Mapping soil pollution risk from riparian land-cover changes in the headwater area of an inter-basin water diversion project

Quanfang Wang¹,*, Mengru Zhang¹+, Yu Zhang¹2,*, Yuqi Yang¹+, Xinsheng Wang¹, Lanfang Yang¹, Longyue Chen¹, and Ying Xiao¹

¹School of Resources and Environmental Science, Hubei University, Wuhan 430062, China
²HuBei TianDiYun Geomatics Technology Group, Wuhan 430010, China
*corresponding: wangqf@hubu.edu.cn; 864711182@qq.com
*these authors contributed equally to this work: Mengru Zhang, Yu Zhang and Yuqi Yang. email: 2061089224@qq.com; 864711182@qq.com; 1159750430@qq.com

Abstract. A novel scheme was proposed for mapping precise land-cover change in riparian zones and the resulted soil pollution risk of non-point source (NPS) from an inter-basin water diversion project on its headwater area. The scheme held three key points: 1) digital visualization of the riparian land-cover’s cyclical features was done using remote sensing data fusion of NDWI at high water level and NDVI and MNDWI at lower water level; 2) riparian zones were identified through color clustering on optimized RGB-fusion image that was picked out by applying Modified Vector-Angular Distance (MVAD) formula to measure difference between riparian and non-riparian zones in RGB space; 3) combining Export Coefficient Model (ECM) and the derived data on riparian land-cover change around Danjiangkou Reservoir in China, time-series loads of the soil NPS pollutants were estimated before and after the Middle Route of South-to-North Water Diversion Project (SNWDP) started running in 2014. The results indicated that risk of the potential soil pollution from land-cover change in the new riparian zones enlarged. Growth rates (GR) of total nitrogen (TN) and total phosphorus (TP) not only had significant spatial variation, but also were higher during initial running period of the SNWDP’s Middle Route.

Keywords: loads of soil NPS pollutants, computer-vision detection, optimized RGB-fusion image color clustering, Export Coefficient Model (ECM), time-series Landsat imageries

Introduction

Large-scale water-transfer projects and their influence on water-diverted regions have become research hotspots. Water shortage hinders socioeconomic growth¹. As a solution, water transfer has received interest amongst policymakers, and numerous water transfer projects had been implemented across the world including in Brazil, China, Germany, Australia and the United States, etc¹³. China, a country with a long history of inter-basin water transfer, started another high-profile water transfer project, i.e. the South-to-North Water Diversion Project (SNWDP) at beginning of the century³. The SNWDP consists of three subproject routes (namely, the East, Middle and West Routes) that will divert 44.8 billion m³ of water annually from the water-rich Yangtze River Basin. Similar to other large-scale water transfer projects elsewhere in the world, the SNWDP aims to relieve water stress in drought-stricken areas (specifically in North China), promote socioeconomic development and help improve the eco-environment in water-deficient regions³⁶. But for the water-diverted region, the activities of the impounded and diverted water and the resultant change of water cover will have complex and profound effects on the regional eco-environment, hydrology, local climate, and patterns of land use and land cover³⁶. So a lot of researches have been keeping eyes on influences from large-scale water-transfer projects on the water-diverted regions.

Now one focus is about how to get reliable information on land-cover changes in riparian zones and the resultant variation of soil pollution risk that’s essential
for objectively evaluating the influences of water-transfer projects. Compared to other areas in water-diverted regions, riparian zones deserve serious attentions\textsuperscript{11-17}. As an interface between land and water, wild riparian zones act like natural biofilters that can remove pollutants from surface runoff before entering the water body. Fennessy and Cronk\textsuperscript{18} showed that a riparian buffer zone of 20 to 30 m width can remove up to 100\% of incoming non-point nitrate. Zhao et al. (2009)\textsuperscript{19} found that the retention rate of agricultural non-point nitrogen pollution by three vegetation types in the riparian zone of the Yellow River wetland in China could exceed 50\%. However, if human activities such as crop plantation exist in the riparian zone, more agricultural NPS pollutants and soil heavy metals are held in the area, which conversely increases the potential risk of soil pollution\textsuperscript{20}; and this risk is greater in new riparian zones resulting from the artificial activities of impounding and diverting water\textsuperscript{21,22}. Thus, many studies have focused on the influence of large-scale inter-basin water-transfer projects on the risk of soil pollution stemming from the associated changes of riparian land cover in the water source area\textsuperscript{20}.

However, obtaining precise information on the coverage and variation of the riparian zones was, and still is, a significant problem in such studies. As is known, riparian zones result from water level fluctuation, and can be viewed as temporary water cover according to definitions of land cover\textsuperscript{23-25}. It thus seems feasible to obtain information on riparian zones using water coverage data extracted at different water levels from remote-sensing imageries, such as through threshold segmentation of modified normalized difference water index (MNDWI) images\textsuperscript{26-28}. But, this popular method for water identification is not always suitable for riparian zones, especially those that emerged from large-scale engineering projects involving the artificial impounding and diverting water projects. The reason lies not only in the difficulties associated with obtaining an accurate threshold with constant water level fluctuation, but also in overemphasizing the hydrological characteristics of riparian zones and failing to thoroughly understand the zone’s nature. It is impossible for a riparian zone to be same to the water cover, even if inundated, which can be proven by the former’s higher heterogeneity\textsuperscript{16} and unique characteristics of seasonal rotation between water and land. Thus, errors often occur when extracting information on riparian zones from remotely sensed classification data of water coverage at different water levels. Additionally, misclassification is still common when extracting water coverage in hilly areas from remote-sensing imageries, because accurately separating water from built-up areas and the shadows of mountains remains a problem\textsuperscript{29-33}. Due to the “error amplification phenomenon”\textsuperscript{34}, larger error rates consequentially emerge when riparian zones are identified using water coverage data at higher levels minus those at lower levels. Therefore, precise identification of riparian zones is indispensable for scientific evaluations of the influence of large-scale water-transfer engineering projects on land-cover change and the potential risk of soil pollution in the water-diverted regions.

The objectives of this study were to propose a scheme for getting reliable data on riparian zone variation and land-cover change, and further exploring the consequent potential risk of riparian soil pollution with water level fluctuation from the SNWDP’s Middle Route on its headwater area. The idea was firstly implemented by digital visualization of the riparian cyclical features of water–land rotation that were recorded by time-series remote sensing data. Then, a computer-vision color clustering algorithm was applied to the picked out thematic imagery from an optimized RGB color-matching option, which ensured to extract precise information on the coverage of the riparian zones at different water levels. Lastly, combining the information on variation of the riparian
zones of Danjiangkou Reservoir at different water levels with land-cover change data up to 2018, an ECM-based estimation of the NPS pollution loads of TN and TP was completed to explore the potential risk of soil pollution during the initial running period of the SNWDP’s Middle Route.

Study area and data sources

Study area. Danjiangkou Reservoir is one of the largest artificial freshwater lakes in Asia, with water surface area of roughly 745 km² and average depth of 23 m². It is located in the middle reaches of the Hanjiang River, the largest tributary of the Yangtze River in China (Fig. 1a). The catchment of Danjiangkou Reservoir encompasses six counties with a combined population of about 3.14 million people, and has approximately 200 tributaries including Han River and Dan River. The study area (Fig. 1b) covers the whole of Danjiangkou Reservoir that mainly lies in Xichuan County and Danjiangkou City.

The catchment of Danjiangkou Reservoir is the only headwater source of the SNWDP’s Middle Route. During the project’s first running phase, an average of 9.5 billion m³ of water per year were planned to be transformed from the reservoir to Beijing city, Tianjin city, and other target areas. To meet the demands, the height of Danjiangkou Dam was raised from 162 m in 2005 to 176.6 m in 2014. On 12 December 2014, the reservoir’s water level was raised to 160.33 m and the SNWDP’s Middle Route commenced transferring water.

Data sources. Water level and Landsat imageries. To explore inherent links between the water level change and the riparian zone variation, real-time data of water level elevation (WLE) in Danjiangkou Reservoir (Fig. 2) were collected from several public websites (http://www.cjmsa.gov.cn, www.moc.gov.cn, http://www.cjh.com.cn and http://www.hbnsbd.gov.cn/zhxx/GQ.aspx). The water level of Danjiangkou Reservoir varied mainly with natural seasonal change of inflow and outflow runoff before the SNWDP’s Middle Route started running. But, since the project went into service on 12 December 2014, the water level has been observably affected by the man-made activities of impounding and diverting water, more than that by natural hydrological processes, and numerous new riparian zones therein came into being. It’s hence important to study relationship between the reservoir’s water level change and its riparian zone variation.

Additionally, for knowing the real-time changes of the riparian land-cover, some specific Landsat imageries at the lowest and highest annual water levels during 2000–2018 were collected (Table 1).

Soil pollution and other data. The collected soil pollution data (0–20 cm) were focused on agricultural NPS pollutants and heavy metals from cropland in the new riparian zones. They comprised 59 NPS samples from field investigation (Fig. 3a) and 169 sample sites of soil heavy metals from existing studies (Fig. 3b), which mostly spread in the riparian zones between water levels of 150 and 170 m.

Other data consisted of land use data during 1990–2010 and Google Earth images at a resolution of 0.5 m, which provided information on the past croplands and improved the remotely sensed interpretation of the present-day croplands in the riparian zone.
Figure 1. (a) overview of the SNWDP’s Middle Route; (b) location of Danjiangkou Reservoir.

Figure 2. Real-time water level of Danjiangkou Reservoir (2011–2018)

Table 1. Adopted Landsat images and real-time water level data of Danjiangkou Reservoir: (a) closer to the dead water level of 139 m; b: closer to normal retaining level of 170 m.

| Image acquisition Date | Satellite Sensor | WLE (m) | Image acquisition Date | Satellite Sensor | WLE (m) |
|------------------------|------------------|---------|------------------------|------------------|---------|
| 2000/03/19             | TM               | 135.12  | 2015/09/05             | OLI              | 152.70  |
| 2006/05/23             | ETM+             | 146.04  | 2015/12/26             | OLI              | 152.67  |
| 2010/05/02             | TM               | 144.25  | 2016/04/16             | OLI              | 150.43  |
| 2010/12/04             | ETM+             | 149.83  | 2016/09/23             | OLI              | 154.51  |
| 2014/01/21             | OLI              | 139.56a | 2017/03/02             | OLI              | 152.59  |
| 2014/07/08             | OLI              | 141.73  | 2017/10/28             | OLI              | 166.94b |
| 2014/12/07             | OLI              | 160.33  | 2017/12/31             | OLI              | 165.45  |
| 2015/04/14             | OLI              | 156.86  | 2018/12/18             | OLI              | 157.08  |
**Methods**

Digital visualization of the riparian zone’s water–land rotation. Riparian zones have unique characteristics of annual cyclical rotation between water and land. Time series of remotely sensed imageries can record the seasonal change in riparian zones, thereby making it possible to gain reliable information on riparian zone variation and landscapes changes in the zone.

Riparian zones have unique characteristics of annual cyclical rotation between water and land. Time series of remotely sensed imageries can record the seasonal change in riparian zones, thereby making it possible to gain reliable information on riparian zone variation and landscapes changes in the zone.

To transform the key properties of riparian zones into available digital and visual information, three steps were implemented:

Firstly, a normalized difference water index (NDWI) was constructed based on the Landsat imagery at high water level of Danjiangkou Reservoir, while a normalized difference vegetation index (NDVI) and MNDWI were employed at the reservoir’s lower water level. These indices are shown as Eq. (1) – (3).

\[
\text{NDVI} = \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + \rho_{\text{red}}} \tag{1}
\]

\[
\text{NDWI} = \frac{\rho_{\text{green}} - \rho_{\text{swir}1}}{\rho_{\text{green}} + \rho_{\text{swir}1}} \tag{2}
\]

\[
\text{MNDWI} = \frac{\rho_{\text{green}} - \rho_{\text{swir}2}}{\rho_{\text{green}} + \rho_{\text{swir}2}} \tag{3}
\]

Where \(\rho_{\text{green}}, \rho_{\text{red}}\) and \(\rho_{\text{nir}}\) represent the reflectance of the green, red and near-infrared bands corresponding to Landsat/OLI bands 3, 4 and 5 (bands 2, 3 and 4 for TM/ETM+), respectively; \(\rho_{\text{swir}1}\) and \(\rho_{\text{swir}2}\) represent the reflectance of different mid-infrared bands corresponding to Landsat/OLI bands 6 and 7 (bands 5 and 7 for TM/ETM+), respectively.

Secondly, these indices were combined into a thematic dataset, in which the recorded information came from characteristics not only of the riparian zone as a temporary water
cover area at the higher water level, but also of its rotation with other landscapes at the lower water level.

Lastly, a composite RGB image was generated from the thematic dataset, with one color (i.e., red, green or blue) given to one of the above indices so as to achieve visualization of the riparian zone’s water–land rotation. For instance, the colorful image shown in Fig. 4a was generated with green given to NDWI at high water level (166.94 m) on Oct. 28, 2017, red to NDVI and blue to MNDWI at lower water level (152.59 m) on Mar. 2, 2017.

Identification of riparian zones based on the optimized RGB-fusion image color clustering. Compared to geometric features (e.g., texture and shape) in images, color has greater stability and has been widely applied in computer-vision detection in the region of interest (ROI)\textsuperscript{42-44}. Generally, if the ROI has its own unique color features in the image, it is certainly feasible to directly identify the ROI based on color difference\textsuperscript{45}.

However, here, six kinds of composite RGB images may be generated from a three-band dataset, the ROI’s color therefore could be changed by setting different color schemes (e.g., various colors of the same riparian zones in Fig. 4). Thus, a problem was inevitably caused—that is, which of the RGB-fusion images was optimized for precise identification of the riparian zone?

To solve the above problem and in view of perceptually accurate retrieval results on prominent colored areas in the RGB domain able to be produced from measures based on the angle of a color vector\textsuperscript{46,47}, a modified vector-angular distance (MVAD) formula\textsuperscript{48} was adopted to measure color difference between riparian and non-riparian zones in RGB color space. The MVAD formula is as shown in Eq. (4):

$$MVAD(x_i, x_j) = \sqrt{\frac{s_r^2 w_r (r_i - r_j)^2 + s_g^2 w_g (g_i - g_j)^2 + s_b^2 w_b (b_i - b_j)^2}{(w_r + w_g + w_b)255^2} + S_\theta \cdot S_{ratio} \cdot \theta^2}$$

(4)

Where $x_i = (r_i, g_i, b_i)$ and $x_j = (r_j, g_j, b_j)$ are two color vectors in RGB space. They all have three color components, i.e. red ($r$), green ($g$), and blue ($b$) with their values in the range [0, 255]. $w_r$, $w_g$, and $w_b$ depicts the sensitivity of human perceptual to the red, green, and blue colors with their values equal to 1, 2, 1, respectively. $S_r$, $S_g$, and $S_b$ are weight coefficients of these color components.

$$S_r = \min\left[\frac{r_i + r_j}{\sum (r, g, b)}, 1\right]$$

(5)
\[ S_g = \min \left[ \frac{g_i + g_j}{\sum (r, g, b)}, 1 \right] \]
\[ S_b = \min \left[ \frac{b_i + b_j}{\sum (r, g, b)}, 1 \right] \]
\[ \sum (r, g, b) = (r_i + r_j + g_i + g_j + b_i + b_j)/3 \]

\[ \theta = \frac{2}{\pi} \arccos \left( \frac{r_i r_j + g_i g_j + b_i b_j}{\sqrt{r_i^2 + g_i^2 + b_i^2} \cdot \sqrt{r_j^2 + g_j^2 + b_j^2}} \right) \]

\[ S_\theta \] represents the contribution from every single color component on the angular between two vectors in RGB space,

\[ S_\theta = S_{\theta_r} + S_{\theta_g} + S_{\theta_b} \]

\[ S_{\theta_r} = \frac{|r_i - r_j|}{|r_1 - r_2|} \cdot S_r^2 \]

\[ S_{\theta_g} = \frac{|g_i - g_j|}{|g_1 - g_2| + |b_1 - b_2|} \cdot S_g^2 \]

\[ S_{\theta_b} = \frac{|b_i - b_j|}{|g_1 + g_2| + |b_1 + b_2|} \cdot S_b^2 \]

\[ S_{\text{ratio}} = \max \left( \frac{r_i r_j + g_i g_j + b_i b_j}{255} \right) \]

**Fig. 5** showed the measuring results on color difference from the two kinds of color matching schemes in **Fig. 4**. And the RGB-fusion image of **Fig. 4b** could be found better with the larger color distance between riparian and non-riparian zones as shown in **Fig. 5b** than that in **Fig. 5a**.

At last, RGB color clustering-based unsupervised classification algorithm was applied to the optimized RGB-fusion images, and clusters of the riparian zone at different water levels were all picked out.

**Classification of land cover in the riparian zone.** Approximate maximum coverage of the riparian zone was firstly derived from Landsat/OLI imageries on 21 January 2014 and 28 October 2017 with water levels of 139.56 m (close to the reservoir’s dead water level of 139 m) and 166.94 m (close to the record highest water level of 167 m), respectively. All land-cover types in the largest riparian zone were then derived from existing land-use data and the late classification of Landsat images during 2010–2018. Using the data in the largest riparian zone (**Fig. 6**), real-time information on land-cover change at any water level could be extracted. Moreover, an error confusion matrix was built for evaluating the accuracy of the land-cover classification by using GPS-based field investigation data and Google Earth images at a resolution of 0.5 m. The resultant Kappa coefficients were all over 90%.

**Estimation of NPS pollution loads.** The ECM has been widely applied in estimating NPS pollution loads. It uses an export coefficient approach to calculate the loads of
TN and TP delivered annually to a water body as the sum of the individual loads exported from each source within a catchment\textsuperscript{49-50}.

The universal ECM\textsuperscript{49} is as follows:

\[ L = \sum_{i=1}^{n} E_i[A_i(I_i)] + p \]  

(Eq.15)

where \( L \) is loss of nutrients (kg), \( E_i \) is the export coefficient for nutrient source \( i \) (kg/km\(^2\)/yr), \( A_i \) is the area of the catchment occupied by land-use type \( i \) (km\(^2\)), or number of livestock type \( i \), or of people, \( I_i \) is the input of nutrients to source \( i \) (kg), and \( p \) is the input of nutrients from precipitation (kg).

Because of the particular nature of Danjiangkou Reservoir’s role as the headwater source of the SNWDP’s Middle Route, nutrient sources (\( A_i \)) in the reservoir’s riparian zone comprised the area of each land-use type through fertilizer applications and nitrogen fixation—that is, people, livestock, and precipitation could not be included in the model for our research.
Table 2. Export coefficients of key land-use types in Danjiangkou Reservoir Area

|              | Rice paddy | Dry land | Woodland | Shrub/Grass land | Uncultivated land |
|--------------|------------|----------|----------|------------------|-------------------|
| (kg/km²/yr) | 655        | 1921     | 2625     | 1032 – 1135      | 1121              |
| TN           | 180        | 79       | 180      | 23 – 38          | 64                |
| TP           |            |          |          |                  |                   |

Figure 7. Temporal changes of the riparian land-cover in the study area during 2000–2018: (a) acreage of the riparian zone based on the water level of 139.56 m; (b) acreage of characteristic land covers in the riparian zones at key water levels

The export coefficient \( E_i \) shows the rate at which nutrients were lost from each identifiable source to the surface drainage network. It has strong spatial heterogeneity and varies among different land-use types. Here, the \( E_i \) of each land-use type in the riparian zone (Table 2) was determined from a literature, in which the export coefficients of different land-use types were determined by local precipitation, terrain, soil and vegetation fraction, etc.

The GR of pollution loads with water level fluctuation was calculated following Eq. (16) to compare the variation of TN and TP loads among different riparian zones.

\[
\Delta L_j = L_j / (j - j_0)
\]  
(Eq.16)

where \( \Delta L_j \) is the GR of TN or TP loads in the riparian zone at water level \( j \), \( L_j \) is the total load of TN or TP in the riparian zone at water level \( j \), and \( j_0 \) is the reservoir’s dead water level (i.e., 139 m).

Results

Spatio-temporal variation of the riparian zone and resultant land-cover change. Due to the water level change in Danjiangkou Reservoir (Fig. 2), the riparian zones varied markedly during 2000–2018. Especially before and after the SNWDP’s Middle Route started transferring water on 12 December 2014, two typical changes of the riparian zones could be clearly found: firstly, the riparian acreage grew considerably compared to that before the water transform project (Fig. 7a). On 28 October 2017, the riparian area increased to 44 481 hm² with the then-water level (i.e., 166.94 m) closer to the project’s normal retention level of 170 m; whereas, on 8 July 2014 with the then-water level of 141.73 m, the riparian area was only 1 625 hm². On 3 January 2000, the reservoir’s water level dropped to 135.12 m (close to its recorded lowest water level), and thus a 6 687 hm² special riparian zone (Fig. 8a) occurred below the reservoir’s 139 m dead water level, which did not occur again after 2014. Secondly, distribution of the riparian zones of...
Danjiangkou Reservoir varied differently at the county scale (Fig. 8a - Fig. 8o). Despite the coverage of the riparian zone having expanded in three counties (Xichuan County, Yunyang County and Danjiangkou City) with the reservoir’s water level rising, always around half of new riparian zones emerged in Xichuan County (Fig. 7a). And as the water level rose from 139 m to 160.33 m, the acreage of new riparian zones in the county enlarged to almost 182 km², whereas the acreage was 80.09 km² and 66.05 km² in Danjiangkou City and Yunyang County, respectively.

Additionally, water level of the reservoir previously fluctuated with the seasonal changes in hydrology with a range from 139 m to 150 m before the SNWDP’s Middle Route (Fig. 2), and a 17 230 hm² riparian zone hence emerged with 716.5 hm² of cultivated land and 101.4 hm² of shrub/grassland in the zone (Fig. 7b). But during the initial water transfer period of the SNWDP’s Middle Route, the reservoir’s water level was always over 150 m (Fig. 2), and the riparian zone apparently enlarged compared to those before the water transfer project started running. On 7 December 2014, the zone’s area was 32 893 hm² and contained around 8 054.65 hm² of cultivated land and 799.67 hm² of shrub/grassland (Fig. 7a). Thereafter, the water level of Danjiangkou Reservoir mostly varied between 150 m and 160 m, although it was raised to 166.94 m on 28 October 2017 with the subsequent area of submerged cultivated land and shrub/grassland reaching 12 896.34 hm² and 6 768.47 hm², respectively (Fig. 7b).
Figure 8. Mapping riparian zones at different water levels in the study area during 2000–2018: (a) at the 135.12 m water level on 3/19/2000; (b) at the 146.04 m water level on 5/23/2006; (c) at the 144.25 m water level on 5/2/2010; (d) at the 149.83 m water level on 12/4/2010; (e) at the 141.73 m water level on 7/8/2014; (f) at the 160.33 m water level on 12/7/2014; (g) at the 156.86 m water level on 4/14/2015; (h) at the 152.67 m water level on 12/26/2015; (i) at the 150.43 m water level on 4/16/2016; (j) at the 154.31 m water level on 9/23/2016; (k) at the 152.59 m water level on 3/2/2017; (m) at the 166.94 m water level on 10/28/2017; (n) at the 165.43 m water level on 12/31/2017; (o) at the 157.08 m water level on 12/18/2018
Table 3. TN and TP content of soil samples from riparian cropland in the study area

| TN (g/kg) | <0.5 | 0.5-0.75 | 0.75-1 | 1.0-1.5 | 1.5-2 | >2.0 |
|-----------|------|----------|--------|---------|-------|------|
| Account of samples | 20   | 13       | 11     | 9       | 4     | 2    |
| TP (g/kg)  | <0.2 | 0.2-0.4  | 0.4-0.6| 0.6-0.8 | 0.8-1 | >1.0 |
| Account of samples | 0    | 8        | 27     | 12      | 8     | 4    |

Figure 9. Timing change of the GR of TN and TP loads with water level fluctuation in the riparian zone: (a) the GR of TN loads with water level fluctuation; (b) the GR of TP loads with water level fluctuation;

Figure 10. Confusion between hill-shadow and water using classification on remote sensing imagery at higher water levels with Label 1 for water and hill-shadow, Label 2 for mostly deeper water and Label 3 for mostly shoal water: (a) Image of OLI bands 6,5,4; (b) Image of MNDWI, (c) Composite RGB image.

NPS pollution before and after the running of the SNWDP’s Middle Route. According to our field samples of soil pollution (Table 3), there were low background contents of TN and TP in the investigated riparian cropland. In all 59 soil samples, about 90% had a TN value less than 1.5 g/kg (Fig. 3a) and 80% less than 0.8 g/kg (Table 3), which is consistent with the findings of Wang (2015)\(^{21}\). Even so, the potential risk of agricultural NPS pollution from the land-cover change would be enlarged, especially from the rotation between water and agricultural lands in the riparian zone. The reason lies in the fact that most of the agricultural lands in the riparian zone are cultivated at lower water levels for one-seasonal crop plantation with popular use of nitrogenous and phosphorus fertilizers\(^{53}\). It is known that nutrient loss from agricultural land use/land cover (e.g., cropland and grassland) is an important aspect of NPS pollution\(^{49}\). Whereas, the utilization rate of fertilizers was less than 40% in the study area, meaning redundant
nitrogen and phosphorus in the soil will ultimately enter the reservoir through runoff as well as underground water.

Besides, the annual GRs of TN and TP loads in the riparian zones were estimated at different water levels. The result (Fig. 9) shows that their GRs were not only always higher than those before the water transfer project started running, but also had more significant spatiotemporal variation during the initial running period of the SNWDP’s Middle Route; e.g., the GRs of TN loads dropped sharply from 11.85 t·a⁻¹ on 28 October 2017 to 1.37 t·a⁻¹ on 16 April 2016, while TP ranged from 0.56 t·a⁻¹ to 0.06 t·a⁻¹. Additionally, the GRs of the two loads in Xichuan County were the highest in the study area, and the GR of TN loads was far higher than that of TP loads regardless of whether the SNWDP’s Middle Route was running.

**Discussion**

** Extraction of riparian zone coverage.** It is crucial to obtain reliable information on the coverage of riparian zones and resultant change of land cover in the zones at different water levels for the precise estimation of the soil pollution risk in water-diverted areas. Previously, a popular method for extracting the coverage information of riparian zones was through the remote sensing of water coverage at high levels minus that at lower levels. However, this approach is prone to error in identification on riparian zones, especially for new riparian zones that have arisen from the engineering projects of artificially impounded and diverted water actions. Fig. 10 shows how the widespread hill-shadow was often misidentified as water in the process of remotely sensed classification and had to be cleared through careful manual revision of the classification results. However, this confusion not only was difficult to be completely solved but also fueled further error in the extraction of riparian zone coverage.

In order to eliminate this misidentification, this paper took advantage of the unique properties of cyclical rotation between water and land in the remote sensing identification of new riparian zones. For example, in spite of greater similarity between water and hill-shadow in multi-band Landsat OLI imagery (Fig. 10a) or single-band MNDWI imagery (Fig. 10b), apparently different coloring of the riparian zone from water and hill-shadow could be observed in the RGB-fusion image of NDWI at higher water levels and NDVI and MNDWI at lower water levels (Fig. 10c). Therein, the riparian zones at different water levels could be easily and accurately extracted with a color clustering algorithm applied to the RGB-fusion image. Based on time-series data of riparian coverage (Fig. 8), a general law of riparian zone’s variation with water level fluctuation in the study area was achieved, e.g., most new riparian zones emerged in Xichuan County and the total acreage of riparian zones enlarged significantly after the SNWDP’s Middle Route started running, which matched well with the reported research results on the dynamic capacity of Danjiangkou Reservoir from Liu (2018)

**Potential risk of soil pollution from change of the riparian zones.** Change of riparian land cover in the study area was characterized by rotation between water and land, which was mostly composed of shoaly land and agricultural land (Fig. 6). According to our field investigation, the shoal land was stable and prohibited cultivation, while most of the agricultural lands in the riparian zone around Danjiangkou Reservoir were cultivated for one-seasonal crop plantation at lower water levels. Because the popular use of mineral fertilizers in local conventional agricultural practices and nutrient loss from agricultural
land use/land cover (e.g., cropland and grassland) are important sources of NPS pollution\textsuperscript{49,52}, the GR of TN and TP loads from the riparian cropland were certainly larger with higher water levels (Fig. 9). Given that some of the shoal land and grassland resulted from abandoned croplands, the true amount of NPS pollution loads in the riparian zone was in fact higher than the calculated amounts in this paper.

Moreover, we didn’t investigate the loads of soil heavy metals and their change in the riparian zone with water level rising. However, an existing study\textsuperscript{22} showed Cd and As were the dominant soil heavy-metal pollutants, in which the content of Cd (1.04 mg/kg) was significantly higher than its background value (0.065 mg/kg) in the sampling area. Because of the reservoir’s role as the only one drinking-water source of the SNWDP’s Middle Route\textsuperscript{27}, close attentions in future should be paid to the potential risk of soil heavy-metal pollution from land-cover change in the riparian zone.

Conclusions

The purpose of this study was to present a scheme to acquire the reliable information of the riparian land-cover change, and to further explore the resultant soil pollution risk in the headwater area of the SNWDP’s Middle Route. Given that riparian zones are characterized by high land-cover heterogeneity and their coverage cannot be precisely extracted from the remotely sensed identification of water coverage at high levels minus that at low levels, this study proposed a scheme for digitizing and visualizing the unique water–land rotational feature of the riparian zone and applied a color clustering algorithm to the optimized RGB-fusion images of NDWI at high water levels and NDVI and MNDWI at lower water levels. The distinguishing performance of the method were showed, especially at accurate identification of the riparian coverage from hill-shadow in the remote sensing imagery. A law of riparian zone variation and the resultant land-cover change with water level fluctuation was also achieved for the study area. The submerged land cover in the riparian zone at the water level of 166.94 m mainly comprised shoal land, cropland and grassland, among which shoal land was dominant. Additionally, most of the new riparian zones emerged in Xichuan County, and the total acreage of the riparian zones and submerged croplands enlarged significantly after the SNWDP’s Middle Route started running.

Based on dynamic information of riparian land cover and field investigation data of soil pollution samples, an ECM-based calculation of NPS pollutant loads in the riparian zone was completed at different water levels during 2006–2018. It was found that the potential risk of soil NPS pollution from land-cover change in the riparian zone rose to a higher level than that before the implementation of this large-scale water-transfer project. The GRs of TN and TP loads were not only higher in the riparian zone during the initial running period of the SNWDP’s Middle Route, but also possessed significant spatial variation insofar as the GRs of TN and TP loads in Xichuan County were always the highest in the study area. Moreover, the GR of TN loads was far higher than that of TP loads, regardless of whether the SNWDP’s Middle Route was kept running.

As a large-scale inter-basin water transfer project, the SNWDP’s Middle Route has the potential to cause complex and profound effects on water-diverted regions from the man-made activities of the impounded and diverted water. This study can provide accurate insights into the project’s influence.
References

[1] Gohari, A., Eslamian, S., Mirchi, A., Abedikoupaei, J., Bavani, A.M. & Madani, K. Water transfer as a solution to water shortage: A fix that can backfire. Journal of Hydrology 491, 23–39 (2013).

[2] Andrade, J.G.P.D., Barbosa, P.S.F., Souza, L.C.A. & Makino, D.L. Interbasin water transfers, the brazilian experience and international case comparisons. Water Resources Management 25, 1915–1934 (2011).

[3] Yu, M., Wang, C., Liu, Y., Olsson, G. & Wang, C.Y. Sustainability of mega water diversion projects, experience and lessons from China. Science of The Total Environment 619–620, 721–731. DOI:10.1016/j.scitotenv.2017.11.006 (2018)

[4] Roozbahani, A., Ghased, H. & Shahdany S.M.H. Inter-basin water transfer planning with grey COPRAS and fuzzy COPRAS techniques, a case study in Iranian Central Plateau. Science of The Total Environment 726, 138499. (2020)

[5] Zhang, L., Li, S.S., Loáiciga, H.A., Zhuang, Y.H. & Du, Y. Opportunities and challenges of interbasin water transfers, a literature review with bibliometric analysis. Scientometrics 105: 279–294 (2015).

[6] Kattel, G.R., Shang, W.X., Wang, Z.J. & Langford, J. China's South-to-North Water Diversion Project empowers sustainable water resources system in the North. Sustainability 13: 3735 (2019).

[7] Li, L.C., Zhang, L.P., Xia, J., Gippel, C.J., Wang, R.C. & Zeng, S.D. Implications of modelled climate and land cover changes on runoff in the Middle Route of the South to North Water Transfer Project in China. Water Resources Management 29: 1–17 (2015).

[8] Quan, Y., Wang, C.X., Yan, Y., Wu, G. & Zhang, H.X. Impact of inter-basin water transfer projects on regional ecological security from a telecoupling perspective. Sustainability 8: 162 (2016).

[9] Li, S.Y., Chen, Y. & Zhang, Q.F. 11-year change in water chemistry of large freshwater Reservoir Danjiangkou, China. Journal of Hydrology 551: 508–517 (2017).

[10] Xin, X.K., Yin, W. & Li, K.F. Estimation of non-point source pollution loads with flux method in Danjiangkou Reservoir area,China. Water Science and Engineering 10: 134–142 (2017).

[11] Naiman, R.J. & Decamps, H. The Ecology of interfaces: riparian zones. Annual Review of Ecology and Systematics 28: 621–658 (1997).

[12] Lü, M.Q., Wu, S.J., Chen, C.D., Jiang, Y., Wen, Z.F., Chen, J.L., Wang, Y., Wang, X.X. & Huang, P. A review of studies on water level fluctuating zone (WLFZ) of the Three Gorges Reservoir (TGR) based on bibliometric perspective. Acta Ecologica Sinica 35: 3504–3518 (2015).

[13] Weisssteiner, C.J., Ickerott, M., Ott, H., Probeck, M., Ramminger, G., Clerici, N., Durfourmont, H. & Sousa, A.M.R.D. Europe’s Green Arteries—a continental dataset of riparian zones. Remote Sensing 8: 925 (2016).

[14] Klemas, V. Remote sensing of riparian and wetland buffers: an overview. Journal of Coastal Research 30: 869–880 (2014).

[15] Gong, Z.N., Cui, T.X., Pu, R.L., Lin, C. & Chen, Y.Z. Dynamic simulation of vegetation abundance in a reservoir riparian zone using a sub-pixel Markov model. International Journal of Applied Earth Observation and Geoinformation 35: 175–186 (2015).

[16] Jeong, S.G., Mo, Y.W., Kim, H.G., Park, C.H. & Lee, D.K. Mapping riparian habitat using a combination of remote-sensing techniques. International Journal of Remote Sensing 37: 1069–1085 (2016).

[17] Fu, B.L., Li, Y., Wang, Y.Q., Campbell, A. & Jin, X.M. Evaluation of riparian condition of Songhua River by integration of remote sensing and field measurements. Scientific Reports 7: 2565. DOI:10.1038/s41598-017-02772-3 (2017)
[18] Fennessy, M.S. & Cronk, J.K. The effectiveness and restoration potential of riparian ecotones for the management of nonpoint source pollution, particularly nitrate. Critical Reviews in Environmental Science and Technology 27: 285–317 (1997).

[19] Zhao, T.Q., Xu, H.S., He, Y.X., Tai, C., Meng, H.Q., Zeng, F.F. & Xing, M.L. Agricultural non-point nitrogen pollution control function of different vegetation types in riparian wetlands: A case study in the Yellow River wetland in China. Journal of Environmental Sciences 21: 933–939 (2009).

[20] Hao, F.H., Zhang, X., Wang, X. & Wei, O.Y. Assessing the relationship between landscape patterns and nonpoint-source pollution in the Danjiangkou Reservoir Basin in China. - Journal of the American Water Resources Association 48: 1162–1177 (2012).

[21] Wang, J., Yin, W., Zhao, X.L., Yang, M., Duan, C., Zhu, D. & Shi, Z.H. The potential risk evaluation of farmland soil from new submerged area in Danjiangkou Reservoir. China Environmental Science 35: 157–164 (2015).

[22] Han, P.P., Xie, J., Wang, J., Qiang, X.Y., Ai, L. & Shi, Z.H. Source apportionment of heavy metals in farmland soil from new submerged area in Danjiangkou Reservoir. China Environmental Science 36: 2437–2443 (2016).

[23] Turner II, B.L., Skole, D., Sanderson, S., Fischer, G., Fresno, L.O. & Leemans, R. Land-Use and Land-Cover Change Science/Research Plan. Global Change Report 43: 669-679 (1995).

[24] Rycroft, M.J. Our Changing Planet: The FY 1995 U.S. Global Change Research Program. - Journal of Atmospheric and Solar-Terrestrial Physics 59: 2247–2248 (1997).

[25] Wang, Q.F., Li, J.Y. & Chen, B.M. Land Cover classification system based on spectrum in Poyang Lake Basin. Acta Geographica Sinica 61: 359–368 (2006).

[26] Makkeasorn, A., Chang, N.B., Li & J.H. Seasonal change detection of riparian zones with remote sensing images and genetic programming in a semi-arid watershed. Journal of Environmental Management 90: 1069–1080 (2009).

[27] Zhang, Y.N., Huang, X., Yin, W. & Zhu, D. Multitemporal landsat image based water quality analyses of Danjiangkou Reservoir. Photogrammetric Engineering and Remote Sensing 83: 643–652 (2017).

[28] Gong, Z.N., Chen, Y.Z., Zhao & W.H. Analysis of fluctuation characteristics and subzone in hydro-fluctuation belt in the Guanting Reservoir during 1979-2013. Acta Ecologica Sinica 37: 5294–5304 (2017).

[29] Sun, P., Wang, Q.F., Zhang, M.R., Yang, H. & Wang, Q.Q. A method to enhance information of water cover based on feature space of NDVI and MNDWI. Journal of Hubei University(Natural Science) 40: 574–579 (2018).

[30] Dai, Y.F., Gao, Y., Zhang, G.Q. & Xiang, Y. Water volume change of the Paiku Co in the Southern Tibetan Plateau and its response to climate change in 2003-2011. Journal of Glaciology and Geocryology 35: 723–732 (2013).

[31] Xia, S., Ruan, R.Z. & Yan, M.C. Study on information extraction method of water body based on TM imagery. Journal of Nanjing Forestry University 36: 29–35 (2012).

[32] Xu, H.Q. Modification of Normalised Difference Water Index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing 27: 3025–3033 (2006).

[33] Chen, H.F., Wang, J.L., Chen, Z., Yang, L. & Xi, W.J. Comparison of water extraction methods in mountainous plateau region from TM image. Remote Sensing Technology and Application 19: 479–484 (2004).

[34] Shao, G.F. The influence of remotely sensed thematic maps on landscape ecology studies. Ecologica Sinica 24: 1857-1862 (2004)

[35] Liu, W.Z., Liu, G.H. & Zhang, Q.F. Shoreline vegetation in the Danjiangkou Reservoir: characteristics, related factors, and differences with adjacent riverine wetlands. CLEAN - Soil, Air, Water 42: 1014–1021 (2014).

[36] Xin, X.K., Li, K.F., Finlayson, B. & Yin, W. Evaluation, prediction, and protection of water quality in Danjiangkou Reservoir, China. Water Science and Engineering 8: 30–39 (2015).
Yin, D.C., Zheng, L.L. & Song, L.R. Spatio-temporal distribution of phytoplankton in the Danjiangkou Reservoir, a water source area for the South-to-North Water Diversion Project (Middle Route), China. Chinese Journal of Oceanology and Limnology 29: 531–540 (2011).

Wu, C., Zhang, Y.L., Xu, X.Z. & Zhang, Q.F. Water surface monitoring of the Danjiangkou Reservoir using remotely-sensed Landsat TM/ETM and HJ-1A/B imagery. - Resources and Environment in the Yangtze Basin 22: 1207–1213 (2013).

Fu, J.Y. Studies on the submergence tolerance and drought tolerance of living suited plants of hydro-fluctuation belt in Danjiangkou Reservoir Area. (Huazhong Agricultural University, Wuhan China, 2013).

Liu, H., Wu, J., Yin, J., Wang, M. & Chen, X.L. Evaluation and change of Danjiangkou Reservoir dynamic capacity. Chinese Journal of Applied Ecology 29: 2658–2666 (2018).

Mefoeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing 17: 1425–1432 (1996).

Sonka, M., Hlavac, V. & Boyle, B. Image Processing, Analysis and Machine Vision (4th ed.). (Tsinghua University Press, Beijing China, 1993).

Meng, Q.K., HE, J., Qiu, R.C., Ma, X.D., Si, Y.S., Zhang, M. & Liu, G. Crop recognition and navigation line detection in natural environment based on machine vision. Acta Optica Sinica 34: 172–178 (2014).

Pundir, A.S. & Raman, B. Dual deep learning model for image based smoke detection. Fire Technology 55: 2419–2442. https://doi.org/10.1007/s10694-019-00872-245 (2019).

Wang, Q.F., Zhang, M.R., Zhang, Y., Wang, Q.Q., Chen, L.Y. & Yang, Y.Q. Intelligent extraction of remote sensing information on large-scale water based on visual attention mechanism. Journal of Computer Applications 40: 1038–1040 (2020).

Syed, K. & Srinivasu, S.V.N. A review of web image mining tools, techniques and applications. International Journal of Computer Trends and Technology 49: 35–43 (2017).

Androutos, D., Platanitis, K. N. & Venetsanopoulos A. N. A novel vector-based approach to color image retrieval using a vector angular-based distance measure. Computer Vision and Image Understanding 75: 46–58 (1999).

Yang, Z.Y., Wang, Y, Yang, Z.D. & Wang, C.D. Vector-angular distance color difference formula in RGB color space. Compute Engineering and Applications 46: 154–156 (2010).

Johnes, P.J. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. Journal of Hydrology 183: 323–349 (1996).

Ding, X.W., Shen, Z.Y., Hong, Q., Yang, Z.F., Wu, X. & Liu, R.M. Development and test of the Export Coefficient Model in the upper reach of the Yangtze River. Journal of Hydrology 383: 233–244 (2010).

Rong, Q.Q., Cai, Y.P., Chen, B., Yue, W.C., Yin, X.A. & Tan, Q. An enhanced export coefficient based optimization model for supporting agricultural nonpoint source pollution mitigation under uncertainty. Science of The Total Environment 580: 1351–1362 (2017).

Fang, N.F., Shi, Z.H. & Li, L. Application of Export Coefficient Model in simulating pollution load of non-point source in Danjiangkou Reservoir Area. Journal of Hydroecology 32: 7–12 (2011).

Cheng, X.L., Yang, Y.H., Li, M., Dou, X.L. & Zhang, Q.F. The impact of agricultural land use changes on soil organic carbon dynamics in the Danjiangkou Reservoir area of China. Plant and Soil 366: 415–424 (2013).

Liu, R.H., Kang, Y.H., Zhang, C., Pei, L., Wan, S.Q., Jiang, S.F., Liu, S.P., Ren, Z.Y. & Yang, Y. Chemical fertilizer pollution control using drip fertigation for conservation of water quality in Danjiangkou Reservoir. Nutrient Cycling In Agroecosystems 98: 295–307 (2014).
Acknowledgements

This research was funded by the Key Technology Innovation Project of Hubei Province (No. 2018ABA078). This work also benefited from the Program of Agricultural and Rural Resources Monitoring of the Ministry of Agriculture, China.

Author Contributions

Quanfang Wang wrote the main manuscript text and conceived the research. Mengru Zhang, Yu Zhang, Yuqi Yang and Ying Xiao conducted the data analysis. Longyue Chen and Lanfang Yang contributed to analysis on field samples of soil pollution. Xinsheng Wang coordinated the research and contributed to the resources and funding acquisition.

Competing Interest

The authors declare no conflict of interest.

Additional information

Correspondence and requests for materials should be addressed to Q. W. or Y. Z. Reprints and permissions information is available at www.nature.com/reprints. Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.