

Spatial Discrimination of Salt-affected Soil Surfaces Using Landsat TM Data : A Preliminary Study in the Middle Reaches of the Heihe River, China

Wu, Yueru

Cold and Arid Regions Environmental and Engineering Institute, CAS

Wang, Weizhen

Cold and Arid Regions Environmental and Engineering Institute, CAS

Mori, Makito

Faculty of Agriculture, Kochi University

Yasutake, Daisuke

高知大学農学部

他

<https://doi.org/10.5109/25206>

出版情報 : 九州大学大学院農学研究院紀要. 57 (2), pp.461-465, 2012-09-20. Faculty of Agriculture, Kyushu University

バージョン :

権利関係 :



Spatial Discrimination of Salt-affected Soil Surfaces Using Landsat TM Data – A Preliminary Study in the Middle Reaches of the Heihe River, China –

Yueru WU¹, Weizhen WANG¹, Makito MORI², Daisuke YASUTAKE², Hiroyuki CHO³,
Tetsuo KOBAYASHI⁴ and Masaharu KITANO*

Laboratory of Agricultural Meteorology, Division of Bioproduction Environmental Sciences,
Department of Agro-environmental Sciences, Faculty of Agriculture,
Kyushu University, Fukuoka 812-8581, Japan

(Received April 27, 2012 and accepted May 10, 2012)

A preliminary study has been conducted to test the effectiveness of the canonical discriminant approach to making spatial discrimination of salt-affected soil surfaces, using the Landsat TM data taken in the middle reaches of the Heihe River, China. A saline area can be considered a mixture of salt-affected patches which differ in degree, and hence we tried to make discrimination of salt-affected soil surfaces by canonical discriminant approach. Three 360×360 m test plots with different land surface conditions, “Bare-soil surface area (strongly saline)”, “Reed area (moderately saline)”, “Barley field (weakly saline)”, were selected as training areas or groups. Five indices, B4 (band 4 DN), *Wetness*, SI, NDSI, and *Fe₂O₃* (composite indices), were determined as discriminating variables, and two canonical discriminant functions, y_1 and y_2 , were obtained for the three groups using these five discriminating variables. When group centroids and individual cases were plotted on the plane made by two canonical discriminant functions, y_1 and y_2 , the groups of cases were quite distinct and their centroids were well separated, which suggest that the canonical discriminant approach is promising for spatial discrimination of salt-affected soil surfaces.

Keywords: canonical discriminant approach, Heihe River, Landsat TM data, soil salinization

INTRODUCTION

Salinity has plagued irrigated agriculture throughout history (Tanji, 1990), and soil salinization is an old-yet-new problem in arid regions. For example, there are approximately 100 million ha of salt-affected soils in China, accounting for one tenth of the total land area (Li, 2009). The development of effective salinity control practices requires an understanding of the causal relationship between soil and water salinization and salt formation and movement in the soil. Thus, the first step in establishing an innovation in the salinity control practices is to develop a means for assessing the degree of soil salinization and its spatial distribution easily and efficiently. Remote sensing is a possible candidate for an effective tool in detecting temporal and spatial changes in soil salinity. However, in general, irrigated lands are studded with saline plots of various sizes (Wang *et al.*, 2008), so that fields with the crops in vigor are intermingled with ones in which the crops lack vigor, or plots almost bare of vegetation. As a result, how to cope with mixed pixels is one of main problems in utilizing remote sensing for this purpose effectively.

Soil salinity is indicated by (a) presence of white crust and efflorescence and (b) soil parameters such as EC and SAR, and also by (c) invasion of salt-tolerant weeds and (d) lack of plant vigor (Metternicht and Zinck,

2009; Araki *et al.*, 2010). The former two, (a) and (b), reflect “soil” conditions, and the latter two, (c) and (d), reflect “vegetation” conditions at each site. In this study we regard a saline area as being composed of salt-affected patches which differ in degree, and try to make discrimination of salt-affected soil surfaces by composing canonical discriminant functions based on the discriminating variables which indicate the “vegetation” and “soil” conditions, and using these functions as salinity indicators. This paper describes the methodology and results of a preliminary study made in the middle reaches of the Heihe River, China, using the Landsat TM data.

CANONICAL DISCRIMINANT APPROACH

We apply a kind of multivariate data analysis, canonical discriminant approach to spatial discrimination of salt-affected soil surfaces. In multiple discriminant analysis, the samples expressed by vector variables of p components are projected from their space into suitable subspace, of which the dimension n is smaller than p . First of all, the framework of this approach is summarized.

Discriminating variables of p components are defined as (Anderson, 1958; Cooley and Lohnes, 1971):

\mathbf{X}_{ki} : vector variable for subject or case i ($i=1\sim N_k$)
and group k ($k=1\sim g$)

where g : number of groups, N_k : number of subjects or cases in group k ,

\mathbf{m} : vector of total mean or ground centroid

\mathbf{m}_k : vector of group mean or group centroid for group k

Then we have

$$\mathbf{X}_{ki} - \mathbf{m} (\equiv \mathbf{x}_{ki}) = (\mathbf{m}_k - \mathbf{m}) + (\mathbf{X}_{ki} - \mathbf{m}_k)$$

Further, matrices of sums of squares and cross products

¹ Cold and Arid Regions Environmental and Engineering Institute, CAS, China

² Faculty of Agriculture, Kochi University, Japan

³ Faculty of Agriculture, Saga University, Japan

⁴ SPA & Water Press, Fukuoka, Japan

* Corresponding author (E-mail: kitano@bpes.kyushu-u.ac.jp)

are defined as:

$$T = \sum_{k=1}^g \sum_{i=1}^{N_k} \mathbf{x}_{ki} \mathbf{x}_{ki}'$$

: matrix of total sums of squares and cross products

$$A = \sum_{k=1}^g N_k (\mathbf{m}_k - \mathbf{m})(\mathbf{m}_k - \mathbf{m})'$$

: matrix of among-groups sums of squares and cross products

$$W = \sum_{k=1}^g \sum_{i=1}^{N_k} (\mathbf{x}_{ki} - \mathbf{m}_k)(\mathbf{x}_{ki} - \mathbf{m}_k)'$$

: matrix of within-groups sums of squares and cross products

Then we have

$$T = A + W \quad (1)$$

The canonical discriminant function is determined in the following way. Let us denote the canonical discriminant functions as:

$$y_j = \mathbf{v}_j' \mathbf{x} \quad (2)$$

where

$$\mathbf{v}_j' \mathbf{v}_j = 1$$

Hereafter we abbreviate the subscripts of \mathbf{x} . The among-groups sum of squares on the function y_i will approximately be

$$\mathbf{v}_j' A \mathbf{v}_j$$

Similarly the within-groups sum of squares on the function y_j will approximately be

$$\mathbf{v}_j' W \mathbf{v}_j$$

Here we maximize the ratio

$$\lambda_j = \mathbf{v}_j' A \mathbf{v}_j / \mathbf{v}_j' W \mathbf{v}_j$$

and get

$$(W^{-1}A - \lambda_j I) \mathbf{v}_j = 0 \quad (3)$$

The rank of matrix $W^{-1}A$, in general, will be that of A . So we have n ($= \min [p, g-1]$) eigenvalues: let these be $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Let \mathbf{v}_1 be the normalized solution of the above equation with $\lambda_i = \lambda_1$, then the vector \mathbf{v}_1 points to the direction along which the among-groups difference will be the largest relative to the within-groups scatter. The vector \mathbf{v}_2 corresponding to $\lambda_i = \lambda_2$ is not perpendicular to \mathbf{v}_1 but along it the among-groups difference will be the largest among those uncorrelated with that along \mathbf{v}_1 , and so forth.

Let us define the vector \mathbf{z} of p components as:

$$X = \frac{1}{\sqrt{N-g}} \begin{pmatrix} W_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & W_{pp} \end{pmatrix}^{\frac{1}{2}} \sqrt{N-g} \mathbf{z} = \mathbf{D}_{diag}^{\frac{1}{2}} \sqrt{N-g} \mathbf{z}$$

where

$$N = \sum_{k=1}^g N_k$$

$$\mathbf{D}_{diag}^{\frac{1}{2}} = \frac{1}{\sqrt{N-g}} \begin{pmatrix} W_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & W_{pp} \end{pmatrix}^{\frac{1}{2}}$$

and W_{ij} is (i, j) element of matrix W .

Then we get

$$y_j = \mathbf{v}_j' \sqrt{N-g} \mathbf{D}_{diag}^{\frac{1}{2}} \mathbf{z} = \mathbf{c}_j' \mathbf{z} \quad (4)$$

where

$$\mathbf{c}_j' = \mathbf{v}_j' \sqrt{N-g} \mathbf{D}_{diag}^{\frac{1}{2}} \quad (5)$$

In general, $\mathbf{u}_j = \mathbf{v}_j \sqrt{N-g}$ is called the unstandardized discriminant function coefficient vector of p components, and $\mathbf{c}_j = \mathbf{D}_{diag}^{\frac{1}{2}} \mathbf{u}_j$ is called the standardized discriminant function coefficient vector (Klecka, 1980).

MATERIALS AND METHODS

Study area

In the grassland which extends around Linze, Gansu, the Lanzhou University Grassland Station and the Atmospheric Weather Station (100°04'E, 39°15'N, 1394 m a.s.l.) are established. The study area was selected near the stations (Fig. 1). Climatic values of annual precipitation, annual maximum and minimum air temperature at the AWS are 145 mm, 35.7°C and -32.4°C, respectively. This study area is located in the middle reaches of the Heihe River, which comes from the Qilianshan Mountains and disappears into the Badain Jaran Desert. Three 360×360 m test plots (representative sample sites or training areas) of various land cover types, bare-soil area (strongly saline), reed area (wet and moderately saline), and barley field (irrigated and weakly saline) are delineated as shown in Fig. 1.

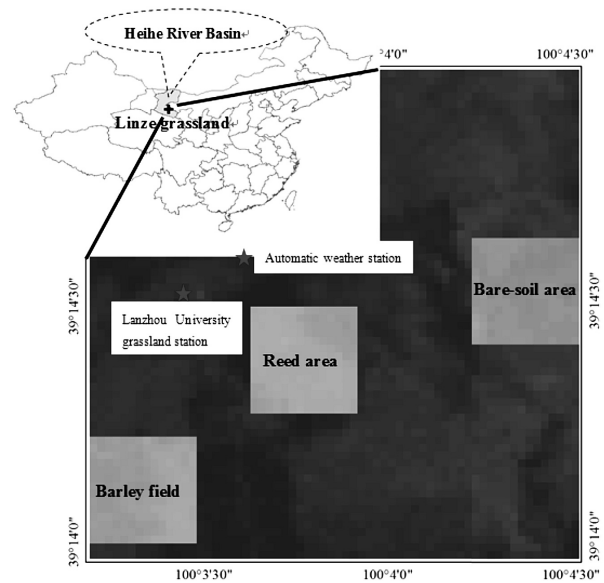


Fig. 1. Study area and polygons defining three test plots or training areas.

Data

Landsat-5 TM images obtained on Sep. 23, 2007 were used in this study. The geometric correction was implemented so that its RMS accuracy was controlled within half a pixel. The resolution of resampled pixel was about 30 m. In atmospheric correction, the mid-latitude summer model was used as the atmospheric profile, and the rural mode of aerosol model was applied. By resampling, 157 pixel values were obtained in and around each test plot, which is called the group in the statistical analysis

and hence 471 pixel values were used in this analysis.

Discriminating variables

Table 1 lists the seven spectral bands of the TM with descriptions of the application related to this study, extracted from the intended principal application of each (Lillesand *et al.*, 2004).

Fourteen candidates for discriminating variables which were devised by various researchers using Landsat TM data (Table 2), including the raw band digital num-

Table 1. Thematic mapper (TM) spectral bands

Band (Wave length, μm)	Nominal Spectral Location	Interested Application
1 (0.45–0.52)	Blue	Soil and vegetation discrimination, and cultural feature identification.
2 (0.52–0.60)	Green	Green reflectance peak of vegetation and vigor assessment, and cultural feature identification.
3 (0.63–0.69)	Red	Chlorophyll absorption region aiding in plant species differentiation, and cultural feature identification.
4 (0.76–0.90)	Near IR	Determining vegetation types, vigor, and biomass content, and soil moisture discrimination.
5 (1.55–1.75)	Mid IR	Vegetation moisture content and soil moisture discrimination.
6 (10.4–12.5)	Thermal IR	Vegetation stress analysis, and soil moisture discrimination.
7 (2.08–2.35)	Mid IR	Discrimination of mineral and rock types, and vegetation moisture content

Table 2. Candidates for discriminating variables (composite indices)

Composite index	Description
$SI = \sqrt{B1 \times B3}$	Salinity index (Khan <i>et al.</i> , 2005)
$NDSI = -NDVI$ $= (B3 - B4) / (B3 + B4)$	Normalized differential salinity index (Khan <i>et al.</i> , 2005)
$BI = \sqrt{B3^2 + B4^2}$	Brightness index (Kahn <i>et al.</i> , 2005)
$Fe_2O_3 = B3/B1$	Weathering index (Dmmeyer and Shukla, 1994)
$Albedo = 0.359B1 + 0.130B3 + 0.373B4$ $+ 0.085B5 + 0.072B7 - 0.018$	Albedo index (Liang, 2000)
<i>Brightness, Greenness, Wetness</i>	Tasseled cap concept extended to Landsat TM data (six bands). The first feature relates to soil reflectance, the second the amount of green vegetation, and the third the canopy and soil moisture (Crist and Cicone, 1984; Peng and Li, 1989)

Table 3. Correlation matrix between each pair of all candidates for discriminating variables for a test plot “Bare–soil area”

	B1	B2	B3	B4	B5	B7	Br*	Gr*	We*	Al*	SI	ND*	Fe2*	BI
B1	1													
B2	0.943	1												
B3	0.918	0.971	1											
B4	0.767	0.755	0.702	1										
B5	0.560	0.670	0.755	0.276	1									
B7	0.476	0.588	0.682	0.145	0.974	1								
Br*	0.828	0.905	0.949	0.593	0.914	0.858	1							
Gr*	–0.362	–0.441	–0.532	0.220	–0.751	–0.816	–0.625	1						
We*	–0.180	–0.301	–0.414	0.137	–0.898	–0.941	–0.654	0.789	1					
Al*	0.974	0.968	0.965	0.796	0.694	0.608	0.920	–0.399	–0.327	1				
SI	0.980	0.977	0.978	0.750	0.670	0.589	0.906	–0.455	–0.301	0.990	1			
ND*	0.715	0.795	0.871	0.267	0.826	0.814	0.873	–0.862	–0.644	0.758	0.808	1		
Fe2*	–0.258	0.017	0.146	–0.199	0.428	0.458	0.240	–0.381	–0.539	–0.082	–0.062	0.336	1	
BI	0.912	0.934	0.919	0.926	0.555	0.444	0.833	–0.161	–0.146	0.953	0.934	0.609	–0.034	1

*Br: *Brightness*; Gr: *Greenness*; We: *Wetness*; Al: *Albedo*, ND: *NDSI*, Fe2: Fe_2O_3

bers (DNs), were tested. Table 3 shows the correlation matrix between each pair of all candidates for discriminating variables using the data obtained in the bare-soil area, the size being 157 (Fig. 1). If two indices are highly correlated to each other, they convey essentially the same information, so it is appropriate to pick up indices that are not correlated to other ones. Fe_2O_3 is regarded as such an index. B4, NDSI and *Wetness* are also included in such indices, and B5, B7 and BI can be removed from the list of candidates. In the contrary, if we select SI as the variable, we can exclude such indices as B1, B2, B3, *Brightness*, and *Albedo*. Finally, *Greenness* seems to be a promising candidate but we judged that NDSI could be a substitute for it. As a result, we got five discriminating variables; that is, B4, *Wetness*, SI, NDSI, and Fe_2O_3 . Since $p = 5$ and $g = 3$ in this study, the number of eigenvalues or canonical discriminant functions is two, because $n = \min [p, g-1] = 2$.

RESULTS AND DISCUSSION

By solving Eq. 3 using 157 pixel values (cases) for each group and hence 471 cases in total, we got two eigenvalues. In this study, incidentally, statistical calculations were made using SPSS16.0 software.

Table 4 shows the eigenvalues and measures of importance. These results show that the first canonical discriminant function (y_1) contains 70% of the total discriminating power in this system of equations. Canonical correlation coefficient is defined as

$$\sqrt{\mathbf{v}_i' \mathbf{A} \mathbf{v}_i / \mathbf{v}_i' \mathbf{T} \mathbf{v}_i} = \sqrt{\lambda_i / (1 + \lambda_i)}$$

where i denotes the relevant canonical discriminant function (Eq.1). This coefficient is a measure of the degree of relatedness between the groups and the function. They are rather large, and hence the two canonical discriminant functions are expected to be powerful discriminators.

Table 4. Eigenvalues and measures of importance

Canonical Discriminant Function	Eigenvalue	Relative percentage	Canonical Correlation coefficient
y_1	4.851	70.0	0.911
y_2	2.080	30.0	0.822

Table 5 shows the standardized discriminant coefficients \mathbf{c}_j (Eq.5) and within-groups structure coefficients. A structure coefficient, or a coefficient of correlation between a discriminating variable and a canonical discriminant function, tells us how closely the variable and the function are related. “Within-groups structure coefficient” s_{ij} for variable i and function j can be calculated as follows:

$$s_{ij} = \sum_{k=1}^5 \frac{W_{ik} C_{kj}}{\sqrt{W_{ii} W_{kk}}}$$

where c_{kj} is k component of the vector of standardized

Table 5. Standardized discriminant function coefficients and within-groups structure coefficients

Discriminating Variable	Standardized Discriminant Function Coefficient (c_{ij})		Within-groups Structure Coefficient (s_{ij})	
	y_1	y_2	y_1	y_2
B4	-0.931	1.191	0.647	0.586
<i>Wetness</i>	0.657	-0.548	0.144	-0.612
SI	3.833	-1.378	0.632	0.511
NDSI	-2.590	0.338	0.338	0.443
Fe_2O_3	0.210	0.483	-0.198	0.640

discriminant function coefficient \mathbf{c}_j and W_{ij} is (i, j) element of matrix \mathbf{W} . From these results it can be deduced that discriminating variables B4 and SI contribute mainly to y_1 , while *Wetness* and Fe_2O_3 contribute mainly to y_2 . This suggests that function y_1 reflects the “vegetation plus salinity” conditions and function y_2 reflects the “soil plus wetness” conditions in each pixel.

Figure 2 shows the two-function plot of group centroids and individual cases. The groups are quite distinct and their centroids are well separated. For getting ground-truth data, we have only made measurements of EC_{SAT} (Kobayashi *et al.*, 2006) of the top soil layer (0–10 cm) at one site for each test plot on 29 June, 2011, which were 180 dS m⁻¹ in “Bare-soil area”, 20 dS m⁻¹ in “Reed area”, and 5 dS m⁻¹ in “Barley field”, and in addition to that the TM images used in this study were obtained on Sep. 23, 2007, when barley had already been harvested. Thus, further study is needed to collect ground-truth data contemporaneously to validate the spectral signatures from the satellite images. However, it turns out that the power of the two functions y_1 and y_2 for discriminating salt-affected soil surfaces is promising, and y_1 seems to reflect the difference in “vegetation plus salinity” conditions between test plots and y_2 the

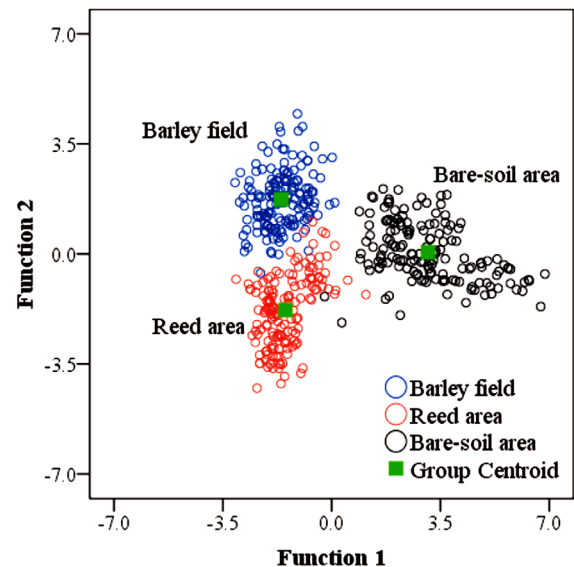


Fig. 2. Two-function plot of group centroids and individual cases. Function 1 and Function 2 represent y_1 and y_2 , respectively.

difference in “soil plus wetness” conditions between them. These estimations accord with the judgment made above from the within-groups structure coefficients.

CONCLUDING REMARKS

A preliminary study has been conducted to test the effectiveness of the canonical discriminant approach to making spatial discrimination of salt-affected soil surfaces, using the Landsat TM data taken in the middle reaches of the Heihe River, China. A saline area can be considered a mixture of salt-affected patches which differ in degree, and hence we tried to make discrimination of salt-affected soil surfaces by canonical discriminant approach. Three 360×360 m test plots with different land surface conditions, “Bare-soil surface area (strongly saline)”, “Reed area (moderately saline)”, “Barley field (weakly saline)”, were selected as training areas or groups. Five indices, B4 (band 4 DN), *Wetness*, SI, NDSI, and Fe_2O_3 (composite indices), were determined as discriminating variables, and two canonical discriminant functions, y_1 and y_2 , were obtained for the three groups using these five discriminating variables. When group centroids and individual cases were plotted on the plane made by two canonical discriminant functions, y_1 and y_2 , the groups of cases were quite distinct and their centroids were well separated, which suggest that the canonical discriminant approach is promising for spatial discrimination of salt-affected soil surfaces.

It can be said, however, that the following subjects which are related closely to this study require further investigation.

- (a) To collect more ground-truth data contemporaneously to validate the spectral signatures from the TM images.
- (b) To test the possibility of microwave sensing of salt-affected areas, because, although the real part of the dielectric constant is insensitive to the presence of salts but strongly responds to soil moisture, the imaginary part is highly sensitive to the salt content in the soil (Metternicht and Zinck, 2009), which seems to be ideal for detecting soil salinity.
- (c) To construct a model of a mixed pixel relating the spectral signatures from the satellite images and soil salinity. If we get such a model, remarkable progress in management practices for salt-affected soils could be accomplished.

ACKNOWLEDGEMENTS

This work is supported by the Project of National

Basic Research Program of China “973” (grant number: 2009CB421305), the key NSFC Project (National Science Foundation of China, grant number: 91125002), and the Intellectual Innovation Project from Chinese Academy of Sciences (grant number: KZCX2-EW-312). The Landsat TM data used in the paper were obtained from the Watershed Allied Telemetry Experimental Research (WATER).

REFERENCES

- Anderson, T. W. 1958 *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, Inc., New York (USA)
- Araki, T., D. Yasutake, W. Wang, Y. Wu, M. Mori, M. Kitano, H. Cho and T. Kobayashi 2011 Saline water seepage from drainage canals induces soil salinization and growth depression in the adjacent cornfields in the upper Yellow River Basin. *Environ. Control Biol.*, **49**: 127–132
- Cooley, W. W. and P. R. Lohnes 1971 *Multivariate Data Analysis*. John Wiley & Sons, Inc., New York (USA)
- Crist, E. P. and R. C. Ciccone 1984 Application of the tasseled cap concept to simulated thematic mapper data. *Photogram. Engineer. Remote Sens.*, **50**: 343–352
- Dmmeyer, P. A. and J. Shukla 1994 Albedo as a modulator of climate response to tropical deforestation. *J. Geophysical Res.*, **99**: 2863–2878
- Khan, N. M., V. V. Rastokuev, Y. Sato and S. Shiozawa 2005 Assessment hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agri. Water Manage.*, **77**: 96–109
- Kobayashi, T., W. Wang, Y. Ikawa, H. Cho and W. He 2006 An easily measurable and practical index of salinity. *J. Japan Soc. Hydrol. & Water Resour.*, **19**: 183–188
- Li, B. 2009 Soil salinization. In “Desertification and Its Control in China”, ed. by L. Ci and X. Yang, High Education Press, Beijing (China), pp. 265–298
- Liang, S. L. Narrowband to broadband conversions of land surface albedo: algorithms. *Remote Sens. Environ.*, **76**: 213–238
- Lillesand, T. M., R. W. Kiefer and J. W. Chipman 2004 *Remote Sensing and Image Interpretation, Fifth ed.* John Wiley & Sons, Inc., New York (USA)
- Lleca, W. R. 1980 *Discriminant Analysis*. SAGE Publications, Newbury Park (USA)
- Metternicht, G. and J. A. Zinck 2009 Remote Sensing of Soil Salinization. CRC Press, Boca Raton (USA)
- Peng, W. Z. and T. J. Li 1989 The effect of Kauth–Thomas transformation of TM data in the analysis of saline soil—An example in Yang–Gao Basin. *Remote Sens. Environ.*, **4**(3): 183–190
- Tanji, K. K. (ed.) 1990 *Agricultural Salinity Assessment and Management*. American Society of Civil Engineers, New York (USA)
- Wang, W., T. Kobayashi, D. Yasutake, M. Kitano, H. Cho, T. Araki and H. Yoshikoshi 2008 Experiments on the control of salinity and sodicity in surface-irrigated fields in the upper Yellow River valley (I) Objectives and Methodology. *J. Fac. Agr., Kyushu Univ.*, **53**: 251–256