

应用时间序列分析模型对大菱鲆选育 F_1 优良家系发育的动态研究

王新安, 马爱军*

(中国水产科学研究院黄海水产研究所, 青岛市海水鱼类种子工程与生物技术重点实验室,
农业部海洋渔业资源可持续利用重点开放实验室, 山东 青岛 266071)

摘要: 根据 Box-Jenkins 建模原理, 采用 ARIMAR(p, d, q) 模型, 以大菱鲆选育 F_1 优良家系为研究对象, 应用时间序列分析方法对在 3~27 月龄间不同发育阶段体重的生长速度进行动态模拟, 建立各家系的预测模型。结果表明, 家系 E $\delta 1 \times E \varphi 2$ 和 F $\delta 2 \times E \varphi 4$ 符合 ARIMAR(2, 0, 0) 模型(2 个家系的模型都不含常数项), 家系 F $\delta 4 \times N \varphi 3$ 、E $\delta 2 \times F \varphi 1$ 和 F $\delta 1 \times F \varphi 4$ 符合 ARIMAR(1, 0, 0) 模型(3 个家系的模型都不含常数项), 且所建模型的残差均为白噪声。由此预测出 27~27.5 及 27.5~28 月龄各家系体重生长速度, 经与相应实测数据的验证说明, 5 个家系各自所建模型在一定程度上能够反映大菱鲆体重生长速度的动态变化过程, 对各家系体重生长速度的趋势预测有一定的适用性。通过对每一家系后期生长速度的预测, 结合前期生长速度的实测值, 综合分析体重生长速度的动态变化, 为在大菱鲆选择育种过程中确定最佳选择时间提供理论依据。

关键词: 大菱鲆; 家系; 时间序列; 模型

中图分类号: S 917

文献标识码: A

自 1992 年大菱鲆 (*Scophthalmus maximus*) 被引入我国以来, 在海水鱼类工厂化养殖领域取得了重大成效^[1]。近年来, 引进大菱鲆在我国种质退化现象较为严重^[1-3]。调研表明, 在 2000 - 2001 年大菱鲆苗种经 10 个月的养成有 30% 可达到上市规格, 现在 10 个月达到上市规格不到 10%, 整批苗种达到上市规格要在 18 个月以上, 甚至 2 年。为保证大菱鲆养殖业健康、稳定、可持续发展, 中国水产科学研究院黄海水产研究所利用大规模家系选育技术对其进行了遗传改良。根据对 F_1 家系不同发育阶段生长性能育种值和表型值的综合评定, 筛选出 5 个生长性能良好且稳定的核心家系, 即 E $\delta 1 \times E \varphi 2$ 、F $\delta 4 \times N \varphi 3$ 、F $\delta 2 \times E \varphi 4$ 、E $\delta 2 \times F \varphi 1$ 和 F $\delta 1 \times F \varphi 4$ 。研究这 5 个家系生长发育的动态规律, 特别是预测未来一定时期的生长发育状况, 为在大菱鲆选育过程中确定最佳选择时间, 以期获得最大的育种成效, 提供重要的指导和参考依据。

有关鱼类的生长发育规律主要通过传统的回归分析方法进行研究^[4-6]。然而, 现代发育遗传学已探明在生物体的不同发育阶段, 基因是按一定的时空秩序有选择地表达的^[7], 控制数量性状的遗传机制在个体发育过程中会发生显著的变化^[8-9]。正是由于生物体数量性状的基因表达受到时间(发育期)和空间(试验地)等外部条件的影响^[9], 不同时空条件下, 鱼类的生长发育速度具有较大的不确定性。因此用传统的回归分析方法对鱼类生长发育速度的变化进行模拟分析, 效果并不理想。20 世纪 70 年代美国统计学家 Box 和英国统计学家 Jenkins 提出了以随机理论为基础的“求和自回归滑动平均模型”(autoregressive integrated moving average, ARIMA), 为随时间而不确定变化的数据进行定量分析提供了很好的方法^[10]。

本研究采用时间序列分析 ARIMA 模型, 对选择家系的生长发育状况进行研究, 拟合生长速

收稿日期:2010-05-07 修回日期:2010-07-08

资助项目:现代农业产业技术体系建设专项资金(nycytx-50);国家支撑计划专题(2006BAD01A12012);农业公益性行业科研专项经费项目(nyhyzx07-046);中国水产科学研究院黄海水产研究所基本科研业务费项目(2009-ts-11)

通讯作者:马爱军, Tel:0532-85835103, E-mail:maaj@ysfri.ac.cn

度指标模型并进行后期生长速度预测。

1 材料与方法

1.1 实验材料

实验所研究的5个全同胞家系E♂1×E♀2、F♂4×N♀3、F♂2×E♀4、E♂2×F♀1和F♂1×F♀4来源于2007年4月12日至7月19日在山东烟台天源水产有限公司所构建的家系。在3~28月龄期间,每隔1.5月龄从5个家系中随机抽取50个样本测量体重(g)。求每个家系的均重,并计算各家系相邻测量阶段的绝对增重率。绝对增重率: $AGR_w = (W_2 - W_1)/(t_2 - t_1)$ 。式中, W_1 (g)和 W_2 (g)分别为时间 t_1 和 t_2 (月)的体重。利用所得到的连续阶段的绝对增重率进行时间序列分析。其中前16个数据用于建模,第17、18个数据用于验证模型的预测效果。

1.2 分析方法

ARIMA模型又称B-J模型,其原理是:某些时间序列是依赖于时间 t 的一组随机变量,构成该时序的单个序列值虽然具有不确定性,但整个序列的变化却有一定的规律性,可以用相应的数学模型(ARIMA)近似描述。通过对该数学模型的分析研究,能够从本质上认识时间序列的结构与特征,达到最小方差意义下的最优预测。ARIMA模型有4种基本类型:自回归(AR)模型、移动平均(MA)模型、自回归移动平均(ARMA)模型以及差分自回归移动平均模型(ARIMA)。基本公式为

$$\hat{y}_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

式中, p 、 d 、 q 分别表示时间序列的自回归阶数、差分阶数和滑动平均阶数, y_t 为时间序列在 t 时期的预测值, y_{t-i} 为 $t-i$ 时期的观测值, ϕ_i 为模型的自回归系数, θ_i 为模型滑动平均系数, e_{t-i} 为时间序列模型在 $t-i$ 时期的误差或偏差。

ARIMA模型建模过程按4个阶段进行:

序列的平稳化判断 建立ARIMA模型的基础必须是平稳性的时间序列。首先,根据时间序列的散点图、自相关函数和偏自相关函数图对序列的平稳性进行识别;其次,为克服主观因素的影响,避免不必要的差分歪曲时间序列模型及降低模型精度,对序列值进行正态分布性和有无趋势性统计检验,判断序列是否平稳。对非平稳序

列要进行差分,直到达到序列平稳。

模型的识别 根据Box-Jenkins,用样本的自相关函数和偏自相关函数的截尾初步识别ARIMA模型的阶数。若平稳序列 $\{x_t\}$ 的 ρ_k 呈 q 步截尾,而 r_k 拖尾,则识别 $\{x_t\}$ 为MA(q)序列;若 ρ_k 拖尾,而 r_k 呈 p 步截尾,则识别 $\{x_t\}$ 为ARI(p)序列;若 ρ_k 、 r_k 均拖尾,则判断 $\{x_t\}$ 为ARIMA(p, q)序列。首先可以经验性地给出 p 、 q 的初步识别,然后通过模型诊断反复识别,找到最优的 p 、 q 组合来确定。模型诊断首先要判断模型参数是否有统计学意义;其次,对备选模型的拟合优度比较,本文采用的拟合优度统计量为赤池信息准则(Akaike information criterion,简记为AIC准则)和施瓦兹贝叶斯信息准则(Schwartz Bayesian criterion,简记为BIC准则),依据AIC和BIC准则确定模型阶数,建立ARIMA预测模型。在不断改变模型的阶数后,AIC与BIC值最小的模型为最佳模型;最后,还要考虑模型参数的简约性。AIC和BIC准则的定义为

$$AIC = N \ln \sigma_a^2(k) + 2(p + q + 1)/N$$

$$BIC = N \ln \sigma_a^2(k) + (p + q + 1)/N$$

其中, N 为样本数, σ_a^2 为残差的方差。

模型参数估计和模型假设检验 根据步骤II模型的识别,确定了各家系体重生长发育速度时间序列模型的结构和阶数,在此基础上可以进行模型参数的估计。模型参数估计的方法很多,常见的有参数的矩估计、最小二乘法估计和极大似然估计等方法。本文采用最小二乘法求解模型参数。估计出时间序列模型参数后还需要进一步检验模型是否合适,通常的做法就是检验其残差 e_t 是否为白噪声。即利用所建模型对历史数据进行递推预测,计算各期的残差及其各阶自相关系数 ρ_k ,并验证博克斯-皮尔斯统计量(Box-Pierce-Q-Statistic) $Q = N \sum_{k=1}^m \rho_k^2$ 是否服从 $X_{0.05}^2(m-p-q)$ 分布。如模型不能通过白噪声检验,则回到上一阶段,重新选定模型。

模型预测及精度评价 用每个家系所建立的模型预测27~27.5及27.5~28月龄体重增长速度,并与相应的实测值相比较,计算预测的相对精度。

采用SPSS 15.0软件进行数据处理与分析。

2 结果与分析

2.1 序列的平稳化检验

由 5 个全同胞家系 E ♂ 1 × E ♀ 2、F ♂ 4 × N ♀ 3、F ♂ 2 × E ♀ 4、E ♂ 2 × F ♀ 1 和 F ♂ 1 × F ♀ 4 体重生长速度时间序列图观察发现(图 1),各家系体重生长速度值均不存在周期性或趋势性。经 SPSS 13.0 软件进一步分析结果表明,各家系均通过无趋势性检验。分别对各家系体重生长速度序列作正态分布检验,结果表明,各家系均很好地符合正态分布,具有随机性。由各家系体重生长速度序列的自相关系数(ρ_k)和偏自相关系数(r_k)(表 1)发现,从总体上看,各家系两数值均呈逐渐减小态势;对于家系 E ♂ 1 × E ♀ 2,偏自相关函数值(r_k)在 $k > 2$ 以后, $|r_k| < 2/\sqrt{16} = 0.5$;对于家系 F ♂ 4 × N ♀ 3、F ♂ 2 × E ♀ 4、E ♂ 2 × F ♀ 1 和 F ♂ 1 × F ♀ 4,偏自相关函数值(r_k)在 $k > 1$ 以后,

$|r_k| < 2/\sqrt{16} = 0.5$,即各家系均为截尾。进一步观察自相关函数图(ACF)和偏自相关函数图(PACF)(图 2)发现,各家系自相关函数图形均呈现正弦线衰减波,而各家系偏自相关函数图形则在滞后 1 期或 2 期后迅速趋于 0。由上述结果可判断各家系序列均具有很好的平稳性。

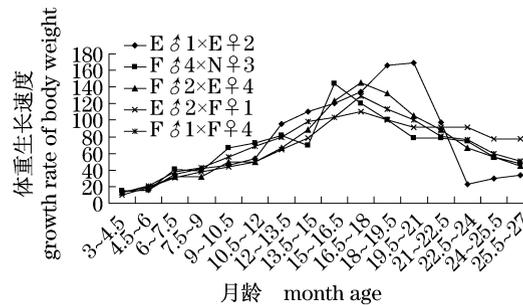


图 1 大菱鲆 5 个优良家系体重生长速度时间序列图
Fig. 1 Change of time series of growth rate of body weight in every selected families

表 1 大菱鲆 5 个优良家系体重生长速度时间序列的各阶自相关系数和偏自相关系数

Tab. 1 Auto-and partial-correlation coefficients of time series of growth rate of body weight of five selected families

| 滞后 lag | E ♂ 1 × E ♀ 2 | | F ♂ 4 × N ♀ 3 | | F ♂ 2 × E ♀ 4 | | E ♂ 2 × F ♀ 1 | | F ♂ 1 × F ♀ 4 | |
|-----------|---------------|--------|---------------|--------|---------------|--------|---------------|--------|---------------|--------|
| | ρ_k | r_k |
| 1 | 0.767 | 0.767 | 0.649 | 0.649 | 0.823 | 0.823 | 0.826 | 0.826 | 0.793 | 0.793 |
| 2 | 0.379 | -0.507 | 0.395 | -0.044 | 0.544 | -0.413 | 0.613 | -0.221 | 0.535 | -0.256 |
| 3 | 0.767 | 0.018 | 0.186 | -0.091 | 0.232 | -0.223 | 0.404 | -0.106 | 0.278 | -0.150 |
| 4 | -0.144 | -0.069 | 0.022 | -0.089 | -0.059 | -0.161 | 0.200 | -0.136 | 0.003 | -0.257 |
| 5 | -0.290 | -0.248 | -0.219 | -0.295 | -0.270 | -0.028 | -0.004 | -0.176 | -0.239 | -0.166 |
| 6 | -0.411 | -0.162 | -0.355 | -0.117 | -0.418 | -0.169 | -0.209 | -0.203 | -0.402 | -0.080 |
| 7 | -0.422 | 0.076 | -0.447 | -0.163 | -0.472 | -0.017 | -0.366 | -0.085 | -0.466 | -0.007 |
| 8 | -0.353 | -0.168 | -0.461 | -0.110 | -0.457 | -0.094 | -0.462 | -0.057 | -0.474 | -0.102 |
| 9 | -0.281 | -0.174 | -0.300 | 0.176 | -0.368 | 0.038 | -0.426 | 0.229 | -0.423 | -0.056 |
| 10 | -0.193 | 0.068 | -0.152 | -0.001 | -0.210 | 0.072 | -0.336 | 0.047 | -0.257 | 0.174 |
| 11 | -0.058 | 0.023 | -0.045 | -0.034 | -0.075 | -0.157 | -0.262 | -0.124 | -0.116 | -0.136 |
| 12 | 0.108 | -0.008 | 0.007 | -0.115 | 0.020 | -0.097 | -0.214 | -0.168 | -0.010 | -0.072 |
| 13 | 0.169 | -0.218 | 0.054 | -0.152 | 0.072 | -0.040 | -0.150 | -0.054 | 0.072 | -0.069 |
| 14 | 0.118 | -0.061 | 0.101 | -0.005 | 0.082 | -0.048 | -0.073 | -0.042 | 0.124 | -0.029 |

注: ρ_k , 自相关系数; r_k , 偏自相关系数。

Notes: ρ_k , Autocorrelations; r_k , partial autocorrelations.

2.2 模型的识别

根据本文平稳时间序列模型结构的识别规则以及本研究各序列的自相关函数图(ACF)和偏自相关函数(PACF)(图 2),家系 E ♂ 1 × E ♀ 2 和 F ♂ 2 × E ♀ 4 符合 ARIMAR(2,0,0)模型(2 个家系的模型都不含常数项),家系 F ♂ 4 × N ♀ 3、E ♂ 2 × F ♀ 1 和 F ♂ 1 × F ♀ 4 符合 ARIMAR(1,0,0)模型(3 个家系的模型都不含

常数项)。

2.3 模型参数估计和模型假设检验

经进行模型参数的估计,得到各家系体重生长速度的时间序列的具体模型如下:

$$\text{家系 E } \delta 1 \times \text{E } \delta 2: \text{ARIMAR}(2, 0, 0) = 1.304X_{t-1} - 0.593X_{t-2} + e_t$$

$$\text{家系 F } \delta 4 \times \text{N } \delta 3: \text{ARIMAR}(1, 0, 0) = 0.793X_{t-1} + e_t$$

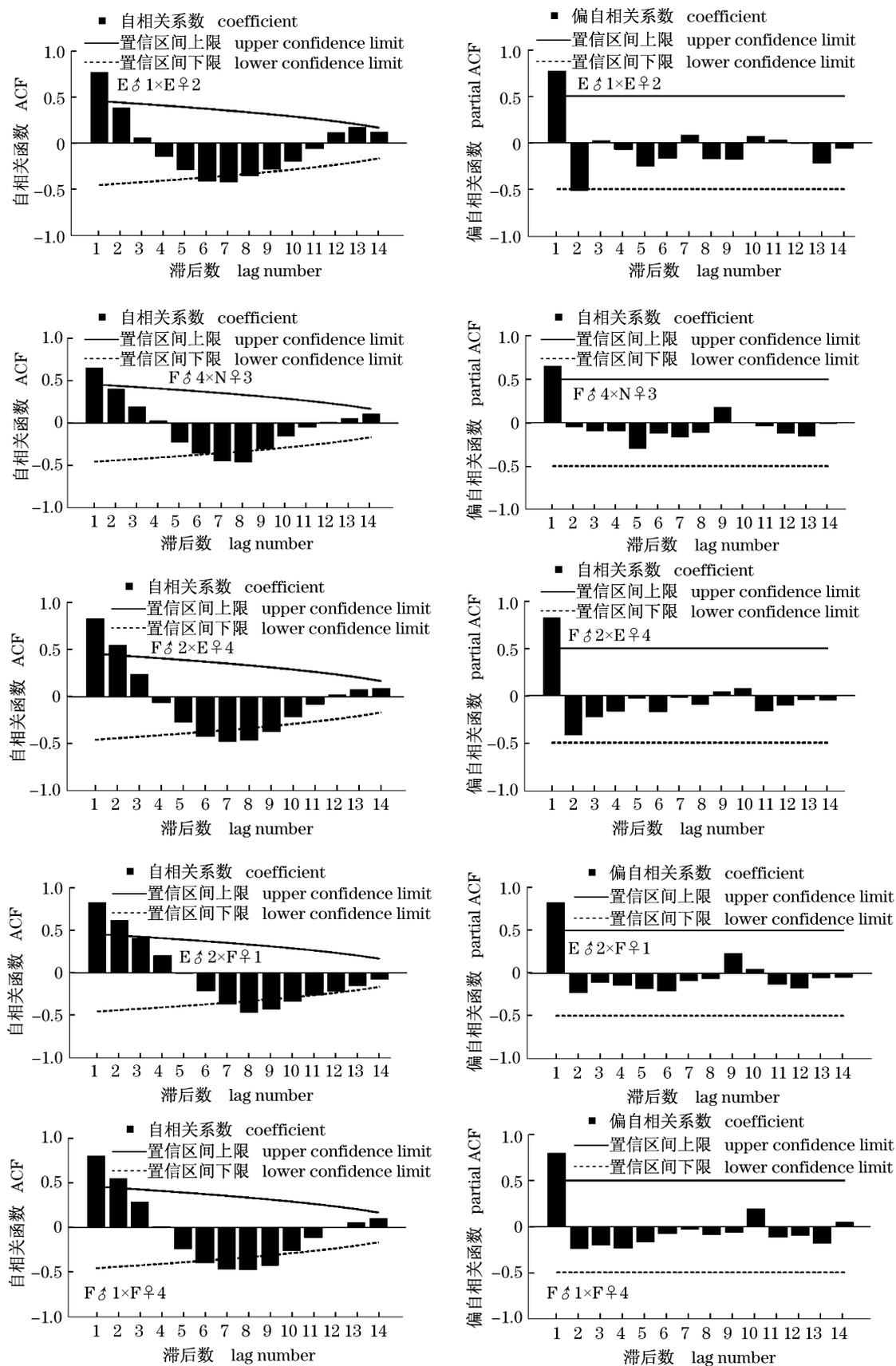


图2 大菱鲂5个优良家系体重增长速度时间序列的ACF和PACF

Fig. 2 Auto- and partial-correlation function of time series of growth rate of body weight of five selected families

家系 F ♂ 2 × E ♀ 4: ARIMAR(2, 0, 0) = 1.517X_{t-1} - 0.665X_{t-2} + e_t

家系 E ♂ 2 × F ♀ 1: ARIMAR(1, 0, 0) = 0.907X_{t-1} + e_t

家系 F ♂ 1 × F ♀ 4: ARIMAR(1, 0, 0) = 0.925X_{t-1} + e_t

对各家系所建模型进行 Q 统计量检验结果表明(表 2),各家系体重生长速度时间序列模型的统计量 Q 值都小于 X_{0.05}² 分布表的临界值,因此

认为在显著水平 0.05 下各模型残差均为白噪声,所建模型可以用于后期体重生长速度的预测。各家系所建的体重生长速度时间序列模型对历史数据进行模拟的结果发现,5 个家系模型预测值的动态趋势与实测值基本一致(图 3);除家系 F ♂ 4 × N ♀ 3 的 1 个实测值外,5 个家系的其它实测值都落入了预测值的可信区间范围之内。说明所建的模型能很好地描述体重生长速度的动态变化。

表 2 大菱鲆 5 个优良家系建立体重生长速度时间序列模型的参数

Tab. 2 Parameters of time series model of growth rate of body weight of five selected families

| 家系 family | 适用模型 adopted model | AIC Akaike's information criterion | BIC Schwarz's Bayesian criterion | 统计量 Q statistic-Q | X _{0.05} ² 临界值 X _{0.05} ² critical value |
|---------------|-----------------------|--|--|----------------------|---|
| E ♂ 1 × E ♀ 2 | ARIMAR(2, 0, 0) | 154.987 | 157.305 | 0.762 | 5.991 |
| F ♂ 4 × N ♀ 3 | ARIMAR(1, 0, 0) | 150.859 | 152.404 | 1.978 | 7.815 |
| F ♂ 2 × E ♀ 4 | ARIMAR(2, 0, 0) | 133.543 | 135.861 | 1.742 | 5.991 |
| E ♂ 2 × F ♀ 1 | ARIMAR(1, 0, 0) | 123.132 | 124.677 | 5.785 | 7.815 |
| F ♂ 1 × F ♀ 4 | ARIMAR(1, 0, 0) | 134.815 | 136.360 | 7.698 | 7.851 |

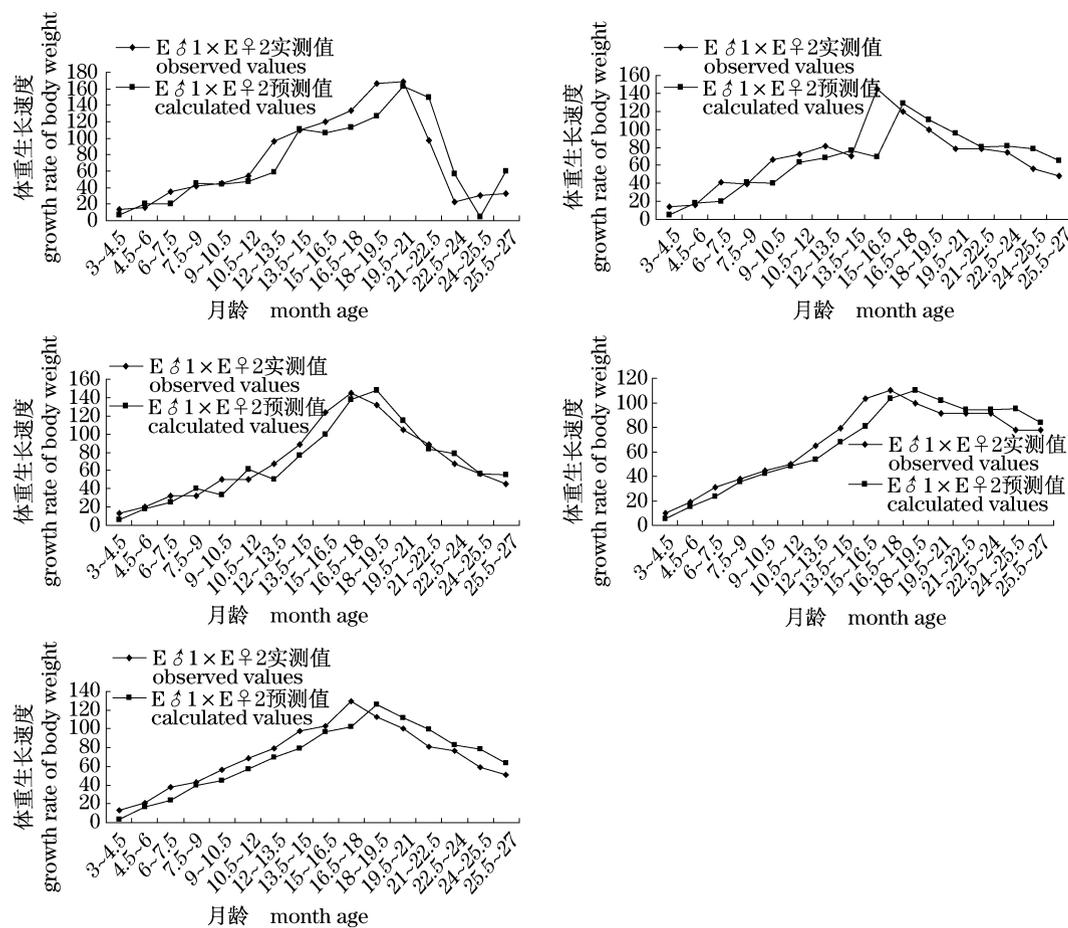


图 3 大菱鲆 5 个优良家系体重生长速度历史实测值与预测值的曲线图

Fig. 3 Matching in growth rate of body weight between observed and calculated values in every selected family

2.4 模型预测及精度评价

利用各家系所建的模型预测 27 ~ 27.5 和 27.5 ~ 28 月龄期间体重的生长速度,并与相应的实测值相比较,计算预测的相对精度。结果表明,5 个家系模型预测值的动态趋势与实测值基本一致(表 3);除家系 F δ 4 \times N φ 3 的相对精度稍低

外,其余 5 个家系的相对精度均大于 80;5 个家系体重生长速度的实测值都落入了预测值的可信区间范围之内。这说明所建模型在一定程度上能反映各家系体重生长速度的变化,可用于后期生长速度的实际预测。

表 3 大菱鲆 5 个优良家系体重生长速度预测结果(g)及精度分析

Tab.3 Results and precision analyse of growth rate of body weight in every selected family

| 月龄 month age | E δ 1 \times E φ 2 | | F δ 4 \times N φ 3 | | F δ 2 \times E φ 4 | | E δ 2 \times F φ 1 | | F δ 1 \times F φ 4 | |
|----------------------------|-------------------------------------|---------|-------------------------------------|---------|-------------------------------------|---------|-------------------------------------|---------|-------------------------------------|---------|
| | 27~27.5 | 27.5~28 | 27~27.5 | 27.5~28 | 27~27.5 | 27.5~28 | 27~27.5 | 27.5~28 | 27~27.5 | 27.5~28 |
| 实测值 observed values | 49.235 | 73.098 | 53.671 | 59.671 | 40.786 | 53.205 | 75.334 | 79.257 | 50.988 | 50.472 |
| 预测值 calculated values | 61.638 | 98.822 | 60.044 | 70.650 | 46.782 | 57.666 | 83.972 | 90.191 | 55.729 | 60.289 |
| 相对精度 relative precision | 74.809 | 64.809 | 88.126 | 81.601 | 85.299 | 91.615 | 88.534 | 86.204 | 90.702 | 80.550 |

注:相对精度 = $[1 - | \text{实测值} - \text{预测值} | / \text{实测值}] \times 100$ 。

Notes: Relative precision = $[1 - | \text{observed values} - \text{calculated values} | / \text{observed values}] \times 100$ 。

3 讨论

发育遗传学研究表明,对于生物的发育数量性状,由于基因表达的时空特性、细胞群体之间复杂的交互及随时间变化的渐成互作的存在,不同发育阶段基因所表达出来的遗传效应是不同的,它存在着一个动态变化^[9],数量性状也伴随着呈现出依时空变化的动态表现过程^[11]。即组成发育数量性状的各分量具有各自的个体发育历史或称“发育轨迹”(即连续发育时间段内任意两点之间的不同特征)^[12-13]。不同发育时间段里,各分量的“发育轨迹”呈现不同的“形状”,性状的遗传和非遗传控制因子有着不同的贡献^[7,13]。在动植物上,分别以番茄、水稻、棉花、甘蓝、小鼠等为遗传材料进行的不同发育时期遗传动态的研究证实了这一论点^[9,14-17]。在对海水名贵鱼类大菱鲆不同发育阶段体重育种值(加性遗传效应)的预测研究中,也同样证实了个体育种值的动态变化^[18]。这样,在对生物体的选育过程中,当利用某一时刻的遗传效应或表型值进行选择时,这一时刻的遗传效应或表型值是否为最佳选择指标需要被判断。最佳选择时间的确定,直接决定了育种工作的成败及能否取得最大的遗传进展。

显然,在鱼类生长性状的选育过程中,由于生长速度的动态变化,选择时间的确定是一项重要的研究内容。它决定了选择的精度,直接影响下

一代遗传进展的获得。尽管传统的因果回归分析方法被广泛应用于鱼类生长发育规律的研究,但影响生长速度因素的复杂性或有关影响因素数据资料无法获得时,使这种方法获得的模型对后期生长速度预测的效果较差,甚至无法完成预测,结果导致最佳选择时间的确定缺乏充分的理论依据。ARIMA 时间序列分析方法可以很好地克服这些难题,该方法以时间(t)综合代替这些影响因素,建立时间序列模型,达到预测的目的。目前,该方法主要应用于气象、农业产量和疾病等研究领域的预测预报^[19-21]。应用这种方法对生长发育规律的研究较少,仅见对儿童和树木直径生长发育时间序列研究的报道^[22-23]。在水产领域的应用仅见零星对水生动物产量的预测研究^[24-25],对水生动物生长发育规律的研究尚未见有报道。

本文根据 Box-Jenkins 建模原理,通过对大菱鲆选育 F₁代 5 个优良家系 E δ 1 \times E φ 2、F δ 4 \times N φ 3、F δ 2 \times E φ 4、E δ 2 \times F φ 1 和 F δ 1 \times F φ 4 在 3~27 月龄期间体重生长速度实测值的分析,分别建立了 ARIMA(2,0,0)、ARIMA(1,0,0)、ARIMA(2,0,0)、ARIMA(1,0,0) 和 ARIMA(1,0,0) 模拟模型(5 个 ARIMA 模型均不含常数项)。由于一个 ARIMA(p,0,0) 过程表示了一个纯 AR(p) 平稳过程,因此本文所构建的 5 个 ARIMA 模型实际上均为纯 AR 模型。对模拟结果的分析表明,5 个家系各自所建模型在一定程

度上能够反映大菱鲆体重生长速度的动态变化过程,对各家系体重生长速度的趋势预测有一定的适用性。通过对每一家系后期生长速度的预测,结合前期生长速度的实测值,综合分析体重生长速度的动态变化,为在大菱鲆选择育种过程中确定最佳选择时间提供理论依据。

参考文献:

- [1] 马爱军,王新安,雷霖霖,等. 大菱鲆(*Scophthalmus maximus*)4个不同地理群体数量形态特征比较[J]. 海洋与湖沼,2008,39(1):24-29.
- [2] 申雪艳,宫庆礼,雷霖霖,等. 进口大菱鲆*Scophthalmus maximus* L. 苗种的遗传结构分析[J]. 海洋与湖沼,2004,35(4):332-341.
- [3] 王新安,马爱军,许可,等. 大菱鲆幼鱼表型形态性状与体重之间的关系[J]. 动物学报,2008,54(3):540-545.
- [4] Cotano U, Alvarez P. Growth of young-of-the-year mackerel in the Bay of Biscay[J]. Journal of Fish Biology,2003,62(5):1010-1020.
- [5] Amoer D, Landry J G, Lambert T C. Growth of juvenile (0-year) Atlantic mackerel (*Scomber scombrus*) in the Gulf St. Lawrence[J]. Canadian Journal of Fisheries and Aquatic Science,1990,47(3):2212-2218.
- [6] Simard P, Castonguay M, Amour D, et al. Growth comparison between juvenile Atlantic mackerel (*Scomber scombrus*) from the two spawning groups of the Northwest Atlantic[J]. Canadian Journal of Fisheries and Aquatic Science,1992,49(11):2242-2248.
- [7] Stewart A D, Hunt D M. The genetics basis of development[M]. Glasgow and London: Blackie, 1982:81-179.
- [8] Atchley W R. Ontogeny, timing of development and genetic variance-co-variance structure[J]. Am Nat, 1984,123(4):519-540.
- [9] Atchley W R, Zhu J. Developmental quantitative genetics, conditional epigenetic variability and growth in mice[J]. Genetics,1997,147(2):765-776.
- [10] 陈超,邹滢. SPSS 15.0 中文版常用功能与应用实例精讲[M]. 北京:电子工业出版社,2008:349-385.
- [11] Atchley W R. Developmental quantitative genetics models of evolutionary change[J]. Developmental Genetics,1994,15(1):92-103.
- [12] Alberch P, Gould S J, Oster D B, et al. Size and shape in ontogeny and phylogen[J]. Paleobiology, 1979,5(3):296-317.
- [13] Atchley W R. Developmental quantitative genetics and evolution of ontogenies[J]. Evolution,1987,41(2):316-330.
- [14] Kherialla A I, Whittington A J. Genetic analysis of growth in tomato: the F₁ generation[J]. Ann Bot, 1962,26(104):489-504.
- [15] Xu Y B, Shen Z T. Diallel analysis of tiller number at different growth stages in rice (*Oryza sativa* L.) [J]. Theor Appl Genet,1991,83(2):243-249.
- [16] Ye Z H, Lu Z Z, Zhu J. Genetic analysis for developmental behavior of some seed quality traits in upland cotton (*Gossypium hirsutum* L.) [J]. Euphytica,2003,129(2):183-191.
- [17] 李云侠,Murali-Tottekkad V,吴建国,等. 甘蓝型油菜不同发育时期角果重量的遗传分析[J]. 浙江大学学报(农业与生命科学版),2009,35(1):45-50.
- [18] 马爱军,王新安,雷霖霖. 大菱鲆(*Scophthalmus maximus*)不同生长阶段体重的遗传参数和育种值估计[J]. 海洋与湖沼,2009,40(2):187-194.
- [19] 王志忠,刘秀菊,王树林,等. 黄河流域棉花品种产量性状时间序列的 ARIMA 模型预测研究[J]. 棉花学报,2007,19(3):220-226.
- [20] 吴家兵,叶临湘,尤尔科. ARIMA 模型在传染病发病率预测中的应用[J]. 数理医药学杂志,2007,20(1):90-93.
- [21] 朱继民,汤林华,周水森,等. 非稳定性疟疾区用时间序列模型预测疟疾发病率的可行性研究[J]. 中国寄生虫学与寄生虫病杂志,2007,25(3):232-236.
- [22] 邓丹,王润华. 时间序列分析在重庆市小学生生长发育状况研究中的应用[J]. 重庆医科大学学报,2008,33(11):1368-1370.
- [23] 罗旭,程承旗,冯仲科,等. 树木直径生长的时间序列分析及模型预测[J]. 中南林业科技大学学报,2007,27(2):7-12.
- [24] 董江水,诸英富. 时间序列分析模型在江苏省河蟹总产量预测中的应用[J]. 金陵科技学院学报,2008,24(3):102-105.
- [25] 李辉华,郭弘艺,唐文乔,等. ARIMA 模型在预测长江靖江段沿岸鱼类渔获量时间格局中的应用[J]. 水产学报,2008,32(6):899-905.

The time series of development of selected families of turbot (*Scophthalmus maximus* L.) selective breeding F_1

WANG Xin-an, MA Ai-jun*

(Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao Key Laboratory for Marine Fish Breeding and Biotechnology, Key Lab for Sustainable Utilization of Marine Fishery Resources, Ministry of Agriculture, Qingdao 266071, China)

Abstract: Turbot, a flatfish of deep water species, is a high-nutrition and economic valued species and currently cultured widely in China. In recent years, however, serious germ plasm degeneration has come into being due to the lack of the long-term and effective broodstock management programmes, production of farmed turbot became highly unstable and total output declined gradually. Therefore, the genetic improvement of turbot will be necessary to sustain the industry development. Within a breeding program in aquaculture, growth rate which determines the total harvest yield is a highly desirable economic trait. However, the growth rate at different development is variational because genes are expressed selectively at different growth stages following a certain sequence of time and place. It is necessary, obviously, to determine a good selection time during selective breeding for obtaining maximal genetic gain. In present study, according to Box-Jenkins theory and ARIMA(p, d, q) model, the time series of growth rate of body weight of five selected families of turbot selective breeding F_1 were identified, simulated and predicted. SPSS 15.0 software was used to construct the ARIMA model. Raw data for model establishment and validation were from 3 to 27 months. The results showed that family E ♂ 1 × E ♀ 2, F ♂ 4 × N ♀ 3, F ♂ 2 × E ♀ 4, E ♂ 2 × F ♀ 1 and F ♂ 1 × F ♀ 4 adapt to ARIMA(2, 0, 0), ARIMA(1, 0, 0), ARIMA(2, 0, 0), ARIMA(1, 0, 0) and ARIMA(1, 0, 0) model, respectively (Constants were not included in all five models). The residual errors of five models were all the white noise series ($P > 0.05$). Relative precisions of the forecasting for growth rate of body weight of five families during months 27 – 27.5, 27.5 – 28 were comparatively high, which indicated those ARIMA models are able to well-describe the temporal and spatial changes in growth rate of body weight of turbot. The development changes of growth rate of body weight would be analysed through past observed values and forecast values, which can provide theoretic basis for determining optimal selection time during selective breeding of turbot.

Key words: turbot (*Scophthalmus maximus* L.); family; the time series; model

Corresponding author: MA Ai-jun. E-mail: maaj@ysfri.ac.cn