Systematic Analysis of Power Dispatching: Using Machine Learning

Yongjin TANG[a], Xuehua TANG[b], Yang YANG[a,1], Zhiyong ZHOU[b] and Yaochao DENG[a]

[a] *Mechanics Institute, Shanghai DianJi University, Shanghai, China*
[b] *School of Design and Art, Shanghai DianJi University, Shanghai, China*

Abstract. With the proposal of the concept of carbon neutrality in China, the market will continue to pay more attention to sustainable energy and energy conversion. The management of the power system is a very important part in this field due to the discontinuity of wind energy and solar energy, which is also called power dispatching. It is particularly important in the process of power system conversion and management, so it is very helpful to understand the development and bottleneck of power system to understand the main research content of power dispatching through literature. Based on machine learning and text mining, this paper analyses the related research of power dispatching. The results show that the related papers of power dispatching mainly include safe operation, power dispatching strategy, power dispatching automation, power dispatching technology analysis and so on. This paper also shows the results by a clear visualization method.

Keywords. Power dispatch, machine learning, text mining, visualization

1. Introduction

Literature review is a comprehensive analysis and research on specific fields. On the early stage, it mainly includes the selection and collection of papers, on the basis of a large number of papers, it summarizes the development trend of this field. The main research area of this paper is power dispatching.

With the rapid development of China's economy and the acceleration of urbanization, people's demand for electric energy is growing rapidly, which requires the State Grid and its subordinate units to make detailed planning for electric power work. Today, with the rapid development of the concept of sustainable development, in addition to the traditional thermal power and hydropower, wind power and solar power generation account for an increasing proportion in the power system. Gao [1] pointed out that the installed capacity of renewable energy increased from 814GW in 2004 to 2195GW in 2017, with an average annual growth of 7.9%, this requires reliable management and distribution of power grid to ensure the security, reliability, stability and economy of power system. In order to improve the overall control performance of nonlinear systems, Song [2] proposed an optimal control method based on hybrid system framework, the decision variables of the method included not only the allowed

---

continuous control, but also the scheduling of subsystem modes. Geev [3] proposes a new method for active distribution network operation in a distribution market environment with multiple wind turbine configurations, which evaluates the effects of multiple wind turbine configurations on the wind capacity available to be injected into the grid and the marginal price of distribution locations across the grid. Hao et al. [4] set up the power dispatching control monitoring communication network, it improves the level of power dispatching and the security of data transmission.

Power dispatching processing needs to be based on the data which are feedback from the information acquisition equipment, combined with the actual power grid execution parameters (load, frequency, current, voltage, etc.), ensure the safe and stable operation of power grid. Therefore, the study of power dispatching network is the basis to ensure the security and stability of power grid.

Because there are many related literatures, it is impossible to sort them out by hand. Therefore, we use machine learning algorithm to analyze the literature. We use "power dispatching" as the key word to search for relevant papers on Web of Science Core collection (WOSCC) and China national knowledge internet (CNKI). There are more than 3000 papers are published every year in WOSCC, In 2020, it's 2908 papers. CNKI published 5845 papers in 2006-2021, and 634 papers in 2017. In the previous paper, machine learning algorithm is mainly used to study the technical problems in power dispatching. Nourianfar et al. [5] introduces a new evolutionary algorithm, namely, the improved exchange market (Mema) algorithm, which solves various problems in the operation of power system. It improves the sustainability of power dispatching and considers the economic problems of power system. Basu [6] adopts the quasi-adversarial differential evolution method, which aims to reduce active power loss, improve voltage distribution and enhance voltage stability by looking for control variables such as generator terminal voltage, transformer tap setting and reactive power output of parallel reactive compensator. Liao et al. [7] established a collaborative optimization model of generation and maintenance of distributed renewable energy intelligent grid based on the analysis of the optimization model of generation plan and maintenance plan. The real-time algorithm based on distributed control strategy can not only realize dynamic compensation for random fluctuation of renewable energy generation, but also satisfy the load curve optimization under the premise of making full use of power supply resources.

2. Algorithm

The main research ideas are as follows: (1) Get the data; (2) Clean up the collected data and delete the incomplete data records; (3) Segment the abstract and key words to get nouns and delete meaningless words; (4) Get word vector by Natural language processing (NLP) technology [8]; (5) Using dimension reduction technology to reduce and visualize keywords; 6. Analyze other information of the article.

NLP is the use of computers to process, understand and use human languages (such as Chinese, English, etc.), it belongs to a branch of artificial intelligence, is the interdisciplinary of computer science and linguistics, also known as computational linguistics. This paper mainly uses text mining technology. Text mining is to obtain valuable information and knowledge from text data. It is a data mining method. The most important and basic application of text mining is text classification and clustering. The former is a supervised mining algorithm, and the latter is an unsupervised mining

algorithm. Text mining is a multidisciplinary field, covering data mining technology, information extraction, information retrieval, machine learning, natural language processing, computational linguistics, statistical data analysis, linear geometry, probability theory and even graph theory.

In the process of text mining, first of all, we need to segment the words in the article, divide the words in the article through the probability of content appear, and distinguish the part of speech of the words, so as to facilitate the subsequent processing. After word segmentation, machine learning technology is used to get the corresponding word vector. Word vector technology is to transform words into dense vectors. For similar words, the corresponding word vectors are also similar. Commonly used techniques include SVD, cbow, word2vec, etc. due to the richness of vocabulary, the dimension of each word vector is very high. In order to facilitate the analysis, this paper uses the dimension reduction method to reduce the word vector and display it. At the same time, we also use word cloud, thematic River map, scatter diagram and other ways to analyze the relevant information of the article.

3. The Specific Method

Firstly, this paper collects relevant articles from web of science core collection (WOSCC), searches with "power dispatch" as the keyword, and searches with "power dispatch" as the keyword in China National Knowledge Infrastructure (CNKI) from 2006 to 2021. The numbers of articles are show in table 1.

Table 1. Numbers of articles (2006-2021).

Years	CNKI (numbers)	WOSCC (numbers)
2006	34	300
2007	160	492
2008	214	588
2009	202	726
2010	197	746
2011	300	846
2012	394	1088
2013	504	1576
2014	565	1864
2015	521	2284
2016	462	2870
2017	634	3102
2018	516	3614
2019	403	3674
2020	290	2908
2021	89	698

The data obtained from WOSCC includes the title, abstract, year, published journal and other information of the article. The data obtained from CNKI includes the title, abstract, keywords, year, published journal, author unit and other information of the article.

In the process of word segmentation and word tagging in English literature, the software we use is TextBlob, which is a Python text processing tool, which is used to

process text data. It provides a simple API for the common NLP tasks, such as word tagging, noun phrase extraction, emotion analysis, classification, translation, etc. In the analysis of Chinese literature, the procedure we adopt is LAC, which is full-scale analysis of Chinese. It is a Lexical Analysis tool developed by Baidu NLP (Natural Language Processing Department), which can realize the functions of Chinese word segmentation, word tagging and special name recognition. In the process of dealing with the literature, we mainly focus on the nouns and the proper nouns in the abstract, and change the plural nouns to nouns and count them.

After obtaining the words, we use the GoogleNews-vectors-negative300 model, which is a model that Google uses a lot of computing power to train and contains 300 million common word vectors based on Google News large corpus. Each word vector has 300 dimensions and can be imported by Python through Gensim package and other tools. When dealing with Chinese word vector, we use Tencent AILab Chinese Embedding file, which provides more than 8 million Chinese words and phrases with 200 dimensional vectors (representation, also known as embedding), has been pre trained on large-scale high-quality data. These vectors capture the semantics of Chinese words and phrases, and can be widely used in many downstream Chinese processing tasks (such as named entity recognition and text classification) and further research.

After the word vector is obtained, we use text mining technology to unsupervised clustering of word vectors, and use Contour method to determine the number of K-means clustering categories. Finally, the word vector is reduced by T-sne technology and displayed through plotly package.

4. Results

After obtaining the nouns and proper nouns in the abstract, we show them in the way of word cloud, as shown in figures 1-2. At the same time, data statistics are carried out, as shown in tables 2-3.

In English literature, except for power, system, energy, dispatch and other nouns, optimization, generation and control appear most frequently and increase year by year. Song et al. [9] propose a novel Multi-objective adaptive Yin-Yang pair optimization algorithm to coordinatively optimize fatigue load distribution and active power dispatch. Hong et al. [10] studied the robust fixed-time consistency problem of nonlinear multi-agent systems with uncertain perturbations under weighted undirected topology, and proposed some nonlinear control protocols under which multi-agent systems could guarantee the fixed-time consistency. Yin et al. [11] propose a deep differential dynamic programming method that can be rejected to solve the real-time integrated generation scheduling and control problem of microgrid.
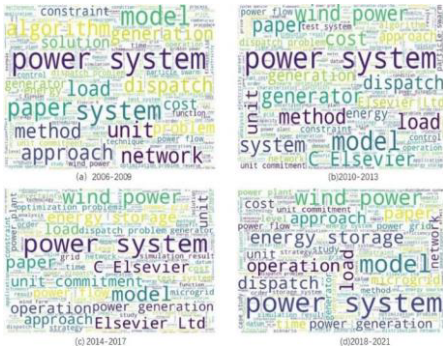
Figure 1. English abstract word cloud.

**Figure 2.** The abstract word cloud of Chinese Literature.

Table 2. Changes in English abstract vocabulary (2006-2021).

Words	2006-2009	2010-2013	2014-2017	2018-2021	All
Power	4782	9634	27030	30632	**72078**
System	4258	8142	21764	24690	**58854**
Dispatch	1804	3542	8320	8492	**22158**
Generation	1576	3264	7896	8028	**20764**
Optimization	1232	2754	6392	7586	**17964**
Load	1184	2212	6538	7352	**17286**
Control	804	1662	5202	6044	**13712**
Network	916	1346	4148	5510	**11920**
Demand	490	1372	4194	4744	**10800**
Electricity	700	1450	3660	4416	**10226**
Distribution	372	810	2854	4090	**8126**
Technique	690	1028	2210	2050	**5978**
Function	710	1242	2436	2594	**6982**
Energy	1084	3252	11794	17082	**33212**

Table 3. Changes of Chinese abstract vocabulary (2006-2021).

Words	2006-2009	2010-2013	2014-2017	2018-2021	All
电网 (Power grid)	411	1530	2306	1720	**5967**
技术 (Technology)	276	747	1302	1093	**3418**
电力系统 (Power system)	71	456	1352	807	**2686**
网络 (Network)	232	448	791	621	**2092**
模型 (Model)	80	139	285	644	**1148**
过程 (Process)	42	203	525	319	**1089**
需求 (Demand)	49	159	463	400	**1071**
功能 (Function)	114	257	279	234	**884**
设备 (Equipment)	93	199	359	362	**1013**
措施 (Measures)	48	217	515	219	**999**
经济(Economics)	57	208	412	241	**918**
能源 (Energy)	23	91	245	307	**666**
安全性 (Security)	24	101	309	194	**628**
结构 (Structure)	123	163	197	179	**662**

In the English word vector processing, the K value is determined as 4 by contour method which is shown in figure 3, and the dimension of the word vector is reduced by T-sne, as shown in figure 4. It can be seen that in the English literature, the research contents can be divided into four categories. The first is system optimization, such as algorithm, optimization, transmission, etc. Such as Papalexopoulos et al. [12] proposed a market design to implement an organized nodal electricity market for distribution networks. Hassan [13] proposed a mathematical model optimization problem and genetic algorithm to solve the system reliability optimization problem of multi-source and multi confluence network under transmission budget constraints. The second type is power system problem solving, such as problem, solution, cost, result and operation. Such as Zou et al. [14] proposed a cooperative method, which makes full use of the advanced characteristics of RL method, greatly reduces the operation and investment costs, and effectively reduces the computational burden. Savelli et al. [15] based on the traditional problem of how non flexible consumers coexist with flexible users in the local distribution area, proposed a method that flexible users will pay the node price, while non flexible users will charge a fixed price from the potential node price. The third is the proper term, such as ED, OPF, Bess. In order to solve the economic scheduling problem of the energy Internet composed of energy routers and smart microgrids, Wu et al. [16] proposed a distributed economic scheduling algorithm based on leader-follower consensus and average consensus protocol to ensure that the economic scheduling of all distributed generators was asymptotically optimal. MAN A et al. [17] proposed a multi-objective security-constrained microgrid energy management system based on the coordinated unit commitment-optimal power flow framework, which aims to minimize the actual power loss, voltage deviation and generation cost. Hinokuma et al. [18] proposed a solar wind hydrogen hybrid system based on fuzzy logic control (FLC) algorithm to reduce the load demand of buildings; Kuang et al. [19] established a dad-ahead economic dispatch model for the integrated electricity and natural gas system (IEGS) with the consider of the high proportion of wind power, and used extreme scenario optimization method to deal with the uncertainty of wind power.
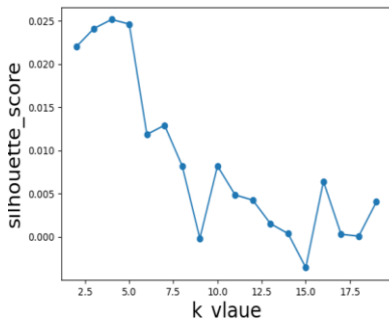
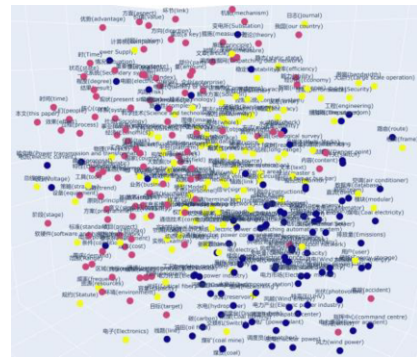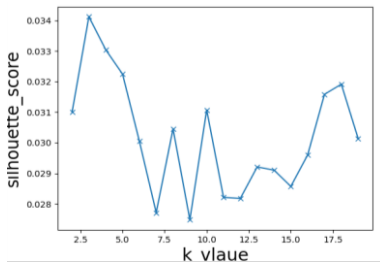**Figure 3.** K-value determination of English word vector.
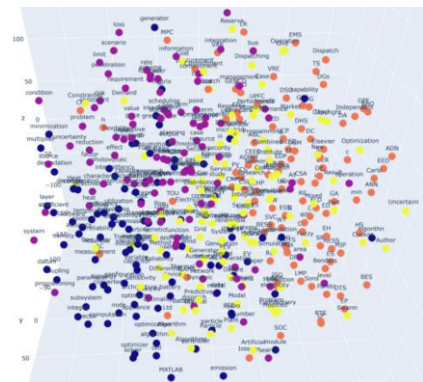
Figure 4. Visualization of corresponding English word vector.

In Chinese word vector processing, K value is determined to be 3 by contour method which is shown in figure 5, and T-sne is used to reduce the dimension of word vector, as shown in figure 6. It can be seen that in the Chinese literature, the research

contents can be divided into three categories. The first category is about the research objects, such as power system, wind power, fault and substation. Ma et al. [20] analyzed and studied the influence of different factors on the output characteristics of wind turbines, and the influence of wind turbine integration on power flow, short circuit and stability of power system. Qiao et al. [21] studied the fault problems in the process of substation operation, deeply analyzed the causes of the fault of reactive power compensation device, and improved the reactive power compensation device to make it fully play its role and operate safely and reliably. The second type is about solutions and measures, such as demand, efficiency, characteristics and market. Zhang et al. [22] deeply analyzed the current situation and challenges of demand response development in China, and combined with the construction path of China's power market, designed the market-oriented trading mechanism of demand response for China's power market, which provided a feasible idea for China's demand response development. Wang et al. [23] analyzed the feasibility, key technology, development demand and possible economic and social benefits of building full DC power supply and distributed power storage. The third category is about power dispatching technology, such as model, algorithm, stability and reliability. For interconnected wide area power system, to solve the time-varying robust stability problem of uncertain load frequency control systems with multiple time delays and external power system disturbances, Wang [24] proposed a time-varying and time-varying stability criterion. Zhang [25] in order to avoid the complexity of short-term load prediction model by traditional methods, adopts improved genetic algorithm and BP hybrid algorithm to train artificial neural network, and makes simulation prediction of short-term load according to the past power load data.

In addition, we analyze the source of the paper. The sources of English literature and Chinese literature are shown in figures 7-8 respectively. It can be seen that most of the English articles related to power dispatching are published in IEEE TRANSACTIONS ON POWER SYSTEEMS, while most of the Chinese articles are published in the Communication World.

Figure 5. K-value determination of Chinese word vector.

Figure 6. Visualization of Chinese word vector.

We classify the keywords and display them in the way of theme river and scatter, the results are shown in figures 9-10. It can be seen that from 2013 to 2018, the research on power dispatching automation is very common, and many articles are related to safe operation, such as, Juang et al. [26] propose a new affine adjustable robust model to solve the safety constrained unit commitment problem considering

these uncertain sources. Huang et al. [27] built a cascade hydropower photovoltaic pumped storage (CH-PV-PS) system model with multiple objectives and constraints to improve the reliability and hydraulic utilization of the system and maximize its economic benefits. Li et al. [28] Considering the multiple uncertainties of wind power forecasting error, load forecasting error and conventional blackout, a method to determine the reserve capacity of power system with wind power is proposed, and the quantitative relationship between the upper limit of reserve capacity and load loss probability (LOLP) is derived, which increases the stability of power system operation.
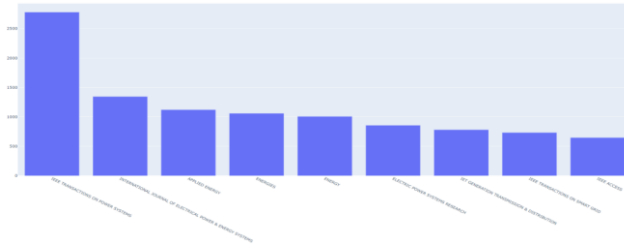
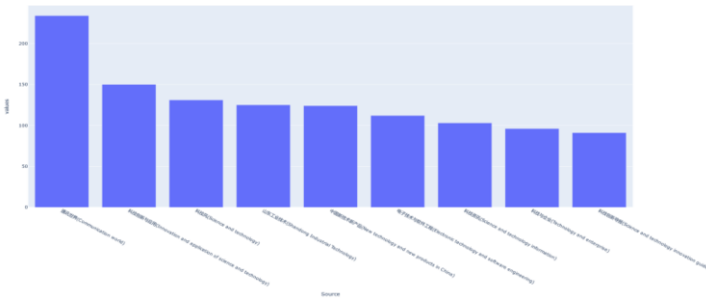**Figure 7.** Distribution of English literature sources.

Figure 8. Distribution of Chinese literature sources.
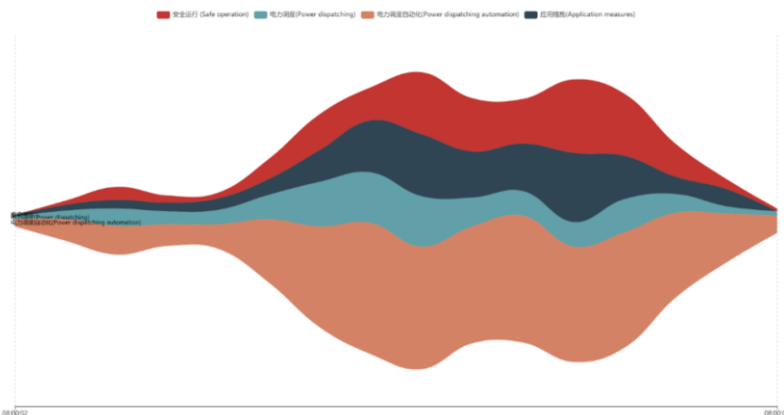

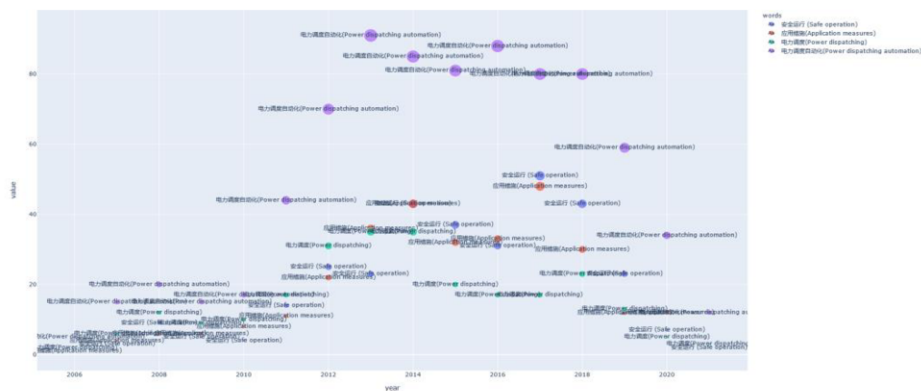**Figure 9**. Chinese keywords theme River Map.

Figure 10. Scatter chart of Chinese keywords.

5. Summary

In this paper, we use machine learning algorithm and a variety of visualization methods to deal with the power dispatching related articles, and show the results according to the relevant data. Through our analysis, we can effectively understand the research content and research trend of power dispatching related articles. Through our analysis, we can see that in English literature, it mainly includes system optimization, power system problem solving, the proper term, research objects. In Chinese literature, it mainly includes research objects, solutions and measures, power dispatching technology. At the same time, we can see that the research of power dispatching automation is more popular. However, we still need to improve the part of text mining, such as using named entity recognition to extract the relationship between nouns, which will be the focus of our next research.

Acknowledgments

References

[1]   Gao M Z, Fan C T and Chen J S 2020 A critical review of the world's first renewable portfolio standard (RPS) for large electricity users in Taiwan: The return of the RPS? *Energy Strategy Reviews* **32** 100585.

[2]   Song C, Bing W and Ping L 2012 A hybrid model-based optimal control method for nonlinear systems using simultaneous dynamic optimization strategies *Journal of Process Control* **22** 852-860.

[3]   Geev M 2017 Active distribution networks operation within a distribution market environment

Sustainable Development in Energy Systems pp 107-118.

[4]   Hao X D, Wu M and Sun C S 2011 Design of the electric-power dispatch network structure of Xinwen mining areas *Proceedings of 2011 International Conference on Computers, Communications, Control and Automation Proceedings (CCCA 2011 V3)* (Institute of Electrical and Electronics Engineers) pp 123-125.

[5]   Nourianfar H and Abdi H 2021 Solving power systems optimization problems in the presence of renewable energy sources using modified exchange market algorithm *Sustainable Energy Grids and Networks* **26** 100449.

[6]   Basu M 2016 Quasi-oppositional differential evolution for optimal reactive power dispatch *International Journal of Electrical Power & Energy Systems* **78** 29-40.

[7]   Liao L and Ji C 2020 Smart grid dispatching optimization for system resilience improvement *Complexity* **2020** 1-12.

[8]   Li W, Zhao Y and Liu N, et al. 2021 Strain gradient induced grain refinement far below the size limit in a low carbon hypoeutectoid steel (0.19 wt% C) via pipe inner surface grinding treatment *Journal of Materials Science & Technology* **78** 155-169.

[9]   Song D 2020 Comprehensive optimization for fatigue loads of wind turbines in complex-terrain wind farms *IEEE Transactions on Sustainable Energy* **99**.

[10] Hong H, Yu W and Wen G 2017 Distributed robust fixed-time consensus for nonlinear and disturbed multiagent systems *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **47** 1464-1473.

[11] Yin L and Zhao L 2021 Rejectable deep differential dynamic programming for real-time integrated generation dispatch and control of micro-grids *Energy* **225** 120268.

[12] Papalexopoulos A, Frowd R and Birbas A 2020 On the development of organized nodal local energy markets and a framework for the TSO-DSO coordination *Electric Power Systems Research* **189** 106810.

[13] Hassan M R 2021 System reliability optimization of multi-source multi-sink stochastic flow networks with budget constraint *International Journal of Reliability Quality and Safety Engineering* **28** 2150025.

[14] Hza B, Jta B and Skec D, et al. 2021 Stochastic multi-carrier energy management in the smart islands using reinforcement learning and unscented transform *International Journal of Electrical Power & Energy Systems* **130** 106988.

[15] Savelli I and Morstyn T 2021 Electricity prices and tariffs to keep everyone happy: A framework for fixed and nodal prices coexistence in distribution grids with optimal tariffs for investment cost recovery *Omega* **103** 102450.

[16] Chen W and Li T 2018 A distributed economic dispatch algorithm based on multi-agent consensus control and incremental power supplying *IFAC-PapersOnLine* **51** 7-12.

[17] Man A, Sn B, Gbg A, Na C and Shh A 2019 A multi-objective voltage stability constrained energy management system for isolated microgrids *International Journal of Electrical Power & Energy Systems* **117** 105646.

[18] Hinokuma T, Farzaneh H and Shaqour A 2021 Techno-economic analysis of a fuzzy logic control based hybrid renewable energy system to power a university campus in Japan *Energies* **14** 1960.

[19] Kuang C, Xiao M and Chen Z, et al. 2021 Distributed optimal dispatch of integrated electricity and natural gas system considering the pipeline storage characteristics *Evolutionary Intelligence* **2021** 1-11.

[20] Ma D and Liu P 2020 Application of wind turbine modeling in power system stability research *Electric Power Survey and Design* **03** 64-69.

[21] Qiao X, Liu F, Tao B, Wang S and Fang L 2018 Fault analysis and improvement measures of reactive power compensation device in substation *Electrical Technology* **19** 83-86.

[22] Zhang G, Xue S, Fan M and Zhang H 2021 Design of market-oriented transaction mechanism for demand response in China's power market *Electric Power Construction* **42** 132-140.

[23] Wang F and Jiang Y 2016 Key technology and benefit analysis of building full DC power supply and distributed power storage *Building Electrical* **35** 16-20.

[24] Ramakrishnan K and Ray G 2015 Improved results on delay-dependent stability of LFC systems with multiple time-delays *Journal of Control Automation & Electrical Systems* **26** 235-240.

[25] Zhang Y 2009 Research of short-term power load forecasting based on improved GA-BP neural network model *Computer Engineering and Applications* **45** 223-226.

[26] Aguilar J E S, Marín-Cano C C, López-Lezama J M and Duque Á J 2021 A new affine1y adjustab1e robust mode1 for security constrained unit commitment under uncertainty *Applied Sciences* **11** 3987.

[27] Huang X, Wang J, Huang T, et al. 2021 An optimal operation method of cascade hydro-PV-pumped storage generation system based on multi-objective stochastic numerical P systems *Journal of Renewable and Sustainable Energy* **13** 16301.

[28] Li Z and Liu C 2021 Research on calculation of spinning reserve capacity of wind power system considering multiple uncertainties *IOP Conference Series: Earth and Environmental Science* **632** 32056.