Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 1 No.2 (2024)

Knowledge Structure and Hotspot Evolution of "Greenwashing": Based on Citespace and Text Analysis Perspectives

Shenglin Ma[1], Huifang Liu[2], Chaoyang Wang[3], Hongjun Zeng[4]*

1.School of Economics and Management, North University of China, Taiyuan, 030051, China.

2.School of Economics and Management, Shandong Youth University of Political Science, Jinan, 250103, China.

3.School of Humanities and Law, Yanshan University, Qinhuangdao, 066000, China.

4.Department of Financial Planning and Tax, School of Accounting, Information Systems and Supply Chain, RMIT University, Melbourne, 3000, Australia.

Corresponding author: Hongjun zeng

Abstract: In the context of the country's promotion of ecological civilization and the fulfillment of the "dual-carbon" goal, green and sustainable development should be an important part of the enterprise's connotative development, rather than just "greenwashing" packaging. Although the research on "greening" related topics has been increasing in China, there are deficiencies in quantitative and systematic research. Therefore, in this paper, by studying the overall trend of articles published on web of science from 2013 to 2023, we use citespace to comprehensively and systematically portray the hotspot evolution law in the field of Green Washing. At the same time, we use python to partition the content of the related news Green Washingwashing text, and with the help of mining methods such as neologism discovery and LDA theme model, we identify the hotspot keywords and carry out the evolutionary analysis. The results show that the greenwashing themes mainly focus on greenwashing behavior, green development, green finance, carbon neutrality, information disclosure, social responsibility, sustainable development and environmental performance. Finally, the paper further outlines the existing research framework on greenwash from the theoretical perspective of sustainable development, and offers a perspective on the future direction of article on greenwash governance from three parts: government, business and individuals.

Keywords: Greenwashing; LDA; Text Mining; Evolutionary Analysis; Sustainable Development

Published: Apr 15, 2024

DOI: https://doi.org/10.62177/apemr.v1i2.266

1.Introduction

"Greenwashing" evolved from "whitewashing, which refers to the behavior of enterprises' false green propaganda and whitewashing for profit, emphasizing their environmental deception (Yang et al. 2020) [1]. This concept originated in 1986, the American environmentalist Jay Westerveld of the United States hotel "recycled towels" environmental protection behavior of reflection, used to describe the enterprise for the purpose of making profit and make false environmental protection propaganda behavior (Tong et al.,2025; Ma et al.,2025) [2,3].

"Greenwashing" is a key concern in addressing climate and environmental change, and despite the increasing diversity and complexity of the phenomenon, the existing literature is still insufficiently in-depth. Formally, enterprises promised to

undertake social responsibility in front of the public, but in essence, they didn't take corresponding actions or even violate their social responsibility, and "greenwashing" was a kind of pseudo-social responsibility behavior (Szabo & Webster, 2021; Lu and Zeng, 2023; Abedin et al., 2024) [4,5,6]. Due to the existence of business competition, it was necessary for companies to establish a green brand image and make eco-friendliness an important asset in their business process (Wen et al., 2024; Zeng et al., 2024) [7,8]. Torelli et al (2020) argued that "greenwashing" advertising was a target strategy to enhance the social awareness and reputation of products and companies, to create a responsible corporate or product image, and was based on semantic manipulation and green worship, symbolising "green"[9]. The marketing strategy was based on semantic manipulation and green worship, which symbolised "green" and obscured the cognition of investors and consumers during the dissemination process (Ma et al.,2025) [10]. Deepening the development of green finance was an important path for China to achieve the aim of "double carbon", and "greenwashing" was the main obstacle to its green development (Lu et al., 2023; Zeng et al., 2023; Zeng et al., 2024) [11,12,13]. Free et al (2024) argued that on the investment side, investors catered to the capital market's preference for ESG investments, but failed to conduct effective ESG-specific reviews and include "non-green" targets [14]. On the issuance side, the funds raised by green bonds were not invested in accordance with the promised framework, the environmental benefits of the invested projects were not disclosed, or the invested projects did not reach the promised targets of green development, resulting in an imbalance between the advancement of green financial projects and the construction of green infrastructure (Duan et al.,2025) [15]. Once the "greenwashing" behaviour was exposed, it weakened investor trust and had a negative effect on the financial market.

So this paper combed the domestic and foreign "greenwashing" related research literature, through the Citespace software to draw the visualization of the knowledge map, the study found that the overall growth of foreign "greenwash" research articles from 2013-2023 tends to be the same, foreign journals are more dispersed. The distribution of foreign journals is relatively decentralized, and the cooperation between authors and institutions is not close enough to form a stable academic cluster. Further comparative analysis has expanded the research review on "greenwashing" to a certain extent. In light of the bibliometric analysis of the existing study, we summarize the characteristics of the article on the topic of "greenwashing" and the places to be expanded. Meanwhile, we make lexical processing on the text related to greenwash and summarize the hot topics in the past ten years by using the LDA model, so as to provide certain references for the deepening of the theoretical research on greenwash in the future.

2.Literature Review

The theoretical study on the topic of "greenwashing" in China started late, but the behavior of "greenwashing" accompanies the whole process of the emergence of climate and environmental problems, which not only impacts the construction of business ethics (Zhang et al.,2025; Wang et al.,2024) [16,17], but also relates to the attainment of the goal of "dual-carbon" and the implementation of sustainable development strategy (Ma et al.,2024) [18]. It not only affects the construction of business ethics, but also relates to the attainment of the "dual-carbon" goal and the implementation of the sustainable development strategy (Zeng et al., 2025; Zhang et al.,2024) [19,20]. Therefore, under the premise of adhering to the problem effect orientation, it is necessary to continue to deepen the research on the topic of "greenwash" in China to serve the national and regional high-quality development (Zou et al.,2024; Wen et al.,2025) [21,22]. On the one hand, the research on the theme of "greening" should innovate the scientific research mode and further deepen the "organised scientific research". Its "organised nature" is manifested in the systematic and oriented nature of scientific research, which represents that scientific research will pay more attention to the selection of research topics, processes, tools and means, and synergistic innovation of different subjects (Li et al.,2025; Shen et al.,2025) [23,24]. Therefore, in determining research topics and in the research process, we cannot limit our vision to environmental science and ecological theory(Tong et al.,2024)[25], but should actively expand interdisciplinary cross-research on the basis of multidimensional disciplines such as behavioral economics, management, journalism and communication, philosophy and ethics, computers, law, etc., to break down barriers between traditional disciplines, and strive to promote the construction of the academic community (Wu et al., 2025;Wen et al.,2025)[26,27].

To prevent the trap of "greening" in the rapid development of green finance, more attention should be paid to the national strategic goal orientation, and the fiscal and taxation policy tools should be actively used to optimise the decision-making and

deployment of "anti-greening", low-carbon transformation. Regarding research methodology, the study of "greenwashing" should insist on linking theory to practice and combining qualitative and quantitative methods (Zeng et al., 2024) [28]. The research method of qualitative analysis alone may not be persuasive enough in terms of the certainty of causal inference, and it needs to be enriched to provide support for its corresponding conclusions and internalise the practice of empirical methods as a kind of academic self-consciousness (Seele & Schultz, 2022) [29].

At present, there were fewer empirical articles in domestic and international research, and one important reason was that there is insufficient sample data of corporate "greenwashing"(Li et al.,2023) [30]. To solve the problem of untrue, untimely and incomplete disclosure of environmental performance data, it was necessary not only to improve the information disclosure system from the policy and regulatory level, but also to encourage and guide the enterprises to build the "soft law system" of corporate green compliance culture, so as to realize the effectiveness of environmental information disclosure (Wang et al., 2024) [31]. As the main body of research, universities, as the main force of scientific research, should be based on the new stage of development, give full play to the advantages of disciplinary level improvement, platform connotation construction, accumulation of innovative talents, transfer and transformation of achievements and reform of the institutional mechanism, and aim at the major needs of the country to conduct scientific research and innovation practices in an oriented and directional manner (Wu et al., 2024) [32]. On the other hand, the content research in the field of "Green Washing" should be progressive from shallow to deep. Firstly, on the basis of clarifying the connotation, motivation, type, characteristics and negative impact of enterprise "greenwash", and in view of the hidden nature of the "greenwash" behavior (Zeng et al., 2025) [33], we should find out how to quickly identify and perceive the effective methods of "greenwash" traps, and find out how to quickly identify and perceive the effective methods of "greenwash" traps. Based on the hidden nature of "greenwashing" behavior, it is important to find out how to quickly identify and perceive "greenwashing" traps, and to quantify the "green content" and "green quality" of project construction (Mateo-Márquez et al, 2022) [34]. Wen et al (2025) explored the design of more effective "greening" governance measures and risk prevention mechanisms from the perspectives of enterprises, government and society, such as the path of corporate internal control, environmental accounting, environmental information disclosure laws and regulations, the construction of quantitative indexes for ESG investment audits, and the channels for improving the emission rights trading market, construction of investor claim mechanism, the supervisory utility of non-profit organizations and news media, etc [35]. Secondly, the combination of internationalization and localization should be done well, such as analyzing the influencing factors of different countries or large-scale multinational enterprises' "greenwashing" phenomenon, such as the willingness of enterprises to disclose environmental information, management's decision-making preference, the intensity of local administrative control or differences in the economic system, etc., or borrowing from other countries and regions to analyze the impacts of "greenwashing" phenomenon(Wu et al.,2025;Wu et al., 2024)[36,37]. It also draws on the experience of other countries and regions in effectively managing "greenwashing", and is dedicated on addressing the practical problems in the process of building China's ecological civilization and synergistic development of the economy and society (Wang et al.,2025; Shen et al.,2024) [38,39].

3.Data Source

3.1 Description

The Web of Science (WOS), a product developed by Clarivate, comprised authoritative journals across all academic disciplines from three principal citation indices (SCI, SSCI and A&HCI). Due to its stringent selection standards and thorough citation indexing methodology, WOS stood as a fundamental assessment tool within bibliometric and scientometric analyses. To deliver a comprehensive global examination regarding greenwashing research status, specific search parameters were established utilising the WOS Core Set database that aligned with the research objectives: (1) Titles containing "greenwashing", "wash green", "wash brown", "wash blue" or "green shine"; (2) Indices including SCI-EXPANDED, SSCI, AHCI, ESCI, CPCI-S, CPCI-SSH, BKCl-S or Book Citation Index-Social Sciences; (3) Document categories encompassing "article", "early access", "review article" or "conference paper"; (4) Publication timeframe spanning "2013-01-01 through 2023-12-31". Furthermore, incomplete manuscripts were excluded from consideration. Following these designated search criteria and subsequent manual verification to eliminate duplicates and irrelevant submissions, a total of 456 publications

were ultimately acquired.

3.2 Analytical tools

Citespace, an information visualisation application constructed upon JAVA programming language, functioned as the most commonly employed knowledge mapping instrument throughout academic literature. This tool rendered visual representations of trending subjects and frontier developments, thereby assisting researchers in grasping both current conditions and future trajectories within scholarly publications. The analytical process incorporated LDA (Latent Dirichlet Allocation), a probabilistic framework extensively applied throughout text mining and natural language processing domains. Its primary function involved uncovering concealed thematic structures within textual information—specifically identifying potential topics present across document collections whilst determining which themes appeared in individual manuscripts and which terminologies constituted each subject area. For this investigation, the researchers utilised the most recent Citespace programme alongside Python to examine structural characteristics and thematic variations among documents concerning "Green Washing", aiming to disclose underlying connections and implicit information embedded within the literature.

4.Results and Discussion

This article shows a visual depiction of greenwashing research is presented through the lens of scientometric analysis. This section delves into an examination and description of the contemporary state and prospective trajectories of greenwashing research, utilizing co-citation analysis, collaborative network analysis, and analysis of emerging trends as analytical frameworks.

4.1 Reference Co-citation

Two references can be considered to have a co-citation linkage when they are jointly cited by a number (n, where n equals 1 or 2) of other references simultaneously, with the intensity of their co-citation being equivalent to n. This co-citation intensity is commonly employed as a metric to gauge the thematic proximity or content-relatedness of the two references in question. Given that co-cited literature typically shares thematic similarities, co-citation intensity essentially quantifies the degree of content-based affinity between the referenced works. The co-citation checks of references was used to assess the mapping of cited references to 456 works in the area of "Green Washing", with a total of 741 nodes and 2043 rows been added. In the reference co-citation network, each node symbolizes a specific reference, with the connections between these nodes illustrating the co-citation relationships. The spatial proximity of these nodes serves as an indicator of the thematic or disciplinary affinity among the referenced literature. Furthermore, the edges linking distinct nodes not only denote the citation interactions between pairs of references but are also color-coded to represent distinct temporal segments (with each color corresponding to a particular year). It is worth noting that the representation of citation links through edges and their color-coding for time slices has been redundantly mentioned in the original context, but here it is streamlined into a single coherent description. The size of each node correlates with the frequency of citations its associated reference has received, thereby reflecting its scholarly impact, as visually demonstrated in Figure 1.

4.2 Institutional Cooperation Cited

To depict both meso- and micro-collaborative frameworks, this paper adapted various network illustrations as presented in Fig.2. Individual reference points appeared merely as isolated dots without forming meaningful linkages, thus contributing negligibly to collaborative structures; consequently, such elements were entirely omitted from our visualisation. The graphical representation concerning dual-subject collaboration patterns displayed remarkably thin connections, indicating minimal partnership intensity; nonetheless, this phase constituted a fundamental foundation for subsequent network expansion. Upon this dual-subject base, triadic participant relationships further developed, ultimately generating actual collaborative webs. Although multi-participant cooperation lacked a definitive central hub, it demonstrated considerably greater maturity compared with earlier collaborative arrangements. Furthermore, as additional participants gradually joined, these interaction systems possessed heightened potential for evolving into comprehensive, large-scale cooperative frameworks.

These modest collaborative structures exhibited notable diversity and geographical dispersion. When examining such networks, conventional mapping techniques typically resulted in visual disarray, hampering clear identification regarding specific information contained within different collaborative categories—particularly concerning quantity, temporal aspects,

and proportional relationships. Therefore, this study sequentially modified relevant illustrations according to both partnership intensity and developmental stages across various collaborative configurations. This methodological approach facilitated intuitive visual comprehension whilst simultaneously ensuring all applicable data remained thoroughly complete and statistically valid.

Figure1: Reference co-citation network

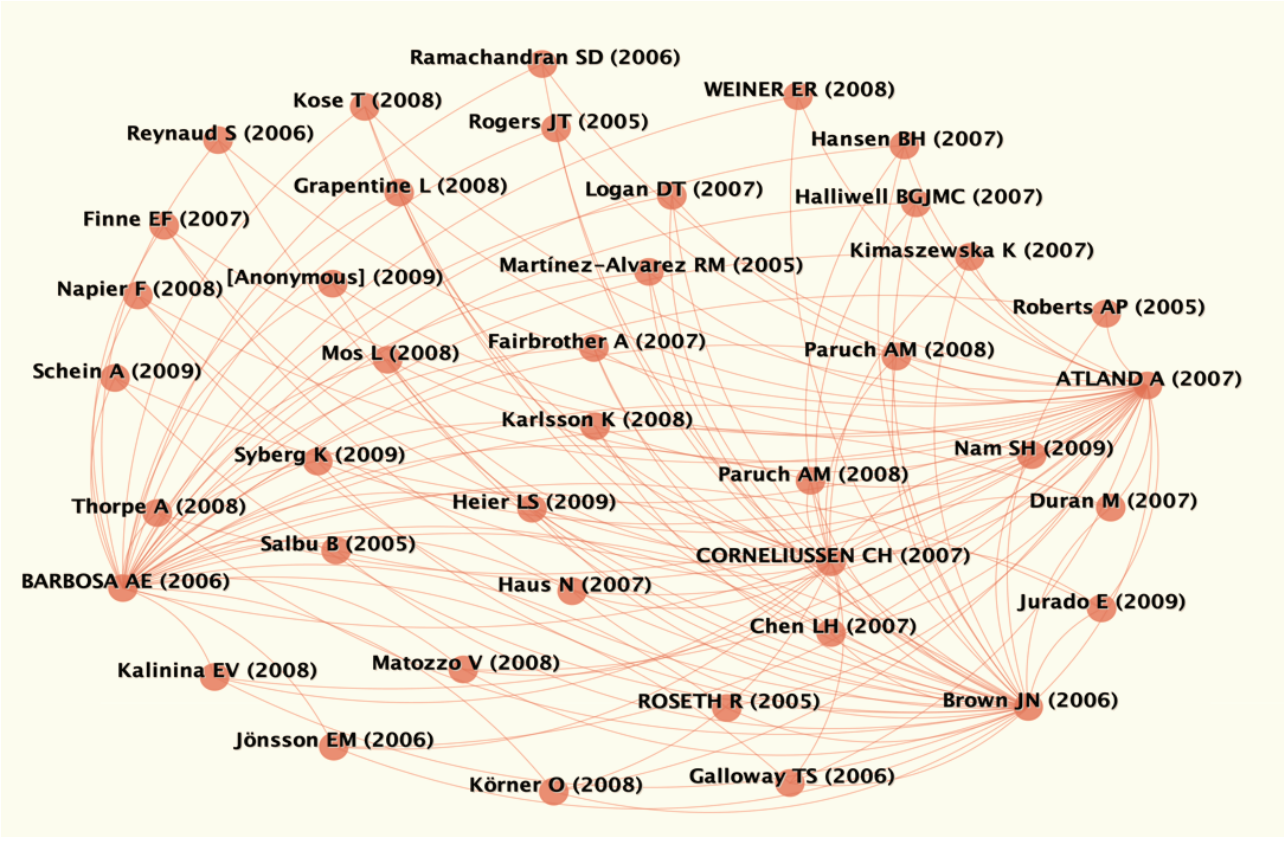
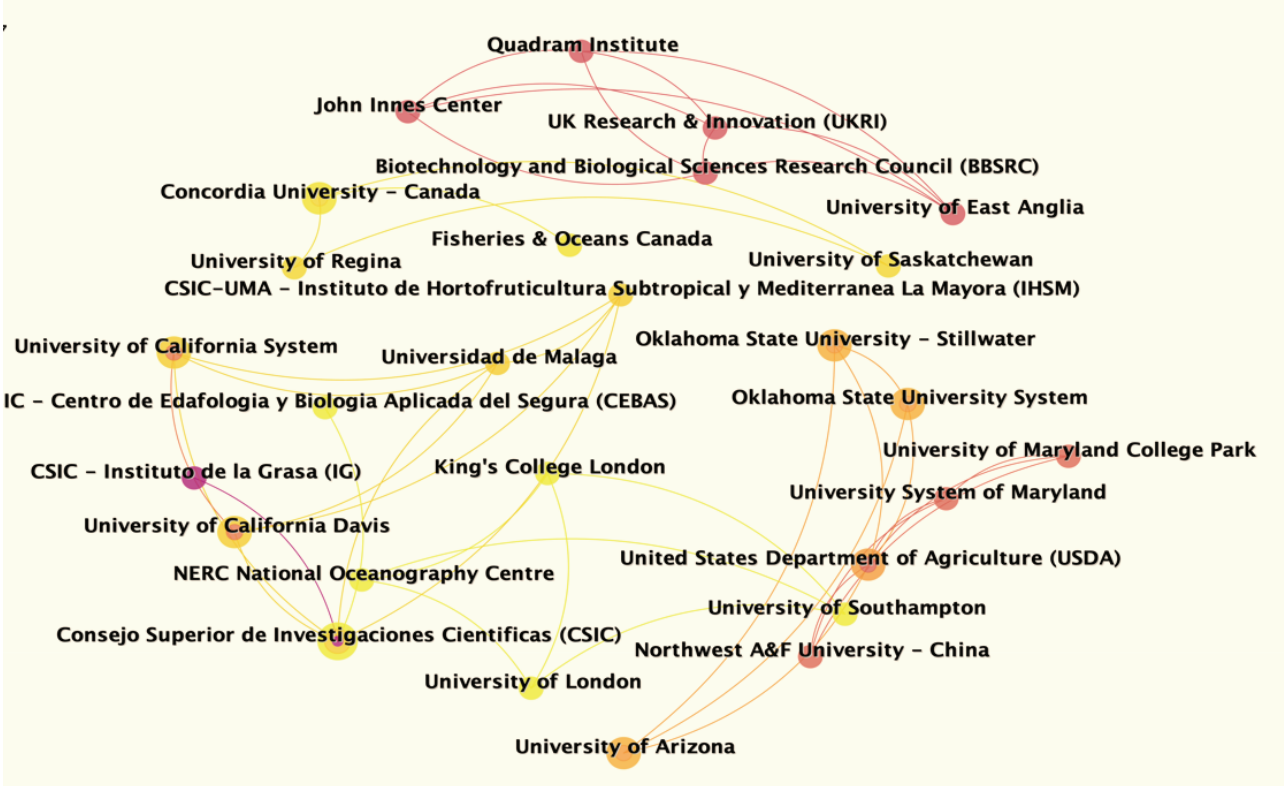


Figure2: Diagram of the institutional cooperation network

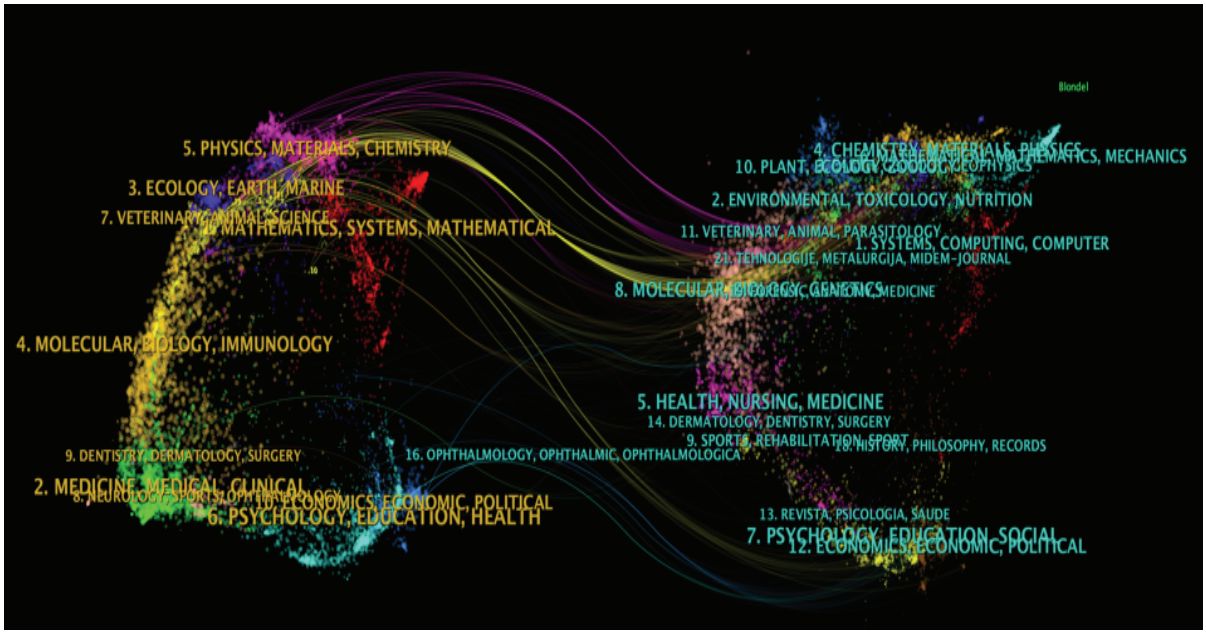


The number of articles released is an important measure of a research institution's research base and research strength in a particular field. Taking the research institutions as the co-occurrence analysis object, the research institutions that show the number of studies published is 20 and above, and get the ecotourism research institutions' cooperative relationship network diagram (as Figure 2). It can be seen that University of Arizona, USDA, CSIC, University of California Davis, University of California system, Oklahoma State University-Stillwater, etc. are the main research institutions. research institutions and colleges as the center of several partnership networks. Institutions with a high number of publications include: University of Arizona (45), USDA (39), CSIC (32), University of California Davis (29), University of California system (23), Oklahoma State University-Stillwater (21). Most of the institutions with a high number of publications have a disciplinary background in forestry, ecology, and environmental studies, while showing an intermingling of the natural and social sciences. The institutional partnerships depicted exhibited a remarkably fragmented structure overall. These cooperative arrangements predominantly manifested as diminutive, autonomous clusters across the landscape. Simultaneously, academic institutions from Western Europe and North America tended to dominate both leadership positions and general participation within these scholarly alliances. Their research endeavours progressed along largely separate trajectories, whilst the established collaborative frameworks that eventually emerged demonstrated both tardy formation and gradual developmental progression. The interconnected webs of organisational cooperation had neither cohered rapidly nor expanded swiftly throughout the examined period.

4.2 Journal dual graph overlay analysis

Journal double-map overlays analysis can visualize the disciplinary distribution of citing and cited papers. In Figure 3, the discipline distribution of cited papers on the left side can be regarded as the application field of Green Washing research, and the discipline distribution of cited papers on the right side represents which disciplines are mainly cited in Green Washing research, which can be regarded as the basis of Green Washing research. In the depicted ellipse, the vertical axis signifies the quantity of papers within the discipline, whereas the horizontal axis denotes the number of authors. The results of the analysis show that during the period 2013-2023, the basic research areas of Green Washing mainly include Systems, Computing, Environmental, Toxicology, Nutrition, Chemistry, Materials, Physics, Health, Medicine and Molecular disciplines. Application areas include Mathematics, Systems, Medicine, Clinical, Ecology, Earth, Marine, Molecular, Biology, Immunology, Physics, Materials, Chemistry, Psychology, Education, Health, Veterinary, Animal, Science, and Neurology. There are more basic research fields in Green Washing, but the overall division between core and marginal fields is more obvious, which shows that the basic research of Green Washing involves a wider range of disciplines, and the application of this field involves more disciplines, and the core fields are clear. The core areas are clear.

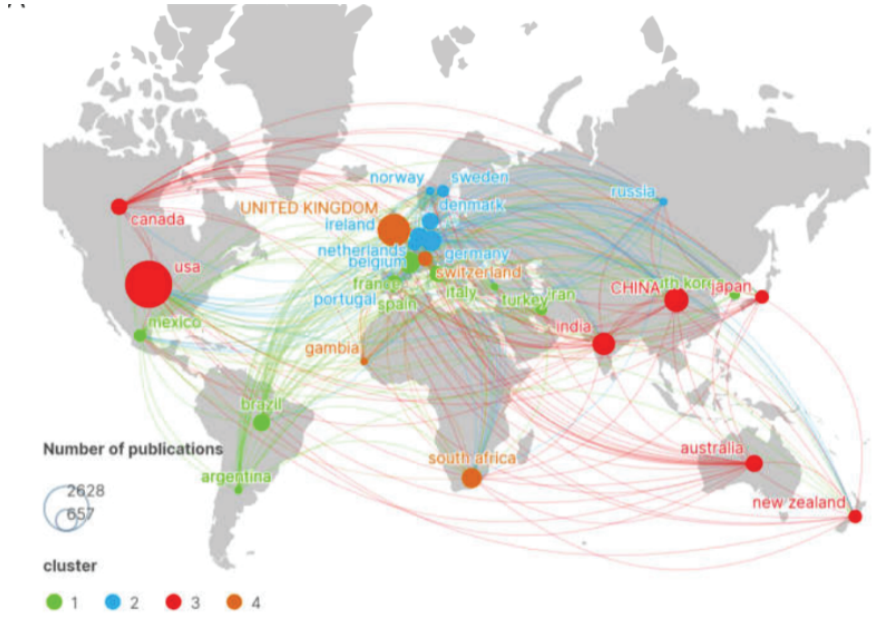
Figure 3: Journal double image overlay



4.3 Journal Dual Chart Stacking Analysis

CiteSpace alongside Scimago Graphica facilitated the creation of geographical distribution maps that illuminated various territories' contributions towards Green Washing scholarship. This investigation categorised 30 nations from amongst 132 countries (each having produced minimally 10 publications) into four distinct groupings, utilising the proportion of participating nations relative to total country count as classification criteria (illustrated in Figure 4). The analytical results revealed that: (1) The initial grouping encompassed nine nations: France, Brazil, Mexico, Italy, Korea, Turkey, Iran, Argentina, and Spain. (2) The second category primarily comprised European territories (Norway, Sweden, Denmark, Ireland, the Netherlands, Portugal, and Germany) alongside Russia. (3) The third classification incorporated the United States, China, India, Australia, Canada, New Zealand, and Japan. (4) The fourth segment featured Great Britain as its predominant constituent, whilst additionally including Switzerland, Gambia, and South Africa. Regarding international research partnerships, American institutions demonstrated the highest collaboration frequency, with British and Chinese establishments following subsequently. Such statistical evidence indicated that America, Britain and France maintained their leadership positions throughout Green Washing academic inquiry concerning both quantitative output and qualitative standards. Concurrently, developing economies such as India and China continued experiencing considerable disparities compared with industrialised nations. The Chinese situation proved particularly noteworthy; despite rapidly accelerating publication quantities, significant improvements regarding scholarly excellence remained necessary.

Figure 4: Country maps



4.4 Study on the Calculation of Green Text

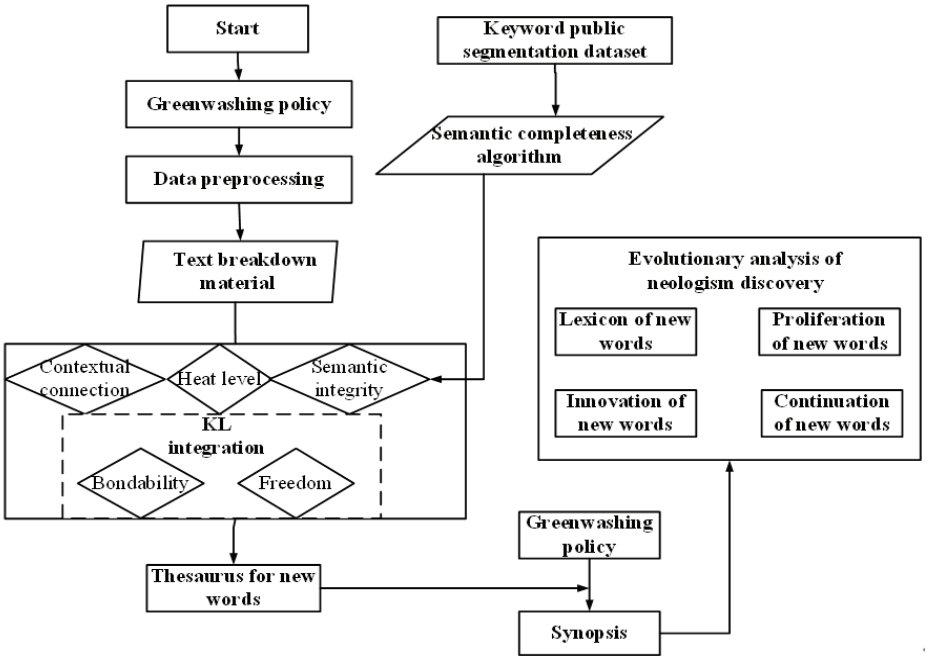
Textual computing is a framework based on the theories of computer science, linguistics, and political science aimed at mining and computational analysis of massive amounts of text, which advocates the use of text encoding, lexicon of textual concepts, and mapping relations between text and language for automatic recognition and processing of textual concepts, involving textual data processing, textual content analysis, and text mining.

Policy text analysis based on text data processing and text mining usually relies on open-source Chinese word segmentation tools. Using the corpus after lexical segmentation for statistical or econometric analysis of themes, types, sentiments, etc., or applying co-occurrence or co-occurrence to analyze the changing rules of policy growth, policy diffusion, policy change, inter-governmental relations, policy evolution, policy evolution, etc., or to carry out the potential discovery of semantic knowledge, association rule discovery, clustering analysis, automatic classification and so on. Open-source Chinese lexical tools are less adaptable to policy texts, less sensitive to unregistered words and new words, and difficult to capture emerging vocabulary and emerging concepts in the field. The lack of specialized thesaurus also limits the application of text mining techniques in policy analysis.

4.5 Analysis of Greenwashing Evolution Based on Neologism Discovery

This paper intends to construct an AI policy thesaurus to assist word division through the neologism discovery algorithm, and conduct an evolutionary analysis of the innovation, continuation, and diffusion of AI policies based on the results of policy neologism discovery, so as to help the policy audience capture the industry development direction and the policy support orientation in a timely manner, and to help the local policy makers follow up on the emerging fields in a timely manner, so as to realize the policy innovation and the targeted development. The research process is shown in Figure 5, including the neologism discovery algorithm and the policy evolution analysis based on neologism discovery.

Figure 5: MFF New Word Discovery and Policy Text Analysis



4.6 MFF Multi-feature New Word Discovery Qlgorithm

The ever-emerging technical neologisms bring a great test to Chinese word segmentation, which affects the accuracy of word segmentation. In order to consider the features of words in different dimensions more comprehensively and capture richer semantic information, this article proposes a new word discovery algorithm (MFF) that fuses multiple features. The MFF multi-feature contains word hotness, contextual relevance, semantic completeness and KL fusion, defined as follows:

(1) Heat (Heat, H): this can be expressed in terms of word frequency. Word frequency is the frequency of occurrence of a word in a text, and high-frequency words appearing in a certain period of time may be candidate neologisms.

(2) Contextual Association (CA): units separated by the same word show high similarity in the vector representation space, and may be potential neologisms when connected subwords show high association in context.

(3) Semantic Completeness (SC): scoring the semantic completeness of a word can determine whether a word and its similar words can be neologisms. In this paper, we use Transformer for global and self-attentive cyclic sequence modeling, and train the semantic completeness judgment algorithm with the help of keywords and publicly available participle data from AI literature.

(4) KL Fusion (KLFusion, KLF): KL Scatter (Kullback-Leibler Divergence) is used to measure the difference between two probability distributions, while KL Fusion utilizes KL Scatter to jointly compute the degree of union and the degree of freedom, which is used to quantify the difference between the internal structure of the word and the external environment, and to obtain more robust and reasonable result of new word discovery with the following formula:

$$\text{KLFusion}° = \begin{cases} KLF\_\text{Value}° & P(°\text{word}°) > 0° \text{and}° Q(°\text{word}°) > 0 \\ 100 & P(°\text{word}°) = 0° \text{or}° Q(°\text{word}°) = 0 \end{cases} \quad (1)$$

$$KLF\_\text{Value}° = \frac{P(°\text{word}°) \times \log\left(\frac{P(°\text{word}°)}{Q(°\text{word}°)}\right) + Q(°\text{word}°) \times \log\left(\frac{Q(°\text{word}°)}{P(°\text{word}°)}\right)}{2.0} \quad (2)$$

Where P(word) is the binding degree, calculated by mutual information. x, y are the left and right subwords of the candidate neologism, P(x,y) is the co-occurrence probability of x and y, P(x)P(y) is the independent occurrence probability of x and y, and PMI(x,y) denotes the probability of word formation of the candidate neologism.

$$PMI(x,y) = \log\left(\frac{P(x,y)}{P(x)P(y)}\right) \tag{3}$$

Q(word) is the degree of freedom, which is calculated by the neighbor entropy. Neighbor entropy can measure the uncertainty of the left and right neighbors of a new word, and the larger its uncertainty, the higher its probability of becoming a word. sl is the set of left neighbors of the candidate new word, Sr is the set of right neighbors, P(Wl|W) denotes the probability of the left neighbor of the candidate new word, P(Wr|W) denotes the probability of the right neighbor of the candidate new word, and N(Wl,W) and N(Wr,W) denote the number of times that the left and right words appear with W, and N(W) denotes the number of times that W appears together. N(Wl,W) and N(Wr,W) denote the number of times the left and right words co-occur with W, respectively, and N(W) denotes the number of times W appears.

$$Q(\text{°word°}) = \begin{cases} -\sum_{w_l \in S_l} P(W_l|W)\log P(W_l|W) \\ -\sum_{w_r \in S_r} P(W_r|W)\log P(W_r|W) \end{cases} \tag{4}$$

$$P(W_l|W) = \frac{N(W_l,W)}{N(W)} \tag{5}$$

$$\text{NewWord} = \begin{cases} 1\text{°if°}f_1 \geq T_1\text{°AND°}f_2 \geq T_2\text{°AND°}f_3 \geq T_3\text{°AND°}f_4 \geq T_4 \\ 0 \qquad\qquad \text{°Otherwise°} \end{cases} \tag{6}$$

Where f1, f2, f3, f4 are the values of H, CA, SC, and KLF, respectively; and T1, T2, T3, T4 are the thresholds of H, CA, SC, and KLF, respectively.

4.6 LDA-based News Topic Recognition Model Construction

In this paper, the LDA topic model is used to construct a topic recognition model for science and technology news, and three logical hierarchical structures are established: document layer, topic layer, and vocabulary layer, and each layer is regulated with corresponding variables and parameters. Combined with the given probability sampling method, the topic variables are used to generate the words in the document, and the topic recognition generation process is as follows: first, for the jth topic, the distribution probability vector of polynomials φzi, which is the feature word on the topic, is computed through the Dirichlet distribution j = Dir(β), and N is selected using the Poisson distribution N = Poisson(ξ) where N represents the length of a single tech news document. Second, θ is determined based on θ~Dir(α), where θ obeys the Dirichlet(α) distribution, and θ represents the probability of occurrence of each topic that α is a parameter of the Dirichlet distribution. Again, a topic zn is selected for each of the k feature words w contained in each science and technology news document, zn obeys the Multinomial(θ) distribution vector θ over the topic distribution θ multinomial distribution, and zn is a randomly selected topic. A feature word wn is randomly selected from the above selected topic zn according to $p(wn|zn; \beta)$. where p is the topic zn of the multinomial distribution; β is a $K \times N$-dimensional matrix; $\beta ij = P(wi = 1|\ zj = 1)$, indicating that β is the probability of generating the feature word wi for the record topic zj. The generation probability $p(wi)$ of wi in the technology news document d is:

$$p(wi) = {}_{j=}^{\sum T} 1P(wi|zi = j)\ P\ (zi = j) \tag{7}$$

Where wi represents the ith theme that appears in theme zi; $p(wi|zi = j)$ denotes the occurrence of the feature word in theme zi wi; $p(zi = j)$ is the probability that topic zi occurs in the document. Therefore, the probability that the feature word w appears in the technology news document d, $p(w|d)$ is:

$$p(w|d) = {}_{w}^{\sum T} {}_{\theta d}^{j=1\varphi j}\ j \tag{8}$$

Finally, the EM algorithm is used to obtain the approximate solution of α, β, which in turn builds the LDA three-layer model:

$$I(\alpha|\beta) = {}_{\sum M}\ i{=}1\log\ p(di|\alpha, \beta) \tag{9}$$

The LDA model is constructed to analyze the internal semantics of the science and technology news document set, and the three-layer model of text-topic-feature word of science and technology news is obtained and presented in the form of a matrix.

4.7 News Record Subject Identification

In this paper, the LDA topic model is utilized to identify the topics of the Green Washing news text set of local websites from 2013-2023. First, the optimal number of topics is determined by the perplexity curve. (as shown in Figure 6). In order to avoid overfitting, this paper combined with the interpretability of the theme after many experiments, the curve appeared a turning point at T=8, so the number of themes of this text dataset was set as T=8. Second, obtain the record theme recognition results. According to the results of theme-feature word distribution obtained from LDA theme modeling analysis, combined with manual screening, each theme is reserved for 8 feature words that are relatively able to reflect the meaning of the theme and summarize the theme name. Finally, it can be found that in the past 10 years, the Green Washing theme focuses on eight areas: Green Washing behavior, green development, green finance, carbon neutrality, information disclosure, social responsibility, sustainable development and environmental performance. The word cloud map of the bleached green text is shown in Figure 7.

Figure 6: Theme-confusion degree model

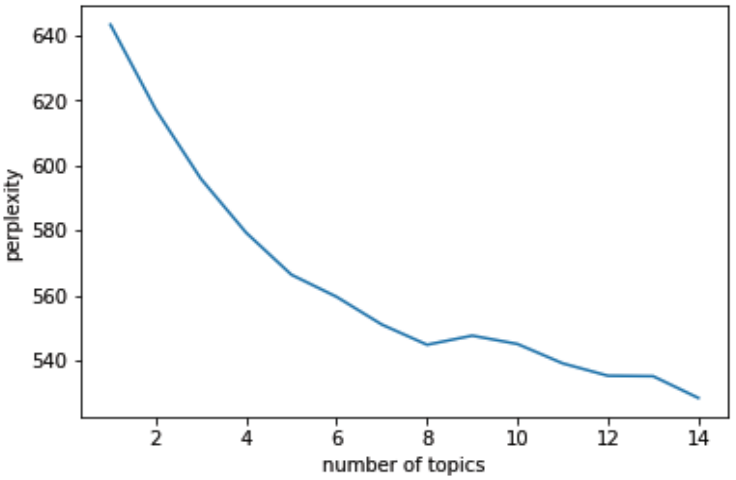


Figure 7: World cloud

Theme 1 is greenwashing behavior. Like a quietly growing vine, greenwashing behavior is gradually spreading in all walks of life. It refers to the false propaganda or misleading tactics adopted by some enterprises or organizations in terms of environmental protection and sustainable development in order to cover up their real environmental impacts. This behavior not only misleads consumers and investors, but also poses a great obstacle to environmental protection. Over time, the forms of greenwash behavior have become more diverse. Some companies have begun to offset their actual emissions by purchasing carbon credits or engaging in carbon trading to achieve the illusion of “zero emissions”. This reminds us that we must remain

vigilant in the pursuit of environmental protection and sustainable development, and avoid being confused by greenwash. As environmental awareness increases and regulation is strengthened, it is believed that greenwash will gradually decrease and true environmental protection and sustainable development will become mainstream.

Theme 2 is green development. There is a growing global call for green development, and people are generally pursuing a sustainable development path that harmonizes the economy with the environment. However, against this backdrop, greenwash is like a dark current, posing considerable challenges to the process of green development. At the same time, greenwashing behavior has weakened people's confidence in green development and affected the consensus and action power of the whole society on environmental protection. Although greenwashing has brought some trouble to the process of green development, we have also seen more enterprises and organizations begin to actively participate in the green cause and practice the concept of green development with practical actions. phase

Theme 3 is green finance. Green finance, as an important force for promoting sustainable development, has attracted much attention globally. Green finance aims to promote the harmonious development of the economy and the environment by directing capital flows to environmentally friendly, low-carbon and sustainable projects and enterprises through financial innovation and financial instruments. However, in the course of the development of green finance, some institutions and enterprises have begun to adopt green-bleaching behaviors in order to pursue short-term economic benefits, i.e., claiming to support green finance on the surface but failing to truly implement the green principles in reality. These green-bleaching behaviors include, but are not limited to, falsely publicizing their green investment results, exaggerating the scale and benefits of green projects, or packaging non-green projects as green projects to attract investors. These behaviors not only mislead investors and consumers, but also have a negative influence on the healthy development of green finance.

Theme 4 is carbon neutrality. With the global awareness of environmental protection, carbon neutrality has become a hot topic. However, in this green wave, there is no lack of "greenwashing". In order to change this situation, more and more companies are beginning to realize that true carbon neutrality requires reducing carbon emissions at the source rather than relying solely on offsets. They have begun to adopt cleaner energy, improve energy efficiency, and promote green travel to reduce their own carbon emissions. At the same time, the government has also stepped up its efforts to combat greenwashing, by enacting stricter environmental regulations and strengthening supervision to regulate the environmental behavior of enterprises.

Theme 5 is information disclosure. As public attention to environmental protection and corporate social responsibility continues to grow, information disclosure has become an important way for companies to demonstrate their green development and environmental practices. However, in this process, the so-called "greenwashing" phenomenon has also emerged, i.e., enterprises exaggerate their environmental achievements or conceal their real environmental problems in information disclosure in order to build up their environmental image. In addition, the development of greenwashing is also reflected in the interpretation and publicity of environmental policies by enterprises, who may package their products and services with labels such as "green innovation" and "sustainable development" in order to attract the attention of consumers and investors. However, with the deepening of social awareness of greenwashing behavior, the public has also put forward higher requirements for corporate environmental disclosure. More and more investors and consumers are paying attention to the real environmental performance of companies rather than just looking at their superficial environmental image. At the same time, regulators are also strengthening their supervision of corporate environmental disclosure, requiring companies to disclose their environmental information truthfully, accurately and completely.

Theme 6 is social responsibility. Social responsibility has become one of the indispensable core values of enterprises. Green Washing development refers to the superficial and formalized practices of some enterprises in assuming social responsibility, using false environmental commitments, non-transparent information disclosure, or over-packaged publicity tactics to build up their positive social image without actually fulfilling their social responsibility. Enterprises may attract the attention of the public and investors by releasing vague and non-specific social responsibility reports that exaggerate their environmental protection commitments, charitable donations and other social responsibility practices. However, behind these flashy figures often lie neglect and concealment of real problems such as environmental damage and labor rights infringement. Behind the

bleached-green development, it reflects the neglect and misunderstanding of social responsibility of some enterprises in the pursuit of economic benefits.

Theme 7 is sustainable development. Sustainable development is now recognized as a global priority, and the phenomenon of greenwashing not only misleads the public and investors, but also hinders the advancement of the cause of true sustainable development. Truly sustainable development requires enterprises to change their business concepts and approaches at the source, and to realize the harmonization of economic, social and environmental benefits through technological innovation, green production and circular economy. In addition, the government, social organizations and the public should also strengthen the supervision and constraints on enterprises' Green Washing development. The government should formulate stricter environmental protection regulations and policies, and strengthen the supervision and punishment of enterprises' environmental protection behaviors.

Theme 8 is environmental performance. With increasing global concern about environmental issues, environmental performance has become an important component of corporate competitiveness. In order to meet the challenges of Green Washing development, enterprises need to abandon the practice of false advertising and report their environmental performance truthfully and accurately. Enterprises should establish a strict environmental management system, strengthen internal supervision and self-restraint, and ensure that all environmental protection measures are effectively implemented. Simultaneously, enterprises should strengthen technological research and development and innovation, promote green, low-carbon and recycling production methods, decrease environmental contamination and material waste, and enhance environmental performance. The government should strengthen the creation and application of environmental regulations and improve the requirements for monitoring and evaluating the environmental performance of enterprises.

Conclusion

With numerous organisations learning and emulating "Green Washing" practices, this phenomenon experienced significant expansion, whilst its negative consequences extended throughout all dimensions of production and everyday existence. Hence, a comprehensive examination of scholarly literature became necessary to analyse this marketing strategy based upon extant research. This investigation employed CiteSpace software to render visual representations of fundamental publications within the greenwashing domain from the WOS database. The research dissected inherent characteristics of the "greenwashing" occurrence and synthesised its evolutionary patterns, thereby facilitating more rational, scientific, efficacious and precise determination, monitoring, assessment and regulation of "greenwashing" activities to bolster environmentally responsible advancement.

Concerning scholarly focus areas, business economics and environmental science disciplines generated the predominant proportion of literature in this field, constituting 77.8%. Regarding publication territories and academic establishments, economically advanced nations maintained supremacy, with American institutions and the European Research Universities Consortium (ERUC) occupying premier positions across geographical regions and research organisations respectively. This circumstance enabled American academics to control global discourse substantially. The co-citation analytical diagram illustrated progressive increases in publications exerting considerable influence on greenwashing investigations and demonstrated relationships between traditional concepts and innovative approaches. Moreover, influential scholars who initiated earlier studies within the greenwashing sphere continuously enhanced theoretical frameworks whilst simultaneously attracting additional researchers to this domain. Network collaboration visualisations revealed robust partnerships amongst developed economies including European nations and America; however, insufficient cooperation existed between emerging economies across Southeast Asia, South America, and African regions. Institutional collaborative arrangements appeared loosely distributed and featured primarily small cooperative clusters. Furthermore, academic collaboration generally displayed fragmentation, whilst individual researchers demonstrated inadequate investigative depth.

Throughout greenwashing research evolution, scholarly attention concentrated on definitional aspects and conceptual elucidation, influential variables, resultant effects and governance structures. Methodological approaches gradually transitioned from normative investigations towards empirical analyses and subsequently combined methodologies. Additionally, research paradigms progressively transcended national boundaries to encompass global perspectives,

emphasising universal applicability whilst diminishing contextual specificity. This progression further enriched and perfected the research framework from defining greenwashing (identifying its nature) to tracking its occurrence (understanding causal factors) to evaluating consequential impacts to controlling its proliferation. Ultimately, greenwashing textual content underwent subdivision utilising Python technology, whilst the LDA thematic model was applied to analyse eight principal greenwashing subject areas over the preceding decade. These focal domains encompassed greenwashing behaviour, environmental advancement, green financial systems, carbon neutrality objectives, information disclosure practices, social responsibility initiatives, sustainable development principles and environmental performance metrics, with detailed explanations provided regarding evolutionary developments within each respective area.

Funding

This work was supported by Guangdong Education Association 2024 Educational Research Project: A Study on the Continuity of Mental Health Education Courses from Primary to University Levels.

Conflict of Interests

The author(s)declare(s) that there is no conflict of interest regarding the publication of this paper.

References

[1] Yang, Z., Nguyen, T. T. H., Nguyen, H. N., Nguyen, T. T. N., & Cao, T. T. (2020). Greenwashing behaviours: Causes, taxonomy and consequences based on a systematic literature review. Journal of business economics and management, 21(5), 1486-1507.

[2] Tong, Z., Ding, Y., Ma, S., & Yan, H. (2025). How to Mitigate Climate Change? Dynamic Linkages between Clean Energy and Systemically Important Banks. Global NEST Journal.

[3] Ma, S. Liu, H., Li, S., Lyu, S.& Zeng, H. (2025). Quantifying the Relative Contributions of Climate Change and Human Activities to Vegetation Recovery in Shandong Province of China. Global NEST Journal.

[4] Szabo, S., & Webster, J. (2021). Perceived greenwashing: the effects of green marketing on environmental and product perceptions. Journal of business ethics, 171, 719-739.

[5] Lu, R., & Zeng, H. (2023). VIX and major agricultural future markets: dynamic linkage and time-frequency relations around the COVID-19 outbreak. Studies in Economics and Finance, 40(2), 334-353.

[6] Abedin, M. Z., Goldstein, M. A., Huang, Q., & Zeng, H. (2024). Forward-looking disclosure effects on stock liquidity in China: Evidence from MD&A text analysis. International Review of Financial Analysis, 95, 103484.

[7] Wen, L., Ma, S., & Lyu, S. (2024). The influence of internet celebrity anchors' reputation on consumers' purchase intention in the context of digital economy: from the perspective of consumers' initial trust. Applied Economics, 1-22.

[8] Zeng, H. (2024). Risk transmission and diversification strategies between US real estate investment trusts (REITs) and green finance indices. Kybernetes.

[9] Torelli, R., Balluchi, F., & Lazzini, A. (2020). Greenwashing and environmental communication: Effects on stakeholders' perceptions. Business strategy and the Environment, 29(2), 407-421.

[10] Ma, S., H. Yan, D. Li, H. Liu and H. Zeng. (2025b). The Impact of Agricultural Mechanisation on Agricultural Carbon Emission Intensity: Evidence from China. Pakistan Journal of Agricultural Sciences. 62: 99-110.

[11] Lu, R., Xu, W., Zeng, H., & Zhou, X. (2023). Volatility connectedness among the Indian equity and major commodity markets under the COVID-19 scenario. Economic Analysis and Policy, 78, 1465-1481.

[12] Zeng, H., Lu, R., & Ahmed, A. D. (2023). Return connectedness and multiscale spillovers across clean energy indices and grain commodity markets around COVID-19 crisis. Journal of Environmental Management, 340, 117912.

[13] Zeng, H., Abedin, M. Z., Zhou, X., & Lu, R. (2024). Measuring the extreme linkages and time-frequency co-movements among artificial intelligence and clean energy indices. International Review of Financial Analysis, 92, 103073.

[14] Free, C., Jones, S., & Tremblay, M. S. (2024). Greenwashing and sustainability assurance: A review and call for future research. Journal of Accounting Literature.

[15] Duan, K., Qin, C., Ma, S., Lei, X., Hu, Q., & Ying, J. (2025). Impact of ESG disclosure on corporate

sustainability. Finance Research Letters, 107134.

[16] Zhang, G., Ma, S., Zheng, M., Li, C., Chang, F., & Zhang, F. (2025). Impact of Digitization and Artificial Intelligence on Carbon Emissions Considering Variable Interaction and Heterogeneity: An Interpretable Deep Learning Modeling Framework. Sustainable Cities and Society, 106333.

[17] Wang, C., Liu, H. and Ma, S. (2024). Analysis of the effect of digital financial inclusion on agricultural carbon emissions in China, Global NEST Journal, 26(8).

[18] Ma, S., Wen, L., and Yuan, Y. (2024). Study on the coupled and coordinated development of tourism, urbanization and ecological environment in Shanxi Province, Global NEST Journal, 26(4).

[19] Zeng, H., Abedin, M. Z., Ahmed, A. D., & Huang, Q. (2025). Extreme risk connection among the European Tourism, energy and carbon emission markets. Research in International Business and Finance, 74, 102693.

[20] Zhang, K., Li, Y., Ma, S., Fu, C. (2024). Research on the Impact of Green Technology Innovation in the Manufacturing Industry on the High-Quality Development of the Manufacturing Industry Under Dual Circulation. Polish Journal of Environmental Studies.

[21] Zou, F. Ma, S. Liu, H. Gao, T. and Li, W. (2024). Do Technological Innovation and Environmental Regulation Reduce Carbon Dioxide Emissions? Evidence from China, Global NEST Journal, 26(7).

[22] Wen, L., Ma, S., Wang, C., Dong, B., Liu, H. (2025). A Study of Green Strategy Choice and Behavioral Evolution of Consumers and Producers under the Double Subsidy Policy. Polish Journal of Environmental Studies.

[23] Li, Y., Yang, X., & Ma, S. (2025). The Efficiency Measurement and Spatial Spillover Effect of Green Technology Innovation in Chinese Industrial Enterprises. Sustainability, 17(7), 3162.

[24] Shen, D., Zhao, X., Lyu, S., Liu, H., Zeng, H., & Ma, S. (2025). Qualification and construction enterprise innovation – quasi-natural experiments based on specialized, high-end and innovation-driven "small giant" enterprises. Journal of Asian Architecture and Building Engineering, 1–19.

[25] Tong, L., Wang, C., Qi, Q., Ma, S., and Mei, J. (2024). Study on the Impact of China's Digital Economy on Agricultural Carbon Emissions, Global NEST Journal, 26(6).

[26] Wu, R., Zeng, H., Abedin, M. Z., & Ahmed, A. D. (2025). The impact of extreme climate on tourism sector international stock markets: A quantile and time-frequency perspective. Tourism Economics, 13548166241311633.

[27] Wen, L., Ma, S., Zhao, G., & Liu, H. (2025). The Impact of Environmental Regulation on the Regional Cross-Border E-Commerce Green Innovation: Based on System GMM and Threshold Effects Modeling. Polish Journal of Environmental Studies, 34(2).

[28] Zeng, H., Abedin, M. Z., & Upreti, V. (2024). Does climate risk as barometers for specific clean energy indices? Insights from quartiles and time-frequency perspective. Energy Economics, 140, 108003.

[29] Seele, P., & Schultz, M. D. (2022). From greenwashing to machinewashing: a model and future directions derived from reasoning by analogy. Journal of Business Ethics, 178(4), 1063-1089.

[30] Li, Y., Cong, R., Zhang, K., Ma, S., & Fu, C. (2024). Four-way game analysis of transformation and upgrading of manufacturing enterprises relying on industrial internet platform under developers' participation. Journal of Asian Architecture and Building Engineering, 1–22.

[31] Wang, Z., Wu, Q. and Ma, S. (2024). Research on Carbon Emission Peaks in Large Energy Production Region in China —Based on the Open STIRPAT Model, Global NEST Journal, 26(5).

[32] Wu, Q., Jin, Y. and Ma, S. (2024). Impact of dual pilot policies for low-carbon and innovative cities on the high-quality development of urban economies, Global NEST Journal, 26(9).

[33] Zeng, H., Wu, R., Abedin, M. Z., & Ahmed, A. D. (2025). Forecasting Volatility of Australian Stock Market Applying WTC-DCA-Informer Framework. Journal of Forecasting.

[34] Mateo-Márquez, A. J., González-González, J. M., & Zamora-Ramírez, C. (2022). An international empirical study of greenwashing and voluntary carbon disclosure. Journal of Cleaner Production, 363, 132567.

[35] Wen, L., Ma, S., Zhao, G., & Liu, H. (2025). The Impact of Environmental Regulation on the Regional Cross-

Border E-Commerce Green Innovation: Based on System GMM and Threshold Effects Modeling. Polish Journal of Environmental Studies, 34(2).

[36] Wu, Y., Zeng, H., Hao, N., & Ma, S. (2025). The impact of economic policy uncertainty on the domestic value added rate of construction enterprise exports—evidence from China. Journal of Asian Architecture and Building Engineering, 1–15.

[37] Wu, R., Li, M., Liu, F., Zeng, H., & Cong, X. (2024). Adjustment strategies and chaos in duopoly supply chains: The impacts of carbon trading markets and emission reduction policies. International Review of Economics & Finance, 95, 103482.

[38] Wang, Z., Wang, F., & Ma, S. (2025). Research on the Coupled and Coordinated Relationship Between Ecological Environment and Economic Development in China and its Evolution in Time and Space. Polish Journal of Environmental Studies, 34(3).

[39] Shen, D., Guo, X., & Ma, S. (2024). Study on the Coupled and Coordinated Development of Climate Investment and Financing and Green Finance of China. Sustainability, 16(24), 11008.