Chin. Geogra. Sci. 2015 Vol. 25 No. 3 pp. 321–336
doi: 10.1007/s11769-014-0718-x

Springer Science Press
www.springerlink.com/content/1002-0063

Mapping of Regional Soil Salinities in Xinjiang and Strategies for Amelioration and Management

WANG Fei [1, 2], CHEN Xi[1], LUO Geping[1], HAN Qifei[1, 2]

(1. *State Key Laboratory of Desert and Oasis Ecology*, *Xinjiang Institute of Ecology and Geography*, *Chinese Academy of Sciences*, *Ürümqi* 830011, *China*; 2. *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*)

Abstract: Information on the spatial distribution of soil salinity can be used as guidance in avoiding the continued degradation of land and water resources by better informing policy makers. However, most regional soil-salinity maps are produced through a conventional direct-linking method derived from historic observations. Such maps lack spatial details and are limited in describing the evolution of soil salinization in particular instances. To overcome these limitations, we employed a method that included an integrative hierarchical-sampling strategy (IHSS) and the Soil Land Inference Model (SoLIM) to map soil salinity over a regional area. A fuzzy c-means (FCM) classifier is performed to generate three measures, comprising representative grade, representative area, and representative level (membership). IHSS employs these three measures to ascertain how many representative samples are appropriate. Through this synergetic assessment, representative samples are obtained and their soil-salinity values are measured. These samples are input to SoLIM, which is constructed based on fuzzy logic, to calculate the soil-forming environmental similarities between representative samples and other locations. Finally, a detailed soil-salinity map is produced through an averaging function that is linearly weighted, which is used to integrate the soil salinity value and soil similarity. This case study, in the Uyghur Autonomous Region of Xinjiang of China, demonstrates that the employed method can produce soil salinity map at a higher level of spatial detail and accuracy. Twenty-three representative points are determined. The results show that 1) the prediction is appropriate in Kuqa Oasis (R^2 = 0.70, RPD = 1.55, RMSE = 12.86) and Keriya Oasis (R^2 = 0.75, RPD = 1.66, RMSE = 10.92), that in Fubei Oasis (R^2 = 0.77, RPD = 2.01, RMSE = 6.32) perform little better than in those two oases, according to the evaluation criterion. 2) Based on all validation samples from three oases, accuracy estimation show the employed method (R^2 = 0.74, RPD = 1.67, RMSE = 11.18) performed better than the multiple linear regression model (R^2 = 0.60, RPD = 1.47, RMSE = 14.45). 3) The statistical result show that approximately half (48.07%) of the study area has changed to salt-affected soil, mainly distributed in downstream of oases, around lakes, on both sides of rivers and more serious in the southern than the northern Xinjiang. To deal with this issue, a couple of strategies involving soil-salinity monitoring, water management, and plant diversification are proposed, to reduce soil salinization. Finally, this study concludes that the employed method can serve as an alternative model for soil-salinity mapping on a large scale.

Keywords: salinity; soil mapping; purposive sampling; Soil Land Inference Model (SoLIM); fuzzy c-means (FCM); multiple linear regression (MLR) model

Citation: Wang Fei, Chen Xi, Luo Geping, Han Qifei, 2015. Mapping of regional soil salinities in Xinjiang and strategies for amelioration and management. *Chinese Geographical Science*, 25(3): 321–336. doi: 10.1007/s11769-014-0718-x

1 Introduction

Salt-affected soils are characterized by the presence of excess levels of soluble salts (saline soils) in the soil solution or at the cation exchange sites. The accumulation of salts in these soils originates in the weathering of parent minerals (causing fossil or primary salinity-sodicity) or from anthropogenic activities involving the

Received date: 2013-12-16; accepted date: 2014-04-14
Foundation item: Under the auspices of Key Program of National Science Foundation of China (No. 41361140361)
Corresponding author: CHEN Xi. E-mail: Chenxi@ms.xjb.ac.cn

inappropriate management of land and water resources (contributing to man-made, or secondary, salinity-sodicity). Excessive salinity negatively affects crop productivity and eventually results in land degradation. Approximately 9.55×10^8 ha of soil has been affected worldwide by primary salinization, whereas secondary salinization affects approximately 7.70×10^7 ha in irrigated areas (Ghassemi *et al*., 1995). In addition, irrigated areas decrease approximately 1% to 2% per year because of salt-affected land surface (FAO, 2002) particularly in arid lands. Uyghur Autonomous Region of Xinjiang, in the furthermost area of northwest China, is the largest arid region in the country, with a total area of approximately 1.66×10^6 km^2, in which the saline and alkaline soil areas represent 8.48×10^4 km^2 (Soil Survey Staff of Xinjiang, 1996) and 31.1% of the salt-affected soils in arable farmland. Soil salinization in this region is a widespread natural phenomenon that constrains sustainable agriculture and social-economic development. In the future, increasing population pressure will require that more drylands be used for agricultural production. Most of this expansion of agriculture into drylands will be achieved by expanding irrigation. This can exacerbate the problem of soil salinization (Metternicht and Zinck, 2003). Regional digital mapping of salt-affected soils can help avoid the continued degradation of land and water resources by providing policy makers with the information they need to make sound policy decisions. Soil mapping is a process in which the spatial distribution of physical, chemical, and descriptive soil properties is evaluated and presented in a form that can be understood and interpreted by various users (Sheng *et al*., 2010).

In recent decades, the soil-science community has expended great effort to develop regional soil databases. Most of the maps in these databases were produced by digitizing older paper maps or by interpolation without sampling current soil resources (Grunwald *et al*., 2011). Moreover, soil bodies are typically represented as discrete, homogeneous entities that lack quantified measures of uncertainty (Goovaerts and Journel, 1995). Despite the shortcomings of conventional soil data, these data are difficult to replace by any mechanical model. Developing and transitional countries typically lack soil information (Mulder *et al*., 2011) such as the extent and characterization of salt-affected soils, on a large scale, in China. No soil-salinity maps quantitatively describe the soil salinization from the northern to the southern Xinjiang after the literature review in the recent 10 years However, in the anthropogenic regions, soil change, and soil formation and degradation, have also accelerated, jeopardizing soil quality and health. The need for up-to-date soil and environmental data that characterize the physicochemical, biological, and hydrological conditions of ecosystems across continents has intensified (Grunwald *et al*., 2011).

Digital soil mapping (DSM) and modeling techniques have proliferated during the past few decades to address these soil-data needs (McBratney *et al*., 2003; Grunwald *et al*., 2011). Two approaches are used in DSM research and practice. One approach aims at truly automatic, objective, and quantitative mapping, taking advantage of techniques in statistics, geostatistics, machine learning, and data mining, and relying heavily on dense sampling from the field or from existing soil maps. The other approach tries to fit within the conventional soil survey and mapping framework, including the conventional process and standard. This approach aims to effectively use the soil scientist′s knowledge, while reducing the inconsistency and cost associated with the traditional manual process (Zhu *et al*., 2001). For example, Soil Land Inference Model (SoLIM) combines the knowledge of local soil scientists with geographic information systems (GIS) under fuzzy logic to map soil at higher attribute accuracy (Zhu *et al*., 1997). The major digital components of this approach include engineering techniques for knowledge acquisition, knowledge representation, and knowledge-based inference. Some authors have argued that making good use of soil scientists′ tacit knowledge is efficient and economical. For instance, MacMillan *et al*. (2010) conducted predictive ecosystem mapping by using a knowledge-based fuzzy semantic import model to assess soils across 8.2×10^4 ha in Canada. Study cost was approximately Canadian $ 0.34/ha, and it had an average accuracy of 69%. Meanwhile, there is little doubt that the number of samples needed by an experienced soil scientist is considerably less than that required by a computationally inductive method. Zhu *et al*. (2008) explores a purposive sampling approach to improve the efficiency of field sampling for digital soil-mapping. These authors believe that unique soil conditions (soil types or soil properties) can be associated with unique combinations (configurations) of environmental conditions. The purposive sampling ap-

proach used the fuzzy c-means (FCM) of classification to identify these unique combinations and their spatial locations. Liu *et al*. (2012) indicated that the purposive sampling approach was effective for mapping the variation of soil textures using Moderate Resolution Imaging Spectroradiometer (MODIS) images. Later, an improved purposive sampling called the integrative hierarchical sampling strategy (IHSS) proved more efficient (Yang *et al*., 2012).

Based on the review mention above, this study employed an integrated method of IHSS and SoLIM with certain repetitive samples to characterize the spatial variability of soil salinity. The case study in the Uyghur Autonomous Region of Xinjiang of China was used to illustrate the efficacy of the employed method at a regional scale. Then the spatial distribution of soil salinity was described based on the prediction and strategies for amelioration and management of salt-affected areas were suggested.

2 Materials and Methods

2.1 Study area

Located in Northwest China, far from oceans, Xinjiang is a typical arid area (34°25′–49°10′N, 73°40′–96°23′E). There are three mountain ranges in Xinjiang. From south to north, these ranges are the Kunlun Mountains, the Tianshan Mountains, and the Altay Mountains. In the middle, the Tianshan Mountains divide Xinjiang into the northern and the southern parts (Fig. 1). Northern Xinjiang has a continental arid and semi-arid climate, with a mean temperature of –13℃ in winter and 22.2℃ in summer. Southern Xinjiang has a continental dry climate, with a winter mean temperature of –5.7℃ and a summer mean temperature of 24.4℃. The annual precipitation is 210 mm in the northern Xinjiang while that in southern Xinjiang is less than 100 mm. Because of the dry climate in Xinjiang, evaporation is vigorous, with a mean annual pan-evaporation between 1000 mm and 4500 mm, which is 500–1000 mm greater than other places in China at the same latitude (Li *et al*., 2011).

The areas denoted by gray in Fig. 1 were selected for study. The study areas are covered by different soil types, including Solonchak, Phaeozems, Gleysols, Fluvisols, and Anthrosols (FAO/UNESCO, 1990). These flat areas have experienced dramatic exploitation since the 1960s (Zhang *et al*., 2003), creating soil salinization that has been a serious ongoing environmental problem (Wang and Li, 2013). In these low-relief areas, each drainage basin has similar geomorphic features, consisting of diluvial-alluvial fans, oases in the middle, and a flat, low-gradient alluvial-diluvial plain (lower reaches) termed a desert-oasis ecotone (Wang and Li, 2013). Inside the oasis, where the major crops are corn and winter wheat, the soil tends to become thicker and sandier. The desert-oasis ecotone is dominated by non-zonal halophilus vegetation, such as *Halocnemum strobilaceum*, *Kalidium foliatum*, *Haloxylon ammodendron*, *Reaumuria soongorica*, *Tamarix chinensis*, *and Nitraria sibirica*. The structure of the vegetation is simple, and its distribution is sparse.

2.2 Data acquisition

2.2.1 Environmental data

It is important to select appropriate environment covariates for mapping soil salinity. Covariates can be selected according to scorpan factors, which can be described as $Sa = f(S, C, O, R, P, A, N)$. S represents other or previously measured attributes of the soil at a point; C represents climatic properties of the environment at a point; O represents organisms; R represents topography; P represents parent material; A represents a time factor; N represents spatial or geographic position. The accuracy of digital soil-mapping also relies on the local relationship between the soil properties and their environmental factors. After the literature review, an understanding of environment factors that may affect (or were affected by) the soil-salinity distribution and the availability of such spatial environmental information (Dehaan and Taylor, 2002; McBratney *et al*., 2003; Masoud and Koike, 2006; Hu *et al*., 2009; Lobell *et al*., 2010; Rodríguez-Pérez *et al*., 2011; Heil and Schmidhalter, 2012; Wang *et al*., 2012; Zhang *et al*., 2012; Wang *et al*., 2013), seven environmental factors were available. These included land surface temperature (LST), enhanced vegetation index (EVI), near-infrared (NIR), slope, topographic wetness index (TWI), and soil texture (silt and sand).

The present study made use of the Moderate Resolution Imaging Spectroradiometer (MODIS) LST product (MOD11A2), which offers daytime and nighttime land-surface temperature data stored on a 1-km grid as the average values of clear-sky LST during an 8-day

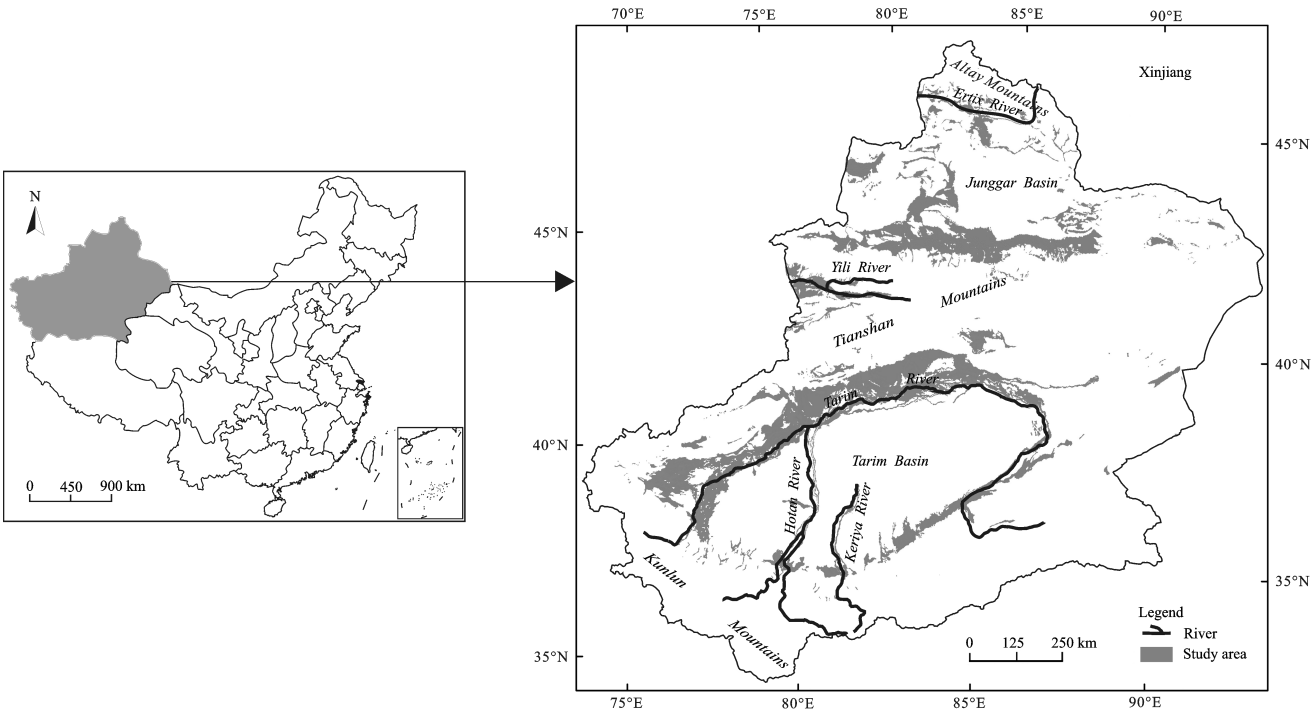

Fig. 1 Distribution of study areas. Study areas (shown in gray) are clipped by soil types that include Solonchak, Phaeozems, Gleysols, Fluvisols, and Anthrosols according to the Harmonized World Soil Database (Nachtergaele *et al*., 2008)

period (Wang *et al*., 2005). The MOD11A2 is derived from two thermal infrared band channels; that is, 31 (10.78–11.28 μm) and 32 (11.77–12.27 μm), using the split-window algorithm, which corrects for atmospheric effects and emissivity using a look-up table based on global land-surface emissivity in the thermal infrared (Snyder *et al.*, 1998). The MOD11A2 products are validated over a range of representative conditions, meaning the product uncertainties are well defined and have been satisfactorily used in a large variety of scientific studies. The EVI is satellite data based on measurement of vegetation greenness produced using blue (0.459–0.479 μm), red (0.620–0.67 μm), and near-infrared (0.842–0.872 μm) bands surface reflectance data from a MODIS instrument. The dataset MOD13A2 contains EVI at 1 km spatial resolution and 16-day frequency. This 16-day frequency arises from compositing; that is, assigning one best-quality EVI value to represent a 16-day period. The MODIS NIR was included in MOD13A2 dataset. Two MODIS products, LST (MOD11A2) and EVI (MOD13A2), and NIR were downloaded from https://lpdaac.usgs.gov/lpdaac/get_data/wist for dates between July 26 and August 11, 2012.

The slope information was derived from a 1-km resolution DEM in ArcGIS 9.3. The topographic wetness index was calculated according to the following equation (Beven and Kirkby, 1979): $\omega = \ln(\alpha/\beta)$, where α is the cumulative upslope area draining through a certain point (per unit contour length) and β is the slope gradient at that point. Because the relief of our study area is flat, multiple-flow strategy MFD-fg was used to calculate the upslope drainage area (α) (Qin *et al*., 2007).

The soil texture (silt and sand) was collected from the Harmonized World Soil Database, a service that is provided by the Institute of Soil Science, Chinese Academy of Sciences. This soil map was generated based on 8595 soil profiles from the Second National Soil Survey in China. The dataset, with a 1-km resolution, is reliable and can be applied to land and climate modeling at a regional scale (Wei *et al*., 2012).

2.2.2 Field sampling

The MODIS measurements and other environmental datasets reflect the average characteristics of a 1 km × 1 km pixel; thus, a comparison of the field measurements at individual points requires some assumptions about how the two scales relate. In this paper, we used a common approach that sampled 16 points at each location, which were averaged at each site to produce the MODIS co-located 1 km × 1 km soil-salinity assessment.

The field-observation data (0–10 cm) that were em-

ployed in this analysis included the soil salt content (g/kg) in the depth of 0–10 cm. In the laboratory, the composite soil samples were air-dried, ground, and sieved, using a 2-mm sieve mesh and then analyzed for the soil-salt content. These measurements were performed at the State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences. The primary ionic concentrations were measured using Ion Chromatography (Dionex 600), following the manufacturer′s protocols. The total soil-salt content was defined as the concentration of the following eight ions: K^+, Ca^{2+}, Mg^{2+}, Na^+, Cl^-, SO_4^{2-}, CO_3^{2-}, and HCO_3^-.

A total 94 samples from Fubei Oasis, Kuqa Oasis, and Keriya Oasis were considered for validation (Fig. 2). Samples from Fubei Oasis (24 validation points) covered major environmental variations within the shortest distance. These were used to evaluate how soil-salinity maps capture spatial variation of soil-salt content. The purpose of Kuqa Oasis sampling (49 validation points) was to validate the overall performance in predicting patterns of soil salinity at catchment scale. Subjective sampling was conducted in Keriya Oasis to assess representativeness of points determined by the study of Yang *et al*. (2012), in which 21 subjective points were collected.

2.3 Methods

2.3.1 Multiple linear regression model

The multiple linear regression (MLR) model (Flury and Riedwyl, 1988), the regression of soil salinity against a set of environmental variables, was developed as a reference to assess the employed method. All of the 94 field points were used to develop the MLR model by the stepwise variable-selection method, with default criteria in SPSS 16.0 (i.e., Probability-of-F-to-enter ≤ 0.050 and Probability-of-F-to-remove ≥ 0.100).

2.3.2 Stratifying study area into relatively homogeneous

To determine the soil-salinity variation through analysis of the spatial differences in the soil surface between geographic locations, it was necessary to control all of the major influencing factors other than the soil condition. Thus, in this research, the entire study area was stratified into vegetation units based on the NDVI.

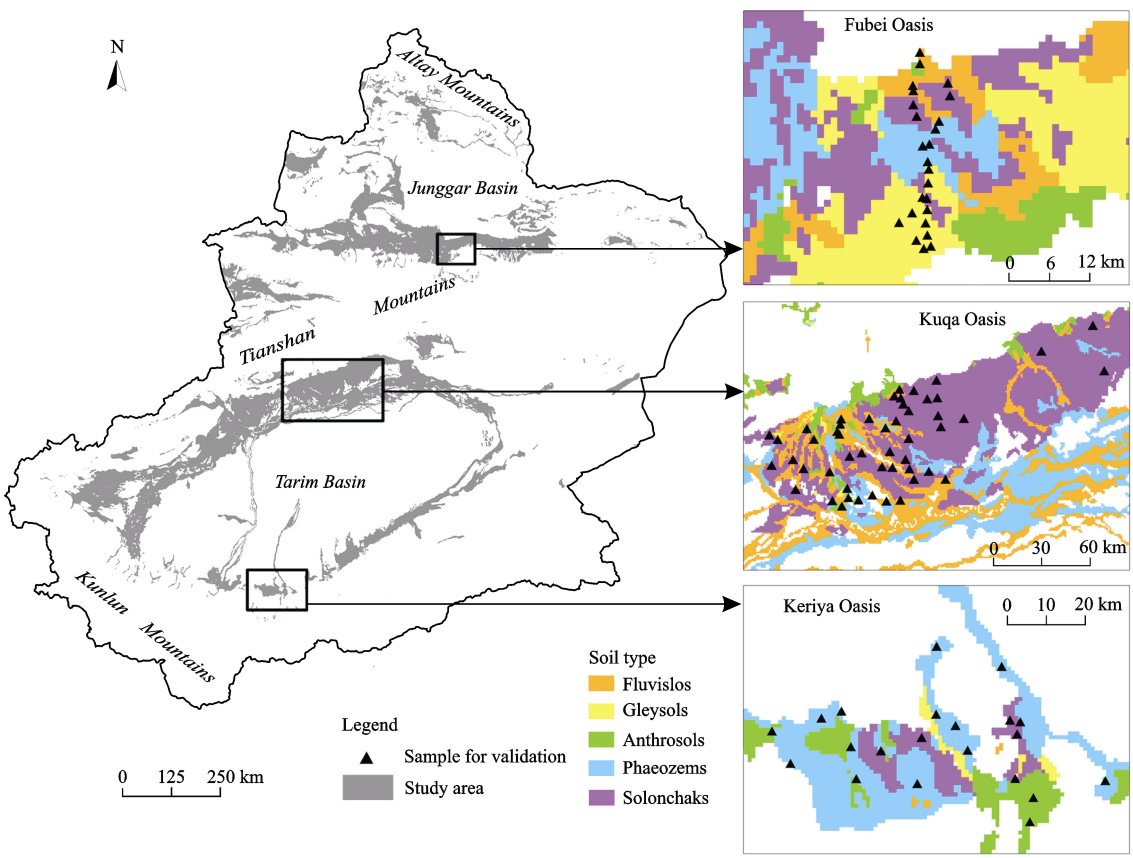

Fig. 2 Location of validation samples

Areas with an NDVI of less than 0.3 were considered as bare land or land with very sparse vegetation coverage (Liu *et al*., 2012). Four units were then classified (Fig. 3). The stratification resulted in relatively homogeneous units. Within any given unit, the soil-surface differences between the locations could be primarily attributed to differences in the soil conditions.

2.3.3 Integrative hierarchical sampling strategy (IHSS)

The basic purpose of IHSS was to sample locations where the soils were typical of the various soil categories (Yang *et al*., 2012). Based on the soil-landscape model theory, this method assumed that typical instances of soil classes corresponded to unique environmental conditions. Fuzzy c-means clustering was used to identify the unique combinations (or environmental classes) that existed in the environmental data set. FCM is an unsupervised classification method that optimally partitions a dataset into a given set of classes and then computes the fuzzy membership of each data element in each class (Bezdek *et al*., 1984).

For each unit, all the seven environmental variables were used for purposive sampling. The hierarchy of a representative grade for candidate samples was determined through a cluster analysis using a fuzzy c-means classifier (Bezdek *et al*., 1984). The FCM classifier was performed in multiple iterations, with each iteration identifying a different number of clusters. For example, FCM could be performed to identify 2 to 20 clusters for a given area. The range of clustering iterations had to cover both large-scale and local patterns of spatial variation for a geographic feature in the study area. To find an appropriate range, one could test a number of clustering iteration ranges and evaluate the scale ranges of the resulting environmental clusters (Yang *et al*., 2012). Two cluster validity measurements, the partition coefficient and normalized entropy, were employed to deter mine the optimal values for this parameter (Bezdek *et al*., 1984). In this case, the optimal value of the

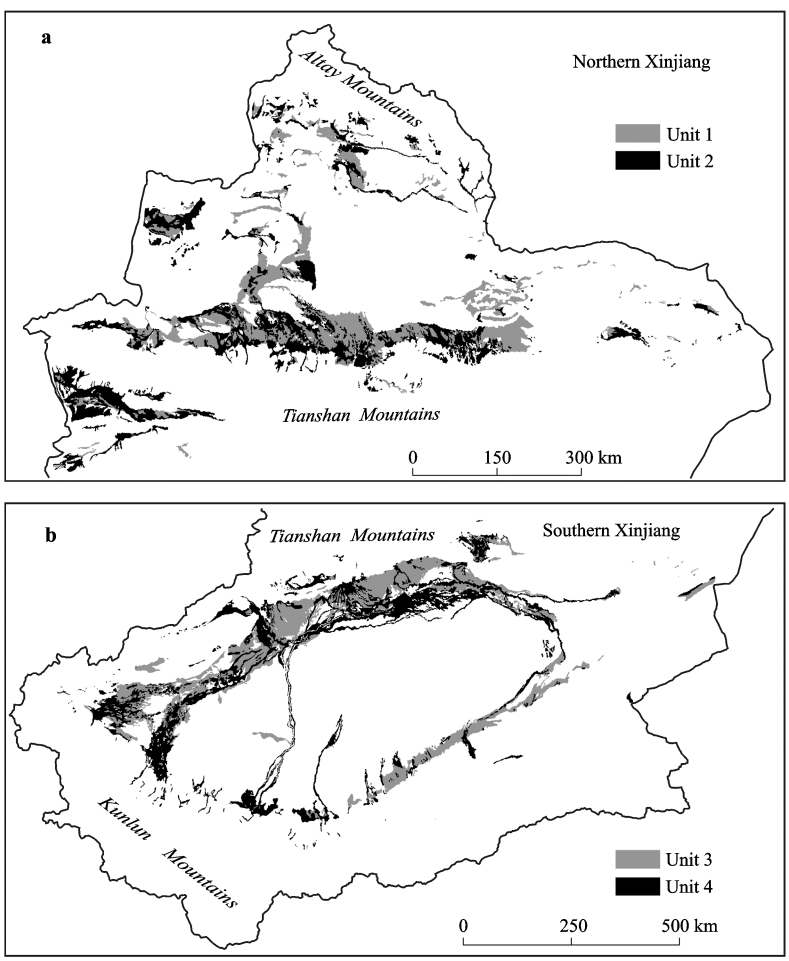

Fig. 3 Stratification of study area into different landform-vegetation units. Unit 1 and unit 3: NDVI < 0.3; unit 2 and unit 4: NDVI > 0.3

fuzzy-weighting exponents was 1.75. For unit 1, 4, 5, 7 and 8 clusters were appropriate. Then, four iterations were performed with iteration 1 to identify four clusters and with iteration 2 to identify five clusters, and so on. Each of these iterations was referred to as an '*x* cluster iteration', where *x* is the number of clusters to be extracted, such as 4 clusters corresponding with 4 classes. A class with high membership values was considered most likely to be typical. Points with membership values greater than 0.8 were regarded as typical (Liu *et al*., 2012) (Fig. 4a). The fuzzy membership maps were then reclassified using binary values, where 1 indicates representative locations (for any class) and 0 indicated non-representative locations. In each iteration, all of the fuzzy membership maps were converted to binary maps to identify representative locations for each class. Figure 4b is a two-valued map showing the representative locations for this class, with a threshold of 0.8. Then, four reclassified binary maps, for example, were overlaid to produce one binary map, and then seven maps were overlaid to produce the next binary map, and so on. Four binary maps were converted. Based on this concept, the reclassified binary maps for all iterations were overlaid once again to produce a frequency map, with values at each location indicating the representative grade.

The process discussed above was repeated until all units were calculated. The grade map is shown in Fig. 4c. For all areas, as we iterated through the five cluster numbers, a value of 5 on the resulting map meant that the pixel was representative of environmental clusters in all five iterations. The minimum value of a pixel was 0, which meant that the pixel was not representative of any environmental clusters during any iteration. Samples were selected at locations with priority given to the

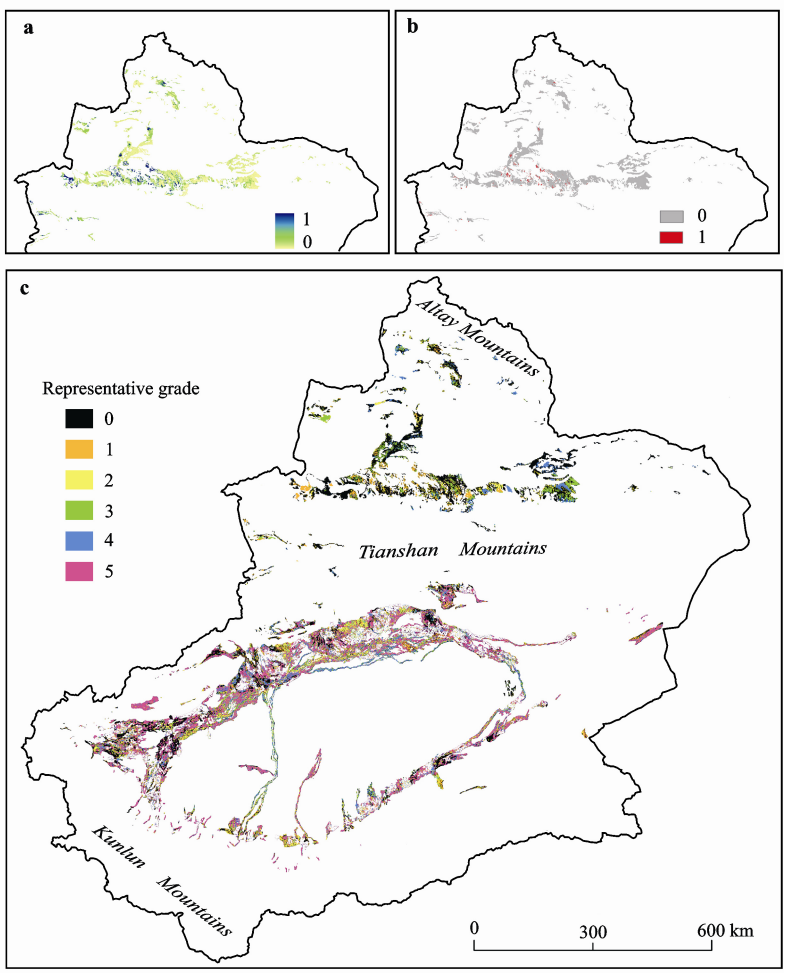

Fig. 4 Fuzzy membership map (a) and reclassified two-valued map (b) of certain class in unit 1, and representative grade map in study area (c). 1 indicates points with membership values greater than 0.8 and 0 indicates points with membership values less than 0.8 in Figs. 4a and 4b. Value 5 indicates 5 representative points overlaid in this pixel and 0 indicates no representative point overlaid in Fig. 4c

highest representative grade relative to the lower grades (Yang *et al*., 2012). In addition, for locations with the same representative grade, those whose environmental class covered a larger area (more pixels) had a higher priority. In our experiment, no sample was drawn from an environmental cluster with less than 40 pixels. Thirty-three typical locations were eventually identified. Based on the representative soil samples, the soil similarity vectors were then calculated using SoLIM.

2.3.4 Soil Land Inference Model (*SoLIM*)

SoLIM is a predictive approach to soil mapping (Zhu *et al*., 2001). The SoLIM states that if one knows the relationship between a soil type and its local environment, then one should be able to infer what soil might be at any particular location by assessing the environmental conditions at that location. The non-linear transformation of the environmental variables by SoLIM could help to improve the performance of linear regression (Zhu *et al*., 2010a). This approach has been used successfully to map the spatial variation of soil organic matter over mountainous areas using topographic variables (Zhu *et al*., 2010b).

The core of SoLIM is similarity (Zhu *et al*., 1996) in representing the soil spatial variation with fuzzy logic. The soil similarity vector is given by:

$$S_{i,j} = (S_{i,j}^1, S_{i,j}^2, ..., S_{i,j}^k, ..., S_{i,j}^n) \quad (1)$$

where n is the number of prescribed soil classes over the area; and $S_{i,j}^k$ is an index that measures the similarity of the local soil at i, j to a representative soil class k.

This measure was predicted, using fuzzy logic based on the similarity between the environmental conditions of a representative soil class k, and those at the local site; that is, this similarity value was the same as the fuzzy membership of the local soil to the soil class. Interested readers are referred to Zhu *et al*. (2001) for details.

2.3.5 Linearly weighted averaging function

Finally, a linearly weighted averaging function (Zhu *et al*., 1997) was used to map the spatial variation of the soil salinity. The modified function is shown in the following:

$$Sa_{i,j} = \frac{\sum_{k=1}^{n} S_{i,j}^k \times Sa^k}{\sum_{k=1}^{n} S_{i,j}^k} \quad (2)$$

where $Sa_{i,j}$ is the predicted value of the soil salt content at pixel (i, j) within any given landform-vegetation unit; Sa^k is the typical value of the soil salinity of environmental class k; $S_{i,j}^k$ is the fuzzy membership value of environmental class k at pixel (i, j); and n is the number of environmental classes within the unit.

2.3.6 Accuracy assessment

The quantitative measurements included the determination coefficient (R^2), root mean square error (RMSE), and ratio of prediction to deviation (RPD) which is the ratio of the standard deviation (SD) and RMSE to quantify the accuracy of the IHSS-SoLIM method and the MLR model. A predictive model was accurate if the R^2 and RPD values were greater than 0.91 and 2.5, respectively. An R^2 value between 0.82 and 0.90 and an RPD value greater than 2 indicated a good prediction, whereas an R^2 value between 0.66 and 0.81 with an RPD value between 1.5 and 2 indicated an approximate prediction. An R^2 value between 0.50 and 0.65 indicated a poor relationship (Farifteh *et al*., 2007).

3 Results and Analyses

3.1 Accuracy of prediction model

The diagrams in Fig. 5 showed the predicted soil-salt contents with measured soil-salt content for different areas and for different soil-salinity classes. The samples from Kuqa Oasis and Keriya Oasis resulted in an approximate relationship (referring to the values of RPD and R^2) between the predicted and measured values. For the Fubei Oasis, the predictions of the soil-salt content were little better than Kuqa oasis and Keriya oasis. Moreover, the field measurements at these sites were further compared with soil subgroups that were obtained from different soil-salinity classes (none-saline soil, < 10 g/kg; slightly saline soil, 8–10 g/kg; moderately saline soil, 10–15 g/kg; severely saline soil, 15–20 g/kg; extremely saline soil, > 20 g/kg) (Szabolcs, 1992; Qiao *et al*., 2012). The results indicate that the non-saline soil group (R^2 = 0.25) had a lower accuracy compared to the saline soil group (R^2 = 0.56).

The quantitative measurements were calculated using all 94 validation samples from three oases, by comparing the IHSS-SoLIM model and MLR model, as shown in Fig. 6. Five variables (NIR, sand, TWI, LST, and EVI) were selected based on stepwise regression model to build the MLR model (Table 1).

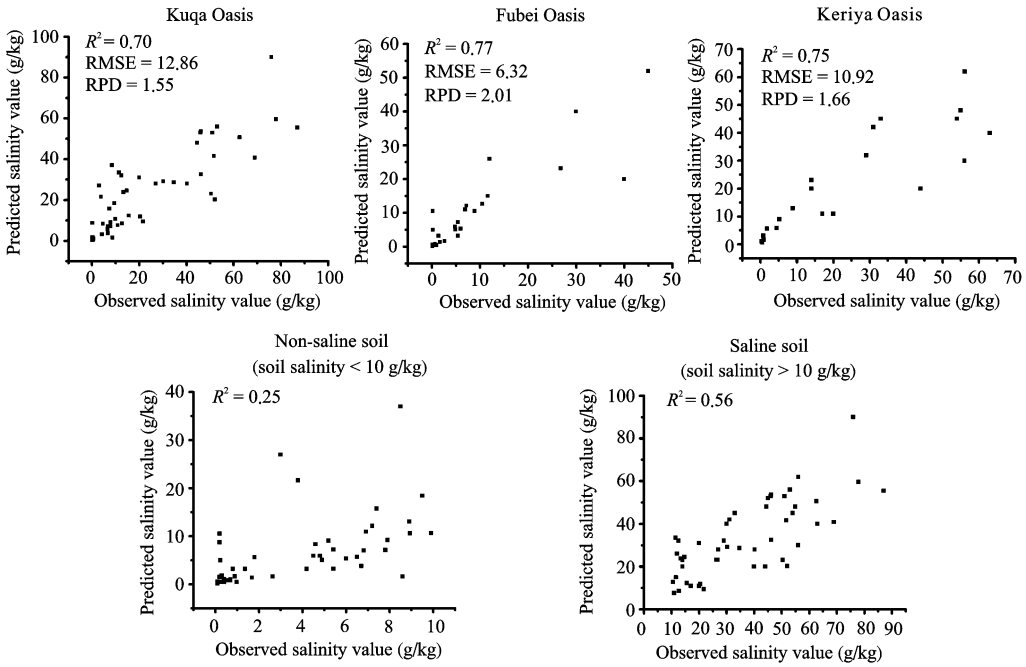

Fig. 5 Comparison of observed and predicted salinity values in different areas and saline-soil classes

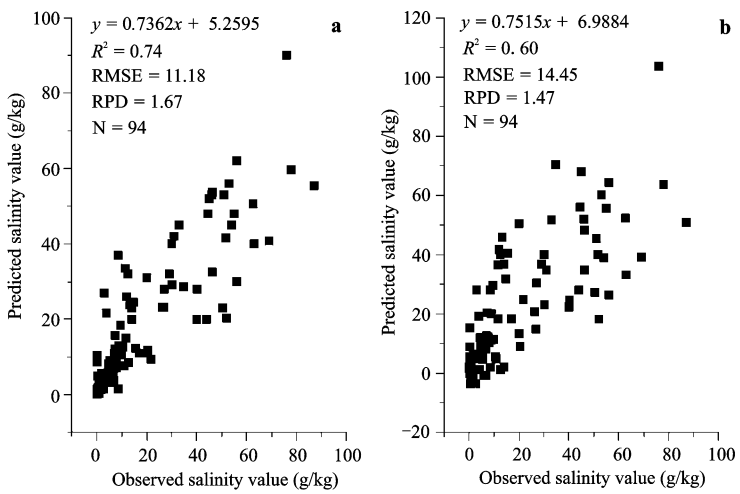

Fig. 6 Comparison of observed and predicted salinity values calculated by IHSS-SoLIM method (a) and MLR model (b)

Table 1 Selection of variables based on stepwise regression model

Model	R	R^2	Adjusted R^2	Equation	S.E. of estimate
1	0.63	0.40	0.39	$y = -220.28 \times NIR + 82.17$	18.83
2	0.75	0.56	0.55	$y = -174.35 \times NIR + 0.580 \times Sand + 43.69$	16.18
3	0.80	0.65	0.64	$y = -159.56 \times NIR + 0.438 \times Sand - 49.86 \times EVI + 64.76$	14.52
4	0.84	0.71	0.70	$y = -154.34 \times NIR + 0.436 \times Sand - 44.99 \times EVI - 27.12 \times Slope + 68.47$	13.26
5	0.87	0.75	0.74	$y = -111.06 \times NIR + 0.381 \times Sand - 42.93 \times EVI - 28.86 \times Slope + 1.24 \times LST + 10.74$	12.26

Notes: NIR is near-infrared; Sand is soil texture (percentage of sand); EVI is enhanced vegetation index; Slope is topographic index (°); LST is land surface temperature (℃); S.E. is standard error

There is no large-scale soil-salinity map that can be used for this aim in the study area. Meanwhile, only a few works in the literature refer to the quantitative study of soil salinity using MODIS data. In the study of Lobell *et al*., 2010, Pearson correlation coefficients were computed between the average-summer 7-year EVI data (MODIS product MOD13Q1 with 250 m resolution) and soil salinity using 118 field points. The result indicated that the multi-year average of EVI was able to explain one-half ($R^2 = 0.53$) of field-measured variations in salinity across Kittson County. In addition, the paper by Bouaziz *et al*. (2011) established the regression empirical model based on the green, red, and near-infrared bands of MODIS Product (MOD09A1 with 500 m resolution), using 112 samples. The R^2 value in the regression output indicated that only 41% of the total variation of the predicted soil salinity can be explained by the predictor variables used in the model. In spite of using MOIDS product and having similar size of validation samples, the study took place in different environmental conditions and used different predicted variables. Therefore, it is little difficult to make the comparison fairly.

In order to compare this result, MLR model was constructed based on 23 representative soil samples. For all 94 samples, R^2 is between 0.66–0.81 and RPD value is greater than 1.5, which indicated an approximated prediction inferred by the IHSS-SoLIM method (Fig. 6a). From Fig. 6b, we can see that the explanatory variables based on SoLIM explain more of the variation in the soil salinity than those (five environmental variables) based on the MLR model. Furthermore, Fig. 6b shows some negative soil-salt content values which relate to the unwanted extrapolation of the regression model (Zhu *et al*., 2010a).

3.2 Predicted map of soil salinity

From Fig. 7 and Fig. 8, several main characteristics of salt-affected soil were identified in a regional perspective. First, using the Tianshan Mountains as the dividing line, the area of soil salinization tended to increase from the northern to the southern Xinjiang. This feature was also proved in the study of Wang *et al*. (2009). In addition, the area around both sides of the Eritx River, located in

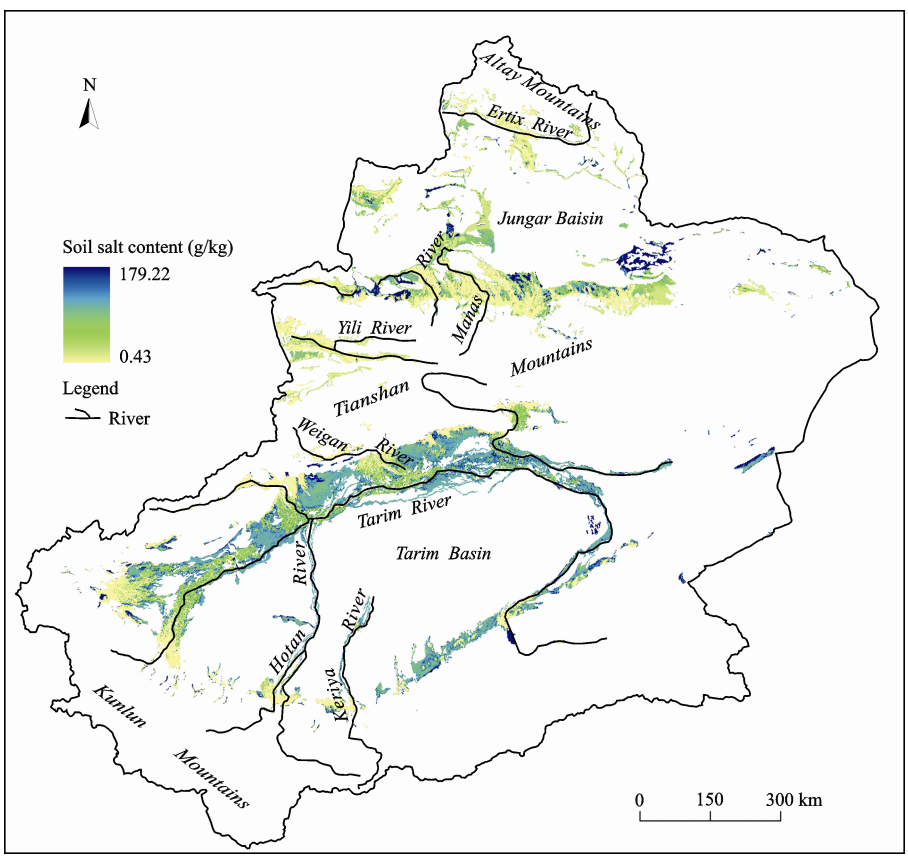

Fig. 7 Predicted map of soil-salt content by ISHH-SoLIM model

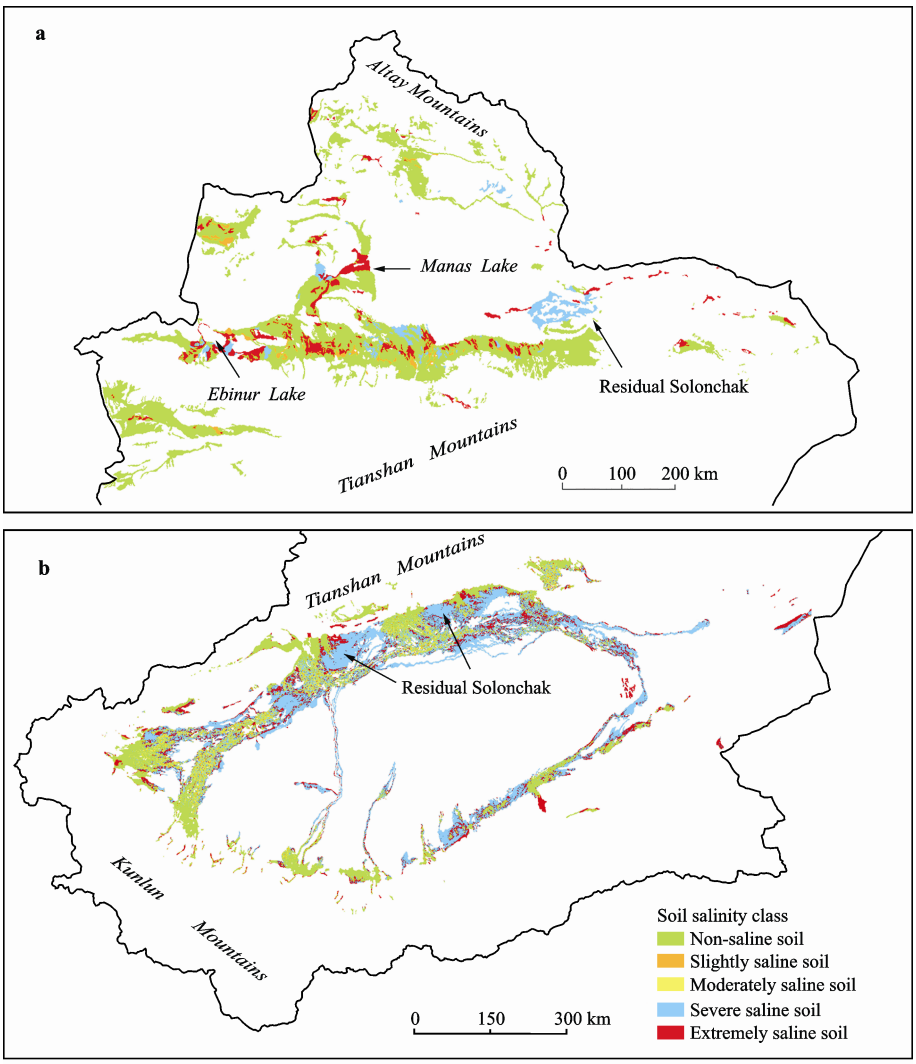

Fig. 8 Distribution of different soil-salinity classes in northern Xinjiang (a) and southern Xinjiang (b), China

the northern Junggar Basin, had lower levels of salinization due to the low degree of water mineralization in the Ertix River (Zhu *et al*., 2011) (Fig. 7). The northern and southern piedmonts of Tianshan Mountains were the main areas of soil salinization. The problem of secondary soil salinization in the oases was more serious in downstream than upstream. This is consistent with the result analyzed in the paper by Qiao *et al*. (2011). Severely saline soils (due to anthropogenic activities) appeared around terminal lakes (such as Ebinur Lake and Manas Lake) (Shi *et al*., 2008; Liu *et al*., 2011) (Fig. 8) and on both sides of rivers (Tarim River, Hotan River, and Keriya River) (Qiao *et al*., 2011) (Fig. 7). Large amounts of residual Solonchak soil were distributed in the southern Xinjiang but were only sporadically present in the northern Xinjiang, as shown in the Xinjiang part of the Harmonized World Soil Database (Nachtergaele *et al*., 2008) (Fig. 8).

From a catchment scale, we used the Kuqa Oasis as an epitome that qualitatively illustrates the spatial patterns of the predicted soil-salt content, by comparing the land use and the soil type (Fig. 9). The observations demonstrated that the major patterns in the predicted results had corresponding changes to the surface features of land use and soil type. Generally, the pattern of surface-soil salinity matched the expected spatial variation of the soil-salt content well, based on our local knowledge (Ding *et al*., 2011) and on a field survey.

The area of different soil-salinity classes was determined based on Fig. 8. In the study area, non-saline soil covered 98 845 km^2, accounting for 48.07% of the total study area. Salinization affected 106 781 km^2 of the

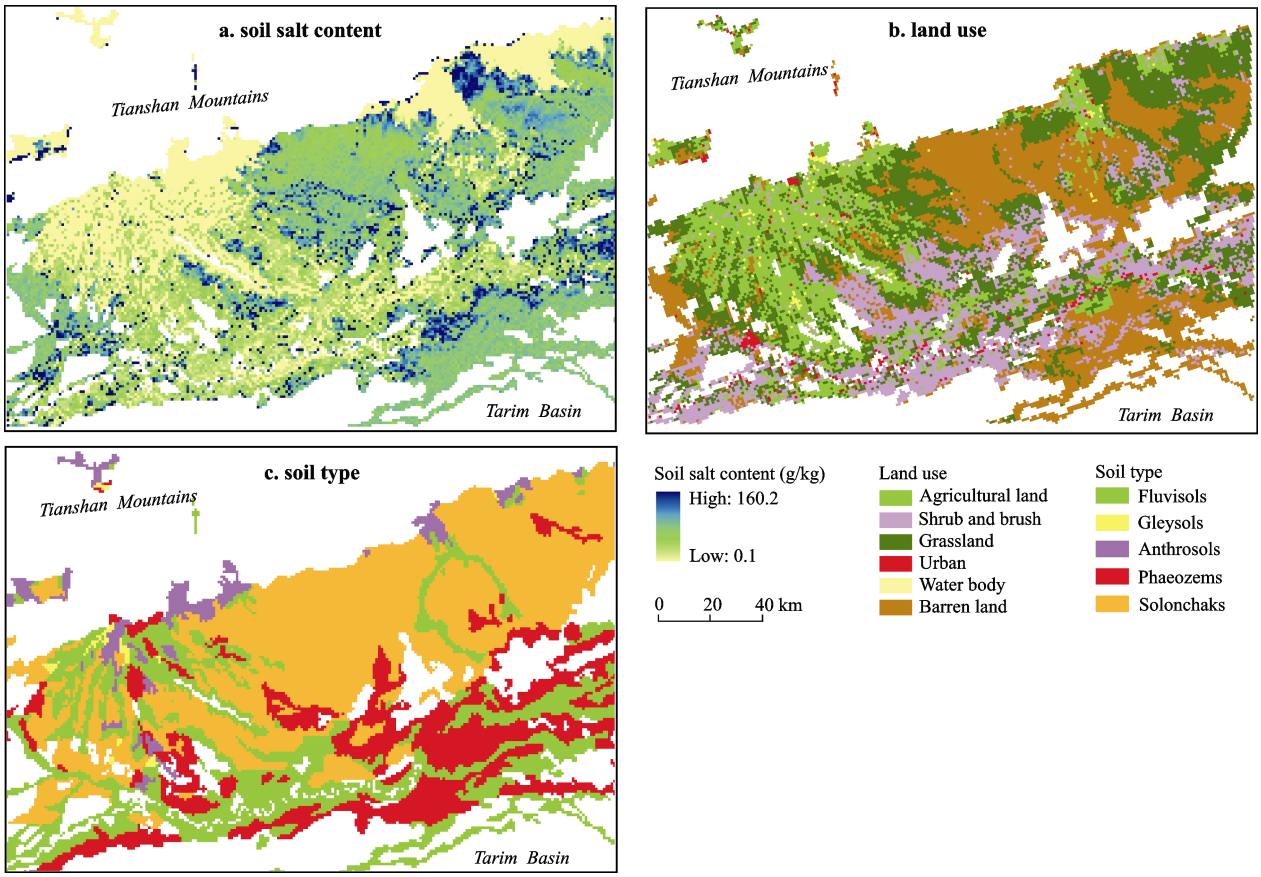

Fig. 9 Predicted map of soil salt content by ISHH-SoLIM method (a), land use type (b) and soil type (c) in Kuqa Oasis, Xinjiang

study area. Among the soils, slightly saline soil covered 2.21%; moderately saline soil covered 4.11%; severely saline soils (the largest area) covered 26.94%; and extremely saline soils covered 18.67%.

4 Discussion

4.1 Model performance

The IHSS-SoLIM method can characterize most of the variation in the soil-salt content over the study area (Fig. 7). The results from three oases provide a promising illustration of the power of combining different environmental data on the variation of soil salinity with IHSS-SoLIM method. The outcome from Fubei Oasis indicates the IHSS-SoLIM method can capture soil-salinity variation over short distances. The spatial pattern of soil salinity in Kuqa Oasis was also described appropriately. The Keriya Oasis is situated away from the main distribution of typical points; however, the result was satisfactory, showing that the representative locations derived from the study of Yang *et al*. (2012) was helpful and efficient.

The employed method can make full use of auxiliary soil-related environmental factors to improve the accuracy of the soil-salinity map, with respect to the MLR model. Moreover, a comparison of the MLR model and the IHSS-SoLIM method indicates that the transformation of simple environmental characteristics to more complicated fuzz membership values, via SoLIM′s fuzzy inference engine, increases the predictive power of those variables (Zhu *et al*., 2010a).

The deficiency of the IHSS-SoLIM method is that the variation of soil salinity was not better distinguished within non-saline soil compared to the saline soil. This is due to the variation of salinity in salt-affected soils is higher than farmland which tends to be more similar or homogeneous (Lobell *et al*., 2010). In addition, the soil conditions in this area were not reflected by the remote sensing at a particular time, due to the overlying vegetation when salts are present during the growing season. The outlier occurred in the non-saline soil (Fig. 5) partly because the area that is affected by salts surrounds

farmland, rather than transitional landscapes, and seems to be non-saline soil. It is difficult to formalize explicit rules for the environmental conditions of an ecotone (Shi *et al*., 2009). This salt-affected soil was mainly distributed at high variables in the ecotone, which acts as an interactive zone between irrigated farmland and the natural desert ecosystem (Wang and Li, 2013). To solve this problem in future work, additional variables will also likely be considered to improve the yield predictions, with information on surface hydrology and depth to groundwater, two potentially useful variables. Moreover, the low spatial resolution limits the discriminating power of the geographic features. The spatial resolution of the environment variables used in this study is not fine enough to describe the role of micro-scale variation, and subsequently to identify transitional landscapes, to create a typical feature. Hence, inadequate resolution of predictors is partially responsible for the unexplained variation. Furthermore, in this research, the temporal difference between the field-plot measurement and remote sensing image is another problem. Ideally, to strengthen the field-to-image correlation, field data and remote sensing data should be collected at the same time. It was not possible for us to collect all the require field data at the time the remote sensing image was acquired, especially for large areas (Song *et al*., 2013). This means that the spectral properties derived from the satellite image are unable to represent the actual spectral feature of the soil and relax the field-to-imagery correlation exactly (Ohmann and Gregory, 2002; Karl and Maurer, 2010).

4.2 Management and amelioration of soil salinization

4.2.1 Construction of soil salinity monitoring system in agricultural land

In order to avoid further deterioration on soil salinization, monitoring systems are required. It should be able to support critical reflection about the management actions, providing both negative and positive feedback in order to iteratively evaluate their effectiveness, and to detect unintended developments of the system being managed in the early stages (Giordano *et al*., 2010). The soil salinity monitoring system plays a fundamental role, especially inside agricultural land with non-saline soils (soil salt content < 8 g/kg) showed in Fig. 8. This system should integrate various types of knowledge, including local knowledge and technical knowledge, to address the issues of complexity and uncertainty of salt-affected environmental systems and to enhance the long term sustainability of the monitoring program.

The first step called the definition of information needs included understanding the current monitoring practices and identifying the main drawbacks. The second step involved eliciting the farmers' understanding of the soil salinization process and identifying key parameters for the assessment. Then, for the local scientists and technicians, this step involved eliciting a technical understanding of the soil salinization process and identifying parameters for the assessment. A GIS-based system was developed to facilitate the analysis and storage of the locally based information. The potential of fuzzy logic to address linguistic variables was exploited to cope with the qualitative nature of the farmers′ information (Giordano *et al*., 2010). At the end of this process, the salinity values were plotted on a map.

4.2.2 Managing water use between upstream and downstream

There is a need to link agricultural water management and the management of downstream groundwater systems to strike soil salinization, where required. Irrigation expansion is warranted in salt-affected area where the water infrastructure is underinvested, such as extremely saline soil distribution in the northern Xinjiang (Fig. 8a) and moderately and severe saline soil in the southern Xinjiang (Fig. 8b). Trade can help alleviate water problems in these areas when economic and political conditions are met by involving stakeholders, who can negotiate unavoidable soil salinization between upstream food production (non-saline soil with yellow color in Fig. 9a) and downstream salt-affected ecosystem (grassland and shrub land in Fig. 9b). There might be a scope to reduce the soil salinization by focusing more on how agriculture alters the dynamics (e.g., timing and variability) of water flow for better adapted to downstream groundwater ecosystem dynamics. The peak growth period of the vegetation is in May, June and July, during which transpiration water loss was tremendous and more water was needed accounting for 47% of the annual requirement. Therefore, the groundwater table as a feedback should be monitored in salt-affected soil after May, to water management of oasis agriculture.

4.2.3 Increasing plant diversification in salt-affected soil

Based on the salt tolerance of plant species, there are emerging examples of plant diversification and management for the optimal utilization of salt-affected soils and saline-sodic waters. The plant species that have shown potential under such environments are divided into five groups: 1) fiber, grain and special crops; 2) forage grass and shrub species; 3) medicinal and aromatic plant species; 4) bio-fuel crops; and 5) fruit trees. An appropriate selection is generally based on the ability of plant species to withstand elevated levels of soil salinity while also providing a saleable product or one that can be used on-farm; however, from an economic perspective, much depends on the local needs. More detail can be found in the studies of Qadir and Oster (2004) and Qadir *et al*. (2008). The stakeholders can formulate plant scheme based on the predicted soil salinity map. Finally, ecological protection and economic development would be balanced with help of soil salinity map in plant layout and the agricultural planning.

5 Conclusions

Increasing soil salinity is responsible for a dramatic decrease of agricultural production. To manage this problem, detailed and up-to-date knowledge about the spatial distribution of soil salinity is needed. Soil-salinity maps for this purpose were produced in large scale through linking the soil salinity values of typical samples to the polygons on the type map, according to the same soil type name. The soil salinity maps produced through conventional direct linking methods usually suffer from low accuracy, lack of spatial details, and poor current knowledge.

This study employed an integrated method (IHSS-SoLIM) to produce a soil-salinity map at a high level of spatial detail and accuracy. The method of application and the evaluation demonstrated that the employed approach can effectively predict the spatial distribution of soil salinity in low-relief areas. With help of 23 representative soil samples, the results indicated the prediction was appropriate in Kuqa Oasis and Keriya Oasis, and was little better in Fubei Oasis than that in those oases, according to the evaluation criterion. Based on all validation samples from three oases, accuracy estimation shows the employed method performed better (R^2 = 0.74, RPD = 1.67, RMSE = 11.18) than the multiple linear-regression (MLR) model (R^2 = 0.60, RPD = 1.47, RMSE = 14.45). The prediction was classified, and the statistical result showed 48.07% of study area was affected by the soil salinization, mainly distributed at downstream oases, around lakes, and on both side of rivers, and was more serious in the southern than the northern parts of Xinjiang. To address this situation, measurements should include constructing a soil-salinity monitoring system, the rational allocation of water resources, multifunctional agro-ecosystems, and plant diversification, with the guidance of predicted soil-salinity maps, to effectively ameliorate soil salinity.

Poor prediction in non-saline soil such as farmland can be accounted for by limitation in spatial resolution, non-synchronism between field sampling, and satellite pass-and-interpretation of ancillary variables in environmental variation. In future work, addition variables such as surface hydrology and groundwater table should be considered to compensate for the error mentioned above. In conclusion, the employed method can serve as an alternative model for soil-salinity mapping on a large scale in arid areas.

Acknowledgements

The authors would like to thank Yang Lin and Liu Feng for their generous help with revising this paper. This work is partly supported by the Xinjiang and Central Asian Scientific Data Sharing Platform.

References

Beven K J, Kirkby M J, 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24(1): 43–69. doi: 10.1080/02626667909491834

Bezdek J C, Ehrlich R, Full W, 1984. FCM: the fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2–3): 191–203. doi: 10.1016/0098-3004(84)90020-7

Bouaziz M, Matschullat J, Gloaguen R, 2011. Improved remote sensing detection of soil salinity from a semi-arid climate in Northeast Brazil. *Comptes Rendus Geoscience*, 343(11–12): 795–803. doi: 10.1016/j.crte.2011.09.003

Dehaan R L, Taylor G R, 2002. Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced soil salinization. *Remote Sensing of Environment*, 80(3): 406–417. doi: 10.1016/S0034-4257(01)00321-2

Ding J L, Tiyip T, Wu M C, 2011. Study on soil salinization information in arid region using remote sensing technique.

Agriculture Science in China, 10(3): 404–411. doi: 10.1016/S1671-2927(11)60019-9

FAO/UNESCO (Food and Agriculture Organization/United Nations Educational, Scientific and Cultural Organization), 1990. *Soil Map of The World Revised Legend*. World Soils Resource Report. FAO, Rome, Iatly.

FAO (Food and Agriculture Organization), 2002. *Crops and Drops*, *Making the Best Use of Water for Agriculture*. Food and Agriculure Organization of the United Nations, Rome, Italy.

Farifteh J, Van der Meer F D, Atzberger C *et al.*, 2007. Quantitative analysis of salt-affected soil reflectance spectra: a comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment*, 110(1): 59–78. doi: 10.1016/j.rse.2007.02.005

Flury B, Riedwyl H, 1988. *Multiple Linear Regression.* Netherlands: Springer, 54–74.

Ghassemi F, Jakeman A J, Nix H A, 1995. *Salinisation of Land and Water Resources*: *Human Causes*, *Extent*, *Management and Case Studies*. Canberra: University of New South Wales Press.

Giordano R, Liersch S, Vurro M *et al.*, 2010. Integrating local and technical knowledge to support soil salinity monitoring in the Amudarya river basin. *Journal of Environmental Management*, 91(8): 1718–1729. doi: 10.1016/j.jenvman.2010.03.010

Goovaerts P, Journel A G, 1995. Integrating soil map information in modelling the spatial variation of continuous soil properties. *European Journal of Soil Science*, 46(3): 397–414. doi: 10.1111/j.1365-2389.1995.tb01336.x

Grunwald S, Thompson J A, Boettinger J L, 2011. Digital soil mapping and modeling at continental scales: finding solutions for global issues. *Soil Science Society of America Journal*, 75(4): 1201–1213. doi: 10.2136/sssaj2011.0025

Heil K, Schmidhalter U, 2012. Characterisation of soil texture variability using the apparent soil electrical conductivity at a highly variable site. *Computers & Geosciences*, 39: 98–110. doi: 10.1016/j.cageo.2011.06.017

Hu S, Zhao R, Tian C *et al.*, 2009. Empirical models of calculating phreatic evaporation from bare soil in Tarim river basin, Xinjiang. *Environmental Earth Sciences*, 59(3): 663–668. doi: 10.1007/s12665-009-0063-z

Karl J W, Maurer B A, 2010. Spatial dependence of predictions from image segmentation: a variogram-based method to determine appropriate scales for producing land-management information. *Ecological Informatics*, 5(3): 194–202. doi: 10.1016/j.ecoinf. 2010.02.004

Li Q, Chen Y, Shen Y *et al.*, 2011. Spatial and temporal trends of climate change in Xinjiang, China. *Journal of Geographical Sciences*, 21(6): 1007–1018. doi: 10.1007/s11442-011-0896-8

Liu D M, Abuduwaili J, Lei J Q *et al*., 2011. Deposition rate and chemical composition of the aeolian dust from a bare saline playa, Ebinur Lake, Xinjiang, China. *Water Air and Soil Pollution*, 218(1–4): 175–184. doi: 10.1007/s11270-010-0633-4

Liu F, Geng X, Zhu A X *et al.*, 2012. Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS. *Geoderma*, 171–172: 44–52. doi: 10.1016/j.geoderma.2011.05.007

Lobell D B, Lesch S M, Corwin D L *et al.*, 2010. Regional-scale assessment of soil salinity in the red river valley using multi-year MODIS EVI and NDVI. *Journal of Environmental Quality*, 39(1): 35–41. doi: 10.2134/jeq2009.0140

MacMillan R A, Moon D E, Coupé R A *et al.*, 2010. Predictive ecosystem mapping (PEM) for 8.2 million ha of forestland, British Columbia, Canada. In: Boettinger J L (eds). *Digital Soil Mapping*: *Bridging Research*, *Environmenatal Application*, *and Oper ation*. Netherlands: Springer, 337–356.

Masoud A A, Koike K, 2006. Arid land salinization detected by remotely-sensed landcover changes: a case study in the Siwa region, NW Egypt. *Journal of Arid Environments*, 66(1): 151–167. doi: 10.1016/j.jaridenv.2005.10.011

McBratney A B, Mendonca S M L, Minasny B, 2003. On digital soil mapping. *Geoderma*, 117(1–2): 3–52. doi: 10.1016/S0016-7061(03)00223-4

Metternicht G I, Zinck J A, 2003. Remote sensing of soil salinity: potentials and constraints. *Remote Sensing of Environment*, 85(1): 1–20. doi: 10.1016/S0034-4257(02)00188-8

Mulder V L, de Bruin S, Schaepman M E *et al.*, 2011. The use of remote sensing in soil and terrain mapping—a review. *Geoderma*, 162(1–2): 1–19. doi: 10.1016/j.geoderma.2010.12.018

Nachtergaele F, Velthuizen H Y, Verelst L *et al*., 2008. *Harmonized World Soil Database* (*version 1.1*). Food and Agriculture Organization of the United Nations, Rome, Italy.

Ohmaan J L, Gregory M J, 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastral Oregon, USA. *Canadian Journal of Forest Research*, 32(4): 725–741. doi: 10.1139/x02-011

Qadir M, Oster J D, 2004. Crop and irrigation management strategies for saline-sodic soils and waters aimed at environmentally sustainable agriculture. *Science of the Total Environment*, 323(1–3): 1–19. doi: 10.1016/j.scitotenv.2003. 10.012

Qadir M, Tubeileh A, Akhtar J *et al.*, 2008. Productivity enhancement of salt-affected environments through crop diversification. *Land Degradation & Development*, 19(4): 429–453. doi: 10.1002/ldr.853

Qiao Mu, Zhou Shengbin, Lu Lei *et al*., 2011. Temporal and spatial changes of soil salinization and improved countermeasures of Tarim Basin Irrigation District in recent 25a. *Arid Land Geography*, 34(4): 604–613. (in Chinese)

Qiao Mu, Zhou Shengbin, Lu Lei *et al*., 2012. Causes and spatial-temporal changes of soil salinization in Weigan River basin, Xinjiang. *Progress in Geography*, 31(7): 904–910. (in Chinese)

Qin C, Zhu A X, Pei T *et al.*, 2007. An adaptive approach to selecting a flow-partition exponent for a multiple-flow-direction algorithm. *International Journal of Geographical Information Science*, 21(4): 443–458. doi: 10.1080/13658810601073240

Rodríguez-Pérez J, Plant R, Lambert J J *et al.*, 2011. Using

apparent soil electrical conductivity (ECa) to characterize vineyard soils of high clay content. *Precision Agriculture*, 12(6): 775–794. doi: 10.1007/s11119-011-9220-y

Sheng J, Ma L, Jiang P A *et al.*, 2010. Digital soil mapping to enable classification of the salt-affected soils in desert agro-ecological zones. *Agricultural Water Management*, 97(12): 1944–1951. doi: 10.1016/j.agwat.2009.04.011

Shi X, Franklin J, Chadwick O A *et al*., 2009. Integrating different types of knowledge for digital soil mapping. *Soil Science Society of America*, 73: 1682–1692. doi: 10.2136/sssaj2007.0158

Shi Xingmin, Li Youlin, Yang Jingchun, 2008. Climatic and tectonic analysis of Manas Lake changes. *Scientia Geographica Sinica*, 28(2): 266–271. (in Chinese)

Snyder W C, Wan Z, Zhang Y *et al*., 1998. Classification-based emissivity for land surface temperature measurement from space. *International Journal of Remote sensing*, 19(14): 2753–2774. doi: 10.1080/014311698214497

Soil Survey Staff of Xinjiang, 1996. *Soil of Xinjiang Province*. Beijing: Science Press, 52–53. (in Chinese)

Song Chuangye, Huang Chong, Liu Huiming, 2013. Predictive vegetation mapping approach based on spectral data, DEM and Generalized Additive Models. *Chinese Geographical Science*, 23(3): 331–343. doi: 10.1007/s11769-013-0590-0

Szabolcs I, 1992. Salinization of soil and water and its relation to desertification. *Desertification Control Bulletin*, 21: 32–37.

Wang F, Chen X, Luo G P *et al*., 2013. Detecting soil salinity with arid fraction integrated index and salinity index in feature space using Landsat TM imagery. *Journal of Arid Land*, 5(3): 340–353. doi: 10.1007/s40333-013-0183-x

Wang Fangfang, Wu Shixin, Qiao Mu *et al*., 2009. Investigation and analysis on the salinization degree of cultivated land in Xinjiang based on 3S technology. *Arid Zone Research*, 26(3): 366–371. (in Chinese)

Wang K, Wang P, Sparrow M *et al*., 2005. Estimation of surface long wave radition and broadband emissivity using Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature/emissivity products. *Journal of Gephysical Research*, 110: D11109. doi: 10.1029/2004JD005566

Wang Q, Li P H, Chen X, 2012. Modeling salinity effects on soil reflectance under various moisture conditions and its inverse application: a laboratory experiment. *Geoderma*, 170: 103–111. doi: 10.1016/j.geoderma.2011.10.015

Wang Y, Li Y, 2013. Land exploitation resulting in soil salinization in a desert-oasis ecotone. *Catena*, 100: 50–56. doi: 10.1016/j.catena.2012.08.005

Wei S, Dai Y, Liu B *et al.*, 2012. A soil particle-size distribution dataset for regional land and climate modelling in China. *Geoderma*, 171–172: 85–91. doi: 10.1016/j.geoderma. 2011.01. 013

Yang L, Zhu A X, Qi F *et al.*, 2012. An integrative hierarchical stepwise sampling strategy for spatial sampling and its application in digital soil mapping. *International Journal of Geographical Information Science*, 27(1): 1–23. doi: 10.1080/13658816.2012.658053

Zhang F, Tiyip T, Ding J *et al.*, 2012. Spectral reflectance properties of major objects in desert oasis: a case study of the Weigan-Kuqa river delta oasis in Xinjiang, China. *Environmental Monitoring and Assessment*, 184(8): 5105–5119. doi: 10.1007/s10661-011-2326-x

Zhang H , Wu J W, Zheng Q H *et al.*, 2003. A preliminary study of oasis evolution in the Tarim Basin, Xinjiang, China. *Journal of Arid Environments*, 55(3): 545–553. doi: 10.1016/S0140-1963(02)00283-5

Zhu A X, Band L E, Dutton B *et al.*, 1996. Automated soil inference under fuzzy logic. *Ecological Modelling*, 90(2): 123–145. doi: 10.1016/0304-3800(95)00161-1

Zhu A X, Band L, Vertessy R *et al.*, 1997. Derivation of soil properties using a Soil Land Inference Model (SoLIM). *Soil Science Society of America Journal*, 61(2): 523–533. doi: 10.2136/sssaj1997.03615995006100020022x

Zhu A X, Hudson B, Burt J *et al.*, 2001. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science Society of America Journal*, 65(5): 1463–1472. doi: 10.2136/sssaj2001.6551463x

Zhu A X, Qi F, Moore A *et al.*, 2010b. Prediction of soil properties using fuzzy membership values. *Geoderma*, 158(3–4): 199–206. doi: 10.1016/j.geoderma.2010.05.001

Zhu A X, Yang L, Li B *et al.*, 2008. Purposive sampling for digital soil mapping for areas with limited data. In: Hartemink A E *et al*. (eds). *Digital Soil Mapping with Limited Data*. Netherlands: Springer, 233–245.

Zhu A X, Yang L, Li B *et al.*, 2010a. Construction of membership functions for predictive soil mapping under fuzzy logic. *Geoderma*, 155(3–4): 164–174. doi: 10.1016/j.geoderma.2009.05.024

Zhu B Q, Yang X P, Rioual P *et al*., 2011. Hydrogeochemistry of three watersheds (the Erlqis, Zhungarer and Yili) in northern Xinjiang, NW China. *Applied Geochemistry*, 26(8): 1535–1548. doi: 10.1016/j.apgeochem.2011.06.018