

DOI: 10-12264/JFSC2022-0089

基于自动机器学习的黄、渤、东海蓝点马鲛渔场丰度指数研究

秦元哲¹, 周振佳¹, 刘阳¹, 田永军¹, 程家骅², 刘旭东³, 张勇³

1. 中国海洋大学水产学院, 渔业海洋学研究室, 山东 青岛 266003;

2. 中国水产科学研究院东海水产研究所, 上海 200090;

3. 山西大地新亚科技有限公司, 山西 太原 030021

摘要: 蓝点马鲛(*Scomberomorus niphonius*)是中国近海重要的大型中上层经济鱼类, 精准预测蓝点马鲛渔场分布对渔业资源评估与管理有重要意义。本研究利用蓝点马鲛捕捞数据与卫星遥感海表温度(sea surface temperature, SST)和海表叶绿素浓度(chlorophyll-a concertation, Chl a)数据, 构建了基于自动机器学习的蓝点马鲛 CPUE 预测模型, 通过与 XGBoost 模型、随机森林模型和广义加性模型(generalized additive models, GAM)对比, 自动机器学习模型的确定系数(coefficient of determination, R^2)分别提高了 51%、107%和 117%, 均方根误差(root mean squared error, RMSE)分别降低了 15%、28%和 32%。通过模型预测的蓝点马鲛 CPUE, 开发了渔场丰度指数, 分析了渔场丰度时空变动规律。结果显示: 蓝点马鲛渔场丰度高值区在春季由于受到 SST 的影响较大, 呈现向北及向近岸移动的趋势, 这种现象与蓝点马鲛索饵及产卵洄游路线一致; 同时, 蓝点马鲛渔场丰度高值区纬度重心的北移, 也与气候变暖影响下蓝点马鲛索饵和产卵的适宜温度区变动有关。通过气候事件指数分析发现, Niño 3.4 指数与蓝点马鲛渔场丰度高值区面积呈显著相关。结合 RCP2.6、RCP4.5、RCP6.0、RCP8.5 4 种情景, 分别预测了 2100 年蓝点马鲛的 CPUE 分布变动, 发现随着全球变暖, 蓝点马鲛 CPUE 高值产区北移, 并相比 2010—2015 年蓝点马鲛 CPUE 预测平均值, 4 种 RCP 情景下分别上升了 0.1、2.2、2.41 和 17.3 kg/h。本研究结果可为中国近海经济鱼种的渔情预报研究提供参考。

关键词: 蓝点马鲛; 渔情预报; 机器学习; 气候变化; 渔场丰度指数

中图分类号: S931

文献标志码: A

文章编号: 1005-8737-(2022)09-1375-13

蓝点马鲛(*Scomberomorus niphonius*)俗名鲅鱼、马鲛、燕尾鲅等, 为暖水性中上层大型经济鱼类, 广泛分布于印度洋、太平洋地区, 我国黄海、渤海、东海亦是其重要产区^[1]。蓝点马鲛具有生长快、个体大等优点, 经济价值高, 是北方海洋渔业重要捕捞对象之一^[2]。由于近年来中国近海捕捞压力增加和海岸带工程等原因, 蓝点马鲛出现了低龄化、小型化, 繁殖期提前等趋势^[3]。在整体近海渔业资源衰退的背景下^[4-5], 蓝点马鲛渔业的产量目前仍较为稳定, 年产量维持在 45 万 t 左右^[6], 在我国海洋捕捞产业中具有重要地位。因此, 预测蓝点马鲛产量及分布动态对于其渔业资源的

可持续利用具有重要意义。

渔情预报在渔业生产和管理中具有重要的作用。刘树勋等^[7]利用判别分析法, 采用 1960—1978 年渔业公司捕捞资料, 对舟山渔场、浙江南部渔场、长江口渔场的中心渔场进行了预报, 并取得理想效果。张孟海^[8]采用多元线性回归, 预测了中国毛虾(*Acetes chinensis*)的产量和中心渔场位置, 并探索在实际生产上的应用。在渔情预报领域, GAM 是最常用的预测模型^[9-10], Lan 等^[9]基于 GAM 模型对热带太平洋黄鳍金枪鱼(*Thunnus albacares*)的潜在捕捞区域进行了预测。牛明香等^[10]采用 GAM 模型预测了东南太平洋智利竹筴鱼

收稿日期: 2022-03-15; 修订日期: 2022-05-11.

基金项目: 国家自然科学基金重点项目(41930534); 中央高校基本科研业务费专项(202242001).

作者简介: 秦元哲(1997-), 男, 硕士研究生, 研究方向为渔业遥感. E-mail: qinyzouc@163.com

通信作者: 刘阳, 教授, 博士生导师, 研究方向为渔业遥感. E-mail: Yangliu315@ouc.edu.cn

(*Trachurus murphyi*)中心渔场。近年来,由于机器学习模型(machine learning, ML)具有更高的拟合精度及更快的运算速度^[11-13],为渔业资源的评估与预报提供了便利条件,在渔业生态学相关研究中表现出比 GAM 更好的预测效果^[14]。但同时存在参数设置复杂、参数调整耗时等缺点,而自动机器学习模型(automated machine learning, AutoML)可以将参数设置步骤自动化,进一步改进和提升机器学习模型的精度和性能^[15]。

全球气候变化及其带动的海洋环境变化可以影响鱼类的分布^[16]。Tian 等^[17]发现了长期气候与海洋变化对秋刀鱼(*Cololabis saira*)资源丰度变动的响应机理。Kim 等^[18]发现鳀(*Engraulis japonicus*)的捕捞产量受到厄尔尼诺现象的影响。Liu 等^[19]发现鳀冬季渔场受到气候变动的影响,随着温度升高鳀产量增加。很多研究发现,蓝点马鲛的产量和分布变动与海洋环境变化密切相关^[20-23]。韦晟等^[20]根据 1972—1980 年水温调查资料,发现等温线的分布与蓝点马鲛中心渔场的位置相关性强。宋超等^[23]研究发现,当水温为 15~19 °C,蓝点马鲛鱼卵丰度最高。IPCC (Intergovernmental Panel on Climate Change)基于 4 种温室气体浓度情景(RCP2.6、RCP4.5、RCP6.0、RCP8.5)对未来全球气候进行推算,至 21 世纪末(2081—2100),相比于 1986—2005 的温度基线,全球温度将上升 1.0~3.7 °C^[24]。因此,未来气候变动对蓝点马鲛的影响成为关注的热点之一。

本研究的主要目的为:(1) 利用蓝点马鲛渔业调查数据与遥感数据,构建精准预测蓝点马鲛丰度指数的自动机器学习模型;(2) 分析蓝点马鲛预测高值产区的分布规律和变动机制;(3) 探明蓝点马鲛预测高值产区变动与环境因素之间的影响机理,预测在未来不同 IPCC-RCP 情景下全球变暖对蓝点马鲛高值产区分布的影响。

1 材料与方法

1.1 数据来源

1.1.1 蓝点马鲛数据 蓝点马鲛渔业数据来源于中国水产科学研究院东海水产研究所在 2010—2015 年 3—5 月的渔业调查数据。模型构建使用调查数

据集中的日期、经纬度及单位捕捞努力渔获量(catch per unit effort, CPUE), CPUE 数据以(kg/h)表示。调查区域主要集中于山东、江苏、浙江近岸海域(图 1),共获取调查数据 489 条,作业方式为双拖。

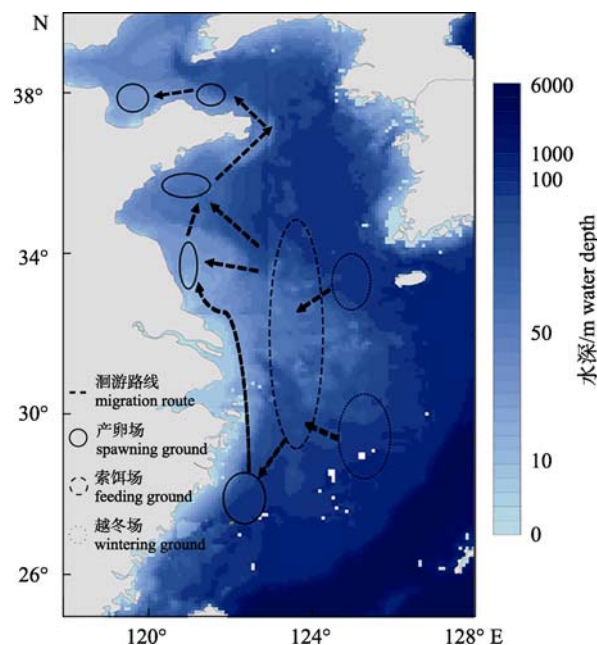


图 1 蓝点马鲛春季洄游路线^[25]

Fig. 1 Spring migration route of *Scomberomorus niphonius* ^[25]

1.1.2 卫星遥感数据 本研究使用的海洋环境数据包括:海表面温度和叶绿素浓度,数据来自美国海洋与大气管理局(NOAA)中分辨率成像光谱仪(moderate resolution imaging spectroradiometer, MODIS)的 L3 级月平均数据,其空间分辨率为(4 km×4 km),网址(<https://oceanwatch.pifsc.noaa.gov/index.html>)。

1.1.3 气候指数 本研究应用的气候指数为 Niño3.4 指数,数据来自美国海洋大气管理局气候预测中心(https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/)。

1.1.4 不同 RCP 情景下温度数据 根据 IPCC 采用的 4 种温室气体浓度情景,代表路径浓度(representative concentration pathway, RCP)由低到高分别为 RCP2.6、RCP4.5、RCP6.0 和 RCP8.5,数字表示到 2100 年辐射强迫水平分别为 2.6 W/m²、4.5 W/m²、6.0 W/m² 和 8.5 W/m²。在 4 种温室气体浓度情景下,对应温度分别上升 1.0 °C、1.8 °C、2.2 °C 和 3.7 °C^[26]。

1.2 研究方法

1.2.1 数据预处理 对蓝点马鲛 CPUE 数据进行对数变换, 公式如下:

$$\text{CPUE}_{\text{tran}} = \lg(\text{CPUE} + \overline{\text{CPUE}})$$

式中, $\overline{\text{CPUE}}$ 为蓝点马鲛全部 CPUE 值的平均值, 通过对数变换消除由于部分生产数据过高带来的影响。

再将所有数据标准化到[0, 1], 公式如下:

$$x_{\text{tran}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

式中, x_i 为第 i 个变量, $x_{\min}$ 为数据集中的最小值, $x_{\max}$ 为数据集中的最大值, 通过数据标准化以统一量纲。

1.2.2 XGBoost 模型构建 本研究使用 python 3.7 软件中的 XGBoost 程序包 (<http://xgboost.readthedocs.io/en/latest/>) 完成模型构建。利用渔业数据及对应的环境变量, 顺序建立多个决策树模型。每个新决策树模型在前一个决策树模型的基础上, 根据损失函数的梯度, 提升模型拟合精度, 将最终模型的预测值作为预测结果^[27]。通过比较模型在不同参数下的拟合精度, 对重要模型参数进行调整, 最终确定模型的最优参数^[28]。

$$R^2 = \left\{ \frac{N \times \sum_{i=1}^N (\text{CPUE}_{i, \text{obse}} \times \text{CPUE}_{i, \text{pred}}) - \sum_{i=1}^N \text{CPUE}_{i, \text{obse}} \times \sum_{i=1}^N \text{CPUE}_{i, \text{pred}}}{\sqrt{\left[N \times \sum_{i=1}^N (\text{CPUE}_{i, \text{obse}})^2 - \left(\sum_{i=1}^N \text{CPUE}_{i, \text{obse}} \right)^2 \right] - \left[N \times \sum_{i=1}^N (\text{CPUE}_{i, \text{pred}})^2 - \left(\sum_{i=1}^N \text{CPUE}_{i, \text{pre}} \right)^2 \right]}} \right\}^2$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\text{CPUE}_{i, \text{obse}} - \text{CPUE}_{i, \text{pred}})^2}{N}}$$

式中, N 为数据集中数据的组数, $\text{CPUE}_{i, \text{obse}}$ 为第 i 组数据的实际值, $\text{CPUE}_{i, \text{pred}}$ 为第 i 组数据的预测值。通过十折交叉验证(10-fold cross validation)^[28] 测试模型的准确性, 将数据集随机均分为 10 份, 轮流将其中的 9 份作为训练集, 另 1 份作为测试集进行测试, 将 10 次测试集 R^2 和 RMSE 的平均值作为结果, 公式如下:

1.2.3 随机森林模型构建 本研究在 python 3.7 软件中应用 scikit-learn 中的函数 (<https://scikit-learn.org/stable/>) 来构建随机森林模型。计算机基于模型输入的环境变量与其对应的渔业数据, 随机建立多个决策树模型进行预测, 并将各个决策树模型预测值的平均值作为随机森林预测结果。模型最优参数的确定与 1.2.2 中 XGBoost 模型的参数调整方法相同^[28]。

1.2.4 广义加性模型构建 本研究中采用 GAM 模型作为对比模型来评估机器学习模型的表现。使用 R version 4.0.5 软件中的 MGCV 包构建 GAM 模型, 表达式为^[29]:

$$\text{CPUE}_{\text{tran}} = \alpha + \sum s(x_i) + \varepsilon$$

式中, $\text{CPUE}_{\text{tran}}$ 为经过 0~1 变换之后的 CPUE, s 为平滑函数, x_i 为第 i 个环境变量, 包括叶绿素浓度、温度等, α 为截距项, ε 为误差项。并采用 AIC^[29] 准则选择最优模型, AIC 值越小, 模型越准确, 其公式为:

$$\text{AIC} = -2\lg(L) + 2m$$

式中, m 为模型中参数个数, L 为最大似然值。

1.2.5 模型检验 本研究使用 R^2 和 RMSE^[28] 对模型进行评价。其公式分别为:

$$R^2_{\text{output}} = \frac{\sum_{i=1}^{10} R_i^2}{10}$$

$$\text{RMSE}_{\text{output}} = \frac{\sum_{i=1}^{10} \text{RMSE}_i}{10}$$

式中, R^2_{output} 和 $\text{RMSE}_{\text{output}}$ 为模型测试结果, R_i^2 为第 i 次验证测试集的 R^2 值, RMSE_i 为第 i 次验证测试集的 RMSE 值。

1.2.6 蓝点马鲛渔场丰度指数模型 本研究通过 R^2 和 RMSE 对已构建模型进行筛选, 选取最优模

型构建蓝点马鲛渔场丰度指数模型(图 2)。使用对数变换对数据集进行预处理后, 导入随机森林模型、XGBoost 模型和 GAM 模型进行模型构建, 对比模型测试集的 R^2 和 RMSE, 选取精度最高模型构建蓝点马鲛 CPUE 预测模型, 通过比较不同参数组合下测试集的 R^2 和 RMSE, 自动选取最优参数组合完成模型构建, 根据蓝点马鲛 CPUE 预测结果, 提出渔场丰度指数换算公式用于评估渔场资源丰度^[14], 高指数预测区域代表渔场丰度较高, 更适宜渔船捕捞作业, 其表达式为:

$$\text{渔场丰度指数} = \frac{\text{CPUE}_i - \text{CPUE}_{\min}}{\text{CPUE}_{\max} - \text{CPUE}_{\min}}$$

式中, CPUE_i 为第 i 个变量, $\text{CPUE}_{\min}$ 为数据集的最小值, $\text{CPUE}_{\max}$ 为数据集的最大值, 根据渔场

丰度指数绘制蓝点马鲛时空分布预测图, 对不同时点蓝点马鲛渔情进行预报。

1.2.7 模型敏感性分析 预测模型环境因子数值范围的变化, 可能导致模型预测精度降低。在环境因子数值范围变化时, 需分析蓝点马鲛 CPUE 预测模型的模型敏感性, 其表达式为^[30]:

$$S = \frac{1}{n} \sum \frac{x_t - x_0}{x_0}$$

式中, S 为模型敏感性指数, n 为数据的组数, x_0 为第 t 组环境数据未发生改变时模型预测结果, x_t 为第 t 组环境数据发生改变后的模型预测结果。通常环境参数改变 10% 时, 若模型敏感性指数 $S < 100\%$, 表明环境因子数值范围的不确定性对模型预测结果精度的影响不显著。

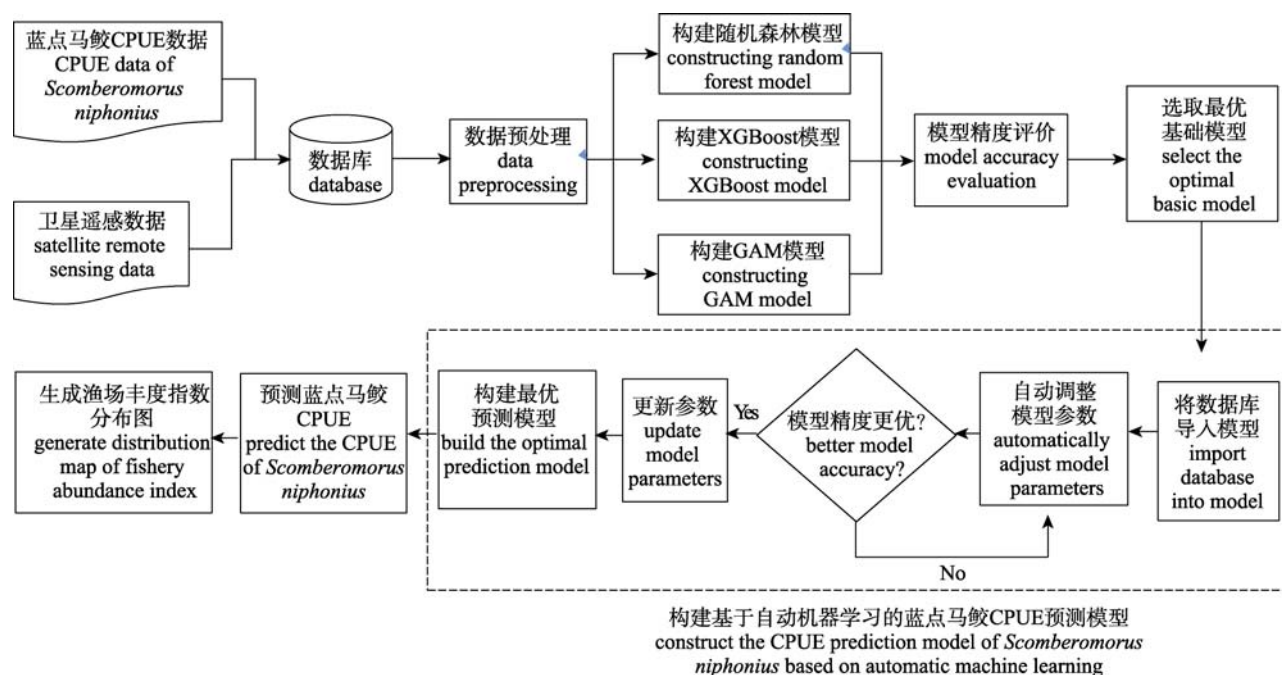


图 2 蓝点马鲛渔场丰度指数模型流程图

Fig. 2 Flow chart of fishery abundance index model of *Scomberomorus niphonius*

2 结果与分析

2.1 自动机器学习模型的筛选及精度评价

基于 Scikit-learn 中的 GridsearchCV 函数, 本研究利用十折交叉验证^[28], 比较模型测试集的 R^2 和 RMSE, 选取表现最优模型构建自动机器学习模型。模型结果显示: XGBoost 模型中 SST 和 Chl a 的分子重要性分别为 69.85、30.95; 随机森

林模型中 SST 和 Chl a 的分子重要性分别为 62.64、37.36。相比于 SST, Chl a 对模型预测结果影响较大, 因此蓝点马鲛 CPUE 更易受 Chl a 影响。

评价指标结果显示(图 3), XGBoost 模型拟合效果最佳($R^2=0.59$)且误差较低(RMSE=18.52), 而 GAM 模型($R^2=0.41$, RMSE=23.20)预测效果相对较差; 且 XGBoost 模型在低值区和高值区预测值的表现结果较为稳定, 而随机森林模型和

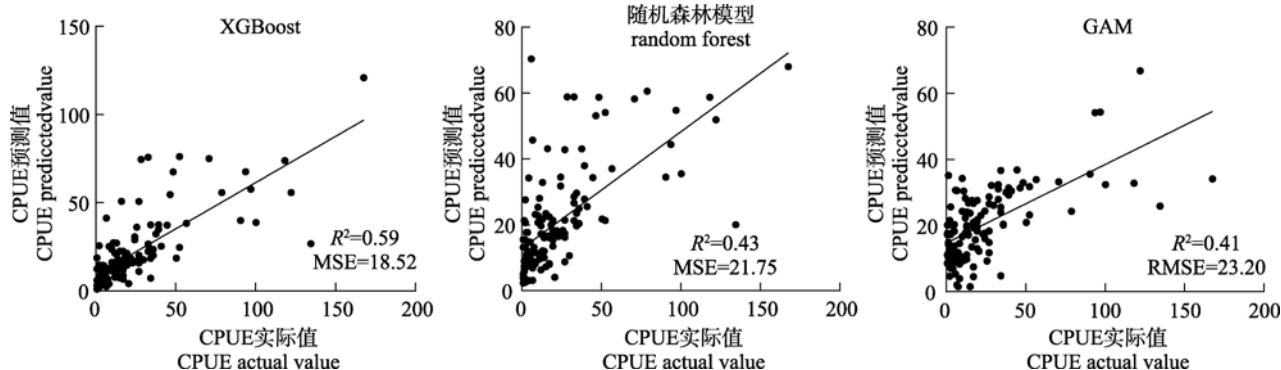


图3 随机森林模型、XGBoost模型和GAM模型的验证结果

Fig. 3 Validation results of random forest model, XGBoost model and GAM model of *Scomberomorus niphonius*

GAM 在低值区域的预测值波动较大。因此, 本研究采用 XGBoost 模型构建蓝点马鲛渔场丰度指数模型。

2.2 蓝点马鲛预测渔场丰度分布变动

蓝点马鲛渔场丰度指数模型预测的 2010—2015 年春季蓝点马鲛渔场丰度指数分布图显示, 蓝点马鲛渔场丰度指数高值区存在向北以及向近岸移动的趋势(图 4)。3 月, 渤海海域蓝点马鲛渔场丰度指数较低, 黄海南部以及东海海域蓝点马鲛渔场丰度指数较高, 适宜蓝点马鲛渔船进行作业; 4 月, 黄海中部以及东海海域蓝点马鲛渔场丰度指数较高, 蓝点马鲛适宜作业海域向北移动; 5 月, 渤海海域以及黄海北部海域蓝点马鲛渔场丰度指数较高, 适宜作业海域继续向黄海北部并向渤海沿岸、中朝韩沿岸移动。随着时间变化, 蓝点马鲛渔场丰度指数高值区存在向朝鲜半岛近岸和向北移动的趋势。

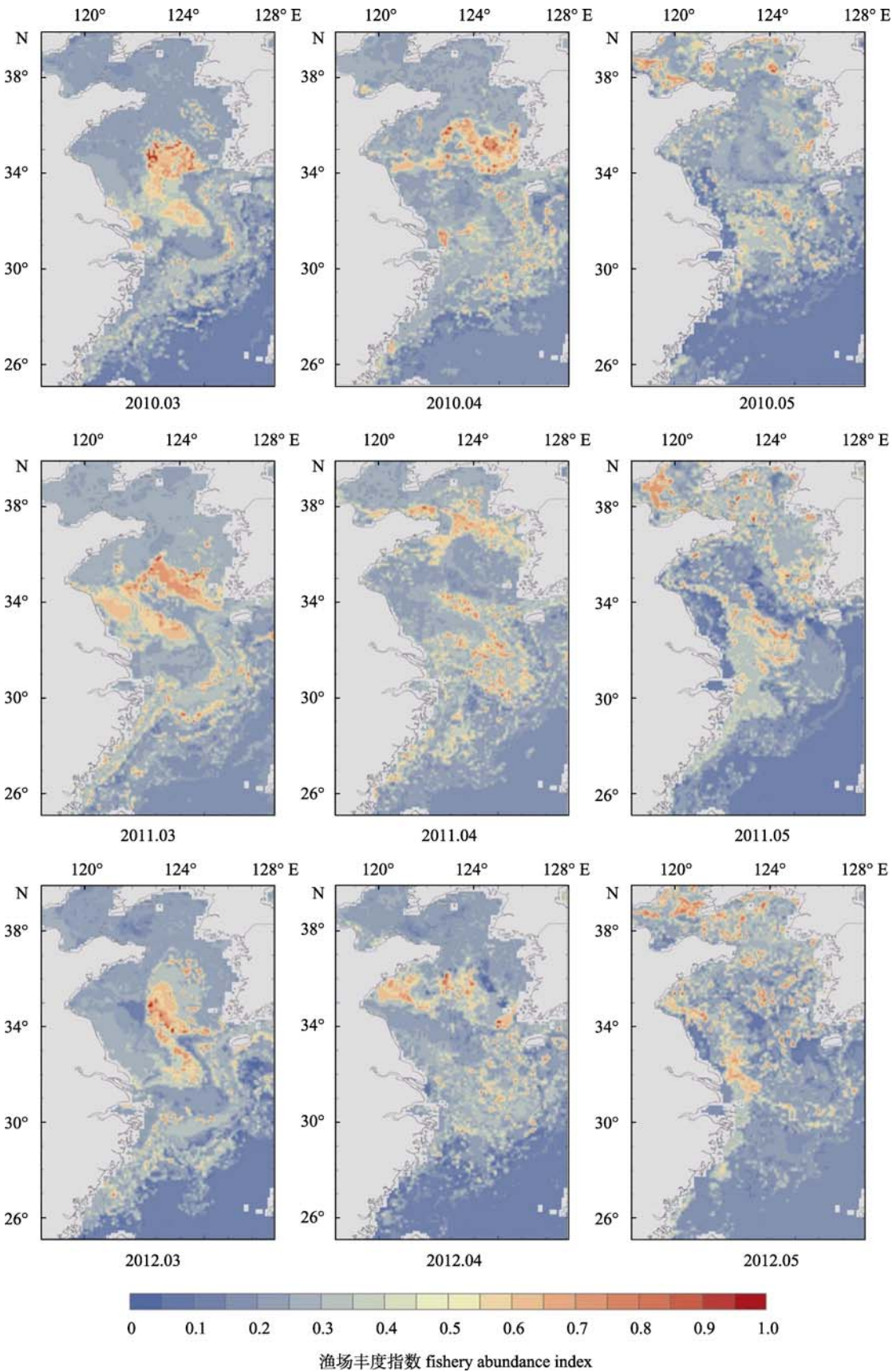
2.3 蓝点马鲛渔场丰度指数年际变动

根据相关性分析发现, 春季渔场丰度指数高值区面积年际变动与 Niño 3.4 指数变动呈显著正相关($r=0.58, P<0.05$)。分析春季高值区面积(3—5 月)、高值区年平均面积, 以及 Niño 3.4 指数的年际变动规律发现(图 5), 2010—2012 年, SST 距平值呈逐渐上升趋势, 渔场丰度指数高值区年平均面积略有增加; 2013 年, 随着 SST 距平值降低, 渔场丰度指数高值区年平均面积随之下降; 2014—2015 年, 当 SST 距平值大幅升高, 渔场丰度指数高值区年平均面积大幅增加。由此发现, 厄尔尼

诺事件能通过水温对蓝点马鲛丰度指数高值区面积的变动产生较大影响。

2.4 未来 RCP 4 种模式的蓝点马鲛 CPUE 预测

在未来 RCP 4 种模式下, 2100 年全球温度将上升 1.0~3.7 °C。因此, 预测 2100 年 SST 升高下的蓝点马鲛 CPUE, 需对蓝点马鲛 CPUE 预测模型进行模型敏感性分析。结果表明, 在 SST 增加 10% 时, 模型敏感性指数为 0.15, 因此, 未来 SST 增加不影响模型预测精度。根据蓝点马鲛 CPUE 预测模型, 分别预测了 2100 年蓝点马鲛在 RCP2.6、RCP4.5、RCP6.0、RCP8.5 4 种情景下的 CPUE 值, 相比于 2010—2015 年春季蓝点马鲛 CPUE 预测值的平均值, 分别上升了 0.1、2.2、2.4 和 17.3 kg/h。其中发现, 在 RCP2.6 情景下, 蓝点马鲛在东海的 CPUE 值明显上升, 27° N 附近海域的 CPUE 值上升最多, 达到 5 kg/h; 黄渤海海域蓝点马鲛 CPUE 值下降。在 RCP4.5 情景下, 蓝点马鲛 CPUE 值上升区域北移至黄海海域, 其中黄海南部 CPUE 值上升最多, 达到 10 kg/h; 东海海域 CPUE 值下降, 渤海海域 CPUE 值未发生变化。在 RCP6.0 时, 蓝点马鲛 CPUE 值上升区域持续北移, 其中黄海中部 CPUE 值上升, 最高达到 12 kg/h; 东海海域蓝点马鲛 CPUE 值下降。在 RCP8.5 情景下, 蓝点马鲛 CPUE 高值区北移至黄海北部以及渤海西部, 最高达到 25 kg/h; 在东海和黄海、南部海域 CPUE 值下降。因此, 长期温度升高使得蓝点马鲛 CPUE 值增加, 且随着温度上升, 蓝点马鲛 CPUE 高值区北移(图 6)。



(续图 4 Fig. 4 continued)

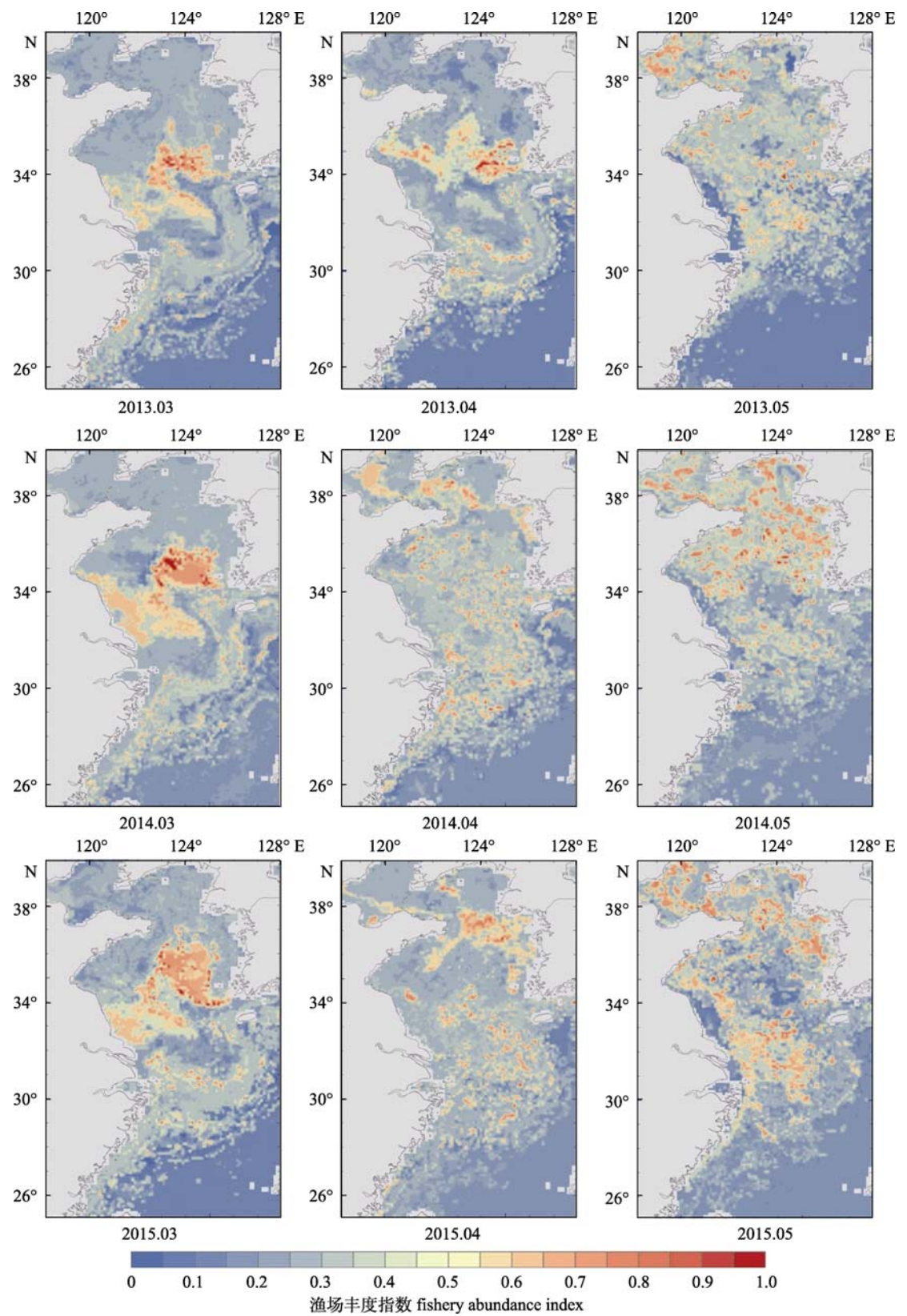


图 4 蓝点马鲛预报渔场丰度指数
Fig. 4 Prediction of *Scomberomorus niphonius* fishery abundance index

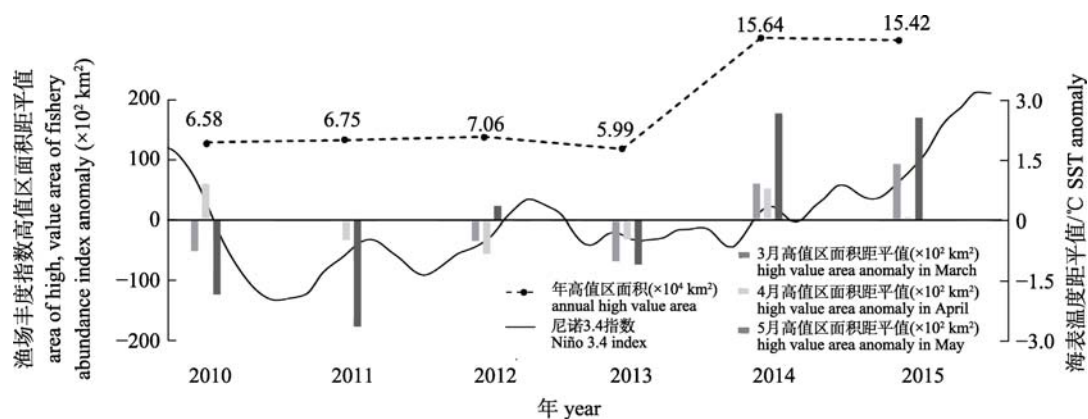


图 5 2010–2015 年 3–5 月蓝点马鲛渔场丰度指数高值区面积距平值、年高值区面积和 Niño3.4 指数

Fig. 5 Area of high value area anomaly of fishery abundance index area of *Scomberomorus niphonius* from March to May 2010–2015, annual high value area and Niño 3.4 index

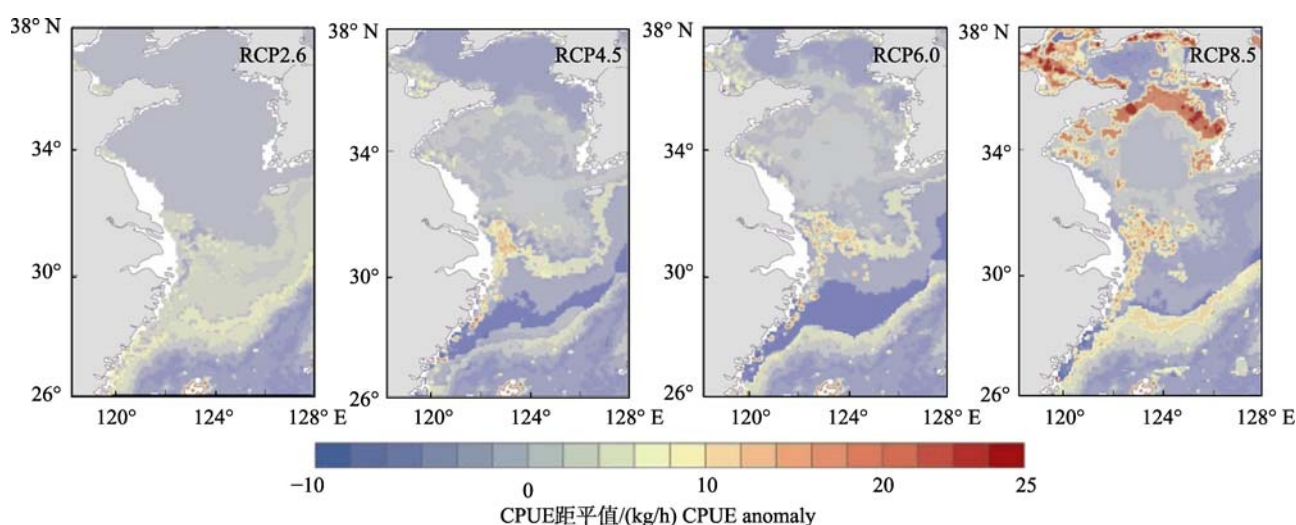


图 6 预测 21 世纪末蓝点马鲛在 RCP2.6、RCP4.5、RCP6.0 和 RCP8.5 情景下 CPUE 距平值

Fig. 6 Predict the anomaly value of CPUE of *Scomberomorus niphonius* under RCP2.6, RCP4.5, RCP6.0, RCP8.5 scenario at the end of the 21st century

3 讨论

3.1 自动机器学习模型与传统预测模型的比较

随机森林、XGBoost 等机器学习模型加入了大量控制参数,使模型拟合精度得以提升,另外通过构建自动机器学习模型的方法,可以进一步提升预测效率。因此,相比传统预测模型具有以下几方面的优势。

精度方面:与传统预测模型相比,自动机器学习模型拟合精度最高($R^2=0.89$)且误差最小(RMSE=15.72),与目前常用的预测模型相比,拟合精度提升了 51%~117%,误差降低了 15%~32%。且传统预测模型在低 CPUE 区预测结果有较大波动(图 3)。

这可能受到极端值的影响,传统预测模型在低值区精度较低,而自动机器学习模型能充分探索渔业数据与环境因子之间的相关性,受到极端值的影响较小,表现更为稳定。

效率方面:传统预测模型需要凭借经验探索参数之间的相关性^[27],这一过程耗时较长,实际研究中效率较低,难以用于实际生产作业。而自动机器学习模型能自动探索各项参数的相互关系,选取最佳参数构建模型,构建速度具有较大优势,将模型构建时间从几个小时缩短至几分钟^[23]。未来可以基于自动机器学习模型,构建自动化渔场预测系统,接收渔船最新作业数据更新预测模型,

结合实时遥感数据用于渔船捕捞作业。

因此在蓝点马鲛渔场预测研究中, 相比于传统预测模型, 自动机器学习模型在稳定性、拟合精度和时效性上均有较大优势。

3.2 环境因子变动对渔场丰度高值区的影响

海洋环境是影响鱼类分布的重要因素之一, 鱼群的分布、洄游均受到海洋环境变化的影响。研究发现, 春汛期间, 蓝点马鲛 CPUE 高值区存在向近岸及向北移动的趋势(图 4)。

这一趋势可能受到 Chl a 的影响, 春汛时蓝点马鲛所需饵料增加, Chl a 高的海域初级生产力较高, 有更充足的饵料供给^[31]。研究发现, 蓝点马鲛渔场丰度指数高值区的 Chl a 随时间上升(图 7, 图 8): 3 月, 蓝点马鲛渔场丰度高值区位于东海海域, 此时 Chl a 为 0~0.7 mg/m³; 4—5 月, 蓝点马鲛渔场丰度高值区向西北移动至黄、渤海海域, 此时 Chl a 为 1.4~2.8 mg/m³。这是受到人为因素的

影响, 近岸营养物质较多且存在春季藻华现象, 适合浮游植物的生长繁殖^[32](图 7), 带动沿岸初级生产力的提高。因此随着蓝点马鲛产卵进入盛期, 渔场丰度高值区向中国沿岸及朝鲜半岛移动。

蓝点马鲛春汛的渔场时空变动也受到 SST 的影响^[3,31]。研究发现, 蓝点马鲛渔场丰度指数高值区北移(图 5)可能与黄、渤海 SST 增加有关(图 8): 3 月, 东海海域蓝点马鲛渔场丰度高值区的 SST 为 18~22 °C (图 7), 与蓝点马鲛产卵最适温度一致^[23], 此时东海 SST 更适宜蓝点马鲛产卵, 最适温度为 19 °C (图 7); 4—5 月, 黄、渤海海域蓝点马鲛渔场丰度指数高值区处的 SST 为 15~20 °C, 适宜蓝点马鲛产卵^[4], 渔场丰度高值区北移至黄、渤海海域(图 5), 此时渔场丰度高值区最适温度为 15 °C (图 7)。因此, 随着 3—5 月 SST 由南向北逐渐升高(图 8), 受到蓝点马鲛产卵时最适水温(15~20 °C)^[33]的影响, 渔场丰度高值区逐渐北移。

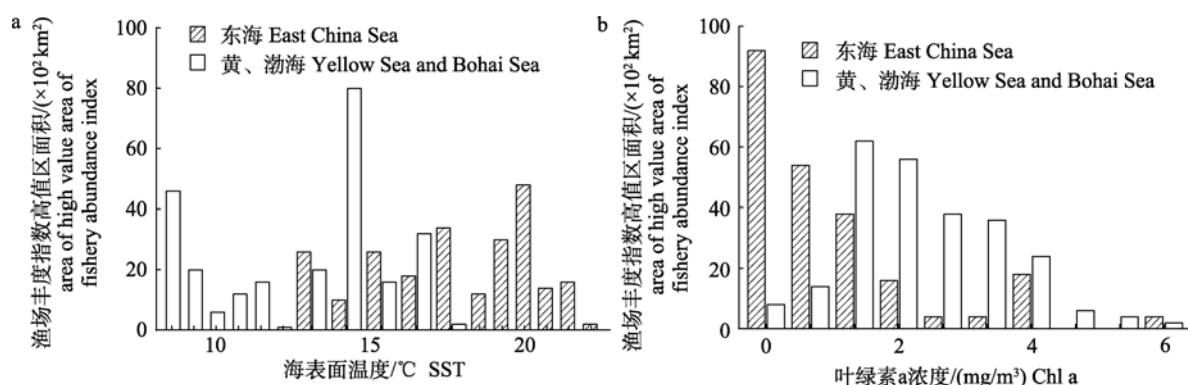


图 7 黄、渤海和东海海域渔场丰度指数高值区(>0.7)海表温度(a)以及叶绿素浓度(b)的分布频率

Fig. 7 Distribution frequency of sea surface temperature (a) and chlorophyll a concentration (b) in the high value area of fishery abundance index (>0.7) in the Yellow Sea, Bohai Sea and East China Sea

3.3 渔场丰度变动与蓝点马鲛生活史的关系

蓝点马鲛在不同生活阶段的分布受到产卵、索饵、越冬等行为的影响^[3]。通过分析渔场丰度指数重心在经度和纬度的年际变动发现, 蓝点马鲛渔场丰度指数高值区重心在每年春季(3—5 月)存在向北及向近岸移动的现象(图 9)。

这一现象可能受蓝点马鲛索饵行为的影响, 鳀是蓝点马鲛春汛时的主要饵料, 在每年的 3—5 月自南向北进行季节性洄游^[32], 与蓝点马鲛渔场丰度指数高值区纬度重心北移一致(图 9); 当 5 月

中下旬鳀向近岸产卵时^[32], 蓝点马鲛渔场丰度指数高值区经度重心同时向近岸移动(图 9)。由此发现, 蓝点马鲛渔场丰度指数高值区的移动规律与鳀的季节性洄游及产卵一致(图 1), 与其索饵行为相关。

蓝点马鲛渔场丰度高值区变动还可能与蓝点马鲛的生殖洄游有关(图 1), 3 月起, 蓝点马鲛随台湾暖流北上^[33], 各渔场渔期自南向北相继开始^[32], 此时渔场丰度指数高值区纬度重心向北(图 9); 5 月, 蓝点马鲛向渤海游动产卵^[32], 此时渔场丰度

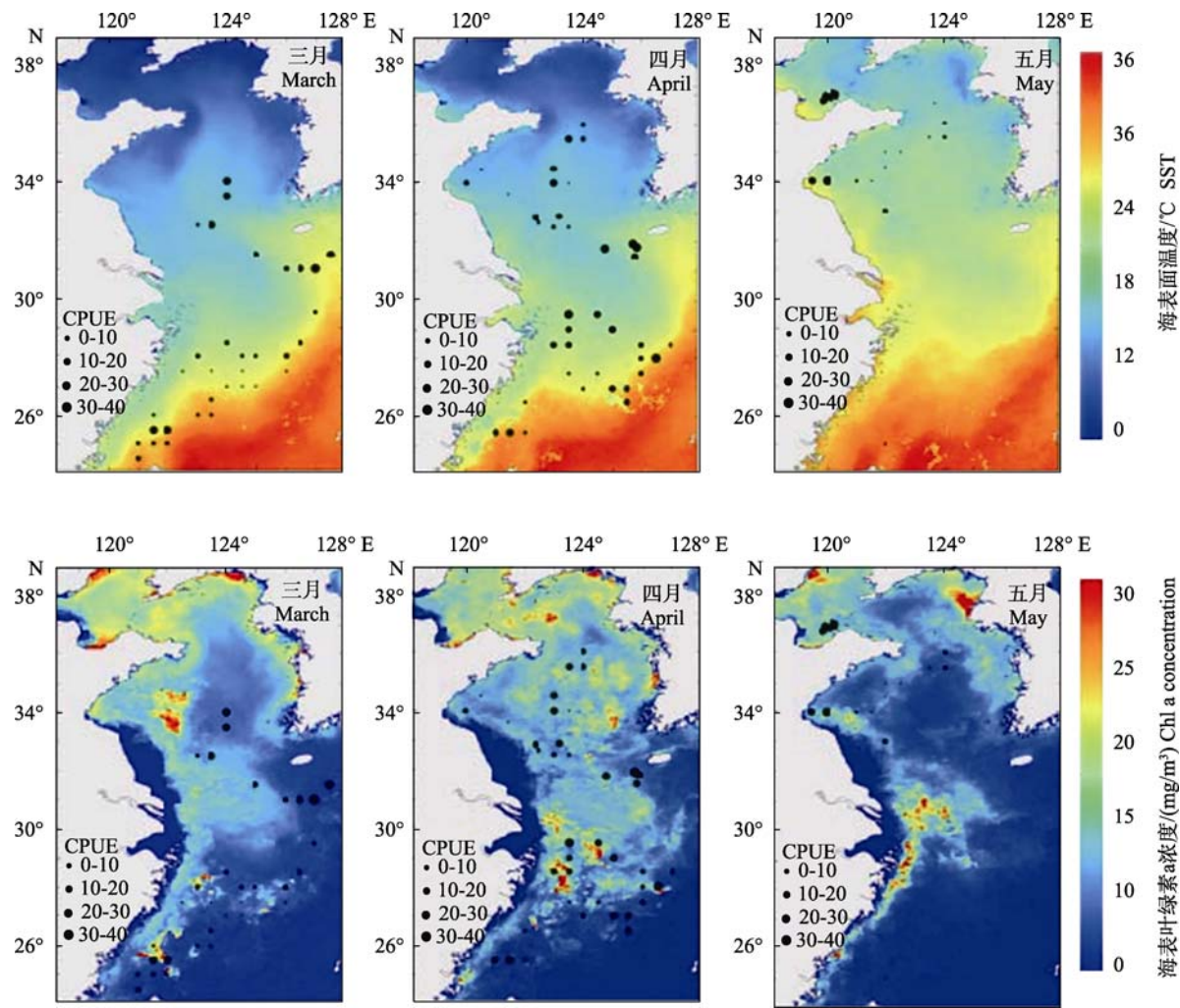


图 8 2010–2015 年 3–5 月卫星遥感海表面温度和叶绿素 a 浓度月累计平均数据及渔船调查点位
Fig. 8 Monthly cumulative averaged sea surface temperature and chlorophyll a concentration based on satellite remote sensing from March to May in 2010–2015 and survey points of fishing vessels

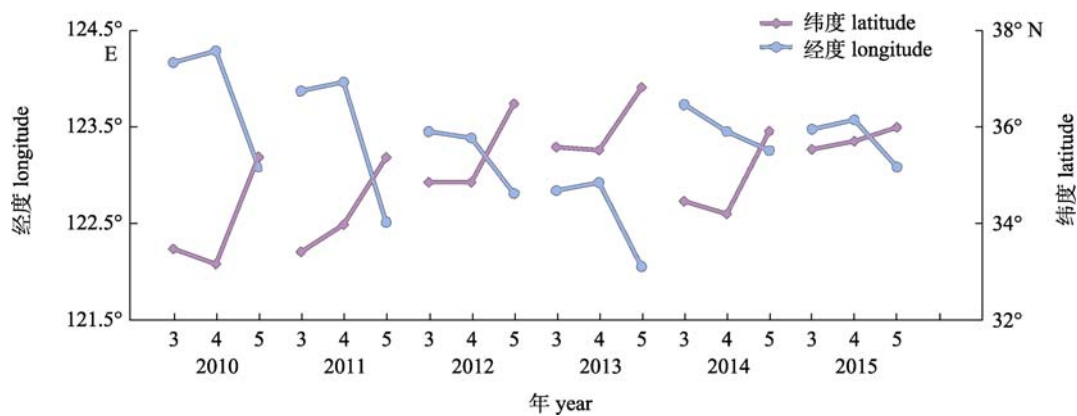


图 9 2010–2015 年 3–5 月渔场丰度指数高值区渔场经纬度重心
Fig. 9 Longitude and latitude center of gravity of fishing ground in high value area of fishery abundance index from March to May in 2010–2015

指数高值区经度重心向西(图9)。由此发现, 蓝点马鲛渔场丰度指数高值区的变化情况与蓝点马鲛生殖洄游路线一致^[31-32]。

3.4 气候变化对渔场丰度高值区的影响

ENSO 能够对中上层鱼类的产卵及索饵海域产生影响^[34]。研究发现, 在 2010—2012 年和 2014—2015 年 SST 距平值增加时期(图5), 蓝点马鲛渔场丰度指数高值区纬度重心也发生北移(图9)。这种现象可能与气候变暖影响蓝点马鲛索饵场变动有关, 有研究表明, 栖息地的纬度^[7]随着 SST 升高, 存在向北迁徙的趋势。蓝点马鲛渔场丰度指数高值区可能受其饵料生物栖息地变动的影响, 纬度重心随着鳀栖息地变动而北移。另外, 温度变动也会影响到蓝点马鲛产卵场的分布, 蓝点马鲛产卵时适宜温度为 15~20 °C^[20], SST 缓慢上升(图5), 导致蓝点马鲛适宜产卵场北移, 使得渔场丰度指数高值区纬度重心呈北移趋势。

蓝点马鲛渔场丰度指数年际变动与 Niño3.4 指数变动呈显著正相关(图5)。研究发现, 气候变动能通过影响 SST 的变化影响蓝点马鲛渔业资源量的变动, 在全球变暖的大背景下, 桡足类的丰度与组成变化可能导致鳀丰度增加^[35], 从而影响蓝点马鲛的资源量。在不同 IPCC-RCP 情景下, 有学者发现 2100 年中国近海鳀的 CPUE 随着气候变暖而增加^[7], 导致蓝点马鲛 CPUE 随其饵料生物资源量一同增加^[19], 该预测结果与本研究预测 2100 年蓝点马鲛 CPUE 随着全球气候变暖而增加的结论一致(图6)。

4 结论

本研究比较了 XGBoost 模型、随机森林模型和 GAM 模型的拟合精度, 选取表现最优的 XGBoost 模型构建了基于自动机器学习的蓝点马鲛渔场丰度指数模型。通过研究预测渔场丰度区域变动发现, 蓝点马鲛渔场丰度高值区呈现向近岸 Chl a 浓度高值区移动的趋势, 并随着黄、渤海 SST 升高逐渐北移, 这与蓝点马鲛索饵和产卵洄游路线一致; 这种现象与 ENSO 影响 SST 进而影响中上层鱼类的产卵场和索饵场的变动有关。结合气候指数分析发现, 蓝点马鲛渔场丰度指数年际变动

与 Niño3.4 指数变动呈显著正相关。另外, 通过预测 RCP2.6、RCP4.5、RCP6.0 和 RCP8.5 4 种情景下蓝点马鲛在 2100 年的 CPUE 发现, 也发生蓝点马鲛高值产区随温度升高继续北移和其 CPUE 不同程度增加的趋势, 本研究结果可对中国近海蓝点马鲛的渔业资源评估与管理提供参考。

致谢: 感谢中国水产科学院东海水产研究所提供的蓝点马鲛渔业数据。

参考文献:

- [1] Qiu S Y, Lyu Z B, Jiao J J, et al. Study on the suitable season for exploitation and utilization of fishery resources in the Bohai Sea[J]. Shandong Fisheries, 2009, 26(3): 18-20. [邱盛尧, 吕振波, 焦金菊, 等. 黄渤海渔业资源适宜开发利用季节的研究[J]. 齐鲁渔业, 2009, 26(3): 18-20.]
- [2] Shui B N, Sun X F, Han Z Q, et al. Morphological variation analysis among populations of Japanese Spanish mackerel *Scomberomorus niphonius* in the Yellow Sea and East China Sea[J]. Journal of Fisheries of China, 2009, 33(3): 445-451. [水柏年, 孙希福, 韩志强, 等. 黄海、东海蓝点马鲛群体的形态特征分析[J]. 水产学报, 2009, 33(3): 445-451.]
- [3] Qiu S Y, Ye M Z. Studies on the reproductive biology of *Scomberomorus niphonius* in the Yellow Sea and Bohai Sea[J]. Oceanologia et Limnologia Sinica, 1996, 27(5): 463-470, 570. [邱盛尧, 叶懋中. 黄渤海蓝点马鲛繁殖生物学的研究[J]. 海洋与湖泊, 1996, 27(5): 463-470, 570.]
- [4] Kim W J, Kim K K, Han H S, et al. Population structure of the olive flounder (*Paralichthys olivaceus*) in Korea inferred from microsatellite marker analysis[J]. Journal of Fish Biology, 2010, 76(8): 1958-1971.
- [5] Gonzalez E B, Aritaki M, Sakurai S, et al. Inference of potential genetic risks associated with large-scale releases of red sea bream in kanagawa prefecture, Japan based on nuclear and mitochondrial DNA analysis[J]. Marine Biotechnology, 2013, 15(2): 206-220.
- [6] Fisheries Bureau of the Ministry of Agriculture, Yellow Sea Fisheries Headquarters of the Ministry of Agriculture. Investigation and Regionalization of Fishery Resources in the Yellow Sea and Bohai Sea[M]. Beijing: Ocean Press, 1990 [农业部水产局, 农业部黄海区渔业指挥部. 黄渤海区渔业资源调查与区划[M]. 北京: 海洋出版社, 1990.]
- [7] Liu S X, Han S X, Wei Y K. Study on the application of discriminant analysis in fishery forecast[J]. Marine Science Bulletin, 1988, 7(1): 63-70. [刘树勋, 韩士鑫, 魏永康. 判别分析在渔情预报中应用的研究[J]. 海洋通报, 1988, 7(1): 63-70.]
- [8] Zhang M H. Studies on the methods of forecasting fishing conditions for *Acetes chinensis* near the Yellow River mouth [J]. Transaction of Oceanology and Limnology, 1986(4): 68-

76. [张孟海. 黄河口附近海区毛虾渔情预报方法的研究[J]. 海洋湖沼通报, 1986(4): 68-76.]
- [9] Lan K W, Shimada T, Lee M A, et al. Using remote-sensing environmental and fishery data to map potential yellowfin tuna habitats in the tropical Pacific Ocean[J]. Remote Sensing, 2017, 9(5): 444.
- [10] Niu M X, Li X S, Xu Y C. Effects of spatiotemporal and environmental factors on the fishing ground of *Trachurus murphyi* in Southeast Pacific Ocean based on generalized additive model[J]. Chinese Journal of Applied Ecology, 2010, 21(4): 1049-1055. [牛明香, 李显森, 徐玉成. 基于广义可加模型的时空和环境因子对东南太平洋智利竹筴鱼渔场的影响[J]. 应用生态学报, 2010, 21(4): 1049-1055.]
- [11] De'ath G, Fabricius K E. Classification and regression trees: A powerful yet simple technique for ecological data analysis[J]. Ecology, 2000, 81(11): 3178-3192.
- [12] Austin M. Species distribution models and ecological theory: A critical assessment and some possible new approaches[J]. Ecological Modelling, 2007, 200(1-2): 1-19.
- [13] Wang J T, Chen X J, Chen Y. Projecting distributions of Argentine shortfin squid (*Illex argentinus*) in the Southwest Atlantic using a complex integrated model[J]. Acta Oceanologica Sinica, 2018, 37(8): 31-37.
- [14] Mugo R, Saitoh S I. Ensemble modelling of skipjack tuna (*Katsuwonus pelamis*) habitats in the western north Pacific using satellite remotely sensed data; a comparative analysis using machine-learning models[J]. Remote Sensing, 2020, 12(16): 2591.
- [15] Yao Q M, Wang M S, Hugo J E, et al. Taking human out of learning applications: A survey on automated machine learning[EB/OL]. 2018: arXiv: 1810.13306. <https://arxiv.org/abs/1810.13306>
- [16] Barber R T, Chávez F P. Ocean variability in relation to living resources during the 1982-83 El Niño[J]. Nature, 1986, 319(6051): 279-285.
- [17] Tian Y J, Akamine T, Suda M. Modeling the influence of oceanic-climatic changes on the dynamics of Pacific saury in the northwestern Pacific using a life cycle model[J]. Fisheries Oceanography, Supplement, 2004, 13(1): 125-137.
- [18] Kim S, Kang S. Ecological variations and El Niño effects off the southern coast of the Korean Peninsula during the last three decades[J]. Fisheries Oceanography, 2000, 9(3): 239-247.
- [19] Liu S H, Liu Y, Alabia I D, et al. Impact of climate change on wintering ground of Japanese anchovy (*Engraulis japonicus*) using marine geospatial statistics[J]. Frontiers in Marine Science, 2020, 7: 604.
- [20] Wei S, Zhou B B. Study on short-term fishing situation forecast of mackerel in Yellow Sea and Bohai Sea[J]. Acta Oceanologica Sinica, 1988, 10(2): 216-221. [韦晟, 周彬彬. 黄渤海蓝点马鲛短期渔情预报的研究[J]. 海洋学报, 1988, 10(2): 216-221.]
- [21] Yuan Y Y, Ye Z J, Liu Q, et al. Fishery oceanography and spatial temporal distribution of *Scomberomorus niphonius* in spring in southern Yellow Sea[J]. Oceanologia et Limnologia Sinica, 2009, (4): 506-510. [袁杨洋, 叶振江, 刘群, 等. 黄海南部春季蓝点马鲛(*Scomberomorus niphonius*)渔场海洋学特征与渔期渔场分布[J]. 海洋与湖沼, 2009, 40(4): 506-510.]
- [22] Jiang Y Q, Fan Y N, Zheng C J, et al. The effect of temperature on embryonic development of *Scomberomorus niphonius*[J]. Journal of Zhejiang Ocean University (Natural Science), 2016, 35(4): 271-275. [姜屹倩, 樊艳楠, 郑春静, 等. 温度对蓝点马鲛胚胎发育的影响[J]. 浙江海洋学院学报(自然科学版), 2016, 35(4): 271-275.]
- [23] Song C, Wang Y T, Liu Z L, et al. Relationship between environmental factors and distribution of *Scomberomorus niphonius* eggs, larvae, and juveniles in Xiangshan Bay[J]. Journal of Fishery Sciences of China, 2016, 23(5): 1197-1204. [宋超, 王宇坛, 刘尊雷, 等. 象山港蓝点马鲛鱼卵、仔稚鱼的时空分布特征及其与环境因子关系[J]. 中国水产科学, 2016, 23(5): 1197-1204.]
- [24] Barber R T, Chávez F P. Ocean variability in relation to living resources during the 1982-83 El Niño[J]. Nature, 1986, 319(6051): 279-285.
- [25] Allen M R, Barros V R, Broome J, et al. Climate change 2014 synthesis report: summary for policymakers[R]. Mitigation of Climate Change. Cambridge, United Kingdom and New York, USA: the Intergovernmental Panel on Climate Change 2014.
- [26] Liu R. The monthly and interannual variability in distribution of fishing grounds of the Spanish mackerel, *Scomberomorus niphonius* (Cuvier et Valenciennes) in Bohai Sea, Yellow Sea and East China Sea[D]. Qingdao: Ocean University of China, 2009. [刘蕊. 渤、黄、东海蓝点马鲛渔场分布的逐月与年间变化[D]. 青岛: 中国海洋大学, 2009.]
- [27] Xu B X. The effect analysis of commercial bank telemarketing based on optimized xgboost model[D]. Lanzhou: Lanzhou University, 2017. [徐彬心. 基于优化的 xgboost 模型的商业银行电话营销效果分析[D]. 兰州: 兰州大学, 2017.]
- [28] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]//Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2. New York: ACM, 1995: 1137-1143.
- [29] Criminisi A. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning[J]. Foundations and Trends® in Computer Graphics and Vision, 2011, 7(2-3): 81-227.
- [30] He T M. Building A dynamic energy budget growth model for cultured *Meretrix meretrix* in shallow tidal flat based on STELLA[D]. Shanghai: Shanghai Ocean University, 2020. [何田妹. 基于 STELLA 的浅海滩涂养殖文蛤(*Meretrix meretrix*)动态能量收支生长模型构建[D]. 上海: 上海海洋大学, 2020.]
- [31] Akbar H I, Supriatna S, Patria M P. The relationship between total suspended matter and sea current for the spread of Chlorophyll-a and fishing area in Jakarta Bay[J]. IOP Confe-

- rence Series: Earth and Environmental Science, 2020, 561(1): 012052.
- [32] Guan W J. Remote-sensing-based assessment of chub mackerel (*Scomber japonicus*) fishing ground and stock dynamics in the East China Sea and Yellow Sea[D]. Shanghai: East China Normal University, 2008. [官文江. 基于海洋遥感的东、黄海鲈鱼渔场与资源研究[D]. 上海: 华东师范大学, 2008.]
- [33] Bi S W, Yang X F. Investigation on fishing season of Spanish mackerel drift-net fishing ground in the Yellow River and Bohai Sea in spring[J]. Acta Oceanologica Sinica, 1986, 8(5): 657-662. [毕庶万, 杨雪舫. 春汛黄、渤海鲈鱼流网渔场渔期调查研究[J]. 海洋学报, 1986, 8(5): 657-662.]
- [34] Zhou X, Sun Y, Huang W, et al. The Pacific decadal oscillation and changes in anchovy populations in the Northwest Pacific[J]. Journal of Asian Earth Sciences, 2015, 114: 504- 511.
- [35] Swartz W, Sala E, Tracey S, et al. The spatial expansion and ecological footprint of fisheries (1950 to present)[J]. PLoS ONE, 2010, 5(12): e15143.

Fishery abundance index of *Scomberomorus niphonius* in the Yellow Sea, Bohai Sea, and East China Sea based on automatic machine learning

QIN Yuanzhe¹, ZHOU Zhenjia¹, LIU Yang¹, TIAN Yongjun¹, CHENG Jiahua², LIU Xudong³, ZHANG Yong³

1. Fisheries Oceanography Laboratory, Fisheries College, Ocean University of China, Qingdao 266003, China;
2. East China Sea Fishery Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090, China;
3. Shanxi Dadi Xinya Technology Co., Ltd., Taiyuan 030021, China

Abstract: The *Scomberomorus niphonius* is an economically important large-scale pelagic fish in the coastal waters of China. The accurate prediction of the fishing distribution of *S. niphonius* is of great significance for the assessment and management of fishery resources. The development of satellite remote sensing and artificial intelligence has provided strong technical support and convenience for fishery forecasting research. In this study, a *S. niphonius* CPUE prediction model based on automatic machine learning was constructed using *S. niphonius* fishing data, satellite remote sensing sea surface temperature (SST), and sea surface chlorophyll-a concentration (Chl-a) data. Through comparison with the XGBoost, random forest, and generalized additive models (GAM), the coefficient of determination (R^2) of the automatic machine learning model increased by 51%, 107%, and 117%, respectively, and the root mean square error (RMSE) was reduced by 15%, 28%, and 32%, respectively. A fishery abundance index was developed through the CPUE predicted by the model, and the temporal and spatial variation laws of fishery abundance were analyzed. The results showed that the high-abundance area of *S. niphonius* presented a trend in moving northward and nearshore because of the influence of SST in spring, which is consistent with the feeding and spawning migration route of *S. niphonius*. The northward shift of the latitudinal center of gravity in the high-abundance area for the *S. niphonius* fishery was related to the change in the suitable temperature area for feeding and spawning of the *S. niphonius* under the influence of climate warming. Through the analysis of the climate event index, it was found that the Niño 3.4 index was significantly correlated with the area of high abundance of *S. niphonius* fishery ($r=0.58$, $P<0.05$). Combined with RCP2.6, RCP4.5, RCP6.0, and RCP8.5 for the four scenarios, the distribution changes in the CPUE of *S. niphonius* by 2100 were predicted. It was determined that with global warming, the high-value production area of CPUE of *S. niphonius* will continue to move northward, and increase by 0.1 kg/h, 2.2 kg/h, 2.4 kg/h, and 17.3 kg/h, respectively, compared with the predicted average value of CPUE of *S. niphonius* from 2010 to 2015. The results of this study provide a reference for predicting the fishing conditions of economically important fish species in the coastal waters of China.

Key words: *Scomberomorus niphonius*; fishery forecast; machine learning; climate change; fishery abundance index

Corresponding author: LIU Yang. E-mail: Yangliu315@ouc.edu.cn