed model.
4. Conclusions
In order to solve the data loss problem of spatial auto-correlation and continuity in the field of structural monitoring, an improved Kriging model based on weighted self-adaptive differential evolution algorithm (KWSDE) is proposed. The significant modification of the DE algorithm is focused on avoiding premature convergence by introduction of the self-adaptive scale factor and crossover rate. It not only enhances local the convergence but also guarantees the global diversity of the algorithm at the end. What’s more, the original fitness function is optimized by a novel weighted least square method to represent the characteristics of monitoring data. Finally, the experimental results demonstrate that the repair precision of the proposed model is greatly enhanced by more than 10%. In general, the KWSDE model has certain reference and engineering application for the repair research of structural monitoring data.

References
[1]. Bonessio N, Lomiento G and Benzoni G, “Damage identification procedure for seismically isolated bridge,” Structural Control and Health Monitoring, 19(5):565–578, 2012.
[2]. Srinivas V and Ramanjaneyulu K, “Multi-stage approach for structural damage identification using modal strain energy and evolutionary optimization techniques,” Structural Health monitoring – An International Journal, 10:5120-5131, 2010.
[3]. Abdeljaber O, Avci O and Kiranyaz M.S, “1-D CNNs for structural damage detection: Verification on a structural health monitoring benchmark data.” Neurocomputing, 275(31): 1308-1317, 2018.
[4]. Allan C.L. Wong and Paul A. Child. “Multiplexed fiber Fizeau interferometer and fiber Bragg grating sensor system for simultaneous measurement of quasi-static strain and temperature using discrete wavelet transform”. Measurement Science and Technology, 17:384-394, 2006.
[5]. Bao YQ and Tang ZY, “Computer vision and deep learning-based data anomaly detection method for structural health monitoring,” Structural Health monitoring – An International Journal, 18: 401-421, 2009.
[6]. Yu Bai and Xing Jichuan, “Detection and identification of external intrusion signals from 33 km optical fiber sensing system based on deep learning.” Optical Fiber Technology, 53, 2019.
[7]. Kyriacos Kalli, Karen E. Carroll, Chi Zhang and J. David, Webb. “Thermal response of Bragg gratings in PMMA microstructured optical fibers,” Optics Express, 15:8844-8850, 2007.

[8]. Liu Y, Hu Y, and Wang H, “Improved Spatio-temporal Kringing and its Application to Regional Precipitation Prediction,” In IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), Metz, France, 2019.

[9]. Yun WY, Lu ZZ and Zhou YC, “AK-SYSi: an improved adaptive Kriging model for system reliability analysis with multiple failure modes by a refined U learning function,” Structural and Multidisciplinary Optimization. 59:263-278, 2019.

[10].Awad NH and Suganthan PN, “An improved differential evolution algorithm using efficient adapted surrogate model for numerical optimization,” Information Science, 451:326-347 2018.

[11].Yang X and Hodler T, “Visual and Statistical Comparisons of Surface Modeling Techniques for Point-Based Environmental Data.” Cartography and Geographic Information Science, 127(2):165-176, 2000.