

Inversion classifications of saline-alkali lands based on the hyperspectral data from the environmental mitigation Satellite

Song Wu^{1, a *} and Qigang Jiang^{1, b}

¹ Jilin University, No.2699, Qianjindajie Street, Changchun, Jilin, China

^a158414215@qq.com, ^bmikewoo55555@163.com,

Keywords: Remote sensing, regression prediction, the least squares support vector machine (SVM), saline-alkali land, Songliao basin, Daqing

Abstract. This study explores the best model for the quantitative inversion of soil salinity index in an area, and adapts the least squares support vector machines (LS - SVM) regression forecast method. A variety of index inversions is made in serious saline-alkali land in Daqing area, and decision-making method of binary tree is adopted to classify the saline-alkali by taking Songliao basin as an example. To have effective preventions and controls of saline-alkali land, based on hyperspectral data of the environmental mitigation Satellite(HJ - 1 A) we contrast the results of predictions between curve regression and least squares support vector machine (SVM) regression which are the two nonlinear regression modes in the prediction effects of salt content inversion. The results shows that the environmental mitigation satellite can be effectively used on the extraction about information saline-alkali land. The inversion model based on least squares support vector machine (SVM) is with highest accuracy; In RS, the method of decision binary tree can effectively be supported by using classifies saline-alkali lands. The results are accurate and reliable; Research shows that The salinization of Daqing area is serious, because the most parts of lands are alkalized soil, including mild alkaline lands , 345.03 km²、 moderate alkaline lands , 1389.03 km²、 and severe alkaline lands , 869.94 km². The research has great significance for the prevention and cure of saline-alkali land and fast extraction information of saline-alkali land.

Introduction

The salinization of lands , a kind of environmental disasters, leads to the degradation of lands, with weakening and destroying their productivities, and agricultural food productions, which is a serious threat to the ecology and the sustainable development of national economy [1-2]. According to incomplete statistics, from the United Nations Educational, Scientific, and Cultural Organization and the food and agriculture organization. The saline-alkali land area is 954.38 million hm², including China for 99.13 million hm². Most of the formations about alkali soil and alkaline soil in China, are associated with the cumulative of carbonates in the soil. Thus the degree of alkalization is generally high. Plants can hardly survive in the region with severe saline-alkali soil. Songliao basin, the salinization of land, gives priority to the soda alkali-saline land, and belongs to semi-humid and semi-arid grasslands meadows-, which is one of the eight saline area in northeast China. In view of the fine spectral resolution of hyperspectral data reflecting features between the nuances of the spectrum, it can make possible for the chemical composition of the remote sensing quantitative analysis [3] Ben-Dor and other people use stepwise regression analysis method, hyperspectral data DAIS – 7915, in the analysis of characteristics of northern Israel on Iran basin including the quantitative graphics about soil moisture, soil conductivity EC organic matter, and pH values [4]. Farifteh and other people use hyperspectral data to have soil classification of salinization. They think they can well distinguish saline-alkali lands and non-saline-alkali lands through the features of the spectral characteristics. Then they put forward the use of PLSR and ANN methods to predict the linear and nonlinear relationship between soil content and salt content, from the standard spectral reflectance from two adaptive methods. [5]. When GuanYuanxiu and other people has the remote sensing investigation of saline-alkali lands on the Yellow River delta, they use the embedded depth of groundwater, soil salt content and normalized difference vegetation index to optimize the supervised and unsupervised classification.

They think they can distinguish salinization degree of saline soils [6].Yong lingweng use Hyperion hyperspectral data, with least-squares regression method to establish relation between the percentage of soil salt index and the soil salt content which has obtained the certain effect [7].

Scholars at home and abroad have made significant achievements in hyperspectral remote sensing data of quantitative inversion [4-9] .But there are rare studies about the environmental mitigation Satellite(HJ - 1 A)applied to the identification and monitoring of saline-alkali research . At the same time ,there are also less studies about using support vector machine (SVM) to make the quantitative inversion of saline-alkali lands. The spectral characteristics of saline-alkali lands are showed in this study on the basis of field investigation and the soil sample gathered in Songliao basin, from the spectral analysis, spectral and correlation analysis about soil physical and chemical characteristics, in Songliao basin. The sensitive wave bands are excavated and identified in saline-alkali lands. The perfect model is explored for the inversion of soil salinity index and a quantitative estimate is established. Finally we realize the quantitative estimates of the spatial distribution for soil salinity index content of environmental mitigation Satellite(HJ - 1 A) are available to extract quantitative information of soil salinity rapidly and economically and to provide basis for forecasting regional soil salinity conditions, formulating comprehensive treatment measures, and policies on using lands reasonable.

The definition and overview of test site

The concept of saline-alkali lands

The Saline-alkali land is the generic terms for all kinds of the solonchak and alkaline soil. Solonchak is refers to the soluble salt content in the soil which reaches significant harm to the crop growth. Salt content indexes are different from different content of salt. Alkali soil is refers to the soil contains sodium that is harmful to the plant growth and the exchangeable of soil properties. Saline-alkali lands are mainly distributed in arid and semi-arid areas, coastal areas. The saline-alkali land can be divided into salt, saline soil, alkaline earth and alkaline soils in our country according to the different of salt and its effect on soil properties[10-11].

The quantitative classification for saline soil

According to the different influences of salt composition and other soil properties, the Chinese usually regards Electrical conductivity (EC) ,the percentage of exchangeable sodium (ESP) and pH value as standard for the division of the soil salinity degree [11].The general accepted quantitative indicators are shown in table 1.

Table 1 Quantitative standard saline soil classification

Type of soils	The percentage of exchangeable sodium (ESP)/%	Electrical conductivity (EC)/(ms·cm ⁻¹)	pH value
saline soil	<15	>4	<8.5
alkaline soil	>15	<4	>8.5
saline-alkali lands	>15	>4	<8.5
Non-saline-alkali lands	>15	<4	>8.5

In the above indexes, pH and EC values are relatively easy to measure, due to the determination of ESP tedious and is wrong easily , the sodium adsorption ratio (SAR) and the ratio of exchangeable sodium (ESR) of the test are much more easily.The relationship between the ESR and the ESP can be estimated according to Gapon - Bolt ion exchange equation [13-14] :

$$ESP/(1-ESP)=KG \cdot SAR=ESR. \quad (1)$$

In above formula KG is proportional constant. We can have this formula

$$ESP = ESR / (1 + ESR) \quad (2)$$

We can get ESP from the testing result of ESR from second formula (2).

Introductions for experimental zones

Experimental zone is located in the west of Heilongjiang province, in Daqing which is the central part of Songliao basin (121°24' - 125° E, 45°46' - 46°55' N). And located in the north temperate zone continental monsoon climate zone, which is effected by the cold air and ocean current monsoon, with snow in winter, and the monsoon in the spring and autumn period, the annual average temperature is 2 ~ 6 °C and the annual average rainfall is 100 ~ 600 mm in the region. However, 1200 ~ 1800 mm, annual evaporation which is far more than the annual precipitation [15]. Daqing city

Analysis and the method for materials

The quantitative inversion method of Saline-alkali lands

Sample analysis of Saline-alkali soil

As you can see from figure 1, the sample point is a strong representative factor with wide ranges. After having the chemical analysis of the soil samples, it shows that in the salt of the soil, the anionic is HCO_3^- , mainly, accounting for 65.03% of the total anion and the mean is 4.066 g/kg. And Cl^- takes the second place, accounting for 12.38% of the total anion, and the mean is 0.774 g/kg. Cation is given priority to Na^+ , accounting for 71.45% of total cation and the mean value is 2.24 g/kg; Ca^{2+} takes the second place, accounting for 18.71% of the total cation, with the mean is 0.589 g/kg. The high pH in the soil, salt content and its coefficient of variation (Cv) indicate that the degree of the surface soil salinization in the study area and the total salt content of surface soil are relative high. the difference between overall samples is bigger. The Saline-alkali land in Songliao basin belongs to inland saline-alkali land like soda. Salt is mainly from soda (Na_2CO_3) and sodium bicarbonate (NaHCO_3),

the introduction of HJ - 1 A and the preprocessing of HSI data

HJ-1A/1B, the first ecological monitoring of environmental satellite, is launched by our country, which includes the HJ - 1 a star with a CCD camera and a hyperspectral imager (HSI). The hyperspectral imager can now completely for 50 km wide, the ground pixel resolution are 100 m, 110 ~ 128 push-broom images of spectral spectrum, which has the ability of side-view in the $\pm 30^\circ$ and the scaling function on the star and the revisit cycle is 4 d, and can observe atmosphere, soil, vegetation and water quality condition quickly and widely, which provides the decision-making basis for the environmental protection [16]. The hyperspectral image data of environmental mitigation Satellite (HJ - 1 A) are mainly concentrated in the range of visible light and near infrared, Its high image quality can show the advantages of high spectral resolution.

1) HDF5 format conversion

The data collected in this study are most from the products of hyperspectral data in level 2, of HJ - 1 A in 2010. And the places which data missing are reinforced by the data in 2009 and 2011 HDF5 format is read with IDL plug-in, and then is converted into common data format.

2) Transition for brightness of standard radiation

The product of hyperspectral data in level 2 of HJ-1A are the brightness products recovered and corrected by the spectrum, radiation calibration and rough radiometric calibration which have absolute calibration. The pixel value must be converted into brightness of standard radiation, the real value into the pupil radiation, when it is used, due to the products is made by the inflation factor to 100.

$$L_\lambda = DN / 100. \quad (3)$$

In above formula L_λ is true radiation brightness values, $\text{W/m}^2 \cdot \text{sr} \cdot \mu\text{m}$; DN is original image detection value, $\text{W/m}^2 \cdot \text{sr} \cdot \mu\text{m}$.

3) Atmospheric corrections

FLAASH model has atmospheric correction for the data of HJ - 1 a with ENVI hyperspectral data. After the atmospheric correction, the reflectivity has no obvious absorption characteristics from atmosphere, water vapor etc., which can better reflect surface features real spectral features.

4) Eliminate the disturbances

Eliminate the necessary disturbances to the atmospheric correction in the image about the clouds shadow, vegetation, water, etc and get the relatively bare images which are convenient the next work.

Analysis of Spectral characteristics

A great differences between the spectrum characteristics of different features exacted in HJ - 1 a hyperspectral images, such as spectral curve, bare soil, saline-alkali land, vegetation spectral differences. It helps to classify different features. Saline soils formed obvious absorption valley in the vicinity of 490, 700, 850 - nm, and formed reflection peak near 590, 750 nm, and also can be used as the band to diagnostic other land, which shows the evidence in related research [7].

The establishment and comparison of the estimation model for the soil alkaline index

Taking the quantitatively estimated HJ - 1, a hyperspectral remote sensing data of soil salt content, as an example, the two regression forecasting models, curve regression and least squares support vector regression, are established respectively. First, the statistical test is carried on, including the degree of fitting test, correlation coefficient test, parameter equation and significance test, to determine the accepted regression equation. Then, the model is evaluated and forecasted through the root mean square error (RMSE) and accuracy (PA).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (4)$$

$$PA = (1 - \frac{RMSE}{mean}) \times 100\% \quad (5)$$

Which x_i is predictive value, \hat{x}_i the measured value, the mean. the scheme for the prediction.

In the 26 tests of soil samples, the 24 tests are divided into two groups except the two outliers caused by artificial operation or machine operation. The first group of 18 measured soil samples is used for statistical analysis about salt content, to set up the inversion model; The second group of 6 soil samples has been established for inspection of inversion model. Modeling and test samples were selected in different regions and the remote sensing images also have changes showing different features of the space and time.

Support Vector machine SVM, a pattern recognition method based on statistical learning theory, shows a lot of unique advantages on solving the small sample, recognition problem about nonlinear and high dimensional pattern. the Support Vector machine (SVM) is a very powerful way in solving the problem of nonlinear function estimation. Least squares support vector machines (LS - SVM) proposed by Suykens and other people is a improved support vector machine (SVM) model. The two-norm is used in its objective function of optimization problem from the loss function of machine learning, and the LS - SVM method which can solve optimization problem is replaced by Kuhn - Tucker conditions for solving a set of linear equations with the equality constraint conditions instead of inequality constraints in the standard SVM algorithm.

The least squares support vector machines can provide a new method and means for inversion of soil salt content. In order to establish a LS - SVM model (with gaussian RBF kernel) two parameters are needed: γ , the regular parameter, which depends on the trade-off between the training error minimization and its smooth. σ^2 is square bandwidth. In addition, the study has analyzed the principal component to the information about original data, given by too much high spectral feature information, which can increase the complexity of analysis. The first four principal components has brought together

most of the information about original data, which can be used for modeling. The LS - SVM model is allowed as the regression modeling with salt rate as the dependent variable, and the main component after the principal component analysis as the independent variable. The results showed that the best effect is the fourth principal component modeling. Fig.1 is the fourth component of salt content and regression modeling effect for principal component. The determination value for regression model γ is 944.7819, the decision value for σ^2 is 0.038955, the regression model of determination coefficient is 0.9217, model validation accuracy is 0.8427, accuracy reached 80.94%, the sample inspection accuracy is 79.81 (table 2), the model can be used to the inversion on the spot.

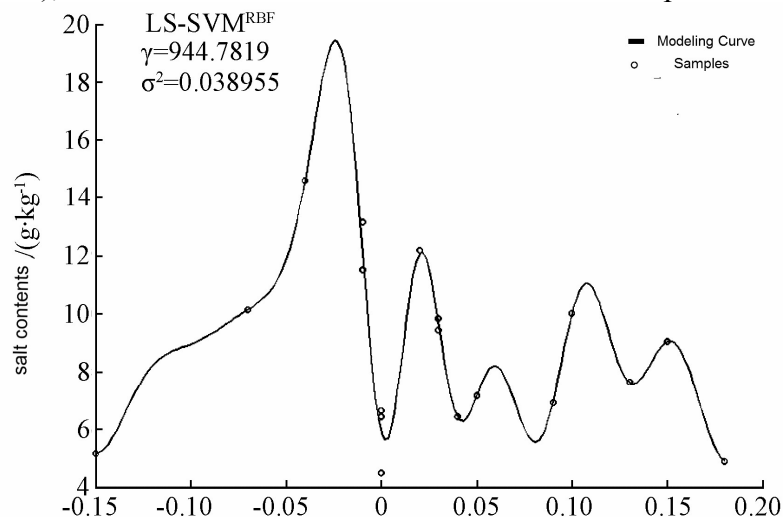


Fig.1 The LS-SVM regression modeling of salt contents and the fourth component of principal component analysis

Table 2 The result of LS-SVM regression modeling and testing by Salt contents and the fourth component of principal component analysis

spectrum transform	Selected band	coefficient of regression model	determination coefficient for modeling	RMSE / PA/%
PCA	Principal component	LS-SVM	0.9217	0.8427
samples measured value	1	2	3	4
predicted value	5	6	7	8
	6.79	10.53	11.49	10.01
	7.75	9.56	11.01	10.49
	10.96	9.29	10.01	8.58
	1.5391	80.94	79.81	

Results of inversion and the analysis

Through the contrast in last section of various modelings on the effect of salt content and soil reflectance, we choose the method of least squares support vector machine (SVM) regression modeling to test the salinization degree of pH, EC, the ESR and so on ,has carried on the quantitative inversion of their quantitative indicators, and has an quantitative exacted and scientific division to the comprehensive indexes, which has achieved the salinization classification to the area.

The regression modeling by LS-SVM and the classification about the result of inversion are presented with the dependent variables, ESR PH and EC, as well as the variable which is the major components.

Results

1) Based on the chemical analysis of the soil in Suliao basin, through the comparison between kinds of regression models, The support vector machine regression (SVMR) on the basis of the least squares method is identified as the best model for the index quantitative inversion of saline soil. The prediction is accuracy and reliable, and the regression effect is best. The accuracy of average inversion is 75.05%, and the precision is improved 14.49% than cubic curvilinear regression. At the same time, the soil in Da Qing region is made the classification according to the quantification standard of salinization, the test shows the distribution of alkali in this area.

2) The advantage of its high spectral spectrum is given full play by using HJ – 1, a hyper spectral remote sensing data, for the quantitative inversion of salinization quantitative indicators, The range of its data is wide, and the source can be downloaded for free in China Resources Satellite Application Center, which can be effectively applied in the quantitative research of salinization.

3)The quantitative model of salinization from this research can be used extendedly to the whole of the songliao basin quantitative inversion .The research has the important practical significance to the fast extraction of information about saline-alkali land in songliao basin, and the prevention and control of saline-alkali land.

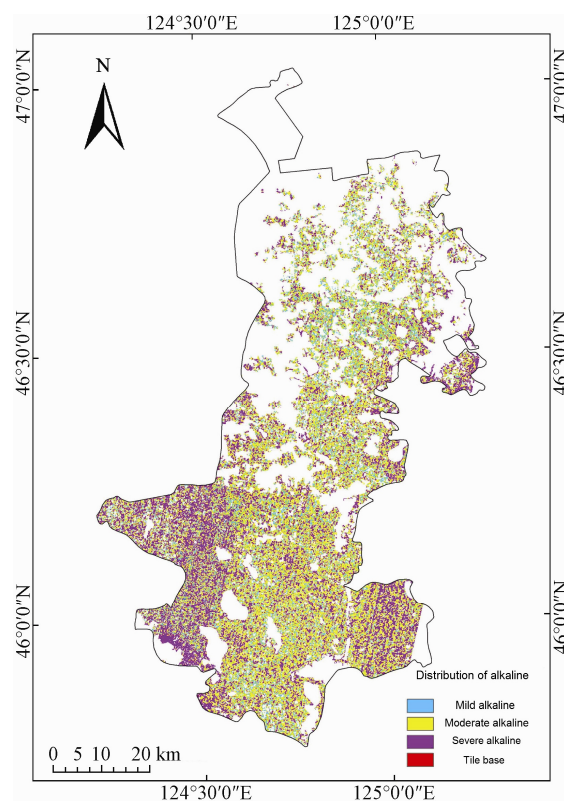


Fig.2 The spatial distribution of alkaline saline soil in test area

Acknowledgement

This work was financially supported by the Program of China Geological Survey (No. 201511078-1).

References

[1] Dehaan, R L, Taylor G R. Field-derived spectra of salinized soils and vegetation as indicators of irrigation induced soil salinization[J]. Remote sensing of Environment, 2002, 80(3): 406—417.

Reference to a book:

[2] Metternicht G I, Zinck J A. Remote sensing of soil salinity: potentials and constraints[J]. Remote Sensing of Environment, 2003, 85(1): 1—20.

[3] Goetz A F H, Vane G, Solomon J E, et al. Imaging Spectroscopy for Earth Remote Sensing[J]. Science, 1985, 228: 1147—1153.

[4] Ben-Dor R, Patkin A, Banin A, et al. Mapping of Several Soil Properties Using DAIS-7915 Hyperspectral Scanner Data: a Case Study over Clayey Soils in Israel[J]. International Journal of Remote Sensing, 2002, 23(6): 1043—1062.

[5] Farifteh J, Vander Meer F, et al. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods(PLSR and ANN). Remote Sensing of Enviroment, 2006, 7(2): 1—20.

[6] Guan Yuanxiu, Liu Gaohuan, Wang Jinfeng. Regionalization of salt-affected land for amelioration in the yellow river delta based on GIS[J]. Acta geographica sinica, 2001, 56(2): 198—205. (in Chinese with English abstract)

[7] Weng Yongling. Study on estimation of salt content in soils of the yellow river delta by space-borne hyperspectral remote sensing[D]. Nanjing: Nanjing University, 2007. (in Chinese with English abstract)

[8] Cheng Bin. Study on Quantitative Retrieval of Soil Organic Matter and Related Elements in the Black Soil in Songliao Plain Based on RS Data[D]. Changchun: Jilin University, 2007. (in Chinese with English abstract)

[9] Xing Yu, Jiang Qigang, Li Yuanhua, et al. Extracting phaeozem information in Songliao plain based on remote sensing and GIS[J]. Transactions of the CSAE, 2010, 26(10): 212—217. (in Chinese with English abstract)

[10] Russel J C. Sodic soil in Iraq[C]//In: Szabolics I. Processdings of the symposium on sodic soils. Budapest, 1964.

[11] Bolt G H, Page A L. Ion-exchange equations based on double-layer theory[J]. Soil Science, 1965, 99: 357—361.

[12] Li Xiaogang, Cao Jing, Li Fengmin, et al. Influence of salinity, sodicity and organic matter on some physical properties of salt-affected soils[J]. Chinese Journal of Soil Science, 2004(1): 64—70. (in Chinese with English abstract)

[13] Liu Dongxing, Gong Weiguang. Evaluation of site quality for saline-alkali soil in Daqing[J]. Science Soil and Water Conservation, 2009, 3(3): 98—103. (in Chinese with English abstract)

[14] Li Chuanrong, Jia Yuanyuan, Hu Jian. An analysis of the prospects of HJ-1 optical satellites in remote sensing application[J]. Remote Sensing for Land and Resources, 2008, 9(3): 1—9. (in Chinese with English abstract)

[15] Wang Gengming. Resarch on Remote Sensing Inversion of Salt Conent in Sline-Alkali Soil Based on BP Neural Network in Songliao Plain[D]. Changchun: Jilin University, 2007. (in Chinese with English abstract)

[16] Suykens J A K, Vandewalle J. Least squares support vector machine classfiers[J]. Neural Processing Letter, 1999, 9(3): 293—300. [5] Farifteh J, Vander Meer F, et al. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods(PLSR and ANN). Remote Sensing of Enviroment, 2006, 7(2): 1—20.

