

企业碳信息披露质量评价及影响因素研究

郭四代, 袁子寒, 雷高文*

西南科技大学 经济管理学院, 绵阳 621010

摘要: 科学有效地评价企业的碳信息披露质量并探寻影响碳信息披露的主要因素, 有助于进一步提升企业的碳信息披露质量水平、发挥企业在我国如期实现碳达峰及碳中和进程中的积极作用。基于碳信息披露的四个维度构建制造业企业碳信息披露质量的评价体系, 综合主成分分析法和熵权法测度我国制造业企业 2018—2020 年真实的碳信息披露质量水平及变化趋势, 有效降低评价指标的维数, 避免单一熵权法权重失真情况, 并通过随机森林回归识别各个指标对企业碳信息披露质量的重要性。研究发现: (1) 主成分和熵权法组合可有效降低信息量少的“干扰”维度对企业碳信息披露的影响, 能客观真实地评价我国制造业企业的碳信息披露质量, 从而获得比传统单独采用单一方法更有效的评价结果。(2) 2018—2020 年, 我国制造业企业的碳信息披露质量整体呈现上升趋势且 2020 年的碳信息披露质量水平提升最大。(3) 我国制造业企业的碳信息披露质量整体偏低, 不同企业间的碳信息披露质量存在较大差异。(4) 企业的研发投入、总资产、总负债等因素对企业碳信息披露质量有较大影响。

关键词: 制造业; 碳信息披露; 主成分分析; 随机森林

Research on quality evaluation and influencing factors of corporate carbon information disclosure

GUO Sidai, YUAN Zihan, LEI Gaowen*

School of Economics and Management, Southwest University of Science and Technology, Mianyang 621010, China

Abstract: Background, aim, and scope Scientific evaluation of enterprise carbon information disclosure quality (ECIDQ) and explore its influencing factors are crucial to achieving carbon peak and carbon neutrality in China as scheduled. The enhancement of ECIDQ is not only an important basis for improving the carbon trading system, but also an important way to promote corporate emission reduction. **Materials and methods** Based on the four dimensions of carbon information disclosure, including disclosure carrier, carbon governance, carbon business and carbon performance, an evaluation index system of ECIDQ is constructed, and the methods of comprehensive principal component analysis, entropy weight method and random forest model are integrated to study the real level and change trend of carbon information disclosure quality of manufacturing enterprises in China from 2018 to 2020, and the importance of each index. **Results** (1) The combination of principal component and entropy weight method can effectively reduce the impact of interfering dimensions with less information

收稿日期: 2022-06-30; 录用日期: 2022-08-10; 网络出版: 2022-08-25

Received Date: 2022-06-30; Accepted Date: 2022-08-10; Online first: 2022-08-25

基金项目: 国家社会科学基金项目 (20XJL013)

Foundation Item: National Social Science Fund of China (20XJL013)

通信作者: 雷高文, E-mail: leigaowen@mails.swust.edu.cn

Corresponding Author: LEI Gaowen, E-mail: leigaowen@mails.swust.edu.cn

引用格式: 郭四代, 袁子寒, 雷高文. 2023. 企业碳信息披露质量评价及影响因素研究 [J]. 地球环境学报, 14(6): 848–860.

Citation: Guo S D, Yuan Z H, Lei G W. 2023. Research on quality evaluation and influencing factors of corporate carbon information disclosure [J]. *Journal of Earth Environment*, 14(6): 848–860.

on ECIDQ, and can objectively and truly evaluate the ECIDQ of China's manufacturing enterprises, so as to obtain more effective evaluation results than the traditional single method alone. (2) From 2018 to 2020, the overall ECIDQ of China's manufacturing enterprises showed an upward trend, and reaching a peak in 2020. (3) The overall ECIDQ of China's manufacturing enterprises was low, and there were great differences in different enterprises. (4) Factors such as research and development (R&D) investment, total assets, and total liabilities have a great impact on ECIDQ. **Discussion** ECIDQ has been greatly improved in 2020, most likely due to that Chinese government put forward the goals of carbon peak and carbon neutrality in 2020, most companies disclose potential opportunities and risks in their social responsibility reports and annual reports. The score of ECIDQ with the highest quality of carbon disclosure remains unchanged, indicating that the content of corporate carbon disclosure is consistent and comprehensive. In the analysis of the influencing factors of ECIDQ, based on the root node splitting frequency and node purity increase, it is found that the asset dimension has the highest impact on ECIDQ, while the management dimension has the lowest impact. **Conclusions** From 2018 to 2020, the overall ECIDQ of China's manufacturing enterprises was low, but showed an upward trend. R&D investment, total assets, and total liabilities have a great impact on ECIDQ. **Recommendations and perspectives** In the future, it is necessary to continue to improve the incentive mechanism for carbon information disclosure, promote the increase of enterprise innovation investment, rationally and efficiently allocate enterprise assets through multiple channels, and improve the quality of corporate carbon information disclosure.

Key words: manufacturing; carbon information disclosure; principal component analysis; random forest

近年来,随着工业经济的飞速发展,环境污染问题有愈演愈烈之势,以CO₂为主的温室气体排放量与日俱增,导致全球气候变暖问题日益严峻。作为世界上最大的碳排放国家,我国2019年的碳排放总量为98.1亿t,约为美国同时期的两倍,表明我国处理气候变化问题所面临的情况更加严峻。2020年,碳达峰和碳中和这一“双碳”目标的提出,不仅展现了我国在应对气候变化上应有的责任和担当,也明确了我国进行碳减排的时间表、路线图和任务书。工业生产作为我国的主要碳源,2019年的排放量约为62亿t,占同期碳排放量的比例超过60%(史丹和李鹏,2021),而类似于金属矿物加工、化工等工业企业不仅是工业生产活动的主要承担者,也是实现“双碳”目标过程中需重点关注的对象(Zhao et al., 2017),如果这些工业企业能主动承担碳减排任务,披露碳排放信息,这将为我国如期实现“双碳”目标提供强大助力。对企业所披露的碳信息质量做出科学评价、探寻企业碳信息披露的影响因素,在实现“双碳”目标的愿景下具有重要价值。

在企业碳信息披露质量评价的相关研究中,学者们采用不同的方法对企业的碳信息披露质量进行综合评价,如直接加法(柳学信等,2021)、相对数值法(李力等,2019;张娇宁等,2021)、因子分析法(刘捷先和张晨,2020)、

功效系数法(刘长奎和边季亚,2020)、物元可拓法(李世辉等,2019)等。虽然直接加法和相对数值法可迅速对企业的碳信息披露质量进行评价,但忽略了指标间的相关性,各指标携带的重叠信息会导致评价结果存在偏误。因子分析法则可将多个变量之间的关联性简化为较少的几个因子之间的关联性,但所使用的统计数据必须服从正态分布(俞立平和刘骏,2018)。功效系数法对指标的数量和取值范围有着严格要求,运用该方法进行评价的结果的稳定性和客观性有待商榷。而物元可拓法在确定经典域及其节域时暂无统一标准。可见,现有研究中关于碳信息披露评价指标的赋权方法相对单一、未充分考虑指标数据之间的内在统计规律,有可能发生权重失真情况,最终导致评价结果无效。

在企业碳信息披露的影响因素方面,学者们大多采用多元线性回归的方式研究企业内外部因素对企业碳信息披露的影响。在企业高管及董事会特征对企业碳信息披露的影响上,唐成林(2017)分析840家上市工业企业的发现,高管的政治关联会正向促进企业的碳信息披露,而郭嵘和陆萍(2017)发现高管的任期和薪酬激励则会产生负向影响。在企业特征对碳信息披露的影响上,众多学者均发现了企业规模、股权性质、大股东持股比例、盈利能力等因素对企业的碳信息披

露有正向影响（杨璐等，2017；童心刚，2019；王志亮和杨媛，2019；温雅丽等，2019）。在企业所处的外部环境上，传统文化、法律制度环境以及公共压力等因素均会促进企业的碳信息披露（李慧云等，2018；唐勇军等，2018），而政府干预这一因素则对企业的碳信息披露产生负向影响（马歆等，2018）。尽管已有较多学者研究了企业内外部因素对企业碳信息披露的影响，但较少有文献对企业碳信息披露各影响因素的重要程度进行系统性分析。

综上所述，要对企业碳信息披露的质量进行科学合理的评价，需充分考虑指标数据之间的内在统计规律，采用组合赋权以弥补单一赋权法的缺陷。主成分分析法可以从收集的众多指标中提取少数几个主成分，不仅可以避免测评指标间的信息重复，又可以抓住主要评估维度，客观高效地对研究对象进行评估与识别（陈远等，2019），且对研究数据的先验分布无特殊要求（俞立平和刘骏，2018）。熵权法则是利用指标数据中所携带的信息熵的大小赋予指标权重，是一种常用的客观赋权方法，避免了主观判断带来的偏差，指标所携带的数据越分散，熵值越小，所包含的信息就越大，对应的权重就越大。然而若要全面综合评价企业碳信息披露质量水平，指标选取一般较多，熵权法则明显不能有效降低评价指标的维数；同时，该方法忽略了指标与指标之间的横向影响，过多的评价指标难免会导致变量之间存在较强的相关性，指标携带的信息出现重叠问题，也会产生数据共线性现象。基于此，本文将主成分分析法和熵权法结合以期降低评价指标的数据维度，同时对评价指标进行客观赋权，从而对我国 A 股上市公司中制造业企业的碳信息披露质量进行科学评价。此外，为了能准确把握企业碳信息披露的主要影响因素，考虑到随机森林回归模型相较于传统的统计方法有着不受变量的正态性、线性性等假设的约束这一优势（张雷等，2014），且能深入分析输入变量的重要程度，因此，本文在评价企业碳信息披露质量的基础上，采用随机森林回归模型进一步分析碳信息披露各影响因素的重要性。

1 研究方法 with 数据来源

1.1 研究方法

1.1.1 企业碳信息披露质量评价——指标体系

参照已有研究（宋晓华等，2019；柳学信

等，2021；张娇宁等，2021；黄炳艺等，2023），将制造业企业的碳信息披露质量评价指标体系分为“披露载体”、“碳治理”、“碳业务”和“碳绩效”4个一级指标和对应的12个二级指标，具体内容见表1。

1.1.2 企业碳信息披露质量评价——评价方法

评价制造业企业碳信息披露的关键步骤是要选择合适的指标以及确定各指标之间的权重。参考已有研究（王莺等，2014；史学飞等，2018；赵宁等，2019；崔丹等，2021），结合主成分分析法和熵权法，对制造业企业碳信息披露的质量进行评价，具体步骤如下：

假设研究所需的数据空间是由 m 个指标和 n 个观测值所组成的 n 行 m 列矩阵 A 。

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \quad (1)$$

(1) 将矩阵 A 中各元素按列进行标准化无量纲化

$$x_{ij} = (a_{ij} - \bar{a}_j) \div \sigma_j, (i=1, 2, \dots, m; j=1, 2, \dots, n) \quad (2)$$

式中： x_{ij} 为矩阵 X 的第 i 行第 j 列标准化后的值， a_{ij} 为指标原始观测值， \bar{a}_j 表示矩阵 A 第 j 列元素的均值， σ_j 表示矩阵 A 第 j 列元素的标准差。标准化后得到新的 n 行 m 列矩阵 X 。

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \quad (3)$$

(2) 计算矩阵 X 相关系数矩阵 R

$$R = \begin{pmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & & \vdots \\ r_{m1} & \cdots & r_{mm} \end{pmatrix} \quad (4)$$

$$r_{jk} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j) \times (x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \times \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}}, (j=1, 2, \dots, m;$$

$$k=1, 2, \dots, m) \quad (5)$$

式中： r_{jk} 表示相关系数矩阵 R 的第 j 行第 k 列的元素。 \bar{x}_j 表示矩阵 X 第 j 列元素的平均值， \bar{x}_k 表示矩阵 X 第 k 列元素的平均值。

(3) 计算主成分得分

通过求解特征方程 $|\lambda E - R| = 0$ ，可得矩阵 R 的特征值 λ_i 及其对应的特征向量 e_i ，由此可计算主成分：

$$C_i = X \times e_i, (i = 1, 2, \dots, m)$$

式中: E 为单位矩阵, C_i 表示第 i 个主成分。

(4) 保留主成分数量

计算主成分 C_i 的方差贡献率:

$$w_i = \lambda_i \div \sum_{j=1}^k \lambda_j$$

(6)

为保证较小的主成分信息损失, 将 w_i 从大到小排序并按照累积和大于 80% 的原则选取前 k 个特征值对应的主成分, 则主成分得分矩阵为:

$$C = \begin{pmatrix} c_{11} & \dots & c_{1k} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nk} \end{pmatrix} \quad (8)$$

表 1 评价指标及其赋值标准
Tab. 1 Evaluation index and its assignment criteria

一级指标 Primary indicator	二级指标 Secondary indicator	赋值标准 Assignment criteria	符号 Symbol
披露载体 Disclosure carrier	独立报告披露状况。 Disclosure status of independent reports.	披露在社会责任报告中为 2 分, 披露在年报中为 1 分, 否则 0 分。 2 points for disclosure in social responsibility, 1 point for disclosure in the annual report, otherwise 0 points.	I_1
	管理层低碳声明及理念。 Management's low carbon statement and philosophy.	披露了管理层低碳声明及理念为 1 分, 否则 0 分。 1 point if the management's low-carbon statement and philosophy are disclosed, otherwise 0 points.	I_2
	员工低碳理念及行为。 Low carbon concept and behavior of employees.	披露理念及行为的具体事项为 2 分, 有理念和行为但无具体事项为 1 分, 否则 0 分。 2 points for disclosing specific items of ideas and behaviors, 1 point for ideas and behaviors but no specific items, otherwise 0 points.	I_3
碳治理 Carbon governance	碳减排相关岗位或机构的设置。 Setting up of jobs or institutions related to carbon emissions reduction.	披露了机构或岗位负责的具体事项 2 分, 设立机构但未披露负责具体事项为 1 分, 否则 0 分。 2 points for disclosing the specific matters that the organization or position is responsible for, 1 point for establishing an organization but not disclosing the specific matters in charge, otherwise 0 points.	I_4
	碳减排考核或激励机制。 Carbon emissions reduction assessment or incentive mechanism.	设立了考核或激励机制为 1 分, 否则 0 分。 1 point for establishing an assessment or incentive mechanism, otherwise 0 points.	I_5
	制定碳减排目标及计划。 Develop carbon reduction targets and plans.	定量描述 2 分; 定性描述 1 分; 否则 0 分。 2 points for quantitative description; 1 point for qualitative description; otherwise 0 points.	I_6
碳业务 Carbon business	与碳排放相关的业务和财务风险。 Business and financial risks associated with carbon emissions.	定量描述 2 分; 定性描述 1 分; 否则 0 分。 2 points for quantitative description; 1 point for qualitative description; otherwise 0 points.	I_7
	与碳排放相关的业务和财务机遇。 Business and financial opportunities related to carbon emissions.	定量描述 2 分; 定性描述 1 分; 否则 0 分。 2 points for quantitative description; 1 point for qualitative description; otherwise 0 points.	I_8
	碳减排技术研发或项目投资。 Carbon emissions reduction technology research and development or project investment.	定量描述 2 分; 定性描述 1 分; 否则 0 分。 2 points for quantitative description; 1 point for qualitative description; otherwise 0 points.	I_9
碳绩效 Carbon performance	碳减排情况。 Carbon emissions reduction.	定量描述 2 分; 定性描述 1 分; 否则 0 分。 2 points for quantitative description; 1 point for qualitative description; otherwise 0 points.	I_{10}
	碳排放交易及资产管理。 Carbon emissions trading and asset management.	定量描述 2 分; 定性描述 1 分; 否则 0 分。 2 points for quantitative description; 1 point for qualitative description; otherwise 0 points.	I_{11}
	碳排放的核算方法或标准。 Accounting methods or standards for carbon emissions.	披露了具体的核算方法或标准为 2 分, 说明按照要求核算 1 分, 否则为 0 分。 2 points for quantitative description; 1 point for qualitative description; otherwise 0 points.	I_{12}

(5) 主成分正向化

由于主成分的值可能存在负数, 需要对主成分进行正向化:

$$f_{ij} = (c_{ij} - \min c_j) \div (\max c_j - \min c_j), (i=1, 2, \dots, n; j=1, 2, \dots, k) \quad (9)$$

式中: f_{ij} 表示正向化后的主成分得分, $\min c_j$ 表示 C 的第 j 列元素的最小值, $\max c_j$ 表示第 j 列元素的最大值。

(6) 计算信息熵

$$E_j = -\frac{1}{\ln n} \times \sum_{i=1}^n p_{ij} \times \ln p_{ij} \quad (10)$$

$$p_{ij} = f_{ij} \div \sum_{i=1}^n f_{ij} \quad (11)$$

式中: E_j 表示第 j 项主成分的信息熵, p_{ij} 表示主成分 f_{ij} 占全部主成分的比重。

(7) 确定各主成分权重

$$Q_j = (1 - E_j) \div \sum_{i=1}^k (1 - E_j) \quad (12)$$

(8) 计算碳信息披露质量

$$G = \sum_{i=1}^n f_{ij} \times Q_j \quad (13)$$

1.1.3 随机森林

准确把握制造业企业碳信息披露的主要影响因素能促进其碳信息披露质量的进一步提升, 通过随机森林回归模型, 可对制造业企业碳信息披露质量影响因素的重要性进行排序, 从而识别主要影响因素。随机森林是机器学习算法中的一类, 基于决策树发展而来, 可以将多棵决策树进行集成, 其中每棵树都依赖于独立采样的随机向量的值且具有相同分布 (Breiman, 2001)。通过 Bootstrap 抽样从原始样本中随机抽取同样数量的样本子集, 利用每个子集建立决策树 (刘敏等, 2015), 结合多棵决策树可以有效解决模型的过拟合问题, 使得结果更加符合现实情况。在本文中, 由于输出变量为企业的碳信息披露质量, 这是一个连续变量, 因此, 需采用基于回归树建立的随机森林回归模型进行分析。相较于传统的统计方法, 回归树不受变量的正态性、线性性等假设的约束 (张雷等, 2014)。参考李航 (2019)、曹桃云 (2022) 的研究, 随机森林回归模型的构建原理如下:

假设数据集 D 有 N 个元素, D 中的每个元素是由输入变量 X 和输出变量 Y 所组成的集合, 可将数据集 D 表示为:

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}, (x_1 \leq x_2 \leq \dots \leq x_N) \quad (14)$$

以输入变量 X 的取值 v 为依据, 将数据集 D 划分为 D_1 和 D_2 两个子集, D_1 中有 M 个元素, 使得输入变量 X 的取值小于等于 v , 有:

$$D_1 = \{(x_1, y_1), (x_2, y_2), \dots, (x_M, y_M), |x \leq v\}, (M < N) \quad (15)$$

$$D_2 = \{(x_{M+1}, y_{M+1}), (x_{M+2}, y_{M+2}), \dots, (x_N, y_N), |x > v\} \quad (16)$$

对于 X 而言, 需要确定划分 D_1 和 D_2 的具体数值 v , 通过求解下式可确定 v 的取值, 以及最优切分点 M^* 。当存在 p 个输入变量时, 遍历每个输入变量, 求解下式可确定回归树的根节点输入变量 s^* 。

$$s^* = \min_s [\min_{c_1} \sum_{i=1}^M (y_i - c_1)^2 + \min_{c_2} \sum_{i=M+1}^N (y_i - c_2)^2], (s=1, 2, \dots, p) \quad (17)$$

$$c_1 = \frac{1}{M^*} \times \sum_{i=1}^{M^*} y_i \quad (18)$$

$$c_2 = \frac{1}{N - M^*} \times \sum_{i=M^*+1}^N y_i \quad (19)$$

式中: c_1 和 c_2 分别表示 D_1 和 D_2 上所有输入变量 x_i 对应的输出变量 y_i 的均值。

在确定根节点的输入变量以及相应的 D_1 和 D_2 后, 对每个区域的数据重复上述过程, 直到满足结束条件为止, 这样就生成了一棵回归树。在建立回归树的基础上, 利用 Bootstrap 抽样, 最终建立随机森林。随机森林建立后, 需要根据特定的原则确定各输入变量 X 以及对输出变量 Y 的影响程度, 本文依据随机森林中决策树根节点的分裂频次和输入变量的节点纯度增加量两个原则对各输入变量的重要性进行分析。

(1) 决策树根节点的分裂频次

决策树的根节点分裂频次是指在随机森林的所有决策树中, 以某一输入变量作为根节点进行分裂的决策树的数量。假设数据中有 n 个输入变量 X_1, X_2, \dots, X_n , 而随机森林由 m 棵决策树组成, 需对数据样本进行 m 次 Bootstrap 抽样, 得到 m 个子样本, 每个子样本按照回归树的构建原理分别生成决策树, 由于有 m 个子样本, 因而会得到 m 棵决策树, 对于每棵决策树而言, 其根节点对应的变量由式 (17) 确定。由此, 统计变量 X_1, X_2, \dots, X_n 中每个变量作为根节点而生成的决策树的

数量 m_1, m_2, \dots, m_n , 变量 $X_j (j=1, 2, \dots, n)$ 对应的 m_j 的值越大, 表明该变量对于输出变量的影响程度越大。

(2) 节点纯度增加量

节点纯度增加量是指决策树中以某一变量为父节点分裂后左右子节点的不纯度下降的程度, 对于数据集 D 而言, 父节点的不纯度为:

$$I = \frac{1}{N} \times \sum_{i=1}^N (y_i - \bar{y})^2 \quad (20)$$

式中: \bar{y} 表示数据集 D 中 y 的均值。

经过分裂后的左右子节点对应的数据集分别为 D_1 和 D_2 , 相应节点不纯度为 I_1 和 I_2 。此时节点纯度增加量为:

$$\Delta I = I - I_1 - I_2 \quad (21)$$

在由 m 棵决策树组成的随机森林中, 以变量 X_j 为父节点进行分裂的节点纯度增加量为:

$$F_j = \frac{1}{p} \times \sum_{i=1}^p \text{sum}(\Delta I_i) \quad (22)$$

式中: F_j 表示随机森林中变量 X_j 的节点纯度增加量, 该值越大, 表明变量 X_j 对于输出变量 Y 的影响程度越大。 p 表示含有变量 X_j 的决策树的数量 ($p \in [0, m]$), $\text{sum}(\Delta I_i)$ 表示决策树 i 中变量 X_j 的所有节点纯度增加量的总和。

1.2 数据来源

选取 2018—2020 年我国 A 股上市公司中的

制造业企业作为研究样本, 根据《GB/T 4754—2017, 国民经济行业分类》(<https://www.mca.gov.cn/n156/n187/n319/index.html>), 选取上市公司中所属行业代码为 C 开头的企业, 同时剔除了 ST、PT 以及同时发行 A、B 股的企业, 最终保留了 751 家制造业企业; 企业的碳信息披露数据来源于企业所披露的年报和社会责任报告, 利用 Python, 在巨潮资讯网 (<http://www.cninfo.com.cn/new/index>) 上爬取所研究企业的相应报告, 之后通过内容分析法, 对照评价的指标体系得到各指标具体数据; 企业的其他相关数据来自中国经济金融研究数据库 (CSMAR) (<https://data.csmar.com/>)。

2 结果与分析

2.1 我国制造业企业碳信息披露质量现状

在对数据进行主成分分析前, 对各变量之间的相关性进行检验, 从而判别主成分分析是否适用。基于 Stata17 平台计算所选指标的 Spearman 相关系数, 同时进行 KMO 检验和 Bartlett 球形检验, 检验结果分别见表 2 和表 3。由表 2 可知: 各指标间基本上都具有显著的相关关系。由表 3 可知: 数据的 KMO 值为 0.714, 大于 0.5, 而且 Bartlett 球形检验的显著性概率 P 值小于 0.001, 表明各变量之间具有较强的相关性。表 2 和表 3 的结果共同表明研究数据适合做主成分分析。

表 2 指标间相关系数

Tab. 2 Correlation coefficient between indicators

指标	Indicator	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}
I_1		1											
I_2		0.69*	1										
I_3		0.45*	0.33*	1									
I_4		0.17*	0.12*	0.16*	1								
I_5		0.13*	0.12*	0.08*	0.21*	1							
I_6		0.27*	0.22*	0.16*	0.12*	0.09*	1						
I_7		0.43*	0.09*	0.06*	0.01	0	0.23*	1					
I_8		0.45*	0.12*	0.05*	0	0.05	0.23*	0.77*	1				
I_9		0.45*	0.36*	0.25*	0.19*	0.13*	0.18*	0.08*	0.12*	1			
I_{10}		0.40*	0.33*	0.29*	0.18*	0.11*	0.23*	0.01	0.04	0.29*	1		
I_{11}		0.20*	0.09*	0.08*	0.09*	0.02	0.05*	0.06*	0.04	0.09*	0.12*	1	
I_{12}		0.20*	0.17*	0.19*	0.01	0.07*	0.09*	0.06*	0.03	0.10*	0.19*	0.14*	1

* 表示在 0.1 的显著性水平下显著。

* indicates significant at a significance level of 0.1.

利用 Stata17 平台对数据进行主成分分析, 得到的特征值和累积贡献率如表 4 所示。依据累积

贡献率超过 80% 的原则, 共提取 7 个主成分。各主成分的得分系数矩阵如表 5 所示, 可知每个主

成分均为各变量的线性组合。

通过表 5 确定每个主成分的表达式并求取相应的值, 对各主成分进行正向化后, 采用熵权法赋予各主成分相应的权重, 如表 6 所示。可知主成分 C_1 所占权重最大, 超过 90%, 而主成分 C_3 所占权重最小, 仅为 0.525%, 各主成分之间的权重比例存在较大差异。

表 3 KMO 和 Bartlett 球形检验结果
Tab. 3 Result of KMO test and Bartlett's test of sphericity

检验方法	指标	检验结果
Testing method	Indicator	Test result
KMO 检验	KMO 值	0.714
KMO test	KMO value	
Bartlett 球形检验	χ^2	7086.224
Bartlett's test of sphericity	自由度 Degrees of freedom	66
	P	0.000

表 5 主成分得分系数矩阵
Tab. 5 Principal component score coefficient matrix

指标	Indicator	C_1	C_2	C_3	C_4	C_5	C_6	C_7
I_1		0.477	-0.006	-0.108	-0.160	-0.207	-0.161	0.097
I_2		0.380	-0.156	-0.092	-0.268	-0.376	-0.207	-0.019
I_3		0.331	-0.195	-0.057	-0.186	0.187	0.200	0.247
I_4		0.168	-0.196	0.511	0.285	0.403	-0.200	0.546
I_5		0.136	-0.126	0.509	0.461	-0.618	0.119	-0.159
I_6		0.278	0.111	0.263	0.037	0.394	0.304	-0.638
I_7		0.251	0.620	0.003	0.080	0.033	0.012	0.150
I_8		0.259	0.611	0.042	0.047	-0.043	-0.010	0.105
I_9		0.323	-0.154	0.023	-0.139	0.038	-0.233	-0.114
I_{10}		0.325	-0.262	0.002	-0.081	0.206	0.153	-0.175
I_{11}		0.137	-0.063	-0.432	0.621	0.173	-0.517	-0.239
I_{12}		0.187	-0.130	-0.446	0.391	-0.076	0.632	0.253

表 6 各主成分权重
Tab. 6 The weight of principal component

主成分	Principal component	权重	Weight/%
C_1		90.914	
C_2		2.186	
C_3		0.525	
C_4		3.587	
C_5		0.531	
C_6		1.142	
C_7		1.115	

求取各主成分的权重后, 可对制造业企业的碳信息披露水平进行综合评价, 表 7 和表 8 分别展示了 2018—2020 年制造业企业碳信息披露质量

表 4 特征值及方差贡献率
Tab. 4 Eigenvalue and variance contribution rate

主成分符号	特征值	方差贡献率	累积贡献率
Principal component notation	Eigenvalue	Variance contribution rate	Cumulative contribution rate
C_1	3.314	0.276	0.276
C_2	1.687	0.141	0.417
C_3	1.139	0.095	0.512
C_4	1.022	0.085	0.597
C_5	0.845	0.070	0.667
C_6	0.835	0.070	0.737
C_7	0.780	0.065	0.802
C_8	0.700	0.058	0.860
C_9	0.645	0.054	0.914
C_{10}	0.619	0.052	0.966
C_{11}	0.225	0.019	0.984
C_{12}	0.189	0.016	1.000

得分均值和排名前十的制造业企业的具体数值。从表 7 中可知: 2018 年和 2019 年 2 a 内制造业企业的碳信息披露质量变化不大, 得分约为 0.09, 处于较低水平, 2020 年约为 0.14, 相较于 2018 年和 2019 年提升较大, 这是由于我国 2020 年提出了碳达峰和碳中和的“双碳”目标, 大部分企业在其社会责任报告和年报中纷纷披露了潜在的机遇和风险。从表 8 可知: 碳信息披露质量第一的企业得分没有变化, 这表明在 2018—2020 年这 3 a 该企业的碳信息披露内容一致, 披露的内容较为全面。结合表 7 和表 8 可知, 我国制造业企业的碳信息披露质量参差不齐, 且整个行业的碳信息披露水平存在巨大的提升空间。

表7 2018—2020年制造业上市公司碳信息披露质量均值
Tab. 7 Average quality of carbon information disclosure of manufacturing listed companies from 2018 to 2020

年份 Year	均值 Mean
2018	0.09006
2019	0.08959
2020	0.14244

表8 2018—2020年制造业上市公司碳信息披露质量前十得分

Tab. 8 Top 10 scores of carbon information disclosure quality of manufacturing listed companies from 2018 to 2020

排名 Ranking	2020	2019	2018
1	0.945424	0.945424	0.945424
2	0.806704	0.817144	0.727500
3	0.760698	0.806704	0.719967
4	0.677623	0.727500	0.648139
5	0.659433	0.559792	0.587007
6	0.653411	0.548885	0.548885
7	0.645758	0.538234	0.532264
8	0.625425	0.532264	0.521047
9	0.613417	0.521047	0.473056
10	0.613210	0.473056	0.443052

此外, 为分析不同区域间企业碳信息披露质量的差异, 分别计算了2018—2020年我国东中西三个区域企业碳信息披露质量的均值, 各区域的变动情况见图1。从图1可知: 2018年和2019年西部地区的碳信息披露质量都略高于其他两个区域, 但在2020年, 东部地区的碳信息披露质量略高于其他两个区域。在变动趋势上, 中部和西部地区的碳信息披露质量呈现先降低后提升的现象, 而东部地区的碳信息披露质量处于稳步提升状态。值得注意的是, 虽然2020年三个区域的碳信息披露质量相较于2018年和2019年均有了较大提升, 但三个区域的碳信息披露均存在较大提升空间。

2.2 制造业企业碳信息披露的影响因素

以往关于碳信息披露影响因素的研究主要聚焦于各因素对企业碳信息披露的促进或者抑制作用, 未能准确把握各因素的重要性。因此, 利用机器学习中的随机森林回归模型, 对影响制造业企业碳信息披露质量的各因素的重要程度进行深入分析。

2.2.1 影响因素的选择

基于数据的可得性和完整性, 结合已有文献,

将企业碳信息披露的内部影响因素划分为创新、资产、营业、管理、员工五个维度并对各因素的重要性进行分析。其中, 创新维度包含了研发人员数量(X_1)、研发人员数量占比(X_2)、研发投入金额(X_3)以及研发投入占营业收入比例(X_4)4个影响因素; 资产维度包含了企业的总资产(X_5)、总负债(X_6)以及无形资产(X_7)3个因素; 营业维度包含了归属于股东的净利润(X_8)、净利润(X_9)、营业收入(X_{10})、营业成本(X_{11})以及营业利润(X_{12})5个因素; 管理维度包含了独立董事占比(X_{13})、监事会规模(X_{14})、高管人数(X_{15})、管理层男性占比(X_{16})以及管理层平均年龄(X_{17})5个因素; 员工维度包含了员工人数(X_{18})、人均创利(X_{19})、超额雇员率(X_{20})以及员工密集度(X_{21})4个因素。五个维度总共21个影响因素的名称以及符号见表9。

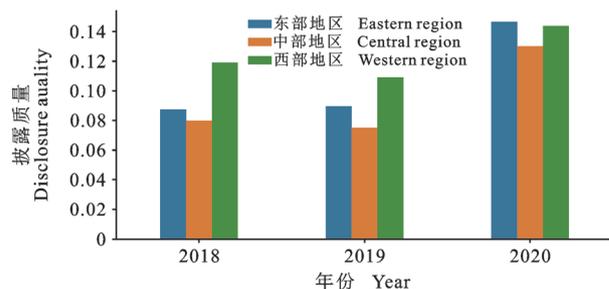


图1 不同区域间碳信息披露质量对比

Fig. 1 Comparison of carbon information disclosure quality among different regions

2.2.2 影响因素重要性分析

为探寻各因素对于制造业企业碳信息披露质量影响的重要程度, 将前文制造业企业碳信息披露质量作为输出变量, 以五个维度的相关指标作为输入变量, 基于Python语言在Jupyter Notebook平台分别进行随机森林回归。按照均方误差最小原则确定的随机森林决策树数量为246棵, 在此基础上, 分别从节点纯度增加量和根节点分裂频次两个方面对企业碳信息披露的内部影响因素的重要性进行分析, 具体结果分别见图2和图3。

图2是以根节点分裂频次为依据, 对我国制造业企业碳信息披露影响因素的重要性进行分析的结果。从图2可知: 在随机森林的246棵决策树中, 以变量研发投入金额(X_3)为根节点生成的决策树的数量超过40棵, 而以总资产(X_5)为根节点生成的决策树超过30棵, 对以各变量为根

节点生成的决策树的数量进行排序, 由多到少依次是研发投入金额 (X_3)、总资产 (X_5)、总负债 (X_6)、营业收入 (X_{10})、无形资产 (X_7)、员工人数 (X_{18})、研发人员数量 (X_1) 和营业成本 (X_{11}) 等, 可知以根节点分裂频次为依据, 对

制造业企业碳信息披露影响程度最大的变量为研发投入金额 (X_3)。此外, 以研发人员数量占比 (X_2)、研发投入占营业收入比例 (X_4) 等 5 个变量为根节点生成的决策树数量均为零, 表明这些变量对制造业企业碳信息披露的影响程度较低。

表 9 企业碳信息披露质量的内部影响因素
Tab. 9 Internal factors influencing the quality of corporate carbon information disclosure

变量维度 Variable dimension	变量名称 Variable name	变量符号 Variable symbol
创新 Innovation	研发人员数量 Number of R&D personnel	X_1
	研发人员数量占比 The proportion of the number of R&D personnel	X_2
	研发投入金额 R&D investment amount	X_3
	研发投入占营业收入比例 R&D investment as a percentage of operating income	X_4
资产 Assets	总资产 Total assets	X_5
	总负债 Total liability	X_6
	无形资产 Intangible assets	X_7
营业 Business	归属于股东的净利润 Net profit attributable to shareholders	X_8
	净利润 Net profit	X_9
	营业收入 Operating income	X_{10}
	营业成本 Operating cost	X_{11}
	营业利润 Operating profit	X_{12}
管理 Management	独立董事占比 Proportion of independent directors	X_{13}
	监事会规模 Size of supervisory board	X_{14}
	高管人数 Number of executives	X_{15}
	管理层男性占比 The percentage of men in management	X_{16}
	管理层平均年龄 Average age of management	X_{17}
员工 Staff	员工人数 Number of workers	X_{18}
	人均创利 Per capita profit	X_{19}
	超额雇员率 Overemployment rate	X_{20}
	员工密集度 Staff density	X_{21}

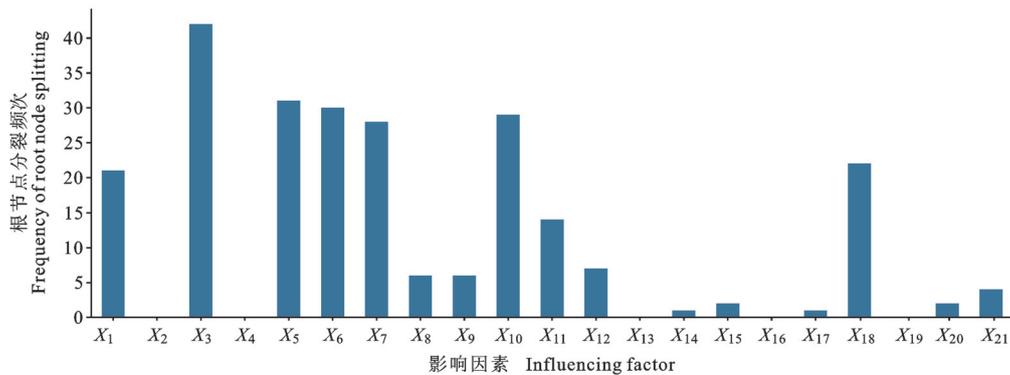


图 2 各影响因素重要性 (根节点分裂频次)

Fig. 2 Importance of each influencing factor (frequency of root node splitting)

图 3 是以节点纯度增加量为依据, 对我国制造业企业碳信息披露影响因素的重要性进行分析的结

果。通过图 3 可知: 研发投入金额 (X_3) 对企业碳信息披露质量的影响程度最大, 之后以变量节点纯

度增加比例为依据, 对变量的重要性进行排序, 依次是研发投入金额 (X_3)、总资产 (X_5)、总负债 (X_6)、无形资产 (X_7)、营业收入 (X_{10})、员工人数 (X_{18})、营业成本 (X_{11}) 以及研发人员数量 (X_1) 等因素。此外, 在创新、资产、营业、管理

以及员工五个维度中, 资产维度的因素对企业碳信息披露质量的影响程度最高, 而管理维度的因素对企业碳信息披露质量的影响程度最低。将图 2 和图 3 进行比较可知: 根据不同原则确定的制造业企业碳信息披露的各影响因素的重要性大致相同。

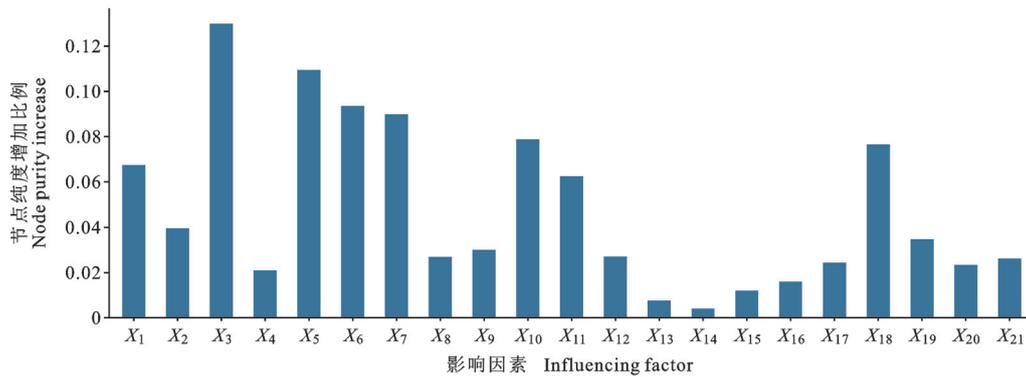


图3 各影响因素重要性(节点纯度增加量)

Fig. 3 The importance of each influencing factor (node purity increase)

为检验上述结果的稳健性, 将随机森林中决策树的数量更换为 500 棵后, 重新计算企业碳信息披露的各影响因素节点纯度增加量, 结果见图 4。从图 4 可知: 研发投入金额 (X_3) 对企业碳信

息披露的影响程度依旧最大, 而其他变量的重要性次序变动不大。表明即使更改了决策树数量, 企业碳信息披露的主要影响因素也不会变动, 证明了结果的稳定性。

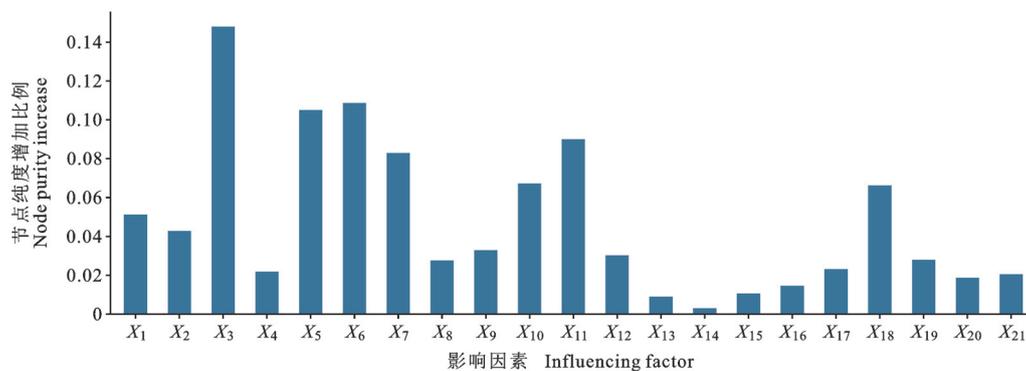


图4 改变决策树数量后变量的节点纯度增加量

Fig. 4 Increase in node purity of variables after changing the number of decision trees

3 结论及启示

3.1 研究结论

以 2018—2020 年我国 A 股上市公司中的制造业企业为研究对象, 通过构建指标体系, 采用主成分分析法和熵权法对制造业企业的碳信息披露质量进行客观评价, 同时利用随机森林回归模型分析制造业企业碳信息披露质量各影响因素的重要性后, 得出如下结论:

(1) 2018—2020 年我国制造业企业的碳信息披露质量呈现上升趋势。通过构建指标体系, 结合主成分分析法和熵权法, 对我国 2018—2020 年 A 股上市公司中 751 家制造业企业的碳信息披露质量进行评价后发现, 2018 年, 我国制造业企业碳信息披露质量的均值为 0.09006, 2019 年均值为 0.08959, 2020 年均值为 0.14244。可知在这 3 a, 我国制造业企业碳信息披露质量整体呈现上

升趋势, 2020 年企业碳信息披露质量较 2018 年和 2019 年有较大提升。

(2) 我国制造业企业的碳信息披露质量整体偏低, 不同企业间的碳信息披露质量存在较大差异。尽管我国制造业企业的碳信息披露质量呈现上升趋势, 但整体上仍处于较低水平, 以 2020 年为例, 当年的企业碳信息披露质量均值仅为 0.14244, 而同年碳信息披露质量最高的企业对应的数值为 0.9454, 可见, 我国制造业企业的碳信息披露质量整体偏低。此外, 在碳信息披露质量排名前十的企业中, 各企业的碳信息披露质量对应的数值波动较大, 不同企业间的碳信息披露质量存在较大差异。

(3) 研发投入、总资产、总负债等因素对企业的碳信息披露质量有较大影响。以制造业企业碳信息披露质量为输出变量, 以企业内部五个维度 21 个影响因素为输入变量, 利用随机森林回归模型分析各因素对企业碳信息披露质量影响的重要性后发现, 以决策树根节点分裂频次为依据和以节点纯度增加量为依据的结果均显示, 研发投入金额、总资产、总负债、无形资产、营业收入、员工人数、营业成本以及研发人员数量等因素对企业碳信息披露质量均有较大影响。

3.2 政策建议

根据上述主要结论, 提出如下政策建议:

(1) 不断完善碳信息披露的激励机制。我国制造业企业的碳信息披露仍然以自愿披露为主要途径, 缺乏相应的激励机制, 从而导致碳信息披露质量参差不齐, 各企业间的碳信息披露水平存在较大差距。值得注意的是, 生态环境部于 2021 年 12 月 31 日颁布的《企业环境信息依法披露管理办法》(https://www.mee.gov.cn/xxgk2018/xxgk/xxgk02/202112/t20211221_964837.html) 中首次提出“重点排污单位等类型的企业需披露其碳排放信息”这一要求, 毫无疑问, 我国企业的碳信息正在逐步从自愿披露迈向强制披露。通过政策法规确立的碳信息强制披露制度在后续的实践中的需不断完善, 一方面, 强制披露的企业范围仍需进一步扩大, 当前规定需强制披露的企业类型较少, 所涉及的企业数量亦是如此, 以重点排污单位为例, 在本文所研究的 751 家制造业企业中, 仅有不到 70 家企业为重点排污单位, 占比不超过 1/10。另一方面, 企业碳信息披露的内容也需要

进一步具体化, 当前要求企业披露的碳信息侧重于其碳排放量、减排设施两项, 而企业的减排计划、减排活动内容亦是促进企业减排的重要推手。在企业碳信息披露的监督方面, 企业披露的碳信息应当向全社会公开, 由政府、股东、公民等利益相关者共同审查评估, 给予碳信息披露质量较好的企业物质奖励或授予相应荣誉, 适当处罚碳信息披露质量较差的企业。

(2) 多渠道促进企业创新投入提升。在影响企业碳信息披露质量的众多因素中, 研发投入金额是最为重要的因素, 要提升企业的碳信息披露质量, 企业的研发投入是必须关注的重点。但是, 对于企业本身而言, 在创新成果尚未转化之前, 加大创新投入不会给企业带来收益, 反而会加重企业负担。因此, 要提升企业碳信息披露质量, 一方面, 政府要疏通企业创新成果转化渠道, 同时给予企业一定的研发补贴; 另一方面, 企业要根据自身情况, 通过重塑管理流程机制, 调整企业的资源配置, 提升要素生产率。

(3) 合理高效配置企业资产。在影响企业碳信息披露质量的五个维度的因素中, 企业资产维度中的总资产、总负债以及无形资产三个变量对企业碳信息披露质量均有较大影响。在全球气候治理进程不断深入的背景下, 合理配置企业资产, 不仅能促进企业碳信息披露质量提高, 更能提升企业应对未来不确定性的能力。随着我国碳交易市场的全面铺开, 企业应进一步将碳资产的管理和配置纳入决策管理流程。

参考文献

- 曹桃云. 2022. 基于随机森林的变量重要性研究 [J]. *统计与决策*, 38(4): 60–63. [Cao T Y. 2022. Study on the importance of variables based on random forest [J]. *Statistics & Decision*, 38(4): 60–63.]
- 陈远, 王超群, 胡忠义, 等. 2018. 基于主成分分析和随机森林的恶意网站评估与识别 [J]. *数据分析与知识发现*, 2(4): 71–80. [Chen Y, Wang C Q, Hu Z Y, et al. 2018. Identifying malicious websites with PCA and random forest methods [J]. *Data Analysis and Knowledge Discovery*, 2(4): 71–80.]
- 崔丹, 卜晓燕, 徐祯, 等. 2021. 中国资源型城市高质量发展综合评估及影响机理 [J]. *地理学报*, 76(10): 2489–2503. [Cui D, Bu X Y, Xu Z, et al. 2021. Comprehensive evaluation and impact mechanism of high-quality

- development of China's resource-based cities [J]. *Acta Geographica Sinica*, 76(10): 2489–2503.]
- 郭 嵘, 陆 萍. 2017. 董事会异质性与企业碳信息披露——基于造纸业上市公司的经验数据 [J]. *财会通讯*, (34): 5–9. [Guo R, Lu P. 2017. Heterogeneity of board of directors and carbon information disclosure of enterprises—based on the empirical data of listed companies in paper industry [J]. *Communication of Finance and Accounting*, (34): 5–9.]
- 黄炳艺, 雷丽娜, 陈春梅. 2023. 碳会计信息披露质量与债务资本成本——基于我国电力行业上市公司的经验证据 [J]. *数理统计与管理*, 42(4): 581–594. [Huang B Y, Lei L N, Chen C M. 2023. Quality of carbon accounting information disclosure and debt capital cost: empirical analysis based on China's electric power industry [J]. *Journal of Applied Statistics and Management*, 42(4): 581–594.]
- 李慧云, 石 晶, 李 航, 等. 2018. 公共压力、股权性质与碳信息披露 [J]. *统计与信息论坛*, 33(8): 94–100. [Li H Y, Shi J, Li H, et al. 2018. Public pressure, equity nature and carbon information disclosure [J]. *Statistics & Information Forum*, 33(8): 94–100.]
- 李 力, 刘全齐, 唐登莉. 2019. 碳绩效、碳信息披露质量与股权融资成本 [J]. *管理评论*, 31(1): 221–235. [Li L, Liu Q Q, Tang D L. 2019. Carbon performance, carbon information disclosure quality and cost of equity financing [J]. *Management Review*, 31(1): 221–235.]
- 李 航. 2019. 统计学习与方法 [M]. 北京: 清华大学出版社: 81–82. [Li H. 2019. *Statistical learning method* [M]. Beijing: Tsinghua University Press: 81–82.]
- 李世辉, 葛玉峰, 王如玉. 2019. 基于改进变权物元可拓模型的碳信息披露质量评价 [J]. *统计与决策*, 35(21): 57–61. [Li S H, Ge Y F, Wang R Y. 2019. Quality evaluation of carbon information disclosure based on improved variable weight matter-element extension model [J]. *Statistics & Decision*, 35(21): 57–61.]
- 刘捷先, 张 晨. 2020. 中国企业碳信息披露质量评价体系的构建 [J]. *系统工程学报*, 35(6): 849–864. [Liu J X, Zhang C. 2020. Construction of China's enterprise carbon information disclosure quality evaluation system [J]. *Journal of Systems Engineering*, 35(6): 849–864.]
- 刘 敏, 郎荣玲, 曹永斌. 2015. 随机森林中树的数量 [J]. *计算机工程与应用*, 51(5): 126–131. [Liu M, Lang R L, Cao Y B. 2015. Number of trees in random forest [J]. *Computer Engineering and Applications*, 51(5): 126–131.]
- 刘长奎, 边季亚. 2020. 外部治理下碳信息披露质量对企业债务成本的影响 [J]. *东华大学学报(自然科学版)*, 46(2): 321–327. [Liu C K, Bian J Y. 2020. The influence of carbon information disclosure quality on cost of debt based on external governance regulation [J]. *Journal of Donghua University (Natural Science)*, 46(2): 321–327.]
- 柳学信, 杜肖璇, 孔晓旭, 等. 2021. 碳信息披露水平、股权融资成本与企业价值 [J]. *技术经济*, 40(8): 116–125. [Liu X X, Du X X, Kong X X, et al. 2021. Carbon information disclosure level, equity financing cost and enterprise value [J]. *Journal of Technology Economics*, 40(8): 116–125.]
- 马 歆, 杨益鹏, 王希胜. 2018. 制度环境、政治关联与碳信息披露水平 [J]. *财会月刊*, (18): 37–45. [Ma X, Yang Y P, Wang X S. 2018. Institutional environment, political connection and carbon disclosure level [J]. *Finance and Accounting Monthly*, (18): 37–45.]
- 史 丹, 李 鹏. 2021. “双碳”目标下工业碳排放结构模拟与政策冲击 [J]. *改革*, (12): 30–44. [Shi D, Li P. 2021. Industrial carbon emission structure simulation and policy impact under the target of “double carbon” [J]. *Reform*, (12): 30–44.]
- 史学飞, 孙 钰, 崔 寅. 2018. 基于熵值-主成分分析法的天津市低碳经济发展水平评价 [J]. *科技管理研究*, 38(3): 247–252. [Shi X F, Sun Y, Cui Y. 2018. Evaluation of low carbon economic development level in Tianjin based on entropy principal component analysis [J]. *Science and Technology Management Research*, 38(3): 247–252.]
- 宋晓华, 蒋 潇, 韩晶晶, 等. 2019. 企业碳信息披露的价值效应研究——基于公共压力的调节作用 [J]. *会计研究*, (12): 78–84. [Song X H, Jiang X, Han J J, et al. 2019. Research on the value effect of enterprise carbon information disclosure—based on the adjustment of public pressure [J]. *Accounting Research*, (12): 78–84.]
- 唐成林. 2017. 政府规制、高管政治关联与碳信息披露 [J]. *财会通讯*, (33): 87–91. [Tang C L. 2017. Government regulation, political connection of executives and carbon information disclosure [J]. *Communication of Finance and Accounting*, (33): 87–91.]
- 唐勇军, 赵梦雪, 王秀丽, 等. 2018. 法律制度环境、注册会计师审计制度与碳信息披露 [J]. *工业技术经济*, 37(4):

- 148–155. [Tang Y J, Zhao M X, Wang X L, et al. 2018. Legal system environment, CPA auditing system and carbon information disclosure [J]. *Journal of Industrial Technological Economics*, 37(4): 148–155.]
- 童心刚. 2019. 基于 CDP 报告的企业经营绩效对碳信息披露水平影响的实证分析 [J]. *生态经济*, 35(6): 39–43. [Tong X G. 2019. Empirical analysis of the impact of business performance on carbon information disclosure level based on CDP report [J]. *Ecological Economy*, 35(6): 39–43.]
- 王 莺, 王 静, 姚玉璧, 等. 2014. 基于主成分分析的中国南方干旱脆弱性评价 [J]. *生态环境学报*, 23(12): 1897–1904. [Wang Y, Wang J, Yao Y B, et al. 2014. Evaluation of drought vulnerability in southern china based on principal component analysis [J]. *Ecology and Environmental Sciences*, 23(12): 1897–1904.]
- 王志亮, 杨 媛. 2017. 企业碳信息披露的内部动因研究 [J]. *企业经济*, 36(2): 14–21. [Wang Z L, Yang Y. 2017. Study on internal motivation of carbon information disclosure of enterprises [J]. *Enterprise Economy*, 36(2): 14–21.]
- 温雅丽, 廖 艳, 王 杰. 2019. 基于低碳农业的公司特征与碳信息披露质量研究 [J]. *农业经济*, (7): 114–116. [Wen Y L, Liao Y, Wang J. 2019. Research on company characteristics and carbon information disclosure quality based on low-carbon agriculture [J]. *Agricultural Economy*, (7): 114–116.]
- 杨 璐, 吴 杨, 唐勇军, 等. 2017. 公司治理特征与碳信息披露——基于 2012—2014 年 A 股上市公司的经验证据 [J]. *财会通讯*, (3): 20–25, 129. [Yang L, Wu Y, Tang Y J, et al. 2017. The study on the characteristics of the corporate governance and carbon disclosure [J]. *Communication of Finance and Accounting*, (3): 20–25, 129.]
- 俞立平, 刘 骏. 2018. 主成分分析与因子分析法适合科技评价吗? ——以学术期刊评价为例 [J]. *现代情报*, 38(6): 73–79, 137. [Yu L P, Liu J. 2018. Are principal component analysis and factor analysis suitable for scientific and technological evaluation? — taking academic journals as an example [J]. *Journal of Modern Information*, 38(6): 73–79, 137.]
- 张娇宁, 孙 慧, 马晓钰. 2021. 碳信息披露对企业债务融资成本的影响——基于环境规制与高管激励的双重调节效应研究 [J]. *中国注册会计师*, (12): 48–54. [Zhang J N, Sun H, Ma X Y. 2021. The impact of carbon information disclosure on corporate debt financing cost— based on the dual regulatory effect of environmental regulation and executive incentive [J]. *The Chinese Certified Public Accountant*, (12): 48–54.]
- 张 雷, 王琳琳, 张旭东, 等. 2014. 随机森林算法基本思想及其在生态学中的应用——以云南松分布模拟为例 [J]. *生态学报*, 34(3): 650–659. [Zhang L, Wang L L, Zhang X D, et al. 2014. The basic principle of random forest and its applications in ecology: a case study of *Pinus yunnanensis* [J]. *Acta Ecologica Sinica*, 34(3): 650–659.]
- 赵 宁, 石 磊, 翟凤勇, 等. 2019. 基于专利信息评价和挖掘智能机器人领域技术创新人才——主成分分析法 (PCA) 的视角 [J]. *科技管理研究*, 39(17): 160–165. [Zhao N, Shi L, Zhai F Y, et al. 2019. Evaluation and excavation of the technological innovation talents in the field of intelligent robot based on patent information: the perspective of principal component analysis (PCA) [J]. *Science and Technology Management Research*, 39(17): 160–165.]
- Breiman L. 2001. Random forests [J]. *Machine Learning*, 45(1): 5–32.
- Zhao R, Min N, Geng Y, et al. 2017. Allocation of carbon emissions among industries/sectors: an emissions intensity reduction constrained approach [J]. *Journal of Cleaner Production*, 142: 3083–3094.