

基于时空演变多重特性建模的近海叶绿素浓度时序预测

王成贺¹ 宋宁¹ 王京禹¹ 刘安安² 聂婕^{1*}

(1.中国海洋大学信息科学与工程学部, 山东青岛 266100; 2.天津大学电气自动化与信息工程学院, 天津 300072)

摘要: 近海环境是沿海地区社会经济发展的关键支撑系统, 近海环境的持续恶化对海洋经济的可持续发展带来了巨大挑战。叶绿素浓度的反映了水体理化性质的演变规律, 对近海生态环境保护具有重要意义。尽管现有时序叶绿素浓度预测方法能从时空数据中挖掘有效信息, 揭示时空数据的发展趋势和变化规律, 但忽略了时空数据的结构化特征以及外界因素/突发因素对叶绿素浓度的影响。因此, 本文提出基于时空演变多重特性建模的近海叶绿素浓度时序预测模型, 并由四部分构成: 自相关时序预测模块预测叶绿素浓度时序变化规律; 多视角空间融合预测模块在构建预测点与其他位置叶绿素浓度空间关联性基础上, 考虑海域气象状况, 提高了空间叶绿素浓度预测的可靠性; 基于环境上下文的突变模块通过对极端因素建模, 挖掘突变因素与的叶绿素浓度变化的关联; 时空动态聚合模块利用结构化模式, 结合时间、空间叶绿素预测结果, 实现不同圈层全要素近海叶绿素浓度建模。在渤海叶绿素浓度数据上的实验结果表明, 该算法模型极大程度提升了近海叶绿素预测模型的准确性与可靠性。

关键词: 叶绿素浓度预测; 时空聚合; 集成学习

中图分类号: X55

文献标识码: A DOI:10.16798/j.issn.1003-0530.****.**.***

Temporal Prediction of Chlorophyll Concentration in Coastal Waters Based on Multi-characteristics Modeling of Spatio-temporal Evolution

Wang Chenghe¹ Song Ning¹ Wang Jingyu¹ Liu Anan² Nie Jie¹

(1.Faculty of Information Science and Engineering,Ocean University of China, Qingdao,Shandong 266100, China;2. School of Electrical and Information Engineering,Tianjin University,Tianjin,300072,China)

Abstract: Chlorophyll plays an important role in the study of marine ecosystem. However, chlorophyll concentration is affected by many coupling factors, so it is a difficult problem to predict chlorophyll concentration accurately. These factors include not only the prediction of the site's own attributes, but also the range of other site attributes that can have an impact on the chlorophyll concentration of the prediction site. At present, the time series prediction of chlorophyll concentration is realized by only considering the self-factor of prediction point, and the effect of space factor on chlorophyll concentration is ignored. Moreover, the combination of flat factors lacks structure, and it is difficult to achieve reasonable feature expression. In this paper, a data-driven spatio-temporal aggregation model is proposed to predict chlorophyll concentration. In this model, autocorrelation time series prediction model and multi-view spatial fusion prediction model are dynamically combined, and the influence of sudden change of environmental factors is considered. This paper evaluates the models using 2017 data from the Bohai Sea Region, and the results are superior to several baseline methods

Key words: chlorophyll concentration; space-time convergence; ensemble learning

基金项目: 国家自然科学基金(62072418)
中国海洋大学创新交叉团队培育计划(202042008)
山东省重大科技创新工程(2019JZZY020705)
中国海洋大学科研启动基金
青岛市科技计划重点研发专项(21-1-2-18-xx)

1 引言

近海海域是经济高质量发展的要地，是社会经济发展的关键支撑。近年来，高强度的人类活动已对海洋生态环境产生巨大的影响，我国近海环境正面临前所未有的环境压力和生态安全风险。叶绿素浓度作为表征海洋初级生产力的关键指标，其演变不仅反应了海域水体富营养化程度，也揭示了海水理化性质综合动态变化规律。因此，近海叶绿素浓度的准确预测对认识水体理化性质的演变规律、预防水体富营养化现象以及保护近海生态环境具有重要意义。

目前，叶绿素浓度预测方法分为两类：基于理化分析叶绿素浓度预测方法和基于数据驱动的叶绿素浓度预测方法^[1]。其中，传统基于理化分析叶绿素浓度预测方法依据水体本身化学性质，以复杂水体动力学模型和水质理化预测模型为基础，实现叶绿素浓度的分析和预测^[2-3]。然而，该方法只针对有限影响因素建模，不能实现多圈层全要素耦合，降低了预测模型精度。此外，其理化模型结构复杂，计算复杂度高，在处理大规模数据上尤为困难。近年来，随着机器学习的不断发展，基于数据驱动的叶绿素浓度预测方法以特征关联挖掘充分、模拟计算简单高效等优势，逐渐取代传统预测方法^[4-5]。在湖泊数据的基础上，MALKE 等人^[6]基于模糊逻辑、递归神经网络、混合进化算法和多元线性回归四种模型，实现了湖水叶绿素浓度的预测，但该研究只考虑了同一时期的水质环境变量的关联，忽略了叶绿素浓度的时序演变预测。因此，WANG 等人^[7]提出非线性自回归神经网络(Nonlinear Autoregressive Neural Network, NARX)，通过检测和考虑时间依赖性，建模叶绿素浓度与其影响因素间的动态关联关系。PARK 等人^[8]分别使用人工神经网络(ANN)与支持向量机(SVM)对韩国淡水河和河口水库进行预测建模，分别采用拉丁-超立方体单因素时间模型(LH-OAT)和模式搜索算法对输入变量进行灵敏度分析，达到了不错的预测效果。LI 等人^[9]采用判别分析(Discriminant Analysis, DA)、主成分分析(Principal Component Analysis, PCA)方法以及随机森林(Random Forests, RF)，实现鄱阳湖静水区和激流区叶绿素浓度的时序演变预测。为了克服单一时序模型预测存在的缺陷，JIA 等人^[1]提出基于聚类-回归堆叠的近岸海域浓度时空分布预测方法，提高了叶绿素

浓度预测的可靠性。ZENG 等人^[10]通过神经网络预测叶绿素浓度，他们使用原理分析的方法确定模型的输入。LIU 等人^[5]通过 2000 年至 2005 年每月对水质的调查数据，研究了水中的化学变量和叶绿素浓度之间的关联关系，并引入了物理因素、有毒物质、营养因素和浮游植物四个潜在变量。XU 等人^[11]运用支持向量机对叶绿素浓度进行了预测，避免了过拟合输出，提高了泛化性能。

尽管基于时序叶绿素浓度预测方法能从时空数据中挖掘有效信息，揭示时空数据的发展趋势和变化规律，但近海时序叶绿素浓度预测仍有以下挑战：其一，时间数据与空间数据的“扁平化”组合。具体来说，上述叶绿素浓度时空分布预测方法在空间数据划分的基础上，对每一部分空间数据进行时间序列建模，导致无法充分探索时间或空间数据内部依赖关系，忽略了数据的结构性特征。其二，叶绿素预测模型过于“理想化”。现实情况下，海域气象状况影响叶绿素浓度，现有预测模型未考虑外界因素对叶绿素浓度的影响，降低了预测模型的准确度。其三，叶绿素预测模型未考虑“突变因素”。暴雨或强风等突发因素使叶绿素浓度剧烈变化，当前预测方法不能挖掘突变因素与的叶绿素浓度变化的关联。因此，如何建立准确、可靠的近海叶绿素浓度预测模型是亟需解决的问题。

对于预测叶绿素浓度这种影响因素较多、影响类别较广的预测任务，通常采用混合动力模型。混合动力模型将不同类型数据的预处理、建模和优化操作集成在一个模型中，充分分析并整合不同因素的影响效果^[12]。ZHENG 等人^[13]通过集成了空间预测器和时间预测器的混合模型对北京天气质量进行了未来 48 小时的预测。BI 等人^[14]提出了一个综合模型，预测各个地理位置云数据中心的工作负载状况，模型的各部分分别表征趋势分量和细节成分的统计特征。混合模型是时空预测的趋势，但是目前仍未应用在叶绿素浓度预测问题上。

因此，在混合动力模型的启发下，针对以上问题本文提出基于时空演变多重特性建模的近海叶绿素浓度时序预测模型。该预测模型由四部分组成，分别为自相关时序预测模块、多视角空间融合预测模块、基于环境上下文突变的叶绿素浓度预测模块和时空动态聚合学习模块。自相关时序预测模块预测叶绿素浓度

时序变化规律；多视角空间融合预测模块在构建预测点与其他位置叶绿素浓度空间关联性基础上，考虑海域气象状况，提高了空间叶绿素浓度预测的可靠性；基于环境上下文的突变模块通过对极端因素建模，挖掘突变因素与的叶绿素浓度变化的关联；时空动态聚合模块利用结构化模式，结合时间、空间叶绿素预测结果，实现不同圈层全要素近海叶绿素浓度建模。本文利用渤海叶绿素浓度数据评估模型的有效性，对比实验与消融实验证明本文提出的模型显著优于现有的最优叶绿素浓度预测方法。

2 方法

2.1 模型架构

本文提出了一个基于时空演变的多重特性叶绿素浓度预测模型，该模型能够实现叶绿素浓度多圈层、全要素、结构化时空建模。整个模型主要分为四个模块，分别是自相关时序预测模块、多视角空间融合预测模块、基于环境上下文突变的叶绿素浓度预测模块和时空动态聚合预测模块。自相关时序预测模块采用 LSTM 方法预测站点自身叶绿素浓度时序变化规律。多视角空间融合预测模块采用随机森林模型，构建预测点与其他位置叶绿素浓度的空间关联性。基于环境上下文的突变模块，通过 Light GBM^[15]模型，对影响叶绿素浓度的极端因素建模。时空动态聚合模块，通过 Light GBM 实现不同圈层全要素结构化建模。

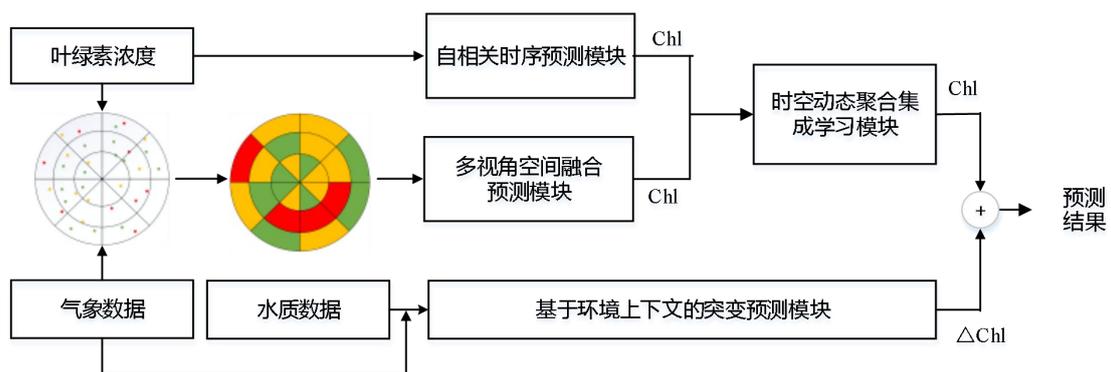


图1 模型整体架构

Fig.1 Framework

2.2 自相关时序预测模块

叶绿素浓度变化具有自相关性,为了研究自相关性对其变化的影响,本文提出了自相关时序预测模块,该模块使用 LSTM 模型挖掘数据的长时依赖信息,以此来学习叶绿素浓度的时序特征,并对叶绿素浓度变化做出预测。如图 2 所示,输入特征为 t_{i-29} 到 t_i 时刻的叶绿素浓度 $Chl_{t_{i-29}}$ 到 Chl_{t_i} , 图中输出的 h_i 即为 t_{i+1} 时刻叶绿素浓度的预测值 $Chl_{t_{i+1}}$ 。

对于每一个细胞,以 t_i 时刻为例,图 3 所示。

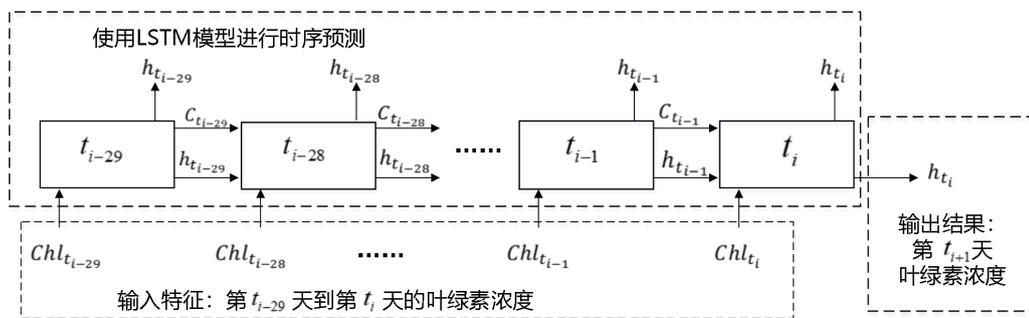


图 2 时序预测说明

Fig.2 Illustration of the temporal predictor

2.3 多视角空间融合预测模块

针对目前方法仅探究理想状态下叶绿素浓度变化趋势而未考虑到外界因素对叶绿素浓度影响的问题,本文提出了多视角空间融合预测模块建模空间影响因素。针对叶绿素浓度数据维度较高、种类复杂的特性,本文采用随机森林实现叶绿素浓度的空间预测。

首先对目标点附近的区域进行划分,如图 3,对于每一个点,以目标点为圆心在其周围划定一个直径为 96km 的圆形区域,将该区域划分为若干个子区域。用不同颜色的点表示周围不同浓度的站点。对每个子区域内所有点的叶绿素浓度取平均值以聚合子区域数据,使用聚合后的平均值代表该子区域内的平均叶绿素浓度,对于风速和风向数据也进行同样的处理。输入特征为将各个区域当天的叶绿素浓度、风速、风

向，输出为目标点下一天叶绿素浓度。

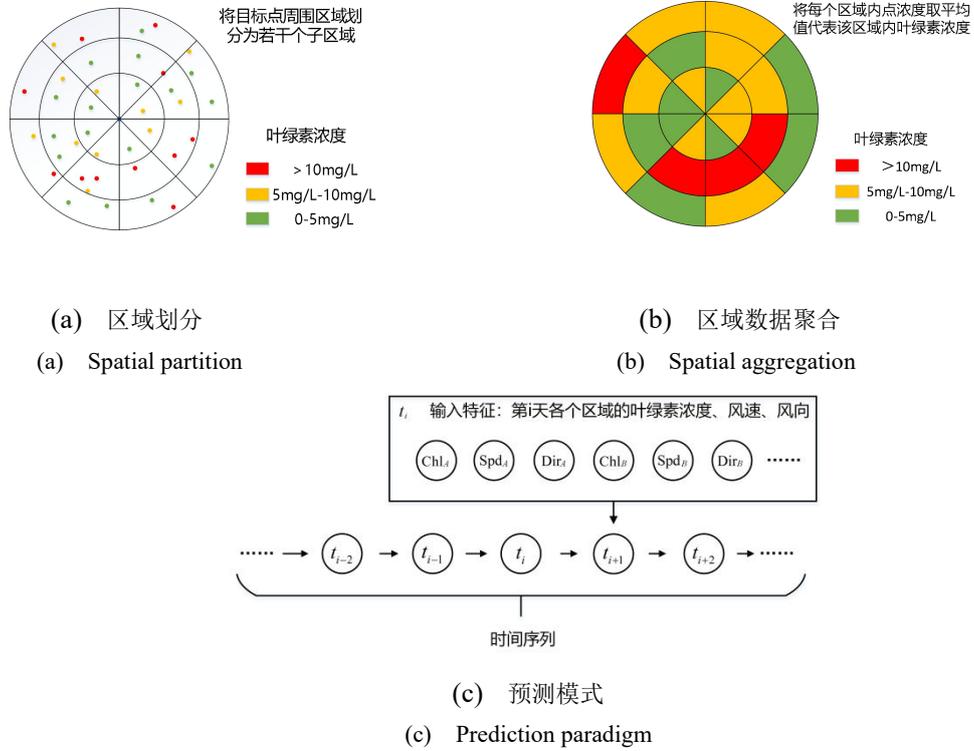


图 3 多视角空间融合预测模型介绍

Fig.3 Illustration of multi view spatial fusion prediction model

2.4 时空动态聚集成学习模块

针对传统模型时空数据组合“扁平化”的问题，本文设计了时空动态聚合学习模块，为时序预测结果和空间预测结果动态分配权重，从而实现对叶绿素浓度变化的结构化预测。利用 Light GBM，进行迭代训练从而得到最后结果。输入特征为当前时刻 t_i 的气象数据 Spd_{t_i} 、 R_{t_i} ，时序模型输出的 t_{i+1} 时刻的叶绿素浓度预测值 $T_{t_{i+1}}$ ，空间模型输出的 t_{i+1} 时刻的叶绿素浓度 $S_{t_{i+1}}$ ，输出为聚合结果 $Chl_{t_{i+1}}$ 。

采用 Light GBM 这种回归树算法，在构建新树时，通过最小化前序结果和真值的残差，找到使目标函数值最小的树，将其确定为新的决策树。在构造新的树时，每次只产生一个分支，用目标函数最小的分支来构建新的树，降低了计算复杂度。其目标函数如公式(1)所示。其中 $\Omega(f_t)$ 为正则项，如公式(2)所示。

$$Obj = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{t=1}^k \Omega(f_t) \quad (1)$$

$$\Omega(f_t) = \gamma T_t + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (2)$$

其中 T_i 是叶子的节点数， ω_j 表示在叶子 j 上节点的权重， γ ， λ 为超参数。

2.5 基于环境上下文突变的预测模块

针对叶绿素预测复杂场景中，天气等因素突变导致的叶绿素浓度在短时间内急剧变化的现象，本文提出了基于环境上下文突变的叶绿素浓度预测模块，以提高叶绿素浓度预测效果。

首先，对水质数据、气象数据以及当天的叶绿素浓度 Chl_{t_i} 与前一天叶绿素浓度 $\text{Chl}_{t_{i-1}}$ 的差值 ΔChl 进行相关性分析，找出影响短时间叶绿素浓度变化的敏感特征。相关性分析结果如表 1 所示。因此，本文选用风速、气温、降水量、化学需氧量、无机氮浓度和海水盐度作为敏感特征进行突变预测。并将其作为特征输入到 Light GBM 模型中进行训练，来预测两天之间叶绿素浓度的变化量。其次，选定一个变化量的阈值，采用异常检测中的三倍标准差法则，计算 ΔChl 的平均值 $\overline{\Delta \text{Chl}}$ 和方差 $\sigma_{\Delta \text{Chl}}$ ，将 $\overline{\Delta \text{Chl}} + 3\sigma_{\Delta \text{Chl}}$ 作为突变阈值，得出阈值为 5mg/L。最后，预测 ΔChl ，当 ΔChl 预测值超过 5mg/L 时，调用突变预测模型，将 ΔChl 与时空聚合集成学习模块所输出的叶绿素浓度相加。

表 1 相关性分析结果

Tab.1 Correlation analysis results

特征名称	权重	特征名称	权重	特征名称	权重	特征名称	权重
风速	0.1507	化学需氧量	0.0491	无机磷	0.0285	腐殖质	0.0125
气温	0.1241	无机氮	0.0380	浮游动物	0.0273	有机氮	0.0119
降水量	0.0493	盐度	0.0314	浮游植物	0.0189	有机磷	0.0108

3 实验

3.1 数据集

本文采用渤海区域叶绿素浓度、水质数据，以及气象观测数据作为数据集，其中，叶绿素浓度数据、水质数据为渤海叶绿素浓度数值模拟三维网格数据和点监测数据进行同化方法^[16-17]后的再分析数据产品。当观测数据达到与之相同体量时，本文训练的模型可直接应用在观测数据中。因此不影响本文的方法论证。

叶绿素浓度数据：使用 FVCOM 海洋数值模型生态模块得到的水质参数数据作为输入，依据适用于渤

海的经验算法 OC2 算法^[18]反演反射率得到的叶绿素浓度作为输出，在渤海选取 1000 个点进行模型的训练和测试，获得叶绿素浓度的仿真数据集。该数据集记录了 2017 年 1 月 1 日到 2017 年 12 月 31 日共 365 天的叶绿素浓度日均数据。图 4 展示了这些站点的地理位置分布，每个点表示一个数据点。叶绿素浓度反演算法 OC2 的表达式如公式(3)所示。

$$Chl_a = e_0 + 10^{e_1 + e_2 \times R + e_3 \times R^2 + e_4 \times R^3}$$

$$R = \log_{10} \left[\frac{R_{rs}(490)}{R_{rs}(555)} \right] \quad (3)$$

$$e_0 = -0.0929, e_1 = 0.2974, e_2 = -2.2429, e_3 = 0.8358, e_4 = -0.0077$$

水质数据：本文采用整个渤海区域的水质仿真数据作为水质数据集，包括海表面温度、海水盐度、溶解无机磷、溶解有机磷、溶解无机氮、溶解有机氮、浮游植物、浮游动物、腐殖质、化学需氧量十个变量。该数据集包括从 2017 年 1 月 1 日到 2017 年 12 月 31 日共 365 天的水质数据。

气象数据：本文采用来自 NOAA 的日均气象监测数据作为气象数据集。该数据集包括渤海周围的九个气象检测点的数据，检测点分别为：承德、惠民、乐亭、龙口、青龙、唐山、潍坊、长岛、周水子，数据类型包括：平均气温、海平面气压、风速、降水数据。该数据集记录了从 2017 年 1 月 1 日到 2017 年 12 月 31 日，共 365 天的气象观测数据。图 5 展示了气象数据监测站点分布情况。

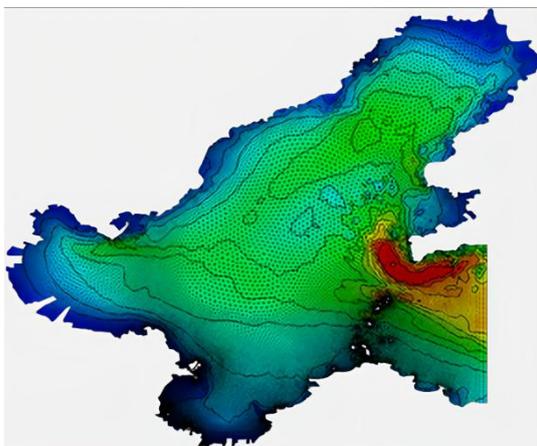


图 4 叶绿素浓度及水质监测点分布

Fig.4 Chlorophyll concentration and distribution of water quality monitoring Station

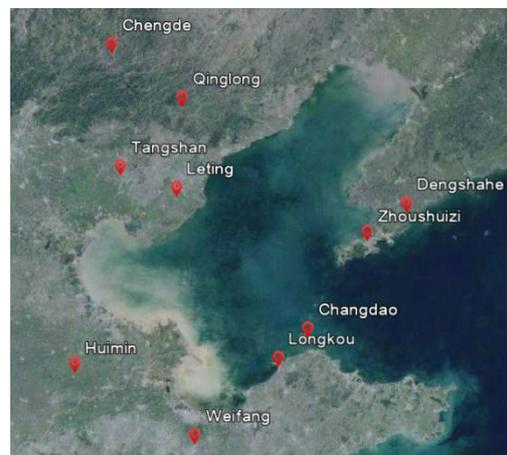


图 5 气象监测站分布

Fig.5 Distribution of meteorological monitoring stations

3.2 评价指标

本文采用两个评价指标，分别为 RMSE 和 R^2 。

RMSE(Root Mean Squard Error)，即均方根误差，用于计算预测值和实际值之间的误差，能够表示预测的准确程度，其公式如(4)所示。

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (4)$$

R^2 (R-Square)，即确定系数，是回归平方和和总平方和的比值，用于表示回归方程的拟合性能，越接近 1，则表示拟合的效果越好，公式如(5)所示。使用 R^2 在时间和空间两个维度下对实验结果进行评价。在时间维度上，计算同一个站点所有时刻的预测结果与真实值的 R^2 ，在空间维度上则计算某一时刻所有站点预测结果与真实值的 R^2 。

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (5)$$

其中， y_i 为 i 时刻叶绿素浓度真实值， \bar{y}_i 表示叶绿素浓度平均值， \hat{y}_i 为 i 时刻叶绿素浓度预测值。

3.3 实验细节

本文中的模型使用 Pytorch 模型框架实现。实验中将数据集按照 4: 1 的比例划分为不重叠的训练集和测试集，进行模型的训练和测试。在自相关时序预测模块中，选用 LSTM 模型。模型迭代次数设置为 100 次，每次迭代训练批次大小为 64。在多视角空间融合预测模块中，选用随机森林模型。其中决策树的数量设置为 200。时空动态聚合集成学习模块中选用 Light GBM 模型。树的最大叶子节点数设置为 1000，树的数量设置为 1000。突变模块中使用 Light GBM 模型。树的最大叶子节点数设置为 100，树的数量设置为 500。

3.4 对比实验

为了验证本文方法得有效性，本节将所提出的叶绿素浓度时序预测方法与现有的先进预测方法进行比较。在数据集上分别测试了 RMSE、时间 R^2 和空间 R^2 三个性能指标，从表 2 可以看出，本文提出的方法在 RMSE、时间 R^2 和空间 R^2 三个评价指标下都具有较好的结果。SVR 和线性回归作为传统的时序预测模

型对叶绿素浓度时序特征进行了有效地学习，但是由于忽略了空间相关性以及缺少突变因素建模，导致预测结果欠佳。同时，数据量庞大时，传统方法难以实现高质量拟合，预测效果远不如本文所提出的模型。

表 2 对比实验结果

Tab.2 Results of comparative experiments

模型	评价指标		
	RMSE	时间 R ²	空间 R ²
SVR	1.4348	-12.1452	0.1255
线性回归	1.3774	-8.96180	0.0677
本文模型	0.3896	0.5066	0.9156

3.5 消融实验

本节通过消融实验来验证每个模块的有效性。在上文提到的数据集上采用如下方式进行消融分析：

1) 仅使用自相关时序预测模块进行预测，用时序预测表示；

2) 仅使用多视角空间融合预测模块进行预测，用空间预测表示；

3) 使用自相关时序预测模块、多视角空间融合模块进行预测，并将预测结果使用时空动态聚合集成学习模块进行聚合，用时序+空间+聚合表示；

4) 使用自相关时序预测模块、多视角空间融合模块进行预测，并将预测结果使用时空动态聚合集成学习模块进行聚合，并加入基于环境上下文突变的叶绿素浓度预测进行预测，用时序+空间+聚合+突变表示。

表 3 所示为时序预测模型、空间模型、突变模型、时序+空间、时序+空间+聚合模型、时序+空间+聚合模型+突变模型的实验效果。实验结果表明，在叶绿素浓度预测中，本文提出的方法可以使预测效果提升。

表 3 消融实验结果

Tab.3 Results of Ablation Experiment

模型	评价指标		
	RMSE	时间 R ²	空间 R ²
时序预测	0.8234	-2.1033	0.6807
空间预测	0.6398	-1.0238	0.8036
时序+空间+聚合模型	0.3902	0.5062	0.9154
时序+空间+聚合模型+	0.3896	0.5066	0.9156

4 结论

本文提出了一种基于时空动态聚合模型实现叶绿素浓度高准确度预测。为了解决时空预测模型过于理想化的问题，本文提出了多视角空间融合预测模块，对影响叶绿素浓度的外界因素进行建模。针对时空数据组合扁平化的问题，本文提出了时空动态聚合集成学习模块，探究时空数据内部依赖关系。为了解决在观察过程中极少存在的突变因素使叶绿素浓度急剧变化的预测问题，本文提出了突变模块，单独训练偏转预测器。本文动态聚合了时间、空间和突变预测模块，对叶绿素浓度进行全要素、多视角、结构化预测建模。与最先进的方法相比，本文的方法更加全面的纳入并融合了多圈层耦合因素预测叶绿素浓度，取得了与真实值更接近的预测结果。

参考文献:

- [1] JIA W, CHENG J, HU H. A cluster-stacking-based approach to forecasting seasonal chlorophyll-a concentration in coastal waters[J]. IEEE Access,2020,vol 8: 99934-99947.
- [2] HAM Y G, JOO Y S,PARK JY . Mechanism of skillful seasonal surface chlorophyll prediction over the southern Pacific using a global earth system model[J]. Climate Dynamics, 2021, 56(12):1-20.
- [3] NAZEER M ,NICHOL J E. Development and application of a remote sensing-based chlorophyll-a concentration prediction model for complex coastal waters of Hong Kong[J]. Journal of Hydrology, 2016, 532(10):80-89.
- [4]YING. W, HONG.J, ZHANG.X, WANG.Y and CHENG.C., Temporal-spatial variations of chlorophyll-a in the adjacent sea area of Yangtze river estuary, [C].2010 18th International Conference on Geoinformatics.Beijing,China; IEEE, 2010: 1-5.
- [5]LIU.Y., GUO.H. and YANG.P.Exploring the influence of lake water chemistry on chlorophyll a: A multivariate statistical model analysis[J]. ECOLOGICAL MODELLING, 2010, 221(4):681-688.
- [6]MALEK S , AHMAD S , SINGH S , et al. Assessment of predictive models for chlorophyll-a concentration of a tropical lake[J]. BMC Bioinformatics, 2011, 12(13):1-11.
- [7]WANG.H, YAN.X,CHEN.H, et al. Chlorophyll-a predicting model based on dynamic neural network[J]. Applied Artificial Intelligence, 2015, 29(10): 962-978.
- [8]PARK.Y, CHO.K.H, PARK.J, CHA.S.M and KIM.J.H, Development of early-warning protocol for predicting chlorophyll-a concentration using machine learning models in freshwater and estuarine reservoirs, Korea, [J]. Science of the Total Environment, 2015, 502:31-41.
- [9]LI B, YANG G, WAN R, et al. Combining multivariate statistical techniques and random forests model to assess and diagnose the

- trophic status of Poyang Lake in China[J]. *Ecological Indicators*, 2017, 83: 74-83.
- [10] ZENG.Y,YANG.Z,LIU J . Prediction of the concentration of chlorophyll-a for Liuhai urban lakes in Beijing City[J]. *Journal of Environmental Sciences*, 2006, 18(4): 827-831.
- [11]XU.Y., et al., Method to predict key factors affecting lake eutrophication—A new approach based on Support Vector Regression model[J]. *International Biodeterioration & Biodegradation*, 2015, 102: 308-315.
- [12]ZHAI and CHEN.J , Development of a stacked ensemble model for forecasting and analyzing daily average PM2. 5 concentrations in Beijing, China[J]. *Science of The Total Environment*, 2018, 635: 644-658.
- [13]ZHENG Y, YI X, LI M, et al. Forecasting fine-grained air quality based on big data[C].*Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. Sydney,Australia:2015: 2267-2276.
- [14]BI.J, YUAN.H., ZHANG.L. and ZHANG.J., SGW-SCN: An integrated machine learning approach for workload forecasting in geo-distributed cloud data centers[J]. *Information Sciences*, 2019, 481: 57-68.
- [15]Ke, Guolin, et al. "Lightgbm: A highly efficient gradient boosting decision tree." *Advances in neural information processing systems* 30 (2017).
- [16] 吕咸青, 樊伟. 海洋生态系统模型中参数空间变化的伴随法反演研究[J]. *中国海洋大学学报(自然科学版)*, 2009, 39(5): 846-854.
- LV Xianqing, FAN Wei. Numerical study on spatially varying parameters of a marine ecosystem model with the adjoint method[J]. *Periodical of Ocean University of China*, 2009, 39(5): 846-854.(in Chinese)[[知网中文](#)][[知网英文](#)]
- [17] 徐青, 刘玉光, 程永存, 等. 海洋生态模型中的伴随同化方法[J]. *海洋通报*, 2005, 24(6): 58-64.
- XU Qing, LIU Yuguang, CHENG Yongcun, et al. Adjoint method in marine ecosystem models[J]. *Marine Science Bulletin*, 2005, 24(6): 58-64.(in Chinese)[[知网中文](#)][[知网英文](#)]
- [18] HOOKER S B, MCCLAIN C R. The calibration and validation of SeaWiFS data[J]. *Progress in Oceanography*, 2000, 45(3/4): 427-465.[[LinkOut](#)]

作者简介



王成贺 男, 1996 年出生, 山东德州人, 中国海洋大学硕士研究生, 主要研究方向: 海洋大数据, E-mail:

wch@stu.ouc.edu.cn



宋宁 男，1996 年出生，山东菏泽人，中国海洋大学博士研究生，主要研究方向：数据挖掘，E-mail: songning@stu.ouc.edu.cn



王京禹 女，1997 年出生，山东德州人，中国海洋大学博士研究生，主要研究方向：海洋环境智能监测，E-mail: wangjingyu3186@stu.ouc.edu.cn



刘安安 男，1982 年出生，黑龙江齐齐哈尔人，天津大学教授，主要研究方向：计算机视觉、机器学习，E-mail: liuanan-tju@163.com。



聂婕（通讯作者） 女，1984 年出生，山东济宁人，中国海洋大学副教授，主要研究方向：多媒体内容分析、人工智能，E-mail: E-mail:niejie@ouc.edu.cn

创新点说明

本文所提出的多视角空间融合预测模块，在构建预测点与其他位置叶绿素浓度空间关联性基础上，考虑海域气象状况，提高了空间叶绿素浓度预测的可靠性；本文所提出的基于环境上下文突变预测模块，通过对极端因素建模，挖掘突变因素与的叶绿素浓度变化的关联；本文所提出的时空动态聚合模块利用结构化模式，结合时间、空间叶绿素预测结果，实现不同圈层全要素近海叶绿素浓度建模。

