

孙霄, 张云雷, 刘笑笑, 等. 两种 GAM 模型对海州湾短吻红舌鲷 (*Cynoglossus joyneri*) 资源分布预测效果的比较研究[J]. 海洋学报, 2020, 42(6): 20–28, doi:10.3969/j.issn.0253-4193.2020.06.003

Sun Xiao, Zhang Yunlei, Liu Xiaoxiao, et al. Evaluation of the prediction effect of two GAMs on the distribution of *Cynoglossus joyneri* in the Haizhou Bay[J]. Haiyang Xuebao, 2020, 42(6): 20–28, doi:10.3969/j.issn.0253-4193.2020.06.003

# 两种 GAM 模型对海州湾短吻红舌鲷 (*Cynoglossus joyneri*) 资源分布预测效果的比较研究

孙霄<sup>1</sup>, 张云雷<sup>1</sup>, 刘笑笑<sup>1</sup>, 程远<sup>3</sup>, 纪毓鹏<sup>1,4</sup>, 任一平<sup>1,2,4</sup>, 薛莹<sup>1,4\*</sup>

(1. 中国海洋大学 水产学院, 山东 青岛 266003; 2. 青岛海洋科学与技术试点国家实验室 海洋渔业科学与食物产出过程功能实验室, 山东 青岛 266237; 3. 近海(大连)生态发展有限公司, 辽宁 大连 116023; 4. 海州湾渔业生态系统教育部野外科学观测研究站, 山东 青岛 266003)

**摘要:** 根据 2011 年及 2013–2018 年春、秋两季在海州湾及其邻近海域进行的底拖网调查数据, 研究该海域短吻红舌鲷 (*Cynoglossus joyneri*) 的资源分布特征及其受环境因子和饵料生物的影响, 并比较了两种模型 (普通 GAM 模型和 PCA-GAM 模型) 对其资源分布的预测效果, 采用交叉验证的方法对模型的预测能力及拟合效果进行评价。结果显示: PCA-GAM 模型的拟合度及预测效果均优于普通 GAM 模型。春、秋两季海州湾短吻红舌鲷资源丰度均呈现南高北低、近岸浅水区大于深水区的分布特征, 因为海州湾南部近岸海域较高的水温利于春、秋季短吻红舌鲷产卵群体性腺发育, 较低的盐度利于其鱼卵及仔鱼的生长发育, 同时, 近岸海域丰富的饵料资源为产卵后的亲体提供大量食物供给。分别应用两种模型预测了 2018 年春季和秋季短吻红舌鲷在海州湾的资源分布, 结果显示, PCA-GAM 模型的预测值与实际调查的结果更为吻合, 预测效果要优于普通 GAM 模型。本研究为今后开展渔业生物空间分布的研究提供了一种新的方法。

**关键词:** GAM 模型; 主成分分析; 海州湾; 黄海中部; 短吻红舌鲷

中图分类号: P714+.5; P931.1

文献标志码: A

文章编号: 0253-4193(2020)06-0020-09

## 1 引言

短吻红舌鲷 (*Cynoglossus joyneri*) 属于鲽形目 (Pleuronectiformes)、舌鲷科 (Cynoglossidae)、舌鲷属 (*Cynoglossus*), 又称焦氏舌鲷、驹舌、乔氏龙舌鱼等, 是亚热带及暖温带浅海底层鱼类, 广泛分布于黄海、渤海、东海和南海 (南至珠江口附近), 日本及朝鲜近海也有分布<sup>[1]</sup>。近年来由于过度捕捞等因素的影响, 导致一些重要经济渔业资源衰退<sup>[2]</sup>, 短吻红舌鲷成为

目前海州湾的优势鱼种之一, 它既是鱼类群落中的优势种<sup>[3-6]</sup>, 同时也是海州湾食物网中的关键种类<sup>[7]</sup>。

GAM 模型 (Generalized Additive Model, GAM) 是研究生物资源量分布的常用方法之一, 但是由于解释变量之间常存在较高的相关性, 因此会导致这些变量存在较高的多重共线性, 从而降低 GAM 模型的预测性能<sup>[8]</sup>。主成分分析 (Principal Component Analysis, PCA) 利用降维的方法, 将多个相关变量转换成若干个主成分 (Principal Component, PCs), 并根据需要选

收稿日期: 2019-08-14; 修订日期: 2019-12-02。

基金项目: 山东省支持青岛海洋科学与技术试点国家实验室重大科技专项 (2018SDKJ0501-2); 国家重点研发计划 (2017YFE0104400); 国家自然科学基金项目 (31772852, 31802301)。

作者简介: 孙霄 (1995—), 女, 山东省菏泽市人, 主要从事渔业生态学研究。E-mail: sunx07102002@163.com

\* 通信作者: 薛莹, 教授, 主要从事食物网营养动力学、鱼类栖息地和空间分布等领域的研究。E-mail: xueying@ouc.edu.cn

择若干主成分能反映大多数原始信息的统计方法<sup>[9]</sup>。因此,基于 PCA 分析得到的 PCs,将其作为 GAM 模型的解释变量,可以消除变量间的共线性,同时能够均衡考虑到所有的重要环境变量。此外,以往的研究多侧重于研究非生物环境因子对物种分布的影响,而忽视了生物因子的重要性。有研究表明,将饵料生物加入解释变量中会提高 GAM 模型的预测能力<sup>[10]</sup>。然而,目前将饵料因子加入解释变量,评估 GAM 模型预测性能的研究较少。

本研究基于 2011 年及 2013–2018 年春、秋季在海州湾及其邻近海域的底拖网调查数据,研究该海域短吻红舌鲷的资源分布特征,并将时空、环境、饵料因子作为解释变量,比较了两种 GAM 模型(普通 GAM 模型和 PCA-GAM 模型)对其资源分布的预测效果,以期为今后开展渔业生物空间分布的研究提供新的方法,也为海州湾短吻红舌鲷的科学管理及资源养护提供基础资料。

## 2 材料与方法

### 2.1 数据来源

样品采自 2011 年及 2013–2018 年春、秋季在海州湾及其邻近海域进行的底拖网调查,采用分层随机取样的方法设计调查站位,调查海域为 34°20′~35°40′N, 119°20′~121°10′E,根据水深、纬度等因素的差异将调查海域分为 5 个区域(图 1),每个航次随机选取一定的站位进行调查(每 10′×10′的方格内设置 1 个站位)。其中,2011 年各航次分别选取 24 个调查站位,其他年份各航次分别选取 18 个调查站位。调查船为 220 kW 单拖渔船,拖速 2~3 kn,拖曳时间约 1 h,囊网网目为 17 mm。在每个调查站位使用 CTD 同步测量底层温度(Sea Bottom Temperature,

SBT)、底层盐度(Sea Bottom Salinity, SBS)、水深等环境数据。样品的采集、处理和分析按照《海洋调查规范》进行。在进行数据分析前对调查数据进行拖速 2 km,拖网时间 1 h 的标准化处理,采用单位面积内的渔获尾数(ind./km<sup>2</sup>)作为相对资源丰度。

### 2.2 GAM 模型构建

GAM 模型是广义线性模型的半参数化扩展形式,可同时针对不同的变量采取不同的策略建模,例如:可对部分预测因子进行线性拟和,而对其他因子通过光滑函数进行拟和,使得变量间的线性和非线性关系得以很好的展现<sup>[11]</sup>。普通 GAM 的构建公式如下:

$$g(Y) = s(\text{Lat}) + s(\text{Lon}) + s(\text{SBT}) + s(\text{SBS}) + s(\text{depth}) + \sum_{i=1}^n s(X_i) + \varepsilon, \quad (1)$$

式中, $g$ 为关联函数, $s$ 为自然立方样条平滑, $\varepsilon$ 为误差项, $Y = \ln(y+1)$ ,其中 $y$ 为相对资源丰度,Lat 为纬度, Lon 为经度, SBT 为底层温度, SBS 为底层盐度, depth 为水深, $X_i$ 为饵料因子。根据短吻红舌鲷摄食生态的研究结果<sup>[12-16]</sup>,春季选择戴氏赤虾(*Metapenaeopsis dalei*)、葛氏长臂虾(*Palaemon graviera*)和日本鼓虾(*Alpheus japonicus*)3 种优势饵料因子,秋季选择葛氏长臂虾和日本鼓虾两种优势饵料因子。

将原始解释变量数据进行 PCA 处理后得到主成分,并选取特征值大于 1 的主成分作为 PCA-GAM 模型的解释变量,新变量间不存在相关性,PCA-GAM 模型的构建如下:

$$g(Y) = \sum_{i=1}^N s(PC_i) + \varepsilon, \quad (2)$$

式中, $PC_i$ 为选取的主成分。

### 2.3 因子筛选与模型拟合

应用方差膨胀因子(Variance Inflation Factor, VIF)

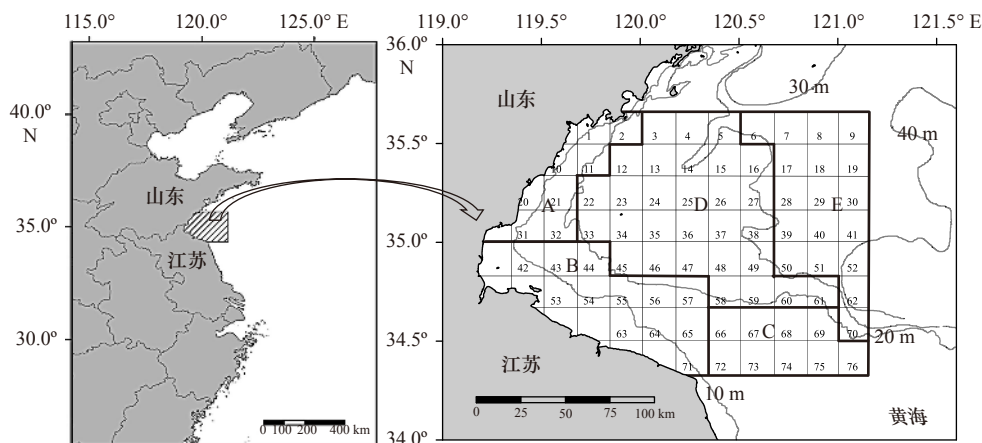


图 1 海州湾及邻近海域调查区域

Fig. 1 Survey areas in Haizhou Bay and adjacent waters

筛选加入模型的解释变量,一般认为 $\sqrt{\text{VIF}}>2$ 即存在多重共线性。利用逐步回归法进行因子筛选,最适模型评估利用赤池信息准则(Akaike Information Criterion, AIC)、方差解释率(Percentage of Variance Explained)、残差偏差(Residual)等指标<sup>[17]</sup>。AIC的计算公式如下:

$$\text{AIC} = 2K - 2\ln(L), \quad (3)$$

式中,  $K$  为模型中参数的数量;  $L$  为模型中极大似然函数值。

#### 2.4 模型比较

采用交叉验证法对两种 GAM 模型的预测性能进行比较。即在总数据中随机选取 80% 的数据作为训练集,剩余的 20% 做为验证集,来验证模型的预测效果,该过程重复 100 次。利用线性回归分析法来描述使用训练集模型的观测值( $O_i$ )与预测值( $P_i$ )之间的线性关系<sup>[18]</sup>,回归方程如下:

$$O_i = a \times P_i + b, \quad (4)$$

式中:  $O_i$  为实际观测值;  $P_i$  为模型预测值;  $a$  为斜率,当  $a$  越接近 1 时,表明模型在预测中的系统偏差越小;  $b$  为截距,当  $b$  越接近 0 时,表明预测值与观测值有相似的空间特征<sup>[19]</sup>;  $R^2$  为决定系数,当  $R^2$  越接近 1 时,表示回归方程的拟合优度越高。

此外还计算了观测值与预测值之间的均方根误差(Root Mean Squared Error, RMSE)<sup>[17]</sup>,该公式的表达形式如下:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2}. \quad (5)$$

#### 2.5 资源分布的预测图

基于 2011 年与 2013–2017 年数据建立模型,然后应用该模型对 2018 年海州湾短吻红舌鳎的相对资源

密度进行预测,利用 Surfer 软件进行绘图,根据预测结果比较两种 GAM 模型预测值与实际调查结果的吻合程度。

上述主成分分析、模型的拟合预测及相关交叉验证均由 R3.2.5 软件完成,其中 PCA 利用 psych 包实现,两种 GAM 均由 gam 包构建。

### 3 结果与分析

#### 3.1 解释变量之间的相关性分析

对海州湾春、秋两季各解释变量进行皮尔逊相关性检验(表 1,表 2),可知春季温度与盐度和水深之间及秋季日本鼓虾与葛氏长臂虾之间都存在极显著的相关性( $p<0.01$ )。利用 VIF 检验的结果均在 1~3 之间(表 3),即春、秋两季所选解释变量均可以加入模型。

#### 3.2 主成分分析

PCA 分析结果表明,海州湾春、秋两季均保留前 3 个主成分,平均累计方差贡献率分别为 65% 和 67%;春季各主成分的方差贡献率分别 31%、19% 和 15%;秋季各主成分的方差贡献率分别为 29%、24% 和 14%。春季 PC1 中,经度与水深存在很高的正载荷,而水温有很高的负载荷;PC2 中戴氏赤虾与葛氏长臂虾存在很高的正载荷;PC3 中纬度有很高的正载荷,而经度有很高的负载荷(表 4)。

秋季,PC1 中纬度、水深存在很高的负载荷,而葛氏长臂虾及日本鼓虾存在很高的正载荷;PC2 中经度与水深存在很高的正载荷,而水温有较高的负载荷;PC3 中盐度存在很高的负载荷(表 4)。

#### 3.3 变量筛选与模型拟合

根据变量筛选结果,春季普通 GAM 最适模型为

表 1 海州湾春季各解释变量之间的皮尔逊相关性检验

Table 1 Pearson correlation test among the interpreted variables in the Haizhou Bay during spring

| 变量    | 经度     | 纬度           | 底温            | 底盐            | 水深            | 戴氏赤虾         | 葛氏长臂虾         | 日本鼓虾          |
|-------|--------|--------------|---------------|---------------|---------------|--------------|---------------|---------------|
| 经度    | -      | <b>0.084</b> | <b>0.000*</b> | <b>0.048*</b> | <b>0.000*</b> | <b>0.060</b> | <b>0.853</b>  | <b>0.478</b>  |
| 纬度    | -0.175 | -            | <b>0.005*</b> | <b>0.060</b>  | <b>0.000*</b> | <b>0.365</b> | <b>0.934</b>  | <b>0.208</b>  |
| 底温    | -0.506 | -0.281       | -             | <b>0.003*</b> | <b>0.000*</b> | <b>0.510</b> | <b>0.950</b>  | <b>0.886</b>  |
| 底盐    | 0.199  | 0.190        | -0.298        | -             | <b>0.007*</b> | <b>0.626</b> | <b>0.854</b>  | <b>0.901</b>  |
| 水深    | 0.510  | 0.429        | -0.695        | 0.272         | -             | <b>0.070</b> | <b>0.435</b>  | <b>0.155</b>  |
| 戴氏赤虾  | 0.190  | 0.092        | -0.067        | 0.050         | 0.183         | -            | <b>0.000*</b> | <b>0.680</b>  |
| 葛氏长臂虾 | 0.019  | 0.008        | -0.006        | 0.019         | 0.079         | 0.489        | -             | <b>0.004*</b> |
| 日本鼓虾  | -0.072 | -0.128       | -0.015        | 0.013         | -0.144        | -0.042       | 0.288         | -             |

注:未加粗数据为相关系数,加粗数据为  $p$  值,\*表示数值在 0.05 置信水平下显著。

表 2 海州湾秋季各解释变量之间的皮尔逊相关性检验

Table 2 Pearson correlation test among the interpreted variables during autumn in the Haizhou Bay

| 变量    | 经度     | 纬度           | 底温           | 底盐           | 水深            | 葛氏长臂虾         | 日本鼓虾          |
|-------|--------|--------------|--------------|--------------|---------------|---------------|---------------|
| 经度    | -      | <b>0.214</b> | <b>0.164</b> | <b>0.195</b> | <b>0.000*</b> | <b>0.792</b>  | <b>0.549</b>  |
| 纬度    | -0.129 | -            | <b>0.874</b> | <b>0.185</b> | <b>0.003*</b> | <b>0.007*</b> | <b>0.000*</b> |
| 温度    | -0.145 | -0.017       | -            | <b>0.133</b> | <b>0.042*</b> | <b>0.531</b>  | <b>0.592</b>  |
| 盐度    | 0.135  | -0.138       | -0.156       | -            | <b>0.110</b>  | <b>0.666</b>  | <b>0.870</b>  |
| 水深    | 0.565  | 0.304        | -0.210       | 0.166        | -             | <b>0.111</b>  | <b>0.120</b>  |
| 葛氏长臂虾 | -0.028 | -0.278       | -0.065       | -0.045       | -0.165        | -             | <b>0.000*</b> |
| 日本鼓虾  | 0.063  | -0.363       | -0.056       | 0.017        | -0.162        | 0.648         | -             |

注:未加粗数据为相关系数,加粗数据为 $p$ 值,\*表示数值在0.05置信水平下显著。

表 3 海州湾春季和秋季各解释变量之间多重共线性的 VIF 检验

Table 3 Multi-collinearity VIF test between explanatory variables during spring and autumn in the Haizhou Bay

| 变量 | 经度    | 纬度    | 底温    | 底盐    | 水深    | 戴氏赤虾  | 葛氏长臂虾 | 日本鼓虾  |
|----|-------|-------|-------|-------|-------|-------|-------|-------|
| 春季 | 2.077 | 1.746 | 2.230 | 1.129 | 2.832 | 1.464 | 1.534 | 1.219 |
| 秋季 | 1.741 | 1.472 | 1.076 | 1.096 | 1.987 | -     | 1.858 | 1.762 |

表 4 海州湾春季和秋季各解释变量的主成分载荷

Table 4 Main component loads of various explanatory variables during spring and autumn in the Haizhou Bay

| 变量    | 春季     |        |        | 秋季     |        |        |
|-------|--------|--------|--------|--------|--------|--------|
|       | PC1    | PC2    | PC3    | PC1    | PC2    | PC3    |
| 经度    | 0.655  | 0.014  | -0.642 | -0.250 | 0.766  | 0.145  |
| 纬度    | 0.432  | -0.166 | 0.815  | -0.623 | -0.248 | 0.434  |
| 底温    | -0.834 | 0.135  | 0.138  | 0.090  | -0.499 | 0.004  |
| 底盐    | 0.485  | -0.074 | 0.058  | -0.081 | 0.450  | -0.758 |
| 水深    | 0.890  | -0.099 | 0.063  | -0.602 | 0.623  | 0.273  |
| 戴氏赤虾  | 0.328  | 0.687  | 0.167  | -      | -      | -      |
| 葛氏长臂虾 | 0.152  | 0.874  | 0.133  | 0.780  | 0.363  | 0.212  |
| 日本鼓虾  | -0.123 | 0.480  | -0.196 | 0.765  | 0.276  | 0.320  |

$\ln(y+1)=s(P. \textit{gravieri})+s(\textit{Lon})+s(\textit{Lat})+\varepsilon$ , 模型中葛氏长臂虾的贡献率最大为 45.039%, 其次是经度和纬度, 分别为 10.155% 和 8.062%(表 5)。春季 PCA-GAM 最适模型为  $\ln(y+1)=s(\textit{PC2})+\varepsilon$ , 模型中 PC2 贡献率为 37.287%(表 5)。

秋季普通 GAM 最适模型为  $\ln(y+1)=s(\textit{Lat})+s(P. \textit{gravieri})+s(\textit{Lon})+s(\textit{SBS})+\varepsilon$ , 模型中纬度有最大的贡献率, 为 35.034%, 其次为经度、盐度和葛氏长臂虾, 分别为 16.395%、15.170% 和 11.156%(表 5)。PCA-GAM 最适模型为  $\ln(y+1)=s(\textit{PC1})+s(\textit{PC2})+s(\textit{PC3})+\varepsilon$ , 模型中

PC1 的贡献率最高为 39.252%, 其次为 PC2 和 PC3, 分别为 12.585% 和 12.449%(表 5)。

春、秋两季普通 GAM 模型的最终累计解释率分别为 63.256%、77.755%; PCA-GAM 模型的最终累计解释率分别为 37.287%、64.286%。

### 3.4 短吻红舌鲷相对资源量与解释变量的关系

春季, 短吻红舌鲷相对资源量随纬度的增加呈现先增加后降低的趋势(图 2a); 秋季, 相对资源量的变化总体随纬度的增加呈现下降趋势(图 2A)。春季相对资源量随经度的变化趋势与秋季相似, 均随经度的

表 5 海州湾春季和秋季两种 GAM 模型拟合结果及各解释变量的重要性

Table 5 The fitting results of two GAMs and the importance of each explanatory variable

| 季节 | 模型      | 因子                      | 累计偏差解释率/% | 贡献率/%  | AIC     |
|----|---------|-------------------------|-----------|--------|---------|
| 春季 | 普通GAM   | +s( <i>P.gravieri</i> ) | 45.039    | 45.039 | 166.936 |
|    |         | +s(Lon)                 | 55.194    | 10.155 | 165.133 |
|    |         | +s(Lat)                 | 63.256    | 8.062  | 163.604 |
| 春季 | PCA-GAM | +s(PC2)                 | 37.287    | 37.287 | 173.255 |
| 秋季 | 普通GAM   | +s(Lat)                 | 35.034    | 35.034 | 186.233 |
|    |         | +s( <i>P.gravieri</i> ) | 46.190    | 11.156 | 184.821 |
|    |         | +s(Lon)                 | 62.585    | 16.395 | 174.697 |
|    |         | +s(SBS)                 | 77.755    | 15.170 | 156.598 |
| 秋季 | PCA-GAM | +s(PC1)                 | 39.252    | 39.252 | 182.904 |
|    |         | +s(PC2)                 | 51.837    | 12.585 | 179.265 |
|    |         | +s(PC3)                 | 64.286    | 12.449 | 172.370 |

增加先下降,到 120.5°E 后又继续上升(图 2b, 图 2B)。

底层盐度是影响秋季短吻红舌鲷相对资源量分布的重要因素,相对资源量随盐度的增加先下降后上升,在盐度达到 28 之后,又急剧下降(图 2C)。

饵料因子中,葛氏长臂虾饵料对春、秋季短吻红舌鲷的相对资源量均有重要影响(图 2)。春季,短吻红舌鲷的相对资源量随葛氏长臂虾生物量的增加先上升后又小幅下降,在生物量达到 65 g/km<sup>2</sup> 之后又继续上升(图 2c)。秋季,其相对资源量随葛氏长臂虾的增加上升幅度较大,在生物量达到 10 g/km<sup>2</sup> 之后变化较平缓(图 2D)。

### 3.5 模型预测效果的比较

两种 GAM 模型交叉验证的结果表明,春季和秋季 PCA-GAM 模型的均方根误差分别为 2.214、2.216,小于普通 GAM 模型(2.401 和 3.171),其斜率及决定系数也都更接近 1,截距更接近于 0(图 3,表 6)。同时,春、秋季 PCA-GAM 模型交叉验证的决定系数  $R^2$  分别为 0.521 和 0.634,高于普通 GAM 模型(0.451 和 0.549)。上述结果均表明,PCA-GAM 模型比普通 GAM 模型具有更优的拟合效果。

两种 GAM 模型对 2018 年春季和秋季海州湾短吻红舌鲷相对资源丰度的预测结果如图 4 所示。秋

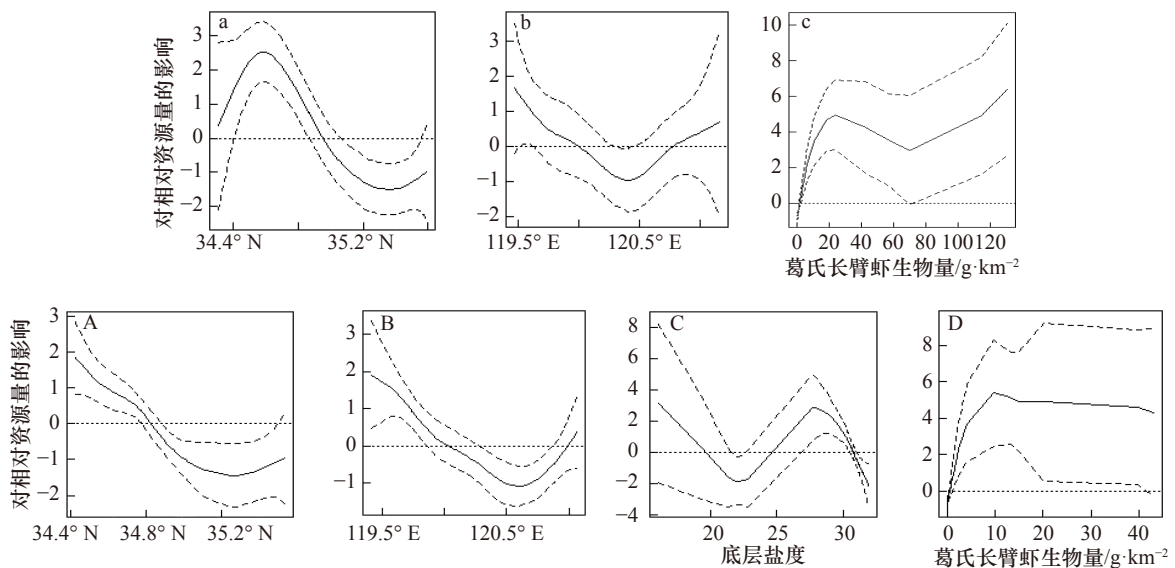


图 2 海州湾春季和秋季各影响因子与短吻红舌鲷相对资源量之间的关系(a-c. 春季; A-D. 秋季)

Fig. 2 Effects of factors on *C. joyneri* relative abundance in regular GAM in Haizhou Bay (a-c. spring; A-D. autumn)

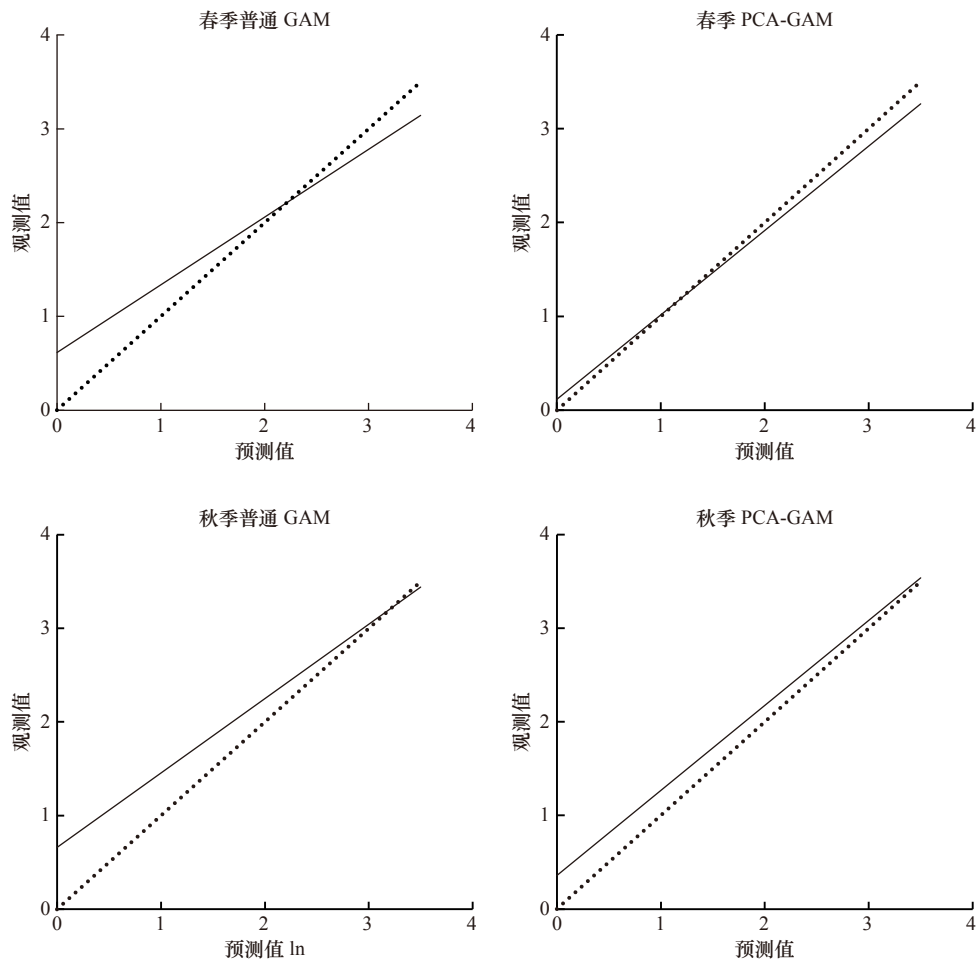


图3 海州湾春季和秋季短吻红舌鲷相对资源丰度预测值与观测值的交叉验证

Fig. 3 Cross validation between predicted and observed abundance of *C. joyneri* during spring and autumn in the Haizhou Bay

实线为交叉验证的平均效应, 点线为对角线

The solid line is the mean of cross validation results, the dashed line is the diagonal line

表6 海州湾春季和秋季两种 GAM 模型交叉验证结果

Table 6 Cross-validation results of two GAM models during spring and autumn in the Haizhou Bay

| 季节 | 模型      | 斜率    | 截距    | 决定系数 $R^2$ | 均方根误差 |
|----|---------|-------|-------|------------|-------|
| 春季 | 普通GAM   | 0.722 | 0.617 | 0.451      | 2.401 |
|    | PCA-GAM | 0.900 | 0.116 | 0.521      | 2.214 |
| 秋季 | 普通GAM   | 0.794 | 0.662 | 0.549      | 3.171 |
|    | PCA-GAM | 0.908 | 0.360 | 0.634      | 2.216 |

季相对资源丰度整体高于春季,且春、秋两季海州湾短吻红舌鲷的分布都呈南高北低的趋势,并与海州湾短吻红舌鲷资源丰度的实际分布大致相符,而且春、秋两季 PCA-GAM 模型预测的资源丰度与观测值更加吻合(图4)。

#### 4 讨论

本研究发现春、秋两季海州湾短吻红舌鲷资源的

分布均呈南高北低、近岸浅水区大于深水区的特征,该分布特征与短吻红舌鲷产卵洄游密切相关。短吻红舌鲷每年1-3月份在深水区越冬,4-10月为繁殖产卵时间<sup>[20]</sup>,春末夏初当海水表层温度水温回升至20℃以上时,开始在近岸10m以浅海域进行产卵,生殖期其亲鱼适宜水温为24.0~28.0℃,盐度为24.0~29.0,鱼卵与仔鱼的适宜温、盐范围分别为20~29℃和20.0~33.0<sup>[21-22]</sup>。本研究的调查数据显示,海州湾

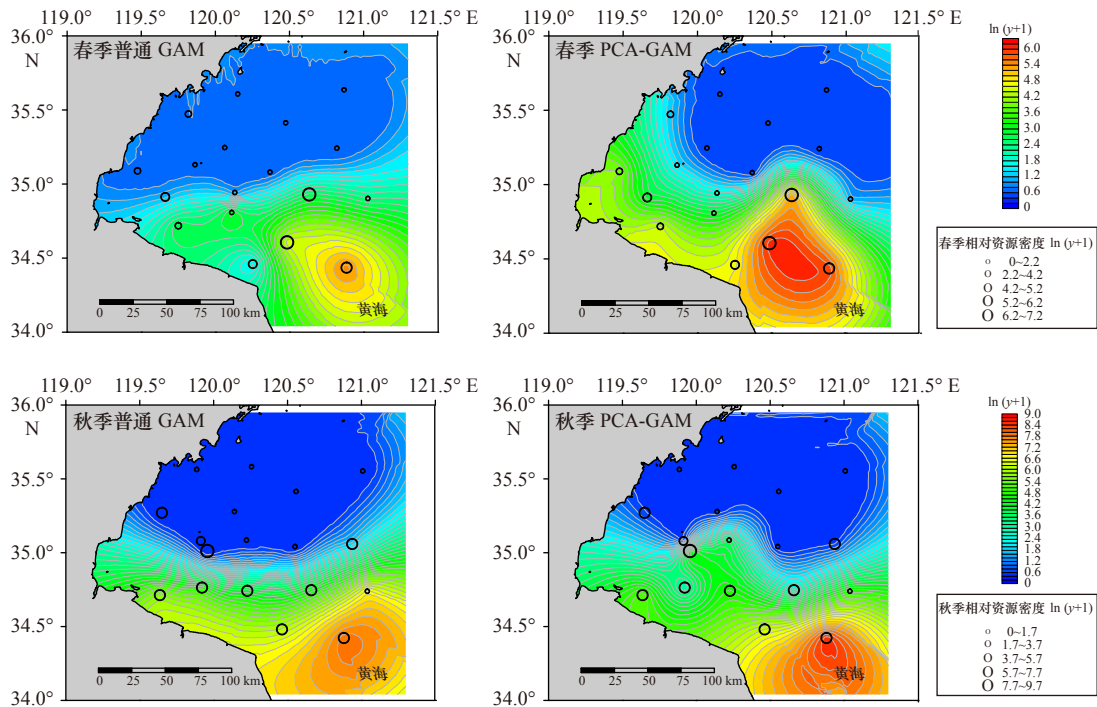


图4 2018年海州湾春、秋季基于普通GAM模型和PCA-GAM模型的短吻红舌鲷相对资源丰度预测值与观测值的叠加图  
Fig. 4 Overlapping maps of prediction and observations of relative abundance of *C. joyneri* in Haizhou Bay based on regular GAM model and PCA-GAM model during spring and autumn in 2018

春、秋季南部近岸水域表层温度范围分别为14.0~18.0℃和18.3~22.1℃,盐度范围分别为29.4~31.8和27.9~30.7,较高的水温利于暖温性鱼类性腺和鱼卵及仔鱼的发育,缩短其发育周期,提高卵子孵化率及仔鱼成活率。春季海州湾南部海域水温回暖较北部早且持续时间长,近岸水域受地表径流、降雨及水团影响大,因此盐度低于北部深水区;此外,其饵料生物,如葛氏长臂虾、日本鼓虾和戴氏赤虾,在近岸浅海水域均有大量分布<sup>[23-25]</sup>,为产卵后的亲鱼提供充足的食物资源。因此,南部近岸水域的生境条件更适宜春季短吻红舌鲷的繁殖。秋季短吻红舌鲷的资源量分布总体也呈现南高北低的趋势,但相较于春季来说,有向北移动的趋势,一方面由于产卵后的亲体陆续进入产卵场附近的深水区索饵,另一方面短吻红舌鲷的仔稚鱼也有向近岸较深水域育幼场索饵洄游的特征<sup>[20]</sup>。综上所述,海州湾短吻红舌鲷产卵洄游的习性对其分布及集群影响较大。

本研究比较分析了两种GAM模型在研究短吻红舌鲷相对资源量与环境及饵料因子关系方面的效果。研究表明,PCA-GAM模型的表现要优于普通GAM模型,说明PCA降低了解释变量间的相关性,提高了模型的稳定性和预测精度;另一方面,在PCA-GAM

模型的解释变量中均包含了具有高正载荷的生物因子主成分,表明饵料因子的加入可以提高模型的预测性能。另外,交叉验证结果显示,加入饵料因子的普通GAM模型预测效果优于仅以环境因子作为变量的普通GAM模型<sup>[17]</sup>,进一步佐证了模型预测精度的提高与饵料因子的加入有关。

由于本研究的资源调查受一定的空间及采样点的限制,模型中响应变量和预测因子之间的关系及物种潜在分布估计可能受到影响<sup>[26-27]</sup>。普通GAM模型的拟合程度不高,反映出春季的某些重要因子没有被选择在原始解释变量中,在今后研究中应考虑更多与鱼类资源分布关系密切的影响因子。

综上所述,本研究比较了两种GAM模型对海州湾短吻红舌鲷资源分布的预测效果,当解释变量之间存在较高的相关性时,建议使用PCA-GAM模型对物种分布进行预测,但在实际应用中,还应根据数据可获得性、研究对象及实验目的的不同选择合适的模型。同时,生物因子的加入有助于提高模型的预测性能,因此,在以后的研究中建议考虑相关生物因素<sup>[10,27-28]</sup>,旨在深入揭示物种分布与影响因子之间的内在机制。

#### 参考文献:

- [1] 胡保存,齐遵利,高文斌.河北沿海焦氏舌鲷资源现状分析[J].河北渔业,2016(12):13-16.

- Hu Baocun, Qi Zunli, Gao Wenbin. Analysis on resource status of the *Cynoglossus joyneri* in coastal waters of Hebei Province[J]. *Hebei Fisheries*, 2016(12): 13–16.
- [2] 张英佳, 张云, 景昕蒂. 中国海湾渔业的现状及问题初探[J]. 长春理工大学学报: 社会科学版, 2012, 25(12): 63–64.  
Zhang Yingjia, Zhang Yun, Jing Xindi. A preliminary study on the current situation and problems of Gulf Fisheries in China[J]. *Journal of Changchun University of Science and Technology: Social Sciences Edition*, 2012, 25(12): 63–64.
- [3] 任晓明, 徐宾铎, 张崇良, 等. 海州湾及邻近海域鱼类群落的营养功能群及其动态变化[J]. *中国水产科学*, 2019, 26(1): 141–150.  
Ren Xiaoming, Xu Binduo, Zhang Chongliang, et al. The composition of and variations in the trophic guilds of fish assemblages in Haizhou Bay and adjacent waters[J]. *Journal of Fishery Sciences of China*, 2019, 26(1): 141–150.
- [4] 牟秀霞, 徐宾铎, 薛莹, 等. 黄海南部近岸海域鱼类群落结构与区系划分[J]. 水产学报, 2017, 41(11): 1734–1743.  
Mou Xiuxia, Xu Binduo, Xue Ying, et al. The fish assemblage structure and fauna discrimination in the coastal waters of southern Yellow Sea[J]. *Journal of Fisheries of China*, 2017, 41(11): 1734–1743.
- [5] 于海婷, 丁月晏, 钱薇薇, 等. 荣成湾渔业资源群落结构季节变化特征[J]. 海洋湖沼通报, 2013(2): 67–75.  
Yu Haiting, Ding Yueyan, Xian Weiwei, et al. Seasonal characteristics of fishery community structures in the Rongcheng Bay[J]. *Transactions of Oceanology and Limnology*, 2013(2): 67–75.
- [6] 张波, 袁伟, 王俊. 崂山湾春季鱼类群落的摄食生态及其主要种类[J]. 中国水产科学, 2015, 22(4): 820–827.  
Zhang Bo, Yuan Wei, Wang Jun. Feeding ecology of the dominant fish species in spring in Laoshan Bay[J]. *Journal of Fishery Sciences of China*, 2015, 22(4): 820–827.
- [7] 麻秋云, 韩东燕, 刘贺, 等. 应用稳定同位素技术构建胶州湾食物网的连续营养谱[J]. 生态学报, 2015, 35(21): 7207–7218.  
Ma Qiuyun, Han Dongyan, Liu He, et al. Construction of a continuous trophic spectrum for the food web in Jiaozhou Bay using stable isotope analyses[J]. *Acta Ecologica Sinica*, 2015, 35(21): 7207–7218.
- [8] Francis M P, Morrison M A, Leathwick J, et al. Predictive models of small fish presence and abundance in northern New Zealand harbours[J]. *Estuarine, Coastal and Shelf Science*, 2005, 64(2/3): 419–435.
- [9] Ahmadi-Nedushan B, St-Hilaire A, Bérubé M, et al. A review of statistical methods for the evaluation of aquatic habitat suitability for in-stream flow assessment[J]. *River Research and Applications*, 2006, 22(5): 503–523.
- [10] Xue Ying, Tanaka K, Yu Huaming, et al. Using a new framework of two-phase generalized additive models to incorporate prey abundance in spatial distribution models of juvenile slender lizardfish in Haizhou Bay, China[J]. *Marine Biology Research*, 2018, 14(5): 508–523.
- [11] Wood S N, Augustin N H. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling[J]. *Ecological Modelling*, 2002, 157(2/3): 157–177.
- [12] 罗秉征, 韦晟, 窦硕增. 长江口鱼类食物网与营养结构的研究[J]. 海洋科学集刊, 1997(38): 143–153.  
Luo Bingzheng, Wei Sheng, Dou Shuozeng. Study on food web and trophic structure of fish in the Changjiang River Estuary[J]. *Studia Marina Sinica*, 1997(38): 143–153.
- [13] 杨纪明. 渤海鱼类的食性和营养级研究[J]. 现代渔业信息, 2001, 16(10): 10–19.  
Yang Jiming. A study on food and trophic levels of Bohai Sea fish[J]. *Modern Fisheries Information*, 2001, 16(10): 10–19.
- [14] 窦硕增, 杨纪明, 陈大刚. 渤海石鲈、星鲈、高眼鲈及焦氏舌鲷的食性[J]. 水产学报, 1992, 16(2): 162–166.  
Dou Shuozeng, Yang Jiming, Chen Dagang. Food habits of stone flounder, spotted flounder, high-eyed flounder and red tongue sole in the Bohai Sea[J]. *Journal of Fisheries of China*, 1992, 16(2): 162–166.
- [15] 张波. 渤海鱼类的食物关系[J]. 渔业科学进展, 2018, 39(3): 11–22.  
Zhang Bo. Feeding ecology of fishes in the Bohai Sea[J]. *Progress in Fishery Sciences*, 2018, 39(3): 11–22.
- [16] 韦晟, 姜卫民. 黄海鱼类食物网的研究[J]. 海洋与湖沼, 1992, 23(2): 182–192.  
Wei Sheng, Jiang Weimin. Study on food web of fishes in the Yellow Sea[J]. *Oceanologia et Limnologia Sinica*, 1992, 23(2): 182–192.
- [17] 栾静, 张崇良, 徐宾铎, 等. 海州湾双斑螯栖息分布特征与环境因子的关系[J]. 水产学报, 2018, 42(6): 889–901.  
Luan Jing, Zhang Chongliang, Xu Binduo, et al. Relationship between catch distribution of Portunid crab (*Charybdis bimaculata*) and environmental factors based on three species distribution models in Haizhou Bay[J]. *Journal of Fisheries of China*, 2018, 42(6): 889–901.
- [18] 吴建辉, 戴黎斌, 戴小杰, 等. GAM模型和BRT模型在长江口鱼类群落多样性预测中的比较[J]. 应用生态学报, 2019, 30(2): 644–652.  
Wu Jianhui, Dai Libin, Dai Xiaojie, et al. Comparison of generalized additive model and boosted regression tree in predicting fish community diversity in the Yangtze River Estuary, China[J]. *Chinese Journal of Applied Ecology*, 2019, 30(2): 644–652.
- [19] Chang J H, Chen Yong, Holland D, et al. Estimating spatial distribution of American lobster *Homarus americanus* using habitat variables[J]. *Marine Ecology Progress*, 2010, 420: 145–156.
- [20] 金显仕, 赵宪勇, 孟田湘, 等. 黄、渤海生物资源与栖息环境[M]. 北京: 科学出版社, 2005: 327–341.  
Jin Xianshi, Zhao Xianyong, Meng Tianxiang, et al. Biological Resources and Habitation Environment in Yellow Sea and Bohai Sea[M]. Beijing: Science Press, 2005: 327–341.
- [21] 杨东莱, 吴光宗, 庞鸿艳. 渤海湾半滑舌鲷及焦氏舌鲷的鱼卵和仔稚鱼的形态[J]. 海洋科学, 1983(2): 29–32.  
Yang Donglai, Wu Guangzong, Pang Hongyan. The morphology of the early stage of longue fishes *Cynoglossus semilaevis* Günther and *C. joyneri* Günther, in the Bohai Bay[J]. *Marine Sciences*, 1983(2): 29–32.
- [22] 杨东莱, 吴光宗, 庞鸿艳. 渤海湾的半滑舌鲷及焦氏舌鲷的鱼卵和仔、稚鱼的季节分布[J]. 生态学杂志, 1984(3): 30–33.

- Yang Donglai, Wu Guangzong, Pang Hongyan. Seasonal distribution of eggs and larvae of tongue sole *Cynoglossus semilaevis* Günther and *C. joyneri* Günther, in the Bohai Bay[J]. Chinese Journal of Ecology, 1984(3): 30–33.
- [23] 赵蒙蒙, 徐兆礼. 海州湾南部海域不同季节虾类数量及其分布特征[J]. 海洋通报, 2012, 31(1): 38–44.  
Zhao Mengmeng, Xu Zhaoli. Abundance and distribution of shrimps in different seasons in the southern Haizhou Bay of Jiangsu Province[J]. Marine Science Bulletin, 2012, 31(1): 38–44.
- [24] 卢衍尔, 朱文斌, 梁君, 等. 韭山列岛海域虾类群落结构与海洋环境因子的关系[J]. 浙江大学学报: 理学版, 2019, 46(1): 65–77.  
Lu Kaner, Zhu Wenbin, Liang Jun, et al. The relationship between shrimps community structure and environmental factors in Jiushan Islands waters[J]. Journal of Zhejiang University: Science Edition, 2019, 46(1): 65–77.
- [25] 卢占晖, 周永东, 朱文斌, 等. 浙江沿岸海域虾类优势种空间生态位分析[J]. 海洋学报, 2018, 40(2): 77–86.  
Lu Zhanhui, Zhou Yongdong, Zhu Wenbin, et al. Spatial niches analysis of dominant shrimp species in the coastal area of Zhejiang Province[J]. Acta Oceanologica Sinica, 2018, 40(2): 77–86.
- [26] Guisan A, Jr. Edwards T C, Hastie T Generalized linear and generalized additive models in studies of species distributions: setting the scene[J]. Ecological Modelling, 2002, 157(2/3): 89–100.
- [27] 许仲林, 彭焕华, 彭守璋. 物种分布模型的发展及评价方法[J]. 生态学报, 2015, 35(2): 557–567.  
Xu Zhonglin, Peng Huanhua, Peng Shouzhang. The development and evaluation of species distribution models[J]. Acta Ecologica Sinica, 2015, 35(2): 557–567.
- [28] Yatsu A, Watanabe T, Ishida M, et al. Environmental effects on recruitment and productivity of Japanese sardine *Sardinops melanostictus* and chub mackerel *Scomber japonicus* with recommendations for management[J]. Fisheries Oceanography, 2005, 14(4): 263–278.

## Evaluation of the prediction effect of two GAMs on the distribution of *Cynoglossus joyneri* in the Haizhou Bay

Sun Xiao<sup>1</sup>, Zhang Yunlei<sup>1</sup>, Liu Xiaoxiao<sup>1</sup>, Cheng Yuan<sup>3</sup>, Ji Yupeng<sup>1,4</sup>, Ren Yiping<sup>1,2,4</sup>, Xue Ying<sup>1,4</sup>

(1. Fisheries College, Ocean University of China, Qingdao 266003, China; 2. Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266237, China; 3. Offshore (Dalian) Ecological Development Co. Ltd., Dalian 116023, China; 4. Field Observation and Research Station of Haizhou Bay Fishery Ecosystem, Ministry of Education, Qingdao 266003, China)

**Abstract:** Based on the bottom trawl surveys in the Haizhou Bay and adjacent waters during spring and autumn of 2011 and 2013–2018, the performance of regular GAM and PCA-GAM was compared, and the distribution of *Cynoglossus joyneri* in this area was predicted. The predictive ability and fitting effect of the two GAMs were evaluated by cross-validation. The results showed that the goodness of fit and prediction effects of PCA-GAM were better than those of regular GAM. In spring and autumn, the abundance of *C. joyneri* in the southern waters was higher than that in the northern waters, and the abundance in the near-shore shallow waters was larger than that in the deep waters. The higher water temperature in the coastal waters of the southern Haizhou Bay was conducive to the development of gonads for the spawning groups during spring and autumn. The lower salinity was conducive to the growth and development of fish eggs and larvae. At the same time, the abundant prey resources in the coastal waters provides a large amount of food for it after spawning period. In this study, two GAMs were used to predict the resource distribution of *C. joyneri* in the Haizhou Bay in spring and autumn of 2018. The results showed that the predicted abundance by PCA-GAM were more consistent with the actual catches, and the performance of PCA-GAM was better than the regular GAM. This study provides a new method for studying the spatial distribution of marine organisms in the future.

**Key words:** GAM; principal component analysis; Haizhou Bay; central Yellow Sea; *Cynoglossus joyneri*