An approach for developing Landsat-5 TM-based retrieval models of suspended particulate matter concentration with the assistance of MODIS

Guofeng Wu,*, Lijuan Cui, Hongtao Duan, Teng Fei, Yaolin Liu

*Shenzhen Key Laboratory of Spatial Information Smart Sensing and Services, College of Life Sciences, Shenzhen University, 518060 Shenzhen, China
School of Resource and Environmental Science, Key Laboratory of Geographic Information System of the Ministry of Education, Wuhan University, 430079 Wuhan, China
Institute of Wetland Research, Chinese Academy of Forestry, 100091 Beijing, China
State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, 210008 Nanjing, China

A R T I C L E   I N F O
Article history:
Received 19 April 2013
Received in revised form 20 August 2013
Accepted 22 August 2013

Keywords:
Water quality
Remote sensing
Model development
Exponential model

A B S T R A C T
It is challenging to develop Landsat-5 TM (T5) image-based retrieval models for estimating the suspended particulate matter concentration ($C_{SPM}$) in water when missing coincident ground $C_{SPM}$ measurements. This study, with the Poyang Lake in China as a case study, proposed an approach for developing T5-based $C_{SPM}$ retrieval models with the assistance of moderate resolution imaging spectroradiometer (MODIS) images. After validation with an independent dataset, a cubic $C_{SPM}$ retrieval model of 250 m MODIS red band was used to estimate the $C_{SPM}$ values at 100 sampling points from the MODIS images (MODIS-based $C_{SPM}$) captured at three time periods. The MODIS-based $C_{SPM}$ values at the time period with the largest $C_{SPM}$ variation were combined with their coincident T5 image reflectance for T5-based model calibrations. The linear, quadratic, cubic, power and exponential models of MODIS-based $C_{SPM}$ against T5 single bands and their combinations were calibrated, respectively. Four best-fitting T5-based $C_{SPM}$ models were selected to retrieve the $C_{SPM}$ values at 100 sampling points from the T5 images (T5-based $C_{SPM}$) at the other two time periods, and the coincident MODIS- and T5-based $C_{SPM}$ values were compared to assess T5-based model performances. Model calibration results showed that the cubic and exponential models of T5 red band (band 3) and red subtracting mid-infrared band (band 5) obtained the best fitting for estimating $C_{SPM}$ from the T5 image on 12 August 2005, and they explained 94–97% of the variation of MODIS-based $C_{SPM}$ values with an estimated standard error of 6.617–8.457 mg/l. Model validations indicated that the exponential model of T5 red band got the best result for estimating $C_{SPM}$ from T5 images when the MODIS-based $C_{SPM}$ values were assumed as ground truths (correlation coefficient between MODIS- and T5-based $C_{SPM}$ values = 0.96, root mean square error = 4.60 mg/l). We concluded that the T5-based $C_{SPM}$ retrieval models could be developed with the assistance of MODIS, and the approach proposed in this study will be helpful for other researchers who also want to retrieve $C_{SPM}$ from T5 image archive but without coincident ground $C_{SPM}$ measurements.

© 2013 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.

1. Introduction

Multiple functions are provided by lakes (Jorgensen et al., 2005), and they are generally influenced by water quality. Suspended particulate matter (SPM) is a key parameter describing water quality (Pozdnyakov et al., 2005; Zhang et al., 2003), and it has important impacts on water bodies, such as transporting nutrients and contaminants as well as reducing light transmission through water column (Cigizoglu and Kisi, 2006; Kirk, 1994), which may influence the whole aquatic ecosystem (Rorslett, 1996). Therefore, obtaining the spatiotemporal information of SPM concentration ($C_{SPM}$) is necessary for understanding, managing and protecting lake ecosystem.

Landsat-5 satellite was launched on 1 March 1984 (Loveland and Dwyer, 2012), and its thematic mapper (TM) instrument obtained about 28-year massive data of the Earth’s surface by the end of 2011 due to its remarkable performance. Such an image archive collected by a single sensor is unique historically, and it provides great opportunities for us to monitor and analyze
long-term spatiotemporal dynamics of the Earth’s surface parameters like \( C_{\text{SPM}} \). Landsat-5 TM (TM5) images have limitations in operational use for regularly monitoring \( C_{\text{SPM}} \) because \( C_{\text{SPM}} \) is highly dynamic and may not be effectively tracked by a satellite with a 16-day repeating cycle. Moreover, frequent cloud cover reduces the availability of suitable images in many parts of the world, which further makes TM5 images inappropriate for regularly monitoring very dynamic variables (Kloiber et al., 2002). However, due to its high spatial resolution, good quality and especially long-term archive, TM5 image has great potentials in retrieving long-term historical \( C_{\text{SPM}} \) values.

Many studies have applied TM5 images to estimate the \( C_{\text{SPM}} \) or suspended sediment concentrations of case II waters (include most inland and coastal waters and are dominated by phytoplankton, inorganic particle and yellow substance together), such as at the Lake Balaton in central Europe (Tyler et al., 2006), Lake Kasumigaura in Japan (Oyama et al., 2009), Bhopal Upper Lake in India (Durga Rao et al., 2009), Taihu Lake (Chen et al., 2010) and Yangtze Estuary (Li et al., 2010) in China, Bombetoka Bay and Betisiboka Estuary in Madagascar (Raharimahafa and Kusky, 2010), Lake Beysehir in Turkey (Nas et al., 2010), Bung Boraphet in Thailand (Sriwongsitanon et al., 2011) and Venice Lagoon in Italy (Volpe et al., 2011). Some TM5-based \( C_{\text{SPM}} \) retrieval models were developed for case II waters; however, no uniform model exists.

Semi-analytical models might be suitable for different water bodies for water quality (including \( C_{\text{SPM}} \) ) retrievals due to their strict theories (Giardino et al., 2007; Ma et al., 2010). They, however, are not frequently employed because of the difficulties or inaccuracies on obtaining inputting parameters (Ma et al., 2010). Therefore, most TM5-based \( C_{\text{SPM}} \) models are empirical or semi-empirical ones, which are often region- and time-dependent because of the different optical properties at different water bodies or different time periods. The developments of empirical or semi-empirical models generally need coincident ground \( C_{\text{SPM}} \) measurements and satellite image (the coincident in this study means that the \( C_{\text{SPM}} \) and satellite image are measured and captured, respectively, nearly on the same day). However, it is very difficult or even impossible to obtain coincident ground \( C_{\text{SPM}} \) measurements and TM5 image with good quality due to the 16-day repeating cycle and retirement of Landsat-5 satellite as well as frequent cloud cover. Therefore, without coincident ground \( C_{\text{SPM}} \) measurements and TM5 image, it is challenging to develop TM5-based \( C_{\text{SPM}} \) retrieval models and further to retrieve and analyze the \( C_{\text{SPM}} \) variation from the long-term historical archive of TM5 images.

The moderate resolution imaging spectroradiometer (MODIS) sensors on Terra and Aqua platforms provide near-daily images with 250 m spatial resolution at red and near infrared bands, and their coincident ground \( C_{\text{SPM}} \) measurements are easier to be obtained than those of TM5 image with 30 m spatial resolution. With its advantages of daily cover, medium spatial resolution, high sensitivity and cost-free distribution (Li and Li, 2004; Miller and McKee, 2004), MODIS was frequently employed for estimating \( C_{\text{SPM}} \) or suspended sediment concentrations (Giardino et al., 2010; Kilham and Roberts, 2011; Miller and McKee, 2004; Nechad et al., 2010; Ondrus et al., 2012; Petus et al., 2010; Zhang et al., 2010). Therefore, the \( C_{\text{SPM}} \) values derived from MODIS images can be assumed as the ground measurements and further be employed to assist the developments of TM5-based \( C_{\text{SPM}} \) retrieval models. Such approach, however, has rarely been explored from our knowledge.

This study, with the Poyang Lake of China as a case study, proposed an approach for developing TM5-based \( C_{\text{SPM}} \) retrieval models with the assistance of MODIS, which may lay foundation for retrieving \( C_{\text{SPM}} \) information from the long-term historical TM5 image achieve in Poyang Lake and provide an approach for retrieving \( C_{\text{SPM}} \) values using TM5 images but without coincident ground \( C_{\text{SPM}} \) measurements.

2. Materials and methods

2.1. Study area

Poyang Lake (115°47’–116°45’E, 28°22’–29°45’N) is located on the south of the middle Yangtze River (Fig. 1A). It is the largest freshwater lake of China with a very fluctuant area from <1000 km² in dry seasons to about 4000 km² in flooding periods. The lake plays important roles in local economic and social developments as well as in global ecological conservation (Huang, 2006). \( C_{\text{SPM}} \) is one of dominant factors affecting the water quality of Poyang Lake (Jin et al., 1990; Wu et al., 2011).

2.2. Approach description

The ground \( C_{\text{SPM}} \) measurements and their coincident MODIS image were first employed to validate a MODIS-based \( C_{\text{SPM}} \) retrieval model, and the validated model was applied to MODIS images to retrieve \( C_{\text{SPM}} \) values. The \( C_{\text{SPM}} \) values derived from MODIS images were then assumed as ground truths, and they were combined with their coincident TM5 image to calibrate TM5-based \( C_{\text{SPM}} \) retrieval models. Finally, the \( C_{\text{SPM}} \) values derived from coincident MODIS and TM5 images were compared to assess the performances of TM5-based models. The details are described within the following sections.

2.3. Fieldwork for MODIS-based model validation

One day of fieldwork was carried out on 31 August 2012, and a total of 54 sampling sites (Fig. 1B) were selected from turbid to clear water. At each site, a global position system receiver (Garmin Ltd., USA) was used to record the location coordinate, and about 300–500 ml of surface water were collected from around 0–50 cm water depth for \( C_{\text{SPM}} \) measurement.

2.4. Laboratory \( C_{\text{SPM}} \) measurement

The \( C_{\text{SPM}} \) value of each water sample was measured according to the investigation criteria about the lakes of China (Huang, 1999) as follows: a 0.45 μm nucleopore membrane filter was weighed first, wetted with distilled water, dried for 1 h at 110 °C in a drying oven, and reweighed after recovering to room temperature in a desiccator; the water sample was then shaken to suspend sediments, and poured through the membrane filter using a vacuum pump filtration apparatus; the membrane filter was dried for 2 h at 110 °C and reweighed after cooling to room temperature in the desiccator; and finally the \( C_{\text{SPM}} \) value was calculated by dividing the weight difference of the membrane filter before and after filtering by the water sample volume.

2.5. Image acquisition and pre-processing

Three good quality TM5 images (path/row = 121/40) captured on 20 August 2002, 12 August 2005 and 30 July 2006 were obtained. All images were pre-processed as follows: the cosine approximation model proposed by Chavez (1988, 1996) was employed for atmospheric correction; the topographic maps with a scale of 1:50,000 were used to register the atmospherically corrected image to Beijing 54/Gauss-Krüger projection using first-order polynomial and nearest-neighbor resampling method, and the root mean square error of positional accuracy was within half a pixel; the projected image was re-projected to WGS 84/UTM zone 50 N; and land regions and small water bodies were removed using a binary mask created from unsupervised classification of images.
MOD09GQ is the surface reflectance product computed from the bands 1 and 2 (centered at 645 and 859 nm) of 250 m MODIS Terra level 1B product (http://modis.gsfc.nasa.gov). The MOD09GQ images captured on the same dates as those of the above-mentioned three TM5 images and one fieldwork were downloaded from NASA EOS data gateway (http://glovis.usgs.gov, accessed 1 October 2012). All downloaded images were projected to WGS 84/UTM zone 50 N using nearest-neighbor resampling. For each image, a sub-image covering Poyang Lake was cut from the original one, and the land regions and small water bodies were removed using a binary mask created by visually interpreting its unsupervised classification image also.

2.6. Sampling point selection for TM5-based model development

Regularly distributed sampling points with vertical and horizontal intervals of 1.5 km were first created within the study area. Based on the visual investigation of three TM5 images, 100 sampling points (Fig. 1B) were then selected for TM5-based C<sub>SPM</sub> model developments, considering that their distances to lake boundary, dredging-related vessel and region covered by aquatic vegetation must be larger than 2 km to avoid the effects of land, lake bottom, vessel or vegetation on the pixel reflectance of MODIS and TM5 images.

2.7. MODIS-based model validation

Several MODIS-based C<sub>SPM</sub> retrieval models have been developed for Poyang Lake (Table 1), including a cubic model of red band (Wu and Cui, 2008), a quadratic model of the sum of blue and green band as well as two power models of red band (Jiang and Liu, 2011), exponential models of the difference of red and mid-infrared band (Feng et al., 2012; Wu et al., 2013) and an exponential model of red band (Cui et al., 2013). One of the power models by Jiang and Liu (2011) (Fig. 2A (2)-B) and the cubic model by Wu and Cui (2008) (Fig. 2 A (3), $C_{SPM} = 86236.23 Red^3 - 15858.70 Red^2 + 1005.29 Red - 15.67, R^2 = 0.92, SE = 12.02 mg/l, F = 154.30, P < 0.001) got very consistent fitting results (Fig. 2A), and thus the latter was validated with an independent dataset for retrieving C<sub>SPM</sub> values from MODIS images.

Fifty-four C<sub>SPM</sub> measurements on 31 August 2012 and the MODIS image on 30 August 2012 (about 21–30 h between the fieldwork and Terra satellite overpass) were used to validate the cubic model developed by Wu and Cui (2008) as follows. The model was first applied to the MODIS image to estimate C<sub>SPM</sub> values. Two samples with the highest C<sub>SPM</sub> values were then removed from the 54 C<sub>SPM</sub> measurements. Sand dredging in Poyang Lake results in very high C<sub>SPM</sub> values in some regions (Wu et al., 2007), and the removed two samples could be located within these areas, but their corresponding MODIS image pixel with 250 m spatial resolution might only cover part of these regions. Such scale difference between field sampling and image pixel might cause that a higher C<sub>SPM</sub> measurement corresponded to a lower MODIS band reflectance and further a lower C<sub>SPM</sub> estimation. Finally, the correlation coefficient ($r$) between the measured and estimated C<sub>SPM</sub> values of remaining 52 sampling points as well as the root mean square error (RMSE, Eq. (1)) and relative root mean square error (RRMSE, Eq. (2)) of estimation were calculated to evaluate model performances. The null hypotheses of intercept equal to zero and slope equal to one for the linear regression line between the measured and estimated C<sub>SPM</sub> values were tested to assess model bias.

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

\[
\text{RRMSE} = \frac{\text{RMSE}}{\sum_{i=1}^{n} X_{\text{Mea},i}/n} \times 100
\]

where $X_{\text{Mea},i}$ and $X_{\text{Est},i}$ are measured and estimated values of sampling point $i$, respectively, and $n$ is the number of sampling points.
Table 1

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Model</th>
<th>$R^2$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu et al. (2013)</td>
<td>CSPM = 0.0365Exp(63.2*(R_s(645)-R_s(1240)))</td>
<td>0.76</td>
<td>42</td>
</tr>
<tr>
<td>Cui et al. (2013)</td>
<td>CSPM = 1.0633Exp(27.859R_s(645))</td>
<td>0.91</td>
<td>54</td>
</tr>
<tr>
<td>Feng et al. (2012)</td>
<td>CSPM = 0.6786Exp(34.366*(R_s(645)-R_s(nearest 1240)))</td>
<td>0.87</td>
<td>38</td>
</tr>
<tr>
<td>Jiang and Liu (2011)</td>
<td>CSPM = 1365.5*(R_s(470) + R_s(555)) - 3969.08*(R_s(470) + R_s(555)) + 27.216</td>
<td>0.81</td>
<td>27</td>
</tr>
<tr>
<td>Liu and Rossiter (2008)</td>
<td>CSPM = 50228.8*(R_s(645))</td>
<td>0.74</td>
<td>18</td>
</tr>
<tr>
<td>Wu and Cui (2008)</td>
<td>CSPM = 86236.23*(R_s(645))^3 - 15858.70<em>R_s(645)^2 + 1005.29</em>R_s(645) - 15.67</td>
<td>0.92</td>
<td>42</td>
</tr>
<tr>
<td>Wu and Cui (2008)</td>
<td>CSPM = 7167*R_s(645)^4 - 42</td>
<td>0.91</td>
<td>25</td>
</tr>
</tbody>
</table>

Fig. 2. (A) Plot of 52 suspended particulate matter (CSPM, mg/L) of Poyang Lake using moderate resolution imaging spectroradiometer (MODIS) images on 30 August 2012 showing the CSPM retrieval models in Poyang Lake: (1) exponential model by Feng et al. (2012); (2) power models by Jiang and Liu (2011) and (3) cubic model by Wu and Cui (2008); (B) validation of the cubic model by Wu and Cui (2008) (the solid line is the regression line between the measured and estimated CSPM values, the dashed lines are 95% prediction interval, and the dotted line is the 1:1 line) ($n = 52$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.8. TM5-based model development

Three pairs of coincident MODIS and TM5 images on 20 August 2002, 12 August 2005 and 30 July 2006 were employed for TM5-based CSPM model development. To each MODIS image, the red band reflectance value at each sampling point was assigned directly with the value of the pixel containing this sampling point. To each TM5 image, the blue (band 1), green (band 2), red (band 3), infrared (band 4) and mid-infrared band (band 5) reflectance values at each sampling point were calculated by averaging the pixel values in a 9 × 9 window with this sampling point as centre, in order to match the spatial resolution of MODIS image as close as possible.

The above validated MODIS-based CSPM model was applied to retrieve the CSPM values at 100 sampling points from the MODIS Terra images (MODIS-based CSPM) at three time periods. Based on the statistical analysis of MODIS-based CSPM values, the CSPM values at the time period with the largest variation and their coincident TM5 band reflectance were used to calibrate TM5-based CSPM models as follows. The correlation of MODIS red band reflectance against TM5 blue, green, red and infrared band values was first explored, respectively. The linear, quadratic, cubic, power and exponential models of MODIS-based CSPM against the single blue, green, red, infrared and mid-infrared band reflectance of TM5 image and their combinations (linear combination or ratio) were then calibrated using the least-squares technique, respectively, and their determination coefficients ($R^2$) and estimated standard errors (SE) were compared to select the best-fitting models.

Four best-fitting models were selected and applied to retrieve the CSPM values from the TM5 images (TM5-based CSPM) on the other two time periods. To each selected model, the correlation between the MODIS- and TM5-based CSPM values at 100 sampling points as well as the RMSE and RRMSE of estimation were calculated to evaluate TM5-based model performance, and the null hypothesis of intercept equal to zero and slope equal to one for the linear regression line between the MODIS- and TM5-based CSPM values was also tested to assess model bias. The r, RMSE and RRMSE values as well as the model bias evaluation results of four selected models were compared to determine the best stable one for estimating CSPM values from TM5 images. Finally, the best one within the four best-fitting models was applied to the three TM5 images to retrieve the CSPM values of Poyang Lake.

3. Results

3.1. MODIS-based model validation

Based on the remaining 52 CSPM measurements and the MODIS image on 30 August 2012, the validation result of the cubic model developed by Wu and Cui (2008) indicated a significantly strong correlation between the measured and estimated CSPM values ($r = 0.68$, $p < 0.01$) (Fig. 2B). The results of null hypothesis tests of intercept ($a = 6.11$, SE = 6.83) equal to zero and slope ($b = 0.90$, SE = 0.14) equal to one for the regression line between the measured and estimated CSPM values revealed that the intercept and slope were not significantly different from zero ($t = 0.89$, d.f. = 50, $p = 0.38$) and one ($t = -0.71$, d.f. = 50, $p = 0.48$), respectively, at a significance level of 0.05, which indicated that this cubic model
obtained an unbiased $C_{SPM}$ estimation. The RMSE and RRMSE values of $C_{SPM}$ estimation were 28.5 mg/l and 79.3%.

### 3.2. TM5-based model development

The above validated cubic model of MODIS red band was applied to retrieve the $C_{SPM}$ values at 100 sampling points from MODIS images at three time periods. The statistical results of these estimated $C_{SPM}$ values (Table 2) showed that the $C_{SPM}$ values on 12 August 2005 had higher values (average $C_{SPM} = 35.69$ mg/L) with a larger variation (variation coefficient (CoeVar) = 98.99%) than those on 20 August 2002 (average $C_{SPM} = 8.33$ mg/L, CoeVar = 88.96%) and 30 July 2007 (average $C_{SPM} = 20.09$ mg/L, CoeVar = 94.52%). Due to their largest variation within three time periods, the MODIS-based $C_{SPM}$ values on 12 August 2005 and their coincident TM5 image were applied to calibrate TM5-based $C_{SPM}$ retrieval models.

The correlation analysis showed that there existed significantly strong correlation between the MODIS red band reflectance on 12 August 2005 and its coincident TM5 blue, green, red and infrared band values (Fig. 3), respectively, and TM5 red band best explained the variation of MODIS red band ($R^2 = 0.96$, SE = 0.0089%, $F = 2605.23, p < 0.01$).

Model calibrations showed that the cubic (Fig. 4A) and exponential (Fig. 4D) models of TM5 red band as well as the cubic (Fig. 4G) and exponential (Fig. 4J) models of red subtracting mid-infrared band obtained the best-fitting in all calibrated models for estimating the $C_{SPM}$ from TM5 image on 12 August 2005, and they explained 94–97% of the variation of the MODIS-based $C_{SPM}$ with an SE of 6.617–8.457 mg/l. The highest and lowest correlation coefficients derived from the four models, $r_{model(2)} = -\sqrt{0.94} = 0.9695$ and $r_{model(3)} = \sqrt{0.97} = 0.9849$, were not significantly different at a significance level of 0.01 ($z$-test for two correlation coefficients: $z = 2.44, p = 0.015$), which indicates that, from the point of view of $R^2$, the fitting accuracies of the four models are not significantly different. The exponential model of TM5 red band (Fig. 4D) and the cubic model of red subtracting mid-infrared band (Fig. 4G) got the highest (8.457 mg/l) and lowest (6.617 mg/l) SE values, respectively, but their difference was small (1.84 mg/l).

The above-selected four best-fitting models were applied to estimate the $C_{SPM}$ values at 100 sampling points from TM5 images on 20 August 2002 and 30 July 2006. The correlation analyses

---

**Fig. 3.** Correlations between the red band of moderate resolution imaging spectroradiometer (MODIS) on 12 August 2005 and its coincident Landsat-5 TM (TM5) blue (A), green (B), red (C) and infrared (D) bands (n = 100). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. 4. Calibration and validation results of Landsat-5 TM (TM5) models (1 and 2 are the cubic and exponential models of red band, and 3 and 4 are the cubic and exponential model of red subtracting mid-infrared band); four best-fitting regression models between the suspended particulate matter (C_{SPM}) derived from moderate resolution imaging spectroradiometer (MODIS) image and the TM5 reflectance on 12 August 2005 (n = 100) (X = red – mid-infrared band) (A, D, G and J); validations of the regression models using the MODIS and TM5 images on 20 August 2002 and 30 July 2006 (the solid line is the regression line between the MODIS- and TM5-based C_{SPM} values, and the dashed line is the 1:1 line) (I = 200) (B, E, H and K); and C_{SPM} estimation residuals (MODIS-based subtracting TM5-based C_{SPM}) vs. TM5-based C_{SPM} (C, F, I and L). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
showed that there existed significantly strong correlation between the MODIS- and TM5-based $C_{SPM}$ values for these four models at a significance level of 0.05 ($r = 0.96, p < 0.01$) (Fig. 4B, E, H and K). The null hypothesis test of intercept equal to zero for the regression line between the MODIS- and TM5-based $C_{SPM}$ values revealed that the intercepts were significantly different from zero at a significance level of 0.05 for all the models but their absolute values were less than 2.5 mg/l (Table 3); while the null hypothesis test showed that the slopes for the regression line between the MODIS- and TM5-based $C_{SPM}$ values were significantly different from one for the cubic model of red band and the exponential model of red subtracting mid-infrared band but not for the exponential model of red band and the cubic model of red subtracting mid-infrared band at a significance level of 0.05 (Table 3). Moreover, no significant correlation was found between the $C_{SPM}$ estimation residual (MODIS-based $C_{SPM}$ subtracting TM5-based $C_{SPM}$) and TM5-based $C_{SPM}$ for the exponential model of red band at a significance level of 0.05 ($r = 0.11, p = 0.126$) (Fig. 4F). These results indicated that the exponential model of TM5 red band obtained the best result for estimating $C_{SPM}$ from TM5 images in this study and its RMSE was 4.60 mg/l when the MODIS-based $C_{SPM}$ values were assumed as the ground measurements. The exponential model of TM5 red band was applied to three TM5 images to retrieve $C_{SPM}$ values, and clear spatial patterns of $C_{SPM}$ were observed (Fig. 5).

### Table 3

Null hypothesis tests of intercept (a) equal to zero and slope (b) equal to one for the regression line between the moderate resolution imaging spectroradiometer (MODIS)-based and Landsat-5 TM-based $C_{SPM}$ values ($n = 205$); (1) and (2) are the cubic and exponential models of red band, and (3) and (4) are the cubic and exponential model of red subtracting mid-infrared band of Landsat-5 TM image.

<table>
<thead>
<tr>
<th>Intercept</th>
<th>Slope</th>
<th>a</th>
<th>SE</th>
<th>$r$</th>
<th>$p$</th>
<th>b</th>
<th>SE</th>
<th>$r$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>1.027</td>
<td>0.438</td>
<td>2.347</td>
<td>0.020</td>
<td>0.869</td>
<td>0.019</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2)</td>
<td>1.736</td>
<td>0.461</td>
<td>3.768</td>
<td>&lt;0.001</td>
<td>1.034</td>
<td>0.022</td>
<td>1.545</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3)</td>
<td>0.950</td>
<td>0.388</td>
<td>2.446</td>
<td>0.015</td>
<td>0.963</td>
<td>0.019</td>
<td>&lt;1.947</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>2.233</td>
<td>0.439</td>
<td>5.086</td>
<td>&lt;0.001</td>
<td>1.131</td>
<td>0.022</td>
<td>5.955</td>
</tr>
</tbody>
</table>

4. Discussion

Several MODIS-based $C_{SPM}$ retrieval models have been developed for Poyang Lake (Table 1). The model by Jiang and Liu (2011) was not validated using an independent dataset, while Feng et al. (2012) applied the in situ $C_{SPM}$ measurements of two fieldworks together to calibrate model, and no independent dataset was further used to validate the developed model. The MODIS-based $C_{SPM}$ retrieval model used in this study was calibrated and validated using independent datasets (validation: $r = 0.94$, $p < 0.01$ (Wu and Cui, 2008); it was further validated using another independent dataset in this study ($r = 0.68$, $p < 0.01$), and both validations obtained acceptable result. Wu et al. (2013) concluded that the exponential model of red minus infrared band obtained stable $C_{SPM}$ estimations in Poyang Lake; while they also reported that cubic model of MODIS red band explained 90% of the variation of $C_{SPM}$, and its validation obtained acceptable result ($r = 0.80$, $p < 0.01$). In addition, Cui et al. (2013) also reported that that cubic model of MODIS red band explained 92% of the variation of $C_{SPM}$. Although the $C_{SPM}$ models generally are time- and site-specific due to different water constituents, especially for very dynamic lakes like Poyang Lake, the above-mentioned results of the cubic $C_{SPM}$ models of MODIS red band indicated that the model we used in this study was relatively stable and applicable.

Water-leaving radiances are captured by remote sensing sensor and used to estimate water quality parameters, and they are affected by the inherent optical properties of water body, the most fundamental parameters of which are absorption and backscattering coefficients (Lee et al., 2002). These coefficients are determined by main water constituents, including phytoplankton, suspended sediment and yellow substance (Kirk, 1994; Ma et al., 2006a). The absorption coefficients of these main water constituents decrease with the increasing wavelength and are close to zero in the red and near infrared regions, especially for the water bodies with low phytoplankton concentration; thus, the water-leaving reflectance is dominated by the backscattering coefficient of SPM in the visible and near infrared bands (Doxaran et al., 2006; Ma et al., 2010). Moreover, there exists a positive correlation between the backscattering coefficient and $C_{SPM}$ in the visible and near infrared bands (Doxaran et al., 2006; Ma et al., 2010, 2006b). Such characteristics of water absorption and backscattering explained why the red band of TM5 image obtained the best result for estimating $C_{SPM}$ in this study.

The exponential model of TM5 red band got the best $C_{SPM}$ estimation in this study, and it is different to the linear models of single band (Chen et al., 2010; Durga Rao et al., 2009; Nas et al., 2010; Tyler et al., 2006), quadratic model of linear combination of several bands (Oyama et al., 2009), linear models of band ratio (Nas et al., 2010; Oyama et al., 2009) and exponential model of band combination (Sriwongsitanon et al., 2011) applied in many studies. Such difference could be explained by the fact that the empirical or semi-empirical models generally varied at different water bodies or different seasons due to their different water constituents (Liu et al., 2003; Ma et al., 2010; Zhang et al., 2008).

Linear models of red band were generally employed when the $C_{SPM}$ values were low while the exponential or power models were usually applied when the $C_{SPM}$ values were high (Wu et al., 2013). There existed positive correlation between the concentration and backscattering of SPM over red spectral region. Such correlation is generally linear when the $C_{SPM}$ Values are low, and becomes exponential or power when the $C_{SPM}$ Values are high because of multiple scattering or shielding effect. These could explain why an exponential model between $C_{SPM}$ and red band was usually employed when the $C_{SPM}$ values were high (Wu et al., 2013).

Many studies have employed TM5 images to estimate $C_{SPM}$ at different water bodies, such as the Lake Balaton in central Europe ($R^2 = 0.89$) (Tyler et al., 2006), Lake Kasumigaura in Japan ($R^2 = 0.96$) (Oyama et al., 2009), Bhopal Upper Lake in India ($R^2 = 0.71$) (Durga Rao et al., 2009), Taihu Lake in China ($R^2 = 0.84$) (Chen et al., 2010), Lake Beysehir in Turkey ($R^2 = 0.67$) (Nas et al., 2010) and Bung Boraphet in Thailand ($R^2 = 0.89$) (Sriwongsitanon et al., 2011). We could not compare the estimation results from the best model found in this study with those of aforementioned studies, because we did not have coincident ground $C_{SPM}$ measurements and TM5 image for model validation. However, we found that the red band of TM5 image explained 96% of the variation of MODIS Terra red band with a linear model ($R^2 = 0.96$, $SE = 0.0089\%$, $F = 2605.23$, $p < 0.01$, $n = 100$), while the cubic model of MODIS Terra red band employed in this study explained 92% of the variation of $C_{SPM}$ measurements in Poyang Lake (Wu and Cui, 2008); thus we deduced reasonably that the TM5 red
band could well explain the variation of $C_{\text{SPM}}$ measurements. Moreover, we also observed in this study that there existed a significant and very strong correlation between the MODIS- and TM5-based $C_{\text{SPM}}$ values at a significance level of 0.05 ($r = 0.96$, RMSE = 4.60 mg/l, $p < 0.01$, $n = 200$), and thus we considered that the MODIS and TM5 image got very consistent results for estimating the $C_{\text{SPM}}$ in Poyang Lake.

The exponential model of TM5 red band best explained the variation of MODIS-based $C_{\text{SPM}}$ values in this study; however, the saturation phenomenon was also observed, namely, TM5 red reflectance will not increase with the increasing $C_{\text{SPM}}$ value when it reaches a certain value (about 18% in this study) (Fig. 4D). This saturation problem indicated that the exponential model of red band was not very suitable for high turbid water, and such result was observed in Fig. 4F, in which the estimation residual increased when the $C_{\text{SPM}}$ value was larger than about 50 mg/l. Therefore, we considered that the exponential model of TM5 red band developed in this study could be more reliable for the water body of Poyang Lake before, but not after 2000, because the dredging activities started in 2001 resulted in a large increasing of $C_{\text{SPM}}$, especially in the northern part (Wu and Cui, 2008).

Although, with the assistance of MODIS, a TM5-based $C_{\text{SPM}}$ retrieval model was developed while it performed well in Poyang Lake, two issues that were not focused on in this study may further improve model performance. TM5 images were atmospherically corrected with the cosine approximation model (Chavez, 1988, 1996) in this study, which was not developed specially for the Case II waters. To date, no standard atmospheric correction method exists for the Case II waters, and other potential methods for the atmospheric corrections of TM5 images should further be explored for improving model accuracy. In addition, only three pairs of MODIS and TM5 images were employed for developing TM5-based $C_{\text{SPM}}$ retrieval model, and the model accuracy and robustness might be improved by employing more images with good quality.

TM5 images hold great potentials in retrieving long-term historical water quality parameters. However, the missing of coincident ground measurements hampered the developments of TM5-based retrieval models, which limited the applications of TM5 images on water quality estimations. We considered that such problem might appear frequently at many water bodies around the world but not only at Poyang Lake. The approach proposed in this study would be helpful for other researchers, who are suffering the same problem, to retrieve water quality parameters ($C_{\text{SPM}}$, suspended sediment and chlorophyll a concentration etc.) from the long-term historical TM5 image archive but without the coincident ground measurements.

5. Conclusion

With the Poyang Lake in China as a case study, we proposed an approach for developing TM5-based $C_{\text{SPM}}$ retrieval models with the assistance of MODIS images in this study. The principal results obtained can be summarized as follows:

1. Based on ground $C_{\text{SPM}}$ measurements and their coincident MODIS image, the cubic model of MODIS red band employed in this study obtained an acceptable validation result.
2. The cubic and exponential models of red band and red subtracting mid-infrared band obtained the best fitting results for estimating $C_{\text{SPM}}$ values from TM5 image.
3. The exponential model of red band ($C_{\text{SPM}} = 2.190 - \exp(21.965\text{Red})$, $R^2 = 0.94$, $SE = 8.457$, $F = 2805.23$) was the best stable for estimating $C_{\text{SPM}}$ from TM5 images when the MODIS-based $C_{\text{SPM}}$ values were assumed as ground truths ($r = 0.96$, $SE = 4.60$ mg/l).

The developed model laid foundation for retrieving $C_{\text{SPM}}$ information from TM5 images in Poyang Lake; while the approach proposed in this study would be helpful for other researchers who also want to retrieve $C_{\text{SPM}}$ or other water quality parameters from the long-term historical TM5 image archive but without the coincident ground measurements.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant Nos. 41171290, 40971191 and 41171271) and the Special Foundation of the Ministry of Finance of China for Nonprofit Research of Forestry Industry (Grant No. 200904001).