IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

A survey of research progress and hot front of natural gas load forecasting from technical perspective

HUIBIN ZENG[1], BILIN SHAO[1], GENQING BIAN[1,2], DAN SONG[1], XIAOJUN LI[1]
[1] School of Management, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China
[2] School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China

Corresponding author: Bilin Shao (e-mail: sblin0462@sina.com)

This work was supported by The National Natural Science Foundation of China under Grant 62072363.

ABSTRACT With economic development and scientific and technological progress, people's requirements for the ecological environment are increasing day by day. As a clean energy, natural gas is favored by many countries all over the world. As an important part of natural gas industry planning, load forecasting plays a vital role in the optimal dispatching and operation of the natural gas network. From the perspective of prediction technology, this paper selects the literature related to natural gas load prediction from the Web of Science and CNKI database as the research object. Firstly, the literature chronology distribution and background, research institutions, and academic communities were analyzed by CiteSpace scientific knowledge mapping software. Secondly, the research hotspots and cutting-edge technologies were presented and analyzed visually, and the hot spot migration process in this field was summarized. In the end, the paper puts forward some problems that should be paid attention to in natural gas load prediction, including data processing in prediction, prediction of prediction model, adaptive prediction method and combination prediction method, which are also the direction of development in the future.

INDEX TERMS Natural gas load forecasting, prediction techniques, CiteSpace, knowledge mapping, visual analysis

I. INTRODUCTION

In the context of the transformation of global energy consumption structure, clean and green energy has aroused a wide concern from all over the world. The government has begun to advocate the development of low-carbon economy, and natural gas, a kind of clean energy has drawn attention of different countries and regions. As is known to all, among the three pillars of the world's primary energy supply (coal, oil, natural gas), coal and oil will bring many problems to the environment, while natural gas is not only efficient, but also can make up for shortcomings of wind and solar energy, for instance, difficult storages and unstable supply. [1] So natural gas is widely used in industrial, residential and commercial areas. In the US, natural gas is expected to account for 40 percent of the nation's energy consumption by 2040 and its production is expected to grow by 65 percent, replacing oil as the main fuel. In Russia, where gas will dominate the energy mix with a 50% share, production will rise to 72 BCF/d [2]. In China, the national consumption of natural gas accounted for 3.4% of the total energy consumption in 2008. By 2018, the consumption of natural gas increased to 7.8% [3]. The consumption of natural gas in China has grown rapidly in decade.

Under the rapid development of natural gas, all countries in the world are accelerating the development of green energy. With the progress of technology, the natural gas industry in the world has also entered a new period of development and entered a new stage of energy adjustment. But in fact, in the world energy market, on the one hand, the contradiction between natural gas supplies and demand is gradually intensifying, and the stable development of natural gas market is facing great challenges. On the other hand, natural gas is non-renewable energy which needs careful management and effective development. The common solutions to these problems include a reasonable evaluation and analysis of users' behavior, an effective construction of natural gas pipe network facilities, and an acceptable price, which are actually closely related to natural gas load forecasting.

In the research of natural gas load prediction, the prediction performance is generally affected by various factors, such as data quality, geographical location, climatic conditions, prediction range, market segmentation and prediction source, etc. [4], so the research Angle of natural gas load prediction is also very diverse. In general, although scholars have done a lot of research on natural gas load forecasting around forecasting technologies, models and methods, there is still a lack of research literatures on the scientific metrology of natural gas load forecasting. The way to predict the future natural gas load is closely concerned with a healthy natural gas market. By combing and analyzing the literatures in this field and summarizing the past development rules of this field, researchers can reasonably explore the future development trend of the natural gas industry based on the existing research results. Web of Science (WOS) almost covers the most comprehensive and authoritative scientific literatures in the world. In this paper, relevant literatures in the WOS database are first collected and organized. In the review of relevant literature in WOS, it is found that China has a lot of research achievements in this field. China has always been becoming a big energy country, and Chinese scholars have also kept eyes on gas load forecasting. However, many high-level papers published of Chinese scholars are published on China National Knowledge Infrastructure (CNKI), the largest academic literature database in China, so it is necessary to bring CNKI, into bibliometric analysis.

In conclusion, relying on the WOS and CNKI database, CiteSpace knowledge mapping software, the literature measurement analysis method, the author analyzed the research status of natural gas load forecasting from the Angle of view of the prediction technology. By combing the research progress in this field, this paper aims to clarify the overall development context of natural gas load forecasting and provides a certain reference for international natural gas load forecasting and the development of the natural gas industry. The main work of the thesis is as follows:

(1) Collect and organize the domestic and foreign literatures on natural gas load forecasting. Summarize the annual publication volume and analyze the chronological background.

(2) Perform visual analysis on research institutions and academic communities in this field, so as to reveal the main research institutions in this field, their categories, geographical distribution, and main research teams.

(3) Construct a keyword co-occurrence network map for natural gas load forecasting, and summarize the research hotspots and cutting-edge technologies in the field of natural gas load forecasting based on the visualization of the map.

(4) Combining the Timeline View and TimeZone View, further excavate the hotspot migration process at each stage, and discuss the issues that should be paid attention to in the theoretical research and practical application of natural gas load forecasting.

II. RESEARCH DESIGN

A. DATA SOURCES

In order to have a detailed understanding of the research content and development trend of natural gas load forecasting, the foreign literatures that are selected in this study come from the three major indexes of SCIE, SSCI and A&HCI in the core Collection of WOS. The retrieval methods are “TS = (Natural gas load OR natgas load) AND TS = (prediction OR forecasting)”. The time span is "all years", and the retrieval time is September 30, 2020. 297 papers are retrieved, and 52 papers are finally obtained through data cleaning and discarding papers irrelevant to the research topic. The Chinese literatures are retrieved from CNKI core database, with "natural gas" and "load forecasting" as the theme. The retrieval time is September 30, 2020. 273 papers are searched. Newspapers and papers irrelevant to the research topic are excluded, and 106 papers are finally obtained.

B. RESEARCH METHODS AND TOOLS

In this paper, by adopting the literature research method and the literature metrology the author analyzes literature data through the scientific knowledge mapping tool; and analyzes the research status of natural gas load forecasting at home and abroad by using the quantitative and qualitative analysis method.

The knowledge mapping tool used in this paper is CiteSpace [5] developed by Dr. Chaomei Chen. This software can be used for visual analysis of new trends of literature and it is also widely used in scientific metrology of literature. CiteSpace combines citation analysis and co-citation analysis to transform a literature data from abstract data into visual and intuitive maps of scientific knowledge through information visualization, helping researchers find research hotspots and frontiers in a scientific field.

III. METROLOGICAL CHARACTERISTICS OF NATURAL GAS LOAD FORECASTING

A. AGE DISTRIBUTION AND BACKGROUND ANALYSIS

Through literature research, it can be found that the research on natural gas load forecasting technology is derived from the research on gas load forecasting technology, because natural gas is a kind of gas, the load forecasting of natural gas and the load forecasting of gas have something in common. At first, people used gas such as coal gas and liquefied petroleum gas, while, natural gas are rarely used. After 2000, with the rapid development of economic globalization, the market demand for clean energy continued to grow, economic and environmentally friendly natural gas was gradually favored by the market, and the related research results of natural gas load forecasting gradually increased. Before 2000, the number of literatures on natural gas load prediction was 0. After 2000, relevant literatures gradually increased, which also confirmed the development history of natural gas. According to the annual publication volume of WOS and CNKI, Origin was used to draw the chronology

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3044052, IEEE Access

distribution of domestic and foreign literatures in this field, and the second-order polynomial fitting was carried out, as shown in Figure 1(Note: As of September 30, there were 7 WOS literatures and 3 CNKI literatures in 2020, so the data of 2020 was not included in the forecast range).

.

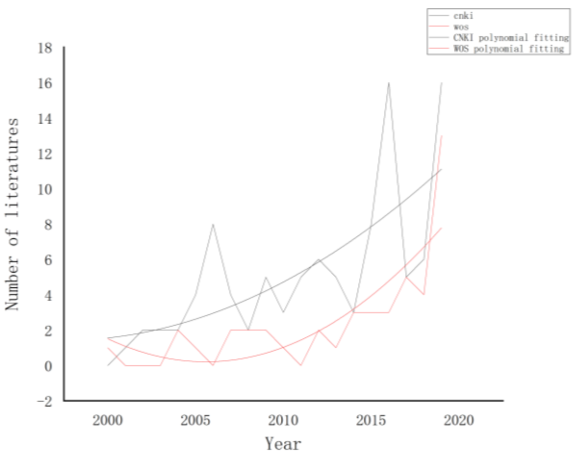

FIGURE 1. Chronological distribution of domestic and foreign literatures

It can be seen from the chronology distribution that the literatures in CNKI mainly focuses on the research of domestic natural gas load forecasting technology, which shows the regular pattern of a stable rising at first and then a dropping down, then it is followed by a stable development and then start gradually rising. Specifically, the number of literatures rose steadily from 2000 to 2006. It started to decline in 2007, while it upswing in 2009. After that, it fell back in 2010. During 2011 and 2013, the development was relatively stable. It rose sharply from 2014 to 2016. In 2017, it pulled back from early highs but managed to stay positive in 2018. Therefore, the rise and fall of the number of literatures have a fluctuation. However, it can be seen from the trend line of second-order polynomial that the number of publications is on the whole rising gradually and will rise steadily in the future. The related literatures of WOS mainly focus on foreign natural gas load forecasting technology. Though it has experienced repeated rise and fall, compared with CNKI, the number of WOS articles is much flatter and increases steadily. The increase in the number of WOS articles from 2018 to 2019 is almost consistent with that of CNKI. It can be seen from the trend line of second-order polynomial that the number of published articles increases year by year on the whole, and it will develop faster in the future.

The quantity of relevant research literatures at home and abroad conforms to the law of logical growth of literature information, that is, literatures will fluctuate lightly after a growth, next it will develop steadily, and finally it will rapidly grow stage after a steady development. Because the research results of natural gas load forecasting gradually increase, this field will develop steadily for a period of time. However, with the development of society and the progress of science and technology, this field is difficult to remain unchanged, and it will encounter some obstacles and new breakthroughs, so the research results of natural gas load forecasting combined with new technologies will increase. According to the number of literatures and the regular patterns of publication, the development of this field at home and abroad can be divided into the embryonic stage (before 2000), the initial stage (2000-2015) and the rapid development stage (2016-present).

1) THE EMBRYONIC STAGE (BEFORE 2000)

Before 2000, the load forecasting of natural gas was still in its beginning stage, and very a few scholars predicted the load of gas by mathematical statistics. In 1964, Zhigang Jin first analyzed and discussed the related problems of gas load in China, and he recommended using mathematical statistics to study gas load [6]. In 1985, Bobrovski of the Soviet Union analyzed and forecasted the consumption of natural gas, and evaluated the forecasting methods such as second-order autoregressive model and exponential smoothing method [7]. These are the documents that have been studied earlier in the field of gas load forecasting through scientific methods and technologies. However, before 2000, the research results of natural gas load forecasting were deficient. Because natural gas industries did not prosper at that time, and natural gas exploitation, transportation and application are underdeveloped.

2) INITIAL STAGE (2000-2015)

In 2000, the reserves and output of natural gas in the world, as well as the international trade volume and consumption increased significantly. The proven reserves of natural gas in the world changed from the stagnant situation in the last century, and increased by 2.57%, with a net increase of 3,743.4 billion cubic meters over the previous year, with both growth rate and increase reaching new highs [8].In China, the construction and application of natural gas is represented by "Shaanxi gas flowing into Beijing and project of natural gas transmission from West to East " indicating that the application of urban gas in China has entered a new era. At the same time, mathematical statistical analysis methods, computer application technology and machine learning algorithm are also in a rapid development stage. Combining the number of articles publication by CNKI and WOS, the number of articles publication in 2000 and 2001 was 0.63%, which was the lowest value in this period, and the number of articles publication in 2015 was 6.92%, which was the highest value in this period. In this stage, the proportion of published articles in 2015 was 6.29 percentage points higher than that in 2000, and the average annual articles publication accounted for 3.14%. Although there is great fluctuation in this stage, the number of articles publication is increasing. The number of articles publication and historical background at this stage means that the research on natural gas load forecasting has officially entered a scientific research stage (from 2000 to 2015).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3044052, IEEE Access

3) RAPID GROWTH STAGE (2016-PRESENT)

2016 and 2017 are the most prosperous two years of global energy transformation [9]. In 2018, the proportion of coal in primary energy consumption reached the lowest since 2004, and natural gas began to attract worldwide attention. It can be predicted that natural gas will become the core force in the energy revolution and the natural gas market will usher in explosive growth. At the same time, the field of natural gas load forecasting will also be developed rapidly, because in the whole industrial chain of natural gas production, storage, transportation and application, natural gas load forecasting is throughout and indispensable. Combined with the number of articles publication by CNKI and WOS, the number of articles publication in 2016 is 12.58%, the number of articles publication in 2017 and 2018 is 6.29%, the number of articles publication in 2019 is 18.24%, and the average annual number of articles publication is 10.85%. In 2016, the number of articles publication increased sharply. Although it declined in 2017, the number of published articles in 2019 increased to a new high, which was the highest among all years. The average number of articles publication in this stage is 7.71 percentage higher than that of in the initial stage. It can be seen that it grows rapidly. Based on the number of publications and historical background of this stage, it can be defined as the rapid development period of natural gas load forecasting (2016 to present).

B. ANALYSIS OF RESEARCH

In this paper, the visual presentation and analysis of research institutions through CiteSpace can help focus on the major research institutions in this field, discover the research priorities of domestic and foreign academic institutions in this field, and help clarify the frontier development of this field. In CiteSpace, the time partition unit is set as one year, and the period from 2000 to 2020 is divided into 21 time partitions, and the Node Types are selected as institutions. Thus, the cooperative network diagram of WOS research institutions (as shown in Figure 2) and CNKI research institutions (as shown in Figure 3) were obtained.

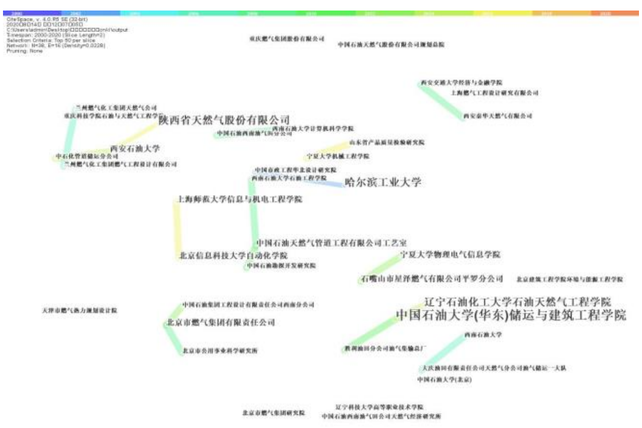

FIGURE 2. **collaboration network diagram of WOS research institutions**

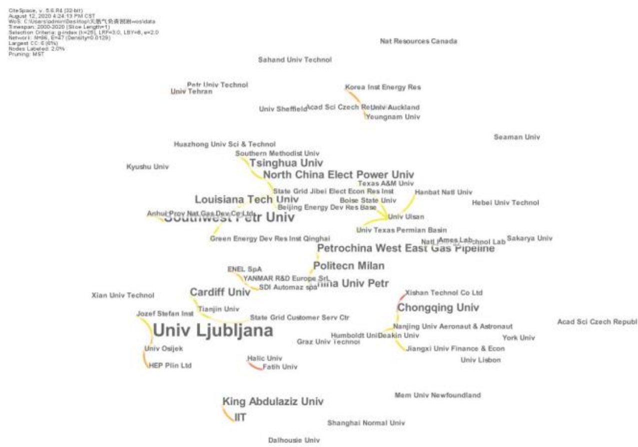

FIGURE 3. **collaboration network diagram of CNKI research institutions**

1) MAIN RESEARCH INSTITUTIONS

According to the collaboration network diagram of research institutions by WOS and CNKI, research institutions with CITATION COUNTS greater than 2 were listed as the main research institutions in this field, as shown in Table I. The node centrality and the sudden increase in the number of articles publication by these research institutions are all 0, so it is impossible to make further analysis. However, on the whole, Ljubljana University, China University of Petroleum (East China) and Southwest Petroleum University are in the leading position in the field of natural gas load forecasting. It also shows that the research focus of natural gas load forecasting depends on the foundation of higher education institutions and petroleum universities.

TABLE I
MAIN RESEARCH INSTITUTIONS

INSTITUTIONS	CITATION COUNTS	YEAR
University of Ljubljana	5	2007
China University of Petroleum (East China)	4	2011
Southwest Petroleum University	4	2007
Harbin Institute of Technology	3	2002
Shaanxi Natural Gas Company Limited	3	2013
Liaoning Shihua University	3	2014
Shanghai Normal University	3	2016
King Abdulaziz University	2	2015
Illinois Institute of Technology	2	2015
PetroChina West-East Gas Pipeline Company	2	2019
Louisiana Tech University	2	2019
Tsinghua University	2	2018
Chongqing University	2	2007
China University of Petroleum (Beijing)	2	2012
North China Electric Power University	2	2019
China National Petroleum Pipeline Engineering Company	2	2009
Beijing Information Science And Technology University	2	2016
Ningxia University	2	2012
Beijing Gas Group Company Limited	2	2009
Xi'an Shiyou University	2	2014

2) CATEGORIES OF RESEARCH INSTITUTIONS

In this paper, the research institutions are classified into four categories: comprehensive universities, petroleum universities, companies, and research institutes. The source institutions of the sample literature (117 names in total) are classified and counted, as shown in Figure 4. According to the data, there are 67 comprehensive universities (57%), 8 petroleum universities (7%), 33 companies (28%) and 9 research institutes (8%). The results show that the publishing units of sample literature mainly come from universities (accounting for 64% of the total), followed by companies, and research institutes. The reason is that universities have the academic platforms, human resources, capital investment, advanced equipment and technology needed for research. What's more they get the support from our government. However, for companies, their main goal is production and operation, therefore they need to raise their own funds and set up their own research teams to carry out research, in this case, their research ability in this field is weak relatively. Besides, some research institutes of gas and petroleum are devoted to the research of natural gas load forecasting, but the number of these research institutes is limited, so the proportion is low.

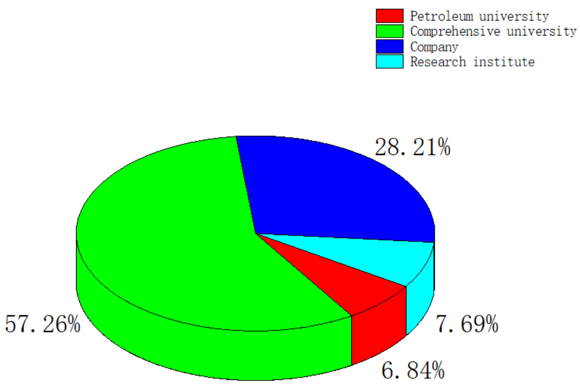

FIGURE 4. Category of research institutions

3) GEOGRAPHICAL DISTRIBUTION OF RESEARCH INSTITUTIONS

Here is the statistical analysis about publishing institutions (81 research institutions in total) based on different countries in WOS, as shown in Table II. The top 7 countries are China, the United States, Italy, Iran, South Korea, Canada and Turkey. China is the country with a large quality of literatures, because we have huge demand for natural gas in China. While, the United States ranks the second, the United States is also a big energy consumption country, and its demand for natural gas is also very high. The United States is a pioneer country in the world to develop and research the natural gas industry. However, with China's growing strength, China's investment in the natural gas industry has gradually increased. More and more Chinese research institutions have invested in natural gas-related research, and the number of results related to natural gas load forecasting has surpassed that of in the United States. It is followed by Italy, Iran, South Korea, Canada, Turkey and other countries. Italy, Iran, South Korea, Canada, Turkey and other countries, some of which are developed countries. Their development is inseparable inevitably linked to natural gas. Some countries are big energy providers, accordingly, there is no shortage of research on natural gas.

TABLE II
COUNTRIES OF RESEARCH INSTITUTIONS

Country	Institutional proportion
China	27.16%
the United States	12.35%
Italy	9.88%
Iran	6.17%
Korea	4.94%
Canada	4.94%
Turkey	4.94%
France	3.70%
the United Kingdom	3.70%
Germany	2.47%
Czech Republic	2.47%
Croatia	2.47%
Japan	2.47%
Slovenia	2.47%

China has the most powerful research ability in this field, so it is necessary to further analyze the distribution law of research related to natural gas load forecasting in China. The regions of China are divided into seven regions: Northeast China, North China, East China, Central China, South China, Southwest China and Northwest China. The research institutions in CNKI literature samples (36 research institutions in total) are classified and counted by region, as shown in Figure 5. In the order from top to bottom, north China accounts for 30%, northwest China 25%, southwest China 17%, East China 17%, Northeast China 11%, South China and central China all equal 0. The results show that the development level of related research on natural gas load forecasting is uneven in the whole country. Few research institutions in South China and Central China are engaged in this research, and the research institutions mainly come from areas with a rich abundance in natural gas resources or areas with rapid economic development. Further analysis of the distribution of cities in various regions shows that the distribution of cities basically conforms to the above laws, that is, they are mainly concentrated in economically developed capitals (such as Beijing) or provinces rich in natural gas resources (such as Shaanxi and Sichuan).

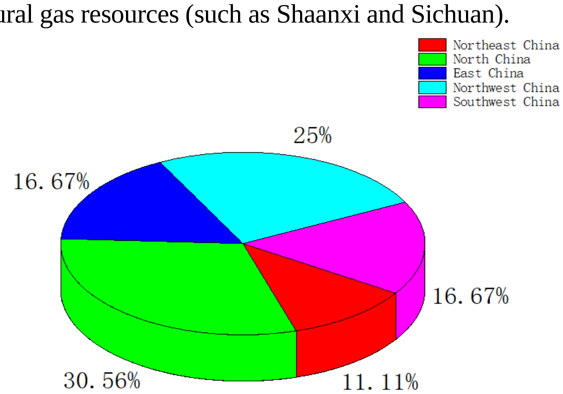

FIGURE 5. Region of domestic research institutions

C. DISTRIBUTION OF ACADEMIC COMMUNITY

Based on CiteSpace's author cooperation network map, this paper further analyzes the main researchers and their teams in this field at home and abroad, which is helpful to grasp the research achievements and development trends in this field and track the latest research achievements and trends of the main research teams in this field in time. Therefore, the sample data of WOS and CNKI are visually presented through CiteSpace. Figure 6 is the author's cooperation network of WOS, and Figure 7 is the author's cooperation network of CNKI. The size of nodes in the network reflects the number of articles publication by authors, and the edges between nodes reflect the cooperative relationship among authors. The more prominent the nodes, the more articles publication will come out. According to the map of the author's cooperation network, we can see that there are many scholars engaged in natural gas load forecasting research in the world, but the author's cooperation network is dispersed, which shows that the research strength and resources of natural gas load forecasting in the world are relatively dispersed.

Furthermore, the author's cooperation network is analyzed through literature retrieval. Although there are academic cooperation groups across regions and institutions, research institutions have very average weights, which indicates that the strength of each research institution is dispersed in this field. Even though the cooperation among scholars is generally carried out in the same region and among the same institutions, a relatively large research group has not been built yet. However, it can still be seen that there are some academic teams at home and abroad who have done more research on natural gas load forecasting. Combined with the above analysis of research institutions, some developed countries such as the United States started earlier in this field. Although China started relatively late, China's natural gas industry has developed rapidly. Therefore, a large number of Chinese scholars have published influential articles in international excellent journals.

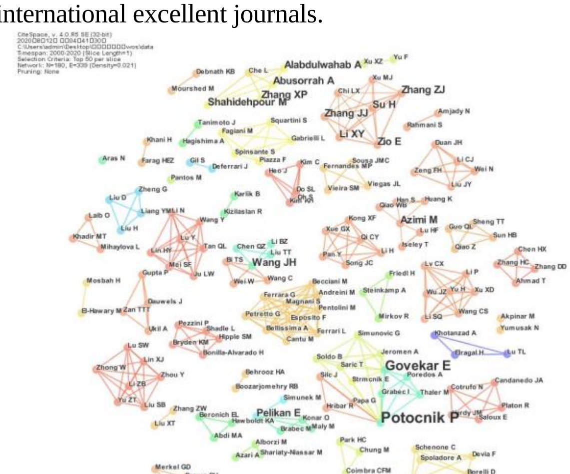

FIGURE 6. WOS author cooperation network map

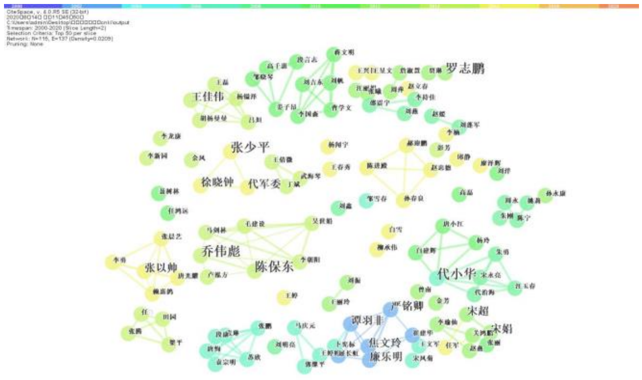

FIGURE 7. CNKI author cooperation network map

1) INTERNATIONAL ACADEMIC COMMUNITY

By analyzing the author's cooperation network of WOS, this paper analyzes four international academic communities and their members in this field.

Firstly, the team from University of Ljubljana in Slovenia, which is under the leadership of P. Potocnik. There are a large number of team members, including E. govekar; B. Soldo; G. Simunovic; A. Jeromen; T. Saric; J. Slic; R. Hribar; G. Papa and others, among which P. Potocnik and E. Govekar have a close cooperation. The research of this team focuses on natural gas load forecasting [10], natural gas consumption forecasting [11], natural gas risk forecasting [12] and natural gas demand forecasting [13], etc. The main methods used are neural network model, support vector machine and linear regression model. In addition, the team also has some research achievements in building temperature prediction [14], heat energy prediction [15] and highway traffic flow prediction [16], which can be used for reference in natural gas load forecasting.

Secondly, Hongfang Lu and M. Azimi's team with K Huang; S Han, etc. also makes achievements. They come from different regions and institutions, such as Louisiana tech university, Southwest Petroleum University, Purdue University, etc., which shows that these universities have close exchanges in the field of natural gas load forecasting. The team applied the nonlinear extension of the improved Arps decline model based on kernel function to forecast natural gas consumption in the United States [17], combined the Fruit fly optimization algorithm with the support vector machine [41], and applied the combined forecasting model to forecast natural gas load. Besides, the team also applied forecasting models to make predictions in areas such as carbon trading volume and power plant [18] [19].

Thirdly, Xiaping. Zhang; M. Shahidehpour, A. Alabdulwahab, and others are another group of academic communities. The team mainly comes from Illinois Institute of Technology and King Abdulaziz University, which are located in Chicago and Saudi Arabia respectively. The team mainly uses data for modeling and simulation, and carries out deployment planning and load forecasting around multiple energy sources such as electricity and natural gas [20] [21] [22].

Finally, the main members of another group of academic communities are Huai Su; E. Zio； Jinjun Zhang and other scholars. They come from China University of Petroleum, Milan Polytechnic University, Paris Sacre University and other institutions. The research mainly focuses natural gas load forecasting, multi-objective optimization and decision-making [23][24][25]. The team uses wavelet transform, deep neural network, data-driven demand side management method and other methods and technologies to realize natural gas load forecasting.

2) CHINESE ACADEMIC COMMUNITY

China has the most powerful research ability in this field at present, it is necessary to analyze the main research teams in China. This paper analyzes the author's cooperation network of CNKI, and finds out four domestic academic communities and their members in this field.

First of all, Xiaozhong Xu of Shanghai Normal University and his graduate students team have systematically studied natural gas load forecasting to forecast natural gas load and optimize and improve it, by using rough set, BP neural network, support vector machine, wavelet transform, artificial bee colony algorithm, cuckoo search, frog leaping algorithm and other methods and models [26] [27]. And they verified the effectiveness of these methods and models. This team spends a lot of time with in-depth research content and scope on this field.

Second of all, the academic team under the leadership Juan Song, with member Chao Song having the closest contact with them. They are mainly from Ningxia University. In order to predict the natural gas load [28] [29], their research is based on artificial intelligence methods such as genetic algorithm, BP neural network, Elman neural network, Fruit fly algorithm, Support vector machine, multiple linear regression, etc.

Thirdly, Baodong Chen and his doctoral student Weibiao Qiao are from China University of Petroleum (East China). The research on natural gas load prediction mainly revolves models based on wavelet theory and RBF (radial basis function)-Elman neural network, based on chaos theory and Volterra adaptive filter, based on Haar wavelet transform and ARIMA-RBF, based on wavelet transform and LSSVM-DE [30][31][32].

Finally, academic team under the leadership of Wenling Jiao, Mingqing Yan, Yufei Tan and Leming Lian [33][34]have an in-depth research on natural gas load forecasting. Their members mainly come from Harbin Institute of Technology and china municipal engineering North China Design and Research Institute, and they are pioneers in China to forecast and study gas load.

IV. RESEARCH HOTSPOTS AND CUTTING-EDGE TECHNOLOGIES

A. REAEARCH HOTSPOTS

In CiteSpace, "Keyword" is used as the node type for clustering, and a keyword co-occurrence network with high frequency and high centrality is drawn. Figure 8 is WOS keyword co-occurrence network, and Figure 9 is CNKI keyword co-occurrence network. The greater the frequency of keyword co-occurrence is, the larger the circle of nodes in the graph is. High-frequency keywords can reflect the research hotspots in a certain period. The analysis of keywords with high frequency and centrality is helpful to grasp the research hotspots in this field. The higher the frequency, the more likely it is to become a research hotspot in this field. The higher the value of centrality, the stronger the central position of the keyword in the co-occurrence network.

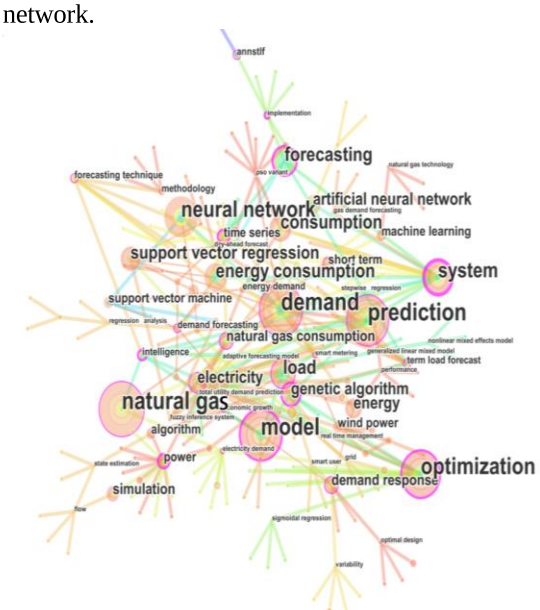

FIGURE 8. WOS keyword co-occurrence network

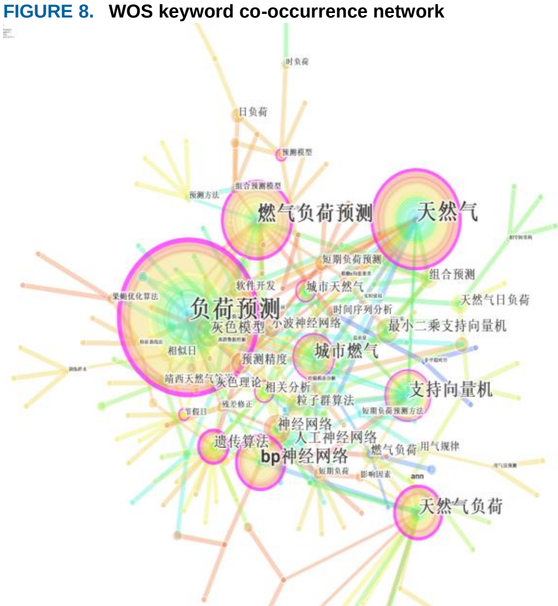

FIGURE 9. CNKI keyword co-occurrence network

TABLE III
FREQUENCY AND CENTRALITY OF HIGH FREQUENCY KEYWORDS

FREQUENCY	CENTRALITY	KEYWORD	FREQUENCY	CENTRALITY	KEYWORD
13	0.28	demand	10	0.13	SVM
12	0.3	model	10	0.08	BP neural network
11	0.14	optimization	9	0.08	city gas
10	0.11	neural network	6	0.16	natural gas daily load
7	0.14	system	6	0.28	genetic algorithm
7	0.11	energy consumption	6	0.04	grey model
7	0.08	consumption	6	0.05	artificial neural network
5	0.07	artificial neural network	5	0.03	least squares support vector machines
5	0.06	genetic algorithm	4	0.09	combination forecast
4	0.21	demand response	4	0.16	correlation analysis
3	0.05	support vector regression	4	0.03	particle swarm algorithm
3	0.02	classification	3	0.05	prediction accuracy
3	0.04	algorithm	3	0.04	grey theory
3	0.01	support vector machine	3	0.02	Fruit fly optimization algorithm
3	0.01	simulation	3	0.02	GM (1, 1)
3	0.02	regression	3	0.06	time series analysis
2	0.32	intelligence	3	0.06	similar day
2	0.09	demand forecasting	2	0.06	gas law
2	0.02	economic dispatch	2	0.02	ann
2	0.05	deep learning	2	0.01	empirical mode decomposition
2	0.01	short-term	2	0.01	short term load forecast
2	0.08	forecasting technique	2	0.01	The neural network
2	0.03	district heating	2	0.01	The holiday season
2	0.04	district	2	0.01	grey prediction method
2	0.01	clustering	2	0	medium and long term load
2	0.03	annstlf	2	0	exponential smoothing model
2	0.02	time series	2	0	software development
2	0	demand side management	2	0	Mathematical model

In the research samples in WOS and CNKI, keywords with co-occurrence frequency higher than 2 (excluding keywords related to search topics) are arranged, as shown in Table III (Note: WOS data in the left column and CNKI data in the right column). By analyzing the keywords with high frequency and high central value in the table, the research hot keywords of natural gas load forecasting can be classified into three categories:①common forecasting techniques of natural gas load forecasting, including neural network, deep learning, support vector machine, time series, regression analysis method, grey theory, genetic algorithm, particle swarm optimization, etc.; ②Classification of natural gas load forecasting can be divided into short-term load forecasting (daily forecasting, including normal days and holidays), medium-term load forecasting (monthly forecasting) and long-term load forecasting (annual forecasting); ③Forecast of natural gas load in actual area, for instance, forecast of the natural gas load of a certain city, prefecture-level city, provincial level and other areas.

B. CUTTING-EDGE TECHNOLOGY

From Figure 8, Figure 9 and Table 3, it can be found that the most frequent keyword hotspots are technologies and methods, which are the basis for carrying out natural gas load forecasting, and may also become the frontier force to promote the development of this field in a certain period of time. Combined with the above research contents, this paper further analyzes the literature contents, and expounds and compares the main technologies of natural gas load forecasting, so as to grasp the research frontier in this field.

1) PREDICTION TECHNOLOGY

Natural gas load forecasting technology is mainly divided into two types of methods, namely traditional forecasting method and artificial intelligence method. Traditional forecasting methods are mainly based on the self-rule of load series, and research is carried out with mathematical thinking and methods, which are often used for short-term and medium-term natural gas load forecasting. Traditional forecasting methods mainly include time series method, regression analysis method and grey theory forecasting method. Artificial intelligence methods have developed with the rise of computer technology and machine learning. Artificial intelligence methods mainly include support vector machine, neural networks, deep learning and so on. This paper mainly analyzes and compares the six main forecasting techniques in the field of natural gas load forecasting, as shown in Table IV.

① Time series method

Time series method is an early mathematical statistical method used in the field of short-term load forecasting of natural gas. In natural gas load forecasting, the establishment of time series model does not need a large amount of historical data, and precise for this date is not that high. it is simple and convenient to operate in short-term load forecasting. However, due to the simplicity of the data, it is impossible to reasonably predict the changing load, and it is difficult to consider the influencing factors of natural gas load change in the process of establishing the forecasting model, so the forecasting model lacks adaptability with load change. Literature [16], [17],[18], [31] is based on time series method to forecast natural gas load.

② Regression analysis method

In natural gas load forecasting, the influential factors of load change should be taken as independent variables and natural gas load as dependent variables. The relationship between independent variables and dependent variables should be explored and used it to build a model, and then the load value can be inferred. Regression analysis method has low requirements on historical data, simple model construction and fast forecasting speed, and it is mostly used for medium and long-term load forecasting. However, this method is difficult to consider various influential factors comprehensively, so it is difficult to select independent variables, and the relationship between variables always changes with the change of internal and external environment, so it lacks self-learning ability. Literature [25],[32] uses regression analysis method to forecast natural gas load scientifically.

③ Grey theory prediction method

In natural gas load forecasting, there are too many factors affecting natural gas load. The relationship among the factors is also very complex, and it is difficult to analyze them one by one, which shows the characteristics of grey theory. Although the appearance of prediction system is fuzzy, the law can be found with the help of suitable models. Therefore, the grey theory forecasting method can effectively predict the natural gas load. Literature [26],[27] uses the grey theory forecasting method to forecast the natural gas load, and it studies and judges the forecasting accuracy.

④ Support vector machine

In the process of natural gas load forecasting, support vector machine (SVM) has some advantages, for example, high running speed, small parameter adjustment range, high fitting accuracy and better classification effect for small sample data sets. Therefore, compared with the traditional forecasting method, the short-term load forecasting accuracy of natural gas by using support vector machine is higher. However, the support vector machine also has the limitations, for instance, high dependence on samples and dependence on model parameters and kernel parameters [35]. Literature [23],[24], [33],[34] uses the hybrid forecasting method based on support vector machine to forecast natural gas. The results show that the forecasting accuracy is more accurate than the traditional forecasting method.

⑤ Neural network

In natural gas load prediction, the neural network algorithm can reasonably deal with the nonlinear correspondence between influencing factors such as weather and temperature and natural gas load, and it can effectively deal with the process and system that are difficult to describe in load prediction. However, neural network algorithm has high requirements on the quantity and quality of sample data, and it is easy to fall into the phenomenon of insufficient learning or over-fitting. Literature [19],[20],[21],[22],[29],[35],[36] uses neural network, the main forecasting method, and it combines other algorithms to optimize and improve it. The results show that the forecasting accuracy is precise.

⑥ Deep learning

In natural gas load forecasting, because the factors affecting natural gas load are complex and changeable, the application of deep learning modeling has stronger self-adaptive ability and self-learning ability, which can help load forecasting better cope with the changes such as environment, temperature and regions. Because there are differences between deep learning and traditional neural networks. Traditional neural networks generally only have two or three layers of neural networks, with limited parameters and calculation units, limited expression ability of complex functions and limited self-learning ability, while deep learning usually has five to ten layers or even more neural networks, and more effective algorithms are introduced. Literature [37],[38] proves that deep learning has a very accurate prediction accuracy in short-term load forecasting of natural gas.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3044052, IEEE Access

TABLE IV
COMPARISON OF FORECASTING TECHNIQUES

Specific technology	CHARACTERISTICS	Defect
Regression analysis method	The requirement for historical data is not high, the model construction is simple; the forecasting speed is fast, and it is suitable for medium and long-term load forecasting;	It is difficult to comprehensively consider various influential factors, the selection of independent variables is difficult, and self-study ability is deficient.
Grey theory prediction method	The requirement for sample data is not high; the calculation is simple, and it is suitable for long-term load forecasting.	The prediction accuracy is not high; The application has great limitations;
Support vector machine	It has a strong forecasting ability, fast convergence speed and global optimization, and is suitable for short-term load forecasting;	It is dependent on sample data and model parameters and kernel parameters.
Neural network	It has a good adaptability, self-organization, fault tolerance, strong learning ability and nonlinear mapping ability, and is suitable for short-term load forecasting.	It puts too much emphasis on overcoming learning errors, weak generalization ability, high demand for sample data and large amount of calculation.
Deep learning	It has a strong self-adaptive ability and self-learning ability, which is suitable for short-term and ultra-short-term load forecasting.	There are high requirements on the quantity and quality of data, large amount of calculation, complicated calculation process, and easy to fall into the phenomenon of insufficient learning or over-fitting.

TABLE V
COMPARISON OF OPTIMIZATION TECHNOLOGIES

Specific technology	CHARACTERISTICS	Defect
Genetic algorithm	Strong adaptability and good compatibility.	It takes a lot of calculation and time.
Particle swarm optimization algorithm	Simple, easy to realize and fast in convergence.	Premature convergence.
Fruit fly optimization algorithm	The algorithm is simple, the code is easy to realize and the calculation efficiency is high.	Premature convergence, slow optimization in the late evolution.

In the research of forecasting methods, there are not only the methods listed in this paper, but also fuzzy theory, expert system method, Kalman filter analysis, cluster analysis, wavelet analysis, data mining and other methods and technologies. These methods have been widely used in power system load forecasting. However, in the process of analyzing the co-occurrence maps of keyword hot spots, it is found that these methods are rarely applied in natural gas load forecasting, so I will not make too many statements here. However, these mature methods and technologies in power load forecasting still have reference significance for natural gas load forecasting. Therefore, when forecasting the natural gas load, we should choose an appropriate method instead of purely pursuing the technology of "new method and new concept".

2) OPTIMIZATION TECHNOLOGY

There are many commonly used neural network optimization algorithms, such as whale optimization algorithm, firefly algorithm, frog leaping algorithm, butterfly optimization algorithm and ant colony optimization algorithm. According to the keyword co-occurrence network map, this paper analyzes the characteristics and defects of three commonly used optimization techniques for natural gas load forecasting, as shown in Table V.

① Genetic algorithm

Genetic algorithm has high parallelism and strong adaptability, and has good compatibility with other algorithms, so it is easy to combine with other algorithms to form a problem-solving method with a better performance [36]. Therefore, in natural gas load forecasting, genetic algorithm is used to optimize other algorithms and make full use of the global search ability of the optimized genetic algorithm, so as to improve the forecasting efficiency and maximize the performance of the forecasting model. However, there are some limitations, such as a large amount of calculation and time-consuming. Literature [29],[35] optimizes the neural network with the help of genetic algorithm, as a result, the prediction accuracy is excellent.

② Particle swarm optimization

Particle Swarm Optimization (PSO) is much simpler and easier, and easy to implement and fast in convergence, but it has the problem of premature convergence. Therefore, many scholars suggest improving strategies to optimize the performance of PSO based on the algorithm parameter adjustment strategy. Therefore, in natural gas load forecasting, the improved PSO can greatly improve the forecasting accuracy. The research results of references [30],[36] show that particle swarm optimization algorithm can effectively improve the prediction accuracy of natural gas load.

③ Fruit fly optimization algorithm

The optimization equation of fruit fly optimization algorithm (FOA) is a first-order differential equation. Compared with particle swarm optimization algorithm, the algorithm is simpler, the program code is easy to implement and the calculation is efficient. In natural gas load forecasting, fruit fly optimization algorithm (FOA) can improve the generalization ability and forecasting accuracy of the model, so optimizing the parameters in the forecasting model with fruit fly optimization algorithm (FOA) can promote an optimal model performance. References [33],[34] use fruit fly optimization algorithm (FOA) to improve the parameters of support vector machine, and the results show that the prediction accuracy is greatly improved.

V. HOT SPOT MIGRATION

In this paper, CiteSpace is used to draw Timeline View for WOS sample documents and TimeZone View for CNKI sample documents. Timeline View and TimeZone View focus on delineating the relationship between clusters and the historical span of documents in a cluster, which can intuitively see the migration process of hot keywords in natural gas load forecasting research and the development context of this field. As shown in Figure 10, it is the Timeline View of WOS; As shown in Figure 11, it is the TimeZone View of CNKI. Combined with the above division of natural gas load forecasting stages (due to the lack of research results of natural gas load forecasting in the embryonic stage, the author will explain other two stages). The evolution and development process of each stage will be set forth in the next part, and the possible research trends in the field of natural gas load forecasting will be put forward in the next part.

A. INITIAL STAGE

It can be seen from Figure 10 and Figure 11 that the artificial neural network (Ann) is the first method to forecast natural gas load in the world, and China is the first country to use time series method to forecast natural gas load. From 2000 to 2015, a large number of forecasting methods and technologies emerged. By using literatures, it was found that most of these forecasting technologies tend to be theoretical. This paper will analyze and interpret the development of natural gas load forecasting in this stage.

Traditional forecasting methods mainly include time series method, regression analysis method and grey theory forecasting method. Among them, time series method is regarded as the most classical, systematic, mature and widely used forecasting method [37]. Time series method was first used in power system load forecasting. With the continuous improvement of this method, it has also been used in short-term load forecasting of natural gas. There are abundant research results on time series method in China. Wenling Jiao et al. [38] established AR time series model for short-term load forecasting of urban natural gas by using least square method based on autocorrelation function, and made an example prediction calculation, which achieved a notable result. Pengjun Liu and Yuan Zhao [39] used the improved chaotic time series to forecast the hourly flow of natural gas load in Xi 'an in winter. Weibiao Qiao and Baodong Chen [31] established a combined forecasting model based on time series analysis, and they used the combined forecasting model of natural gas hourly load based on Haar wavelet transform and ARIMA-RBF to forecast natural gas load. The example results show that its forecasting accuracy is significantly higher than that of the combined forecasting model of self-organizing feature mapping network and multilayer perceptron network (SOFM+MLP). Time series method is often used in short-term load forecasting, and regression analysis method is often used in medium and long-term load forecasting of natural gas. For example, Hong Liu and Yanshuang Zou [40] select the parameters that have great an influence on the load (such as temperature and humidity) as the variable parameters of regression analysis to establish a regression analysis model, and they predict the load by regression analysis. According to the characteristics of the diversity of influencing factors of gas load, they conclude that the combination of piecewise regression forecasting, neural network forecasting and various forecasting methods will achieve more satisfactory forecasting results. Grey theory also can be used in medium and long-term, except for short-term load forecasting of Grey theory forecasting method can be used not only for short-term load forecasting of natural gas, but also for medium-and long-term load forecasting. WANG J-H et al. [41] established a high-precision forecasting model GM (1,1), aiming at overcoming the problems of large population and large fluctuation of urbanization demand in long-term load forecasting. The forecasting results verify that the forecasting method based on grey theory has high accuracy and effectiveness. Jiawei Wang et al. [42] borrowed Matlab to model and simulate the gas load forecasting model based on grey theory. The practical results show that this method can effectively and accurately predict the gas load value. Using grey theory for model construction requires less sample data and it is easy to operate. The accuracy of short-term load forecasting for natural gas will be higher than that of medium and long-term load forecasting.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3044052, IEEE Access

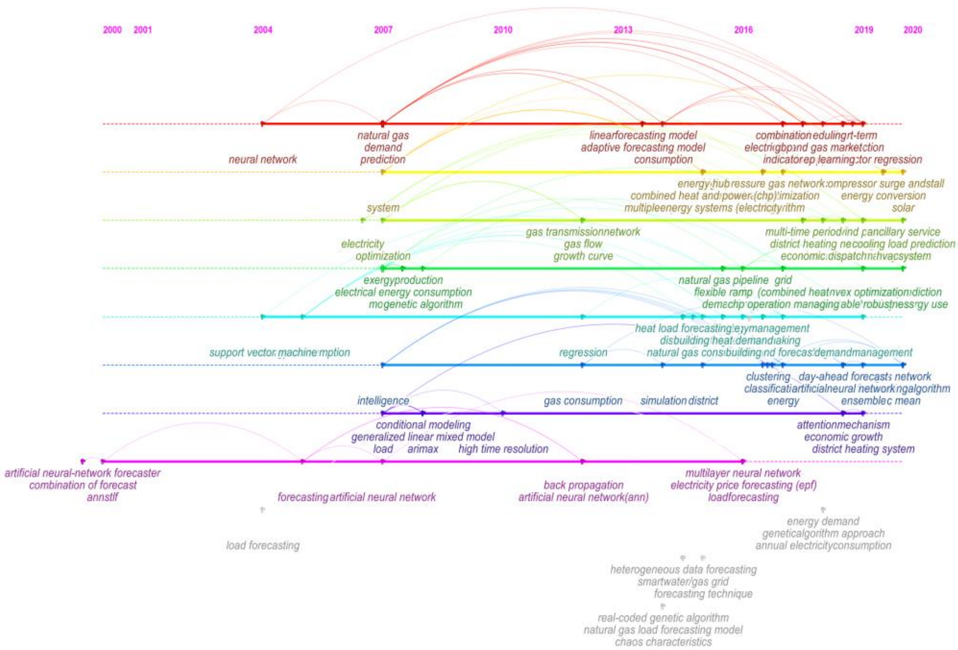

FIGURE 10. Timeline View of WOS

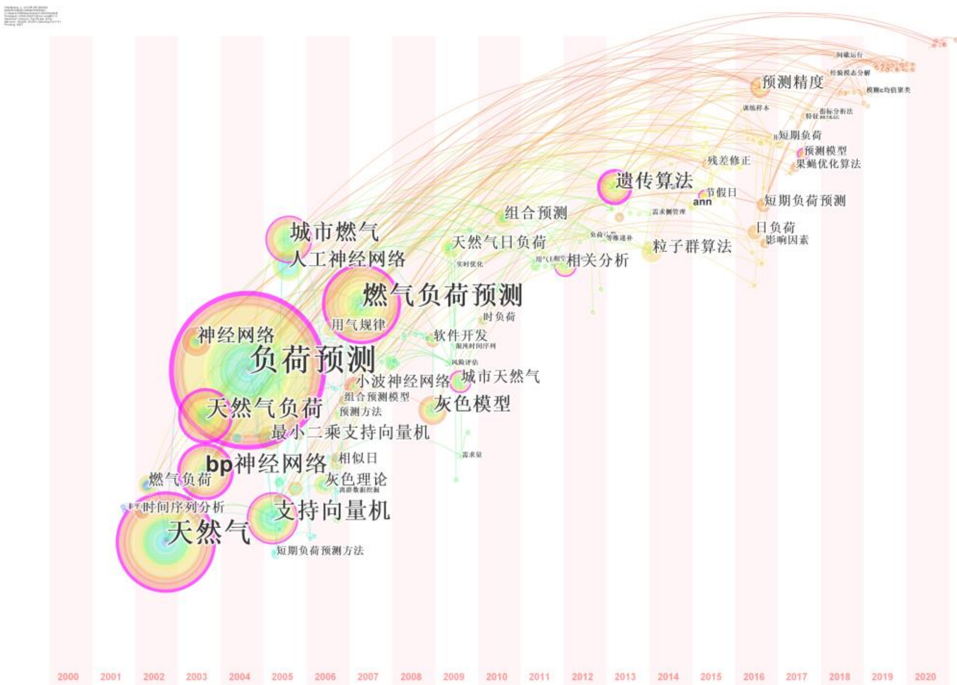

FIGURE 11. TimeZone View of CNKI

Among artificial intelligence prediction methods, people have studied on neural network in the earlier time. In 2000, American scholars A. Khotanzad and H. Elragal [43] applied the artificial neural network (ANN) forecasting technology to forecast the daily natural gas consumption of natural gas enterprises. Seven different combination algorithms based on average method, recursive least square method, fuzzy logic, feedforward neural network, function chain neural network, Karmarkar linear programming algorithm and modular neural network were under research. The results show that the combination strategy of single neural network is superior to other methods. Since then, neural network has become one of the mainstream methods of natural gas load forecasting, and many research literatures have emerged. Chinese scholar Yang Zhao et al. [44] put forward a natural gas load forecasting model based on artificial neural network, and simulated it with Visual C++ and Matlab software. The simulation results show that the method is accurate. R.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3044052, IEEE Access

Kizilaslan, B. Karlik [45] used different neural network algorithm models to forecast the natural gas load in residential and commercial districts of Istanbul, Turkey. The training and test results show that the neural network model based on conjugate gradient descent algorithm is more effective and accurate. A. Azari et al. [46] predicted the short-term and medium-term demand load of natural gas by using trained neural network, and the predicted results conform to the actual results. They suggested that this method can be further applied to predict information in various areas. With the increasing application of machine learning technology in load forecasting, a large number of scholars have studied support vector machine (SVM) regression technology based on statistical learning theory. Support vector machine (SVM) is based on the principle of structural risk minimization, rather than the principle of empirical risk minimization in traditional regression technology. The improved support vector machine is more accurate. Support vector machine (SVM), a new regression algorithm based on statistical learning theory, is one of the early technologies to predict natural gas load by using machine learning technology in this field. The prediction results show that the prediction accuracy of SVM is high [47]. P. Potocnik, B. Soldo [48] studied the short-term load forecasting of natural gas in residential areas of Croatia, and he established linear model, neural network model and support vector regression model Next, he put the models into static version and adaptive version to compare the forecasting performance. The research results show that the self-adaptive model is used to forecast in this area, and the performance is significantly improved.

In order to get a better accuracy of natural gas load forecasting, many scholars have also improved the parameters through some optimization algorithms Genetic algorithm, a robust random search algorithm for solving various forecasting problems, has many remarkable advantages compared with traditional methods. In addition, it has drawn a wide attention in recent years [49]. In order to further improve the prediction accuracy of natural gas load, genetic algorithm is used to optimize the traditional algorithm. Chao Song et al. [50] proposed a natural gas load forecasting method based on genetic algorithm optimization and BP neural network, and the simulation results showed that the accuracy of natural gas load forecasting was improved. Lijuan Wang [51] uses particle swarm optimization algorithm (PSO) to optimize the parameters of weighted least squares support vector regression (WLS-SVR), so as to obtain the prediction model based on PSO-WLS-SVR. The prediction models based on support vector regression (SVR), WLS-SVR and PSO-WLS-SVR are simulated and compared respectively. The results show that PSO can effectively optimize the parameters, and the prediction model based on PSO-WLS-SVR has high prediction accuracy.

Although the number of published papers on natural gas load forecasting fluctuated greatly during this period, many forecasting techniques and methods emerged. They has laid a good theoretical foundation for the next development in this field.

B. RAPID DEVELOPMENT STAGE

The rapid development stage is also a breakthrough period. In this period, the number of articles publication in this area is far higher than that of in the previous period, in addition the results of this period also paid more attention to application and practice. Instead, they don't keep their eyes on theoretical methods, model simulation or dispersed applications, but the prediction data, prediction accuracy and prediction effectiveness.

On the traditional forecasting method of natural gas load, S. Askari et al. [52] used time series decomposition method to train TSK fuzzy system and forecast the high frequency gas consumption of each node of natural gas network. The low frequency data of 3 years are used to predict the gas consumption behavior in the future with the help of node consumption. Xu Duan et al. [53] used regression analysis method to analyze the relationship between gas consumption and temperature in heating season in Hebei Province. The results show that the dual regression effect of using the maximum temperature and the minimum temperature on gas consumption is better than that of using the average temperature on gas consumption. The former one has less error.

In the artificial intelligence forecasting method of natural gas load, many scholars optimize the parameters of SVM with the help of intelligent optimization algorithm to improve the forecasting accuracy. For example, Hongfang Lu et al. [54] combined Fruit Fly Optimization Algorithm (FOA), Simulated Annealing Algorithm (SA), Cross Factor (CF) and Support Vector Machine (SVM) together, and then he used this blending model to forecast the short-term load of urban gas. Compared with BP neural network and autoregressive moving average blending model, which concluded that CF-SA-FFOA-SVM model has higher gas load forecasting accuracy Chinese scholar Juan Song et al. [55] integrated the hybrid optimization strategy FOA-SVM of Fruit fly optimization algorithm and SVM. The results show that the combined optimization method of FOA-SVM has better prediction accuracy than artificial neural network and simple SVM method. Accordingly, many scholars have improved BP neural network with the help of optimization algorithm. Chunxia Liu et al. [56] proposed a natural gas load forecasting method based on genetic algorithm optimization wavelet neural network. Simulation results show that this method improves the prediction accuracy of the model and has certain application value. It can be seen that the prediction method based on genetic algorithm to optimize parameters has a good development prospect. In addition to optimizing parameters by using genetic algorithm, particle swarm optimization (PSO) can also be used to improve prediction technology. Z. Zhang and X. Liu [57] blended the advantages of particle swarm optimization algorithm in optimizing parameters and BP neural network, thus they established a long-term load forecasting model of natural gas based on PSO-BP, and they compared it with BP single forecasting model. The final results verify that the prediction model based on PSO-BP is more effective.

In artificial intelligence methods, deep learning is one of the fastest developing technologies in recent years. In the field of natural gas load forecasting, natural gas load will be affected by many factors. There are a large number of redundant factors in the original data set, which will increase the computational complexity and reduce the accuracy of the forecasting model. Deep learning introduces probabilistic generation model, which can automatically extract features from the training set. Moreover, deep learning can effectively solve the problem of inconsiderate manual characteristics. Combining the blending model of PCCA and LSTM, N. Wei et al. [58] predicts the natural gas load. Finally, the research results verify that the prediction model based on deep learning is accurate. Compared with the linear regression model and the traditional artificial neural network method, many experimental data can prove that the prediction technology based on deep learning can make the short-term load prediction accuracy of natural gas more accurate [59].

In this period, many scholars are still exploring new methods, even though there are many forecasting methods. Most of them use combination forecasting method to carry out natural gas load forecasting research. However, the current research results of natural gas load forecasting are mainly applied in specific areas, and the application results can only show the effect of natural gas load forecasting in this area, and the adaptability of forecasting methods is still limited.

C. RESEARCH TRENDS

Combined with the previous research and the development of natural gas market, this paper will discuss the problems that should be paid attention to in theoretical research and practical application in the field of natural gas load forecasting.

1) DATA PROCESSING IN FORECAST

In recent years, for the research of natural gas load forecasting, scholars attach great importance to the establishment of forecasting model, but there are some problems in analyzing and collecting raw data: ① The large amount of data makes the data arrangement error-prone; ② There are errors in data sources and statistical caliber; ③ An emergency may cause abnormal data. Irregular data will affect the prediction accuracy, even the prediction results, especially the safety of natural gas is unstable. Therefore, we should correctly identify the bad data, so as to deal with the research data more carefully. In this respect, scholars at home and abroad have explored it already. Literature [60] uses a robust extreme learning machine (ELM) algorithm to perform regression analysis on load data. Literature [61] adopts nested hierarchical Dirichlet process for abnormal data, and proposes an unsupervised learning method based on Bayesian model; Literature [62] proposes a data anomaly detection method based on exceptional model mining (EMM) framework for object-relational data. However, on the whole, the research on data processing of natural gas load forecasting is still deficient, which needs a further research.

2) THE PREDICTION OF THE PREDICTION MODEL

The effect of natural gas load forecasting depends on the historical load data and the historical regular pattern of its development. This involves the pre-evaluation of the prediction model. Pre-evaluation of the prediction model can make the prediction accuracy obtain a certain calculation before the actual quantity of the prediction is generated. Since load changes are both regular and random, the goal of pre-evaluation is to find the law of historical load data as much as possible, thereby reducing forecast errors. Therefore, it is necessary to pre-evaluate the regularity of historical load data. With people's high requirements for the accuracy of prediction models, it tend to be the future development directions. However, in the process of literature analysis, the research on the pre-evaluation of natural gas prediction models is still deficient, but it can be used for reference from the pre-evaluation research of power system prediction models. In this respect, Literature [63] studies the factors that affect the performance of K-means clustering algorithm. K-means clustering can be used to group prediction objects before operating prediction, and this method has been proved to improve the accuracy of load forecasting already. Literature [64] puts forward an interval forecasting method based on improved evaluation index, which effectively solves the problem that traditional point-to-point forecasting is difficult to apply to load with large fluctuation and strong uncertainty.

3) SELF-ADAPTIVE FORECASTING METHOD

With the rapid development of information technology, the development of any information industry tends to be intelligent, and the field of natural gas load forecasting is no exception. Artificial intelligence is intelligent, while artificial intelligence methods show a good self-adaptability. In natural gas load forecasting, forecasting technology should pay attention to improving its self-adaptive ability and self-learning ability. Self-adaptation means that the prediction technology automatically adjusts the model parameters according to the region it uses or the latest data, so as to achieve better prediction results. A forecasting method with good self-adaptive ability can learn and adjust itself from its own trial and error, experience and observation of the outside world. Self-adaptive forecasting method is the research direction that should be focused on in the era of artificial intelligence. At present, technologies with good adaptive ability include artificial neural network, fuzzy system method, deep learning, and transferring learning and so on. Literature

[65] analyzes about 50 forecasting methods from three aspects: forecasting accuracy, applicability of time and space, and relevance of policy objectives. It is found that artificial neural network is the most widely used method at present, followed by support vector machine (SVM), and other widely used methods, including autoregressive integrated moving average model (ARIMA), linear regression (LR), genetic algorithm (GA), particle swarm optimization algorithm (PSO) and so on. Literature [66] discusses and classifies adaptive methods in short-term load forecasting. Literature [67] analyzes and summarizes three kinds of artificial intelligence technologies that are commonly used in power systems: artificial neural network, support vector machine and adaptive neuro-fuzzy inference system. Literature [68] analyzes and compares the methods and models of power load forecasting, and the results show that machine learning or forecasting models based on artificial intelligence are the most popular. These research ideas can provide some reference for the adaptive development of natural gas load forecasting model.

4) EXPLORATION OF COMBINATION FORECASTING METHOD

Through bibliometric analysis, it can be found that a variety of technologies and methods have emerged in the field of natural gas load forecasting. In medium and long-term forecasting, the most commonly used forecasting methods are regression analysis, grey theory forecasting and so on. In short-term load forecasting, the most commonly used forecasting methods are time series method, support vector machine, neural network and deep learning, among which artificial neural network is the most studied forecasting method at present. According to the papers published at home and abroad, artificial intelligence prediction methods can achieve good prediction results. With the help of optimization algorithm, it will improve the prediction model, and the prediction accuracy can be further improved. It can be seen that the forecasting method has gradually developed from a single forecasting method to a combined forecasting method, which is the future development trend. In addition to the above literature analysis [10],[18],[19],[29],[33],[34], the combination forecasting method in the field of power load forecasting can also be appropriately referred. Literature [69] combines data preprocessing, hybrid optimization algorithm with local mean decomposition (LMD), gravitational search algorithm (GSA), particle swarm optimization algorithm (PSO) and wavelet neural network (WNN) together. It proposes a new power load forecasting system, which successfully overcomes the defects of traditional single forecasting model and achieves a higher forecasting accuracy. Literature [70] has developed a power combination forecasting model based on decomposition, which adopts dynamic adaptive entropy weight detection method and focuses on the stability of forecasting performance. Literature [71] based on clustering technology, combined with artificial neural network (ANN), wavelet neural network (WNN) and Kalman filter (KF), a power combination forecasting model is proposed, and the results show that the model has high performance. How the combination forecasting method will develop in the future is still worth exploring.

VI. CONCLUSION

In this paper, CiteSpace knowledge map software is used to make bibliometric analysis on the literatures about natural gas load forecasting in WOS and CNKI databases. Firstly, according to the annual output of sample data and its background, the development of natural gas load forecasting at home and abroad is divided into three stages, namely, the embryonic stage (before 2000), the initial stage (from 2000 to 2015) and the rapid development stage (from 2016 to now).

Secondly, visual atlas presentation and quantitative analysis are conducted for research institutions and academic communities in the literature samples. The results show that: ① University of Ljubljana, China University of Petroleum (East China) and Southwest Petroleum University are three universities with strong research power in this field; ② According to research strength, we can rank them from top to bottom, they are: institutions of higher learning, companies, and research institutes; ③ Countries with strong research power, ranked from the top to the bottom, are China, the United States, Italy, Iran, South Korea and so on; ④ China's development level in this field is unbalanced, and research institutions mainly come from areas rich in natural gas resources and areas with rapid economic development; ⑤ The research strength and resources of natural gas load forecasting at home and abroad are relatively dispersed, and eight major academic groups at home and abroad are excavated through literature samples.

Finally, with the help of keywords co-occurrence network, Timeline view and TimeZone view, the research hotspots and cutting-edge technologies in this field are visualized, and the hot spot migration process in this field is discovered, thus the research trends in this field are proposed. Firstly, the research hotspots of natural gas load forecasting can be classified into three categories: ① common forecasting technology; ② classification of natural gas load forecasting according to forecasting time; ③ forecasting in actual areas. Secondly, the cutting-edge technologies in the field of natural gas load prediction are explored and compared. The prediction technologies of artificial intelligence have a better development momentum. Thirdly, according to the division of natural gas load forecasting stages, the evolution and development process of each stage are expounded respectively. The forecasting technology in the initial stage mostly stays at the theoretical level, and the results in the rapid development stage pay more attention to application, focusing on forecasting data, forecasting accuracy and forecasting effectiveness; Fourthly, combined with previous studies, the paper puts forward the problems that should be paid attention to in this field, namely, data processing in forecasting, prediction of forecasting models, adaptive forecasting methods and combination forecasting methods, which are also the research directions that should be vigorously developed in the future.

ACKNOWLEDGEMENT

This work was supported by The National Natural Science Foundation of China（No. 62072363）. The authors are grateful for the help in writing this article by Yu Zhao and Hongbin Dai. They also thank the editor and the anonymous reviewers for their valuable comments.

REFERENCES

[1] S. Y. Kan, B. Chen, J. Meng, and G. Q. Chen, "An extended overview of natural gas use embodied in world economy and supply chains: Policy implications from a time series analysis," (in English), *Energy Policy*, Article vol. 137, pp. 10, Feb 2020, Art no. 111068.
[2] S. Y. Kan, B. Chen, X. F. Wu, Z. M. Chen, and G. Q. Chen, "Natural gas overview for world economy: From primary supply to final demand via global supply chains," (in English), *Energy Policy*, Article vol. 124, pp. 215-225, Jan 2019.
[3] National Bureau of Statistics of China, "2019 China Statistical Yearbook, "China Statistics Press, Sep. 09, 2019. [Online]. Available: http://www.stats.gov.cn/tjsj/ndsj/2019/indexch.htm.
[4] T. Ahmad, H. X. Chen, D. D. Zhang, and H. C. Zhang, "Smart energy forecasting strategy with four machine learning models for climate-sensitive and non-climate sensitive conditions," *Energy*, vol. 198, May 2020, Art no. 117283.
[5] X. T. Li and H. Li, "A Visual Analysis of Research on Information Security Risk by Using CiteSpace," *Ieee Access*, vol. 6, pp. 63243-63257, 2018.
[6] M. Q. Yan, L. M. Lian, W. L. Jiao, Y. F. Tan, and C. H. Zhan, " Gas Load and Progression of Research Works," *Gas & Heat*, vol. 22, no. 6, pp. 490-490, 2002.
[7] Bobrovski, " Natural gas pipeline transportation". *Petroleum industry press*, 2008.
[8] C. Y. Zhou and A. L. Hu, "Overview of the world Natural gas Industry in 2000," *Natural Gas Technology and Economy*, vol. 05, pp. 58-61, 2001.
[9] China Petroleum News, "When will China's natural gas production peak come? "[Online]. Available: https://baijiahao.baidu.com/s?id=1598493793293589622&wfr=spider&for=pc.
Accessed on: Sep. 9,2020.
[10] P. Potocnik and E. Govekar, "Applied short-term forecasting for the Slovenian natural gas market," *in International Conference on the European Energy Market*, 2016.
[11] P. Potočnik, E. Govekar, and I. Grabec, "Short-term natural gas consumption forecasting," *in Iasted International Conference on Applied Simulation & Modelling*, 2007.
[12] P. Potocnik, M. Thaler, E. Govekar, I. Grabec, and A. Poredos, "Forecasting risks of natural gas consumption in Slovenia," *Energy Policy*, vol. 35, no. 8, pp. 4271-4282, 2007.
[13] R. Hribar, P. Potočnik, J. Silc, and G. Papa, "A comparison of models for forecasting the residential natural gas demand of an urban area," *Energy*, vol. 167, no. JAN.15, pp. 511-522, 2019.
[14] P. Potočnik, B. Vidrih, A. Kitanovski, and E. Govekar, "Neural network, ARX, and extreme learning machine models for the short-term prediction of temperature in buildings," *Building Simulation*, vol. 12, no. 6, pp. 1077-1093, 2019.
[15] P. Potočnik, E. Strmčnik, and E. Govekar, "Linear and Neural Network-based Models for Short-Term Heat Load Forecasting," *Strojniski Vestnik*, vol. 61, no. 9, pp. 543-550, 2015.
[16] P. Potocnik and E. Govekar, "Analysis and Short-Term Forecasting of Highway Traffic Flow in Slovenia," *in Adaptive and Natural Computing Algorithms*, Pt I, vol. 6593, A. Dobnikar, U. Lotric, and B. Ster, Eds. (Lecture Notes in Computer Science), 2011, pp. 270-279.
[17] H. F. Lu, X. Ma, and M. Azimi, "US natural gas consumption prediction using an improved kernel-based nonlinear extension of the Arps decline model," *Energy*, vol. 194, pp. 116905-, 2020.
[18] H. F. Lu, X. Ma, K. Huang, and M. Azimi, "Carbon trading volume and price forecasting in China using multiple machine learning models," *Journal of Cleaner Production*, 2020.
[19] H. F. Lu, X. Ma, K. Huang, and M. Azimi, "Prediction of offshore wind farm power using a novel two-stage model combining kernel-based nonlinear extension of the Arps decline model with a multi-objective grey wolf optimizer," *Renewable & Sustainable Energy Reviews*, vol. 127, Jul 2020, Art no. 109856.
[20] X. Zhang, L. Che, M. Shahidehpour, A. S. Alabdulwahab, and A. Abusorrah, "Reliability-Based Optimal Planning of Electricity and Natural Gas Interconnections for Multiple Energy Hubs," *IEEE Transactions on Smart Grid*, pp. 1-10, 2015.
[21] X. Zhang, L. Che, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "Electricity-Natural Gas Operation Planning With Hourly Demand Response for Deployment of Flexible Ramp," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 996-1004, 2017.
[22] X. Zhang, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "Hourly Electricity Demand Response in the Stochastic Day-Ahead Scheduling of Coordinated Electricity and Natural Gas Networks," *IEEE Transactions on Power Systems*, pp. 1-10, 2015.
[23] H. Su et al., "A method for the multi-objective optimization of the operation of natural gas pipeline networks considering supply reliability and operation efficiency," (in English), *Computers & Chemical Engineering*, Article vol. 131, pp. 10, Dec 2019, Art no. Unsp 106584.
[24] H. Su, J. J. Zhang, E. Zio, N. Yang, X. Y. Li, and Z. J. Zhang, "An integrated systemic method for supply reliability assessment of natural gas pipeline networks," (in English), *Applied Energy*, Article vol. 209, pp. 489-501, Jan 2018.
[25] J. Liu and E. Zio, "SVM hyperparameters tuning for recursive multi-step-ahead prediction," (in English), *Neural Computing & Applications*, Article vol. 28, no. 12, pp. 3749-3763, Dec 2017.
[26] S. P. Zhang, X. Z. Xu, and J. W. Dai, "Short-Term Gas Forecasting Based on EMD and PSO-WNN," *Computer Simulation*, vol. 33, no. 003, pp. 305-309, 2016.
[27] X. X. Wang, X. Z. Xu, T. Zhang, and C. W. Gao, " Gas Load Forecasting Method Based on Integrated Deep Learning Algorithms," *Computer Systems & Applications*, vol. v.28, no. 12, pp. 51-58, 2019.
[28] C. Song, J. Song, and J. Ren, " Modelling and Simulation of Elman Neural Network on Short − term Gas Load Prediction," *Industrial Control Computer*, vol. 29, no. 04, pp. 87-88, 2016.
[29] J. Song and S. T. Liao, " Short-term Gas Load Forecasting Based on BP Neural Network and Multivariable Linear Regression," *Ningxia Engineering Technology*, 2019.
[30] W. B. Qiao, B. D. Chen, S. J. Wu, C. Y. Li, J. S. Mao, and J. L. Ma, "A forecasting model of natural gas daily load based on wavelet transform and LSSVM-DE," *Natural Gas Industry*, vol. 34, no. 009, pp. 118-124, 2014.
[31] W. B. Qiao and B. D. Chen, "Hourly Load Prediction for Natural Gas Based on Haar Wavelet Transforming and ARIMA-RBF," *Journal of Petrochemical Universities*, no. 04, pp. 77-82, 2015.
[32] W. B. Qiao, B. D. Chen, and H. F. Lu, " Gas load forecasting based on chaotic theory and Volterra adaptive filter," *Scientia Sinica(Technologica)*, vol. 000, no. 001, pp. 91-102, 2015.
[33] W. L. Jiao, L. M. Lian, J. H. Cui, and M. Q. Yan, "Time Series Model of Predicting Short Period Cit" Gas Load," *Natural Gas Industry*, vol. 22, no. 001, pp. 92-94, 2002.
[34] L. W. Jiao, Y. K. Q, and L. B. Zhao, "Study on Forecasting System of City Gas Load," *Natural Gas Industry*, no. 01, pp. 155-157, 2005.
[35] V. N. Vapnik, "The Nature of Statistical Learning Theory," 1995.
[36] D. W. Gong, J. Sun, and Z. Miao, "A Set-Based Genetic Algorithm for Interval Many-Objective Optimization Problems," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 99, pp. 47-60, 2018.
[37] S. N. Dodamani, V. J. Shetty, and R. B. Magadum, "Short term load forecast based on time series analysis: A case study," *in International Conference on Advancements in Power & Energy*, 2015.

[38] W. L. Jiao, J. B. Jin, L. M. Lian, and M. Q. Yan, "Application of time series analysis to load forecast of natural gas in town," *Journal of Harbin University of Civil Engineering and Architecture*, no. 04, pp. 79-82, 2001.
[39] P. J. Liu and Y. Zhao, "Short-term Load Prediction of Natural Gas in Winter in Xi'an City," *Gas & Heat*, vol. 29, no. 010, pp. 37-42, 2009.
[40] H. Liu and Y. S. Zou, " Gas load prediction based on regression analysis," *Inner Mongolia Petrochemical Industry*, vol. 38, no. 04, pp. 39-40, 2012.
[41] J. H. Wang, B. Z. Li, Q. Z. Chen, and T. T. Liu, "Forecast of long-term gas load in small town based on grey theory," *Journal of Central South University of Technology*, vol. 14, pp. 26-29, Dec 2007.
[42] J. W. Wang, L. Wang, Y. M. M. Hu, T. L, and Z. Y. Yang, " Urban gas load forecasting based on grey theory," *Public Communication of Science & Technology*, vol. 000, no. 004, pp. 130-130,127, 2014.
[43] A. Khotanzad and H. Elragal, "Combination of artificial neural-network forecasters for prediction of natural gas consumption," *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 464-73, 2000.
[44] Y. Zhao, Y. Liu, Z. B. Miao, and Z. N. Liu, " Application of Neural-network in Natural Gas Load Forecasting," *Gas & Heat*, vol. 23, no. 006, pp. 331-332, 2003.
[45] R. Kizilaslan and B. Karlik, "Combination of neural networks forecasters for monthly natural gas consumption prediction," *Neural Network World*, vol. 19, no. 2, pp. 191-199, 2009.
[46] A. Azari, M. Shariaty-Niassar, and M. Alborzi, "Short-term and Medium-term Gas Demand Load Forecasting by Neural Networks," *Iranian Journal of Chemistry & Chemical Engineering International English Edition*, vol. 31, no. 4, pp. 77-84, 2012.
[47] H. Liu, D. Liu, G. Zheng, Y. Liang, and Y. Ni, "Research on natural gas load forecasting based on support vector regression," *in World Congress on Intelligent Control & Automation*, 2004.
[48] P. Potocnik, B. Soldo, G. Simunovic, T. Saric, A. Jeromen, and E. Govekar, "Comparison of static and adaptive models for short-term residential natural gas forecasting in Croatia," *Applied Energy*, vol. 129, no. sep.15, pp. 94-103, 2014.
[49] N. Aras, "Forecasting Residential Consumption of Natural Gas Using Genetic Algorithms," *Energy Exploration & Exploitation*, vol. 26, no. 4, pp. 241-266, 2008.
[50] C. Song and J. Song, " Parameter Optimization for BP Neural Network with GA on Short − term Gas Load Prediction," *Industrial Control Computer*, vol. 25, no. 010, pp. 82-84, 2012.
[51] J. L. Wang, "Research on Forecasting Methods of Natural Gas Load Based on Particle Swarm Optimization-Least Squares Support Vector Machine," *East China University of Science and Technology*, 2013.
[52] S. Askari, N. Montazerin, and M. H. F. Zarandi, "Modeling energy flow in natural gas networks using time series disaggregation and fuzzy systems tuned by particle swarm optimization," *Appl. Soft. Comput.*, vol. 92, p. 106332, 2020.
[53] X. Duan, Z. Q. Wang, J. Ma, L. Han, X. L. Zhao, and J. Gong, "Analysis of relationship between natural gas load and air temperature in heating season in Hebei Province," *Chemical Engineering of Oil & Gas*, vol. 048, no. 005, pp. 42-48, 2019.
[54] Hongfang, Lu, Mohammadamin, Azimi, Tom, and Iseley, "Short-term load forecasting of urban gas using a hybrid model based on improved fruit fly optimization algorithm and support vector machine," *Energy Reports*, vol. 5, pp. 666-677, 2019.
[55] J. Song and H. Pan, " Gas Daily Load Forecasting Based on Fruit Fly Optimization Algorithm and SVM," *Control Engineering of China*, no. 10, pp. 29-36, 2017.
[56] C. X. Liu, J. Li, W. C. Dang, S. W. Bai, and X. B. Wang, " Short-Term Natural Gas Load Forecasting Based on Wavelet Neural Network Optimized by Genetic Algorithm," *Computer Systems & Applications*, vol. 029, no. 004, pp. 175-180, 2020.
[57] Z. Zhang and X. Liu, "Prediction of long-term gas load based on particle swarm optimization and gray neural network model," *Advances in Mechanical Engineering*, vol. 9, no. 7, p. 168781401771108, 2017.
[58] N. Wei, C. J. Li, J. H. Duan, J. Y. Liu, and F. H. Zeng, "Daily Natural Gas Load Forecasting Based on a Hybrid Deep Learning Model," *Energies*, vol. 12, no. 2, Jan 2019, Art no. 218.
[59] G. D. Merkel, R. J. P. Id, and R. H. Brown, "Short-Term Load Forecasting of Natural Gas with Deep Neural Network Regression," *Energies*, vol. 11, no. 8, pp. 2008-, 2018.
[60] X. G. Peng, W. Q. Zheng, L. X. Lin, and Y Liu, " Short-term Load Forecasting Method Based on Outlier Robust Extreme Learning Machine Considering Adaptive Load Detection and Repair," *Proceedings of the CSEE*, no. 23, pp. 6409-6417, 2016.
[61] H. Chen, K. Damevski, D. Shepherd, and N. A. Kraft, "Modeling hierarchical usage context for software exceptions based on interaction data," *Automated Software Engineering*, vol. 26, no. 4, pp. 733-756, Dec 2019.
[62] F. Riahi and O. Schulte, "Model-based exception mining for object-relational data," *Data Mining and Knowledge Discovery*, pp. 1-42, 2020.
[63] X. Wang, W. J. Lee, H. Huang, R. L. Szabados, and P. V. Olinda, "Factors that Impact the Accuracy of Clustering Based Load Forecasting," *IEEE Transactions on Industry Applications*, vol. 52, no. 5, pp. 1-1, 2016.
[64]S. H. Xu et al., " Improved Evaluation Index Based Short-term Interval Prediction of Fluctuation Load," *Automation of Electric Power Systems*, vol. 44, no. 2, pp. 155,2020.
[65] K. B. Debnath and M. Mourshed, "Forecasting methods in energy planning models," (in English), Renewable & Sustainable Energy Reviews, Review vol. 88, pp. 297-325, May 2018.
[66] D. Singh and S. P. Singh, "Self Organization and Learning Methods in Short Term Electric Load Forecasting: A Review," *Electric Machines & Power Systems*, vol. 30, no. 10, pp. 1075-1089, 2002.
[67] K. Zor, O. Timur, and A. Teke, "A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting," i*n 2017 6th International Youth Conference on Energy (IYCE)*, 2017.
[68] M. A. Hammad, J. Borut, R. Bojan, and D. Dejan, "Methods and Models for Electric Load Forecasting: A Comprehensive Review," *Logs & Sustainable Transport*, vol. 11, 2020.
[69] R. Wang, J. Y. Wang, and Y. Z. Xu, "A novel combined model based on hybrid optimization algorithm for electrical load forecasting," (in English), *Applied Soft Computing*, Article vol. 82, pp. 21, Sep 2019, Art no. 105548.
[70] Z. Hu, J. Ma, L. Yang, X. Li, and M. Pang, "Decomposition-Based Dynamic Adaptive Combination Forecasting for Monthly Electricity Demand," *Sustainability*, vol. 11, no. 5, 2019.
[71] H. H. H. Aly, "A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid," *Electric Power Systems Research*, vol. 182, May 2020, Art no. 106191.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3044052, IEEE Access

HUIBIN ZENG was born in China in 1993.She received the B.M. degree from Guangxi Teachers Education University in 2016 and the M.S degree from Nanning Normal University in 2019.She is currently pursuing the Ph.D. degree with the School of Management, Xi'an University of Architecture and Technology. She research interests include intelligent algorithm, data mining, information technology and management. Email: zenghuibin@163.com.

Genqing Bian received the PhD degree at School of Management, Xi'an University of Architecture and Technology, Shaanxi, China. He is a member of China Computer Federation (CCF) and Association for Computing Machinery (ACM). He is an associate professor at XAUAT. His primary research interests include the areas of Information Security, Knowledge Graph and Recommendation system.

Bilin Shao received the M.S. degree at School of Management, Xi'an University of Architecture and Technology, Shaanxi, China. He is professor at XAUAT. He is a member of China Computer Federation (CCF). His primary research interests include the areas of Knowledge Graph, Artificial intelligence and Recommendation system. *The corresponding author. Email: sblin0462@sina.com.

Dan Song is currently pursuing PhD at School of Management, Xi'an University of Architecture and Technology, Shaanxi, China. She is a member of China Computer Federation (CCF). Her primary research interests include the areas of Deep learning, Data mining, and Knowledge graph. Email: dandans125@163.com.

Xiaojun Li received the M.S. degree at School of Software, Shandong University. Shandong, China in the year 2008. He is currently pursuing PhD at School of Management, Xi'an University of Architecture and Technology, Shaanxi, China. He is a member of China Computer Federation (CCF). His primary research interests include the areas of Data mining, Knowledge Graph and Recommendation system. He is a Reviewer of Chinese Journal of Computers, and Systems Engineering-Theory & Practice. He is the editor-in-chief of the "Python Language" tutorial which published by Tsinghua University Press. Email: xiaojunlee@foxmail.com.