remote sensing

MDPI

Article

Mining Coastal Land Use Sequential Pattern and Its Land Use Associations Based on Association Rule Mining

Zhi Ding [1,2,3], Xiaohan Liao [1,*], Fenzhen Su [1] and Dongjie Fu [1]

1 State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; dingzhi11@mails.gucas.ac.cn (Z.D.); sufz@lreis.ac.cn (F.S.); fudongjie@gmail.com (D.F.)
2 Department of Resources and Environment; University of Chinese Academy of Sciences, Beijing 100049, China
3 School of Geography; Beijing Normal University, Beijing 100875, China
* Correspondence: liaoxh@igsnrr.ac.cn; Tel.: +86-10-6488-9502

Academic Editors: Liping Di, Qian Du, Peng Liu, Lizhe Wang, Xiaofeng Li and Prasad S. Thenkabail
Received: 23 October 2016; Accepted: 24 January 2017; Published: 29 January 2017

Abstract: Research on the land use of the coastal zone in the sea–land direction will not only reveal its land use distribution, but may also indicate the interactions between inland land use and the ocean through associations between inland land use and seaward land use indirectly. However, in the existing research, few have paid attention to the land use in sea–land direction, let alone the sequential relationship between land-use types. The sequential relationship would be useful in land use planning and rehabilitation of the landscape in the sea–land direction, and the association between land-use types, particularly the inland land use and seaward land use, is not discussed. Therefore, this study presents a model named ARCLUSSM (Association Rules-based Coastal Land use Spatial Sequence Model) to mine the sequential pattern of land use with interesting associations in the sea–land direction of the coastal zone. As a case study, the typical coastal zone of Bohai Bay and the Yellow River delta in China was used. The results are as follows: firstly, 27 interesting association patterns of land use in the sea–land direction of the coastal zone were mined easily. Both sequential relationship and distance between land-use types for 27 patterns among six land-use types were mined definitely, and the sequence of the six land-use types tended to be tidal flat > shrimp pond > reservoir/artificial pond > settlement > river > dry land in sea–land direction. These patterns would offer specific support for land-use planning and rehabilitation of the coastal zone. There were 19 association patterns between seaward and landward land-use types. These patterns showed strong associations between seaward and landward land-use types. It indicated that the landward land use might have some impacts on the seaward land use, or in the other direction, which may help to reveal the interactions between inland land use and the ocean. Thus, the ARCLUSSM was an efficient tool to mine the sequential relationship and distance between land-use types with interesting association rules in the sea–land direction, which would offer practicable advice to appropriate coastal zone management and planning, and might reveal the interactions between inland land use and the ocean.

Keywords: coastal zone; land use; spatial sequence pattern; from the land to the sea; association rules; the anisotropy

1. Introduction

The coastal zone is a strip region interacts with the sea, atmosphere, and Earth, and has gradient characteristics, such as physiographic gradient, biogenic gradient, and physiochemical gradient, that are linear and roughly parallel to the shore. The gradient could be demonstrated by the areas of the coastal zone. In geomorphologic and hydrological terms, the coastal zone is divided into five reference units as uplands, coastal plains, and tidelands, in the terrestrial portion, and shoreface entrainment and offshore entrainment volume, in the marine portion [1]. For discussion in light of natural and social characteristics in the coastal zones, a "three-zone" conception was put forward, including the featured core (the strongest land-ocean interaction) zone, adjacent (stronger land-ocean interaction) zone, and the peripheral (weak land-ocean interaction) zone [2], with increasing distance from the sea. Affected by the physiographic factors, biogenic factors, and physiochemical factors [1] in the areas of the coastal zone, the land use in the coastal zone also shows a tendency of gradient with increasing distance from the sea. Based on the coastal land use data in China, remarkable sea-land gradient characteristics of land use was found; particularly, water bodies and residential areas showed the most notable seaward characters [3]. The gradient of land use was also seen in the Jiangsu coastal zone. Additionally, the land use diversity tended to be low-high-low in the land–sea direction [4], and in the Bohai Sea region, significant differences of land use were also found among different administrative regions and from the sea to the land [5]. In this sense, research on the gradient of land use in the coastal zone may not only help to reveal the land use pattern from land to sea, but also indicates the characteristic of zones of the coastal zone. However, the relevant studies focus on the zones [1,2,6] and structures [3–5] and changes [7] of land use in them. They could not reveal the sequential relationship (we name it sequential pattern) between land-use types in the sea–land direction in the coastal zone, let alone one area of the coastal zone. For example, giving three land-use types, such as shrimp pond, forestland, and cultivated land in the coastal zone, the sequential pattern of them in the sea–land direction may be shrimp pond > forestland > cultivated land, or forestland > shrimp pond > cultivated land, or shrimp pond > cultivated land > forestland, or the other combinations. From the sequential pattern, the landscape of the coastal zone in sea–land direction is explicitly revealed. The patterns with settlement or cultivated land behind the forestland in the sea–land direction indicate that settlement or cultivated land may be protected by forestland from seawater salt spray and even storms [8]. This would help in landscape planning of the coastal zone, particularly the rehabilitation of the landscape in the land–sea direction.

Apart from the gradient characteristic of land use in the coastal zone, anthropogenic land use, particularly near to the sea, may impact aquatic ecosystems in several ways, including increased frequency and intensity of floods, sedimentation, and loss of microtopography [9,10]. A study was done to test whether human activities on land can increase subsidies of terrigenous nitrogen to open rocky coasts and whether these differences can predict apparent deforestation of kelp forests, and the results were positive [11]. It was found that land reclamation had a significant impact on the M_2–M_4 tidal-duration asymmetry between 1935 and 1966 in Jiaozhou Bay [12]. It also suggested that benthic conditions were positively correlated with watershed forestation in Chesapeake Bay [13]. In contrast, the ocean conditions, such as sea level rise (SLR), storms, and waves [14–17] also have some impacts on the land in the coastal zone. Especially with respect to SLR, it is suggested that coastal wetlands, such as saltmarshes and mangroves, will decline unless they have a sufficient sediment supply to keep pace with SLR [18]. These mutual impacts between land use mainly distribute around the interface between the land and the ocean, indicating the process of land-ocean interaction, which is essential for understanding the coastal dynamics [14] and has become an important research area [19]. However, research on the interaction directly needs vast data such as land use, rainfall, sediments, and chemical elements. Some of these data, like chemical elements, are hard to obtain for land-use-related researchers who have little knowledge about these data. Therefore, a new and simple way to study the land-ocean interactions is welcomed.

To fill the gaps, this study firstly simplified the banded coastal zone as adjacent lines that were perpendicular to the coastline. Each line was used to represent one small coastal region. The sequential relationship between land-use types in the small coastal region could be represented by the relationship on the line, yet the relationship on the line could be easily extracted by numbering the segments of land-use types on the line. In this way, the sequential relationship between land-use types was explicitly extracted in one small coastal region. Simultaneously, the sequential relationship for the whole coastal region could be mined from these small regions. Secondly, the interaction between the ocean and seaward land was well studied above. Then we attempted to through the association between inland and seaward land use to indirectly indicate the interaction between inland land use and the ocean. This might offer a way to study the land-ocean interactions. Therefore, combining with typical data mining method, association rule mining, we developed a new model named ARCLUSSM (association rules-based coastal land-use spatial sequence model) to mine (1) the major land use sequential patterns that indicate the sequential relationship between land-use types in the sea–land direction; (2) associations between inland and seaward land-use; and (3) and the distribution of the major sequential pattern. We used the model to mine the land use sequential pattern of the coastal zone of Bohai Bay and Yellow River delta in China as a case study.

2. Model of ARCLUSSM

The ARCLUSSM is a data mining model basing on the association rules, and consists of three parts:

1. Simplifying the coastal zone.

This is the base of the model. The coastal zone is a strip-shaped region located at the junction area of the land and the ocean. In order to explore the gradient of land use in the coastal zone, traditional methods often divide the coastal zone into several zones based on the distance to the coastline [3,20]. However, this method has two obvious weaknesses. First, the gradient of land-use in one zone is difficult to distinguish. Second, the width of one zone is usually defined by the user through personal experience, which possesses definite subjectivity. Therefore, this study simplifies the coastal zone as equidistant lines that are perpendicular to the coastline. In this way, each line is represented as one small coastal region in the study. Then, the distance between lines represents the length of the coastline of the small coastal region and the length of the line represents the width of the coastal zone. In this sense, the strip of the coastal zone is still retained. In addition, the land-use types in the small region would be represented by segments segmented by the location of these land-use types in the line. The sequential relationship between land-use types can be easily distinguished by numbering the segments in the line. In the end, this method not only divides the whole coastal zone into independent lines that are easily explored, but it also distinguishes the sequential relationship between land-use types in the small coastal region.

2. Mining interesting land use association rules

This part aims to mine interesting associations between land-use types, such as shrimp pond, forestland, and cultivated land that are contained by the line based on association rule mining. Association rule mining (ARM) is one typical data mining method that aims to extract interesting correlations, frequent patterns, associations, or casual structures among sets of items in the transaction databases or other data repositories, first introduced by Agrawal [21,22]. Before mining the interesting land use association rules, we should interpret the land use on the simplified lines mentioned above. In order to make sure that each land-use type is unique in one line, we preprocess the land-use types contained by each line, and the detailed process is described below (Section 2.2). In this sense, the land-use types contained by one line are represented as one transaction and each land-use type on the line is regarded as one item.

Traditionally, the association rule mining is to find the association rules that satisfy predefined minimum support (S) and confidence (C) from a given database. However, due to the total number of transactions and null-invariant transactions, the result often contains uninteresting rules. In other words, only the support and confidence are not sufficient to filter out uninteresting association rules. There is an index, named Kulczynski (Kulc), that was proved to be an efficient index to extract interesting rules [23], first introduced by Kulczynski in 1972 [24]. Therefore, the study chooses the Kulc as the subsidiary index to extract interesting rules. Above all, one interesting land use association rule is regarded as several different land-use types that are contained by most of lines (judged by support) and have interesting associations between each other (judged by confidence and Kulc).

3. Mining land use sequential patterns

This part aims to extract the sequential relationships and distance between different land-use types. The detailed information of the two indexes will be described below in Section 2.3.

2.1. Simplifying the Coastal Zone

There are three steps to simplify the coastal zone below:

Step 1 Coastline extraction.

The coastline is defined as instant water line of the ocean, and it is interpreted by a human in this study.

Step 2 Determining the length of the line and distance between two lines.

The length of the line is equivalent to the width of the coastal zone. To maintain consistency with the length of the line, the width of the coastal zone is defined as one uniform value. In order to cover the whole sequential pattern of land use from sea to land, we develop a method. Firstly, the coastal zone is divided into several zones of equivalent width. Secondly, we calculate the number of land-use types whose total areas are greater than the average area of all land-use types in one zone. Third, the numbers of zones in sequence from sea to land are compared. If there is one number that the value of the numbers behind it start to retain stabilization, the landward boundary of the zone that contains the turning number is the land boundary of the coastal zone, and the length of the line is equivalent to the distance from the coastline line to the land boundary determined above. Of course, the length could also be determined depending on the user's requirements.

The distance between two lines largely depends on the smallest patch of land use in the study area. In general, the distance should not be smaller than the resolution of the data, or it would lead to redundant data. Too small a distance may lead to the overlapping of lines, which may cause chaotic information in the result. However, overly long distances may lose some land-use types that are in small size. Therefore, the distance is defined as the average value of the width and length of the smallest land use patch in the study area, which aims to assure that most land-use types are covered.

Step 3 Constructing simplified lines.

This step generates the simplified lines based on the width and distance defined above. In this study, it is realized through GIS (geographic information system).

2.2. Mining Interesting Land Use Association Rules

Firstly, the purpose is to interpret the land use data of simplified lines. The data is interpreted from remote sensing images if there is no land use data, or it could be extracted from existing land use data.

Secondly, in order to distinguish a sequential relationship between land-use types, the land-use types in one line are numbered orderly from sea to land. Considering one land-use type may occur more than twice, and ensuring the uniqueness of one land-use type in one line, we choose the average number of all numbers of the land-use type in the line as the final number of the land-use type. In the end, we number the land-use types in the line again depending on the final number. For example, there is one line that contains land-use types of tidal flat, shrimp pond, settlement, cultivated land, forestland, grassland, and settlement, respectively, from sea to land. The numbers of these land-use types are successively 1, 2, 3, 4, 5, 6, and 7. Then the sequential indexes of land uses in the line are 1(tidal flat), 2 (shrimp pond), 4 (cultivated land), 5 (settlement), 5 (forestland), 6 (grassland). In addition, the sequential pattern of land use in the line is tidal flat > shrimp pond > cultivated land > settlement = forestland > grassland.

Third, according to the numbered land use data, it is to mine the interesting land use association rules with indices of support, confidence, and Kulczynski degree.

(1) Support and confidence

Support of an association rule is defined as the percentage/fraction of records that contain X∪Y to the total number of records in the database [22]. In this study, the support is used as the indicator of the distribution of land-use types appearing in the coastal area. The higher support, the wider distribution of land use. Confidence of an association rule is defined as the percentage/fraction of the number of transactions that contain X∪Y to the total number of records that contain X [22]. It helps us to know the probability of distribution between land-use types. For example, as the rule of cultivated land -> beach with a support of 75% and confidence of 90%, the support of 75% means that 75% of the area has cultivated land and beach and the confidence of 90% means that if cultivated land appears in the area, the existent probability of beach is about 90%.

(2) The Kulczynski degree

The Kulczynski degree is the average probability of itemset A and itemset B, and is calculated by Equation (1) [23] below:

$$\mathrm{K}ulc(A,B) = \frac{1}{2}(P(A|B) + P(B|A)) \tag{1}$$

where Kulc(A,B) is the Kulczynski degree of itemsets A and B; P(A | B) is the probability when itemset A does not exist and when itemset B does exist in the same time; P(B | A) is contrary to P(A | B); the range of the Kulczynski degree is (0, 1). If the value is in the range of (0, 0.5), the association between itemset A and itemset B is a negative association. If the value is in the range of (0.5, 1), the association is positive, and the larger the value, the stronger the association. If the value is 0.5, the association is neutral [23].

In the end, to extract interesting land use association rules, the rules with support and confidence that satisfy predefined minimum values (e.g. support >= 40% and confidence >=60) and the values of Kulc of them are not equal to 0.5, are mined as strong association rules.

2.3. Sequential Index and Distance Index

The sequential index named average sequence (AS) is used to represent the sequence of the land-use type in one strong association rule. The AS is the average value of numbers of one type in one line, which is calculated by Equations (2)–(3).

$$AS_{lu} = \frac{\sum_{k=1}^{c} ave\_num_k(lu)}{c} \tag{2}$$

$$ave\_num_k(lu) = \frac{\sum_{x=1}^{y} num\_lu_x}{y} \tag{3}$$

where AS_{lu} is the value of average sequence of land use lu; c is the count of the lines that contain the strong association pattern (the land use lu belongs to the pattern); ave_num$_k$ (lu) is the average number of land use lu on the line k; y is the count of land use lu on the line k.

The AS index largely indicates the number of land use types between two types, which has limits to reveal the specific distance between two types. Thus, the model develops one distance index—average length (AL)—to represent the distance between different types. The AL is the distance between two types in one line, which is calculated by Equations (5)–(7) below:

$$AL_{lu1\_lu2} = \frac{\sum_{k=1}^{c} D_{lu1_k} - D_{lu2_k}}{c} \tag{5}$$

$$D_{lu1_k} = \frac{\sum_{i=1}^{n} len\_lu1_i}{n} \tag{6}$$

$$D_{lu2_k} = \frac{\sum_{j=1}^{m} len\_lu1_j}{m} \tag{7}$$

AL_{lu1_lu2} is the value of the average length between land use $lu1$ and land use $lu2$; c is the count of the lines that contain the strong association pattern (land use 1 and land use 2 belong to the pattern); D_{lu1_k} is the distance of land use $lu1$ from the coastline on the line k; D_{lu2_k} is the distance of land use $lu2$ from the coastline on the line k; n and m are the counts of land use $lu1$ and land use $lu2$ on the line k.

With the AS and AL above, the sequential relationships and distance between land use types in the interesting association rules above would be calculated. Then, with the support from GIS, the distribution of different association-based sequential patterns is easily mapped. The distribution not only helps to reveal the spatial distribution characteristics of the patterns, but may also help to reveal the driving factors of the patterns.

3. A Case Study in the Coastal Zone of Bohai Bay and the Yellow River Delta

3.1. Study Area

The study area (37°28′–39°29′N, 117°15′–119°17′E) (Figure 1) is located in the west of the Bohai Sea in North China, including Bohai Bay and the Yellow River delta. It ranges from Laoting county in the north of Bohai Bay to Kenli county in the south of the Yellow River delta, covering 16 coastal counties and districts belonging to three provinces. The area is the central and most developed area of the Bohai Rim, and it is also a key element in China's national strategy. The type of coast is mainly silt and the length of the coastline is about 1716 km. The area is characterized by a typical continental climate (monsoon and warm-temperate climate), and the annual average precipitation and evaporation are approximately 600 mm and 1944 mm, respectively, with 70% of the total precipitation occurring between July and August [25]. The main soil types are alluvial soil and saline soil in the lower part of the coastal area, bog soil in depressions, and cinnamon soil in hilly regions mainly in the east of the area. The natural vegetation in the study area is composed of halophytic plant communities predominated by herb and shrub species. The wetland hydrological characteristics are affected by the interactions between freshwater and seawater and between groundwater and surface water due to the low elevation (generally below 10 m except for roads and dikes) and being near the sea. According to the analysis of the change of numbers of land-use types in 30 zones with 1 km width in the sea–land direction, it was found that the 20-th zone was the turning point. Therefore, the width of the coastal zone is defined as 40 km, with 20 km distance inland and to the sea from the coastline.

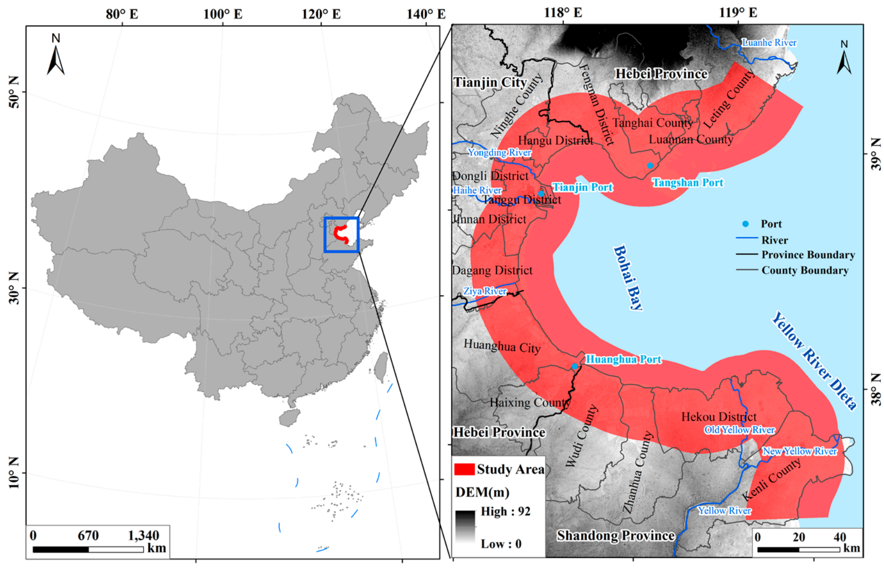
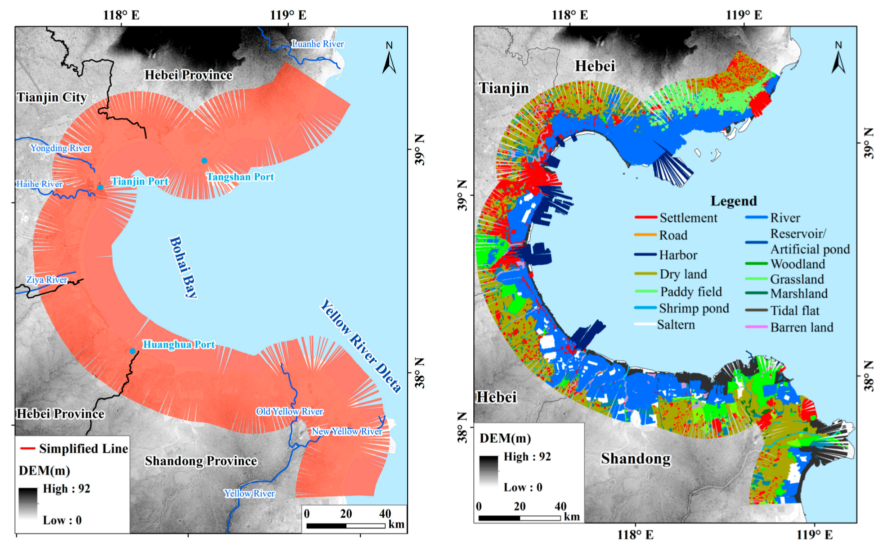


Figure 1. Study area.

3.2. Data

The data mainly includes the simplified line data and their land use data. The simplified line data (Figure 2A) is produced by a software program. The length of the line is 40 km, which is equivalent to the width of the coastal zone defined above, and the average value of the width and length of the smallest land-use patch is about 200 meters (m), so the distance between two lines is defined as 200 m.



(**A**) (**B**)

Figure 2. Simplified lines (**A**) and their land use (**B**).

The original land use data comes from the land-use database of the Resources and Environment Scientific Data Center, Chinese Academy of Sciences (RESDC) [26–28]. The land-use database is constructed from remotely-sensed digital images by the US Landsat TM/ETM satellite with a spatial resolution of 30 m × 30 m. The period of the land use data in this study is 2010. The classification system of the study area is defined with 14 land use types, including barren land (BL), tidal flat (TF), marshland (ML), grassland (GL), woodland (WL), reservoir/artificial pond (Re/AP), river (Ri), saltern (Sa), shrimp pond (SP), paddy field (PF), dry land (DL), harbor (HB), road (R), and settlement (Se). The land use of simplified lines (Figure 2B) is processed from the land use data of RESDC by intersecting analysis through GIS. Land-use types in the lines are numbered and processed for uniqueness based on the method above (Section 2.2).

4. Results

Basing on the model of ARCLUSSM, the land use association rules with support exceed 40%, confidence exceeded 60%, and Kulc not equivalent to 0.5, were mined as interesting rules. We had mined a total of 64 interesting land use association rules, including 25 rules for two-kind land-use types (Table 1, which indicates associations between two different land-use types, and three-kind and four-kind below are similar with it), 31 rules for three-kind land-use types, and eight rules for four-kind land-use types. Simultaneously, 27 land use sequential patterns were extracted based on the rules, including 14 patterns for two-kind land-use types, 11 patterns for three-kind land-use types, and two patterns for four-kind land-use types. In the end, the distributions of these patterns were mapped, respectively.

Table 1. Interesting association rules for two-kind land-use types (S > 40, C > 60, Kulc unequal to 0.5).

ID	Rules	S (%)	C (%)	Kulc	AS		AL (m)
	T1 -> T2				T1	T2	T1-T2
1	Se -> DL	77.22	92.97	0.89	29	32	4798
2	DL -> Se	77.22	84.41	0.89	32	29	4798
3	SP -> DL	77.13	90.05	0.87	14	32	11,809
4	DL -> SP	77.13	84.31	0.87	32	14	11,809
5	Se -> SP	73.04	87.93	0.87	26	13	9383
6	SP -> Se	73.04	85.27	0.87	13	26	9383
7	TF -> DL	64.00	96.25	0.83	1	28	18,826
8	DL -> TF	64.00	69.96	0.83	28	1	18,826
9	Ri -> DL	63.95	90.84	0.80	27	34	5674
10	DL -> Ri	63.95	69.91	0.80	34	27	5674
11	Ri -> SP	59.39	84.36	0.77	22	9	8530
12	SP -> Ri	59.39	69.34	0.77	9	22	8530
13	Ri -> Se	57.93	82.29	0.76	26	25	6204
14	Se -> Ri	57.93	69.75	0.76	25	26	6204
15	TF -> SP	56.80	85.42	0.76	1	11	6918
16	SP -> TF	56.80	66.32	0.76	11	1	6918
17	TF -> Se	53.32	80.18	0.72	1	28	15,914
18	Se -> TF	53.32	64.19	0.72	28	1	15,914
19	TF -> Ri	48.61	73.11	0.71	1	24	15,235
20	Ri -> TF	48.61	69.05	0.71	24	1	15,235
21	Re/AP -> DL	48.42	93.63	0.73	23	31	5482
22	Re/AP -> SP	43.86	84.80	0.68	19	8	11,121
23	Re/AP -> Se	43.01	83.17	0.67	23	23	6031
24	Re/AP -> Ri	42.59	82.35	0.71	24	26	7004
25	Ri -> Re/AP	42.59	60.49	0.71	26	24	7004

("->" indicates that the land-use type on the right of it has strong associations with the land-use type on the left. If the land-use type on the left is appearing in one line, the land-use type on the right will appear on the same line with high probability (defined as confidence); the larger the AS, the longer the distance from the coastline.)

4.1. Interesting Land Use Association Rules and Their Sequential Patterns

There were 25 interesting land-use association rules between two land-use types among six different land-use types, including dry land, settlement, river, reservoir/artificial pond, shrimp pond, and tidal flat. Excepting the dry land, each land-use type among the six land-use types has strong association with the others. Dry land only has a weak association with the reservoir/artificial pond type among the six land-use types. According to the AS shown in Figure 3A, the six land-use types could be divided into two classes, which are the seaward land-use type, like tidal flat, and landward land-use type, like dry land. The baseline is the AS with a value of 15. The seaward land-use type is the land-use type that tends to distribute inland of the coastal zone with AS greater than 15. The landward land-use type is the land-use type that tends closer to the coastline with AS less than 15. In this sense, tidal flat and shrimp pond belong to the seaward land-use types, and the dry land, settlement, river and reservoir/artificial pond belong to the landward land-use types. From Figure 3A, each landward land-use type has strong associations with the other landward land-use types, except dry land, among the four landward land-use types. The two seaward land-use types also have strong association between each other. The tidal flat has strong association with landward land-use types, except reservoir/artificial pond. The shrimp pond has strong association with the four landward land-use types. Above all, there are six strong association patterns between landward land-use types, one strong association pattern between seaward land-use types, and seven strong association patterns between seaward and landward land-use types.

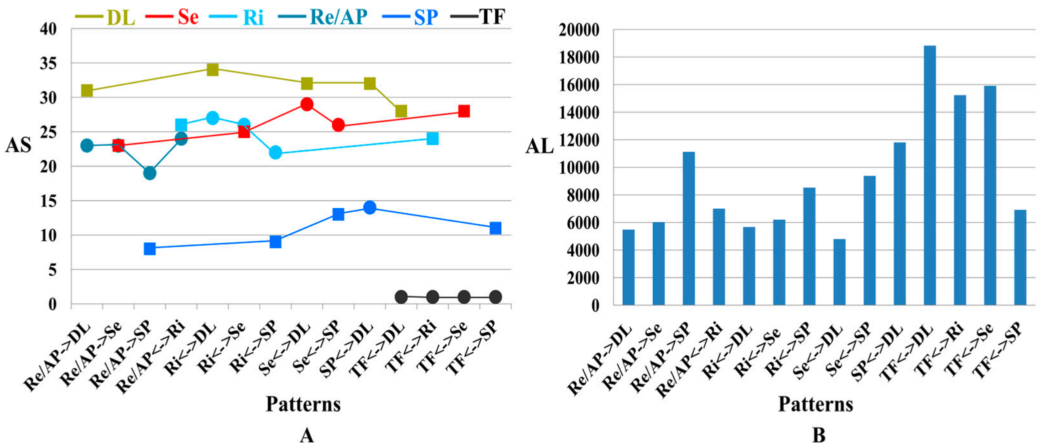


Figure 3. The AS (**A**) and AL (**B**) of land-use types for two-kind association pattern. ("<->" indicates that the land-use type both on the right and left has strong associations with each other. If the land-use type on the left or right appears in one line, the land-use type on the right or left will appear on the same line with high probability; the rectangle (**A**) indicates the land-use type on the right of "->" or "<->" (T2 in Table 1), and the circle indicates the land-use type on the left (T1).

Comparing the AS and AL (Figure 3A) between land-use types in strong association patterns above, the sequential relationship and distance between land-use types has some obvious characteristics. Firstly, the seaward land-use types tend to be in the front in the sea–land direction among the strong association patterns containing both seaward and landward land-use types. Simultaneously, the landward land-use types tend to be to the rear. For example, in the pattern of reservoir/artificial pond -> shrimp pond, the sequential pattern is shrimp pond> reservoir/artificial pond in the sea–land direction. Dry land tends to be to the rear of land-use types that have strong association with it. In contrast, tidal flat tends to be in the front. The sequential patterns between land-use types among settlement, river and reservoir/artificial pond are reservoir/artificial = settlement, reservoir/artificial > river and settlement > river. As mentioned above, the sequential pattern for the six land-use types is tidal flat > shrimp pond > reservoir/artificial pond = settlement > river > dry land. Secondly, the distance between landward and seaward land-use types tends to be longer, and the distance between landward or seaward land-use types tends to shorter (Figure 3B). The value of the difference of AS between landward and seaward land-use types is in accordance with the distance between them, but distance between landward land-use types is not in accordance with the AS.

There were 31 interesting land-use association rules (Table 2) between three land-use types among the six different land-use types mentioned above. The six land-use types have four landward land-use types and two seaward land-use types, which are the same as the two-kind land-use types. Among these association rules, most of them refer to associations between landward and seaward land-use types (the next are described as landward type or seaward type), including three association patterns between one landward type and the two seaward types (tidal flat and shrimp pond) and six association patterns between two landward types and the one seaward type (tidal flat or shrimp pond) (Figure 4A). The landward types in association patterns that contain two seaward types are river, settlement and dry land, and among the association patterns that contain one seaward type, there are four patterns that contain shrimp pond and two patterns that contain tidal flat. There are only two association patterns between landward types.

Table 2. Interesting association rules for three-kind land-use types (S > 40, C > 60, Kulc not equal to 0.5).

ID	Rule	S (%)	C (%)	Kulc	AS			AL (m)		
	T1 T2 -> T3				T1	T2	T3	T1-T2	T1-T3	T2-T3
1	Se SP -> DL	67.20	92.01	92.01	25	12	30	9668	4770	12,200
2	DL SP -> Se	67.20	87.13	87.13	30	12	25	12,200	4770	9668
3	DL Se -> SP	67.20	87.02	87.02	30	25	12	4770	12,200	9668
4	SP TF -> DL	54.31	95.61	95.61	11	1	26	6727	11,888	18,355
5	DL TF -> SP	54.31	84.85	84.85	26	1	11	18,355	11,888	6727
6	DL SP -> TF	54.31	70.41	70.41	26	11	1	11,888	18,355	6727
7	Se Ri -> DL	53.88	93.01	93.01	27	27	33	6271	4788	5800
8	DL Ri -> Se	53.88	84.25	84.25	33	27	27	5800	4788	6271
9	DL Se -> Ri	53.88	69.77	69.77	33	27	27	4788	5800	6271
10	SP Ri -> DL	52.94	89.14	89.14	9	23	30	8771	12,166	5864
11	DL Ri -> SP	52.94	82.78	82.78	30	23	9	5864	12,166	8771
12	DL SP -> Ri	52.94	68.64	68.64	30	9	23	12,166	5864	8771
13	Se TF -> DL	52.66	98.76	98.76	28	1	27	15,983	4428	18,486
14	DL TF -> Se	52.66	82.28	82.28	27	1	28	18,486	4428	15,983
15	DL Se -> TF	52.66	68.19	68.19	27	28	1	4428	18,486	15,983
16	Se Ri -> SP	50.12	86.52	86.52	21	22	9	6254	9195	8733
17	SP Ri -> Se	50.12	84.39	84.39	9	22	21	8733	9195	6254
18	Se SP -> Ri	50.12	68.62	68.62	21	9	22	9195	6254	8733
19	Se TF -> SP	47.95	89.94	89.94	24	1	12	15,314	9603	6568
20	SP TF -> Se	47.95	84.42	84.42	12	1	24	6568	9603	15,314
21	Se SP -> TF	47.95	65.66	65.66	24	12	1	9603	15,314	6568
22	Ri TF -> DL	46.16	94.97	94.97	25	1	31	15,223	5862	19,330
23	DL Ri -> TF	46.16	72.19	72.19	31	25	1	5862	19330	15,223
24	DL TF -> Ri	46.16	72.13	72.13	31	1	25	19,330	5862	15,223
25	Se Re/AP -> DL	41.41	96.28	96.28	24	23	30	6195	4878	4982
26	DL Re/AP -> Se	41.41	85.52	85.52	30	23	24	4982	4878	6195
27	Ri TF -> SP	41.36	85.09	85.09	21	1	9	14,366	8258	6818
28	SP TF -> Ri	41.36	72.83	72.83	9	1	21	6818	8258	14,366
29	SP Ri -> TF	41.36	69.65	69.65	9	21	1	8258	6818	14,366
30	SP Re/AP -> DL	40.56	92.49	92.49	8	20	27	11,295	12619	5054
31	DL Re/AP -> SP	40.56	83.77	83.77	27	20	8	5054	12619	11,295

The sequential relationship of three-kind sequential pattern is explicitly shown in Figure 4A, which is similar with the two-kind pattern, except that of the pattern of "settlement, tidal flat <-> dry land". In the pattern of "settlement, tidal flat <-> dry land", dry land is in front of settlement. The distance between land-use types in the three-kind patterns is also in accordance with the AS, except the distance between settlement and dry land. In patterns that contains settlement and dry land, Figure 4A,B show the larger difference of AS between them and the shorter distance between them, respectively.

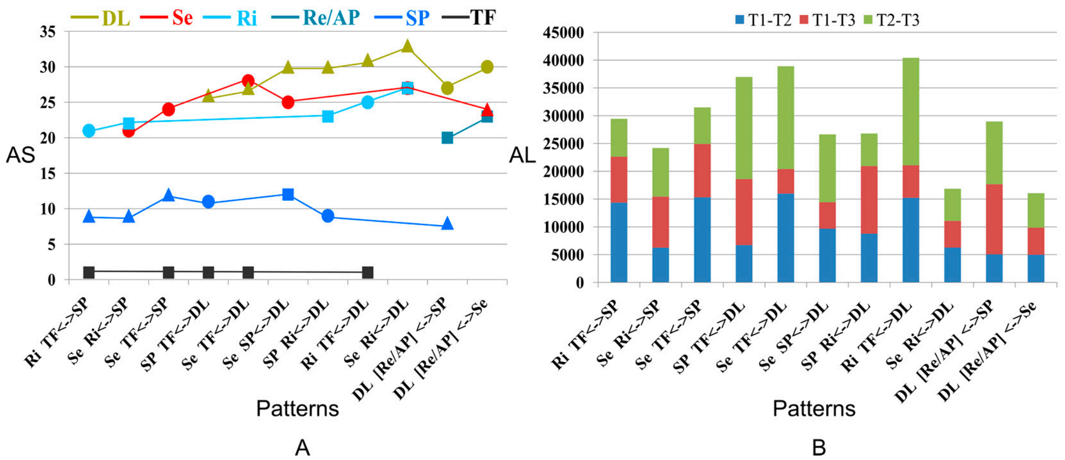


Figure 4. The AS (**A**) and AL (**B**) of land-use types for three-kind association pattern. ("T1 T2 <-> T3"; one association pattern indicates that any two land-use types among T1, T2, and T3 have strong association with the other one; "T1 [T2] <-> T3"; one association pattern indicates that T2 and T1 or T3 has a strong association with T3 or T1; the circle indicates T1. The rectangle (**A**) indicates T2. The triangle (**A**) indicates T3.

There were eight interesting land-use association rules (Table 3) between four land-use types among the five different land-use types, including dry land, settlement, river, shrimp pond and tidal flat. The landward land-use types are dry land, settlement and river. The seaward land-use types are still tidal flat and shrimp pond, and there are only two association patterns (Figure 5). One is the "settlement, shrimp pond, tidal flat <-> dry land" with the sequence of tidal flat > shrimp pond> settlement > dry land. The other one is the "settlement, shrimp pond, river <-> dry land" with sequence of shrimp pond > settlement > river> dry land. In this sense, the sequential relationships between land-use types in these two patterns are similar with the two-kind and three-kind patterns, and the distance between land-use types in these two patterns also tend to be in accordance with the AS.

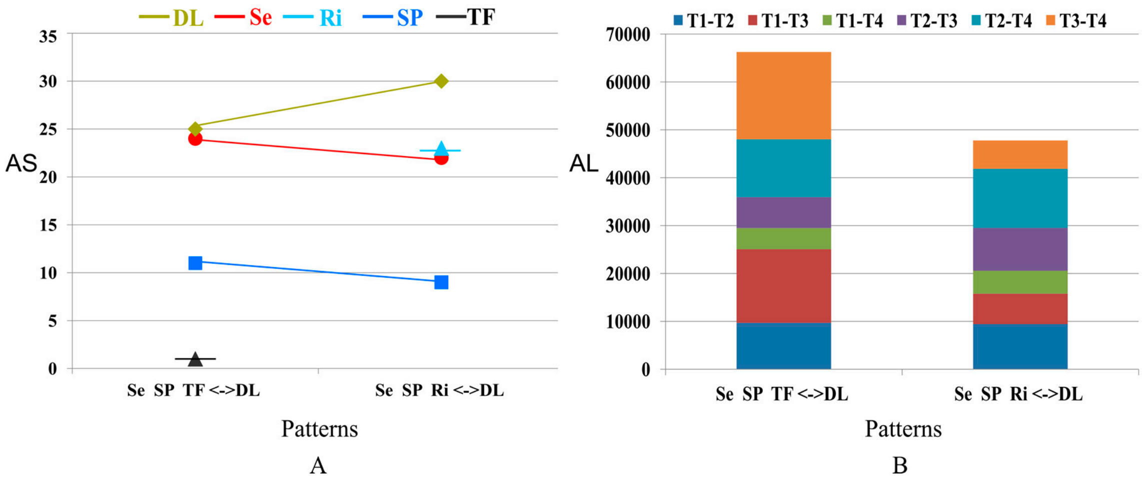


Figure 5. The AS (**A**) and AL (**B**) of land-use types for the four-kind association pattern. ("T1 T2 T3 <-> T4"; one association pattern indicates that any three land-use types among T1, T2, T3, and T4 have strong association with the other one. The circle indicates T1. The rectangle (**A**) indicates T2. The triangle (**A**) indicates T3. The rhombus indicates T4).

Table 3. Interesting association rules for four-kind land-use types (S > 40, C > 60, Kulc not equal to 0.5).

ID	Rule	S (%)	C (%)	Kulc
1	Se SP TF -> DL	47.29	98.63	0.75
2	DL TF -> SP	47.29	89.81	0.73
3	DL SP TF -> Se	47.29	87.09	0.72
4	DL Se SP -> TF	47.29	70.38	0.71
5	Se SP Ri -> DL	46.07	91.92	0.71
6	DL SP Ri -> Se	46.07	87.02	0.71
7	DL Se Ri -> SP	46.07	85.50	0.70
8	DL Se SP -> Ri	46.07	68.56	0.67

4.2. The Distribution of the Sequential Patterns

In Figure 6, the distributions of two-kind sequential patterns are shown, respectively. The sequential patterns with only landward land-use types are shown in Figure 6A–F. The patterns with reservoir/artificial pond (Figure 6A–C) tend to distribute around the west of Bohai bay and the Yellow River delta and few distribute around Leting county, Luannan county, and Tanghai county. The patterns with river have similar distribution, which is near to the port or river (Figure 6D,E). The pattern of “settlement > dry land” is widely distributed around the study area (Figure 6F). The sequential patterns with shrimp pond are shown in Figure 6G–J. These patterns are also widely distributed around the area except, the area in Dagang district and the area around the old Yellow River estuary. The sequential patterns with tidal flats are shown in Figure 6K–N. These patterns also have wide distribution around the area, except the port areas, such as Tangshan port, Tianjin port, and Huanghu port.

In Figure 7, the distributions of three-kind sequential patterns are shown, respectively. The sequential patterns with only landward land-use types are shown in Figure 7A,B. These patterns show a scattered distribution around the area and, especially in the area around the Yellow River delta and Tangshan port, there are few. The sequential patterns that contain both shrimp pond and tidal flat are shown in Figure 7C–E. These patterns have segmented distribution around the area, and the distribution is segmented by port and estuary. The distributions of sequential patterns with one seaward type—shrimp pond—are shown in Figure 7F–I. These patterns are also segmented by port and estuary and tend to be more scattered. The distributions of the sequential patterns with one seaward type—tidal flat—are similar with patterns shown in Figure 7C–I, but the old Yellow River estuary has some distributions.

The sequential pattern of “settlement, shrimp pond, tidal flat <-> dry land” is obviously segmented by port and estuary areas. The pattern of “settlement, shrimp pond, river <-> tidal flat” has fragmented distribution around the area and is mainly distributed around the areas of Hangu district, Ninghe county, Huanghua city, and Haixing county.

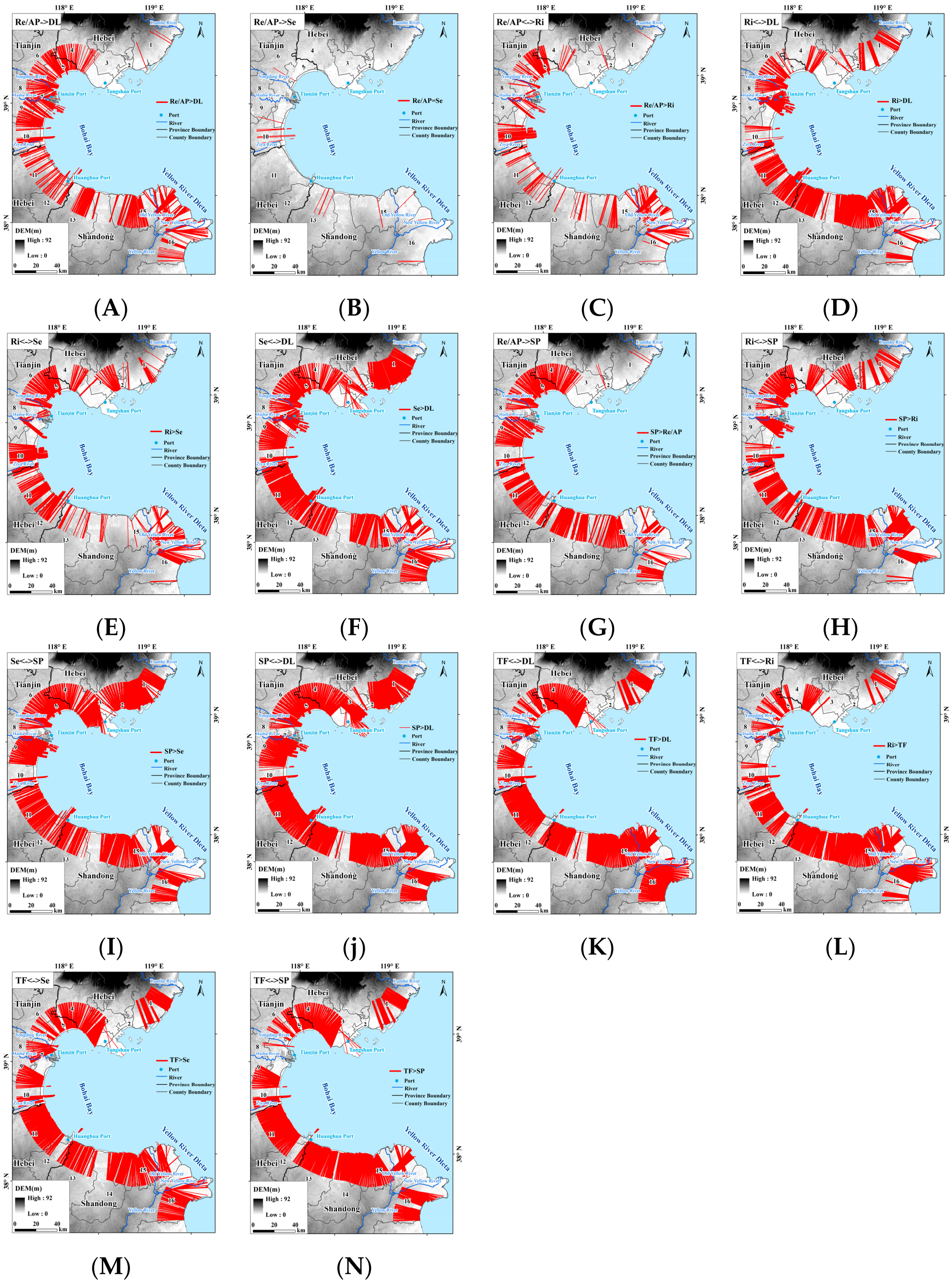


Figure 6. Distributions of two-kind sequential patterns. "T1 > T2"; one sequential pattern indicates that T1 is in the front of T2 starting from the sea direction; "T1 =T2"; one sequential pattern indicates that T1 and T2 nearly distribute around the same region in the sea-land direction. (**A**) distribution of Re/AP > DL ; (**B**) distribution of Re/AP = Se ; (**C**) distribution of Re/AP > Ri; (**D**) distribution of Ri > DL; (**E**) distribution of Ri > Se; (**F**) distribution of Se > DL; (**G**) distribution of SP > Re/AP; (**H**) distribution of SP > Ri; (**I**) distribution of SP > Se (**J**) distribution of SP > DL; (**K**) distribution of TF > DL; (**L**) distribution of TF > Ri; (**M**) distribution of TF > Se; (**N**) distribution of TF > SP; (In the map, the number of 1 indicates Leting county; 2 indicates Luannan county; 3 indicates Tanghai county; 4 indicates Feangnan district; 5 indicates Hangu district; 6 indicates Ninghe county; 7 indicates Tanggu district; 8 indicates Dongli district; 9 indicates Jinnan district; 10 indicates Dagang district; 11 indicates Huanghua city; 12 indicates Haixing county; 13 indicates Wudi county; 14 indicates Zhanhua county; 15 indicates Hekou district; 16 indicates Kenli county.)

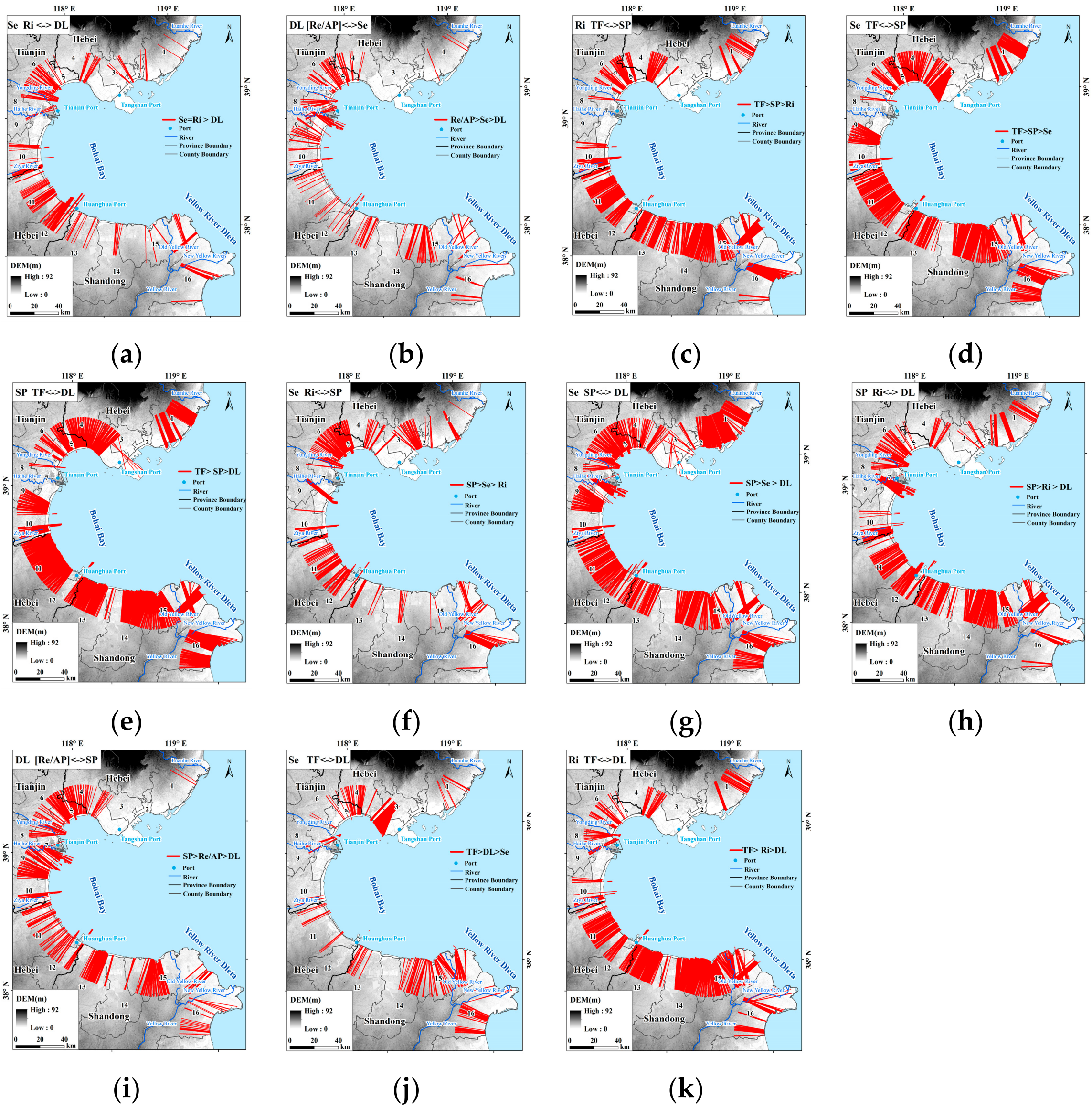


Figure 7. Distributions of three-kind sequential patterns. "T1 > T2 > T3"; one sequential pattern indicates that the sequence of T1, T2 and T3 is T1, T2, T3 orderly starting from the sea direction; (**A**) distribution of Se=Ri > DL ; (**B**) distribution of Re/AP > Se > DL ; (**C**) distribution of TF > SP > Ri; (**D**) distribution of TF > SP > Se; (**E**) distribution of TF > SP > DL; (**F**) distribution of SP > Se > Ri; (**G**) distribution of SP > Se > DL; (**H**) distribution of SP > Ri > DL; (**I**) distribution of SP > Re/AP > DL (**J**) distribution of TF > DL > Se; (**K**) distribution of TF > Ri > DL; (In the map, the number of 1 indicates Leting county; 2 indicates Luannan county; 3 indicates Tanghai county; 4 indicates Feangnan district; 5 indicates Hangu district; 6 indicates Ninghe county; 7 indicates Tanggu district; 8 indicates Dongli district; 9 indicates Jinnan district; 10 indicates Dagang district; 11 indicates Huanghua city; 12 indicates Haixing county; 13 indicates Wudi county; 14 indicates Zhanhua county; 15 indicates Hekou district; 16 indicates Kenli county.)

5. Discussion

With the case study of the coastal zone of Bohai bay and the Yellow River delta, the sequential relationship between land-use types ranging from two-kinds to four-kinds was extracted (Figures 3–5). Associations between land-use types, such as landward type and landward type, landward and seaward type, were mined.

As the results mentioned in Section 4.1, the seaward land-use types had obvious boundaries with landward land-use types. In this study, the boundary was defined by the AS with a value of 15 (Figures 3A, 4A and 5A). The seaward land-use types were mainly tidal flat and shrimp pond, which were in accordance with the fact that these land-use types mainly distribute around the coastline. The seaward land-use types were located in the front of the landward land-use types. The landward land-use types mainly were reservoir/artificial pond, settlement, river and dry land. Although the sequential relationship between dry land and settlement varies with patterns, the sequence of these four land-use types tended to be reservoir/artificial pond > settlement > river > dry land in the sea–land direction. The sequential relationship between land-use types were both landward and seaward types is tidal flat > shrimp pond > reservoir/artificial pond > settlement > river > dry land. Comparing with results of the traditional method [3,20,29], the ARCLUSSM could mine the sequential relationship between land-use types definitely and easily, and the distance between land-use types was also mined (Figures 3B, 4B and 5B). With the sequential relationship and distance between land-use types, the distribution of different land-use types is explicitly shown. For example, for the pattern of "settlement, shrimp pond, tidal flat <-> dry land", the sequence of it in the sea–land direction was tidal flat > shrimp pond > settlement > dry land (Figure 5) and the distance between adjacent land-use types are 6483 m, 9692 m, and 4396 m, respectively (Table 4). In this sense, it indicated that tidal flat was nearer to the coastline, and then the shrimp pond was located in the area, 6483 meters to the coastline, and the same to settlement and dry land types. This detailed information would offer specific support for land use planning and rehabilitation of the coastal zone.

Table 4. AS and AL of four-kind interesting association rules.

ID	Rule	AS				AL (m)					
	T1 T2 T3 ->T4	**T1**	**T2**	**T3**	**T4**	**T1-T2**	**T1-T3**	**T1-T4**	**T2-T3**	**T2-T4**	**T3-T4**
1	Se SP TF -> DL	24	11	1	25	9692	15,383	4396	6483	12,058	18,243
2	DL Se TF -> SP	25	24	1	11	4396	18,243	12,058	15,383	9692	6483
3	DL SP TF -> Se	25	11	1	24	12,058	18,243	4396	6483	9692	15,383
4	DL Se SP -> TF	25	24	11	1	4396	12,058	18,243	9692	15,383	6483
5	Se SP Ri -> DL	22	9	23	30	9431	6337	4791	8962	12,322	5937
6	DL SP Ri -> Se	30	9	23	22	12,322	5937	4791	8962	9431	6337
7	DL Se Ri -> SP	30	22	23	9	4791	5937	12,322	6337	9431	8962
8	DL Se SP -> Ri	30	22	9	23	4791	12,322	5937	9431	6337	8962

With respect to the result described in Section 4.1, the landward land-use types have strong association with the seaward land-use types. With a total of 27 association patterns mined above, there were 19 association patterns that included seaward land-use types, including eight patterns in 14 two-kind patterns, nine patterns in 11 three-kind patterns, and two four-kind patterns (Figure 3A, Figure 4A, and Figure 5A). These association patterns were also in accordance with the natural environment and economic conditions in the areas. Of the patterns that contained seaward land-use types, few were distributed around the areas of ports and estuaries (Figure 6G–N, Figure 7C–I, and Figure 8) because the port areas generally have few tidal flats or shrimp ponds and the areas near to the estuary have few land ward land-use types [30]. According to these associations, the relation between inland land use and land use near to the coastline was linked. For example, as the strong association between dry land and tidal flat (Table 1), which indicated that the dry land may have some impacts on the tidal flat, or the other way around. These associations may help to reveal the interactions between inland land use and the ocean.

Above all, the ARCLUSSM developed in this study is an effective tool to mine the sequential relationship and association between land-use types in the sea–land direction. The results will offer specific support for land use planning and rehabilitation of the coastal zone and may help to reveal the interactions between inland land use and the ocean. However, there are also some insufficiencies.

Firstly, the actual sequence of land use in certain regions of the coastal zone is uncovered. Secondly, the information that implied associations between landward and seaward land-use types should have been deeply studied. These insufficiencies are also the focus of our future work.

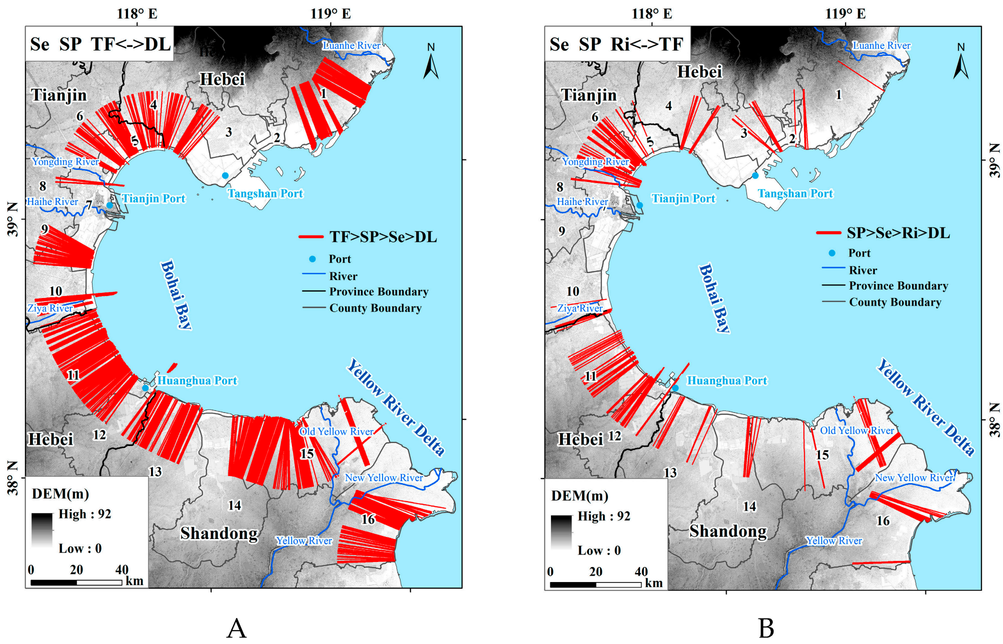


Figure 8. Distributions of four-kind patterns of TF > SP > Se > DL (**A**) and SP > Se > Ri > DL (**B**).

6. Conclusions

In this study, the ARCLUSSM was presented, which can mine the associations between land-use types, especially the landward and seaward land-use types, and the sequential relationship and distance between land-use types in the sea–land direction in the coastal zone. In ARCLUSSM, the coastal zone was simplified in a novel manner as equidistant lines that are perpendicular to the coastline. Appropriate indexes of AS and AL were defined to calculate the sequence and distance between land-use types with the association rule mining.

Evidently, many assumptions underlie the application for the coastal zone of Bohai Bay and the Yellow River delta. Firstly, 27 interesting association patterns of land use in the sea–land direction of the coastal zone were mined easily. Both sequential relationship and the distance between land-use types for 27 patterns among six land-use types were mined definitely, and the sequence of the six land-use types tended to be tidal flat > shrimp pond > reservoir/artificial pond > settlement > river > dry land in sea–land direction. These patterns could offer specific support for land use planning and rehabilitation of the coastal zone. There were 19 association patterns between seaward and landward land-use types. These patterns showed strong associations between seaward and landward land-use types. This indicated that the landward land use might have some impacts on the seaward land use, or in the other direction, which may help to reveal the interactions between inland land use and the ocean.

Thus, the ARCLUSSM was an efficient tool to mine the sequential relationship and distance between land-use types with interesting association rules in the sea–land direction, which would offer practicable advice to appropriate coastal zone management and planning, and might reveal the interactions between inland land use and the ocean.

Acknowledgments: This research was supported by the National Natural Science Foundation of China (No. 41271409) and Youth Science Funds of LREIS (O8R8A080YA), CAS, grants (41501473) funded by the National Science Foundation of China.

Author Contributions: Zhi Ding, Xiaohan Liao and Fenzhen Su conceived of and designed the study; Zhi Ding made substantial contributions to the acquisition, analysis, and interpretation of the data; Dongjie Fu reviewed and edited the draft. All authors have read and approved the submitted manuscript, have agreed to be listed as authors, and have accepted the final version for publication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ray, G.C. Coastal-zone biodiversity patterns. *Bioscience* **1991**, *41*, 490–498. [CrossRef]
2. Yongming, L. Sustainability associated coastal eco-environmental problems and coastal science development in China. *Bull. Chin. Acad. Sci.* **2016**, *31*, 1133–1142.
3. Hou, X.Y.; Xu, X.-L. Spatial patterns of land use in coastal zones of China in the early 21st century. *Geogr. Res.* **2011**, *30*, 1370–1379.
4. Xu, Y.; Pu, L.-J.; Zhang, R.S.; Zhu, M. Spatial-temporal dynamics of land use and land cover change in the coastal zone of Jiangsu Province. *Resour. Environ. Yangtze Basin* **2012**, *21*, 565–571.
5. Wu, L.; Hou, X.Y. Analysis of land use change in the coastal zone of Circum the Bohai sea region during 2000 to 2010. *Mar. Sci.* **2015**, *39*, 101–110.
6. Poulos, S.E.; Chronis, G.T.; Collins, M.B.; Lykousis, V. Thermaikos Gulf coastal system, NW Aegean Sea: An overview of water/sediment fluxes in relation to air–land–ocean interactions and human activities. *J. Mar. Syst.* **2000**, *25*, 47–76. [CrossRef]
7. Serra, P.; Pons, X.; Saurí, D. Land-cover and land-use change in a Mediterranean landscape: A spatial analysis of driving forces integrating biophysical and human factors. *Appl. Geogr.* **2008**, *28*, 189–209. [CrossRef]
8. De Zoysa, M. Casuarina coastal forest shelterbelts in Hambantota City, Sri Lanka: Assessment of impacts. *Small Scale For.* **2008**, *7*, 17–27. [CrossRef]
9. Alix, D.M.; Anderson, C.J.; Grand, J.B.; Guyer, C. Evaluating the effects of land use on headwater wetland amphibian assemblages in coastal Alabama. *Wetlands* **2014**, *34*, 917–926. [CrossRef]
10. Correll, D.L.; Jordan, T.E.; Weller, D.E. Nutrient flux in a landscape: Effects of coastal land use and terrestrial community mosaic on nutrient transport to coastal waters. *Estuar. Coasts* **1992**, *15*, 431–442. [CrossRef]
11. Gorman, D.; Russell, B.D.; Connell, S.D. Land-to-sea connectivity: Linking human-derived terrestrial subsidies to subtidal habitat change on open rocky coasts. *Ecol. Appl.* **2009**, *19*, 1114–1126. [CrossRef] [PubMed]
12. Gao, G.D.; Wang, X.H.; Bao, X.W. Land reclamation and its impact on tidal dynamics in Jiaozhou Bay, Qingdao, China. *Estuar. Coast. Shelf Sci.* **2014**, *151*, 285–294. [CrossRef]
13. Dauer, D.M.; Ranasinghe, J.A. Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay. *Estuar. Coasts* **2000**, *23*, 80–96. [CrossRef]
14. Syvitski, J.P.M.; Harvey, N.; Wolanski, E.; Burnett, W.C.; Perillo, G.M.E.; Gornitz, V.; Arthurton, R.K.; Bokuniewicz, H.; Campbell, J.W.; Cooper, L. Dynamics of the coastal zone. In *Coastal Fluxes in the Anthropocene*; Global Change—The IGBP Series; Springer: Berlin/Heidelberg, Germany, 2006; pp. 39–94.
15. Hadley, D.; Beddington, J. Land use and the coastal zone. *Land Use Policy* **2009**, *26* (Suppl. 1), S198–S203. [CrossRef]
16. Cao, W.; Wong, M.H. Current status of coastal zone issues and management in China: A review. *Environ. Int.* **2007**, *33*, 985–992. [CrossRef] [PubMed]
17. Islam, M.A.; Mitra, D.; Dewan, A.; Akhter, S.H. Coastal multi-hazard vulnerability assessment along the Ganges deltaic coast of Bangladesh—A geospatial approach. *Ocean Coast. Manag.* **2016**, *127*, 1–15. [CrossRef]
18. Nicholls, R.J.; Cazenave, A. Sea-level rise and its impact on coastal zones. *Science* **2010**, *328*, 1517–1520. [CrossRef] [PubMed]
19. Xu, X.; Li, X.; Chen, M.; Li, X.; Duan, X.; Zhu, G.; Feng, Z.; Ma, Z. Land-ocean-human interactions in intensively developing coastal zone: Demonstration of case studies. *Eur. J. Neurosci.* **1995**, *133*, 962–971. [CrossRef]
20. Wang, J.; Liu, Y. Tourism-led land-use changes and their environmental effects in the southern coastal region of Hainan Island, China. *J. Coast. Res.* **2016**, *29*, 1118–1125. [CrossRef]
21. Agrawal, R. Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD Conference, Washington, DC, USA, 25–28 May 1993; Volume 22, pp. 207–216.

22. Kotsiantis, S.; Kanellopoulos, D. Association rules mining: A recent overview. *Int. Trans. Comput. Sci. Eng.* **2005**, *32*, 71–82.
23. Han, J. *Data Mining: Concepts and Techniques*; Morgan Kaufmann Publishers: Burlington, MA, USA, 2006; pp. 1–18.
24. Tan, P.N.; Kumar, V.; Srivastava, J. Selecting the right interestingness measure for association patterns. In Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, AB, Canada, 23–26 July 2002; pp. 32–41.
25. Zhang, T.T.; Zeng, S.L.; Gao, Y.; Ouyang, Z.T.; Li, B.; Fang, C.M.; Zhao, B. Assessing impact of land uses on land salinization in the Yellow River Delta, China using an integrated and spatial statistical model. *Land Use Policy* **2011**, *28*, 857–866. [CrossRef]
26. Deng, X.; Huang, J.; Rozelle, S.; Uchida, E. Cultivated land conversion and potential agricultural productivity in China. *Land Use Policy* **2006**, *23*, 372–384. [CrossRef]
27. Liu, J.; Liu, M.; Zhuang, D.; Zhang, Z.; Deng, X. Study on spatial pattern of land-use change in China during 1995–2000. *Sci. China Ser. D Earth Sci.* **2003**, *46*, 373–384.
28. Liu, J.; Zhang, Z.; Xu, X.; Kuang, W.; Zhou, W.; Zhang, S.; Li, R.; Yan, C.; Yu, D.; Wu, S. Spatial patterns and driving forces of land use change in China during the early 21st century. *J. Geogr. Sci.* **2010**, *20*, 483–494. [CrossRef]
29. Wu, C.; Cai, Q.; Zhu, J. *China's Coastal Zone Land Use*; Ocean Press: Beijing, China, 1993. (In Chinese)
30. Zhu, G.; Xu, X.; Ma, Z.; Xu, L.; Porter, J.H. Spatial dynamics and zoning of coastal land-use change along Bohai Bay, China, during 1979–2008. *J. Coast. Res.* **2016**, *284*, 1186–1196. [CrossRef]