CADERNOS DE DEREITO ACTUAL
www.cadernosdedereitoactual.es

© ***Cadernos de Dereito Actual*** Nº 31. Núm. Ordinario (2026), pp. 257-283
·ISSN 2340-860X - ·ISSNe 2386-5229

Institutional embedding of ecological civilization discourse in environmental justice: A textual thematic evolution analysis based on Supreme People's court cases

Anqian Wang[1,*]
Macau University of Science and Technology

Abstract: To address the disconnect between ecological civilization discourse and environmental judicial practice, and the difficulty of traditional thematic models in capturing the deep semantic evolution of judgment texts, this paper systematically examines the presentation and evolutionary characteristics of ecological civilization discourse in judicial decisions, using 450 publicly available environmental-related judgments from the Supreme People's Court as its research object. This study provides an empirical path for understanding ecological civilization discourse through judicial interpretation and case institutionalization, and offers practical suggestions for judicial standardization and evidence chain construction. Methodologically, at the

[1] An Qian Wang is a Ph. D. candidate in the Faculty of Law at Macau University of Science and Technology，Macau, China. His research interests lie in fundamental theories of environmental law, environmental justice, and environmental policy. E-mail: anqianwang@hotmail.com (corresponding author).

level of the judgment reasoning paragraphs, this paper employs the BERTopic thematic model framework—based on Chinese SBERT semantic embedding combined with UMAP dimensionality reduction, HDBSCAN clustering, and c-TF-IDF keyword extraction—to conduct time-slice analysis of judicial discourse, revealing the institutionalization path of ecological civilization discourse. The results show that the thematic weights of ecological restoration and restoration to the original state have been continuously increasing (0.193 in 2012, 0.228 in 2024); the themes of ecological civilization value and legal application appear explicitly in 63.6% of the cases, and are significantly correlated with the citation of guiding cases (OR=2.62, p=0.001). The BERTopic model performs well in terms of topic consistency (c_v=0.62), cluster consistency (NMI=0.74), and interpretability (0.84), improving the accuracy and interpretability of identifying the evolution path of ecological civilization discourse in the judicial system.

Keywords: Ecological Civilization Discourse, Environmental Justice, Supreme People's Court, Evolutionary Analysis, BERTopic Model

1. Introduction

In recent years, ecological civilization has become a core category in China's national governance and policy discourse, playing a strong normative guiding role in environmental protection, resource utilization, and sustainable development[2,3,4]. Meanwhile, courts bear important functions in the judicial handling of environmental issues, including supervising administrative law enforcement, providing relief to victims, and establishing judicial rules[5,6]. Although the discourse on ecological civilization at the policy level is becoming increasingly prominent[7,8], existing observations show that it remains unclear whether and how this discourse is embedded in judicial practice, especially in the precedents and judicial interpretations of the Supreme People's Court. If the discourse on ecological civilization cannot be effectively institutionalized in the judicial field, its influence in rule-of-law governance can be difficult to realize. Therefore, systematically examining the emergence, evolution, and institutionalization path of this discourse from judicial corpora has significant theoretical and practical value.

Existing research has explored the relationship between ecological civilization discourse and law and the judiciary from multiple perspectives. At the historical and

[2] ZHANG J, FU B. Eco-civilization: A complementary pathway rooted in theory and practice for global sustainable development. Ambio, 2023, 52(12): 1882-1894. https://doi.org/10.1007/s13280-023-01902-8

[3] ZHAO W, ZHOU A, Yin C. Unraveling the research trend of ecological civilization and sustainable development: A bibliometric analysis. Ambio, 2023, 52(12): 1928-1938. https://doi.org/10.1007/s13280-023-01947-9

[4] ZHANG Z, XU H, SHAN S, LU Y, DUAN H. The impact of ecological civilization construction on environment and public health—Evidence from the implementation of ecological civilization demonstration area in China. International Journal of Environmental Research and Public Health, 2022, 19(9): 5361. https://doi.org/10.3390/ijerph19095361

[5] KOMOLIDDIN H. The role and importance of courts in environmental protection. International Journal of Law And Criminology, 2024, 4(1): 71-76. https://doi.org/10.37547/ijlc/Volume04Issue01-13

[6] HOU X, YANG J, HOU C. Can strengthening environmental justice promote carbon reduction? Evidence from environmental courts in China. Environmental Science and Pollution Research, 2024, 31(46): 57081-57098. https://doi.org/10.1007/s11356-024-31907-z

[7] HUAN Q, CHEN Y, HUAN X. A frugal eco-innovation policy? Ecological poverty alleviation in contemporary China from a perspective of eco-civilization progress. Sustainability, 2022, 14(8): 4570. https://doi.org/10.3390/su14084570

[8] ZHOU C, ZHANG W, RICHARDSON-BARLOW C. Navigating ecological civilisation: Polycentric environmental governance and policy regulatory framework in China. Energy Research & Social Science, 2025, 128(1): 1-12. https://doi.org/10.1016/j.erss.2025.104347

policy level, Westman L traced the historical roots of ecological civilization in China's environmental policy, pointing out that it relies on existing governance practices and has a substantial impact on areas such as resource protection and ecological red lines, but its transformative power in environmental litigation is limited[9]. At the jurisprudential and legislative level, An R analyzed the value basis of environmental crime legislation from the perspective of environmental philosophy, advocating the absorption of diverse environmental ethics to build a robust system[10]. In the discussion on constitutionalization, Wang E L believed that although the inclusion of ecological civilization in the constitution has strengthened the governance discourse, the operation of the system is still constrained by administrative enforcement and coordination[11]. The study on the transformation of legal thought and judicial practice points out that ecological priority is penetrating from political slogans to legislation and the judiciary, but judicial professionalization and enforcement are insufficient[12]. In the study of criminal and sentencing, Li S revealed the role and discretionary limitations of restorative justice in the sentencing of green crimes[13]. He X found that courts are more active in public interest litigation but tend to be more restrained in administrative and private litigation[14]. Wang S pointed out that the criminal law regulation of illegal hunting is out of touch with the concept of ecological civilization[15]. Zhang K assessed the controversy over the application of punitive damages in ecological torts[16]. In summary, there is a disconnect between the discourse on ecological civilization and environmental judicial practice. Although the discourse on ecological civilization has been widely discussed at the policy and legal level and has gradually permeated some legal texts and practices, there is still a lack of systematic and quantifiable evidence regarding its institutionalized embedding path in judicial judgments, especially in large-scale case corpora, how the discourse is transformed into reasoning and institutional rules, and the temporal characteristics of discourse evolution.Specifically, existing research has not yet systematically measured the frequency of occurrence, semantic type, and relationship between ecological civilization discourse in judgment reasoning and judicial interpretation and judgment basis within a unified judicial corpus framework, and also lacks comparable quantitative analysis across years and case types.

---

[9] WESTMAN L, HUANG P. Ecological civilisation a historical perspective on environmental policy narratives in China. International Quarterly for Asian Studies, 2022, 53(2): 181-206. https://orcid.org/0000-0003-4599-4996

[10] AN R, LIU P. Research on the environmental philosophy of China's environmental crime legislation from the perspective of ecological civilization construction. International Journal of Environmental Research and Public Health, 2023, 20(2): 1517. https://doi.org/10.3390/ijerph20021517

[11] WANG E L. Ecological civilisation constitutionalised: has the time come for China to constitutionalise environmental rights?. Asia Pacific Journal of Environmental Law, 2025, 28(1): 60-86. https://doi.org/10.4337/apjel.2025.01.03

[12] TAN Y, ZHANG W. The evolution and transformation of legal thinking under the concept of "Ecological Priority". Law and Humanities, 2025, 1(2): 25-33. https://doi.org/10.63313/LH.9012

[13] LI S, YOU M. Mapping restorative justice for green crime offenders: An empirical study of sentencing outcomes in China. Deviant Behavior, 2025, 46(1): 1-23. https://doi.org/10.1080/01639625.2024.2327555

[14] HE X. Judicialisation for environmental public interest protection in China: the faces of court in different forms of environmental litigation. Journal of Energy & Natural Resources Law, 2025, 43(2): 179-202. https://doi.org/10.1080/02646811.2024.2365070

[15] WANG S, AN R. Challenges and Solutions: China's Illegal Hunting Crime from the Perspective of Ecological Civilization Development. Ecological Civilization, 2024, 2(1): 1-15. https://doi.org/10.70322/ecolciviliz.2024.10016

[16] ZHANG K. The Application of Punitive Liability in the Field of Ecological Environmental Infringement. International Journal of Frontiers in Sociology, 2024, 6(12):1-6. https://doi.org/10.25236/IJFS.2024.061201

In the field of environmental law and judicial text analysis, traditional research relies on statistical text mining methods such as topic modeling to identify the topic structure and evolution of legal corpora. Raghupathi W et al. combined machine learning and LDA model to conduct topic analysis on more than 400 climate change lawsuits, extracting core issues such as forests, land, water and emissions, but the analysis was limited to keyword co-occurrence and did not reveal the semantic structure and value orientation of judicial reasoning in depth[17]. Zhao M et al. used LDA (Latent Dirichlet Allocation) model to analyze Chinese environmental policy texts and found that the policy focus shifted from "resource development" to "ecological restoration", providing a quantitative basis for policy evolution, but did not reflect the institutional transformation in the judicial context[18]. In order to improve the accuracy of topic identification, Zhang T et al. proposed an evolution path identification method based on LDA2vec, which combined word vector semantics and topic model to enhance semantic coherence and interpretability, but its application in legal corpora is limited[19]. Abeywardena S Y compared the judgments of the Supreme Court of Sri Lanka and the United Kingdom, revealing the differences in judicial priorities in different jurisdictions and demonstrating the potential of theme modeling in judicial pattern recognition, but lacked analysis of temporal evolution[20]. Hemmat Z et al. applied LDA to assess the fit between Iranian legislative texts and sustainable development goals, pointed out institutional deficiencies in the fields of education and environmental protection, and proposed a monitoring mechanism[21]. Although these studies have made significant progress in theme identification and policy analysis, they mainly focus on static theme models such as LDA and have failed to effectively capture the dynamic characteristics of legal semantics evolving over time. At the semantic level, they have neglected the semantic connections of the unique institutional logic, value orientation, and ecological discourse system in judicial judgments; at the application level, most studies focus on policy texts or cross-border legislative comparisons, while less attention is paid to the evolutionary structure and discourse embedding mechanism of case data from the highest judicial institutions.

Although the above studies have deepened the understanding of the relationship between ecological civilization and environmental justice from multiple levels such as policy, legislation, jurisprudence and case analysis, there are still three relatively clear research gaps: (1) Existing studies are mostly based on normative analysis or small sample case discussions, lacking systematic evidence based on large-scale judgment texts of the Supreme People's Court, and have not yet answered whether, how and to what extent ecological civilization discourse is institutionalized in judicial reasoning; (2) Existing text analysis mostly stays at the static or keyword level, making it difficult to depict the structural characteristics of the evolution of ecological civilization discourse over time, especially ignoring the connection between judicial interpretations, guiding cases and specific judgments; (3) Research on the

[17] RAGHUPATHI W, MOLITOR D, RAGHUPATHI V, SAHARIA A. Identifying key issues in climate change litigation: A machine learning text analytic approach. Sustainability, 2023, 15(23): 1-30. https://doi.org/10.3390/su152316530

[18] ZHAO M, WANG G, DAI D, XIE J. Policy Topic Research on Chinese Environmental Protection by an LDA Model of Text Mining. Polish Journal of Environmental Studies, 2025, 34(4): 3933-3946. https://doi.org/10.15244/pjoes/189479

[19] ZHANG T, CUI W, LIU X, JIANG L, LI J. Research on topic evolution path recognition based on LDA2vec symmetry model. Symmetry, 2023, 15(4):820. https://doi.org/10.3390/sym15040820

[20] ABEYWARDENA S Y. Reading Supreme Courts from afar: Topic modelling judgements of the Supreme Courts of Sri Lanka and the United Kingdom. University of Colombo Review, 2023, 4(1):3-30.

[21] HEMMAT Z, MEHRAEEN M, FATTAHI R, SHIRANI F. Aligning Iran's Legislation with Sustainable Development Goals: A Topic Modeling Approach. Masaryk University Journal of Law and Technology, 2025, 19(2): 127-149.

relationship between discourse and specific institutional results (such as relief types and interpretation basis) is still mainly qualitative inference, lacking quantifiable and comparable tests. This paper is aimed at the above gaps and attempts to systematically measure and analyze the evolution of ecological civilization discourse and its institutional embedding in the judgment corpus of the Supreme People's Court.

This paper aims to identify and measure the institutionalization process of ecological civilization discourse in the environmental judicial texts of the Supreme People's Court. Based on the reasoning paragraphs of 450 judgments from 2012 to 2024, the paper first performs standardized preprocessing and synonym unification. It then uses Chinese SBERT (Sentence-BERT) to generate sentence vectors, UMAP (Uniform Manifold Approximation and Projection) for dimensionality reduction, and HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) clustering to construct the BERTopic (Bidirectional Encoder Representations from Transformers Topic) model. The paper uses c-TF-IDF (Class-based Term Frequency-Inverse Document Frequency) to extract topic keywords and calculates document-topic probabilities using soft assignment, tracking weights by year. Subsequently, the topic probabilities are combined with manually coded institutionalization indicators (such as explicit_eco, cites_guidance, relief type, etc.). Contingency statistics, odds ratios, bipartite networks, and heterogeneity analysis are used to test the embedding path, and LDA/NMF (Nonnegative) is applied. Matrix Factorization/TextRank are used for robust comparisons.

2. Theoretical framework

The theoretical framework aims to clarify the three core concepts involved in this paper and their internal logical relationships, providing conceptual guidance for subsequent textual empirical analysis. "Ecological civilization discourse" is regarded as a value orientation and normative discourse embedded in the national governance discourse system, which includes governance goals and policy wording, and is also reproduced in practice through legal interpretation and judicial decisions[22,23]. "Environmental justice" refers to the court's fact-finding, legal application and relief decisions in environmental disputes. It is not only an important area of discourse practice, but also an important test of the results of discourse institutionalization[24,25,26]. In this paper, "institutional embedding" is defined as a dynamic process, referring to how the discourse on ecological civilization gradually enters the reasoning structure of judicial decisions from the expression of policies and values, and is transformed into observable rules of judgment and interpretive basis, including discourse being cited as a reason for judgment, affecting the application of clauses as an interpretation standard, and facilitating relief or procedural adjustments. Correspondingly, "institutionalization" is understood as the stable result of this

[22] WANG X, ZHAO X. From growth obsession to ecological promotion: The discursive construction of party image in Chinese political discourse on ecological civilization. Discourse & Communication, 2023, 17(6): 741-763. https://doi.org/10.1177/17504813231181080

[23] XIUFENG Z. A Social-cognitive Study of Ecological Civilization Discourse. Journal of Beijing International Studies University, 2023, 45(3): 29-45. https://doi.org/10.12002/j.bisu.451

[24] ZHENG W. Research on the Design of Multi-Objective Decision Optimization Model and Legal Suitability in Environmental Legal Dispute Resolution. Journal of Combinatorial Mathematics and Combinatorial Computing, 2025, 127(1): 1327-1345. https://doi.org/10.61091/jcmcc127b-073

[25] HELLNER A, EPSTEIN Y. Allocation of Institutional Responsibility for Climate Change Mitigation: Judicial Application of Constitutional Environmental Provisions in the European Climate Cases Arctic Oil, Neubauer, and l'Affaire du siècle. Journal of environmental law, 2023, 35(2): 207-227. https://doi.org/10.1093/jel/eqac024

[26] INDRESWARI T L, LUMBANRAJA A D, NATALIS A. Enhancing Environmental Law Enforcement in Indonesia: Integrating Climate Justice into Judicial Legal Reasoning. Pakistan Journal of Life & Social Sciences, 2024, 22(1):1836-1844. https://doi.org/10.57239/PJLSS-2024-22.1.00131

embedding process, namely, the formation of relatively fixed and repeatable adjudication patterns and institutional effects of the discourse on ecological civilization in judicial practice.Based on this framework, this paper divides the process of discourse institutional embedding into three levels: clarity, explanatory function, and institutional outcome[27]. It also clarifies the observable indicators in case texts corresponding to each level to measure the degree of institutionalization of ecological civilization discourse in the judicial field, laying a theoretical foundation for subsequent thematic modeling and coding work.

In terms of methodology and theory, this paper employs a theme modeling approach centered on semantic embedding to capture the micro-level changes in discourse within judicial texts. The model output is then juxtaposed with institutional indicators to reveal embedding paths. The theme model identifies recurring semantic clusters in judgments and their evolution over time. These semantic patterns are then paired with institutional variables such as judicial interpretations, clause citations, and types of remedies to identify which discourse components are more readily absorbed by the judicial system and under what judicial contexts discourse is more deeply embedded. This combination of theory and methodology not only responds to the empirical need in discourse research to move from declaration to institutionalization but also provides an operational analytical path for interpreting the institutionalization process of ecological civilization discourse at the highest judicial level in China.

This study concretizes the three-tiered discourse institutionalization framework into measurable indicators. "Explicitness" corresponds to topic weights and the frequency of explicit terms: quantified by document-topic probabilities obtained from BERTopic and their annually aggregated topic weights, supplemented by manually coded explicit_eco (whether terms like "ecological civilization" explicitly appear in the judgment); "Interpretive Function" corresponds to the intensity of citations of higher-level judicial interpretations/guiding cases in the text, the co-occurrence rate of topics and guiding precedents, and the Jensen-Shannon similarity of topic word distribution, used to measure the juxtaposition of discourse and interpretative texts; "Institutional Outcome" represents actual judicial arrangements through manually coded types of remedies (such as remedy_restore, remedy_compensation, and revocation/rectification orders) and changes in the application of clauses. Based on the aforementioned variables, this study employs contingency analysis, chi-square test, and advantage ratio (including 95% confidence interval and Wald test) to assess the strength of association, and uses bipartite network analysis (node weighting degree and betweenness centrality) to depict the co-occurrence structure of discourse-institutionalization indicators. Simultaneously, the robustness of the results is tested through parameter perturbation, cross-model comparisons (LDA/NMF/TextRank), and consistency indices (C_v, NMI, Cohen's κ). In this way, the theoretical three-tiered path of "explicitness—explanatory function—institutional outcome" can be concretely quantified and empirically tested at the level of judgment texts, rather than remaining merely a conceptual inference.

3. Research design and methodology

3.1. Data sources and sample selection

The corpus for this study comes from the publicly released judgment documents database of the Supreme People's Court, the guiding cases and judicial interpretations published on the official website of the Supreme People's Court, and the China Judgments Online website as a supplementary search channel. The data collection period spans from 2012 to 2024. Initial screening involved full-text searches using

[27] XIAO Y, CHEN J, YIN P. Institutionalization of Ecological Civilization Construction in China: Measurements and Influencing Factors. Sustainability, 2025, 17(4): 1719-1740. https://doi.org/10.3390/su17041719

keywords such as "ecological civilization," "ecological protection," "ecological restoration," "green development," "environmental protection," and "ecological damage compensation." The document types obtained included judgments, rulings, guiding case texts, and key points of trial guidance. After deduplication and manual review, approximately 450 documents were ultimately selected. The main unit of analysis was the paragraphs of the judgment reasoning. Guiding cases and judicial interpretations were separately marked and retained as institutionalized evidence for juxtaposition and analysis.

The approximately 450 Supreme People's Court judgments used in this paper are not a complete collection of all environmental or ecological issues during that period, but rather a conditional sample formed under specific search strategies and inclusion/rejection rules. This study focuses on "ecological civilization" and its related core concepts, emphasizing texts with explicit normative arguments and value expressions in the reasoning of judgments. Therefore, cases involving only environmental fact-finding, procedural judgments, or lacking relevant discourse were not included in the sample. This selection aims to improve conceptual consistency between the corpus and the research question, rather than providing a statistical description of the number of environmental cases handled by the Supreme People's Court. To reduce potential selection bias, this paper transparently explains the search keywords, manual screening criteria, and rejection rules. In the empirical analysis, heterogeneity tests are conducted by combining different case types and levels of adjudication to enhance the explanatory power and generalizability of the conclusions within the judicial context.

The inclusion criteria for the samples are: (1) The document is publicly available and the full text is accessible. (2) The text contains paragraphs explaining the reasoning behind the judgment related to the determination of environmental facts or the application of law.

The document contains one of the above-mentioned search keywords or is determined by manual analysis to be closely related to the discourse on ecological civilization.

The sample rejection rules include: (1) Procedural documents unrelated to environmental matters, or short documents that are merely formal explanations of case acceptance or documents without substantive reasons. (2) Duplicate archived copies and samples with text lengths significantly lower than the information requirements of the analysis unit (judgment reasoning paragraphs of less than one hundred words).

All selected documents also record necessary metadata, including case number or name, judgment date, level of jurisdiction, court name, and cause of action category, for use in subsequent grouping and heterogeneity analysis.

3.2. Data preprocessing

In the text preprocessing stage, formatting and cleaning are performed first, including removing webpage templates, headers and footers, fixed format tags in judgment documents, and non-text annotations; unifying the encoding of simplified and traditional Chinese characters and punctuation marks; and correcting common encoding errors. Subsequently, the judgment documents are segmented according to logical paragraphs, prioritizing the retention of the judgment reasoning paragraphs and merging semantically continuous short paragraphs within the same case to ensure that the modeling units possess sufficient semantic information. Proper nouns, institution names, and place names in the documents are standardized, and a thesaurus is established to unify synonyms such as "ecological restoration" and "ecological recovery," improving the consistency of subsequent theme extraction.

The word segmentation and term standardization adopts a Chinese word segmentation and dictionary expansion strategy, adding legal terms, policy terms and

ecological professional terms to the word segmentation dictionary to reduce segmentation errors[28]. At the same time, a stop word list is applied to remove non-informative functional words, but keywords such as "ecology", "environment" and "restoration" are explicitly retained. To control noise, deduplication and outlier removal are also implemented to remove documents with highly repetitive content or those containing only formal legal clause citations but lacking substantive semantics[29,30]. Finally, the processed text is merged and saved with the previously recorded metadata to form a clean corpus that can be used for subsequent semantic embedding and topic modeling, and version records are established for the preprocessing process and key dictionary changes to ensure reproducibility.

3.3. BERTopic model construction and time slicing strategy

BERTopic is a topic modeling method based on a combination of deep semantic representation and clustering. Its core idea is to introduce contextual embedding representations generated by pre-trained language models on the basis of traditional topic models in order to capture the semantic similarity between texts[31,32,33]. The model first uses the sentence vector embedding model SBERT to convert the text into high-dimensional dense semantic vectors, and then uses the unified manifold approximation and projection algorithm to reduce the dimensionality to preserve local semantic structure features. Then, the density-based spatial clustering method HDBSCAN is used to identify potential topic clusters. Finally, the c-TF-IDF method based on intra-class TF-IDF weighting is used to extract high-frequency feature words for each topic. This paper introduces the BERTopic model for text topic evolution analysis. The study uses the Chinese sentence vector model to generate embedding vectors for each preprocessed judgment reason paragraph. The embedding model uses the verified Chinese/multilingual SBERT checkpoint to ensure the consistency of sentence-level semantics and context sensitivity. In order to eliminate the scale effect, all vectors are subjected to L2 normalization, as shown in Equation (1).

$$\tilde{\mathbf{e}}_i=\frac{\mathbf{e}_i}{\|\mathbf{e}_i\|_2} \quad (1)$$

In Equation (1), $\tilde{\mathbf{e}}_i$ represents the normalized vector and $\mathbf{e}_i$ represents the embedded vector.

In order to preserve the local semantic structure and reduce noise, dimensionality reduction is performed. UMAP is applied to the normalized embedding to obtain a low-dimensional representation $\mathbf{y}_i=f_{UMAP}(\tilde{\mathbf{e}}_i)$[34]. The reason for choosing UMAP is that it

---

[28] ZHANG C. Improved word segmentation system for Chinese criminal judgment documents. Applied Artificial Intelligence, 2024, 38(1): 1-17. https://doi.org/10.1080/08839514.2023.2297524

[29] Henderson P, Krass M, Zheng L, Guha N, Manning CD, Jurafsky D,et al. Pile of law: Learning responsible data filtering from the law and a 256gb open-source legal dataset. Advances in Neural Information Processing Systems, 2022, 35(1): 29217-29234.

[30] BORROHOU S, FISSOUNE R, BADIR H. Data cleaning survey and challenges–improving outlier detection algorithm in machine learning. Journal of Smart Cities and Society, 2023, 2(3): 125-140. https://doi.org/10.3233/SCS-230008

[31] RAWAT A J, GHILDIYAL S, DIXIT A K. Topic modelling of legal documents using NLP and bidirectional encoder representations from transformers. Indones. J. Electr. Eng. Comput. Sci, 2022, 28(3): 1749-1755. https://doi.org/10.11591/ijeecs.v28.i3.pp1749-1755

[32] REJEB A, REJEB K, ZAHER H F, SIMSKE S. Blockchain and Smart Cities: Co-Word Analysis and BERTopic Modeling. Smart Cities, 2025, 8(4): 111-143. https://doi.org/10.3390/smartcities8040111

[33] AOUICHATY S, MALEH Y, MOHTADI M T, HAJAMI A, AlLali H. Sustainable topic modeling for legal Moroccan Arabic language: A challenging study on bertopic technique. Procedia Computer Science, 2024, 236(1): 582-588. https://doi.org/10.1016/j.procs.2024.05.069

[34] RAMAN R, PATTNAIK D, LATHABAI H H, KUMAR C, GOVINDAN K, NEDUNGADI P. Green and sustainable AI research: an integrated thematic and topic modeling analysis. Journal of Big Data, 2024, 11(1): 1-28. https://doi.org/10.1186/s40537-024-00920-x

can maintain the local neighborhood structure while taking into account the global coherence, thereby improving the density-based clustering recognition rate.

Density-based clustering is used to identify semantic topics. HDBSCAN is used to cluster low-dimensional representations. HDBSCAN constructs a cluster tree structure by calculating the core distance and mutual reachability distance of points and automatically selects stable clusters[35]. The mutual reachability distance between two points a and b is shown in Equation (2).

$$mreach_k(a,b) = \max\{core_k(a), core_k(b), dist(a,b)\} \quad (2)$$

In Equation (2), $core_k(a)$ represents the k-neighbor core distance, and $dist$ represents the Euclidean distance.

HDBSCAN searches for the cluster with the greatest persistence on the minimum spanning tree of mutually reachable distances, thus eliminating the need for a preset number of clusters and identifying noise points (labeled as -1). The clustering result assigns a cluster label to each document $z_i \in \{1,\ldots,K\} \cup \{-1\}$, where K is the automatically determined number of topics.

Now, this paper constructs topic representations and extracts keywords. For each cluster, it treats the set of documents and contains as a class set and use c-TF-IDF to generate topic word distributions[36]. The c-TF-IDF representation is shown in Equation (3).

$$\text{c-TF-IDF}(t,c) = \frac{tf_{t,c}}{\sum_u tf_{u,c}} \cdot \log\frac{C}{df_t} \quad (3)$$

In Equation (3), $tf_{t,c}$ represents the class frequency of term t in class c, C represents the total number of classes in the corpus, and df_t represents the number of classes in which the term appears. The first term of c-TF-IDF is the intra-class normalized term frequency, and the second term is the class-level inverse document frequency. This can emphasize terms that are highly specific and concentrated within the topic cluster, generating a semantically distinct set of topic keywords. Each topic is sorted by c-TF-IDF value and manually reviewed to name the topic. Figure 1 shows the BERTopic text topic evolution analysis framework.

In Figure 1, the overall framework begins with the data layer, using Supreme People's Court judgments, guiding cases, and judicial interpretations as the core corpus. After preprocessing stages including text cleaning, paragraph segmentation, synonym standardization, and word segmentation denoising, the system moves to the semantic embedding stage. Sentence vectors are generated using Chinese SBERT and normalized to ensure semantic consistency. UMAP is then used for dimensionality reduction to preserve local semantic structure, followed by density clustering and topic segmentation using HDBSCAN. The clustering results are then processed using the c-TF-IDF algorithm to extract topic keywords, and manually reviewed for naming to obtain semantically consistent topic sets. In the soft allocation and temporal aggregation module, the document-topic probability distribution is obtained by calculating the cosine similarity between documents and topic centers and normalizing it using Softmax. Time-slice analysis is then performed by year or policy stage to reveal trends in discourse explicitness. Furthermore, the topic weight results are juxtaposed with manually coded institutional indicators (such as judicial interpretation citations, relief types, etc.) to construct contingency analysis and co-occurrence matrices to identify the institutional embedding paths of ecological civilization discourse. Finally, the robustness and interpretability of the results are ensured through text topic evolution analysis.

---

[35] GOKCIMEN T, DAS B. Exploring climate change discourse on social media and blogs using a topic modeling analysis. Heliyon, 2024, 10(11):1-16. https://doi.org/10.1016/j.heliyon.2024.e32464

[36] UMAMAHESWARAN S, DAR V, SHARMA E, KURIAN JS. Mapping climate themes from 2008-2021—An analysis of business news using topic models. IEEE Access, 2023, 11(1): 26554-26565. https://doi.org/10.1109/ACCESS.2023.3256530

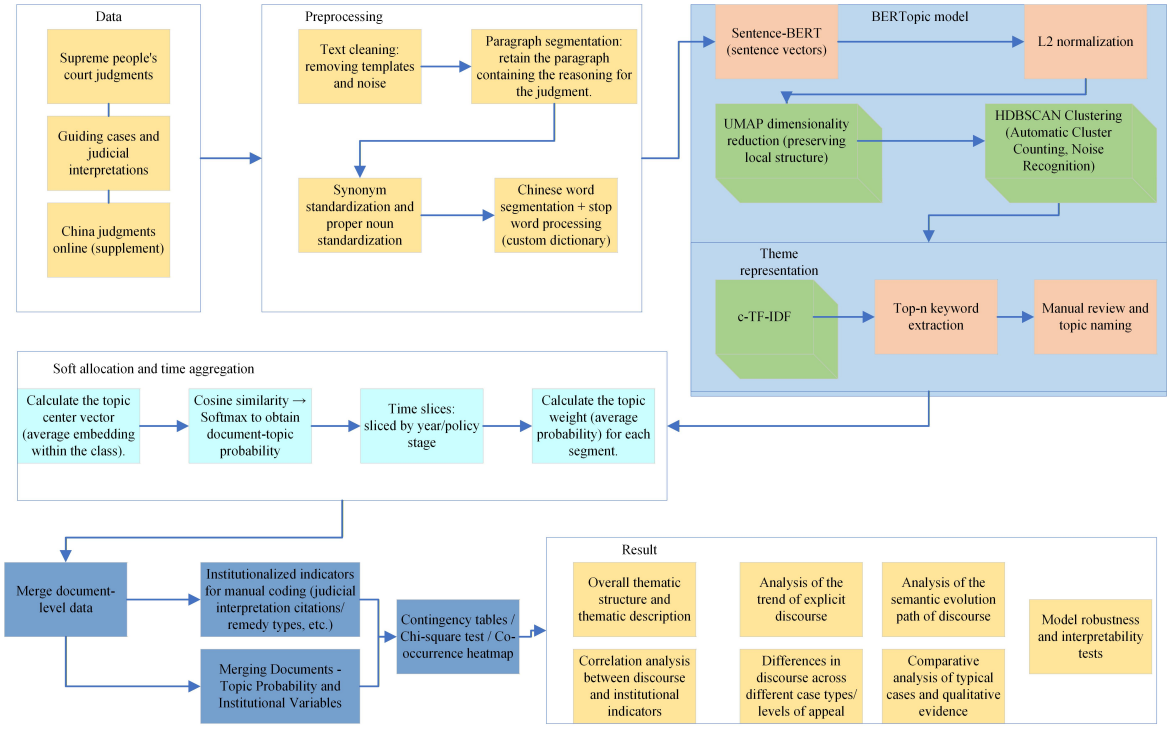


Figure 1. BERTopic text topic evolution analysis framework diagram.

To measure topic weights over time, a soft assignment strategy based on cosine similarity is used to associate documents with topic center vectors. The center vector of topic k is defined as the intra-cluster normalized embedding average, as shown in Equation (4).

$$\hat{\mathbf{c}}_k = \frac{1}{|S_k|} \sum_{i \in S_k} \tilde{\mathbf{e}}_i \tag{4}$$

In Equation (4), S_k represents the set of document indexes for topic k, and $\hat{\mathbf{c}}_k$ represents the center vector of topic k.

The similarity between document i and topic k is cosine similarity. A softmax scale is performed using the temperature parameter to obtain the document-topic probability, as shown in Equation (5). This probability characterizes the membership strength of a document across multiple topics, avoiding the loss of semantic mixing information through hard assignment.

$$p_{i,k} = \frac{\exp\left(sim(\tilde{\mathbf{e}}_i, \hat{\mathbf{c}}_k)/\tau\right)}{\sum_{j=1}^{K} \exp\left(sim(\tilde{\mathbf{e}}_i, \hat{\mathbf{c}}_j)/\tau\right)} \tag{5}$$

In Equation (5) τ represents the temperature parameter, and $p_{i,k}$ represents the document-topic probability.

The weight of each topic in the time slice at time point t is defined as the average probability of the document for that topic within that time period, as shown in Equation (6).

$$w_{t,k} = \frac{1}{|T_t|} \sum_{i \in T_t} p_{i,k} \tag{6}$$

In Equation (6), $w_{t,k}$ represents the weight of the topic, and T_t represents the set of document indexes for the time period t.

To maintain consistency of topic identifiers across different time slices, two types of similarity measures are used for cross-period matching. Jensen-Shannon divergence based on word distribution measures the overall difference in topic word distribution; Cosine similarity based on topic center embedding measures the semantic continuity of vectors. Jensen–Shannon divergence is defined as the symmetric Kullback–Leibler distance between two distributions P and Q with respect to the intermediate distribution $M = \frac{1}{2}(P + Q)$, as shown in Equation (7). A smaller JSD (Jensen-Shannon Divergence) and a higher cosine similarity together indicate the semantic continuity of topics, connecting cross-period topics into an evolutionary

chain.

$$\begin{cases} JSD(P\|Q)=\frac{1}{2}KL(P\|M)+\frac{1}{2}KL(Q\|M) \\ KL(P\|M)=\sum_x P(x)\log\frac{P(x)}{M(x)} \end{cases} \quad (7)$$

In Equation (7), $JSD(P\|Q)$ represents the Jensen–Shannon divergence.

To ensure the quality of the topics and the reliability of the evolution conclusions, the topic coherence index (coherence based on point mutual information) and the cluster consistency index (normalized mutual information) are calculated. The coherence calculation relies on word pair statistics. The UMass (University of Massachusetts) coherence approximation is defined on the topic word set using co-occurrence count and word occurrence count, as shown in Equation (8).

$$Coherence_k=\sum_{i<j}\log\frac{D(w_i,w_j)+\varepsilon}{D(w_j)} \quad (8)$$

In Equation (8), $D(w_i,w_j)$ represents the co-occurrence count, $D(w_j)$ represents the word occurrence count, and ε represents the smoothing constant to avoid zero count.

Cluster stability was quantified by repeating experiments with different embedding models, different UMAP and HDBSCAN parameters, and calculating normalized mutual information. Finally, time series weights $w_{t,k}$ and manually coded institutional indicators (such as judicial interpretation citations and relief types) were combined at the document level, and nonparametric statistical methods such as contingency tables and chi-square tests were used to help explain the relationship between topic evolution and institutional embedding.

4. Empirical results and analysis

4.1. Overall thematic structure and thematic description

In this paper, the topic-keyword table is shown in Table 1. Table 1 lists the topic number, topic size, topic name, and the top 15 keywords for each topic. Keywords are high-weight terms from the c-TF-IDF output, representing judgment paragraphs that are extractive restates of the semantics of the document in the model.

Table 1 presents eight themes covering different levels of environmental justice, from liability determination to institutional value orientation. Regarding the theme weight distribution, themes with strong practical orientation, such as ecological restoration and restoration to the original state, pollution liability and damage compensation, and environmental administrative law enforcement and judicial supervision, occupy a large share of documents, reflecting the Supreme People's Court's focus on practical remedies and liability allocation in case handling. Meanwhile, discursive themes, represented by "ecological civilization values and legal application," also have a significant presence, indicating that the policy discourse of ecological civilization has been expressed and introduced to a certain extent at the level of judicial interpretation and judgment reasoning. The keyword structure shows that different themes include both technical terms (such as monitoring, environmental impact assessment, and discharge permits) and value-based and institutional terms (such as ecological civilization, public interest, and guiding cases), indicating that judicial discourse uses technical language when dealing with specific environmental facts and also reflects policy values in the application of law.

To facilitate the display of semantic strength differences, c-TF-IDF weights (normalized to the range of 0–1) are shown based on topics and keywords (the top 10 keywords for each topic). The semantic saliency distribution of keywords for each topic is shown in Figure 2.

Table 1. Topic-keyword table.

No.	Topic size	Subject name	Top 15 keywords
1	85	Ecological restoration and restoration to original state	Ecological restoration, restoration to original state, ecological damage assessment, restoration costs, ecological compensation, restoration measures, restoration period, monitoring, restoration plan, ecological function, natural resource damage, restoration responsibility, restoration standards, expert evaluation, restoration effect
2	75	Pollution Liability and Damage Compensation	Pollution, tort liability, damages, scope of compensation, calculation of compensation, joint and several liability, pollution sources, corporate liability, environmental monitoring, evidence, compensation standards, property damage, personal injury, compensation fund, pollution control
3	60	Environmental Administrative Enforcement and Judicial Supervision	Administrative enforcement, administrative penalties, judicial supervision, administrative inaction, enforcement procedures, administrative discretion, enforcement, enforcement bodies, administrative reconsideration, enforcement evidence, legality of enforcement actions, supervisory responsibilities, judicial remedies, administrative compensation, procedural justice
4	55	Ecological Civilization Values and Legal Application	Ecological civilization, green development, sustainable development, legal interpretation, constitutional principles, value orientation, judicial application, public interest, guiding cases, judicial interpretation, principle provisions, legislative purpose, policy coherence, ecological priority, institutionalization
5	50	Environmental Public Interest Litigation and Social Organization Participation	Public interest litigation, social organizations, litigation qualifications, evidence preservation, litigation remedies, case acceptance, procuratorate, pre-litigation procedures, public participation, public interest purpose, use of compensation, environmental restoration, information disclosure, representativeness, legal representation
6	40	Ecological Protection Red Lines and Regional Planning	Ecological protection red lines, nature reserves, land use, resource development, restricted development, regional planning, ecological function zones, protection measures, approval procedures, ecological compensation mechanisms, forest land protection, sensitive areas, control, access restrictions, supervision
7	40	Environmental Impact Assessment and Project Approval Compliance	Approval compliance, environmental impact assessment, project approval, environmental impact assessment report, public participation, expert review, approval procedures, environmental protection acceptance, falsification of documents, monitoring data, supervision and inspection, rectification requirements, approval conditions, information disclosure, compliance
8	45	Pollution Control Standards and Monitoring Compliance (Air/Water)	Air pollution, water pollution, discharge permits, emission standards, excessive emissions, monitoring reports, pollutants, emergency response, pollution control projects, pollution source investigation, cost of violations, environmental quality, monitoring stations, compliance, investigation and rectification.

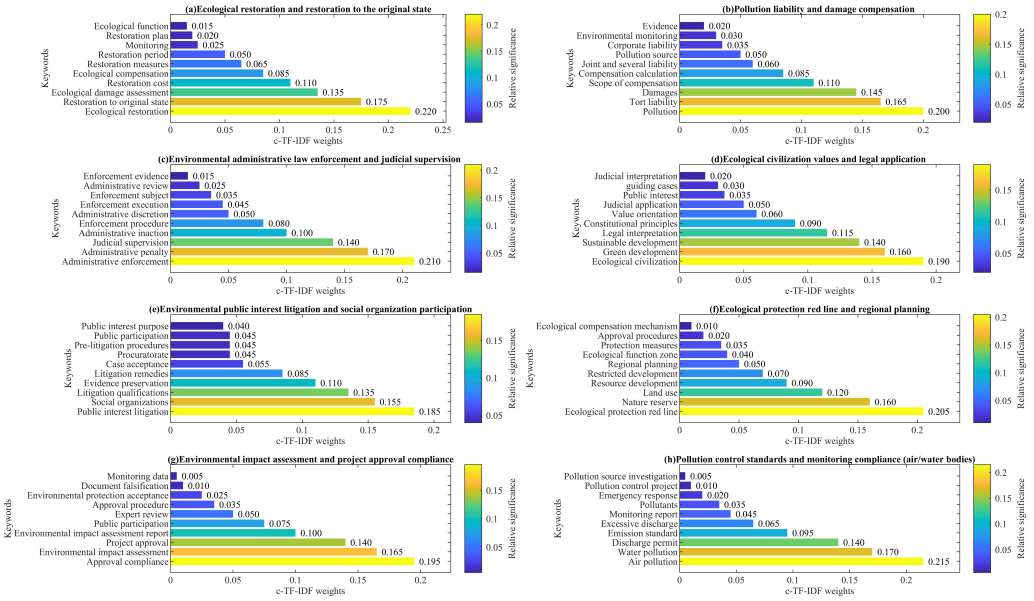


Figure 2. Semantic salience distribution of keywords for each topic. Figure 2(a) Ecological restoration and restoration to the original state; Figure 2(b) Pollution liability and damage compensation; Figure 2(c) Environmental administrative law enforcement and judicial supervision; Figure 2(d) Ecological civilization value and legal application; Figure 2(e) Environmental public interest litigation and social organization participation; Figure 2(f) Ecological protection red line and regional planning; Figure 2(g) Environmental impact assessment and project approval compliance; Figure 2(h) Pollution control standards and monitoring compliance (atmosphere/water body).

Figure 2 is constructed based on publicly released judgments, guiding cases, and judicial interpretations of the Supreme People's Court from 2012 to 2024. The unit of analysis is the semantic paragraph of the reasoning section of the judgment. After deduplication, cleaning, segmentation, and terminology standardization, the texts were analyzed using the BERTopic framework to identify eight stable topics, and the semantic saliency of keywords within each topic was calculated based on the c-TF-IDF method. Each subplot in the figure shows the top ten keywords in the corresponding topic, whose c-TF-IDF weights have been normalized to 0–1 within the topic to compare the relative importance of different keywords in the same topic. This figure reflects the relative distinctiveness and representativeness of keywords in a specific topic context, rather than their absolute frequency of occurrence in all judicial texts.

Figure 2 shows the semantic saliency distribution (c-TF-IDF weights) of the top 10 keywords in eight ecological civilization discourse themes, reflecting the core focus and weight hierarchy of each theme in the Supreme People's Court judgments. Themes with higher overall weight concentration include "ecological restoration and restoration to the original state" and "pollution control standards and monitoring compliance," with the highest keyword weights reaching 0.220 and 0.215 respectively, indicating that the judicial texts emphasize ecological restoration responsibility and pollution control mechanisms most prominently. In the themes of "ecological civilization value and legal application" and "pollution liability and damage compensation," the first keyword weights are 0.190 and 0.200 respectively, showing that ecological value orientation and damage compensation rules have become important discourse support for judicial decisions. In contrast, the weight distribution of "environmental public interest litigation and social organization participation" and "environmental impact assessment and project approval compliance" is more gradual, indicating that judicial discourse in these areas is still in a diversified exploration stage.

The overall weight distribution shows a long-tail pattern ranging from 0.220 to 0.005, reflecting that there are obvious semantic salience levels among the keywords within each theme. The top three keywords account for an average of 59% of the total weight of the theme, revealing that the discourse on ecological civilization in environmental judicial texts exhibits the characteristic of core semantic concentration.

4.2. Trend of explicit discourse

In order to conduct a trend analysis of the explicitness of discourse, the time series changes of topic weights are statistically analyzed, and the results are shown in Figure 3.

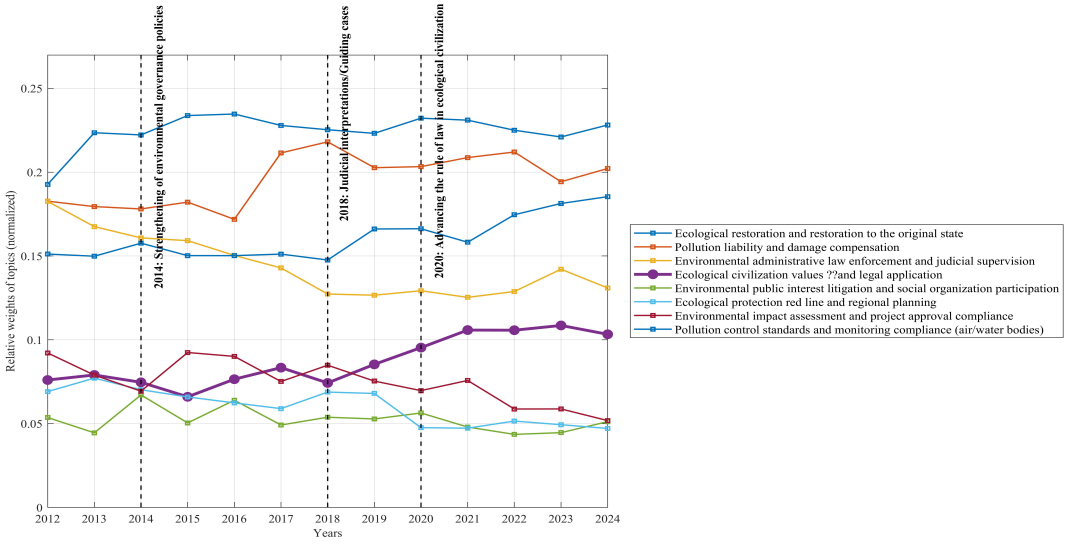


Figure 3. Results of the analysis on the trend of explicit discourse.

Figure 3 is constructed based on publicly released judgments, guiding cases, and judicial interpretations from the Supreme People's Court between 2012 and 2024, analyzing semantic paragraphs in the reasoning sections of judgments. After cleaning, segmentation, and terminology standardization, eight ecological civilization discourse themes identified using the BERTopic framework are summarized by year, showing their relative weights in the judgment texts. The theme weights shown in the figure are annually normalized results, meaning the sum of all theme weights for each year is 1, used to characterize the relative prominence of different themes in the judicial discourse structure of the same year and their changing trends over time. This figure reflects the dynamic process of ecological civilization-related discourse gradually shifting from peripheral expressions to core reasoning in judgments at the highest judicial level.

As shown in Figure 3, the weights of each theme exhibit a clear dynamic evolution trend from 2012 to 2024, reflecting the gradual expansion and structural reshaping of the discourse on ecological civilization in environmental judicial practice. Overall, "Ecological Restoration and Restoration to Original State" (Theme 1) maintained the highest weight for a long period, rising from 0.193 in 2012 to 0.228 in 2024, demonstrating the continuously strengthening dominant position of the ecological restoration concept in judicial decisions. "Pollution Liability and Damage Compensation" (Theme 2) also showed an upward trend, increasing from 0.183 in 2012 to 0.202 in 2024, indicating that the courts' determination and application of ecological damage compensation are becoming more normalized. " Ecological

Civilization Value and Legal Application " (Theme 4) significantly increased after 2018, rising from 0.076 in 2016 to a peak of 0.109 in 2023, showing the increasing explicitness of the discourse on ecological civilization in judicial reasoning and legal interpretation. While the weight of "Environmental Public Interest Litigation and Social Organization Participation" (Theme 5) was generally low (fluctuating between 0.04 and 0.06), it showed a slight increase between 2014 and 2018, which closely coincides with the establishment of the environmental public interest litigation system and the release of guiding cases by the Supreme People's Court. All eight themes experienced slight fluctuations around 2020, reflecting the impact of judicial policy adjustments and the process of legalizing ecological civilization on the discourse structure.

This difference in explicitness primarily stems from different stages of policy guidance and institutional embedding. Since the implementation of the new Environmental Protection Law in 2014 and the promulgation of the "Overall Plan for the Reform of the Ecological Civilization System," judicial practice has begun to systematically absorb the concept of ecological civilization, promoting the use of terms such as "restoration priority," "damage compensation," and "prevention principle" in judgments. In 2018, the Supreme People's Court issued a series of guiding cases and judicial interpretations on ecological and environmental protection, strengthening the normative nature of ecological civilization discourse at the level of legal application, leading to a significant increase in the explicitness of Theme 4 (the value of ecological civilization and its legal application) and Theme 1 (ecological restoration and restoration to the original state). In contrast, the explicitness of discourse on environmental impact assessment and project approval compliance, and pollution control standards and monitoring compliance (atmosphere/water bodies) (Themes 7 and 8) has increased more slowly, as they are mostly focused on administrative compliance and procedural review. Overall, the judicial system, policy nodes, and social environmental governance issues have jointly shaped the hierarchical evolution and explicitness trend of ecological civilization discourse in environmental judicial texts.

4.3. Semantic evolution path of discourse

This paper presents the keyword migration of the same theme (the value of ecological civilization and the application of law) across three time periods (2012–2016, 2017–2020, and 2021–2024) to visualize the semantic evolution path of discourse, as shown in Figure 4. In Figure 4, the node labeling format is "keyword (weight)," where the weight represents the normalized semantic saliency of the word calculated by the c-TF-IDF algorithm within the corresponding time period (the sum of weights for each time period is 100). A higher value indicates a higher discourse weight for the keyword in judicial texts.

Figure 4 is constructed based on publicly released judgments, guiding cases, and judicial interpretations from the Supreme People's Court between 2012 and 2024, focusing on the semantic paragraphs of the reasoning sections of judgments. Using the stable theme of "the value of ecological civilization and the application of law" as the core of the analysis, the sample is divided into three periods: 2012–2016, 2017–2020, and 2021–2024. Representative sets of keywords are extracted from each period. Keyword weights are calculated using the c-TF-IDF algorithm and normalized within each time period (the sum of keyword weights in each period is 100) to characterize the relative semantic salience of keywords in the judicial discourse structure of the same period. Nodes in the Sankey diagram represent core keywords in different time periods (with corresponding weights in parentheses). The flow direction and width represent the relative strength of semantic association and continuity of keywords in adjacent time periods, used to visualize the evolution of ecological civilization-related discourse from a legal interpretation-oriented approach

to a value integration-oriented approach at the highest judicial level.

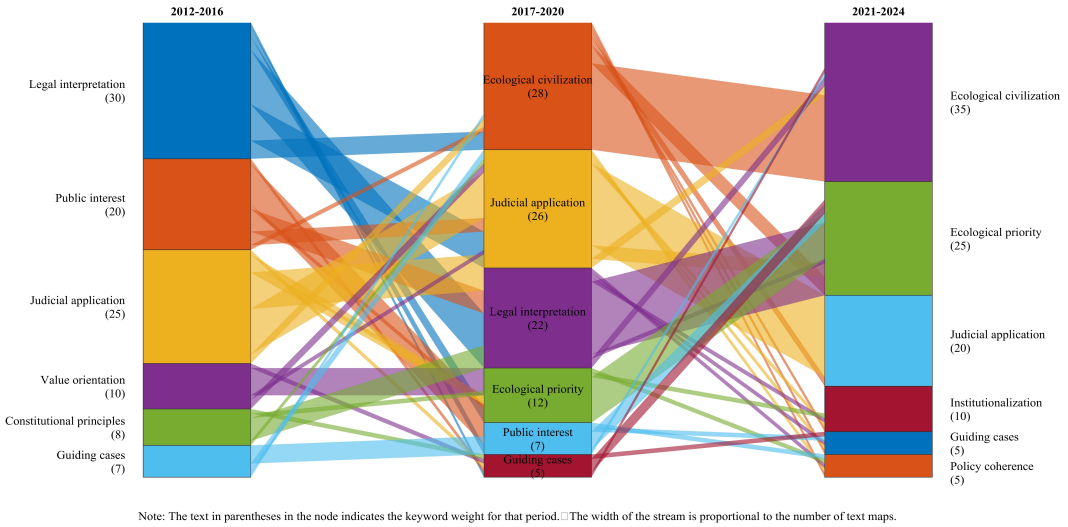


Figure 4. The path of discourse semantic evolution.

As can be seen from Figure 4, the themes of ecological civilization value and legal application underwent significant semantic evolution between 2012 and 2024. In the early stage (2012–2016), the discourse focus was concentrated on normative expressions such as "legal interpretation (30)", "judicial application (25)", and "public interest (20)", indicating that environmental justice was still centered on traditional legal logic and the relationship between rights and responsibilities. In the middle stage (2017–2020), the emergence and expansion of "ecological civilization (28)" and "ecological priority (12)" reflected that the concept of ecological civilization began to become a value orientation explicitly cited in judicial decisions; among them, the weight of the keyword "judicial application" increased from 25 to 26, showing that judges had incorporated ecological value considerations when applying the law. In the later stage (2021–2024), "ecological civilization (35)" and "ecological priority (25)" became the main keywords, and the original discourse power of "legal interpretation" declined significantly, indicating that ecological value gradually replaced simple legal logic and became the core adjudication concept in environmental justice.

This semantic shift primarily stems from the institutional impetus and judicial response following the elevation of ecological civilization construction to a national strategy. The release of the "*Overall Plan for the Reform of the Ecological Civilization System*" in 2015 and the guiding cases issued by the Supreme People's Court in 2018 prompted the systematization and institutionalization of the ecological civilization concept into judicial decision-making logic. Simultaneously, the increase in environmental public interest litigation and ecological damage compensation cases has led judicial organs to emphasize ecological restoration, public interest, and long-term sustainability in their judgments, rather than solely focusing on the division of rights and responsibilities. This discourse evolution reflects the gradual transformation of ecological civilization from a political declaration into an intrinsic value principle of legal adjudication, realizing a shift from "interpretive" to "value-based" judicial expression.

4.4. Correlation analysis between discourse and institutionalization indicators

Based on a corpus of 450 judgment documents (with initial screening numbers of 85, 75, 60, 55, 50, 40, 40, and 45 documents per topic), this paper presents the

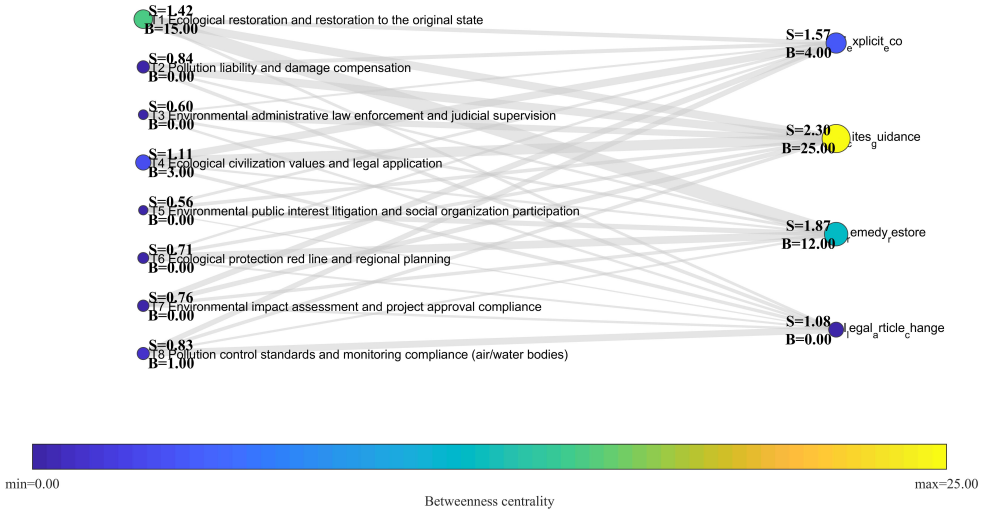
sample size for each topic, the proportion (percentage and count) of the four institutionalization indicators within documents, and the odds ratio, 95% confidence interval, and two-sided p-value for cites_guidance (whether or not the Supreme People's Court's judicial interpretations/guiding cases are cited). The odds ratio is calculated based on the comparison between "documents citing guidance within the topic" and "documents citing guidance outside the topic," using a logarithmic OR approximation and its standard error estimate. The topic-institutionalization indicator correlation matrix results are shown in Table 2.

Table 2. Results of the correlation matrix between themes and institutional indicators.

Theme	N	explicit_eco (%, n)	cites_guidance (%, n)	OR_cites_guidance (95% CI)	p (Wald)	remedy_restore (%, n)	remedy_compensation (%, n)	legal_article_change (%, n)
1	85	32.9% (28)	40.0% (34)	1.63 (0.999, 2.657)	0.051	70.6% (60)	47.1% (40)	11.8% (10)
2	75	29.3% (22)	40.0% (30)	1.61 (0.962, 2.682)	0.070	13.3% (10)	73.3% (55)	16.0% (12)
3	60	13.3% (8)	30.0% (18)	0.94 (0.521, 1.702)	0.842	13.3% (8)	50.0% (30)	10.0% (6)
4	55	63.6% (35)	50.9% (28)	2.62 (1.479, 4.643)	0.001	9.1% (5)	18.2% (10)	14.5% (8)
5	50	20.0% (10)	16.0% (8)	0.39 (0.177, 0.847)	0.018	24.0% (12)	40.0% (20)	4.0% (2)
6	40	17.5% (7)	15.0% (6)	0.36 (0.149, 0.887)	0.026	37.5% (15)	20.0% (8)	7.5% (3)
7	40	30.0% (12)	22.5% (9)	0.62 (0.286, 1.336)	0.221	15.0% (6)	37.5% (15)	10.0% (4)
8	45	28.9% (13)	15.6% (7)	0.38 (0.164, 0.866)	0.022	8.9% (4)	66.7% (30)	33.3% (15)

Table 2 shows significant differences in the correlations among the themes regarding institutionalization indicators. " Ecological Civilization Values and Legal Application " (Theme 4, N=55) had the highest proportion of explicit_eco indicators at 63.6%, and showed a strong positive correlation with citations of Supreme People's Court judicial interpretations/guidance cases. In this theme, cites_guidance accounted for 50.9%, with an OR of 2.62 (95% CI (Confidence Interval) 1.479 – 4.643) compared to literature outside the theme, p = 0.001. This result indicates that, in the sample corpus, judgment texts containing the expression "ecological civilization" appear more frequently alongside guiding documents from the Supreme People's Court.Conversely, several topics primarily focused on procedural or technical compliance (such as topics 5, 6, and 8) showed significantly lower odds ratios (OR) < 1 in cites_guidance (e.g., topic 8 OR = 0.38, 95% CI 0.164–0.866, p = 0.022), indicating that documents on these topics cited fewer Supreme Court guiding cases and leaned more towards technical remedies or administrative procedures. The types of remedies also showed thematic differentiation. Topic 1, primarily focused on " ecological restoration and restoration to the original state, " had the highest proportion of restoration-related remedies (70.6%), while topics 2 and 8 had high proportions of compensation (topic 2 compensation 73.3%; topic 8 compensation 66.7%), this indicates that there is a textual correspondence between discourse types and institutionalized outcomes.

The discourse-institutionalization indicator network diagram is shown in Figure 5. Figure 5 is a two-part network diagram. The nodes on the left represent 8 discourse topics (T1-T8), and the nodes on the right represent 4 institutionalization indicators. The thickness of the connecting lines reflects the strength of the association between the topics and the indicators (the larger the value, the stronger the association). The "S" and "B" labels next to the nodes represent the weighted degree (the sum of the node connection strength) and betweenness centrality (the strength of the network bridging effect), respectively.



Note: Edge width = association strength; node size = weighted degree (strength); color = betweenness centrality. Labels show Strength (S) and Betweenness (B).

Figure 5. Discourse-institutionalization indicator network diagram.

Figure 5 illustrates the bidirectional correlation network between discourse themes and institutionalized indicators. Overall, the network structure is relatively concentrated, with some themes showing strong connections across multiple indicators. Regarding weighting, "T4 Ecological Civilization Value and Legal Application" (strength=1.11) and "T1 Ecological Restoration and Restore to Original State" (strength=1.42) rank highly. In terms of institutionalization, "citing guidance documents" (strength = 2.30) and "restoration measures" (strength = 1.87) have high connection density, indicating that they are common accompanying elements of ecological civilization-related discourse in judicial texts. Betweenness centrality further reveals the structural position of discourse within this network. Specifically, T4 (betweenness centrality = 3.00) and T1 (ecological restoration and restoration to original state = 15.00) exhibit high structural bridging characteristics, reflecting their strong connecting or bridging role between different discourse-institutional indicator combinations. Overall, the discourse-institutionalization indicator network presents several frequently co-occurring combinations of "core discourse-institutional elements," providing empirical clues for understanding the embedding of ecological civilization discourse in judicial texts.

4.5. Differences in discourse across different case types and levels of jurisdiction (heterogeneity analysis)

To assess the differences between different case types and between different levels of judicial discourse, a heterogeneity analysis was conducted. The results of the differences in discourse across different case types and levels of judicial authority are shown in Figure 6. Based on the *Law of the People's Republic of China on the Organization of the People's Courts* and environmental judicial practice, this study categorizes case types into four types: civil, administrative, criminal, and public interest litigation; the judicial hierarchy corresponds to a four-tiered court system (basic level, intermediate level, advanced level, and Supreme People's Court). The analysis sample shows the following distribution of cases across the four types: 150 civil cases, 120 administrative cases, 80 criminal cases, and 100 public interest litigation cases; and the distribution across the four levels of judicial authority: 200 basic level cases, 120 intermediate level cases, 80 advanced level cases, and 50 Supreme People's Court cases.The levels of trial marked in the diagram—basic,

intermediate, advanced, and Supreme People's Court—are strictly distinguished according to the court level information stated in the judgments. These texts are all derived from the Supreme People's Court's public database and its systematic collection of lower court judgments.

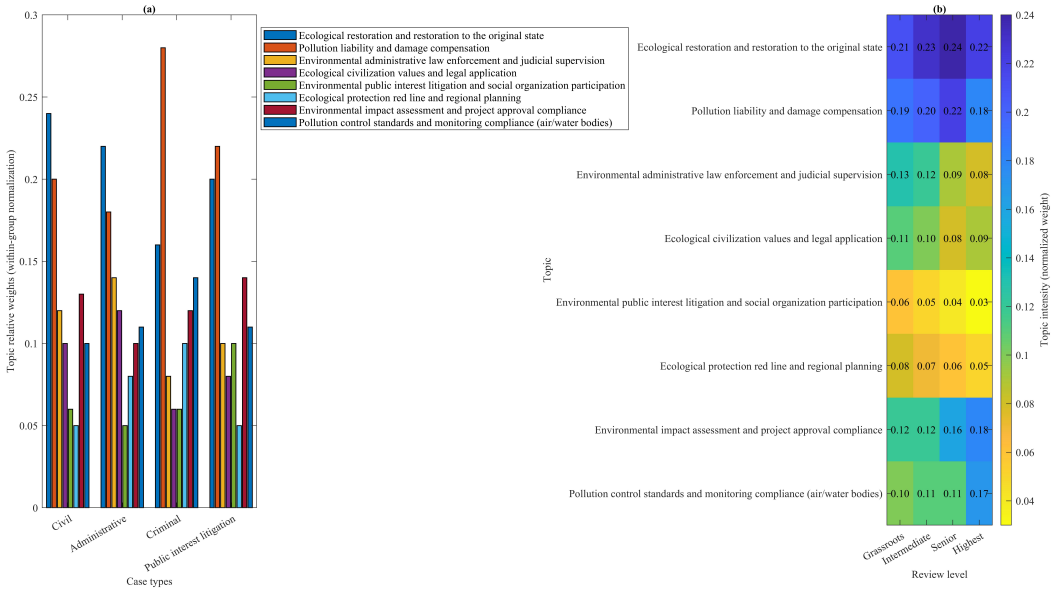


Figure 6. Differences in discourse across different case types and levels of adjudication. Figure 6(a) Distribution of subject matter intensity for different case types; Figure 6(b) Distribution of subject matter intensity for different levels of adjudication.

Figure 6(a) shows the distribution of thematic intensity for different case types. In civil cases, "ecological restoration and restoration to the original state" (0.24), "pollution liability and damage compensation" (0.20), and "environmental impact assessment and project approval compliance" (0.13) have relatively high weights. In administrative cases, "ecological restoration and restoration to the original state" (0.22), "pollution liability and damage compensation" (0.18), and "environmental administrative law enforcement and judicial supervision" (0.14) are the main themes. In criminal cases, "pollution liability and damage compensation" (0.28) has the highest weight, indicating that criminal cases focus more on accountability. In public interest litigation cases, "ecological restoration and restoration to the original state" (0.20), "pollution liability and damage compensation" (0.22), and "environmental impact assessment and project approval compliance" (0.14) are dominant. Figure 6(b) shows the distribution of thematic intensity at different levels of trial. The primary and intermediate courts have higher weights on "ecological restoration and restoration to the original state" (0.21 and 0.23) and "pollution liability and damage compensation" (0.19 and 0.20), while the higher and supreme courts have increased their attention to "environmental impact assessment and project approval compliance" (0.16 and 0.18) and "pollution control standards and monitoring compliance" (0.11 and 0.17), indicating that higher-level cases place greater emphasis on institutional norms and standard implementation.

The aforementioned differences stem from the varying nature of the cases and the different responsibilities of the adjudicators. Civil cases primarily involve the adjustment of rights and obligations, thus ecological restoration and compensation are the dominant themes. Administrative cases emphasize law enforcement supervision, reflecting the importance of administrative enforcement and judicial oversight. Criminal cases stress accountability, hence "pollution liability and damage compensation" carries the highest weight. Public interest litigation cases involve the

protection of public interests, thus the various ecological themes are relatively balanced. Regarding differences in the level of jurisdiction, lower-level courts handle a larger number and wider range of cases, focusing more on direct ecological restoration and compensation. Higher-level courts and the Supreme Court typically handle complex or exemplary cases, requiring attention to environmental impact assessments, approval compliance, and pollution control standards, reflecting a trend towards institutionalization and standardization.

4.6. Comparative analysis of typical cases and qualitative evidence

Table 3 shows a comparative analysis of typical cases and qualitative evidence, including basic case information, model themes and probabilities, and institutionalized evidence coding. Excerpts from the judgment (representative paragraphs) are shown in Table 4.

Table 3. Comparative analysis of typical cases and qualitative evidence.

No.	Case No.	Date	Level	Topic Tags	Topic Probability (top topic)	Explicit_eco	Cites_guidance	Relatives/Applicable Measures
1	SPC2019-001	2019-06-12	Supreme people's court (review)	T1	0.72	No	Yes (Cited from "Several opinions on environmental civil public interest litigation")	Restoration to original state; monitoring report
2	SPC2020-004	2020-11-20	Higher people's court	T2	0.68	No	No	Compensation + remediation
3	SPC2021-002	2021-03-15	Supreme people's court (guiding case)	T4	0.66	Yes	Yes (Cited as a guiding case)	Guiding judgments/instructions on the application of the system
4	SPC2018-007	2018-09-05	Intermediate people's court	T3	0.63	No	Yes (Cited from relevant judicial interpretations of administrative law)	Revocation of administrative decision; order for re-approval
5	SPC2022-009	2022-02-28	Supreme people's court (demonstration)	T7	0.71	No	Yes (Cited from relevant judicial interpretations/guidance on environmental impact assessment)	Revocation of approval + order for rectification
6	SPC2017-005	2017-12-10	Higher people's court	T5	0.62	No	No	Repair + information disclosure
7	SPC2023-011	2023-07-04	Intermediate people's court	T6	0.58	No	No	Stop work + restore
8	SPC2016-003	2016-05-20	Basic people's court	T8	0.60	No	No	Fine + remediation

Table 4. Excerpts from the judgment (representative paragraphs).

No.	Representative Judgment Excerpt
1	"The defendant is ordered to restore the original condition within a specified period and submit a remediation plan and subsequent monitoring report; the remediation costs shall be borne by the responsible party."
2	"Enterprises that discharge pollutants causing groundwater pollution should bear tort damages, remediation costs, and corresponding civil liability; joint and several liability is established."

No.	Representative Judgment Excerpt
3	"In applying the law, this court, referring to the concept of ecological civilization, takes 'ecological priority and green development' as the value basis for legal interpretation and adjusts the remedy structure accordingly."
4	"The administrative agency violated procedures during the approval process. The court, in accordance with the law, revoked the administrative permit and ordered the agency to make a new decision and fulfill its corresponding supervisory responsibilities."
5	"The project failed to publicize its environmental impact assessment report according to legal procedures and the report was found to be falsified. The approval should be revoked, and the project should be ordered to rectify the situation, supplement the environmental impact assessment, and disclose the information."
6	"When a public interest lawsuit is filed by a procuratorate/social organization, the court recognizes its standing to accept the case, orders the disclosure of information, and mandates that the defendant restore the damaged ecosystem within a specified period and publicly disclose the restoration progress."
7	"The land in question is located within the ecological protection red line area; therefore, the project approval was denied, and construction was ordered to cease and the original ecological environment restored."
8	"The defendant exceeded emission standards and was fined and ordered to rectify the situation based on the emission standards, and was required to submit a qualified monitoring report and rectification plan."

The eight typical cases in Tables 3 and 4 show that the topic labels output by the BERTopic model are highly consistent with the core arguments of the judgment texts in most cases (topic probability range: 0.58–0.72, with 7/8 cases having a probability ≥0.60), indicating that the model can effectively capture the clustering of semantic units surrounding remediation, compensation, environmental impact assessment compliance, and law enforcement supervision in the judgments. Judgments involving " environmental impact assessment and project approval compliance " (Case 5, p=0.71) and " ecological restoration and restoration to the original state " (Case 1, p=0.72) all contain explicit procedural wording and specific remedial orders, directly verifying the model's reliability in identifying institutionalized enforcement elements (such as "revocation of approval," "order to rectify," and "remediation plan"). Meanwhile, only a few judgments (Case 3 in Table 3) explicitly use the term "ecological civilization," indicating that the use of "ecological civilization" as an explicit legal value expression in judgment texts is still relatively rare. However, when it does appear, it often runs parallel to judicial interpretations/guiding cases, promoting the embedding of discourse into institutionalized interpretations.

This high consistency stems from the action of a triple mechanism. The BERTopic model accurately captures the contextual semantics of professional terms in judicial texts through sentence vectors generated by Chinese SBERT. For example, it automatically associates "order to rectify" and "revoke approval" with T7 (environmental impact assessment and project approval compliance), avoiding the misjudgment of the word "rectification" in different topics (such as T2 pollution liability/T7 environmental impact assessment compliance) by traditional word frequency models. The Supreme People's Court documents have highly standardized reasoning logic, such as the fixed argument chain of "remediation plan → monitoring report → cost bearing" in topic T1 (Case 1), which forms a semantic resonance with the keyword sequence extracted by the model through c-TF-IDF (Table 1). The unified implementation of synonyms (such as "ecological restoration/ecological recovery") and the enhancement of legal terminology enable the model to identify the equivalence of "restoring the original state" (civil relief term) and "ecological restoration" (policy term) in the judicial context. The accurate classification of "stop construction + restore the original state" into T6 (ecological protection red line) in Case 7 is an example. The results confirm the theoretical assumption: the institutional characteristics of judicial texts (standardized expressions + logical chains) are

structurally compatible with the semantic embedding capabilities of BERTopic, achieving stable identification of topic probabilities ≥ 0.58 without relying on manual rules.

4.7. Model robustness and interpretability tests

This study constructs a three-dimensional evaluation system to verify the robustness and interpretability of the model. Parameter robustness is assessed by perturbing the key parameters of UMAP and HDBSCAN, observing the fluctuation range of topic coherence (C_v) and stability (NMI), and measuring the model's resistance to noise in judicial texts. Method robustness is assessed by using NMI (Normalized Mutual Information) to quantify the differences in topic distribution between BERTopic and LDA/NMF/TextRank on the same corpus, evaluating the cross-model consistency of the results. Interpretability is assessed by calculating the Cohen's κ coefficient through double-blind manual annotation (three environmental law scholars independently annotated topics) and introducing a 5-level Likert scale to score the semantic clarity of topics (1 = completely incomprehensible, 5 = highly consistent with judicial logic), directly examining the fit between the model output and judicial cognition. Among these, the parameter sensitivity test is the core test of parameter robustness, reflecting the model's adaptive boundary in judicial contexts; the manual consistency test is the gold standard for interpretability, resolving the dilemma of traditional topic models being "technically effective but judicially ineffective."

The method selection follows the logic of "benchmark-evolution-application". LDA serves as a classic benchmark for topic modeling, NMF represents the non-probabilistic decomposition paradigm, and TextRank reflects commonly used graph algorithms in legal texts. These three constitute the traditional method family. BERTopic represents the forefront of deep semantic embedding. The comparison of these four methods can fully verify the adaptation requirements of semantic depth and institutional logic in judicial texts. The results of the model robustness and interpretability tests are shown in Figure 7. Figure 7(a) shows the model performance with three indicators (number of topics, coherence, and stability) under the change of UMAP/HDBSCAN parameters as axes. Figure 7(b) compares the mean values of BERTopic, LDA, NMF, and TextRank in terms of topic coherence (C_v), stability (NMI), and interpretability (human scoring). Figure 7(c) shows the consistency between human annotation and model-assigned topics (Cohen's k).

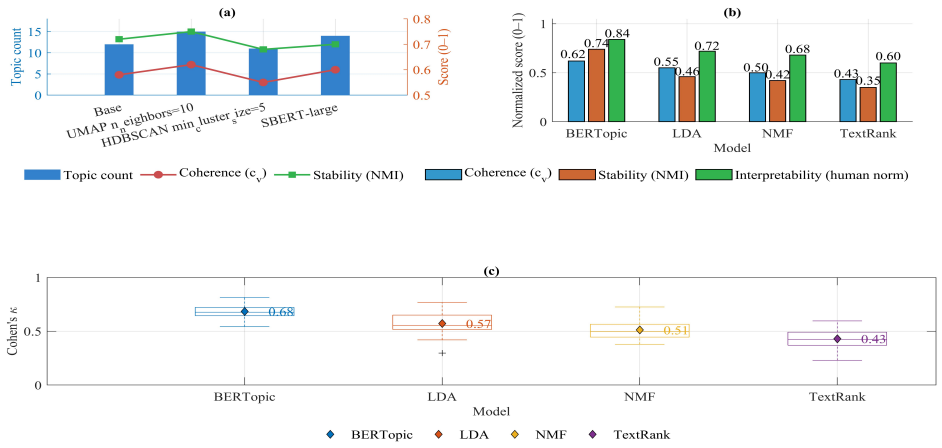


Figure 7. Model robustness and interpretability test results. Figure 7 (a) Parameter sensitivity; Figure 7 (b) Model performance comparison; Figure 7 (c) Human consistency test.

In Figure 7, the BERTopic model performs best overall in parameter sensitivity, clustering stability, and human consistency test. In the parameter sensitivity test in Figure 7(a), the number of topics is 12 under the basic configuration. When UMAP n_neighbors=10, it increases to 15, and the corresponding topic coherence and stability (NMI) reach 0.62 and 0.75, respectively, indicating that local embedding parameters can effectively enhance semantic aggregation. In the model performance comparison in Figure 7(b), BERTopic's average coherence coefficient is 0.62, stability is 0.74, and the normalized human interpretability score is 0.84, all of which are higher than LDA (0.55/0.46/0.72), NMF (0.50/0.42/0.68), and TextRank (0.43/0.35/0.60). Figure 7(c) shows that in the human consistency test, the average Cohen's κ value of BERTopic is 0.68, which is significantly better than other models (LDA: 0.57, NMF: 0.51, TextRank: 0.43), indicating that the discourse topics generated by the model are highly consistent with the human annotation results.

The main reason for this performance difference lies in the fact that the BERTopic model comprehensively utilizes SBERT semantic embedding and HDBSCAN hierarchical density clustering, enabling it to capture implicit discourse structures in a high-dimensional semantic space and achieve high-resolution segmentation of topic semantics. In contrast, traditional models such as LDA and NMF rely on word frequency statistics or linear decomposition, which are insufficient to fully represent contextual relationships and implicit value discourse in legal texts. On the other hand, the interpretability module of the BERTopic model (based on c-TF-IDF) improves the concentration of keywords, making human judgments and model clustering results more consistent, demonstrating significant advantages in both robustness and interpretability.

5. Discussion

Based on the above analysis, this paper draws the following conclusions: (1) The discourse on ecological civilization has gradually become explicit and has evolved in a stratified manner. From the perspective of the theme weights from 2012 to 2024, the representative discourse "ecological restoration and restoration to the original state" (T1) has been dominant for a long time and has increased from 0.193 to 0.228, this reflects the consistently high frequency of this repair concept appearing in the sample judgment texts; "pollution liability and damage compensation" (T2) has also steadily increased from 0.183 to 0.202, this indicates that statements related to compensation and governance occupy a relatively stable position in judicial discourse. (2) Value-based discourse (ecological civilization) shows a significant correlation with several institutional indicators in a statistical sense. Taking “Ecological Civilization Value and Legal Application” (T4) as an example, 63.6% (35/55) of the documents on this topic explicitly contain ecological civilization terms, and the OR of cites_guidance is 2.62 (95% CI 1.479–4.643, p=0.001), indicating that when the text explicitly uses ecological civilization discourse, it is more likely to be accompanied by the Supreme People’s Court’s judicial interpretation or guiding case citations—that is, value discourse enters the judicial application level through the path of “Supreme People’s Court → interpretation/case”. (3) There is a significant statistical correlation between discourse and institutional outcomes (with significant differences in relief types). For example, 70.6% (60/85) of the documents in T1 adopt the relief of restoration to the original state; while the themes mainly based on compensation/technical compliance (T2, T8) account for 73.3% and 66.7% of compensation respectively, showing that the discourse type is directly related to the institutionalized results. (4) There is a heterogeneous absorption mechanism between case types and trial levels. Civil/administrative cases place greater emphasis on restoration and compensation (civil T1=0.24, T2=0.20), while criminal cases place higher weight on accountability and compensation (T2=0.28). Meanwhile, higher/highest court levels pay more

attention to institutional norms and standards (T7, T8) (higher/highest T7=0.16/0.18, T8 increases to 0.17 at the highest court level), this reflects the structural differences in the focus of discourse in judicial texts at different levels of court.

This paper theoretically divides the framework of "discourse institutionalization" into three levels: clarity, interpretive function, and institutional outcome. This analytical perspective has been validated through textual statistics and correlation analysis. Observations show: (1) clarity of discourse (the explicit appearance of T4); (2) interpretive function (T4 documents more frequently cite Supreme Court guidance); (3) institutional outcome (different themes correspond to different types of remedies). These results support the theoretical chain of discourse, from policy statements through judicial interpretation to judgment reasoning, and ultimately manifested in remedy arrangements. Further analysis shows that the ways in which discourse is embedded in judicial texts exhibit significant differences: technical or procedural discourse (such as environmental impact assessment and monitoring) more often appears alongside administrative compliance or procedural review in the sample, but less frequently alongside Supreme People's Court interpretative documents; value-oriented discourse is statistically more concentrated in texts that appear simultaneously with high-level judicial interpretations or guiding cases. This finding suggests the need to expand the single-path understanding of institutionalization into multiple parallel or intertwined embedding modes.

Based on the above analysis, this paper summarizes a typical but non-exclusive textual association chain at the explanatory level: policy statement → (legislative/administrative text) → Supreme People's Court judicial interpretation/guiding case → exemplary judgment → specific relief arrangements. This chain aims to conceptually describe the high-frequency co-occurrence order of different text types in the sample. The analysis also shows that factors such as evidence availability, administrative-judicial division of labor, and local enforcement costs may be related to the low frequency of certain technical or procedural discourses in judicial texts, but their specific mechanisms of action require further research.

Policy and judicial recommendations:

（1）The Supreme People's Court has clarified the application framework of the principle of ecological civilization in its judicial interpretations (for example, providing operational guidance on the application conditions, evidentiary requirements, and order of remedies for "ecological priority" and "restoration priority"), in order to reduce the uncertainty of lower courts in the transformation of values.

（2）Standardize ecological discourse and remedy wording in judicial documents (provide templated judgment reasoning statements and remedy wording examples) to facilitate the traceability of value expression and institutional consistency.

（3）Strengthen the chain of evidence and link technical standards: Encourage courts to establish a regular cooperation mechanism with environmental monitoring/expert assessment agencies, clarify the probative value standards of environmental impact assessment/monitoring data in judicial determination, and reduce technical obstacles.

（4）Linking judicial practice with policy-making: It is recommended to establish a joint judicial-administrative mechanism or a regular discussion mechanism to promote timely judicial response to new policies and transparency in environmental impact assessments and acceptance processes, thereby accelerating the transformation of discourse into institutional outcomes.

The methods and data presented in this paper still have several unavoidable limitations, particularly in the area of uncertainty quantification. Topic weights, advantage ratios (OR), and their 95% confidence intervals are based on conditional samples rather than overall population data. Their statistical results reflect the correlation structure within the sample, rather than an overall inference of all environmental judicial decisions. Therefore, the width of the confidence intervals is influenced to some extent by the sample size and the stability of topic segmentation.

The semantic embedding, UMAP dimensionality reduction, and HDBSCAN clustering relied upon by BERTopic all involve parameter selection and random initialization. Although this paper assesses the robustness of the results through parameter perturbation, cross-model comparison, and consistency tests, topic boundaries may still drift slightly under different parameter combinations, introducing unobservable model uncertainty. The association analysis between topics and institutionalization indicators mainly uses contingency tables, advantage ratios, and nonparametric tests, without introducing causal inference or hierarchical regression frameworks. Therefore, correlation coefficients and significance levels cannot be interpreted as the magnitude of causal effects. Although manually coded indicators (such as explicit_eco and relief type) have undergone consistency tests, subjective judgment errors may still exist. The aforementioned uncertainties mean that the conclusions of this paper should be understood as probabilistic and structural discoveries under given corpus and methodological conditions, rather than precise measurements or causal estimates. This also provides a clear direction for the future introduction of Bayesian uncertainty modeling, larger-scale corpora, and causal inference methods.

From a comparative law perspective, the "discourse → interpretation → remedy" chain identified in this article is not unique to the Chinese judicial system, but rather shares a high degree of structural isomorphism with the common path by which value principles enter judicial decisions in civil law countries. In civil law systems, which are centered on codified law and interpretative law, policy or constitutional discourses (such as sustainable development, the preventive principle, intergenerational justice, or climate neutrality) typically do not directly generate judicial outcomes. Instead, they are first "legalized" through high-level norms or authoritative interpretations, and then transformed into interpretative benchmarks and grounds for discretion in specific cases, ultimately manifesting in remedies, liability allocation, or procedural adjustments. Whether it is the German Federal Constitutional Court reshaping the application of the proportionality principle through constitutional environmental clauses, the French administrative court system adjusting the legality review of administrative licenses through "public interest" and "environmental protection goals," or other European civil law countries coordinating policy objectives with existing legal provisions through guiding judgments or interpretative precedents, the operational logic is that value discourses do not bypass the legal structure, but are embedded in the existing normative system through interpretative mechanisms, and gain institutionalized expression at the remedy level. Therefore, the analytical framework and methodology proposed in this paper do not rely on the uniquely Chinese concept of "ecological civilization," but rather provide an operational tool for comparative research to examine how environmental or climate values in different civil law countries are transformed into specific institutional outcomes through judicial interpretation mechanisms, thereby expanding the scope of application of this study in comparative environmental justice and discourse institutionalization research.

6. Conclusion

This paper employs the BERTopic model for topic evolution analysis, constructing a reproducible topic evolution process based on Chinese SBERT, UMAP, HDBSCAN, and c-TF-IDF. It then uses soft assignment to juxtapose document-topic probabilities with institutionalized indicators of manual coding, identifying eight stable topics (T1 N=85, T4 N=55). Empirical results show that the weight of T1 increased from 0.193 in 2012 to 0.228 in 2024, and T4 appeared explicitly in 63.6% of documents and was significantly correlated with Supreme Court citations (OR=2.62, p=0.001). The BERTopic model outperforms traditional methods in coherence (c_v = 0.62), cluster consistency (NMI = 0.74), and manual consistency (Cohen's k = 0.68). The analysis supports the institutionalized path of the diffusion of ecological discourse to local

courts through Supreme Court interpretations/guiding cases and exemplary judgments. The main limitations are that the sample mainly consists of documents from the Supreme People's Court, the manual coding is subjective, and the model is sensitive to embedding and short texts. Future research should expand the corpus, strengthen multi-model comparison, and introduce causal or chain of evidence tests to consolidate the findings.

7. References

ABEYWARDENA S Y. Reading Supreme Courts from afar: Topic modelling judgements of the Supreme Courts of Sri Lanka and the United Kingdom. University of Colombo Review, 2023, 4(1):3-30.

AN R, LIU P. Research on the environmental philosophy of China's environmental crime legislation from the perspective of ecological civilization construction. International Journal of Environmental Research and Public Health, 2023, 20(2): 1517. https://doi.org/10.3390/ijerph20021517

AOUICHATY S, MALEH Y, MOHTADI M T, HAJAMI A, AlLali H. Sustainable topic modeling for legal Moroccan Arabic language: A challenging study on bertopic technique. Procedia Computer Science, 2024, 236(1): 582-588. https://doi.org/10.1016/j.procs.2024.05.069

BORROHOU S, FISSOUNE R, BADIR H. Data cleaning survey and challenges–improving outlier detection algorithm in machine learning. Journal of Smart Cities and Society, 2023, 2(3): 125-140. https://doi.org/10.3233/SCS-230008

GOKCIMEN T, DAS B. Exploring climate change discourse on social media and blogs using a topic modeling analysis. Heliyon, 2024, 10(11):1-16. https://doi.org/10.1016/j.heliyon.2024.e32464

HE X. Judicialisation for environmental public interest protection in China: the faces of court in different forms of environmental litigation. Journal of Energy & Natural Resources Law, 2025, 43(2): 179-202. https://doi.org/10.1080/02646811.2024.2365070

HELLNER A, EPSTEIN Y. Allocation of Institutional Responsibility for Climate Change Mitigation: Judicial Application of Constitutional Environmental Provisions in the European Climate Cases Arctic Oil, Neubauer, and l'Affaire du siècle. Journal of environmental law, 2023, 35(2): 207-227. https://doi.org/10.1093/jel/eqac024

HEMMAT Z, MEHRAEEN M, FATTAHI R, SHIRANI F. Aligning Iran's Legislation with Sustainable Development Goals: A Topic Modeling Approach. Masaryk University Journal of Law and Technology, 2025, 19(2): 127-149.

Henderson P, Krass M, Zheng L, Guha N, Manning CD, Jurafsky D,et al. Pile of law: Learning responsible data filtering from the law and a 256gb open-source legal dataset. Advances in Neural Information Processing Systems, 2022, 35(1): 29217-29234.

HOU X, YANG J, HOU C. Can strengthening environmental justice promote carbon reduction? Evidence from environmental courts in China. Environmental Science and Pollution Research, 2024, 31(46): 57081-57098. https://doi.org/10.1007/s11356-024-31907-z

HUAN Q, CHEN Y, HUAN X. A frugal eco-innovation policy? Ecological poverty alleviation in contemporary China from a perspective of eco-civilization progress. Sustainability, 2022, 14(8): 4570. https://doi.org/10.3390/su14084570

INDRESWARI T L, LUMBANRAJA A D, NATALIS A. Enhancing Environmental Law Enforcement in Indonesia: Integrating Climate Justice into Judicial Legal Reasoning. Pakistan Journal of Life & Social Sciences, 2024, 22(1):1836-1844. https://doi.org/10.57239/PJLSS-2024-22.1.00131

KOMOLIDDIN H. The role and importance of courts in environmental protection. International Journal of Law And Criminology, 2024, 4(1): 71-76. https://doi.org/10.37547/ijlc/Volume04Issue01-13

LI S, YOU M. Mapping restorative justice for green crime offenders: An empirical study of sentencing outcomes in China. Deviant Behavior, 2025, 46(1): 1-23. https://doi.org/10.1080/01639625.2024.2327555

RAGHUPATHI W, MOLITOR D, RAGHUPATHI V, SAHARIA A. Identifying key issues in climate change litigation: A machine learning text analytic approach. Sustainability, 2023, 15(23): 1-30. https://doi.org/10.3390/su152316530

RAMAN R, PATTNAIK D, LATHABAI H H, KUMAR C, GOVINDAN K, NEDUNGADI P. Green and sustainable AI research: an integrated thematic and topic modeling analysis. Journal of Big Data, 2024, 11(1): 1-28. https://doi.org/10.1186/s40537-024-00920-x

RAWAT A J, GHILDIYAL S, DIXIT A K. Topic modelling of legal documents using NLP and bidirectional encoder representations from transformers. Indones. J. Electr. Eng. Comput. Sci, 2022, 28(3): 1749-1755. https://doi.org/10.11591/ijeecs.v28.i3.pp1749-1755

REJEB A, REJEB K, ZAHER H F, SIMSKE S. Blockchain and Smart Cities: Co-Word Analysis and BERTopic Modeling. Smart Cities, 2025, 8(4): 111-143. https://doi.org/10.3390/smartcities8040111

TAN Y, ZHANG W. The evolution and transformation of legal thinking under the concept of "Ecological Priority". Law and Humanities, 2025, 1(2): 25-33. https://doi.org/10.63313/LH.9012

UMAMAHESWARAN S, DAR V, SHARMA E, KURIAN JS. Mapping climate themes from 2008-2021—An analysis of business news using topic models. IEEE Access, 2023, 11(1): 26554-26565. https://doi.org/10.1109/ACCESS.2023.3256530

WANG E L. Ecological civilisation constitutionalised: has the time come for China to constitutionalise environmental rights?. Asia Pacific Journal of Environmental Law, 2025, 28(1): 60-86. https://doi.org/10.4337/apjel.2025.01.03

WANG S, AN R. Challenges and Solutions: China's Illegal Hunting Crime from the Perspective of Ecological Civilization Development. Ecological Civilization, 2024, 2(1): 1-15. https://doi.org/10.70322/ecolciviliz.2024.10016

WANG X, ZHAO X. From growth obsession to ecological promotion: The discursive construction of party image in Chinese political discourse on ecological civilization. Discourse & Communication, 2023, 17(6): 741-763. https://doi.org/10.1177/17504813231181080

WESTMAN L, HUANG P. Ecological civilisation a historical perspective on environmental policy narratives in China. International Quarterly for Asian Studies, 2022, 53(2): 181-206. https://orcid.org/0000-0003-4599-4996

XIAO Y, CHEN J, YIN P. Institutionalization of Ecological Civilization Construction in China: Measurements and Influencing Factors. Sustainability, 2025, 17(4): 1719-1740. https://doi.org/10.3390/su17041719

XIUFENG Z. A Social-cognitive Study of Ecological Civilization Discourse. Journal of Beijing International Studies University, 2023, 45(3): 29-45. https://doi.org/10.12002/j.bisu.451

ZHANG C. Improved word segmentation system for Chinese criminal judgment documents. Applied Artificial Intelligence, 2024, 38(1): 1-17. https://doi.org/10.1080/08839514.2023.2297524

ZHANG J, FU B. Eco-civilization: A complementary pathway rooted in theory and practice for global sustainable development. Ambio, 2023, 52(12): 1882-1894. https://doi.org/10.1007/s13280-023-01902-8

ZHANG K. The Application of Punitive Liability in the Field of Ecological Environmental Infringement. International Journal of Frontiers in Sociology, 2024, 6(12):1-6. https://doi.org/10.25236/IJFS.2024.061201

ZHANG T, CUI W, LIU X, JIANG L, LI J. Research on topic evolution path recognition based on LDA2vec symmetry model. Symmetry, 2023, 15(4):820. https://doi.org/10.3390/sym15040820

ZHANG Z, XU H, SHAN S, LU Y, DUAN H. The impact of ecological civilization construction on environment and public health—Evidence from the implementation of ecological civilization demonstration area in China. International Journal of Environmental Research and Public Health, 2022, 19(9): 5361. https://doi.org/10.3390/ijerph19095361

ZHAO M, WANG G, DAI D, XIE J. Policy Topic Research on Chinese Environmental Protection by an LDA Model of Text Mining. Polish Journal of Environmental Studies, 2025, 34(4): 3933-3946. https://doi.org/10.15244/pjoes/189479

ZHAO W, ZHOU A, Yin C. Unraveling the research trend of ecological civilization and sustainable development: A bibliometric analysis. Ambio, 2023, 52(12): 1928-1938. https://doi.org/10.1007/s13280-023-01947-9

ZHENG W. Research on the Design of Multi-Objective Decision Optimization Model and Legal Suitability in Environmental Legal Dispute Resolution. Journal of Combinatorial Mathematics and Combinatorial Computing, 2025, 127(1): 1327-1345. https://doi.org/10.61091/jcmcc127b-073

ZHOU C, ZHANG W, RICHARDSON-BARLOW C. Navigating ecological civilisation: Polycentric environmental governance and policy regulatory framework in China. Energy Research & Social Science, 2025, 128(1): 1-12. https://doi.org/10.1016/j.erss.2025.104347