

辽河流域水质监测指标间的相关性分析及主成分分析

秦小昂, 陈子璇, 朱 娇, 殷 丽, 王木新, 金美松

北部战区疾病预防控制中心卫生监督监测科, 辽宁 沈阳

收稿日期: 2025年3月10日; 录用日期: 2025年5月10日; 发布日期: 2025年6月12日

摘 要

为快速监测辽河流域水质, 有效追踪并控制水质污染情况, 本文选用辽河流域14个水质监测站2023年12期监测数据, 分析常规水质指标与核心水质指标间的相关性。使用时间变化趋势检验和主成分综合指数效价评估各断面指标变化, 确定核心指标关联度参数, 且与常规水质类别进行方差分析。结果表明, 不同检测站的核心指标与常规指标的关联度各异, 主成分1 (浊度、高锰酸盐指数、总磷) 反映悬浮物与有机污染负荷, 主成分2 (电导率、氨氮、总氮) 表征氮营养盐富集, 主成分3 (pH、溶解氧) 指示水体自净能力。流域污染呈现“上游物理性浑浊 - 中下游氮负荷突出 - 工业点源复合污染”的空间分异格局, 建议优先实施通江口截污控磷、于家房河段脱氮工程及工业区酸碱度监管, 以针对性改善水质。

关键词

主成分分析法, 皮尔逊相关系数法

Analysis of Correlations and Principal Components among Water Quality Monitoring Indicators in the Liao River Basin

Xiao'ang Qin, Zixuan Chen, Jiao Zhu, Li Yin, Muxin Wang, Meisong Jin

Health Supervision and Monitoring Section, Center for Disease Control and Prevention in the Northern Theater Command, Shenyang Liaoning

Received: Mar. 10th, 2025; accepted: May 10th, 2025; published: Jun. 12th, 2025

Abstract

To enable rapid water quality monitoring and effective pollution tracking and control in the Liao

文章引用: 秦小昂, 陈子璇, 朱娇, 殷丽, 王木新, 金美松. 辽河流域水质监测指标间的相关性分析及主成分分析[J]. 环境保护前沿, 2025, 15(6): 893-899. DOI: 10.12677/aep.2025.156100

River Basin, this study analyzed correlations between conventional and core water quality indicators using 2023 monitoring data from 14 stations across 12 monthly periods in the Liao River Basin. Temporal trend analysis and principal component comprehensive index evaluation were applied to assess cross-section parameter variations, determine core indicator correlation coefficients, and conduct ANOVA with conventional water quality categories. Results revealed station-specific variability in core-conventional indicator correlations, with Principal Component 1 (turbidity, permanganate index, total phosphorus) reflecting suspended solids and organic pollution loads, Principal Component 2 (conductivity, ammonia nitrogen, total nitrogen) characterizing nitrogen nutrient salt enrichment, and Principal Component 3 (pH, dissolved oxygen) indicating self-purification capacity of water body. Spatial differentiation displayed distinct pollution patterns: “physical turbidity dominant in upper reaches, elevated nitrogen loading in mid-lower reaches, and compound industrial pollution in point-source zones”. Targeted mitigation strategies include prioritizing phosphorus interception at Tongjiangkou, implementing nitrogen removal in Yujiafang river sections, and enhancing pH regulation in industrial areas.

Keywords

Principal Component Analysis, Pearson Correlation Coefficient

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

河流对城市发展具有重要意义，不仅为城市提供丰富的水源，而且是经济发展及生命安全的命脉，随着工业化与城市化的加快，人口出现高聚集性的时空分布，以至于生活和工业废水的排放日益增多，水质逐渐恶化，进而引发一系列环境污染和经济问题[1]。工业污染、城镇生活污染、农村面源污染、畜禽养殖污染和历史遗留污染等多种因素叠加，突发水环境污染事件成为环境应急工作、水质卫生监督管理部门工作的重点难点。辽河流域流经河北、内蒙古、吉林、辽宁四省，流域内物产丰富，地形以平原为主，同时盛产石油和天然气，是中国重要的能源产区之一。辽河全流域总人口 3383 万人，国内生产总值高达 9172 亿元，是中国重要的工业基地和商品粮基地[2]。因此，自“十一五”至“十三五”期间水专项流域研究，国家实施水体污染控制与专项，但在目前的研究报告中未有明确的断面及趋势研究。

目前国内外研究聚焦于快速高效的水质监测，依托实验室化验技术，这种方法精准但耗时耗力且无法及时追踪污染源。当下在多学科交流及遥感智能系统的更新、信息技术和大数据等领域的飞速发展，我国水质监测基础设施日益完善，水质自动监测系统的布设也日益广泛，基于当下水质自动监测系统，获得实时数据，基于水质监测指标值，建立相关数学模型，对水环境质量的好坏进行定性或者定量描述水质现状[3]-[5]。常用的水质评价方法较多，代表性方法有单因子评价法、综合污染指数法、模糊综合评判法、神经网络法、质量指数法、主成分分析法等。当下水质预测仍处于发展阶段，如何有效利用已有庞大的水质在线监测数据，实现水质精确预测，成为构建和掌握区域流域水质发展的关键[6]-[8]。

为了摸清辽河流域水质情况及综合指数，健全流域各断面水质分析，切实守住水质安全底线，本文利用 2023 年辽河流域 14 个断面水质监测数据，采取主成分分析法和皮尔逊相关系数法，得出水质评价结果，掌握水质现状、变化趋势及主要污染物种类，建立综合评价指数以期为水环境质量评分，为未来

评价体系及风险预测研究提供科学依据和理论支撑。

2. 材料与方法

2.1. 研究区域概况

辽河是中国七大河流之一，位于中国东北地区，地处北纬 39°~45°、东经 118°~125°之间，干流全长约 1390 公里，流域面积达 22.9 万平方公里，涉及辽宁、吉林、内蒙古 3 省区。辽河发源于河北省平泉县山区，由东辽河和西辽河在辽宁省昌图县三江口汇合而成，最终注入渤海。

东辽河发源于吉林省辽源市帽山，自东向西流经辽源市、四平市、铁岭市等地，于昌图县与西辽河交汇。西辽河发源于内蒙古自治区克什克腾旗的石砬沟，流经赤峰市、通辽市等地。干流从三江口汇合处开始，向南经调兵山市和辽阳市，最后在盘锦市汇入渤海。

2.2. 大数据来源

辽河流域设置 14 个水质监测断面，本研究采用 2023 年各断面水质监测数据(数据来源于国家地表水水质自动监测实时数据发布系统 <https://szzdjc.cnemc.cn:8070/GJZ/Business/Publish/Main.html>), 选取水质、水温、pH、溶解氧(mg/L)、电导率(μS/cm)、浊度(NTU)、氨氮(mg/L)、总氮(mg/L)、总磷(mg/L)、高锰酸盐指数(mg/L)等 10 种监测指标作为评价因子，各断面每月监测一次。

2.3. 相关分析

水质参数指标间存在互相影响的效果，不同区域之间的水质指标间影响程度存在差异性。本文采用 Pearson 统计相关系数来度量水质指标参数间的相关程度，其值介于-1 与 1 之间。

表达式如下：

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

公式中， n 为各断面水质监测平台月份水质样本的化验次数； x_i 为水质参数指标浓度。以上指标可视化及相关性分析均在 R 4.4.1 中完成。

主成分分析法

主成分分析法是利用降维思想，通过线性组合将多个相互关联的数值变量转化为主要影响变量。主要计算步骤是首先将原有的变量标准化，随后计算各变量之间的相关矩阵、该矩阵的特征根和特征向量，由大到小排列，分别计算出对应的主成分。确定是否进行因子分析，需要进行 KMO 统计值和 Bartlett 球形检验，当 KMO 值小于 0.6，不能进行因子分析；Bartlett 检验对应 P 值小于 0.05，则可进行主成分分析。确定影响疾病发生的主要因子，主要从两个方面确定，累计贡献率：当 k 个主成分的累计贡献率达到 70%以上时，则保留前 k 个主成分；特征根：选取特征大于 1 的主成分。

3. 结果

3.1. 辽河流域水质参数相关性整体分析

辽河流域浊度与总磷呈现显著正相关，相关系数为 0.774，高锰酸盐指数与总磷的相关系数为 0.747。水温与溶解氧呈明显负相关，相关系数为-0.514，符合热力学规律下溶解氧随温度升高而减少的特征。电导率与总氮相关性达到 0.555，高锰酸盐指数相关性达到 0.52，表明水体中离子强度与有机污染的综合表征因子存在相关性；溶解氧与总氮呈现微弱相关性，相关系数达 0.07，两者受不同环境过程主导(见图 1)。

3.2. 辽河流域水质参数趋势变化研究

辽河流域水质总体变化趋势在一月到二月，多数断面水质较好，Ⅰ类和Ⅱ类占比高，如本溪市凤鸣电站、南太子河水库入口保持Ⅰ类水质。夏季到秋季，每年的五月份至八月份，水质类别波动显著，部分断面污染加重，如铁岭市通江口水质从1月份Ⅲ类恶化至7月的Ⅴ类，铁岭市凡河一号桥5月、7月出现Ⅴ类水质，总磷、氨氮指标明显上升。九月至十一月，水质逐步恢复，但部分断面(如通江口)仍维持较高污染等级(见图2)。

3.3. 辽河流域水质参数主成分分析

选取除去水温、pH 的其他 8 项指标进行主成分分析，辽河流域的 KMO 值为 0.626，Bartlett 球形检验值均小于 0.05，流域主成分可选出 3 个主成分。流域主成分旋转载荷平方和方差百分比为 76.319%，总方差解释表展示了主成分分析(PCA)的结果。根据凯撒准则(特征值 > 1)，前三个成分被提取，其初始特征值分别为 2.852、1.924 和 1.329，共解释总方差的 76.319% (累积%)，表明这三个成分能够有效概括原始数据的主要变异。通过方差最大旋转法调整后，各成分的方差贡献趋于均衡：成分 1 的贡献从 35.65% 降至 33.211%，成分 2 从 24.052% 升至 25.676%，成分 3 从 16.617% 微增至 17.432%，但总解释方差(76.319%)保持不变。这种调整优化了成分载荷的分布，增强了潜在因素的可解释性。其余成分(4 至 8)因特征值均小于 1 且累积解释力有限未被保留(见表 1)。

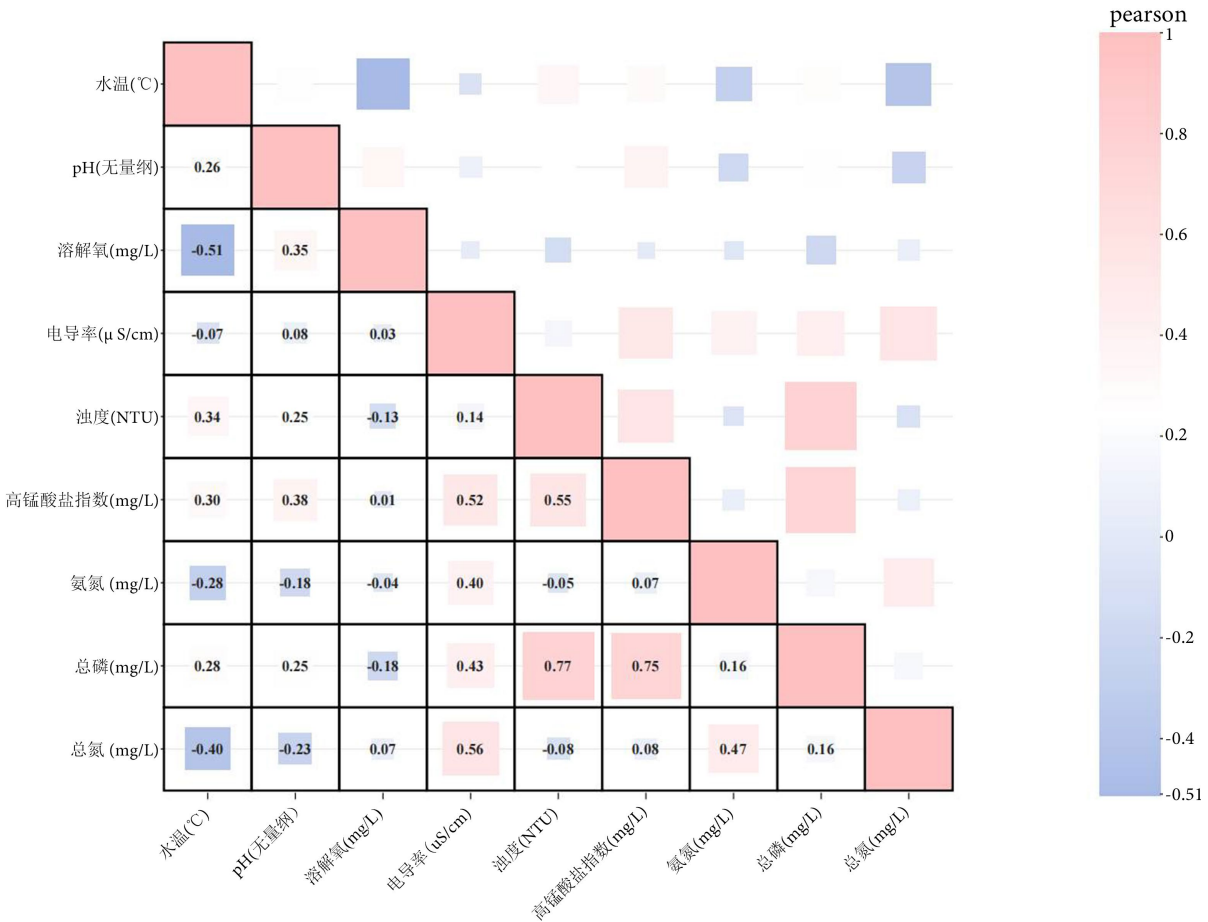


Figure 1. Heat map of correlation analysis of water quality parameters in the Liao River Basin
图 1. 辽河流域水质参数相关性分析热图

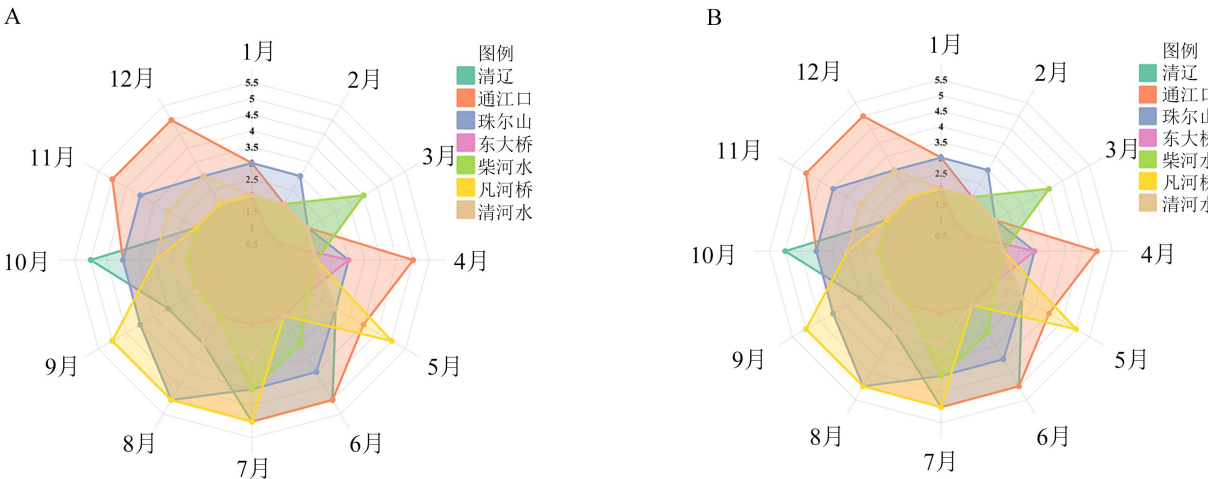


Figure 2. Radar chart of the temporal trend of water quality parameters in the Liao River Basin
图 2. 辽河流域水质参数时间趋势雷达图

Table 1. Explanation of eigenvalues and variance
表 1. 特征值及方差解释

成分	初始特征值			提取载荷平方和			旋转载荷平方和		
	总计	方差百分比	累积%	总计	方差百分比	累积%	总计	方差百分比	累积%
1	2.852	35.65	35.65	2.852	35.65	35.65	2.657	33.211	33.211
2	1.924	24.052	59.702	1.924	24.052	59.702	2.054	25.676	58.886
3	1.329	16.617	76.319	1.329	16.617	76.319	1.395	17.432	76.319

根据断面主成分分析结果(见表 2), 通江口断面综合污染指数最高, 主成分 1 (浊度、高锰酸盐指数、总磷)和主成分 2 (电导率、氨氮、总氮)均显著超标, 表明该区域存在悬浮物富集与富营养化双重污染。于家房断面主成分 2 (电导率、氨氮、总氮)达到峰值, 反映其受农业面源氮污染突出。清辽和兴国桥同样呈现显著氮污染特征。柴河水库入库口主成分 1 与总磷指标相对突出, 需防范沉积物内源释放风险。主成分 3 (pH、溶解氧)整体波动较小, 各断面溶解氧水平尚可维持水体自净能力。污染空间分异显示, 电导率、总氮等主成分 2 污染物是流域主要超标因子。

Table 2. Principal component score coefficients and comprehensive indices of water quality in each cross-section
表 2. 各断面水质主成分得分系数及综合指数

断面	主成分 1	主成分 2	主成分 3	综合指数
柴河水库入库口	22	302	14	338
东大桥	9	267	14	291
凡河一号桥	26	286	16	328
凤鸣电站	5	206	13	224
戈布桥	8	379	15	403
南太子河入库口	4	93	14	112
蒲石河大桥	6	215	14	235
清河水库入库口	27	216	13	256

续表

清辽	14	411	15	440
通江口	68	576	13	657
兴国桥	3	417	14	434
于家房	40	632	15	688
于台	6	617	11	635
珠尔山	81	485	16	582

4. 讨论

辽河流域是中国七大河流之一，流经河北省、内蒙古自治区、吉林省、辽宁省，由于近年来极端气候影响，辽河流域内洪水频繁，平均 7~8 年发生一次较大规模的洪涝灾害[9]。中下游城市，水资源短缺严重，污染现象未受到遏制。自 1993 年起开始整治辽河工程，重点环保项目，依法打击未达标排污单位，水资源得到控制。当下通过单一如溶解氧、化学需氧量、总磷、总氮和氨氮等关键水质参数衡量水质状况和处理效果，已经无法满足现实情况需要。通过主成分分析我们进行键位处理，将多个相关水质指标浓缩为少数独立的主成分，通过简化数据复杂性解释潜在污染驱动因子。主成分 1 (浊度、高锰酸盐、总磷)反映悬浮颗粒物与有机污染物的协同输入，高浊度与总磷关联表明泥沙中存在含磷载体，高锰酸盐指数则指向有机物降解耗氧。针对主成分 1 高的地区，建议实施通江口截污控磷工程，例如建设生态护坡减少泥沙冲刷，并强化农业面源磷肥管控。主成分 2 (电导率、氨氮、总氮)表征氮素营养盐富集，电导率升高暗示工业废水或生活污水排放导致离子浓度增加，氨氮与总氮的高负荷则指向化肥施用、养殖废水等非点源污染。主成分 3 (pH、溶解氧)指示水体自净能力，pH 异常波动(如酸性工业废水排放)会抑制微生物活性，而溶解氧降低(如有机物分解耗氧)直接威胁水生生物生存。

辽河流域水质存在显著的时空差异性，局部区域污染问题突出。铁岭市通江口断面 6 月高锰酸盐指数达 11.1 mg/L (超 III 类标准 3 倍)，珠尔山断面 8 月总磷浓度升至 0.473 mg/L (超 V 类标准)，主要受农业面源污染、汛期径流冲刷及工业排放叠加影响；沈阳市于台、于家房区域全年总氮浓度偏高(7.05~10.45 mg/L)，暴露出污水管网覆盖率不足、生活污水直排等问题。季节性水质恶化特征明显：夏季高温导致溶解氧骤降(如抚顺戈布桥 7 月溶解氧仅 0.23 mg/L)，加剧水体缺氧风险；冬季冰封期电导率异常升高(如沈阳于台 1 月达 1033 μ S/cm)，可能与融雪剂盐分积累相关。相比之下，本溪市通过严格水源地保护措施(如限制开发、生态修复)，凤鸣电站等断面污染物指标稳定达标，治理经验值得推广。建议针对性施策：铁岭市重点强化通江口、珠尔山周边农业面源管控(如建设生态缓冲带)及凡河流域工业排污排查；沈阳市加快于台、于家房区域污水管网改造与雨污分流工程，控制氮磷输入；全流域需建立跨区域预警机制，汛期增设移动湿地应对有机污染，冰封期前实施底泥清淤以减少内源释放，系统性改善水质生态。河流健康通过长期有效的监测，多因素发展变化得到明确监管，通过生物多样性、河流纵向连通性、水资源开发利用效率改善等措施，水质得到了明显改善[10]。通过即时监控各断面变化，建立水质预警平台，社会和公众共同参与，加强合作，制定有效的管理措施和政策，使流域得到可持续发展及生态保护的目标[11]。

本研究从整体上探究了辽河流域水质变化情况，为研究同一水质参数在不同地理区域的变化特征，在未来的研究过程中可对多点位数据进行处理，分析河流不同区域水质数据的变化特征，实现时间和空间两个维度上的环境指标评估。其次，在生态要素的综合分析过程中，降水变化、气候影响、人类活动等多方因素影响，在今后的研究分析中需要综合考虑，提升计算的精准性。

参考文献

- [1] 冯家成, 曹学康, 陶理, 等. 长江安徽段水质时空变化特征及影响因素研究[J/OL]. 长江科学院院报: 1-9. <http://kns.cnki.net/kcms/detail/42.1171.TV.20240914.1417.006.html>, 2025-05-01.
- [2] 张雯雯. 基于 PCA 分析的奉贤区地表水水质评价[J]. 陕西水利, 2025(2): 77-79.
- [3] 周泽仪, 蒋少璐, 赵义勇, 等. 滇池流域捞鱼河湿地水质评价与时空变化特征[J]. 昆明学院学报, 2024, 46(3): 68-75.
- [4] 闫淼, 陈峰. 广州市地下水水质评价方法研究——以主成分分析法和综合指数评价法为例[J]. 黑龙江环境通报, 2024, 37(9): 77-80.
- [5] 王思依, 邹顺瑛, 孙文豪, 等. 基于水质标识指数与主成分分析法的复州河流域水质质量评价研究[J]. 干旱环境监测, 2024, 38(4): 172-178.
- [6] 从俊强, 付祥元, 徐李军. 基于相关性主成分分析的铸坯质量缺陷成因追溯方法[J]. 冶金自动化, 2024, 48(6): 48-56.
- [7] 张贺玉, 杨莉园, 卢少勇, 等. 基于主成分分析的城市河流水质时空分布特征研究——以沧州市为例[J]. 环境工程技术学报, 2024, 14(4): 1273-1283.
- [8] 昌红, 刘达, 肖耿锋, 等. 广东省某大型水库水质时空变化特征综合评价[J]. 广东水利电力职业技术学院学报, 2025, 23(1): 21-27.
- [9] 陈军, 胡高宇, 陈飞, 等. 温岭隘顽湾海水养殖区池塘水质状况调查与评价[J]. 浙江海洋大学学报(自然科学版), 2024, 43(3): 260-267.
- [10] 江泓, 杨红, 葛黄群. 黄冈市遗爱湖水质综合评价及富营养化分析[J]. 水利技术监督, 2025(3): 135-137+201+215.
- [11] 严超群, 牛子航, 贾慧如, 等. 基于主成分分析的株洲市典型城市水体水质影响因子研究[J]. 广东蚕业, 2025, 59(1): 62-67.