VISUALIZATION ANALYSIS OF FUTURE CITY RESEARCH BASED ON CITESPACE

Jing Han [1]

[1] Tongji University

1. Introduction

Humans' exploration of "future cities" has never stopped. The experimental history of the future city is a development history in which human beings gradually meet their own needs and dreams through technological tools. With the great development of science and technology in recent decades, the economic structure and social organization of cities are undergoing unprecedented reorganization and transformation. In addition, the global COVID-19 pandemic has caused a lot of confusion and suffering. Thinking and answering about the future direction of cities, and what adaptive changes will be adopted in urban planning, is the key challenge that current urban researchers, planners and even all sectors of society are facing. It is also a lasting proposition that needs to be continuously deepened and improved in the future. The international academic community has always attached great importance to the research and practice of future cities. Comprehensively examining the outcomes of existing scientific papers on *future city* and clarifying its development context and evolution trend are necessary prerequisites for constructing future city theory and helping urban decision makers to formulate future urban development paths.

2. Methods and data

2.1 CiteSpace

Bibliometrics is a scientific quantitative analysis method that integrates mathematics, statistics and philology. CiteSpace is a bibliometric software that focus on exploring the potential knowledge contained in scientific carriers. It was developed by Professor Chaomei Chen of the Drexel University and his team. It can measure relevant literature through co-citation analysis theory, etc., and draw a series of visual maps, so as to analyze the research status and trends of a given field. This paper uses CiteSpace to quantitatively analyze the literature and information related to future city, in order to grasp the overall situation and key points of global future city research.

2.2 Data retrieval process and resulting dataset

Comprehensively comparing the data quality and the degree of matching with the CiteSpace tool, this study selected the Web of Science Core Collection (WoSCC), the world's largest and most comprehensive scientific publication database, as the data source. The following literature search strategies were applied on 30 May 2022 (Table 1):

Table 1. Literature search strategy.

Data sources	Science Citation Index Expanded (SCI-Expanded) Social Sciences Citation Index (SSCI) Arts & Humanities Citation Index (A&HCI)
Timespan	From 2000.01.01 to 2022.05.30
Retrieved content	Full records and cited references
Query	(TS = ("future* city*" OR "city* future*" OR "future* urban*" OR "urban* future*") AND LA=(English) AND DT = (Article OR Proceedings Paper OR Review)) AND PY = (2000-2022)

After deduplication, a total of 1933 valid records were obtained, including 1854 papers and 79 review articles.

3. Macro overview: disciplines and topics involved in future city

3.1 Category co-occurrence network

What scientific fields are involved in research on cities of the future? Web of Science assigns one or more subject category labels to each article it indexes, making it easy for readers to know what fields of research each article describes. When a document cites research results from different disciplines, it means that there is a cross-border combination relationship between these different scientific fields of knowledge, and this relationship must be recognized by the author. When enough authors agree on the same relationship, then the relationship can be considered to have stable and reliable significance in this research direction.

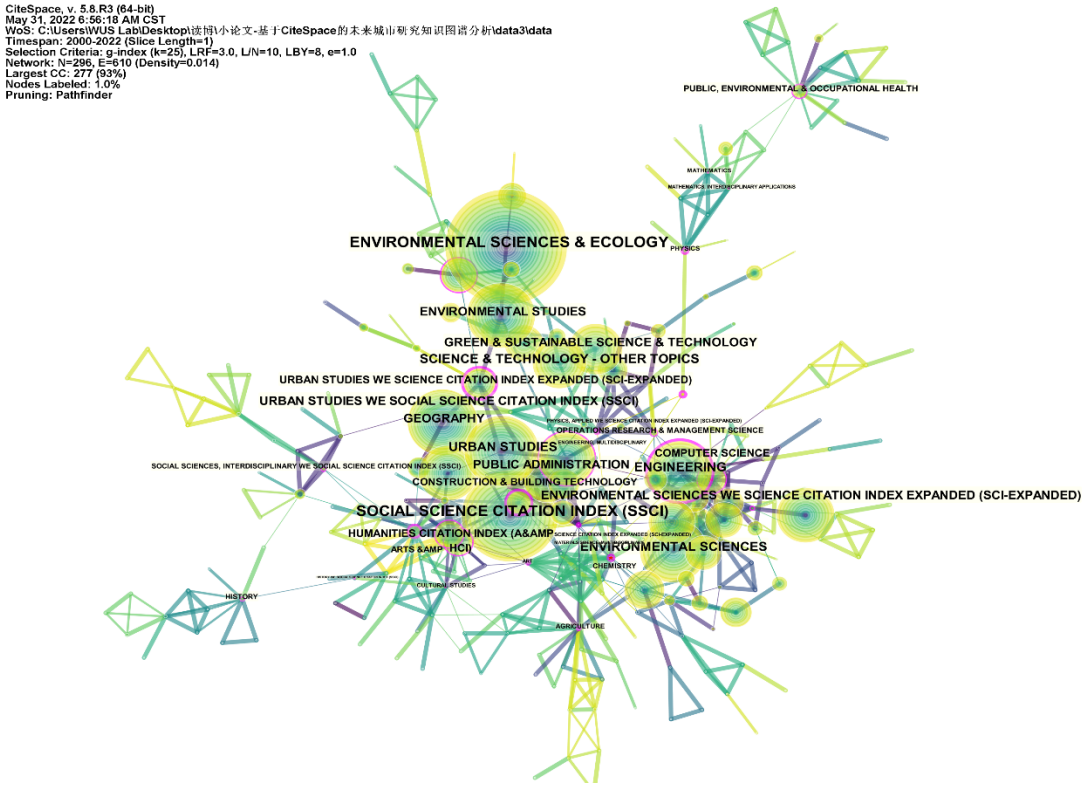

Figure 1. Disciplines involved in future city

Figure 1 shows a simplified category co-occurrence network by Pathfinder that preserves the most prominent connections. Each node is depicted as a series of tree rings, the darker the color, the earlier it was referenced. The size of the node indicates how many times the relevant scientific field has been cited. As we can see, the most common categories are Environmental Science & Ecology and Social Science, with the largest circles, followed by Urban Studies, Science Technology – Other Topics, Green and Sustainable Science & Technology, Geography, and Public Administration. Although Chemistry, Physics & Applied, Multidisciplinary Engineering, Agriculture, Art, Construction & Building Technology, Material Science, Culture Studies, Humanities, Computer Science, HCI, History, Mathematics, Operations Research & Management Science were cited much less often, they are all structurally significant nodes. Such nodes are shown as purple rings, the thickness of which indicates their degree of betweenness centrality, a measure related to the transformative potential of scientific contributions. Such nodes tend to bridge different stages of development in the scientific field. Engineering is an important subject area with both large scale and high centrality, and it is a major core position for future urban research.

3.2 Dual-map overlay

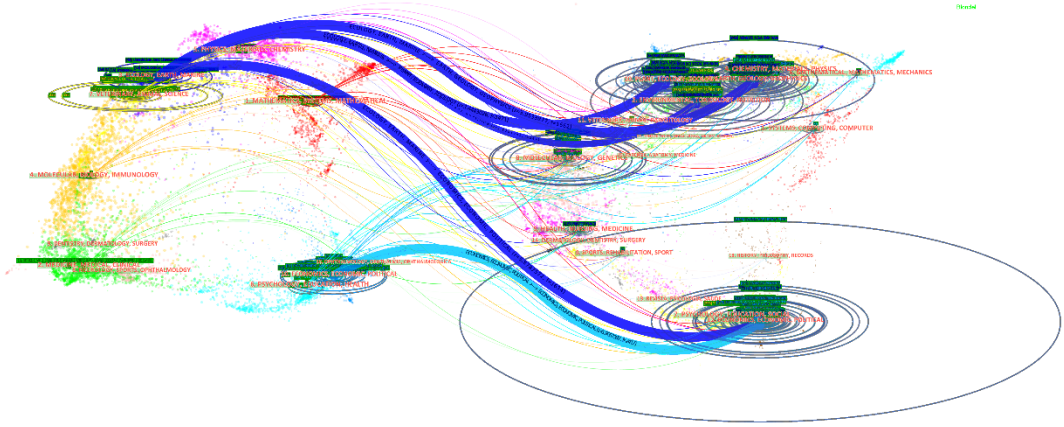

Figure 2. Dual-map overlays of both the citing and cited articles related to future city

Dual map overlays depict both citing and cited objects in one uninterrupted view, making it easy to see where citations are coming from and where they are pointing at, helping to study interdisciplinarity at both the source and the destination. On the basis of the global scientific journal map, the publication status and citation relationship of 1,933 reference articles collected in this study and their 49,591 citating paper were superimposed to form an analysis map of future urban research fields based on double-map overlay (Figure 3).

In this figure, the left side is the distribution of the journals where the citing documents are located, which represents the main discipline to which future city belongs, and the right side is the distribution of the journals corresponding to the cited documents, which represents which disciplines the future city mainly cites. The former can be regarded as the field application of future city, and the latter can be regarded as the research basis of future city. The vertical axis of the ellipse represents the number of publications, and the horizontal axis represents the number of authors.

It shows that scholars from many disciplines such as Mathematics & Systems, Medicine & Clinical, Molecular & Biology & Immunology, Physics & Materials & Chemistry, Psychology & Education & Health have published papers related to future cities (Figure 3). Veterinary & Animal Science has the highest number of publications. The adjacent Ecology & Earth & Marine disciplines are the second. The dark blue curve shows that the research output of the discipline of Ecology & Earth & Marine is mainly based on the four disciplines of Earth & Geology & Geophysics, Plant & Ecology & Zoology, Environment & Toxicology & Nutrition, and Economics & Political on the right side of the map. The cyan curve shows that research in the Economics & Political discipline has largely derived from the existing knowledge base of the discipline itself.

4. Micro understanding: the intellectual structure of future city

4.1 Co-cited reference network and clustering

The literature co-citation map can help people analyze the evolution of the inquiring subject through the key nodes, clusters and colors in the map, and the co-occurrence word map is more conducive to people to see clearly the hotspots and their evolution, especially with the burst term detecting functions. Figure 3 shows a hybrid network of co-cited references and burst terms on future city. Each cluster involves both citing and cited articles.

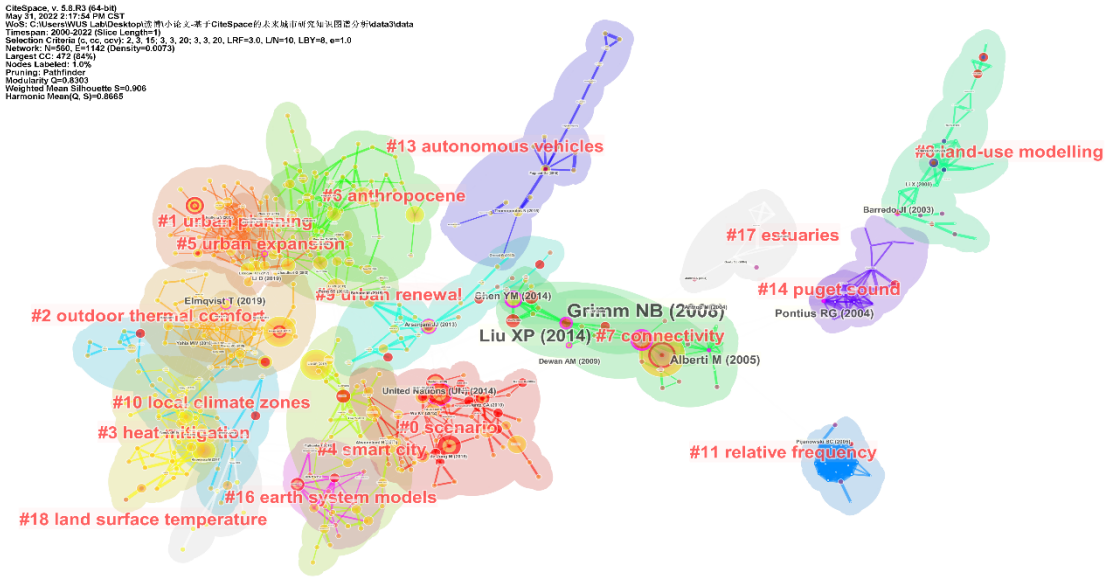

Figure 3. a hybrid network of co-cited references and burst terms from titles and abstracts. Clusters are labeled in red text. Articles are labeled in black. Burst terms are displayed in light red text. Red circles indicate articles with citation bursts, that is, rapid increases of citation counts.

The size of the node (and label) represents the frequency of the reference, and the larger the node, the higher the article's citied frequency. The most cited articles are often considered landmark articles because of their seminal contributions. Cluster #7 and #0 each has 2 articles in the top 7 landmark articles. The most cited article in our dataset is Seto KC (2012) with 59 citations, followed by Arsanjani JJ (2013) with 35 citations and Liu XP (2001) with 34 citations (Table 2).

If the rings are surrounded by purple circles, it indicates that the node occupies an important centrality position in the network - they are highly connected to other nodes such as hubs, or positioned between different groups of nodes. The articles with high betweenness centrality scores (Table 3) can be seen as landmark works in our broadly defined context of future city.

If some annual rings of nodes are filled in red, it indicates that these references have experienced sudden changes (like burst) during the research period. Burst has two properties: the intensity of the burst and the duration of the burst. Table 4 lists the six references with the strongest citing burst in

the entire dataset. The first article with a strong citation burst is from Cluster #7 on connectivity. Interestingly, two 1998 articles (both in cluster #8) are by the same author.

If a node excels in above two measures, then it has both structure centrality and citation burst, making it the most important pillar article in the entire map. It is essential in the structure of the entire knowledge system, and has strong heuristic, foundational, and heritage significance, as measured by the Sigma indicator (Table 5).

Table 6 lists the top 6 main clusters in the co-cited reference network by size, that is, the number of references in each cluster. Other small clusters are not as representative as the larger ones and are not listed. The quality of a cluster is also reflected in its silhouette score, which is a measurement of cluster homogeneity or consistency. Homogeneous clusters tend to have silhouette values close to 1, ensuring tight connections within the same cluster but loose connections between different clusters.

Table 2. Most cited references.

Citation counts	References	Cluster #
59	Seto KC, 2012, P NATL ACAD SCI USA, 109, 16083	7
35	Arsanjani JJ, 2013, INT J APPL EARTH OBS, 21, 265	2
34	Liu XP, 2017, LANDSCAPE URBAN PLAN, 168, 94	4
27	Seto KC, 2011, PLOS ONE, 6, 0	0
22	Grimm NB, 2008, SCIENCE, 319, 756	7
20	United Nations, 2018, WORLD URB PROSP 2018, 0, 0	1
20	Moghadam HS, 2013, APPL GEOGR, 40, 140	0

Table 3. Cited citations with the highest betweenness centrality.

Rank	Centrality	Reference	Cluster #
1	17	Pijanowski BC, 2005, INT J GEOGR INF SCI, 19, 197	11
2	16	Elmqvist T, 2019, NAT SUSTAIN, 2, 267	2
2	16	Bagan H, 2014, ENVIRON RES LETT, 9, 0	1
2	16	Al-Hathloul S, 2004, HABITAT INT, 28, 609	11
2	16	Acevedo MF, 2008, GEOFORUM, 39, 846	11
2	16	Alkheder S, 2005, 3 INT S REM SENS DAT,0,0	11

Table 4. References with the strongest citation bursts.

Citation bursts	References	Cluster #
11.66	Grimm NB, 2008, SCIENCE, 319, 756	7
7.97	Sante I, 2010, LANDSCAPE URBAN PLAN, 96, 108	0
7.61	Seto KC, 2011, PLOS ONE, 6, 0	0
7.47	Landis J, 1998, ENVIRON PLANN B, 25, 657	8
6.65	Landis J, 1998, ENVIRON PLANN B, 25, 795	8
6.61	Clarke KC, 1998, INT J GEOGR INF SCI, 12, 699	8

Table 5. Structurally and temporally significant references.

Sigma	Burst	Centrality	Citations	References	Cluster #
43.12	11.66	0.38	22	Grimm NB, 2008, SCIENCE, 319, 756	7
3.85	4.54	0.35	8	Liu XP, 2014, INT J GEOGR INF SCI, 28, 148	7
2.37	5.05	0.19	18	Alberti M, 2005, INT REGIONAL SCI REV, 28, 168	0
1.98	4.41	0.17	7	Pontius RG, 2004, ECOL MODEL, 179, 445	8
1.59	7.97	0.06	16	Chen YM, 2014, INT J GEOGR INF SCI, 28, 234	0
1.59	6.61	0.07	10	Elmqvist T, 2019, NAT SUSTAIN, 2, 267	8

Table 6. Major clusters of co-cited references.

Cluster ID	Size	Silhouette	Label (LLR)	Year Ave.
0	51	0.804	scenario	2012
1	48	0.82	urban planning	2016
2	43	0.853	outdoor thermal comfort	2016
3	39	0.867	heat mitigation	2015
4	36	0.857	smart city	2015
5	32	0.924	urban expansion	2015

4.2 Observation in time dimension

The timeline view can show the time span, peak interval, and contextual correlation of the development and evolution of each cluster. Figure 4 is clear that clusters #8 on land use modelling, #11 on relative frequency, #14 on puget sound, and #17 on estuaries were all earlier (1990s to early 2000s) clusters in the future city field. In these research paths, although there are individual citation burst nodes, they basically no longer receive attention after 2010. Clusters #7 on connectivity, #16 on earth system models, #9 on urban renewal, #10 on local climate zones, #4 on smart city and #5 on urban expansion emerged after the disappearance of the previous batch of clusters, and although there have been more concentrated burst nodes and structural centrality nodes, high-profile outputs still lack after about 2018 and 2019. Cluster #0 on scenario has been discussed since 2017, while the popularity has declined in recent years. Other clusters like #13 on autonomous vehicles, #6 on Anthropocene, and #1 urban planning, etc. keep hot and increasing.

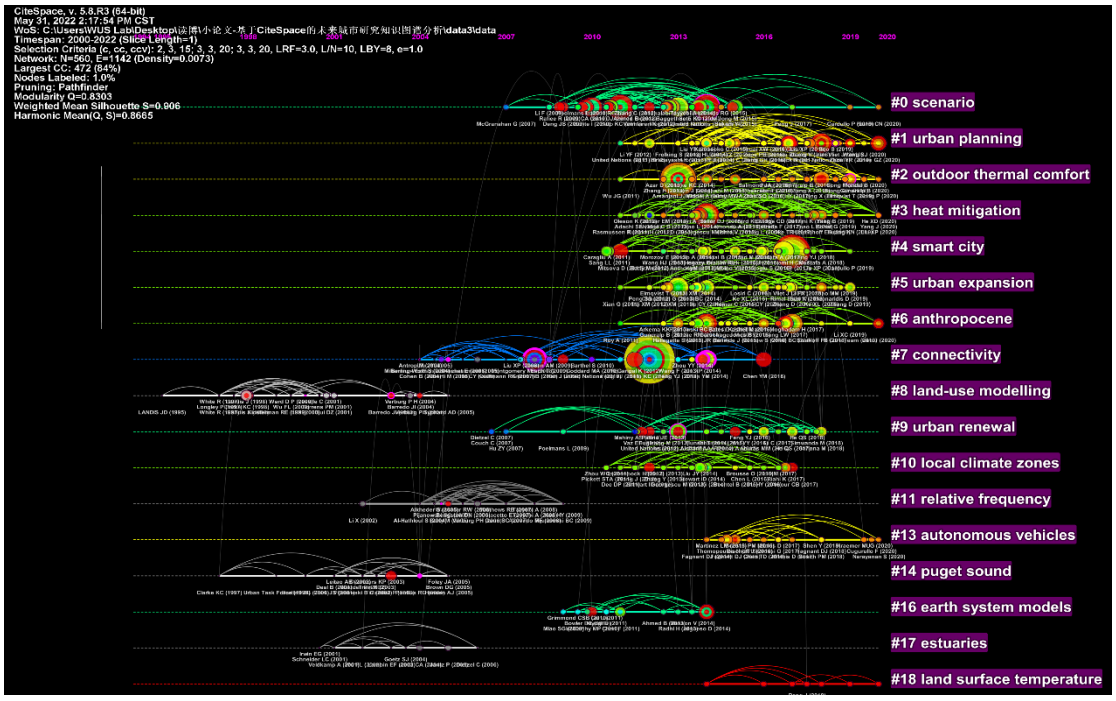

Figure 4. Timelines of co-citation clusters.

5. Structural variation Analysis： emerging trends

The modularity of a network measures the degree to which nodes in a network can be divided into groups, such that nodes within the same group are more closely connected than nodes between different groups. The collective knowledge structure of a scientific field can be represented as an associative network of co-cited references. Such networks will evolve over time. Newly published articles may introduce profound structural changes or have little or no effect on structure.

Figure 5 shows the change of modularity of networks over time. Each network is constructed based on a 2-year sliding window. The number of publications per year increased considerably. It is

noticeable that the modularity dipped in 2007 and bounced back to the previous level before it dropped even deeper in 2009. Based on this observation, it is plausible that groundbreaking works appeared in 2007 and 2009. We will, therefore, specifically investigate potential emerging trends in these 2 years.

Figure 5 shows the change in network modularity over time. Each network is built based on a 1-year sliding window. The number of publications increases steadily each year. It is worth noting that the modularity indicator gradually increased in the first 3 years, and then decreased slowly after 2003. Based on this observation, we consider that innovative work does not emerge in a blowout fashion, but in gentle force. The newly published paper has led to a decrease in the modularity of the network, that is, the originally scattered clusters are "integrated" together to a certain extent. That is to say, the literature published in these years, after accumulation, gradually builds bridges between originally unrelated clusters, thereby causing a slight structural variation in the entire field.

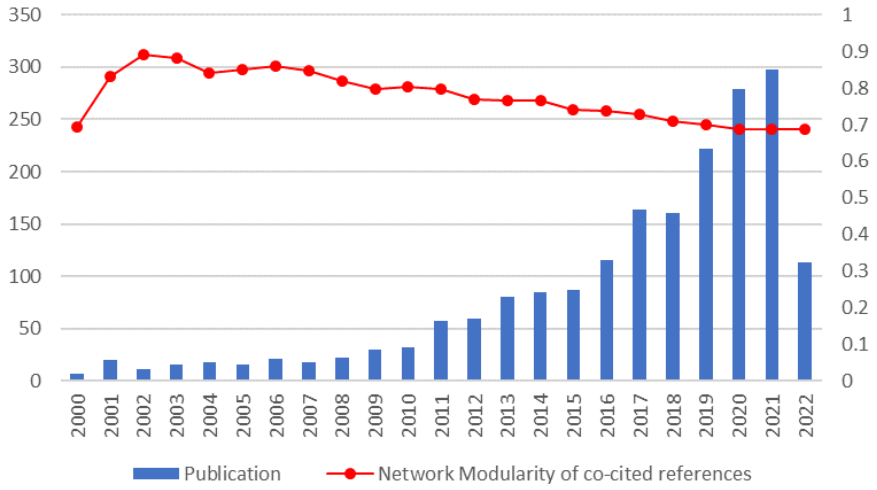

Figure 5. The modularity of the network climbed at the beginning and dropped considerably since 2003, suggesting that nearly no structural changes took place in the long downhill stage.

6. Discussions and conclusions

While bibliometric work is common in other fields, it is relatively rare in the field of future cities. The object of this study is the research status of future cities by different scholars around the world in the past two decades. By conceptualizing the macro-viewpoint of the overall development context and trend of the academic community, as well as clarifying the mainstream research directions, turning points, and important literature of concern, this research hopes to bring some useful references to scholars.

The data for this study comes from WoS, a relatively high-quality paper database, but not all papers are included. Many research results published in other journals were not included in the WoS database and were not collected in this study. Therefore, the results must have certain limitations. To

expand the breadth and depth of the research, it is suggested that other scholars can use different analytical tools to collect data from other databases than WoS. Results can be compared to identify similarities and differences in order to provide scholars with more detailed and valuable research.

In conclusion, the analysis and citation-based extension of the future urban literature outlines the evolutionary trajectory of collective knowledge over the past two decades and highlights areas of active pursuit. Emerging trends and patterns identified in the analysis are based on CiteSpace's computed properties. In a more ideal way, it is necessary to combine bibliometric tools with traditional literature reading and review work, in order to obtain a comprehensive, in-place and accurate grasp of the knowledge base of future cities. In short, there is still a lot of research space for future cities. Scholars need to use different paradigms, methods, and perspectives to explain and measure future cities, and further expand the research field.

References

Chen, Y., Chen, C. and Wang, Z.Y. (2015). *The Methodology Function of CiteSpace Mapping Knowledge Domains*. Studies in Science of Science. 33, 242-253.

Chen, C. (2006). *CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature*. Journal of the American Society for information Science and Technology, 57(3), 359-377.

WU, Z., HAN, J., TENG, Y.W., HUANG, Y.Z., & GAN, W. (2018). *Qingdao Sino-German Future City: Absorption of Innovative Concepts in Line with Laws and Devotion to Homeland Contribution*. Eco-City and Green Building, 2, 27-33.

Chen, X., & Liu, Y. (2020). *Visualization analysis of high-speed railway research based on CiteSpace*. Transport Policy, 85, 1-17.

Wu, Y.Z., YANG J.C., CHEN Q.H. (2020). *Research Progress and Key Dimension Exploration of Healthy Community Construction: Based on the Analysis of International Knowledge Map*. Urban Planning International, 35(05), 80-90.