Evaluation of the Crosta method for the retrieval of water quality parameters from remote sensing data in the Pearl River estuary
Feng Gao, Yunpeng Wang and Yuanzhi Zhang

ABSTRACT
In recent decades, many algorithms have been developed for the retrieval of water quality parameters using remotely sensed data. However, these algorithms are specific to a certain geographical area and cannot be applied to other areas. In this study, feature-orientated principal component (PC) selection, based on the Crosta method and using Landsat Thematic Mapper (TM) for the retrieval of water quality parameters (i.e., total suspended sediment concentration (TSM) and chlorophyll a (Chla)), was carried out. The results show that feature-orientated PC TSM, based on the Crosta method, obtained a good agreement with the MERIS-based TSM product for eight Landsat TM images. However, the Chla information, selected using the feature-orientated PC, has a poor agreement with the MERIS-based Chla product. The accuracy of the atmospheric correction method and MERIS product may be the main factors influencing the accuracy of the TSM and Chla information identified by the Landsat TM images using the Crosta method. The findings of this study would be helpful in the retrieval of spatial distribution information on TSM from the long-term historical Landsat image archive, without using coincident ground measurements.

Key words | Crosta method, Landsat TM, Pearl river estuary, water quality parameters

INTRODUCTION
In recent years, global studies have shown that it is possible to pair in-situ data collected at discrete stations with remotely sensed data, with the aim of retrieving water quality parameters. Many algorithms have been developed using different kinds of methods (Miller & McKee 2004; Dall’Olmo et al. 2005; Gitelson et al. 2007, 2008; Gilerson et al. 2010; Shen et al. 2010; Xi & Zhang 2010; Mao et al. 2012; Ondrusek et al. 2012; Song et al. 2012; Wu et al. 2013; Chen et al. 2014; Knaeps et al. 2015; Shi et al. 2015; Wu et al. 2015). In recent years, machine learning-based regression algorithms have been applied in the retrieval of water quality parameters from satellite remote sensing images (Tang et al. 2013; Blix et al. 2018; DeLuca et al. 2018; Peterson et al. 2018; Ruescas et al. 2018). These algorithms ranged from linear relationships between satellite remote sensing reflectance and field measurements of water quality parameters (e.g., total suspended sediment concentration (TSM) and chlorophyll a (Chla)) in the estuary and inland waters to region-specific algorithms, explicitly developed for coastal waters, using in-situ measured reflectance and TSM or Chla. For example, Xi & Zhang (2010) established an empirical two-band model using the ratio of remote sensing reflectance at 629 and 671 nm to retrieve the TSM concentration in the Pearl River estuary (PRE). In addition, they used a MERIS image and in-situ remote sensing reflectance to map the distribution of the TSM in the PRE. Xing et al. (2013) used...
in-situ remote sensing reflectance and a combination of Hyperspectral bands using an exponential regression model to estimate the TSM concentration in the PRE, and they obtained a good performance. However, all these algorithms have no uniform model, because they lack a physical foundation, and most of them are therefore geographically dependent and cannot be applied to other areas (Zhang et al. 2010).

In the field of geology, the Crosta method has been used to extract remote sensing alteration information for mineral exploration. The idea of this method is to take advantage of the information contained in Landsat Thematic Mapper (TM) imagery for mineral exploration purposes. The Crosta method is essentially principal component analysis (PCA). An appropriate band combination can be used to conduct PCA, based on the specific spectral characteristics of different minerals in different spectral ranges, with the aim of selecting the principal component (PC) with specific mineral information. In 1989, the Crosta method was proposed and successfully applied to extract iron oxide and mudding anomalies from Landsat TM data (Crosta & Moore 1989). Each PC obtained by PCA using Landsat TM data often has a certain geological significance, which it does not share with any other component, meaning that each PC has unique characteristics. For example, the criterion for judging the PC of iron-stained minerals is based on the four bands of TM1, TM3, TM4 and TM5. The characteristic vector of the PC should be constituted by, and its TM3 coefficient should be the opposite of that of TM1 and TM4. The coefficient symbol of TM3 is generally the same as that of TM5. Based on these criteria, the information on iron dyeing contained in the PC can be selected. Thus, this selected PC can be called the anomalous PC of iron dyeing. Crosta uses these spectral characteristics and combinations of several TM bands to diagnose remote sensing alteration information for iron exploration (Crosta & Moore 1989).

By the end of 2012, the Landsat TM instrument had obtained a large amount of data on the Earth's surface over 28 years (Wu et al. 2013). The image archive collected by the Landsat TM is historically unique, and it provides excellent opportunities for us to monitor and analyze the long-term spatiotemporal dynamics of the Earth's surface parameters, like TSM and Chla.

The objective of this study is to use Landsat TM imagery to evaluate the Crosta method in the retrieval of water quality parameters (e.g., TSM and Chla) in the PRE. Firstly, the spectral characteristics of TSM and Chla in the visible and near-infrared (NIR) band are described. Secondly, appropriate Landsat TM bands for diagnosing the TSM and Chla of the water components are selected. Thirdly, the effect of the combinations of eigenvectors with reflectance values of spectral bands on PC images is analyzed. Finally, an analysis of the correlation between the feature-orientated PC and corresponding MERIS products (TSM and Chla) is presented.

**MATERIALS AND METHODS**

**Study area**

The Pearl River is the second largest river in China and is considered one of the most complicated fluvial networks in the world. It is comprised of three major tributaries (the western, northern and eastern rivers) and other small rivers, draining into the PRE, which occupies an area of ~17,200 km². It plays a vital role in supplying fresh water to the large cities in the Pearl River Delta region, such as Macau, Hong Kong, Zhuhai, and Guangzhou. The annual suspended sediment load of the river is 88.7 Mt/y. The annual average flowrate is $2.281 \times 10^8$ m³/y, and the annual suspended-sediment discharge is $6,567 \times 10^4$ t/y for the western river. The annual average water and suspended sediment discharges for the northern river are $449 \times 10^4$ m³/y and $864 \times 10^4$ t/y, respectively, while those for the eastern river are $234 \times 10^6$ m³/y and $236 \times 10^4$ t/y, respectively (Wu et al. 2016). The total water and sediment discharges of these three major rivers account for more than 80% and 95% of the total load entering the sea, respectively. As a result, there has been a seaward extension of the mouth region of the PRE at a rate of 40 m per year (Wu et al. 2016). Soil erosion and soil weathering are very serious in the upstream of Pearl River. Sampling points were randomly selected in the PRE (Figure 1), which can be used for validation and statistical analysis.

**Image acquisition and pre-processing**

Landsat TM scenes with no or minimal cloud cover (Row/Path: 44/122) were downloaded from the United States

Downloaded from http://iwaponline.com/wqrj/article-pdf/55/2/209/709563/wqrj0550209.pdf
Geological Survey (USGS) website (https://earthexplorer.usgs.gov/). Standard MERIS FR level 2 products were obtained from the Europe Space Agency (ESA) website (https://merisFR-merci-ds.eo.esa.int/). Detailed information can be found in Table 1.

In this study, atmospheric correction was performed on Landsat TM images using the FLAASH atmospheric correction model with the ENVI 5.3.1 software. The pixel surface reflectance for all Landsat bands was retrieved. Due to the limitation of in-situ measurements, the results of the atmospheric correction were not analyzed and evaluated in this study. The MERIS sensor is designed mainly for ocean and coastal water remote sensing, with 15 narrow spectral bands in a range of 390–1,040 nm and a revisit period of one to three days. The eight bands centered at 412, 442, 490, 510, 560, 620, 665 and 708 nm were used, along with a neural network, to derive the MERIS level 2 products (Kratzer et al. 2013). MERIS FR level 2 products were projected into Geographic Lat/Log (WGS84). The sample points that matched with the feature-orientated PC were extracted using Beam 4.8 software (http://www.brockmann-consult.de/cms/web/beam/welcome). The average PC values that matched with the sample points were extracted using a 3 by 3 window on specific PC images. Detailed information on the bands and spectral range can be found in Table 2.

A single MERIS FR level 2 product has many bands, including TSM, Chla, yellow substance and other parameters. Thus, TSM and Chla images in the PRE could be extracted from MERIS FR level 2 products on Oct. 17, 2003 and Oct. 17, 2009 using the Beam 4.8 software. These datasets are assumed as the ground measurements and are further employed to validate the performance of the feature-orientated PCs.

### Spectral characteristics of TSM and Chla

Water components, TSM and Chla, have unique spectral characteristics in the electromagnetic spectrum, and water reflectance is highly variable over the visible and NIR spectral regions. The main source of TSM is suspended sediment, full of organic and inorganic particles (Xing et al. 2013). The spectral characteristic of TSM are very similar to those of suspended sediment. There are two typical reflectance peaks in the range of 500–600 nm and around 800 nm (Cai et al. 2013). The level of phytoplankton activity within

| Date       | Sensor     | Region | Row/Path | Purpose  |
|------------|------------|--------|----------|----------|
| Oct. 17, 2003 | Landsat TM | PRE    | 44/122   | PCA      |
| Oct. 17, 2003 | MERIS      |        |          | Validation |

### Table 2 | Bands and spectral information on TM and MERIS data

| Band number | Landscat TM Band Information | MERIS FR Band Information |
|-------------|------------------------------|---------------------------|
| 1           | 450–520 nm                   | 407.5–417.5 nm            |
| 2           | 520–580 nm                   | 437.5–447.5 nm            |
| 3           | 630–690 nm                   | 485–495 nm                |
| 4           | 760–900 nm                   | 505–515 nm                |
| 5           | 1,550–1,750 nm               | 555–565 nm                |
| 6           | 1,040–1,250 nm               | 615–625 nm                |
| 7           | 2,080–1,350 nm               | 660–670 nm                |
| 8           | 677.5–685 nm                 | 677.5–685 nm              |
| 9           | 700–710 nm                   | 700–710 nm                |
| 10          | 750–757.5 nm                 | 750–757.5 nm              |
| 11          | 758.75–761.25 nm             | 758.75–761.25 nm          |
| 12          | 767.5–782.5 nm               | 767.5–782.5 nm            |
| 13          | 855–875 nm                   | 855–875 nm                |
| 14          | 885–895 nm                   | 885–895 nm                |
| 15          | 895–905 nm                   | 895–905 nm                |
a water body, which can be represented by phytoplankton biomass, is one of the most important considerations in water monitoring (Augusto-Silva et al. 2014). The changes of phytoplankton biomass are related to the concentration of phytoplankton’s main and major pigment: Chla. The phytoplankton pigment exhibits a unique spectral absorption signature, with two marked peaks: one in the blue region of the spectrum (~443 nm) and another one in the red region of the spectrum (~686 nm) (Bernardo et al. 2016). Its main reflectance peak is in the green of the spectrum (~550 nm), with a secondary peak at around 715 nm, which is associated with a high scattering of the phytoplankton cells (Gitelson 1992; Gitelson et al. 2007). According to the nature of the Crosta method and characteristics of TSM and Chla, the information on TSM and Chla contained in the PC can be selected appropriately.

**Principal component analysis**

Essential to the Crosta method is PCA. PCA is designed to transform the original variables into transformed, uncorrelated variables, called PCs, which are linear combinations of the original variables. In digital image analysis, PCA is an important technique that can have several functions. It is used to reduce the volume of large datasets, while retaining the majority of the image information; to extract a series of uncorrelated spectral components from multispectral and hyper-spectral imagery; and to enhance image features for interpretation (Bierman et al. 2011). Different water constituents in coastal waters and estuaries are contained in different PCs. The PCs are ordered from the largest variance (PC1) to the next largest variance (PC2), and so on. Algebraically, for p original variables, \(x_1, x_2, \ldots x_p\),

\[
PC1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p = \sum_{j=1}^{p} a_{1j}x_j \\
PC2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2p}x_p = \sum_{j=1}^{p} a_{2j}x_j
\]

This sequence continues for all p PCs. The first few PCs will tend to contain (or explain) a large percentage of the total variance and may be used to describe multivariate patterns or variance in water quality. Often these patterns are related to specific sources of contamination (Olsen et al. 2012). For example, PCA is carried out in six bands of TM1, TM2, TM3, TM4, TM5 and TM7. The criteria for judging the PC of TSM is that the TM1, TM2, TM3, TM4, TM5 and TM7 coefficients of the PC should generally be positive, and the coefficients of TM1, TM2, TM3 and TM4 should be significantly larger than those of TM5 and TM7. The criterion for judging the PC of Chla is that the TM1 and TM3 coefficients of the PC should be significantly larger than those of the other bands, and the TM1 coefficient of the PC should generally be the opposite of that of the TM3 coefficient symbols.

**Statistical analysis and accuracy assessment**

An analysis of the correlation between the feature-orientated PC and ground truth TSM and Chla was conducted using Pearson’s correlation coefficient, with a statistical significance of \(p < 0.0001\). The significance level of 0 (less than 0.0001) indicated, in this study, that there are strong correlations between the feature-orientated PC and satellite-based TSM and Chla.

The root mean square error (RMSE, Equation (1)) was calculated to evaluate the performance of the feature orientated PC, where \(X_{\text{Mea},i}\) and \(X_{\text{Est},i}\) are the MERIS-based and the feature-orientated PC values for sample i. n is the total number of samples. Smaller RMSE values indicated a higher overall accuracy and lower predicted error of the Crosta method.

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (X_{\text{Mea},i} - X_{\text{Est},i})^2}{n}}
\]

**RESULTS AND DISCUSSION**

**Feature-orientated principal component selection**

PCA was conducted using pre-processed Landsat TM images (excluding the TM6 band), with only water body pixels included. Feature-orientated PC selection involves examining the eigenvectors used in calculating the PCs and relating each PC image to the two or three original bands that contribute most data to it (Crosta & Moore 2011).
Specific water components (e.g., TSM or Chla) can then be selected based on the major contributions (negative or positive eigenvectors) from the original bands that are likely to display the desired water quality parameters. The eigenvectors calculated for Landsat TM on October 17, 2003 are presented in Table 3. In Table 3, the eigenvectors are expressed as a percentage of the loading from the original bands. The transformation of the original eigenvalue values to percentages makes the understanding of the original band contributions to each PC more straightforward. The knowledge on the spectral characteristics of TSM and Chla in the Landsat TM bands was used to define and select the PCs containing spectral information due to these water components.

As shown in Table 3, PC1 reveals a significant spectral feature, since it is composed of a positive mixture of TM1–TM4 bands, which, in this case, has a greater proportion of TM2 (29.91%) and TM3 (33.88%) showing TSM information. In the PRE, the reflectance curves of suspended sediment at different concentrations showed that with the increase of the suspended sediment concentration, an increase of the reflectance in the visible and NIR range can be clearly defined. At around 580 nm, the reflectance curves of the suspended sediment concentration can effectively differentiate different suspended sediment concentrations. Additionally, Xi and Zhang demonstrated that the red band is sensitive to changes of TSM and can be used to retrieve TSM concentrations successfully in the PRE (Xi & Zhang 2010). Thus, PC1 can be regarded as a measure of TSM information, according to the water volume. The distribution of the PC1 image is presented in Figure 2(a). PC2 is dominated by the contribution of TM4 (−25.70%), TM5 (−24.04%) and TM7 (−19.09%), showing a water body spectral response due to the high absorption in the NIR and short-wave infrared range. PC5 is dominated mainly by the contribution of TM1 (33.16%) and TM5 (−24.41%), showing Chla spectral information due to the high absorption in the blue and red range, since absorption features can be found in the blue and red ranges due to the high absorption in Chla (Bernardo et al. 2016). The distribution of the inverse PC5 image is presented in Figure 2(b). PC4 does not reveal any significant spectral feature, since it is composed of a positive and negative mixture of all bands, and several original bands contribute balanced proportions. PC5 is mainly formed by two negative loadings, from TM1 (−13.65%) and TM3 (−23.48%), and two positive loadings, from TM2 (24.76%) and TM4 (22.41%). As Chla has a high absorption in the blue and red range and a high reflectance in the green and NIR range (Bernardo et al. 2016), we are able to consider PC5, as a measure of Chla is left unexplained by PC3. PC6 is mainly formed by the contribution of TM5 (−43.13%) and TM7 (49.56%), which is considered to be spectral information on minerals, according to the water volume (Crosta & Moore 1989).

### Effect of combinations of eigenvectors with reflectance values of spectral bands on PC images

Each PC image is generated as the sum of the products of the eigenvectors and reflectance values for respective spectral bands at each pixel (Gupta et al. 2015), i.e.,

\[
\text{PC image pixel value} = \text{SUM} (\text{Eigenvectors for the band } i \times \text{Pixel reflectance value of the band } i)
\]

From the formula of the PC image pixel value, a substantial numeric value of an eigenvector (positive or negative) has a higher influence on the PC image. A low numeric value of the eigenvector means that the particular band has a relatively low significance for that PC. Here, we consider the effects of large (positive or negative) eigenvectors, as combined with a high or low reflectance value of the image (which, in turn, imply a strong reflectance or strong absorption, respectively). Low values of the eigenvectors are not considered, since these would have a minimal influence on the target PC image.

### Table 3 | PC loadings, drawn from Landsat TM reflectance bands on Oct. 17, 2003, expressed in percentages

|   | TM1 | TM2 | TM3 | TM4 | TM5 | TM7 |
|---|-----|-----|-----|-----|-----|-----|
| PC1 | 16.59 | 29.91 | 33.88 | 14.36 | 3.02 | 2.24 |
| PC2 | 13.62 | 13.37 | −4.18 | −25.70 | −24.04 | −19.09 |
| PC3 | 33.16 | 7.44 | −24.41 | −2.30 | 17.78 | 14.91 |
| PC4 | 21.60 | −21.54 | −0.21 | 26.03 | −15.55 | −15.08 |
| PC5 | −13.65 | 24.76 | −23.48 | 22.41 | −6.62 | −9.08 |
| PC6 | −0.43 | 0.97 | −1.81 | 4.10 | −45.13 | 49.56 |
An image data set (Figure 3(a)) and PC1 subset (Figure 3(b)) of Landsat TM, taken on Oct. 17, 2003, are presented in Figure 3. Figure 3(c) shows the eigenvector of the PC1 image, and Figure 3(d) presents the spectral curve for the pixel under consideration. The spectral bands, TM1, TM2, TM3 and TM4, have very high positive eigenvectors, and all of the other bands have small positive eigenvectors (Figure 3(c)). The pixel in the crosshair, presented in Figure 3(a), is characterized by a very strong reflectance in TM2 and TM3 and a weaker reflectance in TM5 and TM7. Thus, in the resulting PC1 image, due to the combination of the large positive eigenvector in TM2 and TM3 and the very high reflectance value in the sub-scene of Figure 3(a), the pixel in the crosshair appears bright (Figure 3(b)). This means that the TSM concentration of the pixel is very high in the crosshair shown in Figure 3(a).

A second pixel, selected in the sub-scene (Figure 4(a)) and PC1 image (Figure 4(b)) of Landsat TM on Oct. 17, 2003, is presented in Figure 4. Figure 4(c) presents the spectral curve for the pixel under consideration. The eigenvectors of the PC1 image can be found in Figure 3(c). The eigenvectors of PC1 are the same as those of the pixel shown in Figure 4, while the reflectance values of all bands are smaller than the pixel shown in Figure 3. Thus, in the resulting PC1 image, the pixel in the crosshair appears dark (Figure 4(b)). In this case, the TSM concentration of the pixel is relatively lower than that of the pixel shown in Figure 3.

**Validation of the feature-orientated principal component**

To evaluate the performance of the feature-orientated PC, 66 satellite-based samples from the MERIS FR level 2 TSM product were used to directly validate the results of the Crosta method. The distribution of the TSM and Chla extracted from the MERIS FR level 2 product are presented in Figure 5. The results show that the PC1 generated from the Landsat TM, taken on Oct. 17, 2003, and the MERIS TSM product are strongly correlated, with a highly significant linear relationship ($R^2 = 0.83$, $p$-value $< 0.0001$, RMSE $= 4.77$ g/m$^3$) (Figure 6(a)). This indicates an acceptable level of consistency for the Landsat TM, taken on Oct. 17, 2003.

As for another water quality parameter, Chla, PC3, selected from the Landsat TM, is regarded as a measure of Chla, according to the water volume. To validate the performance of PC3, 66 satellite-based samples from the MERIS FR level 2 Chla product were used to validate
the results of the Crosta method directly. The results show that there is a reliable correlation between PC3 and the MERIS Chla product ($R^2 = 0.74$, $p$-value $< 0.0001$, RMSE $= 2.88$ mg/m$^3$) (Figure 6(b)).

Considering the distribution of PC1 (Figure 2(a)), selected based on the feature-orientated PC from the Landsat TM, the spatial characteristics of the PC1 image are consistent with the MERIS TSM product (Figure 5(a)). From atmospherically corrected images, taken on Oct. 17, 2003 (Figure 3(a) and Figure 4(a)), the spatial distribution of suspended sediment can clearly be observed from visual inspection, which is also consistent with the spatial distribution characteristics of the PC1 image. The high values for the TSM concentration occurred on the west coast of the PRE, while the low values for the TSM concentration were present in the middle and on the east coast of the PRE. These spatial characteristics of TSM can be observed in both the PC1 image and the MERIS TSM product. The spatial distribution characteristics of the PC3 image are consistent with those of the MERIS Chla product. The validation results show a good correlation between the PC3 image and the MERIS Chla product. Overall, TSM and Chla in the PRE present similar spatial distribution characteristics, high values of TSM and Chla always occur on the west coast of the PRE, and low values are mainly distributed in the middle and on the east coast.
In order to test the stability of the Crosta method in identifying the water quality parameters (TSM and Chla) from the Landsat TM images, another seven Landsat TM images, which matched with the acquisition date of the MERIS FR level 2 product, were tested. The correlation coefficient ($r$), determination coefficient ($R^2$) and RMSE of the feature-orientated PC and MERIS product (TSM and Chla) were calculated and are listed in Table 4. TSM, identified by the feature-orientated PC (TSM) using Landsat TM images, demonstrated that the Crosta method presented a good performance in the retrieval of the spatial distribution of TSM in the PRE. However, a poor performance (except for Mar. 4, 2008 and Oct. 17, 2009) was obtained for the feature-orientated PC (Chla) and MERIS-based Chla product. Overall, the feature-orientated PC (TSM) can effectively
obtain the spatial distribution of TSM, while Chla is unstable using the Crosta method and Landsat TM images. It should be noted that the reflectance in the visible and NIR spectral range on surface water is very small in the remote sensing imagery. Additionally, the effect of the atmosphere makes the reflectivity of the water weaker. In the optically complex estuary and coastal waters, suspended sediment is the strongest constituent of scattering and reflection, which may characterize more information on remote sensing images. On the contrary, a high reflectance of suspended sediment masks the Chla information due to the strong absorption of Chla and pigments in the PRE.

**Limitations**

In this study, the Crosta method, essential for PCA, was used to identify and select water quality parameters (TSM and Chla) from Landsat TM images and validate the results using MERIS FR level 2 TSM and Chla products. PCA is a dimension reduction technique, which can transform many original interrelated variables into fewer, uncorrelated variables, called PCs. The first few PCs provide meaningful information on some specific parameters. Each PC can represent the most meaningful parameter, according to the coefficients of the loadings on each original variable. Thus, each PC may contain the most meaningful parameter and
other unremarkable parameter information due to the PC being a linear combination of original variables. In the visible and NIR spectral range, spectral absorption and a reflectance signature, with two or three marked peaks in TSM and Chla, exist. In addition, there is also an impact on unmarked peaks for TSM and Chla. Therefore, PC contained the most meaningful TSM and Chla information identified by the Crosta method and also included other information on water quality parameters, i.e., it did not only contain information on TSM. This information on TSM and Chla cannot be used for the spatial and temporal analysis of TSM and Chla in the PRE. Instead, it can only be used to analyze the spatial distribution of TSM and Chla qualitatively, when in-situ measurements are not available for deriving satellite-based water quality parameters. While previous studies demonstrated that the FLAASH atmospheric correction method was successfully applied in the retrieval of surface reflectance on Landsat TM images in the estuary and coastal waters, where in-situ measurements of TSM are not available in the study period. The findings presented in this study would be helpful in the retrieval of TSM information from the long-term historical Landsat image archive. In future studies, other atmospheric correction methods should be conducted and validated in relation to the Landsat TM/OLI images using in-situ measurements in the PRE.

**CONCLUSIONS**

In this study, the Crosta method was used to investigate the applicability of water quality parameters (TSM and Chla) identified by Landsat TM images. Feature-orientated PC TSM obtained a good performance, compared with the MERIS-based TSM product. However, the Chla information selected from the feature-orientated PC demonstrated that the Chla information derived from the Landsat TM using the Crosta method is unstable. A poor agreement can be observed between the feature-orientated PC Chla and the MERIS-based Chla product. Chla and pigments in the estuary and coastal waters are dominated by absorption, and the information obtained by remote sensing data is not rich. Additionally, this may be caused by the poor accuracy of the MERIS-based Chla product. Atmospheric correction may be another factor associated with this problem.

Overall, the Crosta method is an alternative method for the retrieval of spatial distribution information on TSM from Landsat TM images in the optically complex estuary and coastal waters, where in-situ measurements of TSM are not available in the study period. The findings presented in this study would be helpful in the retrieval of TSM information from the long-term historical Landsat image archive. In future studies, other atmospheric correction methods should be conducted and validated in relation to the Landsat TM/OLI images using in-situ measurements in the PRE.

**ACKNOWLEDGEMENTS**

This study was supported by the China NSFC (U1901215) and 863 Program (2006AA06A306), and the Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (2019L0485). We would like to extend our gratitude to Jamie Robert Beagle for his help in editing the English language of this paper.

---

**Table 4** Correlation coefficient (r), determination coefficient (R²) and RMSE for the feature-orientated PC and MERIS product (TSM and Chla)

| Date       | WQ | r   | R²  | RMSE |
|------------|----|-----|-----|------|
| Dec.20,2003| TSM| 0.91| 0.82| 4.57 |
|            | Chla| -0.32| 0.11| 3.46 |
| Mar.4,2008 | TSM| 0.82| 0.66| 3.42 |
|            | Chla| -0.62| 0.39| 1.82 |
| Jul.26,2008| TSM| 0.84| 0.7 | 6.13 |
|            | Chla| -0.1| 0.01| 5.37 |
| Oct.17,2009| TSM| 0.9 | 0.8 | 3.47 |
|            | Chla| -0.64| 0.41| 3.17 |
| Nov.2,2009 | TSM| 0.9 | 0.81| 4.14 |
|            | Chla| -0.47| 0.22| 4.12 |
| Dec.4,2009 | TSM| 0.78| 0.61| 4.17 |
|            | Chla| 0.12| 0.01| 2.34 |
| Dec.23,2010 | TSM| 0.78| 0.61| 5.38 |
|            | Chla| -0.47| 0.22| 4.02 |

Note: WQ means water quality parameters.
REFERENCES

Augusto-Silva, P., Ogashawara, I., Barbosa, C., de Carvalho, L., Jorge, D., Fornari, C. & Stech, J. 2014 Analysis of MERIS reflectance algorithms for estimating chlorophyll-a concentration in a Brazilian Reservoir. Remote Sensing 6 (12), 11689–11707.

Bernardo, N., Watanabe, F., Rodrigues, T. & Alcântara, E. 2016 Evaluation of the suitability of MODIS, OLCI and OLI for mapping the distribution of total suspended matter in the Barra Bonita Reservoir (Tietê River, Brazil). Remote Sensing Applications: Society and Environment 4, 68–82.

Bierman, P., Lewis, M., Ostendorf, B. & Tanner, J. 2011 A review of methods for analysing spatial and temporal patterns in coastal water quality. Ecological Indicators 11, 103–114.

Blix, K., Pålly, K., Töth, V. & Eltoft, T. 2018 Remote sensing of water quality parameters over Lake Balaton by using Sentinel-3 OLCI. Water 10 (10), 1428.

Cai, L., Tang, D. & Li, C. 2015 An investigation of spatial variation of suspended sediment concentration induced by a bay bridge based on Landsat TM and OLI data. Advances in Space Research 56, 293–303.

Chen, J., Cui, T., Qiu, Z. & Lin, C. 2014 A three-band semi-analytical model for deriving total suspended sediment concentration from HJ-1A/CCD data in turbid coastal waters. ISPRS Journal of Photogrammetry and Remote Sensing 93, 1–13.

Crosta, A. & Moore, J. 1989 Enhancement of Landsat Thematic Mapper imagery for residual soil mapping in S W Minais Gerais State, Brazil: a prospecting case history in Greenstone belt terrain. In: Proceedings of the 7th ERIM Thematic Conference: Remote Sensing for Exploration Geology, pp. 1173–1187.

Dall’Olmo, G., Gitelson, A. A., Rundquist, D. C., Leavitt, B., Barrow, T. & Holz, J. C. 2005 Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. Remote Sensing of Environment 90, 176–187.

DeLuca, N., Zaitchik, B. & Curriero, F. 2018 Can multispectral information improve remotely sensed estimates of total suspended solids? A statistical study in Chesapeake Bay. Remote Sensing 10 (9), 1393.

Gilerson, A. A., Gitelson, A. A., Zhou, J., Gurlin, D. & Moses, W. 2010 Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. Optical Society of America 18, 24109–24125.

Gitelson, A. 1992 The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. International Journal of Remote Sensing 13 (17), 3367–3373.

Gitelson, A. A., Schalles, J. F. & Hladik, C. M. 2007 Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. Remote Sensing of Environment 109, 464–472.

Gitelson, A. A., Dall’Olmo, G., Moses, W., Rundquist, D. C., Barrow, T., Fisher, T. R., Gurlin, D. & Holz, J. 2008 A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. Remote Sensing of Environment 112, 3582–3593.

Guanter, L., Gonzalez-Sanpedro, M. D. & Moreno, J. 2007 A method for the atmospheric correction of ENVISAT/MERIS data over land targets. International Journal of Remote Sensing 28, 709–728.

Guanter, L., Ruiz-Verdu, A., Odermatt, D., Giardino, C., Simis, S. & Estelles, V. 2010 Atmospheric correction of ENVISAT/MERIS data over inland waters: Validation for European lakes. Remote Sensing of Environment 114, 467–480.

Gupta, R. P., Tiwari, R. K., Saini, V. & Srivastava, N. 2013 A simplified approach for interpreting principal component images. Advances in Remote Sensing 02, 111–119.

Knaeps, E., Ruddick, K. G., Doxaran, D., Dogliotti, A. L., Nechad, B., Raymackers, D. & Sterckx, S. 2015 A SWIR based algorithm to retrieve total suspended matter in extremely turbid waters. Remote Sensing of Environment 166, 66–79.

Krätzer, S., Brockmann, C. & Moore, G. 2008 Using MERIS full resolution data to monitor coastal waters—a case study from Himmerfjarden, a fjord-like bay in the northwestern Baltic Sea. Remote Sensing of Environment 112, 2284–2300.

Mao, Z., Chen, J., Pan, D., Tao, B. & Zhu, Q. 2012 A regional remote sensing algorithm for total suspended matter in the East China Sea. Remote Sensing of Environment 124, 819–831.

Miller, R. L. & McKee, B. A. 2004 Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters. Remote Sensing of Environment 93, 259–266.

Olsen, R. L., Chappell, R. W. & Loftis, J. C. 2012 Water quality sample collection, data treatment and results presentation for principal components analysis—literature review and Illinois River Watershed case study. Water Research 46, 3110–3122.

Ondrusk, M., Stengel, E., Kinkade, C. S., Vogel, R. L., Keegstra, P., Hunter, C. & Kim, C. 2012 The development of a new optical total suspended matter algorithm for the Chesapeake Bay. Remote Sensing of Environment 119, 243–254.

Peterson, K., Sagan, V., Sidike, P., Cox, A. & Martinez, M. 2018 Suspended sediment concentration estimation from Landsat Imagery along the Lower Missouri and Middle Mississippi Rivers using an extreme learning machine. Remote Sensing 10 (10), 1503.

Ruescas, A., Hieronymi, M., Mateo-Garcia, G., Koponen, S., Kallio, K. & Camps-Valls, G. 2018 Machine learning regression approaches for colored dissolved organic matter (CDOM) retrieval with S2-MSI and S3-OLCI simulated data. Remote Sensing 10 (5), 786.

Shen, F., Verhoef, W., Zhou, Y., Salama, M. S. & Liu, X. 2010 Satellite estimates of wide-Range suspended sediment
concentrations in Changjiang (Yangtze) Estuary using MERIS data. *Estuaries and Coasts* **33**, 1420–1429.

Shi, K., Zhang, Y., Zhu, G., Liu, X., Zhou, Y., Xu, H., Qin, B., Liu, G. & Li, Y. 2015 Long-term remote monitoring of total suspended matter concentration in Lake Taihu using 250 m MODIS-Aqua data. *Remote Sensing of Environment* **164**, 43–56.

Song, K., Li, L., Wang, Z., Liu, D., Zhang, B., Xu, J., Du, J., Li, L., Li, S. & Wang, Y. D. 2012 Retrieval of total suspended matter (TSM) and chlorophyll-a (Chl-a) concentration from remote-sensing data for drinking water resources. *Environmental Monitoring and Assessment* **184**, 1449–1470.

Stevenson, R. J., Qi, J., Novitski, L. N. & Lin, S. 2018 Landsat TM/ETM+ and machine-learning algorithms for limnological studies and algal bloom management of inland lakes. *Journal of Applied Remote Sensing* **12** (02), 1.

Tang, S., Larouche, P., Niemi, A. & Michel, C. 2013 Regional algorithms for remote-sensing estimates of total suspended matter in the Beaufort Sea. *International Journal of Remote Sensing* **34** (19), 6562–6576.

Wu, G., Cui, L., Duan, H., Fei, T. & Liu, Y. 2015 An approach for developing Landsat-5 TM-based retrieval models of suspended particulate matter concentration with the assistance of MODIS. *ISPRS Journal of Photogrammetry and Remote Sensing* **85**, 84–92.

Wu, G., Cui, L., Liu, L., Chen, F., Fei, T. & Liu, Y. 2015 Statistical model development and estimation of suspended particulate matter concentrations with Landsat 8 OLI images of Dongting Lake, China. *International Journal of Remote Sensing* **36**, 343–360.

Wu, C. S., Yang, S., Huang, S. & Mu, J. 2016 Delta changes in the Pearl River estuary and its response to human activities (1954–2008). *Quaternary International* **392**, 147–154.

Xi, H. & Zhang, Y. 2010 Total suspended matter observation in the Pearl River estuary from in situ and MERIS data. *Environmental Monitoring and Assessment* **177**, 563–574.

Xing, Q., Lou, M., Chen, C. & Shi, P. 2013 Using in situ and satellite hyperspectral data to estimate the surface suspended sediments concentrations in the Pearl River Estuary. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **6** (2), 731–738.

Zhang, M., Tang, J., Dong, Q., Song, Q. & Ding, J. 2010 Retrieval of total suspended matter concentration in the yellow and east China seas from MODIS imagery. *Remote Sensing of Environment* **114**, 392–403.

First received 27 May 2018; accepted in revised form 22 January 2020. Available online 19 February 2020.