

彭伟图, 陈峰, 李建雄, 等. 基于随机森林模型的秘鲁外海茎柔鱼资源丰度与环境因子关系研究[J]. 海洋学报, 2024, 46(7): 41–50, doi:10.12284/hyxb2024078

Peng Weituo, Chen Feng, Li Jianxiong, et al. Relationship between abundance of *Dosidicus gigas* and environment factors in the offshore waters of Peru based on random forest model[J]. Haiyang Xuebao, 2024, 46(7): 41–50, doi:10.12284/hyxb2024078

基于随机森林模型的秘鲁外海茎柔鱼资源丰度与环境因子关系研究

彭伟图^{1,2}, 陈峰², 李建雄^{1,2}, 赵艺翔^{1,2}, 董钊江^{1,2}, 戴乾², 徐汉祥^{1,2}, 朱文斌^{2*}

(1. 浙江海洋大学海洋与渔业研究所, 浙江舟山 316021; 2. 浙江省水产研究所/农村农业部重点渔场渔业资源科学观测实验站/浙江省海洋渔业资源可持续利用技术研究重点实验室, 浙江舟山 316021)

摘要: 秘鲁外海茎柔鱼是大洋性经济鱼类, 其资源丰度受环境因素变化的影响较大。研究基于 2018–2021 年 9–12 月秘鲁外海茎柔鱼的渔捞日志数据, 并结合卫星遥感获取海表温度 (SST)、海表盐度 (SSS)、海表高度 (SSH) 和叶绿素 *a* 浓度 (Chl *a*) 4 个环境数据, 运用 ArcGIS 和随机森林模型分析秘鲁外海茎柔鱼资源丰度与环境因子之间的关系。研究表明, 2018–2021 年 9–12 月茎柔鱼渔场重心分布范围集中于 13°~21°S, 76°~87°W 海域, 9–12 月渔场重心呈现西北向东南方向移动趋势。随机森林模型分析结果显示, 各月环境因子对秘鲁外海茎柔鱼资源丰度和分布的影响存在一定的差异性, 最适 SST 范围 16.3~18.5°C, 最适 SSS 范围 35.1~35.4, 最适 SSH 范围 0.55~0.60 m, 最适 Chl *a* 浓度范围 0.18~0.46 mg/m³, 通过随机森林模型的 CPUE 预测值与名义 CPUE 随时间变化趋势基本一致, 表明随机森林模型适用于分析秘鲁外海茎柔鱼与环境因子之间的关系。本研究对了解该海域茎柔鱼资源变动规律、指导茎柔鱼生产具有重要意义。

关键词: 茎柔鱼; 随机森林模型; 资源丰度; 环境因子

中图分类号: P714⁺.5

文献标志码: A

文章编号: 0253-4193(2024)07-0041-10

1 引言

茎柔鱼 (*Dosidicus gigas*) 隶属于头足纲 (Cephalopoda), 枪形目 (Teuthoidea), 柔鱼科 (Ommastrephidae), 茎柔鱼属 (*Dosidicus*)^[1]。广泛分布于加利福利亚北部至智利南部的东太平洋地区^[2]。茎柔鱼资源丰富, 分布范围也较广泛, 营养价值较高, 是远洋鱿钓渔业的主要捕捞对象^[3], 了解茎柔鱼资源时空变动对茎柔鱼资源可持续利用与保护起着重要作用。茎柔鱼生命周期短, 生长快, 其资源量变动与环境变化密

切相关^[4]。在渔场资源分布的研究中, 资源丰度与环境因子之间的关系一直备受关注^[5], 历史研究发现茎柔鱼资源丰度主要受到海表温度 (Sea Surface Temperature, SST)、海表盐度 (Sea Surface Salinity, SSS)、海表高度 (Sea Surface Height, SSH)、叶绿素 *a* 浓度 (Chlorophyll *a*, Chl *a*) 等环境因子影响^[6-9]。在以往的研究中, 广义线性模型 (Generalized Linear Model, GLM)、广义可加模型 (Generalized Additive Models, GAM)、栖息地模型、灰色系统、协同克里金法等广泛用于渔场资源分布与环境因子之间的关系分析^[10-13], 但模型结果

收稿日期: 2023-12-05; 修订日期: 2024-04-01。

基金项目: 国家重点研发计划项目 (2019YFD0901505); 浙江省重点研发计划项目 (2018C02026); 2022–2023 浙江省远洋渔业资源常规检测 (2023HZ024)

作者简介: 彭伟图 (1997—), 男, 湖南省娄底市人, 主要研究领域为渔业资源。E-mail: 1367101299@qq.com

* 通信作者: 朱文斌 (1982—), 男, 浙江省义乌市人, 正高级工程师, 主要从事渔业资源与管理研究。E-mail: foolse@126.com

容易受异常值和冗余数据的影响,模型对渔业调查数据中的零值问题的处理也存在不足,选择合适的模型方法对茎柔鱼资源丰度与环境因子的关系进行研究至关重要。

随机森林模型(Random Forest, RF)可用于回归和分类,是以一定数量的决策树为基础,根据相应的准则对随机生成的决策树进行组合生成随机森林,由于随机森林模型是通过随机构成的决策树组成,因此可以用来模拟多重非线性关系^[14]。随机森林模型从提出开始,就被许多学者在各个领域使用。随机森林引入了随机性概念,随机抽取训练样本,随机选取特征子集,有效提高了分类能力和抗噪能力,使随机森林不容易陷入过度拟合^[15]。采用随机森林模型进行回归分析时,受到的离群值影响较小,在随机干扰较多的情况下表现较为稳定^[16]。随机森林算法可以模拟自变量之间的复杂交互作用,且对异常值的容忍度较强^[17],可有效降低环境因子交互性对模型结果的影响,提高模型分析的准确率。本文根据中国远洋鱿钓船2018–2021年9–12月在秘鲁外海的渔业生产统计数据 and 卫星遥感数据,采用随机森林模型对秘鲁外海茎柔鱼丰度与环境因子的关系进行研究,结果可为秘鲁外海茎柔鱼资源可持续开发和利用提供科学依据。

2 数据和方法

2.1 数据来源

渔业数据来源于中国2018–2021年9–12月秘鲁外海鱿钓渔船商业捕捞渔捞日志,渔捞日志记录数据

包括捕捞日期、经度、纬度、手钓人数、手钓渔获量、钓机数量、机钓渔获量等。图1为2018–2021年9–12月鱿钓船作业区域,研究区域位于东南太平洋8°~22°S, 75°~95°W海域。

SSS、SST、SSH数据来源于哥白尼海事服务网站(<https://data.marine.copernicus.eu>),时间分辨率为月,空间分辨率为 $0.25^{\circ} \times 0.25^{\circ}$; Chl *a*数据来自美国国家海洋和大气管理局(<https://oceanwatch.pifsc.noaa.gov/>),时间分辨率为月,空间分辨率为 $0.25^{\circ} \times 0.25^{\circ}$ 。图2为SST、SSS、SSH、Chl *a*月平均值的空间分布。

由于中国远洋鱿钓船的作业方式和作业时间相同,单位努力捕捞渔获量(CPUE)可以作为一种可靠的渔场柔鱼丰富度指标^[18]。渔业生产数据均为点数数据,遥感数据空间分辨率为 $0.25^{\circ} \times 0.25^{\circ}$,因此需要统一匹配生产数据和环境数据,定义 $0.25^{\circ} \times 0.25^{\circ}$ 为1个渔区,生产统计数据按经纬度 $0.25^{\circ} \times 0.25^{\circ}$ 空间分辨率进行汇总统计,并按月进行处理。按每个月计算每个渔区内的CPUE,计算公式如下:

$$CPUE = \frac{\sum C_i}{\sum F_i}, \quad (1)$$

式中, $\sum C_i$ 表示1个月内,1个 $0.25^{\circ} \times 0.25^{\circ}$ 渔区内的总产量; $\sum F_i$ 表示1个 $0.25^{\circ} \times 0.25^{\circ}$ 渔区内所有船舶作业天数。

2.2 研究方法

2.2.1 模型构建

随机森林是使用重抽样(Bootstrap)方法从原始训练样本中随机抽取多组训练样本,并对抽取的多组样本分别建立一个决策树的随机森林模型,然后综合

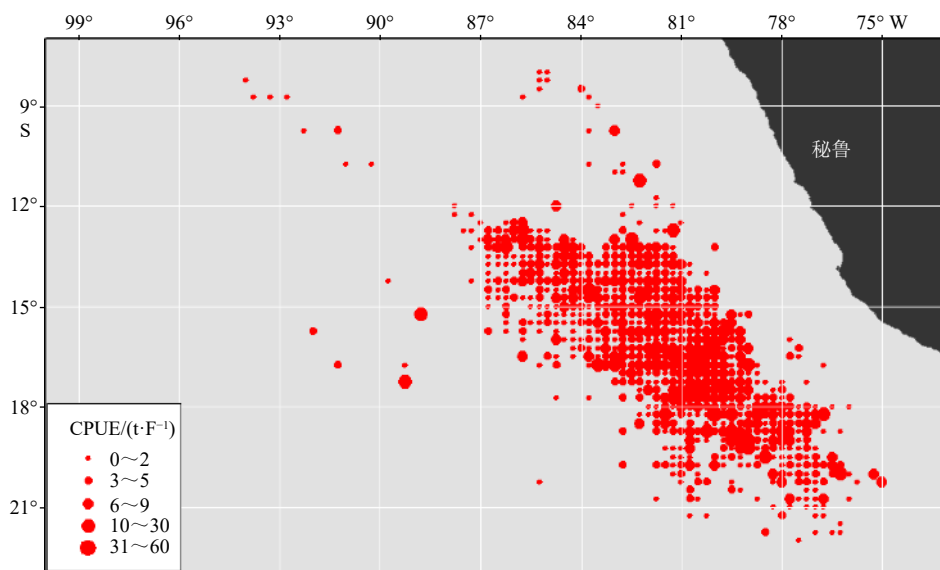


图1 2018–2021年9–12月茎柔鱼CPUE日平均分布

Fig. 1 The average daily distribution of CPUE of *Dosidicus gigas* from September to December in 2018–2021

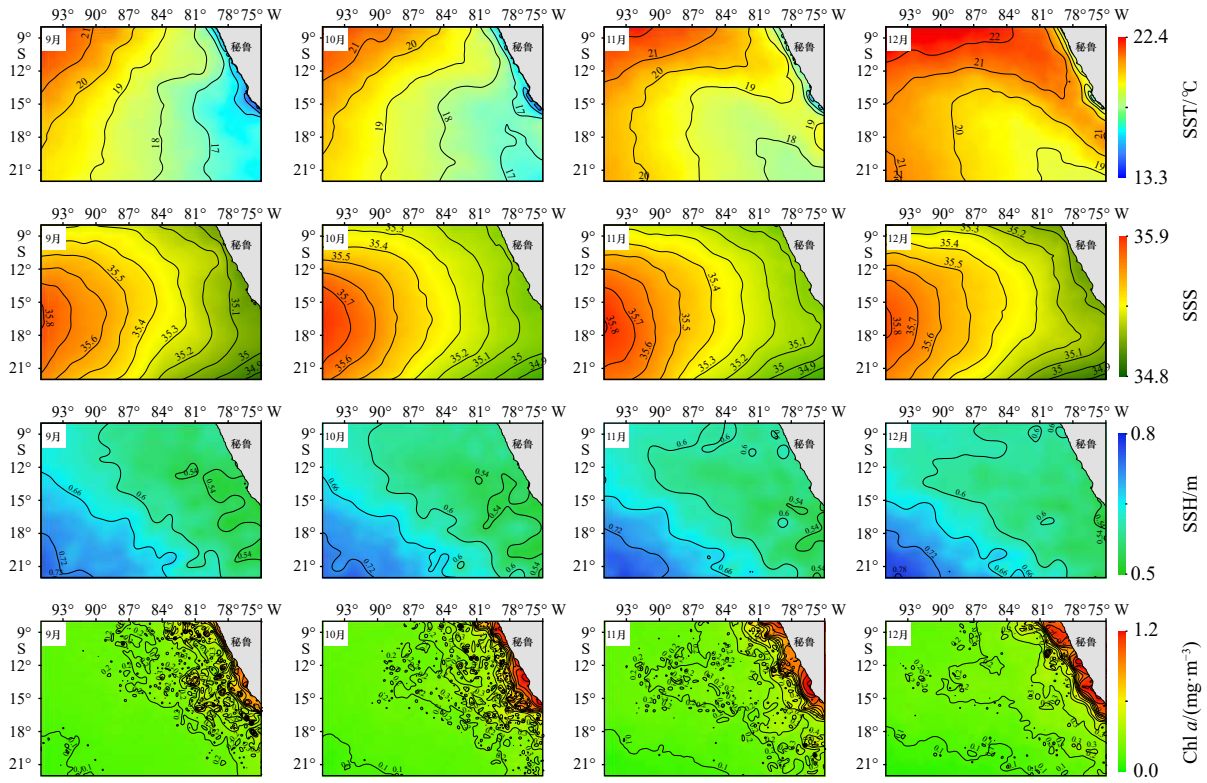


图2 9-12月秘鲁外海环境因子月平均值空间分布

Fig. 2 The spatial distribution map of monthly average environmental factors in offshore waters of Peru from September to December

多个决策树的预测,采用均值或者投票的方式得出最终的预测结果^[19]。选取的测试集评估每个特征的重要性,以验证评估模型的预测性能。每个特征的重要性采用特征重要性得分 (Importance Score, IS) 来表示。在随机森林中选择特定变量作为分裂节点降低错误率,将每棵树的特征重要性得分进行平均,得到最终模型中每个特征的重要性得分,将特征重要性得分进行规格化,使其落在(0,1)的范围内,且所有特征重要性得分之和为1。特征重要性得分的计算方法如下^[20]:

$$IS = \sum \frac{(T_1 - T_2)}{N}, \quad (2)$$

式中, N 为训练样本数量; T_1 为随机森林模型中每个决策树的袋外数据误差; T_2 是所有袋外数据样本的特征添加随机噪声干扰后重新计算误差。

随机森林的构建与模拟过程是使用 Python 3.12.0 scikit-learn 模块进行,该过程主要包括训练样本的选择、随机森林结构创建、输入参数的确定、训练和模拟等过程。将2018-2021年9-12月秘鲁外海 SST、SSS、SSH、Chl a 值作为模型的输入变量,同时期鱿钓作业 CPUE 为输出变量,构建随机森林模型。

2.2.2 回归评价指标

方均根偏移 (RMSD) 是预测值与实际值偏差的平

方和与观测次数 n 比值的平方根,能够体现数据集的离散程度^[19],决定系数 (R^2) 是自变量 X 引起的平方和在因变量 Y 的总平方和中所占的比例^[21]。本研究从数据样本随机抽取 70% 的数据作为训练集进行建模,将剩下的 30% 作为测试集进行验证,将该过程重复 100 次,得出 RMSD 和 R^2 来判断模型的预测效果。RMSD 和决定系数 R^2 计算公式如下:

$$RMSD(y, p) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2}, \quad (3)$$

$$R^2(y, p) = 1 - \frac{\sum_{i=0}^n (y_i - p_i)^2}{\sum_{i=0}^n (y_i - \bar{y})^2}, \quad (4)$$

式中, y 为原来的值; p 为预测值; n 为样本数。

RMSD 值越小,模型预测的准确性越高, R^2 越接近 1,模型的参考价值越高, R^2 越接近 0,模型的参考价值越低^[22]。

2.2.3 核密度估计法分析

本研究基于秘鲁外海茎柔鱼丰度数据,运用核函数^[23]进行分析秘鲁外海 9-12 月茎柔鱼相对资源丰度的时空分布情况,计算公式如下:

$$f(x) = \frac{1}{nb^d} \sum_{i=1}^n K \left[\frac{1}{b} (x - x_i) \right], \quad (5)$$

式中, K 是核函数; b 为带宽; n 是在带宽范围内的已知点数目; d 是数据的维度。

3 结果与分析

3.1 环境因子的相关系数

通过构建 Seaborn 相关性热力矩阵研究年份 (Year)、月份 (Month)、经度 (Lon)、纬度 (Lat)、SSS、SST、SSH、Chl a 等各变量之间的相关性。图 3 下三角为显著性水平 (*表示 $p < 0.05$; **表示 $p < 0.01$; ***表示 $p < 0.001$), 方框大小与相关系数的绝对值成正比, 上三角为相关系数。SSS 与 SST、SSH、Chl a 相关系数分别为 0.31、0.22、-0.11; SST 与 SSH、Chl a 相关系数分别为 0.32、-0.17; SSH 与 Chl a 相关系数分别为 -0.25。

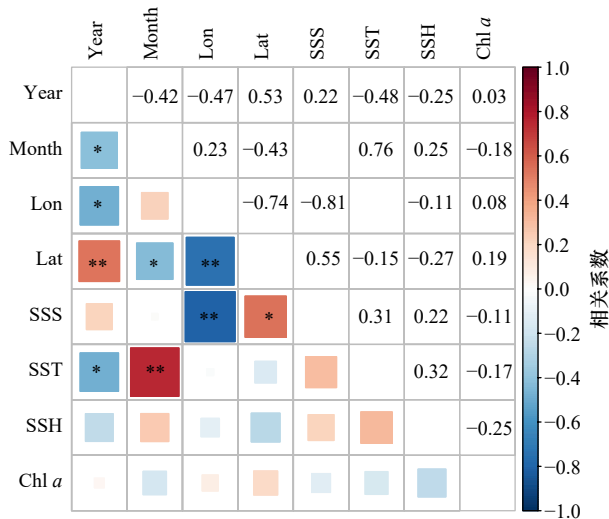


图 3 各因子之间的相关系数
Fig. 3 Correlation coefficient between factors

3.2 模型的性能评估

本研究选取了 70% 数据作为训练集, 30% 数据作为测试集, 通过交叉验证得到 10 次预测结果。相关系数 (correlation coefficient) 反映了预测值和实际值之间的线性关系强度, 可以看出 9-12 月的相关系数取值范围在 0.75~0.88 之间, 9-12 月的 RMSD 范围在 0.54~0.71 之间, 9-12 月份 R^2 的取值范围在 0.54~0.75 之间 (图 4)。

3.3 茎柔鱼资源分布

核密度模型计算结果显示 9-12 月秘鲁外海茎柔鱼资源分布呈现一定的聚集性, 渔场主要分布在 $13^{\circ}\sim 21^{\circ}\text{S}$, $76^{\circ}\sim 87^{\circ}\text{W}$ 区域。各月份茎柔鱼资源分布又存在一定的差异, 资源核密度值随着月份的增加呈现先增大后减小的趋势, 10 月份最高, 核密度最大值

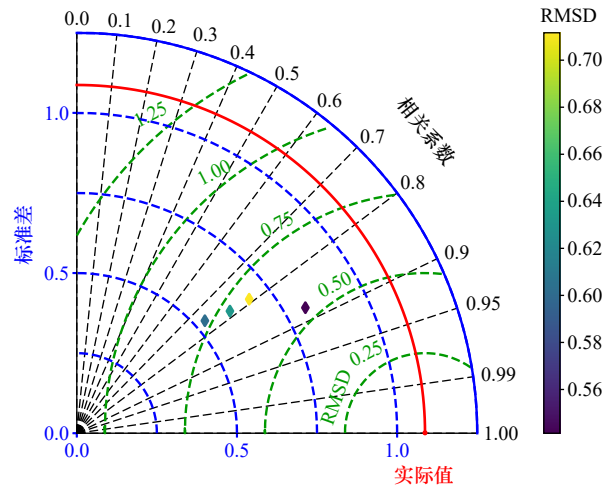


图 4 9-12 月模型交叉验证结果泰勒图
Fig. 4 Taylor plot of model cross-validation results from September to December

大于 40 t/m^2 。其中 9 月份茎柔鱼资源核密度高值区域主要集中在 $13^{\circ}\sim 16^{\circ}\text{S}$, $81^{\circ}\sim 83^{\circ}\text{W}$ 。10 月份逐渐向东南方向迁移, 主要集中在 $16^{\circ}\sim 18^{\circ}\text{S}$, $79^{\circ}\sim 82^{\circ}\text{W}$ 。11 月茎柔鱼资源核密度呈西北-东南分布, 集中在 $13^{\circ}\sim 15^{\circ}\text{S}$, $83^{\circ}\sim 87^{\circ}\text{W}$ 与 $13^{\circ}\sim 19^{\circ}\text{S}$, $77^{\circ}\sim 82^{\circ}\text{W}$ 海域。12 月茎柔鱼资源核密度呈西北-东南分布, 集中在 $16^{\circ}\sim 18^{\circ}\text{S}$, $81^{\circ}\sim 82^{\circ}\text{W}$ 与 $18^{\circ}\sim 21^{\circ}\text{S}$, $77^{\circ}\sim 79^{\circ}\text{W}$ 海域。12 月与 11 月份相比, 茎柔鱼核密度高值继续向东南方向迁移 (图 5)。

3.4 模型结果

在随机森林中, 特征的贡献率通常基于特征在决策树中的节点分裂次数和分裂所获得的信息增益来计算。随机森林不同月份各环境因子对 CPUE 贡献率如图 6 所示, 各月份各环境因子贡献率存在差异, 9-11 月环境因子贡献率最高的均为 SST, 12 月环境因子贡献率与前 3 个月存在显著差异, 贡献率最高的为 Chl a 。

由图 7 可知, 模型 CPUE 预测值随 SSS 的增大呈现先增大后减小的趋势, 最适 SSS 范围为 $35.1\sim 35.4$; CPUE 随 SST 上升不断减小, 最适 SST 范围为 $16.3\sim 18.5^{\circ}\text{C}$; CPUE 随 SSH 上升先增大后减小, 最适 SSH 范围为 $0.55\sim 0.60 \text{ m}$; CPUE 随 Chl a 上升先增大后减小, 最适 Chl a 范围为 $0.18\sim 0.46 \text{ mg/m}^3$ 。

3.5 茎柔鱼分布预报

利用随机森林模型对茎柔鱼各月资源丰度进行预测, 模型 CPUE 预测值与名义 CPUE 变化趋势如图 8 所示。除 2019 年 9-10 月、11-12 月外, CPUE 预测值与名义 CPUE 变化趋势基本一致。CPUE 预测

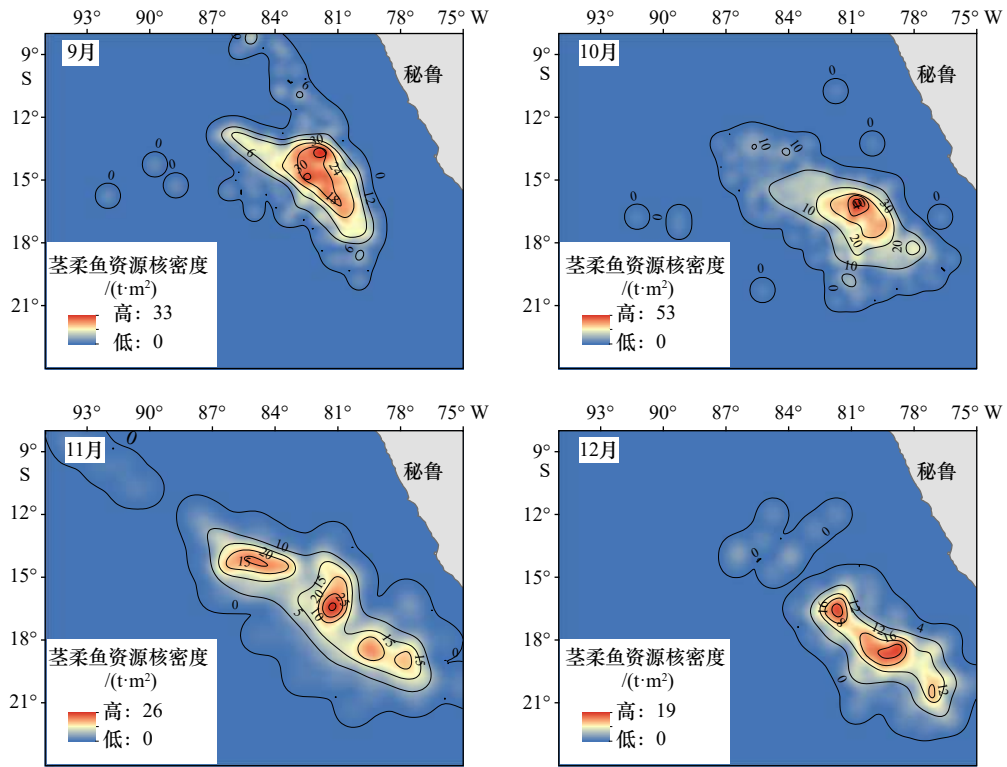


图 5 9-12月茎柔鱼资源核密度分布

Fig. 5 Kernel density distribution of *Dosidicus gigas* resources from September to December

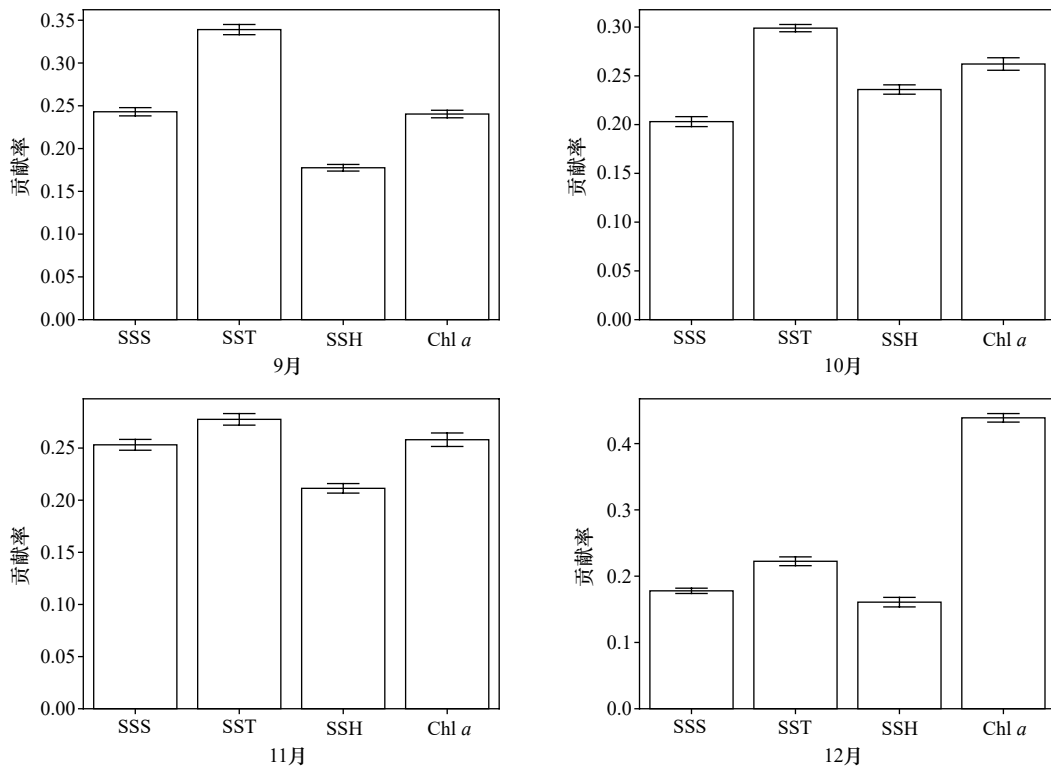


图 6 随机森林模型中环境因子贡献率

Fig. 6 Contribution rate of environmental factors in Random forest model

值与名义 CPUE 相比,前者波动幅度更为平缓。CPUE 预测值最大值在 2021 年 9 月,最小值在 2018 年

12 月;名义 CPUE 最大值在 2021 年 9 月,最小值在 2021 年 12 月。

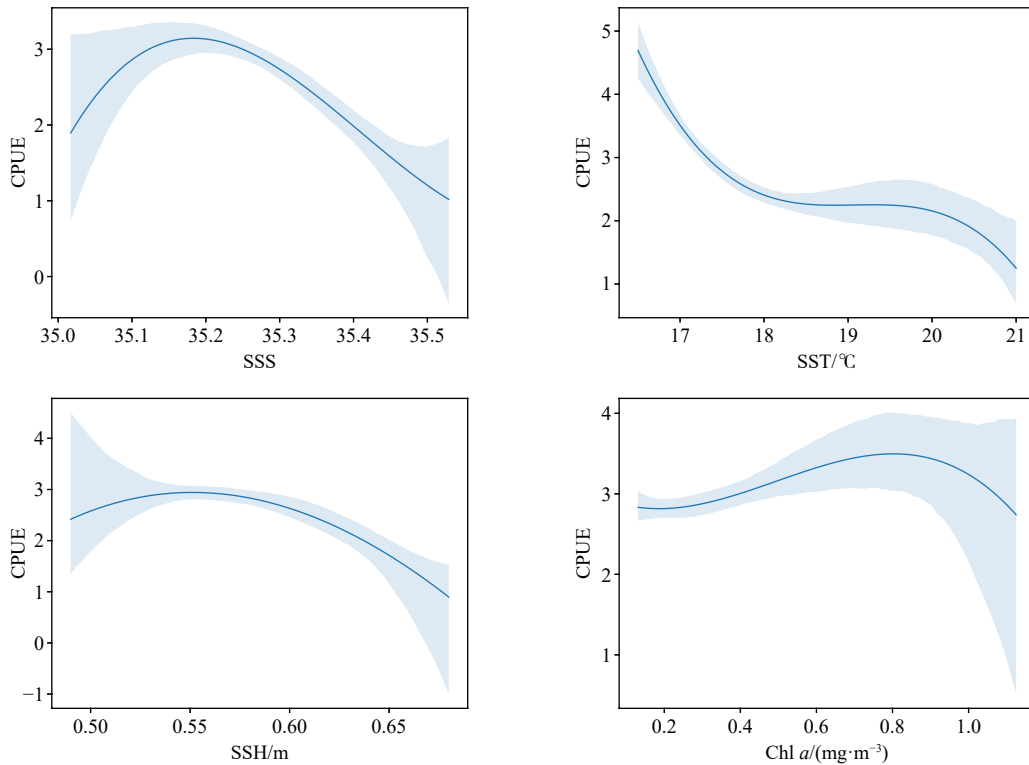


图 7 基于随机森林模型分析环境因子对茎柔鱼资源量的影响

Fig. 7 Effects of environmental factors based on factor Random forest model on the biomass of *Dosidicus gigas*

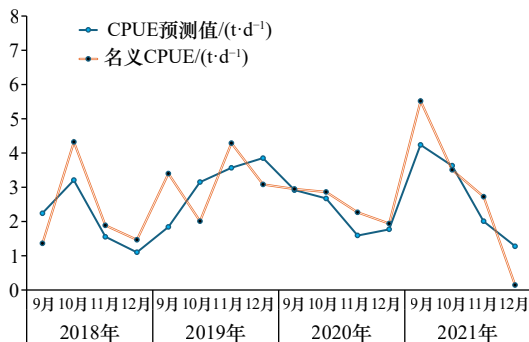


图 8 2018–2021 年 9–12 月秘鲁外海茎柔鱼名义 CPUE 与 CPUE 预测值

Fig. 8 Nominal and predicted CPUE for *Dosidicus gigas* in offshore waters of Peru from September to December 2018 to 2021

4 讨论与分析

4.1 随机森林模型

随机森林模型在生物资源与环境因子的关系应用研究较多^[24–26]。本研究基于 2018–2021 年 9–12 月中国鱿钓船秘鲁外海作业数据,并选取 SSS、SST、SSH、Chl *a* 共 4 个环境因子进行建模。在 100 次 7 折交叉验证中得出的预测值与实际值之间存在正相关(相关系数大于 0.75),且模型的准确性较高(RMSD < 0.71),能较好地解释变量之间的关系($R^2 > 0.54$),表

明采用随机森林模型能较好地分析资源丰度与环境因子之间的关系。此外本文利用随机森林模型对茎柔鱼各年份的 9–12 月资源量进行预测,预测结果表明秘鲁外海茎柔鱼 CPUE 预测值大部分要低于名义 CPUE,这与张弼强等利用 GAM 对西北印度洋鳶乌贼标准化结果一致^[27]。两者随时间变化趋势大体一致,CPUE 预测值变化幅度相比名义 CPUE 小,即通过随机森林模型处理后的 CPUE 预测值,在一定程度上能排除环境因子和时空因子对于 CPUE 的影响,能够更好地解释茎柔鱼资源丰度变动规律^[28]。因此,利用随机森林模型来分析秘鲁外海茎柔鱼与环境因子之间的关系具有较高的可靠性,可以为秘鲁外海茎柔鱼的可持续利用提供依据。

4.2 秘鲁外海茎柔鱼渔场产量重心变动分析

秘鲁外海茎柔鱼资源分布,因为受到茎柔鱼洄游特性和环境因素的影响,导致茎柔鱼资源分布在时空上出现变动^[29]。本研究发现秘鲁外海茎柔鱼渔场重心在 9–12 月由西北向东南方向移动。本研究认为渔场重心变化主要受到捕捞船作业路线和 SST 的影响。9–12 月为南半球的春季和夏初,随着月份的增加,SST 上升,茎柔鱼为寻找适宜的 SST 海域,进行水平洄游,此外有经验的渔业作业者会选择合适的 SST 的海域进行生产作业,故茎柔鱼产量重心变化与

不同月份渔船的转场作业及 SST 有关。本研究发
现秘鲁外海茎柔鱼资源分布重心在 11 月、12 月时出现
双重心分布, 12 月相对 11 月双重心之间的距离变小,
这可能因为 11 月为秘鲁茎柔鱼产卵高峰期^[10], 茎柔
鱼的卵和仔鱼受到洪堡海流的影响向北流去, 使茎柔
鱼资源重心出现向北偏移(图 9)。

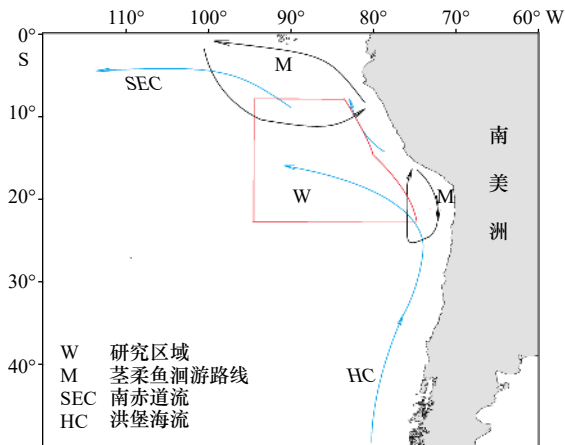


图 9 秘鲁外海洋流及茎柔鱼洄游路线
Fig. 9 Currents distribution of offshore waters of Peru
and the migration routes of *Dosidicus gigas*

4.3 秘鲁外海茎柔鱼资源与环境因子关系

9–12 月各环境因子贡献率结果显示, 各环境因
子对秘鲁外海茎柔鱼资源丰度和分布的影响存在一
定的差异性(图 6)。SST 是影响茎柔鱼分布的重要环
境因子, 通常与茎柔鱼渔场变动密切相关^[5,8], 其对茎
柔鱼的生长发育、摄食、洄游会产生直接或者间接的
影响。如陈新军等^[29]认为 SST 对秘鲁外海茎柔鱼资
源产量以及分布具有显著影响; 刘杉等^[30]利用栖息地
模型探究秘鲁海域水温变化对茎柔鱼栖息地的影响,
当海水温度升高时茎柔鱼渔场向南移动, 且适宜栖息
地面积逐渐减小, 这与本文的研究结果基本一致。本
研究结果显示 12 月份 SST 对茎柔鱼资源丰度影响低
于 Chl *a*, 这与方星楠^[8]利用最大熵模型分析茎柔鱼
资源与环境因子关系时所得出的结果 10–12 月 SST 对
资源丰度的影响较大有所差异, 这可能与 9–11 月海
表平均温度低于 20°C, 而 12 月份海表温平均值高于
20°C 有关。本研究发
现秘鲁外海茎柔鱼最适 SST 为 16.3~18.5°C, 这与李莉等^[6]认为的 18~20°C, 方星楠
等^[10]认为的 17.2~20.6°C, 刘必林等^[31]认为的 16~
20°C, 方学燕等^[13]认为的 17~23°C, 胡振明等^[7]认为
的 18~23°C 结果有些差异, 这可能因为本文研究区
域与研究时间与各研究者之间存在差异。9–11 月为
南半球春季, 随着月份的增加, 南半球温度整体逐渐
上升, 相应区域内对应的 SST 范围也发生改变; 各类

遥感数据库中数据存在偏差, 本文对比各数据库相同
时空的 SST, 发现 SST 存在差异, 这对各研究结果也
会产生影响; 气候变化会改变秘鲁外海茎柔鱼的生存
环境^[32]和上升流的强弱也会影响秘鲁外海茎柔鱼栖
息地变化。

Chl *a* 浓度的高低可以反应初级生产力状况^[33], 初
级生产力越高的地方浮游生物越丰富, 这能为茎柔鱼
提供充足的食物, 茎柔鱼资源量相对大, 因而 Chl *a*
浓度高低与茎柔鱼产量相关^[34]。本研究发
现 Chl *a* 在 9–12 月份时贡献率都处在相对高的状态,
11–12 月份 Chl *a* 贡献率均在 0.25 左右。这可能是因为 10 月到
翌年 2 月是茎柔鱼的高产卵期, 此时需要大量饵料,
而 Chl *a* 是浮游植物进行光合作用的主要色素, 造成
9–12 月份 Chl *a* 贡献率较高。

盐度对茎柔鱼分布范围存在重要影响, 高盐度锋
面可能制约茎柔鱼的迁移和洄游过程^[35]。SSS 的适
宜范围均以带状分布, 且与适宜栖息地较好重叠^[6]。
本研究发
现, 在选取的 4 个环境因子中, SSS 贡献率
也相对较高, 这可能是受到海流的影响, 近岸低盐水和
外海的高盐水相互补偿, 使得该区域盐度适中, 导致
SSS 在对茎柔鱼资源量的影响密切。

陈新军等^[36]认为秘鲁外海茎柔鱼资源分布主要
受加利福尼亚和秘鲁海流的影响, 海流的入侵会对
SSH 带来影响^[37], 本文研究发
现 9–12 月 SSH 对秘
鲁外海茎柔鱼资源分布贡献率在 0.2 上下波动, 这可
能与秘鲁外海茎柔鱼分布在秘鲁上升流区有关, 海流
上升时可以将营养物质和盐从海底运输到海面, 促进
了茎柔鱼的生长和繁殖, 从而增加了 SSH 与资源量之
间的关系。本研究选用了 SSS、SST、SSH、Chl *a* 4 个
环境因子进行分析, 但还有影响秘鲁外海茎柔鱼资源
变动的因子没考虑进去, 如海流、涡旋^[8]都可能对茎
柔鱼分布产生影响。因此, 在今后的研究中, 可以在
模型中加入其他因子进行分析, 以更加全面地了解秘
鲁外海茎柔鱼资源丰度与环境因子的关系。

5 结论

本研究基于随机森林模型, 对秘鲁外海茎柔鱼资
源丰度与环境因子之间的关系进行了分析。研究发
现: 通过随机森林模型得出的预测值与实际值之间存
在正相关, 模型的准确性较高, 能较好地解释变量之
间的关系; 基于随机森林模型的 CPUE 预测值与名义
CPUE 随时间变化趋势基本一致, 表明随机森林模型
适用于分析秘鲁外海茎柔鱼与环境因子之间的关系;
9–12 月秘鲁外海茎柔鱼资源分布呈现一定的聚集

性, 渔场主要分布在 $13^{\circ}\sim 21^{\circ}\text{S}$, $76^{\circ}\sim 87^{\circ}\text{W}$ 区域, 资源核密度值随着月份的推移呈现先增大后减小的趋势, 渔场重心由西北向东南方向转移; 渔场最适 SST 范围为 $16.3\sim 18.5^{\circ}\text{C}$, 最适 SSS 范围为 $35.1\sim 35.4$, 最适 SSH 为 $0.55\sim 0.60\text{ m}$, 最适 Chl *a* 浓度范围为 $0.18\sim 0.46\text{ mg/m}^3$, SST 和 Chl *a* 对该海域茎柔鱼资源丰度影

响较大。由于秘鲁海域气候条件复杂, 茎柔鱼本身还存在昼夜垂直移动的现象, 涡旋和垂直水温等环境因素对茎柔鱼资源也会带来影响, 因此在今后的研究中可以加入更多的环境因素, 以便能够更全面解析茎柔鱼资源丰度与环境因子的关系, 为秘鲁外海茎柔鱼资源可持续开发提供科学依据。

参考文献:

- [1] 王尧耕, 陈新军. 世界大洋性经济柔鱼类资源及其渔业 [M]. 北京: 海洋出版社, 2005: 240–264.
Wang Yaogeng, Chen Xinjun. World Oceanic Economic Soft Fish Resources and Their Fisheries[M]. Beijing: China Ocean Press, 2005: 240–264.
- [2] Markaida U, Quiñónez-Velázquez C, Sosa-Nishizaki O. Age, growth and maturation of jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) from the Gulf of California, Mexico[J]. *Fisheries Research*, 2004, 66(1): 31–47.
- [3] 陈新军, 陆化杰, 刘必林, 等. 大洋性柔鱼类资源开发现状及可持续利用的科学问题 [J]. 上海海洋大学学报, 2012, 21(5): 831–840.
Chen Xinjun, Lu Huajie, Liu Bilin, et al. Current exploitation and some scientific issues in the sustainable utilization of *Ommastrephidae*[J]. *Journal of Shanghai Ocean University*, 2012, 21(5): 831–840.
- [4] Nigmatullin C M, Nesis K N, Arkhipkin A I. A review of the biology of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae)[J]. *Fisheries Research*, 2001, 54(1): 9–19.
- [5] 玄文丹, 崔国辰, 李哲, 等. 东太平洋赤道海域茎柔鱼 (*Dosidicus gigas*) 小型群体资源分布及其渔场环境特征 [J]. *海洋与湖沼*, 2022, 53(5): 1234–1241.
Xuan Wendan, Cui Guochen, Li Zhe, et al. Distribution and environmental dependency of small schools of squid *Dosidicus gigas* and in the equator of eastern Pacific Ocean[J]. *Oceanologia et Limnologia Sinica*, 2022, 53(5): 1234–1241.
- [6] 李莉, 吕翔, 张硕, 等. 水温变化对秘鲁外海茎柔鱼栖息地的影响 [J]. 大连海洋大学学报, 2016, 31(6): 685–691.
Li Li, Lü Xiang, Zhang Shuo, et al. Impact of changes in water temperature on habitat of Jumbo flying squid *Dosidicus gigas* off the coast of Peru[J]. *Journal of Dalian Ocean University*, 2016, 31(6): 685–691.
- [7] 胡振明, 陈新军, 周应祺, 等. 利用栖息地适宜指数分析秘鲁外海茎柔鱼渔场分布 [J]. 海洋学报, 2010, 32(5): 67–75.
Hu Zhenming, Chen Xinjun, Zhou Yingqi, et al. Forecasting fishing ground of *Dosidicus gigas* based on habitat suitability index off Peru[J]. *Haiyang Xuebao*, 2010, 32(5): 67–75.
- [8] 方星楠. 东太平洋茎柔鱼资源丰度时空分布差异及其海洋动力因素分析 [D]. 上海: 上海海洋大学, 2022.
Fang Xingnan. Tempo-spatial distribution differences of abundance index of *Dosidicus gigas* and the influences of marine dynamic factors in the Southeast Pacific Ocean[D]. Shanghai: Shanghai Ocean University, 2022.
- [9] 魏裙倚, 崔国辰, 玄文丹, 等. 海表面温度及叶绿素 *a* 浓度对西北印度洋鸕乌贼时空分布的影响 [J]. *中国水产科学*, 2022, 29(3): 388–397.
Wei Qunyi, Cui Guochen, Xuan Wendan, et al. Effects of SST and Chl-*a* on the spatiotemporal distribution of *Sthenoteuthis oualaniensis* fishing ground in the northwest Indian Ocean[J]. *Journal of Fishery Sciences of China*, 2022, 29(3): 388–397.
- [10] 方星楠, 何妍, 余为, 等. 秘鲁外海茎柔鱼栖息地时空分布及对环境因子的响应差异 [J]. *中国水产科学*, 2021, 28(5): 658–672.
Fang Xingnan, He Yan, Yu Wei, et al. Spatio-temporal distribution of the jumbo flying squid *Dosidicus gigas* off Peru and differences in the effects of environmental conditions[J]. *Journal of Fishery Sciences of China*, 2021, 28(5): 658–672.
- [11] 徐国强, 朱文斌, 张洪亮, 等. 基于 GAM 模型分析印度洋大眼金枪鱼和黄鳍金枪鱼渔场分布与不同环境因子关系 [J]. 海洋学报, 2018, 40(12): 68–80.
Xu Guoqiang, Zhu Wenbin, Zhang Hongliang, et al. Relationship between fishing grounds of *Thunnus obesus* and *Thunnus albacores* with environmental factors in the Indian Ocean based on generalized additive model[J]. *Haiyang Xuebao*, 2018, 40(12): 68–80.
- [12] 王言丰, 陈新军, 陈芑, 等. 基于灰色系统西南大西洋阿根廷滑柔鱼资源丰度预测模型的构建 [J]. 海洋学报, 2019, 41(4): 64–73.
Wang Yanfeng, Chen Xinjun, Chen Peng, et al. Prediction of abundance index of *Argentine shortfin* squid in the southwest Atlantic Ocean based on grey system model[J]. *Haiyang Xuebao*, 2019, 41(4): 64–73.
- [13] 方学燕, 陈新军, 冯永玖, 等. 基于综合环境因子的协同克里金法分析茎柔鱼资源丰度空间分布 [J]. 海洋学报, 2017, 39(2): 62–71.
Fang Xueyan, Chen Xinjun, Feng Yongjiu, et al. Study of spatial distribution for *Dosidicus gigas* abundance off Peru based on a comprehensive environmental factor[J]. *Haiyang Xuebao*, 2017, 39(2): 62–71.
- [14] 游介文, 邹滨, 赵秀阁, 等. 基于随机森林模型的中国近地面 NO_2 浓度估算 [J]. *中国环境科学*, 2019, 39(3): 969–979.
You Jiewen, Zou Bin, Zhao Xiuge, et al. Estimating ground-level NO_2 concentrations across mainland China using random forests regression modeling[J]. *China Environmental Science*, 2019, 39(3): 969–979.
- [15] 董师师, 黄哲学. 随机森林理论浅析 [J]. 集成技术, 2013, 2(1): 1–7.
Dong Shishi, Huang Zhexue. A brief theoretical overview of random forests[J]. *Journal of Integration Technology*, 2013, 2(1): 1–7.

- [16] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5–32.
- [17] Xia Jiangzhou, Ma Minna, Liang Tiangang, et al. Estimates of grassland biomass and turnover time on the Tibetan Plateau[J]. *Environmental Research Letters*, 2018, 13(1): 014020.
- [18] Cao Jie, Chen Xinjun, Chen Yong. Influence of surface oceanographic variability on abundance of the western winter-spring cohort of neon flying squid *Ommastrephes bartramii* in the NW Pacific Ocean[J]. *Marine Ecology Progress Series*, 2009, 381: 119–127.
- [19] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述 [J]. *统计与信息论坛*, 2011, 26(3): 32–38.
Fang Kuangnan, Wu Jianbin, Zhu Jianping, et al. A review of technologies on random forests[J]. *Statistics & Information Forum*, 2011, 26(3): 32–38.
- [20] Meng Yanrong, Yang Mingxia, Liu Shan, et al. Quantitative assessment of the importance of bio-physical drivers of land cover change based on a random forest method[J]. *Ecological Informatics*, 2021, 61: 101204.
- [21] 李进文, 陈朝辉, 孙燕, 等. 多元线性回归模型变量选择的总偏回归平方和法 [J]. *数理医药学杂志*, 2007, 20(2): 126–127.
Li Jinwen, Chen Zhaohui, Sun Yan, et al. A method of total partial regression sum of squares for selecting variables in the linear model[J]. *Journal of Mathematical Medicine*, 2007, 20(2): 126–127.
- [22] Xu Binduo, Zhang Chongliang, Xue Ying, et al. Optimization of sampling effort for a fishery-independent survey with multiple goals[J]. *Environmental Monitoring and Assessment*, 2015, 187(5): 252.
- [23] 陈仁丽, 吴晓青, 刘柏静, 等. 基于 AIS 的海上渔船捕捞活动聚集区提取方法研究 [J]. *地球信息科学学报*, 2021, 23(12): 2163–2173.
Chen Renli, Wu Xiaoqing, Liu Baijing, et al. Mapping method of fishing grounds based on marine AIS data[J]. *Journal of Geo-information Science*, 2021, 23(12): 2163–2173.
- [24] 崔晏华, 刘淑德, 张云雷, 等. 海州湾春季短蛸的栖息分布特征及其与环境因子的关系 [J]. *应用生态学报*, 2022, 33(6): 1686–1692.
Cui Yanhua, Liu Shude, Zhang Yunlei, et al. Habitat characteristics of *Octopus ocellatus* and their relationship with environmental factors during spring in Haizhou Bay, China[J]. *Chinese Journal of Applied Ecology*, 2022, 33(6): 1686–1692.
- [25] 魏广恩, 陈新军. 基于产卵场和索饵场适宜性的西北太平洋柔鱼丰度预测 [J]. *海洋学报*, 2020, 42(12): 14–25.
Wei Guang'en, Chen Xinjun. Forecasting northwest Pacific Ocean neon flying squid abundance based on suitability of spawning and feeding grounds[J]. *Haiyang Xuebao*, 2020, 42(12): 14–25.
- [26] 陈雪忠, 樊伟, 崔雪森, 等. 基于随机森林的印度洋长鳍金枪鱼渔场预报 [J]. *海洋学报*, 2013, 35(1): 158–164.
Chen Xuezhong, Fan Wei, Cui Xuesen, et al. Fishing ground forecasting of *Thunnus alalunga* in Indian Ocean based on random forest[J]. *Haiyang Xuebao*, 2013, 35(1): 158–164.
- [27] 张弼强, 陆化杰, 赵懋林, 等. 基于 GAM 模型西北印度洋鸚乌贼 CPUE 标准化 [J]. *海洋与湖沼*, 2023, 54(1): 259–265.
Zhang Biqiang, Lu Huajie, Zhao Maolin, et al. Standardization of catch per unit effort (CPUE) in northwest Indian Ocean *Sthenoteuthis oualaniensis* based on generalized additive model[J]. *Oceanologia et Limnologia Sinica*, 2023, 54(1): 259–265.
- [28] 林泓羽, 汪金涛, 陈新. 利用 BP 神经网络标准化东南太平洋茎柔鱼 (*Dosidicus gigas*) 资源丰度 [J]. *海洋与湖沼*, 2022, 53(5): 1279–1287.
Lin Hongyu, Wang Jintao, Chen Xin. The CPUE standardization of *Dosidicus gigas* in the southeastern Pacific Ocean using BP neural network[J]. *Oceanologia et Limnologia Sinica*, 2022, 53(5): 1279–1287.
- [29] 徐冰, 陈新军, 钱卫国, 等. 秘鲁外海茎柔鱼渔场时空分布分析 [J]. *中国海洋大学学报*, 2011, 41(11): 43–47.
Xu Bing, Chen Xinjun, Qian Weiguo, et al. Spatial and temporal distribution of fishing ground for *Dosidicus gigas* in the offshore waters of Peru[J]. *Periodical of Ocean University of China*, 2011, 41(11): 43–47.
- [30] 刘杉, 张子健, 吴佳峰, 等. 海表面温度变化下秘鲁外海茎柔鱼栖息地的时空变动 [J]. *水产科学*, 2022, 41(4): 636–644.
Liu Shan, Zhang Zijian, Wu Jiayi, et al. Spatial and temporal variations of potential habitats of jumbo flying squid *Dosidicus gigas* off Peru under increasing sea surface temperature[J]. *Fisheries Science*, 2022, 41(4): 636–644.
- [31] 刘必林, 陈新军, 田思泉. 秘鲁外海茎柔鱼索饵场栖息地研究 [J]. *水产科学*, 2023, 42(3): 480–487.
Liu Bilin, Chen Xinjun, Tian Siqian. Feeding habitat of jumbo flying squid *Dosidicus gigas* off the Coast of Peru[J]. *Fisheries Science*, 2023, 42(3): 480–487.
- [32] 徐冰. 秘鲁外海茎柔鱼渔场时空分布及资源补充量与环境的关系 [D]. 上海: 上海海洋大学, 2012.
Xu Bing. Spatial and Temporal distribution of Fishing Ground for *Dosidicus gigas* and the relationship between Recruitment and Environment in the Offshore Waters of Peru[D]. Shanghai: Shanghai Ocean University, 2012.
- [33] Chen Xinjun, Li Jianhua, Liu Bilin, et al. Age, growth and population structure of jumbo flying squid, *Dosidicus gigas*, off the Costa Rica Dome[J]. *Journal of the Marine Biological Association of the United Kingdom*, 2013, 93(2): 567–573.
- [34] Robinson C J, Gómez-Gutiérrez J, de León D A S. Jumbo squid (*Dosidicus gigas*) landings in the Gulf of California related to remotely sensed SST and concentrations of chlorophyll a (1998–2012)[J]. *Fisheries Research*, 2013, 137: 97–103.
- [35] Ichii T, Mahapatra K, Watanabe T, et al. Occurrence of jumbo flying squid *Dosidicus gigas* aggregations associated with the counter-current ridge off the Costa Rica Dome during 1997 El Niño and 1999 La Niña[J]. *Marine Ecology Progress Series*, 2002, 231: 151–166.
- [36] 陈新军, 龚彩霞, 田思泉, 等. 栖息地理论在海洋渔业中的应用 [M]. 北京: 海洋出版社, 2019: 82–90.
Chen Xinjun, Gong Caixia, Tian Siqian, et al. Application of Habitat Theory to Marine Fisheries[M]. Beijing: China Ocean Press, 2019: 82–90.

- [37] 宋婷婷. 基于海面高度数据研究西北太平洋巴特柔鱼 (*Ommastrephes bartrami*) 渔场分布 [D]. 上海: 上海海洋大学, 2014.
Song Tingting. Relationship between fishing grounds of *Ommastrephes Bartrami* and satellite altimeter data in northwestern Pacific[D]. Shanghai: Shanghai Ocean University, 2014.

Relationship between abundance of *Dosidicus gigas* and environment factors in the offshore waters of Peru based on random forest model

Peng Weitu^{1,2}, Chen Feng², Li Jianxiong^{1,2}, Zhao Yixiang^{1,2}, Dong Yijiang^{1,2},
Dai Qian², Xu Hanxiang^{1,2}, Zhu Wenbin²

(1. Marine and Fishery Institute of Zhejiang Ocean University, Zhoushan 316021, China; 2. Zhejiang Marine Fisheries Research Institute/Scientific Observing and Experimental Station of Fishery Resources for Key Fishing Grounds, Ministry of Agriculture and Rural Affairs/Key Laboratory of Sustainable Utilization of Technology Research for Fishery Resources of Zhejiang Province, Zhoushan 316021, China)

Abstract: *Dosidicus gigas* is an oceanic economic fish in the offshore waters of Peru, and its abundance is greatly affected by environmental factors. This study obtained the abundance of *D. gigas* through fishing logs from September to December in 2018–2021, combined with environmental factors, including surface temperature (SST), sea surface salinity (SSS), sea surface height (SSH) and chlorophyll *a* concentration (Chl *a*) acquired from satellite remote sensing. Random forest model and ArcGIS were applied to analyze the correlations between the abundance of *D. gigas* and environmental factors. Results showed that the distribution of the center of gravity of fishing ground was concentrated in the range of 13°–21°S, 76°–87°W during 2018–2021, and the center of gravity shifted from northwest to the southeast from September to December. The results of random forest model analysis showed that, the impact of environmental factors on the abundance and distribution of *D. gigas* in the offshore waters of Peru varied among different months. The optimal SST ranged from 16.3°C to 18.5°C, The optimal SSS ranged from 35.1 to 35.4, the optimal SSH ranged from 0.55 m to 0.60 m, the optimal Chl *a* concentration ranged from 0.18 mg/m³ to 0.46 mg/m³. The predicted CPUE values derived from the random forest model were generally consistent with the nominal CPUE distribution, indicating the suitability of the random forest model for analyzing the relationship between *D. gigas* and environmental factors in the offshore waters of Peru. This study is of great significance for understanding the resource dynamics of *D. gigas* and guiding its production.

Key words: *Dosidicus gigas*; random forest model; abundance; environmental factors