

Replies to the comments and suggestions

Dear editor and Anonymous Referee #2:

Thank you for your letter and for the reviewers' comments concerning our manuscript entitled "Recognition of spatial framework for water quality and its relation with land use/cover types from a new perspective: A case study of Jinghe Oasis in Xinjiang, China". Those comments are all valuable and very helpful for revising and improving our paper, as well as important guiding sense to our researches. The authors have studied comments carefully and have made correction which we hope meet with approval. Revised portions are marked in red in the paper and authors have tried our best to revise the manuscript based on their suggestions. The point-by-point responses to each of the comments are presented as follows.

Anonymous Referee #2

This manuscript conducted a very detailed work on investigating the water quality and its relation with land use/cover types. While the data and samples have been processed carefully and rigorously, the results look reasonable and trustful. However, I have several major concerns regarding to this work.

Reply: Thank you for your positive evaluation and encouragement. We have carefully revised the paper according to your comments and suggestions. The responses to the specific comments are presented below

Q: There are two key words in the title that I cannot find strong evidence from the context to support, the "spatial" and "a new perspective". I cannot see how the authors make their findings 'spatial', since there is nothing that has been shown as a map, except for the study area (Figure 1) and land cover (Figure 6). From the Introduction, I also cannot see why the authors labeled their work as "a new perspective"? It is new because of what? This should be made clear in the Introduction.

Reply: Thanks to Reviewers for their suggestions in improving the manuscript, based on reviewer 1, the title is revised; 'spatial' is changed "spatial pattern". "a new perspective" is deleted by authors.

The title has changed into "Recognizing spatial patterns in water quality and their relation with land use/cover types: A case study of the Jinghe Oasis in Xinjiang, China"

Q: I am also unclear about the role that remote sensing data played in this study. What was it for, only for revealing the land use/cover types of each sample? How was it used? It was said there was a 1km buffer zone established for each sample. Then how? And why 1km? Besides, in line 246, there is the accuracy of land use/cover classification. I am wondering how these accuracy numbers were derived? Using what as reference

(ground truth)?

Reply: In order to extract the area of land use/cover types, therefore, remote sensing data is selected in this study.

1km buffer zone established for each sample in this study, based on most research (Wang et al., 2016; wang et al., 2017) and regional natural situation, Wang et al., (2016) reports a correlation between croplands areas and water quality in rivers, but this correlation isn't significant in 500 m. Wang et al., (2017) reports a significant correlation between land use and cover areas and water quality in kilometer scale in rivers. Because the study area is located in the arid areas of Xinjiang, desert dust and salt dust are major environmental hazards in this study area. Desert dust and salt dust seriously affect the atmosphere and water quality and accelerate the degradation of vegetation and threaten ecological security in the oasis. However, with economic development of and population growth, human improvement of desert land and saline land increases the area of croplands (Yu et al., 2017), and small changed in landscape. Therefore, 1km buffer zone established for each sample in this study.

The accuracy of land use/cover classification is required by confusion matrix, historical data, high-resolution Google Earth images, and field survey data, selected to verify that more than 100 pixels of each land cover type were used for the training data, and the confusion matrix to verify the classification results. From the classification results obtained in May and October 2015 (Figure 6), precision has increased to 89.9750% and 86.2848%, and the kappa coefficients are 0.8681 and 0.8184, respectively by confusion matrix (Table 2). Please see the article on page 6-7, lines to246-250 in revised manuscript with obviously marked as follows:

In this study, historical data, high-resolution Google Earth images, and field survey data were selected to verify that more than 100 pixels of each land cover types were used for the training data, and the confusion matrix was used to verify the classification results. For the classification results obtained in May and October 2015 (Figure 6), precision increased to 89.9750% and 86.2848%, respectively, and the kappa coefficients were 0.8681 and 0.8184, respectively, based on the confusion matrix (Table 2),

Table 2 The calculation of a confusion matrix by a maximum likelihood supervised classification

	LULC	Water body	Saline land	Farmland	Forest grassland	Other land	Total	User's accuracy (%)
May	Water body	144	0	0	0	0	144	100
	Saline land	0	77	0	0	16	93	82.79
	Farmland	0	36	101	0	0	137	73.72
	Forest-Grass land	0	36	0	101	0	137	73.72
	Other land types	1	0	0	0	87	88	98.96
	Total	145	149	101	101	103	Overall=89.9750%	
	Producer's accuracy (%)	99.31	51.67	100	100	84.46	Kappa=0.8681	

October	Water body	144	0	0	0	0	144	100
	Saline land	0	57	0	0	26	83	86.67
	Farmland	0	16	101	0	0	117	86.32
	Forest-Grass land	4	16	0	101	0	117	86.32
	Other land types	0	0	0	0	77	77	100
	Total	148	89	101	101	103	Overall=86.2848%	
	Producer's accuracy (%)	97.29	64	100	100	74.75	Kappa=0.8184	

Reference

Wang XP, Zhang F, Li XH, Cao C, Guo M. 2017. Correlation analysis between the spatial characteristics of land use/cover-landscape pattern and surface-water quality in the Ebinur Lake area. *Acta Ecologica sinica*, 2017, 37(22):7438~7452

Wang J, Zhang F, Zhang Y, Ren Y, Yu HY, 2016. Correlation between the spatial water quality and land use/cover in the Ebinur Lake area *Acta Ecologica sinica*, 2016,36(24):7971~7980.

Yu, H., Zhang, F., Kung, H. T., Johnson, V. C., Bane, C. S., & Wang, J., et al. 2017. Analysis of land cover and landscape change patterns in ebinur lake wetland national nature reserve, china from 1972 to 2013. *Wetlands Ecology & Management*(3), 1-19.

Q: The grammar and usage have pervasive problems. It needs to be comprehensively edited by someone with strong English language skills before publication. There are some words, such as ‘framework’ in the title, and ‘layers’ in line 19, that are very confused to me. I have no idea what they are referring to.

Reply: Thanks to Reviewers for their suggestions in improving the manuscript, based on reviver 1, the title is revised, ‘spatial framework’ is changed “”spatial pattern”. But then Professor **Hsiang-te Kung (University of Memphis)** has helped us improve the grammar and edit the manuscript carefully. Please see the revised version of the paper.

Q: Line 132, Davies-Bouldin index just comes out suddenly. What is this? What is the reference for it? What do you mean by “Through the K-means :”? These content needs to be rewritten.

Reply: The Davies-Bouldin clustering index was used to determine the optimal number of the clusters for a dataset. The sections of about Davies-Bouldin clustering index is rewritten by authors. Please see the article on page 4, lines to 129-133 in revised manuscript with obviously marked. As follows:

The Davies-Bouldin clustering index was used to determine the optimal number of clusters for the dataset (An et al., 2016; Park et al., 2014). The lower the Davies–Bouldin index value is, the better the clusters are

differentiated. The K-means cluster analysis was combined with the Davies–Bouldin index (DBI) to select the clustering number. (Zhou et al., 2016).

Reference

- An, Y., Zou, Z., & Li, R. 2016. Descriptive characteristics of surface water quality in hong kong by a self-organising map. *International Journal of Environmental Research & Public Health*, 13(1):115.
- Park, Y. S., Kwon, Y. S., Hwang, S. J., & Park, S. 2014. Characterizing effects of landscape and morphometric factors on water quality of reservoirs using a self-organizing map. *Environmental Modelling & Software*, 55(5):214-221.

Q: What is the DCA gradient axis in line 163? Why do you say “Therefore, the redundancy analysis (RDA) method was applied to”. I cannot see connections between the sentences before the word ‘Therefore’ and those after it.

Reply: The section is revised by authors. Please see the article on page 4, lines to 161-165 in revised manuscript with obviously marked. As follows:

The results showed that the DCA gradient shaft length was less than 3. Based on the results of Wang et al (2017), when the DCA gradient shaft length is less than 3, the redundancy analysis (RDA) method can explore the relationship; therefore, the redundancy analysis (RDA) method was applied to determine the influence trend of land use/cover changes on the water quality within the Ebinur Lake buffer area.

Reference

- Wang XP, Zhang F, Li XH, Cao C, Guo M. 2017. Correlation analysis between the spatial characteristics of land use/cover-landscape pattern and surface-water quality in the Ebinur Lake area. *Acta Ecologica sinica* 2017,37(22):7438~7452

Q: Sentence in line184-185 needs to be rewritten.

Reply: Thanks to Reviewers for their suggestions in improving the manuscript, the sentence is rewritten, Please see the article on page 5, lines to 183-184 in revised manuscript with obviously marked. As follows:

Six clusters were formed according to the DBI, where minimal value is at six clusters as Figure 3a.

In addition, authors have revised the figures and tables as well as words expression in the entire manuscript, please see the text.

Authors tried our best to improve the manuscript and made some changes in the manuscript. Authors appreciate for Editors/Reviewers’ warm work earnestly, and hope that the correction will meet with approval. Once again, thank you very much for your comments and suggestions.

All in all, if you have any questions about our paper, please contact with me as follow address:

E-mail:zhangfei3s@163.com

Thanks very much.

Best wishes and warmly regards for you.

Sincerely yours Fei ZHANG

11th, Jan., 2017

Recognizing spatial patterns in water quality and their relation with land use/cover types: A case study of the Jinghe Oasis in Xinjiang, China

Fei ZHANG^{1,2,3*} Juan WANG⁴ Xiaoping WANG^{1,2,3}

1. College of Resources and Environment Science, Xinjiang University, Urumqi, Xinjiang 830046

2. Key Laboratory of Oasis Ecology, Xinjiang University, Urumqi, Xinjiang 830046

3. Key Laboratory of Xinjiang wisdom city and environment modeling, Urumqi, Xinjiang 830046

4. College of Geography and Remote Sensing Science, Beijing Normal University, 100875, Beijing

Abstract: To understand the relation between spatial water quality patterns and changes in land use/cover types in the Jinghe Oasis, this study divided 47 water sampling sites, which were measured in May and October 2015, into 6 cluster layers using the self-organizing map (SOM) method based on non-hierarchical *k*-means classification. Next, it determined the distribution characteristics of water quality indices for the time sequence. The water quality indices included the chemical oxygen demand (COD), biological oxygen demand (BOD), suspended solids (SS), total phosphorus (TP), total nitrogen (TN), ammonia nitrogen (NH₃-N), chromaticity (SD), and turbidity (NUT). Based on the results, we collected data on the changes in the farmland, forest-grass land, water body, salinized land, and other land types during the wet and dry seasons and combined these data with the classification results of the GF-1 remote sensing satellite data obtained in May and October 2015. We then discussed the influences of land use/cover type on water quality for different layers and seasons. The results indicate that Clusters 1 to 3 included monitoring samples from the wet season (May 2015), whereas Clusters 4 to 6 included monitoring samples from the dry season (October 2015). In general, the COD, SS, NUT, TN, and NH₃-N values were high in Clusters 1 and 2. The SD values for Clusters 1, 4, and 6 were high. Moreover, high BOD and TP values were mainly concentrated in Clusters 4 and 6. In the discussion on the relation between the different water quality parameters and land use/cover type changes, we determined that farmland, forest-grassland, and salinized land significantly influenced the water quality parameters in the Jinghe Oasis. In Clusters 1, 2, and 6, the size of the water area, to a certain extent, also influenced changes in the water quality parameters. In addition, the influences of various land use/cover types on the water quality parameters in the research zone during the different seasons exhibited the following descending order of magnitude: farmland → forest-grass land → salinized land → water body → others. Moreover, their influences were lower during the wet season than the dry season. In conclusion, developing research on the relation between the spatial framework of the water quality in the Jinghe Oasis and land use/cover type changes is significant for the time sequence distribution of water quality in arid regions from both theoretical and practical perspectives.

Key words: SOM; Water quality spatial distribution; Land use/cover; Correlation analysis; GIS

0 Introduction

Water quality is of great importance to the study of water resources in arid regions. Accurate information on the spatial distribution of surface water quality is imperative for assessing environmental monitoring, land-surface water management and watershed changes (NRC, 2008; Sun et al., 2012). Land use/cover changes in drainage basins significantly influence the water quality of rivers, lakes, river mouths, and coastal areas (Huang et al., 2013a; Bu et al., 2014; Hur et al., 2014). Surface water resources, through runoff or infiltration, will always carry a large amount of pollutants (Swaney et al., 2012). Therefore, the spatial allocation of land use and land cover changes in drainage basins frequently influences or even

* Corresponding author: Zhangfei Tel: +86 13579925126 E-mail address: zhangfei3s@163.com

endangers water quality through non-point source pollution (Swaney et al., 2012). However, the regional differences and complexity of land use/cover types result in various relations between land use/cover and water quality in different regions (Yang et al., 2016). Therefore, it is very important to explore the relationship between land use/cover types and water quality for the development and management of the basin. (Uuema et al., 2005; Xiao et al., 2007; Wan et al., 2014). At present, numerous scholars have extensively applied statistical methods to determine the mutual relations between land use/cover changes and water quality in various research zones (Céréghino et al., 2009; Bierman et al., 2011; Huang et al., 2013b). These methods include correlation analysis (Lee et al., 2009; Li et al., 2015), multiple regression (Park et al., 2014), and redundancy analysis (De et al., 2008; Shen et al., 2015).

A self-organizing map (SOM) is a type of artificial neural network algorithm; it is a self-organizing and self-learning network visual method that can express multi-dimensional spatial data in low-dimensional points through non-linear mapping (Kohonen., 2001). A SOM is an all-purpose classification tool that can connect samples with variables (Kohonen., 2013; Zhou et al., 2016). In recent years, SOMs have become increasingly popular in environmental research because of their capacity to address non-linear relations. Kalteh (2008) and Céréghino (2009) discussed the application of the SOM method in environmental science, particularly in water resource classification. Chon (2011) evaluated the application of SOM technology in the field of ecology. The high-dimensional, non-linear, and uncertain features of water quality monitoring data result in a certain complexity during the analysis and evaluation of surface water quality data. Therefore, data mining and the modern mode recognition method have been introduced to analyse and explain water quality monitoring data, which can, to a certain extent, offset the deficiency of the traditional method (Li et al., 2013). The Jinghe Oasis, which comprises an oasis and a desert, is a typical mountainous zone in an arid region and an important part of the northern slope of the Tianshan Mountains. Under the influences of the drainage basin climate and human activities, the pollution of the regional ecological environment, which results from agricultural and domestic wastewater sources around the Jinghe Oasis that are directly discharged or discharged through the river, has become an urgent problem related to sustainable socioeconomic development in Xinjiang. Therefore, a typical section of the Jinghe Oasis in the plain area of the arid region was selected as the research object of this study. The SOM method was applied to recognize the spatial distribution of water quality in the Jinghe Oasis. Based on the result, this study offers a tentative exploration of the relation between water quality and land use/cover changes in different clusters and provides new insights on controlling, managing, and protecting the ecological environment in the Jinghe Oasis.

The main objectives of this study were to (1) analyse the spatial framework of water quality using the self-organizing map (SOM) method based on non-hierarchical k-means classification; (2) explore the relationship between the water quality parameters and land use/cover types in different clusters; and (3) analyse the relationship between water quality parameters and land use/cover types in different stages.

1 Materials and methods

1.1 Study area

The Jinghe Oasis is in the centre of Eurasia in the northwest Xinjiang Uygur Autonomous Region at 44°02'~45°10'N and 81°46'~83°51'E. The Jinghe Oasis is composed of wetland and desert oasis vegetation and wildlife and is a national desert ecological reserve. The study area has a unique wetland ecological environment, and it has been listed as the Xinjiang Uygur autonomous region "wetland nature reserve". The region has 385 kinds of desert plants, approximately 64% of the vast amount of desert plants

in China. The Jinghe Oasis was once fed by 12 river branches belonging to three major river systems, the Bortala River, Jing River and Kuytun River, which were mainly rivers connected with Ebinur Lake. Due to natural environmental changes and human activities (i.e., modern oasis agricultural development), many rivers gradually lost their hydraulic connections with Ebinur Lake, and only Bortala River and Jing River currently supply water to the lake. The climate in the Jinghe Oasis is a typical continental arid climate, with an annual average temperature of 7.36 °C, an average precipitation of 100~200 mm, and an average evaporation of 1500~2000 mm (Zhang et al., 2015). In recent years, under the dual influence of natural and human factors, the water resources of the Jinghe Oasis have degraded seriously, causing an extreme decrease in the natural oasis and water area, desertification of the land, salinization of the farmland, serious grassland degradation, and water quality salinization (Jilil et al., 2002). At the same time, under the effect of the strong winds in Alashankou, the region has become a main source of dust; this affects the ecological environment of northern Xinjiang. This study, by using actual data mining, established the research area shown in Figure 1.

Fig. 1 Location of the study area

1.2 Data acquisition and processing

(1) We applied GF-1 remote sensing images obtained in May and October 2015 as the data sources (see <http://www.cresda.com/CN/>). These images were not influenced by clouds, fog, or snow cover, and their quality was good. We conducted radiation and orthographic corrections for the remote sensing image data combined with 1:50,000 scale digital elevation model (DEM) data. We established five land use/cover types by using the Environment for Visualizing Images software (ENVI Version 5.0), namely, farmland, forest-grassland, water body, salinized land, and others, based on the actual conditions of the research zone. Finally, we generated a vector data map of the land use/cover types for two stages of the research zones.

(2) The pillar industries in the Jinghe Oasis include salt production and *Artemia* breeding. No heavy industry is present; thus, point source pollution from industrial wastewater was not considered in the research zone. Research samples were collected from agricultural land in Jinghe County and Tuotuo Village, which surround Ebinur Lake, a national ecological zone in Ebinur Lake Bird Isle, and the Ganjia Lake *Haloxylon* natural conservation area. In total, we collected 47 water samples, with 23 collected in May and 24 collected in October 2015. The monitoring indices used included chemical oxygen demand (COD), five-day biological oxygen demand (BOD₅), suspended solids (SS), total phosphorus (TP), total nitrogen (TN), ammonia nitrogen (NH₃-N), chromaticity (SD), and turbidity (NUT). All polyethylene bottles were used to store the samples. The bottles were cleaned, dried, and sealed with deionized water before sampling. The samples were taken to the laboratory for the measurements and analyses after collection. We applied dichromate titration, dilution and inoculation, gravimetry, ammonium molybdate spectrophotometry, alkaline potassium persulfate decomposition UV spectrophotometry, and Nessler reagent spectrophotometry to measure the COD, BOD₅, SS, TP, TN, and NH₃-N, respectively. The analyses of all the samples were entrusted to and completed by Urumqi Jincheng Measurement Technology Co., Ltd.

1.3 Recognition of water quality spatial characteristics based on the SOM method with non-hierarchical *k*-means classification

At present, classifications based on the SOM neural network are mainly unsupervised and applied to

fault analyses, text clustering, and water quality evaluations. The method, which does not require a consistent data distribution, is simple and can address detailed information without the influence of minor local problems. The results are the distribution features of the input mode and topological structures (Li et al., 2010). A typical SOM network generally consists of input and output clusters. All the nerve cells in the input cluster and the weight vectors in the output cluster are connected and classified as typed data using the SOM via a learning process. Accordingly, the *k*-means algorithm is applied to keep each cluster compact and to separate the clusters from each other as much as possible. The Davies-Bouldin clustering index was used to determine the optimal number of clusters for the dataset (An et al., 2016; Park et al., 2014). The lower the Davies-Bouldin index value is, the better the clusters are differentiated. The K-means cluster analysis was combined with the Davies-Bouldin index (DBI) to select the clustering number. (Zhou et al., 2016).

The SOM method based on non-hierarchical *k*-means classification was applied to the spatial framework of the water quality in the research zone by implementing the following steps. (1) We input the water sample data for clustering from May and October 2015 to the SOM network. We applied the topological values for calculating the network size to select the quantity of nerve cells and determine the output results based on the minimum values of the Quantization Error (QE) and Topological Error (TE). The QE was used to determine the capacity of the established neural network in distinguishing the original input data, whereas the TE was used to measure the neural network quality, i.e., to evaluate whether the network was applicable for training (Kohonen, 2001). After determining the network size, we conducted network training and obtained a set of weight values. (2) The weight value obtained from the SOM clustering results was considered the initial cluster centre, and the *k*-means algorithm was initialized to execute this algorithm and, combined with the DBI index, select the clustering number. This clustering combination algorithm maintains the self-organizing features of the SOM network, inherits the high efficiency of the *k*-means algorithm, and offsets the poor clustering effects that result from the excessive convergence time of the SOM network and the inappropriate selection of the initial clustering centre for the *k*-means algorithm. The SOM requires a SOM toolbox and some basic functions in Matrix Laboratory (MATLAB) (Zhang, 2015). This study used MATLAB 2013a as the calculation platform.

1.4 Spatial analysis of the influences of land use/cover change on water quality

As an artificial system disturbance, land use/cover type is the second major boundary condition that directly or indirectly influences hydrologic processes and exerts a considerable effect on the drainage water environment. First, we obtained information on the land use/cover types within the 1 km buffer zone of the water quality sampling points in the research area using the spatial analysis function of ArcGIS 9.3. Based on the results, we then discussed and analysed the correlation between water quality and land use/cover type changes at different levels and periods. For different levels, we established the correlation between water quality and land use/cover types in each layer and discussed the influence of land use/cover type changes on water quality. For different periods, analysed the land use/cover type information and eight types of water quality indices during the dry and wet seasons. The land use/cover type information and eight water quality indices were imported into Canoco 4.5 (Ter Braak and Smilauer, 2002) to test the DCA gradient axis. The results showed that the DCA gradient shaft length was less than 3. Based on the results of Wang et al (2017), when the DCA gradient shaft length is less than 3, the redundancy analysis (RDA) method can explore the relationship; therefore, the redundancy analysis (RDA) method was applied to determine the influence trend of land use/cover changes on the water quality within the Ebinur Lake buffer area. This method indicates the contribution rate of a single land use/cover variable on water quality and

can directly demonstrate the correlation between land use/cover type and water quality parameters via a 2D ordination graph. The methodology is explained in the following section, and a conceptual flow chart describing the methodology is shown in Figure 2.

Fig. 2 Conceptual model of the methodology

2 Results and analysis

2.1 Spatial framework of water quality in the Jinghe Oasis

Regarding the network structure selection, neural networks with a more complicated structure generally have a better capability to address complicated non-linear problems but require a longer training time (Kohonen, 2013). Increasing the number of water quality indices can provide more abundant information; however, the correlations among indices will also increase. The topological values were selected to determine the grid size in this study, and the *k*-means clustering method was adopted to obtain the results. Overall, after the standard processing of the water quality data, the best network training effect was obtained from 35 (7×5) nerve cells, and the QE and TE values were 1.033 and 0.001, respectively.

When the average variance values are less than 5% for different clusters, the DBI is low; thus, the corresponding clustering number can be regarded as the best clustering result. Therefore, this study input the trained weights of the neuron nodes through the K-means cluster analysis combined with the DBI to select the clustering number. The results are shown in Figure 3a. **Six clusters were formed because this number yielded the minimum DBI value (Figure 3a).**

Fig. 3 (a) Davies–Bouldin index plot. (b) Results of the SOM clustering of the cells on the map plane (distribution of the sampling sites on the SOM according to the eight water quality parameters and clustering of the trained SOM).

Figure 3b presents the results of the SOM clustering of the cells on the map plane, which exhibited similarities among the different monitoring stations. In particular, Cluster 1 included the sampling points around the Ganjia Lake *Haloxylon* natural conservation area in the southern Ebinur Lake region and points east of Ebinur Lake and around the Kuitun River during the wet season. Cluster 2 included the monitoring stations in the Jing River and around the agricultural ditch in the western Ebinur Lake region. Cluster 3 comprised samples of water from melted ice in the southern–western corner of the research zone, which were grouped into only one type. Cluster 4 included the sampling points within the Ganjia Lake *Haloxylon* natural conservation area during the dry season. Cluster 5 included the sampling points in the Jing River, the agricultural ditch and around Ebinur Lake. Cluster 6 contained points located around the Kuitun River and Ebinur Lake Bird Isle, which had more pools. In general, although individual points may interfere with the explanation of the results, the classification results can identify the time sequence features in the research zone. Clusters 1 to 3 entirely included samples from the wet season (May 2015), whereas Clusters 4 to 6 contained the monitoring samples from the dry season (October 2015). To further observe the information on the water quality parameters of the Jinghe Oasis based on the response of different nerve cells, the water quality information from the various cluster groups was visualized. The results are shown in Figure 4.

Fig. 4 The patterning results for the water quality parameters on the SOM plane

Figure 4 shows the distribution of different water quality parameters on the SOM. High COD, TN, NH₃-N, and SD values were recorded in the right corner of the SOM network, thereby indicating a declining trend in the southern Ebinur Lake region and the surrounding Kuitun River. The values during the wet season were higher than those during the dry season. High values of SS and NUT were observed in the left corner of the SOM network, which falls within the scope of the agricultural ditch and the Jing River areas during the dry season. This region is mainly distributed around the agricultural land area and is significantly influenced by human activities. High TP values were observed within the scope of the lower left corner, which mainly focuses on the agricultural ditch and Jing River during the dry season. Crops were mature during the dry season, and the farmland area was increased, which thereby increased the TP value. In contrast, the distribution of the BOD₅ was different, indicating a declining trend from the centre to the surrounding area. High BOD₅ values were mainly observed downstream of the Ganjia Lake *Haloxylon* natural conservation zone, which is surrounded by a salt field, thereby exerting a certain influence on the surrounding water quality. The regional change in the water quality of the Jinghe Oasis was reflected clearly and directly through the SOM. To observe the distribution of the water quality parameters directly, we collected different values of the water quality parameters at various layers (Figure 5).

Fig. 5 Average values for the water quality parameters

Figure 5 shows that the distribution of water quality parameters varies in different clustering layers. Among the six clusters, water quality was generally relatively better in Cluster 3. However, Cluster 3 only had one sampling point, which comprised ice and snow water. Therefore, its water quality was not considered. In addition, the COD, SS, NUT, TN, and NH₃-N contents were high in Clusters 1 and 2, which indicates a relatively lower water quality compared with the water downstream of the Ganjia Lake *Haloxylon* natural conservation zone, the surrounding Ebinur Lake region, and the agricultural land. The SD values were high in Clusters 1, 4, and 6. Meanwhile, high BOD₅ values were mainly concentrated in Clusters 4 and 6, around the Ganjia Lake *Haloxylon* natural conservation zone and Ebinur Lake Bird Isle. Many pools in these clusters are, to a certain extent, influenced by the water area. The concentration difference in the TP of each layer was minimal and considerably influenced by human activities, particularly changes in the agricultural land area. Based on these results and Chinese Environmental Quality Standards for Surface Water (GB3838-2002), we evaluated the grades of the water quality parameters, including the COD, BOD₅, TN, NH₃-N, and TP, at different layers (Table 1).

Table 1 The classes of water quality parameters in each cluster

As shown in Table 1, according to Chinese Environmental Quality Standards for Surface Water (GB 3838-2002), clusters 1 to 6 did not satisfy the potable water quality level. Clusters 1 and 2 had an identical water quality classification level, and their COD and TN contents were higher than the standard values. In particular, the COD content exceeded level V. Meanwhile, the BOD₅ and COD contents were excessive in Clusters 4 and 5, respectively. In Cluster 6, both the COD and BOD₅ contents were excessive, but the COD content was higher.

2.2 Analysis of land use/cover type and its relation to water quality at different layers

In this study, historical data, high-resolution Google Earth images, and field survey data were selected to verify that more than 100 pixels of each land cover type were used for the training data, and the confusion matrix was used to verify the classification results. For the classification results obtained in May

and October 2015 (Figure 6), precision increased to 89.9750% and 86.2848%, respectively, and the kappa coefficients were 0.8681 and 0.8184, respectively, based on the confusion matrix (Table 2), which indicates an accurate classification result that satisfies the research requirements. Accordingly, ArcGIS was used to establish a 1 km buffer zone around the water sampling points. The composition of land use/cover at different layers was analysed according to the hierarchical results of the water quality parameters, and the results are presented in Figure 7.

Fig. 6 The change in land use/cover of the Ebinur Lake area during the rainy (May) and dry (October) seasons in 2015 (a: May; b: October)

Fig. 7 The area of land use/cover for each cluster

Table 2 The confusion matrix calculation with a maximum likelihood supervised classification

Figure 7 shows the land use/cover mode at different layers. In general, the salinization phenomenon was serious across the entire research zone. Among the six clusters, Cluster 1 mainly included the sampling points around the Ganjia Lake *Haloxylon* natural conservation area, the eastern Ebinur Lake region, and the Kuitun River. The major land type in this cluster was forest-grass land. The monitoring site in Cluster 2 was in the irrigation ditch of the Jinghe Oasis and in the Jing River. In May, crops do not grow abundantly in the research zone, and the major land type in this cluster was forest-grass land. Cluster 4 mainly included the Ganjia Lake *Haloxylon* natural conservation area and the sampling points in Tuotuo Village during the dry season, which include farmland and forest-grassland. The sampling points in Cluster 5 were in the agricultural ditch, the Jing River, and the surrounding Ebinur Lake. The land types mainly included forest-grassland and farmland, which was larger than the other land type. Cluster 6 was mainly located in the Kuitun River and the surrounding Ebinur Lake Bird Isle. A considerable number of plants, such as reed, grow in some of its pools. The percentages of the forest-grass land and salinized land within the 1 km buffer were large, based on the actual conditions in the research zone and the distribution of the sampling points. Based on these results, the correlation between land use/cover type and water quality parameters from Clusters 1 to 6 was analysed. The results are presented in Table 3.

Table 3 The correlation coefficients between the land use/cover and water quality parameters in each cluster

In Cluster 1, the forest-grass land exhibited a negative correlation with the SS and NUT under the significance level of 0.01, with coefficients reaching up to -0.710 and -0.724 , respectively. At a significance level of 0.05, the water body type exhibited an obvious positive correlation with the COD and a negative correlation with the TN, with coefficients of 0.986 and -0.721 , respectively. At a confidence level of 0.01, the salinized land demonstrated a positive correlation with the NUT, with a coefficient of 0.756 . In Cluster 2, the farmland presented a negative correlation with the COD and a positive correlation with the NUT at a confidence level of 0.05, and the coefficients were -0.581 and 0.639 , respectively. Under the same conditions, the forest-grass land exhibited a positive correlation with the COD, and the coefficient was 0.613 . At a confidence level of 0.01, the water body type exhibited an evident negative correlation with the SS and NUT, and the correlation coefficients were -0.983 and -0.990 , respectively. In Cluster 4, several water quality parameters were mainly influenced by the farmland, forest-grass land, and salinized land types. At a confidence level of 0.05, the farmland exhibited a negative correlation with the COD, with a coefficient of -0.652 . At a confidence level of 0.01, the farmland demonstrated an evident positive correlation with TP, and the coefficient is 0.872 . At a confidence level of 0.05, salinized land shows a clear negative correlation with TP and NUT, and the coefficients were -0.791 and -0.819 , respectively. In this layer, the others land type exhibited a positive correlation with the TP, with a

coefficient of 0.868. The sampling points in this layer were mainly located around Tuotuo Village, where the influences of human activities are considerable; therefore, in Cluster 4, the correlation percentage of the others land type with TP was high. In Cluster 5, farmland demonstrated an evident negative correlation with the BOD₅ at a 0.01 confidence level, with a correlation coefficient reaching up to -0.881. At a confidence level of 0.05, the salinized land showed a positive correlation with the BOD₅, with a correlation coefficient of 0.774. In Cluster 6, the forest-grass land showed a clear negative correlation at a confidence level of 0.01, and the correlation coefficient reached -0.884. At a confidence level of 0.05, the water area exhibited a positive correlation, with a correlation coefficient of 0.980.

From the results of the comprehensive analysis, the farmland, forest-grass land, and salinized land types have a considerable influence on the water quality parameters in the Jinghe Oasis. In Clusters 1, 2, and 6, the size of the water area, to a certain extent, also influenced changes in the water quality parameters. Given the unbalanced distribution of the sampling points at different layers, the effect of the land use/cover composition on the water quality in the research zone varied, so it can only indicate influences to a certain extent. Therefore, considering the actual conditions in Ebinur Lake, different land use/cover types and water quality influences must be understood as a whole, and the correlation between the land use/cover types and water quality in the Jinghe Oasis at different periods is further discussed.

2.3 Analysis of land use/cover changes in the Jinghe Oasis and their correlation with water quality at different seasons

The constituents of the land use/cover types at different seasons exerted diverse influences on the water quality. Therefore, an analysis of the influences of land use/cover type changes on water quality was conducted. The results are presented in Figure 8.

Fig. 8 RDA analyses of comprehensive land use/cover and water quality (a: wet season; b: dry season)

As shown in Figure 8, the farmland exhibited a negative correlation with the COD at a confidence level of 0.01 during the wet season, with a correlation coefficient of -0.543. In contrast, it showed a positive correlation with the NUT, with a correlation coefficient of 0.555. At a confidence level of 0.05, the farmland demonstrated a negative correlation with the NH₃-N, and the correlation coefficient was -0.461. At a confidence level of 0.05, the forest-grass land showed a positive correlation with the BOD₅ and TP, with correlation coefficients of 0.470 and 0.518, respectively. They exhibited a negative correlation with the SS and NUT, with correlation coefficients of -0.529 and -0.498, respectively. At a confidence level of 0.05, the salinized land demonstrated a positive correlation with the BOD₅ and TP, with correlation coefficients of -0.503 and 0.518, respectively. In contrast, it presented a negative correlation with the SS and NUT, with correlation coefficients of 0.449 and 0.449, respectively. During the dry season, the influence of the farmland on various water quality parameters evidently increased because of crop growth. At a confidence level of 0.01, the farmland presented a clear negative correlation with the COD and an evident positive correlation with the TP, with correlation coefficients of -0.620 and 0.616, respectively. At a confidence level of 0.05, the farmland showed a positive correlation with the TN and a negative correlation with the BOD₅, NH₃-N and SD, with correlation coefficients of 0.543, -0.495, -0.522, and -0.526 for TN, BOD₅, NH₃-N and SD, respectively. At a confidence level of 0.01, the salinized land demonstrated a negative correlation with the NUT and TP, and the correlation coefficients were -0.543 and -0.603, respectively. In contrast, it presented a positive correlation with the BOD₅ at a confidence level of 0.05 and a correlation coefficient of 0.522. Similarly, during the wet and dry seasons, the correlation between the water body and others land types with the water quality parameters was small. Based on these

results, the influences of the various land use/cover types in the research zone on the water quality parameters exhibited the following order, in descending order of influence: farmland → forest–grass land → salinized land → water body → others. Moreover, the influence was lower during the wet season than during the dry season.

3 Discussion

Given the seasonal differences, the unbalanced distribution of precipitation resulted in an apparent variation in the surface runoff and further imbalanced the spatial distribution of the water quality in the research zone (Fan et al., 2012; Prathumratana et al., 2008; Li et al., 2015). During the wet season (May) in the Jinghe Oasis, melted water from mountain ice and snow promote the flow in the Jing River, thereby resulting in a significant increase in surface runoff leading to an improved water quality in the rainy season compared to the dry season. During the dry season, the aquatic plants in rivers and lakes grow as the temperature rises, which can, to a certain degree, absorb and purify part of the water quality parameters. Therefore, significant and seasonal changes in the surface runoff at the research zone are important factors resulting in noticeable differences in the spatial distribution of water quality characteristics during the wet and dry seasons. Another major factor that causes differences in the spatial distribution of water quality is the change in land use/cover, especially farmland. During the dry season, farmland areas have a greater influence on more water quality variables than they do during the rainy season because of their intensive fertilization and agricultural runoff from soil erosion (Ngoye et al., 2004; Li et al., 2009; Tran et al., 2010). Multiple factors threaten the ecological safety of the Jinghe Oasis system. Especially in recent years, the lakeside desertification zone has rapidly expanded because of the decrease in the Ebinur Lake area and the degradation of lakeside vegetation under influence of the strong winds in Alashankou. In the current overall situation, the human activities influencing land use/cover changes are directly related to the development of the vulnerable ecological area that surrounds Ebinur Lake.

Recent statistics indicate that the annual growth rate of the population in the Jinghe Oasis is approximately 2.49%, which is slightly higher than previous growth rates (Li, 2006). Under the stress of a large population, the amount of inappropriate activities that negatively impact land use/cover in the Jinghe Oasis will increase. For the last 30 years, cotton has been the major crop in the Jinghe Oasis. The results of the current study indicate that the sampling points surrounding the farmland in the research zone have lower water quality values than the other studied regions. The primary livelihoods of the urban residents around the Ebinur Lake area are agricultural and animal husbandry industries. Pollutants that result in high TP and $\text{NH}_3\text{-N}$ contents in water include the excessive application of chemical fertilizers on farmlands, the production of livestock manure in rural villages, randomly stocked garbage, and domestic wastewater. The improper application of chemical fertilizers and pesticides to a vast area leads to high water nitrogen and phosphorus contents, resulting in the spread of algae in river sections. Consequently, the amount of dissolved oxygen in the river may decrease, the water quality may deteriorate, and eutrophication may occur. Furthermore, this scenario poses a serious salinization problem. Certain measures have been implemented for the ecological protection of Ebinur Lake, such as returning farmland to forest, cultivating ecological forests, and promoting efficient irrigation and water-saving technologies. However, these measures promote the gradual expansion of the lake area and result in different degrees of negative consequences. The most apparent result has been the rise of the underground water level, which has aggravated land salinization in the lowland areas and resulted in vast expanses of uncultivated former agricultural lands. Statistics indicate that soil salinization in the Ebinur Lake area mainly occurs in Bortala River, Jing River, the villages and towns surrounding Ebinur Lake, areas downstream of the Daheyanzi

River, and areas north of Bole City (Mi et al.,2010). Severe soil salinization has seriously affected the farming of crops; therefore, some farmers have increased the amount of chemical fertilizers they apply to increase yield, which also increases the pollution of the water and soil. Others have even abandoned their land, thereby causing land use/cover change.

Most rivers in Xinjiang are characterized by a low water yield, short flow, small water environmental capacity, poor self-cleaning capability, and low tolerance to pollution. Hence, an artificial change in the land use and exploration of resources in lake regions lead to an evident correlation between land use/cover types and water quality. In addition, the scientific utilization and protection of the water resources of Ebinur Lake and the scientific application of chemical fertilizers and improvement of their application rates are important actions and should be addressed to achieve sustainable development in the agricultural irrigation zones of the Jinghe Oasis and rivers of Xinjiang.

4 Conclusions

The spatial distribution characteristics of water quality in the Jinghe Oasis and their correlation with land use/cover types were analysed, and the following conclusions were drawn.

(1) Using the SOM method based on non-hierarchical *k*-means classification, 47 water quality sampling points were divided into 6 types, and the time sequence characteristics of the research zone were better recognized in the classification results. Clusters 1 to 3 comprised samples from the wet season (May 2015), whereas Clusters 4 to 6 comprised monitoring samples from the dry season (October 2015). In general, the COD, SS, NUT, TN, and NH₃-N contents were high. The SD value was high in Clusters 1, 4, and 6. In addition, high BOD and TP values were mainly concentrated in Clusters 4 and 6. Based on these findings, the water quality at different layers of the research zone was further evaluated. The results show that Clusters 1 to 6 do not satisfy potable water quality standards.

(2) The correlations between the land use/cover types and water quality parameters for Clusters 1 to 6 were analysed according to the hierarchical results of the water quality parameters. The comprehensive analysis indicates that the farmland, forest-grass land, and salinized land exerted significant influences on the water quality parameters of the Jinghe Oasis. In Clusters 1, 2, and 6, the size of the water area, to a certain extent, also influenced changes in the water quality parameters.

(3) During the wet and dry seasons, the influences that various land use/cover types in the research zone had on the water quality parameters exhibited the following order, in descending order of influence: farmland → forest-grass land → salinized land → water body → others. Moreover, the influences were lower during the wet season than during the dry season.

In general, the land use/cover type, area percentage, and water quality in the Jinghe Oasis demonstrated apparent correlations. The results of this study can tentatively explain the relationship between water quality and land use/cover types in different clusters by the SOM method. This work provides new insight for further studies on the correlation between land use/cover and water quality in the Jinghe Oasis, as well as a scientific reference for formulating regulations and control policies for the spatial development and water environmental protection of the Jinghe Oasis.

Acknowledgements: Thanks to the National Meteorological Information Center data provided meteorological data. The research was carried out with the financial support provided by the Natural Science Foundation of Xinjiang Uygur Autonomous Region, China (2016D01C029); National Natural

Science Foundation of China (41361045), the Scientific and technological talent training program of Xinjiang Uygur Autonomous Region (grant No. QN2016JQ0041).The authors wish to thank the referees for providing helpful suggestions to improve this manuscript.

References

- An, Y., Zou, Z., & Li, R. 2016. Descriptive characteristics of surface water quality in hong kong by a self-organising map. *International Journal of Environmental Research & Public Health*, 13(1):115.
- Bu H M, Meng W, Zhang Y, Wan J.2014. Relationships between land use patterns and water quality in the Taizi River basin, China. *Ecological Indicators*, 41: 187-197.
- Bierman P, Lewis M, Ostendorf B, Tanner J.2011.A review of methods for analysing spatial and temporal patterns in coastal water quality. *Ecological Indicators*, 11(1): 103-114.
- Cérégino R, Park Y S.2009. Review of the self-organizing map (SOM) approach in water resources: commentary. *Environmental Modelling & Software*, 24(8): 945-947.
- Chon T S. 2011. Self-organizing maps applied to ecological sciences. *Ecological Informatics*, 6(1): 50-61.
- De Jonge M, Van de Vijver B, Blust R, Bervoets L.2008.Responses of aquatic organisms to metal pollution in a lowland river in Flanders: a comparison of diatoms and macroinvertebrates. *Science of the Total Environment*, 407(1): 615-629.
- Fan X Y, Cui B S, Zhang K J, Zhang Z M.2012.Water quality management based on division of dry and wet seasons in Pearl River Delta, China. *Clean-Soil,Air,Water*, 40(4):381-393.
- Huang J L, Li Q S, Huang L,Zhang Z F,Mu J L, Huang Y L.2013a. Watershed-scale evaluation for land-based nonpoint source nutrients management in the Bohai Sea Bay, China. *Ocean & coastal management*, 71: 314-325.
- Huang J L, Li Q S, Pontius Jr R G, Hong H S. 2013b. Detecting the dynamic linkage between landscape characteristics and water quality in a subtropical coastal watershed, Southeast China. *Environmental management*, 51(1): 32-44.
- Hur J, Nguyen H V M, Lee B M.2014. Influence of upstream land use on dissolved organic matter and trihalomethane formation potential in watersheds for two different seasons. *Environmental Science and Pollution Research*, 21(12): 7489-7500.
- Jilil A, Mu G J.2002.Analysis on the dust storms and their disasters in the lakebed region of Ebinur Lake,Xinjiang.*Arid Land Geography*, 25(2):122-126.
- Kohonen T. 2001.*Self-Organizing Maps*. Berlin: Springer.
- Kohonen T.2013. Essentials of the self-organizing map. *Neural Networks*, 37: 52-65.
- Kalteh A M, Hjorth P, Berndtsson R.2008.Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling & Software*, 23(7): 835-845.
- Lee S W, Hwang S J, Lee S B,Sung H C. 2009. Landscape ecological approach to the relationships of land use patterns in watersheds to water quality characteristics. *Landscape and Urban Planning*, 92(2): 80-89.
- Li Y Y, Jiao J X, Wang Y,Yang W, Meng C, Li B Z, li Y, Wu J S. 2015. Characteristics of nitrogen loading and its influencing factors in several typical agricultural watersheds of subtropical China. *Environmental Science and Pollution Research*, 22(3): 1831-1840.
- Li W, Yao X Y, Liang Z W, Wu Y M, Shi J Y, Chen Y X. 2013. Assessment of surface water quality using self-organizing map and Hasse diagram technique. *Acta Scientiae Circumstantiae*, 33 (3) :893-903.
- Li J, Hao Z G.2010.Tourism planning based on Google Earth virtual earth platform.*Remote Sensing for Land & Resources*, (1):130-133.
- Li K,Wang L,Li Z H,Wang X R,Chen H B,Wu Z,Zhu P. 2015.Spatial variability characteristics of water quality and its driving forces in honghu lake during high water-level period. *Environmental Science*, 36(4):1285-1292.
- Li Y H. 2006.Water resources carrying capacity in the Ebinur Lake basin of Xinjiang.Shanghai:East China Normal University.
- Li S, Gu S, Tan X, Zhang Q.2009. Water quality in the upper Han River basin, China: the impacts of land use/land cover in

- riparian buffer zone. *Journal of Hazardous Materials*, 165, 317–324.
- Mi Y, Chang S L, Shi Q D, Gao X, Huang C. 2010. Study on the effect of agricultural non-point source pollution to water environment of the Ebinur Lake basin during high flow period. *Arid Zone Research*, 27(2):278-283.
- Ngoye E, Machiwa J F. 2004. The influence of land use patterns in the Ruvu river watershed on water quality in the river system. *Phys. Chem. Earth*, 29, 1161-1166.
- National Research Council. 2008. *Integrating Multiscale Observations of U.S. Waters*; The National Academies Press: Washington, DC, USA.
- Park Y S, Kwon Y S, Hwang S J, Park S. 2014. Characterizing effects of landscape and morphometric factors on water quality of reservoirs using a self-organizing map. *Environmental Modelling & Software*, 55: 214-221.
- Prathumratana L, Sthiannopkao S, Kim K W. 2008. The relationship of climatic and hydrological parameters to surface water quality in the lower Mekong River. *Environment International*, 34(6): 860-866.
- Swaney D P, Humborg C, Emeis K, Kannen A, Silvert W, Pastres R, Solidoro C, Yamamuro M, Hénocque Y, Nicholls R. 2012. Five critical questions of scale for the coastal zone. *Estuarine, Coastal and Shelf Science*, 96: 9-21.
- Shen Z, Hou X, Li W, Aini G, Chen L, Gong Y. 2015. Impact of landscape pattern at multiple spatial scales on water quality: A case study in a typical urbanised watershed in China. *Ecological Indicators*, 48: 417-427.
- Sun F, Sun W, Chen J, Gong P. 2012. Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery. *International journal of remote sensing* 33(21), 6854-6875.
- Tran C P, Bode R W, Smith A J, Kleppel G S. 2010. Land-use proximity as a basis for assessing stream water quality in New York State (USA). *Ecological Indicators*, 10, 727–733.
- Ter Braak C J F, Smilauer P. 2002. *CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5)*. In: Section on Permutation Methods. Microcomputer Power, Ithaca, NY.
- Uuemaa E, Roosaare J, Mander Ü. 2005. Scale dependence of landscape metrics and their indicatory value for nutrient and organic matter losses from catchments. *Ecological Indicators*, 5(4): 350-369.
- Wan R, Cai S, Li H, Yang G, Li Z, Nie X. 2014. Inferring land use and land cover impact on stream water quality using a Bayesian hierarchical modeling approach in the Xitaoxi River Watershed, China. *Journal of Environmental Management*, 133: 1-11.
- Wang XP, Zhang F, Li XH, Cao C, Guo M. 2017. Correlation analysis between the spatial characteristics of land use/cover-landscape pattern and surface-water quality in the Ebinur Lake area. *Acta Ecologica sinica* 2017,37(22):7438~7452.(In Chinese)
- Xiao H, Ji W. 2007. Relating landscape characteristics to non-point source pollution in mine waste-located watersheds using geospatial techniques. *Journal of Environmental Management*, 82(1): 111-119.
- Yang Y N, Wang J L, Cheng G J, Xi X H, Wang C. 2016. Relationship between land use pattern and water quality change in Fuxian Lake basin. *Remote Sensing for Land and Resources*, 28(1):159-165.
- Zhang, Y., Liang, X., Wang, Z., & Xu, L. 2015. A novel approach combining self-organizing map and parallel factor analysis for monitoring water quality of watersheds under non-point source pollution. *Scientific Reports*, 5(3), 16079.
- Zhou P, Huang J L, Pontius Jr R G, Hong H S. 2016. New insight into the correlations between land use and water quality in a coastal watershed of China: Does point source pollution weaken it? *Science of the Total Environment*, 543: 591-600.
- Zhang F, Tiyyip T, Johnson V C, Ding J L, Zhou M, Chan N W. 2015. The influence of natural and human factors in the shrinking of the Ebinur Lake, Xinjiang, China, during the 1972–2013 period. *Environmental Monitoring and Assessment*, 187(1): 1-14.

Figure captions

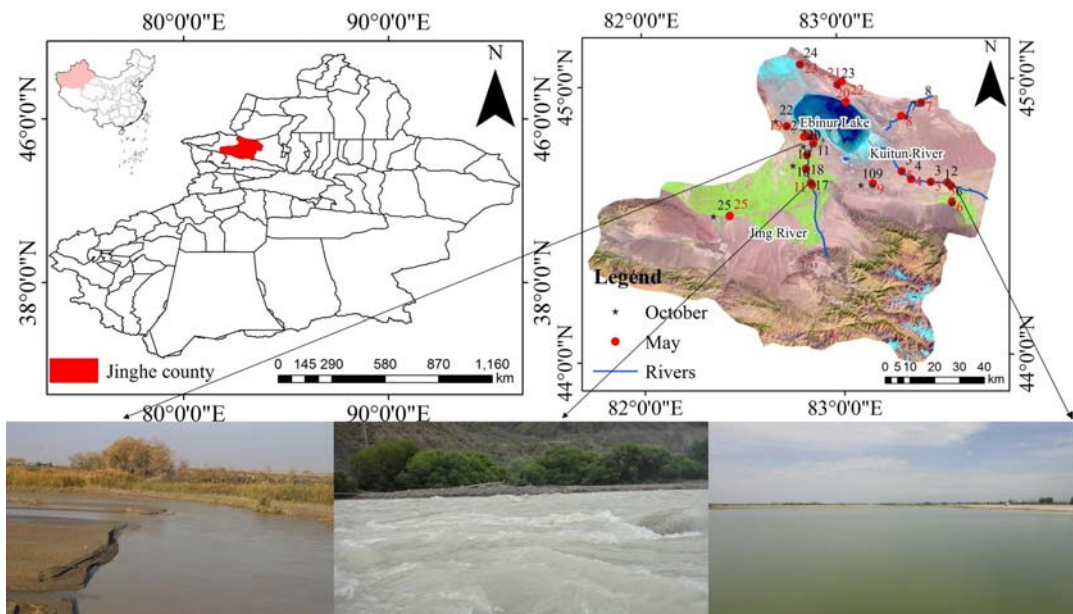


Fig. 1 Location of the study area

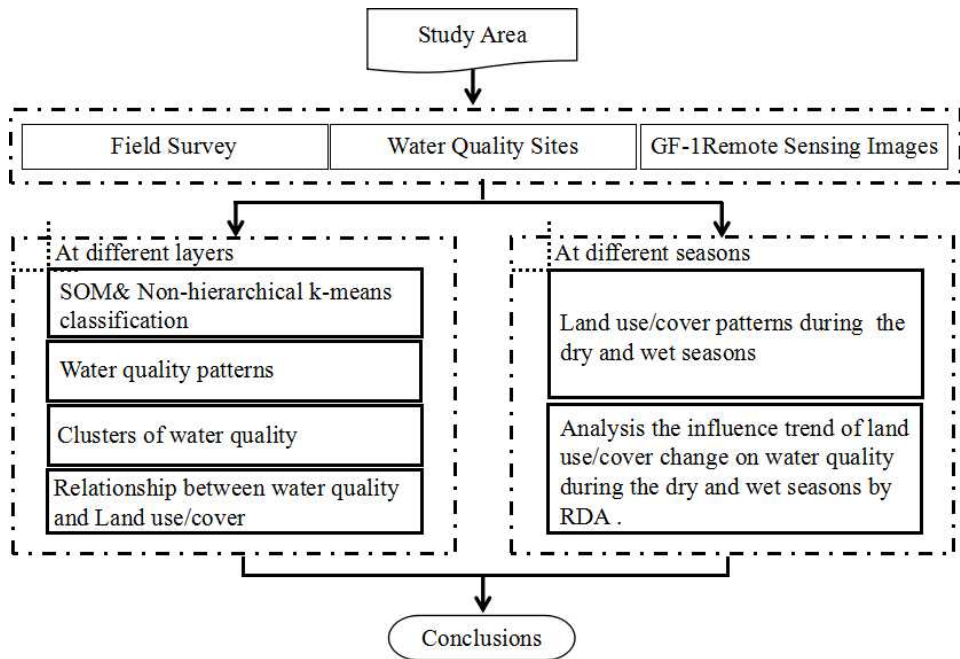


Fig. 2 Conceptual model of the methodology

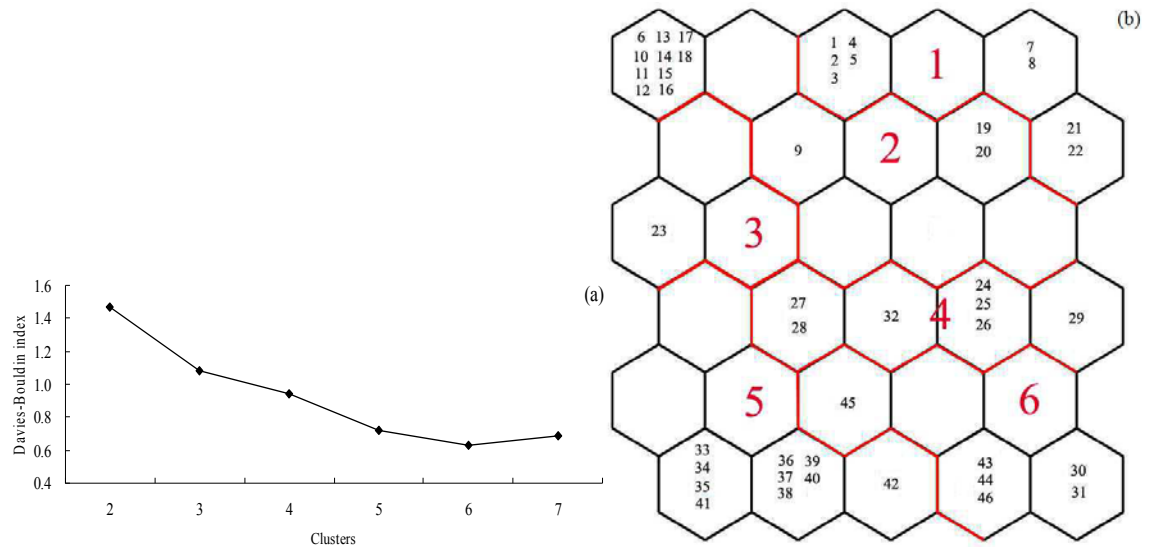


Fig. 3 (a) Davies–Bouldin index plot. (b) Results of the SOM clustering of the cells on the map plane (distribution of the sampling sites on the SOM according to the eight water quality parameters and clustering of the trained SOM).

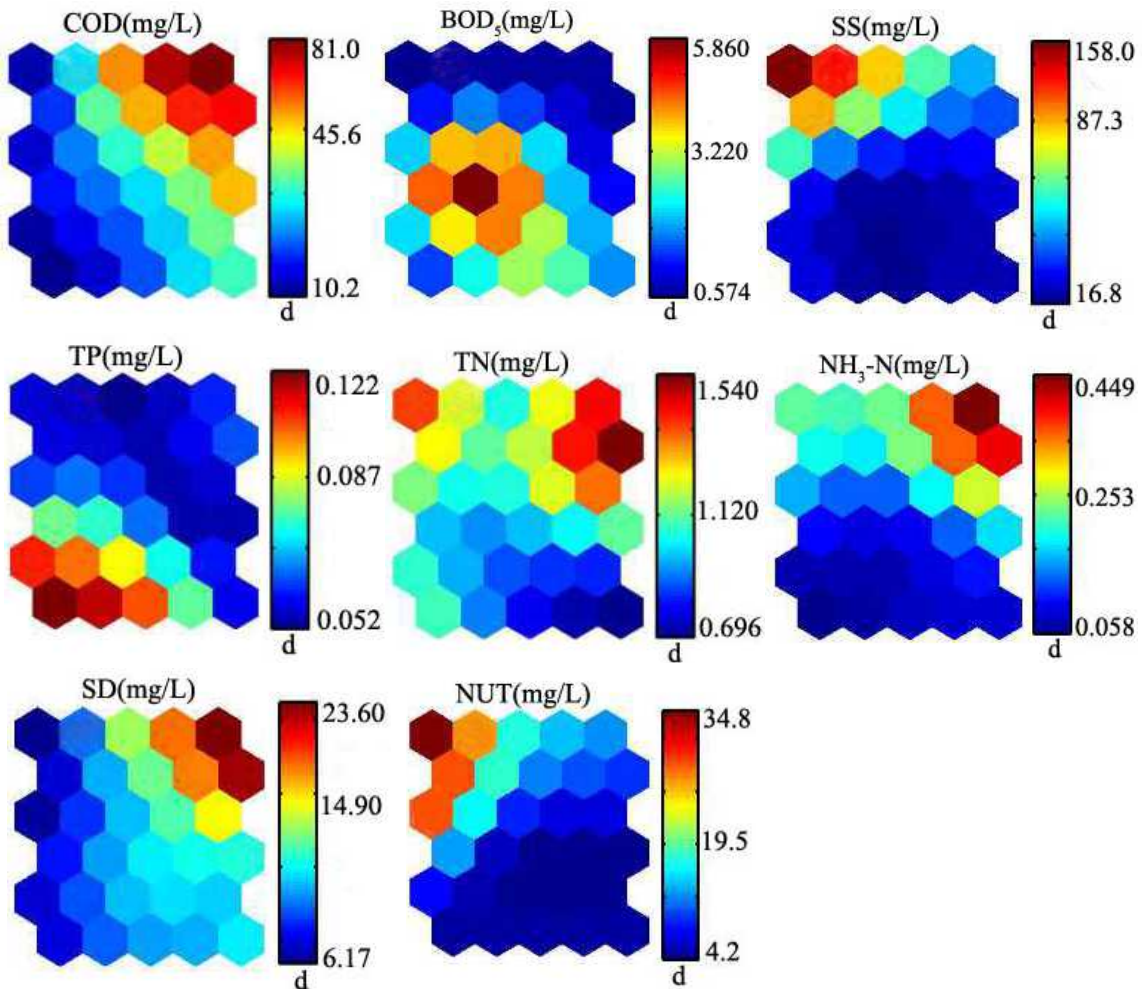


Fig. 4 The patterning results for the water quality parameters on the SOM plane

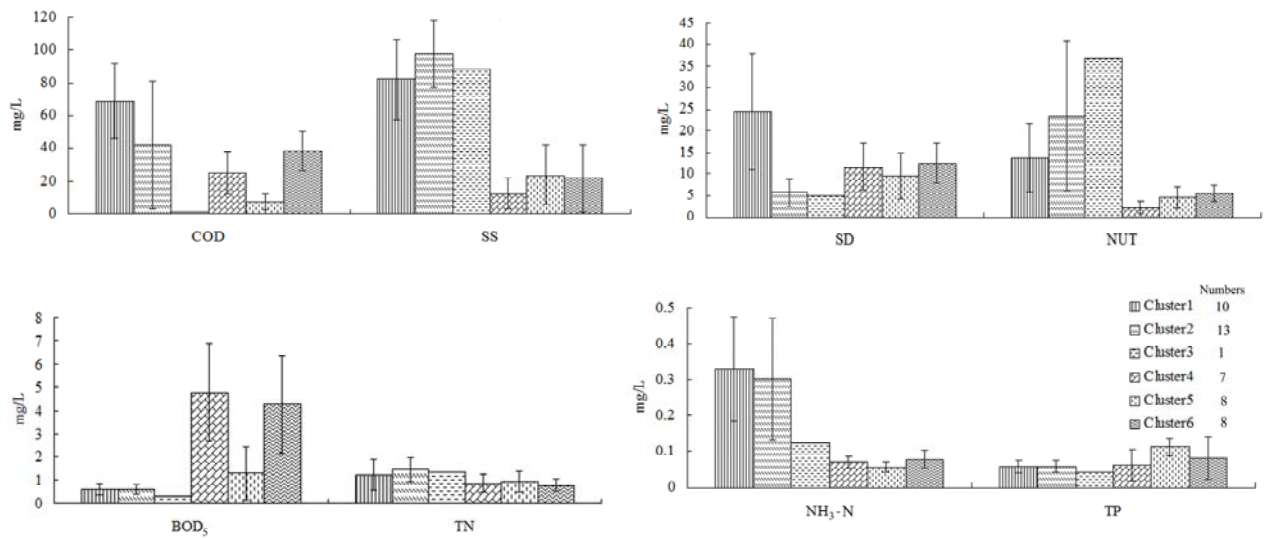


Fig. 5 Average values for the water quality parameters

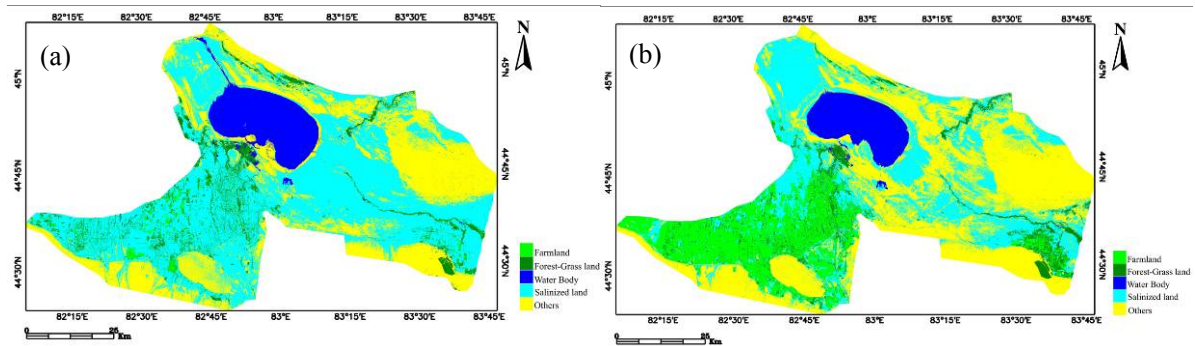


Fig. 6 The change in land use/cover of the Ebinur Lake area during the rainy (May) and dry (October) seasons in 2015 (a: May; b: October)

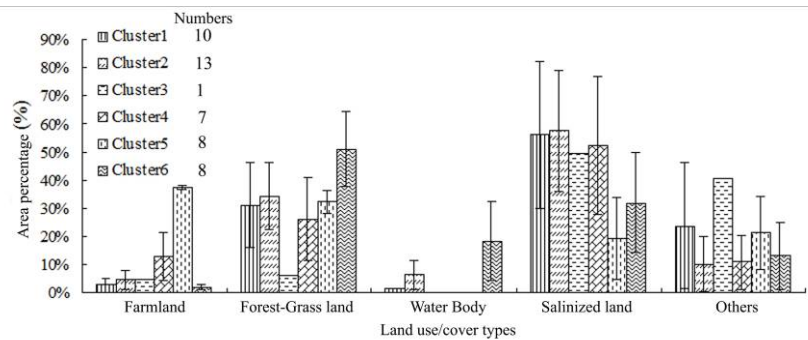


Fig. 7 The area of land use/cover for each cluster

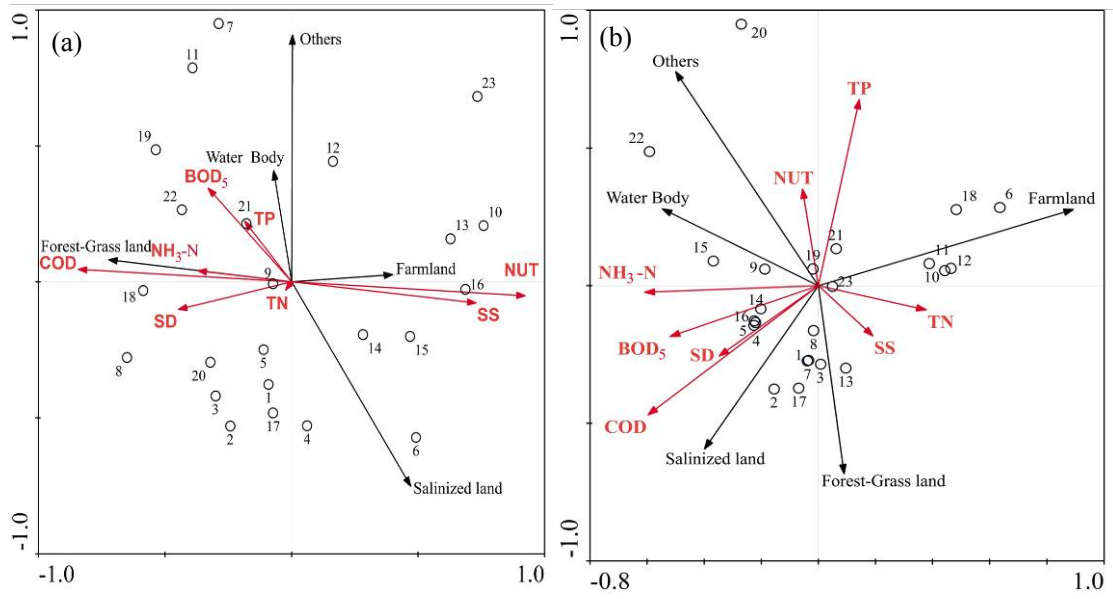


Fig. 8 RDA analyses of comprehensive land use/cover and water quality (a: wet season; b: dry season)

Tables caption

Table 1 The classes of water quality parameter in each cluster

	Cluster1	Cluster2	Cluster4	Cluster5	Cluster6
COD	Exceed V	Exceed V	I	IV	V
BOD ₅	I	I	IV	I	IV
TN	IV	IV	III	III	III
NH ₃ -N	II	II	I	I	I
TP	II	II	II	III	II

Table 2 The calculation of a confusion matrix by a maximum likelihood supervised classification

	LULC	Water body	Saline land	Farmland	Forest grassland	Other land	Total	User's accuracy (%)
May	Water body	144	0	0	0	0	144	100
	Saline land	0	77	0	0	16	93	82.79
	Farmland	0	36	101	0	0	137	73.72
	Forest-Grass land	0	36	0	101	0	137	73.72
	Other land types	1	0	0	0	87	88	98.96
	Total	145	149	101	101	103	Overall=89.9750%	
	Producer's accuracy (%)	99.31	51.67	100	100	84.46	Kappa=0.8681	
October	Water body	144	0	0	0	0	144	100
	Saline land	0	57	0	0	26	83	86.67
	Farmland	0	16	101	0	0	117	86.32
	Forest-Grass land	4	16	0	101	0	117	86.32
	Other land types	0	0	0	0	77	77	100
	Total	148	89	101	101	103	Overall=86.2848%	
	Producer's accuracy (%)	97.29	64	100	100	74.75	Kappa=0.8184	

Table 3 The correlation coefficients between land use/cover and water quality parameters in each cluster

	Parameters	Farmland	Forest-Grass land	Water Body	Salinized land	Others
Cluster1	COD	-0.161	0.240	0.986*	-0.110	-0.361
	BOD ₅	0.074	0.492	-0.439	-0.613	0.552
	SS	-0.271	-0.710**	0.801	0.619	-0.384
	TP	-0.195	0.453	0.371	0.444	0.623
	TN	0.464	0.524	-0.721*	-0.224	0.121
	NH ₃ -N	-0.491	0.039	0.071	-0.066	0.291
	SD	-0.296	0.448	0.415	-0.426	0.396
	NUT	-0.261	-0.724**	0.550	0.756**	-0.612
Cluster2	COD	-0.581*	0.613*	0.916	-0.693**	0.442
	BOD ₅	-0.004	0.455	0.055	-0.545	0.242
	SS	0.493	-0.512	-0.983**	0.386	0.047
	TP	-0.222	0.531	0.850	-0.129	0.382

	TN	0.351	0.415	-0.867	-0.356	-0.311
	NH ₃ -N	-0.467	0.121	0.122	-0.269	0.284
	SD	-0.226	-0.073	-0.051	-0.217	0.473
	NUT	0.639*	-0.446	-0.990**	0.513	-0.236
Cluster4	COD	-0.652*	0.484	/	0.375	-0.048
	BOD ₅	-0.482	-0.402	/	0.505	0.688
	SS	-0.155	0.658	/	-0.179	-0.167
	TP	0.872**	-0.398	/	-0.791*	0.868*
	TN	0.336	0.468	/	-0.571	-0.124
	NH ₃ -N	-0.202	-0.540	/	0.398	0.352
	SD	-0.543	0.825*	/	0.214	-0.549
	NUT	0.578	0.469	/	-0.819*	-0.129
Cluster5	COD	0.094	-0.372	/	0.325	-0.400
	BOD ₅	-0.881**	0.503	/	0.774*	-0.044
	SS	0.621	-0.533	/	-0.284	-0.380
	TP	0.587	-0.702	/	-0.565	0.136
	TN	0.735	-0.588	/	-0.184	-0.604
	NH ₃ -N	-0.675	0.108	/	0.487	0.308
	SD	-0.632	0.330	/	0.208	0.576
	NUT	0.311	0.076	/	-0.459	0.154
Cluster6	COD	0.489	0.401	0.980*	-0.454	0.289
	BOD ₅	-0.256	-0.884**	-0.660	0.367	0.327
	SS	-0.481	0.194	0.341	-0.150	-0.062
	TP	-0.545	-0.656	0.269	-0.060	0.206
	TN	-0.158	-0.364	-0.022	0.516	0.313
	NH ₃ -N	-0.553	-0.366	0.517	-0.090	0.603
	SD	0.811	-0.037	-0.857	0.249	-0.282
	NUT	0.450	0.165	0.764	-0.497	0.636

*p<0.05(2-tailed) **p<0.01(2-tailed)