Retrieval of Suspended Sediment in Songhua River Based on Sentinel-2 Satellite

Xu Han and Chong Du*

College of Water Conservancy and Electric Power, Heilongjiang University, Harbin 150080 China
E-mail: 149596471@qq.com

Abstract. In view of the current situation that the remote sensing inversion of suspended sediment is mainly based on low spatial resolution satellite data, sentinel-2 satellite data with resolution of up to 10m is used to retrieve the suspended sediment content in Harbin section of Songhua River. And the inversion model is established based on the measured data and remote sensing data. The results show that the fitting degree of b3/ (B2 + B4) in the multi band ratio and the measured suspended sediment concentration is the highest, reaching 0.904. The results show that the average relative error between the inversion value and the measured value is 19.88%, and the root mean square error is 0.039g/l. The results show that the suspended sediment content of Songhua River is lower than that of other six rivers in China, and belongs to the less sand river. The sediment content on both sides of the bank is higher than that of the center of the river. In the wide river, the suspended sediment content is low, and the suspended sediment content is higher in the narrow area.

Keywords: Suspended sediment; Remote Sensing Retrieval; Sentinel-2

1. Introduction

Suspended sediment is one of the most important water quality parameters of rivers. The quantitative study of suspended sediment remote sensing inversion is of great significance to the ecological environment, bank geology and geomorphology, and channel construction of rivers. The traditional method of measuring suspended sediment is to monitor the sampling points one by one in the sampling area for a long time. This method is not only time-consuming, but also has many disadvantages, such as high cost and unable to monitor large areas of water for a long time. Remote sensing technology has a wide range of monitoring, short time-consuming, and the cost of remote sensing image acquisition is much lower than the traditional sampling method, so the use of remote sensing technology to monitor river suspended sediment has become a new development trend [1].

Restricted by the resolution of remote sensing image and other factors, the previous retrieval of suspended sediment using remote sensing technology mostly focused on the surface water bodies such as coastal second-class water bodies and lakes, but less on the linear water bodies such as rivers. In this paper, the remote sensing image resolution of sentinel-2 is as high as 10m after resampling by snap software, and the revisiting period of sentinel-2 is shorter than that of Landsat series [2].

2. Data and Methods

2.1. Overview of the Study Area

Songhua River is one of the seven rivers in China. Located in the north of Northeast China, it flows...
through Inner Mongolia, Heilongjiang and Jilin provinces. The specific latitude and longitude range is north latitude 41° 42′~51° 38′, east longitude 119° 52′~132°, it has a total length of 2000 km. In recent years, the average precipitation is about 500 mm. The precipitation in the southeast mountainous area is 700-900 mm, while in the arid Inner Mongolia basin it is only 400 mm. The general trend is that the precipitation in the mountainous area is larger than that in the plain; It is slightly larger in the south central part, next in the East, and the smallest in the northwest. The precipitation from June to September in flood season accounts for 50% ~ 70% of the whole year, and the precipitation from December to February in winter is only about 5% of the whole year [3].

Harbin section of Songhua River is selected as the research area. In recent years, Harbin has made great efforts to develop shipping. The water transport capacity of Songhua River is more than 10 million tons. The study of sediment distribution in this area can provide data support for channel construction and route planning.

2.2. Measured Data
The field sampling time was April 12. There were 14 sampling points, 8 for the establishment of inversion model and 6 for the verification of model accuracy. Before the field measurement, the quality of the container and filter membrane corresponding to each sampling point should be measured [4]. The container is 500ml wide mouth plastic bottle, and the field sampling depth is 0.5m. Three bottles of water samples are collected at each sampling point to take the average value, so as to reduce the error and later data screening. After collecting water samples, immediately return to the laboratory for treatment, the experimental process is as follows. First shake the sample well, measure the total mass of the container and water with a balance accurate to 0.001g, and then use 0.45µm The filtered membrane and sediment are dried in a 40 ℃ dryer for 5h and then taken out. The sediment mass is measured with a balance accurate to 0.001g. Calculation formula:

Sediment content = sediment mass / sample volume

2.3. Remote Sensing Data
The remote sensing data used in this paper are all from sentinel-2 satellite, which is downloaded on the website of ESA. The satellite was launched on June 23, 2015, with a wide range of 290km, a revisit period of 5D and 13 bands. The parameters of each band are shown in table 1.

| Band   | Wavelength/µm | Spatial resolution /m | Band width/nm |
|--------|---------------|-----------------------|---------------|
| Band1  | 0.433         | 60                    | 20            |
| Band2  | 0.490         | 10                    | 65            |
| Band3  | 0.560         | 10                    | 35            |
| Band4  | 0.665         | 10                    | 30            |
| Band5  | 0.705         | 20                    | 15            |
| Band6  | 0.740         | 20                    | 15            |
| Band7  | 0.783         | 20                    | 20            |
| Band8  | 0.842         | 10                    | 115           |
| Band8A | 0.865         | 20                    | 20            |
| Band9  | 0.945         | 60                    | 20            |
| Band10 | 1.375         | 60                    | 20            |
| Band11 | 1.610         | 20                    | 90            |
| Band12 | 2.190         | 20                    | 180           |

The measured time is about 14:00 p.m. on April 12. Since the labeling time of sentinel-2 satellite image is Greenwich mean time and the difference between Greenwich mean time and Beijing time is 8h, the time of sentinel-2 satellite image should be as close as possible to 6:00 a.m. to reduce the error
caused by time difference. Finally, the image of 6:32 a.m. on April 12 is selected [5]. The cloud coverage is less than 10% and the visibility is high, which is suitable for the establishment of inversion model.

Because sentinel-2 satellite data download is L1C product, after geometric correction, there is no radiometric calibration and atmospheric correction, so it is necessary to preprocess remote sensing data. There are many methods of atmospheric correction, including the atmospheric correction of the flash sh module in envi, the atmospheric correction of the 6S model, and the sen2cor method. Because the sentinel-2 satellite remote sensing image is selected in this paper, the sen2cor provided by ESA is used to process the L1C level data to obtain the L2 a level atmospheric bottom reflectance data [6].

The spatial resolution of each band of sentinel-2 satellite is different, so resampling is needed to unify the spatial resolution. Using snap software, the spatial resolution of each band is resampled to 10m, and the resampling method is bilinear interpolation method. The output data is in the format that envi can open. After band fusion, the layerstacking tool in envi is used for image clipping and stitching, and the remote sensing reflectance data of measured points in each band is obtained [7].

3. Construction of Inversion Model

3.1. Sensitive Band Selection

Sentinel-2a has 13 bands, and the tenth band is used for atmospheric correction, so it does not participate in the correlation study. From the first wave break, the central wavelength increases continuously, and the wavelength range from the first band to the twelfth band can reach 0.433 µm~2.190 µm. Compared with other water systems in China, Songhua River is a river with less sand. In clear water, the reflectance of the river is inversely proportional to the wavelength, that is, the larger the wavelength, the lower the reflectance. Through image analysis, at the wavelength of 1.0 µm~2.190 µm In addition, band5, Band6 and band7 are red edge bands of vegetation, which are less sensitive to water. Therefore, we finally choose band2 blue band, band3 green band, Band4 red band and band8 near infrared band as research bands [8].

3.2. Band Correlation Analysis

At present, there are many methods for retrieving suspended sediment, such as single band, band ratio, multi band combination and so on. According to previous research results, band ratio and multi band combination can reduce the impact of chlorophyll and other suspended matter on the retrieval results. In this paper, we first do single band Pearson correlation analysis on the selected four bands, then make the ratio of these four bands to each other, and do Pearson correlation analysis with the concentration of suspended sediment. Finally, we use band math tool in envi to calculate the combined data of three bands and four bands, and do Pearson correlation analysis with the concentration of suspended sediment. The correlation coefficients of each band are shown in table 2.

| Band   | Correlation coefficient | P     |
|--------|------------------------|-------|
| B2     | 0.162                  | 0.695 |
| B3     | 0.224                  | 0.155 |
| B4     | 0.463                  | 0.087 |
| B8     | 0.096                  | 0.831 |
| B2/B3  | 0.621                  | 0.038 |
| B2/B4  | 0.567                  | 0.041 |
| B2/B8  | 0.373                  | 0.124 |
| B3/B4  | 0.574                  | 0.053 |
| B3/B8  | 0.377                  | 0.126 |
| B4/B8  | 0.369                  | 0.130 |
According to Pearson correlation analysis, the correlation of single band is relatively low, especially the correlation of band 8 near infrared band is only 0.096, the correlation is very low, only the p value of band 2 blue band and band 3 green band is less than 0.5, and the correlation coefficient is small. In the band ratio analysis, the overall correlation is higher than the single band correlation, and the correlation of B2 / B3, B2 / B4, B3 / B4 is greater than 0.5, which shows good performance. In the multi band correlation analysis, compared with the band ratio and single band, the correlation of B2 / (B3 + B4) and B3 / (B2 + B4) is greater than 0.6, which can meet the correlation requirements of inversion model. In general, whether the correlation analysis of single band 8 near infrared band or band ratio and multi band combination involves band 8 near infrared band correlation analysis, the correlation is much lower than the same level. It may be that the Songhua River is a river with less sediment, and the ice and snow on the river surface has just melted, so the sediment on the river surface is less impurity, so band8 is not sensitive to it.

3.3. Establishment of Inversion Model

According to band correlation analysis, this paper selects band ratio and band combination with correlation coefficient greater than 0.6 to establish inversion model. Linear model, exponential model, quadratic polynomial model and logarithmic model are selected for inversion model. The SPSS curve estimation module is used to establish the inversion model [9]. The band ratio and the remote sensing reflectance of multi band combination are used as independent variables, and the suspended sediment concentration is used as dependent variables. The model with the highest fitting degree is selected as the final inversion model. The specific model relationship and fitting degree are shown in Table 3.

| Independent variable | Empirical model | Empirical model relations | $R^2$ |
|----------------------|----------------|--------------------------|-------|
| B2/B3                | Linear         | SSC=0.073x+0.028          | 0.686 |
|                      | Index          | SSC=0.070ln(x)+0.103      | 0.724 |
|                      | Quadratic      | SSC=0.181x²-0.055x-0.022 | 0.751 |
|                      | Logarithm      | SSC=0.046e^0.768x         | 0.642 |
| B2/(B3+B4)           | Linear         | SSC=0.216x-0.011          | 0.763 |
|                      | Index          | SSC=0.102ln(x)+0.169      | 0.865 |
|                      | Quadratic      | SSC=0.123x²+0.095x+0.011  | 0.795 |
|                      | Logarithm      | SSC=0.029e^2.326x         | 0.818 |
| B3/(B2+B4)           | Linear         | SSC=0.286x-0.047          | 0.783 |
|                      | Index          | SSC=2.968ln(x)+0.021      | 0.841 |
|                      | Quadratic      | SSC=0.556x²-0.259x-0.116  | 0.904 |
|                      | Logarithm      | SSC=0.200e^0.147x         | 0.806 |
The fitting degree analysis results show that $B_3 / (B_2 + B_4)$ has the highest fitting degree and the correlation coefficient is 0.904

$$f(x) = 0.556x^2 - 0.259x - 0.116$$

### 3.4. Model Accuracy Verification

Two methods, average relative error and root mean square error, are used to evaluate the accuracy of the model.

1. **Average relative error**: calculate the absolute value of the relative error of each point and then take the average value. The smaller the average relative error is, the higher the accuracy of the model is. On the contrary, the larger the average relative error is, the lower the accuracy of the model is [10].

   $$MAPE (\%) = \frac{1}{n} \times \sum_{i=1}^{n} \left| \frac{Y_i - f_i}{Y_i} \times 100\% \right|$$

2. **Root mean square error**: also known as standard error, it is the square root of the ratio of the square of the difference between the simulated value and the measured value and the number of measurements $n$. The smaller the root mean square error, the higher the accuracy of the model. On the contrary, the larger the root mean square error, the lower the accuracy of the model.

   $$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - f_i)^2}$$

In the above formula: $n$ is the number of samples; $Y_i$ is the measured suspended sediment content; $f_i$ is simulate suspended sediment content.

In this model validation, 6 verification points of 14 points were collected. The maximum relative error was 25.13%, the minimum relative error was 7.19%, the average relative error was 19.88%, and the root mean square error was 0.039 g/L. The model accuracy basically met the requirements. Figure 1 shows the comparison between the measured value and the simulated value. It can be seen from the figure that the simulated value is generally higher than the measured value. Except for individual points, the size of the simulated value and the measured value are basically in a fitting state. The error may be caused by the error of nearly 8 hours between the remote sensing image time and the measured time, and the change of hydrological conditions such as river velocity. Remote sensing image has a small amount of cloud cover, which affects the subsequent atmospheric correction results, resulting in errors. There are chlorophyll and other substances in the river, which affect the remote sensing reflectance of sensitive bands and cause errors. Finally, the experimental process of filtration, drying and other steps in the operation process will also cause errors.

![Figure 1. Comparison of simulated and measured values.](image-url)
4. Conclusion
Taking Harbin section of Songhua River as the research area, based on sentinel-2 satellite remote sensing data and field measured data, this paper studies the inversion model of suspended sediment concentration in Harbin section of Songhua River, constructs the suspended sediment concentration model of Harbin section of Songhua River, and draws the following conclusions according to this model [11].

Among 20 single band, band ratio and multi band combination inversion models, the B3 / (B2 + B4) multi band combination quadratic model can better reflect the relationship between remote sensing reflectance and suspended sediment concentration in Harbin section of Songhua River, and is suitable for suspended sediment concentration inversion in Harbin section of Songhua River.

By calculating the average inversion value, the average sediment concentration of Harbin section of Songhua River in April is 0.531 g / L. compared with the Yellow River, Yangtze River, Pearl River and other major rivers in China, the sediment concentration is less, which belongs to the less sediment river [12].

Through the inversion model, the sediment distribution in the study area is generally characterized by high concentration of suspended sediment on the bank, low concentration of sediment in the center of the river, high concentration of suspended sediment in the narrow channel and low concentration of suspended sediment in the wide channel.

Reference
[1] Lee Z P, Shang S L, Hu C M, et al. 2015 Secchi disk depth: A new theory and mechanic model for underwater visibility Remote Sensing of Environment.
[2] Dang Y Y and Wang X 2020 Application of CPU-GPU heterogeneous system in optical remote sensing image processing Infrared and Laser Engineering 49(51): 177-185. (in Chinese)
[3] Jiang X W, Liu J Q, Zou B, et al. 2009 Satellite remote sensing emergency monitoring system for Enteromorpha prolifera disaster and its application Acta Oceanographica Sinica 31(01): 52-64.
[4] Li D R 2000 The current situation and development trend of photogrammetry and remote sensing Journal of Wuhan University of Surveying and Mapping (01): 1-6.
[5] Wang R Sh, Xiong Sh Q, Nie H F, et al. 2011 Remote sensing geological exploration technology and application research Acta Geologica Sinica 85(11): 1699-1743.
[6] Chang G 2001 Research and Practice of Urban Landscape Modeling Technology Based on Remote Sensing Data PLA University of Information Engineering.
[7] Wu B F 2000 Operational remote sensing method for national agricultural monitoring and yield estimation Acta Geographica Sinica (01): 25-35.
[8] Lin M S, Zhang Y G, yuan X Zh 2015 Development history and trend of ocean remote sensing satellite Acta Oceanographica Sinica 37(01): 1-10.
[9] Li Zh B 1997 Resource characteristics and development of island areas in Zhejiang Province: A case study of Zhoushan archipelago Journal of Natural Resources (02): 44-50.
[10] Philpot W D, Ackleson S G 1981 Remote Sensing of Optically Shallow, Vertically Inhomogeneous Waters: A Mathematical Model NASA. Langley Research Center Chesapeake Bay Plume Study.
[11] Gao Sh W, Wang B F, Zhu L Y, et al. 1998 Monitoring and evaluation index system of sandy desertification land in China Forestry Science (02): 3-12.
[12] Buhe A, Liu J Y and Wang Ch Y 1997 The role of remote sensing technology in the study of resources and environmental change in China Science and Technology Guide (08): 14-17.