Power Sequencial Data - Forecasting Trend

Lejie Li[1], Lu Zhang[2], Bin Sun[2](✉), Benjie Dong[1], and Kaining Xu[1]

[1] Jinan Lingsheng Info Tech. Co. Ltd., Huaiyin District, Jinan 250000, China
[2] School of EE, University of Jinan, Jinan 250022, China
cse_sunb@ujn.edu.cn

Abstract. In reduce the use of non-renewable energy, the use of renewable energy is increasing day by day. In recent years, with the strong support of the state, renewable energy has been applied in various industries. Renewable energy generates a considerable amount of electricity, which brings us huge economic benefits but also brings certain problems. For example, the instability of the power generation system, the scheduling, and distribution of power, etc. Therefore, the analysis of the massive power data generated by the power system has become particularly important. Effective processing and forecasting of these power data can not only improve the efficiency and performance of the power system but also enable effective power dispatching and deployment. At the same time, it can ensure the safety of industrial and family users and ensure social stability. Machine learning has been widely used in various fields and achieved good performance in recent years. Therefore, many researchers have begun to use machine learning to predict power data. Therefore, we provide a preliminary overview of the history and evolution of machine learning-based power data analysis and forecasting from the perspective of bibliometrics.

Keywords: Electricity · Data analytics · Data prediction · Machine learning

1 Introduction

In recent years, machine learning has been widely used in various industries and achieved good results. Electricity is one of the indispensable energy sources in our lives. At present, due to the increase in supply and demand, the diversification of power generation methods (wind power, water power, solar energy, etc.), and the wide application of new energy vehicles, a large number of power systems have been generated [1]. Data, power generation, transmission, distribution, and electricity consumption all have a large amount of data waiting to be processed and analyzed. Accurate analysis of power data can not only grasp the user's demand for energy consumption but also avoid energy loss. Moreover, it has a more important significance for protecting equipment and personnel safety and promoting social stability [2]. In the research direction of power data analysis and prediction, it is also closely related to research in other fields.

B. Wang et al. (Eds.): ICMTEL 2023, LNICST 535, pp. 369–379, 2024.
https://doi.org/10.1007/978-3-031-50580-5_32

2 Background

At present, researchers at home and abroad have made in-depth explorations in the field of power data prediction. Including some wind forecasts, energy consumption forecasts, solar forecasts, etc. The de Alencar, DB [3] team proposed a hybrid model for wind prediction, and conducted short-, medium-, and long-term verifications on real meteorological data sets. Hanifi, Shahram [4] et al. provide a brief analysis of the history of wind forecasting systems and systematically analyze the latest methods in various fields (physics, statistics, etc.). Alhussein, Musaed [5] predicted the electricity consumption of households by combining the deep learning framework, which assisted in the work and operation of the power grid. Lee, Donghun [6] et al. used three deep learning algorithms to predict hourly solar power generation and verified it on real datasets.Bin Sun et al. proposed a method for emergency data prediction [7]. In the field of fault diagnosis, Yiming Xiao et al. proposed an unsupervised cross-domain fault diagnosis method [8]. Mingzhi Chen et al. significantly improved the accuracy of fault diagnosis tasks by improving GAN [9]. Bin Sun et al. [10] propose a data-driven approach to anomaly detection.

3 Related Papers and Sources

We selected 499 articles on the processing of power system time series data and their references in the core database of Web Of Science (WOS), and the research is mainly based on machine learning algorithms to predict these data. These articles are mainly composed of journal papers, including 472 articles, 16 reviews, 6 proceedings papers, and so on. Article information was imported into R library bibliometrix to generate and analyze bibliometrics [11]. Its main information is shown in Table 1.

Table 1. Main information

Main information about data	Result
Timespan	2011:2023
Sources	129
Documents	499
Average citations per doc	26.32
References	16189
Keywords Plus (ID)	677
Author's Keywords (DE)	1676
Authors	1719

From Fig. 1, we can see that the prediction of power system data based on machine learning algorithms has been widely studied by scholars around the world. Especially after 2017, research in related directions and the output of articles have grown rapidly. This shows that the massive data generated by the power system will be increasingly inseparable from machine learning in future processing and prediction.

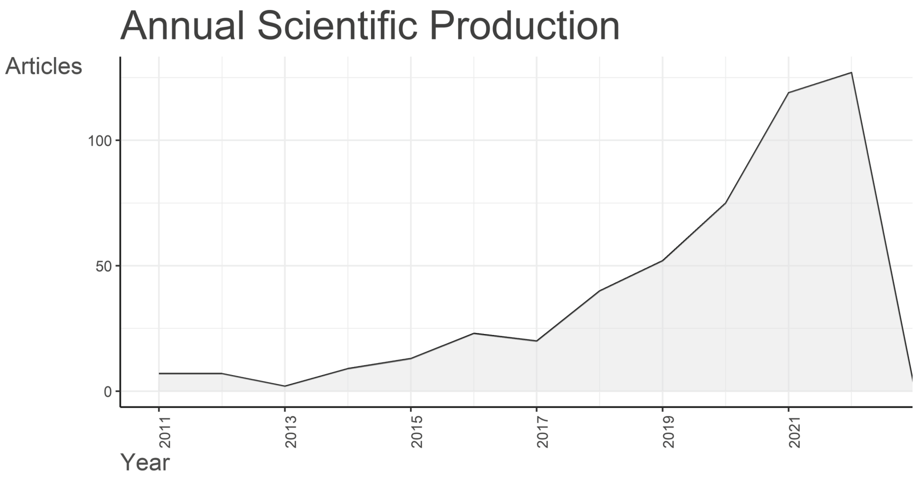

Fig. 1. Annual scientific production

Figure 2 shows some of the magazines from which the papers come. We can see that Energies and Applied Energy selected the most important sources in the database for us, with 64 and 42 papers respectively. At the same time, journals in computer and other fields also published more papers on this topic.

Although some journals publish more papers than others, it does not prove that it is also far superior to other journals in terms of the quality of papers. As shown in Fig. 3, we sort the magazines according to their influence (G-index). We can see that Energies, which produces the most output, is not the most influential journal, and although Applied Energy does not produce the most articles, it ranks among the best in terms of influence. The journal published a total of 42 papers in our collected database. Some of them provide a powerful help to the research of subsequent scholars. For example, the Wang, HZ team predicted the wind speed by improving the deep belief network [12], and applied it to the power energy system, which was cited 316 times in Web Of Science (WOS). Liu, ZK et al. [13] used the combination model for short-term wind speed prediction and achieved good results, which were cited 127 times in Web Of Science (WOS). Wang, JZ et al. [14] proposed a hybrid model to forecast wind speed, and verified the reliability of the model on a real data set in a certain area of Gansu Province, China, and was cited 118 times in Web Of Science (WOS).

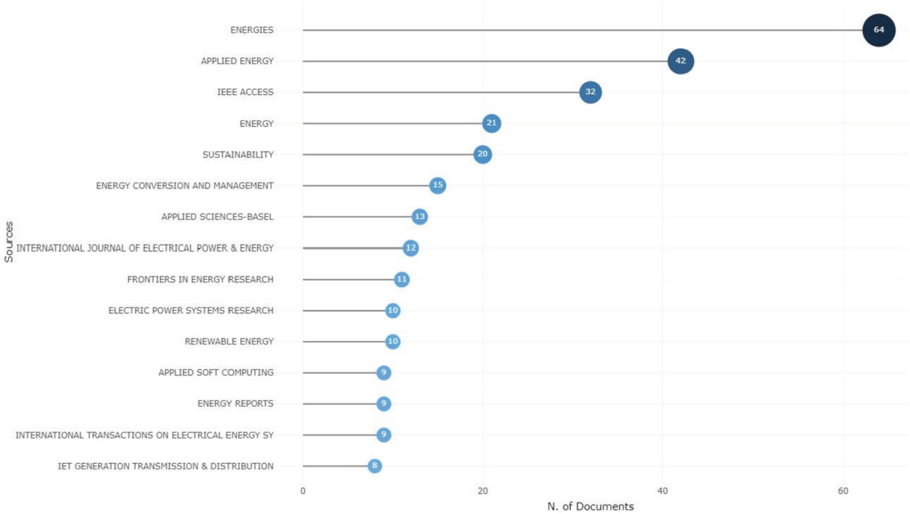

Fig. 2. Journals from which the main papers are sourced

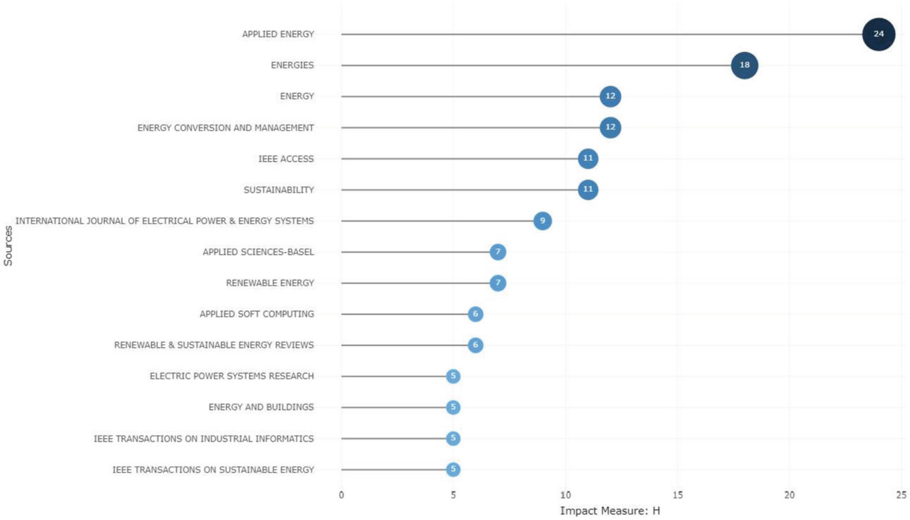

Fig. 3. Source impact (Source G-index)

4 Citations and References

We sort the articles in the selected database according to the number of citations, as shown in Fig. 4. We can see that the most cited article is that Inman, RH team conducted a survey and analysis of solar forecasting methods, introduced some successful applications of forecasting methods [15], and reviewed the successful application of some forecasting methods in power plants. This article has been cited 591 times in WOS. The second place is Das, UK team [16]. They conducted a systematic review of photovoltaic power generation forecasting models and briefly analyzed their advantages and disadvantages, which were cited 387 times in WOS. The third most cited article is from Applied Energy magazine Wang, HZ's team [12], which shows that Applied Energy magazine does have a great influence in this field.

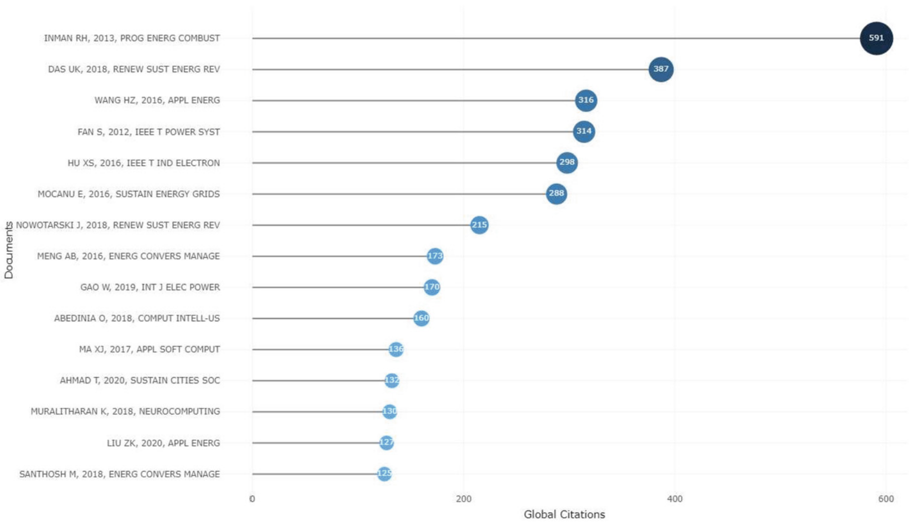

Fig. 4. Most cited documents

We divide the authors by region and analyze the scientific productivity of the region, as shown in Table 2. We can see that the main source of these articles is China. The reason may be that China's power system is widely distributed and the amount of data generated is also very large, so the demand for data analysis has also increased. Secondly, the scientific productivity of the United States and India is also among the best.

Table 2. Regional scientific production

Region	Freq
CHINA	471
USA	86
INDIA	74
SOUTH KOREA	64
IRAN	35
PAKISTAN	32
BRAZIL	29
AUSTRALIA	28

According to the literature statistics (Fig. 5), we can see that the more cited papers have a greater impact on the research of other scholars. At the same time, we found that in the papers with a high number of citations, most authors chose the deep learning method when dealing with power system data (including charge, wind speed, etc.). This shows that in recent years and the next few years or even decades, deep learning has always been an important technical difficulty and hot spot in the analysis and prediction of power system data.

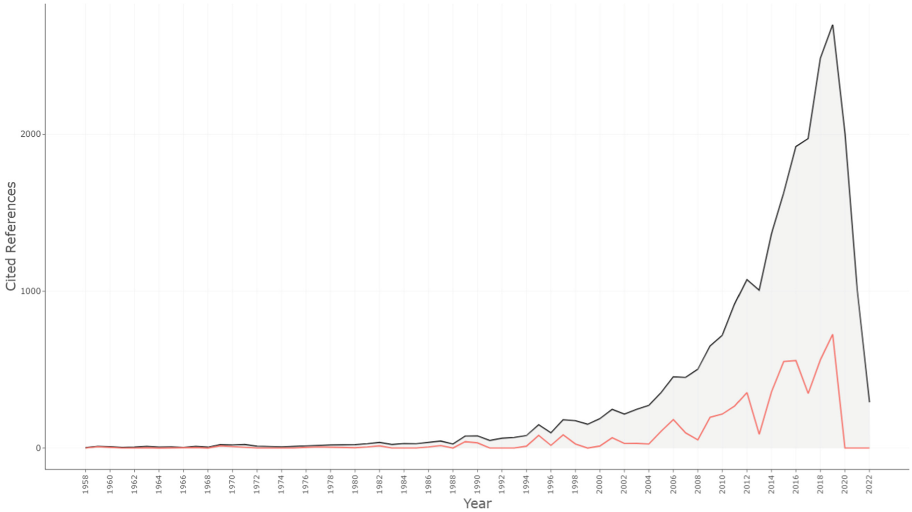

Fig. 5. References spectroscopy (RPYS)

We analyze the relationship between authors, papers, and keywords, as shown in Fig. 6. From the figure, we can see the more important subject keywords in recent years, as well as the literature of some scholars in this direction. Among them, the more important research topics include deep learning, hybrid models, and the prediction of power consumption.

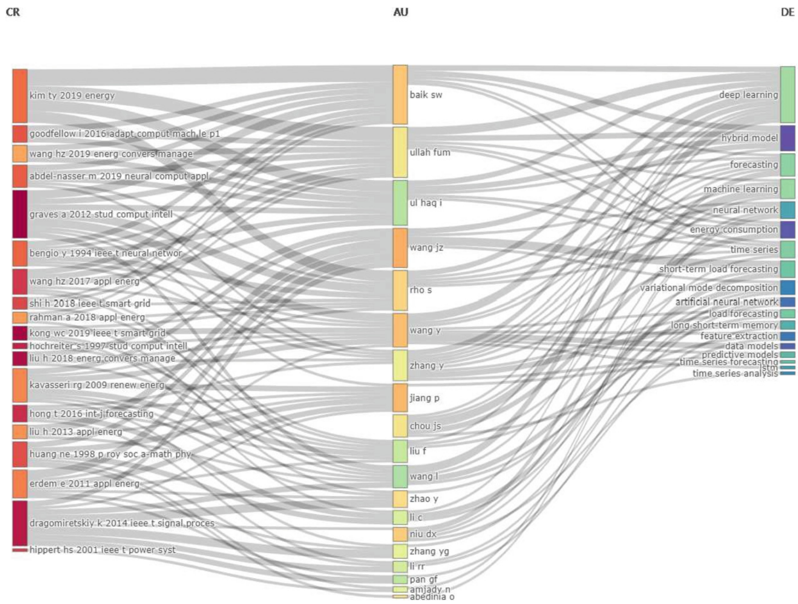

Fig. 6. Three-Field Plot

5 Trending Topics and Theme Evolution

In order to further understand the development of the field of power system forecasting, we analyze the trend changes of topic keywords and the evolution of topics, as shown in Fig. 7 and Fig. 8. From the subject keywords, we can see that the data processing of the power system is mainly in the prediction of wind energy, power consumption, and solar energy. At the same time, great changes have taken place in research methods. Before 2018, most scholars chose support vector machine (SVM) as the main method for prediction. After 2018, the research methods have shifted to a certain extent, and neural network algorithms have surpassed SVM and other algorithms and become the main choice in power system data analysis and prediction. The reason may be that with the improvement of computer hardware, some breakthroughs have been made in deep learning, and deep learning has gradually been applied in power systems and has shown good results. At the same time, we can see from the evolution of the theme that before 2020, the prediction system that occupies the main position has gradually flowed to neural networks, power systems, energy consumption, wind speed, etc.

We divide trend themes into four different types for analysis, as shown in Figs. 9 and 10. The four quadrants represent four different types respectively. The horizontal axis represents centrality, and the vertical axis represents density. First of all, the first type is Motor Themes. Mainly distributed in the upper right corner, the centrality and density of this part are relatively high, so the

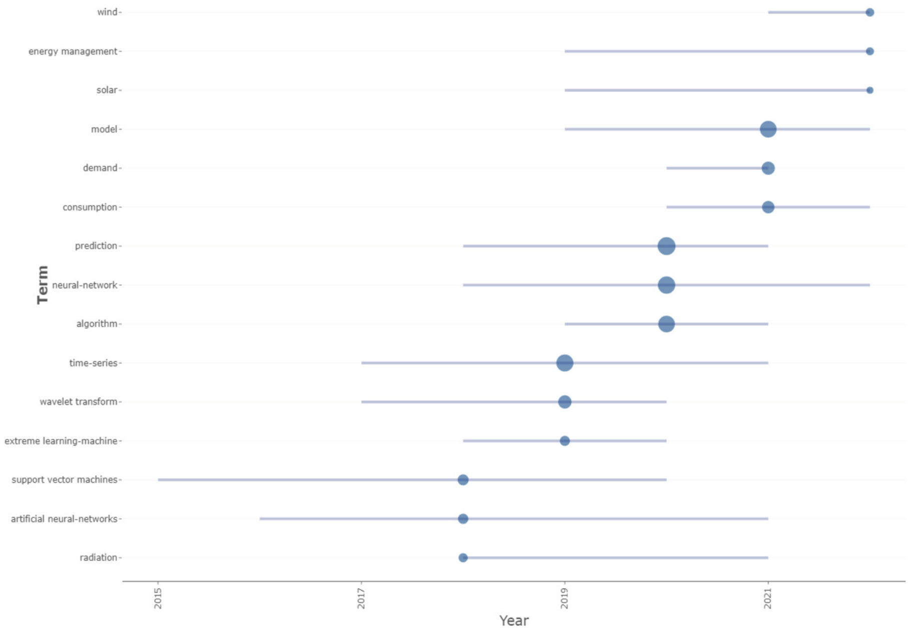

Fig. 7. Trend changes in Topic keywords

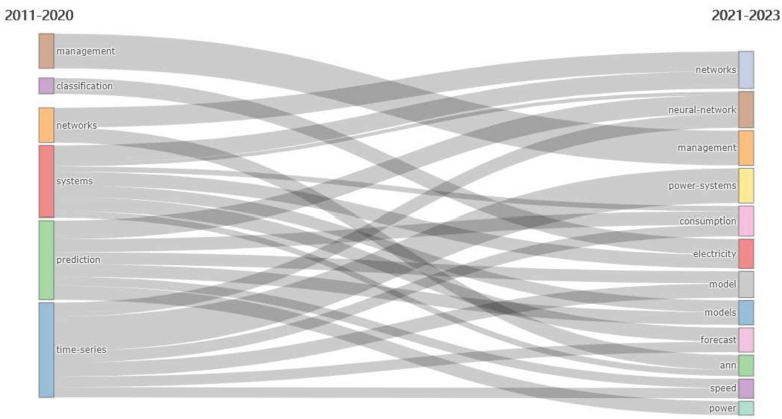

Fig. 8. Thematic evolution

topics included are relatively important and have developed to a certain extent. The second type is Niche Themes. This part has a high density of themes, but has fewer applications in this field, and is mainly distributed in the upper left corner. The third type has relatively low centrality and density. This type mainly contains some emerging or about to decline themes, mainly distributed in the lower left corner. The last type is Basic Themes. The themes in this part are more suitable for the development of this field, but the density is low, mainly some relatively basic themes.

From Fig. 9, we can see that before 2020, in the first quadrant, time series and wavelet neural transformation are more important, but they have not been well developed. The prediction algorithm model, power system, etc. are distributed in the fourth quadrant as some basic topics. Management is a marginal topic, which is the main research work in the third quadrant.

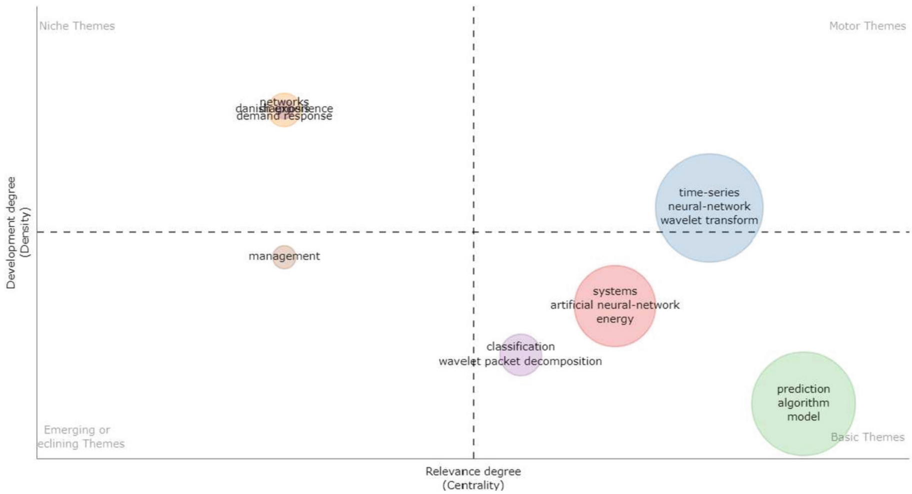

Fig. 9. Thematic map until 2020

After 2020, the thematic map has undergone some changes, as shown in Fig. 10. Among them, the obvious corners are the appearance of wind speed, the hybrid model, and some deep learning algorithms in the first quadrant. This shows that in the power system, data analysis, especially in data prediction, deep learning algorithms, and some hybrid models have been widely used. In the fourth quadrant, there are mainly big data, power demand, etc., which provide great help for data prediction as a research basis. There are also some topics that have slowly been removed from the research field after 2020, such as state estimation in the third quadrant.

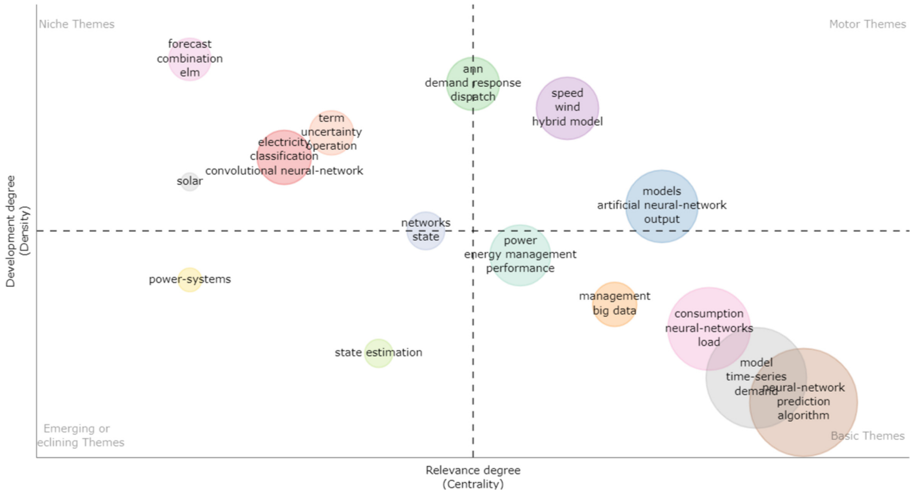

Fig. 10. Thematic map after 2020

6 Conclusions and Future Directions

According to the evolution of topics and changes in topic trends, we can see that artificial intelligence has gradually been integrated into the power system. Deep learning has become indispensable in the processing and prediction of power system data [17]. As a topic that has attracted attention in recent years, neural networks, we believe that more and more researchers will devote themselves to the research of neural networks in the future. In particular, the processing of power system data based on neural networks can be said to be the focus and difficulty of research in the next few years or even ten years.

References

1. Sun, B., Geng, R., Zhang, L., Li, S., Shen, T., Ma, L.: Securing 6G-enabled IoT/IoV networks by machine learning and data fusion. EURASIP J. Wirel. Commun. Netw. **2022** (2022)
2. Sun, B., Geng, R., Yuan, X., Shen, T.: Prediction of emergency mobility under diverse IoT availability. EAI Endorsed Trans. Pervasive Health Technol. **8**(4), e2 (2022)
3. de Alencar, D.B., et al.: Different models for forecasting wind power generation: case study. Energies **10**(12), 2017 (1976)
4. Hanifi, S., Liu, X., Lin, Z., Lotfian, S.: A critical review of wind power forecasting methods-past, present and future. Energies **13**(15), 3764 (2020)
5. Alhussein, M., Aurangzeb, K., Haider, S.I.: Hybrid CNN-LSTM model for short-term individual household load forecasting. IEEE Access **8**, 180544–180557 (2020)
6. Lee, D., Kim, K.: Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information. Energies **12**(2), 215 (2019)

7. Sun, B., Geng, R., Shen, T., Xu, Y., Bi, S.: Dynamic emergency transit forecasting with IoT sequential data. Mob. Netw. Appl. 1–15 (2022)
8. Xiao, Y., Shao, H., Han, S.Y., Huo, Z., Wan, J.: Novel joint transfer network for unsupervised bearing fault diagnosis from simulation domain to experimental domain. IEEE/ASME Trans. Mechatron. **27**(6), 5254–5263 (2022)
9. Yan, S., Shao, H., Xiao, Y., Liu, B., Wan, J.: Hybrid robust convolutional autoencoder for unsupervised anomaly detection of machine tools under noises. Robotics Comput. Integr. Manuf. **79**, 102441 (2023)
10. Sun, B., Ma, L., Shen, T., Geng, R., Zhou, Y., Tian, Y.: A robust data-driven method for multiseasonality and heteroscedasticity in time series preprocessing. Wirel. Commun. Mob. Comput. **2021**, 6692390:1–6692390:11 (2021)
11. Aria, M., Cuccurullo, C.: Bibliometrix: an R-tool for comprehensive science mapping analysis. J. Informet. **11**(4), 959–975 (2017)
12. Wang, H.Z., Wang, G.B., Li, G.Q., Peng, J.C., Liu, Y.T.: Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl. Energy **182**, 80–93 (2016)
13. Liu, Z., Jiang, P., Zhang, L., Niu, X.: A combined forecasting model for time series: application to short-term wind speed forecasting. Appl. Energy **259**, 114137 (2020)
14. Wang, J.-Z., Wang, Y., Jiang, P.: The study and application of a novel hybrid forecasting model - a case study of wind speed forecasting in china. Appl. Energy **143**, 472–488 (2015)
15. Inman, R.H., Pedro, H.T.C., Coimbra, C.F.M.: Solar forecasting methods for renewable energy integration. Prog. Energy Combust. Sci. **39**(6), 535–576 (2013)
16. Das, U.K., et al.: Forecasting of photovoltaic power generation and model optimization: a review. Renew. Sustain. Energy Rev. **81**, 912–928 (2018)
17. Sun, B., Cheng, W., Bai, G., Goswami, P.: Correcting and complementing freeway traffic accident data using mahalanobis distance based outlier detection. Tehnicki Vjesnik-Technical Gazette **24**, 1597–1607 (2017)