

Research Article

A New Method for Identifying Key and Common Themes Based on Text Mining: An Example in the Field of Urban Expansion

Yanwei Zhang ¹, Xinhai Lu ^{1,2}, Chaoran Lin ^{3,4}, Feng Wu ¹ and Jinqiu Li ⁴

¹College of Public Administration, Huazhong University of Science and Technology, Wuhan 430074, China

²College of Public Administration, Central China Normal University, Wuhan 430079, China

³College of Computer Science and Technology, Harbin Engineering University, Harbin, Heilongjiang 150001, China

⁴School of Economics & Management, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Chaoran Lin; hcab@qq.com

Received 3 August 2021; Accepted 26 September 2021; Published 14 October 2021

Academic Editor: Jorge E. Macias-Diaz

Copyright © 2021 Yanwei Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Urban land use is a core area of multidisciplinary research that involves geography, land science, and urban planning. With the rapid progress of global urbanization, urban expansion has become a research focus in recent years. Therefore, how to scientifically and accurately identify key and common themes in the urban expansion literature has become crucial for scientific research institutions in various countries. This paper proposes a new framework for identifying such themes based on an analysis of scientific literature and by using text mining and thematic evolutionary analysis. First, the latent Dirichlet allocation algorithm is used to capture the thematic clustering of scientific literature. Second, the key degree of the thematic node in the thematic evolution transfer network is used to represent the key feature of a theme, and the PageRank algorithm is employed to measure the critical score of this theme. When recognizing common themes, the common features of various themes are digitized and mapped to a specially selected quadratic function to measure the degree of commonness. Finally, the hidden Markov model is used to build a thematic prediction model. This method can efficiently identify key and common themes from the literature and provide theoretical and technical support for future research in related fields.

1. Introduction

The increasingly drastic land-use changes during the process of urbanization are important factors that affect the global social economy and ecological stability [1, 2]. For half a century, many areas in the world have undergone rapid urbanization, thereby resulting in the continuous emergence of large cities and megacities [3, 4]. Accordingly, urban expansion has become a research hotspot in the fields of geography, economics, ecology, environmental science, and sociology [5]. Urban expansion refers to an increase in the total area of urban land and the outward development of land use under the influence of economic development, population growth, urban planning, and urbanization [6]. However, due to the lack of reasonable planning and guidance for urbanization, the process of urban expansion is often associated with an excessive

demand for urban land and the disorderly development of urban fringe areas [7], both of which have a negative impact on regional economic growth, the layout of production and living spaces, residential types, urban morphology, and ecological environment [8–10]. Therefore, how to guarantee a continuous demand for urban land for social-economic development in the new era and how to clarify the formation mechanism of urban growth in order to reasonably control the urban scale, delimit the urban growth boundary, and optimize the spatial pattern of urban land have become the working emphases of current urban management [11, 12]. Scholars have carried out extensive research to clarify those factors that drive urban expansion [13, 14], strengthen the development and application of dynamic monitoring technologies for urban expansion [15, 16], and accelerate sustainable urban spatial planning [17].

The scientific literature is an important and authoritative knowledge carrier. Using bibliometrics and text mining methods to study the thematic evolution and thematic prediction of the massive urban expansion literature can help trace the development trajectory and grasp the flow of knowledge in the field of urban expansion. In recent years, some scholars have qualitatively combed the findings of urban expansion research from multiple levels, dimensions, and perspectives [18, 19]. Some scholars have also used knowledge maps and bibliometric methods to quantify and visualize the results and topics in urban expansion research [20, 21]. However, as the number of documents has increased exponentially, the types of these documents have also become increasingly abundant. Accordingly, thematic identification is increasingly being used in big data scientific literature identification. When automatic thematic identification is faced with high data dimensions and complex data types, traditional thematic identification methods may become ineffective.

This article divides themes into key and common themes. Key themes play relatively important roles in the urban expansion field. These themes have attracted significant concern, are mature, and have development potential in each time window. The evolutionary process of key themes plays an important role in describing the future development of a theme [22]. In addition, in the globalization context, cities have increasingly become the power centers of global social-economic development. Whether in developed or developing countries, the status and role of cities in national development have become increasingly important and have begun to represent countries in a global competition. Therefore, urban development and expansion play key roles in the competition among countries in the context of globalization. This article then defines common themes as “themes that have received equal amounts of attention from scholars in developed and developing countries.”

How to accurately identify themes has always been a challenge in the field of bibliometrics. Statistical methods based on word frequency and co-occurrence frequency are widely used in thematic identification. However, these methods only simulate the literature as a language package and do not fully consider the relationship among themes, and revealing the rich thematic information contained in the literature is not an easy task [23]. The topic model represented by latent Dirichlet allocation (LDA) uses Dirichlet distribution to describe the literature generation process and obtains vocabulary clusters by maximizing the co-occurrence probability of keywords. This model can avoid parameter explosion and overfitting problems and can effectively extract hidden themes from the literature [24]. However, this method requires predefined empirical values and can only reveal the potential semantic relationships among themes. To solve these problems, scholars have recently examined thematic identification by constructing networks and comprehensively evaluating certain indicators, such as network centrality. For example, by constructing a citation network, Shibata et al. [25] demonstrated the novelty of themes from the time and function

dimensions and detected the emerging themes in regenerative medicine. Small [26] and Lee and Choe [27] not only considered the novelty of themes in their identification method but also employed network time series analysis and structural hole theory to measure the characteristics of thematic growth and influence. However, these studies conduct static analysis based on historical literature data, which are unable to reflect the dynamic development characteristics of a network in real-time. Therefore, how to construct a dynamic network of themes in the field of urban expansion and how to accurately identify key and common themes that can help scholars achieve breakthroughs in this field are crucial.

Identifying themes is not the end. By expanding research on thematic evolution, this article aims to predict the future development of these themes. Researchers have divided thematic prediction methods into qualitative and quantitative analyses based on different theoretical foundations. While qualitative methods are often limited by subjective judgment, quantitative methods are highly scientific. The most commonly used forecasting methods include the gray forecasting method [28], the life cycle method [29], time series analysis [30], and the neural network forecasting method [31]. However, uncertainty, ambiguity, and randomness are essential phenomena in scientific research, and the above models generally ignore the randomness in the development of technological innovation. The hidden Markov model (HMM) with a double stochastic process can describe the Markov stochastic process of mutual transfer among various themes, reveal the potential evolution path, and provide a basis for predicting future thematic development [32].

This article first applies the LDA model for topic modelling based on the title and abstract of urban expansion literature to obtain a detailed thematic classification. Second, in identifying key themes, the key degree of the thematic node in the thematic evolutionary transfer network is regarded as the key feature of themes. Third, the PageRank algorithm is employed to measure the criticality score of each thematic node in the thematic relationship network. When identifying the common theme, the common features of various themes are digitized and mapped to a specially selected quadratic function to measure the degree of commonality. Finally, by using the HMM, the future development trend of each theme is predicted from the microcosmic angle of the thematic evolution, and a visual display is given.

2. Materials and Methods

2.1. Data Sources. In order to ensure the quality and completeness of the sample, this paper selects the Web of Science core collection database for retrieval. Web of Science includes articles, reviews, editorials, letters, and other document types. Considering that the article is more creative and the results are more complete, this article only selects the article for retrieval. The search formula is $TI = \text{“Urban expansion” or “Urban extension” or “Urban Growth Boundaries” or “Urban land growth” or “Urban land expansion” or “Urban sprawl”}$ [21, 33], and the time span is

“1985–2020” (November 4, 2020). A total of 1,933 papers are retrieved, and the bibliographic information is downloaded and summarized in the form of full records (including references). The bibliographic items used in this article include article titles, abstracts, publication years, and reprint addresses, which provide the name, organization, and country information of the corresponding author. Given that international collaborative papers contain information from multiple countries, the corresponding author is used to determine the country of each paper (when the corresponding author is not specified, the first author is used). Each article is attributed to only one country to prevent international co-authored papers from affecting the accuracy of national distinction. After the data deduplication, cleaning, and sorting, a total of 969 documents from developed countries and 1,045 documents from developing countries are obtained.

2.2. Research Methods

2.2.1. Thematic Extraction Module. Scholars have investigated the concept of thematic identification by using the topic model. The current mainstream model adopted in thematic identification is the LDA model proposed by Blei et al. [34]. As a text mining method based on unsupervised machine learning, LDA can dig out potential themes from documents while overcoming the shortcomings of traditional methods in calculating text similarity. In addition, the LDA model can express scientific literature in the form of thematic probability vectors, thereby greatly reducing the dimensionality of the literature data and improving the accuracy of text classification and thematic identification. LDA and its improved models have been widely used in text analysis. The output of these models is usually obtained based on the distribution of words under each theme in order to extract high-frequency keywords to describe the themes and achieve excellent thematic classification results [35]. The hidden themes in the urban expansion literature are assumed to follow the distribution

$$\text{Dir}(\theta_d | \alpha) = \frac{\Gamma(\sum_{k=1}^k \alpha_k)}{\prod_{k=1}^k \Gamma(\alpha_k)} \prod_{k=1}^k \theta_{dk}^{\alpha_k - 1}, \quad (1)$$

where θ_{dk} represents the distribution of theme k in scientific literature d . The thematic term distributions $\varnothing_k \sim \text{Dir}(\beta)$ and $\theta_d \sim \text{Dir}(\alpha)$ are generated for themes k and d , respectively, and the thematic term $Z_{dn} \sim \text{Multinomial}(\varnothing_{Z_{dn}})$ is generated for the n -th term in each literature. Therefore, the LDA likelihood model can be described as follows:

$$P(W | \alpha, \beta) = \prod_{d=1}^D \int P(\theta_d | \alpha) \prod_{n=1}^{N_d} \sum_{Z_d} P(Z_{dn} | \theta_d) P(w_{dn} | \varnothing_{Z_{dn}}) d\theta_d. \quad (2)$$

This paper uses Heinrich's parameter estimation method, where $\alpha = 50/k$ and $\beta = 0.1$, and Gibbs sampling to obtain the theme set $K = \{k_1, \dots, k_h\}$ and theme attribution set $D_k = \{j_1, \dots, j_n\}$ of each paper.

2.2.2. Key Theme Identification Model. Based on the connotation of key themes, during the model construction, the thematic evolution is regarded as a hidden Markov process to obtain the thematic transfer network. There are two dynamics for thematic evolution in the field of urban expansion research: one is the inspiration of historical research results and the emergence of new ideas in the process of thematic evolution. However, due to the lack of a record carrier, this process is an unobservable hidden sequence; the second is that under the impetus of the first kind of driving force, as the research environment changes and unexpected research results appear, scholars constantly adjust their research thinking and then change the research direction. The professional literature effectively records the research results into an observable sequence. The latter constitutes the microfoundation of the former, and the former is the macroscopic manifestation of the latter. Therefore, the thematic evolution in the field of urban expansion can be seen as the superposition of these two processes. This research uses HMM to describe the evolution process of urban expansion theme. By inferring the state transition matrix and the probability distribution of the initial state in the HMM, the confusion and transition matrices between the themes in the evolution of themes are determined, and then the evolution history and future evolution trends of the themes are determined. Afterward, the criticality of these themes is measured based on their network relationship. The PageRank algorithm is then used to calculate the scores of network nodes in the thematic transfer network and serve as the foundation of key theme identification. This process is specifically described as follows:

- (1) Set the hidden state random transition sequence set of HMM to $S = \{s_1, \dots, s_h\}$, where h is the number of themes generated in the LDA model. Suppose that the hidden state sequence generated by the random process is $Q = \{q_1, \dots, q_t\}$, where $q_t \in S$.
- (2) The probability distribution of the transition state is $A = \{\alpha_{ij}\}$, where $\alpha_{ij} = P\{(q_{t+1} = S_j | q_t = S_i)\}$, $1 \leq i$, and $j \leq N$, and satisfies $\alpha_{ij} \geq 0$, $\sum_{j=1}^N \alpha_{ij} = 1$, which suggests that during the development of the urban expansion field, the themes will shift from state S_i to S_j .
- (3) When the state is S_i , the probability distribution of the observed variable is $B = \{b_j(v)\} = \{f(Q_t = v | q_t = S_i)\}$, where Q_t is the t -th observation variable. The observation sequence is $O = \{O_1, \dots, O_t\}$ or the proportion of each theme over the years.
- (4) The probability distribution of the initial state of the system is $\pi = \{\pi_i, 1 \leq i \leq N\}$, where π_i is the occurrence probability of state S_i . Given that a higher frequency of theme co-occurrence will facilitate the shift and evolution among themes, this paper uses the thematic co-occurrence matrix as the initial iteration value of the state transition matrix π_i .
- (5) Set the initial value of model training to $O = Q = \pi$. This paper uses the Baum–Welch algorithm [36] to

obtain the following single optimal state transition matrix:

$$U = \pi' = \begin{bmatrix} \pi'_{11} & \cdots & \pi'_{1h} \\ \vdots & \ddots & \vdots \\ \pi'_{h1} & \cdots & \pi'_{hh} \end{bmatrix}, \quad (3)$$

where π'_{nm} is the probability of transition from theme n to theme m . By extracting all π'_{nm} that exceed a certain threshold, a directed graph of the topological relationship between themes in the transfer network can be established. Key theme identification has always been an important research problem in thematic network analysis. PageRank processes the search results of thematic matching based on a web page link analysis. As the most famous web page ranking algorithm, the PageRank algorithm has been widely used to monitor key nodes in various directed, undirected, weighted, or unweighted networks [37, 38]. Applying this algorithm to compute for the centrality of thematic network nodes presents a very meaningful research problem.

When calculating the PageRank value of theme i at each moment in a dynamic thematic network, the network topology structure of the current snapshot and the influence of the previous centrality on the existing network should both be considered. One effective method for achieving this goal is network reconstruction, where the previous network topology relationships are weighed into the current network to construct a new network. To describe the dynamic network G , this dynamic network needs to be sampled at different times. The sampling results are then arranged in a time sequence to obtain the time sequence network $G = \langle G_1, G_2, \dots, G_{t-1}, G_t, \dots, G_n \rangle$, where G_t represents the sampling result at time t or the snapshot at time t . The analysis of the dynamic network is transformed into the analysis of the sequential network [39]. Let $G'_1 = G_1$, $G'_i = \alpha * G_{i-1}' + (1 - \alpha) * G_i$, $i \geq 2$. In this case, $G' = \langle G'_1, G'_2, \dots, G'_n \rangle$ can be obtained. The parameter $\alpha \in [0, 1]$ is used to balance the contribution of the current and previous topologies to the centrality of the network node. The PageRank centrality of node i at time t is treated as the PageRank value of node i in the construction network G'_i . The key theme in this article refers to any mature theme that has development potential in scientific research. Therefore, key themes can easily achieve migrating power in the process of thematic network transfer. Given that the G score can measure the importance of nodes in the process of directed network migration, this paper takes the standardized score G' as the critical score for each theme.

2.2.3. Common Theme Identification Model. To identify the degree of commonness of themes, the selected model should be able to measure the common skewness of different themes in the field of urban expansion. Skewness refers to the numerical characteristics of the asymmetric degree in the statistical data distribution [40–42]. Common

skewness in this article refers to the measurement of the direction and degree of skewness of each thematic distribution. Let $k \in K = \{k_1, \dots, k_m\}$, and theme k corresponds to the number of documents p_k . A type documents (developed countries) correspond to a_k themes, and B type documents (developing countries) correspond to b_k themes. Let $p_k = a_k + b_k$, and define U_{ka} and U_{kb} as the common skewness of A and B type documents in theme k as follows:

$$U_{ka} = \frac{a_k / \sum_{k=1}^m a_k}{(a_k / \sum_{k=1}^m a_k) + (b_k / \sum_{k=1}^m b_k)}, \quad (4)$$

$$U_{kb} = \frac{b_k / \sum_{k=1}^m b_k}{(a_k / \sum_{k=1}^m a_k) + (b_k / \sum_{k=1}^m b_k)}, \quad (5)$$

where $U_{ka} + U_{kb} = 1$. Formulas (4) and (5) eliminate the influence of the number of A and B documents on common skewness to prevent the difference in the number of documents from affecting the calculation of the co-occurrence degree of themes.

When the common skewness is $U_{ka} = U_{kb} = 1/2$, that is, theme t comprises the themes of A and B type documents, such skewness indicates the highest degree of commonness. By mapping U_{ka} and U_{kb} to the $[0, 1]$ range of the inverted quadratic function, the common function can be monotonized. Without loss of generality, this quadratic function relationship is set to

$$\begin{aligned} C_{ka} &= 1 - 4 * (U_{ka} - 1/2)^2, \\ C_{kb} &= 1 - 4 * (U_{kb} - 1/2)^2. \end{aligned} \quad (6)$$

By learning $C_{ka} = C_{kb}$ from the symmetry of the quadratic function, $C_k = C_{ka} = C_{kb}$ can be obtained. The C_k function is then used to measure the degree of commonness. The highest and lowest degrees of commonness are measured when $C_k = 1$ and $C_k = 0$, respectively. The logic structure is shown in Figure 1.

3. Results and Analysis

3.1. Data Preprocessing and Thematic Extraction from the Urban Expansion Literature. The preprocessing work in this article mainly involves word segmentation, removal of stop words, root restoration, and marked information removal. In view of the language characteristics of English articles, the words in a text can be directly divided by spaces and punctuation. Removal of stop word removes those words that do not provide useful information for the text analysis, such as auxiliary words, pronouns, conjunctions, and adverbs. According to the characteristics of the collected urban expansion literature, this article expands the stop words to include some additional words (e.g., “data,” “study,” and “use”) that are unique in the field and appear frequently yet have no effects on the experimental results. Root restoration restores words to their corresponding roots. After such processing, the number of feature items in the sample set can be greatly reduced, and the efficiency of thematic extraction can be improved.

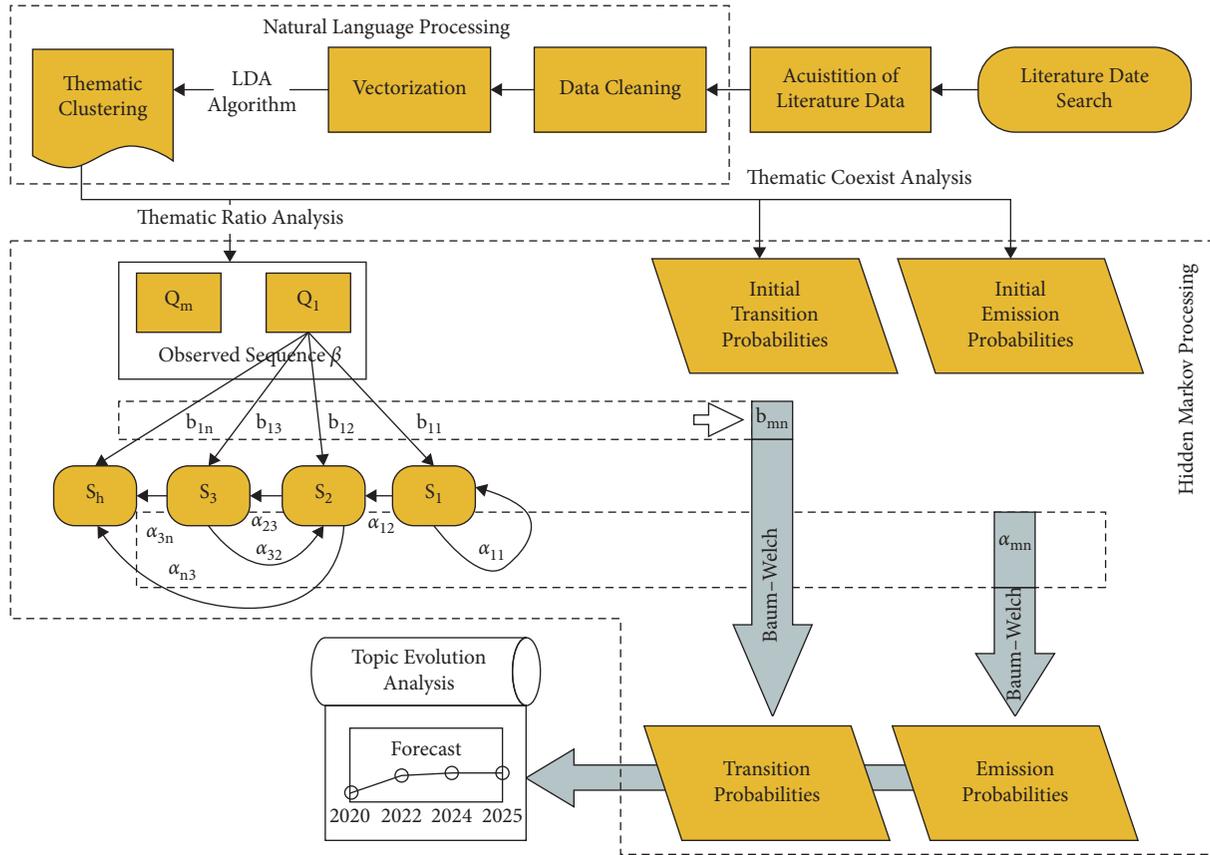


FIGURE 1: Logic structure flowchart.

Themes are abstract concepts, and the number of themes in the corpus can be quantified by dividing them into different granularities. The number of themes in the LDA model should be specified in advance. A larger corpus corresponds to a greater number of themes, and such number dynamically changes across different time windows. This article uses perplexity to determine the optimal number of themes [43]. Perplexity gradually decreases along with an increasing number of themes. A lower perplexity corresponds to a stronger generalization ability and better performance of the model.

By calculating the perplexity of each theme, the optimal number of themes in the LDA model employed in this work is 29. Experts in the field of urban expansion have read the sample of thematic classification literature and observed a relatively high accuracy (with a classification error rate of less than 3%). The boundaries between the themes are clear, and the division effect is ideal. For ease of reference, these themes are named based on keywords (Table 1).

3.2. Identification of Key Themes in the Urban Expansion Literature. The confusion matrix in the HMM indicates the possibility of transforming a hidden state into an observable state. This probability in turn can measure the threshold barriers for the transition of 29 themes in urban expansion research and can characterize the direction and extent of thematic evolution. The dark (light) squares in the confusion

matrix heat map represent those themes that are easy (difficult) to transfer in the innovation process (Figure 2). Most themes in the field of urban expansion show limited movement, and the thematic evolution is relatively stable. Varying degrees of transfer possibilities are also observed among different themes. To highlight the transfer relationship among these themes, this paper draws a confusion relationship network diagram (Figure 3) where the direction of arrows indicates the direction of thematic transfer.

Figure 3 shows that certain themes, including themes 26 (temperature), 15 (urban agglomeration), and 11 (economic development), have a high proportion of transfer inflow and a small proportion of transfer outflow. These themes are identified as core themes in the field of urban expansion. Therefore, transfer inflow and outflow are important manifestations of the criticality of a theme. To measure such criticality, this paper uses the PageRank link analysis algorithm, which obtains the critical evaluation of each node based on PageRank scores. A higher score corresponds to a higher criticality of a theme. The results are shown in Table 2.

Table 2 shows the key themes in the field of urban expansion, including themes 26 (temperature), 15 (urban agglomeration), 11 (economic development), 13 (housing development policy), 17 (surface change), and 9 (population density), of which temperature is the most critical. One obvious feature of urban expansion is the continuous increase in the area and density of various buildings in urban construction, which leads to the transformation of many

TABLE 1: Keywords and theme naming.

No.	KW1	KW2	KW3	KW4	KW5	ID
1	expansion	urban	china	urbanization	city	urban expansion
2	growth	urban	spatial	pattern	model	spatial pattern
3	city	growth	urban	infrastructure	state	urban infrastructure
4	urban	plan	planning	growth	management	urban planning management
5	urban	scenario	future	growth	prediction	scenario prediction
6	land	change	agricultural	development	urban	agricultural land change
7	sprawl	urban	measure	study	index	urban sprawl
8	urban	study	remote	city	datum	remote
9	population	city	density	large	high	population density
10	metropolitan	region	urban	period	paper	metropolitan region
11	development	spatial	urban	pattern	economic	economic development
12	factor	urban	distance	variable	datum	distance variable
13	policy	development	housing	effect	local	housing development policy
14	public	social	cost	population	economic	social cost
15	urban	level	regional	agglomeration	beijing	urban agglomeration
16	space	development	green	indicator	paper	green space
17	datum	surface	image	change	map	surface change
18	urban	dynamic	analysis	coastal	result	coastal urban
19	transportation	effect	transport	emission	center	transportation emission
20	model	simulation	simulate	cellular	automaton	cellular automata model
21	change	increase	climate	effect	result	climate change
22	associate	health	form	location	physical	physical health
23	forest	increase	zone	urbanization	loss	forest loss
24	landscape	pattern	metric	soil	patch	landscape pattern
25	risk	flood	level	present	environment	flood risk
26	increase	effect	temperature	decrease	degree	temperature
27	urban	global	carbon	time	project	urban carbon
28	ecosystem	natural	service	loss	river	ecosystem service
29	water	impact	quality	environmental	concentration	water quality

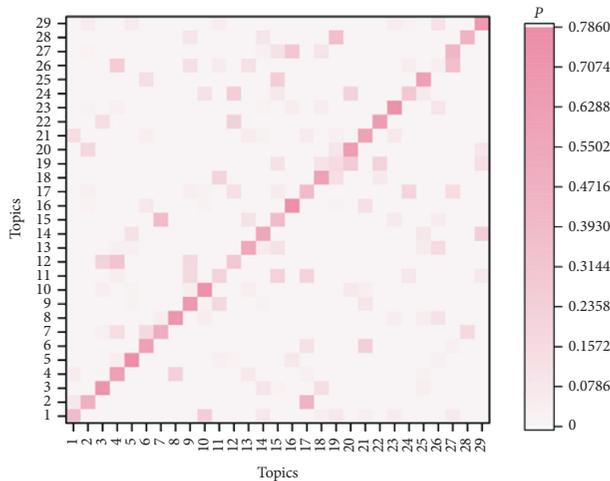


FIGURE 2: Confusion matrix estimated by HMM.

natural surfaces into impervious surfaces. The changes in the type and spatial structure of land cover affect the storage and transmission of surface temperature, thereby generating urban heat island effects [44]. Using remote sensing technology in analyzing surface thermal infrared information makes the result of urban spatial temperature distribution more accurate than the traditional calculations based on surface meteorological data. Therefore, such information provides a reliable basis for quantitatively studying the

spatial distribution of urban thermal environments [45]. Studies on the relationship between urban expansion and surface radiant temperature based on remote sensing technology are of great significance for improving urban thermal environments.

Scholars have also investigated those factors that drive urban expansion and find that economic development (theme 11), population density (theme 9), and housing development (theme 13) have important effects on urban expansion [46, 47]. Urban expansion and economic development conform to the Kuznets curve. During the initial stage of urbanization, economic development requires the development of a large amount of construction land and infrastructure land, thereby resulting in the outward expansion of cities. However, with the adjustment of the industrial structure, the improvement of infrastructure, and the increasing intensiveness of land use, the rate of urban expansion will decline [48]. Moreover, urban land is the main place that supports human life, work, and study. An increase in the urban population will inevitably increase the pressures on housing, transportation, and public facilities. Therefore, the demand of the urban population for space will generate momentum for urban expansion. For example, by studying the law of urban expansion and population growth in the metropolitan regions of the USA, Marshall [49] found that the average land area needed to support a new urban population is twice larger than the per capita land area of the existing city. Moreover, due to the agglomeration economy

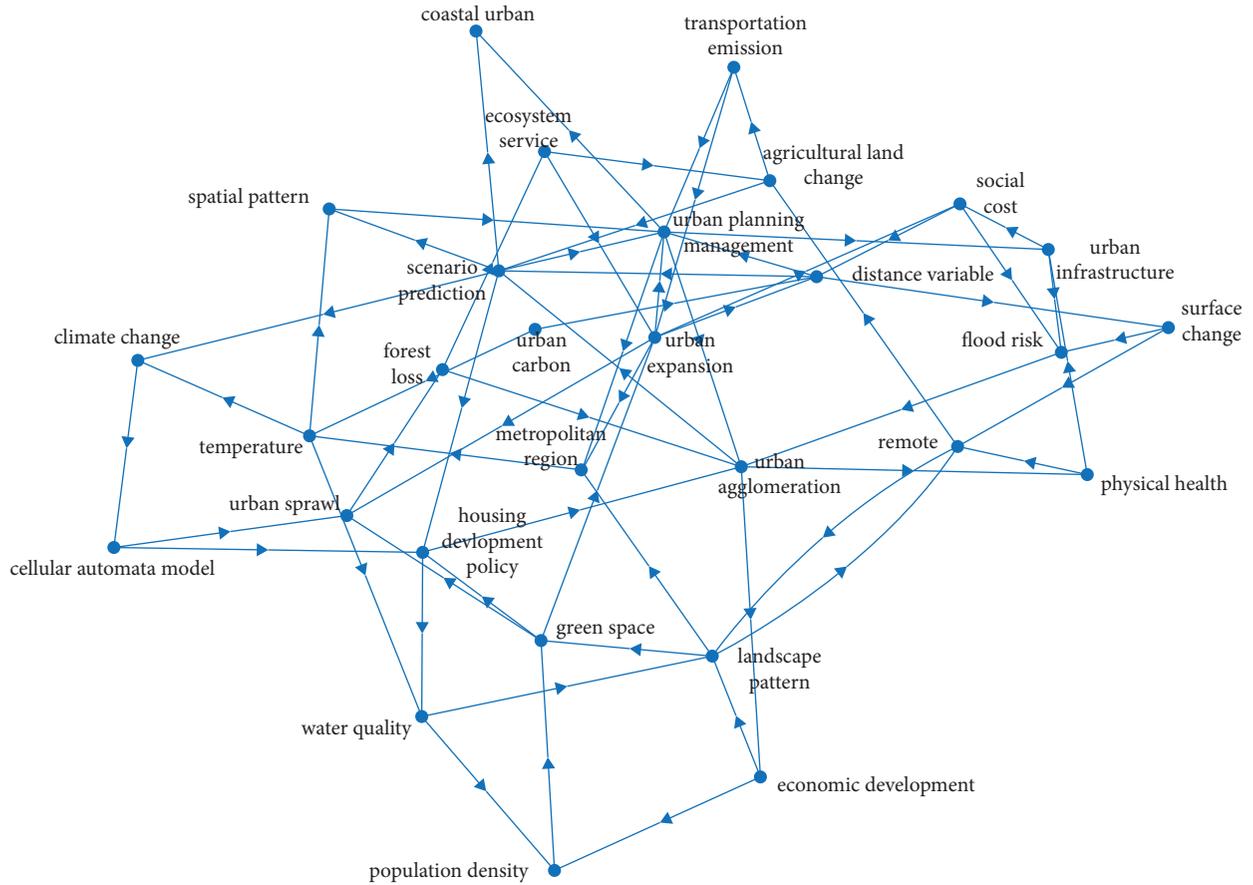


FIGURE 3: Confusion relationship network diagram among themes in the field of urban expansion.

TABLE 2: Critical scores of each theme.

No.	ID	Critical score	No.	ID	Critical score
26	temperature	1.0000	18	coastal urban	0.3262
15	urban agglomeration	0.8994	16	green space	0.2226
11	economic development	0.7654	2	spatial pattern	0.2154
13	housing development policy	0.6528	24	landscape pattern	0.2012
17	surface change	0.6197	3	urban infrastructure	0.1993
9	population density	0.6014	6	agricultural land change	0.1835
8	remote	0.5052	20	cellular automata model	0.1624
4	urban planning management	0.5035	28	ecosystem service	0.1449
10	metropolitan region	0.5025	29	water quality	0.1182
27	urban carbon	0.4330	22	physical health	0.0796
7	urban sprawl	0.4210	5	scenario prediction	0.0386
1	urban expansion	0.4025	14	social cost	0.0000
21	climate change	0.3833	23	forest loss	0.0000
19	transportation emission	0.3778	25	flood risk	0.0000
12	distance variable	0.3565			

of sharing, matching, and learning in urban areas, enterprises and laborers are constantly attracted to these areas. However, the urban space is limited, and the constant gathering of the labor force has increased both housing

prices and living costs. People are also forced to settle further away from the city center and pay high commuting costs. When the costs of living and commuting are high enough, these laborers will move elsewhere due to the low net utility

of living in urban areas. In this case, the government invests in the conversion of land into urban transportation infrastructure [50]. By substituting commuting and housing costs [51], the negative impact of rising housing costs is weakened, thereby facilitating a continued urban expansion.

3.3. Identification of Common Themes in the Urban Expansion Literature.

Based on the abovementioned thematic distribution, the proportion of each theme in the documents of developed and developing countries after unitization is calculated, and the degree of commonness of these themes is measured using formulas (4) and (5). The above results are then used to plot the degree of commonness of each theme in a graph as shown in Figure 4. The red and blue bars indicate the proportion of relevant documents in developing and developed countries after unitization, respectively, whereas the folding line indicates the degree of commonness of themes. The common themes in the field of urban expansion include themes 16 (green space), 26 (temperature), 4 (urban planning management), 2 (spatial pattern), 18 (coastal urban), and 5 (scenario prediction). With the transformational improvement of research data and technical research methods over the last few years, the available methods for urban expansion research have further expanded to scenario prediction [52], 3S spatial analysis [53], spatial econometrics [54], cellular automata [55], and multiagent simulation [56]. Using these methods to explore the spatial-temporal pattern distribution of urban expansion and effectively describe, simulate, analyze, and predict the process of urban evolution can provide decision-making support for urban planning and management. In addition, urban expansion research in developed and developing countries has mainly focused on coastal urban areas [57, 58] because compared with other cities, coastal cities have unique geographical locations and resource advantages, and urban expansion is highly susceptible to economic development, land-use policies, and regional development policies. Some significant differences in future land-use change are also observed under different development strategies. To expand living and development spaces, coastal areas are reclaiming land from the sea to address the increasingly serious problem of scarcity of land resources [59]. Reclaiming land from the sea is a large-scale human process that greatly disturbs the geographic processes of coastal zones. On the one hand, such land reclamation can increase food supply, attract more investments, and provide a new development space for urban areas. On the other hand, this reclamation can also reduce the service functions of marine ecosystems, destroy the ecological security of bay landscapes, result in marine sedimentation, degrade the quality of marine environments and habitats, and reduce coastal biodiversity [60, 61]. Therefore, how to protect coastal zones during their development has become a research hotspot.

By combining the aforementioned key and commonness indices, a key and commonness bubble for themes in the urban expansion field can be drawn (see Figure 5). This bubble chart is divided into the following quadrants based on the mean values of key and commonness: high degrees of

key and commonness (first quadrant), high degree of key and low degree of commonness (second quadrant), low degrees of key and commonness (third quadrant), and low degree of key and high degree of commonness (fourth quadrant). The first quadrant has eight themes, namely, themes 26 (temperature), 15 (urban agglomeration), 11 (economic development), 17 (surface change), 4 (urban planning management), 7 (urban sprawl), 27 (urban carbon), and 19 (transportation emission), of which themes 7 (urban sprawl) and 4 (urban planning management) have more documents than the median. In other words, these themes have received much attention in urban expansion research and are considered key research directions in this field. The connotations of urban sprawl include the following: (1) urban sprawl is a unique way of urban growth that usually occurs when the land development rate exceeds the population growth rate; (2) urban sprawl is characterized by low density, fragmentation, unsustainability, single-form development, excessive reliance on motor vehicles, and massive consumption of agricultural and ecological lands [62]; and (3) urban sprawl has a series of negative effects on traffic flow, plant and animal habitat, the ornamental nature of natural landscapes, and water circulation mechanisms [63, 64]; An in-depth study of urban sprawl has resulted in the formulation of three main theories in the field of urban expansion, namely, compact city theory [65], smart growth theory [66], and new urbanism theory [67]. Urban sprawl control methods can also be divided into two categories. The first category includes the urban planning methods that are implemented by the government and have attributes of administrative orders, such as urban growth boundary, zoning, planned unit development, transfer of development rights, traditional neighborhood development, and transit-oriented development [12, 68, 69]. These measures are based on the best spatial structure and scale of urban areas and directly affect the development decisions of landowners and developers. The second category includes guided regulation measures that are based on market orientation, including land development, fuel, property, and split-rate taxes. These measures do not compulsorily regulate the behavioral choices of people and have indirect control over the urban sprawl. In curbing urban sprawl, the pure market mechanism has a very limited influence on the development of compact cities. Therefore, the government needs not only to formulate various urban sprawl control measures but also to ensure that the relevant policies match the legal and political environment while restraining rapid urban expansion [70, 71].

3.4. Forecast on Thematic Evolution in the Field of Urban Expansion.

This paper uses 2020 as the forecast base period and imports the confusion matrix parameters into the HMM module to obtain the hidden Markov forecast results for the evolution of themes in the urban expansion literature from 2020 to 2025 (Figure 6). The proportion of landscape patterns in the prediction results has rapidly increased from 3.09% to 5.14%. The natural landscape is an important environmental resource in the urban ecosystem that has

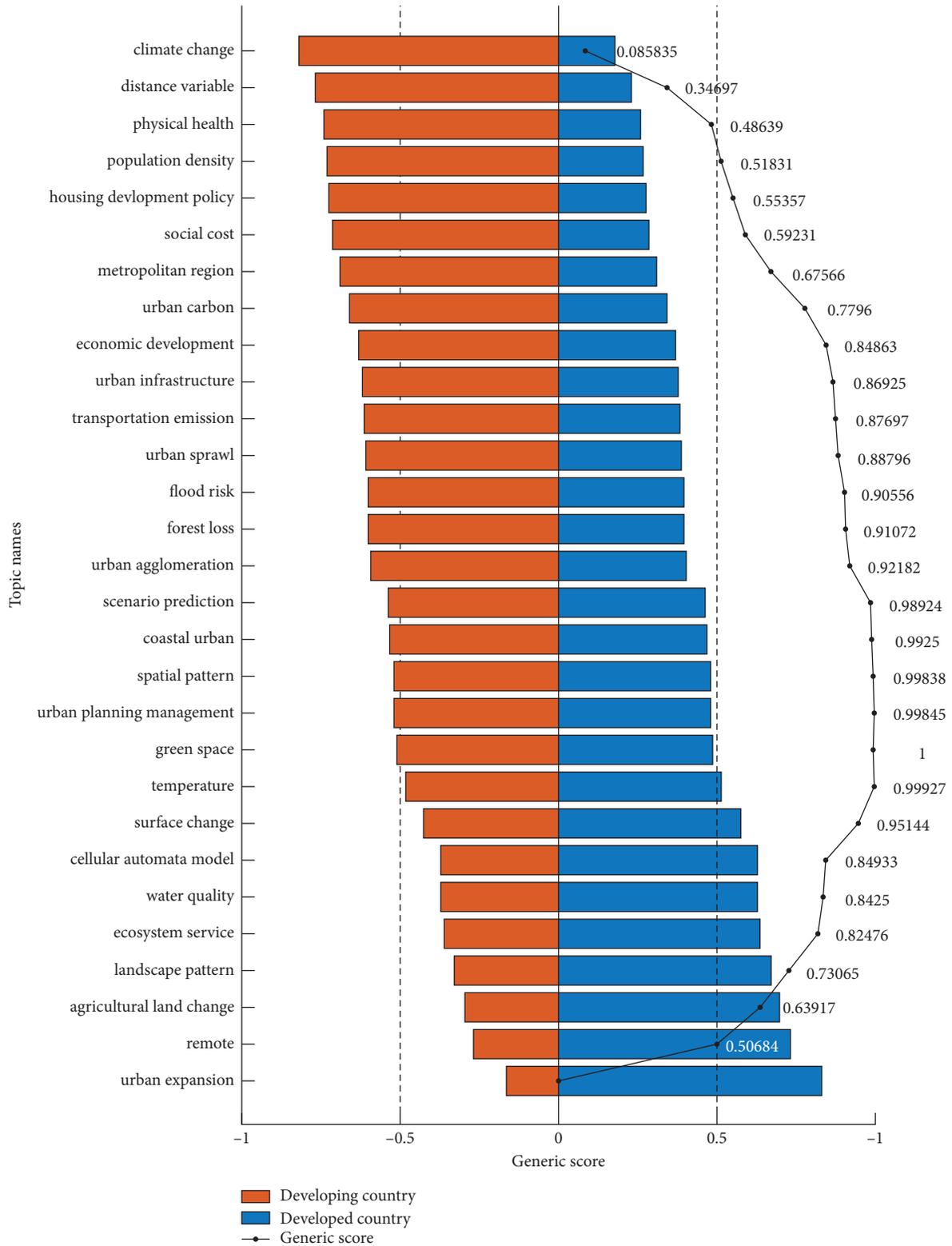


FIGURE 4: Degree of commonality of each theme.

significant ecological and social functions. Meanwhile, rapid urban expansion is a process in which man-made landscapes gradually erode, occupy, and transform natural landscapes, including forest land, cultivated land, lakes, and grassland,

under the influence of human disturbance. Therefore, rapid urban expansion not only reduces the natural landscape area but also results in the fragmentation of natural landscape patterns. A landscape tends to be a complex, heterogeneous,

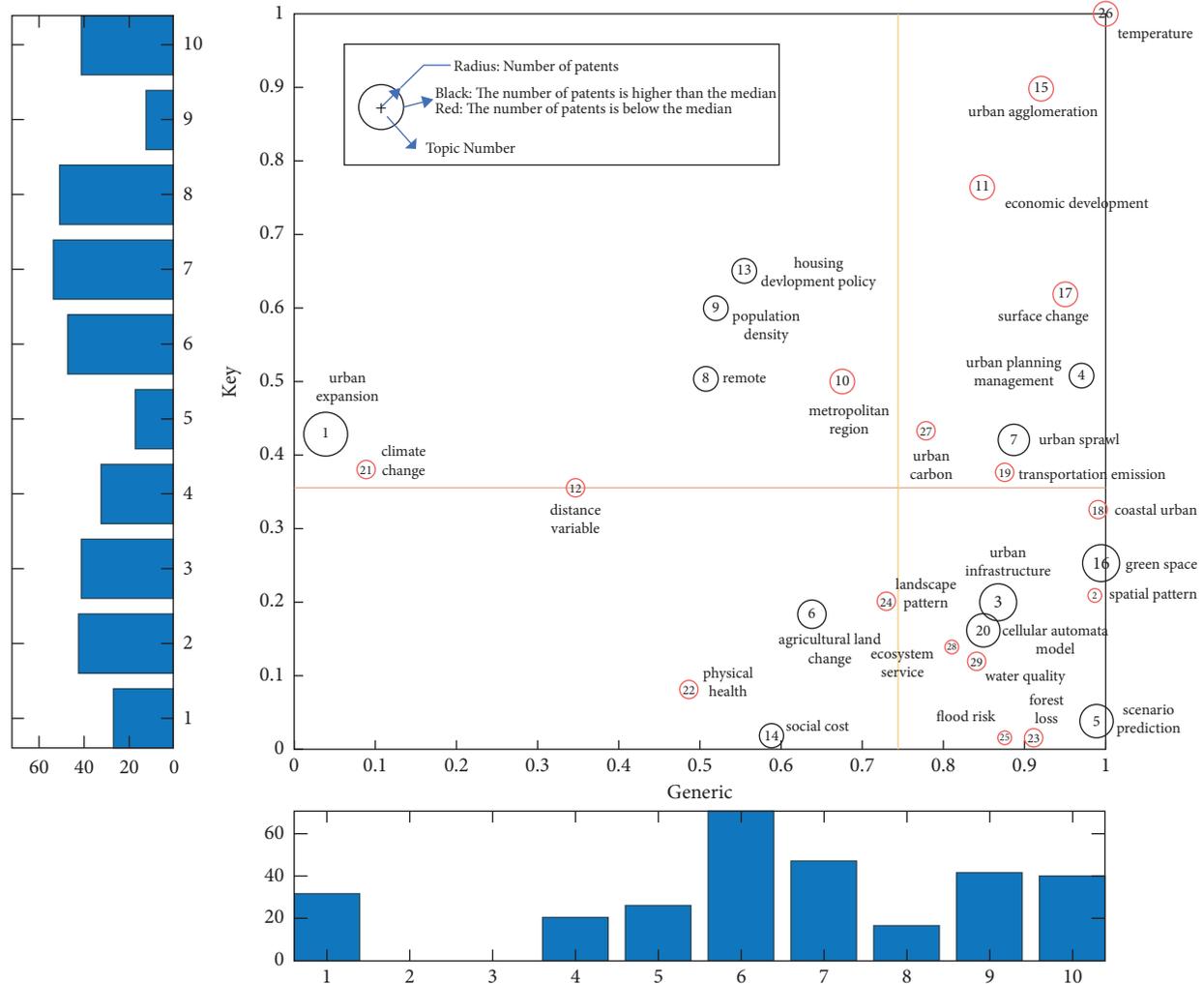


FIGURE 5: Thematic strategic diagram based on critical score and degree of commonness.

and discontinuous patch mosaic from a single, homogeneous, and continuous relative whole [72, 73]. The fragmentation of the urban landscape not only reduces the quality of the living environment of residents but also seriously endangers the urban ecosystem and urban sustainable development. Therefore, quantitatively identifying the urban landscape based on a remote sensing index (e.g., vegetation, impervious, and water indices) and exploring the responses of natural landscapes to urban expansion have become important ways of understanding the ecological effects of urban landscape evolution [74, 75] and provide valuable references for regional urban planning and ecological construction.

Agriculture land change remains the main direction in urban expansion research. The cultivated land occupied by urban expansion faces not only a decreasing quantity but also changes in its quality. Those areas that surround cities have excellent conditions, topography, water conservancy, and transportation. Urban expansion often encroaches on high-quality cultivated land [76] and affects cultivated land-use intensity in two ways. On the one hand, urban expansion easily results in the scarcity of cultivated land resources. The

intensity of cultivated land increases along with the continuous growth of population and demand for food. In addition, the rapid increase in the degree of intensification of agricultural production also promotes the application of chemical fertilizers and pesticides per unit area of cultivated land, thereby bringing agricultural nonpoint source pollution and ecological damage to the environment [77]. On the other hand, the increasingly open labor market promotes the transfer of agricultural labor and consequently reduces or abandons agricultural labor input. After the abandonment of cultivated land, the natural succession of farmland ecosystems destroys species habitats and degrades traditional agricultural landscapes with a high conservation value [78]. In addition, some species that live in the farmland system, especially birds and arthropods, will begin to disappear. The natural succession after abandonment also homogenizes the vegetation on abandoned land, thereby increasing the risk of fire and reducing biodiversity by promoting the growth of pyrophytes [79]. Therefore, studying the evolution of the spatial-temporal pattern of cultivated land occupied by urban expansion can provide technical support and a decision-making basis for handling the relationship between

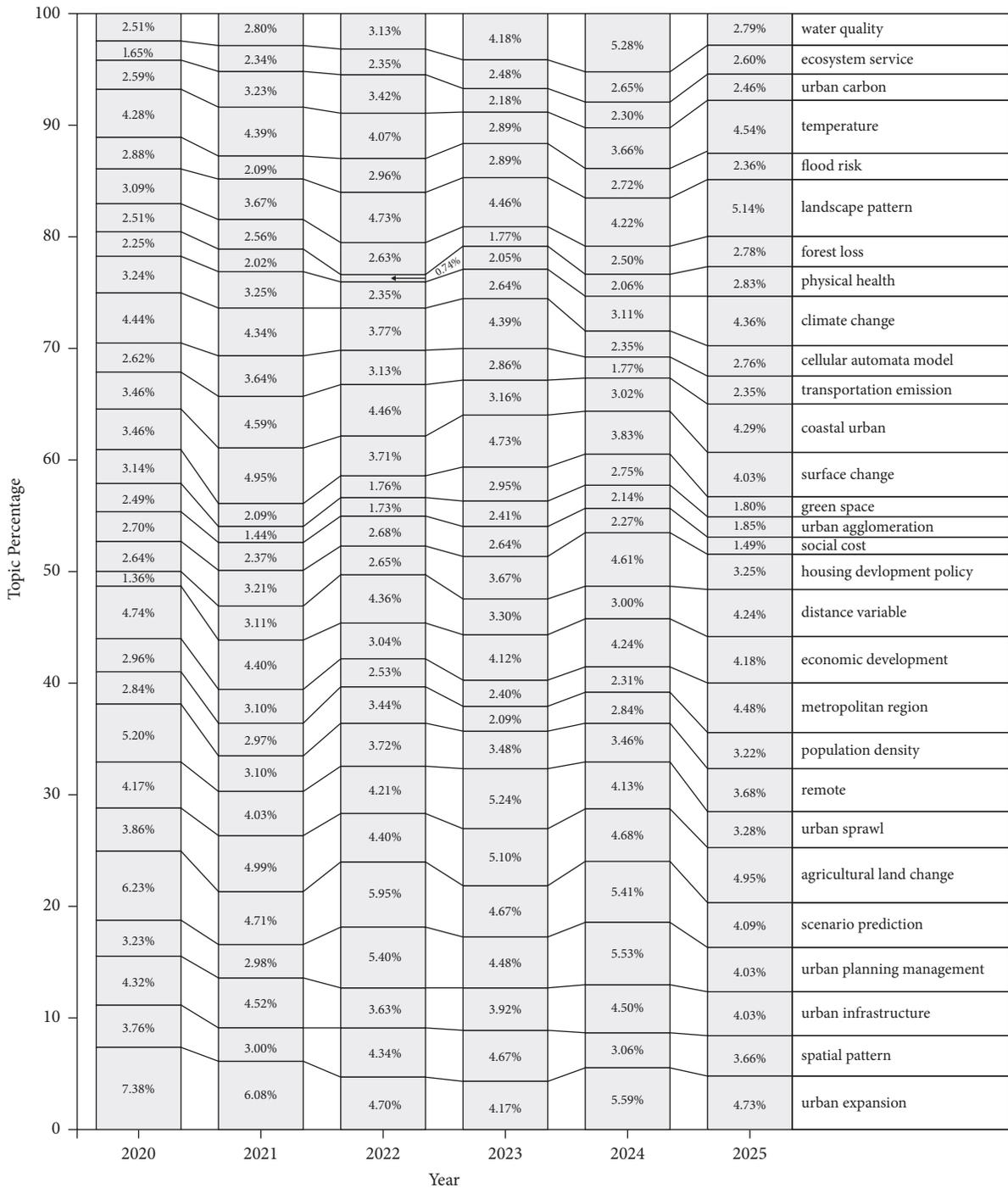


FIGURE 6: Forecasted thematic evolution trend from 2020 to 2025.

urban expansion and cultivated land protection and for scientifically coordinating urban development. Examining such evolution also has important practical significance in realizing sustainable land use.

4. Conclusion

Our study combines the LDA topic model with HMM to develop a new method for identifying key and common themes from the urban expansion literature. This method

overcomes the subjectivity of traditional methods. By applying text mining in a large number of studies in the field of urban expansion, an accurate thematic classification can be achieved, and the identified themes meet the empirical expectations. This study provides theoretical and operational support for identifying key and common themes in the field of bibliometrics.

To study the development trends in the field of urban expansion, this paper divides the scientific literature into 29 themes. By considering both the critical score and degree of

commonness, a total of eight important themes for developed and developing countries are identified, of which six themes (i.e., temperature, urban agglomeration, economic development, surface change, urban carbon, and transportation emission) have documents less than the median number. Future works should focus on these themes in light of the practical problems being faced in the urban expansion field.

The key and common theme identification methods proposed in this paper have good clustering effects, clear thematic boundaries, and accurate recognition results, all of which fully demonstrate their effectiveness and practicability. Future research may consider increasing the scope of the literature collection and including multisource heterogeneous documents to achieve a more comprehensive identification of key and common themes. However, this article also has shortcomings. The data only come from the core database of Web of Science, so the comprehensiveness of the data cannot be guaranteed. This may have a certain impact on the accuracy of the analysis results. Therefore, in future research, various databases should be combined to broaden the data sources in order to more accurately identify the key themes and common themes in the field of urban expansion.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this study.

Authors' Contributions

Yanwei Zhang and Xinhai Lu contributed equally to this work and share the first authorship.

Acknowledgments

The authors would like to thank Jingwen Liao for her contribution to the photo editing in the article. This research was supported by the National Natural Science Foundation of China (no. 71673096), the National 985 Project of Nontraditional Security at Huazhong University of Science and Technology, P.R. China, the Fundamental Research Funds for the Central Universities, HUST (no. 2021WKZDJC001) and the Social Science Foundation of Heilongjiang Province (no. 20JYC156).

References

- [1] H. Wang, Q. He, X. Liu, Y. Zhuang, and S. Hong, "Global urbanization research from 1991 to 2009: a systematic research review," *Landscape and Urban Planning*, vol. 104, no. 3-4, pp. 299-309, 2012.
- [2] P. Zhao, "Too complex to be managed? New trends in peri-urbanisation and its planning in Beijing," *Cities*, vol. 30, pp. 68-76, 2013.
- [3] B. Anasuya, D. Swain, and V. Vinoj, "Rapid urbanization and associated impacts on land surface temperature changes over Bhubaneswar Urban District, India," *Environmental Monitoring and Assessment*, vol. 790, 2019.
- [4] Y. Anker, V. Mirlas, A. Gimburg et al., "Effect of rapid urbanization on Mediterranean karstic mountainous drainage basins," *Sustainable Cities and Society*, vol. 51, Article ID 101704, 2019.
- [5] S. Angel, J. Parent, D. L. Civco, A. Blei, and D. Potere, "The dimensions of global urban expansion: estimates and projections for all countries," *Progress in Planning*, vol. 75, pp. 53-107, 2011.
- [6] S. M. Richter, "Revisiting urban expansion in the continental United States," *Landscape and Urban Planning*, vol. 204, Article ID 103911, 2020.
- [7] H. Wang, B. Zhang, Y. Liu et al., "Urban expansion patterns and their driving forces based on the center of gravity-GTWR model: a case study of the Beijing-Tianjin-Hebei urban agglomeration," *Journal of Geographical Sciences*, vol. 30, no. 2, pp. 297-318, 2020.
- [8] Y. Tu, B. Chen, and L. Yu, "How does urban expansion interact with cropland loss? a comparison of 14 Chinese cities from 1980 to 2015," *Landscape Ecology*, vol. 36, 2020.
- [9] L. Tang, X. Ke, Y. Chen et al., "Which impacts more seriously on natural habitat loss and degradation? Cropland expansion or urban expansion," *Land Degradation and Development*, vol. 32, 2020.
- [10] W. H. Lee, "How to identify emerging research fields using scientometrics: an example in the field of Information Security," *Scientometrics*, vol. 76, no. 3, pp. 503-525, 2008.
- [11] D. Huang, J. Huang, and T. Liu, "Delimiting urban growth boundaries using the CLUE-S model with village administrative boundaries," *Land Use Policy*, vol. 82, pp. 422-435, 2019.
- [12] S. Chakraborti, D. N. Das, B. Mondal, H. Shafizadeh-Moghadam, and Y. Feng, "A neural network and landscape metrics to propose a flexible urban growth boundary: a case study," *Ecological Indicators*, vol. 93, pp. 952-965, 2018.
- [13] Y. Feng, J. Wang, X. Tong et al., "Urban expansion simulation and scenario prediction using cellular automata: comparison between individual and multiple influencing factors," *Environmental Monitoring and Assessment*, vol. 191, p. 291, 2019.
- [14] A. Colsaet, Y. Laurans, and H. Levrel, "What drives land take and urban land expansion? A systematic review," *Land Use Policy*, vol. 79, pp. 339-349, 2018.
- [15] A. Awotwi, G. K. Anornu, J. A. Quaye-Ballard, and T. Annor, "Monitoring land use and land cover changes due to extensive gold mining, urban expansion, and agriculture in the Pra River Basin of Ghana," *Land Degradation & Development*, vol. 29, pp. 3331-3343, 2018.
- [16] C. Zeng, M. Zhang, J. Cui, and S. He, "Monitoring and modeling urban expansion-A spatially explicit and multi-scale perspective," *Cities*, vol. 43, pp. 92-103, 2015.
- [17] J. Yao, A. T. Murray, J. Wang, and X. Zhang, "Evaluation and development of sustainable urban land use plans through spatial optimization," *Transactions in GIS*, vol. 23, pp. 705-725, 2019.
- [18] Z. Zhang, F. Liu, X. Zhao et al., "Urban expansion in China based on remote sensing technology: a review," *Chinese Geographical Science*, vol. 28, no. 5, pp. 727-743, 2018.

- [19] A. Abu Hatab, M. E. R. Cavinato, A. Lindemer, and C.-J. Lagerkvist, "Urban sprawl, food security and agricultural systems in developing countries: a systematic review of the literature," *Cities*, vol. 94, pp. 129–142, 2019.
- [20] V. Saini and R. K. Tiwari, "A systematic review of urban sprawl studies in India: a geospatial data perspective," *Arabian Journal of Geosciences*, vol. 13, pp. 1–21, 2003.
- [21] H. Xie, Y. Zhang, and K. Duan, "Evolutionary overview of urban expansion based on bibliometric analysis in Web of Science from 1990 to 2019," *Habitat International*, vol. 95, Article ID 102100, 2020.
- [22] M. Lacey-Barnacle, R. Robison, and C. Foulds, "Energy justice in the developing world: a review of theoretical frameworks, key research themes and policy implications," *Energy for Sustainable Development*, vol. 55, no. 138, 2020.
- [23] G. W. Ryan and H. R. Bernard, "Techniques to identify themes," *Field Methods*, vol. 15, no. 1, pp. 85–109, 2003.
- [24] D. M. Blei and J. D. Lafferty, "A correlated topic model of science," *Annals of Applied Statistics*, vol. 1, pp. 17–35, 2007.
- [25] N. Shibata, Y. Kajikawa, Y. Takeda, I. Sakata, and K. Matsushima, "Detecting emerging research fronts in regenerative medicine by the citation network analysis of scientific publications," *Technological Forecasting and Social Change*, vol. 78, no. 2, pp. 274–282, 2011.
- [26] H. Small, "Tracking and predicting growth areas in science," *Scientometrics*, vol. 68, no. 3, pp. 595–610, 2006.
- [27] D. Lee and H. Choe, "Estimating the impacts of urban expansion on landscape ecology: forestland perspective in the greater Seoul metropolitan area," *Journal of Urban Planning and Development*, vol. 137, no. 4, pp. 425–437, 2011.
- [28] Z. Zhang, W. Sun, and Y. Yu, "Research on the development of marine regional E-commerce based on the analysis of cloud computing and grey prediction method," *Journal of Coastal Research*, vol. 115, pp. 333–337, 2020.
- [29] Y. P. Tsang, W. C. Wong, G. Q. Huang, C. H. Wu, Y. H. Kuo, and K. L. Choy, "A fuzzy-based product life cycle prediction for sustainable development in the electric vehicle industry," *Energies*, vol. 13, no. 15, 3918 pages, 2020.
- [30] J. Wang, G. Nie, and C. Xue, "Landslide displacement prediction based on time series analysis and data assimilation with hydrological factors," *Arabian Journal of Geosciences*, vol. 13, no. 460, 2020.
- [31] F. Almomani, "Prediction the performance of multistage moving bed biological process using artificial neural network (ANN)," *The Science of the Total Environment*, vol. 744, Article ID 140854, 2020.
- [32] L. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *The Annals of Mathematical Statistics*, vol. 37, pp. 554–1563, 1966.
- [33] C. Zeng, Y. Liu, Y. Liu, and L. Qiu, "Urban sprawl and related problems: bibliometric analysis and refined analysis from 1991 to 2011," *Chinese Geographical Science*, vol. 24, no. 2, pp. 245–257, 2014.
- [34] D. Blei, A. Ng, and M. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
- [35] H. Jelodar, Y. Wang, C. Yuan et al., "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15169–15211, 2018.
- [36] R. Welch, "Hidden Markov models and the Baum–Welch algorithm," *IEEE Information Theory Society Newsletter*, vol. 53, pp. 194–211, 2003.
- [37] K. M. Frahm and D. L. Shepelyansky, "Ising-PageRank model of opinion formation on social networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 526, Article ID 121069, 2019.
- [38] F. A. Massucci and D. Docampo, "Measuring the academic reputation through citation networks via PageRank," *Journal of Informetrics*, vol. 13, no. 1, pp. 185–201, 2019.
- [39] P. Holme and J. Saramäki, "Temporal networks," *Physics Reports*, vol. 519, no. 3, pp. 97–125, 2012.
- [40] G. Abramo, C. A. D'Angelo, and A. Soldatenkova, "An investigation on the skewness patterns and fractal nature of research productivity distributions at field and discipline level," *Journal of Informetrics*, vol. 11, no. 1, pp. 324–335, 2017.
- [41] L. Bornmann and L. Leydesdorff, "Skewness of citation impact data and covariates of citation distributions: a large-scale empirical analysis based on Web of Science data," *Journal of Informetrics*, vol. 11, no. 1, pp. 164–175, 2017.
- [42] J. Ruiz-Castillo and R. Costas, "The skewness of scientific productivity," *Journal of Informetrics*, vol. 8, no. 4, pp. 917–934, 2014.
- [43] D. M. Blei, "Probabilistic topic models," *Communications of the ACM*, vol. 55, no. 4, pp. 77–84, 2012.
- [44] I. Buo, V. Sagris, I. Burdun, and E. Uuemaa, "Estimating the expansion of urban areas and urban heat islands (UHI) in Ghana: a case study," *Natural Hazards*, vol. 105, pp. 1299–1321, 2021.
- [45] J. Du, X. Xiang, B. Zhao, and H. Zhou, "Impact of urban expansion on land surface temperature in Fuzhou, China using Landsat imagery," *Sustainable cities and society*, vol. 61, 102346 pages, 2020.
- [46] W. Admasu, S. Van Passel, A. Minale, E. Tsegaye, H. Azadi, and J. Nyssen, "Take out the farmer: an economic assessment of land expropriation for urban expansion in Bahir Dar, Northwest Ethiopia," *Land Use Policy*, vol. 87, Article ID 104038, 2019.
- [47] M. Jia, Y. Liu, S. Lieske, and T. Chen, "Public policy change and its impact on urban expansion: an evaluation of 265 cities in China," *Land Use Policy*, vol. 97, Article ID 104754, 2020.
- [48] Y. Zhang and H. Xie, "Interactive relationship among urban expansion, economic development, and population growth since the reform and opening up in China: an analysis based on a vector error correction model," *Land*, vol. 8, 153 pages, 2019.
- [49] J. Marshall, "Urban land area and population growth: a new scaling relationship for metropolitan expansion," *Urban Studies*, vol. 44, 2007.
- [50] P. Zhao, B. Lü, and G. de Roo, "Urban expansion and transportation: the impact of urban form on commuting patterns on the city fringe of Beijing," *Environment and Planning: Economy and Space*, vol. 42, no. 10, pp. 2467–2486, 2010.
- [51] R. J. Arnott and J. E. Stiglitz, "Aggregate land rents, expenditure on public goods, and optimal city size," *Quarterly Journal of Economics*, vol. 93, no. 4, pp. 471–500, 1979.
- [52] Z. Lei, Y. Feng, X. Tong, S. Liu, C. Gao, and S. Chen, "A spatial error-based cellular automata approach to reproducing and projecting dynamic urban expansion," *Geocarto International*, vol. 16, pp. 1–21, 2020.
- [53] H. Lamphar, "Spatio-temporal association of light pollution and urban sprawl using remote sensing imagery and GIS: a simple method based in Otsu's algorithm," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 253, Article ID 107068, 2020.

- [54] S. Mathur, "Impact of an urban growth boundary across the entire house price spectrum: the two-stage quantile spatial regression approach," *Land Use Policy*, vol. 80, pp. 88–94, 2019.
- [55] S. Chen, Y. Feng, Z. Ye et al., "A cellular automata approach of urban sprawl simulation with Bayesian spatially-varying transformation rules," *GIScience & Remote Sensing*, vol. 57, 2020.
- [56] A. Mustafa, M. Cools, I. Saadi, and J. Teller, "Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (HUEM)," *Land Use Policy*, vol. 69, pp. 529–540, 2017.
- [57] Y. Deng, W. Qi, Bo. Fu, and K. Wang, "Geographical transformations of urban sprawl: exploring the spatial heterogeneity across cities in China 1992–2015," *Cities*, vol. 105, Article ID 102415, 2019.
- [58] Y. Kim and G. Newman, "Climate change preparedness: comparing future urban growth and flood risk in Amsterdam and Houston," *Sustainability*, vol. 11, 1048 pages, 2019.
- [59] Y. Li, Y. Shi, X. Zhu, H. Cao, and T. Yu, "Coastal wetland loss and environmental change due to rapid urban expansion in Lianyungang, Jiangsu, China," *Regional Environmental Change*, vol. 14, no. 3, pp. 1175–1188, 2014.
- [60] D. S. Hammond, V. Gond, C. Baider, F. B. V. Florens, S. Persand, and S. G. W. Laurance, "Threats to environmentally sensitive areas from peri-urban expansion in Mauritius," *Environmental Conservation*, vol. 42, no. 3, pp. 256–267, 2015.
- [61] T. M. Coxon, B. K. Odhiambo, and L. C. Giancarlo, "The impact of urban expansion and agricultural legacies on trace metal accumulation in fluvial and lacustrine sediments of the lower Chesapeake Bay basin, USA," *The Science of the Total Environment*, vol. 568, pp. 402–414, 2016.
- [62] F. Bidandi and J. Williams, "Understanding urban land, politics, and planning: a critical appraisal of Kampala's urban sprawl," *Cities*, vol. 106, Article ID 102858, 2020.
- [63] S. Hamidi and R. Ewing, "Is sprawl affordable for Americans?" *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2500, no. 1, pp. 75–79, 2015.
- [64] S. M. Mccoshum and M. A. Geber, "Land conversion for solar facilities and urban sprawl in southwest deserts causes different amounts of habitat loss for *Ashmeadiella* bees," *Journal of the Kansas Entomological Society*, vol. 92, no. 2, pp. 468–478, 2020.
- [65] M. Jun, "The effects of polycentric evolution on commute times in a polycentric compact city: a case of the Seoul metropolitan area," *Cities*, vol. 98, Article ID 102587, 2020.
- [66] Å. Gren, J. Colding, M. Berghauser-Pont, and L. Marcus, "How smart is smart growth? Examining the environmental validation behind city compaction," *Ambio*, vol. 48, no. 6, pp. 580–589, 2018.
- [67] A. Stanislav and J. Chin, "Evaluating livability and perceived values of sustainable neighborhood design: new Urbanism and original urban suburbs," *Sustainable Cities and Society*, vol. 47, Article ID 101517, 2019.
- [68] A. H. Whittemore, "The new communalism," *Journal of Planning History*, vol. 14, no. 3, pp. 244–259, 2015.
- [69] W. Lang, E. Hui, T. Chen, and X. Li, "Understanding livable dense urban form for social activities in transit-oriented development through human-scale measurements," *Habitat International*, vol. 104, Article ID 102238, 2020.
- [70] B. Fernandez Milan and F. Creutzig, "Municipal policies accelerated urban sprawl and public debts in Spain," *Land Use Policy*, vol. 54, pp. 103–115, 2016.
- [71] D. Coq-Huelva and R. Asián-Chaves, "Urban Sprawl and Sustainable Urban Policies. A review of the cases of Lima, Mexico City and Santiago de Chile," *Sustainability*, vol. 11, 5835 pages, 2019.
- [72] P. Hien, N. Men, P. Tan, and M. Hangartner, "Impact of urban expansion on the air pollution landscape: A case study of Hanoi, Vietnam," *Science of The Total Environment*, vol. 702, Article ID 134635, 2020.
- [73] B. Rimal, H. Keshtkar, R. Sharma, N. Stork, S. Rijal, and R. Kunwar, "Simulating urban expansion in a rapidly changing landscape in eastern Tarai, Nepal," *Environmental Monitoring and Assessment*, vol. 191, p. 255, 2019.
- [74] Y. Al-husban, "Urban expansion and shrinkage of vegetation cover in Al-Balqa Governorate, the Hashemite Kingdom of Jordan," *Environmental Earth Sciences*, vol. 78, 620 pages, 2019.
- [75] S. Bonilla-Bedoya, A. Mora, A. Vaca, A. Estrella, and M. Herrera, "Modelling the relationship between urban expansion processes and urban forest characteristics: an application to the Metropolitan District of Quito," *Computers, Environment and Urban Systems*, vol. 79, Article ID 101420, 2019.
- [76] B. Güneralp, M. Reba, B. U. Hales, E. A. Wentz, and K. C. Seto, "Trends in urban land expansion, density, and land transitions from 1970 to 2010: a global synthesis," *Environmental Research Letters*, vol. 15, no. 4, Article ID 044015, 2020.
- [77] H. Xie, Y. Zhang, and Y. Choi, "Measuring the cultivated land use efficiency of the main grain-producing areas in china under the constraints of carbon emissions and agricultural nonpoint source pollution," *Sustainability*, vol. 10, Article ID 1932, 2018.
- [78] J. Fischer, T. Hartel, and T. Kuemmerle, "Conservation policy in traditional farming landscapes," *Conservation Letters*, vol. 5, no. 3, pp. 167–175, 2012.
- [79] C. Vega-Garcia and E. Chuvieco, "Applying local measures of spatial heterogeneity to Landsat-TM images for predicting wildfire occurrence in Mediterranean landscapes," *Landscape Ecology*, vol. 21, pp. 595–605, 2006.