Research on China's Green Finance Policies Based on Text Mining

Yuqing Liao¹, Jingliang Chen²

¹ University of Shanghai for Science and Technology, Shanghai, 200092, China

² University of Shanghai for Science and Technology, Shanghai, 200092, China

Abstract. Based on the green finance policies in China from 2017 to 2019, this paper extracts feature and high-frequency words from policy documents, uses word cloud diagram, co-occurrence matrix and social network analysis techniques to quantitatively analyse the information contained in the green finance policies over the past three years and highlights the hot issues in question, thus providing a multi-layered and wide-ranging pathway for facilitating the orderly development of green finance industries across China.

1 Introduction

Since the reform and opening up, China's economy has perpetuated robust growth and topped the world for this. However, in the wake of this, a large number of natural resources have been exploited and the ecological environment has been seriously damaged. In an effort to cope with climate change and conserve and efficiently utilize natural resources, China has begun exploring the road to green economy. At the core of modern economy, finance is of vital importance in resource allocation. Green finance can steer funds into environmental protection industries gradually from high energy consumption and high pollution industries, vigorously promoting sustainable economic development and industrial upgrading and transformation. [1] In recent years, with the rollout of the "Belt and Road Initiative" and the rural revitalization strategy, the policy system of green finance has gradually been amplified. On this basis, this paper mainly analyzes and studies the green finance policies from 2017 to 2019, explores the development emphases of green finance industries under the guidance of the policies, to demonstrate the development pattern of green finance policies in last three years.

2 Literature Review

Green finance, in essence, organically integrates a portfolio of financial products, market systems and government regulatory measures concerning green investment and financing projects and risk control. [2] Nowadays, in face of heavy environmental quality tasks and large demand for environmental protection investment, it is the right time to construct a green financial system in China. Combining relevant green finance policies and systems in China, it can be seen that the green finance has been ascended as the national

strategic goal and underpinned by relevant policies and regulations that have been promulgated by the nation for green finance development. [4] But compared with developed countries, problems like backward laws and regulations system, limited subsidy of financial funds and insufficient role of policy institutions are still prevalent. [5] For this, it can be a pivotal step to enhance measurable quality supervision indicators [6] and perfect green financial system by stimulating government to push forward green industries and green finance and guaranteeing efficient information communication and consistent green finance policies.

2.1 Review on Text Mining Techniques

In the quantitative researches of policies, text mining techniques have been widely applied, and have contributed to a wealth of research results especially based on diverse policy documents. Computer methodology was employed later and proves to be quite helpful in increasing the number of materials and processing accuracy handled by text mining techniques, along with new approaches and concepts. [7] Text mining techniques can extract hidden knowledge from vast policy data, interpret and learn policy stances and tendencies, and analyze policies in broad sense. [8] Content analysis method is characterized by systematicness, objectivity and rationality (Hu Yanran) [9]. Based on text mining, content analysis method has been used for quantitative analysis and research of the fiscal and taxation policies for railway transportation enterprises in China (Zhao Gongmin) [3]. Grounded theory and text mining research methods have been leveraged to analyze the word frequency, semantic network and centrality of policy texts, and to study the operation mode of fintech in Guangdong Province.

Through analysis of the existing literature, it is found that most of the researches on green finance policies are qualitative yet fewer are quantitative. Therefore, this

*Corresponding author's e-mail: liaoYuqing@usst.edu.cn

paper explores the development emphases of China's green finance industries and demonstrates the development patterns in the past three years.

3 Data Source and Research Methodology

3.1 Data Source

This paper mines its data from the green finance policy documents in the website of the Lawstar using the "Bazhuayu" data scraping tool that have collected the required data from green finance policy documents spanning from 2017 to 2019. But because of the access restrictions set by the website to prevent individuals or enterprises from gaining benefits illegally, only 1500 documents are collected, including central policy documents, departmental regulations and provincial/municipal policy documents. After combing and removing the overlapped texts, 1487 are available. Some of the policy documents are listed in table 1.

table 1. Example of the Green Finance Policy Documents

No.	Policy Documents
1	Notice of Nantong Municipal People's Government on Forwarding the Implementation Opinions of Nantong Banking Regulatory Bureau and Other Six Departments on Accelerating the Development of Green Finance and Bolstering the Construction of Ecological Civilization in Nantong
2	Notice of the Ministry of Agriculture and the Agricultural Bank of China on Promoting Financial Support for Green Agriculture Development, No.63 [2017] of the Ministry of Agriculture, for all provinces, autonomous regions and direct-administered municipalities
3	Implementation Opinions of the People's Government of Jiangxi Province on Accelerating the Development of Green Finance, No.37 [2017] of the People's Government of Jiangxi Province, for Municipal and County (District) People's Governments and Provincial Government
4	Notice of the People's Government of Jiangxi Province on Issuing the Plan for Establishing a Green Financial System in Jiangxi during the 13th Five-Year Plan, No. 79 [2017] of the People's Government of Jiangxi Province
5	Notice of Sanming Municipal People's Government on Issuing the Implementation Plan for Establishing a Green Financial System in Sanming, No. 12 [2017] of Sanming Municipal People's Government, Green Financial System in Sanming
6	Notice of Chengde Municipal People's Government on Issuing the Plan for Establishing a Green Financial System in Chengde, No.176 [2017] of Chengde Municipal People's Government
7	Notice of Beijing Municipal Bureau of Finance on Issuing the Implementation Measures for Establishing a Green Financial System in Beijing, No.152 [2017] of Beijing Municipal Bureau of Finance
8	Implementation Opinions of Ulanqab Municipal People's Government on Establishing a Green Financial System, No. 106 [2017] of Ulanqab Municipal People's Government, for administrative divisions of Inner Mongolia
9	Notice of Zhangzhou Municipal People's Government on Issuing the Implementation Plan for Establishing a Green Financial System in Zhangzhou, No.112 [2017] of Zhangzhou Municipal People's Government, Green Financial System in Zhangzhou
10	Guiding Opinions of Nanping People's Government on the Implantation Plan for Establishing a Green Financial System in Fujian Province, No.145 [2017] of Nanping People's Government

3.2. Research Methodology

It can be effort-taking and time-starved to comb the key points of substantial documents simply by reading and recording them one by one personally. The research leverages ROST CM6 and UCINET to extract high-frequency words from China's green finance policy documents issued between 2017 and 2019 and form a co-occurrence matrix and social semantic network, while visualizing text data with word cloud to demonstrate the development patterns of China's green finance policies in recent three years.

4 Extraction and Stocktaking of Key Words from Policy Documents in Recent Three Years

First, the author deleted all the meaningless words such as issuing agencies, attachment processes and names after excluding overlapped documents. Second, because the customized and stop vocabulary in the present content mining system have little correlation with the research subject of this paper, the author added "green credit", "credit information" and "carbon finance" into the customized vocabulary, and repeatedly revised and updated the vocabulary using the ROST with word classification function based on the understanding of the policy documents collected earlier, so as to ensure the validity of the analysis results. Finally, by classifying the words and calculating word frequency in the policy

documents of recent three years, and the keywords of co-occurrence are obtained as shown in Table 2

Table 2. Examples of Word Co-occurrence in Green Finance Policy Documents

2017		2018		2019	
Key words	Word Frequency	Key Words	Word Frequency	Key Words	Word Frequency
Innovation	6327	Innovation	6242	Innovation	5172
Agriculture	4484	Resources	4984	Resources	4314
Reform	4082	Technology	4975	Industry	3928
Financial institutions	2526	Industry	4127	Supervision	3727
Agricultural products	1375	Agriculture	3722	Environment	3645
Financial services	1263	Ecology	3529	Technology	3502
Elderly care services	1163	Supervision	3526	Safety	2926
Big data	1033	Development	3409	Agriculture	2839
Environmental protection	998	Financial institutions	2347	Ecology	2737
SMEs	921	Tourism	2336	Education	2621

5 Density Analysis of Co-occurrence Matrix in Recent Three Years

The author tries to show the correlation between the keywords in the policy documents of recent three years to underlie the core content of these policy documents. Through statistics of the co-occurrence frequency of the keywords in the policy documents of the same year, the co-occurrence matrix is built and shows that the higher the

co-occurrence frequency is, the closer the two keywords will be. However, it is far from enough to reflect the correlation between words simply by frequency, so it's necessary to convert native co-occurrence matrix into similarity matrix. The numerical range in the similarity matrix is [0, 1], and the closer the numeric is to 1, the more correlated the two keywords will be. A good example is 2017 Policy Text-based Similarity Matrix, as shown in Table 3.

Table 3. Similarity Matrix of the Green Finance Policy Documents in 2017

	Innovation	Agriculture	Environment	Reform	Insurance	Planning	Social security	Funds	Supervision	Green	Technology	Operation
Innovation	1.000	0.439	0.943	0.943	0.500	0.884	0.986	0.928	0.841	0.885	0.608	0.797
Agriculture	0.439	1.000	0.466	0.466	0.887	0.481	0.433	0.459	0.522	0.496	0.737	0.551
Environment	0.943	0.466	1.000	1.000	0.530	0.938	0.929	0.985	0.892	0.939	0.644	0.845
Reform	0.943	0.466	1.000	1.000	0.530	0.938	0.929	0.985	0.892	0.939	0.644	0.845
Insurance	0.500	0.887	0.530	0.530	1.000	0.547	0.493	0.522	0.594	0.565	0.769	0.627
Planning	0.882	0.481	0.938	0.938	0.547	1.000	0.901	0.955	0.920	0.936	0.665	0.872
Social security	0.986	0.433	0.929	0.929	0.493	0.901	1.000	0.944	0.829	0.873	0.599	0.786
Funds	0.928	0.459	0.985	0.985	0.522	0.955	0.944	1.000	0.879	0.925	0.635	0.832
Supervision	0.841	0.522	0.892	0.892	0.594	0.920	0.829	0.879	1.000	0.951	0.722	0.911
Green	0.885	0.496	0.939	0.939	0.565	0.936	0.873	0.925	0.951	1.000	0.686	0.900
Technology	0.608	0.737	0.644	0.644	0.769	0.655	0.599	0.635	0.722	0.686	1.000	0.762
Operation	0.797	0.551	0.845	0.845	0.627	0.872	0.786	0.832	0.911	0.900	0.762	1.000

Further, the keyword network is calculated with a density range of [0, 1]. The closer the density is to 1, the more closely the nodes will be linked and the more complex the network will be. For this, UCINET's density tool is leveraged to measure the network density of policy documents from 2017 to 2019, deriving 0.6120, 0.5138 and 0.4691 respectively, which shows that the green finance policy documents in the past three years feature a complex keyword network, wherein the nodes are closely

linked and the document topics are rich. The mean variance is 0.2039, 0.4818 and 0.4729 respectively, which indicates that compared with 2017, the keyword networks of 2018 and 2019 maintain a small-range density, and the green finance policies have clear development themes.

6 Analysis of Green Finance Policies Based on Text Feature Words by Word Cloud Diagram and Social Semantic Network

Given the large number of policy documents and the higher thematic dimensions, in order to better visualize the

high-frequency keywords, this paper first extracts and quantifies text feature words from policy documents to convey the information in the documents and removes meaningless keywords, as shown in Table 4. The text feature words from 2017 to 2019 are then imported into the WordArt to build the corresponding word cloud diagram.

Table 4. Example of Text Feature Words in Green Finance Policy Documents from 2017 to 2019

2017	Entrepreneurship, innovation, housing, strengthen, direct financing, energy conservation, green, pollution, information sharing
	Government-led, innovation, strengthen, water conservation, intellectual property, green, pollution, information sharing, consumption
	Financing service, positioning, reform and innovation, social security, clustering, insurance, protection, financial industry, water, environment
2018	Fiscal, perfect, innovation, investment, energy conservation, green, pollution, SMEs, services
	Certification, encourage, loan, coordination, financing, environmental protection, monitoring, supervision
	Monitoring, supervision, approval, green development, green finance, training, fund, China, reform
2019	Housing, agricultural and rural, banking, environment, supervision, green development, green finance, reform, ecology
	Housing, green, information sharing, services, agricultural and rural, banking, new energy, energy conservation and environmental protection, financing
	Environment, supervision, green building, dynamic, fund, reform, ecology, digital, resources

After the collated policy text feature words are imported into the WordArt, the keywords of highest frequency will be highlighted in the word cloud diagram. The “UCINET → NETDAW” function is employed to create the social semantic networks for the green finance policy documents in recent three years respectively, in

which nodes represent keywords and the lines represent the co-occurrence relationship between keywords.

6.1 Analysis of Green Finance Policies in 2017 by Word Cloud Diagram and Social Semantic Network

Figure 1. Word Cloud Diagram of Green Finance Policies in 2017

Figure 2. Social Semantic Network of Green Finance Policies in 2017

Table 5. High Frequency Words of Green Finance Policies in 2017

Tourism	Assessment	Reform
Pollution	China	Sound
Pension	Sharing	Green
Intelligent	Supervise	Supervision
Environmental protection	Training	Operation
Forestry	Science	Fund
Housing	Specification	Internet
Agricultural products	Evaluation	Capital
Energy conservation	Public service	Planning
Consumption	Ecology	Financial institutions
Credit	Specialization	Talent
Financial services	Social capital	Credit loan
Breeding	Improve	Financing
Farm	Incentive	Banking
Elderly care services	Innovation	Education
Strengthen	Strengthen	Land
Finance	Environment	Guarantee
Big data	Cultivation	Urban and rural
Environmental protection	Renovation	Entrepreneurship
Reward	Technology	Agriculture
SMEs	Transformation and Upgrading	Industry
Intellectual property	Social security	Culture

Early in 2017, China’s green finance policies had a clear infrastructure. It can be seen from Figure 1 that the words like “innovation”, “Internet”, “big data”, “technology”, “supervision”, “evaluation and assessment”, “financial institutions”, “insurance”, “credit loan”, “fund”, “guarantee”, “education” and “agriculture” are all highlighted in the word cloud diagram. Combined with Figure 2, it is shown that with the development and

maturity of IT, the green financial system also needs to cater to the development of society and give priority to reform and innovation. Based on the previous leading green financial products such as green credit, green insurance, green bonds and green intermediary services, Internet and big data technology are more often emphasized for innovation of green financial products. The government plays an important role in the development of green finance. A scientific constraint

mechanism is required more by policy finance than the traditional one, to strengthen the supervision on the industrial institutions involved and advance credit information sharing, and promote the assessment of green finance industries in view of participation qualifications and performance. All these keywords display the leading products of China's green finance. Fueled by the "rural revitalization" strategy put forward in 2017, China's green finance policy makers began setting their sight towards rural agriculture, greatly improving the allocation of land

and resources. With the established and perfected specialized mechanism for financial institutions as the underpinning, it's important to strengthen financial talent education primarily by the scientific research centers in universities.

6.2 Analysis of Green Finance Policies in 2018 by Word Cloud Diagram and Social Semantic Network

Figure 3. Word Cloud Diagram of Green Finance Policies in 2018

Figure 4. Social Semantic Network of Green Finance Policies in 2018

Table 6. High Frequency Words of Green Finance Policies in 2018

Industry	Support	Form
Building	Supervision	Features
Loan	Demonstration	People's government
Talent	Priority	Finance
Banking	Renovation	Agriculture
Transportation	Financing	Investment
Tourism	Fiscal	Strengthen
Training	Service	Finance
Credit	Perfect	Environment
Approval	Coordination	Safety
Rural	Innovation	China

It can be seen from Figure 3 that such words as “innovation”, “service”, “optimization”, “resources”, “coordination”, “environment”, “approval”, “supervision”, “assessment”, “investment”, “support”, “cultivation”, “insurance” and “agriculture” are all highlighted in the word cloud diagram. Combining Figure 4, we know that in 2018, no new contents of green financial products were added in the green finance policies based on the year earlier. However, the government played a more prominent role of guidance to motivate green financial systems at all levels to perpetuate innovation, increase investment in green financial enterprises and institution, bolster the green transformation of SMEs, and require enterprises and institutions engaged in green finance to provide premium financial services; attach importance to the environment-

6.3 Analysis of Green Finance Policies in 2019 by Word Cloud Diagram and Social Semantic Network

Figure 5. Word Cloud Diagram of Green Finance Policies in 2019

Figure 6. Social Semantic Network of Green Finance Policies in 2019

Table 7. High Frequency Words of Green Finance Policies in 2019

Pollution	SMEs	Service system	Supervision	Education
Elderly care services	Rural revitalization	Resources	Investment	Agricultural and Rural
Farmer	Belt and Road	Pilot	Sharing	Agriculture
Energy	Agricultural industry	Services	Green	Protection
Water conservation	Credit information	Capital	Economy	Culture
Energy conservation	Element	Financing	Cultivation	Intelligent
Digital	Public service	Agricultural products	Ecology	Training
Environmental protection	Specialization	Assessment	Health	Rural
Financial services	Financial Institutions	Social security	Tourism	
Demonstration zone	Big Data	Innovation	Reward	
New energy	Internet	Finance	Public	
AI	Assessment	Investment	High quality	
Credit loan	Fund	guarantee	Planning	
Transformation and upgrading	Intellectual property	Housing	Safety	
Funding	bank	Environment	Technology	
Innovation and entrepreneurship	Insurance	Reform	Sound	

It can be seen from Figure 5 that the words like “innovation”, “resources”, “services”, “supervision”, “environment”, “technology”, “training”, “financial institutions”, “education”, “public services”, “high quality”, “agricultural and rural”, “agricultural products”, “intellectual property”, “banking” and “sharing” are all underlined in the word cloud diagram. Combining Figure 6, we know that the policies of 2019, compared with those of 2017 and 2018, illustrate that green finance, as a kind of policy finance, focuses more on social benefits than traditional finance, along with more targeted policies, which seek to innovate and upgrade the green financial system by virtue of new technological instruments such as big data, Internet and AI, establish and perfect the information sharing platform for the enterprises and

institutions involved, and enhance the interconnection effect; strengthen government supervision, and motivate the social community to urge green financial enterprises and institutions to accelerate the improvement of the financial service system; prioritize green transformation of intellectual property with the popularity of intellectual property (IP); intensify efforts on financing in agricultural and rural areas, and carry out brand marketing campaigns for agricultural products and cultural tourism to substantially contribute to the rural revitalization strategy.

7 Summary

This paper elaborates on the green finance policies in China from 2017 to 2019, discussing the development

emphases of the present green financial policy system, and exhibiting the development patterns of green financial policies system in recent three years using a mixture of word cloud diagram and social semantic network. The key points are shown below:

On one side, from the perspective of policy formulation, combining the comparative analysis of the word cloud diagrams and the social semantic networks from 2017 to 2019 above, it can be seen that the green finance policies vary over time. The high-frequency keywords and the social semantic networks of the green finance policy documents of 2017 and 2019 are highlighted by “innovation”, “supervision”, “services”, “green fund” and “rural agriculture”. There are both keywords of policy guidance, and targeted measures for specific green financial products in the green financial system. While the high-frequency words and social semantic networks of the policy documents of 2018 mainly concentrate on policy guidance words such as “coordination”, “innovation” and “supervision”. The green finance policies of 2018 failed to set a clear development strategy for green financial products, but rather consolidated the policies of 2017 and prepared for the policies of 2019.

On the other side, considering the policy content, first of all, green finance, as a kind of policy finance, is greatly bolstered by the government in system construction and development. The policy documents of recent three years fully demonstrate the government’s endeavors to motivate the enterprises and institutions engaged in green finance practice, perfect the supervision system, spearhead financial institutions to increase investment in green finance, provide better green financial service, and work out sound green financial product catalog. Second, China’s green financial policy system has preliminarily taken shape. Suited to the development level of all regions and the advancement of information technology, green finance will be built into a system with Chinese characteristics, and a credit information sharing platform will be created for green enterprises to cement the bridge between all regions, and search for new social development drivers, such as intellectual property and AI to energize China’s green financial system. Third, with the introduction of the “rural revitalization” strategy, the transformation and upgrading of rural agriculture has become a part of the green financial system construction, developing branding strategy for green agricultural products and developing signature green tourism products like cultural tourism and sports tourism. Fourth, universities and scientific research institutions will be bolstered to cultivate top financial talents for sustainable development of green financial system.

This paper demonstrates the development patterns of China’s green financial policies in recent three years, but failed to analyse all the policy documents of each year due to data access restrictions, or to classify the policies available in more detailed way, for example, comparing the policies issued by different institutions. Looking forward to the future, the author will continue the research on green financial policy documents from multiple perspectives.

Reference:

1. Zou J. (2017), Green Financial Policy, Policy Coordination and Industrial Pollution Intensity: From the Perspective of Policy Document Analysis [J]. *Financial Theory and Practice*, 12: 71-74.
2. Qin Y. (2019) Wang Jing. Research on the Current Situation and Problems of China’s Green Finance [J]. *Tianjin Economy*, 12: 20-25.
3. Zhao G.; Liu J.; Wu Y, Yang F. (2019) Research on Co-word Network of Guangdong Science and Technology Financial Policy Based on Grounded Theory and Text Mining [J]. *Science and Technology Management Research*, 39 (03): 51-57.
4. Yang G., Li J., Liu H. (2017) Thinking on the Construction of the Laws and Policies on Green Finance in China [J]. *West China Finance*, (12): 90-92.
5. Zhou Z. (2017) Construction of the Green Financial System in China [J]. *CO-Operative Economy & Science*, DOI: 10.13665/j.cnki.hzjjykj.17.023
6. Gui B. (2019) Research on the Laws and Policies on Green Finance [J], *Finance & Economy*. DOI: 10.15880/j.cnki.zsjj.2019. 21.035
7. Pei L.; Sun J.; Zhou Z. (2016) Policy Document Calculation: A New Way to Interpret Policy Documents [J]. *Library & Information*, (06): 47-55.
8. Zhang X.; Zhou X.; Wang Y. (2018) Quantification and Evolution of China's Science and Technology Service Policy: Based on Grounded Theory and Text Mining Analysis [J]. *China Science and Technology Forum*, (06): 6-13.
9. Hu Y. (2016) Research on Fiscal and Tax Support Policies of Chinese Railway Transportation Enterprises Based on Text Mining [D]. Beijing Jiaotong University.