RESEARCH ON GPS HEIGHT FITTING BASED ON ABC OPTIMIZED LSSVM

Zhang Yan,1,2, Huang Jingjin3, Tang Shihua1,2, Xing Pengwei1,2

1College of Geomatics and Geoinformation, Guilin University of Technology, 319 Yanshan Street, Guilin, Guangxi, 541006, China-2902502765@qq.com
2Guangxi Key Laboratory of Spatial Information and Geomatics, 319 Yanshan Street, Guilin, Guangxi, 541006, China-zy458974@163.com
3Geography National Monitoring Institute of Guangxi, No.5 Jianzheng Road, Nanning, Guangxi, 530023, China-L756419@163.com

Commission IV, WG IV/8

KEY WORDS: Height Fitting, Artificial Bee Colony, Least Squares Support Vector Machine, Regularization Parameter, Kernel Parameters

ABSTRACT:

Aiming at the problem that the fitting parameters of the least squares support vector machine fitting method are difficult to select, a method of introducing the artificial bee colony algorithm into the least squares support vector machine to establish a high-precision region fitting model is proposed. The artificial bee colony algorithm can perform global tracking search on the parameters in the least squares support vector machine, imitate the honey collecting process of the bees, and use the primary value of the parameters as the honey source, and the average square error predicted by the least squares support vector machine as the target. The function value is determined by iterative update within a certain range to determine the optimal parameters, and finally a GPS height fitting model is proved, which provides a certain reference value for the establishment of GPS height fitting model.

1. INTRODUCTION

With the rapid development of modern measurement technology, the application of GPS technology in surveying and mapping work has been widely promoted, but GPS directly measures the earth's height, while the actual engineering commonly used is the normal high, there is an elevation abnormal value before the two(Tan, L. P., 2013; Zuccari, C., 2013). Therefore, how to achieve high-precision elevation anomaly values is a problem to be solved in GPS measurement work. At present, intelligent algorithms such as particle swarm optimization, genetic algorithm and artificial neural network are often applied to the construction of GPS height fitting model, and the accuracy of the elevation anomaly value to be sought is further improved(Liu, L. L., 2014). Liu, et al (2015) has proposed to use the genetic algorithm to optimize the RBF neural network to achieve the purpose of global search for the optimal radial basis function center value, so that the fitted model can better predict the elevation. Liu, et al (2016) has used the idea of “moving-fitting-recovery” to introduce a method based on the EGM2008 model and considering the influence of topographic factors on the quadratic surface fitting GPS elevation, and the accuracy of the model is improved. Pu, et al (2018) has proposed the particle swarm optimization algorithm to optimize the traditional multi-face function fitting method, and further improved the search efficiency of eigenvalues in multi-faceted functions, and provided reasonable necessary fitting parameters for the construction of the fitting model. The least squares support vector machine (LSSVM) can achieve a good height fitting effect by using a small number of data samples. The key point is the regularization parameters c and kernel parameters σ in the least squares support vector machine. The optimization of parameters in LSSVM fitting method often uses cross-validation method, but it has the defects of large calculation and long time. It is often difficult to select the optimal parameter value, and the accuracy of fitting model is difficult to reach the expected value. Finding the right method to extract the best parameter values can improve the accuracy of the fitted model (Lin, D., 2019; Shao, Y. H., 2019). Compared with other intelligent algorithms, artificial bee colony (ABC) has the characteristics of less parameters, simple calculation and global search for optimal values (Hajimirzaei, B., 2019). This paper proposes to use artificial bee colony algorithm to optimize least squares support vector. The GPS height fitting method of the machine can use the artificial bee colony algorithm to quickly and effectively find the advantage of the optimal value, find the optimal parameter value for the least squares support vector machine within the specified range, and finally use a small amount of GPS level to coincide. Point to build a GPS height fitting model.

2. BASIC THEORY

2.1 Artificial Bee Colony Algorithm

The artificial bee colony optimization algorithm is a bio-intelligence optimization algorithm that mimics the process of

* Corresponding author: Huang Jingjin Email:L756419@163.com
bees searching for honey (Gergin, Z., 2019). A bee colony is often divided into three types: lead bee, accompanying bee, and scout bee depending on the work in the group. The work of guiding the bee is to transfer and share the information of the discovered food source to other partners waiting in the bee. The information transmitted includes the direction, distance and probability of existence of the honey source (Bajer, D., 2019; Wang, S. W., 2019). After receiving the message leading the bee, the accompanying bee tries to find the specific location of the honey source, and the detection bee is looking for a new source of honey. In the honey source search work, the lead bee and the accompanying bee each account for half of the bee colony, and the whole process can be understood as the process of obtaining the optimal value of the objective function (Szczepanski, R., 2019). The specific principles are as follows: The ABC first randomly generates an initial value \( x_i \) \((i=1,2,\ldots,n)\) similar to the optimal position of the honey source, and \( n \) is the number of bee swarms. The size of the space to be solved is consistent with the range of honey source, and \( n \) is the number of bee swarms. The size of the range, and each accompanying bee decides whether to follow the leading bee around the initial source of honey, and the new source of honey is not equal to the initial source of honey. As shown in formula (1):

\[
x_i = m_i + \text{rand}(0, 1)(m_i - M_i)
\]

where \( x_i \) = initial value of the honey source location \( m_i \) = minimum value of the search range \( M_i \) = maximum value of the search range

Among them, the initial value of the honey source is randomly distributed in the range of \((m_i, M_i)\). The bee is led to find a new source of honey around the initial source of honey, and the new source of honey is not equal to the initial source of honey. As shown in formula (2):

\[
X_i = x_i + \varphi(X_i - x_i)
\]

where \( \varphi \) = random number between \([-1,1]\] \( X_i \) = new honey source \( x_i \neq x_i \) \( j=1,2,\ldots,n \)

Use the greedy algorithm to evaluate the fitness of the two honey sources and decide to update or retain the original honey source. As shown in formula (3):

\[
\text{fit} = \begin{cases} 
1 / (1 + f_i), & f_i \geq 0 \\
1 + \text{abs}(f_i), & \text{other}
\end{cases}
\]

where \( f_i \) = objective function value

A further step is needed to calculate the probability that the accompanying bee will follow the lead to find the source of honey:

\[
p_i = \text{fit} / \sum_{i=1}^{n} \text{fit}_i
\]

where \( p_i \) = Probability

The accompanying bee joins to further narrow down the search range, and each accompanying bee decides whether to follow the lead bee according to a random number \( a \) generated in \([0,1]\). When \( p_i > a \), the accompanying bee will follow the formula (2) Generate a new source of honey, use the greedy algorithm to determine the honey source that should be retained. If a better source of honey has not been found after a limited number of iterations, then turn it into a detection bee and seek a new source of honey according to formula (1). Give up, make a judgment and keep the best value.

### 2.2 Least Squares Support Vector Machine

The least squares support vector machine is proposed by Suykens on the basis of the original support vector machine (Baghban, A., 2018). The support vector machine has the characteristics of simple structure, global optimization and generalization ability. LSSVM also improves these advantages (Sedaghat, M., 2018). Convergence speed and simplification of the calculation process. The specific steps are as follows:

For a given training set:

\[
A = \{ (x_i, y_i), \ldots, (x_l, y_l) \} \in \mathbb{R}^n \times \mathbb{Y}^l
\]

where \( A \) = training set \( x_i \in \mathbb{R}^n \) \( y_i \in \mathbb{Y} = \mathbb{R} \)

The nonlinear regression map is used to map the sample input to the high-dimensional feature space, and the linear regression function of the high-dimensional feature space is constructed:

\[
g(x, w) = w^T \varphi(x) + b
\]

where \( w \) = weight vector \( b \) = the threshold

Based on the principle of structural risk minimization, it is transformed into the following optimization problems:

\[
\min_{w, \xi, \xi} \frac{1}{2} \|w\|^2 + \frac{c}{2} \sum_{i=1}^{n} \xi_i^2
\]

where \( c \) = the regularization parameter \( \xi \) = the errors

The function model of LSSVM can be obtained by using Lagrange function and KKT (Karush-Kuhn-Tucker) optimization condition:

\[
Z(x) = a_i \left[ \varphi(x_i) \varphi(x) \right] + b
\]

According to the Hilbert-Schmidt principle, the kernel function (see Formula 9) is introduced, and the calculation of the inner product of the high-dimensional space is converted into the calculation of the function in the original space, which solves the calculation problem of the high-dimensional feature space, namely:

\[
Z(x) = \sum_{i=1}^{n} a_i K(x_i, x) + b
\]

In this paper, the radial basis kernel function is chosen as the kernel function of LSSVM to meet the GPS fitting modeling requirements, namely:

\[
K(x, x_i) = \exp\left(- \frac{\| X - x \|^2}{2\sigma^2} \right)
\]
It can be seen from the above process that using the least squares support vector machine to construct the fitting model only needs to determine the kernel function parameter $\sigma$ and the regularization parameter $c$. Compared with the original, the calculation process is much simpler.

3. ABC OPTIMIZED LSSVM

Compared with other intelligent algorithms, the artificial bee colony algorithm is not easily affected by its own parameters in the process of searching for the optimal value of the objective function, and can perform global search, avoiding the defects of falling into the local optimal value. The algorithm selects the optimal parameter value for the least squares support vector machine and establishes a higher precision GPS height fitting model. The specific steps of the ABC-LSSVM combination method are as follows:

1. normalize the collected height fitting data;
2. Initialization of ABC algorithm parameters: the size of the bee colony $N$, the number of honey sources $N/2$, the limit of the number of times the honey source is not updated, the maximum number of iterations $max$, the number of parameters to be optimized $D$, the search interval of the parameters;
3. Determine the initial value: randomly set the initial value within the effective range, and calculate the root mean square error (RMSE) of the LSSVM prediction corresponding to each initial value, and use it as the objective function;
4. Finding new sources of honey: lead the bees to use the formula (2) to find new sources $x_i$ of honey for each source of honey. And ensure that the selected adjacent honey source is not the current honey source and is not out of range, calculate the objective function value and the adaptation function value of the mutated honey source, compare with the current honey source, determine whether the current honey source needs to be updated, if the current honey source is not updated If the limit is exceeded, the honey source is abandoned and a new source of honey is sought;
5. Develop a new honey source again: lead the bee to pass the updated honey source information to the accompanying bee, and the accompanying bee will determine whether to update the new honey source again after the corresponding probability comparison. If the conditions for continued development are met, the accompanying bee The steps for redevelopment are the same as (4), and the optimal honey source (solution) is retained;
6. The number of iterations reaches the maximum value: When the number of iterations of the accompanying bee reaches the maximum value, no new better honey source is found, and the accompanying bee is converted into a detection bee, and the detection bee will generate a new honey source through formula (1);
7. It is judged whether the updated honey source satisfies the condition, and when the new honey source (solution) satisfies the final requirement, the optimal value is retained, otherwise it returns to step (4) to continue searching;
8. construct the GPS height fitting model by LSSVM: using the optimal parameter value of ABC optimization, substituting LSSVM to establish the height fitting model, and obtaining the internal and external matching accuracy of the fitted model through the test data, and conducting comparative analysis.

After several iterations to find the best value, ABC-LSSVM can perform global search to the optimal parameters, further speed up the convergence, and finally obtain the best fit model (See Figure 1 above).

4. EXPERIMENTAL ANALYSIS

4.1 Collection of research data

In order to verify whether the bee colony optimization least squares support vector machine fitting method is effective and feasible in constructing the GPS height fitting model, and the accuracy of the model constructed by the previous method is compared, the urban suburb of eastern-north of Guangxi is selected as the research area. The coverage area is about 400km$^2$, the terrain is relatively flat, mostly residential areas, the traffic is convenient, the GPS level coincidence point is convenient to collect, the point distribution is relatively uniform, the accuracy meets the research requirements, and the necessary basic data conditions are provided for the bee colony algorithm optimized least squares support vector machine fitting method. In the 600 GPS level coincidence points collected, the ant colony algorithm is used to select representative GPS level coincidence points, and 70 points with uniform distribution are used as check points (Seleuk, T., 2019). The optimized least squares support vector machine fitting method is used to establish a fitting model based on the same batch of GPS level coincidence points with the other three methods, and the fitting results are compared and analyzed (See Figure 2 above).
4.2 Fitting model accuracy analysis

When using the artificial bee colony algorithm to optimize the least squares support vector machine fitting method to construct the model, the initial value and the number of iterations need to be set, the bee colony population size is set to 20, and the honey source is not updated by the update number limit to 100. The maximum number of iterations is 80. Since the parameters that need to be optimized for the construction of the fitting model by LSSVM are only the regularization parameter \( c \) and the kernel parameter \( \sigma \), the dimension \( D \) of the problem to be solved is 2, the range of values is \((0.01, 50)\), and finally the fitting will be performed. The results were compared with the results of BP neural network fitting method, conventional LSSVM fitting method and particle swarm optimization algorithm (PSO). In order to preliminarily verify the stability and effectiveness of the fitted model, multiple experiments were performed using four different fitting methods. Table 1 shows the internal accuracy of the fitted model in 10 experiments.

| Number | BP/mm | LSSVM/mm | PSO/mm | ABC/mm |
|--------|-------|----------|--------|--------|
| 1      | 6.0   | 4.4      | 3.2    | 2.0    |
| 2      | 4.6   | 5.0      | 3.8    | 2.1    |
| 3      | 5.8   | 3.9      | 4.3    | 2.0    |
| 4      | 6.0   | 4.2      | 3.6    | 2.3    |
| 5      | 3.9   | 5.2      | 4.4    | 2.7    |
| 6      | 5.6   | 4.6      | 3.7    | 2.0    |
| 7      | 6.0   | 5.3      | 3.2    | 2.2    |
| 8      | 5.8   | 4.2      | 4.0    | 2.1    |
| 9      | 5.6   | 3.9      | 2.9    | 2.2    |
| 10     | 6.2   | 5.4      | 3.4    | 1.9    |
| 11     | 4.6   | 4.4      | 3.7    | 2.2    |
| 12     | 4.8   | 4.9      | 4.0    | 2.1    |
| **Average** | **5.4** | **4.6** | **3.7** | **2.1** |

Table 1. Comparison of internal precision of constructed model

The results of the table statistics show that in the 12 experiments, the mean accuracy of the model established by the artificial bee colony optimization is 2.1 mm, the average model accuracy value of the checkpoints in 12 experiments were counted. The accuracy of the fitted model in 10 experiments.

The results of 70 checkpoints, and the accuracy of the model was best fit model construction method was determined by comparing the average of all experiments with a single measurement (See Table 3 above).
Table 3. Comparison of detection accuracy of different methods

| Number | BP/mm | LSSVM/mm | PSO/mm | ABC/mm |
|--------|-------|----------|--------|--------|
| 1      | 18.6  | 19.2     | 14.4   | 10.8   |
| 2      | 19.4  | 19.7     | 15.6   | 10.6   |
| 3      | 19.2  | 18.9     | 14.2   | 10.8   |
| 4      | 18.5  | 18.6     | 16.1   | 10.4   |
| 5      | 18.8  | 19.9     | 15.7   | 10.7   |
| 6      | 18.6  | 19.5     | 16.6   | 10.3   |
| 7      | 19    | 19.6     | 14.4   | 10.6   |
| 8      | 18.6  | 19.5     | 15.4   | 10.4   |
| 9      | 17.9  | 19.9     | 15.6   | 10.6   |
| 10     | 18.4  | 19.3     | 15.8   | 10.7   |
| 11     | 18.9  | 18.8     | 14.4   | 10.6   |
| 12     | 18.4  | 19.4     | 15.2   | 10.5   |
| Average| 18.7  | 19.4     | 15.3   | 10.6   |

Statistical experimental data can be seen that whether it is the accuracy of each experiment or the accuracy of the 12 experiments, the artificial bee colony improved fitting method is better than the other three methods, the improved bee colony improved fitting method The average accuracy of the fitted model is ±10.6mm. On the basis of selecting the same set of test data, the overall fitting accuracy of the particle swarm optimization fitting method is ±15.3mm, and the bee colony optimization algorithm is increased by 30.7%. The fitting accuracy of the conventional LSSVM fitting method is ±19.4mm, and the accuracy of the bee colony optimization method is 45.4% higher. Compared with the BP neural network fitting method, the artificial bee colony optimization least squares support vector machine fitting method is constructed. The accuracy of the model is improved by 43.3%, which fully demonstrates that the artificial bee colony optimization fitting method can improve the accuracy of the GPS fitting model.

5. CONCLUSIONS

In order to solve the problem that the parameter selection is difficult when establishing the GPS height fitting model by the least squares support machine fitting method, the artificial bee colony algorithm is used to optimize the parameters of the least squares support vector machine fitting method, and through experiments and other methods. A comparative analysis was carried out and the conclusions were as follows:

1. The conventional least squares support vector machine fitting method often seeks to fit the model parameters through the cross-validation method, and this time, the artificial bee colony algorithm is used to optimize the parameters, and the internal and external conformity accuracy statistics table and residual trend charts of the multiple experiments are both It shows that the artificial bee colony algorithm optimization fitting method has better convergence effect, and it is more conducive to searching the optimal value efficiency within the effective search range;

2. Compared with the particle swarm optimization algorithm, the artificial bee colony algorithm increases the diversity of the organism through different division of labor, ensuring that the final acquisition result is globally optimal, avoiding the trouble of falling into the local optimal solution, and working in the optimization process. The constant conversion of the property improves the convergence speed, and provides the best fitting parameters for the least square support vector machine to construct the fitting model in a short time;

3. According to the experimental results, the model accuracy of the bee colony optimization fitting method is higher than that of the other three methods. The overall accuracy of the fitted model is ±10.6mm, which is 45.4% higher than the conventional least squares support vector machine fitting method and 43.3% higher than the BP neural network fitting method, which fully proves the stability and effectiveness of the artificial bee colony optimization fitting method.

The bee colony optimization least squares support vector machine fitting method proposed in this paper is fully validated in the establishment of the model, but the experimental range is limited. For a large range of fitted regions, the stability and effectiveness of the proposed model to construct a fitted model need to be further verified.

ACKNOWLEDGEMENTS

National Natural Science Foundation of China (41864002); Guangxi Key Laboratory of Spatial Information and Surveying and Mapping (16-380-25-25, 16-380-25-13); Guangxi Middle School and Young Teachers’ Basic Ability Enhancement Project (KY2016YB823);

REFERENCES

Tan, L. P., Li, Y., 2013: Application and Research of BM Model in the Elevation Fitting of the Express Way[J]. Advanced Materials Research,2013,2385.

Zuccari, C., 2013: Fitting Height and Diameter of Character Degree Graphs[J]. Communications in Algebra,2013,41(8).

Liu, L. L., Zhang, T. X., Zhou, M., et al, 2014: The Research of GPS Elevation Fitting Considering the Influence of Covariance Function[J]. Applied Mechanics and Materials,2014,3252.

Liu, J., Wang, Q. J., Wang, X. H., et al, 2013:Application of RBF neural network based on genetic algorithm in GPS height fitting[J].Science and Mapping Science,2013,38(02):143-145.

Liu, B., Guo, B. M., Shi, J. B., et al, 2016: GPS Elevation Fitting Using EGM2008 Model and Terrain Correction[J].Journal of Wuhan University (Information Science Edition),2016, 41(04):554-558.

Pu, L., Tang, S. H., Luo, J., et al, 2018: Analysis of GPS Height Fitting Accuracy Based on Particle Swarm Optimization[J].Journal of Surveying and Mapping Science and Technology,2018,3-5(06):563-568.

Lin, D., Liao, J., 2019: Wavelet kernel function based multiscale LSSVM for elliptic boundary value problems[J]. Neurocomputing,2019,356.

Shao, Y. H., Li, C. N., Huang, L. W., et al, 2019: Joint sample and feature selection via sparse primal and dual LSSVM[J]. Knowledge-Based Systems,2019.

Hajimirzaei, B., Navimipour, N. J., 2019: Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm[J]. ICT Express,2019,5(1).

Gergin, Z., Tunçbilek, N., Esnaf, S., 2019: Clustering Approach Using Artificial Bee Colony Algorithm for Healthcare Waste
Disposal Facility Location Problem [J]. *International Journal of Operations Research and Information Systems (IJORIS)*, 2019, 10(1).

Bajer, D., Bruno, Z., 2019: An effective refined artificial bee colony algorithm for numerical optimisation [J]. *Information Sciences*, 2019, 504.

Wang, S. W., Guo, X. P., Liu, J., 2019: An efficient hybrid artificial bee colony algorithm for disassembly line balancing problem with sequence-dependent part removal times [J]. *Engineering Optimization*, 2019, 51(11).

Szczepanski, R., Tarczewski, T., Grzesiak, L. M., 2019: Adaptive state feedback speed controller for PMSM based on Artificial Bee Colony algorithm [J]. *Applied Soft Computing Journal*, 2019, 83.

Baghban, A., Sasanipour, J., Habibzadeh, S., et al., 2018: Sulfur dioxide solubility prediction in ionic liquids by a group contribution — LSSVM model [J]. *Chemical Engineering Research and Design*, 2018.

Sedaghat, M., Rouhibakhsh, K., 2018: On prediction of asphaltene precipitation in different operational conditions utilization of LSSVM algorithm [J]. *Petroleum Science and Technology*, 2018, 36(16).

Selcuk, T., Alkan, A., 2019: Detection of microaneurysms using ant colony algorithm in the early diagnosis of diabetic retinopathy [J]. *Medical Hypotheses*, 2019, 129.

**APPENDIX**

| Numb er | Known elevation anomaly /m | LSSVM Residual /mm | BP Residual /m | Residual /mm | PSO Residual /mm | ABC Residual /m | Residual /mm |
|---------|---------------------------|--------------------|---------------|--------------|-----------------|-----------------|--------------|
| 1       | -20.993                   | -20.9639           | 29.1          | 23.9         | -20.9694        | 23.6            | -20.9815     | 11.5         |
| 2       | -20.8093                  | -20.8130           | -3.7          | 0.6          | -20.8278        | -18.5           | -20.8179     | -8.6         |
| 3       | -20.831                   | -20.8130           | 18.0          | 22.3         | -20.8287        | 3.2             | -20.8179     | 13.1         |
| 4       | -20.8585                  | -20.8269           | 31.6          | 27.6         | -20.8285        | 30.0            | -20.8412     | 17.3         |
| 5       | -20.711                   | -20.6905           | 20.5          | 22.0         | -20.6923        | 18.7            | -20.6999     | 11.1         |
| 6       | -20.5277                  | -20.5050           | 22.7          | 21.1         | -20.5035        | 24.2            | -20.5155     | 12.2         |
| 7       | -20.4387                  | -20.4214           | 17.3          | 18.9         | -20.4219        | 15.8            | -20.4289     | 9.8          |
| 8       | -20.2807                  | -20.2582           | 22.5          | 24.9         | -20.2639        | 16.8            | -20.2694     | 11.3         |
| 9       | -20.0178                  | -19.9834           | 34.4          | 31.2         | -19.9866        | 23.4            | -20.0041     | 13.7         |
| 10      | -19.954                   | -19.9486           | 5.4           | 16.7         | -19.9373        | 18.4            | -19.9427     | 11.3         |
| 11      | -20.537                   | -20.5209           | 16.1          | 17.5         | -20.5195        | 12.5            | -20.5288     | 8.2          |
| 12      | -20.4243                  | -20.3999           | 24.4          | 19.6         | -20.4047        | 21.4            | -20.4126     | 11.7         |

The mean error /mm: 20.4, 20.5, 18.9, 11.7

Table 2. Comparison of fitted residuals of detection points Comparison of fitting residuals of detection point