International Journal of *Geo-Information*

Article

Evaluating School Location Based on a Territorial Spatial Planning Knowledge Graph

Xiankang Xu [1,2], Jian Hao [1,2] and Jingwei Shen [1,2,*]

[1] Chongqing Jinfo Mountain Karst Ecosystem National Observation and Research Station, School of Geographical Sciences, Southwest University, Chongqing 400715, China; xxkgis2021@email.swu.edu.cn (X.X.); hj1358168939@email.swu.edu.cn (J.H.)
[2] Key Laboratory of Monitoring, Evaluation and Early Warning of Territorial Spatial Planning Implementation, Ministry of Natural Resources, Chongqing 400715, China
* Correspondence: sjwgis@swu.edu.cn

Abstract: The reasonable spatial planning of primary and secondary schools is an important factor in education development. In spatial planning, there are many models for the locations of primary and secondary schools; however, few quantitative evaluation models are available. Therefore, based on the many factors affecting the layout planning of primary and secondary schools, a knowledge graph of territorial spatial planning that considers the topological relationship, direction relationship and metric relationship in spatial planning is designed and constructed. A school location evaluation model based on the knowledge graph of territorial spatial planning is proposed. The model combines many factors of the locations of schools, such as the service population, the impact of factories on schools, the adjacency and centrality of school plots, terrain and existing schools in the region, to quantitatively evaluate whether schools are reasonably located within a region. This study focuses on the Guangyang Island area in Chongqing, China, exploring the superiority and rationality of the planned land use for primary and secondary schools within the region. By analyzing the top three and bottom three ranked schools in conjunction with the actual conditions of the site, and comparing them with AHP hierarchical analysis and ArcGIS modelling research, the study concludes that the results of this model are highly reasonable within the scope of China's territorial spatial planning.

Keywords: school location; knowledge graph; territorial spatial planning; topological relationship

check for updates

Citation: Xu, X.; Hao, J.; Shen, J. Evaluating School Location Based on a Territorial Spatial Planning Knowledge Graph. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 173. https://doi.org/10.3390/ijgi13060173

Academic Editors: Jiangfeng She, Jun Zhu, Min Yang and Wolfgang Kainz

Received: 28 March 2024
Revised: 19 May 2024
Accepted: 23 May 2024
Published: 24 May 2024

1. Introduction

Spatial planning refers to the long-term planning and overall arrangement of land resources and spatial layout within the jurisdiction of a national, regional government or at local levels [1,2]. Local spatial planning is essential for the detailed organization of communities, including the planning of schools, residential areas and local infrastructure. While national and regional levels often focus on strategic plans, local planning ensures that development meets the specific needs and context of individual communities [3,4]. Spatial planning can be considered the soul for the development of a country or region. Reasonable spatial planning is an important tool for promoting regional coordinated development and sustainable development [4]. As primary and secondary schools are an important part of the educational infrastructure, their layout planning is an important part of territorial spatial planning and urban spatial structure. Reasonable primary and secondary school layout planning is crucial to the development of national education and sustainable development. As important public facilities, primary and secondary schools emphasize the social benefits and simultaneously pay attention to the equal rights enjoyed by citizens [5]. Schools symbolize the quality of life of urban residents and have an important role in improving their livelihood and quality of life. A reasonable layout of schools considers the vital interests of the public, directly affects the study, work and life of teachers and students in schools, and it is critical to the process of urban construction and development [6].

However, the current layout of many primary and secondary schools may be unreasonable, such as being too close to factories and having an insufficient population to serve.

The influencing factors of school layout involve many aspects, and each factor affects the others. The following points should be considered when choosing the location of a school [7–10]: (1) the school should have enough service to cover as many residents as possible within a certain commuting range, and residents can choose to enroll in nearby schools; (2) the straight-line distance between the school and factories should be short, and the 500 m around the school should not be built on; (3) the school should be equipped with existing public service facilities that are easily accessible; (4) the school should be located on flat terrain; (5) there should not be too many schools of the same type nearby, so resources should be reasonably allocated. If there is a chemical plant or other dangerous factory, it will be farther. The adverse effects of factory gas and noise on schools should be avoided. All of the criteria mentioned are for areas that are in a position to meet these criteria, whereas in the case of school siting in areas with steeper slopes, the importance of a slope in the final siting analysis is weaker compared to the other conditions, and it is possible to adjust the weighting of each criterion in the model according to the actual situation. There are many geographical elements involved in school location evaluation, which have complex and poorly related interrelationships that are often disregarded, so it is difficult to conduct a comprehensive analysis. Therefore, it is necessary to consider which model to use to evaluate whether the layout of primary and secondary schools is reasonable. For example, GIS location-allocation models for improving the accessibility to primary schools in Mansura city, Egypt [9] use the spatial statistical analysis tools to model and analyze school locations. A multi-objective optimization model for school location-allocation coupling of demographic changes [11] proposes a multi-objective optimization model of school location considering the population changes. An evaluation of primary schools and their accessibility using GIS techniques [12] uses parameters including literacy rate, student gender ratio, teacher taught ratio and student to toilet ratio to evaluate schools and uses network analysis to optimize the path.

In the site selection analysis process, there are typically multiple factors that are mutually conflicting or interdependent. The application of multi-criteria analysis (MCA) as a decision support tool in site selection analysis is crucial. It can better deal with multiple conflicting or interdependent factors, helping decision makers to provide a structured approach to complex decision problems when faced with multiple choices, making the decision-making process more transparent, rational and scientific. In recent years, there have been many scholars who used MCA to deal with the conflicts between the multiple factors for industrial site selection [13] and for wind energy power site selection [14]. Prasetyo et al. [7] have proposed a weighted superposition of six spatial factors: administration, population, transportation, land use, student mobility and public preference for school location, and the Analytic Hierarchy Process (AHP) weights each factor in the multi-criteria decision analysis to select school sites. Additionally, other scholars have developed a primary school site selection model using Geographic Information Systems (GIS) and a multi-criteria evaluation model (MCEM) [15]. These existing applications provide important ideas for the site selection evaluation in this study. In general, terrain, service population, existing schools, land planning, spatial pattern, demand analysis and other factors are considered by many models to select the locations of schools, so how to deal with these factors to establish an effective school siting evaluation model still needs further exploration. The school location model is employed for selecting the optimal geographical location of a school before planning and constructing a school, whereas the school evaluation model is based on the planned school in territorial spatial planning to evaluate the geographical location of the school. Due to the poor correlation of various factors, there is no effective integration and comprehensive analysis. In addition, the adjacency, centrality and influence of the factories of school plots should be considered in spatial planning, and the relevant qualitative research discussion is more comprehensive than quantitative research, which makes it difficult to provide the necessary technical means and measurement standards for local school layout planning.

The knowledge graph can efficiently express the intricate relationship between two geographic element entities, while the building of a knowledge graph of territorial spatial planning and the optimization of the school site selection model are new ideas. The term "knowledge graph" has been referred to in the literature since at least 1972 [16], but the modern and publicly known knowledge graph was proposed by Google in 2012 and extensively applied in the field of information search [17]. The knowledge graph has become a widely utilized term in academia and industry and plays an important role in intelligent question answering, intelligent decision-making and other applications [18]. A knowledge graph is essentially a knowledge base that is referred to as a semantic net as proposed by Richens [19], that is, a knowledge base with a directed graph structure, and it is defined as a graph of data intended to accumulate and convey the knowledge of the real world [20]. The nodes of the graph represent entities or concepts, and the edges of the graph represent the potentially different relations between these entities.

The advantage of the knowledge graph [18–20] is that it can effectively integrate different types of data, express knowledge and share data in a way that is easier for humans and machines to read. A knowledge graph has a strong relationship expression ability and can handle complex and diverse association analysis based on graph models. A knowledge graph has a fast response speed for computing queries and efficient feedback mechanisms and uses efficient graph algorithms, such as centrality algorithms, path search algorithms and community discovery algorithms. The entity or concept in a knowledge graph can be any real-world object or abstract concept, and the relationship can be any relationship between them that contains types and properties with a well-defined meaning [21]. For example, in recent years, knowledge was extracted from electronic medical records and stroke diseases to build medical knowledge graphs [22,23], spatial topological relationships were extracted to build spatial scene knowledge graphs [24], and geographic entities were aligned from historical maps to build knowledge graphs [25]. It is a major challenge and opportunity to effectively acquire knowledge from territorial spatial planning in a structured mode to construct a knowledge graph to form a knowledge representation model that can be utilized and understood by computers.

Therefore, an effective approach to evaluate the locations of primary and secondary schools is attempted to be proposed in this article. The innovation of this article is that a territorial spatial planning knowledge graph considering the relations of topology, direction and metrics in spatial relations is constructed. Based on this knowledge graph, a school location evaluation model that integrates various factors, such as terrain, direction, distance, centrality and adjacency, is proposed. The remainder of the article is outlined as follows: The study area and data sources are introduced in Section 2. The construction of the model and knowledge graph is introduced in Section 3. The application results of this model are discussed in Section 4. The conclusions and future research directions are provided in Section 5.

2. Materials and Methods

2.1. Study Area and Data Sources

2.1.1. Study Area

Guangyang Island is located between Mingyue Mountain and Tongluo Mountain on the south bank of Chongqing, China. It is a sandbar island in the upper reaches of the Yangtze River and the largest green island in the main urban area of Chongqing. Guangyang Island, which is only 11 km from the city centre with a population of over 7000 residents, boasts a superior natural ecological environment with dense vegetation, characterized by a subtropical monsoon climate. Rainfall is abundant during the rainy season, with an annual average precipitation of 1163.3 mm. The average sunshine duration per year reaches 1233 h, and the average temperature hovers around 18 degrees Celsius. The frost-free period extends for approximately 342 days. Guangyang Island stands as a crucial nexus for Chongqing's unique river vistas, rich ecological endowments and modern urban conveniences, highlighting the importance of strategic development, sustainable utilization and the robust preservation of this invaluable, yet scarce, resource. In 2019, recognizing the significance of the Guangyang Island area, the Chongqing municipal government conducted land planning for Guangyang Island and its surrounding areas. This included the formulation of the overall urban planning for the Guangyang Island area in Chongqing (https://www.cqrd.gov.cn/, accessed on 20 January 2024). Hence, this paper opts for the ecological and civilizational spatial planning of Guangyang Island, delving into the pertinent issues surrounding the assessment of school site locations. Figure 1 illustrates the geographic location of the Guangyang Island Yangtze River Ecological Civilization Innovation Experimental Zone.

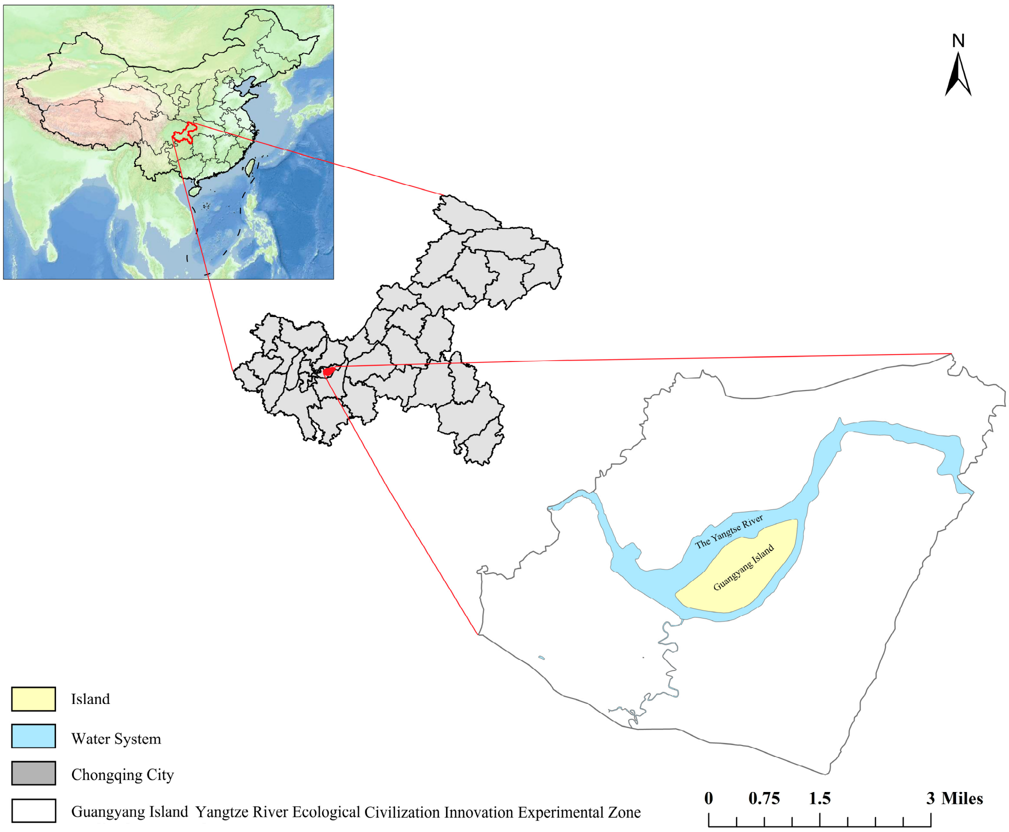

Figure 1. Guangyang Island Yangtze River Ecological Civilization Innovation Experimental Zone.

2.1.2. Data Sources

The data include the digital resources of territorial spatial planning of the study area and 12.5 m DEM data. The sources are as shown in Table 1.

Table 1. Data sources.

Data	Format	Data Sources
Guangyang Island territorial spatial planning	Vector	https://www.cqna.gov.cn (accessed on 20 January 2024)
DEM	Raster (12.5 m)	ALOS (https://search.asf.alaska.edu, accessed on 21 January 2024)

Figure 2a shows the territorial spatial planning map of the study area. Figure 2b shows the road network data of the study area. The road network data in the study area were extracted from the centre line of the digital resource road. Figure 3 shows 12.5 m DEM data of the study area. It can be seen that within the study area, there is a certain difference in elevation due to the presence of mountains in the west and east, with an overall flat location in the centre. Figure 4 shows the primary and secondary schools and the residential and industrial land in the study area. The area of residential land is about 12 km^2, the area of primary and secondary schools is about 0.6 km^2, and the area of industrial land is about 5 km^2.

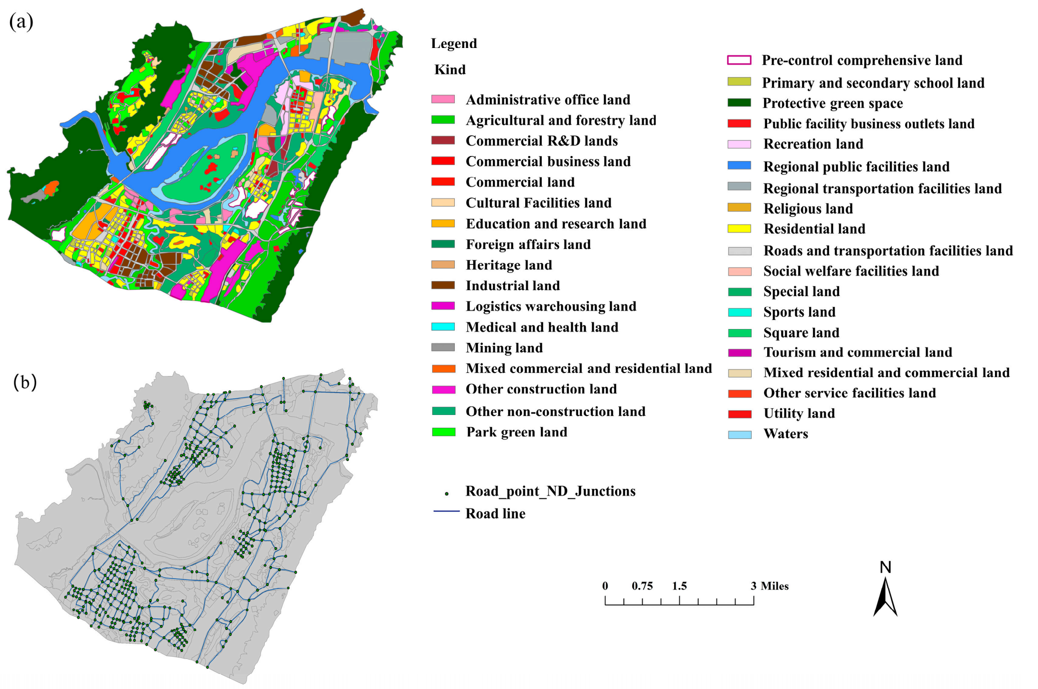

Figure 2. Digital resources of the study area: (**a**) territorial spatial planning map; (**b**) road network.

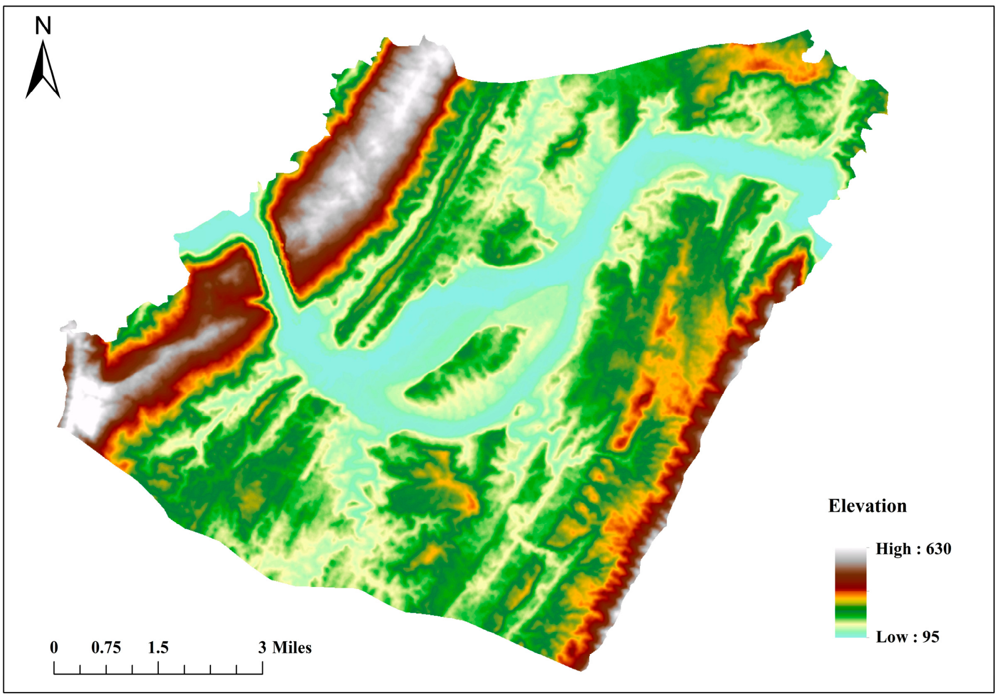

Figure 3. DEM of the study area.

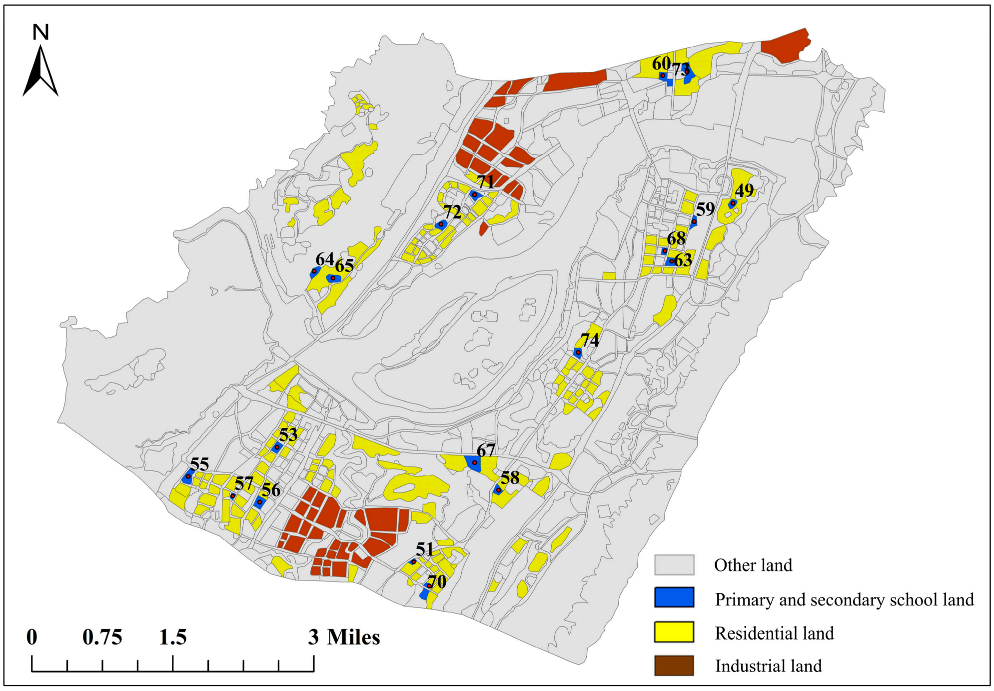

Figure 4. Primary and secondary schools, residential and industrial land in the study area.

2.2. Methods

2.2.1. School Location Evaluation Model

Based on the summary of the five major categories of influencing factors mentioned earlier, a school location evaluation model based on the knowledge graph considering the topological relationships, direction relationships and distance relationships of territorial spatial planning is proposed, as shown in Equation (1).

$$S = \alpha \cdot \frac{m}{M} + \beta \cdot \frac{W_1 \cdot \sum(S_1 \cdot d) + W_2 \cdot \sum k_1}{n_1 \cdot (W_1 + W_2)} + \gamma \cdot C + \delta \cdot S_t + \lambda \cdot \frac{\sum(S_2 \cdot k_2)}{n_2} \quad (1)$$

where S is the final evaluation measure of a school. The measurement range is (0–1). The larger the value, the better the location of the school based on the evaluation of the model. m is the population that the school can cover regarding specific needs; M is the total number of people living in the region; S_1 is the Euclidean distance between the school and the factory; d is the radian value of the absolute value of the wind direction minus the angle between the school and the factory; k_1 is the K-order neighbour value between the factory and the school; n_1 is the total number of factories in the region; C is the Harmonic centrality of the school entity in the topological relation subgraph; S_t is the reciprocal of the slope at which the school is located; S_2 is the shortest road network path between other schools; k_2 is the K-order neighbour value between two schools; n_2 is the total number of schools in the region; $W_1 + W_2 = 1$ and $\alpha + \beta + \gamma + \delta + \lambda = 1$; and $S_1 \cdot d$, $S_2 \cdot k_1$, k_2, C and S_t are the standardized data.

In this article, the winter north wind is used as an example. According to the research on the importance of the influence of each factor [7–10], the weights of α, β, γ, δ, λ, W_1, and W_2 temporarily are determined to be 0.3, 0.2, 0.25, 0.125, 0.125, 0.5 and 0.5, respectively in this article (Table 2).

The adjacency, association and inclusion relations among points, lines and surfaces in space are described as spatial topological relations [26–28]. These relations are important for the storage and expression of spatial data, spatial analysis and practical application. Two-dimensional geographic entities that are not covered and do not overlap are the main concept or entity type in territorial spatial planning. Therefore, only three topological relations—Touches, Contains and Within—regarding surface–surface relationships are considered. Equations (2) and (3) show that the topological relationship between two different two-dimensional geographic objects can be judged by calculating the distance between them and determining whether common edges exist [24]. Figure 5 shows the three spatial topological relationships—Touches, Within and Contains—among different two-dimensional geographic objects a, b and c. Equations (4) and (5) show the symmetry of some topological relations, which can improve the computational efficiency of topological relations. In Equations (2)–(5), a and b are two different two-dimensional geographic entity objects, and L is the length of the common edge. Topological relationships are considered as one type of relationship in the knowledge graph.

$$L = a \cap b \neq \varnothing \neq La \neq Lb \Rightarrow \ a.Touches\ (b) = \ True \tag{2}$$

$$L = a \cap b \neq \varnothing \neq La \neq Lb \Rightarrow \ a.Within\ (b) = \ True \tag{3}$$

$$a.Touches\ (b) = \ True \ \Rightarrow \ b.Touches\ (b) = \ True \tag{4}$$

$$a.Contains\ (b) = \ True \ \Rightarrow b.Within\ (a) = \ True \tag{5}$$

Table 2. Values of weights.

Weights	**Value**
W_1	0.5
W_2	0.5
α	0.3
β	0.2
γ	0.25
δ	0.125
λ	0.125

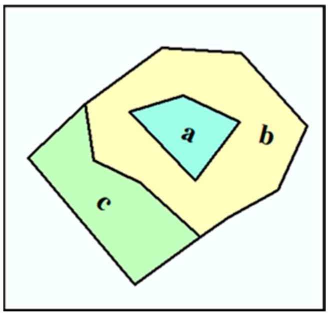

Figure 5. Touches, Contains and Within topological relations between different two-dimensional geographic objects of a, b, c.

The direction relationship, which describes the cardinal directions of a target object with respect to a given reference object, is an important binary spatial relation that describes the spatial location of two different geographical objects in space. It is also an important part of spatial reasoning [29,30]. In this article, when calculating the direction relationship between two different two-dimensional geographical objects, the centre point of the geographical object is used to replace itself, and the direction relationship is measured by 0–360°. The calculation method is shown in Equation (6), where D is the angle, and (x_1, y_1) and (x_2, y_2) are the coordinates of the centre point of the two-dimensional geographic objects. Figure 6 shows the direction relationship among different two-dimensional geographic objects a, b and c. Object d in Equation (1) is obtained by using Equation (6) to calculate the angle between two different geographical objects, subtracting the value and angle of the wind direction, and then calculating the absolute value and converting it to radians.

$$D = \tan^{-1}\left(\frac{y_1 - y_2}{x_1 - x_2}\right) \quad (6)$$

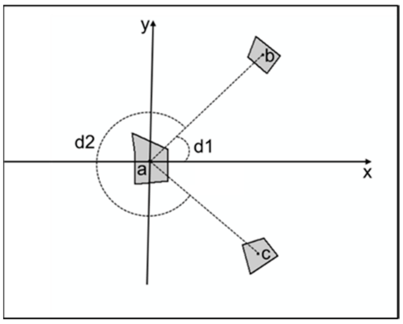

Figure 6. Touches, Contains and Within topological relations between different two-dimensional geographic objects of a, b and c.

Euclidean Distance, Manhattan Distance, Network Distance, etc., are often used to express spatial metric relations. The Euclidean distance in 2D and 3D spaces is the straight-line distance between two points. The network distance is the path distance or cost distance between two points based on an actual network, such as a road network [31]. In this article, only two-dimensional space is involved, so the Euclidean distance is selected to express the distance between two different two-dimensional geographical objects, and the straight-line distance between them is calculated by using the centre point.

In the context of school siting, the dispersion of factory pollutants is not constrained by road networks, making the Euclidean distance a more appropriate and relevant metric. S_1 in Equation (1) is calculated using this method. However, concerning the traffic features in school siting, road networks play a crucial role. Network distance is utilized to represent the commuting distance between two different geographic objects. In a road network, the shortest path distance is the shortest network path length from the starting point through the road network to the end point. S_2 and m in Equation (1) are obtained by using the above method to calculate the network distance. m is the sum of the population served by the accessible residential areas via the shortest path network within a given minimum commuting time.

Two objects are K-order neighbours if one object passes at least k adjacent objects en route to another object [24]. Figure 7 shows the adjacent objects of order 1–3 of a two-dimensional geographic object a. Everything is interrelated, but nearby objects have a greater interrelation [32], that is, the smaller is the k of two geographical objects that are K-order neighbour objects, the greater the influence between them and vice versa. In this model, k is used to measure the adjacency between the different objects, which represents the spatial influence between them. k_1 and k_2 in Equation (1) are computed via the graph shortest path algorithm based on the topological subgraph.

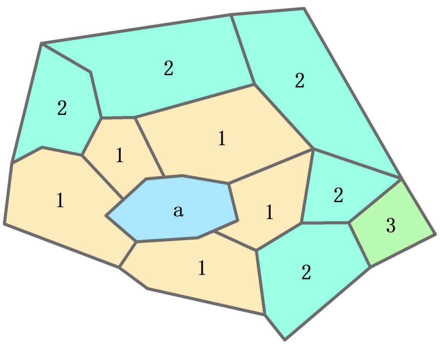

Figure 7. 1–3-order neighbours of geographic object a.

Centrality is applied to determine the importance of distinct nodes in a network [33]. The centrality of a single node is mainly divided into degree, PageRank, betweenness, closeness, etc. Harmonic centrality [34] is a variant of closeness centrality and is proposed to solve the problem caused by the processing of unconnected graphs by closeness centrality. Harmonic centrality is a way to detect nodes that are able to efficiently spread information through a graph. The closeness centrality of a node measures its average farness (inverse distance) to all other nodes. Rather than summing the distances of a node to all other nodes, the harmonic centrality algorithm sums the inverse of those distances. This approach enables it to handle infinite values. Nodes with a high score have the shortest distances to all other nodes. Equation (7) shows the standardized formula of harmonic centrality. In Equation (7), $Nhar$ is a normalized harmonic centrality, and $\sum_n Dis$ is the sum of the reciprocal of the distance from the node to every other node, excluding itself. n is the number of nodes in the graph. Harmonic centrality in the topological relations subgraph of plots is used to measure the difficulty of accessing each plot and to identify plots that have a critical impact. For example, harmonic centrality can be used to determine whether the locations of public services in a city are superior or to reselect the locations of public services. In this model, harmonic centrality is a very important factor in evaluating the location of a school. If the value of harmonic centrality is large, the school can be more easily accessed and more compatible with other plots. The C in Equation (1) is computed via the harmonic centrality algorithm based on the topological relations subgraph of the plots.

$$Nhar = \frac{\sum_n Dis}{n-1} \tag{7}$$

After considering the five influence principles of school location selection and evaluation discussed in Section 1, the topological relationship, direction, metric relationship and proximity are integrated. The slope is used to measure the smoothness of the terrain of the school location. In this study, we use the average slope of plots as the measurement standard. The K-order neighbourhood value and the shortest path between two schools are utilized to measure whether the educational resources are reasonably allocated. The K-order neighbourhood value is calculated based on the topological relationships between plots. The serviceable population during a commuting time is applied to measure the serviceable population of a school. In this article, we assume the coverage range within a 15 min walking distance from the school and use the area of the residential zone within this coverage range to represent the population. The harmonic centrality of the school plots in the topological relation subgraph is used to measure the convenience of accessibility between schools and other plots. Harmonic centrality is calculated based on the graph structure of territorial spatial planning knowledge graphs. When calculating the harmonic centrality, factory plots are excluded because of the negative correlation between factories and schools. The straight-line distance between a school and a factory and the difference between their direction and the wind direction are employed to measure the impact of factory emissions on the school, and the K-order neighbourhood value is used to measure the proximity between the factory and the school. Equation (1) is proposed by weighted linear superposition of five parts. The weight of each part is determined by the influence of each factor on school location. The input of the model is composed of the above five parts and their influence weights. The output is the final evaluation score S. For the model, the larger the S, the better the evaluation.

The implementation process of the quantitative school location evaluation model based on the territorial spatial planning knowledge graph involves the following steps: (1) extraction of land parcel entities and their attributes from territorial spatial planning data; (2) identification of relationships among land parcels, including topological, directional and metric relationships; (3) construction of the territorial spatial planning knowledge graph; (4) storage of the knowledge graph data; (5) based on the knowledge graph, the formulas needed to build the assessment model.

2.2.2. Construction of the Knowledge Graph

The knowledge graph schema is the core of the knowledge graph. It stores the conceptual model abstracted from facts [24]. First, the knowledge graph schema of territorial spatial planning, including the classes, entities, attributes and relationships, should be defined. Class is an abstract concept about geographic objects; entities are instances of classes; attribute is the characteristic of a class; and relationship represents how the classes and entities are associated. The schema of the binary system of the plot class and plot type class is defined. Each plot and plot type is an instance of its own class. Each entity has all the attributes of its class. The relationship between the plot class and the plot type class is that the plot class belongs to the plot type class. Plot class and plot type class also have certain relations.

Knowledge acquisition is aimed at constructing knowledge graphs from unstructured text and other structured or semi-structured sources, completing an existing knowledge graph, and discovering and recognizing entities and relations [21]. It is necessary to acquire the entities and relationships in territorial spatial planning before constructing a knowledge graph.

The entity knowledge should be extracted from the digital resources of the territorial spatial planning map and mean slope map in Section 2 to acquire entities and the attributes of entities. Every plot has the attributes of ID, Plot Type, Mean_slope and Code. The plot serves as the main entity, with entity properties such as ID, Shape_area, Shape_length and Mean_slope. Plot type serves as the information entity for the plot entity, with entity properties such as type code and type. Table 3 shows the entities and entity attributes of the territorial spatial planning knowledge graph.

Table 3. Entities and entity attributes.

Entity	Attributes
Plot	ID, Shape_area, Shape_length, Mean_slope
Plot Type	ID, Code, Type

A knowledge graph is defined as G = (E, R and F), where E, R and F are sets of entities, relations and facts [21]. The triple is a general representation of a knowledge graph. The basic forms of a triple mainly include (entity A, relation and entity B) and (entity, attribute and attribute value). In this article, the entities and their properties are represented by triples such as (a, Mean_slope and 12), which means that the average slope of a is 12°. After acquiring the entity, the various items of knowledge of the relationships in the digital resources of territorial spatial planning need to be extracted.

The specific association between two entities is defined as the entity relationship, which can be regarded as the edge connecting two nodes in the knowledge graph. Spatial relationships refer to relationships with spatial characteristics between two different spatial objects, which are mainly divided into topological, directional and metric relationships. The calculation methods discussed in Section 2.2.1 are employed to acquire the various relationships between two entities of the digital resources of the study area. There is a subordinate relationship with the plot type; this relationship is acquired according to the land-type classification code data. The triple form is also used to store various relationships and relationship attributes.

The Neo4j graph database is selected to store and visualize the territorial spatial planning knowledge graph in the form of a graph. The Neo4j native graph database is an important, open source and efficient NOSQL graph database based on the Property Graph Model. The basic structure includes nodes, relationships and properties. Nodes are connected via relationships to form a network structure [35].

To store data in the Neo4j graph database, the mapping rules between the schema and Neo4j's data structures need to be designed for standardizing the data storage process. The mapping rules are as follows: Nodes in the Neo4j graph database represent specific entity objects in the digital resources of territorial spatial planning. The relationships in the Neo4j graph database are used to connect not only different nodes but also independent nodes to form a knowledge network. The relationship between two entities of the digital resources in territorial spatial planning can be transformed into a directed relationship between two nodes in the graph database. The data attributes of entities are generally saved in the Neo4j database as attributes of nodes. Additionally, the relationship between two nodes can also establish attributes, which are derived from the relationship between two entities. For example, the attribute value of the "Touches" relationship between Plot a and Plot b is "100", indicating that the adjacent common edge length between Plot a and Plot b is 100 m. The Cypher language and the py2neo tool library are used to store nodes, node attributes and node relationships in Neo4j according to the mapping rules [36]. The entities include "Plot" and "Plot Type", with relationships such as "Belong" and "Belong kind" linking them. Additionally, various relationships exist between plots, including "Touches", "Within", "Contains", "Degree", "Euclidean distance" and "Short distance". Node attributes follow Table 3.

3. Results

By reference to the various entities and entity relationships obtained in Section 2, a knowledge graph of the Guangyang Island territorial spatial planning considering topological relationships, directional relationships and distance relationships was constructed. An interactive visualization client for territorial spatial planning based on the knowledge graph of the study area was developed. The client can quickly query different types of plots in the area, and various centralities, such as the harmonic centrality and closeness centrality of each plot, can be calculated to measure the impact of the plot in the area. Based on the school evaluation model of Equation (1), the locations of schools in the area were evaluated, and then it was determined whether the planning location of a school was reasonable, which helps residents choose primary and secondary schools and assists in the spatial layout construction planning of primary and secondary schools.

This visualization client can interactively query the serviceable and unserviceable residential areas of a school regarding the specific needs in the study area, such as walking or driving for n minutes. Figure 8 shows that all the schools in the area are located within 15 min of walking, which is both the serviceable range and unserviceable range. The area of serviceable residential land is about 6.9 km^2.

Equation (1) was used to evaluate the locations of schools in the study area and to determine whether the planned location of a school is superior and suitable. Table 4 shows the top three measurement scores in the evaluation model under the conditions of the region being affected by the winter north wind and within 15 min of walking. The primary and secondary school plots with the last three scores of 56, 51 and 70 under the same conditions are shown in Table 5. The locations are shown in Figure 9.

Table 4. Top three measurement scores.

OBJTECTID	Score
59	0.861
74	0.806
58	0.765

Table 5. Last three measurement scores.

OBJTECTID	Score
56	0.211
51	0.254
70	0.307

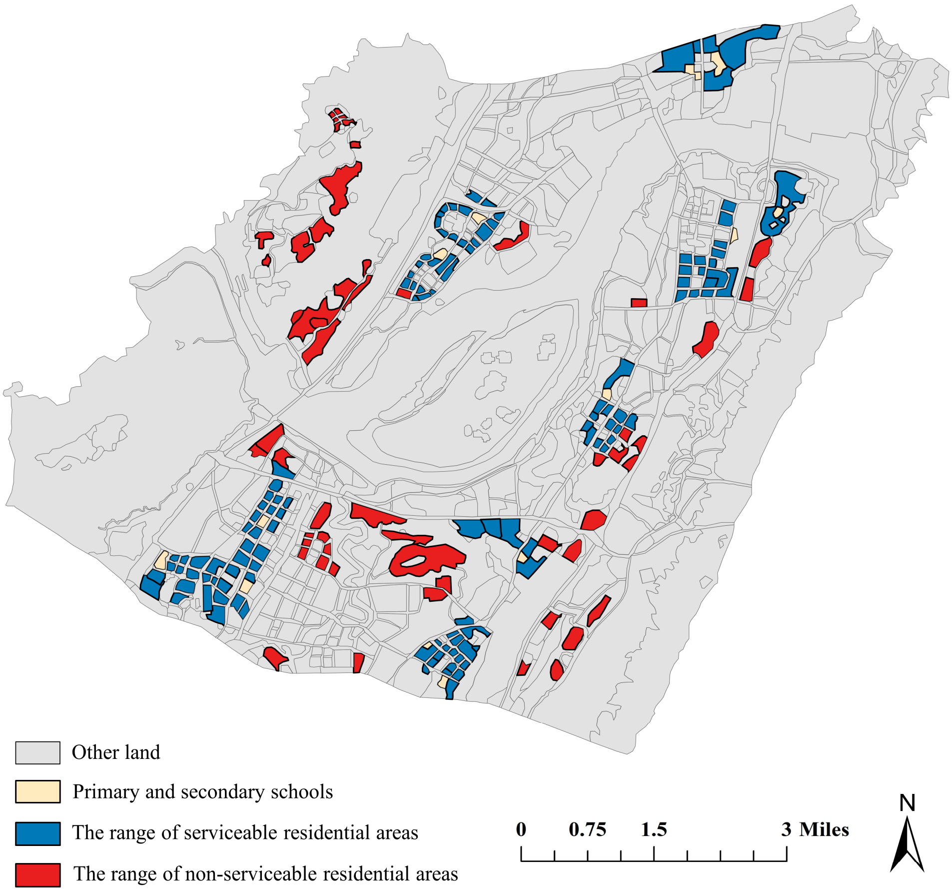

Figure 8. All schools in the area are serviceable or non-serviceable within a residential area with a walking commute time of 15 min.

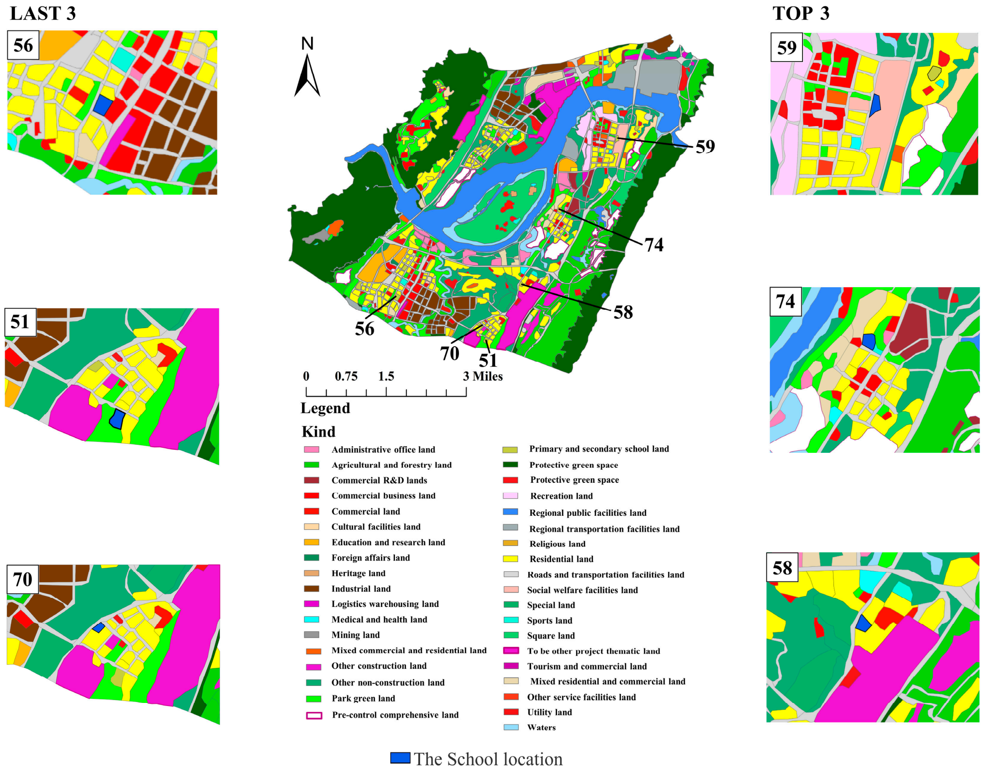

Figure 9. The locations of the top three and bottom three ranked schools.

The school location of ID 59 can serve about 0.054 km^2 of residential land under the condition of a model 15 min walk. It is in a central position in the topological knowledge map and can be easily reached for most public service facilities. The terrain is flat with a mean slope of 7.4° and the distance from all factories in the area and the K-order neighbourhood value are large, which means that the factory has the least impact on the school. The school location of ID 56 is the worst-performing location in this model. Although it can serve about 0.042 km^2 of residential land under the condition of a model 15 min walk, this location has a mean slope of 12.7° and a five-order neighbourhood with the nearest factory, is near a factory in a straight line, and there are three same-type schools nearby. The location should be moved to a flatter plot that is farther from the factory in a straight line, and the surrounding schools are not dense. Through the above analysis, the location of the highest and lowest schools in the study area is in line with the actual situation. Therefore, the model is reliable in this study area.

4. Discussion

Presently, there are many models of school location, such as location models based on AHP hierarchy analysis, location models based on linear programming, location models based on ArcGIS modelling analysis and location models based on multifactor grid overlay analysis. Although there are many school location models, there are few models for evaluating the locations of existing schools in territorial spatial planning. In this article, a school location evaluation model based on a knowledge graph, which considers the terrain of the school location, the serviceable population, the connection with existing schools and factories, and the proximity and degree of influence of school plots in the region, is constructed. The quantitative evaluation method is employed to measure the advantages and disadvantages of a school planning location in the study area.

In the same research area, we utilized the same five major influencing factors and conducted studies using both AHP hierarchical analysis and ArcGIS modelling for school site selection, as shown in Figures 10 and 11. Tables 6 and 7 present a comparison of the specific normalized scores obtained from the three different models. According to the result graphs, the outcomes of AHP analysis and ArcGIS modelling highly coincide with the findings of our study. For instance, the top three positions identified by AHP are ID 59, 74 and 71, while those from ArcGIS modelling are ID 71, 73 and 59. However, the results obtained from our proposed school quantitative site selection model based on knowledge graphs are ID 59, 74 and 48. It is evident that our results align to some extent with other methods, but there are also discrepancies. For example, our optimal result is ID 59, which matches the AHP analysis, but differs from the ArcGIS modelling result, which is ID 71. The proximity of School ID 71 to factories, as depicted in Figure 10, raises concerns. Additionally, the second position according to AHP is also ID 71. Conversely, the locations of IDs 59, 74 and 48 from our model are advantageous, serving a larger population and having complete supporting facilities. The population distribution and government planning for the functional areas are also considered to be the two most important factors in other school site selection studies [7,11]. The accessibility of the school also has a similar impact on the final choice of site [37,38]. The population distribution, functional area planning around schools and school accessibility all have an important place in the model proposed in this paper. Topographical factors are often considered in school siting issues, but have different importance in different studies, which is closely related to the differences in the overall topographical characteristics of the study area [38,39]. The topography of the study area can provide a great deal of flat space for school siting, so we considered topography as an important factor in this study. The spatial location of the final site selection result has great similarity to the spatial location of the sites selected in other studies, thus further demonstrating the validity and reasonableness of our model.

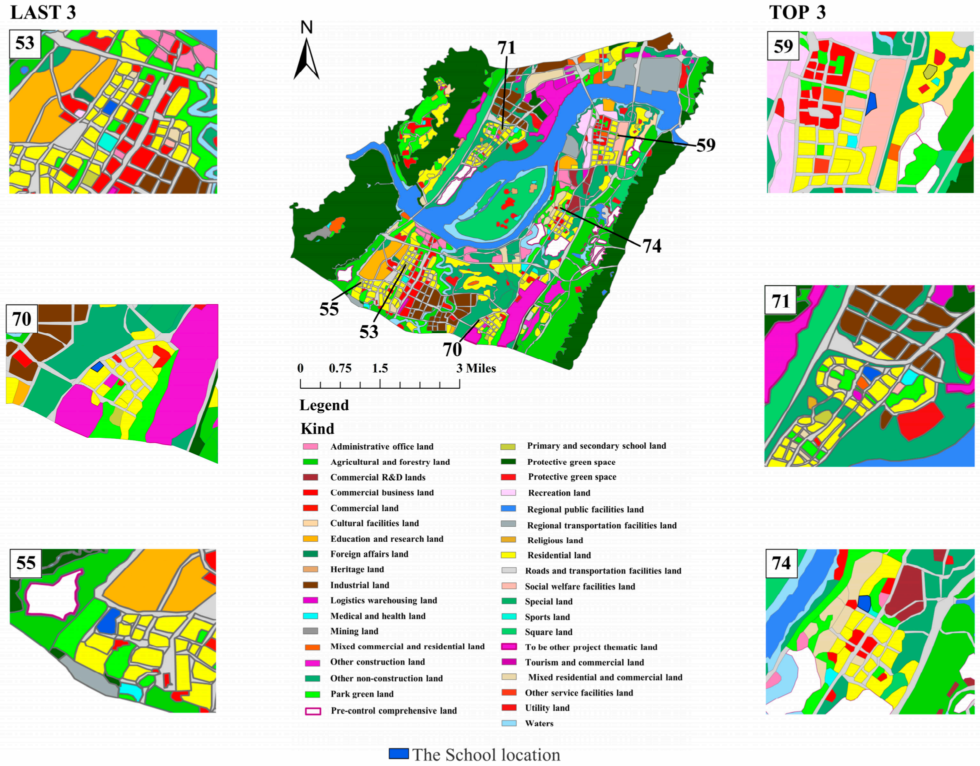

Figure 10. The locations of the top three and bottom three ranked schools in AHP.

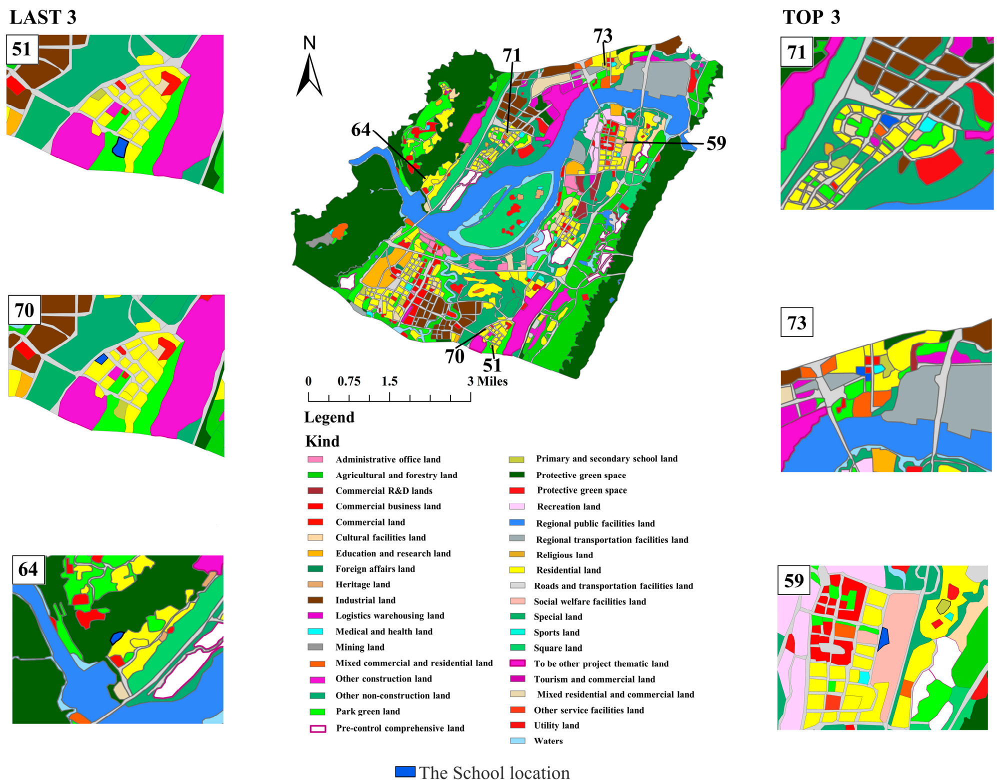

Figure 11. The locations of the top three and bottom three ranked schools in ArcGIS.

Table 6. Top three measurement scores with the three models.

Origin		AHP		ArcGIS	
ID	**Score**	**ID**	**Score**	**ID**	**Score**
59	0.861	59	0.832	71	0.841
74	0.806	71	0.788	73	0.817
58	0.765	74	0.761	59	0.789

Table 7. Last three measurement scores with the three models.

Origin		AHP		ArcGIS	
ID	**Score**	**ID**	**Score**	**ID**	**Score**
56	0.211	53	0.165	51	0.205
51	0.254	70	0.237	70	0.248
70	0.307	55	0.292	64	0.336

The school evaluation model proposed in this article refers to the influencing factors and some analysis methods of the school location model, but it is different from the analysis method of the location model. After considering many influencing factors, the model combines the adjacency and centrality of topological subgraphs to quickly, effectively and quantitatively evaluate a school's location. The model considers the existing factory's assessment of a school's location and combines the wind direction to measure the impact of polluting gases. The shortest path network distance is applied instead of the distance buffer [8] to analyze the coverage when calculating the serviceable population, but population density, population change and other factors are not considered [11]. The quantitative integration of various influencing factors replaces the use of overlay analysis of multiple grids [8] to obtain the final result. The shortest network distance is used to consider the commuting convenience of a school, but the road obstacles and traffic conditions are not considered [12]. Parameters such as K-order neighbours and centrality, which cannot be obtained by conducting spatial analysis, are innovatively added to measure the proximity and influence of plots instead of grid superposition and spatial statistical methods.

By and large, based on the characteristics that a knowledge graph is convenient to express and store relationships and can calculate their connections very efficiently, this article proposes a new multi-factor evaluation model for primary and secondary school location selection, quantitatively evaluates the advantages and disadvantages of school location selection in a certain area, and innovatively integrates the topology, centrality, K-order neighbours and other factors of the spatial inland block. This method can be extended from the research area of this paper to other areas in need, and can provide suggestions for the location of primary and secondary schools in the region.

5. Conclusions

Reasonable land space planning has an important role in the development of cities, and school location planning is a major problem. It is particularly important to evaluate the planned location of a school. In this article, the construction of a territorial spatial planning knowledge graph considering three major relationships in spatial relationships—spatial topological relationships, direction relationships and metric relationships—is investigated. A school location evaluation model based on territorial spatial planning knowledge graphs is proposed, and the results of using the model are discussed. The definitions of the territorial spatial planning knowledge graph schema and the acquisition and storage of knowledge about the topological relations, distance relations and direction relations were explored to construct a knowledge graph. Based on the knowledge graph of land and space planning, a school location evaluation model that combines terrain, spatial relationship, centrality and adjacency is proposed. The knowledge was acquired from the digital resources territorial spatial planning map of the Guangyang Island Yangtze River

Ecological Civilization Innovation Experimental Zone. The knowledge of the plot and plot type was used to build the entity. Touches, containing nine topological relations, the angle relation in the direction relation, the Euclidean distance and the network distance in the matric relation were utilized to construct the relationships between two entities. The Neo4j graph database was selected for knowledge graph storage and visualization. The model was applied to the study area for the experiment, and it is concluded that the study area evaluation results scored the best and worst schools.

This study focuses on the territorial spatial planning of Guangyang Island in Chongqing, China, employing a quantitative site selection model based on knowledge graphs. It assumes the monsoon direction to be northward and travel by foot, with the service population set within a 15 min walking distance. The results reveal the top three and bottom three ranked plot IDs in the planned spatial area to be 59, 74, 58 and 56, 51, 70, respectively. Comparing these findings with AHP hierarchical analysis and ArcGIS modelling, similarities and differences are noted. Through comparative analysis, it is concluded that the school location evaluation model proposed in this article is feasible, realistic and comprehensive, integrating multiple factors. The results also show that the constructed spatial planning knowledge graph is correct and meaningful and that the knowledge graph is an intuitive and efficient method to represent the relationship between two spatial entities or concepts.

This article has certain drawbacks. In the acquisition of direction relationship knowledge, the centre point of the plot is used to replace the plot. However, due to the irregularity of the plot, the use of the centre point of the plot instead does not always fully represent the characteristics and location of the plot, and better models of computing direction and metric relationships will be employed in the future. Although knowledge graphs have the advantages of the rapid integration and rapid analysis of knowledge, the knowledge of nodes and relationships in knowledge graphs is acquired in advance, and spatial analysis cannot be dynamically utilized to acquire knowledge in spatial scenes. Therefore, if the spatial planning pattern of the region changes, it is necessary to update the knowledge graph to use the model. The model also has some shortcomings. When calculating the service population, the model does not consider the housing volume rate and population density. Whether the model can effectively measure the service population needs to be addressed. This model only considers five factors: terrain, population, other schools, factories, land importance and proximity. Additional factors, such as economic factors, entertainment facilities and public facilities location will be considered in future analyses.

In this study, the model is only applicable to a small area in China, and the effectiveness of evaluating schools in other countries or regions remains unclear. Additionally, due to data limitations, we did not differentiate between primary and secondary schools but grouped them together for analysis. Moreover, only public schools were considered in territorial spatial planning, and private schools were not taken into account. In future research, if the data permit, we should focus more on precise studies, such as exclusively examining secondary schools or incorporating private primary and secondary schools into the analysis.

Furthermore, this study only presents a school evaluation model based on a knowledge graph, aiming to select better locations. However, proposing a comprehensive school location model based on a knowledge graph will be crucial in future research. Overall, future studies should prioritize enhancing accuracy, expanding the scope of analysis and optimizing the methodologies to provide more robust insights into school location planning.

Author Contributions: Xiankang Xu implemented the experiments and drafted the manuscript. Jian Hao processed the data. Jingwei Shen provided overall guidance to the project and obtained funding to support this research. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Natural Science Foundation of Chongqing (No. CSTB2022NSCQ-MSX0336), and Chongqing Postgraduate Education and Teaching Reform Research Project (No. yjg213035).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Hersperger, A.M.; Oliveira, E.; Pagliarin, S.; Palka, G.; Verburg, P.; Bolliger, J.; Gradinaru, S. Urban land-use change: The role of strategic spatial planning. *Glob. Environ. Change-Hum. Policy Dimens.* **2018**, *51*, 32–42. [CrossRef]
2. Liu, Y.; Zhou, Y. Territory spatial planning and national governance system in China. *Land Use Policy* **2021**, *102*, 105288. [CrossRef]
3. Witte, P.; Punt, E.; Geertman, S. Smart governance in the making: Integrating 'smart' in local spatial planning. In *Handbook of Planning Support Science*; Edward Elgar Publishing: Cheltenham, UK, 2020; pp. 226–237.
4. Högström, J.; Brokking, P.; Balfors, B.; Hammer, M. Approaching sustainability in local spatial planning processes: A case study in the Stockholm region, Sweden. *Sustainability* **2021**, *13*, 2601. [CrossRef]
5. Sakti, A.D.; Rahadianto, M.A.E.; Pradhan, B.; Muhammad, H.N.; Andani, I.G.A.; Sarli, P.W.; Abdillah, M.R.; Anggraini, T.S.; Purnomo, A.D.; Ridwana, R.; et al. School Location Analysis by Integrating the Accessibility, Natural and Biological Hazards to Support Equal Access to Education. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 12. [CrossRef]
6. Barbara, M.; Rey, D.; Akbarnezhad, A. Optimizing Location of New Public Schools in Town Planning Considering Supply and Demand. *J. Urban Plan. Dev.* **2021**, *147*, 04021057. [CrossRef]
7. Prasetyo, D.H.; Mohamad, J.; Fauzi, R. A GIS-based multi-criteria decision analysis approach for public school site selection in Surabaya, Indonesia. *Geomatica* **2018**, *72*, 69–84. [CrossRef]
8. Ghazal, N.; Jaleel, Q.A. Using GIS and Python Language to Choose the Best School Location (Study on the Urban Areas of the Husseiniya District in Karbala Province). *Int. Rev. Phys. IREPHY* **2017**, *11*, 112–119.
9. Al-Sabbagh, T.A. GIS location-allocation models in improving accessibility to primary schools in Mansura city-Egypt. *Geojournal* **2022**, *87*, 1009–1026. [CrossRef]
10. Al Quhtani, S. Spatial distribution of public elementary schools: A case study of Najran, Saudi Arabia. *J. Asian Archit. Build. Eng.* **2022**, *22*, 705–725. [CrossRef]
11. Lotfi, R.; Pilehforooshha, P.; Karimi, M. A multi-objective optimization model for school location-allocation coupling demographic changes. *J. Spat. Sci.* **2023**, *68*, 225–244. [CrossRef]
12. Meena, D.K.; Tripathi, R.; Agrawal, S. An evaluation of primary schools and its accessibility using GIS techniques: A case study of Prayagraj district, India. *Geojournal* **2023**, *88*, 1921–1951. [CrossRef] [PubMed]
13. Rikalovic, A.; Cosic, I.; Lazarevic, D. GIS based multi-criteria analysis for industrial site selection. *Procedia Eng.* **2014**, *69*, 1054–1063. [CrossRef]
14. Shorabeh, S.N.; Firozjaei, H.K.; Firozjaei, M.K.; Jelokhani-Niaraki, M.; Homaee, M.; Nematollahi, O. The site selection of wind energy power plant using GIS-multi-criteria evaluation from economic perspectives. *Renew. Sustain. Energy Rev.* **2022**, *168*, 112778. [CrossRef]
15. Ahmed Ali, K. Multi-criteria decision analysis for primary school site selection in Al-Mahaweel district using GIS technique. *J. Kerbala Univ.* **2018**, *14*, 342–350.
16. Shen, T.; Zhang, F.; Cheng, J. A comprehensive overview of knowledge graph completion. *Knowl. Based Syst.* **2022**, *255*, 109597. [CrossRef]
17. Singhal, A. Introducing the Knowledge Graph: Things, Not Strings. Google Blog 2012. Available online: https://www.blog.google/products/search/introducing-knowledge-graph-things-not/ (accessed on 11 February 2024).
18. Fensel, D.; Şimşek, U.; Angele, K.; Huaman, E.; Kärle, E.; Panasiuk, O.; Toma, I.; Umbrich, J.; Wahler, A. Introduction: What is a knowledge graph? In *Knowledge Graphs: Methodology, Tools and Selected Use Cases*; Springer: Cham, Switzerland, 2020; pp. 1–10.
19. Richens, R.H. Preprogramming for mechanical translation. *Mech. Transl. Comput. Linguist.* **1956**, *3*, 20–25.
20. Hogan, A.; Blomqvist, E.; Cochez, M.; d'Amato, C.; Melo, G.D.; Gutierrez, C.; Kirrane, S.; Gayo, J.E.L.; Navigli, R.; Neumaier, S. Knowledge Graphs. *Acm Comput. Surv.* **2021**, *54*, 1–37. [CrossRef]
21. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 494–514. [CrossRef] [PubMed]
22. Cheng, B.J.; Zhang, J.; Liu, H.; Cai, M.L.; Wang, Y. Research on Medical Knowledge Graph for Stroke. *J. Healthc. Eng.* **2021**, *10*, 5531327. [CrossRef]
23. Li, L.; Wang, P.; Yan, J.; Wang, Y.; Li, S.; Jiang, J.; Sun, Z.; Tang, B.; Chang, T.-H.; Wang, S. Real-world data medical knowledge graph: Construction and applications. *Artif. Intell. Med.* **2020**, *103*, 101817. [CrossRef]
24. Shen, J.; Shi, K.; Ma, M. Exploring the construction and application of spatial scene knowledge graphs considering topological relations. *Trans. GIS* **2022**, *26*, 1531–1547. [CrossRef]
25. Sun, K.; Hu, Y.; Song, J.; Zhu, Y. Aligning geographic entities from historical maps for building knowledge graphs. *Int. J. Geogr. Inf. Sci.* **2020**, *35*, 2078–2107. [CrossRef]
26. Clementini, E.; Di Felice, P. Approximate topological relations. *Int. J. Approx. Reason.* **1997**, *16*, 173–204. [CrossRef]
27. Zhang, Y.; Zhang, A.; Gao, M.; Liang, Y. A Spatial Relation Model of Three-Dimensional Electronic Navigation Charts Based on Point-Set Topology Theory. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 259. [CrossRef]

28. Strobl, C. Dimensionally Extended Nine-Intersection Model (DE-9IM). In *Encyclopedia of GIS*; Shekhar, S., Xiong, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 240–245.
29. Izmirlioglu, Y.; Erdem, E. Qualitative reasoning about 2D cardinal directions using answer set programming. *J. Artif. Intell. Res.* **2023**, *77*, 1371–1453. [CrossRef]
30. Shekhar, S.; Liu, X.; Chawla, S. An object model of direction and its implications. *GeoInformatica* **1999**, *3*, 357–379. [CrossRef]
31. Buczkowska, S.; Coulombel, N.; de Lapparent, M. A comparison of euclidean distance, travel times, and network distances in location choice mixture models. *Netw. Spat. Econ.* **2019**, *19*, 1215–1248. [CrossRef]
32. Miller, H.J. Tobler's First Law and spatial analysis. *Ann. Assoc. Am. Geogr.* **2004**, *94*, 284–289. [CrossRef]
33. Freeman, L.C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1978**, *1*, 215–239. [CrossRef]
34. Marchiori, M.; Latora, V. Harmony in the small-world. *Phys. A Stat. Mech. Its Appl.* **2000**, *285*, 539–546. [CrossRef]
35. Vagner, A. Route planning on GTFS using Neo4j. *Ann. Math. Informaticae* **2021**, *54*, 163–179. [CrossRef]
36. Chen, Q.; Xia, J.; Feng, J.; Tong, M. Research on Knowledge Graph Construction for Python Programming Language. In *Resilience and Future of Smart Learning*; Yang, J., Liu, D., Kinshuk, A., Tlili, M., Chang, E., Popescu, D., Burgos, Z.A., Eds.; ICSLE; Lecture Notes in Educational Technology; Springer: Singapore, 2022.
37. Talen, E. School, community, and spatial equity: An empirical investigation of access to elementary schools in West Virginia. *Ann. Assoc. Am. Geogr.* **2001**, *91*, 465–486. [CrossRef]
38. Baser, V. Effectiveness of school site decisions on land use policy in the planning process. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 662. [CrossRef]
39. Ahmad, M.B.A.; Bahiah, B.; Rasib, A.W.; Saifullizan, M.B. GIS and Multi-criteria Analysis for School Site Selection (Study Case: Malacca Historical City). *Int. J. Integr. Eng.* **2021**, *13*, 234–241.