International Journal of
Geo-Information

MDPI

Article

Identifying the Production–Living–Ecological Functional Structure of Haikou City by Integrating Empirical Knowledge with Multi-Source Data

Bingbing Zhao [1,2], Xiaoyong Tan [1], Liang Luo [3], Min Deng [1] and Xuexi Yang [1,*]

1 Department of Geo-Informatics, Central South University, Changsha 410083, China; zbingbing@csu.edu.cn (B.Z.); bariomax@csu.edu.cn (X.T.); dengmin@csu.edu.cn (M.D.)
2 Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong 999077, China
3 Guangdong Polytechnic of Industry and Commerce, Guangzhou 510510, China; luoliang_csu@163.com
* Correspondence: yangxuexi@csu.edu.cn

Abstract: The inefficient use of urban resources and the imbalance of spatial structures make optimizing land use management a top priority in urban environmental management. Traditional land use classification systems that prioritize only natural features while disregarding human activity can result in redundancy and conflicts in urban planning. The Production–Living–Ecological Space (PLES) approach was developed as an integrated method for territorial spatial classification. However, most existing studies on PLES are conducted at provincial scales, largely overlooking fine-scale usage within cities. In addition, the existing concept of PLES has been vaguely defined, resulting in linear and simple identification methods that are not applicable to complex urban environments. To address these issues, this study proposes a method to identify urban PLES based on supervised classification using random forest models, which integrate empirical knowledge and multi-source heterogeneous information. The experiments conducted in Haikou reveal the regional aggregation of living and production spaces and the scarcity of ecological space in the city. Our study proposes a concrete concept of PLES and a method for identifying PLES that can be applied to multiple regions, providing an effective tool for the coordinated management of urban production, living, and ecological environments.

Keywords: Production–Living–Ecological Space; random forest; spatial structure; multi-source data

Citation: Zhao, B.; Tan, X.; Luo, L.; Deng, M.; Yang, X. Identifying the Production–Living–Ecological Functional Structure of Haikou City by Integrating Empirical Knowledge with Multi-Source Data. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 276. https://doi.org/10.3390/ijgi12070276

Academic Editors: Wolfgang Kainz and Godwin Yeboah

Received: 12 May 2023
Revised: 9 July 2023
Accepted: 11 July 2023
Published: 12 July 2023

1. Introduction

The frequency of socio-environmental issues, such as urban land use conflicts, ecological pressure, water pollution, and declining food production, is increasing due to rapid urbanization. These issues pose a significant threat to the sustainable development of our society [1,2]. A series of studies related to sustainable development have been conducted to address these challenges, most of which have focused on the analysis of land use structure [3–5]. Land use classification relies on the division of physical space and refers specifically to objects on land, such as grasslands, cultivated areas, and built-up areas. However, this hard division fails to capture the relationship between human activity and nature, as it attempts to portray human activity in terms of a single object. For instance, greenery, residential areas, and commercial premises often coexist within the same area of the city, and describing this area as a single object is insufficient. To address this issue, the concept of "Production–Living–Ecological Space" (PLES) was developed at the 18th National Congress of the Communist Party of China [6], which provides a new perspective for understanding and governing urban space. The PLES concept divides urban space into living, production, and ecological spaces from the perspective of human activities. Production space typically refers to land used for agricultural, industrial, and commercial activities aimed at obtaining or generating products, while living space encompasses land for human habitation, consumption, and entertainment. Conversely, ecological space refers to land used for regulating, maintaining, and securing ecological functions [4]. The PLES concept is an integrated functional zoning approach that can be oriented towards multiple spatial scales and is compatible with diverse functional needs. It facilitates the formulation of regional development strategies in accordance with local conditions and has been applied in various fields such as environmental protection [7], regional spatial planning integrating multiple regulations [8], ecological services management [9], and land use structure analysis [10].

The prospect of a wide range of applications has inspired research on PLES, which includes concepts of identification and optimization [11,12] and evolutionary analysis [13], among which the identification of PLES contributes to the scientific knowledge of spatial distribution patterns and forms the cornerstone of subsequent simulation and optimization [8,14,15]. Initially, owing to the limited availability of data, scholars focused on macro-scale studies with cities, counties, and towns as the basic research units. For example, Liao et al. [16] analyzed the spatial distribution trends and structural differentiation of PLES in the hilly areas of Sichuan Province at the county scale, while Duan et al. [11] analyzed the spatial distribution characteristics and quantitative transformation process of PLES at the village scale. Bian et al. [17] analyzed the functional characteristics and equilibrium of the study area, with streets and townships as research units, and proposed corresponding optimization suggestions for the coordinated and sustainable development of spatial functions. The subsequent rise of multi-source spatio-temporal big data (e.g., remote sensing imagery data, social media data, and travel data) has provided scholars with the opportunity to identify functional patterns at the microscopic scale. Fu et al. [18] used point of interest (POI) data and hierarchical analysis to identify the patterns of urban PLES at a 300 m grid scale. Li et al. [10] constructed a functional classification system for urban PLES by coupling biophysical process measurement and value conversion methods. The emergence of these studies has enhanced the implementation of the PLES approach at the human scale, which is useful for the fine-grained planning and management of cities. However, these studies mostly described the functions of urban space in terms of physical space, neglecting the significance of human activity space in representing the functions of land parcels. Urban space provides a place for human activities and constrains them in turn. Therefore, identifying urban functions while considering human activity characteristics is vital for bridging the gap between physical space and real functional space.

Existing spatial identification methods for PLES can be categorized into the merge, value measurement, and indicator synthesis methods. The merge method refers to the grouping of categories into living, ecological, and production functions based on an existing classification system [19], which is limited by the classification accuracy of the existing classification systems and ignores the intensity of the functions. The value measurement method refers to the representation of land functions through the biophysical process and value conversion calculations [10,20], which are more refined but complex to calculate and difficult to obtain. The indicator synthesis method indicates the construction of a functional evaluation indicator system based on the interpretation of PLES and then quantifies the functional intensity by applying the entropy weight method, expert scoring method, and hierarchical analysis to synthesize the indicators [8,17,21]. Indicator synthesis methods are often preferred due to the computational simplicity and flexibility of built-on-demand indicator systems. Methods such as entropy weighting and hierarchical analysis provide ways to establish weights among indicators, while such linear calculation rules (applying a set of weighting coefficients to all indicators) are too simple to model complex rules for the interaction of multiple variables in geographic space. It is crucial to solve this problem by mining non-linear rules based on human knowledge. The rise of supervised learning methods has provided a method to achieve this from the limited empirical knowledge of humans, allowing training to find rules between features and labels based on existing data [22,23].

In summary, to identify PLES more precisely, two problems need to be solved: (I) how to understand the functional structure of urban space by combining human activity characteristics with physical environment characteristics, and (II) how to mine the rules of non-linear spatial cognition from empirical knowledge data. To address the above problems, a description system of PLES from two perspectives of human activities and physical environment was constructed by integrating the data of online taxi trips, POI data, and AOI (area of interest) data. The annotation set was then constructed using empirical perception to annotate living, production, and ecological spaces that can be easily distinguished. Finally, a random forest model was used to mine non-linear rules between the indicator set, serving to describe the urban function and the annotated set, and the rules were applied to other unlabeled regions to obtain the full spatial structure of the urban PLES. An experimental study and analysis were conducted in the main urban area of Haikou City to provide a theoretical basis for the optimization of spatial patterns in the city.

This paper consists of five sections: Section 1 comprises the study background and introduction. Section 2 describes the study area and the data selected. The research method is described in detail in Section 3. In Section 4, the results of the identification of the PLES in the study area are presented. The discussions and conclusions are presented in Section 5.

2. Study Area and Data

2.1. Study Area

Haikou, the capital of Hainan Province, is located in the northeastern part of the province and is the first pilot area for provincial spatial planning reform in China. The identification of the urban functional structure is a prerequisite for the implementation of its spatial planning strategy. In response to the need for urban spatial planning, the main urban area of Haikou, which is densely populated, economically prosperous, and has a variety of land uses, was selected as the study area to provide a proxy for the identification of PLES in other areas. As the road network density in Haikou is roughly 5.41 km/km^2, this corresponds to an average neighborhood scale of about 300 m [24], totaling 2794 regular grids. The study area is shown in Figure 1.

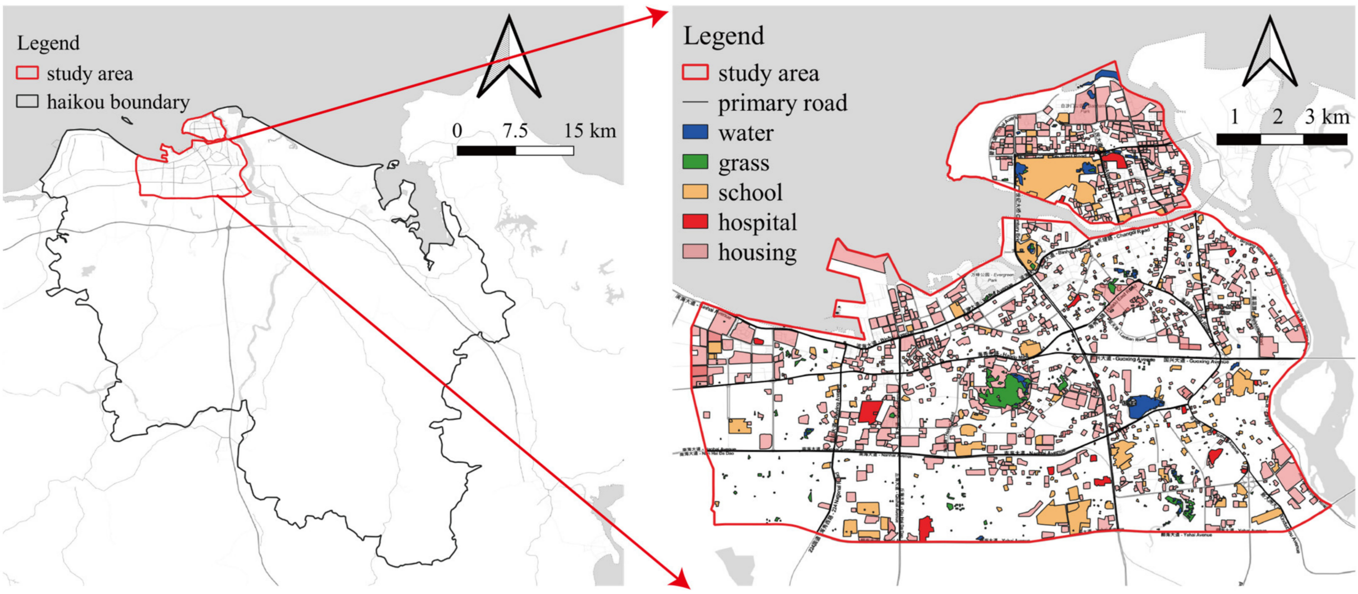


Figure 1. Map of the study area.

2.2. Data and Pre-Processing

POIs refer to geographical entities that are closely related to people's lives, such as restaurants, banks, and restaurants. We obtained 129,608 POIs in December 2017 from Gaode Maps (https://lbs.amap.com/ (accessed on 29 December 2017)), a leading Chinese navigation and location service solution provider, comprising 10 categories such as restaurants, living services, and companies. Subsequently, 67,339 POIs were maintained after removing any data sourced from beyond the study area. The POIs were aggregated into three general categories to eliminate coupling between the original categories. The aggregation rules are listed in Table 1.

Table 1. The aggregation rules of POIs.

General Category	Detailed Category	Description with Examples
Entertainment	shopping	shopping centers, stores, etc.
	restaurants	snack bars, dessert stores, etc.
	leisure places	cinemas, theaters, cabarets, etc.
	sports	stadiums, fitness centers, etc.
	hotels	star-rated hotels, budget hotels, B&Bs, etc.
Industry	companies	-
	factories	-
Service	living services	post offices, communication offices, laundromats, photo studios, etc.
	financial services	banks, credit unions, pawnshops, etc.
	government agencies	Administrative units, public prosecutors and law enforcement agencies, welfare agencies, etc.

AOIs refer to area-like geographical entities on a map that do not comprise a single point but rather a polygon area, such as schools. A total of 1481 AOIs were obtained from Baidu Maps (https://lbsyun.baidu.com/ (accessed on 28 December 2017)) through web crawling techniques and consisted of three categories: education, medical, and residential areas. The boundary of each AOI was stored as a polygon so that its area could be calculated.

The land cover data for Haikou City in 2017 were downloaded from the 30 m annual land cover dataset [25], including a total of nine land use categories: cropland, forest, grassland, shrub, wetland, water, impervious, barren, and snow/ice. Forests, shrubs, and grasslands were combined into the greenery category, given their relatively similar functions in ecological regulation.

The road network data of Haikou City were obtained from OpenStreetMap (https://www.openstreetmap.org (accessed on 12 February 2020)), including the road number, road type, road name, and other attributes. Road types were categorized as primary, secondary, and tertiary roads. Data outside the study area were removed, leaving 509 rows of data for use in analyses.

The travel data from vehicles with online tracking were obtained from the publicly available Gaia dataset of DiDi ChuXing (https://gaia.didichuxing.com (accessed on 22 March 2021)), which is a diversified travel platform in China that provides users with travel-booking services. Based on the passenger transport data released by the Ministry of Transport of the People's Republic of China [26], taxis accounted for around 20% of the total passenger transport in Haikou City in January 2020, serving as a significant addition to public transportation. Additionally, the Didi company's user base mainly consists of individuals aged between 20 and 40 in the urban area, which represents the primary demographic engaged in productive activities within the city. Understanding their activity patterns provides valuable insights for comprehending the functional structure of the city. Data were collected on weekdays between 4 September and 29 September 2017, with 1,338,144 travel records. Weekday data were chosen because the difference between people's living and production behaviors is more pronounced on weekdays. Each travel record includes five attribute fields: order ID, departure time, pick-up position, arrival time, and drop-off location. Because the coordinates and time recorded by the vehicle GPS positioning sensor may contain some errors due to meteorological and environmental influences, travel records with excessive travel time (>2 h) or a travel time record of 0 were deleted. Duplicate travel records and travel records with origins or destinations outside the study area were also removed. In total, 1,326,473 travel records were included.

3. Methodology

To address the biases caused by neglecting human activity and the limitations caused by the simple linear combination of indicators in the existing research on urban PLES, we first constructed a set of indicators for the description of PLES by enriching the existing indicators with human activity and AOI data. Then, the living, production, and ecological spaces were identified based on artificial empirical experience to obtain the annotation set of PLES. Finally, the random forest algorithm was applied to identify the PLES, and a comprehensive analysis of the PLES was performed based on the identification results. The workflow of the research process is shown in Figure 2.

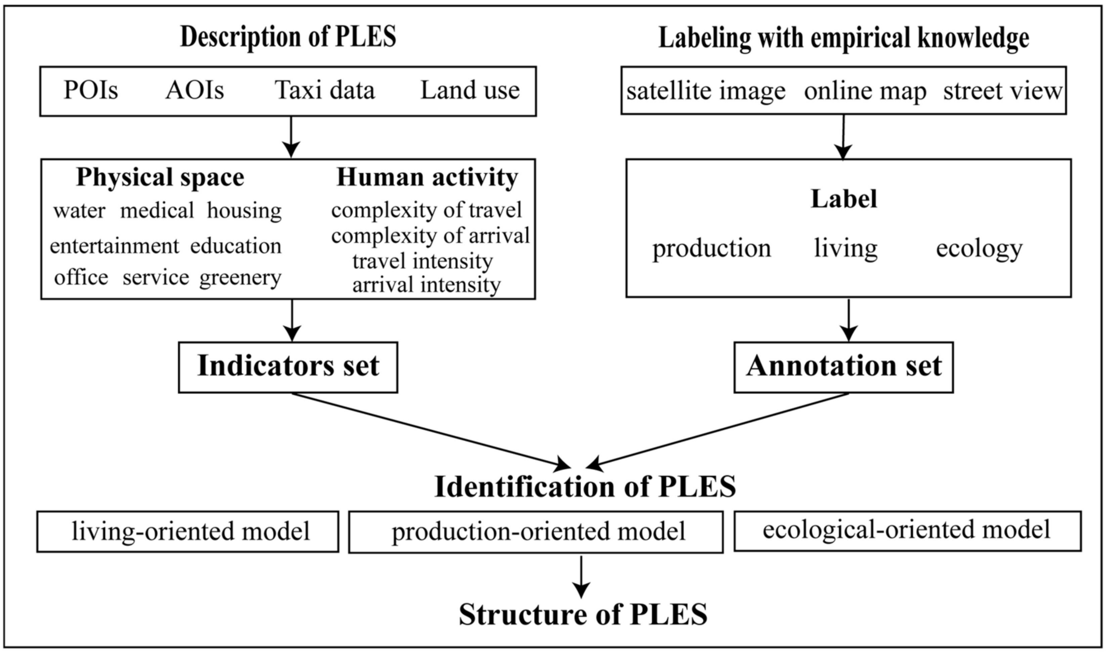


Figure 2. The workflow of the research process.

3.1. Description of PLES: Construction of a Set of Indicators

Among the existing indicators describing urban PLES, living space mainly focuses on residential functions, ecological space concerns green areas, and production space concerns functions such as transport and industry [27]. These studies demonstrate valid experimental results by concretizing human knowledge in the understanding of concepts and consequently measuring living, production, and ecological functions separately. However, this concretization process is prone to disagreement because people are usually sensitive to the results and are not sensitive to the judgment process used to obtain them. For instance, while it is straightforward to categorize a hospital as a living area, it becomes challenging to quantify the extent to which we disregard the dispersed office spaces and ecological functions surrounding the hospital, opting to define it solely as having a living function. Instead of specifying definitions of living space, production space, and ecological space, these multiple sources of data are combined to provide as rich a description of the function of this region as possible.

AOI and land cover data were used to characterize the functional intensity of educational resources, medical resources, and residential and office buildings, considering that the influence of these entities cannot be expressed by a single coordinate point. The functional intensity of some entities that can be abstracted as point entities were characterized by their density. To date, eight functional descriptions of physical space have been established, including housing, healthcare, education, industry, green space, and other aspects of daily life.

The ability to attract human activity varies significantly across urban functional areas, influenced by the rhythm of human life. Areas providing commercial services typically maintain high levels of population dynamics throughout the day, while residential and industrial production areas show differences in population dynamics at night and during the day, influenced by the work rhythms of residents [28]. The important role of travel activity differences in the identification of the functional structure of a city has also been demonstrated in many applications [29,30]. Four indicators were constructed using online taxi data based on the pick-up and drop-off behaviors of users to enrich the functional description of the area. Detailed indicator descriptions and calculations are presented in Table 2.

3.2. Labeling PLES in Combination with Empirical Knowledge

A region where a landmark is located was selected as the priority label region because we can determine the functional properties of such regions. Satellite imagery and street maps were then integrated to label the functions of the study units, thus establishing a set of labels for the identification of PLES. Figure 3 shows the labeled map, and in Table 3, we have listed some typical examples: schools, hospitals, and housing estates were labeled as living spaces because they mainly serve human activities, such as housing, medical care, and education. Woods and lakes were labeled as ecological spaces, and production workshops and commercial offices as the main workplaces were labeled as production spaces. A total of 524 grid cells were obtained for the annotation set, of which 236, 102, and 186 corresponded to living, production, and ecological spaces, respectively.

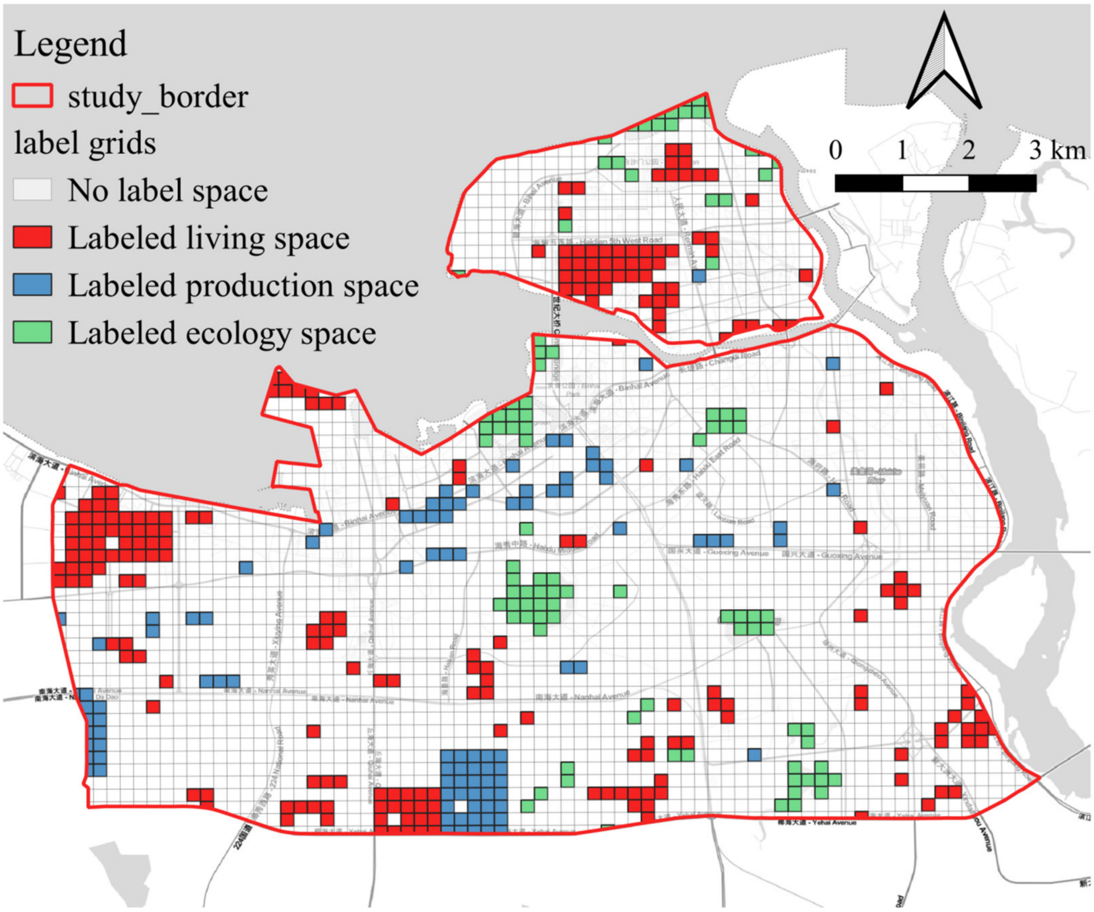


Figure 3. Distribution of the annotation set on the map (Road names are shown on the map in both English and Chinese).

Table 2. Definition of indicators.

Dimensions	Indicators	Calculation Formula	Explanation	Mean	Min	Max
The functional intensity of physical space	Housing function intensity	$x_1 = S_{\text{residence}} / S_{unit}$	Area of AOIs in the residential category as a proportion of the grid area. (%)	7.7	0	77.5
	Entertainment function intensity	$x_2 = n_{\text{entertainment}} / S_{unit}$	The density of the POI category entertainment in the grid. (pcs/km^2)	177.2	0	4477.8
	Educational function intensity	$x_3 = S_{\text{education}} / S_{unit}$	Area of AOIs in the education category as a proportion of the grid area. (%)	2.3	0	48.8
	Medical function intensity	$x_4 = S_{\text{hospital}} / S_{unit}$	Area of AOIs in the residential category as a proportion of the grid area. (%)	0.3	0	44.4
	Service facility intensity	$x_5 = n_{\text{service}} / S_{unit}$	The density of the POI category service in the grid. (pcs/km^2)	67.3	0	1022.2
	Office building intensity	$x_6 = n_{\text{industry}} / S_{unit}$	The density of the POI category industry in the grid. (pcs/km^2)	23.3	0	688.9
	Water coverage	$x_7 = S_{\text{water}} / S_{unit}$	Area of water as a proportion of the grid area. (%)	0.3	0	43.7
	Greenery coverage	$x_8 = S_{\text{green}} / S_{unit}$	Area of greenery as a proportion of the grid area. (%)	0.4	0	44.4
Travel characteristics of residents	Complexity of travel	$x_9 = \sqrt{\sum_{t=1}^{24} (x_{O,t} - x_{O,t-1})^2}$	$x_{O,t}$ is the number of cabs departing from the grid, where t represents the hour ranging from 1 to 24. ($x_{10} \times 10^4$)	36.4	0	1917.9
	Travel intensity	$x_{10} = \sum_{t=1}^{24} (x_{O,t})^2$	$x_{D,t}$ is the number of cabs arriving at the grid, where t represents the hour ranging from 1 to 24.	4.0302	0	1524.5
	Complexity of arrival	$x_{11} = \sqrt{\sum_{t=1}^{24} (x_{D,t} - x_{D,t-1})^2}$	Complexity is an estimate of the fluctuation level of the time series [31].	38.2	0	1864.3
	Arrival intensity	$x_{12} = \sum_{t=1}^{24} (x_{D,t})^2$	($x_{12} \times 10^4$)	8.5199	0	6864.7

Table 3. Examples of labeling.

ID	Name	Description	Label
1	Hainan University, Haidian Campus	A campus of Hainan University	Living space
2	Wanlv Garden	The largest open tropical seaside eco-garden in Haikou	Ecological space
3	Hainan Overseas Chinese High School	A high school	Living space
4	Hainan Haima Automobile Limited company	Automotive companies responsible for the development and manufacture of automotive components	Production space
5	Jinniuling Park	A large landscaped area in Haikou with a 96% greenery rate	Ecological space
6	Jing Rui Building	A commercial office building	Production space
7	The Second Affiliated Hospital of Hainan Medical College	A general hospital	Living space
8	Hongchenghu Park	An open park with a green area of 64,567.16 square meters	Ecological space

3.3. Identification of PLES by Random Forests

Traditional regression methods are mostly based on linear or curve assumptions to model the relationship between variables, which cannot model complex rules under the interaction of multiple variables [32]. Random forest models have been widely used for mining non-linear rules between variables because of their ability to prevent overfitting and tolerate outliers [33–35]. The random forest classification model [36] consists of many decision tree classifiers based on the idea of ensemble learning. Sub-datasets are constructed by sampling the original dataset with multiple inputs, and the data that are not sampled are called the out-of-bag data of the tree. Separate classifiers were constructed based on the sub-datasets used for training, and the classification result of the model was determined by the vote of all classifier output categories.

Three random forest classification models were built on these three datasets to identify the living, production, and ecological functions of the study area. Given that many geographical areas have multiple functions, it would not be practical to classify an area that contains both office buildings and residential areas as a single function using a multi-classification model. To address this, we developed three datasets for single-function recognition, using the functional labels provided in the annotated sets. For instance, to construct the production-oriented dataset, we labeled all areas in the annotation sets as non-production spaces, except for those that were explicitly identified as production spaces.

A total of 70% of each dataset was used as the training set and 30% was used as the validation set. Four important hyperparameters need to be considered in a random forest classification model, namely, n_{trees} indicating the number of decision trees, $t_{\text{criterion}}$ indicating the type of decision tree, max_depth indicating the maximum depth of the tree, and $max\_features$ indicating the maximum number of features considered when training. The hyperparameters for each classification model are determined by a grid search method, where n_{trees} ranges from 10 to 200 in intervals of 20, $t_{\text{criterion}}$ ranges from {gini, entropy}, and max_depth and $max\_features$ range from {3,5,7,9}. Ultimately, each grid can obtain classification results from the production-oriented, living-oriented, and ecological-oriented classification models, which are combined to form the grid's functional labels, as shown in Figure 4.

In addition, we analyze the identification rules for PLES using feature importance and radar plots obtained from a random forest model. Feature importance was determined by calculating the average decrease in impurity [37] for each PLES indicator, utilizing the Gini index. This helps us to identify the key indicators in the identification of PLES. Radar plots provided a visual representation of the average value of each PLES indicator. To validate our findings, given the challenges in obtaining precise spatial distribution data for PLES, we compared our results with previous studies, specifically examining the consistency and discrepancies. We used jobs–housing space [30] as a point of reference due to its significant correlation with living/production activities.

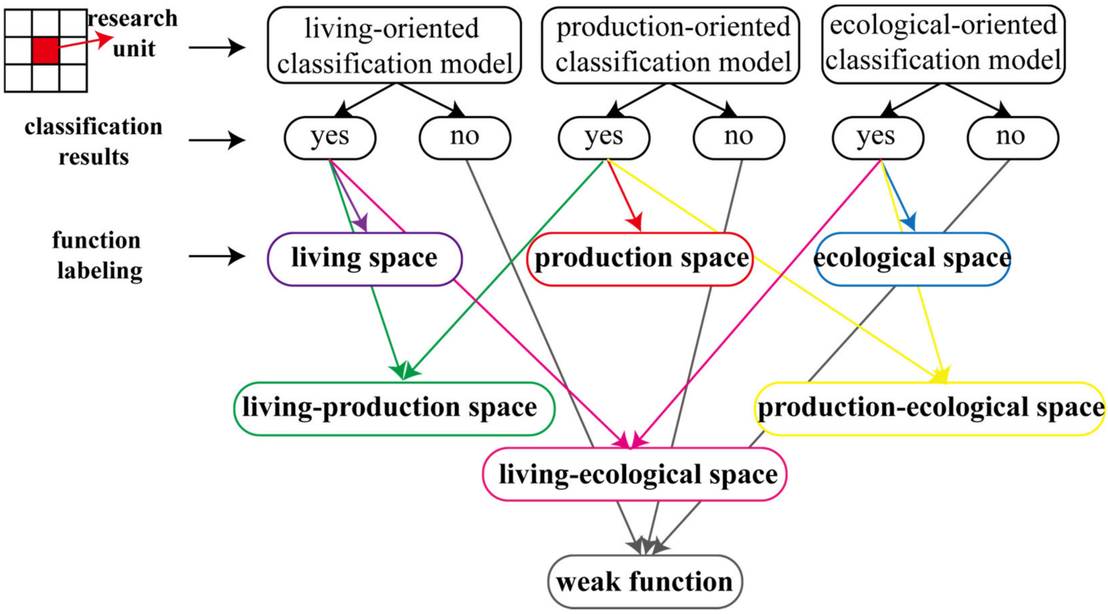


Figure 4. Rules applied when combining functional labels.

4. Results

4.1. Results of PLES Classification

To assess the effectiveness of the random forest classification models, we employed two metrics: classification accuracy and out-of-bag error. The classification accuracy measures the agreement between the true labels of the grid and the classification results, with higher accuracy indicating better model performance. Meanwhile, the out-of-bag error reflects the error rate of the classification model on out-of-bag data, and it is desirable for this metric to be as small as possible to ensure the model's reliability. According to Table 4, the accuracy of all three classification models exceeded 92%, and the maximum out-of-bag error was 0.1503, indicating that the three classification models have good generalization ability and can accurately identify the functional labels of the study units.

Table 4. Evaluation of random forest classification models.

Model	Accuracy	Out-of-Bag Error	n_{trees}	$t_{criterion}$	*max_dep*	*max_features*
Living-oriented classification	0.9338	0.1209	30	entropy	7	7
Production-oriented classification	0.9269	0.1503	70	entropy	9	9
Ecological-oriented classification	0.9292	0.1503	50	entropy	9	9

To further analyze the identification rules of the random forest model for PLES, the feature importance and radar maps for each functional space are plotted in Figure 5.

The values of housing function intensity and educational function intensity are much higher than those of the other indicators in Figure 5a, revealing the importance of these three factors for living space identification. Combined with the distribution of these three factors in Figure 5b, the definition of living space in the model can be interpreted as a place used by residents to live and receive an education. A similar phenomenon appears in Figure 5c, where only three indicators appear to be particularly important: office building intensity, housing function intensity, and educational function intensity. However, the strength of the educational function was low, as shown in Figure 5d. The resulting production spaces typically comprise clustered office buildings and fewer schools but may include residential areas. Figure 5e, showing the case with ecological space, is completely different; specifically, multiple metrics are important, but greenery coverage is extremely important. In addition, travel characteristics are highly important, but travel complexity and intensity are low (as shown in Figure 5f). This may be because we selected travel data of residents on weekdays, when people rarely visit ecological spaces. We deduced that ecological spaces refer to places where water and vegetation are located and are less visited on weekdays.

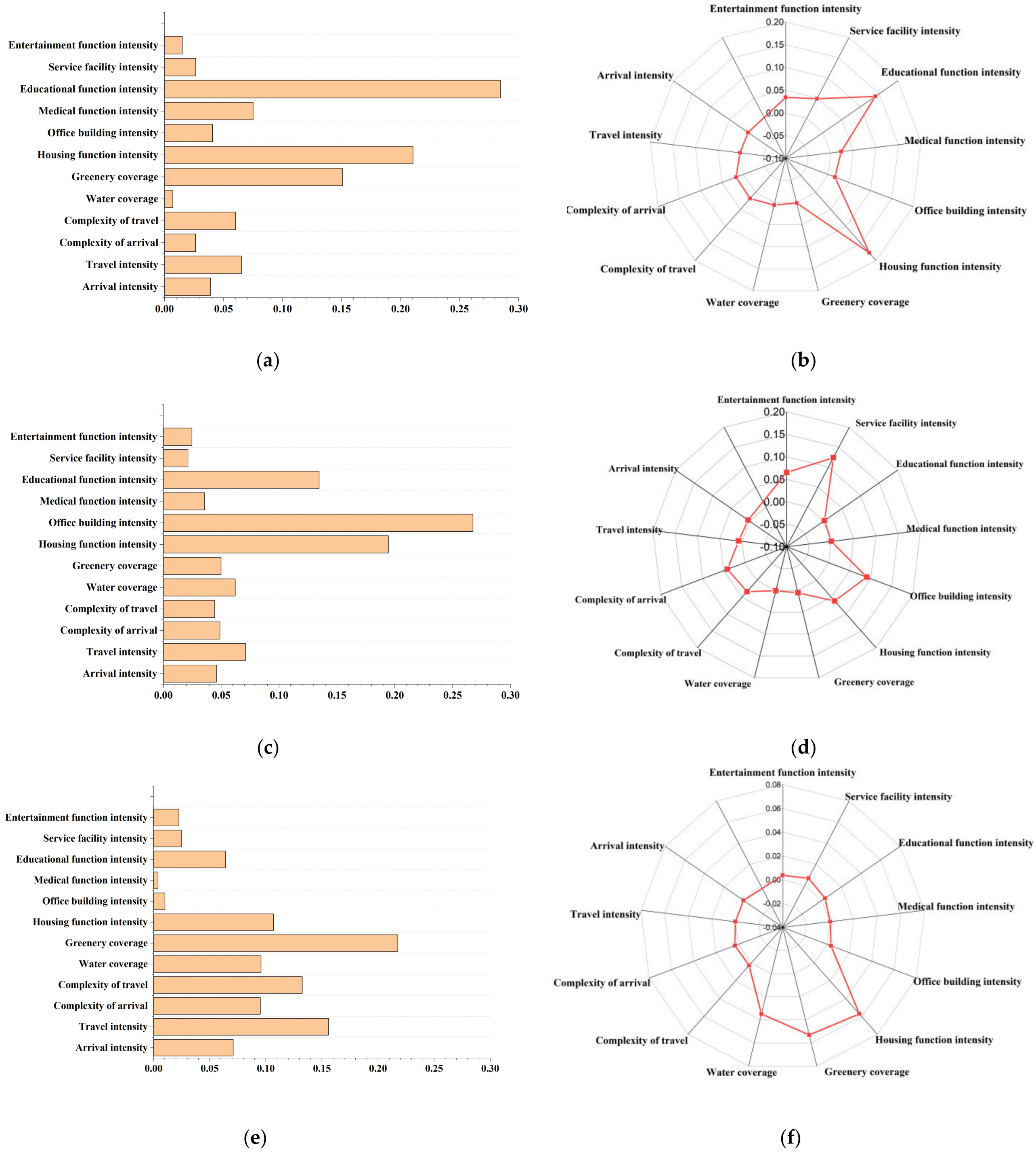


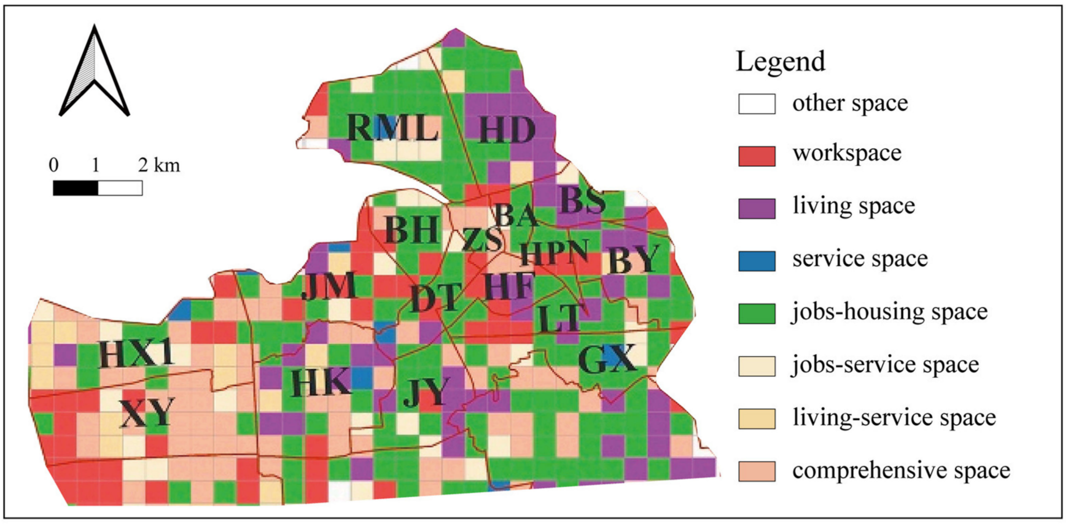
Figure 5. Feature importance and radar plots for the three classification models. (**a**) Feature importance of the living−oriented model; (**b**) Radar plot of the living−oriented model; (**c**) Feature importance of production−oriented model; (**d**) Radar plot of the production−oriented model; (**e**) Feature importance of ecological−oriented model; (**f**) Radar plot of the ecological−oriented model.

4.2. Comparison with Jobs–Housing Spaces

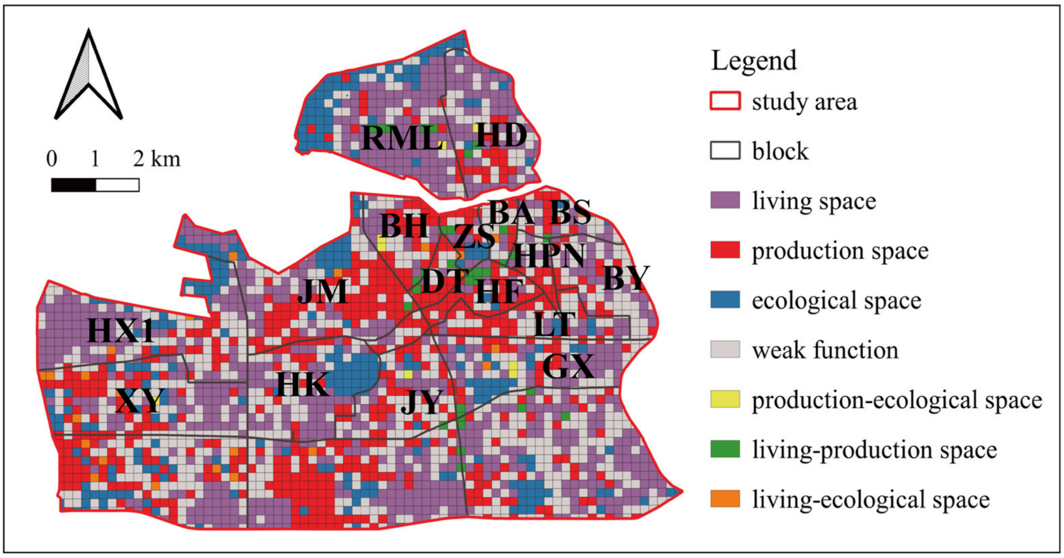
The results were compared with the jobs–housing space detected by Zhang et al. [30] to verify its correctness. In contrast to our approach, which uses supervised learning to learn rules from human cognitive labels, Zhang et al. [30] constructed rules to identify region functions based on human activity patterns and distributional features of POIs. The identification results of Zhang et al. and our method are shown in Figure 6, with similar blocks divided for ease of comparison. Because working is the main production activity of people in urban central areas, it is assumed that the workplace mentioned in Zhang et al.'s research is consistent with the content of our production space. As we can see, the clustering of production spaces occurs in the XY, JM, and LT blocks in Figure 6b, which is in line with the distribution of the workspace in Figure 6a. However, an inconsistency appears in the HD block, where some production spaces appear in Figure 6b while no workspace is shown in Figure 6a. This may be due to the fact that the production space we probe appears negligible on larger grids. The living spaces in both figures are represented in purple, but it should be noted that their definitions of living space are somewhat different. While the living spaces detected by Zhang et al. refer specifically to dwellings, our detected living spaces refer to places where people carry out activities related to their lives, including housing and education. Thus, it can be seen that the living space in Figure 6b is much larger than that in Figure 6a. In Figure 6a,b, the living areas are scattered in blocks such as BS, BY, HF, HK, and JY, and clustered in the HD block, which demonstrates the consistency of our identification results. However, the wide area of the jobs–housing space in the RML block in Figure 6a is identified as a living space in Figure 6b. According to the real online map, the RML block is the site of Hainan University surrounded by a large number of residential areas and services, which facilitates the identification of this area as a living space in this study. However, the residential function specified in Zhang et al.'s research is weakened here so that it is identified as a jobs–housing space. Similarly, the HX1 and XY blocks in both figures show different identification results affected by service facilities. In summary, we still find that the PLES identified in this study is plausible, despite the difference in the identification results between the two figures.

4.3. Analysis of Production–Living–Ecological Spaces

The detailed PLES identification results are listed in Table 5 and their spatial distributions are shown in Figure 7. A weak function refers to the fact that this grid polygon was not identified as a corresponding function in any of the three classification models; such areas represent the second largest percentage at 30.77%. Figure 7 demonstrates that weak functional areas are often located in the vicinity of secondary or tertiary roads. Secondary roads play a crucial role in connecting different parts of the city, such as residential areas and commercial districts. Similarly, tertiary roads, including footways and cycleways, support the daily travel needs of pedestrians and cyclists. It is, therefore, reasonable to infer that weak functional areas act as transitional zones between different functional spaces. These areas may lack significant functions of their own, but their location near secondary and tertiary roads means that they can serve as gateways to other parts of the city. This highlights the importance of well-planned transportation infrastructure in shaping the spatial organization of a city, as it can affect how people move through and interact with different areas. According to Table 5, the most significant daily needs of people are living and production, accounting for 31.25% and 21.87%, respectively. These functions tend to cluster within the area enclosed by primary roads, as indicated in Figure 7. This suggests that the road network plays a crucial role in shaping the activity space for people in the study area. Interestingly, only 13.71% of the grid in the study area was identified as an ecological space, which is much lower than the proportion of living space. According to the national garden city standards proposed by China's Ministry of Housing and Urban–Rural Development, the total area of all types of green areas and waters in the built-up area should account for more than 43% of the total area of the built-up area [38]. This indicates that the ecological space in the main urban area of Haikou still needs to be built. It is crucial to balance the development of living and production spaces with ecological preservation to ensure a sustainable and healthy living environment for the residents. Contrary to what was stated earlier, our results indicate that mixed spaces appear extremely rarely, at just below 3%. This is likely due to the resolution of our research units, as the study area was divided into a grid of polygons with 300 m sides, which may have limited the identification of hybrid features. As shown in Figure 7, a higher mix of the three functions is laid out in the central areas, meaning that the residents living there can satisfy their daily needs more easily, while the functions in the peripheral areas are mostly composed of a single function or a mix of two functions, which reveals the need to optimize the space in the peripheral areas.



(**a**)



(**b**)

Figure 6. The comparison between jobs–housing space and PLES. (**a**) Jobs–housing space identification results of the travel flow model considering the spatial distribution of public facilities. Reproduced from Zhang et al. [30]; (**b**) Identification results of PLES. BS, Baisha Street; BH, Binhai Street; BA, Boai Street; BY, Binjiang Street & MeiYuan Street; DT, Datong Street; GX, Guoxing Street; HD, Haidian Street; HF, Haifu Street; HK, Haiken Street; Hx1, Haixiu Street; HPN, Hepingnan Street; JM, Jinmao Street; JY, Jinyu Street; LT, Lantian Street; RML, Renminglu Street; XY, Xiuying Street; ZS, Zhongshan Street.

Table 5. Results of the identification of PLES.

Function Labeling	Number	Density
Living Space	873	31.25%
Production Space	611	21.87%
Ecological Space	383	13.71%
Living-production Space	32	1.15%
Living-ecological Space	8	0.29%
Production-ecological Space	27	0.96%
Undefined function	860	30.77%

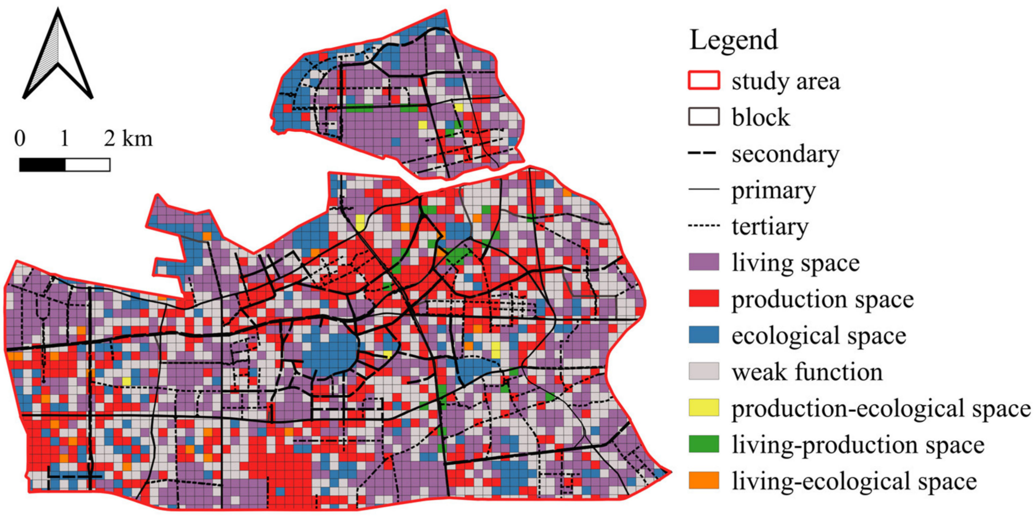
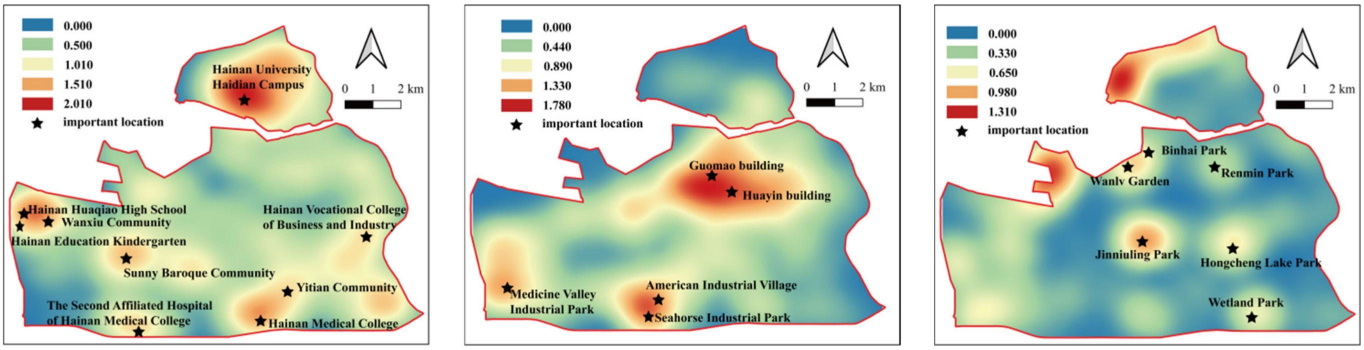


Figure 7. Results of the identification of PLES.

The awareness of the distribution characteristics of the three functions facilitates the development of targeted optimization policies. To this end, we computed kernel density estimates for the living space, production space, and ecological space, the results of which are shown in Figure 8. The living spaces in the main urban areas of Haikou show a well-defined circular structure, with low-density and scattered areas in the center and patches of high-density areas in the periphery. A comparison with real-world maps reveals that high-density living areas are usually located around schools, such as Hainan University, Hainan Overseas Chinese High School, and Hainan Ninth Primary School, suggesting that the construction of educational resources can promote the formation of urban living space clusters. Production space gathers towards the southwest, mainly formed around industrial parks in the southwest (e.g., Yakult Industrial Park, Haima Industrial Park) and office clusters in the center (e.g., Huayin Building and China World Trade Center), with this layout probably resulting in long-distance commuting within the city. The ecological space shows a sparse dotted distribution, mostly in park-based scenic areas such as Jinniuling Park and Wetland Park, indicative of a lack of green space in terms of residential areas and industrial estates. The Haikou government needs to promote a balanced layout of educational resources and industrial clusters, as well as to improve residents' well-being by increasing vegetation cover in the vicinity of residential areas and industrial parks.



(**a**) (**b**) (**c**)

Figure 8. Kernel density analysis of PLES: (**a**) living space; (**b**) production space; (**c**) ecological space.

5. Discussions and Conclusions

As an important part of the urban spatial structure, urban PLES have an important impact on people's daily activities. Accurate identification of urban PLES can help provide a theoretical basis for understanding human mobility and further guide the resolution of conflicts between human activities and functional layouts. In this regard, this study proposes a city-oriented approach to the identification of PLES using supervised learning techniques that fuse empirical knowledge and multi-source data. An accuracy of more than 92% and an out-of-bag error below 0.1503 guaranteed the soundness of the random forest model. In addition, a comparison with the jobs–housing pattern also confirms the validity of the method in this study. By analyzing the results of the PLES identification, the spatial optimization of Haikou was guided by the discovery of the regional aggregation of living and production spaces and the scarcity of ecological space in the major cities of Haikou.

Specifically, to address the biased nature of physical space features for functional recognition, we introduced human activity features to enrich the feature dimension. Although none of these human activity features occupy the top position in the feature importance results (Figure 5), it can be seen that travel characteristics are more important than arrival characteristics in the living space, whereas the difference is not obvious in the production space, which confirms the need to consider the human activity. At the same time, the insignificance of these features of human activity may also be related to the methods used to establish the indicators. By relying solely on the volume of travel within a 24 h period to represent the travel characteristics of a region, this approach may inadvertently conceal the temporal patterns associated with morning and evening peak occurrences, thereby impeding the accurate identification of production and living spaces. The oversimplification of travel data in this manner disregards the intricate dynamics that play out during different times of the day and might hinder the ability to discern distinct activity patterns within the analyzed dataset. Furthermore, the study's construction of human activity characteristics is limited in its scope as it exclusively focuses on two dimensions: the complexity of travel and the volume of travel. While these dimensions do offer valuable insights into overall activity patterns, the exclusion of other pertinent dimensions, such as the temporal variation in travel intensity, restricts the depth of understanding regarding human activity.

To address the limitations of linear weighting rules in indicator synthesis, the method of random forest classification was applied to mine the non-linear rules between empirical perception and objective description, which makes the identification of functional spaces quicker and more flexible, as we only need to use the human experience to modify the annotated set when oriented to different optimization goals, reducing human effort. In addition, we analyzed the definition of PLES based on feature importance. Huang et al. [15] pointed out that production space mainly provides people with products and services; living space is the space for human activities to meet the needs of housing, consumption, and entertainment, while ecological space is the space to provide ecological products. Zhang et al. [39] defined PLES from the perspective of land function cognition, that is, production space is the land that carries out agricultural, industrial, and commercial activities to obtain products and supply functions; living space provides the function of carrying and securing human habitation; and ecological space focuses on the function of regulating, maintaining, and securing ecological security. Compared with the generality of these definitions, this study provides the concept of PLES concretely through the random forest model, which enriches the theoretical knowledge of related studies. Additionally, we constructed learning units based on grid polygons with a side length of 300 m, which encouraged the study of functional patterns at fine scales. Meanwhile, we found that more mixed functions may appear when the study cell becomes larger than the jobs–housing space identified by Zhang et al. [30], emphasizing the important influence of the size of the study unit on the identification results.

Future research can be improved by considering further the following aspects: (1) Consider more human activity characteristics. In addition to the functional characteristics identified in our study, future research can consider more human activity characteristics such as morning and evening peak travel patterns. Such patterns can help to distinguish between residential areas and workplaces more effectively. (2) Conduct multi-scale studies. Future research can conduct multi-scale studies to identify the functional characteristics of different regions at various levels of granularity. Integrating functional identification results across multiple scales can provide domain knowledge for spatial optimization, as the optimization of a target region usually requires considering its neighboring regions.

Author Contributions: Conceptualization, Bingbing Zhao and Xiaoyong Tan; methodology, Bingbing Zhao; data curation, Xiaoyong Tan; writing—original draft preparation, Bingbing Zhao; writing—review and editing, Xuexi Yang; supervision, Xuexi Yang, Liang Luo, and Min Deng. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by grants from the National Natural Science Foundation of China (42271485 and 42171459), the Natural Science Foundation of Hunan Province (2022JJ40585), the Scientific Research Project of Education Department of Hunan Province (22B0015), the Research Foundation of the Department of Natural Resources of Hunan Province (20230121XX), and the Scientific Research Project of Natural Resources Department of Hunan Province (2013-17, 2014-12, 2015-09, and 2017-15).

Data Availability Statement: Publicly available datasets were analyzed in this study and links of them were shared in the manuscript.

Acknowledgments: We are grateful for the support of the Didi Chuxing GAIA Initiative (https://gaia.didichuxing.com (accessed on 22 March 2021)), and we would like to thank the editors and anonymous reviewers for their time in reviewing this manuscript. This work was carried out in part using computing resources at the High-Performance Computing Platform of Central South University.

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

References

1. Xie, W.; Huang, Q.; He, C.; Zhao, X. Projecting the Impacts of Urban Expansion on Simultaneous Losses of Ecosystem Services: A Case Study in Beijing, China. *Ecol. Indic.* **2018**, *84*, 183–193. [CrossRef]
2. Zou, L.; Liu, Y.; Wang, J.; Yang, Y.; Wang, Y. Land Use Conflict Identification and Sustainable Development Scenario Simulation on China's Southeast Coast. *J. Clean. Prod.* **2019**, *238*, 117899. [CrossRef]
3. Franco, L.; Magalhães, M.R. Assessing the Ecological Suitability of Land-Use Change. Lessons Learned from a Rural Marginal Area in Southeast Portugal. *Land Use Policy* **2022**, *122*, 106381. [CrossRef]
4. Wang, B.; Tian, J.; Wang, S. Process and Mechanism of Transition in Regional Land Use Function Guided by Policy: A Case Study from Northeast China. *Ecol. Indic.* **2022**, *144*, 109527. [CrossRef]
5. Fang, F.; Zeng, L.; Li, S.; Zheng, D.; Zhang, J.; Liu, Y.; Wan, B. Spatial Context-Aware Method for Urban Land Use Classification Using Street View Images. *ISPRS J. Photogramm. Remote Sens.* **2022**, *192*, 1–12. [CrossRef]
6. Wang, D.; Jiang, D.; Fu, J.; Lin, G.; Zhang, J. Comprehensive production–living–ecological space assessment based on the coupling coordination degree model. *Sustainability* **2020**, *12*, 2009. [CrossRef]
7. Li, C.; Wu, J. Land Use Transformation and Eco-Environmental Effects Based on Production-Living-Ecological Spatial Synergy: Evidence from Shaanxi Province, China. *Environ. Sci. Pollut. Res.* **2022**, *29*, 41492–41504. [CrossRef]
8. Lyu, Y.; Wang, M.; Zou, Y.; Wu, C. Mapping Trade-Offs among Urban Fringe Land Use Functions to Accurately Support Spatial Planning. *Sci. Total Environ.* **2022**, *802*, 149915. [CrossRef]

9. Liu, Y.; Liu, X.; Zhao, C.; Wang, H.; Zang, F. The Trade-Offs and Synergies of the Ecological-Production-Living Functions of Grassland in the Qilian Mountains by Ecological Priority. *J. Environ. Manag.* **2023**, *327*, 116883. [CrossRef]
10. LI Guangdong, F.C. Quantitative Function Identification and Analysis of Urban Ecological-Production-Living Spaces. *Acta Geogr. Sin.* **2016**, *71*, 49. [CrossRef]
11. Duan, Y.; Wang, H.; Huang, A.; Xu, Y.; Lu, L.; Ji, Z. Identification and Spatial-Temporal Evolution of Rural "Production-Living-Ecological" Space from the Perspective of Villagers' Behavior—A Case Study of Ertai Town, Zhangjiakou City. *Land Use Policy* **2021**, *106*, 105457. [CrossRef]
12. Chen, H.; Yang, Q.; Su, K.; Zhang, H.; Lu, D.; Xiang, H.; Zhou, L. Identification and Optimization of Production–Living–Ecological Space in an Ecological Foundation Area in the Upper Reaches of the Yangtze River: A Case Study of Jiangjin District of Chongqing, China. *Land* **2021**, *10*, 863. [CrossRef]
13. Chen, Y.; Zhu, M. Spatiotemporal Evolution and Driving Mechanism of "Production-Living-Ecology" Functions in China: A Case of Both Sides of Hu Line. *Int. J. Environ. Res. Public Health* **2022**, *19*, 3488. [CrossRef]
14. Fan, Y.; Jin, X.; Gan, L.; Jessup, L.H.; Pijanowski, B.C.; Yang, X.; Xiang, X.; Zhou, Y. Spatial Identification and Dynamic Analysis of Land Use Functions Reveals Distinct Zones of Multiple Functions in Eastern China. *Sci. Total Environ.* **2018**, *642*, 33–44. [CrossRef]
15. Huang, A.; Xu, Y.; Lu, L.; Liu, C.; Zhang, Y.; Hao, J.; Wang, H. Research Progress of the Identification and Optimization of Production–Living–Ecological Spaces. *Prog. Geogr.* **2020**, *39*, 503. [CrossRef]
16. Liao, G.; He, P.; Gao, X.; Deng, L.; Zhang, H.; Feng, N.; Zhou, W.; Deng, O. The Production–Living–Ecological Land Classification System and Its Characteristics in the Hilly Area of Sichuan Province, Southwest China Based on Identification of the Main Functions. *Sustainability* **2019**, *11*, 1600. [CrossRef]
17. Bian, Z.; Cheng, X.; Yu, M.; Li, H.; Cui, W. The Proportionality of the Functions of Production, Life and Ecology in Connection Zone between Shenyang and Fushan. *Chin. J. Agric. Resour. Reg. Plan.* **2016**, *12*, 84–92. [CrossRef]
18. Fu, C.; Tu, X.; Huang, A. Identification and Characterization of Production–Living–Ecological Space in a Central Urban Area Based on POI Data: A Case Study for Wuhan, China. *Sustainability* **2021**, *13*, 7691. [CrossRef]
19. Lin, G.; Fu, J.; Jiang, D. Production–Living–Ecological Conflict Identification Using a Multiscale Integration Model Based on Spatial Suitability Analysis and Sustainable Development Evaluation: A Case Study of Ningbo, China. *Land* **2021**, *10*, 383. [CrossRef]
20. Zou, L.; Liu, Y.; Yang, J.; Yang, S.; Wang, Y.; Zhi, C.; Hu, X. Quantitative Identification and Spatial Analysis of Land Use Ecological-Production-Living Functions in Rural Areas on China's Southeast Coast. *Habitat Int.* **2020**, *100*, 102182. [CrossRef]
21. Zhang, X.; Xu, Z. Functional Coupling Degree and Human Activity Intensity of Production–Living–Ecological Space in Underdeveloped Regions in China: Case Study of Guizhou Province. *Land* **2021**, *10*, 56. [CrossRef]
22. Fragou, S.; Kalogeropoulos, K.; Stathopoulos, N.; Louka, P.; Srivastava, P.K.; Karpouzas, S.; Kalivas, D.P.; Petropoulos, G.P. Quantifying Land Cover Changes in a Mediterranean Environment Using Landsat TM and Support Vector Machines. *Forests* **2020**, *11*, 750. [CrossRef]
23. Berhane, T.M.; Lane, C.R.; Wu, Q.; Autrey, B.C.; Anenkhonov, O.A.; Chepinoga, V.V.; Liu, H. Decision-Tree, Rule-Based, and Random Forest Classification of High-Resolution Multispectral Imagery for Wetland Mapping and Inventory. *Remote Sens.* **2018**, *10*, 580. [CrossRef]
24. Ministry of Housing and Urban-Rural Development of China; China Academy of Urban Planning and Design, and NavInfo Corporation. Annual Road Network Density Monitoring Report for Major Cities in China. Available online: http://www.199it.com/archives/719480.html (accessed on 22 March 2021).
25. Yang, J.; Huang, X. The 30 m Annual Land Cover Dataset and Its Dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 3907–3925. [CrossRef]
26. Ministry of Transport of China. Passenger Traffic in Central Cities in January 2020. Available online: https://xxgk.mot.gov.cn/2020/jigou/zhghs/202006/t20200630_3321315.html (accessed on 29 June 2023).
27. Li, J.; Sun, W.; Li, M.; Meng, L. Coupling Coordination Degree of Production, Living and Ecological Spaces and Its Influencing Factors in the Yellow River Basin. *J. Clean. Prod.* **2021**, *298*, 126803. [CrossRef]
28. Tang, S.; Ta, N. How the built environment affects the spatiotemporal pattern of urban vitality: A comparison among different urban functional areas. *Comput. Urban Sci.* **2022**, *2*, 39. [CrossRef]
29. Liu, H.; Xu, Y.; Tang, J.; Deng, M.; Huang, J.; Yang, W.; Wu, F. Recognizing Urban Functional Zones by a Hierarchical Fusion Method Considering Landscape Features and Human Activities. *Trans. GIS* **2020**, *24*, 1359–1381. [CrossRef]
30. Zhang, Y.; Liu, J.; Wang, Y.; Cao, Y.; Bai, Y. Research on the Method of Urban Jobs-Housing Space Recognition Combining Trajectory and POI Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 71. [CrossRef]
31. Batista, G.E.A.P.A.; Keogh, E.J.; Tataw, O.M.; de Souza, V.M.A. CID: An Efficient Complexity-Invariant Distance for Time Series. *Data Min. Knowl. Discov.* **2014**, *28*, 634–669. [CrossRef]
32. Xue, F.; Yao, E. Adopting a Random Forest Approach to Model Household Residential Relocation Behavior. *Cities* **2022**, *125*, 103625. [CrossRef]
33. Wu, R.; Wang, J.; Zhang, D.; Wang, S. Identifying Different Types of Urban Land Use Dynamics Using Point-of-Interest (POI) and Random Forest Algorithm: The Case of Huizhou, China. *Cities* **2021**, *114*, 103202. [CrossRef]
34. Xu, F.; Ho, H.C.; Chi, G.; Wang, Z. Abandoned Rural Residential Land: Using Machine Learning Techniques to Identify Rural Residential Land Vulnerable to Be Abandoned in Mountainous Areas. *Habitat Int.* **2019**, *84*, 43–56. [CrossRef]

35. Xu, L.; Huang, Q.; Ding, D.; Mei, M.; Qin, H. Modelling Urban Expansion Guided by Land Ecological Suitability: A Case Study of Changzhou City, China. *Habitat Int.* **2018**, *75*, 12–24. [CrossRef]
36. Reiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
37. Louppe, G.; Wehenkel, L.; Sutera, A.; Geurts, P. Understanding Variable Importances in Forests of Randomized Trees. In *Proceedings of the Advances in Neural Information Processing Systems*; Burges, C.J., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2013; Volume 26.
38. Ministry of Housing and Urban-Rural Development of China. National Garden City Selection Criteria. Available online: https://www.gov.cn/zhengce/zhengceku/2022-01/14/content_5668177.htm (accessed on 29 June 2023).
39. Zhang, H.; Xu, E. An Ecological-Living-Industrial Land Classification System and Its Spatial Distribution in China. *Resour. Sci.* **2015**, *37*, 1332.