

文章编号: 1674 - 5566(2012)03 - 0425 - 06

基于信息增益法选取柔鱼中心渔场的关键水温因子

易倩¹, 陈新军^{1,2,3}

(1. 上海海洋大学 海洋科学学院, 上海 201306; 2. 上海海洋大学 大洋生物资源开发和利用上海市高校重点实验室, 上海 201306; 3. 上海海洋大学 大洋渔业资源可持续开发省部共建教育部重点实验室, 上海 201306)

摘要: 西北太平洋海域是我国鱿钓渔船重要的作业渔场之一。影响渔场的环境因子是渔业海洋学的重要研究内容, 它有助于理解渔场形成机理和指导渔业生产。根据 1998 - 2008 年的 8 - 10 月在北太平洋海域(150°E ~ 170°E)不同水层的温度和我国鱿钓船的生产资料, 对柔鱼作业渔场分布及其与不同水层(海表面、55 m、105 m)的温度、水温垂直结构(0 ~ 55 m 和 55 ~ 105 m)进行分析, 并利用信息增益技术计算作业次数对应的各分类属性的信息增益值。结果表明, 8 - 10 月中心渔场的最适海表面温度呈现季节性变化, 55 m 水层和 105 m 水层最适温度在 9 月份相对较低, 而对应的 0 ~ 55 m 的温度梯度较大。信息增益值结果表明, 影响中心渔场的关键水温因子依次是 0 ~ 55 m 水层温度梯度、55 m 水层温度、海表面温度、105 m 水层温度和 55 ~ 105 m 水层温度梯度。

研究亮点: 柔鱼是我国远洋鱿钓渔业的重要捕捞对象。利用信息增益技术分析了影响中心渔场的主要环境因子, 影响程度依次为 0 ~ 55 m 水层温度梯度、55 m 水层温度、海表面温度、105 m 水层温度和 55 ~ 105 m 水层温度梯度, 研究结果对指导渔业生产具有重要的意义。

关键词: 信息增益; 水温因子; 柔鱼; 北太平洋

中图分类号: S 932.8

文献标志码: A

柔鱼(*Ommastrephes bartramii*)作为大洋性种类, 广泛分布于北太平洋整个海域, 该海域是中国、日本、韩国等国的重要鱿钓作业渔场。北太平洋鱿钓渔业已成为我国远洋渔业的支柱^[1]。西北太平洋柔鱼捕捞季节通常从 5 月开始, 8 月进入盛渔期, 11 - 12 月渔汛结束^[2]。北太平洋柔鱼渔场由黑潮和亲潮交汇形成, 因此在影响渔场分布的环境因素中, 温度是一项重要的因素。水温不仅显著地影响个体性腺发育速度, 同时也约束群体的行动和空间分布, 是重要的非生物性预报指标^[3]。国内外学者对分布在西北太平洋海域的柔鱼资源分布、渔场与环境条件(特别是表层水温)的关系进行过研究^[4-10]。研究表明, 各海区柔鱼群体大小及其资源分布不一, 作业渔场

的形成与表温、各水层温度和温度梯度存在一定的关系。

柔鱼属于大洋性暖水种, 存在昼夜垂直洄游现象, 因此水温分布及其垂直结构可能对柔鱼的分布及渔场形成会产生一定的影响^[1,8], 但是, 有关水温因子(包括不同水温结构)对北太平洋海域(150°E ~ 170°E)柔鱼中心渔场分布的影响, 还没有相关报道。为此, 本文根据 1998 - 2008 年 8 - 10 月我国鱿钓船在北太平洋海域生产数据及其不同水层的温度、温度梯度资料, 利用信息增益法计算各温度因子对应的信息增益值的大小, 来确定影响中心渔场的关键温度因子, 为科学寻找中心渔场提供依据。

收稿日期: 2011-11-11 修回日期: 2011-12-17

基金项目: 国家高技术研究发展计划项目(2007AA092201, 2007AA092202); 国家发改委专项(2060403); 上海市捕捞学重点学科建设项目(S30702)

作者简介: 易倩(1985—), 女, 硕士研究生, 研究方向为渔业资源学。E-mail: m100302460@stmail.shou.edu.cn

通讯作者: 陈新军, E-mail: xjchen@shou.edu.cn

1 材料与方法

1.1 材料

本文的数据采用 1998 - 2008 年我国北太平洋柔鱼捕捞生产日志,内容包括日期、经纬度、产量、作业次数和平均日产量,空间分辨率为 $0.5^\circ \times 0.5^\circ$ (即一个渔区)。

水温资料,包括海表面温度 (T_0)、55 m 水层温度 (T_{55}) 和 105 m 水层温度 (T_{105}) 来自美国哥伦比亚海洋环境数据库,网址是 <http://iridl.ldeo.columbia.edu/>。时间跨度为 1998 - 2008 年,时间分辨率为月,空间分辨率为 $0.5^\circ \times 0.5^\circ$ 。

1.2 分析方法

将水温数据进行预处理,计算各年度 8 ~ 10 月的 0 ~ 55 m 水层温度梯度 (T_{0-55}), 55 ~ 105 m 水层温度梯度 (T_{55-105})。作业次数通常可作为鱼类出现或鱼类利用情况的指标^[11]。分析作业次数与各月各渔区的 T_0 、 T_{55} 、 T_{105} 、 T_{0-55} 和 T_{55-105} 之间的关系。

1.3 信息增益的计算

信息增益技术 (IGT) 是决策树分类方法的一种算法,也是最常用的一种算法^[12]。基于信息增益法决策树生成的主要计算集中在属性的选择上,即为信息增益值的计算。在决策支持系统中,一般用户更希望看到简单而又条理清晰的决策树,所以采用信息增益法生成决策树能取得较好的效果^[13]。本文利用 IGT,计算柔鱼作业次数对应的各分类属性 (水温环境因子) 的信息增益值。根据水温环境数据 (T_0 、 T_{55} 、 T_{105} 、 T_{0-55} 、 T_{55-105}) 计算作业次数对应环境属性的信息增益值,以此来反映每个环境因子对渔场的影响程度。其具体步骤为^[14]:

(1) 求出作业船次的信息期望 I :

$$I(S_1, S_2, \dots, S_m) = - \sum_{i=1}^m \frac{S_i}{S} \log \frac{S_i}{S} \quad (1)$$

式中: m 为作业船次不同的属性值个数; S_i 为作业船次为第 i 个属性值的记录条数; S 为样本总数。

(2) 求出每个属性对应于作业船次分类的熵 $E_1(A)$:

$$E_1(A) = \sum_{j=1}^v \frac{S_{1j} + S_{2j} + \dots + S_{mj}}{S} I(S_{1j}, S_{2j}, \dots, S_{mj}) \quad (2)$$

$$I(S_{1j}, S_{2j}, \dots, S_{mj}) = - \sum_{i=1}^m \frac{S_{ij}}{S_j} \log \frac{S_{ij}}{S_j} \quad (3)$$

式中: v 为属性 A 不同属性值的个数; S_{ij} 为属性 A 值为 A_j 且作业船次为第 i 个属性值的记录条数; S 为样本总数; $I(S_{1j}, S_{2j}, \dots, S_{mj})$ 为属性 A 取值 A_j 时对应的作业船次分类的信息期望。

(3) 计算每个属性对应于作业船次分类信息增益 $G_1(A)$ 。

$$G_1(A) = I(S_1, S_2, \dots, S_m) - E_1(A) \quad (4)$$

2 结果

2.1 各月份作业次数与各温度因素的关系分析

统计结果分析发现,8 月柔鱼作业渔场的适宜水温因子: T_0 为 16 ~ 20 $^\circ\text{C}$, 约占总作业次数的 79.06%, 其中 17 ~ 18 $^\circ\text{C}$ 约占总作业次数的 28.62%; T_{55} 为 4 ~ 10 $^\circ\text{C}$, 约占总作业次数的 74.57%; T_{105} 为 4 ~ 10 $^\circ\text{C}$, 约占总作业次数的 78.86%; T_{0-55} 为 0.16 ~ 0.22 $^\circ\text{C}/\text{m}$, 约占总作业次数的 63.22%; T_{55-105} 为 0.02 ~ 0.06 $^\circ\text{C}/\text{m}$, 约占总作业次数的 79.91% (图 1)。

9 月柔鱼作业渔场的适宜水温因子: T_0 为 15 ~ 19 $^\circ\text{C}$, 约占总作业次数的 88.02%; T_{55} 为 4 ~ 9 $^\circ\text{C}$, 约占总作业次数的 74.14%; T_{105} 为 2 ~ 6 $^\circ\text{C}$, 约占总作业次数的 89.23%; T_{0-55} 为 0.19 ~ 0.24 $^\circ\text{C}/\text{m}$, 约占总作业次数的 50.16%; T_{55-105} 为 0.01 ~ 0.06 $^\circ\text{C}/\text{m}$, 约占总作业次数的 82.75% (图 2)。

10 月柔鱼作业渔场的适宜水温因子: T_0 为 13 ~ 15 $^\circ\text{C}$, 约占总作业次数的 78.54%; T_{55} 为 6 ~ 11 $^\circ\text{C}$, 约占总作业次数的 75.23%; T_{105} 为 3 ~ 6 $^\circ\text{C}$, 约占总作业次数的 76.16%; T_{0-55} 为 0.09 ~ 0.14 $^\circ\text{C}/\text{m}$, 约占总作业次数的 58.05%; T_{55-105} 为 0.04 ~ 0.10 $^\circ\text{C}/\text{m}$, 约占总作业次数的 88.73% (图 3)。

2.2 信息增益分析

各属性分类结果见表 1。 T_0 、 T_{55} 、 T_{105} 按 1 $^\circ\text{C}$ 间隔进行划分区间, T_{0-55} 、 T_{55-105} 和作业次数分别按 0.1 $^\circ\text{C}/\text{m}$ 、0.01 $^\circ\text{C}/\text{m}$ 和 100 次的间隔进行划分区间。

信息增益分析结果认为 (表 2), T_0 、 T_{55} 、 T_{105} 、 T_{0-55} 和 T_{55-105} 对应与作业次数的信息增益分别是 0.454 8、0.526 9、0.410 0、0.569 3 和 0.387 2。

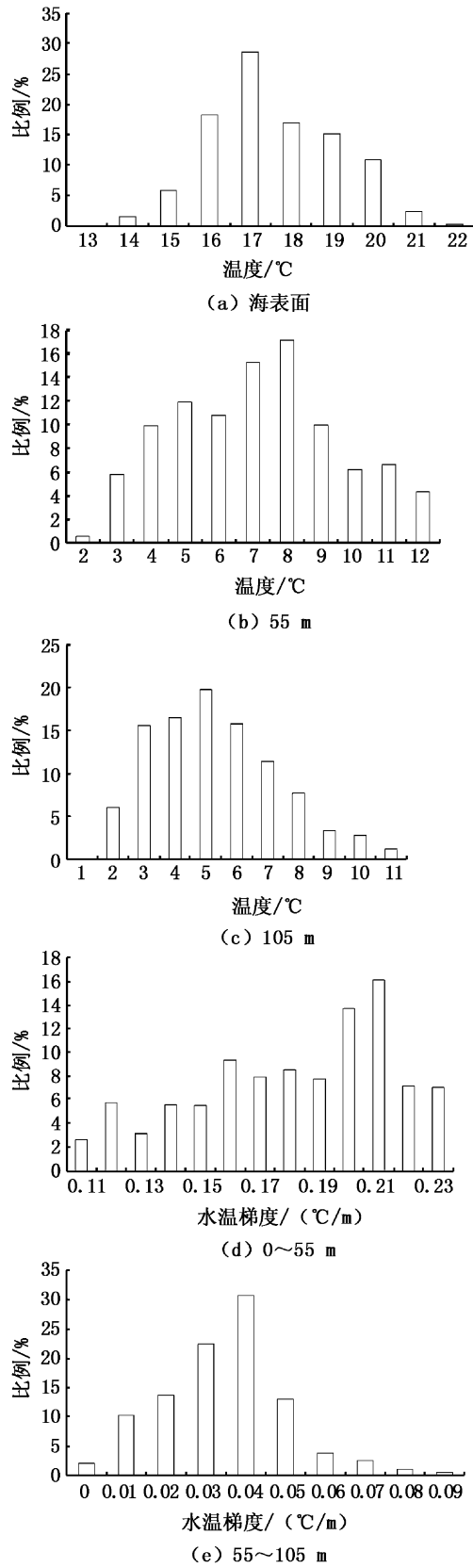


图 1 8 月柔鱼渔场作业次数与水温的关系
Fig. 1 The relationship between fishing effort for targeting *O. bartramii* and water temperature in August

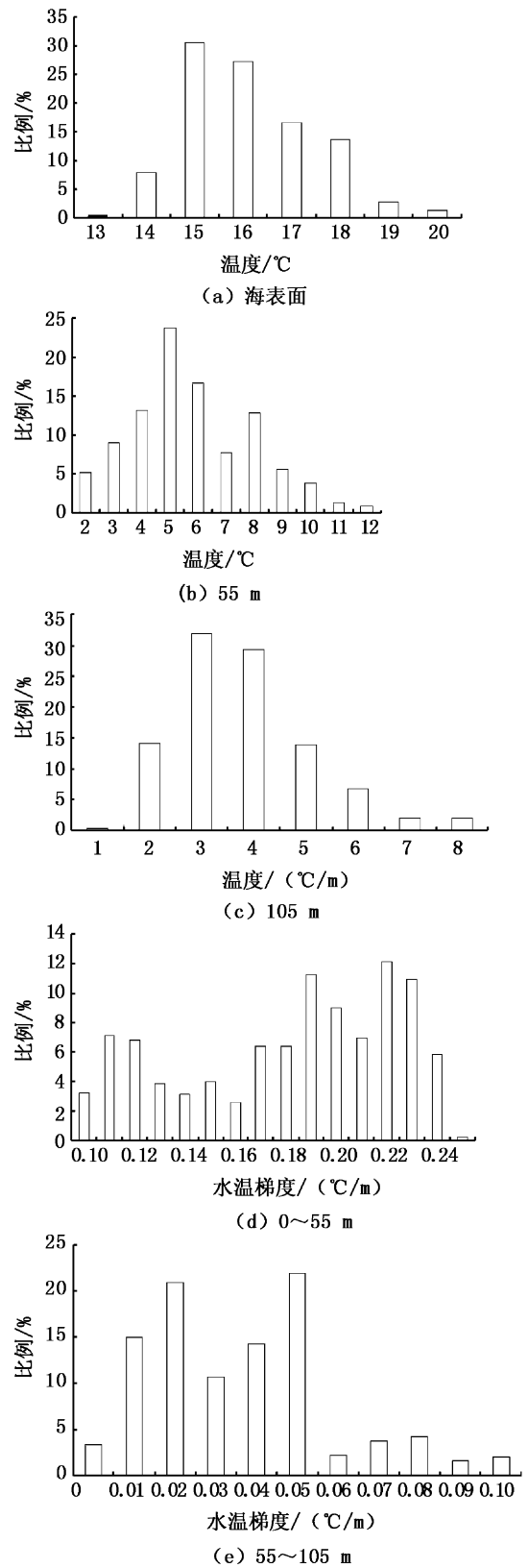


图 2 9 月柔鱼渔场作业次数与水温的关系
Fig. 2 The relationship between fishing effort for targeting *O. bartramii* and water temperature in September

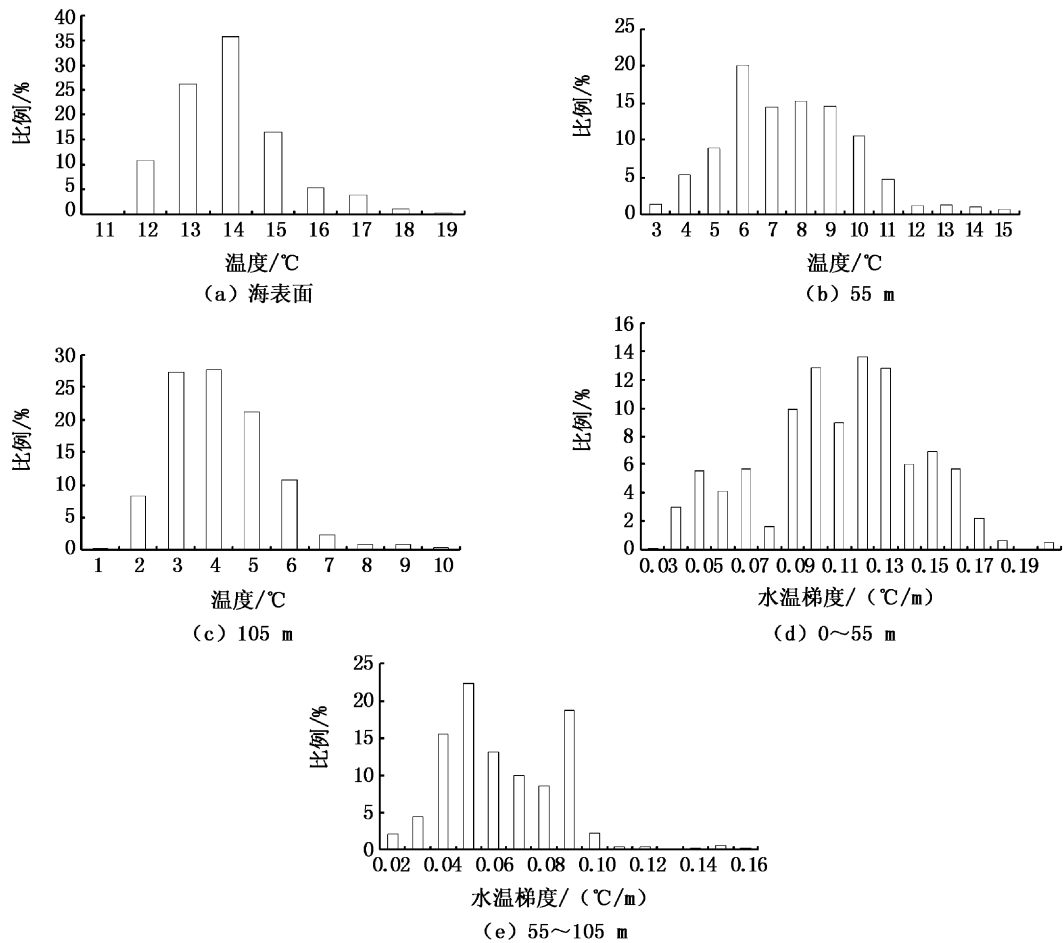


图3 10月柔鱼渔场作业次数与水温的关系
 Fig.3 The relationship between fishing effort for targeting *O. bartramii* and water temperature in October

表1 各属性对应的分类区间
 Tab.1 The corresponding classification sectors of each attribute

属性(环境和生产因子)	分类区间	区间总数
$T_0/^\circ\text{C}$	<14, 14~15, 15~16, 16~17, 17~18, 18~19, 19~20, 21~22, >22	9
$T_{55}/^\circ\text{C}$	<2, 2~3, 3~4, 4~5, 5~6, 6~7, 7~8, 8~9, 9~10, 10~11, 11~12, 12~13, >1313	11
$T_{105}/^\circ\text{C}$	<2, 2~3, 3~4, 4~5, 5~6, 6~7, 7~8, 8~9, 9~10, 10~11, >11	11
$T_{0-55}/(^\circ\text{C}/\text{m})$	<0.12, 0.12~0.13, 0.13~0.14, 0.14~0.15, 0.15~0.16, 0.16~0.17, 0.17~0.18, 0.18~0.19, 0.19~0.20, 0.20~0.21, 0.21~0.22, 0.22~0.23, >0.23	13
$T_{55-105}/(^\circ\text{C}/\text{m})$	<0.01, 0.01~0.02, 0.02~0.03, 0.03~0.04, 0.04~0.05, 0.05~0.06, 0.06~0.07, 0.07~0.08, 0.08~0.09, >0.09	10
作业次数	0~100, 100~200, 200~300, 300~400, 400~500, 500~600	6

表2 各属性分别对应于作业次数的信息增益值
 Tab.2 The corresponding information gain values of fishing effort of each attribute

属性(环境因子)	对应作业船次信息增益值
$T_0/^\circ\text{C}$	0.454 8
$T_{55}/^\circ\text{C}$	0.526 9
$T_{105}/^\circ\text{C}$	0.410 0
$T_{0-55}/(^\circ\text{C}/\text{m})$	0.569 3
$T_{55-105}/(^\circ\text{C}/\text{m})$	0.387 2

3 分析与讨论

对1998-2008年8-10月柔鱼中心渔场分布与各水层水温的分析结果表明:柔鱼中心渔场基本上分布在155°E~157°E和42°N~44°N海域,处于黑潮与亲潮的交汇区域。统计分析表明各水层(表层、50 m和105 m)各月的水温有所差

异,海表面温度随月份而下降,从8月的17~18℃,下降到9月的15~16℃,再下降至10月的14~15℃,但 T_{55-105} 基本稳定在0.04~0.06℃/m的范围内。此结果与陈新军和刘必林^[5]研究认为的北太平洋柔鱼作业渔场适宜海表面温度呈现出季节性变化的结论相一致。但中心渔场中 T_{55} 和 T_{105} 的最适温度没有呈现明显的季节性变化,其中9月最适温度范围偏低,对应的 T_{0-55} 偏大。则该水层有一较显著的温跃层,等温线分布较密集,符合中心渔场形成的条件^[3]。

柔鱼渔场形成与海洋环境关系密切,特别是海表面温度^[5,9-10,15]。陈新军和许柳雄^[6]利用灰色关联分析表明,在北太平洋(150°E~165°E)水温对柔鱼作业渔场的影响依次是海表面温度、0~50 m 温度梯度、0~100 m 温度梯度和纬度。本文研究结果认为,各水温因子对作业次数影响重要性依次是 T_{0-55} 、 T_{55} 、海表面温度、 T_{105} 和 T_{55-105} ,与陈新军和许柳雄^[6]稍有差异,但基本一致,即表温和0~50 m 水层的温度梯度极为重要。这一研究结果与柔鱼生活习性基本一致,柔鱼白天栖息在300~600 m 的水层,夜晚栖息在表层至50 m 的水层^[1],其中心渔场由黑潮和亲潮交汇形成,因此,温跃层是中心渔场形成的主要环境因子^[16-17]。

在利用IGT分析海洋环境因子对中心渔场的影响时,关键是对环境因子(即属性)的概化。由于环境因子是一个连续的属性,确定这一属性的最佳分裂点并没有一个严格的标准,要想获得比较理想的结果,需要结合专家的意见和经验来进行。文中主要选取了温度(海表面温度、各水层温度和温度梯度)对中心渔场分布进行了研究,今后还需要结合更多的环境因素,如海面高度、盐度、叶绿素、浮游生物等进行分析,选取最适宜的一个或者几个环境因子对中心渔场进行预测分析,进一步提高预测的精度。

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,2005:124-155.
- [2] CHEN C S, CHIU T S. Variations of life history parameters in two geographical groups of the neon flying squid, *Ommastrephes bartramii*, from the North Pacific [J]. Fisheries Research, 2003, 63: 349-366.
- [3] 陈新军. 渔业资源与渔场学[M]. 北京:海洋出版社,2004:116-129.
- [4] YATSU A, WATANABE T, MORI J, et al. Interannual variability in stock abundance of the neon flying squid, *Ommastrephes bartramii*, in the North Pacific Ocean during 1979-1998: impact of driftnet fishing and oceanographic conditions[J]. Fisheries Oceanography, 2000, 9(2):163-170.
- [5] 陈新军,刘必林. 2004年北太平洋柔鱼产量分析及作业渔场与表温的关系[J]. 海洋通报:英文版,2006(2):83-91.
- [6] 陈新军,许柳雄. 北太平洋150°E~165°E海域柔鱼渔场与表温及水温垂直结构的关系[J]. 海洋湖沼通报,2004(2):36-44.
- [7] 陈新军,田思泉. 西北太平洋海域柔鱼的产量分布及作业渔场与表温的关系研究[J]. 中国海洋大学学报:自然科学版,2005(1):101-107.
- [8] 陈新军. 北太平洋150°E以西海域柔鱼渔场与时空、表温及水温垂直结构的关系[J]. 上海水产大学学报,2004,13(1):78-83.
- [9] 陈新军. 关于西北太平洋的柔鱼渔场形成的海洋环境因子的分析[J]. 上海水产大学学报,1997,6(4):263-267.
- [10] 陈新军. 北太平洋(160°-170°E)大型柔鱼渔场的初步研究[J]. 上海水产大学学报,1999,8(3):1-5.
- [11] ANDRADE H A, GARCIA A E. Skipjack tuna in relation to sea surface temperature off the southern Brazilian coast[J]. Fisheries Oceanography, 1999,8(4):245-254.
- [12] 董立岩,李真,周灵艳. 特定信息增益决策森林分类器研究[J]. 计算机工程与应用,2010,46(26):111-113.
- [13] 胡智喜,唐学忠. 基于信息增益法的决策树构造方法[J]. 计算机与现代化,2006(3):28-30.
- [14] 郭亚宁,冯莎莎. 基于决策树方法的数据挖掘分析[J]. 软件导刊,2010,9(9):103-105.
- [15] 陈新军,刘必林,田思泉,等. 利用基于表温因子的栖息地模型预测西北太平洋柔鱼渔场[J]. 海洋与湖沼,2009,40(6):707-713.
- [16] 马金,陈新军,刘必林,等. 北太平洋柔鱼渔业生物学研究进展[J]. 上海海洋大学学报,2011,20(4):567-570.
- [17] 陈峰,陈新军,刘必林,等. 西北太平洋柔鱼渔场与水温垂直结构关系[J]. 上海海洋大学学报,2010,19(6):748-754.

Selection of key factors of water temperature in the fishing ground of *Ommastrephes bartramii* based on the information gain technology

YI Qian¹, CHEN Xin-jun^{1,2,3}

(1. College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Shanghai Education Commission for Oceanic Fisheries Resources Exploitation, Shanghai Ocean University, Shanghai 201306, China; 3. Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources, Shanghai Ocean University, Ministry of Education, Shanghai Ocean University, Shanghai 201306, China)

Abstract: *Ommastrephes bartramii* distributed in the northwest Pacific is one of the important fishing grounds. The environmental factors are the important content of the study on fisheries oceanography, which is also helpful to understand the forming mechanism of fishing ground and creating some possibilities for providing certain scientific advice for the production of Chinese squid jigging fishery. Based on the fishing data collected by Chinese squid jigging fleets in the north Pacific Ocean (150°E – 170°E), the temperature data between August and October from 1998 to 2008, the water temperature at different layers (T_0 , 55 m, 105 m) and temperature structure (0 – 55 m and 55 – 105 m) in the fishing ground of *O. bartramii* are analyzed. The information values of different classification attributes for environment factors are calculated by means of the information gain technology (IGT). The results show that the T_0 in fishing ground seems to be varied seasonally from August to October. The water temperatures at 55 m and 105 m are lower and the temperature gradient at 0 – 55 m is higher in September. Based on the IGT, the results show that the key water temperature factors are the temperature gradient between 0 – 55 m, the water temperature at 55 m, the T_0 , the water temperature at 105 m and the temperature gradient at 55 – 105 m in turn.

Key words: information gain; water temperature factor; *Ommastrephes bartramii*; Pacific Ocean