

A Systematical Model to Search for the Main Factors Controlling the Spatial Distribution of Artificial Oasis and Inartificial Oasis in the Arid Area

Jing Wang^{1*}, Hailong Liu²

¹Xinjiang Institute of Water Resources and Hydropower Research, Urumqi, China

²School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu, China

Email: *2399743736@qq.com, liu_hai_tiger@163.com

How to cite this paper: Wang, J., & Liu, H. L. (2020). A Systematical Model to Search for the Main Factors Controlling the Spatial Distribution of Artificial Oasis and Inartificial Oasis in the Arid Area. *Journal of Geoscience and Environment Protection*, 8, 255-275. <https://doi.org/10.4236/gep.2020.85017>

Received: April 17, 2020

Accepted: May 24, 2020

Published: May 27, 2020

Abstract

The distribution and the affecting factors of the artificial oasis and inartificial oasis have become a serious and widespread problem in arid lands. Understanding its controlling factors is vital for environmental governance, improvement, and optimization. The study aimed to identify the crucial factors affecting the distribution of artificial oasis and inartificial oasis in arid land through the NDTG (the union of deep learning method, the modified a three-band maximal gradient method, Geodetector method) Model. Environmental factors include meteorological factors, chemical compositions, salinities, groundwater depth and time-series of Landsat images. The results show that Salinity factor was the dominant factor which explained 32.95% of the spatial variation of the artificial oasis distribution. Nonlinear enhancements were observed for the interactions between salt content and Evaporation ($q = 0.93$), salt content and Precipitation ($q = 0.78$). It indicated that Meteorological factors, and Salinity were the main factors determining the spatial pattern of the artificial oasis distribution. Salt, precipitation, evaporation, Mg, Cl, Na explained 37%, 26%, 25%, 24%, 23%, 20% of the spatial pattern of the inartificial oasis in arid lands, respectively. The results indicated that salinity, meteorological factors and chemical composition were the main factors determining the spatial distribution of inartificial oasis in arid lands. Moreover, the NDTG Model provided evidence to explore the factors controlling spatial patterns of the distribution of artificial oasis and inartificial oasis in arid lands.

Keywords

The Arid Lands, Multiple Sources Data, The Crucial Factors, The NDTG Model

1. Introduction

The arid inland plain of northwest China has rare rainfall, dry climate, strong surface evaporation, extremely weak recharge of groundwater by atmospheric precipitation, shortage of water resources, serious soil erosion and extremely fragile ecological environment (Yang et al., 2018; Akiyama et al., 2018). The increasingly prominent contradiction between human-water-land and ecological environment has intensified the deterioration of ecological environment, and affected the sustainable development of social economy in northwest China. The changes of oasis in arid lands are the most common and dangerous natural factors in ecologically fragile areas, which leads to serious damage to people's lives and property (Petley, 2012). In the arid regions of northwest of China, desertification is becoming more and more serious. Especially in the remote areas of arid regions, oasis distribution and the problem of water resources distribution is very prominent. The distribution of oasis and environmental factors has become the biggest constraints to socio-economic development. Therefore, it is of great significance for the protection and development of artificial oasis, oasis and brackish water resources to understand accurately and in detail the spatial distribution.

Generally, the identification and detection of oasis area in arid area of desert and the evaluation of main factors affecting oasis area in arid area are of great significance to local social and economic development. However, there is no systematical model to search for the main factors controlling the spatial distribution of artificial oasis and inartificial oasis in the arid area. Previous studies pay more attention to the accuracy of one single algorithm. We choose some studies showed that the classification method of deep learning is fairly effective when classifying artificial oasis (Petley, 2012; Semwal et al., 2017; Yabuki et al., 2018; Garg et al., 2018; Wang et al., 2019; Yu et al., 2017; Fawaz et al., 2019; Tyr et al., 2018). The classification method of Modified a three-band maximal gradient difference Algorithm (TGDVI) originating from the field of band math is more effective when we classified the inartificial oasis (Gutman & Ignalov 1998; Tang et al., 2003; Jiapaer et al., 2011). Remote sensing images are often used to generate oasis indexes (Huo et al., 2010; Wilford et al., 2016). Moreover, land-use and vegetation indexes change over time, especially in rapidly changing semi-arid and arid lands. Using images from a specific time may not reflect the effects of temporal variation in these factors on the accumulation of the distribution of artificial oasis and natural oasis. Therefore, the Landsat series provides image records of the Earth's surface for more than 5 years, and these images are suitable for extracting long term artificial vegetation indexes and land-use information. Principal component analysis (PCA) and cluster analysis (CA) are applied to assist in the identification of environmental factors controlling oasis accumulation (Shi et al., 2018; Chen et al., 1997; Guo et al., 2012; Lee et al., 2006; Li et al., 2004; Ordonez et al., 2003; Sun et al., 2013; Lin et al., 2002; Navas & Machin 2002; Luo et al., 2015). A geographical detector method, namely geodetector, may be a better choice for exploring the factors controlling artificial oasis and

inartificial oasis accumulation in arid lands. Geodetector is based on the spatial stratified heterogeneity of geographical phenomena; its key underlying assumption is that if a geographical factor A is controlled by another geographical factor B, then B will exhibit a spatial distribution similar to that of A (Wang et al., 2010; Wang et al., 2016; Li et al., 2013). Geodetector has been applied to analyze the factors controlling the spatial patterns of various geographical phenomena. For instance, Luo et al. (2015) employed geodetector to identify the dominant factors of dissection density over the entire conterminous United States (Wang et al., 2010), and Li et al. (2004) applied geodetector to investigate the spatial relationship between planting patterns and residual fluoroquinolones in soil.

In our study, it can be found that oasis is divided into two parts: artificial oasis and inartificial oasis. An artificial oasis is defined as human cropland and other oasis places that are made by humans. Except artificial oasis, we define other vegetation land as natural oasis. Given the importance of Water to Oasis in Arid lands, this study aimed to identify the factors controlling Oasis using the NDTG Model, especially time-series of remotely sensed data. The results of this study are expected to reveal the internal regularities affecting the spatial pattern of artificial oasis and natural oasis in Arid lands, and to provide a geo-statistical way to explore their factors controlling these spatial patterns by combining geographical information science and remote sensing methods.

2. Materials and Methods

2.1. Study Area

The Alar Basin is located in the Tarim River watershed, Alar County, XinJiang Province, China (**Figure 1**). Its geographic coverage is approximately 40°22' - 40°57'N and 80°30' - 81°58'E, with an area of 4161 km², including Alar City (9th

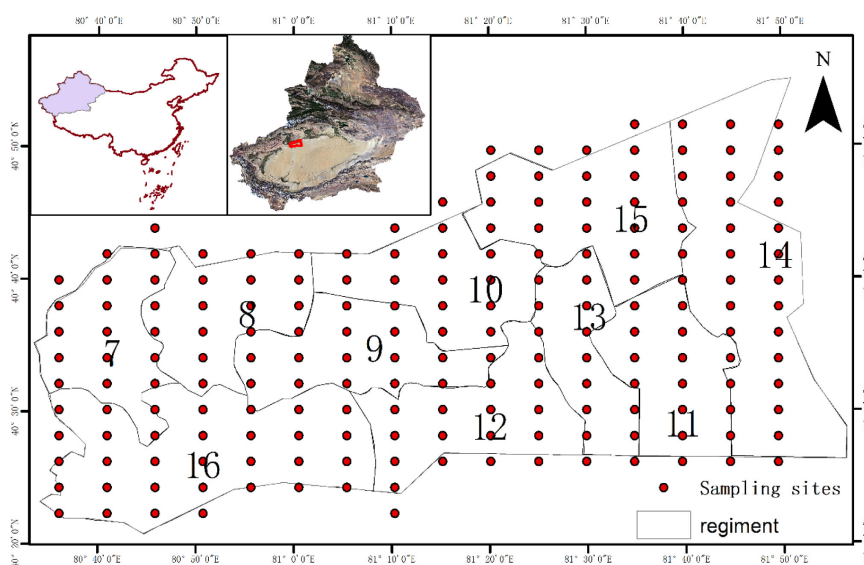


Figure 1. Study area and sampling locations.

Regiment, 7th Regiment, 8th Regiment, 10th Regiment, 11th Regiment, 12th Regiment, 13th Regiment, 14th Regiment, 15th Regiment and 16th Regiment). The distribution diagram of the study area is shown in the figure.

2.2. Water Sampling

This study concentrated on the 8, 9, 10, 11, 12, 13, 14, 15 and 16 regiments in the Alar region on the upper reaches of the Tarim River. The study area was divided into regular grids of 2 × 2 km for sampling, and a sampling site was randomly selected in each grid. The geographical coordinates of sampling sites were recorded using a global positioning system receiver, and information on land-use, oasis cover, and landform were also recorded. A total of 163 underground soil samples and underground water samples were collected in September 2017 (Figure 1), and artificial deposits, such as rubbles, concrete debris, and wastes, were avoided. Approximately 1.0 litre of water in wells were collected during two sampling campaigns(Tang et al., 2003) after removing plant residues, roots, and stones. The collected soil samples were kept in polyethylene bags and brought to a laboratory for Chemical Composition and Salinity content analysis.

2.3. Data Source and Processing

The main data used in this research include Landsat 8 image data of Alar region from 2013 to 2018, the groundwater quality and quantity, the vector map of Alar administrative boundary and other spatial data. Using remote sensing image data carry out supervised classification with deep learning law and Modified a three-band maximal gradient difference model (TGDVI), spatial registration and mosaic with reference to topographic map, and using ENIV carry out image interpretation to obtain the land use classification map of artificial oasis and in-artificial oasis from 1990 to 2015. Combined with the correct measured data and the actual sampling point data, the image interpretation accuracy of the six years is above 85%, which can meet the analyzing requirements with the domain factors including Chemical Composition (sodium ion, magnesium ion, calcium ion, potassium ion, sulfate radical, chloride ion, bicarbonate radical), Salinity (salt content and mineralization degree), Groundwater (groundwater depth) (Table 1).

Table 1. Environmental factors for analyzing controlling factors of soil heavy metals.

Environmental factors		Data source
Meteorological data	Evaporation, Precipitation	https://www.ecmwf.int/
Chemical Composition	sodium ion, magnesium ion, calcium ion, potassium ion, sulfate radical, chloride ion, bicarbonate radical	Measured data
Salinity	salt content and mineralization degree	
Groundwater	groundwater depth	

Due to the fact that evaporation data were not collected easily during the actual sampling process. Relevant research found that ERA-interim data was used to replace the measured rainfall and evaporation in this study. Rainfall and evaporation data can well reflect the measured rainfall and evaporation data, as shown above. Scholars have verified the usability of the small areas by analyzing the small areas of the data (Hodges et al., 2010; Sylla et al., 2010; Weedon et al., 2015), so this paper also chooses them as the parameter input items.

Meteorological stations in catchments are fairly few in arid regions like Tarim River basin due to the high and steep terrain, backward economic and technological situations. Besides, most of the stations are located in the mountain-passes. Even in the irrigation districts, meteorological stations observing temperature and precipitation for more than 30 years are few. Therefore, the ERA-Interim reanalysis dataset was selected as the substitution due to its good performance (Hodges et al., 2010; Sylla et al., 2010; Weedon et al., 2015). Its spatial and temporal resolution is $0.125^{\circ} \times 0.125^{\circ}$.

2.3.1. The NDTG Model Construction

In this paper, the NDTG model (**Figure 2**) was constructed according to the deep learning method, the modified a three-band maximal gradient method. Then the mathematical model of NDTG was generalized as follows.

1) Extraction of Artificial Oasis by Deep Learning Algorithm

The deep learning Algorithm selected in this paper is built based on Tensor Flow framework. Through the sample set construction of data, the identification and testing of test set and verification set based on depth learning, and through the selection and testing of classifiers, the activation function with the best segmentation effect for artificial oasis vegetation in arid areas is found. The optimal model is obtained through several trainings by modifying parameters (i.e. layer number, epoch and other parameters) of the training model (**Figure 3**).

2) Extraction of Inartificial oasis by modified a three-band maximal gradient difference Algorithm (TGDVI)

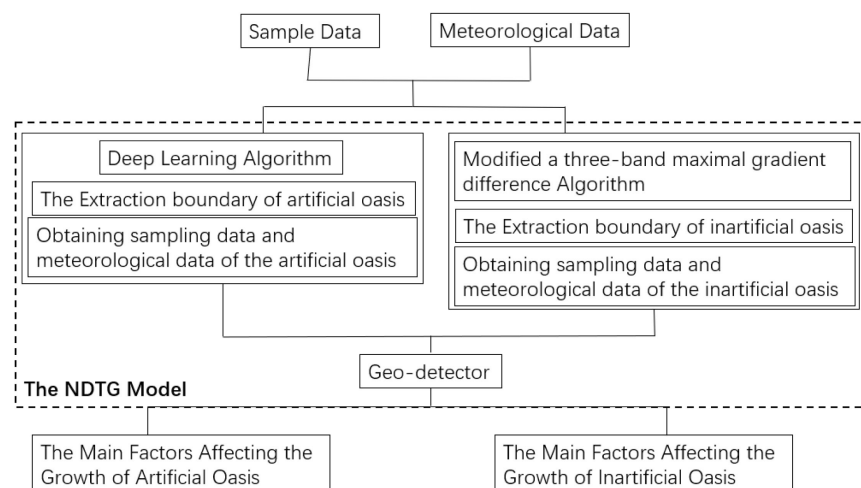


Figure 2. The NDTG model construction.

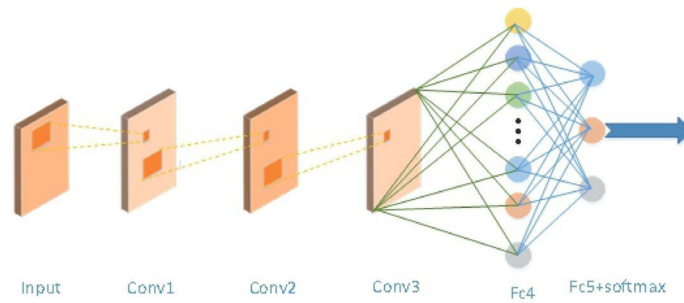


Figure 3. The deep learning algorithm construction.

The difference between oasis and non-oasis is most significant in the wavelengths SWIR 1650 nm and SWIR 2220 nm, NIR 830 nm and red 660 nm, suggesting that green wavelengths can be replaced by shortwave infrared wavelengths of 1650 nm or 2220 nm. The TGDVI is adjusted as follows:

$$A = \frac{d}{d_{\max}}, d = \frac{R_{ir} - R_r}{\gamma_{ir} - \gamma_r} - \frac{R_r - R_g}{\gamma_r - \gamma_g} \text{ if } A < 0, A = 0 \quad (1)$$

where R_{ir} , R_r and R_{swir} are defined as the reflectivity in the NIR, red and SWIR bands, respectively, and ir , r and $swir$ are the wavelengths of the NIR, red and SWIR bands, d_{veg} stands for the d of highly dense oasis.

3) The geographical detector

The geographical detector method was developed to measure the spatially stratified heterogeneity of geographic variable Y (for example the Normalized Difference Oasis Index (NDVI) of Artificial Oasis and Inartificial oasis in this study) and to explore how factor X explains the spatial pattern of Y . Formally, Y is composed of N samples, and X is stratified into L strata; stratum $h \in [1, 2, \dots, L]$ is composed of N_h samples; y_i is the value of sample i in the whole sample population; and y_{hi} denotes the value of sample i in stratum h . The concept of spatially stratified heterogeneity is the q-statistic, which is defined as follows (Li et al., 2013):

$$q = 1 - \frac{\sum_{h=1}^L \sum_{i=1}^{N_h} (y_{hi} - \bar{y}_h)^2}{N\sigma^2} = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N\sigma^2} = 1 - \frac{SSW}{SST} \quad (2)$$

$$SSW = \sum_{h=1}^L \sigma_h^2 N_h \quad (3)$$

$$SST = N\sigma^2 \quad (4)$$

where: $h = 1, \dots, L$ is the Strata of variable Y or factor X , and so on. Which may be classification or partition; N_h and N are the unit numbers of layer H and the whole region, respectively; σ_h^2 and σ^2 are the variances of the whole region (Y) and values of layer h , respectively. SSW and SST are the sum of intra-layer variances (Within Sum of Squares) and Total Sum of Squares, respectively. q has a range of $[0, 1]$.

The larger the value, the more obvious the spatial difference of Y is. If the stratification is generated by the independent variable X , the larger the q value, the stronger the explanatory power of the independent variable X to the attribute

Correct citation for Geographical detector:
 [1] Wang JF, Li XH, Christakos G, Liao YL, Zhang T, Gu X & Zheng XY. 2010. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China. *International Journal of Geographical Information Science* 24(1): 107-127.
 [2] Wang JF, Zhang TL, Fu BJ. 2016. A measure of spatial stratified heterogeneity. *Ecological Indicators* 67: 250-256.

Y , and the weaker the other way round. In extreme cases, The q value of 1 indicates that factor X completely controls the spatial distribution of Y , the q value of 0 indicates that factor X has no relationship with X , and the q value indicates that X explains $100 \times q\%$ of Y . A simple transformation of Q value satisfies non-central F distribution.

$$F = \frac{N-L}{L-1} \frac{q}{1-q} \sim F(L-1, N-L; \gamma) \quad (5)$$

$$F = \frac{N-L}{L-1} \frac{q}{1-q} \sim F(L-1, N-L; \gamma) \quad (6)$$

where: λ is a non-central parameter; Y_h is the average of layer H , According to Equation (2), it is possible to check whether the Q value is significant by looking up the table or using the geographic detector software.

Furthermore, an “interaction detector” was defined to assess the interaction between two different factors, namely X_1 and X_2 , by comparing $q(X_1 \cap X_2)$ with $q(X_1)$ and $q(X_2)$. $X_1 \cap X_2$ indicates a new stratum created by overlaying factors X_1 and X_2 .

2.3.2. The NDTG Model Solution

In this paper, A complete model in Python was selected to solve the NDTG. The Artificial Oasis was Extracted by Deep Learning Algorithm. The Inartificial oasis was extracted by Modified a three-band maximal gradient difference Algorithm. The result was measured by geographical detector, and the NDTG model was finished.

3. Results

3.1. The Extraction of Artificial Oasis by Deep Learning Algorithm

As shown in the following **Figure 4**, artificial vegetation was extracted through deep learning Algorithm. This paper selected the time series images from 1990 to 2015 to get the artificial vegetation area.

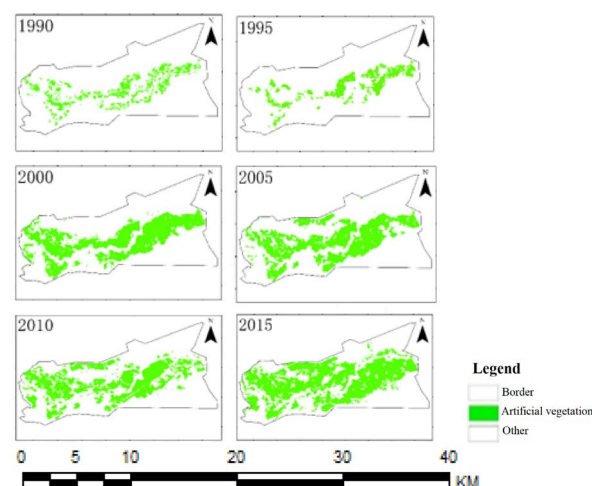


Figure 4. Time-series normalized difference oasis index (NDVI) of study area.

It shows that the square of artificial vegetation increased evidently. In the Alar region, the area of cultivated land in each regiment can be obtained by consulting the Alar yearbook of Agricultural First Division, and the area accuracy can be acquired. The accuracy of the area of cultivated land in each regiment of the Alar can be obtained by consulting the Alar yearbook of Agricultural First Division.

According to the existing Alar Yearbook (1990-2015) of Agricultural First Division, the oasis area in 1990, 1995, 2000, 2005, 2010 was calculated. The accuracy was verified through deep learning Algorithm, which can meet the data accuracy in our study (**Table 2**).

3.2. Modified a Three-Band Maximal Gradient Difference Algorithm (TGDVI)

Compared with extracting artificial oasis using deep learning Algorithm. Because the non-artificial vegetation in arid areas is sparse, the deep learning method cannot extract the non-artificial vegetation in arid areas effectively and efficiently. According to the previous research (Jiapaer et al., 2011), the modified three-band gradient method is more adopted to extract the non-artificial oasis (especially spare forest, grassland and desert vegetation), so in this study, we chose this Algorithm to extract inartificial oasis (**Figure 5**).

All the measured points were collected in August and visually interpreted through the 1990-2015 high-resolution images in Google Earth's research area. The interpretation results were corrected in combination with the actual investigation of field samples and compared with the measured values (**Table 3**).

Table 2. The accuracy using deep learning algorithm.

year	Statistical data/km ²	Extract data/km ²	Accuracy
1990	419.33	444.52	94.33%
1995	601.4	527.49	87%
2000	1110.4	1246.78	88%
2005	1120.4	1245.13	89%
2010	915.4	1235.98	74%
2015	1895	1911	99.16%

Table 3. The accuracy of oasis using TGDVI.

Year	Actual sampling points	Classification error point	Accuracy
1990	100	10	90
1995	100	12	88
2000	100	8	92
2005	100	11	89
2010	100	8	92
2015	100	7	93

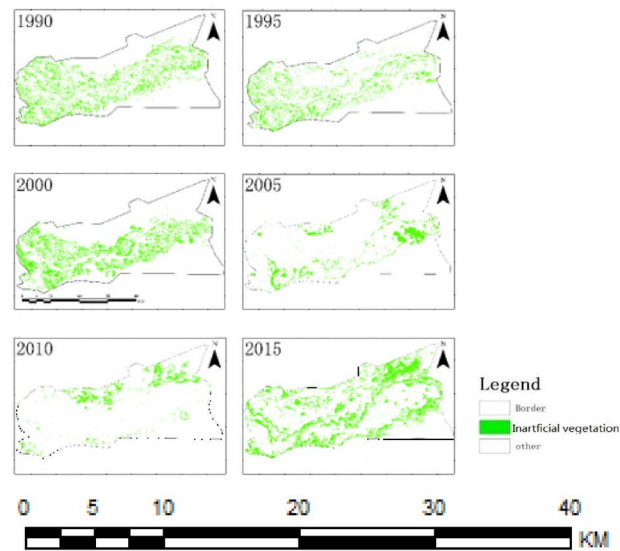


Figure 5. Time-series normalized difference oasis index (NDVI) of study area.

At 100 points of actual sampling, the accuracy range from 1990 to 2015 is 88% - 93%, which basically meet the classification accuracy requirements. In the later period of actual investigation, it was found that the classification error points were mostly distributed beside rivers, lakes and canals, which may involve inaccurate longitude and latitude positioning during sample collection and may be the cause of inaccurate classification.

3.3. The Analysis of Vegetation Growth

In the Section 3.1 and 3.2 we have known that deep learning and TGDVI can meet the requirements of extracting vegetation area. There are differences in index calculation due to different data types (for example, the processing of land-sat 7 ETM strips). When calculating the normalized difference vegetation index, the June-August partly cloudy or cloudless data of unified lands at 8 from 2013 to 2018 are used. In this article, we always use Normalized Difference Vegetation Index (NDVI) to analysis the vegetation growth. According to the former extraction of artificial oasis and inartificial oasis, we can divide this area NDVI with two different parts: artificial oasis and inartificial oasis using artificial oasis and inartificial oasis border.

3.3.1. The Artificial Vegetation Growth

The normalized difference vegetation oasis index (NDVI) of artificial oasis (cultivated land) is obtained through band calculation, and the NDVI of artificial oasis (cultivated land) from 2013 to 2018 is proposed.

As can be seen from the **Figure 6**, the maximum value in 2013 is 0.64, mainly distributed around the 9th and 12th regiments, and the minimum value is 0, mainly scattered in non-oasis areas. In 2014, the maximum value was 0.62, mainly distributed in 9 regiments and the minimum value was 0, mainly scattered in the vicinity of non-oasis areas. The maximum value in 2015 is 0.53,

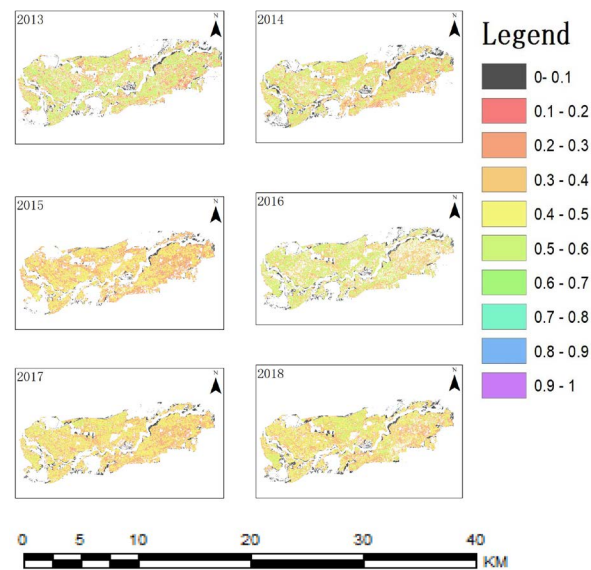


Figure 6. Time-series normalized difference vegetation index (NDVI) of study area.

mainly distributed in the 9th and 16th regiments, and the minimum value is 0, mainly scattered in the vicinity of non-vegetation areas. The maximum value in 2016 is 0.64, mainly distributed in the 12th and 13th regiments, and the minimum value is 0, mainly scattered in the vicinity of non-vegetation areas; the maximum value in 2017 is 0.56, mainly distributed in the 7th, 8th and 10th regiments, and the minimum value is 0, mainly scattered in the vicinity of non-vegetation areas. The maximum value in 2018 is 0.55, mainly distributed in the 9th and 10th regiments, and the minimum value is 0, mainly scattered in the vicinity of non-vegetation areas.

3.3.2. The Inartificial Vegetation Growth

The normalized difference vegetation index (NDVI) of inartificial oasis is obtained through band calculation, and the NDVI of artificial oasis from 2013 to 2018 is proposed.

As can be seen from **Figure 7**, the maximum value in 2013 is 0.45, mainly distributed in the 9th, 16th and 5th regiments, and the minimum value is 0, mainly distributed in the 14th and 15th regiments of non-vegetation areas. In 2014, the maximum value was 0.61, mainly distributed in the 9th and 16th regiments, and the minimum value was 0, mainly scattered in the 14th and 15th regiments. The maximum value in 2015 is 0.49, mainly distributed in the 8th and 16th regiments, and the minimum value is 0, mainly distributed in the 14th and 15th regiments. The maximum value in 2016 is 0.63, mainly distributed in the 15th and 16th regiments, and the minimum value is 0, mainly distributed in the 14th and 15th regiments. The maximum value in 2017 is 0.55, mainly distributed in the 7th and 8th regiments, and the minimum value is 0, mainly distributed in the 14th and 15th regiments. In 2018, the maximum value was 0.54, mainly distributed in the 12th regiment, the minimum value was 0, mainly scattered in the 14th regiment and the 15th regiment.

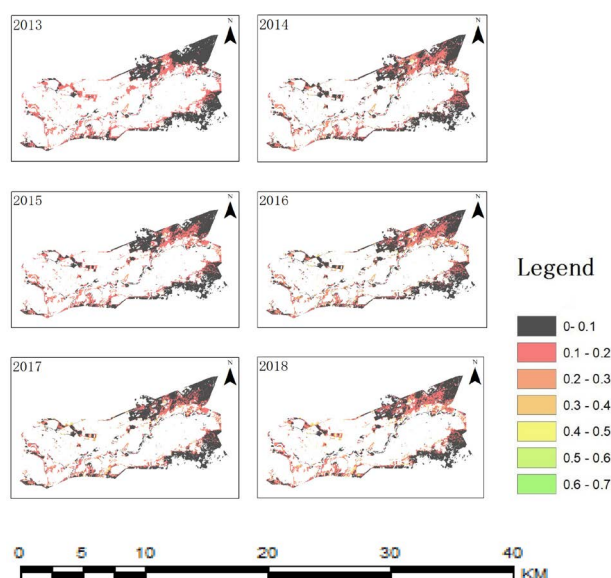


Figure 7. The Time-series normalized difference vegetation index (NDVI) of study area.

3.3.3. The Normalized Difference Vegetation Index (NDVI) of Artificial Oasis and Inartificial Oasis (Figure 8 and Table 4)

The descriptive statistics for Normalized Difference Vegetation Index (NDVI) of Artificial Oasis and Inartificial oasis are shown in **Table 2**. Normalized Difference Vegetation Index (NDVI) of Artificial Oasis and Inartificial vegetation contents varied in the ranges of 0 - 0.64 and 0- 0.54 with average values of 0.4 and 0.12, respectively. This suggests the differences in the study area of Artificial Oasis and Inartificial oasis in terms of Normalized Difference Vegetation Index (NDVI).

3.3.4. The Environmental Factors

From the remote sensing image extraction of long-time series, it can be seen that the distribution area and range of artificial oasis and natural oasis change greatly with the extension of time, and the relevant image factors affect the distribution area. Here, we consider the climate, salinity, Chemical ion, groundwater as the main influencing factors, and analyze which factor has the highest degree of influence. For long-time series data, except climate data is easier to obtain, the rest of the measured data is hard to obtain, so 2017 with sufficient data is chosen as the representative year for research.

Ten major ion types were diagnosed from actual selection and measurement, including SO_4 ion, K ion, Na ion, Ca ion, CL ion, Mg ion, HCO_3 ion. The relief factors, including underground water depth, water type.

From the above figure, it can be seen that the groundwater depth (the first picture) in the study area ranges from 0.32 to 14.99 meters. The TDS in the study area ranges from 0.3 to 25.8 g/kg. The salt content in the study area ranges from 0.23 g/kg to 2.85 g/kg.

The SO_4^{2-} content in the study area is in the range of 0.53 - 30.13 g/kg. The K content in the study area ranges from 0.02 to 7.8g/kg. The Na content in the

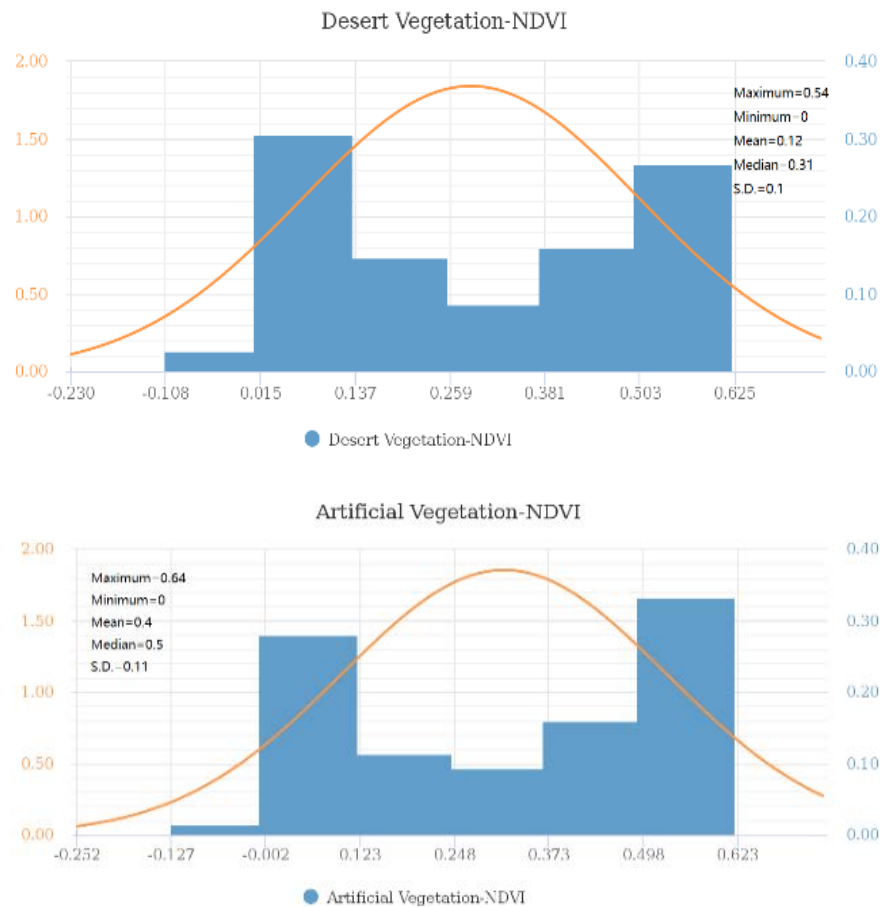


Figure 8. Frequency distributions of artificial Oasis and Inartificial oasis; S.D. is standard deviation.

Table 4. Descriptive statistics of Normalized Difference Vegetation Index (NDVI) of Artificial Vegetation and Inartificial vegetation.

type	Minimum	Maximum	Mean	Median	S.D
Artificial Vegetation	0	0.64	0.4	0.5	0.11
Inartificial vegetation	0	0.54	0.12	0.31	0.1

S.D. is standard deviation.

study area ranges from 0.21 g/kg to 46.2 g/kg. The Ca content in the study area ranges from 0.008 g/kg to 2.01 g/kg. The Cl content in the study area ranges from 0.29 g/kg to 39.98g/ kg. The Mg content in the study area ranges from 0.02 to 1.07 g/kg. The HCO_3^- content in the study area ranges from 0.23 g/kg to 2.85 g/kg.

From all the images in **Figure 9**, we can see the volume of precipitation and evaporation, the precipitation content in the study area ranges from 52 mm to 186 mm. The evaporation content in the study area ranges from 74 mm to 344 mm, which indicates that the climatic characteristics of arid areas are large evaporation and low precipitation.

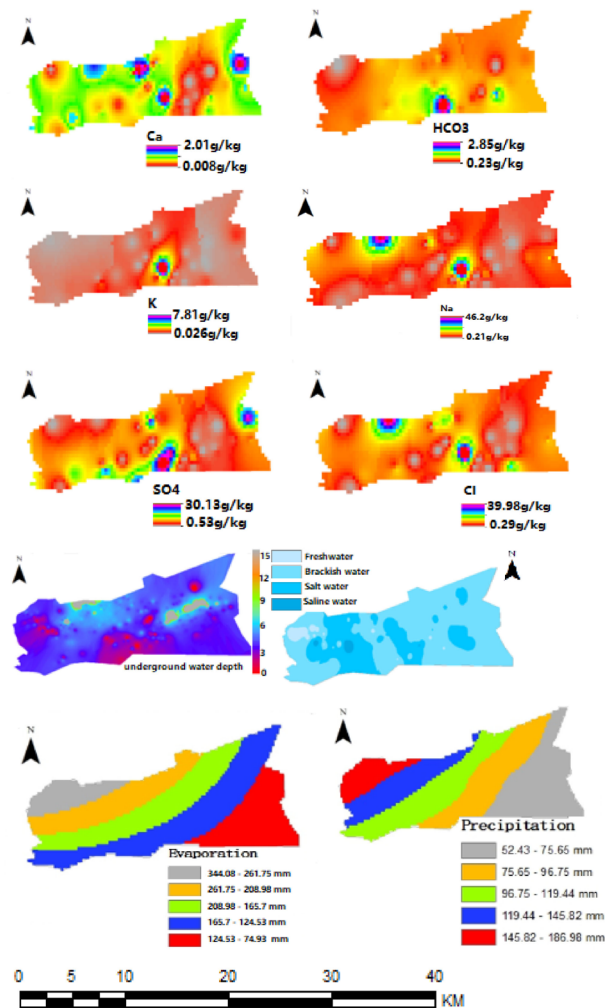


Figure 9. The environmental factors.

3.4. q Statistics

All environmental factors in **Table 1** and their interactions were considered in q statistics for the distribution of artificial oasis and inartificial oasis. For the distribution of artificial oasis, dominant factors, including sodium ion, magnesium ion, calcium ion, potassium ion, sulfate radical, chloride ion, bicarbonate radical aspect did not pass the significance test at a significance level of 0.05.

It can be seen from **Table 5** that salt has the greatest explanation to artificial vegetation. The q statistic results indicated that the Salinity factor was the dominant factor for the distribution in the study area, and it explained 33% of the spatial variation in the distribution of artificial oasis. The second is evaporation, which can explain the spatial distribution of 21% artificial vegetation. Artificial vegetation is considered to be greatly affected. Therefore, precipitation does not explain it very well. The size of salt content mainly depends on the amount of ions, so this study also tested the relationship between salt ions and artificial vegetation, so it is convenient to analyze its mechanism, as shown below.

Table 5. The Environmental factors and factor interactions with q values in artificial oasis.

NDVI	Component	q/%
Artificial oasis	Salt content	33%
	Evaporation	21%
	Precipitation	13%
	Tds	11%
	Underwater	1%
	Salt content \cap Evaporation	93%

This paper re-established the relationship between ions and artificial vegetation. As shown in **Figure 10**, it is found that Na, K, Cl and SO_4 have higher interpretation of artificial vegetation, accounting for 19%, 19%, 17% and 17% respectively. Among them, Na and K ions have higher significance on vegetation, which reflect the role of Na and K ions in artificial vegetation. Na, Cl and SO_4 ions will combine into salt ions, which will inhibit the respiration of artificial vegetation and reflect the influence of water resources on the distribution of artificial vegetation in arid areas. The data shows that the remaining ions are low which do not affect the vegetation of the oasis, thus no further analysis is needed.

It can be seen from the table above that salt has the greatest explanation to natural vegetation. The q statistic results indicated that the Salinity factor was the dominant factor for the distribution in the study area, and it explained 33% of the spatial variation in the distribution of artificial oasis (**Table 6**). The second is rainfall and evaporation, which can explain the spatial distribution of 26% and 25% natural vegetation. Due to the small relationship between natural vegetation and man-made, besides salt content, rainfall and evaporation can explain the distribution of natural vegetation to the greatest extent. The size of salt content mainly depends on the amount of ions, so this study also tested the relationship between salt ions and natural vegetation, so it is convenient to analyze its mechanism, as shown below.

The relationship between the ion distribution and the natural vegetation distribution is re-established. As shown in **Figure 11**, it is found that Mg, Cl, Na and SO_4 have higher degree of explanation for natural vegetation, 24%, 22%, 20% and 17% respectively. Among them, Mg and Cl ions have the highest degree of explanation to vegetation, reflecting the role of Mg and Cl ions in natural vegetation in the growth of natural vegetation. Mg ions play a positive role in the growth of vegetation. Magnesium is an essential nutrient element of plants and plays an important role in the physiological and biochemical processes of plants. The interpretation of the remaining ions to vegetation are low and are not the main factor, so no further analysis is needed.

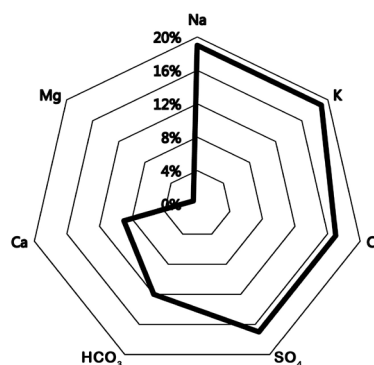


Figure 10. The environmental factors and factor interactions with q values in artificial oasis.

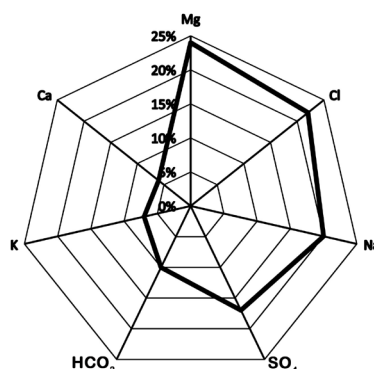


Figure 11. The environmental factors and factor interactions with q values in inartificial oasis.

Table 6. The Environmental factors and factor interactions with q values in inartificial oasis.

NDVI	Component	q statistic/%
Inartificial oasis	Salt	37%
	Precipitation	26%
	Evaporation	25%
	Tds	11%
	underwater	4%
	Salt \cap Precipitation	85%

4. Discussion

4.1. The Extraction of Artificial Oasis

Many researchers found that the accuracy of using deep learning classification to classify the remote sensing data to get the criterion are over 88% (Garg et al., 2018; Wang et al., 2019; Yu et al., 2017; Fawaz, 2019). So through comparing with other algorithm, we choose deep learning algorithm to classify remote sensing data into artificial vegetation (Goodrich et al., 2009).

From the **Table 2**, it can be seen that the accuracy has been rising. Although

our data only test and certify the 3 years data. When we find this question, through the analysis of deep learning, we can find the using deep learning algorithm every samples have renewed, so the accuracy has been approved.

In the previous part, we have extracted artificial oasis and inartificial oasis. Firstly, Landsat data has a good effect on the extraction of artificial oases into small and medium-sized regions (Zhang et al., 2005; Dee et al., 2011). The resolution of 30m can fully reflect the extraction of artificial oasis in small and medium-sized regions. The deep learning algorithm has a high accuracy on the extraction of artificial oasis, which can meet the time scale and spatial scale of the study. In this paper, the extraction accuracy of artificial oasis is more than 85%. Secondly, the deep learning algorithm can adapt to the extraction of artificial oasis in arid area, especially the extraction of artificial oases on the edge of deserts, which were not available in previous studies. In addition, for the extraction of artificial oasis in Alar region, the areas of artificial oases from 2013 to 2018 were 1846.33, 2018.59, 1911.84, 1708.38, 2073.492 and 2024.03 km² respectively, showing that the area of artificial oasis was in minor dynamic and stable change from 2013 to 2018, basically consistent with the area of the existing yearbooks, and there was no significant change as a whole.

4.2. Modified a Three-Band Maximal Gradient Difference Algorithm (TGDVI)

According to the modified a three-band maximal gradient difference Algorithm (TGDVI), the inartificial oasis growth area is extracted using Landsat8 data. The results show that, Landsat data has a good extraction effect on natural oases in small and medium-sized regions. The resolution of 30 meters can fully reflect the extraction of inartificial oasis in small and medium-sized regions. TGDVI has a high extraction accuracy on artificial oasis, which can meet the time scale and spatial scale of the study. In this paper, the extraction accuracy of natural oasis is more than 85%. Secondly, TGDVI can adapt to the extraction of artificial oasis in arid lands. In addition, the extraction of inartificial oasis in Alar region shows that the area of artificial oasis is in a dynamic and stable change from 2013 to 2018, and there is no significant change as a whole.

4.3. Geodetector

4.3.1. Analysis of Applicability of the Model

In this study, geodetector was employed to explore the factors controlling Oasis distribution of artificial oasis and non-artificial oasis accumulation in arid lands. Compared with PCA and CA, geodetector provided more convincing evidence to explain controlling factors by measuring the consistency of their spatially stratified heterogeneity with Oasis distribution of artificial oasis and non-artificial oasis. Wang declared that the consistency of spatial distribution between two geographic variables was more difficult to obtain than linear correlation of the two variables (Wang et al., 2017). Therefore, compared with Pearson correlation analysis, geodetector offered stronger statistics to reveal the causal relationship

between independent and dependent variables. Moreover, geodetector is designed for processing categorical variables, such as the geological type in this study, whereas Pearson correlation analysis is not suitable for processing this kind of data.

4.3.2. Analysis of Model Result

Through analysis, we can see that the main factors affecting the artificial oasis area in arid lands are salinity and evaporation. Q values are 32.95% and 21.74% respectively. As this area is located in a dry area and at the edge of the desert, which is dry all year round, salt content and evaporation are the main factors that affect the growth trend of artificial oasis. The main reason why precipitation has not become the main factor is that the water supply source of artificial oasis area is river basin and artificial reservoir. The study area is located near the confluence of the three rivers in the upper reaches of Tarim River, and the main reason for the small proportion of soil ion influencing factors is artificial.

The distribution of sodium ions can explain the distribution of 21.47% artificial oases. As shown in **Figure 9**, the maximum sodium ion content is 46.22 g/kg and the minimum is 0.39 g/kg. The more sodium ions in the soil, the greater the possibility of soil hardening. Soil hardening will lead to a decrease in air permeability, which will reduce the oxygen absorbed by the roots. The direct result is that the aerobic respiration of the roots is weakened or even anaerobic respiration is carried out, and then the energy provided by the roots is reduced, which is very unfavorable to the growth of plants. Therefore, too much sodium ion can inhibit vegetation. All other factors are below 20%, so we will not discuss them here.

Compared with artificial oases, it can be seen that there are many factors affecting natural oases in arid lands, with salt, evaporation, rainfall, magnesium ions, chloride ions and sodium ions accounting for the main influencing factors. q values are 36.9%, 26.27%, 25.96%, 24%, 22.72% and 20.26% respectively. Compared with artificial oases, there is no reservoir water supply in natural oases. As this area is located in arid lands and the edge of deserts, it is dry all the year round. Salt content and evaporation are the main factors influencing the growth trend of artificial oasis. The main reason why rainfall has not become the main factor is that the water supply source in artificial oasis areas is artificial reservoir water supply (Ye et al., 2010; Chen et al., 2004; Chen et al., 2009). However, the water supply of natural oasis mainly comes from rainfall and alkali drainage channels, so one of the main factors affecting the growth of natural oasis is rainfall, which also shows that this analysing model is very suitable for arid lands. However, magnesium ion, chloride ion and sodium ion, as the main other major influencing factors, can reflect the actual situation of the natural oasis region and also reflect the characteristics of the natural oasis region according to the relevant literature (Wang et al., 2017; Tokola, 2015).

In relation to salt content, magnesium ion can explain 24% of artificial oasis distribution. As shown in **Figure 9**, magnesium ion content is 0.97 g/kg at the

maximum and 0.05 g/kg at the minimum. Magnesium is one of the essential elements for plant growth. It is not only a component of chlorophyll, but also an activator of various enzymes in plants. According to its required amount, magnesium, calcium, sulfur and silicon are collectively referred to as medium elements. In recent years, many scholars have even listed magnesium as the fourth most essential element of plants after n, p and k. This shows the importance of magnesium ion to natural plant growth. The second is chloride ion and sodium ion, which can explain 22.72% and 20.26% respectively. All other factors are below 20%, so we will not discuss them here.

5. Conclusions

This study explored the factors controlling the distribution of artificial oasis and inartificial oasis accumulation in arid lands using a geographical detector method and multiple data sources. The most important conclusions were as follows.

1) By using Deep Learning algorithm and Modified a three-band maximal gradient difference Algorithm (TGDVI) to extract artificial oasis and natural oasis respectively, the accuracy meets the requirements, which can offer the significant Reference to other places in arid area.

2) Geodetector provided evidence to explore the factors controlling the spatial patterns of the distribution of artificial oasis and inartificial oasis accumulation in arid lands.

3) The NDTG model can offer a complete vision to get the main factors controlling the spatial distribution of artificial oasis and inartificial oasis in arid area.

4) Through the analysis, the main factors affecting the artificial oasis area in arid lands are salinity and evaporation; the main factors affecting the desert oasis area in arid lands are salinity, evaporation, rainfall, magnesium ions, chloride ions and sodium ions.

Conflicts of Interest

The author declares that there is no conflict of interest.

References

- Akiyama, T., Kubota, J., Fujita, K. et al. (2018). Use of Water Balance and Tracer-Based Approaches to Monitor Groundwater Recharge in the Hyper-Arid Gobi Desert of Northwestern China. *Environments*, 5, 55.
<https://doi.org/10.3390/environments5050055>
- Chen, T. B., Wong, J. W. C., Zhou, H. Y., & Wong, M. H. (1997). Assessment of Trace Metal Distribution and Contamination in Surface Soils of Hong Kong. *Am. J. Infect. Control*, 96, 61-68. [https://doi.org/10.1016/S0269-7491\(97\)00003-1](https://doi.org/10.1016/S0269-7491(97)00003-1)
- Chen, Y. J., Chen, Y. N., Liu, J. Z. et al. (2009). Influence of Intermittent Water Releases on Groundwater Chemistry at the Lower Reaches of the Tarim River, China. *Environmental Monitoring & Assessment*, 158, 251-264.
<https://doi.org/10.1007/s10661-008-0579-9>

- Chen, Y., Zhang, X., Zhu, X. et al. 2004 (). Analysis on the Ecological Benefits of the Stream Water Conveyance to the Dried-Up River of the Lower Reaches of Tarim River, China. *Science in China Series D: Earth Sciences*, 47, 1053-1064.
<https://doi.org/10.1360/03yd0101>
- Dee, D. P. et al. (2011) The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Quarterly Journal of the Royal Meteorological Society*, 137, 553-597.
- Fawaz, H. I., Forestier, G., Weber, J. et al. (2019). Deep Learning for Time Series Classification: A Review. *Data Mining and Knowledge Discovery*, 33, 917-963.
<https://doi.org/10.1007/s10618-019-00619-1>
- Garg, D., Goel, P., Kandaswamy, G., Ganatra, A., & Kotecha, K. (2018). A Roadmap to Deep Learning: A State-of-the-Art Step Towards Machine Learning. Advanced Informatics for Computing Research, Springer, Cham.
https://doi.org/10.1007/978-981-13-3140-4_15
- Goodrich, B. A., Koski, R. D., & Jacobi, W. R. (2009). Condition of Soils and Vegetation Along Roads Treated with Magnesium Chloride for Dust Suppression. *Water, Air, and Soil Pollution*, 198, 165-188. <https://doi.org/10.1007/s11270-008-9835-4>
- Guo, G., Wu, F., Xie, F., & Zhang, R. (2012). Spatial Distribution and Pollution Assessment of Heavy Metals in Urban Soils from Southwest China. *J. Environ. Sci.*, 24, 410-418.
- Gutman, G., & Ignalov, A. (1998). The Derivation of the Green Vegetation Fraction from NOAA/AVHRR Data for Use in Numerical Weather Prediction Models. *Int. J. Remote Sens.*, 19, 1533-1543. <https://doi.org/10.1080/014311698215333>
- Hodges, K. I. et al. (2010). A Comparison of Extratropical Cyclones in Recent Reanalyses ERA-Interim, NASA MERRA, NCEP CFSR, and JRA-25. *Journal of Climate*, 24, 4888-4906. <https://doi.org/10.1175/2011JCLI4097.1>
- Huo, X. N., Li, H., Sun, D. F., Zhang, W. W., Zhou, L. D., & Li, B. G. (2010). Spatial Autogression Model for Heavy Metals in Cultivated Soils of Beijing. *Trans. CSAE.*, 26, 78-82. (In Chinese)
- Jiapaer, G., Chen, X., & Bao, A. (2011). A Comparison of Methods for Estimating Fractional Vegetation Cover in Arid Regions. *Agricultural and Forest Meteorology*, 151, 1698-1710. <https://doi.org/10.1016/j.agrformet.2011.07.004>
- Lee, C. S., Li, X. D., Shi, W. Z., Cheung, S. C., & Thornton, I. (2006). Metal Contamination in Urban, Suburban, and Country Park Soils of Hong Kong: A Study Based on GIS and Multivariate Statistics. *Sci. Total Environ.*, 356, 45-61.
<https://doi.org/10.1016/j.scitotenv.2005.03.024>
- Li, X. D., Lee, S. L., Wong, S., Shi, W. Z., & Thornton, I. (2004). The Study of Metal Contamination in Urban Soils of Hong Kong Using a Gis-Based Approach. *Environ. Pollution.*, 129, 113-124. <https://doi.org/10.1016/j.envpol.2003.09.030>
- Li, X. W., Xie, Y. F., Wang, J. F., Christakos, G., Si, J. L., Zhao, H. N., Ding, H. N., & Li, J. (2013). Influence of Planting Patterns on Fluoroquinolone Residues in the Soil of an Intensive Vegetable Cultivation Area in Northern China. *Sci. Total Environ.*, 458-460, 63-69. <https://doi.org/10.1016/j.scitotenv.2013.04.002>
- Lin, Y. P., Teng, T. P., & Chang, T. K. (2002). Multivariate Analysis of Soil Heavy Metal Pollution and Landscape Pattern in Chuanghuacounty in Taiwan. *Landsc. Urban Plan.*, 62, 19-35. [https://doi.org/10.1016/S0169-2046\(02\)00094-4](https://doi.org/10.1016/S0169-2046(02)00094-4)
- Luo, W., Jasiewicz, J., Stepinski, T., Wang, J. F., Xu, C. D., & Cang, X. Z. (2015). Spatial Association between Dissection Density and Environmental Factors over the Entire

- Conterminous United States. *Geophys. Res. Lett.*, *43*, 692-700.
<https://doi.org/10.1002/2015GL066941>
- Navas, A., & Machin, J. (2002). Spatial Distribution of Heavy Metals and Arsenic in Soils of Aragon (Northeast Spain): Controlling Factors and Environmental Implications. *Appl. Geochem.*, *17*, 961-973. [https://doi.org/10.1016/S0883-2927\(02\)00006-9](https://doi.org/10.1016/S0883-2927(02)00006-9)
- Ordóñez, A., Loredó, J., De Miguel, E., & Charlesworth, S. (2003). Distribution of Heavy Metals in the Street Dusts and Soil of an Industrial City in Northern Spain. *Arch. Environ. Contam. Toxicol.*, *44*, 160-170. <https://doi.org/10.1007/s00244-002-2005-6>
- Petley, D. (2012). Global Patterns of Loss of Life from Landslides. *Geology*, *40*, 927-930. <https://doi.org/10.1130/G33217.1>
- Semwal, V. B., Mondal, K., & Nandi, G. C. (2017). Robust and Accurate Feature Selection for Humanoid Push Recovery and Classification: Deep Learning Approach. *Neural Computing and Applications*, *28*, 565-574. <https://doi.org/10.1007/s00521-015-2089-3>
- Shi, T. Z., Hu, Z. W., Shi, Z. et al. (2018). Geo-Detection of Factors Controlling Spatial Patterns of Heavy Metals in Urban Topsoil Using Multi-Source Data. *Science of the Total Environment*, *643*, 451-459. <https://doi.org/10.1016/j.scitotenv.2018.06.224>
- Sun, C. Y., Liu, J., Wang, Y., Sun, L., & Yu, H. (2013). Multivariate and Geostatistical Analyses of the Spatial Distribution and Sources of Heavy Metals in Agricultural Soil in Dehui. *Northeast China. Chemosphere*, *92*, 517-523. <https://doi.org/10.1016/j.chemosphere.2013.02.063>
- Sylla, M. B. et al. (2010). Multiyear Simulation of the African Climate Using a Regional Climate Model (RegCM3) with the High Resolution ERA-Interim Reanalysis. *Climate Dynamics*, *35*, 231-247. <https://doi.org/10.1007/s00382-009-0613-9>
- Tang, S.-H., Zhu, Q.-J., Zhou, Y.-Y., Bai, X.-H., & Shuai, Y.-M. (2003). A Simple Method to Estimate Crown Cover Fraction and Rebuild the Background Information. *J. Image Graphics*, *8*, 1304-1309. (In Chinese)
- Tokola, T. (2015). Remote Sensing Concepts and Their Applicability in REDD+ Monitoring. *Current Forestry Reports*, *1*, 252-260. <https://doi.org/10.1007/s40725-015-0026-4>
- Tyr, W. H., Stewart, E. L., Nicholas, K. et al. (2018). Image Set for Deep Learning: Field Images of Maize Annotated with Disease Symptoms. *BMC Research Notes*, *11*, Article Number: 440. <https://doi.org/10.1186/s13104-018-3548-6>
- Wang, J. F., & Xu, C. D. (2017). Geodetector: Principle and Prospective. *Acta Geograph. Sin.*, *72*, 116-134.
- Wang, J. F., Li, X. H., Christakos, G., Liao, Y. L., Zhang, T., Gu, X., & Zhang, X. Y. (2010). Geographical Detectors-Based Health Risk Assessment and its Application in the Neural Tube Defects Study of the Heshun Region, China. *Int. J. Geogr. Inf. Syst.*, *24*, 107-127. <https://doi.org/10.1080/13658810802443457>
- Wang, J. F., Zhang, T. L., & Fu, B. J. (2016). A Measure of Spatial Stratified Heterogeneity. *Ecol. Indic.*, *67*, 250-256. <https://doi.org/10.1016/j.ecolind.2016.02.052>
- Wang, K. M., Qian, J., Huang, S. Y., Wang, Y., Yang, X. B., & Duo, B. (2019). Through-the-Wall Radar Imaging Based on Deep Learning. *Social Informatics and Telecommunications Engineering*, Springer, Cham. https://doi.org/10.1007/978-3-030-19156-6_58
- Weedon, G. P. et al. (2015). The WFDEI Meteorological Forcing Data Set: WATCH Forcing Data Methodology Applied to ERA-Interim Reanalysis Data. *Water Resources Research*, *50*, 7505-7514. <https://doi.org/10.1002/2014WR015638>

- Wilford, J., de Caritat, P., & Bui, E. (2016). Predictive Geochemical Mapping Using Environmental Correlation. *Appl. Geochem.*, 66, 275-288.
<https://doi.org/10.1016/j.apgeochem.2015.08.012>
- Yabuki, N., Nishimura, N., & Fukuda, T. (2018). Automatic Object Detection from Digital Images by Deep Learning with Transfer Learning. Workshop of the European Group for Intelligent Computing in Engineering. Springer, Cham.
https://doi.org/10.1007/978-3-319-91635-4_1
- Yang, Q., Mu, H., Wang, H. et al. (2018). Quantitative Evaluation of Groundwater Recharge and Evaporation Intensity with Stable Oxygen and Hydrogen Isotopes in a Semi-Arid Region, Northwest China. *Hydrological Processes*.
<https://doi.org/10.1002/hyp.11474>
- Ye, Z. X., Chen, Y. N., & Li, W. H. (2010). Ecological Water Demand of Natural Vegetation in the Lower Tarim River. *Journal of Geographical Sciences*, 20, 261-272.
<https://doi.org/10.1007/s11442-010-0261-3>
- Yu, S., Yuan, L., Guan, W. et al. (2017). Deep Learning for Plant Identification in Natural Environment. *Computational Intelligence and Neuroscience*, 2017, 1-6.
<https://doi.org/10.1155/2017/7361042>
- Zhang, R., Li, Z., Sun, X. et al. (2005). On the Applicability of Kirchoff's Law and the Principle of Heat Balance in Thermal Infrared Remote Sensing: A Non-Isothermal System. *Science in China*, 48, 53-64. <https://doi.org/10.1360/03YD0108>