J. Lake Sci.(湖泊科学), 2027, 39(4):000-000
DOI 10.18307/2027.0423
©2027 *by Journal of Lake Sciences*

基于卫星观测资料初步分析太湖蓝藻水华与太阳黑子活动的相关性*

李旭文[1,2]，姜　晟[1,2]，王甜甜[1,2]，张　悦[1,2]，李京航[3]，颜　瑾[1,2]，于方圆[1,2]，郭金金[1,2]，徐晓悦[1,2]，周钰菊[1,2]

（1: 江苏省环境监测中心，南京 210019）
（2: 江苏省生态环境保护遥感监测重点实验室，南京 210019）
（3: 南京信息工程大学遥感与测绘工程学院，南京 210044）

摘　要：利用开源 Prophet 模型及天文学数据分析常用的 Lomb-Scargle 周期图算法，对 2000－2025 年间长达 26a、以 MODIS 资料为主的太湖蓝藻水华面积遥感监测数据进行周期检测和分析，识别到该时间序列存在~11a 和~1a 周期且具有明确的信号显著性。结果表明，在太阳黑子活动高年，蓝藻水华发生面积、频次、物候均偏小、偏少、偏迟，太阳黑子低年蓝藻水华面积、频次、物候均偏大、偏多、偏早，存在较明显的消长现象,具有明显的“同频反相共振”关联性。因此，太阳黑子数是非常重要的蓝藻水华指示因子，可考虑纳入太湖蓝藻水华预测预报模型构建。对水华面积、主要气象因子、太阳黑子数等响应与特征变量开展皮尔逊相关系数、偏相关分析，结果表明气象因子与太阳黑子无共变周期性。多元线性回归、XGBOOST 和随机森林等 3 种机器学习模型的分析结果均显示蓝藻水华面积与太阳黑子数存在负向关联。然而，现有研究深度尚未明确太阳活动影响太湖蓝藻水华面积的机理作用与传导路径。部分推测认为，太阳黑子暴发会扰动地球大气层，改变大气环流进而引发气候波动变化，叠加地磁波动因素，最终对太湖蓝藻水华规模产生较强抑制效应。太湖蓝藻水华与太阳黑子活动的相关性作为有待厘清的科学问题，仍需开展全面系统的观测与深入研究。

关键词：卫星遥感；太湖；蓝藻水华；太阳黑子；周期性；Prophet 模型；Lomb-Scargle 周期图；偏相关分析；回归模型

Preliminary analysis of the correlation between cyanobacterial bloom in Lake Taihu and sunspots activity based on satellite remote sensing data*

Li Xuwen[1,2**],Jiang Sheng[1,2],Wang Tiantian[1,2],Zhang Yue[1,2],Li Jinghang[3],Yan Jin[1,2],Yu Fangyuan[1,2],Guo Jinjin[1,2], Xu Xiaoyue[1,2] & Zhou Yuju[1,2]

(1: *Jiangsu Provincial Environmental Monitoring Center,Nanjing* 210019, *P.R.China*)
(2: *Jiangsu Provincial Key Laboratory of Remote Sensing Monitoring for Ecological and Environmental Protection, Nanjing* 210019, *P.R.China*)
(3: *School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing* 210044, *P.R.China*)

Abstract: This study uses the open-source Prophet model and the Lomb-Scargle periodogram algorithm to identify and analyze periodic variations in a 26-year time series (2000–2025) of cyanobacterial bloom areas in Lake Taihu. The time series data mainly come from MODIS remote sensing data. The analysis identifies statistically significant periodic cycles of approximately 11 years and 1 year in the time series. The results show clear interannual variations in cyanobacterial blooms. In solar maximum years, cyanobacterial blooms have smaller coverage areas, lower occurrence frequencies and delayed phenological periods. In solar minimum years, blooms have larger coverage areas, higher occurrence frequencies and advanced phenological periods. This study confirms a significant correlation with consistent frequency and opposite phase. Sunspot number can serve as a key indicator of cyanobacterial bloom dynamics and can

* 2026-05-02 收稿；2026-08-22 收修改稿；2026-08-24 录用。
** 通信作者；E-mail: lxw241@163.com

be included in bloom prediction models for Lake Taihu. Pearson correlation and partial correlation analyses of bloom area, major meteorological factors and sunspot number show no consistent periodic covariation between meteorological factors and sunspot activity. Three machine learning models, including multiple linear regression, XGBoost and random forest, all confirm a negative relationship between cyanobacterial bloom area and sunspot number. Existing studies have not clarified the specific mechanisms and transmission pathways through which solar activity affects cyanobacterial bloom area in Lake Taihu. Preliminary inferences indicate that sunspot activity disturbs atmospheric conditions and circulation, triggers climate fluctuations, and interacts with geomagnetic variations. These combined effects strongly suppress the scale of cyanobacterial blooms in Lake Taihu. The linkage between cyanobacterial bloom dynamics and sunspot activity in Lake Taihu requires further systematic observation and in-depth scientific investigation.
Keywords: Satellite remote sensing; Lake Taihu; Cyanobacterial blooms; Sunspots; Periodicity; Prophet model; Lomb-Scargle periodogram; partial correlation analysis; regression model

湖泊蓝藻水华的成因机制在生态学研究领域一直受到关注，暴发面积的监测预测是湖泊治理和保护的重要依据。近年来，研究学者围绕太湖蓝藻水华的主要驱动因素开展了大量研究。张运林等[1]通过气温、水文、风速、水位和透明度指标分析了 1980－2020 年太湖剧烈的湖泊物理环境变化特征，指出气候变暖和城市化扩张导致湖泊增温、风速下降，有利于蓝藻水华暴发。邹伟等[2]阐述了湖泊藻类的涵盖氮、磷、光照、温度、水动力及食物网调控的多重限制机制及其耦合关系。秦伯强等[3]研究了 2000 年以来太湖蓝藻水华的生消变化及其效应，发现 2017 年出现历史最严重的蓝藻水华，高强度持续至 2020 年，2021 年起强度逐渐减弱，至 2024 年蓝藻水华大幅减少，气候变化尤其是暖冬因素是 2015－2020 年间蓝藻水华反弹的主要原因。杨芷萱等[4]分析了 2003－2023 年太湖蓝藻水华面积总体呈现周期性波动，水文气象因素和营养盐共同影响蓝藻水华的发生。严照江等[5]对浮游藻类物候的遥感监测研究进展开展了梳理分析，从气候要素及其变化、城市化与人类活动维度系统阐述了上述因子对浮游藻类物候的影响，并得出浮游藻类物候与水体表面温度、光照、溶解态无机氮和磷酸盐浓度呈正相关关系，与风速、降水、混合层深度呈负相关关系的结论。朱广伟等[6]分析了太湖水质达Ⅲ类条件下藻情与水质变化特征，分析了总磷浓度与藻情、气象水文条件及其它营养盐之间的协同变化特征，发现 2024 年总磷浓度低值的出现与 2022 年以来蓝藻水华强度下降、外源氮磷入湖负荷降低密切相关，2024 年秋冬季偏暖使得藻情反弹。

当前既有的太湖蓝藻水华研究多聚焦于环境因子及短临尺度－亚年尺度－气候尺度（3 到 5a）时间周期。如富营养化程度与气象水文条件（温度、风、光照、流场）在 1d、1week 等短临尺度或较短周期（<1a）内的季节性协同作用。少量研究与厄尔尼诺、拉尼娜等气候波动时间尺度有关联，着重从两个方面研究了太湖蓝藻水华的影响（驱动）因素：一是侧重分析湖体环境（生境、水文）等物质基础条件，主要是水温、水质、底泥，营养盐、水位、水生植被等指标；二是时间尺度上侧重分析外部气象气候条件影响因素。总体而言，基于日地关系的更长周期时间尺度的影响研究只有少量开展，如李未等[7]利用小波变换方法分析了 Chl.*a*、水温、TP、TN 四种因子不同时间尺度上的时间格局特征，识别出在年际和年代际尺度上都具有清晰的多时间尺度特征，采用时域-频域联合分析的小波变换方法，对 1992－2010 年近 20 a 太湖梅梁湾 Chl.*a*、水温、TP 和 TN 浓度时间序列的变化特征和规律进行分析，明确各因子的周期特征和突变点。

太阳黑子活动是调控地球气候周期性变化的关键长周期因子之一，1843 年德国天文学家施瓦贝（Schwabe）首次发现太阳活动周期平均约为 11a，这是太阳活动中最早被确认的周期。该活动可通过日-地系统耦合和响应的传递链，对地球系统的动力学过程及湖泊等生态系统均产生影响[8-9]，Hu 等[10]首次对太阳黑子数与澳大利亚昆士兰州 Fred Haigh 水坝蓝藻水华关联性开展研究，研究结果显示太阳黑子数与蓝藻细胞计数存在正相关关系。但太湖蓝藻水华的已有研究中，未见将太阳黑子活动周期作为蓝藻水华发生驱动因子或解释变量。

多年来，江苏省环境监测中心依据《水华遥感与地面监测评价技术规范（试行）》（HJ-1098－2020）[11]，基于 MODIS 等卫星传感器获取的逐日遥感影像数据开展太湖蓝藻水华监测，形成了 2000－2025 年间共 26a 的长时间序列蓝藻水华面积数据集。初步分析表明，在 2006－2008 年、2017－2020 年两个时

段有较高频次的大面积蓝藻水华发生，而 2000－2005 年、2010－2015 年、2022－2025 年三个时段大面积蓝藻水华的相对发生频次较少，隐约存在多年尺度周期特征，与太阳黑子活动周期及时间节点有一定的吻合性。为此，本文利用天文学领域常用的 Lomb-Scargle 周期图分析技术[12-13]，探究太湖蓝藻水华面积是否存在~11a 周期，并进一步与 2000－2025 年间第 23、24、25 太阳活动周进行对比，分析是否存在谐振响应关系。

1 数据和研究方法

1.1 卫星影像及太阳黑子数据

太湖蓝藻水华可回溯至上世纪 70 年代，1977 年 7 月 1 日的 Landsat-2 MSS 影像显示，太湖西部沿岸、湖心区、梅梁湖、竺山湖等多处水域均出现蓝藻水华（图 1），面积约 210 km^2，这是公开卫星遥感数据记录到的最早太湖蓝藻水华。20 世纪 80、90 年代的 Landsat-5 TM 数据表明，太湖蓝藻水华呈现出频发趋势。但由于卫星重访周期较长（16d）以及阴雨天气云覆盖等因素，该时期可用卫星影像稀少，有效观测频次不足，对太湖蓝藻水华发生规律性研究造成了较大影响。1999 年和 2002 年，EOS/Terra 和 Aqua 卫星相继发射，具备每日重访太湖的能力，可提供每日 1~2 次重访观测的丰富数据资源，为实现太湖蓝藻水华的日连续性解译提供了关键性支撑。本文以 MODIS 数据的解译结果为主，辅以 FY-3F MERSI 数据，构建 2000－2025 年的以日为时间单元的太湖蓝藻监测卫星影像序列。

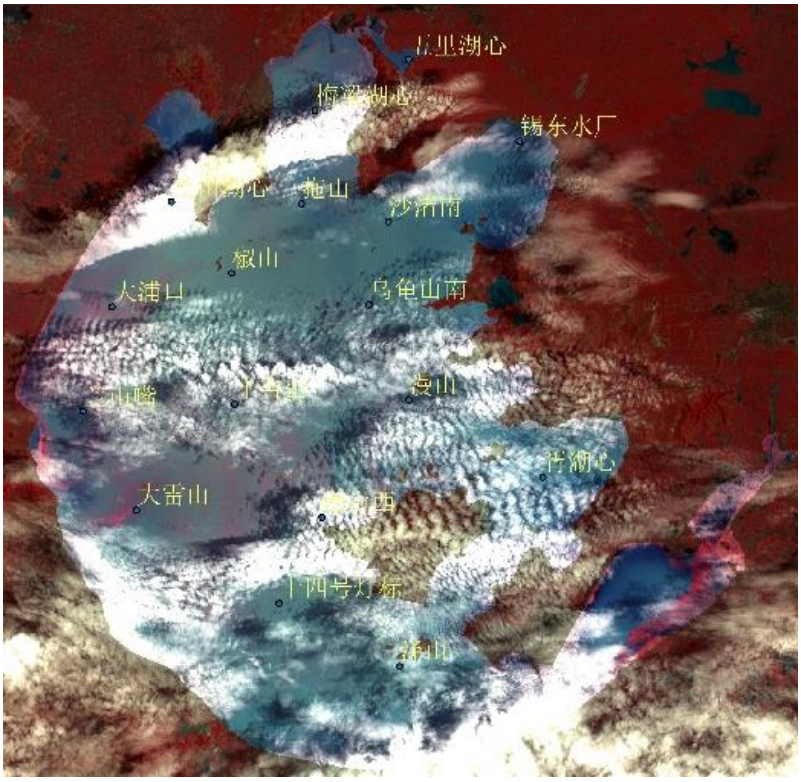

图 1 太湖蓝藻水华最早的卫星遥感记录（Landsat-2 MSS 1977 年 7 月 1 日，面积约 210km^2）

Fig.1 The first satellite remote sensing observation of cyanobacterial blooms in Lake Taihu dates back to July 1, 1977 (Landsat-2 MSS), covering an area of about 210 km²

本文采用的国际标准太阳黑子数据来自比利时皇家天文台太阳影响数据分析中心（Solar Influences Data Analysis Center, SIDC）维护的太阳黑子指数与长期太阳观测项目（Sunspot Index and Long-term Solar Observations, SISLO），SIDC 是全球最权威的太阳黑子数据汇总和审定机构，自 1818 年开始记录太阳

黑子数，形成汇编数据集并保持每日滚动更新，具有日观测、月均值、年均值三个时间尺度的太阳黑子数时间序列[14],本文选取与蓝藻水华日观测结果形成对应时间序列的 V2.0 版日观测值开展分析。

1.2 数据处理方法

本文的数据处理流程和采用的模型与算法见图 2。

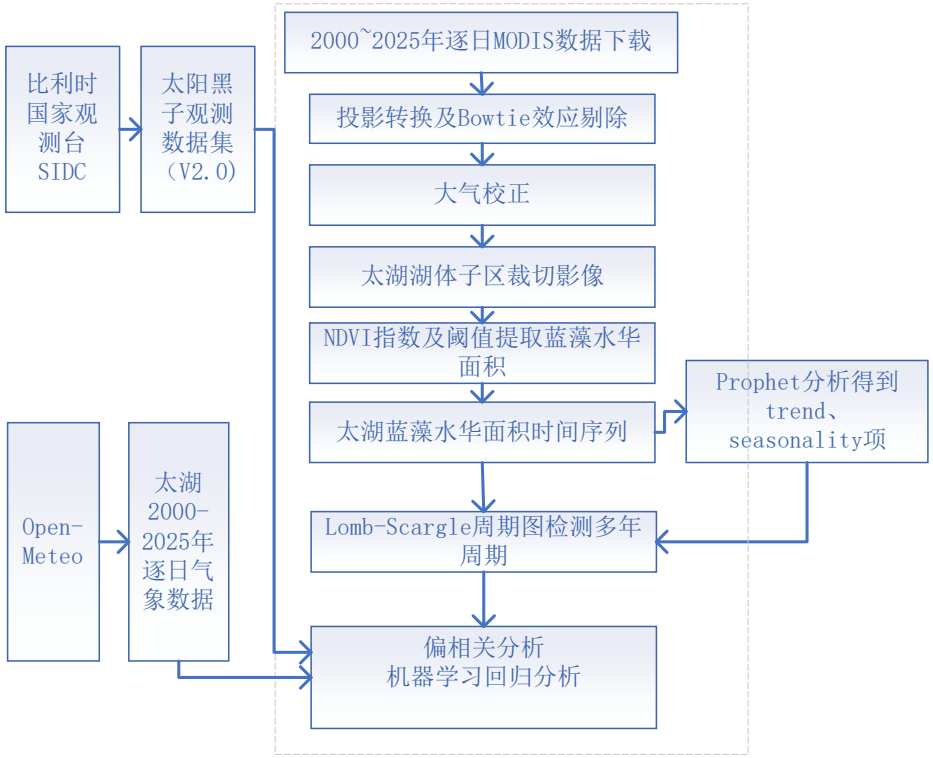

图 2 卫星遥感解译太湖蓝藻水华面积及与太阳黑子活动周期关系数据处理流程

Fig.2 Data processing workflow for satellite remote sensing interpretation of cyanobacterial bloom areas in Lake Taihu and its relationship with sunspot activity cycles

1.2.1 Prophet 模型 Prophet 模型为 Meta（原 Facebook）公司推出的开源数据智能分析 Python 包[15]，针对业务场景中普遍存在的时间上不规则观测数据处理而设计（如湖体有云覆盖无法监测蓝藻水华而缺失数据、卫星传感器故障或接收系统故障而未能获得有效影像）。相较于 ARIMA 等经典时间序列分析模型，其核心优势在于无需把时间上不规则观测数据预处理或重采样成等间隔序列，而是可以方便直观地从原始数据序列中解析出相应的组分（贡献项），特别是在分析长期趋势（如气候变化、生态过程、污染治理成效）和揭示季节性模式时（如每年 5－6 月蓝藻水华强度高峰）。

Prophet 模型有加性（additive）和乘性（multiplicative）两种算法，缺省使用加性模型，见下式：

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t) \quad (1)$$

其中，$g(t)$代表趋势性贡献（组分）项（trend），可基于分段线性模型建模，只依赖于实际的观测时刻（即时间戳）,去除所有季节性和人为活动如节假日效应后的“长期趋势”（包含潜在的多年变化）。但 Prophet 关注的是提取 yearly_seasonality 等短周期，时间序列中的长周期信号（如 5a、11a 周期）没有得到解析。虽然$g(t)$项可能蕴含了>1a 的周期信号，但 Prophet 没有提供>1a 的周期检测分析功能。$s(t)$代表季节性模式的贡献（组分）项（seasonality），可通过傅里叶级数在连续时间域上建模，不要求观测值逐日完整覆盖。Prophet 通过$s(t)$项来分析年内的周期性，包括小时/日/周/月/季/半年。$h(t)$代表人为活动带来的极短期性贡献项（holiday），如节假日、体育赛事等，对于研究大尺度生态系统现象和过程，该

项基本可以忽略。但对于分析年内特定时段如引江济太、梅雨时段，可以将其模拟为$h(t)$项以评估预测其带来的驱动影响。$\varepsilon(t)$为随机噪声误差项。

1.2.2 Lomb-Scargle 周期图　Prophet 模型获得的$g(t)$项，形成了一个新的以平滑、低频主导的数据序列，用于进一步分析其是否存在长时间、多年尺度（>1a）的周期波动规律。

快速傅里叶变换（FFT）是信号处理中常用来检测时间序列是否存在显著周期的方法，但 FFT 要求原始观测结果为时间等间隔采样、无缺失值的均匀时间序列。生态学、天文学领域的观测参数如太湖蓝藻面积、恒星亮度等通常受到观测及实验条件的制约，因此有水华面积记录的日期难以做到均匀连续。如使用 FFT 提取长时间周期就会受到严重的频谱“堆叠”（aliasing）影响，带来偏畸，引发结果偏差。

Lomb-Scargle 周期图（Periodogram）[16]是专为处理不规则采样或含缺失值的时间序列而设计，适用于分析时间非等间距观测的频谱分析方法。目前已在天文学[17-18]、生态学等领域得到了广泛应用[19]，尤其适合处理像卫星观测、野外监测、历史记录等缺失数据或采样不均匀的时间序列。该方法由 Lomb（1976）提出，Scargle（1982）完善，其数学形式已被证明等价于对非均匀数据的傅里叶展开（VanderPlas，2015,2018）。天文学领域数据处理常用的开源 Python 工具包 AstroPy[20]提供 Lomb-Scargle 周期图分析功能，本文利用 AstroPy 对太湖蓝藻水华面积时间序列中的长周期信号进行特征提取。Lomb-Scargle 周期图算法如下[21]：

本文将蓝藻水华面积时间序列记为$\{(t_i, y_i)\}_{i=1}^{N}$，每个观测点由一对数据组成,t_i为第 i 个观测的日期，y_i为日期t_i处的蓝藻水华面积，时间序列共包含 N 个观测点。

对一定频率范围[f_{min}，f_{max}]的频率 f，通过最小二乘法拟合由正弦和余弦函数组成的“正弦曲线模型”（sinusoidal model）：

$$y(t) = y_0 + Acos[2\pi f(t-\tau)] + B\,sin[2\pi f(t-\tau)] \tag{2}$$

其中，τ 为用来保证基函数正交性的时间偏移量，由下式计算：

$$\tau = \frac{1}{4\pi f} tan^{-1}\left(\frac{\sum_{i=1}^{N} sin(4\pi f t_i)}{\sum_{i=1}^{N} cos(4\pi f t_i)}\right) \tag{3}$$

计算 Lomb-Scargle 功率谱$P_{LS}(f)$，定义为频率 f 时正弦曲线模型解释方差的度量：

$$P_{LS}(f) = \frac{1}{2}\left\{\frac{\left(\sum_{i=1}^{N} y_i\, cos(2\pi f[t_i-\tau])\right)^2}{\sum_{i=1}^{N} cos^2(2\pi f[t_i-\tau])} + \frac{\left(\sum_{i=1}^{N} y_i\, sin(2\pi f[t_i-\tau])\right)^2}{\sum_{i=1}^{N} sin^2(2\pi f[t_i-\tau])}\right\} \tag{4}$$

AstroPy 中功率谱计算有 4 种模式，最常用的是 standard 模式，即采用 Scargle 归一化形式，其值介于 0 至 1 之间：

$$P_{norm}(f) = \frac{\left(\chi_0^2 - \chi^2(f)\right)}{\chi_0^2} \tag{5}$$

其中，$\chi_0^2 = \sum_{i=1}^{N}(y_i - \overline{y})^2$为“零模型”（null model）的残差平方和，即用均值$\overline{y}$拟合数据的误差；$\chi_0^2(f) = \sum_{i=1}^{N}[y_i - \hat{y}(t_i, f)]^2$为正弦曲线模型在频率 f 下的残差平方和，$\hat{y}(t_i, f)$为$t_i$时刻处利用式（2）正弦曲线模型的拟合预测值。

$P_{norm}(f)$越大，表示该频率 f 下的周期性越显著。在[f_{min}，f_{max}]频率范围搜索功率谱峰值对应的频率后，为评估检测到的峰值是否具有统计意义，需进一步计算 FAP（False Alarm Probability, 假警报概率），当 FAP<0.05 时，视为该频率对应的周期（$\frac{1}{f}$）存在显著性。本文使用了 Baluev 法[22]检验太湖蓝藻水华面积时间序列的功率谱峰，以确保识别的周期（或者说频率）具有显著性。

在频率范围设定上，由于太湖蓝藻面积观测数据的时间跨度已达 26a，考虑到蓝藻暴发可能受气候振荡变化（如 ENSO，2~7a 周期）、太阳活动（~11a 周期）、流域治理政策、规划与措施（~5a 周期）等自然或人为因素的多时间尺度过程影响，同时排除小于 1a 的日/周/季小尺度周期干扰，本文聚焦于周

期上下限为$^{1}/_{12}\,a$（即 1 个月）至 20a 的区间，时间 t 统一转换为“年”单位，频率（单位为 a^{-1}）网格在频率区间[1/20,12]采用线性间距扫描，间距点数设为 10000。根据 Nyquist 采样定律，均匀间隔的时间序列可有效检测的周期不宜超过数据的时间跨度的 1/2，我们解译得到了的太湖蓝藻水华面积数据序列跨度为 26a，考虑到该序列为以日为单位的准均匀间距，因此可以支持最高达~13 a 的多年周期检测。

2 结果

2.1 2000－2025 各年度水华面积

图 3 为 2000－2025 年各年度的逐日遥感监测蓝藻水华面积时间序列图。由图可见，2005 年以前蓝藻水华年发生面积、频次与物候分别呈现面积偏小、频次偏低、物候期偏迟的特征；2006－2008 年蓝藻水华整体表现为发生面积偏大、频次偏多、物候偏早；2010－2015 年蓝藻水华回到发生面积偏小、频次偏低、物候期偏迟的状态；2017－2020 年间蓝藻水华再次呈现发生面积偏大、频次偏多、物候偏早；2022－2025 年蓝藻水华又转为发生面积偏小、频次偏少、物候偏迟。此外，2008 年与 2025 年 1 月份连续发生小面积水华，呈现出上一年 12 月较频发水华的“拖尾”效应，2008 年的春季蓝藻水华首次暴发时间为 4 月中旬，2025 年则在 4 月底。

2.2 Prophet 模型解析结果

Prophet 模型解析的太湖蓝藻水华面积的多年趋势显示（图 4a），2000-2008 年水华面积呈增加态势，2007－2008 年段处于高位；2009－2013 年呈下降态势，2012－2014 年段处于低谷；2014－2019 年又呈增加态势，2017－2019 年段处于次高位，相较于 2007－2008 年的峰值水平，水华面积约下降 20 km^2 左右，2020 年以来呈显著下降趋势，2025 年水华强度回落至 2003 年的水平。

Prophet 模型提取的太湖蓝藻水华面积年内季节性波动特征（图 4b）显示，蓝藻水华面积自 4 月中下旬起呈现增长趋势，5－6 月上旬为年内大面积峰值时段（初夏峰期），随后的梅雨时段（6 月中旬至 7 月上旬）有所回落，7 月至 8 月中旬呈现小幅波动，8 月下旬至 10 月初时段形成秋季峰期，值得注意的是 11 月中旬至 12 月上旬也出现冬季峰期。

2.3 Lomb-Scargle 周期图

分别对水华面积原始数据序列、Prophet 模型分解得到的趋势项序列进行 Lomb-Scargle 分析，以识别太湖蓝藻水华面积是否存在更长周期性，所得周期结果汇总于表 1，采用 Baluev 法 FAP 检验太湖蓝藻水华面积时间序列的功率谱峰的显著性。

基于原始数据序列的P_{norm}最高谱峰为 0.05337，对应 10.61a 的周期（图 5a），该峰高较弱，但FAP_{Baluev}为 2.22275×10^{-27}（<0.05），检验判断 10.61a 周期性为“显著”。第二高谱峰强为 0.02106，对应 1.09a 的周期，接近于 1a，FAP_{Baluev}均小于 0.05，该~1a 的周期性为“显著”。

基于 Prophet 模型趋势项的P_{norm}为 0.6953，对应该功率谱峰的周期为 11.33a（图 5b），谱峰很强，FAP_{Baluev}为 0（<0.05），检验判断 11.33a 周期性为“显著”。第二、第三高谱峰强为 0.1056、0.1032，分别对应 1.10a、0.92a 的周期，均接近于 1a，FAP_{Baluev}均小于 0.05，该~1a 的周期性为“显著”。

表 1　太湖蓝藻水华面积时间序列的 Lomb-Scargle 周期图结果

Tab.1 Lomb-Scargle periodogram results for the time series of cyanobacterial bloom area in Lake Taihu

时间序列	周期（a）	归一化功率谱峰强	FAP (Baluev) 检验	周期性判断
原始数据	10.61	0.05337	2.26×10^{-28}（<0.05）	显著
	1.09	0.02106	9.24×10^{-9}（<0.05）	显著
Prophet 模型趋势项	11.33	0.6953	0（<0.05）	显著
	1.10	0.1056	9.49×10^{-62}（<0.05）	显著
	0.92	0.1032	3.65×10^{-60}（<0.05）	显著

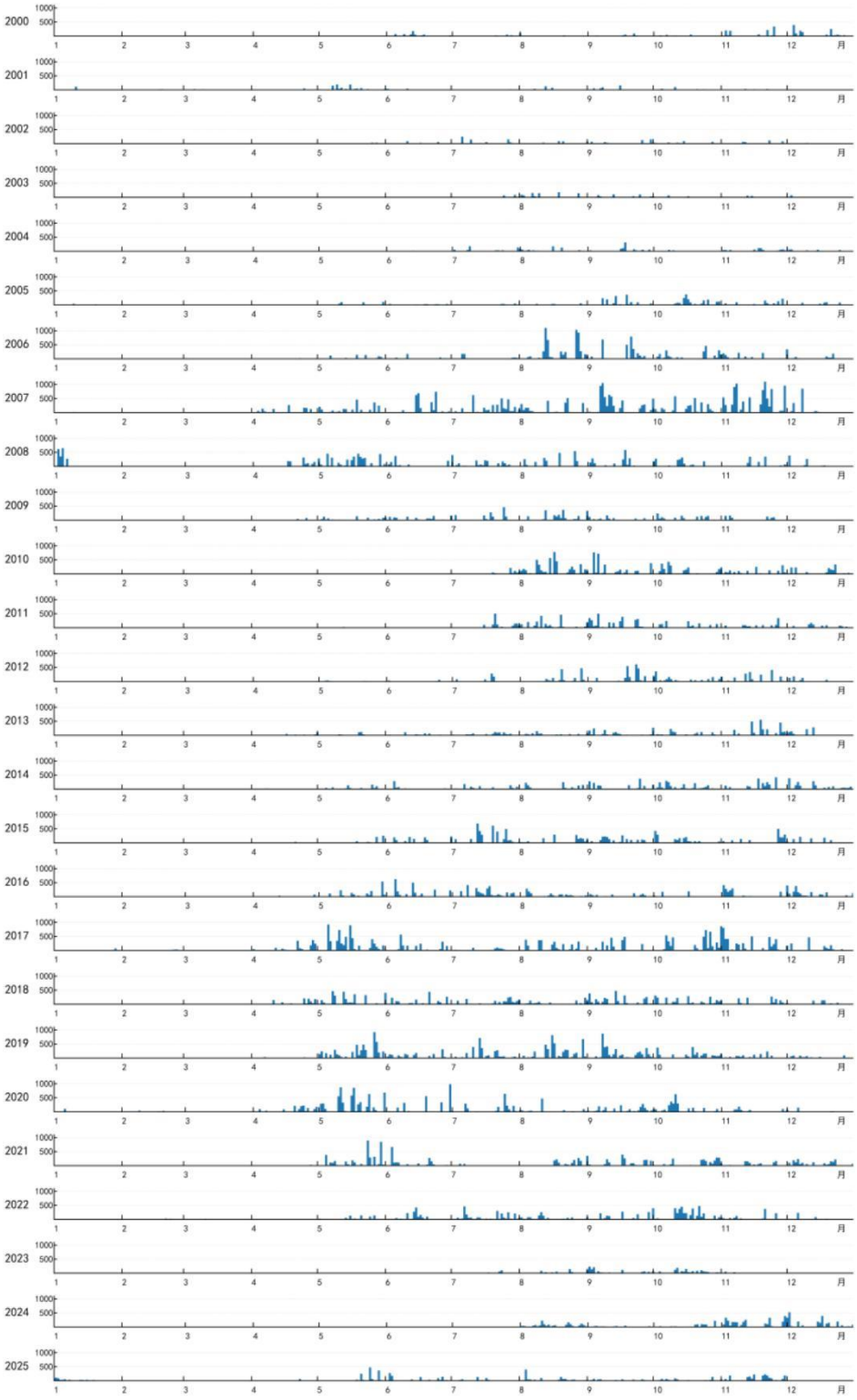

图 3 2000－2025 年逐年遥感监测蓝藻水华面积

Fig. Daily cyanobacterial bloom areas monitored via remote sensing,2000-2025

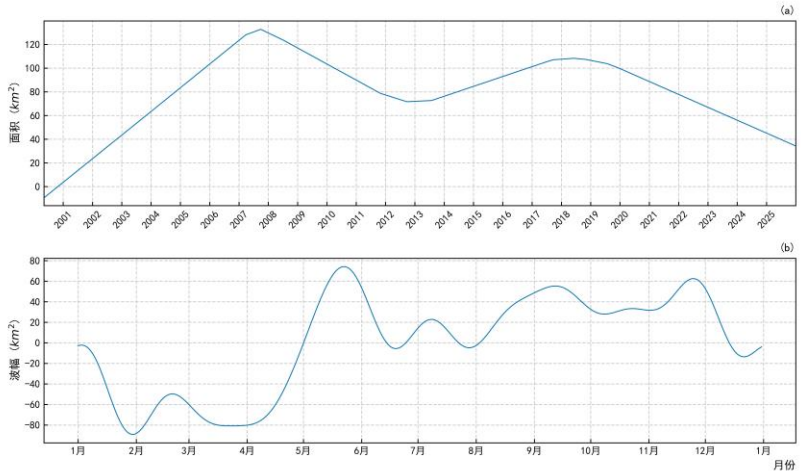

图 4 Prophet 模型解析的太湖蓝藻水华面积多年趋势（a）、年内季节性波动特征（b）

Fig.4 Multi-year trend (a) and intra-annual seasonal fluctuation characteristics (b) of cyanobacterial bloom area in Lake Taihu fitted by the Prophet model

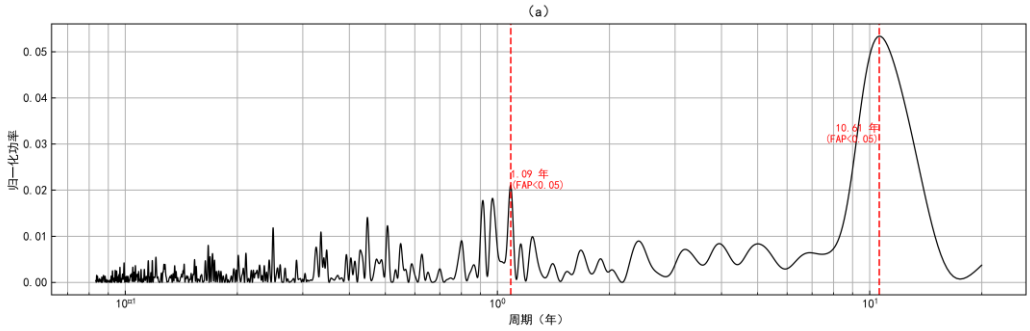

图 5 Lomb-Scargle 周期图及功率峰值对应的周期，分别基于原始数据(a)、Prophet 模型的趋势项(b)得到

Fig.5 Lomb-Scargle periodograms and the periods corresponding to power peaks, derived from the original time series (a), and the trend component of the Prophet model (b), respectively

3 分析与讨论

无论是基于原始数据还是趋势项的时间序列，都能检测到存在~11a 和~1a 两个周期，且两者的 FAP 均<0.05，周期性显著。~11a 的周期与太阳黑子活动周期高度吻合，表明太阳黑子与太湖蓝藻水华规模性暴发存在较为明显的周期一致性，太阳黑子可能是通过日—地系统耦合作用而影响太湖蓝藻发生的不可忽略的长周期性驱动因素。~1a 的周期的形成可归因于蓝藻春季复苏、夏秋频发、越冬休眠的年度物候波动性特征。

本文基于跨度达 26a 的长时间序列太湖蓝藻水华面积卫星遥感监测数据，结合 Prophet 模型与 Lomb-Scargle 周期图提取到的长周期信息，开展蓝藻水华与太阳黑子活动观测数据的关联分析，结果表明蓝藻水华面积与太阳黑子活动存在较明显的负相关现象，如图 6 所示，第 23、24 周期交接（以 2008 年 1 月为界）过渡的 2006－2008 年太阳黑子活动强度较低，大面积水华较多；第 24、25 周期交接时段（以 2019 年 1 月为界）,即 2017－2020 年，同样呈现大面积水华频发的特征；两个交接时段之间的太阳黑子活动高值年，大面积水华的发生频次相对较低。

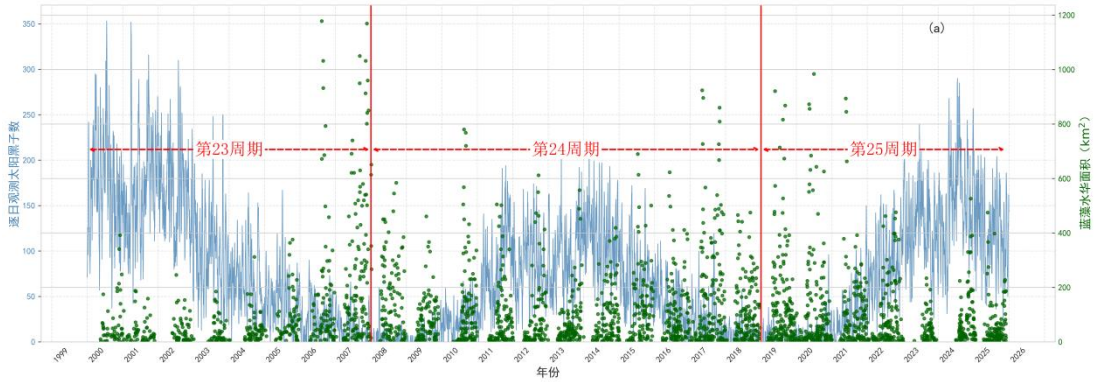

图 6 2000－2025 年太阳黑子日观测数与太湖蓝藻日监测结果的对比

Fig.6 Comparison between daily sunspot observations and daily remote sensing monitored cyanobacterial blooms areas in Lake Taihu for years 2000 to 2025

此外，第 24 周期的太阳黑子强度比第 23、25 周期要弱，因此 2011－2015 年间太湖仍多次出现一定规模的蓝藻水华。第 25 太阳活动周的 2024 年上半年太阳黑子活动持续偏强，黑子数量超过 160，与 2002 年的黑子水平相当，这可能是 2024 年卫星遥感首次监测到蓝藻水华日期最迟（7 月 3 日）的原因之一。

表 2 为 2000－2025 年间太湖蓝藻水华面积按 100 km^2 分箱，以当日及 4 日（前 3 天+当日）时间窗口滑动，得到太阳黑子数均值及标准差统计结果。结果显示，蓝藻水华暴发超过 700 km^2 时太阳黑子数整体处于极低水平，均值约为 13；蓝藻水华暴发 500~600 km^2 时，太阳黑子数仍处于较低水平，均值不超过 30。部分大面积蓝藻水华暴发过程对应太阳黑子数甚至连续多日为 0，表明太阳黑子暴发可能对蓝藻生长造成重要胁迫，甚至抑制蓝藻水华的暴发。图 7 散点图显示，水华面积较小时，太阳黑子数存在较宽的动态范围；但随着水华面积增大，太阳黑子数处于较低水平，表明太阳黑子数可“指示”蓝藻水华面积的上限规模。通过线性模式计算得到一条将 99.9%的散点落在下侧的包络线（图 7），得到蓝藻水华面积上限 y（km^2）与太阳黑子数 x 的经验计算式如下：

表 2　太湖蓝藻水华面积分箱统计太阳黑子数均值及标准差

Tab.2 Sunspot number means and standard deviations within bins of cyanobacterial bloom area in Lake Taihu

水华面积分箱（km^2）	水华当日太阳黑子数均值	水华当日太阳黑子数标准差	4 日窗口滑动太阳黑子数均值	4 日窗口滑动太阳黑子数标准差
0-100	66.2	63.5	66.4	62.3
100-200	51.2	54.0	51.1	53.1
200-300	43.3	49.1	43.3	49.3
300-400	39.1	47.2	40.9	50.7
400-500	34.4	42.	33.6	36.6
500-600	24.5	37.6	28.1	44.3
600-700	20.0	21.8	21.2	24.
>=700	13.2	16.3	12.6	14.3

$$y = 1145.89 - 3.35x \quad (6)$$

太阳黑子数按每 20 个单位进行分箱，提取各分箱内的蓝藻水华面积峰值数据点，基于上述数据点构建蓝藻水华面积上限与太阳黑子数之间指数衰减关系的经验计算式如下：

$$y = 1234e^{-0.008x} \quad (7)$$

根据 SILSO 的预测结果(https://www.sidc.be/SILSO/)，2026 年 6 月上旬，预测太阳黑子数（estimated international sunspot number，EISN）在 140 左右，以上两式得到的蓝藻水华面积上限为 677 km^2、404 km^2。据此推断，若该时段内营养、气象、水文、光照条件适宜，则存在发生规模为 400~680 km^2 蓝藻水华的风险，发生更大面积蓝藻水华的概率相对较低。

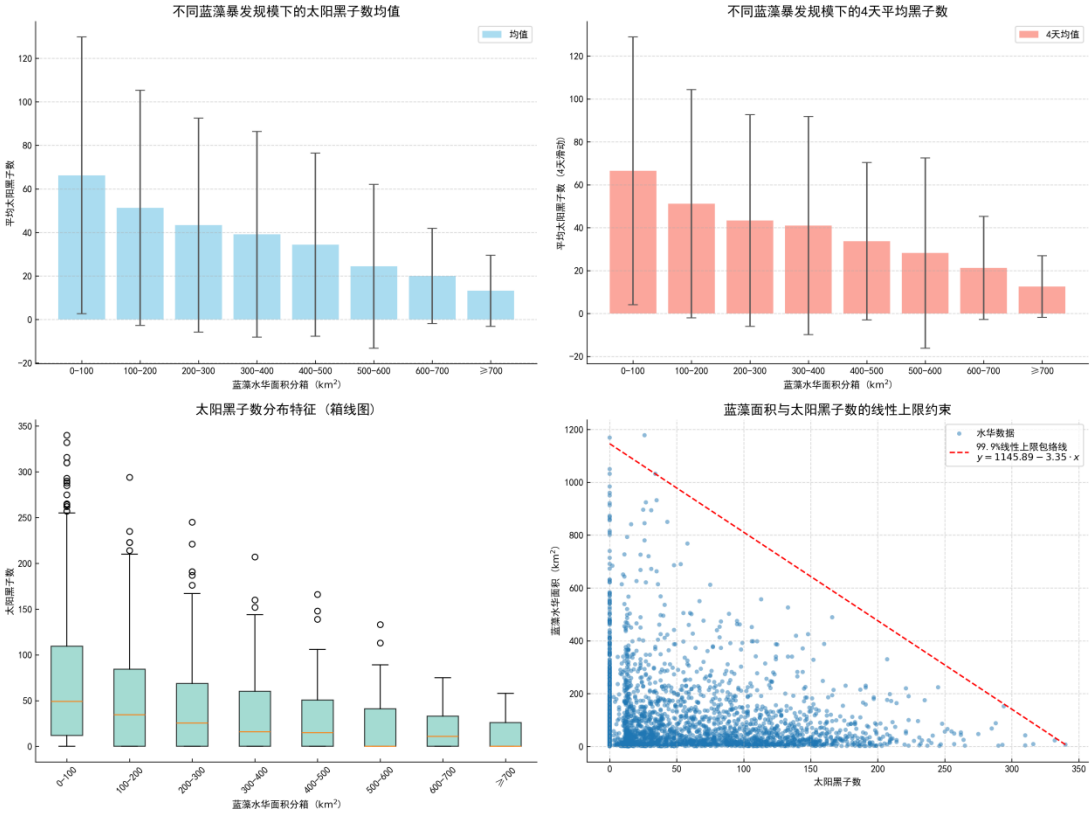

图 7　2000-2025 年间蓝藻水华面积分箱统计太阳黑子数均值及标准差绘图
和对太湖蓝藻水华面积的线性上限约束

Fig.7 Plotting of sunspot number means and standard deviations within bins of cyanobacterial bloom area from 2000 to 2025, and the linear upper bound constraint on the cyanobacterial bloom area in Lake Taihu

太阳黑子与太湖水华面积的消长关联性表明，太阳黑子活动强度可作为评估太湖蓝藻水华暴发规模风险的重要指示因子。在太阳黑子活动高年，太湖蓝藻水华的暴发规模及频次均降低明显，该现象的潜在机制尚不清晰，初步推测可能为太阳黑子活动引发空间环境变化，进而扰动地球大气和环境，促使浮游植物等生命活动产生适应性调整。蓝藻是具有光合能力的原核生物，也是地球上最古老的产氧生物，对地球表层大气环境从无氧状态变为有氧起到巨大作用，在地球环境从还原性向氧化性的长期演化过程中，蓝藻可能演化出对太阳活动变化的敏感响应机制，通过主动减少在水体表层集聚，以规避不利自身生命习性的影响。因此，在太阳黑子活动高值年卫星监测到的表观水华面积有所降低，而在太阳黑子低值年，太湖的氮磷营养盐供应充足、未成为影响蓝藻生长的关键限制因子，蓝藻则更倾向于在水体表层增殖和漂移，形成卫星可观测到的大面积水华，暴发频次也相应增加。在太阳黑子低值年，大面积水华可延续至冬季,如 2007 年 11 月 30 日、2007 年 12 月 8 日、2008 年 12 月 2 日、2017 年 12 月 11 日，甚至冬季 1 月都曾发生较大面积水华，如 2008 年 1 月 3 日和 1 月 5 日。

Hu 等[11]研究指出，澳大利亚昆士兰州 Fred Haigh 蓝藻水华指标与太阳黑子数为正向关联，该结论

与本研究相反。但该研究把蓝藻细胞计数而非蓝藻水华面积作为响应变量（Response Variable）,检测样品是利用带软管的采样泵（Integrated hose-pipe sampler）采集的，所得结果表征的为深水侧（水下 20~50 cm 水柱）的藻密度水平，而卫星遥感获取的是水体表面（0~10 cm 水柱）的蓝藻反射光谱信息，如果太阳黑子多伴随深水侧蓝藻密度增高，则存在蓝藻向深部迁移、缩小表层水华面积的可能性。

为了解蓝藻水华面积与气象因子、太阳黑子数等特征变量的关联关系，本文从开源网站 Open-Meteo(https://open-meteo.com/)下载了以太湖为中心经纬度的 2000–2025 年逐日气象数据，开展皮尔逊相关系数计算（表 3）。结果表明，蓝藻水华面积与 10 m 高度日均风速、太阳黑子数存在一定程度的负相关性，与前 10 日均温存在正相关性，与其他特征变量相关性极弱。

表 3 太湖蓝藻水华面积(A_{bloom})与特征变量的皮尔逊相关系数及显著性

Tab.3 Pearson correlation coefficients and significance between cyanobacterial bloom area (A_{bloom}) of Lake Taihu and feature variables

特征变量	指标意义及单位	皮尔逊相关系数	*P* 检验值	显著性（*P*<0.01）
T_{air} (temperature_2m_mean)	2m 高度日均气温（℃）	0.0008	0.96678	否
S_{wind} (wind_speed_10m_mean)	10m 高度日均风速（m/s）	-0.2234	5.0938×10^{-32}	极强
H_r (relative_humidity_2m_mean)	2m 高度日均相对湿度（%）	0.0525	0.00625	弱
P_{sur} (surface_pressure_mean)	日均地表气压（hPa）	0.0024	0.90212	否
R_{sw} (shortwave_radiation_sum)	日累积短波辐射（MJ/m²）	0.0052	0.78617	否
C_{cloud} (cloud_cover_mean)	日均云量（%）	-0.0432	0.02436	否
SN (sunspot_number)	每日太阳黑子数（个）	-0.1903	1.5356×10^{-23}	很强
T_{accu}	前 10 日均温（℃）	0.0798	3.1685×10^{-5}	较强

基于皮尔逊相关系数计算的变量间偏相关系数相关性热力图（图 8）显示，太阳黑子数与气象类变量基本不存在相关关系，说明太阳活动周期并未导致上述表征地表气象过程的特征变量发生同步周期性变化。

为了解各特征变量对蓝藻水华面积的影响程度，本文分别采用多元线性回归、XGBOOST、随机森林等机器学习回归模型算法开展建模分析（表 4）。多元线性回归模型中， Tair、Taccu、Swind 与太阳黑子数偏回归系数分别为-51.7544、45.4388、-35.6146、-30.8833，反映了这些特征变量对水华面积的正向或负向影响，其中前 10 日均温、相对湿度对水华形成具有促进作用，气温、风力、太阳黑子数则不利于水华形成。在 XGBOOST、随机森林回归模型中，特征重要性排名前 3 位的特征均为太阳黑子数、前 10 日均温、风力三项。上述结果表明，三种机器学习回归模型的结果均表明蓝藻水华面积与太阳黑子数存在负向关联。

表 4 太湖蓝藻水华面积(A_{bloom})与三种机器学习回归模型结果

Tab.4 Cyanobacterial bloom area (A_{bloom}) and results of three machine learning regression models for Lake Taihu

回归模型	R^2	RMSE（km^2）	MAE（km^2）	特征偏回归系数/重要性
多元线性回归	0.1129	129.8	90.0	A_{bloom}=-51.7544* T_{air} -35.6146* S_{wind}+14.9693* H_r +2.7155* P_{sur}-30.8833* SN +45.4388* T_{accu} +20.7948* C_{cloud}
XGBoost	0.1416	127.71	87.07	前 3 位重要特征：SN (0.1773)、T_{accu} (0.1676)、S_{wind} (0.1595)
随机森林	0.1893	124.11	85.88	前 3 位重要特征：S_{wind} (0.2017)、T_{accu} (0.157)、SN(0.1264)。

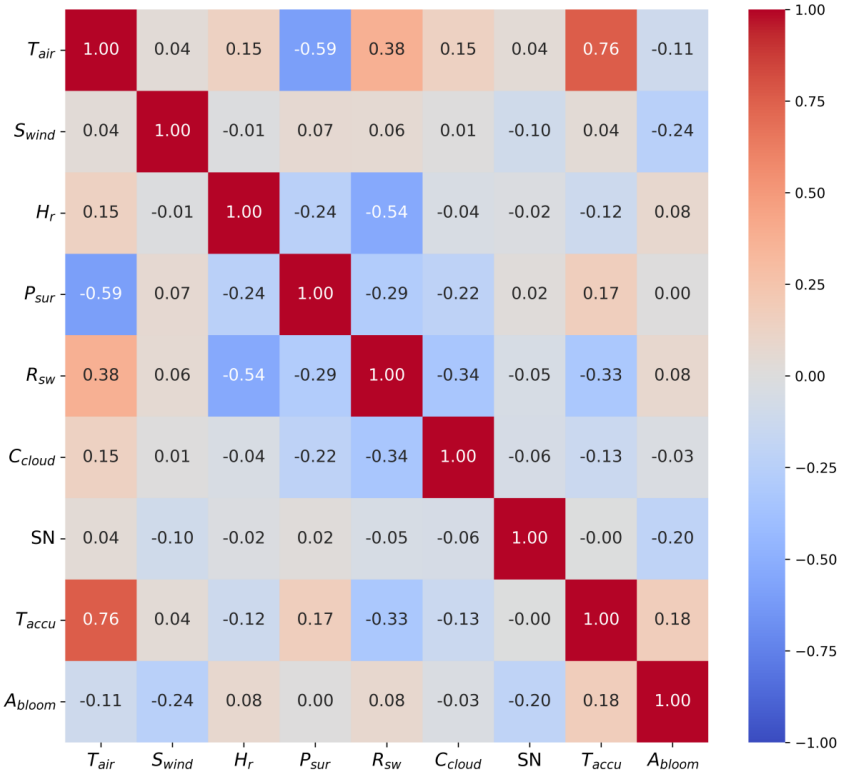

图 8 太湖蓝藻水华面积及特征变量间偏相关矩阵热力图

Fig.8 Heat map of the partial correlation matrix among A_{bloom} and all feature variables in Lake Taihu

4 结语

湖泊蓝藻水华的发生机制具有高度复杂性，尽管国内外已开展大量研究，目前仍缺乏具备确定性与定量化效力的驱动因子识别方法及预测模型。总体而言，蓝藻水华是物理、化学、生理、生境、生态多因子共同驱动、耦合作用的综合结果。本文基于 Lomb-Scargle 周期图分析识别出太阳黑子活动周期与太湖蓝藻水华面积规模存在潜在的负相关关系，该结果揭示了太阳黑子活动与蓝藻水华面积存在一定的消长关联性，在太湖蓝藻水华预测模型构建中可作指示因子将其纳入。随着卫星遥感监测太湖蓝藻水华面积数据的持续积累，未来仍需积累长期观测资料以进一步确证该负相关关系的稳定性、提供更清晰的科学解释。在此提出若干研究方向建议：一是太阳黑子活动引发的太阳辐射变化是否影响了大气热层并向下传递、影响大气环流产生气候/气象/水文效应，进而对蓝藻生长产生影响或调控效果。二是开展太阳活动对地磁场的影响及其诱发的生物效应研究，借助生命科学实验手段，了解地磁场变化与蓝藻水华发生机制之间是否存在关联。三是本研究基于太湖观测得到的初步结果提出了藻型浅水湖泊日地关联的可能性，未来仍需在全球范围对更多典型湖泊进行数据分析，更有必要综合叠加气象、水文、水质、生物等多学科、多专业、多层次指标，构建蓝藻水华机理模型及机器学习/深度学习模型开展分析。

声明:(a)署名作者的各自贡献陈述。李旭文作为第一作者及通信作者负责论文的整体研究框架、Lomb-Scargle 周期图算法及 Astropy 应用、主要分析结论和学术观点总体把握、与编辑部及与审稿专家意见互动交流、内容修改。姜晟负责协调 2000–2026 年长时间序列太湖卫星遥感数据的处理、会商复核，气象类特征变量数据的获取等工作。王甜甜负责蓝藻水华面积、太阳黑子数、主要气象特征变量的相关分析、偏回归分析、热力图等处理工作。张悦承担基于 Prophet 模型分析 2000–2025 各年度水华面积趋势、规

律等研究工作。李京航承担太阳黑子数据、太阳极紫外、UVB、VIS、NIR、SWIR 辐射数据下载、编制 Python 代码进行统计、制图，机器学习回归模型等处理工作。颜瑾承担 2007–2019 年时间序列太湖蓝藻水华卫星遥感数据解译结果复核工作，负责全文文字、英译、参考文献引用梳理。于方圆承担 2000–2007 年 6 月时间序列太湖蓝藻水华卫星遥感数据预处理、解译以及数据复核工作。郭金金参加 2020–2024 年时间序列太湖蓝藻水华卫星遥感数据预处理、解译以及数据复核工作。徐晓悦参加 2025 年时间序列太湖蓝藻水华卫星遥感数据预处理、解译以及数据复核工作。周钰菊参加 2025 年时间序列太湖蓝藻水华卫星遥感数据预处理、解译以及数据复核工作。(b)中文写作及修改全部由团队自主完成，过程中未使用生成式 AI 或 AI 辅助技术，英文翻译借助了“千问”AI 工具辅助提高语法、语态时态、术语的严谨性和专业性。承诺所有 AI 工具生成内容均经作者逐一审阅、按需修订，论文全部内容由署名作者承担完整学术责任。(c)论文关联数据读者获取：本团队不提供论文关联数据。本文的长时序蓝藻水华面积数据均基于国际上公开共享的 MODIS 等卫星遥感数据处理得到，感兴趣的研究者可以自己加工提取，太阳黑子数据可以自行下载。

5 参考文献

[1] Zhang YL, Qin BQ, Zhu GW. Long-term changes in physical environments and potential implications for the eco-environment of LakeTaihu in the past four decades. *J Lake Sci*, 2020, **32**(5): 13481359. DOI: 10.18307/2020.0503. [张运林, 秦伯强, 朱广伟. 过去 40 年太湖剧烈的湖泊物理环境变化及其潜在生态环境意义. 湖泊科学, 2020, **32**(5): 1348-1359.]

[2] Zou W, Qin BQ, Wu TH *et al*. Revisiting the limiting factors of phytoplankton proliferation:A scientific basis for cost-efficient management of lake blooms. *J Lake Sci,* 2026, **38**(3):875-884. DOI: 10.18307/2026.0300. [邹伟, 秦伯强, 吴天浩等. 浮游植物生长限制因子的审视：湖泊水华精准治理的科学基础. 湖泊科学, 2026, **38**(3): 875-884.]

[3] Qin BQ, Pan T, Cui CX *et al*. Recent changes and dynamics of cyanobacterial bloom intensities in Lake Taihu. *Chinese Science Bulletin*, 2026, **71**(13): 3018-3025. DOI: 10.1360/csb-2025-5815. [秦伯强,潘婷,崔驰潇等. 太湖蓝藻水华近期的生消变化及其效应. 科学通报, 2026, **71**(13): 3018-3025.]

[4] Yang ZX, Li YX, Zhu GW *et al.* Control factors of cyanobacterial bloom area in Lake Taihu,China(2003–2023). *J Lake Sci*, 2025, **37**(3): 734-751. DOI: 10.18307/2025.0311. [杨芷萱, 黎云祥, 朱广伟等. 2003—2023 年太湖蓝藻水华面积变化的影响因子分析. 湖泊科学, 2025, **37**(3): 734-751.]

[5] Yan ZJ, Fang C, Song KS *et al.* Advances in remote sensing monitoring of phytoplankton phenology. *J Lake Sci*, 2025, **37**(3): 1-13. DOI: 10.18307/2025.0101. [严照江, 房冲, 宋开山等. 浮游藻类物候遥感监测研究进展. 湖泊科学, 2025, **37**(3): 1-13.]

[6] Zhu GW, Guo CX *et al*. Dynamics of cyanobacterial bloom and water quality in Lake Taihu under National Class Ⅲ water quality achieved in 2024. *J Lake Sci*, 2025, **37**(3): 705-715. DOI: 10.18307/2025.0300. [朱广伟, 国超旋, 康丽娟等. 太湖水质达Ⅲ类背景下藻情与水质变化特征. 湖泊科学, 2025, **37**(3): 705-715].

[7] Li W, Qin BQ. Multiple temporal scale analysis of main driving factors of eutrophication in Meiliang Bay, Lake Taihu. *J Lake Sci*, 2012, **24**(6): 865-874. DOI: 10.18307/2012.0609. [李未, 秦伯强. 太湖梅梁湾富营养化主要驱动因子的多时间尺度分析. 湖泊科学, 2012, **24**(6): 865-874.]

[8] Deng YY, Zhou GP, Dai SW *et al*. Solar Polar-orbit Observatory. *Science Bulletin*, 2023, **68**(4): 298-308. DOI: 10.1360/TB-2022-0674. [邓元勇, 周桂萍, 代树武等. 太阳极轨天文台. 科学通报, 2023, **68**(4): 298-308.]

[9] Yin ZQ, Ma LH, Han YB *et al.* Variation of very long-term periodicity in solar activity. *Science Bulletin*, 2007 (16): 1859-1863. DOI: 10.1360/csb2007-52-16-1859. [尹志强, 马利华, 韩延本等. 太阳活动的甚长周期性变化. 科学通报, 2007 (16): 1859-1863.]

[10] Hu WB, Connell D, Mengersen K *et al*. Weather variability, sunspots, and the blooms of cyanobacteria. *EcoHealth*, 2009, **6**(1): 71-78. DOI: 10.1007/s10393-009-0223-3.

[11] 生态环境部.《水华遥感与地面监测评价技术规范（试行）》(HJ 1098-2020). 北京：中国环境出版集团, 2020.

[12] Lomb NR. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 1976, **39**(2): 447-462. DOI: 10.1007/BF00648343.

[13] Scargle JD. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 1982, **263**: 835. DOI: 10.1086/160554.

[14] Clette F, Lefèvre L. The new sunspot number: Assembling all corrections. *Solar Physics*, 2016, **291**(9): 2629-2651. DOI: 10.1007/s11207-016-1014-y.

[15] Taylor SJ, Letham B. Forecasting at scale. *The American Statistician*, 2018, **72**(1): 37-45. DOI: 10.1080/00031305.2017.1380080.

[16] Press WH, Rybicki GB. Fast algorithm for spectral analysis of unevenly sampled data. *The Astrophysical Journal*, 1989, **338**: 277-280. DOI:10.1086/167197.

[17] Lu L, Zhang HJ, Zhang H *et al.* Studying the Properties of Long-period Light Variation about Blazar 3C 454.3 with the Methods for Lomb-Scargle Periodogram. *Astronomical Research & Technology*, 2022, **19**(4): 283-290. DOI:10.14005/j.cnki.issn1672-7673.20210707.002. [陆林, 张皓晶, 张欢等. Lomb-Scargle Periodogram 方法研究耀变体 3C 454.3 长周期光变特性. 天文研究与技术, 2022**, 19**(4): 283-290.]

[18] Zhang GQ, Tu ZL, Wang FY. Possible periodic activity in the short bursts of SGR 1806-20: Connection to fast radio bursts. *The Astrophysical Journal*, 2021, **909**(1): 83. DOI: 10.3847/1538-4357/abdd27.

[19] Ye l, Qu ZN, Wang M *et al.* The Periodogram analysis of El Campo solar radar observational Data. ACTA ASTRONOMICA SINICA, 2015, **56**(6): 616-627. DOI:10.15940/j.cnki.0001-5245.2015.06.007. [叶林, 屈智宁, 汪敏等. 基于 El Campo 太阳雷达数据对太阳活动的周期性分析. 天文学报, 2015, **56**(6): 616-627.]

[20] Collaboration TA, Robitaille TP, Tollerud EJ *et al*. Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, 2013, **558**: A33. DOI: 10.1051/0004-6361/201322068.

[21] VanderPlas JT. Understanding the lomb–scargle periodogram. *The Astrophysical Journal Supplement Series*, 2018, **236**(1): 16. DOI: 10.3847/1538-4365/aab766.

[22] Baluev RV. Assessing the statistical significance of periodogram peaks. *Monthly Notices of the Royal Astronomical Society*, 2008, **385**(3): 1279-1285. DOI: 10.1111/j.1365-2966.2008.12689.x.