A Study on Retrieval Algorithm of Black Water Aggregation in Taihu Lake Based on HJ-1 Satellite Images

ZOU Lei, ZHANG Bing, LI Junsheng, SHEN Qian, ZHANG Fangfang, WANG Ganlin
Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100190, China
lzou@ceode.ac.cn

Abstract. The phenomenon of black water aggregation (BWA) occurs in inland water when massive algal bodies aggregate, die, and react with the toxic sludge in certain climate conditions to deprive the water of oxygen. This process results in the deterioration of water quality and damage to the ecosystem. Because charge coupled device (CCD) camera data from the Chinese HJ environmental satellite shows high potential in monitoring BWA, we acquired four HJ-CCD images of Taihu Lake captured during 2009 to 2011 to study this phenomenon. The first study site was selected near the Shore of Taihu Lake. We pre-processed the HJ-CCD images and analyzed the digital number (DN) gray values in the research area and in typical BWA areas. The results show that the DN values of visible bands in BWA areas are obviously lower than those in the research areas. Moreover, we developed an empirical retrieving algorithm of BWA based on the DN mean values and variances of research areas. Finally, we tested the accuracy of this empirical algorithm. The retrieving accuracies were 89.9%, 58.1%, 73.4%, and 85.5%, respectively, which demonstrates the efficiency of empirical algorithm in retrieving the approximate distributions of BWA.

1. Introduction
After a large number of blue algae aggregate and die in eutrophic water under conditions of high water temperature, slow wind, low air pressure, and a weak reoxygenation rate, they decompose and release substances such as volatile sulfide and dim ethyl trisulfide, which combine chemically with heavy metals in sediments. As a result, a foul-smelling brownish black cluster forms that leads to anoxic conditions and the deterioration of water quality. This phenomenon is known as black water aggregation (BWA) [1].

Such clusters appear on a small scale and generally persist for less than five days. BWA is responsible for causing severe damages to the environment. The toxins released pollute water resources and threaten people’s health. In addition, BWA leads to a rapid decrease of dissolved oxygen concentrations in water, which can destroy aquatic flora and fauna, affecting the natural environment and aquaculture. Taihu Lake is a typical inland eutrophic water body that was selected for research because BWA occurs there occasionally. Lu and Ma (2009) and Sheng et al. (2010) researched the characteristics and formation factors of BWA in Taihu Lake [1, 2]. Lu and Ma (2010) also analyzed the phenomenon in 2009 at Taihu Lake with focus on an area susceptible to BWA [3]. Wang et al. (2010) presented an early warning method for BWA in Taihu Lake based on
meteorological factors \[^{[4]}\]. Li et al. (2012a) proposed a remote sensing retrieval algorithm of BWA based on Thermal Mapping (TM) images. However, in general, research on the BWA phenomenon remains novel worldwide \[^{[6]}\].

Chinese remote sensing satellites use micro-satellite constellations to forecast and monitor environmental changes and disasters. These satellites are equipped with two wide-coverage multispectral charge coupled device (CCD) cameras. Known as HJ-CCD, they are important tools because they are capable of covering 700km per image with medium-high spatial resolution (30 m) and high temporal resolution (two days). Therefore, the HJ-CCD satellite is more suitable for BWA monitoring than are the Moderate-Resolution Imaging Spectroradiometer (MODIS) and TM methods with special and temporal resolutions of 250 m and one day, and 30 m and 16 days, respectively.

Li et al. (2012b) studied visual interpretations of BWA by using HJ-CCD images \[^{[7]}\]. The present study uses the same data to explore an automatic retrieval algorithm of BWA to provide a new approach for operational monitoring of the BWA phenomenon.

2. Data Processing

Li (2012) reported that BWA occurred on May 11, 2009; August 21, 2010; July 28, 2011; and September 24, 2011 \[^{[7]}\]. Therefore, we acquired the corresponding HJ-CCD images for our research. After composition, clipping, and application of geometric precision correction to each HJ-CCD image, we obtained four digital number (DN) images of Taihu Lake. We chose the northwest near shore region of Taihu Lake as our research area because it is known for BWA occurrence (Figure 1).

![Figure 1. True color images of research area after preprocessing.](a) May 11, 2009 (b) August 21, 2010 (c) July 28, 2011 (d) September 24, 2011

On the basis of previous research on the formation conditions of BWA, we masked off the part of this region that is susceptible to BWA, represented by a red frame in Figure 2. According to the results of Li (2012), we determined the central parts of typical BWA areas by visual interpolation, which are shown as yellow frames in the figure.

![Figure 2. Research areas and typical areas of black water aggregation (BWA), marked off by red and yellow frames, respectively.](a) May 11, 2009 (b) August 21, 2010 (c) July 28, 2011 (d) September 24, 2011
In Tables 1–4, the statistical results of DN values of research areas and BWA areas in visual bands of HJ-CCD images are presented. DN_min, DN_max, and DN_mean represent the minimum, maximum, and mean values of gray values in different bands. \( \sigma \) represents the variance of the gray values in each band.

**Table 1.** Statistical results of gray values in visible light bands captured on May 11, 2009.

| Band | Object        | DN_min | DN_max | DN_mean | \( \sigma \) |
|------|---------------|--------|--------|---------|-------------|
| 1 (blue) | BWA areas     | 35     | 40     | 37.28   | 0.71        |
|       | Research areas | 32     | 53     | 45.07   | 3.27        |
| 2 (green) | BWA areas     | 26     | 32     | 28.54   | 0.87        |
|       | Research areas | 23     | 47     | 38.72   | 4.11        |
| 3 (red)  | BWA areas     | 23     | 31     | 26.03   | 1.20        |
|       | Research areas | 20     | 49     | 38.27   | 4.73        |

**Table 2.** Statistical results of gray values in visible light bands captured on August 21, 2010.

| Band | Object        | DN_min | DN_max | DN_mean | \( \sigma \) |
|------|---------------|--------|--------|---------|-------------|
| 1 (blue) | BWA areas     | 49     | 56     | 50.73   | 1.29        |
|       | Research areas | 45     | 69     | 60      | 3.69        |
| 2 (green) | BWA areas     | 29     | 41     | 32.80   | 2.08        |
|       | Research areas | 29     | 62     | 45.54   | 4.49        |
| 3 (red)  | BWA areas     | 23     | 37     | 27.03   | 2.09        |
|       | Research areas | 21     | 61     | 43.72   | 6.54        |

**Table 3.** Statistical results of gray values in visible light bands captured on July 28, 2011.

| Band | Object        | DN_min | DN_max | DN_mean | \( \sigma \) |
|------|---------------|--------|--------|---------|-------------|
| 1 (blue) | BWA areas     | 62     | 67     | 64.06   | 0.93        |
|       | Research areas | 60     | 79     | 70.39   | 2.37        |
| 2 (green) | BWA areas     | 44     | 53     | 47.23   | 1.54        |
|       | Research areas | 43     | 65     | 56.43   | 2.94        |
| 3 (red)  | BWA areas     | 36     | 46     | 39.59   | 1.59        |
|       | Research areas | 36     | 61     | 49.94   | 3.79        |

**Table 4.** Statistical results of gray values in visible light bands captured on September 24, 2011.

| Band | Object        | DN_min | DN_max | DN_mean | \( \sigma \) |
|------|---------------|--------|--------|---------|-------------|
| 1 (blue) | BWA areas     | 38     | 41     | 39.37   | 0.53        |
|       | Research areas | 34     | 55     | 44.71   | 2.47        |
| 2 (green) | BWA areas     | 28     | 33     | 30.32   | 0.62        |
|       | Research areas | 20     | 55     | 40.86   | 4.66        |
| 3 (red)  | BWA areas     | 26     | 30     | 27.24   | 0.69        |

The DN values in visible light bands of BWA areas were lower than those in the research areas. However, DN values of BWA areas varied significantly among images, partly because no atmospheric correction was applied to the images. Previous research indicates that extraction of BWA areas from HJ-CCD data by constant threshold values was unsuccessful. Based on data and the above analysis, we examined the relationships between gray values of BWA and research areas to eliminate atmospheric effects and to determine BWA areas accurately.

3. Retrieval Algorithm of Black Water Aggregation

In the present study, we selected the former three HJ-CCD images, from May 11, 2009; August 21, 2010; and July 28, 2011 as training data to develop the retrieval algorithm. All four images were used to test the algorithm.
3.1. Identification of Black Water Aggregation from HJ-CCD Images

Because the visually interpreted BWA areas contained abnormal values other than those associated with BWA, our results did not reflect the actual ranges of BWA’s gray values in the HJ-CCD images. To eliminate these abnormal values, we employed a range of \( \text{DN}_{\text{min}} \) (research area), \( \mu + 2.58 \sigma \) to replace the range of \( \text{DN}_{\text{min}} \) (BWA), \( \text{DN}_{\text{max}} \) (BWA) as the ranges of BWA’s DN values. The symbols, \( \mu \) and \( \sigma \), are mean value and variance of DN values in BWA areas, respectively. The results are shown in Table 5.

### Table 5. Band threshold values of black water aggregation.

| Image Date | Band 1 | Band 2 | Band 3 |
|------------|--------|--------|--------|
| May 11, 2009 | (32, 39) | (23, 30) | (18, 29) |
| August 21, 2010 | (45, 54) | (29, 37) | (21, 32) |
| July 28, 2011 | (60, 66) | (43, 51) | (36, 43) |

Moreover, to eliminate boundary effect, pixels were identified as BWA only when more than 200 linked pixels met the band thresholds stated above. The extracted results, shown in Figure 3, indicate that these threshold values could precisely extract the areas of BWA.

![Figure 3. Results of extraction by threshold values](image)

3.2. Building Retrieval Algorithm of Black Water Aggregation

Because the gray values in previous research were obtained by visual interpolation, which is inconvenient and limits the operational application of HJ-CCD data, we developed a method for automatically determining the gray values of BWA areas from the DN values in research areas. We built an empirical algorithm to extract BWA by utilizing mean values and variance of gray values in the research areas. We determined that the pixels are BWA areas when at least 200 pixels in the research areas are linked together and all of the gray values in bands 1–3 meet the criteria in formulas (1) to (3).

\[
\begin{align*}
\text{DN}(1) < (\mu \text{ research area (DN)} - n^1 \sigma \text{ research area (DN)}) & \quad (1) \\
\text{DN}(2) < (\mu \text{ research area (DN)} - n^2 \sigma \text{ research area (DN)}) & \quad (2) \\
\text{DN}(3) < (\mu \text{ research area (DN)} - n^3 \sigma \text{ research area (DN)}) & \quad (3)
\end{align*}
\]

\( \text{DN}(1), \text{DN}(2), \) and \( \text{DN}(3) \) are the gray values of one pixel in the research area in bands 1 to 3, respectively. \( \mu \text{ research area (DN)} \), \( \mu \text{ research area (DN)} \), and \( \mu \text{ research area (DN)} \) are the mean gray values of pixels in research areas in bands 1 to 3. \( \sigma \text{ research area (DN)} \), \( \sigma \text{ research area (DN)} \), and \( \sigma \text{ research area (DN)} \) are the variances of DN values of the pixels in research areas.
areas in bands 1 to 3. \( n_1, n_2, \) and \( n_3 \) are extracting coefficients of bands 1 to 3. By formula (4), we calculated empirical values of \( n_1, n_2, \) and \( n_3 \), as shown in Table 6.

\[
n = \frac{\mu \text{ research area} - (\mu_{\text{BWA}} + 2.58\sigma_{\text{BWA}})}{\sigma \text{ research area}}
\]

Table 6. Extracting coefficients of training images

| Extracting Coefficients | May 11, 2009 | August 21, 2010 | July 28, 2011 | Mean Value |
|-------------------------|--------------|-----------------|--------------|------------|
| \( n_1 \)              | 1.82         | 1.61            | 1.65         | 1.69       |
| \( n_2 \)              | 1.93         | 1.67            | 1.77         | 1.79       |
| \( n_3 \)              | 1.93         | 1.76            | 1.65         | 1.78       |

Table 6 shows that the coefficients are relatively stable. This study used mean values as coefficients to extract BWA. The retrieval algorithm is as follows:

\[
\text{DN}(1) < \mu \text{ research area} (\text{DN}(1)) - 1.69\sigma \text{ research area} (\text{DN}(1)) + 0.5
\]

\[
\text{DN}(2) < \mu \text{ research area} (\text{DN}(2)) - 1.79\sigma \text{ research area} (\text{DN}(2)) + 0.5
\]

\[
\text{DN}(3) < \mu \text{ research area} (\text{DN}(3)) - 1.78\sigma \text{ research area} (\text{DN}(3)) + 0.5
\]

More than 200 linked pixels which meet the criteria in upper formulas are identified as BWA.

4. Results

We first tested the retrieval accuracies of the training images (Figures 4(a) to 4(c)), which were 89.9%, 58.1%, and 73.4%. We then tested the retrieval accuracies of testing images (Figure 4(d)), which was 85.5%. The red frames in the figures represent BWA areas extracted by the algorithm, and the yellow frames represent those extracted by visual interpretation.

This method worked effectively for May 11, 2009; July 28, 2011; and September 24, 2011. However, some areas identified as BWA were not accurate, particularly for August 21, 2010. For example, in Figure 4(b), some small areas judged as BWA areas at the right top of the image were not identified as such in previous research. A possible reason is that previous research estimated the BWA by visual interpolation, which is difficult when the areas are small and discrete. Therefore, previous researchers may amend their studies to relatively larger areas that are easily identified.

![Results of extraction by using the retrieval algorithm.](a) May 11, 2009  (b) August 21, 2010  (c) July 28, 2011  (d) September 24, 2011

5. Discussion and Conclusion

5.1. Discussion
The algorithm was effective for data of May 11, 2009; July 28, 2011; and August 21, 2010. The reason for the errors is likely visual interpretation. The results illustrate that the algorithm realized automatic monitoring of BWA areas. This approach has many advantages. The results are easily achieved with a high computation speed for successful orientation and assessment the BWA areas. However, disadvantages include stability issues and common interference by algal blooms.

5.2. Conclusion
In this study, we masked off research areas susceptible to BWA by using HJ-CCD images. We then analyzed the mean values and variances of gray values in the images of typical BWA areas and research areas. The results showed that gray values of visual bands in the BWA areas are obviously lower than those in research areas. Moreover, we developed a retrieval algorithm (Formulas (5)–(7)) of BWA based on mean values and variances of gray values in the visual bands in the research areas, and we tested the method against experimental satellite images. The results show that the retrieval algorithm can be used to effectively and accurately monitor the locations and regions of BWA. This study is a meaningful step in the operational monitoring of BWA.

However, this research is a preliminary study because no standard exists on the methods used to identify BWA. In addition, errors in visual interpretation limit the accuracy assessment of the retrieval algorithm. Therefore, we will further our research in the following aspects. It is essential to measure the apparent and inherent optical properties to investigate the spectral characteristics of BWA, and we will research the differences among the spectra of BWA in normal water bodies and in algal bloom areas. A more accurate and rational retrieval model of BWA is expected to be developed on the basis of this proposed research.

6. Acknowledgement
The work was jointly supported by National High Technology Research and Development Program of China (Grant No. 2013AA12A302), Key Projects in the National Science & Technology Pillar Program (Grant No. 2011BAH12B07), and China major projects on High Resolution Earth Observation Systems (Grant No. 05-Y30B02-9001-13/15-5).

References
[1] Guihua LU, Qian MA, 2009, Analysis on the Cause of Forming Black Water Cluster in Taihu Lake, *Advances in Water Science*, 20(3): 438-442. (in Chinese)
[2] Dong SHENG, et al., 2010, Cause and Impact Analysis of Black Water Cluster in Taihu Lake, *Water Resource Protection*, 26(3): 41-44 (in Chinese)
[3] Guihua LU, Qian MA, 2009, Monitoring and Analysis on ‘Black Water Aggregation’ in Lake Taihu, *Journal of Lake Science*, 22(4): 481-487 (in Chinese)
[4] Chenglin WANG, et al., 2010, The Research on Warning of Black Water Cluster in Taihu Lake Based on Meteorological Factors, *Environmental Monitoring and Forewarning*, 2(5): 1-4. (in Chinese)
[5] Xuwen LI, Zhichun NIU, 2012, Satellite Remote Sensing Monitoring of Black Color Water Blooms in Lake Taihu, *Management and Technology of Environmental Monitoring*, 24(2): 12-17. (in Chinese)
[6] Min YANG, et al., 2008, Taihu Lake Not to Blame for Wuxi’s Woes, *Science*, 319: 158.
[7] Xuwen LI, Zhichun NIU, 2012, Remote Sensing Monitoring of Black Color Water Blooms in Lake Taihu Based on HJ Satellite CCD Data, *Environmental Monitoring and Forewarning*, 4(3): 1-9. (in Chinese)