

Knowledge mapping of tourism demand forecasting research

Chengyuan Zhang^a, Shouyang Wang^b, Shaolong Sun^{c,*}, Yunjie Wei^b

^a School of Economics and Management, Beihang University, Beijing 100191, China

^b Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

^c School of Management, Xi'an Jiaotong University, No.28, Xianning West Road, Beilin District, Xi'an, 710049, China



ARTICLE INFO

Keywords:

Bibliometric analysis
Tourist arrival
Hospitality demand
Knowledge map
CiteSpace
Infographic

ABSTRACT

Utilizing a scientometric review of global trends and structure from 388 bibliographic records over two decades (1999–2018), this study seeks to advance the building of comprehensive knowledge maps that draw upon global travel demand studies. The study, using the techniques of co-citation analysis, collaboration network and emerging trends analysis, identified major disciplines that provide knowledge and theories for tourism demand forecasting, many trending research topics, the most critical countries, institutions, publications, and articles, and the most influential researchers. The increasing interest and output for big data and machine learning techniques in the field were visualized via comprehensive knowledge maps. This research provides meaningful guidance for researchers, operators and decision makers who wish to improve the accuracy of tourism demand forecasting.

1. Introduction

The tourism industry, a significant sector for economic growth, helps develop catering, accommodation, transportation, entertainment, retail and other consumer tourism industries (Liu, Liu, Wang, & Pan, 2018; Pai, Hung, & Lin, 2014; Sun, Wei, Tsui, & Wang, 2019). Global international tourists (overnight visitors) grew by 6% to an estimated 1.4 billion in 2018, and tourism revenue that year reached US\$5.34 trillion, equivalent to over 6% of global GDP. In emerging economies, tourism income is outside, equivalent to 15% of their GDP. The Middle East (+10%) and Africa (+7%) led 2018 tourism growth, with the increase in arrivals to Asia and the Pacific and to Europe (both +6%) remaining in line with the global average. (United Nations World Tourism Organization (UNWTO), 2019; World Tourism Cities Federation (WTCF), 2019).

Tourism's contributions to the economic growth of many countries and regions make tourism management and, consequently, the forecasting of tourism demand fertile areas for scholarly research (Athanasopoulos, Song, & Sun, 2018; Coshall & Charlesworth, 2011; Song & Li, 2008). Considering the disconnect between limited tourism resources and the steady growth in demand, accurate forecasting of demand is a vital research direction that can enable practitioners and decision makers in the tourism industry to optimize resource allocation (Gunter & Önder, 2015; Pan & Yang, 2017; Song & Li, 2008). In the past decades, a large number of forecasting studies and methods have been

proposed, applied and tested in the field of tourism demand forecasting (Chen, 2011; Li, Song, & Witt, 2006; Yao et al., 2018). With the globalization of world economy and rapid development of modern information technology, new topics and insight into tourism demand forecasting are emerging (Li, Ma, & Qu, 2017; Pan & Yang, 2017). Given the sharp increase in research interest in tourism demand forecasting, traditional tourism journals have developed and academic journals in related disciplines publish relevant academic papers, traditional narrative reviews and systematic quantitative review (Claveria, Monte, & Torra, 2015; Liu et al., 2018; Song & Li, 2008; Yang, Khoo-Lattimore, & Arcodia, 2017a). Despite extensive publications and methods for forecasting tourism demand, previous review articles have infrequently visualized knowledge maps in the field, leaving little known about such maps' potential to show overall structure (Fang, Yin, & Wu, 2018; Li et al., 2017a).

A priority in this paper is to express the concept, method and hot topic of tourism demand forecasting research through visual/graphic (infographic) formats to demonstrate how the field develops over time and to see the future research directions (Fang et al., 2018; Li et al., 2017a). Scientometric analysis is both a visual knowledge figure and a serialized knowledge pedigree for discovering knowledge domains, identifying research trends, and revealing complex relationships among authors and organizations (Olawumi & Chan, 2018; Li et al., 2017a; Si, Shi, Wu, Chen, & Zhao, 2018). Scientometric analysis also can evaluate and examine the research development and performance of institutions,

* Corresponding author at: School of Management, Xi'an Jiaotong University, No.28, Xianning West Road, Beilin District, Xi'an 710049, China.

E-mail addresses: cyzhang@buaa.edu.cn (C. Zhang), sywang@amss.ac.cn (S. Wang), sunshl@amss.ac.cn (S. Sun), weiyunjie@amss.ac.cn (Y. Wei).

countries and journals in a specific research field (Olawumi & Chan, 2018). To provide one-stop approach for scientometric analysis, CiteSpace, an especially popular tool for knowledge mapping, integrating co-citation analysis, collaboration network and evolutionary trend detection (Yang et al., 2018) is useful for identifying among papers their knowledge domains and emerging trends. Nodes and links are the building blocks of CiteSpace visualization graphs which use concentric circles of different colors in a node to denote researches in time slices with networked relationships represented by links (Fang et al., 2018).

By using the CiteSpace and reviewing tourism demand forecasting research published between 1999 and 2018, it is possible to display the evolution of a knowledge domain on a network map and to identify research frontiers. In particular, four major questions should be focused on: 1) What are the statistical key figures of the development status of subject classification, high-yield journals, high-yield authors, high-yield institutions and highly cited papers? 2) What are the citation status and influence of documents, journals and authors in co-citation analysis? 3) How prominent are individual countries/regions, institutions and authors in the corresponding collaboration network? 4) What research phases and opportunities for future research seem promising? Accordingly, three main objectives of this study are: (1) summarize the tourism forecasting research published during 1999–2018, by statistical analysis subject categories, high-yield journals, high-yield authors, high-yield institutions, and highly cited papers; (2) provide an analysis of the overall research status for tourism demand forecasting from the perspective of document, journal, author co-citation analysis and collaboration network; (3) present, based upon keywords analysis and documents co-citation cluster analysis, a new integrated, holistic knowledge map that includes knowledge domains, evolutionary trends, and future research directions.

In Section 2, we elaborate on methodology involving research framework and database construction. In Section 3, six statistical summaries of findings are demonstrated. The results of the scientometric analysis are reported in Section 4, including the co-citation analysis, collaboration network and emerging trends analysis. Section 5 gives a comprehensive conclusion and corresponding discussion based on scientometric analysis, and suggestions for future research.

2. Methodology

This section gives our methodology of scientometric analysis for tourism demand forecasting. Section 2.1 presents the methodology's general framework. Section 2.2 shows the collection of empirical data.

2.1. Framework

In this study, we constructed an integrated analysis framework to analyze and visualize tourism demand forecasting research published in 1999–2018, as shown in Fig. 1. This study encompassed a statistical summary of findings and scientometric analysis. Based on a descriptive statistical summary, six interesting general findings regarding the distribution of topical articles published in 1999–2018, subject categories, statistics about high-yield journals, authors and institutions, and most highly cited papers were deduced for the existing tourism research in the first part. As for the scientometric analysis, author and institution influence, the cooperation network between countries/regions, the research phases, hot topics and future research directions can be obtained by identifying and visualizing the evolution of the co-citation analysis, collaboration network and emerging trends. While the co-citation analysis incorporates co-cited authors, co-cited documents, and co-cited journals, the collaboration network mainly analyzes cooperation among countries/regions, institutions and authors, while the emerging trends analysis includes keywords analysis and documents co-citation cluster analysis.

2.2. Database construction

For the purpose of this study, the core collection database from Web of Science (WoS), which contains some 12,000 high-impact journals and more than 160,000 conference proceedings, was used to obtain research data in the field of tourism demand forecasting (Fang et al., 2018; Li, Ma, & Qu, 2017). The use of WoS pinpointed high-quality journal articles that guaranteed the reliability of the study's data source.

When retrieving articles from the WoS Core Collection database, the criteria were: (1) themes = “tourism demand forecasting” or “tourism demand prediction” or “tourist arrivals forecasting” or “tourist arrivals prediction” or “tourist flow forecasting” or “tourist flow prediction” or “hotel demand forecasting” or “hotel demand prediction” or “hotel room forecasting” or “hotel room prediction”; (2) database = Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index-Science (CPCI-S) and Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH); (3) time span = “1999–2018”; (4) document types = “article” or “review” or “proceedings paper.” Literature type was limited to journal articles in English. The volume of the research literature on tourism demand forecasting and the need for accuracy necessitated setting those four screening criteria (Ghoddusi, Creamer, & Rafizadeh, 2019). When each article's bibliographic information was downloaded from the WoS Core Collection database, the different variables (e.g., author(s), title, abstract, keywords, source publications, research direction and references) were retained. Articles unrelated to research theme were removed through a manual screening process. Proceedings papers with fewer than five citations were deleted. The search with these parameters retrieved 388 publications that were downloaded May 4, 2019.

3. Statistical summary of findings

This section provides the articles' statistical account by demonstrating the distribution of published papers, subject categories, high-yield journals, high-yield authors, high-yield institutions, and highly cited papers.

3.1. Distribution of published papers

The time distribution of literature from 1999 to 2018 (illustrated in Fig. 2.), and some periodical characteristics of article evolution can be seen. The study shows a steady global increase in number of tourism demand forecasting articles, falling into roughly three phases. In the first (development) phase, the number of articles published from 1999 to 2007 was under 10 per year, a sluggish rate that corresponds to the rate of development of the global economy. In the second (rapid development) phase, more than 20 articles were annually published during 2008 to 2013, and in this third (steady development) phase, the number of annually published articles increased to more than 35 during 2014 to 2018.

Particularly, the published article numbers do show that, before 2008, the top three countries each published an equal number of articles, but since then, China (including Mainland, Hong Kong, Taiwan and Macao) has gradually become the most prolific country. Specifically, in the rapid development phase (2008 to 2013), the number of articles in the top three countries rose and then declined to varying degrees. In the steady development phase (2014 to 2018), the number of articles from the United Kingdom (England, Scotland, Wales and Northern Ireland) and the United States were in a steady state, while Chinese articles experienced an upward trend of rising excepting 2016. For the period 1999–2018, China accounted for 39.17% in overall volume of publications, the US for 16.49% and the United Kingdom for 12.88%.

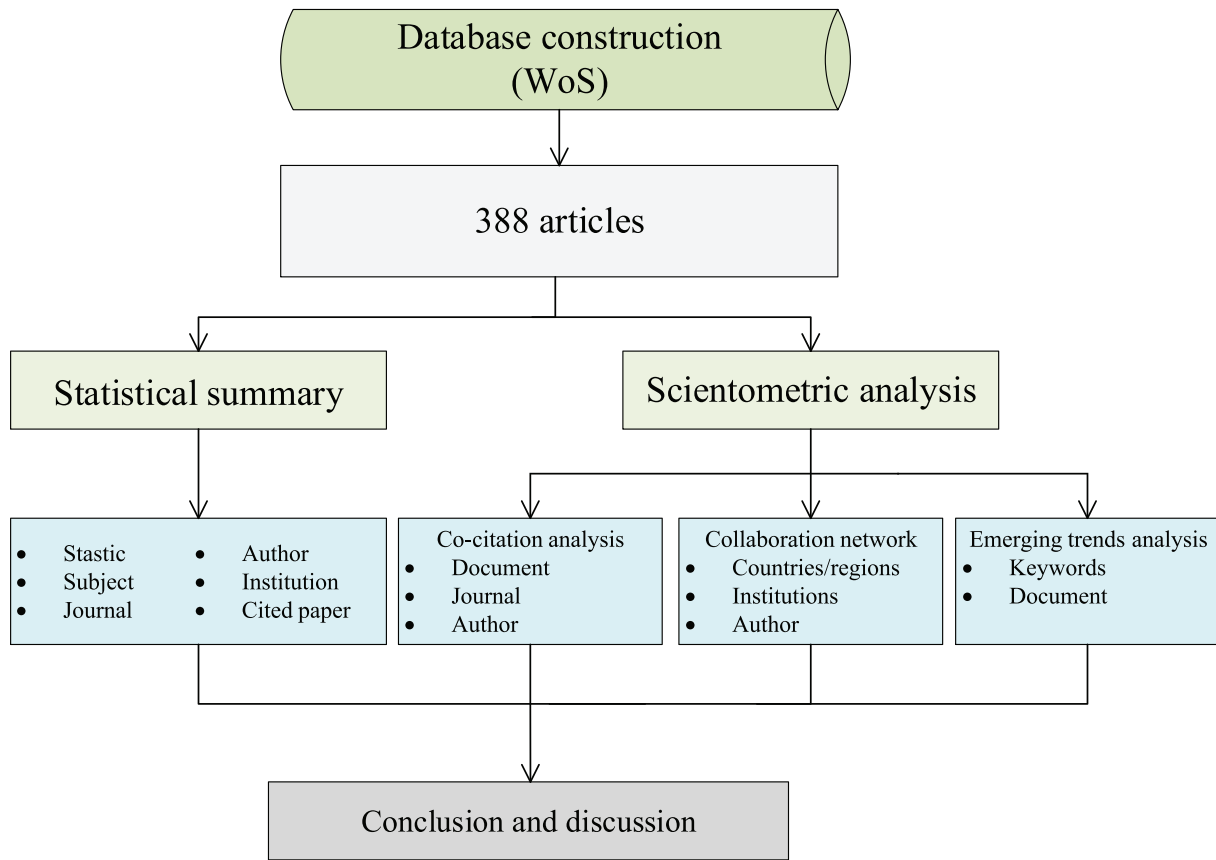


Fig. 1. The integrated analysis framework.

3.2. Subject categories

From the perspective of subject categories distribution, as shown in Fig. 3, tourism demand forecasting is a multidisciplinary research field mainly involving hospitality leisure sport tourism (56.44% of the total), management (30.93%), economics (22.86%), environmental studies (18.30%) and computer science artificial intelligence (7.47%).

3.3. High-yield journals

From the distribution, as shown in Table 1, it can be seen that high-yield journals primarily come from tourism and hospitality, and from forecasting and operations research. *Tourism Management* leads with 65 publications, followed by *Tourism Economics* (41), *Annals of Tourism Research* (21) and *Journal of Travel Research* (21). Journals about forecasting and hospitality management also issued highly influential articles. These ten journals shown in Table 1 are fundamental to tourism

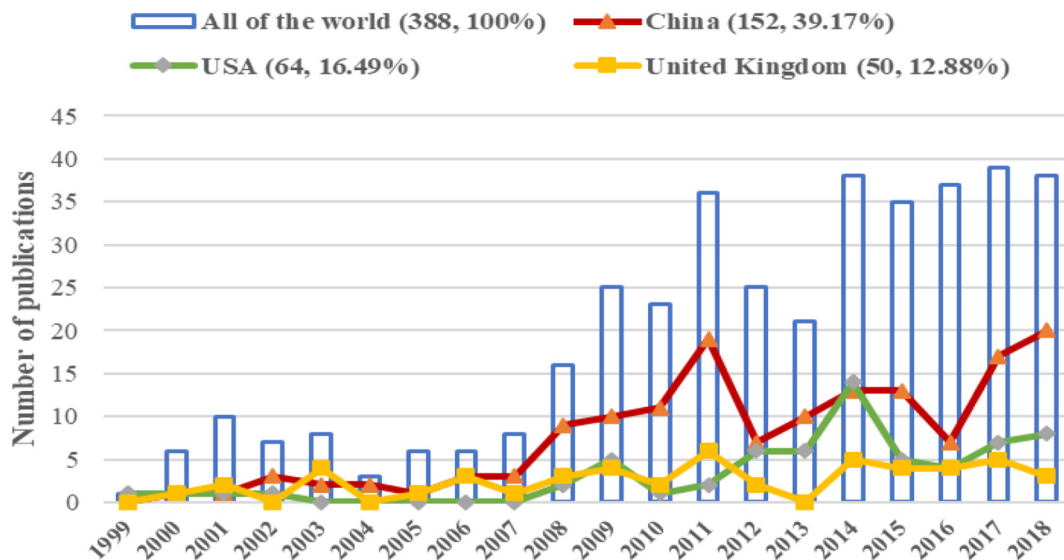


Fig. 2. Time distribution of tourism demand forecasting literature, 1999–2018.

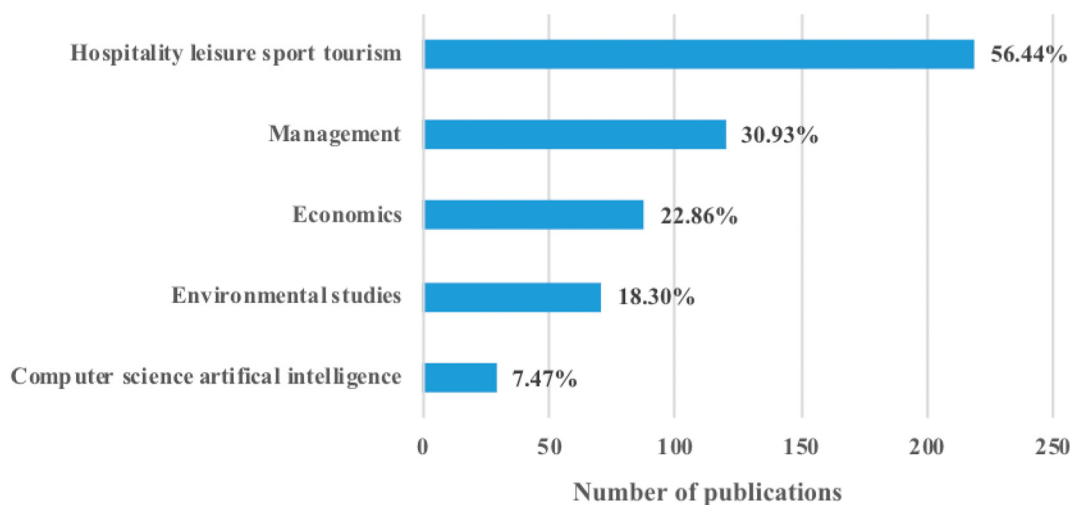


Fig. 3. Distribution of the main subjects of tourism demand forecasting literature.

Table 1
High-yield journals in tourism demand forecasting.

Num.	Journal	Number of publications	Proportion /%	Impact factor (2018)	Country
1	Tourism Management	65	16.75%	6.012	England
2	Tourism Economics	41	10.56%	1.098	England
3	Annals of Tourism Research	21	5.41%	5.493	USA
4	Journal of Travel Research	21	5.41%	5.338	USA
5	International Journal of Forecasting	16	4.12%	3.386	Netherlands
6	International Journal of Contemporary Hospitality Management	12	3.09%	3.957	England
7	Expert Systems with Applications	10	2.57%	4.292	USA
8	International Journal of Hospitality Management	10	2.57%	4.465	England
9	International Journal of Tourism Research	10	2.57%	2.278	England
10	Journal of Travel Tourism Marketing	7	1.80%	2.988	England

forecasting research publishing, but it is difficult for readers to visualize their influence and citation status by only using data represented in Table 1.

3.4. High-yield authors

It can be seen that the high-yield authors mainly come from China, England, Spain, Australia and the USA (presented in Table 2). Notably, Authors may have a cross-country cooperative relationship (Liu et al., 2018), but from the above table it is not possible to draw the co-operative relationships between them.

3.5. High-yield institutions

As for the research institution, institutions that in 1999–2018 spawned the most published articles are based in Western countries and China (Table 3). The Hong Kong Polytechnic University ranked first in

Table 2
High-yield authors in the field of tourism demand forecasting, 1999–2018.

Author	Country	Total number of published literatures	Total number of citations	Mean citations per article	Maximum number of citations	H index
Song, H. Y.	China	36	1628	45.22	512	43
Law, R.	China	14	1092	78.00	399	53
Witt, S. F.	England	14	807	57.64	121	35
Li, G	England	9	746	82.89	512	20
Athanasopoulos, G.	Australia	8	299	37.38	70	13
Pan, B.	USA	8	165	20.63	62	23
Claveria, O.	Spain	7	90	12.86	60	7
McAleer, M.	Australia	7	256	36.57	131	40
Torra, S.	Spain	7	90	12.86	60	6
Monte, E.	Spain	6	20	3.33	15	10

the number of such publications during study period. However, although the number and type of academic institutions could be partially reflected in the table, the knowledge map can visualize the current development status of research institutions to make it easier for researchers to understand the influence of research institutions in the field of tourism demand prediction.

3.6. Highly cited papers

The top ten highly cited papers found in the 1999–2018 period (Table 4) were regarded as the knowledge base for the tourism demand forecasting research. Among them, Song and Li (2008) published in *Tourism Management* had the highest value of citations (512 citations), and played a significant role in this field.

Table 3
High-yield institutions in tourism demand forecasting.

Institution	Country	Number of articles	Share within the country /%	Share in the world /%
Hong Kong Polytechnic University	China	62	40.79%	15.98%
University of Surrey	England	17	36.96%	4.38%
Monash University	Australia	10	24.39%	2.58%
Bournemouth University	England	9	19.57%	2.32%
University of Barcelona	Spain	9	22.50%	2.32%
National Chi Nan University	China	8	5.26%	2.06%
Victoria University	Australia	8	19.51%	2.06%
Griffith University	Australia	7	17.07%	1.80%
Penn State University	USA	7	10.94%	1.80%
Polytechnic University of Catalonia	Spain	7	17.50%	1.80%

4. Scientometric analysis

The scientometric analysis that the study conducted to visualize the review of the tourism demand forecasting knowledge area is explained in the following section in terms of co-citation analysis, collaboration network, and emerging trends analysis.

4.1. Co-citation analysis

The analytical tool, co-citation analysis, is often employed to examine a large amount of literature and to reveal a scientific discipline's knowledge maps (Acedo & Casillas, 2005). Such analysis examines the frequency that two items of related literature (such as document, journal, or author) are co-cited, meaning jointly cited, by the literature (Olawumi & Chan, 2018). By producing and analyzing the document co-citation network, journal co-citation network, and the author co-citation network, the scientific knowledge structures for the field can be obtained.

4.1.1. Document

The study's document co-citation analysis evaluated references cited by 388 bibliographic records that were oriented towards understanding the intellectual structures of tourism demand forecasting knowledge domain. In the network of documents co-citation, each node represented one document, the links of the connected nodes represented co-citation relationships, and the different colors expressed corresponding time slices (each time slice represents one year). Larger nodes indicate articles cited by numerous different scholars, suggesting the article's relatively more substantial influence to-date, as shown in Fig. 4.

To place a numerical value on influence, the study determined that at the end of 2018, 35 documents each received 100 or more citations. Within the network of document co-citation, 13 articles (48.1% percent) of the top 27 cited articles were published in *Tourism Management*, as anticipated, since it is the source journal with most publications on tourism demand forecasting topics. Among the most highly cited papers, Song and Li (2008), the authors who received the highest number

of citations, had reviewed studies published since 2000 on tourism demand modeling and forecasting, and had proposed forecast combination and integration of qualitative and quantitative approaches as research directions for improving accuracy.

In terms of network connections and, therefore, potential influence, as visually shown in Fig. 4, after Song and Li in size, comes Wong, Song, Witt, and Wu (2007) and Cho (2003) Based on a combination of four popular methods for forecasting tourist arrivals, Wong et al. (2007) had suggested that a combined forecasting method is likely to be better than a single method in many practical situations. Cho (2003) investigated the application of three time-series methods (exponential smoothing, univariate ARIMA, and Elman's Model of Artificial Neural Networks) to forecast tourist arrival demand, concluding that Neural Networks is the best way to predict visitor arrivals, especially for a series without an obvious pattern. These review articles could have influence because of their role in the past and the future, but pressing topics represented by new forecasting technology applications have also drawn researchers' attention.

Three phases seem obvious from the perspective of time dimension. Of the top 27 cited articles, 13 high-frequency cited papers belonged to the development phase (1999 to 2007), and 4 of the top 5 were in this phase, i.e., Li, Song, & Witt, 2005 (frequency = 38), Wong et al. (2007) (frequency = 33), Cho (2003) (frequency = 30) and Song and Witt (2006) (frequency = 29). In the second phase (2008 to 2013), there were 9 high-frequency cited papers. During the two-phase, the knowledge structure began taking shape. The top high-frequency cited paper, Song and Li (2008), with a frequency of 84, was in this phase. In the third phase (2014 to 2018), the number of high-frequency cited articles was 5; three of them introduced web search data into tourism demand forecasting and proved that the introduction of search engine data was trending in big data. Notably, to improve the prediction accuracy, effective quality management of search engine data is an obvious research direction and includes the issue of language bias (Bokelmann & Lessmann, 2019; Dergiades, Mavragani, & Pan, 2018) and platform bias (Dergiades et al., 2018).

Table 4
Highly cited papers in the field of tourism demand forecasting.

Authors	Published year	Journal	Total number of citations	Average citations per year
Song & Li	2008	Tourism Management	512	42.67
Ye, Law, & Gu(Ye et al., 2009)	2009	International Journal of Hospitality Management	399	36.27
Lee & Taylor(Lee and Taylor, 2005)	2005	Tourism Management	166	11.07
Goh & Law(Goh and Law, 2002)	2002	Tourism Management	166	9.22
Chen & Wang(Chen and Wang, 2007)	2007	Tourism Management	154	11.85
Law	2000	Tourism Management	147	7.35
Lim & McAleer(Lim and McAleer, 2002)	2002	Tourism Management	131	7.28
Cho	2003	Tourism Management	123	7.24
Shan & Wilson(Shan and Wilson, 2001)	2001	Applied Economics Letters	123	6.47
Song & Witt	2006	Tourism Management	121	8.64

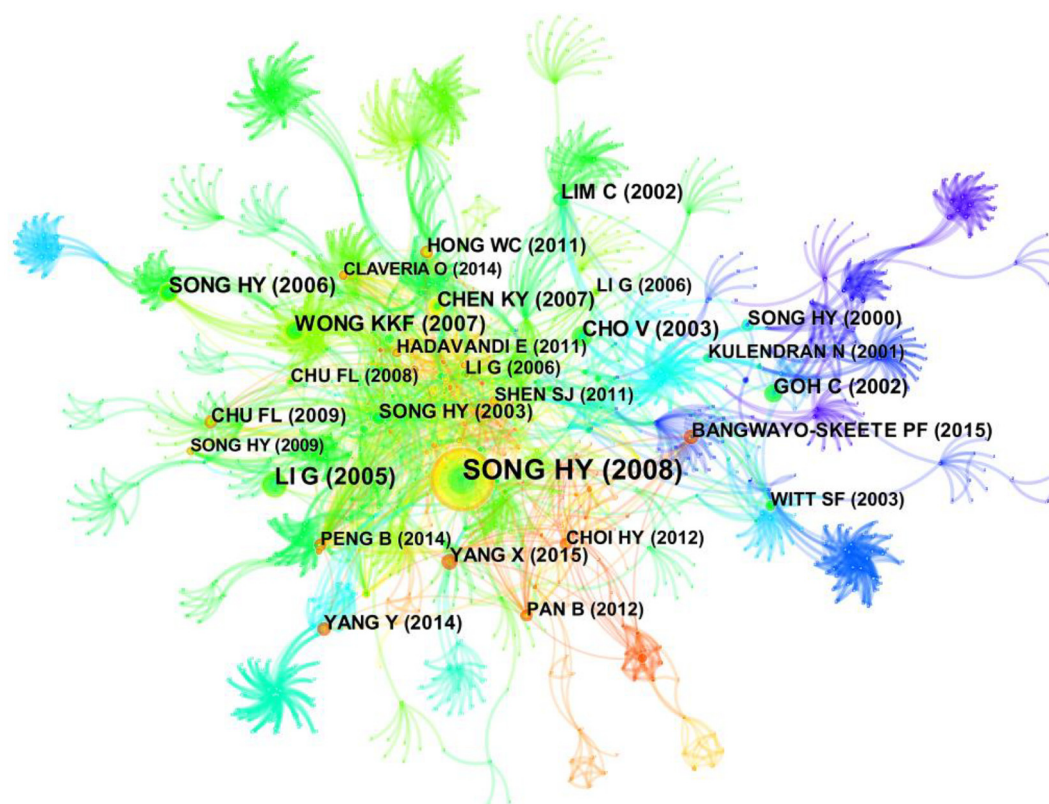


Fig. 4. Network of documents' co-citation: 1999–2018.

4.1.2. Journal

Co-citation analysis of journals enables researchers to better understand mainstream journals and their relative influence. The 388 WoS bibliographic records accessed for this study were sourced from a hundred and thirty-four (134) journals, among them eight journals with ten or more records in the research corpus. Notably, Betweenness Centrality (BC), which is a measure of the centrality of a node in a network by calculating the fraction of shortest paths between node pairs that pass through the node of interest, is a reliable index of measure node (Leydesdorff, 2007; Newman, 2005). Therefore, the BC is employed in co-citation analysis for finding and measuring journal importance, allowing the study to highlight such journals with a purple circle in the co-cited journal network.

Based on a summary of 20 years' worth of journals (1999–2018) serving tourism demand forecasting research, Fig. 5 shows the co-citation network at journal level. *Tourism Management* is most prominent with 65 co-citations, followed by *Tourism Economics* (41), and *Annals of Tourism Research* and *Journal of Travel Research* (21 each). These three international journals have high impact factors. The popular travel journal *Tourism Economics* also has published a substantial amount of research literature. Since the literature spans several scientific disciplines, top academic journals such as *International Journal of Forecasting* and *Journal of Forecasting* have published related research papers. *Tourism Management* has the highest BC ratio (0.20) but *Journal of Travel Research* is statistically close (0.17), and the two journals can be diagrammed as core nodes connecting other nodes in the journal co-citation network.

By using the journals cited to generate a network of co-cited journals, demonstrating 918 nodes and 3337 links in entirety, the most significant cited journals can easily be seen in an infographic (Fig. 5), with the relative size of the node representing the frequency of co-citation in each journal within the dataset. Co-citation among the network's six most co-cited journals was found to have frequency ranging from 296 to 103 citations: *Tourism Management* (296); *Journal of Travel*

Research (234), *Annals of Tourism Research* (231), *International Journal of Forecasting* (219), *Tourism Economics* (162) and *Journal of Forecasting* (103), confirming from the perspective of co-citation frequency their importance and influence to tourism demand forecasting studies.

4.1.3. Author

The development and improvement of a discipline, including tourism demand forecasting, mainly relies on the cooperation of researchers in related disciplines (Li, Ma, & Qu, 2017; Liu et al., 2018). Regarding the co-citation author network, co-authors' knowledge mapping analysis can visualize data about authors with different influences in their field. Accordingly, the author's co-citation analysis can not only obtain the distribution of highly cited authors, but also identify scholars' relative influence. In the network, the node size is also representative of each author's co-citation frequency within the dataset, and the links indicate an indirect cooperative alliance of authors based on their co-citation frequency.

The author co-citation network that contributes to tourism demand forecasting is shown in Fig. 6, which contains 1588 nodes and 4993 co-citation links. It should be pointed out that in this analysis only the first author is considered and all published articles from a particular author are combined into one. The 11 most cited authors were identified from the network by setting the frequency threshold of 60. These authors, listed with their first co-citation year, were: Song HY (frequency = 194, 2000), Lim C (frequency = 149, 2000), Witt SF (frequency = 124, 2000), Li G (frequency = 100, 2006), Kulendran N (frequency = 98, 2000), Law R (frequency = 92, 2000), Goh C (frequency = 83, 2003), Chu FL (frequency = 71, 2001), Wong KKF (frequency = 69, 2000), Box GEP (frequency = 64, 2000), and Cho V (frequency = 63, 2005). From the first time of reference, all the authors were engaged in tourism demand forecasting research in the field's first phase (1999 to 2007), meaning that they have carried out the related study early and made a consistent contribution.

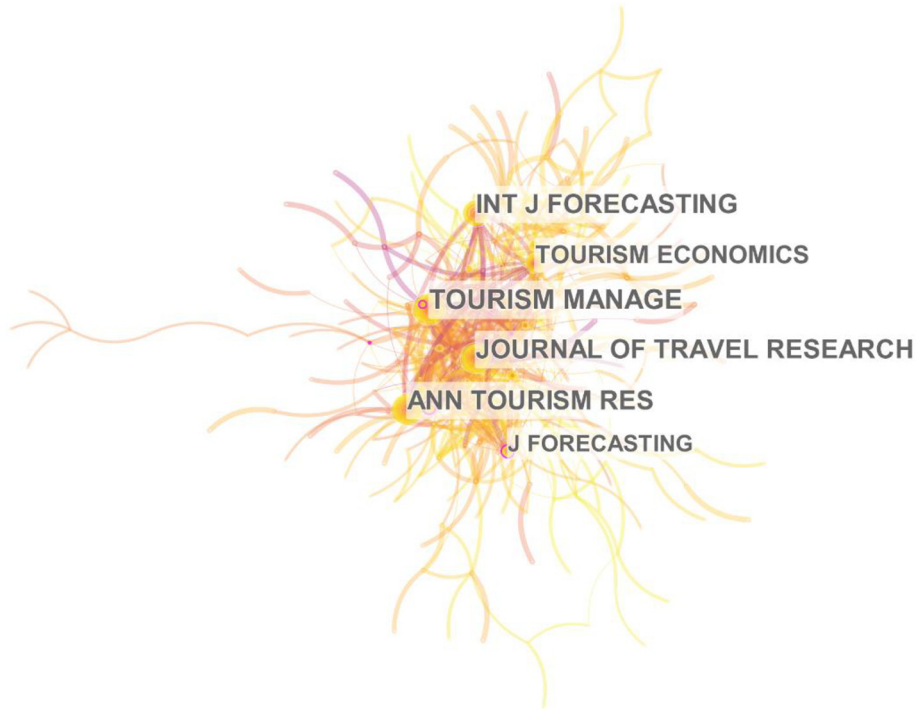


Fig. 5. Network of journals' co-citation: 1999–2018.

4.2. Collaboration network

Collaboration analysis will help to understand the trend of academic knowledge spreading among countries/regions, institutions, and authors, and to locate influential research institutions and authors. Therefore, Collaboration analysis is critical to understand scholarly

communication and knowledge diffusion. Our study included analysis of countries/regions collaboration network, institutions collaboration network, and author collaboration network.

4.2.1. Countries/regions

The network of collaborating countries generated 61 nodes and 129

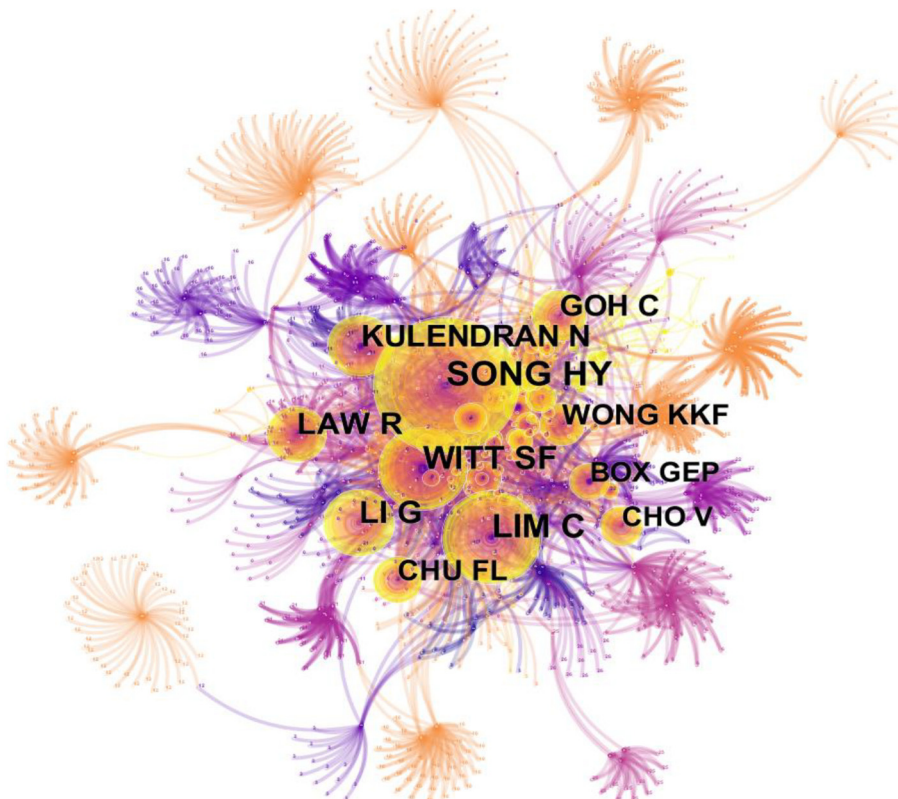


Fig. 6. Network of author co-citation: 1999–2018.

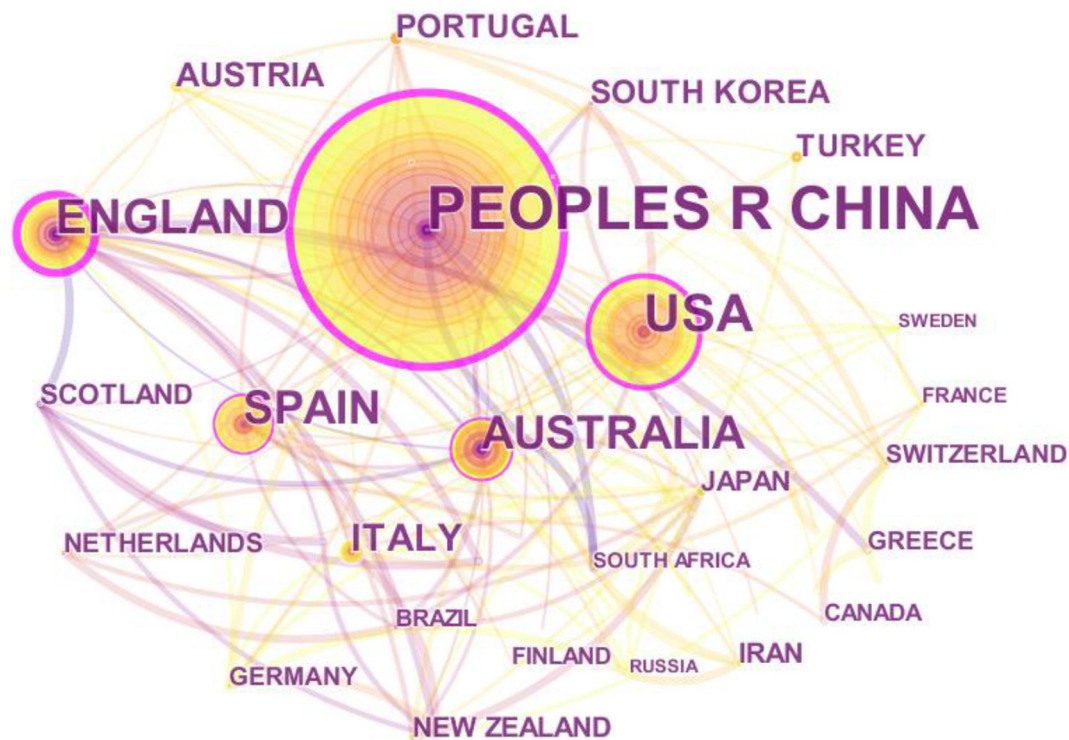


Fig. 7. The network of countries for tourism demand forecasting research.

links, as shown in Fig. 7, for the years studied. In-network, 25 countries were identified by relative contribution (more than 5 articles) to this research area. Countries with the most publications included China (146 articles, 0.51); USA (64 articles, 0.21); England (46 articles, 0.44); Spain (40 articles, 0.16) and Australia (40 articles, 0.11). These top five countries are core nodes establishing links with other nodes in the countries' collaboration network.

The output is related to tourism market demand, number of research institutions, research funding, and proportion of tourism-focused institutions (Fang et al., 2018). China (including Mainland, Hong Kong, Taiwan and Macao) was the largest contributor to the study of tourism demand forecasting, publishing 146 papers, followed by the United States (64 articles). Among the reasons for this are that China and the United States have relatively large numbers of citizens who travel, sizable tourism resources and numerous researchers. The fact that the study uses only articles written in English is another important factor to be kept in mind, as Chinese academics are encouraged to publish in the English language. European countries play a crucial role in making connections with other countries according to their high BC, beginning with England (46 articles, 0.44), Spain (40 articles, 0.16), and Switzerland (0.09). Using lines' color as the visual indicator, cooperative relationships among the top five countries, among authors writing in English, that is, can easily be seen to have been established during the 21st century, while cooperative relationship between the other European countries (such as France and Germany) was established around 2010.

A low level of cooperation is implied by thinner links between countries (Fang et al., 2018). East Asia, China, Japan and South Korea have relatively less close cooperation as indicated among published articles written in English. Among Western countries, Spain and the United States cooperate less closely with the United Kingdom. This also indicates that under the conditions of large tourism resources, the study of bilateral or multilateral tourism cooperation between China, Japan and South Korea could be strengthened. Meanwhile, the US and European countries could also strengthen research on bilateral or multilateral tourism cooperation. From an intercontinental perspective,

though South America and Africa have large tourism resources, tourism management research represented by tourism demand forecasting, as far as sharing knowledge in the English language, is under-represented at present.

4.2.2. Institutions

An analysis of the institution collaboration network can help display relationships involved in the cooperation between key institutions in tourism demand forecasting research and can reveal institutions' influence. In the cooperative network, node size represents number of articles published by the corresponding institution. A time zone map of institutions is also demonstrated to show different institutions with their first-ever publication of research results taken as nodes and arranged in order from far to near. In the time dimension, the visual effect clearly shows the evolution of tourism demand forecasting research institutions.

The institution collaboration network consisted of 443 nodes and 553 collaboration links for the years 1999–2018 (Fig. 8). After assigning 5 articles as threshold, 32 research institutions rate being listed, among them, in rank: Hong Kong Polytechnic University (61), University of Surrey (17), Monash University (10), and Bournemouth University (9). From the quantitative perspective among institutions, Hong Kong Polytechnic University, closely followed by Monash University and the University of Surrey, is the largest contributor and represents the leader in tourism demand forecasting research. Therefore, these institutions are unique in their outputs of research in the field of tourism demand forecasting. Meanwhile, from a national perspective, Hong Kong Polytechnic has links with mainland China institutions, specifically the Chinese Academy of Sciences, Beijing Union University, and Shanxi Normal University. UK is the country that has the next top institutions; its University of Surrey ranked 2 and Bournemouth University ranked 3. Bournemouth University in the UK has connections with the University of Portsmouth, the University of Pretoria, and the University of the Arts London. Griffith University is closely linked to fellow Australian Institutions University of Queensland and University of Western Australia, giving an indication that cultural atmosphere and

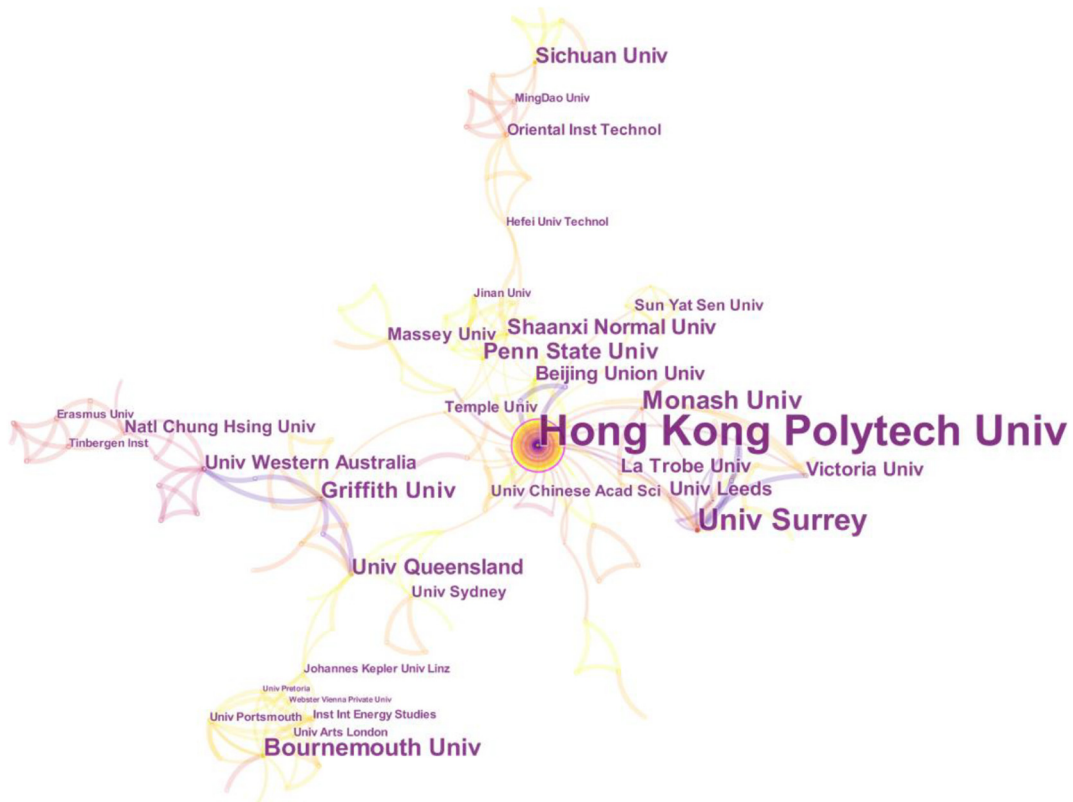


Fig. 8. The network of institutions for tourism demand forecasting research.

country of origin may affect cooperation within tourism demand forecasting research.

This section presents a time zone view of an institution's participation in tourism demand forecasting, as shown in Fig. 9. Over time, the number of research institutions involved has shown a significant upward trend because of attention from researchers who have published. The studied period can also be divided into three distinct phases

from 1999 to 2018 according to the order of publication of institution research results. In the development phase (1999 to 2007), research institutions represented by Hong Kong Polytechnic University and the University of Surrey led in conducting tourism demand forecasting studies, followed by Australia's Griffith University, University of Queensland, and Monash University. In the rapid development phase (2008 to 2013), a group of research institutions, represented by

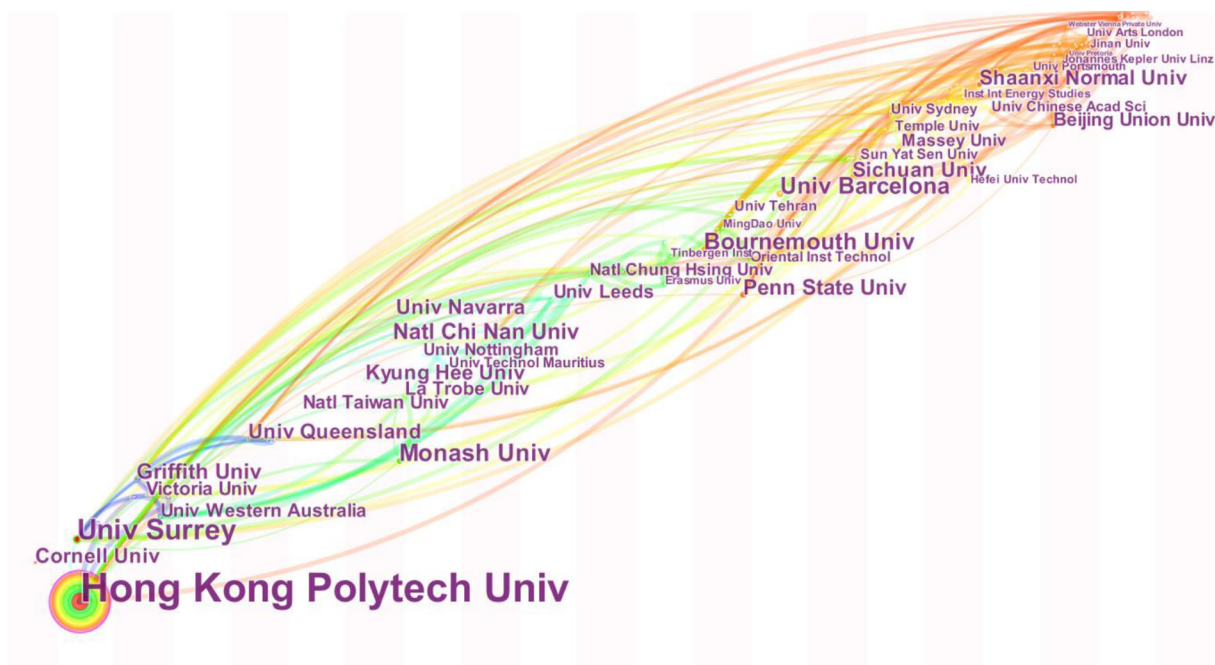


Fig. 9. Time zone view of institutions.

Pennsylvania State University (Penn State) in the US, Bournemouth University in the UK, and University of Barcelona, Spain, joined the tourism demand forecasting field around 2011. During the steady development phase (2014 to 2018), mainland China institutions Sichuan University, Jinan University, Shanxi Normal University and the Chinese Academy of Sciences joined the field. New Zealand's Massey University and University of Sydney also began to be involved, and in England, the University of Portsmouth and the University of the Arts London, also began to invest in research. It may be worth noting that research institutions involved at this phase seemed to be mostly influenced by the top first-phase traditional institutions.

4.2.3. Author

Collaboration among researchers is necessary for the development of tourism demand forecasting (Li, Ma, & Qu, 2017). In the collaboration network, the field's influential authors can be identified by generating a knowledge mapping analysis of the co-author network. Mapping such a network can help researchers establish cooperative relationships.

The collaboration network for authors who contributed to tourism demand forecasting research consisted of 803 nodes and 1190 collaboration links (presented in Fig. 10). The network, with its numerous participants and wide-ranging collaborations, shows the field's interdisciplinary nature and reveals that cooperation between researchers has promoted the development and improvement of tourism demand forecasting. The most networked cooperative relationship was the work published by Song HY (frequency = 34) of the Hong Kong Polytechnic University, followed by Witt SF (frequency = 14) from the University of Surrey, Law R (frequency = 14) from Hong Kong Polytechnic University, and Li G (frequency = 9) from University of Surrey, who are leading scholars in tourism demand modeling and forecasting. Thereafter came Athanasopoulos G from Monash University, and Pan B from Penn State University—each with a frequency of eight—that eventually formed a network of cooperation centered on Song HY and Law R.

The analytical results suggest how institution intimacy and mentoring relationships could contribute to research cooperation over the long term, given that much collaboration is initiated by researchers' doctoral students, with the next most common collaboration type occurring between colleagues at the same university or institution, followed by collaboration between researchers with a past working relationship (Jiang, Ritchie, & Benckendorff, 2017).

4.3. Emerging trends analysis

Keywords are a clear sign of understanding the research paper, and cluster analysis of documents co-citations can highlight researcher-recognized information (Fang et al., 2018; Liu et al., 2018). Therefore, the evolution of hot topics and emerging trends in tourism demand forecasting research can be obtained by analyzing keywords with co-occurrence, time zone, clustering, and documents co-citation clustering

4.3.1. Keywords analysis

Keywords are an obvious marker for understanding research articles' focal content. Keywords can accurately locate research hotspots of tourism demand forecasting (such as research methods and forecasting directions) and, among the keywords, burst words that represent trends emerging at a certain period of time can indicate the research field's frontier (Gunter, Önder, & Smeral, 2019; Yao et al., 2018). Analyzing keyword evolution in tourism demand forecasting will clearly show its trending process on the time zone map. In addition, a cluster analysis was performed for analyzing significant topics, content and inter-relationships. Since the keywords of each group must be related to one another and differ from other groups, the log likelihood ratio (LLR) is selected as the principle of classification statistics. LLR can generate high-quality clusters with intra-class similarity and low inter-class similarity (Fang et al., 2018)

Keywords with the strongest citation bursts were detected and analyzed using CiteSpace as shown in Fig. 11. To more intensely

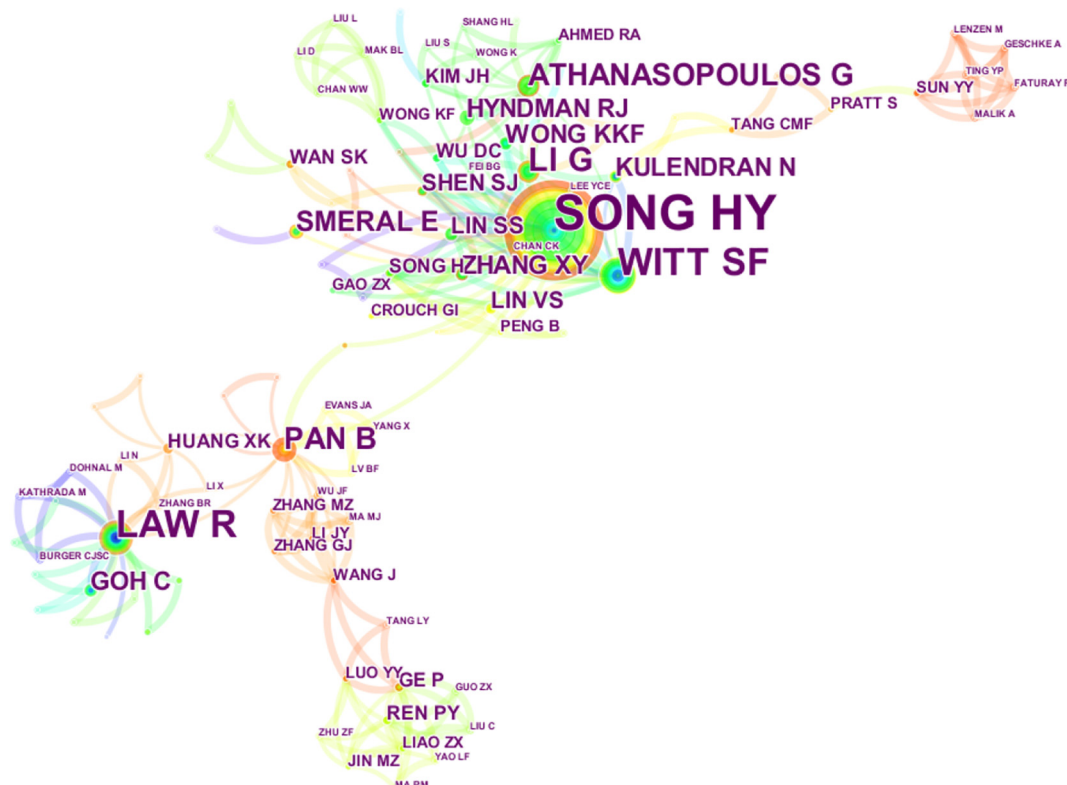


Fig. 10. Visualization of the author collaboration network.

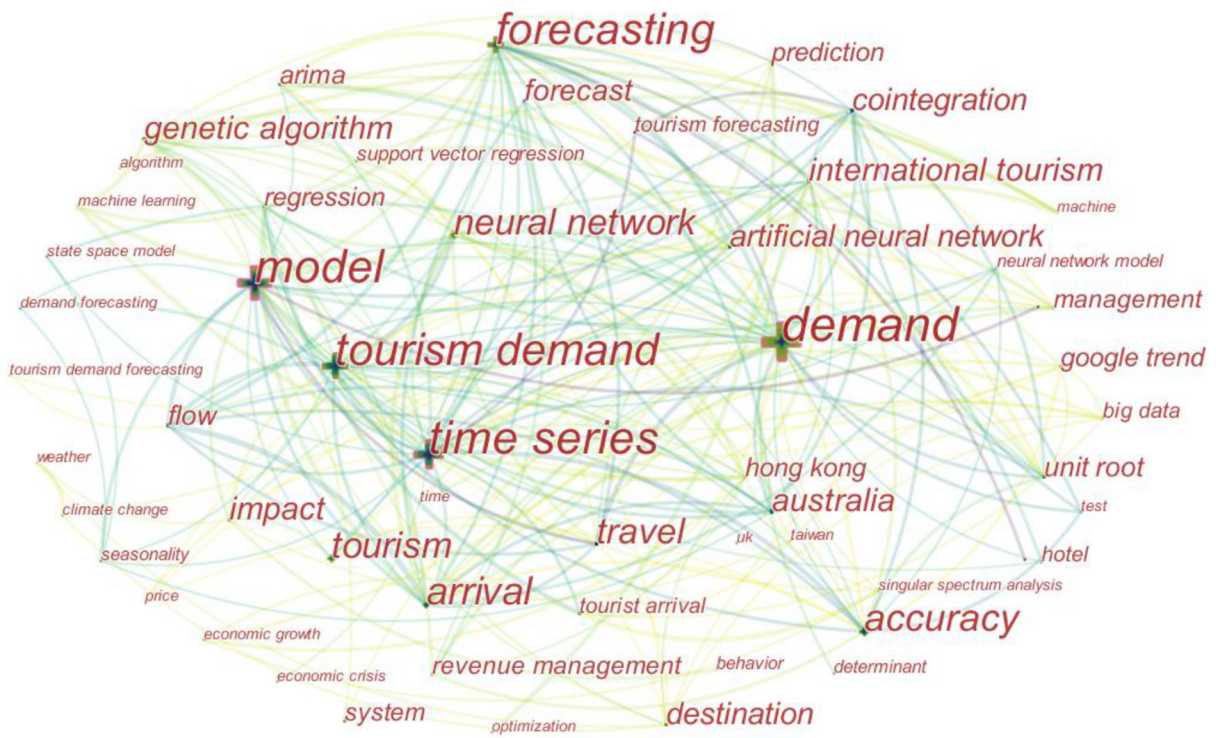


Fig. 11. The network of keywords.

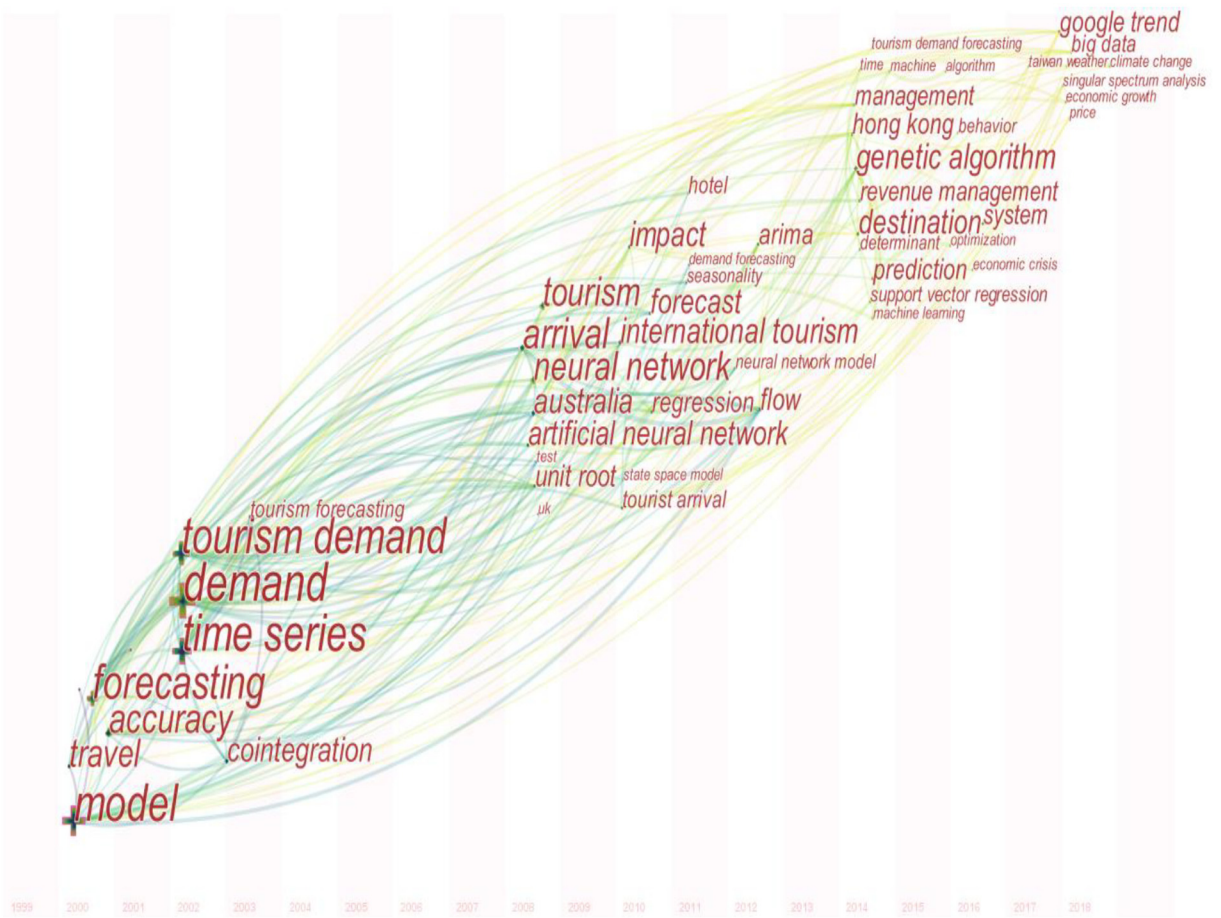


Fig. 12. Time zone map of keywords on tourism demand forecasting.

explore research direction a time zone view of keywords (illustrated in Fig. 12) arranges keywords in correspondence to the time of their publication or their peak time. Based on the network of keywords three interesting results are derived from Fig. 11 and 12. First as mentioned above the size of node point represents the frequency of keywords appearing in the 388 articles. “Tourism demand,” “demand,” “model,” “time series” and “forecasting” are the most frequent keywords which is consistent with the theme research. Second by visualizing the keyword map the time series model (i.e. Arima) econometric model (i.e. state space model) and intelligence model (i.e. artificial neural network) are three common predictions method by counting the frequency of occurrence (Law, 2000; Li, Wong, Song, & Witt, 2006; Song & Li, 2008). Third exogenous variables that affect tourism demand forecasting also appear such as weather climate change behavior big data and Google trend (Gössling, Scott, Hall, Ceron, & Dubois, 2012; Huang, Zhang, & Ding, 2017; Li, Xu, Tang, Wang, & Li, 2018). In the era of big data people’s internet search behavior is obviously helpful for improving forecasting accuracy (Gunter & Önder, 2016; Li, Chen, Wang, & Ming, 2018; Li, Xu, et al., 2018; Padhi & Pati, 2017; Rivera, 2016; Yang, Pan, Evans, & Lv, 2015). For example a very interesting review Li et al., 2018 summarized a comprehensive review on different types of big data in tourism and classified the data sources. Meanwhile Li, Pan, Law, and Huang (2017) proposed a composite search index which collected from Google trends and Baidu index for improving the prediction accuracy. The upcoming 5th generation mobile networks (5G) will generate a greater number of news comments website searches website traffic and other data that may be mined to improve the prediction accuracy. Future research in tourism demand forecasting will not be limited to traditional institutional data as the role of unstructured data in early warning and prediction becomes more prominent

In this study the 388 articles were arranged into six highlighted research clusters based on keywords. Fig. 13 shows the clustering results and the relative importance rank through the LLR test. Accordingly, cluster IDs with the largest group size are cluster #0 “big data” (13 members) and #1 “machine learning” (11 members); cluster #5 “tourist arrivals” was the smallest sized cluster with 5 members. The majority of relationships in clusters #0, #1, #4 and #5 were formed between 2011 and 2014, while some links in clusters #2 and #3 were formed between 2005 and 2008. From the keyword cluster network, it is evident that recent development in tourism demand forecasting research has centered on clusters #0 and #1, as shown by the cluster ID (Fig. 13). As for the cluster #0, the big data application represented by search engine data has opened up new fields for tourism demand forecasting, and has obtained better forecasting accuracy (Huang et al., 2017; Sun et al., 2019). However, overall correction of search engine data for forecasting, as a likely turning point in the research, is receiving attention (Bokelmann & Lessmann, 2019). As for the cluster #1,



Fig. 13. Keyword clusters network.

Figs. 11 and 12 suggest that Support Vector Regression (SVR), Artificial Neural Network (ANN), and Genetic Algorithm (GA) are among the most popular techniques of that cluster. Given ongoing developments in deep network architecture, the deep learning technique based on machine learning will be a promising direction (Law, Li, Fong, & Han, 2019; Lv, Peng, & Wang, 2018). For example, as a very interesting research, Law et al. (2019) proposed deep network architecture for tourism demand forecasting by using the long-short-term-memory (LSTM). Furthermore, in the big data era (i.e., #0) as shown in Fig. 13, text data has become one of the main formats of tourism big data (Li, Li, Zhang, Hu, & Hu, 2019). As a direct reflection of tourists’ opinions, text mining with the text-based natural language processing (NLP) techniques (such as LSTM which is the extension of cluster #1) on unstructured data can effectively capture tourists’ emotions, and their applications in market demand and destination image analysis (Li et al., 2019).

4.3.2. Documents co-citation cluster analysis

Cluster analysis of documents co-citation can effectively divide the research field, method, and branches of the literature into manageable clusters, so that other researchers can objectively grasp the information of each group or cluster (Si et al., 2018; Yang et al., 2018). Moreover, the geographic proximity (Fig. 14) and the label (Table 5) which were both produced by documents co-citation clusters indicate sharing of similar ideas and information exchange for tourism demand prediction (Kuntner & Teichert, 2016). The silhouette score represents the homogeneity of the cluster, which has higher the value mean the more consistent the members in the cluster (Li, Ma, & Qu, 2017). In particular, forty-seven (47) document co-citation clusters were generated from the research power network using the LLR algorithm (Fig. 14), but only 25 clusters are significant. Because the other 22 clusters have zero silhouette scores and only one cluster member, they are not counted as salient clusters in tourism demand forecasting research. The 25 salient and significant clusters, sorted by size, are shown in Table 5. Cluster #0, “tourism demand,” with 133 members, is proportionally the largest cluster, and cluster #46, “seasonal time series,” with five members is smallest in size.

Table 5 benefits from the visualization of the grouping structure and lists the 25 largest clusters in terms of group size. The silhouette scores for clusters range from 0.691 to 0.999. It’s worth noting that every salient group has representative literatures which are the journal articles with highest citation frequency. Furthermore, each cluster label is influenced by the co-cited literature, showing that it is well referenced in the field and therefore worth paying attention to what it represents (Olawumi & Chan, 2018). In this paper, the existing clusters label indicate the main research topic in the field of tourism demand forecasting, and it can also be roughly divided into four categories, including data feature (i.e., #2, #3, #8 and #46), forecasting methods (i.e., #4, #6, #7, #9, #11, #14, #21, #24, #30, #35 and #37), destination target (i.e., the #15 and #19) and significant term (i.e., #0, #5, #10, #12, #16, #22, #33 and #38) (Kuntner & Teichert, 2016).

As for the data feature, the data selected in the literature have obvious characteristics such as seasonality, time series, and monthly frequency that are unique to the tourism market, and are accompanied by the characteristics of big data. Notably, the semi-structured and unstructured data, which are promising data sources in the field of prediction (Colladon, Guardabascio, & Innarella, 2019; Li et al., 2019). Secondly, the forecasting methods mainly involve statistical methods and artificial intelligence methods (Song & Li, 2008). Combining the data characteristics in big data era, how to make data-driven model selection to improve prediction accuracy is an important research direction in the future (Li, Xu, et al., 2018). More sophisticated techniques such as multivariate singular spectrum analysis have also been introduced to this research area, which is consistent with the conclusions of Wu, Song, and Shen (2017). Thirdly, focusing on the destination target, Prideaux, Laws, and Faulkner (2003) and Lee, Song, and

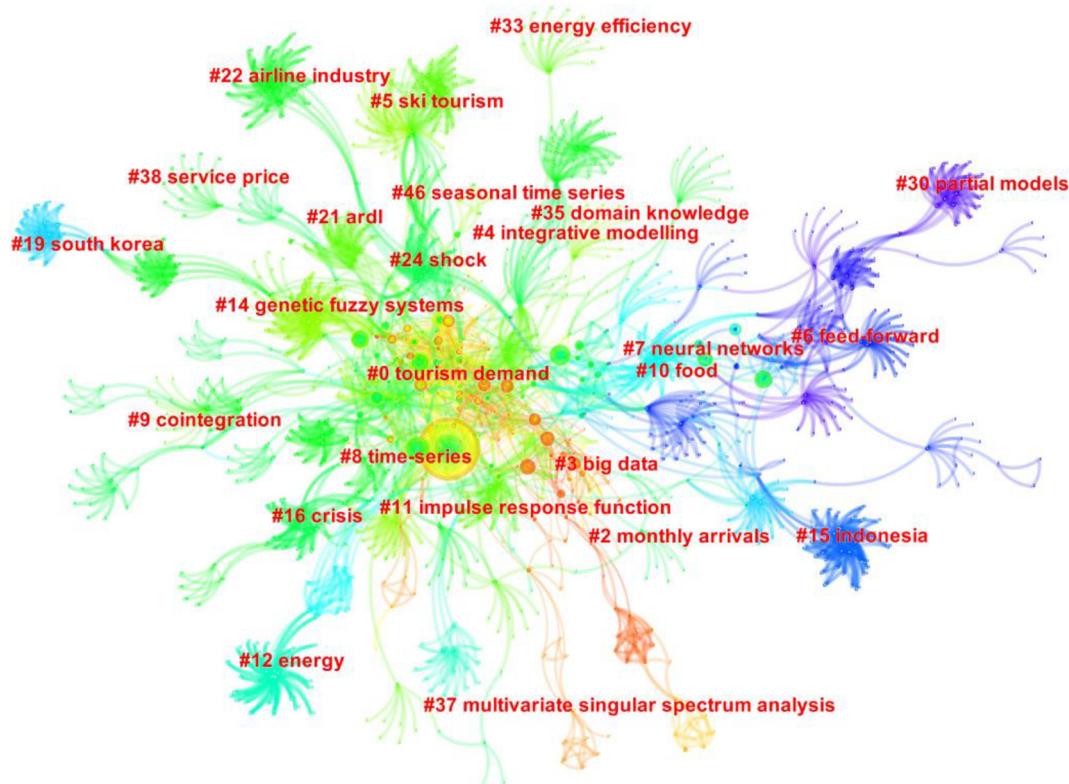


Fig. 14. Document co-citation clusters network.

Table 5
Documents co-citation clusters (1999–2018).

Cluster ID	Size	Silhouette	Mean (year)	Labels
0	133	0.691	2008	Tourism demand
2	80	0.912	2002	Monthly arrivals
3	78	0.938	2011	Big data
4	72	0.964	2006	Integrative modeling
5	71	0.961	2007	Ski tourism
6	65	0.956	1996	Feed-forward
7	59	0.941	1999	Neural networks
8	58	0.926	2009	Time-series
9	57	0.984	2005	Co-integration
10	56	0.934	1999	Food
11	52	0.984	2005	Impulse response function
12	50	0.985	2003	Energy
14	43	0.968	2008	Genetic fuzzy systems
15	39	0.985	1999	Indonesia
16	39	0.97	2007	Crisis
19	33	0.991	2003	South Korea
21	30	0.983	2006	ARDL
22	30	0.994	2005	Airline industry
24	27	0.926	2004	Shock
30	19	0.986	1995	Partial models
33	17	0.999	2008	Energy efficiency
35	15	0.999	2007	Domain knowledge
37	13	0.994	2013	Multivariate singular spectrum analysis
38	13	0.999	2006	Service price
46	5	0.997	2006	Seasonal time series

Bendle (2010) have studied the impact of different events on tourism demand forecasting, such as crisis and visa-free policy. In the future, the impact of the gray rhino incident on tourism demand, such as the length of time, the rate of change in tourist volume, and spillover effects, need to be further strengthened. For example, the impact of political issues between South Korea and Japan on tourism demand, the influence of

Brexit on European countries, and the spillover effects on Macau and Singapore caused by the Hong Kong problem. Lastly, according to the significant term, the research field and branch of literature in the field of tourism demand forecasting can be highlighted, so that other researchers can better review the highlighted relevant researches and expand research directions. Notably, more and more researches are shifting from the analysis of overall tourism demand to the research of market segments, such as ski demand (i.e., #5) (King, Abrahams, & Ragsdale, 2014). Meanwhile, the influence of exogenous variables on the tourism market, such as crisis analysis (i.e., #16) (Smeral, 2009; Song, Lin, Zhang, & Gao, 2010), policy shock (i.e., #12) (Balli, Shahzad, & Uddin, 2018) and online behavior of tourists (i.e., #3) (Li et al., 2019; Li, Chen, et al., 2018), have been favored by researchers. Notably, the effects of different crises on the tourism market deserve attention. For example, the spillover effect of the surge in Singaporean arrivals caused by the Hong Kong issue, and the sharp decline in tourism demand caused by global public health emergencies (such as novel coronavirus pneumonia) has impacted the tourism industry.

5. Conclusion and discussion

A scientific visualization analysis framework is proposed in this paper to depict the 388 articles. Scientometric analysis (i.e., CiteSpace) was used with co-citation analysis, collaboration network and emerging trends analysis in order to present an integrated knowledge map of the tourism demand forecasting field, and to capture hot topics with emerging trends.

Based on the analysis framework, four basic conclusions are drawn. First, according to the statistical data from 1999 to 2018, the field's research literature strongly increased throughout three apparent development phases. The overall trend of tourism in the 21st century was one of continued expansion and diversification (Hassani, Silva, Antonakakis, Filis, & Gupta, 2017), attracting scholars and institutions to the study of tourism demand forecasting. Second, research written in

English on tourism demand forecasting is dominated by China (including Mainland, Hong Kong, Taiwan and Macao), USA, UK, Australia and Spain, based upon the productivity of published authors and institutions. Third, in terms of specific institutions, the Hong Kong Polytechnic University, the University of Surrey, and the University of Barcelona contributed most in the field of tourism demand forecasting based on the articles studied. Hong Kong Polytechnic University is at the center of the most important collaboration network, having established close cooperative networks with other institutions including the University of Surrey and Monash University. Fourth, in terms of the research authors, Song HY (34 articles), Law R (14 articles), Witt SF (14 articles), and Li G (9 articles) were the most productive contributors in this field.

Two conclusions about emerging trends were obtained based on scientometric analysis using CiteSpace, which included methodologies used in tourism demand forecasting research and hot research topics. As for the methodology, in line with Song, Qiu, and Park (2019), time series models, econometric models and intelligence models can also be found to be the primary methods through co-citation cluster analysis of documents in knowledge map of tourism demand forecasting research (see the forecasting method cluster labels #4, #6, #7, #9, #11, #14, #21, #24, #30, #35 and #37 in Table 5). More important, although artificial intelligence methods represented by neural network, genetic algorithms and support vector regressions have been applied by scholars to tourism demand forecasting, deep learning has been less widely applied than in some other fields. Were deep learning's application to tourism demand forecasting to reach full swing, there is greater possibility for improving the prediction accuracy. Furthermore, based on the keyword clusters network, researchers have attempted to combine forecasts generated from different models to try to improve accuracy, but additional advanced individual forecasting methods and multiple forecasting horizons should be explored (Shen, Li, & Song, 2011).

As for the hot research topics, in big data era, a more reasonable integration of web-based data is the focus and direction of future research in forecasting tourism demand. While surveying contemporary topics and trends to understand tourists' thinking and motives, several researchers have used search engine data and tourist emotion data as input variables to improve prediction accuracy. The studies revealed how web-based data from search engines, website traffic and tourist emotion can be shaped into an exogenous variable to better forecast demand. Big data quality management has the potential to improve forecasting accuracy (Bokelmann & Lessmann, 2019; Dergiades et al., 2018), as well. Moreover, the influence of tourism related events on demand, such as data applications generated by 5G technology, high-speed railway construction, policy measures, and thematic tourism (such as agritourism, parent-child travel) are, so far, untapped research directions. Furthermore, some research topics in tourism forecasting, such as the use of mixed methods, have systematically reviewed articles (Khoo-Lattimore, Mura, & Yung, 2019), but there is no scientometric review study, which can be used as a research direction in the future.

A limitation of this study is that its data was retrieved only from the core database of WoS. Although WoS is considered the most authoritative source of data for most publications, some worthwhile literature solely found in other databases may have been overlooked and literature in languages other than English would not have been included, such as Scopus, EBSCO Host (Hospitality and Tourism Complete), ProQuest, Science Direct (Elsevier), Sage and Emerald (Perkins, Khoo-Lattimore, & Arcodia, 2020; Yung & Khoo-Lattimore, 2019). As for the language, the search strategy beyond English language will provide a comprehensive insight in the field of tourism demand forecasting (Yang, Khoo-Lattimore, & Arcodia, 2017b). By taking into account other types of databases and documents, further research can extend the boundaries of this study to integrate a more comprehensive knowledge map for tourism demand forecasting.

Authors' contributions list

Chengyuan Zhang and Shaolong Sun conceived of the presented idea. Chengyuan Zhang developed the integrated analysis framework and performed the experiment. Chengyuan Zhang and Shaolong Sun contributed to the interpretation of the results. Shouyang Wang and Yunjie Wei encouraged Chengyuan Zhang and Shaolong Sun to process the bibliographic records and supervised the findings of this work. Chengyuan Zhang took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript. All authors read and approved the manuscript.

Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgements

This research work was partly supported by the National Natural Science Foundation of China under Grant No. 71801213, No. 71988101, No. 71642006 and by the Fundamental Research Funds for the Central Universities under Grant No. xpt012020022. This research work was also partly supported by the Academic Excellence Foundation of BUAA for PhD Students. The authors would like to express their sincere appreciation to the editor and the referees for their very valuable comments and suggestions. Their comments and suggestions have improved the quality of the paper immensely.

References

- Acedo, F. J., & Casillas, J. C. (2005). Current paradigms in the international management field: an author co-citation analysis. *Int. Bus. Rev.* 14(5), 619–639.
- Athanasopoulos, G., Song, H., & Sun, J. A. (2018). Bagging in tourism demand modeling and forecasting. *J. Travel Res.* 57(1), 52–68.
- Balli, F., Shahzad, S. J. H., & Uddin, G. S. (2018). A tale of two shocks: What do we learn from the impacts of economic policy uncertainties on tourism? *Tour. Manag.* 68, 470–475.
- Bokelmann, B., & Lessmann, S. (2019). Spurious patterns in Google trends data: An analysis of the effects on tourism demand forecasting in Germany. *Tour. Manag.* 75, 1–12.
- Chen, K. (2011). Combining linear and nonlinear model in forecasting tourism demand. *Expert Syst. Appl.* 38(8), 10368–10376.
- Chen, K., & Wang, C. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tour. Manag.* 28(1), 215–226.
- Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. *Tour. Manag.* 24(3), 323–330.
- Claveria, O., Monte, E., & Torra, S. (2015). Common trends in international tourism demand: are they useful to improve tourism predictions? *Tour. Manag. Perspect.* 16, 116–122.
- Colladon, A. F., Guardabascio, B., & Innarella, R. (2019). Using social network and semantic analysis to analyze online travel forums and forecast tourism demand. *Decis. Support. Syst.* 123. <https://doi.org/10.1016/j.dss.2019.113075>.
- Coshall, J., & Charlesworth, R. (2011). A management orientated approach to combination forecasting of tourism demand. *Tour. Manag.* 32(4), 759–769.
- Dergiades, T., Mavragani, E., & Pan, B. (2018). Google trends and tourists' arrivals: Emerging biases and proposed corrections. *Tour. Manag.* 66, 108–120.
- Fang, Y., Yin, J., & Wu, B. (2018). Climate change and tourism: A scientometric analysis using CiteSpace. *J. Sustain. Tour.* 26(1), 108–126.
- Ghoddusi, H., Creamer, G., & Rafizadeh, N. (2019). Machine learning in energy economics and finance: a review. *Energy Econ.* 81, 709–727.
- Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tour. Manag.* 23(5), 499–510.
- Gössling, S., Scott, D., Hall, C. M., Ceron, J. P., & Dubois, G. (2012). Consumer behaviour and demand response of tourists to climate change. *Ann. Tour. Res.* 39(1), 36–58.
- Gunter, U., & Önder, I. (2015). Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data. *Tour. Manag.* 46, 123–135.
- Gunter, U., & Önder, I. (2016). Forecasting city arrivals with Google Analytics. *Ann. Tour. Res.* 61, 199–212.
- Gunter, U., Önder, I., & Smeral, E. (2019). Scientific value of econometric tourism demand studies. *Ann. Tour. Res.* 78. <https://doi.org/10.1016/j.annals.2019.06.005>.
- Hassani, H., Silva, E., Antonakakis, N., Filis, G., & Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Ann. Tour. Res.* 63, 112–127.

- Huang, X., Zhang, L., & Ding, Y. (2017). The Baidu index: Uses in predicting tourism flows—a case study of the Forbidden City. *Tour. Manag.* 58, 301–306.
- Jiang, Y., Ritchie, B., & Benckendorff, P. (2017). Bibliometric visualisation: an application in tourism crisis and disaster management research. *Current Issues in Tourism* doi: 13683500.2017.1408574.
- Khoo-Lattimore, C., Mura, P., & Yung, R. (2019). The time has come: a systematic literature review of mixed methods research in tourism. *Curr. Issue Tour.* 22(13), 1531–1550.
- King, M. A., Abrahams, A. S., & Ragsdale, C. T. (2014). Ensemble methods for advanced skier days prediction. *Expert Syst. Appl.* 41(4), 1176–1188.
- Kuntner, T., & Teichert, T. (2016). The scope of price promotion research: An informetric study. *J. Bus. Res.* 69(8), 2687–2696.
- Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tour. Manag.* 21(4), 331–340.
- Law, R., Li, G., Fong, D., & Han, X. (2019). Tourism demand forecasting: a deep learning approach. *Ann. Tour. Res.* 75, 410–423.
- Lee, C., & Taylor, T. (2005). Critical reflections on the economic impact assessment of a mega-event: The case of 2002 FIFA World Cup. *Tour. Manag.* 26(4), 595–603.
- Lee, C. K., Song, H. J., & Bendle, L. J. (2010). The impact of visa-free entry on outbound tourism: a case study of South Korean travelers visiting Japan. *Tour. Geogr.* 12(2), 302–323.
- Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *J. Am. Soc. Inf. Technol.* 58(9), 1303–1319.
- Li, G., Song, H., & Witt, S. F. (2005). Recent Developments in Econometric Modeling and Forecasting. *Journal of Travel Research*, 44, 82–99.
- Li, G., Song, H., & Witt, S. F. (2006). Time varying parameter and fixed parameter linear AIDS: An application to tourism demand forecasting. *Int. J. Forecast.* 22(1), 57–71.
- Li, G., Wong, K. K., Song, H., & Witt, S. F. (2006). Tourism demand forecasting: A time varying parameter error correction model. *J. Travel Res.* 45(2), 175–185.
- Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tour. Manag.* 68, 301–323.
- Li, Q., Li, S., Zhang, S., Hu, J., & Hu, J. (2019). A review of text corpus-based tourism big data mining. *Appl. Sci.* 9(16), <https://doi.org/10.3390/app9163300>.
- Li, S., Chen, T., Wang, L., & Ming, C. (2018). Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index. *Tour. Manag.* 68, 116–126.
- Li, X., Ma, E., & Qu, H. (2017). Knowledge mapping of hospitality research: A visual analysis using CiteSpace. *Int. J. Hosp. Manag.* 60, 77–93.
- Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tour. Manag.* 59, 57–66.
- Lim, C., & McAleer, M. (2002). Time series forecasts of international travel demand for Australia. *Tour. Manag.* 23(4), 389–396.
- Liu, H., Liu, Y., Wang, Y., & Pan, C. (2018). Hot topics and emerging trends in tourism forecasting research: A scientometric review. *Tour. Econ.* 25(3), 448–468.
- Lv, S. X., Peng, L., & Wang, L. (2018). Stacked autoencoder with echo-state regression for tourism demand forecasting using search query data. *Appl. Soft Comput.* 73, 119–133.
- Newman, M. E. (2005). A measure of betweenness centrality based on random walks. *Soc. Networks*, 27(1), 39–54.
- Olawumi, T., & Chan, D. (2018). A scientometric review of global research on sustainability and sustainable development. *J. Clean. Prod.* 183, 231–250.
- Padhi, S. S., & Pati, R. K. (2017). Quantifying potential tourist behavior in choice of destination using Google trends. *Tour. Manag. Perspect.* 24, 34–47.
- Pai, P., Hung, K., & Lin, K. (2014). Tourism demand forecasting using novel hybrid system. *Expert Syst. Appl.* 41(8), 3691–3702.
- Pan, B., & Yang, Y. (2017). Forecasting destination weekly hotel occupancy with big data. *J. Travel Res.* 56(7), 957–970.
- Perkins, R., Khoo-Lattimore, C., & Arcodia, C. (2020). Understanding the contribution of stakeholder collaboration towards regional destination branding: A systematic narrative literature review. *J. Hosp. Tour. Manag.* 43, 250–258.
- Prideaux, B., Laws, E., & Faulkner, B. (2003). Events in Indonesia: Exploring the limits to formal tourism trends forecasting methods in complex crisis situations. *Tour. Manag.* 24(4), 475–487.
- Rivera, R. (2016). A dynamic linear model to forecast hotel registrations in Puerto Rico using google trends data. *Tour. Manag.* 57, 12–20.
- Shan, J., & Wilson, K. (2001). Causality between trade and tourism: Empirical evidence from China. *Appl. Econ. Lett.* 8(4), 279–283.
- Shen, S., Li, G., & Song, H. (2011). Combination forecasts of international tourism demand. *Ann. Tour. Res.* 38(1), 72–89.
- Si, H., Shi, J., Wu, G., Chen, J., & Zhao, X. (2018). Mapping the bike sharing research published from 2010 to 2018: A scientometric review. *J. Clean. Prod.* 213, 415–427.
- Smeral, E. (2009). The impact of the financial and economic crisis on European tourism. *J. Travel Res.* 48(1), 3–13.
- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tour. Manag.* 29(2), 203–220.
- Song, H., Lin, S., Zhang, X., & Gao, Z. (2010). Global financial/economic crisis and tourist arrival forecasts for Hong Kong. *Asia Pac. J. Tourism Res.* 15(2), 223–242.
- Song, H., Qiu, R. T., & Park, J. (2019). A review of research on tourism demand forecasting. *Ann. Tour. Res.* 75, 338–362.
- Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tour. Manag.* 27(2), 214–224.
- Sun, S., Wei, Y., Tsui, K. L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tour. Manag.* 70, 1–10.
- United Nations World Tourism Organization (UNWTO) (2019). International tourist arrivals reach 1.4 billion two years ahead of forecasts. <http://www2.unwto.org/press-release/2019-01-21/international-tourist-arrivals-reach-14-billion-two-years-ahead-forecasts>.
- Wong, K. K., Song, H., Witt, S. F., & Wu, D. C. (2007). Tourism forecasting: To combine or not to combine? *Tour. Manag.* 28(4), 1068–1078.
- World Tourism Cities Federation (WTCF) (2019). Report on world tourism economy trends. <http://en.wtcf.org.cn/Research/WTCFReports/2019011516448.html>.
- Wu, D. C., Song, H., & Shen, S. (2017). New developments in tourism and hotel demand modeling and forecasting. *Int. J. Contemp. Hosp. Manag.* 29(1), 507–529.
- Yang, E. C. L., Khoo-Lattimore, C., & Arcodia, C. (2017a). A systematic literature review of risk and gender research in tourism. *Tour. Manag.* 58, 89–100.
- Yang, E. C. L., Khoo-Lattimore, C., & Arcodia, C. (2017b). A narrative review of Asian female travellers: Looking into the future through the past. *Curr. Issue Tour.* 20(10), 1008–1027.
- Yang, S., Sui, J., Liu, T., Wu, W., Xu, S., Yin, L., ... Liang, G. (2018). Trends on PM_{2.5} research, 1997-2016: A bibliometric study. *Environ. Sci. Pollut. Res.* 25(13), 12284–12298.
- Yang, X., Pan, B., Evans, J. A., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tour. Manag.* 46, 386–397.
- Yao, Y., Cao, Y., Ding, X., Zhai, J., Liu, J., Luo, Y., ... Zou, K. (2018). A paired neural network model for tourist arrival forecasting. *Expert Syst. Appl.* 114, 588–614.
- Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *Int. J. Hosp. Manag.* 28(1), 180–182.
- Yung, R., & Khoo-Lattimore, C. (2019). New realities: A systematic literature review on virtual reality and augmented reality in tourism research. *Curr. Issue Tour.* 22(17), 2056–2081.



Chengyuan Zhang received the M.S. degree from School of Transportation, Jilin University, Changchun, China, in 2017. He is currently pursuing the Ph.D. degree in Management Science and Engineering at School of Economics and Management, Beihang University, Beijing, China. His research interests include big data analysis, artificial intelligence, tourism management, knowledge management and forecasting. He has published three journal papers in *Applied Energy*, *Tourism Economics* and *Advances in Mechanical Engineering*.



Shouyang Wang received his PhD degree in Operations Research from the Institute of Systems Science, Chinese Academy of Sciences, China, in 1986. He is currently a Bairen Distinguished Professor of Management Science at the Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He has received many research related awards and honors. He has published 35 monographs and published more than 330 papers in international academic journals. He is/was a co-editor of 16 journals and a guest editor of special issues/volumes of more than 15 journals. His research interests include decision analysis, risk management, economic analysis and forecasting.



Shaolong Sun received the Ph.D. degree majoring in Management Science and Engineering from the Institute of Systems Science, Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing, China. He is currently a Professor with the Department of Management Science, School of Management, Xi'an Jiaotong University, Xi'an, China. His research interests include artificial intelligence, big data mining, machine learning, social networks analysis, knowledge management, and economic and financial forecasting. He has authored or coauthored more than 20 papers in journals including *Tourism Management*, *Applied Energy* and *Journal of Environmental Management*.



Yunjie Wei received her PhD degree in Management Science and Engineering at Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China in 2017 and also received a PhD degree in Management Science at City University of Hong Kong in 2018. She is currently an assistant professor at Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China. Her research interests include economic modeling, analysis and forecasting. She has published over 16 papers in journals including *Applied Energy*, and *Tourism Management*.