

文章编号:1000-0615(2009)02-0220-09

应用一般线性模型估算鲈、鳎大型 灯光围网渔业的捕捞效率

官文江^{1,2,3}, 陈新军^{1,2,3}

(1. 上海海洋大学海洋科学学院, 上海 201306;

2. 大洋生物资源开发和利用上海市重点实验室, 上海海洋大学, 上海 201306;

3. 大洋渔业资源可持续开发省部共建教育部重点实验室, 上海海洋大学, 上海 201306)

摘要:大型灯光围网是我国近海捕捞鲈、鳎的主要作业方式之一,探讨和估算捕捞效率对其渔业资源评估和管理有着重要的意义。根据1998-2003年东黄海我国鲈、鳎大型机轮灯光围网生产统计数据,利用一般线性模型对其捕捞效率进行了估算,讨论了线性模型误差结构的选择及数据变换对结果的影响。根据鲈、鳎单位捕捞努力量的渔获量(CPUE)数据呈正偏,以及CPUE均值与方差在对数尺度下的线性关系,选择了负二项分布、伽马分布与对数正态分布作为误差分布。研究表明,由于CPUE零值的存在,其对误差分布结构有很大影响。当采用最大似然估计时,对数正态分布与伽马分布的CPUE需加一常数(δ), δ 取值对结果有较大影响,随 δ 增大将使估计的捕捞效率对比度得到压缩。为了避免 δ 取值的影响,同时采用了Delta-GLM方法。通过不同模型比较,选择了Delta-负二项或Delta-伽马方法。根据Delta-负二项或Delta-伽马方法估算结果,各渔业公司间的捕捞效率存在很大的差异,且具有明显的区域性。32°N以北海区,捕捞效率高低依次为苏渔、辽渔、青渔、舟渔、宁渔、沪渔;台湾东北部海区,捕捞效率高低依次为苏渔、辽渔、舟渔、沪渔、宁渔、青渔。

关键词:鲈;鳎;大型灯光围网;捕捞效率;一般线性模型;德耳塔方法的一般线性模型

中图分类号:S 931

文献标识码:A

鲈(*Pneumatophorus japonicus*)、鳎广泛分布在我国近海,是我国近海重要的中上层鱼类,主要为灯光围网所捕捞。各渔船性能差异和捕捞技术水平会导致捕捞效率的不同,估算捕捞效率是开展渔业资源评估和管理的基础性工作。捕捞效率通常有两种估算方法^[1]:一是通过渔业调查试验,使捕捞渔船在相同资源密度区进行捕捞,对比各渔船单位捕捞努力量的渔获量(CPUE)确定^[2];二是利用一般线性模型(generalized linear model, GLM)进行估计^[3-7],此法现已成为商业性捕捞数据估算捕捞效率的基本方法。应用GLM模型估算捕捞效率需要根据CPUE分布选择相应的误差结构,并对数据进行适当的变换^[8-9]。国外对

误差结构的选择^[7]及数据变换^[8-9]已有相关文献进行了讨论,国内应用GLM模型对CPUE进行分析已有一些尝试^[10-11]。但国内外文献对GLM模型估算捕捞效率及误差结构与数据变换对捕捞效率估算的影响讨论较少,为此,本文根据东黄海鲈、鳎大型机轮灯光围网生产统计数据,利用GLM模型估算各渔业公司的捕捞效率,同时分析不同误差结构和数据变换对捕捞效率的影响。

1 材料与方法

1.1 数据来源

1998-2003年东黄海鲈、鳎大型机轮灯光围网生产统计数据来自上海海洋大学鱿钓技术组,

收稿日期:2007-09-20 修回日期:2008-05-05

资助项目:国家科技支撑计划(2006BAD09A05);卫星海洋环境动力学国家重点实验室开放研究基金资助(200510);上海市重点学科建设项目编号(S30702)

通讯作者:陈新军, E-mail: xjchen@shou.edu.cn

包括生产日期、船名、作业位置、捕捞网次、捕捞产量等数据记录。根据生产统计的空间分布,将作业海区分为3个子区域:32°N以北称为1区,台湾东北部为2区,长江口及舟山外海为3区。由于3区数据较少,本文不作讨论。渔船归属辽宁海洋渔业公司(辽渔),江苏海发渔业公司(苏渔),上海海洋渔业公司(沪渔),宁波海裕渔业公司(宁渔),舟山海洋渔业公司(舟渔)和青岛海裕渔业公司(青渔)。对不同渔业公司的生产数据进行分组,以估计各渔业公司捕捞效率。

1.2 数据分布类型

1区与2区的CPUE概率分布呈正偏(图1-

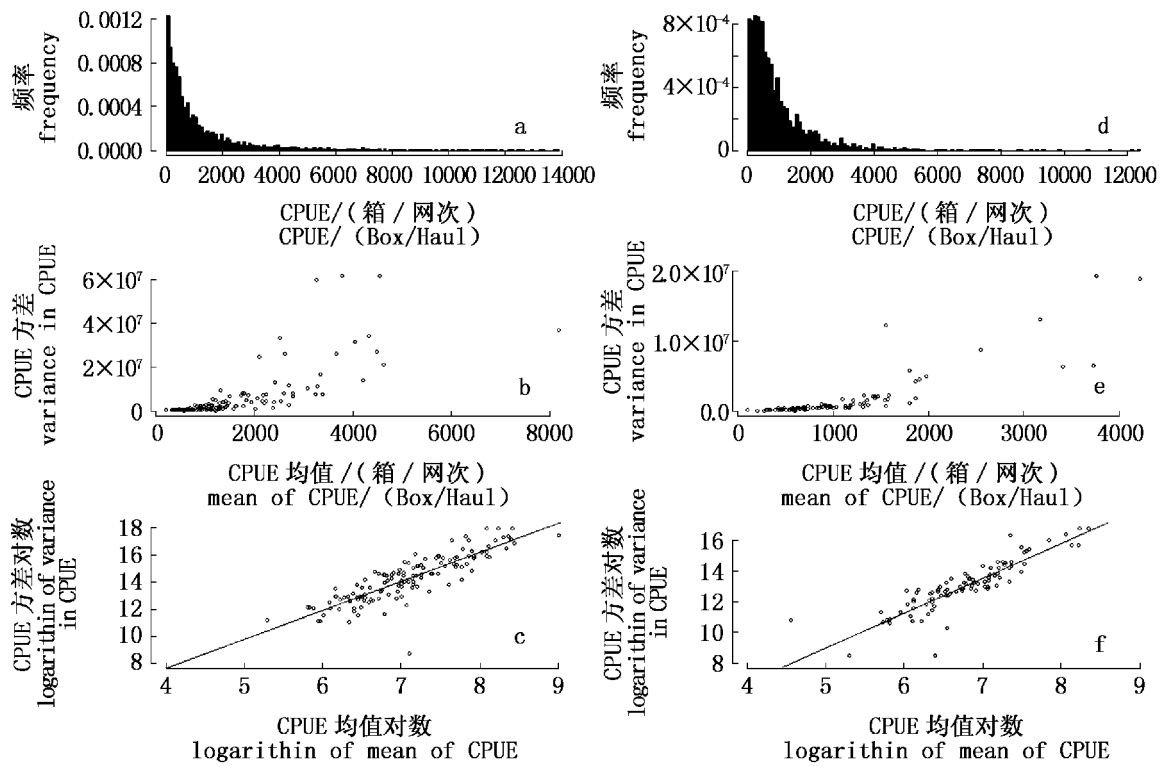


图1 CPUE频率分布图(a,d),CPUE均值与其方差(b,e), CPUE均值对数与其方差对数(c,f)关系(a~c为1区,d~f为2区)
Fig. 1 Frequency distribution of CPUE (a, d), the relationship between mean and variance of CPUE (b, e), and the relationship between log-mean and log-variance of CPUE(c, f) (a-c for area 1, d-f for area 2)

表1 1区与2区均值与方差对数回归系数及区间估计
Tab.1 Linear regression coefficients of mean and variance of CPUE transformed by logarithm and corresponding interval estimation

区域 fishing area	系数 coefficient	估计值 estimator	95%置信区间 95% confidence limit	相关系数 correlation coefficients	样本数 number of sample
区1 area 1	斜率 slope	2.14	[1.97, 2.33]	0.85	131
区1 area 1	截距 intercept	-0.90	[-2.35, 0.29]	0.85	131
区2 area 2	斜率 slope	2.27	[1.85, 2.59]	0.87	96
区2 area 2	截距 intercept	-2.33	[-4.63, 0.54]	0.87	96

a, d), 其均值与方差的关系呈曲线状(图1-b, e), 经过对数变换能得到较好的线性关系(图1-c, f; 表1) ($P < 0.05$)。回归系数斜率为1.849与2.594之间 ($P < 0.05$), 为此, 选择对数正态分布、负二项分布(CPUE取最近整数)、伽马分布作为误差类型^[5,7], 其均值与方差的关系见表2。由于存在CPUE为零的数据点, 所以也采用Delta-GLM模型, 其方法为: 以二项分布估计各渔业公司出现零CPUE的概率, 并求其平均概率(P_D), 对大于零的CPUE采用对数正态分布、负二项分布与伽马分布估计捕捞效率(P_t), 则实际估算捕捞效率为($P_D \times P_t$)。

表 2 对数正态分布、伽马分布与负二项分布及其均值与方差

分布类型 type of distribution	概率密度函数 function of probability density	均值 mean	方差 variance
对数正态 lognormal	$\frac{1}{y\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\text{Log}y-\mu}{\sigma}\right)^2\right]$	$\exp(\mu + 0.5\sigma^2)$	$\exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]$
伽马 gamma	$\frac{1}{\Gamma(\alpha)}\left(\frac{\alpha}{\mu}\right)^\alpha y^{\alpha-1} \exp\left(-\frac{\alpha y}{\mu}\right) \quad (\mu > 0, \alpha > 0)$	μ	$\frac{\mu^2}{\alpha}$
负二项 negative binomial	$\frac{\Gamma(r+k)}{k! \Gamma(r)} p^r (1-p)^k \quad (0 < p < 1, r > 0)$	$r \frac{1-p}{p}$	$r \frac{1-p}{p^2}$

1.3 GLM 模型

GLM 模型表达式为

$$\eta(u) = \text{Company} + \text{Year} + \text{Month} + \text{Year} \times \text{Month} \quad (1)$$

$$u = E[f(\text{CPUE})] \quad (2)$$

$$\text{CPUE} = \frac{\text{Catch}}{\text{Nets}} \quad (3)$$

其中, *Catch* 为捕捞产量, *Nets* 为网次, *Company* 为公司效应, *Year* 为年效应, *Month* 为月效应, *Year* × *Month* 为年与月的交互效应。f 为函数变换, E 为期望, η 为连接函数。

对于对数正态分布、伽马分布、负二项分布, f 形式为式(4)。式中 δ 为一常数, 为了比较所加常数对估算捕捞效率的影响, δ 选择如下四种情形: ① δ = 1^[12]; ② δ = 10% Mean (CPUE 均值的 10%)^[6]; ③ δ = Mean (CPUE 的均值)^[8]; ④ δ = 10 Mean (CPUE 均值的 10 倍)^[8]。对负二项分布, 也同时计算 δ 为零 (δ = 0) 的情形。Delta-GLM 模型中的 δ 为 0。

$$f = \text{CPUE} + \delta \quad (4)$$

对于二项分布, f 为式(5)。

$$f = \begin{cases} 0 & \text{CPUE} = 0 \\ 1 & \text{CPUE} > 0 \end{cases} \quad (5)$$

对数正态分布、伽马、负二项分布以自然对数为连接函数; 二项分布的连接函数为 logit。残差采用式(6)计算, 如果 η 为对数且 f(y_i) 为零, 则 f(y_i) 加 0.001。

$$\text{Residual} = \eta[f(y_i)] - \hat{\eta}(u) \quad (6)$$

依 Dick^[7] 与 Stefansson^[13] 的方法, 采用最大似然估计求取各参数。

2 结果

2.1 捕捞效率估算

对数正态分布情形下, CPUE 加 1 时, 捕捞效率估计值较大, 随 δ 值的增大, 所估计的捕捞系数的差异逐步缩小。当 δ 为 10 倍的 CPUE 均值时, 所估计的捕捞系数几乎相等, 而采用 Delta 对数正态分布的结果与 δ 为 CPUE 均值 10% 的结果相似(图 2-c, f)。

对于伽马与负二项分布, 其估计的捕捞效率几乎相等, 随 δ 的增大, 其估计的捕捞效率对比度同样得到压缩; 对于伽马与负二项分布, 采用 Delta 方法估计的捕捞效率几乎等于 δ 为零或为 1 时的结果(图 2-a, b, d, e; 表 3)。

2.2 方差与均值关系的比较

从图 3 可知, δ 为 1 时(图 3-a, d), 对数正态分布大幅偏离拟合曲线, 而伽马及负二项分布接近拟合曲线; δ 取 10% Mean 时(图 3-b, e), 对数正态分布结果有很大改善, 对方差较大的数据点, 对数正态分布很接近拟合曲线, 而对方差较小的数据点, 伽马及负二项分布较接近拟合曲线; CPUE 大于 0 的情况下(图 3-c, f), 伽马及负二项分布能更好表达方差较小数据点, 对数正态分布则能较好拟合方差较大的数据点。

2.3 模型选择

对数正态分布模型, 虽然 CPUE 加常数 δ 有利于减少偏度和峰度, 使其分布接近正态分布(对数变换后), 但依然存在较大偏差(图 4-a, c)。相比较而言, 正值 CPUE 更接近对数正态分布(图 4-b, d)。

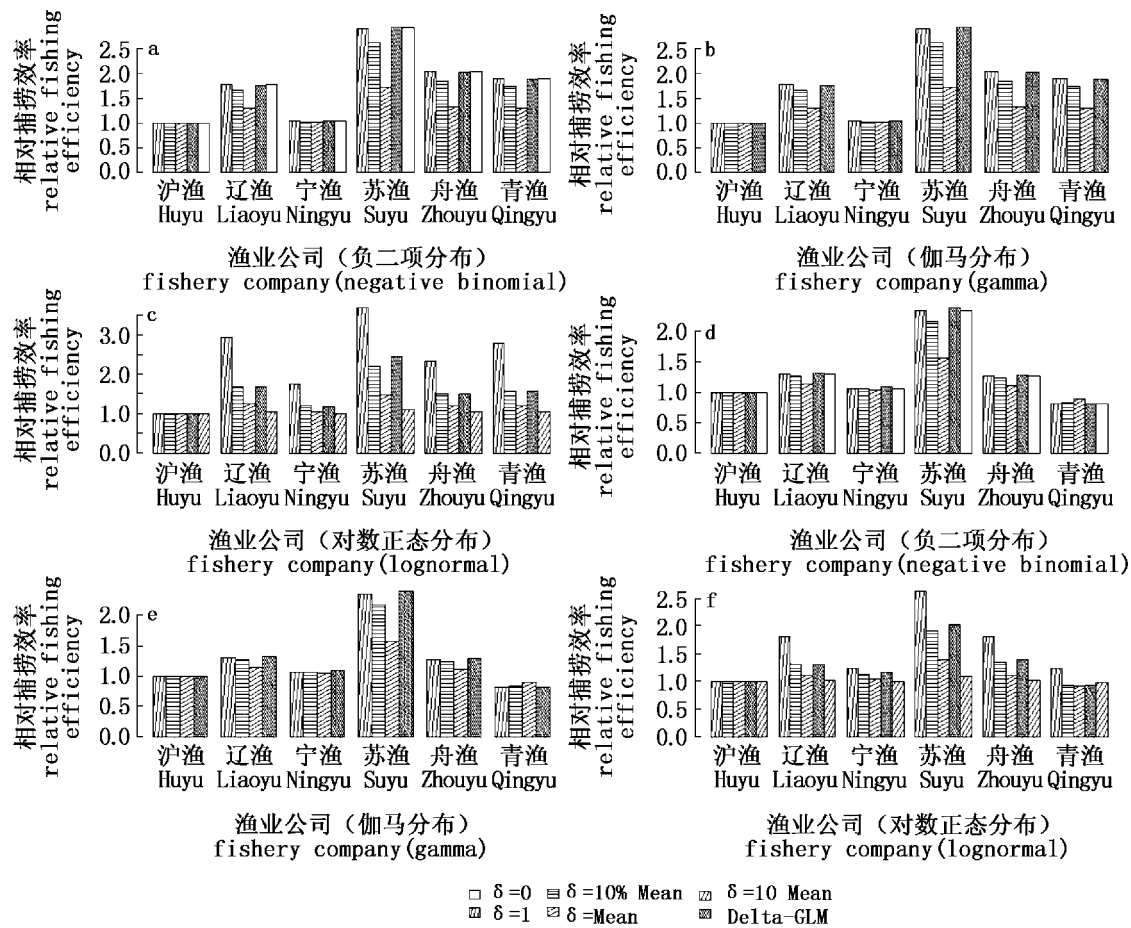


图 2 各模型估计的各渔业公司捕捞效率

a~c 为 1 区; d~f 为 2 区

Fig. 2 Fishing efficiency for different company estimated from different model

a-c for area 1 ; d-f for area 2

表 3 Delta-GLM 估算的捕捞效率的比较

Tab. 3 Fishing efficiency estimated by Delta-GLMs

模型 model	沪渔 Huyu	辽渔 Liaoyu	宁渔 Ningyu	苏渔 Suyu	舟渔 Zhouyu	青渔 Qingyu	区域 fishing area
二项分布 binomial distribution	0.89	1	0.97	0.99	1	1	1
CPUE = 0	246	0	10	11	0	0	1
对数正态 lognormal	1	1.68	1.19	2.43	1.52	1.58	1
伽马 gamma	1	1.77	1.04	2.95	2.02	1.88	1
负二项 negative binomial	1	1.77	1.04	2.95	2.02	1.88	1
负二项* negative binomial*	1	1.79	1.04	2.93	2.05	1.91	1
二项分布 binomial distribution	0.93	1	0.96	1	1	1	2
CPUE = 0	84	0	15	0	0	0	2
对数正态 lognormal	1	1.30	1.17	2.03	1.39	0.92	2
伽马 gamma	1	1.32	1.09	2.39	1.29	0.81	2
负二项 negative binomial	1	1.32	1.09	2.40	1.30	0.81	2
负二项* negative binomial*	1	1.31	1.06	2.35	1.27	0.81	2

注: * 为 δ 为零时的结果; 二项分布估算的是各渔业公司 CPUE 不为零的概率; CPUE = 0 是各渔业公司 CPUE 为零值的次数

Notes: The result was derived by assigning δ to be zero for the row signed by *; the rows for binomial distribution showed the probability where CPUE is more than zero; the row signed by CPUE = 0 showed the frequency of CPUE when its value was zero

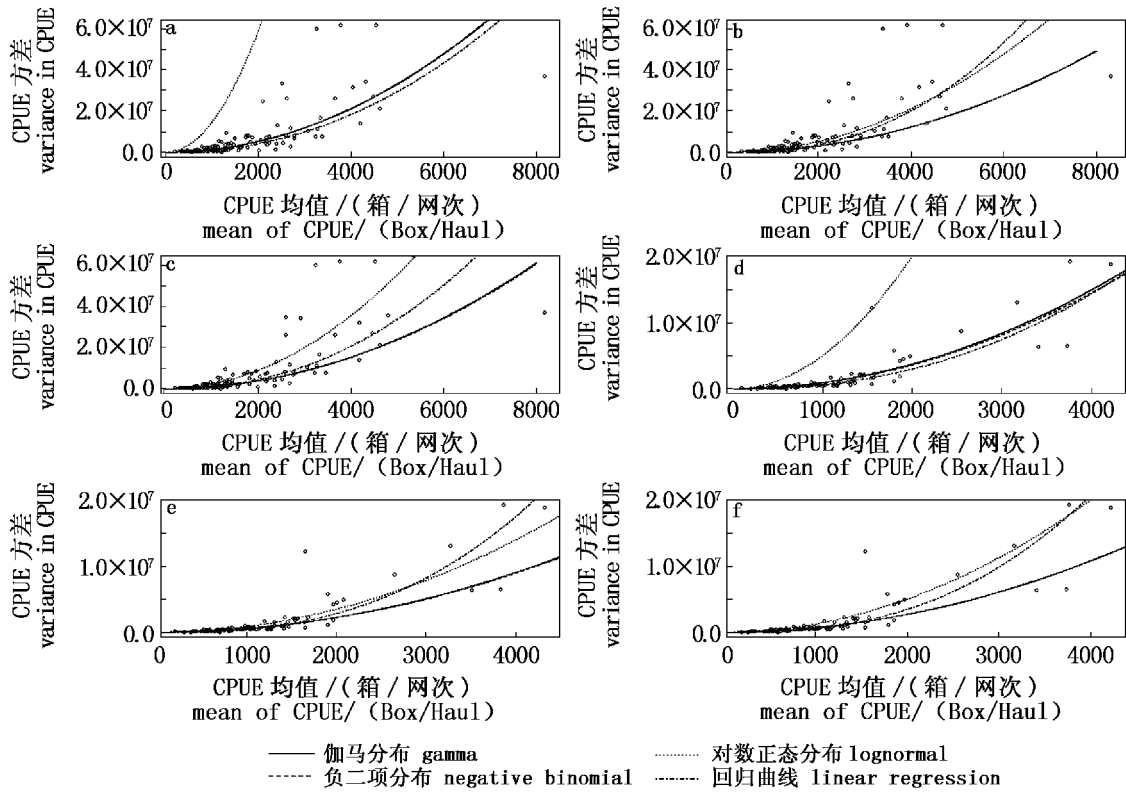


图3 CPUE 方差与均值的关系

a ~ c 为 1 区, d ~ f 为 2 区; a, d 中 $\delta = 1$; b, e 中 $\delta = 10\%$ Mean; c, f 中 CPUE > 0

Fig. 3 The relationship between mean and variance of CPUE

a-c drew from area 1 and figure d-f drew from area 2. $\delta = 1$ for the figures a and d and $\delta = 10\%$ Mean for the figures b and e. The data are deleted when its CPUE is zero for the figure signed by c and f, i. e. CPUE > 0

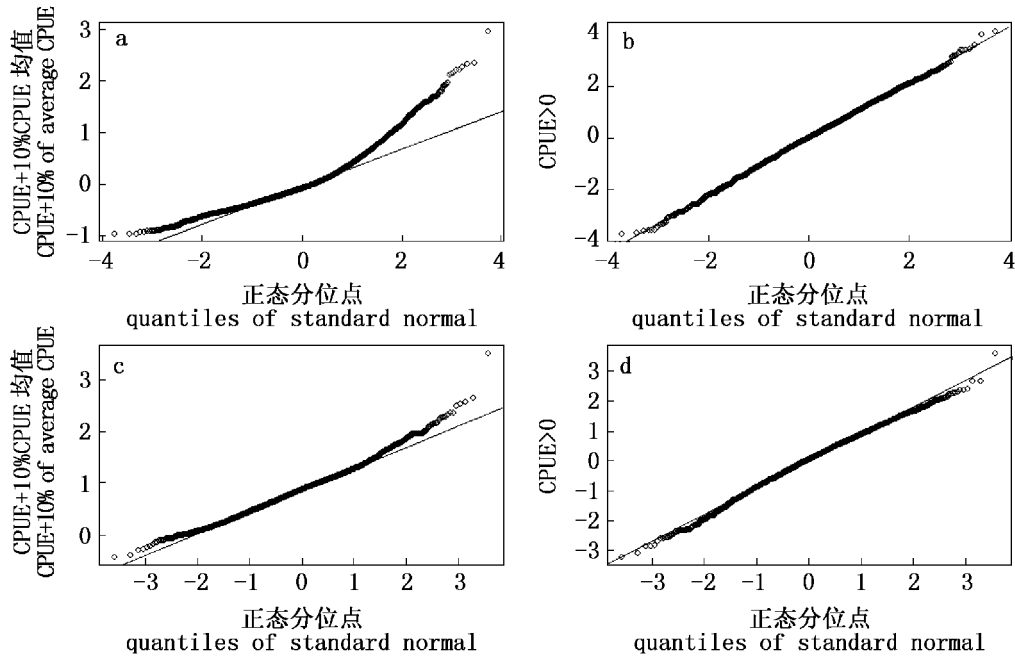


图4 对数正态分布残差的 QQ 图

a, b 为 1 区; c, d 为 2 区

Fig. 4 A normal quantile-quantile plot of the residuals

a ~ b for area 1; c ~ d for area 2

对伽马与负二项分布模型, δ 为零(负二项)或 1(伽马)时的残差分析可知,对于存在零的数据,其残差更偏离其期望残差结构(图 5-a, c, e, g),对正值 CPUE,其误差分布类型则接近假设的分布类型(图 5-b, d, f, h),因此 Delta-GLM 方法把 CPUE 为零与大于零的数据分开处理具有优

势。但采用 Delta-GLM 方法时,二项分布会引起相应的误差,如果忽略二项分布估计所引起的误差,可认为 Delta-GLM 方法较优^[5]。当采用 Delta-GLM 方法时,选择 AIC 标准,则对数正态分布在两区域都优于伽马或负二项分布(表 4)。

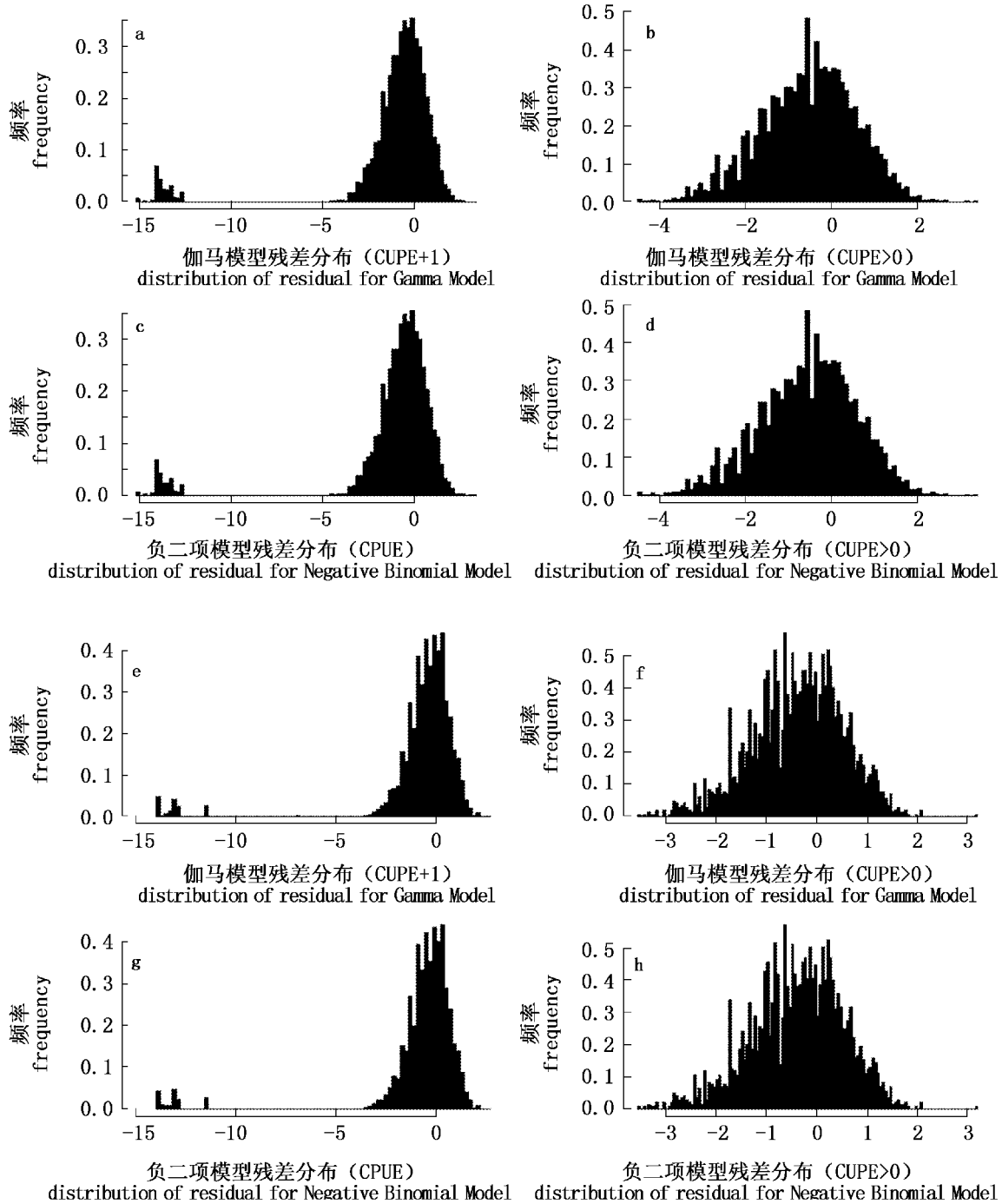


图 5 伽马与负二项分布的残差

a ~ d 为 1 区; e ~ h 为 2 区

Fig. 5 Histograms of residual of Gamma model and negative binomial model

a-d for area 1; e-h for area 2

表 4 Delta-GLM 方法的 AIC 值
Tab. 4 AIC calculated from Delta-GLM models

δ 值 value of δ	对数正态 lognormal	伽马 gamma	负二项 negative binomial	区域 fishing area
CPUE > 0 (Delta)	86 291.58	86 837.22	86 842.08	1
CPUE > 0 (Delta)	45 454.72	45 566.84	45 568.92	2

为进一步比较各模型估算结果与实际的符合程度,对 1 区与 2 区各渔业公司的捕捞数据,按年、月次计算其平均 CPUE,分选出与沪渔同年同月的捕捞数据(1 区不考虑网次小于 25 次的数据点,其中青渔为 10 网次;2 区不考虑网次小于 20 的数据点,其中舟渔为 15,由于青渔仅有 13 网次,所以不去除),计算各渔业公司月平均 CPUE 与沪渔月平均 CPUE 的比率,同时采用 Bootstrap 方法估算各渔业公司月平均 CPUE 同沪渔月平均

CPUE 比率的经验分布,分位数为 2.5%、50%、97.5% 的比率值见表 5。比较各月平均 CPUE 比率,及表 5 的结果,表明 Delta-GLM 方法的计算结果较好,其中 Delta-伽马和 Delta-负二项分布的结果优于 Delta 对数正态分布。此外,交叉验证(cross validation)及决定系数(determination coefficient)的结果也以上述二分布为佳,所以采用这两个模型的估计值。

表 5 分位数为 2.5%、50%、97.5% 时 Bootstrap 估算各渔业公司 CPUE 与沪渔的比率
Tab. 5 Monthly averaged CPUE ratio corresponding to the quantile 2.5%, 50% and 97.5% respectively estimated from Bootstrap between Huyu and the other companies

分位数 quantile	沪渔 Huyu	辽渔 Liaoyu	宁渔 Ningyu	苏渔 Suyu	舟渔 Zhouyu	青渔 Qingyu	渔区 fishing area
2.5%	1	1.64	0.92	2.44	1.04	1.28	1
50%	1	2.03	1.06	3.08	1.53	1.84	1
97.5%	1	2.50	1.22	3.79	2.02	2.48	1
2.5%	1	1.20	0.85	1.95	1.07	0.75	2
50%	1	1.33	0.99	2.55	1.21	0.89	2
97.5%	1	1.47	1.14	3.10	1.38	1.03	2

3 结论与分析

3.1 不同公司捕捞效率比较

由 Delta-负二项分布(或 Delta-伽马)结果可知(表 3),各渔业公司捕捞效率存在较大差异。在 1 区,苏渔捕捞效率最高,而沪渔最低,苏渔捕捞效率为沪渔的 3 倍;在 3 区,苏渔捕捞效率仍最高,而最低捕捞效率则为青渔,同时苏渔与沪渔的效率比为 2.5 左右。从上述计算结果,捕捞效率在作业区域和渔业公司之间的差异明显,因此,渔业资源评估时必须对捕捞努力量进行标准化,才能正确估算 CPUE。

3.2 GLM 模型比较

由于存在 CPUE 为零的数据点,采用对数正态分布时,必须加一常数,选择常数的标准通常是使数据能减少偏度和峰度,使残差(对数条件下)相对接近正态分布^[14]。但从上述结果来看,常数

数值较大,将使各变量的效应得到压缩,减少对比,因此必须综合考虑。尽管 δ 取 CPUE 均值的 10% 时,其对数正态分布的假设存在很大问题(图 4-a, c),但其结果很接近 Delta 对数正态分布,同时, δ 也存在其它取值情况^[8]。必须指出,选择 δ 值使其有利于线性模型的拟合,但对捕捞效率的估算则可能是破坏性的。

尽管伽马及二项分布允许 CPUE 值为零(采用最大似然估计时,伽马分布不能取零值),从本文结果看,零值同样对其残差结构产生影响,使其偏离期望分布(图 5-a, c, e, g),但采用伽马与负二项分布时,这种残差错误对其捕捞效率估计的影响不大。同时, δ 取值较小时,对其结果影响较小,而对数正态分布则会使估计系数产生较大的变化,因此对数正态分布对其分布假设更敏感^[15],而伽马或负二项分布则相对稳定、有效^[16]。

采用 Delta-GLM 方法可避免零值对假设误差结构的影响,从本文看其结果较优,Delta-GLM 方法也是目前常用的数据处理方法^[5,7,10,17],但是否是处理这类数据的必选模型则要具体分析所研究的数据^[17]。

对比 Delta 对数正态、Delta 伽马及 Delta 负二项分布估计值,1 区有较大差异,表现为采用负二项(或伽马)分布模型估计的苏渔、舟渔与青渔的捕捞效率相对偏大,而对宁渔捕捞效率的估计相对偏小;2 区差异相对较小,主要表现为对数正态分布对苏渔捕捞效率估计相对偏小(表 3),这表明不同误差结构的选择对结果有较大影响。

3.3 存在的问题

本文采用不同渔业公司对数据进行分组,忽略了同一渔业公司渔船间的差异,也没有考虑公司与年的交互以及温度、叶绿素、温度距平、叶绿素距平值等环境变量,模型所能解释的 CPUE 的变化也相对有限(大约为 13%)。是否存在更合理的因子^[18-19]或模型^[8,19-20]来提高对 CPUE 变化的解释,需要进一步研究,但对于数据量大的商业渔业数据,大量不可解释的变化可能是合理的。

尽管从实际情形来看,上述估算结果存在合理性,如捕捞效率较低的沪渔,主要从事金枪鱼围网,近海鲈鳕鱼围网为其金枪鱼围网培养船员,且其渔船船龄长,设备老化等严重影响其捕捞效率。但是,捕捞效率高低受很多方面的影响,如企业性质、网船船型、探渔设备、灯光配置、船员和船长的捕捞经验和水平等,本文缺少这方面的详细数据,很难做出更合理的比较,这需要进一步的调查研究与对比分析。

本文大部分工作在马里兰大学完成,感谢马里兰大学郑全安博士,美国 NOAA 李晓锋博士及缅因州立大学陈勇博士,他们对本论文提出了宝贵的修改意见。

参考文献:

- [1] Quinn T J, Deriso B R. Quantitative fish dynamics [M]. New York: Oxford University Press, 1999: 18-20.
- [2] 陈新军, 赵小虎, 刘金立. 印度洋西北海域鳕乌贼钓捕技术试验[J]. 海洋渔业, 2005, 27(3): 201-205.
- [3] Robson D S. Estimating of the relative fishing power of individual ships [J]. ICNAF Res Bull, 1966, 3:5-14.
- [4] Gavaris S. Use of a multiplicative model to estimate catch rate and effort from commercial data [J]. Canadian Journal of Fisheries and Sciences, 1980, 37: 2272-2275.
- [5] Punt A E, Walker T I, Taylor B L, et al. Standardization of catch and effort data in a spatially-structure shark fishery [J]. Fisheries Research, 2004, 70:129-145.
- [6] Campbell R A. CPUE standardisations and the construction of indices of stock abundance in a spatially varying fishery using general linear models [J]. Fisheries Research, 2004, 70:209-227.
- [7] Dick E J. Beyond lognormal versus gamma: distribution among error distributions for generalized linear models [J]. Fisheries Research, 2004, 70:351-366.
- [8] Porch C E, Scott G P. A numerical evaluation of GLM methods for estimating indices of abundance from West Atlantic bluefin tuna catch per trip data when a high proportion of the trips are unsuccessful [J]. Col Vol Sci Pap, ICCAT, 1994, 42(1):240-245.
- [9] Quinn T J, Deriso B R. Quantitative fish dynamics [M]. New York: Oxford University Press, 1999: 30-35.
- [10] 田思泉. 西北太平洋柔鱼资源评价及其与海洋环境关系的研究[D]. 上海:上海水产大学海洋学院, 2006.
- [11] 陈新军, 郑波, 李纲. GLM 模型和 GAM 模型研究东黄海鲈鱼资源渔场与环境因子的关系[J]. 水产学报, 2008, 32(3):379-386.
- [12] Damalas D, Megalofonou P, Apostolopoulou M. Environmental, spatial, temporal and operational effects on swordfish (*Xiphias gladius*) catch rates of eastern Mediterranean Sea longline fisheries [J]. Fisheries Research, 2007, 84: 233-246.
- [13] Stefansson G. Analysis of groundfish survey abundance data combining the GLM and delta approaches [J]. ICES Journal of Marine Science, 1996, 53: 577-588.
- [14] Berry D A. Logarithmic transformation in ANOVA [J]. Biometrics, 1987, 43:439-456.
- [15] Penning M. Efficient estimators of abundance, for fish and plankton surveys [J]. Biometrics, 1983, 39:281-286.
- [16] Firth D. Multiplicative errors: log-normal or

- gamma? [J]. *Journal of the Royal Statistical Society*, 1988, B2: 266–268.
- [17] Maunder M N, Punt A E. Standardizing catch and effort data: a review of recent approaches [J]. *Fisheries Research*, 2004, 70:141–159.
- [18] Wang J, Pierce G J, Sacau M, *et al.* Remotely sensed local oceanic thermal features and their influence on the distribution of hake (*Merluccius hubbsi*) at the Patagonian shelf edge in the SW Atlantic [J]. *Fisheries Research*, 2007, 83:133–144.
- [19] Zagaglia C R, Lorenzetti J A, Stech J L. Remote sensing data and longline catches of yellowfin tuna (*Thunnus albacares*) in the equatorial Atlantic[J]. *Remote Sensing of Environment*, 2004, 93: 267–281.
- [20] Venables W N, Dichmont C M. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research [J]. *Fisheries Research*, 2004, 70: 319–337.

Fishing efficiency for mackerel and scad in the large light purse seine fishery estimated by generalized linear model

GUAN Wen-jiang^{1,2,3}, CHEN Xin-jun^{1,2,3}

(1. *College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China;*

2. *The Key Laboratory of Oceanic Fisheries Resources Exploitation of Shanghai Education Commission, Shanghai Ocean University, Shanghai 201306, China;*

3. *The Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources, Ministry of Education, Shanghai Ocean University, Shanghai 201306, China)*

Abstract: Reliable estimation of effective fishing effort, which is proportional to fishing mortality, can provide information critical to the assessment and management of fisheries resources. To estimate effective fishing effort, we need to understand fishing efficiency and factors that may influence it. In this study fishing efficiency was estimated for mackerel and scad purse seine fisheries using generalized linear model. This was done for different companies involved in the fisheries. Different choices of error structures were considered in the estimation and their impacts on the estimation were discussed. The negative binomial distribution, gamma distribution, and log-normal distribution were chosen as error distributions according to log-linear regression of variance versus log-mean of CPUE (catch per unit effort). Zero CPUE values had great impacts on the assumed error structure and adding a constant (δ) to CPUE was needed for gamma distribution and log-normal distribution in maximum likelihood estimation. As δ increased, the contrast of estimated fishing efficiency was reduced greatly. Delta approaches were also chosen as an alternative way to deal with zero CPUE value in this study. Comparing the results of different models, we considered Delta-negative binomial and Delta-gamma as most appropriate error distributions for this study. The results showed that the fishing efficiency differed greatly among fisheries companies and among different areas.

Key words: mackerel; scad; large light purse seine; fishing efficiency; generalized linear model (GLM); Delta-GLM