



基于LSTM模型的中西太平洋鲣栖息地预测

周成¹, 周想¹, 胡媛媛¹, 刘力文¹

上海海洋大学海洋生物资源与管理学院, 上海 201306

College of Marine Living Resource Sciences and Management, Shanghai Ocean University, Shanghai 201306, P. R. China

摘要

为了解决传统栖息地预测模型中无法捕捉具有时间序列信息的环境因子对金枪鱼空间分布滞后影响的不足。本研究采用2022至2024年金枪鱼围网渔捞日志数据通过构建滞后天数为1d、5d、10d、15d的长短期记忆神经网络模型分别对单位捕捞单位努力量渔获量和经纬度进行了预测。研究表明, 滞后10d的模型精度最高, 其均方误差为0.0187, 平均绝对误差为0.0776, 表明鲣空间分布受过去短期内环境累计效应的影响。通过对最佳模型进行验证, 结果表明预测纬度与实际纬度之间的 R^2 为0.97, 预测经度与实际经度之间的 R^2 为0.65, 说明空间分布预测范围与实际基本吻合。本研究为揭示鲣栖息地特征及其生态过程的动态机制提供了新的理解, 同时为中西太平洋鲣围网渔业的科学管理提供了重要参考依据。

背景

鲣是全球分布最广、经济价值最高的金枪鱼之一, 产量中超 60% 来自围网渔业。传统模型难以反映鲣对环境变化的时滞响应。本研究基于中西太平洋渔捞日志与多层海洋环境数据, 构建不同滞后天数的 LSTM 模型, 预测鲣的 CPUE 与空间分布。结果表明, 鲣对环境因子存在明显的时滞响应, LSTM 模型在刻画其动态分布方面优于传统模型, 为渔业管理与栖息地研究提供了新思路。

材料和方法

- 本研究渔业数据来自国家远洋渔业数据中心 (2012年12月—2024年1月), 包括作业时间、位置、投网次数、捕捞类型及渔获量。海洋环境数据源自欧盟哥白尼中心, 涵盖0–200 m水层的温度、盐度、溶解氧、叶绿素 a、初级生产力等要素, 分辨率为 $0.25^\circ \times 0.25^\circ$ 、1 d。
- 渔捞日志与环境数据依据日期和经纬度以最小欧氏距离匹配, 将作业点映射至对应渔区单元, 并合并同日同单元数据以获取日均渔获量和环境均值。构建含时间步长d的三维样本矩阵($n \times d \times N$)以适配LSTM输入。
- 为避免多重共线性, 本研究采用弹性网络模型筛选变量, 并通过交叉验证确定最优正则化参数 λ , 保留系数非零的关键特征作为LSTM输入。
- 基于时间序列构建LSTM模型预测鲣的CPUE及经纬度分布。数据按时间划分为训练集 (80%)、验证集 (20%) 和测试集 (20%)。设定1、5、10、15 d四种时间步长评估滞后效应, 通过网格搜索与交叉验证优化超参数, 并结合早停和学习率调度防止过拟合。

结果

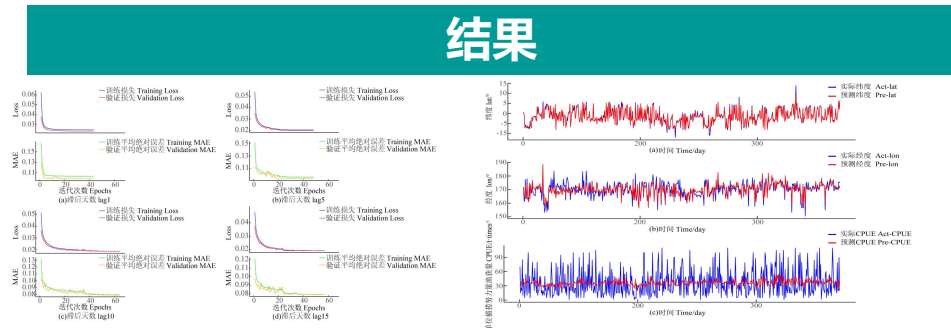


图1 不同时间步长LSTM的最佳模型 LOSS函数曲线图

图2 滞后0~10d 条件下海洋环境-CPUE时空分布 LSTM模型预测结果和实际值的时序对比图

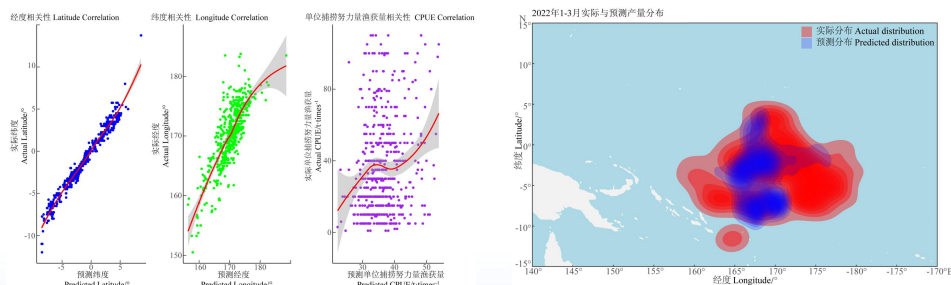


图3 滞后10d条件下海洋环境-CPUE时空分布LSTM 模型预测结果和实际值的相关性分析

图4 2020年1—3月基于LSTM模型预测的鲣在栖息地和实际渔场分布对比图

- 通过 EN 模型筛选出 26 个关键海洋环境变量, 包括 DO、Chl.a、PP、SS、ST、GCV、MLD 和 SSH 等, 对鲣 CPUE 的变化具有显著影响。
- 对比 1d、5d、10d、15d 四种滞后时间模型, 滞后 10d 的 LSTM 模型表现最佳, 具有最低的 MSE (0.0187) 和 MAE (0.0776)。
- 预测与实际结果拟合良好: 纬度 $R^2 = 0.97$; 经度 $R^2 = 0.65$; CPUE $R^2 = 0.03$; 表明模型在空间定位预测上精度较高。

讨论

- LSTM在鲣栖息地预测中显示出空间位置的趋势一致性, 虽受样本量限制, 仍具参考价值。
- 滞后10天模型性能最佳, 表明鲣对环境变化存在短期滞后响应。
- 相较静态模型, LSTM可揭示环境因子的时滞效应, 未来可引入气候指标提升模型解释力与预测精度。