Estimation CODMn in Guangzhou section of Pearl River based on GF-1 images

Y. B. Feng 1,2, Y. Q. He 1,2, Q. H. Fu 1,2, C. Q. Liu 1,2, H.Z. Pan 1,2, B. Yin 1,2

1 Key Laboratory of dynamics and associated process of the Pearl River Estuary, Ministry of water resources, China – (robinut, heyinqing}@foxmail.com
2 Pearl River Hydraulic Research Institute, Guangzhou, China

KEY WORDS: CODMn, GF-1, water color remote sensing, Pearl River, Guangzhou

ABSTRACT:

Due to the way that remote sensing works, it has natural advantage to detect optical constituents in waters. And many kinds of inversion models were constructed based on the three main optical constituents, namely chlorophyll-a (Chl-a), suspended particulate matter (SPM), colored dissolved organic matter (CDOM). Except Chl-a used as an indicator of eutrophication, however, the public generally cares less about other two parameters and is more familiar with Grade 1 ~ V scheme for utilization and protection purposes. Notice the three main optical constituents are also organic-related to some extent. It offers a possible way to estimate CODMn via remote sensing. According to field measurement conducted along the Guangzhou section of Pearl River (GPR for short), the spatial variation of CODMn in GPR shows some kinds of geographical feature, so does the correlation between CODMn and water color constituents. It indicated the complicated contribution of CODMn in GPR or some other urban rivers. Based on the band setting of GF-1 satellite, two kinds of inversion model of CODMn in GPR were finally constructed. One directly achieved CODMn from regression models of which predictors were different band combinations in different channels of GPR. To make the study more practical, the other one first provided empirical models of the three optical constituents, and then estimated CODMn of GPR based on its relationship with optical constituents. After all, Chl-a, SPM and CDOM could be distinguished optically, and remote sensing models of these three constituents in other studies may also be available.

1. INTRODUCTION

1.1 General Instructions

In the field of water colour remote sensing, we centre on optical properties of water. Due to the longstanding civilization from mass media and elementary education, however, the public pays more attention to the water quality. The most striking distinction is that the Case 1 and Case 2 classification of waters is commonly used for bio-optical modelling while most citizens are familiar with Grade 1 ~ V scheme for utilization purposes and protection objectives. It leads to the three main optical constituents as chlorophyll-a (Chl-a), suspended particulate matter (SPM), colored dissolved organic matter (CDOM) retrieved by remote sensing are not aligned with pollution indicators such as chemical oxygen demand (COD), biochemical oxygen demand (BOD), total phosphorus (TP), ammonia nitrogen and heavy metals. This embarrassing mismatch restricts the application of remote sensing in the field of water monitoring.

Take permanganate index (CODMn) as an example, it is one of important parameters used to assess organic pollution in surface water and ground water (Tian et al., 2008). Before remote sensing comes into view, field sampling and indoor testing form the primary method and provide us the credible CODMn values (Baker et al., 1999; Udovichenko and Nabivanets, 2001; Udovichenko et al., 2001). But it is a time-consuming and laborious work, and is impossible to acquire real-time and large-scale data for monitoring. That’s instead the advantage where remote sensing lives by. So remote sensing has been applied more and more in the water quality monitoring recently.

For instances, Fu et al (2007) took the Grand Canal through southern Jiangsu as the studied area and found Band 1 of Landsat TM images was highly correlated with CODMn; Wang et al (2003) constructed a BP neural network model to inverse CODMn of Poyang Lake from TM data; Yang et al (2007) employed TM images of Tai Lake to retrieve CODMn based on the empirical relationship with Chl-a which is computed by a semi-empirical method; still the Tai Lake, Tao et al (2014) proposed an Advanced CODMn Forecast Index (ACFI) to estimate CODMn using Landsat-8; Hao et al (2011a, 2011b) found the ratio of TM3 and TM5 had high relevance with CODMn of Daliangdian Reservior; Wang et al (2011) applied a support vector regression (SVR) method to predict CODMn in the Weihe River by SPOT-5 data.

Obviously, there are not unified methods to estimate CODMn of inland waters. Despite the different band response of satellite sensors, the most important reason is that the CODMn is a synetical index rather than a certain substance like Chl-a. It refers to the amount of oxygen consumed when the organic matter in a given volume of water is chemically oxidized to CO2 and H2O by permanganate (Xia, 2005). And this organic matter is constructed with organic particulate matter and dissolved organic matter in any proportions. Considering the various spectral characteristics of organics in different aquatic environments, it’s necessary to re-establish the inversion method of CODMn in regional studies.

The Pearl River is an extensive river system in southern China and supplies water to numerous cities. Among them, the Guangzhou City which serves as a political-economic-cultural centre of Guangdong province is the most populous one. So pollutants like industrial waste, sewage runoff and agriculture discharges all make the water quality of this section declined. Nevertheless, only few studies have been published on the
application of remote sensing in contamination assessment of
the Guangzhou section of the Pearl River (GPR for short).
Wang et al (2001) recognized water pollution from TM images
by qualitatively analysing the gray scale variation of each band
due to organic pollutants. The result was practical because it
visually displayed different levels of water quality in line with
GB3838-88 where the Grade 1 ~ V scheme of surface water is
proposed. Fan and Chen (2009, 2012) computed comprehensive
pollution index of water quality based on referenced values of
the Grade III water, and then established regression models of
this index and TM bands to present water pollution. It’s a
quantitative attempt about this area, but the robust of the
multiple linear regression and the practice that a single grade
was referenced are worth discussing.

Therefore, this study is an extension of the previous effort,
targeting on quantitatively estimating CODMn of the GPR
and emphasizing on the regional relationship between CODMn
and water colour constituents in order to combine well-developed
water colour models. After all, Chl-a, SPM and CDOM could
be distinguished optically. A remote sensing model for the
retrieval of these three constituents may achieve more credible
result. To make the study more practical, we also provide an
empirical method to get CODMn directly from remote sensing
images, like GF-1.

2. DATA ACQUISITION

In strict sense, GPR ranges from the mountain areas in the north
(Baiyun Mountain) to sea level at the confluence of Pearl River
in the south.

On 5~6 August 2015, field measurement was conducted along
the Guangzhou section of Pearl River. As shown in Figure 1,
there were 27 sampling sites where water samples were
collected for Chl-a, SPM, CDOM and CODMn test in laboratory;
bottom sediments were collected for water-tank experiment;
and water surface spectral were recorded in situ according to
above-water method (Tang et al., 2004).

![Figure 1. Sampling sites distribution of GPR](image)

3. RESULTS AND DISCUSSIONS

3.1 Spatial variations of CODMn in GPR

From the field measurement, the lowest CODMn (2.8 mg/L)
appeared at B3 and B4 located along the back-channel (Hou
Hangdao) of GPR while the highest (7.2 mg/L) was found at
A11, A14 and A15 all in the west-channel (Xi Hangdao). It
indicated that the degree of organic contamination of GPR may
have some kinds of regional feature. Thus all the sampling
points were divided into 3 groups geographically: a) the west
channel represented by A11~A17; b) the front-channel
represented by A01~A10; c) the back-channel represented by
B01~B10.

As shown in the boxplot (Figure 3), the median values of
CODMn decreased from west-channel, front-channel to back-
channel, so do the data ranges and other statistical values. The
maximum CODMn of front-channel denoted by the upper edge
was found at A06 which was around LieDe, the downtown area
of Guangzhou. For the outlier of back-channel, which was
greater than the threshold exceptions q3 + w × (q3 – q1) in a
boxplot (q1, the lower quartile; q3, the upper quartile; w, the
1.5 standard deviations), it’s recorded in B01 where was just the
downstream of west-channel. These spatial features of CODMn
were basically in line with previous studies (Wang et al., 2001;
Ma et al., 2003; Wang et al., 2009). So we regrouped sampling
sites and the B01 was moved to the west-channel for
subsequent analysis.

![Figure 2. Comparison of CODMn in different parts of GPR](image)

3.2 Relationships between CODMn and optical constituents

As a synthetic index of water pollutions, CODMn indicates the
organic content in water. The three kinds of water colour
constituents which remote sensing concerned are also organic-
related to some extent. Generally speaking, Chl-a is considered
to be the proxy for phytoplankton; SPM is divided into organic
and inorganic components; CDOM as the abbreviation of
colored dissolved organic matter is definitely organic. So the
correlation coefficients between CODMn and water colour
constituents were computed based on the regional division in
section 3.1, and the result was shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Chl-a</th>
<th>SPM</th>
<th>ag440</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front-channel</td>
<td>0.613</td>
<td>0.436</td>
<td>0.546</td>
</tr>
<tr>
<td>Back-channel</td>
<td>0.820</td>
<td>0.068</td>
<td>0.550</td>
</tr>
<tr>
<td>West-channel</td>
<td>-0.063</td>
<td>0.717</td>
<td>0.933</td>
</tr>
<tr>
<td>All field sites</td>
<td>0.860</td>
<td>0.520</td>
<td>0.743</td>
</tr>
</tbody>
</table>

Table 1. Correlation coefficients between CODMn and water
colour constituents in GPR

This contribution has been peer-reviewed.

https://doi.org/10.5194/isprs-archives-XLII-3-349-2018 | © Authors 2018. CC BY 4.0 License.
3.2.1 West-channel: For the west-channel, CODMn was closely tied to ag440, the absorption coefficient of CDOM at 440 nm normally as its concentration, and the correlation coefficient R was up to 0.933. The correlation between CODMn and SPM was also relatively high (R = 0.717), however, the concentration variation of SPM under a high concentration of CODMn (about 7.2 mg/L) was quite large compared with those of CDOM (Figure 3). As for Chl-a, the correlation was negligible. Therefore, the CODMn in the west-channel of GPR could be estimated from ag440 and their relationship is constructed as Eq.1

\[ y = 8.108x^{0.4241}, \quad R^2 = 0.8512 \quad (1) \]

where \( y = \text{CODMn} \)
\( x = \text{ag440} \)

At this point, different regression models were tried to quantify this relationship. From the numerical perspective, a complicated regression function may generate high value to evaluate fitting precision (e.g. \( R^2 \)), but the possibility of sudden changes outside the range of training data (over-fitting) could make this model far away from the actual. So only linear polynomial, one-term exponential model and one-term power model were considered based on the sampling data. As shown in Figure 4, all the three regression functions reflect the CODMn variation with ag440 quite well and achieve high R2. However, only the power function gives a reasonable trend when the CODMn concentration is below 5 mg/L.

![Figure 4. Different fitting curves of CODMn and ag440 in west-channel of GPR](image)

3.2.2 Front-channel: For the front-channel, the correlation coefficients between CODMn and three optical constituents are all around 0.5, an embarrassing value. It says that 1) the CODMn of this channel could hardly be obtained from either one of optical constituents; 2) the hydrodynamic condition (or water environment) here is more complicated so that some measurements are unsuited to analysis together; 3) the main contribution to CODMn here is non-optical ingredients. The latter possibility stands outside the discussion owing to the required data is beyond the scope of this paper.

For the first possibility, multivariate linear regression was used to evaluate the combined contribution of optical constituents to CODMn. We found that the \( R^2 \) statistic decrease from 0.768 to 0.428, and the predictor variables were respectively all-the-three-optical-constituents and Chla-ag440. So the CODMn of this channel may be predicted by Eq.2.

\[ y = 5.7687+1.6545x_1+1.7073x_2+0.7260x_3, \quad R^2 = 0.7682 \quad (2) \]

where \( y = \text{CODMn} \)
\( x_1 = \text{Chl-aN} \)
\( x_2 = \text{SPMN} \)
\( x_3 = \text{ag440N} \)

The subscript N indicated that the predictors were all scaled between -1 and 1 by normalizing the minimum and maximum values of each optical constituent. The minimum values were all set to 0, and the maximum values were respectively 120 \( \mu \text{g/L} \) of Chl-a, 80 mg/L of SPM and 0.8 \( \text{m}^{-1} \) of ag440 on a comprehensive basis of several studies in GPR (Fan, 2012; Wang et al., 2009; Ma et al., 2003; Li et al., 2013; Jiang et al., 2010).

For the second possibility, the scatter plots of CODMn v.s. each optical constituent are displayed in Figure 3. It’s observed that 1) the correlation between CODMn and Chl-a would become significant (R = 0.903) if sample points A3 and A5 were removed, whose Chl-a concentrations were respectively 93.2 \( \mu \text{g/L} \) and 70.3 \( \mu \text{g/L} \); 2) the correlation coefficient between CODMn and SPM would be raised to 0.894 if sample points A8–A10 were removed, which SPM concentration were 15 mg/L, 24 mg/L, 23 mg/L successively; 3) the correlation coefficient between CODMn and ag440 would increase to 0.752 if A1 and A5 were not considered, which ag440 were 0.368 \( \text{m}^{-1} \) and 0.415 \( \text{m}^{-1} \) respectively. Upon these assumptions, we located the points removed and found they were all near docks or factories (Table 2). However, it’s hard to identify whether the discharge of industrial wastewater was normal condition because the removal of deferent sites would generate different sensitive parameter to CODMn. In consequence, the analysis of this possibility didn’t bring about any quantitative results.

### Table 2. Abnormal sites of the front-channel and the buildings around

<table>
<thead>
<tr>
<th>Sampling sites</th>
<th>Nearby buildings</th>
</tr>
</thead>
<tbody>
<tr>
<td>A01</td>
<td>Nanhai shipyard</td>
</tr>
<tr>
<td>A03</td>
<td>Yuzhu shipyard, Jiali wharf</td>
</tr>
<tr>
<td>A05</td>
<td>Yongxing tile factory, Guangzhou</td>
</tr>
<tr>
<td></td>
<td>Paian Concrete Ltd</td>
</tr>
<tr>
<td>A08</td>
<td>Sha Tau cruise terminal</td>
</tr>
<tr>
<td>A09</td>
<td>Tianzi Wharf</td>
</tr>
<tr>
<td>A10</td>
<td>Xidi Wharf</td>
</tr>
</tbody>
</table>

This contribution has been peer-reviewed.

https://doi.org/10.5194/isprs-archives-XLI-3-349-2018 | © Authors 2018. CC BY 4.0 License.
3.2.3 Back-channel: For the back-channel, Chl-a became the highly referential constituent of COD$_{Mn}$ since their correlation coefficient was 0.820, way above other parameters. As in the west case, the COD$_{Mn}$ here could be simply modelled in Eq-3.

\[
y = 1.2333x^{0.3137}, \quad R^2 = 0.6432 \quad (3)
\]

where \( y = \text{COD}_{Mn} \) and \( x = \text{Chl-a} \)

![Figure 5. The fitting curve of COD$_{Mn}$ and Chl-a in back-channel of GPR](image)

3.2.4 All sites of GPR: Given all the sites of GPR, each optical constituent behaved certain correlation against COD$_{Mn}$, and the links decreased as follow: Chl-a, ag440, TSM (Figure 4). Despite different grades, all the scatter trends shared the same feature that the correlation became impaired as COD$_{Mn}$ increased. As a result, even if Eq-4 employed Chl-a to acquire COD$_{Mn}$ concentration of GPR and it worked out fine (\( R^2 = 0.767 \)), Eq-1 was still recommended to evaluate high COD$_{Mn}$ (i.e. those above 6 mg/L for GPR) which was more likely to be found in west-channel. This further idicated the complicated contribution of COD$_{Mn}$ in GPR or some other urban rivers.

\[
y = 0.8832x^{0.4317}, \quad R^2 = 0.7666 \quad (4)
\]

where \( y = \text{COD}_{Mn} \) and \( x = \text{Chl-a} \)

![Figure 6. The fitting curve of COD$_{Mn}$ and Chl-a in back-channel of GPR](image)

3.3 Remote sensing retrieval of COD$_{Mn}$

Benefiting from high spatial resolution and low revisit period, the GF-1 satellite is now widely used in many domestic industries. The PMS optical sensor on-board could acquire 2m resolution images in visible and near-infrared bands (Blue, 450-520nm; Green, 520-590nm; Red, 630-690nm; NIR, 770-890nm), and offer a new way to monitor water qualities in urban.

On basis of regional features of GPR COD$_{Mn}$ discussed above, two kinds of inversion models were constructed. One of them directly achieved COD$_{Mn}$ from regression models of which predictors were different band combinations in different channels of GPR. The combinations include common forms such as \( x_1 \), \( x_1/x_2 \), \( x_1/(x_1+x_2) \), \( x_1-x_2)/(x_1+x_2) \), \( x_1/(x_1+x_2+x_3) \), \( x_1x_2)/(x_1+x_2+x_3) \), etc. And each form was constructed by different bands. After comparing the fitting performance, regression equations with the highest \( R^2 \) was shown in Figure 8. For the front and west channels, the \( R^2 \) were both above 0.65 and the fitting results were acceptable. For the back-channel, the statistic \( R^2 \) was just 0.27 although the fitting curves made a roughly good prediction. In order to make up for the lack of measurements, all sites of GPR were assembled and the band combination \( B3/(B1+B4) \) was selected to get COD$_{Mn}$ not only in the whole GPR but also the back-channel.

![Figure 7. The COD$_{Mn}$ regression models based on GF-1 PMS band setting](image)

To make our study more practical, the other one first provided empirical models of the three optical constituents, and then estimated COD$_{Mn}$ of GPR based on its relationship with optical constituents discussed above. After all, Chl-a, SPM and CDOM could be distinguished optically, and remote sensing models for the retrieval of these three constituents in other studies may also be available. According to measurement, the fitting equation of Chla, SPM and ag440 of each channel was constructed and shown in Table 3~5.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Band Combination</th>
<th>Fitting equation</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>B3/(B2+B4)</td>
<td>( y = -236.15x+219.91 )</td>
<td>0.700</td>
</tr>
<tr>
<td>Back</td>
<td>B2</td>
<td>( y = 0.001x+2.618 )</td>
<td>0.554</td>
</tr>
<tr>
<td>West</td>
<td>B1/(B3+B4)</td>
<td>( y = 40.358\exp(1.092x) )</td>
<td>0.679</td>
</tr>
<tr>
<td>GPR</td>
<td>(B1+B3)/(B1+B2+B4)</td>
<td>( y = 11.322x^{5.653} )</td>
<td>0.686</td>
</tr>
</tbody>
</table>

![Table 3. The Chl-a regression models of GPR channels based on GF-1 PMS band setting](image)
CODMn of back-channel were lower than those of other two in inversion result basically concided with field measurments. The GF-1 images.

It's found that 1) the fitting $R^2$ of ag440 was generally low, so SPM and CDOM) in the remote sensing domain, then the fitting $R^2$ of SPM in the three channels were all above 0.79 the first kind of CODMn inversion model (the direct one) may be belonging to Grade Ⅲ while this value fell to 0.22 considering all sites of GPR, it showed different.

Upon these fitting equations, four GF-1 PMS images (imaging date: Dec. 7, 2016) of GPR were employed to get the CODMn distribution. As displayed in Figure 8, the spatial trend of inversion result basically concided with field measurements. The CODMn of back-channel were lower than those of other two parts of GPR in general. And it showed that the CODMn values of GPR ranged mainly between 4 mg/L to 8 mg/L, roughly belonging to Grade Ⅲ─Ⅳ waters according to Environmental Quality Standards for Surface Water (GB 3838-2002).

Table 4. The SPM regression models of GPR channels based on GF-1 PMS band setting

<table>
<thead>
<tr>
<th></th>
<th>x</th>
<th>Fitting equation</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>(B1+B3)/(B1+B2+B3)</td>
<td>$y=3.17*10^{-10}$ (\exp(42.80x))</td>
<td>0.810</td>
</tr>
<tr>
<td>Back</td>
<td>(B1+B3)/(B1+B2+B4)</td>
<td>$y=0.264\exp(4.965x)$</td>
<td>0.835</td>
</tr>
<tr>
<td>West</td>
<td>(B2+B4)/(B1+B2+B3)</td>
<td>$y=9.892*10^{22}x^{-9.84}$</td>
<td>0.792</td>
</tr>
<tr>
<td>GPR</td>
<td>B2/(B3+B4)</td>
<td>$y=25.094x^{-2.149}$</td>
<td>0.224</td>
</tr>
</tbody>
</table>

Table 5. The ag440 regression models of GPR channels based on GF-1 PMS band setting

It’s found that 1) the fitting $R^2$ of ag440 was generally low, so the first kind of CODMn inversion model (the direct one) may be more suitable to achieve CODMn in west and front channels; 2) the fitting $R^2$ of SPM in the three channels were all above 0.79 while this value fell to 0.22 considering all sites of GPR, it indicated the dominant component of SPM in each channel may be different.

In view of the complicated contribution of CODMn in GPR, this study hold the point that water qualities retrievals via remotely sensed images should take regional characteristics and composition of target parameter into consideration.

ACKNOWLEDGEMENTS

Acknowledgements of support for the projects: Estimation and Forecast of Suspended Sediment Concentration in the Pearl River Estuary from Remote Sensing Images based on Data Assimilation Algorithm (Natural Science Foundation of Guangdong, China, NO. 2017A030313232)

Remote Sensing Identification and Dynamic Monitoring of Urban Black-Odor Water Bodies based on Chroma Parameters: A Case Study in Guangzhou (Science and Technology Program of Guangzhou, China, NO. 201804010270).

REFERENCES


Li, P., Gui, C., & Guan, X., 2013. Investigation on the Eutrophication Status of the Pearl River Guangzhou Section 4. CONCLUSION

Based on field measurement of GPR, this study divided the urban river into three channels and analysed the correlation between CODMn and the three water colour constituents (Chl-a, SPM and CDOM) in the remote sensing domain, then constructed two kinds of CODMn inversion models suitable for GF-1 images. These analysis indicated that: 1) the CODMn concentration of west-channel was lower than other part of GPR in general; 2) the CODMn of west-channel and back-channel was respectively dominated by CDOM and Chl-a while that of front-channel was contributed by all the three optical constituents; 3) the CODMn showed higher correlation with Chl-a on the whole of GPR; 4) the CO2dh inversion directly from band combinations of GF-1 images was applied here, but the acquisition of CODMn from optical constituents also recommended because it provided an open port and could be referenced from other studies; 5) the inversion CODMn based on GF-1 images (Dec. 7, 2016) showed the water quality of GPR mainly belonged to Grade Ⅲ─Ⅳ.


