

利用 BP 神经网络标准化东南太平洋茎柔鱼 (*Dosidicus gigas*)资源丰度*

林泓羽¹ 汪金涛^{1, 2, 3, 4, 5①} 陈新军^{1, 2, 3, 4, 5}

(1. 上海海洋大学海洋科学学院 上海 201306; 2. 农业农村部大洋渔业开发重点实验室 上海 201306; 3. 国家远洋渔业工程技术研究中心 上海 201306; 4. 大洋渔业资源可持续开发教育部重点实验室 上海 201306; 5. 农业农村部大洋渔业资源环境科学观测实验站 上海 201306)

摘要 东南太平洋茎柔鱼(*Dosidicus gigas*)是短生命周期大洋性经济鱼类,其资源量受环境因素变化的影响较大。根据我国鱿钓船队 2013~2017 年在东南太平洋的生产统计数据,结合海洋环境数据包括海表面温度(SST)、海表面盐度(SSS)、叶绿素 *a* 浓度(chl *a*),运用 BP 神经网络(back propagation network)模型来标准化单位捕捞努力量渔获量(catch per unit effort, CPUE, 也称名义 CPUE)。以均方误差(mean square errors, MSE)和平均相对变动值(average relative variances, ARV)为最优模型判断依据,比较隐含层节点 3-10 的神经网络模型,发现 6-9-1 结构为最优模型。用 Garson 算法解释模型结果,发现各输入层因子对东南太平洋茎柔鱼资源丰度影响重要度排序为 chl *a*、SST、经度(Lon)、SSS、纬度(Lat)、月份(Month)。并作名义 CPUE 和标准化 CPUE 资源丰度对比分布图,结果显示 CPUE 与标准化 CPUE 总体分布状况基本一致,但局部区域存在明显差异,80°~85°W 及 10°~20°S 海域适宜鱿钓生产,表明 BP 神经网络模型可以适用于东南太平洋茎柔鱼的 CPUE 标准化,从而为鱿钓渔业生产提供一定参考依据。

关键词 BP 神经网络; 东南太平洋茎柔鱼; 资源丰度; 环境因子
中图分类号 S931; S932 **doi:** 10.11693/hyhz20211200343

茎柔鱼隶属头足纲(Cephalopoda)、柔鱼科(Ommastrephidae)、茎柔鱼属(*Dosidicus*) (王尧耕等, 2005)。在柔鱼科中,茎柔鱼是个体最大、资源量最为丰富的种类之一(Nigmatullin *et al.*, 2001)。茎柔鱼广泛分布于东太平洋海域,特别是加利福尼亚(37°~40°N)至智利(45°~47°S)海域(Taibe *et al.*, 2001),是中国、日本、韩国、秘鲁、越南等国鱿钓船的重要捕捞对象(陈新军, 2019)。茎柔鱼生长快速、生命周期短,产卵后即死去,为一年生种类,具有垂直和水平迁移的能力,洄游范围广,能快速响应不断变化的环境条件,资源丰度较易受到环境因素的影响(Benoit-Bird *et al.*, 2012),包括海表面温度(SST)、海表面盐度(SSS)、叶绿素 *a* 浓度(chl *a*)等环境因素。茎柔鱼在南半球的产

卵场主要分布在秘鲁沿岸的北部区域(3°~8°S)和南部区域(12°~17°S) (Tafur *et al.*, 2001),其在秘鲁沿岸的洄游主要受洪堡海流和南赤道流的影响(胡贯宇等, 2018)。据联合国粮农组织(FAO)统计,中国在东南太平洋捕捞的茎柔鱼(*Dosidicus gigas*)的捕捞量在 2001~2017 年间有一定的波动,但总体呈现上升趋势,从 2011 年的 1.78 万 t,增加到 2014 年的历史最高产量,达到 33.25 万 t; 2015~2016 年捕捞量略有下降,2015 年为 32.36 万 t, 2016 年为 22.33 万 t; 2017 年捕捞量又有所回升,达到 29.61 万 t,已成为中国远洋渔业的重要组成部分。

单位捕捞努力量渔获量(catch per unit effort, CPUE)指某渔场在一定时期内,平均每一个捕捞努力

* 国家自然科学基金项目, NSFC41876141 号。林泓羽, 硕士研究生, E-mail: m210200683@st.shou.edu.cn

① 通信作者: 汪金涛, E-mail: jtwang@shou.edu.cn

收稿日期: 2021-12-24, 收修改稿日期: 2022-02-18

量所能捕获的渔获量,常被认为与渔业资源丰度成正比(Hilborn *et al.*, 1992),是最常用的资源丰度指标(邵帽瑛等, 2006)。CPUE 与环境因子间存在关系,如柔鱼(*Ommastrephes bartrami*) CPUE 与海表温度,即表温存在着经验关系,CPUE 与温度之间的关系是非线性的(田思泉, 2003)。然而,CPUE 除与资源丰度有关以外,还会受到栖息地面积(Salthaug *et al.*, 2003)、环境效应(Ziegler *et al.*, 2003)及渔民行为(官文江, 2008)等众多因素的影响,导致 CPUE 与资源丰度间的正比关系有时难以成立(Harley *et al.*, 2001)。因此,研究人员常对 CPUE 进行标准化,消除影响进而重构 CPUE 与资源丰度间的正比关系,达到合理有效利用 CPUE 数据的目的(Ye *et al.*, 2009)。CPUE 标准化时渔业资源评估研究的基础性工作(杨胜龙等, 2015),因此国内外学者对 CPUE 的标准化方法进行了大量研究和总结(杨胜龙等, 2015),包括传统的统计学模型和机器学习方法,传统的统计学模型包括广义线性模型(GLM) (官文江等, 2009)、基于贝叶斯的广义线性模型(GLBM) (陆化杰等, 2013)、广义加性模型(GAM) (戴小杰等, 2011)、基于栖息地的 CPUE 标准化方法(HBS) (Hinton *et al.*, 1996); 机器学习方法包括回归树模型(regression trees, RT) (官文江等, 2014a)、人工神经网络(artificial neural network, ANN) (Maunder *et al.*, 2006)、随机森林(random forest, RF)模型 (Breiman, 2001), 支持向量机(support vector machine, SVM) (杨胜龙等, 2015)等。

在传统的统计学模型中, GLM 广泛应用于 CPUE 标准化(Chang *et al.*, 2011), 但缺点在于无法处理非线性关系(Maunder *et al.*, 2004), 在误差假设、变量选择、缺失数据、交叉项及异常值处理等方面缺乏灵活性(官文江等, 2009); GAM 模型可用于处理非线性问题,但其假设检验等推断过程可能存在问题(Venables *et al.*, 2004)。尽管传统的统计学方法通过 GLM 和 GAM 的互补能够较好地解决线性和非线性问题,但在具有非线性关系的问题映射上,传统的统计学方法远不如机器学习方法的表现,与前者相比,如人工神经网络和回归树模型为例的机器学习方法可以获得较好的预测结果。官文江等(2014b)比较了 GLM 和回归树模型用于 CPUE 标准化的效果,认为回归树模型具有阶跃函数特征,适用于当渔业数据存在非线性关系及异常值时;而当渔业数据不存在非线性关系与异常值时,则 GLM 模型更具优势。虽然机器学习方法在渔业研究上具有较好的表现,杨胜龙等(2015)通

过不同模型在 CPUE 标准化问题上的对比,发现 SVM 和 RF 方法应该被优先考虑用于渔业数据 CPUE 数据标准化。SVM 模型在简单分类问题上具有先进的性能,并且能够高效地解决小样本问题等等,但其无法扩展到较大的较复杂的数据集;神经网络是一种黑盒子技术,其运行时具有不稳定的特点,因而也常常无法解释 CPUE 与自变量间的关系(Maunder, 2001)。虽然神经网络是一种黑盒模型,但其在分类、学习、误差容忍度和非线性关系处理方面的能力使其受到关注(Chen *et al.*, 1990),且研究人员已提出了几种解释神经网络中不同神经元之间权重的方法,如 Garson 算法(Garson, 1991)、Olden 算法(Olden *et al.*, 2004)和灵敏度分析(Özesmi, 1999)等。

1 数据与方法

1.1 渔业数据

东南太平洋茎柔鱼渔业统计数据来自中国远洋渔业数据中心,生产数据包括作业日期、作业位置(经纬度)、作业船数和渔获量,时间范围为 2013~2017 年,空间范围为 75°~90°W、5°~30°S,空间分辨率为 0.5°×0.5°。海洋环境数据包括 SST、SSS、chl *a*, 其中 SST、SSS 来自哥伦比亚大学气候数据实验室网站(<http://iridl.ldeo.columbia.edu>), chl *a* 来自美国国家海洋和大气管理局(<http://oceanwatch.pifsc.noaa.gov>),时间分辨率为月。

生产数据按空间分辨率 0.5°×0.5°、时间分辨率为月进行时空融合统计,并计算单位捕捞努力量渔获量(CPUE)作为其局部资源丰度指标。渔业数据和环境数据采用基于局部均值的伪最近邻法按照作业地点和作业日期进行匹配(Gou *et al.*, 2014)。

$$CPUE_{(i,j,m,y)} = \frac{C_{(i,j,m,y)}}{E_{(i,j,m,y)}}, \quad (1)$$

式中, $CPUE_{(i,j,m,y)}$ 表示 m 月 y 年, 位置 i, j 的平均 CPUE; $C_{(i,j,m,y)}$ 表示 m 月 y 年, 位置 i, j 的总产量; $E_{(i,j,m,y)}$ 表示 m 月 y 年, 位置 i, j 的总的作业船数。

1.2 BP 神经网络模型

BP 神经网络是一种按误差反向传播算法训练的多层前馈神经网络,包括输入层(input layer)、隐含层(hidden layer)和输出层(output layer)三部分,运算过程包括信号的正向传播和反向传播(图 1)。BP 神经网络模型建立过程一般包括三个阶段:训练(training)、验证(validating)和测试(testing)。

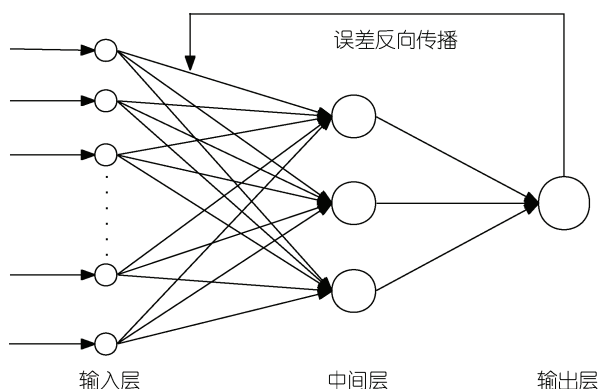


图 1 BP 神经网络的结构

Fig.1 The structure of the BP (back propagation) network

1.2.1 模型构建 将 2013~2017 年共 5 073 个样本利用最小-最大标准化法(min-max scale)进行归一化处理, 使 CPUE 值处于 0~1 之间, 随机分配 70% 的样本集作为训练样本, 30% 的样本集作为测试样本。BP 神经网络的输入层因子为月份、纬度、经度、SST、SSH、chl *a*, 输出层因子为 CPUE, 最大计算次数设置为 1 000, 权值和阈值的更新的学习速率设置为 0.1, 结束条件为误差函数的绝对偏导数小于 0.01。隐含层神经元个数通过经验公式(刘维群等, 2005)设置在 3~10 的范围内, 其具体确定需要在训练过程中合理地增加或减少隐藏层中节点的数量(刘维群等, 2005), 这个过程是相当重要和困难的, 这也需要通过使用反向传播的算法。在最后, 一般可利用交叉验证, 计算其均方误差(mean square errors, MSE)和平均相对变动值(average relative variances, ARV)来量化比较结果, 选出神经网络模型的最佳模型结构(Wang *et al.*, 2018)。

$$n_1 = \sqrt{n+m} + a, \quad (2)$$

式中, n_1 为隐含层个数, n 为输入层个数, m 为输出层个数, a 为 [1, 10] 之间的整数。

1.2.2 模型评价 BP 模型以均方误差和泛化能力作为判断最优模型的标准。均方误差是将预测值与实际值进行比较所得, 其用 MSE 表示, 函数定义式见式 3。泛化能力是指神经网络学习算法对新样本的适应能力, 其目的是学习隐含在数据样本中的规律, 经过训练后的网络对非训练样本也能给出合理的输出, 是衡量预测值与实测值之间差别的变量, 用 ARV 表示, ARV 值越小, 网络泛化能力越强, 其表达式为式 4 (Cholewo *et al.*, 1997; 潘玉民等, 2012):

$$MSE = \frac{\sum_i^n (\hat{X}_i - X_i)^2}{n}, \quad (3)$$

$$ARV = \frac{\sum_{i=1}^n (X_i - \hat{X}_i)^2}{\sum_{i=1}^n (X_i - \bar{X}_i)^2}, \quad (4)$$

式中, n 表示 CPUE 的个数, X_i 为 CPUE 的实际值, $\hat{X}_i$ 为 CPUE 的预测值, $\bar{X}_i$ 为 CPUE 实际值的平均值。

1.2.3 模型解释 在神经网络解释图(neural interpretation diagram, NID)中, 输入层到隐含层、隐含层到输出层之间的连接权重用直线表示, 直线的粗细表示连接权重绝对值的大小, 即信号作用的强弱。Garson 算法是一种借助神经网络获得权重所做的灵敏度分析方法, 它可以利用连接权重的乘积得到输入变量对输出变量的重要度(蔡毅等, 2008)。

2 结果

2.1 最优模型结构及其解释性

由经验公式(式 4)选取 3~10 作为隐藏层的叠加个数, 建立相对应的 6-N-1 神经网络模型, 共计 8 个神经网络模型, 使用交叉验证的方法测试训练模型, 通过对每个模型进行 100 次的循环, 计算其 MSE 和 ARV, 并绘制箱线图如下图所示, MSE 见图 2a, ARV 见图 2b。隐含层节点数 3~10 的网络结构, 对比其 MSE 和 ARV 结果(图 2), 隐藏层为 3~7 的神经网络模型, 整体的 MSE 和 ARV 呈现数值变小的趋势, 于隐藏层节点为 7 时较为稳定, 而后在节点为 8~10 时, 其 MSE 有小幅波动, 于节点为 9 时整体为较小的模型结果, 可见 6-9-1 结构的 MSE 和 ARV 最小, 为最优结构。

构建 6-9-1 结构的 BP 神经网络模型结果(图 3), I 为输入层, 输入的变量依次为月份(Month)、经度(Lon)、纬度(Lat)、海表面盐度(SSS)、海表面温度(SST)、叶绿素 *a* 浓度(chl *a*)。H 为隐藏层, 共有 9 个隐藏层, 为 H1~H9。O 为输出层, 输出变量为 CPUE。B1 和 B2 表示在每一步添加的偏差。图中线的粗细表示各因子所占权重的大小, 黑色线为神经网络传递的正信号, 灰色线为负信号。采用 Garson 算法比较各输入层因子对输出预测的重要度进行灵敏度分析, 其结果, 输入层变量的重要度排序为 chl *a*、SST、Lon、SSS、Lat、Month(图 4)。

2.2 标准化茎柔鱼 CPUE

东南太平洋茎柔鱼的生产期为 10~12 月及次年 1 月的月平均 CPUE 较高, 其中 12 月份 CPUE 最高, 为 3.80 t/d; 其余各月为生产淡季, 3 月份 CPUE 最低, 为 2.30 t/d(图 5)。

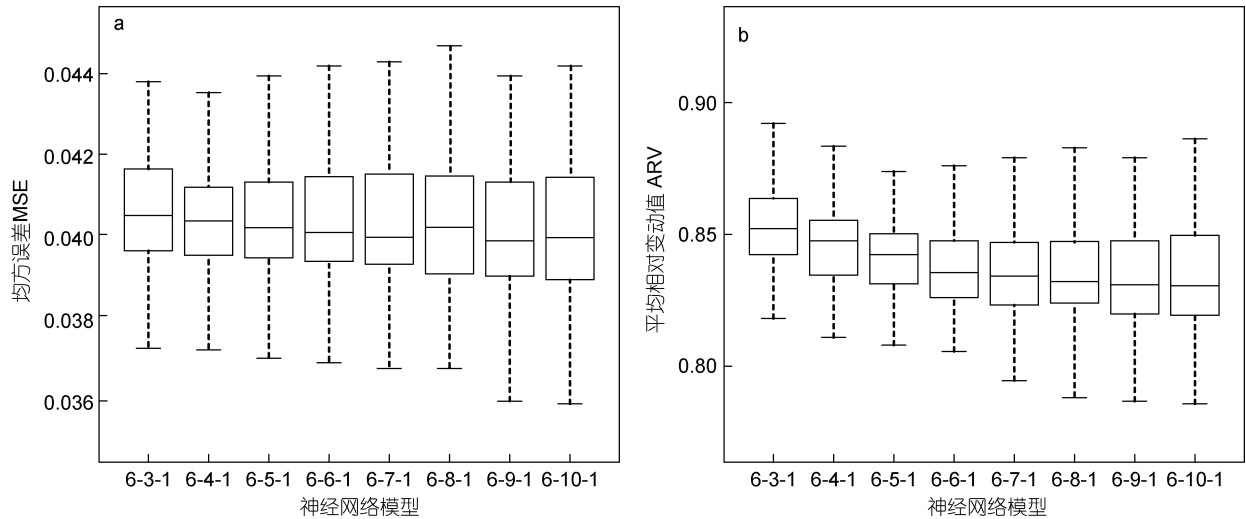


图2 不同节点神经网络模型的均方误差 MSE 和平均相对变动值 ARV

Fig.2 The mean square error and the average relative variance of different node neural network models
注: a. MSE; b. ARV

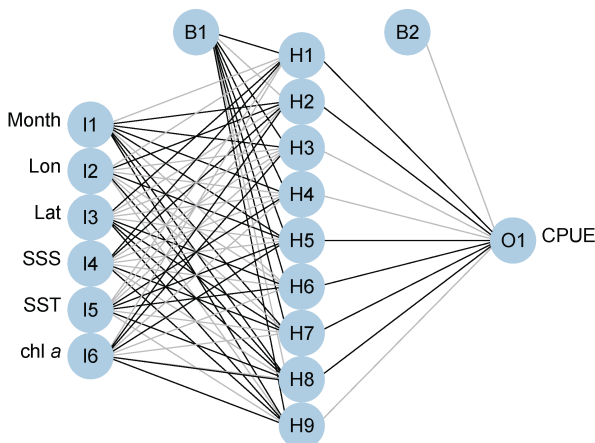


图3 神经网络解释图

Fig.3 The neural interpretation diagram

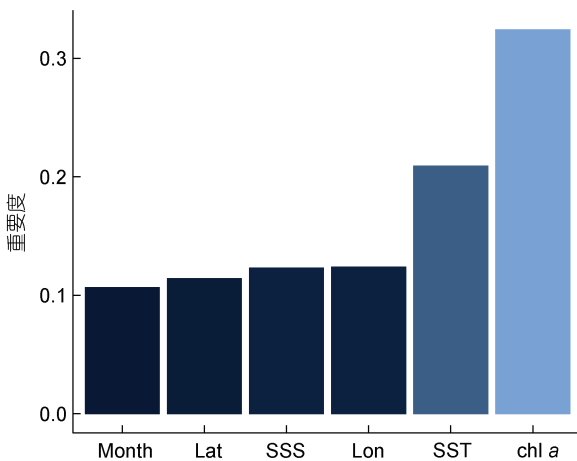


图4 Garson 算法计算输入参数对输出预测的重要度
Fig.4 The importance of input layer for output layer by the Garson algorithm

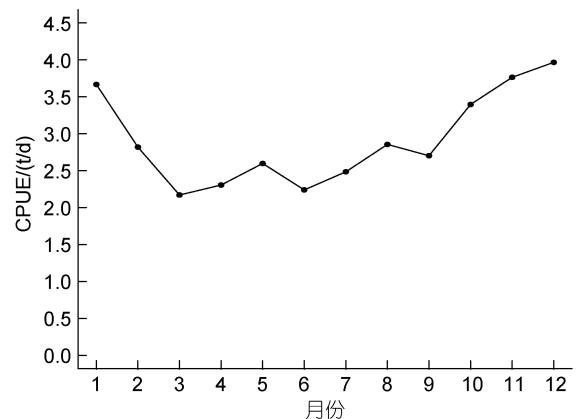


图5 2013~2017 年东南太平洋茎柔鱼月平均 CPUE 分布

Fig.5 Monthly average CPUE distribution of *D. gigas* in the southeastern Pacific Ocean from 2013 to 2017

用选择的最佳 BP 神经网络模型, 预测 2013~2017 年 75°~90°W、5°~30°S 海域的 CPUE 值(记为 CPUE-BP), 将小于 0, 大于 1 的预测值进行二次归一化处理, 其月平均标准化 CPUE 分布(图 6), 年平均 CPUE 标准化分布(图 7)。在时间上, 于月份尺度上整体变化趋势较大, 不论是实际的还是预测的; 而在 2013~2017 年间的年均变化上, 变化的尺度较小。预测的 CPUE (CPUE-BP)同样于 1~3 月开始呈现下降趋势, 并在 3 月达到最小值, 而后呈现上升趋势, 与实际的 CPUE 值的月均变化趋势较一致。2013~2017 年间的年平均标准化 CPUE 分布, 发现 2013~2017 年 CPUE 在 0.50~0.56 内呈小幅度波动, 2015 年达到最大值, 2016 年又有所下降。整体来说, 名义的 CPUE 月

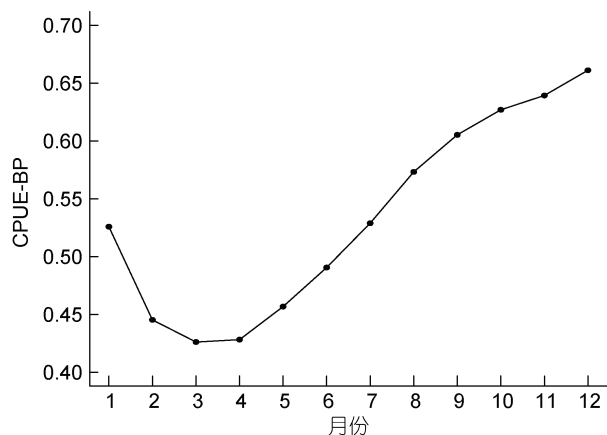


图6 2013~2017年东南太平洋茎柔鱼月平均预测 CPUE 分布

Fig.6 Monthly average predicted CPUE distribution of *D. gigas* in the Southeast Pacific Ocean from 2013 to 2017

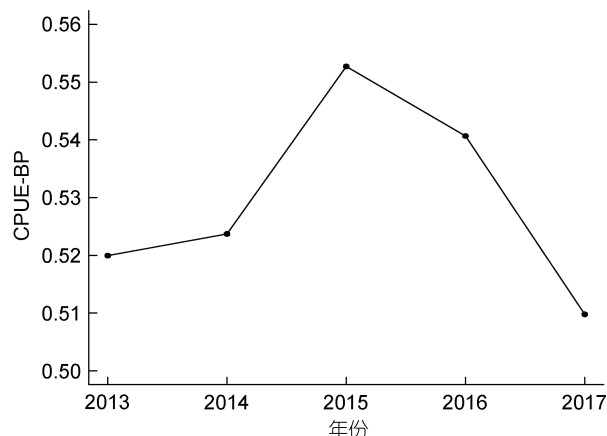


图7 2013~2017年东南太平洋茎柔鱼年平均预测 CPUE 分布

Fig.7 Annual average predicted CPUE distribution of *D. gigas* in the southeastern Pacific Ocean from 2013 to 2017

平均变化与归一化后的 CPUE 月平均变化趋势较一致,说明标准化茎柔鱼的 CPUE 能够较好解释其变化规律,并且标准化后的 CPUE 能够消除捕捞因素等外部因素对于茎柔鱼资源丰度的影响,从而使得标准化后的 CPUE 的变动幅度较小。

2.3 茎柔鱼资源时空变化规律

作 2013~2017 年各年各月标准化 CPUE 分布图(图 8),图中颜色代表预测后的标准化 CPUE 的大小,颜色越深的区域,其预测后的标准化 CPUE 越大。东南太平洋茎柔鱼 CPUE 的时空分布在 2013~2017 年间整体呈现变化明显的状态,各年间的较高值 CPUE 的分布主要集中于 80°~85°W 及 10°~20°S 海域。于月间变化也较为显著,于 3~6 月整体资源丰度较低,而后

升高,常常在 11、12、1 月达到较高值,即颜色分布较深的区域。而年间的分布上以 2015 年为 CPUE 集中分布最为显著的年份,2015 年的 1 月为中心渔场范围最大的月份。东南太平洋茎柔鱼 CPUE 的时空分布整体呈现一个带状分布趋势,由南向北资源分布开始减少。

3 讨论

目前,有关东南太平洋茎柔鱼的研究大多集中在资源和渔场的时空分布变化与环境因子的关系,对其资源丰度的标准化研究则较少。本文利用 2013~2017 年东南太平洋茎柔鱼的生产数据和海洋环境数据,建立了基于东南太平洋茎柔鱼 CPUE 的 BP 神经网络模型,对东南太平洋茎柔鱼的资源丰度进行标准化,并以标准化 CPUE 作其资源丰度分布图。以月份、纬度、经度、SST、SSS、chl *a* 为输入层因子,以 CPUE 为输出层因子,隐含层节点选用 3~10 个,建立了不同的 BP 神经网络模型,对东南太平洋茎柔鱼资源丰度进行标准化研究,并经过 100 次循环计算的交叉验证进行对比。交叉验证在选择平均绝对相对误差较低的模型时表现良好,但并不能总是选择出误差最小的模型(Maunders *et al*, 2006)。交叉验证的结果表明 6-9-1 结构的神经网络模型 MSE 最小,ARV 最小(图 2),泛化能力最强,为最优模型。

根据本文所建立的神经网络模型在各层之间的权重连接情况(图 3),输入层节点 chl *a* 所连出的直线较粗,对隐藏层有较大的贡献,同时通过用 Garson 算法比较各输入层因子对 CPUE 的贡献率(图 4),结果显示对 CPUE 贡献最大的分别是 chl *a*、SST、Lon、SSS、Lat、Month。chl *a* 代表初级生产力,主要影响茎柔鱼的摄食状况,由于茎柔鱼摄食中大型浮游植物,高浓度的叶绿素 *a* 为其提供了丰富的食物(Schwing *et al*, 1999),本研究表明茎柔鱼偏好分布于叶绿素浓度高 SST 低的区域,这与 Robinson 等(2013)证实了茎柔鱼主要分布在海温低、叶绿素 *a* 浓度高的地区的结论一致。

2013~2017 年东南太平洋茎柔鱼的月平均名义 CPUE 分布(图 5)在 10~12 月及 1 月的月平均 CPUE 较高,其中 12 月份 CPUE 最高,为 7.05 t/d。4 月份 CPUE 最低,为 3.91 t/d。这说明 10~12 月及 1 月为捕捞东南太平洋茎柔鱼的盛产期,其余各月为生产淡季。由于 CPUE 可能因除丰度以外的其他因素而变化。因此,通过去除与丰度无关的影响捕集能力的因

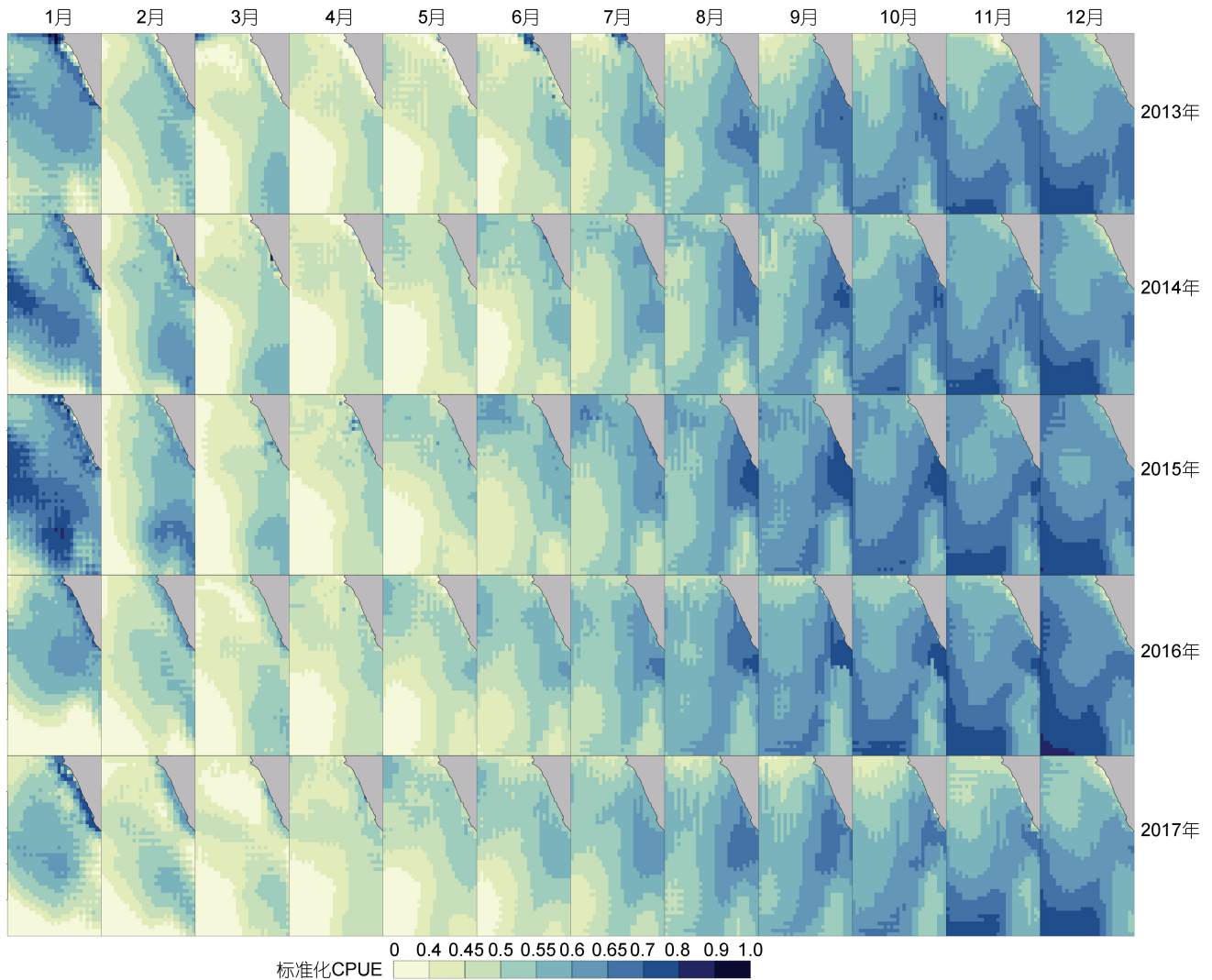


图 8 2013~2017 年间各月标准化 CPUE 分布

Fig.8 Standardized CPUE distribution in each month of 2013~2017

注: 每张子图的空间范围为 30°S~0°, 90°~75°W

素来标准化 CPUE 是很重要的(Maunders *et al.*, 2006)。图 6 直观地显示了 2013~2017 年间 75°~90°W、5°~30°S 海域的月平均 CPUE 分布, 发现 9~12 月为捕捞东南太平洋茎柔鱼的盛产期; 由 2013~2017 年间的年平均标准化 CPUE 分布(图 7), 发现 2013~2017 年 CPUE 在 0.50~0.56 内呈小幅度波动, 2015 年达到最大值, 2016 年又有所下降。由此可见, 标准化后的 CPUE 变化波动幅度小, 这也可能是在通过神经网络模型进行标准化后, 剔除了与丰度无关的因素包括年月份、经纬度等对其造成的影响。由 2013~2017 年间的标准化 CPUE 的时空分布图(图 8), 在时间分布上, 呈现明显的月份季度以及年份尺度上的变化, 于 2015 年的中心渔场范围为分布最为广泛密集, 其他年份

较 2015 年为分散和分布少, 而月份上也以 11、12 以及 1 月份的分布最为密集, 在 3~6 月的分布上较为稀疏, 资源丰度较低, 这可能也是与茎柔鱼的秋生冬生群体以及其摄食行为有关; 从空间分布上, 可见预测后标准化 CPUE 的分布, 主要集中于 80°~85°W 及 10°~20°S 海域, 与汪金涛等(2014)的结论 80°W 以西及 15°S 以南区域适宜鱿钓生产; Robinson 等(2013)结论 1~7 月份产量集中在 80°~84°W 及 14°~20°S 海域; Nash 等(1970)结论 6 月份渔场重心在 81°~84°W 及 10°~15°S 海域; 胡振明等(2008)结论年渔场重心在 79.5°~82.7°W 和 11.9°~17.1°S 海域, 月渔场重心在 81.0°~81.2°W 和 14.3°~15.4°S 基本一致。尽管, 由于受到初始值的影响, 神经网络可能存在多解, 不能给

出不确定性估计(官文江等, 2014a), 但上述结果表明 BP 神经网络对于东南茎柔鱼的 CPUE 标准化结果, 即在时空分布上都有较高度的一致性, 因此能够较好地贴合实际生产状况以及前人在研究中的结论。结果表明 BP 神经网络适合用于 CPUE 的标准化。

4 结论

本文通过采用 6-9-1 结构的 BP 神经网络进行东南太平洋茎柔鱼的 CPUE 标准化, 其结果与前人利用其他统计学方法等研究的渔场重心时空位置基本一致, 结果表明 BP 神经网络适合用于 CPUE 的标准化。但本文只考虑了 SST、SSS、chl *a* 三个环境因素, 没有考虑海表高度异常(SSHA)、海表温度异常(SSTA)、海面高度(SSH)等因素的影响, 也没有考虑到厄尔尼诺事件和拉尼娜事件的影响, 而研究表明厄尔尼诺和拉尼娜事件会影响茎柔鱼幼体和成熟体的生长、栖息等生活史过程。此外, 茎柔鱼具有昼夜垂直移动的现象(Nigmatullin *et al.*, 2001), 还需要考虑到海洋水温的垂直结构对其的影响。因此, 研究过程可以加入更多环境因素, 以便进一步提高模型精度。另外, 本文所用环境数据来自卫星遥感, 由于云层遮挡等原因造成了环境数据的缺失, 在对一整块海域进行 CPUE 的标准化时, 会有缺失值。由于神经网络模型是一种黑盒模型, 其对于 CPUE 标准化表现的性能可能并不完全准确, 通过对神经网络的修改, 包括额外的隐藏层、约束训练或不同的似然函数, 可能会提高它的性能。在判断选择最优神经网络模型的方法上, 采取了交叉验证的方法, 但交叉验证可能并不总是选择误差最小的模型, 这也会造成最终结果具有误差。此外, 本文仅分析了单隐含层的 BP 神经网络结构模型, 对于多隐含层的神经网络结构是否能用于标准化资源丰度或者提高预报精度, 有待进一步研究。

参 考 文 献

- 王尧耕, 陈新军, 2005. 世界大洋性经济柔鱼类资源及其渔业 [M]. 北京: 海洋出版社.
- 田思泉, 2003. 西北太平洋柔鱼资源、作业渔场时空分布及其与表温关系的研究[D]. 上海: 上海水产大学.
- 刘维群, 李元臣, 2005. BP 网络中隐含层节点优化的研究[J]. 交通与计算机, 23(2): 83-86.
- 杨胜龙, 张禹, 张衡, 等, 2015. 不同模型在渔业 CPUE 标准化中的比较分析[J]. 农业工程学报, 31(21): 259-264.
- 汪金涛, 高峰, 雷林, 等, 2014. 基于神经网络的东南太平洋茎柔鱼渔场预报模型的建立及解释[J]. 海洋渔业, 36(2): 131-137.
- 陆化杰, 陈新军, 曹杰, 2013. 基于 GLBM 模型的中国大陆阿根廷滑柔鱼鱿钓渔业 CPUE 标准化[J]. 生态学报, 33(17): 5375-5384.
- 陈新军, 2019. 世界头足类资源开发现状及我国远洋鱿钓渔业发展对策[J]. 上海海洋大学学报, 28(3): 321-330.
- 邵帼瑛, 张敏, 2006. 东南太平洋智利竹荚鱼渔场分布及其与海表温关系的研究[J]. 上海水产大学学报, 15(4): 468-472.
- 官文江, 田思泉, 王学昉, 等, 2014a. CPUE 标准化方法与模型选择的回顾与展望[J]. 中国水产科学, 21(4): 852-862.
- 官文江, 陈新军, 2008. 利用元胞自动机探讨商业性 CPUE 与资源量之间的关系[J]. 中国海洋大学学报, 38(4): 561-566.
- 官文江, 陈新军, 2009. 应用一般线性模型估算鲈、鲆大型灯光围网渔业的捕捞效率[J]. 水产学报, 33(2): 220-228.
- 官文江, 陈新军, 高峰, 等, 2014b. GLM 模型和回归树模型在 CPUE 标准化中的比较分析[J]. 上海海洋大学学报, 23(1): 123-130.
- 胡贯宇, 方舟, 陈新军, 2018. 东太平洋茎柔鱼生活史研究进展[J]. 水产学报, 42(8): 1315-1328.
- 胡振明, 陈新军, 2008. 秘鲁外海茎柔鱼渔场分布与表温及表温距平值关系的初步探讨[J]. 海洋湖沼通报(4): 56-62.
- 蔡毅, 邢岩, 胡丹, 2008. 敏感性分析综述[J]. 北京师范大学学报(自然科学版), 44(1): 9-16.
- 潘玉民, 赵立永, 张全柱, 2012. 瓦斯涌出量的混合 pi-sigma 模糊神经网络预测模型[J]. 计算机应用, 32(4): 1045-1049.
- 戴小杰, 马超, 田思泉, 2011. 印度洋中国大眼金枪鱼延绳钓渔业 CPUE 标准化[J]. 上海海洋大学学报, 20(2): 275-283.
- BENOIT-BIRD K J, GILLY W F, 2012. Coordinated nocturnal behavior of foraging jumbo squid *Dosidicus gigas* [J]. Marine Ecology Progress Series, 455: 211-228.
- BREIMAN L, 2001. Random forests [J]. Machine Learning, 45(1): 5-32.
- CHANG S K, HOYLE S, LIU H I, 2011. Catch rate standardization for yellowfin tuna (*Thunnus albacares*) in Taiwan's distant-water longline fishery in the Western and Central Pacific Ocean, with consideration of target change [J]. Fisheries Research, 107(1/2/3): 210-220.
- CHEN S, BILLINGS S A, GRANT P M, 1990. Non-linear system identification using neural networks [J]. International Journal of Control, 51(6): 1191-1214.
- CHOLEWO T J, ZURADA J M, 1997. Sequential network construction for time series prediction [C] // Proceedings of International Conference on Neural Networks. Houston, TX, USA: IEEE: 2034-2038.
- GARSON G D, 1991. Interpreting neural-network connection weights [J]. AI Expert, 6(4): 46-51.
- GOU J P, ZHAN Y Z, RAO Y B, *et al.*, 2014. Improved pseudo nearest neighbor classification [J]. Knowledge-Based Systems, 70: 361-375.
- HARLEY S J, MYERS R A, DUNN A, 2001. Is catch-per-unit-effort proportional to abundance? [J]. Canadian Journal of Fisheries and Aquatic Sciences, 58(9): 1760-1772.
- HILBORN R, WALTERS C J, 1992. Quantitative Fisheries Stock

- Assessment: Choice, Dynamics and Uncertainty [M]. New York: Chapman and Hall.
- HINTON M G, NAKANO H, 1996. Standardizing catch and effort statistics using physiological, ecological, or behavioral constraints and environmental data, with an application to blue marlin (*Makaira nigricans*) catch and effort data from Japanese longline fisheries in the Pacific [J]. Inter-American Tropical Tuna Commission Bulletin, 21: 171-200.
- MAUNDER M N, 2001. A General Framework for integrating the Standardization of Catch Per Unit of Effort into Stock Assessment Models [J]. Canadian Journal of Fisheries and Aquatic Sciences, 58.10.1139/f01-029.
- MAUNDER M N, HINTON M G, 2006. Estimating relative abundance from catch and effort data, using neural networks [R]. Inter-American Tropical Tuna Commission Special Report 15.
- MAUNDER M N, PUNT A E, 2004. Standardizing catch and effort data: a review of recent approaches [J]. Fisheries Research, 70(2/3): 141-159.
- NASH J E, SUTCLIFFE J V, 1970. River flow forecasting through conceptual models part I-A discussion of principles [J]. Journal of Hydrology, 10(3): 282-290.
- NIGMATULLIN C M, NESIS K N, ARKHIPKIN A I, 2001. A review of the biology of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) [J]. Fisheries Research, 54(1): 9-19.
- OLDEN J D, JOY M K, DEATH R G, 2004. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data [J]. Ecological Modelling, 178(3/4): 389-397.
- ÖZESMI S L, ÖZESMI U, 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction [J]. Ecological Modelling, 116(1): 15-31.
- ROBINSON C J, GÓMEZ-GUTIÉRREZ J, DE LEÓN D A S, 2013. 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, 137: 97-103.
- SALTHAUG A, AANES S, 2003. Catchability and the spatial distribution of fishing vessels [J]. Canadian Journal of Fisheries and Aquatic Sciences, 60(3): 259-268.
- SCHWING F B, MOORE C S, RALSTON S, *et al*, 1999. Record coastal upwelling in the California Current in 1999 [R]. CalCOFI Rep, 41, 148-160.
- TAFUR R, VILLEGAS P, RABÍ M, *et al*, 2001. Dynamics of maturation, seasonality of reproduction and spawning grounds of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in Peruvian waters [J]. Fisheries Research, 54(1): 33-50.
- TAIPE A, YAMASHIRO C, MARIATEGUI L, *et al*, 2001. Distribution and concentrations of jumbo flying squid (*Dosidicus gigas*) off the Peruvian coast between 1991 and 1999 [J]. Fisheries Research, 54(1): 21-32.
- VENABLES W N, DICHMONT C M, 2004. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research [J]. Fisheries Research, 70(2/3): 319-337.
- WANG J T, CHEN X J, CHEN Y, 2018. Projecting distributions of Argentine shortfin squid (*Illex argentinus*) in the Southwest Atlantic using a complex integrated model [J]. Acta Oceanologica Sinica, 37(8): 31-37.
- YE Y M, DENNIS D, 2009. How reliable are the abundance indices derived from commercial catch-effort standardization? [J]. Canadian Journal of Fisheries and Aquatic Sciences, 66(7): 1169-1178.
- ZIEGLER P E, FRUSHER S D, JOHNSON C R, 2003. Space-time variation in catchability of southern rock lobster *Jasus edwardsii* in Tasmania explained by environmental, physiological and density-dependent processes [J]. Fisheries Research, 61(1/2/3): 107-123.

THE CPUE STANDARDIZATION OF *DOSIDICUS GIGAS* IN THE SOUTHEASTERN PACIFIC OCEAN USING BP NEURAL NETWORK

LIN Hong-Yu¹, WANG Jin-Tao^{1, 2, 3, 4, 5}, CHEN Xin-Jun^{1, 2, 3, 4, 5}

(1. College of Marine Sciences of Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Oceanic Fisheries Exploration, Ministry of Agriculture and Rural Affairs, Shanghai 201306, China; 3. National Engineering Research Center for Oceanic Fisheries, Shanghai 201306, China; 4. Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources, Ministry of Education, Shanghai 201306, China; 5. Scientific Observing and Experimental Station of Oceanic Fishery Resources, Ministry of Agriculture and Rural Affairs, Shanghai 201306, China)

Abstract *Dosidicus gigas* is a short-lifespan economic fish, and the resources depends greatly on environmental factors. The production statistics of China's squid fishing fleet in the southeastern Pacific Ocean from 2013 to 2017 were used to standardize the catch per unit effort (CPUE) in BP neural network (BACK Propagation Network) model in combination with marine environmental data (sea surface temperature (SST), sea surface salinity (SSS), chlorophyll *a* concentration). Based on MSE (mean square errors) and ARV (average relative variances), hidden layer nodes 3~10 from the neural network model were compared, from which the 6-9-1 structure was found the optimal model. The Garson algorithm was used to interpret the model results, and it was found that the order of importance of each input layer factor on the abundance of *D. gigas* was chl *a*, SST, longitude, SSS, latitude, and the month. Results show that the distributions of CPUE and standardized CPUE are largely the same. The areas of 80°~85°W and 10°~20°S are suitable for squid fishing operation. Therefore, the BP neural network model can be applied for the CPUE standardization of *D. gigas*, and provide a reference for squid fishing operation.

Key words BP neural network; *Dosidicus gigas* in the southeast Pacific Ocean; resource abundance; environmental factors