水生所等研发出基于大数据挖掘和深度学习的有害藻类水华预警系统

作者：writer 来源：中国科学院

本文原地址：https://www.iikx.com/news/progress/26562.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

近日，中国科学院水生生物研究所毕永红团队联合德国卡尔斯鲁厄工学院，研发出基于大数据挖掘和深度学习的有害藻类水华预警系统。相关研究成果作为封面文章，发表在《环境科学与技术》（*Environmental Science Technology*）上。

该研究致力于创建高时间分辨率的水柱垂向维度有害藻华（HABs）早期预警系统，如创建大数据平台、采集高时空分辨率的水生态环境大数据、开发自主学习并能跨时间点聚类捕获Chl *a* 动态的先进聚类算法、设计专门处理复杂多变量时间序列的预测模型等。具体工作以太湖为研究水域，构建高频垂直剖面监测系统（VAMS）进行水生态大数据采集，基于大数据开发了具有深度学习功能的Bloomformer-2模型，通过“ DeepDPM-光谱聚类 ”和Bloomformer-2模型优化建模策略，智能识别水柱分层并进行不同水深藻类生物量的精确预测，将预测结果与世界卫生组织的“ 警戒级别框架 ”有机结合，构建了有害藻类水华的预警系统。该系统能够及时预测有害藻华的发生，并能够识别出藻华的主要驱动因子。这提高了藻华的预测能力以及预测模型的环境适应性，为藻华预警及防控提供了新的解决方案。

Bloomformer-2模型具有时空多头自注意力机制。该机制使用时间和空间自注意力层的“ 查询 ”、“ 关键词 ”和“ 值 ”输出，综合了时间和空间维度。同时，优化这一机制，可更好进行多变量时间序列预测中的上下文学习。利用时空多头自注意力机制，以识别和提取各种水生态环境参数间的时空相互关系，用于Chl *a*的精确预测。

不同于传统方法，该模型为特定时间框架内的每个变量设置了独特的上下文优先级；利用基于Transformer的编码器-解码器架构，处理历史和目标数据生成的双时空序列，且整个过程包括并行双序列输入、时空嵌入、时空注意力和效率优化。为了阐明Bloomformer-2性能的稳健性，该研究使用LSTM作为比较模型；为了帮助识别和优化模型，使其在短期和长期预测中均可以提供准确预测结果，该研究进行单步预测和多步预测，更全面地评估模型在不同时间尺度上的表现和适用性，为不同预测需求提供支持。

该研究证明了预测模型在单步和多步预测中的优越性，可以准确识别藻华动态的主要驱动因素，为提前采取预防措施提供了依据。此外，该系统利用大数据的5Vs特性，提高了数据采集的体积、速度、多样性及分析数据的价值，增强了早期预警系统的性能可靠性。未来，有望通过扩大VAMS监测网络和增加传感器进一步提升系统的大数据容量和应用范围，使其更广泛地应用于HABs监测和预警。

论文链接

研究团队单位：水生生物研究所

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发