doi:10.3772/j.issn.2095-915x.2026.04.009

绿色航运政策的主题演变与区域差异分析

开放科学（资源服务）标识码（OSID）

赵凯星　于卫红　马一鸣

大连海事大学航运经济与管理学院 大连 116026

摘要：[目的 / 意义] 为应对环境挑战、推动航运业绿色转型，本研究旨在通过分析绿色航运政策文本数据，揭示政策文件中的主要主题及其时间演变趋势和地区性差异，为政策制定者提供科学的数据支持。[方法 / 过程] 研究采用结构化主题模型 (STM) 对政策文本数据进行分析。首先，收集绿色航运相关基础法规的政策文本数据。其次，对文本数据进行预处理并确定主题数量。最后使用 STM 对预处理后的文本数据进行主题建模，提取主要主题并分析其随时间的演变趋势和在不同地区的分布差异。[结果 / 结论] 研究发现，绿色航运政策的各主题在不同时间和地区的流行度差异显著，显示出政策制定者在不同历史阶段和地区对绿色航运的不同关注点。

关键词：政策文本挖掘；绿色航运政策；结构化主题模型；区域政策差异

中图分类号：G35

Thematic Evolution and Regional Differences of Green Shipping Policy

ZHAO Kaixing　YU Weihong　MA Yiming

School of Maritime Economics and Management, Dalian Maritime University, Dalian 116026, China

Abstract: [Objective/Significance] To address environmental challenges and promote the green transformation of the shipping industry, this study analyzes green shipping policy texts to identify major themes and their temporal evolution and regional differences, providing scientific data support for policymakers. [Methods/Processes] The study employs Structural Topic Modeling (STM) to analyze policy texts. First, it collects relevant policy texts on green shipping regulations. Then, it preprocesses the data and determines the number of topics. Finally, it uses the STM model to extract major themes and analyze their trends over time and regional distribution. [Results/Conclusions] The findings show significant differences in the popularity of green shipping policy themes across time and regions, reflecting varying policy priorities in different historical stages and regions.

Keywords: Policy Text Mining; Green Shipping Policy; Structural Topic Modeling (STM); Regional Policy Differences

引言

在全球环境问题日益严峻的背景下，航运业作为全球经济的重要支柱，其对环境的影响愈发受到关注[1]。绿色航运政策的制定与实施，成为推动航运业可持续发展的关键手段，然而

基金项目　2025 年大连海事大学研究所教育教学改革项目：大数据技术与应用（YJG2025503）。

作者简介　赵凯星（2001-），通信作者，硕士研究生，主要研究方向为文本分析、政策量化，E-mail：zhaokaixing2001@163.com；于卫红（1972-），博士，副教授，主要研究方向为智能信息处理、文本挖掘、政策量化分析等；马一鸣（1998-），硕士研究生，主要研究方向为文本分析。

引用格式　赵凯星，于卫红，马一鸣 . 绿色航运政策的主题演变与区域差异分析究 [J]. 情报工程，2026, 12(4): 119-127.

doi:10.3772/j.issn.2095-915x.2026.04.009

在绿色航运政策的制定与执行过程中还面临着诸多挑战。例如要提高政策实施效果需要对政策进行深入分析，而政策理解本身存在诸多困难，一方面政策文本是非结构化数据，内容复杂且表述多样，难以直接量化分析；另一方面政策文件包含大量专业术语、背景信息和目标描述，提取有用信息需要花费大量时间进行深入解读[2]。此外，政策工具种类繁多，包括法规、标准、补贴、税收等，不同工具在政策文本中的体现形式各异，增加了分析的复杂性。

为应对政策文本分析中的挑战，学术界进行了诸多尝试。传统的内容分析法通过对文本的观察、精读和逻辑推理，对文本中各要素的概念、联系及结构进行描述和推理性分析，以把握文本内容的复杂背景和思想，发掘作者的意图和文本的真正意义[3]。然而，内容分析法也存在一定的局限性，如依赖人工编码[4]，主观性强，耗时费力[5]等。

为了克服这些局限性，主题建模作为一种新兴的分析方法逐渐受到关注。主题建模利用统计模型自动识别文本主题，常见的模型包括潜在狄利克雷分配（Latent Dirichlet Allocation，LDA）和结构化主题模型（Structural Topic Model，STM）。LDA 算法通过贝叶斯推断估计文档的主题分布和主题的词汇分布，能够有效处理非结构化文本数据，自提出以来，凭借其强大的文本主题提取能力，在众多领域得到了广泛应用。LDA 常被应用于政策文本分析、情感分析和信息检索等领域。例如，杨慧等[6]利用 LDA 模型对国际气候领域的政策文本进行了量化分析。Wang 等[7]利用 LDA 主题模型对内部的政务文本数据进行挖掘，从而帮助相关部门快速发现问题与解决问题。然而，LDA 具有一定局限性，其假设主题之间是相互独立的，无法捕捉主题之间的相关性，因此难以同时考虑其他维度的信息，如政策制定的地区、政策类型等。而在研究政策文本时，由于政策文本中的主题往往存在内在联系且会随着时间等改变[8]，仅依靠 LDA 进行政策文本分析可能无法全面揭示政策主题的复杂性和动态性。

STM 算法作为 LDA 的扩展，通过在主题模型中引入文档级协变量（如时间、地区等），打破了 LDA 主题相互独立的假设，允许主题之间存在相关性，并能够通过协变量建模分析主题的动态变化。例如，Bai 等[9]利用 STM 分析了 1991—2020 年间航运领域的 3199 篇学术文章，揭示了研究主题从政策管理向高效、集成和可持续发展的演变趋势。类似地，靳彤等[10]通过政策文本计量法分析了我国科研诚信政策的特征，发现政策主题随热点事件波动，且存在多部门管理和政策不足等问题。

在航运业绿色转型的背景下，已有研究多从定性角度分析绿色航运政策的演变趋势，但对政策主题的量化分析和动态演变规律的探讨相对不足。而 STM 能够通过引入协变量分析主题的动态变化，适用于揭示政策主题的演变趋势与区域差异。因此，本文应用 STM 对绿色航运政策文本进行深度挖掘，系统提取政策中的主要主题，并对其进行时间与地理区域上的比较分析。此外，本文还将绿色航运政策按地区分类，分析各地区政策的异同，进一步深化现有文献中关于区域差异的研究。

1 相关概念与理论

1.1 绿色航运政策

绿色航运政策是指为应对航运业对环境的影响，推动航运业可持续发展而制定的一系列政策措施。这些政策涵盖了多个方面，如表 1 所示。绿色航运政策的制定与实施旨在减少航运业对环境的负面影响，提高航运业的资源利用效率，促进航运业与环境的协调发展。

表 1 绿色航运政策类型

类型	政策示例
船舶污染防治	上海市船舶污染防治条例
河流与湖泊保护	贵州省赤水河流域保护条例
海域与海洋保护	中华人民共和国海洋环境保护法
港口管理	天津市港口条例
航道管理	山东省黄河保护条例
地方性环境保护条例	安康市汉江水质保护条例

1.2 结构化主题模型

结构化主题模型是一种先进的主题建模方法，通过引入文档级别的协变量信息（如时间戳、作者、地区等），能够更精确地推断文本主题，同时捕捉主题之间的相关性，适用于从大规模文本数据中提取潜在主题结构。

STM 的核心在于其能够处理主题之间的相关性，并通过协变量建模分析主题的动态变化。这使得其在处理复杂文本数据时更具优势。同时，STM 通过引入协变量，能够更灵活地处理文本数据中的结构信息，从而提供更准确的主题推断。其生成过程可以概括为以下几个步骤：

（1）主题分布的生成：对于每个文档 d，生成主题分布 θ_d，该分布是文档特定协变量 X 的函数。

（2）词语分布的生成：对于每个主题 k，生成词语分布 β_k，该分布是主题特定协变量 Y 的函数。

（3）文档中词语的主题分配：对于文档 d 中的每个词语 n，从主题分布 θ_d 中随机选择一个主题 $z_{d,n}$。

（4）词语的生成：根据选定的主题 $z_{d,n}$ 和词语分布 $\beta_{zd,n}$，生成词语 $w_{d,n}$。

STM 的生成过程可以用以下公式表示：

$$\theta_d \mid X,\gamma,\Sigma \sim LogisticNormal(\mu = X\gamma,\Sigma) \quad (1)$$

$$\beta_{d,k,v} \propto \exp\left(m_v,\kappa_{k,v} + \kappa_{yd,v} + \kappa_{yd,k,v}\right) \quad (2)$$

其中，μ 为逻辑正态分布的均值向量，γ 为对应的回归系数矩阵，$\beta_{d,k,v}$ 表示文档 d 中主题 k 的词语 v 的概率。m_v 是词语 v 的基线对数频率。$\kappa_{k,v}$ 是主题 k 对词语 v 的偏差。$\kappa_{yd,v}$ 是协变量 Y 对词语 v 的偏差。$\kappa_{yd,k,v}$ 是主题 k 和协变量 Y 的交互项。

2 研究思路

任何政策的制定和实施都受到时间和地区因素的显著影响 [11]。政策的演变反映了社会、经济、环境等多方面需求的变化，而不同地区的政策重点则受到当地发展水平、地理条件和特定需求的制约。同样的，对于绿色航运政策，随着时间的推移，环境问题日益突出，其政策重点必然随之调整。而各地区的经济发展水平、地理条件、环境压力不同，对绿色航运政策的关注点也有所区别。这种时间和地区的双重影响，使得绿色航运政策的演变呈现出复杂性和多样性。为了深入理解这一演变过程及其背后的驱动因素，本文提出以下假设：

假设 1：绿色航运政策的主题演变与时间

doi:10.3772/j.issn.2095-915x.2026.04.009

因素显著相关。

假设 2：不同地区的绿色航运政策主题存在显著差异。

本文主要步骤如图 1 所示。

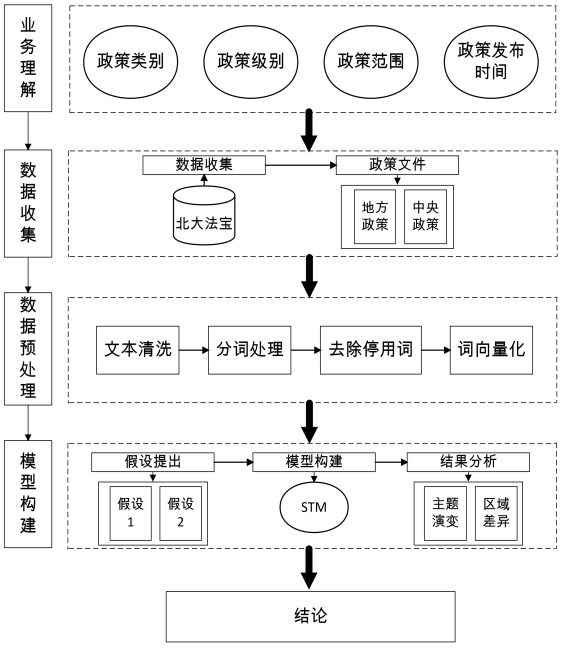

图 1　技术路线图

（1）业务理解：确定研究的对象，即绿色航运政策文本，确定政策文本的类别、级别、范围和发布时间。

（2）数据收集：从北大法宝上收集与绿色航运政策有关的中央和地方层面的政策文本数据。

（3）数据预处理：对收集到的文本数据进行清洗，去除无关内容，进行分词、去除停用词和词向量化，以确保数据的质量和适用性。

（4）主题建模与分析：基于研究问题提出假设 1 和假设 2，使用 STM 对预处理后的文本数据进行主题建模，并分析这些主题随时间的变化趋势以及在不同地区的分布情况。

（5）结论：总结研究发现。

3　实证分析

3.1　数据收集和处理

本文的数据来源于中国多个省市的地方性法规，数据覆盖华东、华南、华中、华北、东北、西北和西南地区。政策数据通过北大法宝汇编收集，研究共收集了从 2003—2025 年的 30 篇中央层面的基础法规和 83 篇地方层面的政策法规。表 2 展示了数据的基本结构内容示例，包括法规的标题、制定机关及其所在区域、施行日期、时效性、效力位阶以及具体内容等信息。法规的时效性以“现行有效”为主。效力位阶以“省级地方性法规”和“设区的市地方性法规”为主，并根据制定机关所在区域进行分类。

表 2　数据的基本结构内容示例

变量名称	数据类型	描述	示例
标题	字符型	法规的名称，描述法规的主题和内容	安徽省水路运输条例
制定机关	字符型	制定法规的机构名称	安徽省人大
制定机关所在区域	字符型	制定机关所在的地理区域	华东地区
施行日期	日期型	法规的生效日期	2017-08-01
时效性	字符型	法规的法律状态，表示法规是否仍然有效	现行有效
效力位阶	字符型	法规的法律效力层级，描述法规的法律地位	省级地方性法规
内容	字符型	法规的具体内容	水路运输发展规划应当与航道规划、港口规划、环境保护规划等相衔接……

在数据预处理阶段，首先对文本进行清洗，去除 HTML 标签、特殊字符和数字等无关内容，保留有效文本信息。接着，利用中文分词工具将文本切分为有意义的词语，并去除如“的”“是”等无意义的停用词。最后通过词袋模型将文本数据转化为数值型特征向量，以便进行主题模型分析。

3.2 主题提取

在本文中采用 STM 对绿色航运政策文本数据进行深入分析，以提取政策文本中的主要主题。首先对收集到的绿色航运政策文本数据进行了预处理。由于 STM 模型的一个重要优势是能够引入文档级别的协变量信息以帮助模型更准确地推断文本主题，并捕捉主题之间的相关性。因此在绿色航运政策领域，本文使用了以下协变量信息：（1）时间戳，即政策文本的发布或施行日期，用于分析主题随时间的演变趋势；（2）地区，即政策文本的制定机关所在区域，用于分析不同地区的主题差异。

为了确定最佳主题数量，本文参考 Chen 等 [12] 的建议，比较了不同主题数下文档拟合度、残差、语义一致性和收敛下限指标的情况，如图 2 所示。结果显示，当主题数量 K=7 时，文档拟合度达到较高水平，残差相对较低，模型拟合效果较好；语义一致性较高，主题解释性较强；收敛下限随 K 值增加而逐渐提高，表明模型收敛性良好。因此，本文选择 K=7 作为最佳主题数量。

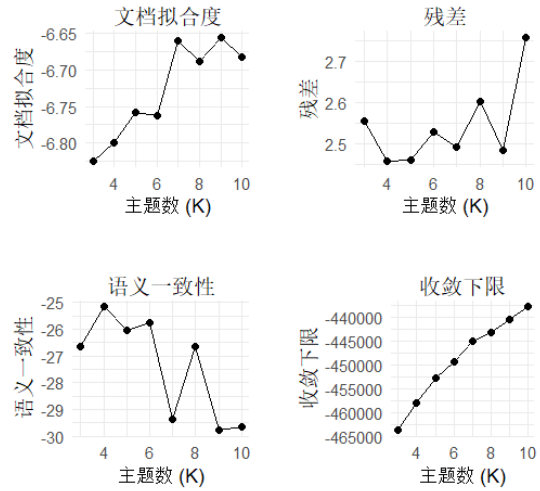

图 2 主题模型评估

通过拟合 STM，最终提取了 7 个主题，并根据高频词汇人工对每个主题进行了命名和解释，如表 3 所示。其中，7 个主题所占比重如图 3 所示。从图中可以发现主题 7（航道与港口管理）和主题 6（船舶污染与防治）是整个语料库中最突出的主题，分别占 27% 和 18%。

表 3 主题及其主要内容

主题序号	主题名称	高频词	主要涉及内容
主题 1	航运服务与设施完善	发展、船舶、建设、企业、完善、港口、服务	关注于提升航运服务的可靠性和效率，改善港口设施，优化用户体验，具体包括船舶调度优化、货物装卸效率提升等。
主题 2	船舶与航运发展	船舶、登记、国际、机构、规定、管理、证书	关注航运市场健康发展、航运企业竞争力提升、船舶制造技术创新，具体包括绿色船舶制造、企业规模扩张、航线优化、市场规范等。
主题 3	海洋与船舶服务	海洋、发展、港口、建设、国家、资源、海域	关注海洋资源合理开发、船舶配套服务优化、航运金融服务创新，具体包括海洋渔业、海域管理、船舶维修等。
主题 4	航道与港口建设	建设、发展、海洋、航运、港口、航道、区域	关注航道规划与建设、港口基础设施升级、综合运输能力提升，具体包括航道升级改造、港口码头建设、物流园区配套等。
主题 5	物流与交通运输	建设、交通运输、物流、发展、港口、联运	关注物流系统优化、运输效率提升、供应链整合，具体包括物流信息系统建设、多式联运发展、运输网络优化等。
主题 6	船舶污染与防治	船舶、污染、防治、污染物、接收、规定	关注船舶污染物排放控制、海洋环境保护、污染事故应急处理，具体包括排放标准制定、清洁能源推广、污染事故应急响应等。
主题 7	航道与港口管理	航道、港口、船舶、管理、建设、安全、通航	关注航道规划与管理、港口运营管理、航运安全监管，具体包括航道科学规划、港口智能化管理、航运安全标准制定等。

doi:10.3772/j.issn.2095-915x.2026.04.009

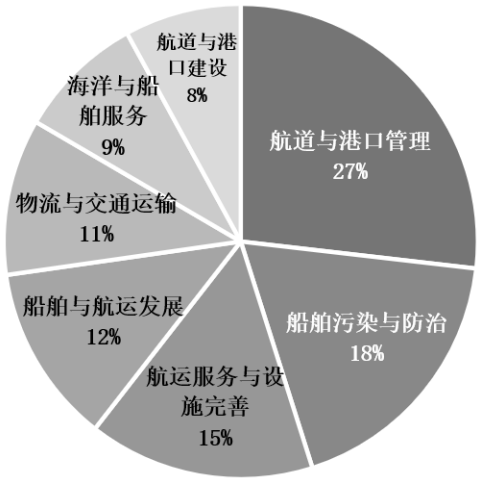

图 3 绿色政策中的主题比例

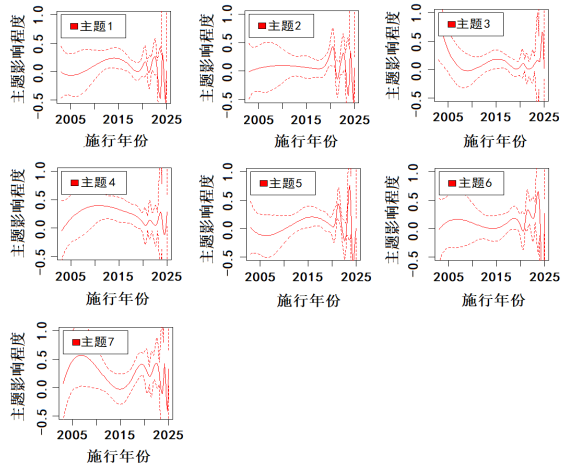

图 4 各主题流行度的变化趋势

3.3 政策主题的流行度演变分析

在主题模型分析中，主题流行度（Topic Prevalence）是衡量某一主题在语料库中被提及频率的核心指标，反映了政策制定者对该主题的关注程度。通过分析主题流行度随时间和地区的变化，可以揭示政策重点的动态演变与区域差异。在本文中，主题流行度是通过分析政策文本中各主题在整个语料库中所占比例来衡量的，流行度越高，表明该主题在政策文本中被提及的频率越高，政策制定者对其关注度也越高。为了验证假设 1，即绿色航运政策的主题演变与时间因素显著相关，本文将政策施行时间设置为协变量，分析不同主题随时间的流行度变化 [13]。图 4 展示了 2003—2025 年间各主题的流行度趋势变化，图表的横轴表示年份，纵轴表示预期主题比例（范围从 –0.5 到 1.0），其中正值表示主题在该年份的流行度较高，负值则表示流行度较低。实线依据各主题在不同时间点所占比例绘制，反映主题流行度的实际变化趋势；虚线则依据流行度的 95% 置信区间绘制，表示流行度变化的不确定性范围。

从图中可以看出，主题 4（航道与港口建设）与主题 2（船舶与航运发展）在 2003—2015 年保持高位，随后开始下滑；而主题 7（航道与港口管理）与主题 6（船舶污染与防治）自 2015 年起同步陡升，主题 3（海洋与船舶服务）、主题 5（物流与交通运输）则持续低位徘徊。这一时序特征恰好契合 Baz K 等 [14] 提出的绿色技术 - 政策生命周期萌芽、发展、扩散、成熟的四阶段模型：2003—2015 年对应“萌芽 – 发展”阶段，政策重心在于大规模基建扩张；2015 年《船舶污染防治行动计划》的出台与 2020 年“双碳”目标则构成由“发展”向“扩散”阶段跃迁的拐点，主题 7、6 的流行度陡升正是“规制与治理”阶段的政策信号；主题 3、5 的低稳态势则预示其尚处“扩散”边缘，可能成为下一轮“成熟”阶段的政策补位议题。因此，我国绿色航运政策主题的演变呈现了增量建设、存量治理、潜在整合的三阶段轨迹。由此可以看出，绿色航运政策的主题演变与时间因素显著相关，验证了假设 1 的成立。

3.4 政策主题的区域差异分析

为了验证假设 2，即不同地区的绿色航运政

策主题存在显著差异，本文对不同地区的政策文本进行了比较分析[15]。图 5 展示了各主题在不同地区的流行度分布，图表的横轴表示主题影响程度，其中正值表示该主题在相应地区的政策文本中较为突出，负值则表示相对较少被提及。纵轴列出了中国的不同地区颁布的政策，包括华北地区、华东地区、华南地区、华中地区、西北地区、西南地区、东北地区的政策以及中央颁布的政策。由于篇幅限制，以地区差异较为明显的主题 3 和主题 7 为例进行分析。从图中可以看出，主题 3 在华南沿海地区更受关注，这可能是因为它们直接依赖海洋资源和船舶服务来支持其经济活动，如渔业、旅游业和海上运输。而主题 7 在华东和华南等经济发达地区的流行度显著高于其他地区。这表明这些地区对航道与港口管理的关注度较高，这可能与这些地区的贸易活动频繁、港口经济发达有关。

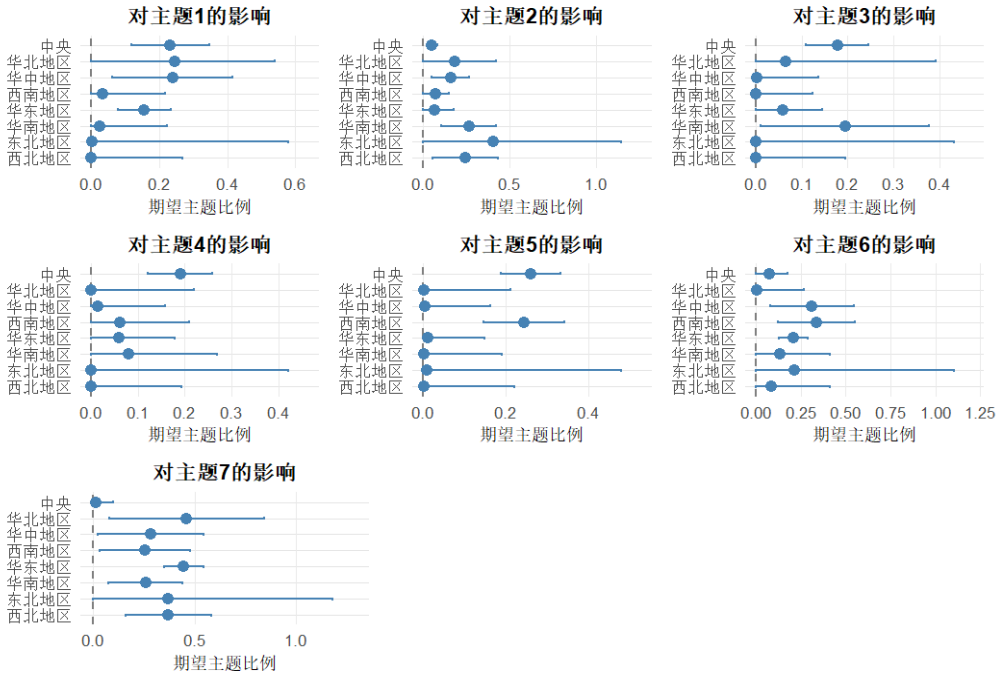

图 5　区域差异对各主题的影响

此外，分析还揭示了不同地区对绿色航运政策主题的关注点存在显著差异。华北地区的政策主题主要集中在航道与港口管理上，这可能与华北地区拥有多个重要港口（如天津港）和繁忙的航运活动有关。华东地区的政策主题主要集中在航道与港口管理和船舶污染与防治上，华东地区经济发达，拥有多个重要港口（如上海港、宁波港），因此对这些主题的关注度较高。华南地区的政策主题主要集中在航道与港口管理、船舶与航运发展和海洋与船舶服务上，华南地区沿海，海洋资源丰富，海洋经济发达，因此对海洋与船舶服务的关注度最高。华中地区的政策主题主要集中在航道与港口管理和船舶污染与防治上，这可能与该地区内河航运较为发达有关。西北地区深居内陆，在各主题上的关注度普遍较低。西南地区的政策主

doi:10.3772/j.issn.2095-915x.2026.04.009

题主要集中在航道与港口建设和船舶污染与防治上。西南地区多山，内河航运较为发达，因此对航道与港口建设的关注度较高。东北地区的政策主题主要集中在船舶与航运发展和船舶污染与防治上。东北地区是老工业基地，航运业发展较早，因此对这些主题的关注度最高。

中央与地方在绿色航运政策的总体目标上保持高度一致，但不同于地方政策根据当地实际情况，制定符合自身发展需求的政策措施，中央政策更侧重于制定全国性的宏观规划，如航道与港口的总体布局。中央政策的主题主要集中在航道与港口建设和物流与交通运输上。这表明中央政府在绿色航运政策中，重点关注航道与港口建设及物流与交通运输等方面，旨在推动全国航运业的绿色发展和可持续发展。

由此可以看出，不同地区的绿色航运政策主题存在显著差异，验证了假设 2 的成立。

4 结论

本文采用 STM 作为分析工具，对中国多个省市的地方性法规及相关中央层面的基础法规文本数据进行了深入挖掘。通过对 2003—2025 年共 113 份政策文本进行预处理、主题提取与分析，识别出 7 个主要政策主题，且以政策施行时间及地区为自变量，深入探讨了这些主题在时间演变和区域分布上的差异性。研究结果表明，绿色航运政策的主题演变与时间因素显著相关：2010 年以前的早期政策更侧重于航道与港口的基础设施建设，2015 年以后政策主题逐渐转向船舶污染与防治、航运服务优化等环保和可持续发展领域。这一演变趋势反映了政策制定者在不同历史阶段对绿色航运不同关注重点的动态调整。同时，不同地区的绿色航运政策主题存在显著差异。经济发达地区（如华东、华南）的政策更关注与航运服务优化和环保技术创新相关的主题，这与其较高的经济发展水平、繁忙的贸易活动以及对高效航运服务的迫切需求密切相关；而经济欠发达地区（如东北）则更侧重于基础设施建设和基础航运能力提升的主题，反映了这些地区通过发展基础航运业以促进经济增长和区域发展的政策诉求。

鉴于绿色航运政策主题随时间变化以及不同地区间存在显著差异，政策制定者应定期评估并灵活调整政策重点，确保政策与当前环境和社会需求同步更新。同时，考虑到地区间经济发展水平和特定需求的差异，应制定定制化的政策措施，并加强区域合作，同时关注新兴领域的政策需求以促进最佳实践的传播和区域均衡发展，从而更有效地推动航运业的可持续发展。

参 考 文 献

[1] 工业和信息化部，国家发展改革委，财政部，生态环境部，交通运输部 . 船舶制造业绿色发展行动纲要（2024—2030 年）[EB/OL]. (2023-12-26) [2025-05-07]. https://www.gov.cn/zhengce/zhengceku/202312/content_6923175.htm.

[2] 杨慧 . 社会科学研究中的政策文本分析：方法论与方法 [J]. 社会科学 , 2023(12): 5-15.

[3] 高美香 , 朱虹 , 陈红微 , 等 . 基于“政策工具 - 利益相关者”框架的我国儿童药物政策文本分析 [J]. 医药导报 , 2025, 44(7): 1179-1184.

[4] 刘青 , 赵岫岫 . 我国数字素养政策脉络演进及文本量化分析 [J]. 图书馆建设 , 2026(2): 128-136.

[5] 欧秋琦 , 刘兰萱 , 杨曦 , 等 . 基于内容分析法的多学科诊疗政策文件的可视化分析 [J]. 医学信息 , 2025, 38(6): 55-60.

[6] 杨慧 , 杨建林 . 融合 LDA 模型的政策文本量化分析 —— 基于国际气候领域的实证 [J]. 现代情报 , 2016, 36(5): 71-81.

[7] WANG D, GUO J.The Big Data Analysis and Visualization of Mass Messages under “Smart Government Affairs” Based on Text Mining [J]. Mathematical Problems in Engineering, 2022: 1-18.

[8] 刘琦岩 , 张泽玉 , 张运良 , 等 . 从主题词词频变化看我国科技政策演变 [J]. 情报工程 , 2016, 2(4): 14-19.

[9] BAI X, ZHANG X, LI K X, et al. Research topics and trends in the maritime transport: A structural topic model[J]. Transport Policy, 2021(102): 11-24.

[10] 靳彤 , 张红伟 , 赵勇 . 政策文本计量视角下我国科研诚信治理的特征与启示 [J]. 情报工程 , 2018, 4(5): 116-126.

[11] 王怡 . 立法社会效果考量的方式及其决策思维 [J]. 中国法学 , 2025(2): 188-206.

[12] CHEN S, MIAO C, ZHANG Q. Understanding the evolution of China’s green shipping policies: Evidence by social network analysis[J]. Journal of Cleaner Production, 2024, 482: 144204.

[13] 陈美 , 曹语嫣 , 张峥峥 . 同质与差异：基于结构主题模型的美国政府数据开放政策叙事研究 [J]. 情报理论与实践 , 2025, 48(1): 193-204.

[14] BAZ K, ZHU Z. Life cycle analysis of green technologies: assessing the impact of environmental policies on carbon emissions and energy efficiency[J]. Geoscience Frontiers, 2025, 16(2): 102004.

[15] PARK Y S, WANG X, OH Y, et al. Application of structural topic modeling in a literature review of air transport[J]. Journal of Air Transport Management, 2025, 122: 102708.

（责任编辑：刘泸颖）