PAPER • **OPEN ACCESS**

Influence and remedial measures of missing test data on magnetic field source location of underwater vehicle

To cite this article: Ping Hu *et al* 2023 *J. Phys.: Conf. Ser.* **2486** 012061

Influence and remedial measures of missing test data on magnetic field source location of underwater vehicle

Ping Hu[1*[0000-0002-1807-9779]], **Dawen Jiao**[1], **Jiahui Qi** [1] **and Sidi Chen**[1]

[1] Science and Technology on Underwater Test and Control Laborary, Dalian, China

Email:dlinstitute@vip.163.com

Abstract. Due to the incomplete data collected, it will increase the difficulty of inversion location in the field source inversion of the static magnetic field data of underwater vehicles. Therefore, it is very necessary to study the influence of data loss on the static magnetic source location of underwater vehicles and the field source location for timely remedy of missing data. In this paper, the static magnetic equivalent source model of underwater vehicle was established. Then the inversion positioning algorithm was used to locate the magnetic field source of the underwater vehicle obtained from the test, and the influence of the location result under the incomplete data information was analyzed. We used deleting method, linear interpolation method and sliding filling method to populate the data to fill the remedy of missing data. Using the localization algorithm to carried out the inversion location of source, then we took the location of source inversion and the size of the magnetic dipole moment to the ellipsoid model to get fitting curve of magnetic field. The simulation curve of contrast verification test data with experimental data and the lack of remedial method was feasible. The results show that the missing data can not be ignored for the location of the underwater vehicle's magnetic field source, and the data pairs filled with missing data can be effectively located for the equivalent static magnetic field source of the underwater vehicle through using deletion method, linear interpolation method and sliding filling method.

1. Introduction

Magnetic signal is the main exposure and recognition sources of underwater vehicles. Detection technology based on magnetic anomalies is an ideal new development direction for underwater vehicle positioning, with strong penetration ability, not easy to be interfered by non-magnetic environment and short execution time[1-6].In addition, magnetic anomaly detection technology is a kind of passive detection, and there is no need to worry about exposing its position in the detection process. It is a key research direction of underwater vehicle detection. The magnetic field signal of underwater vehicle mainly comes from geomagnetic field magnetization, which is several orders of magnitude smaller than geomagnetic background field. The magnetic anomaly technology is an ideal way to realize the real-time, accurate, continuous and concealed positioning of underwater vehicles by detecting and analyzing magnetic anomalies caused by magnetic characteristics of the target.

The current research is based on the positioning analysis of the target under the condition of complete test data, but the magnetic field data quality greatly affects the positioning of the underwater vehicle and its related applications, and the data cannot be effectively used if the data is of poor quality [7-14].Therefore, it is necessary to pay attention to the location and remedy method of underwater vehicle magnetic field data source under data loss. The data loss of underwater vehicle test includes data loss in experimental collection process and data loss in transmission process. Missing data will affect the quality of data analysis and the accuracy of final positioning decision, so it is very important to preprocess missing data. Aiming at the problem of how to locate the equivalent magnetic field source of underwater vehicles more accurately, this paper constructed a remedial method for missing data filling based on deletion method, linear interpolation method and sliding filling method, and preprocessed data based on this method to achieve effective location of the equivalent magnetic field source of underwater vehicles.

2. Static magnetic field source model construction of underwater vehicle

The static magnetic field source and magnetic dipole of underwater vehicle are very similar in nature, and the magnetic field generated is equivalent. According to Biot-Savart law, the magnetic field generated by magnetic material micro-current element at a certain point of underwater vehicle is expressed as (1):

$$d\boldsymbol{B} = \frac{\mu_0}{4\pi} \cdot \frac{Id\vec{l} \times \vec{r}}{r^3} \tag{1}$$

Where, I is the vector, represents the current and $\vec{r}$ represents the position vector between the micro current element and the detection point. Where $\mu_0 = 4\pi \times 10^{-7}\mathrm{H/m}$ represents the vacuum permeability, and I represents the current, r represents the distance between the space coordinate point and the current element. The direction of magnetic induction intensity can be determined according to the right hand criterion, and generally lies in the plane which is perpendicular with both space current plane and r-plane. Through definite integral operation, the vector expression of magnetic induction intensity in the whole space can be obtained in (2):

$$\boldsymbol{B} = \frac{\mu_0}{4\pi} \int \frac{Id\vec{l} \times \vec{r}}{r^3} \tag{2}$$

3. Positioning principle and steps of underwater vehicle based on static magnetic field

The magnetic field localization of underwater vehicle is to locate the magnetic field signal of the magnetic hull in the geomagnetic field. If the magnetic field is equivalent to a dipole model, the magnetic field signal generated by the magnetic field at any point in space can be calculated under the condition of known magnetic couple moment. Conversely, if the magnetic field signal distribution of dipole source excitation in space is known, the location parameters of the dipole source can also be calculated by the algorithm of model inversion. According to the formula of static magnetic field generated by magnetic dipole, it can be concluded that the relationship between magnetic field values of multiple measuring points, coordinates of measuring points and magnetic dipole moment can be simplified into linear overdetermined equations, which can be written in the matrix form in (3):

$$B = FM \tag{3}$$

$$B = \left(B_{x1}, B_{y1}, B_{z1}, \cdots, B_{xi}, B_{yi}, B_{zi}, \cdots B_{xL}, B_{yL}, B_{zL}\right)$$

Where L is the total number of measured points. F is the matrix composed of known space functions. $M = IL$ is the magnetic dipole moment, and F is represented in the model as a matrix composed of space functions including three-dimensional coordinates of the dipole source, velocity, heading and other factors. Powell local search algorithm, which is more mature in geological inversion, is used to complete the final inversion and location of the objective function through local optimization algorithm.

4. Processing method of missing magnetic field data

Based on the missing magnetic field test data, deletion method, linear interpolation method and sliding filling method were used in this paper to fill and remedy the missing data. Based on this method, the data was preprocessed and the effective positioning of the equivalent magnetic field source of underwater vehicle was realized. The following three filling remedies are briefly introduced.

4.1. Deletion method

Deletion methods include list deletion and pair deletion. The deletion method is very simple to deal with the problem of missing data. If the missing data is MCAR (Missing Completely at Random) type, the sample size is large enough, and the proportion of missing data is less than 5%, the deletion method can be used. However, when there are a lot of missing values or a small sample size in the dataset, the deletion method may discard too many samples, resulting in a large error [11]. List deletion: All cases with incomplete observations are eliminated for all variables to be used in the analysis. Samples such as in table 3, 2, 3, 5 are missing data, if you want to variable X, Y, Z correlation analysis between the two, the list is deleted, all only sample 1, 4, can participate in the analysis process, the output of the correlation coefficient of variables between X, Y, Y, Z, and between the correlation between the variables X and Z, are only sample 1, 4, to participate in the analysis. Paired deletion: Include sample responses to target variables for all variables to be used in the analysis.

4.2. Linear interpolation

Compared with the deletion method, which removes missing data, the linear interpolation method adds data and gives some alternative values to each missing data to make the data set complete. The interpolation method ensures the data sample size for analysis. Interpolation method is to use the principle of similar triangle to calculate the data of interpolation points. According to the $f(x)$ function value of the unknown function at some points in a certain interval, the specific function which value is equal to the value of $f(x)$ at these points is derived to approximate the original function$f(x)$, then the specific function can be used to calculate the approximate value of the original function at other points in the interval. According to the properties of specific functions, there are linear interpolation, nonlinear interpolation, etc. According to the number of arguments (independent variables) points, there are single interpolation, double interpolation and three interpolation. Linear interpolation refers to the interpolation method in which the interpolation function is a polynomial, and the interpolation error on the interpolation node is zero. Compared with other interpolation methods, such as parabolic interpolation, linear interpolation is simple and convenient. The geometric meaning of linear interpolation is to outline the approximation of an anti-derivative function using a line passing through points A and B. Linear interpolation can be used to approximate the original function or to calculate values that are not in the table during table lookup.

4.3. Sliding window filling method

The sliding window algorithm reduces the complexity of the problem by operating on a string or array of a particular size, rather than the entire string or array, and thus reduces the nesting depth of the loop. You can actually see here that sliding windows are mostly used for arrays and strings. The algorithm performs the required operation on an array or string given a particular window size. This technique can turn a part of the problem's nested loop into a single loop, so it can reduce time complexity. The sliding window algorithm reduces the complexity of the problem by operating on a string or array of a particular size, rather than the entire string or array, and thus reduces the nesting depth of the loop. Figure 1 shows the principle of this method.

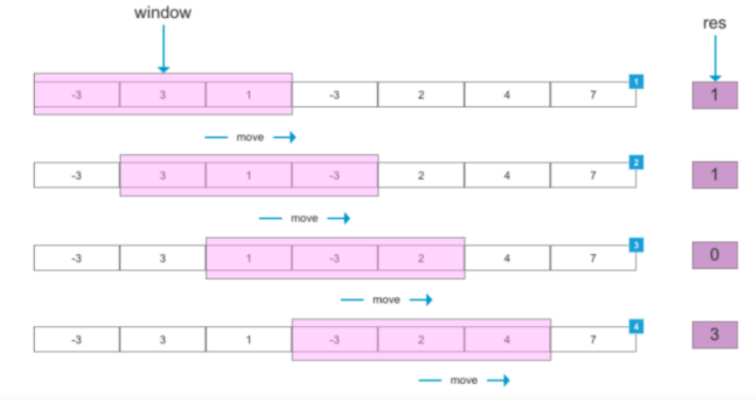

Figure 1. Sliding window filling method

5. Experimental verification analysis

The positioning results of test data recovery after missing validation is on the basis of simulation data, the following three parts: first of all, based on the numerical simulation to obtain complete field data, we used the positioning results given localization algorithm and magnetic dipole moment, and obtained the positioning of the results and the magnetic dipole moment value was taken into the static magnetic field. Each component of the simulation model and the curve of the PLD were given, which was called the simulation curve. Then the100 points were replaced by complete magnetic field data randomly, forming the missing data. Using the missing data given localization algorithm to get location results and the magnetic dipole moment, the positioning of the results and the magnetic dipole moment value were taken into the static magnetic field simulation model, then the curve of each component and the total field were given, called the fitted curve. The difference between the missing data curve and the simulation curve was obtained by comparing the fitted curve with the simulation curve. Through using the method of three kinds of data missing remedy respectively, the missing data was deleted, interpolated, or smooth filled. Filling data source location results and the magnetic dipole moment were given, the positioning of the results and the magnetic dipole moment value were taken into the static magnetic field simulation model, then the curve of each component and the total field were given. And the curve was compared with the simulation curve of complete data to confirm feasibility of remedy.

5.1. Static magnetic field source location and magnetic couple moment inversion in the case of complete data

Simulation test was done in the static magnetic field model, the real field source point was set as [0,0,0] and the real magnetic moment was set as [2e5,1e5,2e5], generating a group of sensor coordinate points. The measuring range in the x direction was set as -500~500, the measuring range in the y direction was set as -50~50, and the coordinate in the z direction was fixed as 46. The length and width of the ship were respectively set as 80 and 8. Substituting the above parameters into the static electric field model to achieve the static electric field data.

Then the source point fix and the magnetic moment of inversion were done, the x direction search range was set to [100100], y direction search range was set to [- 20, 20], z direction coordinate was fixed to 0. Using matching search optimization inversion in five dimensions, the population size was set to 50, the number of iterations was set to 50. Algorithm search results are as in table 1.

Table 1. Algorithm search results.

source point positioning(m)	inverse magnetic moment X component （A.m^2）	inverse magnetic moment Y component （A.m^2）	inverse magnetic moment Z component （A.m^2）

6.6201, -0.2208, 0	1.2298e+05	9.9322e+04	2.1877e+05

Tab1 gives the Loc is the source point of the positioning field, and mMx1, mMy1 and mMz1 are the inverse magnetic moments, Fig2. The compare curves between simulated field and fitted field.

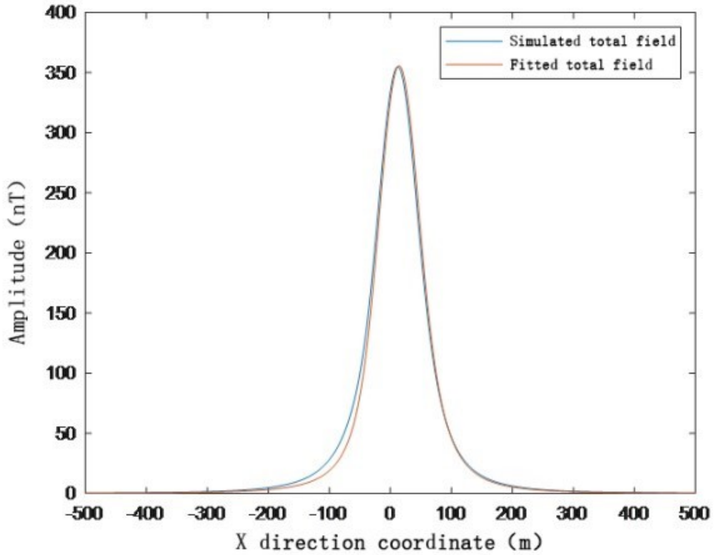

Figure 2. The compare curves between simulated field and fitted field

In the case of complete simulation data, the field source location and magnetic moment inversion can be realized, and the fitted field strength curve is basically consistent with the simulated field strength curve.

5.2. Static magnetic field source location and magnetic moment inversion in the case of missing data

Taking 100 points arbitrarily from the simulated field intensity curve and making the 100 points non-number (NaN), the simulated field intensity curve with missing data can be obtained as follows in fig3. As seen from the figure above, there is an obvious breakpoint on the curve. Under the same parameter setting, matching inversion is used to search for five-dimensional optimization, and the search results of the algorithm are as in table 2.

Table 2. Algorithm search results.

source point positioning(m)	inverse magnetic moment X component （A.m^2）	inverse magnetic moment Y component （A.m^2）	inverse magnetic moment Z component （A.m^2）
13.2202, 2.6440, 0	3.9661e+04	3.9661e+04	3.9661e+04

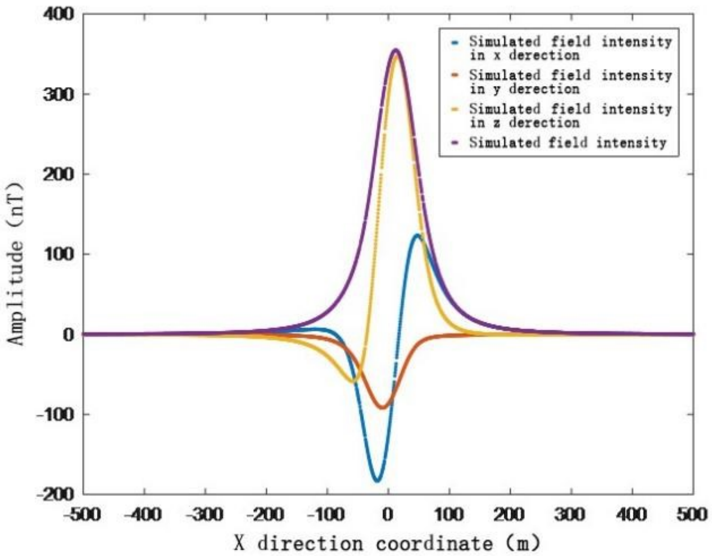

Figure 3. The simulated field intensity curve with missing data

Substituting the above results into the static magnetic field model, the compared curve can be obtained as in figure 4.

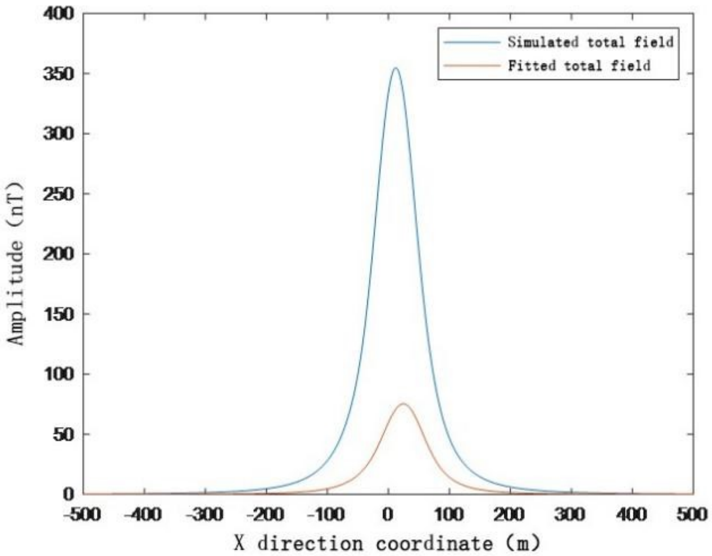

Figure 4. The compare curves between simulated field and fitted field

As seen from the above simulation results, data loss does affect the positioning and inversion effect, and the error of the results obtained from missing data increases, and there is a great difference between the fitted total field curve and the simulated total field curve. Next, there are three ways to solve the problem of missing data.

5.3. Remedies for missing data

5.3.1. Delete the missing data corresponding to the missing data. After deleting the point with NaN data directly from the data set, the field source point location and magnetic moment inversion were carried out. Under the same parameter setting, matching inversion was used for five-dimensional search and optimization. The search results of the algorithm are as in Table 3, the compare curvers between simulated field and fitted field is in Fig5.

Table 3. Algorithm search results.

source point positioning(m)	inverse magnetic moment X component （A.m^2）	inverse magnetic moment Y component （A.m^2）	inverse magnetic moment Z component （A.m^2）
1.1942, -0.6619, 0	1.7350e+05	1.1977e+05	2.0606e+05

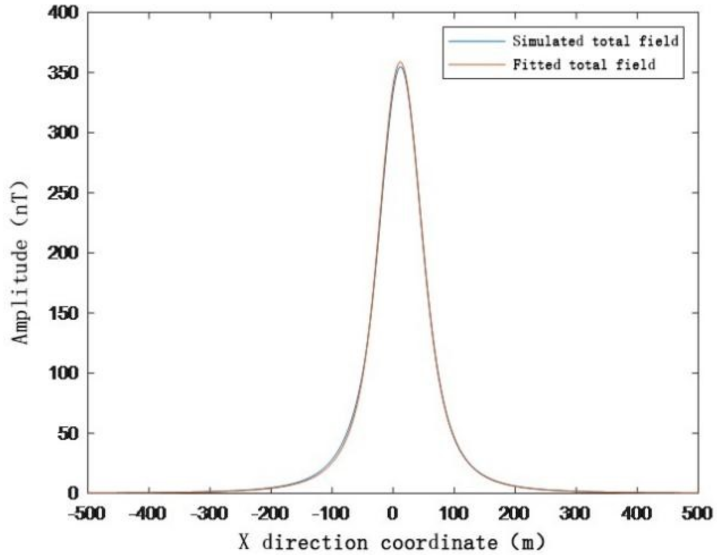

Figure 5. The compare curves between simulated field and fitted field

According to the above simulation results, the location of field source points and inversion of magnetic moment can be realized after the corresponding points of missing data are deleted directly, and the fitted field intensity curve obtained is basically consistent with the simulated field intensity curve.

5.3.2. Linear interpolation method to fill missing data. The missing data were filled with linear interpolation method,after the data was filled with linear interpolation method, the breakpoints on the simulated field strength curve were basically filled, and then the location of the field source point and the inversion of the magnetic moment were carried out. Under the same parameter setting, the matching inversion was used for five-dimensional search and optimization. And the search results of the algorithm are as in table 4, the compare curvers between simulated field and fitted field is in Fig6.

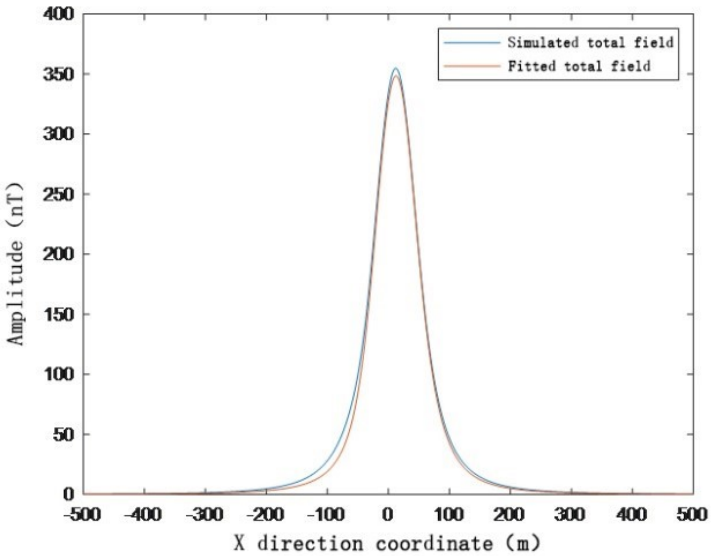

Figure 6. The compare curves between simulated field and fitted field

It can be seen from the above simulation results that the location of field source points and inversion of magnetic moment can be realized after filling data with linear interpolation method, and the fitted field intensity curve obtained is basically consistent with the simulated field intensity curve.

Table 4. Algorithm search results.

source point positioning(m)	inverse magnetic moment X component ($A.m^2$)	inverse magnetic moment Y component ($A.m^2$)	inverse magnetic moment Z component ($A.m^2$)
4.8695, -0.2926, 0	1.1764e+05	9.3388e+04	2.1558e+05

5.3.3. Missing data filled by sliding filling method. The missing data were filled with the sliding filling method, which using window median length of 10 mobile to replace NaN values of the data set. It can be seen as in table 5.

Table 5. Algorithm search results.

source point positioning(m)	inverse magnetic moment X component ($A.m^2$)	inverse magnetic moment Y component ($A.m^2$)	inverse magnetic moment Z component ($A.m^2$)
2.7621, 2.9745, 0	1.7661e+05	7.2411e+04	2.1570e+05

Substituting the above results into the static magnetic field model, the compared curves can be obtained as in Fig7:

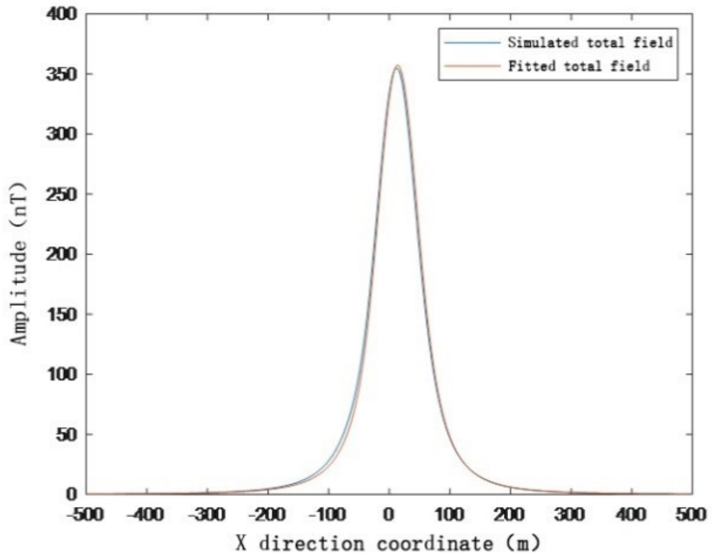

Figure 7. Compare curves between simulated field and fitted field

According to the above simulation results, the location of field source points and inversion of magnetic moment can be realized by using the sliding filling method after filling data, and the fitted field intensity curve obtained is basically consistent with the simulated field intensity curve.

6. Summary

Firstly, in this paper the source equivalent magnetic dipole model of static magnetic field was established. The static magnetic field was generated by magnetic dipole source. And the position of static magnetic field data was input to inversion software for inversion static magnetic field of underwater bodies, the equivalent source location and size of the dipole moment. And then 100 positions of the magnetic field data were randomly chose, they were set to NaN. After the missing data was located again, the results showed that the missing data affect the positioning and inversion effect, the error of the results obtained from the missing data increased, and the fitted total field curve was quite different from the simulated total field curve. Finally, this paper used the deletion method, linear interpolation method and smooth window filling method to invert the static magnetic field source with missing data to locate and invert the magnetic moment, so as to achieve the location of the field source point and obtain the magnetic moment effectively. The fitted field intensity curve obtained was basically consistent with the simulated field intensity curve.

Meanwhile, it should be noted that the method proposed in this paper is mainly verified based on simulation data, and the missing data is randomly selected. However, in the actual marine environment, the data acquisition performance and data transmission of sensors are affected by factors such as waves, ocean currents and human interference. The applicability of the application of remedial methods and positioning methods based on the missing magnetic field data of actual underwater vehicles still needs further verification.

7. References

[1] Dong Peng, Sun Zhe, Zou Nianyang and Li Yueyang. *Ship Science and Technology*, 2018, 40(11): 166-169
[2] CORFIELD S and YOUNG J. Unmanned surface vehicles-game changing technology for naval operations[J]. *IEEE Control Engineering Series*, 2006, 69:311
[3] Zheng Huarong, WEI Yan and QU Fengzhong. Research status of surface unmanned vehicle [J]. *Shipbuilding of China*,2020,61(S1):228-240
[4] YAN Rujian, PANG Shuo and SUN Hanbing, et al. Development and missions of unmanned surface vehicle[J]. *Journal of Marine Science and Application*, 2010, 9(4): 451-457

[5] JAMES E T, *The navy unmanned surface vehicle (USV) master plan*[R]. Washington: U.S. Department of Defense,2007
[6] U.S. *unmanned system roadmap* 2007-2032[R]. Washington: U.S. Department of Defense, 2007
[7] Allison P D. Multiple Imputation for Missing Data: A Cautionary Tale[J].*Sociological Methods and Research*, 2000,28(3)
[8] Zhang Chengping. *Incomplete Data Filling* [D]. Changsha: Central South University, 2006
[9] Yang Xiaoqian. *Research on selection of missing data interpolation method* [D]. Lanzhou: Dissertation of Lanzhou University of Finance and Economics, 2016
[10] Zhang Lijun. Nanjing: Nanjing University of Science and Technology,2018. (in Chinese)
[11] Li Shengyu. Shijiazhuang: Hebei University of Economics and Business, 2015. (in Chinese)
[12] Pang Xinsheng. Implementation of Multiple Interpolation Method for Missing Data [J]. *Statistics and Decision*,2012, (11)
[13] Ye Sha. Overview of Missing data and its Processing Methods [J]. *Electronic Measurement*, 2017, (18)
[14] Little R J A, Rubin D B. Statistical Analysis With Missing Data[M].*New York: Wiley and Sons Inc*, 1987

Acknowledgments
This paper was supported by basic scientific research project (JCKY2019207A019).