

文章编号:1000-0615(2009)03-0438-07

栉孔扇贝夏季大规模死亡的神经网络预测模型

邓登¹, 麦康森¹, 李晓明², 马洪明¹, 谭北平¹

(1. 中国海洋大学教育部海水养殖重点实验室, 山东 青岛 266003;

2. 中国海洋大学海洋信息科学与工程学院, 山东 青岛 266003)

摘要:根据2002年和2003年对山东荣成桑沟湾栉孔扇贝养殖海区的水温、盐度、pH、氨氮浓度、亚硝氮浓度等环境因子和扇贝血清中的蛋白浓度、酸性磷酸酶活力、碱性磷酸酶活力、超氧化物歧化酶活力和过氧化氢酶活力等免疫学指标及栉孔扇贝养殖密度和死亡率的监测数据,运用人工神经网络(artificial neural network, ANN)的原理和误差反相传播(back propagation, BP)网络的方法,利用MATLAB软件初步建立养殖栉孔扇贝夏季大规模死亡的BP人工神经网络预测模型。预测模型经过300次的学习训练,误差平方和由67.46下降至0.0091。该预测模型对未参与模型构建的样本预测的结果与实际监测结果的符合率达到87.5%。首次将人工神经网络与水产动物病害死亡的预测相结合,建立的预测模型具有对数据适应能力强,可适时学习,预测结果准确等突出优点,为水产养殖动物病害死亡程度的预测提供了一个新的研究方法。

关键词:栉孔扇贝;大规模死亡;人工神经网络;预测模型

中图分类号:S 917

文献标识码:A

栉孔扇贝(*Chlamys farreri*)为我国北方海域的特色贝种,是人工养殖的主要贝类之一,具有巨大的经济价值和生态价值,至今已有二十多年的人工养殖史^[1-4]。近十年来随着扇贝养殖规模和养殖密度的急剧增加,加上全球气候变暖、水质恶化及生态平衡紊乱等各种环境胁迫的加剧,导致病害频频发生,栉孔扇贝大规模暴发性死亡越来越频繁^[5-7]。仅1997年、1998年,山东省养殖扇贝因大规模死亡而带来的直接经济损失就达40亿元以上,给我国北方的扇贝养殖业带来了巨大的经济损失和负面影响^[8-9]。夏季栉孔扇贝的大规模死亡已成为制约栉孔扇贝人工养殖业发展的主要原因。如果能够根据以往监测的扇贝养殖历史数据,建立栉孔扇贝夏季大规模死亡预测模型,对栉孔扇贝夏季死亡情况进行预测,则可以使栉孔扇贝的收获和渡夏工作得以有计划 and 可超前地进行^[10-11],以减少扇贝死亡的损失。

建立基于时间序列数据的预测模型是利用历

史数据,通过某种途径,使之生成一个系统模型,能够预测未来的行为^[12]。人工神经网络(artificial neural network, ANN)的出现为预测模型的建立提供了创新性的方法,历史数据可以作为“训练”输入,在过去、现在、将来的数据之间建立一个经验的修正关系,从而建立预测模型^[13]。近年来,根据疾病发生情况、环境因子和研究对象特征指标的调查结果,利用人工神经网络建立疾病发生与影响因素的关系模型和疾病发生情况的预测模型,该研究方法在人类疾病和农作物病虫害暴发预测上的应用已逐步兴起,成为国际流行病学研究中的热门领域^[14-21]。

本研究运用人工神经网络原理和误差反相传播(back propagation, BP)模型的方法,对建立养殖栉孔扇贝夏季大规模死亡预测模型进行初步研究,成功地建立了栉孔扇贝夏季死亡率与相关环境因子和扇贝免疫指标之间复杂的非线性关系,为栉孔扇贝和其它水产养殖动物病害发生的预测

收稿日期:2008-04-18

修回日期:2008-10-05

资助项目:国家“九七三”重点基础研究发展规划项目(G1999012012)

通讯作者:麦康森, Tel: 0532-82032495; E-mail: kmai@ouc.edu.cn

提供了一个新方法。

1 研究方法

ANN 是受生物神经系统的结构和信息处理方式启发而设计的信息处理系统^[22]。1987 年, Kohonen 在 IEEE 第一届神经网络国际会议上将其定义为“大规模并行的,用简单而又常常是自适应元件内联而成的网络;它们具有分层结构,并以与生物神经系统同样的方式与外部世界对象相互作用”^[23]。现实观测到的数据很难通过线性统计模型来刻画其特性,ANN 通过对非线性数据的编码来充分刻画这些数据具有的特征^[13]。通过神经网络中连接权重所包含的信息,可以建立适合数据变化趋势的神经网络模型,从而得到预测值^[22]。

本研究采用目前应用最广泛的一种 ANN 方法——BP 网络^[24-27],即反向传播网络。其主要思想是把学习过程分为两个步骤:(1)正向过程,给出输入信息,通过输入层经隐含层逐层处理并计算每个单元的实际输出值;(2)反向过程,若在输出层未能得到期望的输出值,则逐层递归地计算实际输出与期望输出之差值即误差,据此调节权值^[28-30]。图 1 是基于 BP 网络的 3 层网络模型的结构示意图,图中的圆圈代表神经元,X 为神经元输入信号,O 为输出信号。

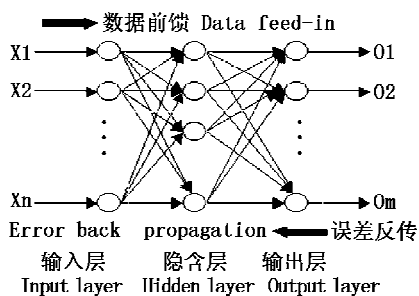


图 1 BP 算法的 3 层网络模型结构示意图

Fig.1 The structure of BP neural network model with 3 layers

其中,输入层接收信号,隐含层实现人工神经网络的非线性特性,所得信号在输出层形成系统输出。这些层由若干结点组成,层与层之间的结点通过权重连接(从节点 i 到节点 j 连接权重为 w_{ij})。正向传播时,样本信号经过隐含层处理后传向输出层,输出信号将与目标信号相比较,形成一个误差能量函数(公式 1),如果这个误差函数

E (均方误差)没有达到设定的精度要求,则转入误差的逆向传播过程。将输出误差按某种形式,通过隐含层向输入层逐层返回,并“分摊”给各层的所有结点,从而获得各层结点的参考误差或称误差信号,以作为修改各层结点连接权重的依据,权重修改见公式 2。

$$E = \frac{1}{2} \sum_{j=1}^m (Y_j - \hat{Y}_j)^2 \quad (1)$$

式中, Y_j 为目标值, \hat{Y}_j 为网络计算输出值, m 为输出层神经元个数。

$$w_{ij}(p+1) = w_{ij}(p) + \alpha \Delta w_{ij}(p-1) - \eta \frac{\partial E(p)}{\partial w_{ij}(p)} \quad (2)$$

式中, α 和 η 分别为动量系数和学习系数,均在 0、1 之间取值, E 为误差能量函数。

图 2 是 3 层 BP 网络计算的流程图。可以看出,BP 网络信号的正向传播与误差的逆向传播对各层权重矩阵的修改过程是周而复始地进行。权重不断修改的过程就是网络的学习(或称训练)过程。此过程一直进行到网络输出的误差减少到可接收的程度或达到设定的学习次数为止。实践证明,对于任意闭合区间连续函数都可以用含有一个隐含层的 BP 网络来逼近^[17]。

2 模型构建

2.1 资料处理

资料选取 模型选取的资料来源于 2002 年和 2003 年对山东桑沟湾栉孔扇贝养殖海区的水质因子(水温、盐度、pH、氨氮浓度、亚硝酸氮浓度)、扇贝免疫指标(血清中的蛋白浓度、酸性磷酸酶活力、碱性磷酸酶活力、超氧化物歧化酶活力和过氧化氢酶活力)和扇贝死亡情况每月监测的部分结果。由于养殖扇贝的死亡集中于 7-9 月,因此选择两年中 6-10 月的数据为模型样本。其中,以扇贝周死亡率(%)为模型的预测对象。首先对各监测指标与扇贝周死亡率进行相关分析和针对养殖密度的偏相关分析,排除与扇贝死亡相关性较低的指标参与模型构建。分析结果见表 1,统计软件为 SPSS 11.0。

选取水温、pH、血清过氧化氢酶活力、血清蛋白浓度和养殖密度为模型输入指标。

样本归一化处理 由于模型样本数据的单位和数量级不相同,为满足 BP 网络节点函数的

条件,利用公式(3)将预测样本的数据映射到[0, 1],结果见表2^[19]。

$$S_i = (R_i - R_{\min}) / (R_{\max} - R_{\min}) \quad (3)$$

式中, R_i 为实测数据; R_{\min}, R_{\max} 分别为最小、最大的实测数据。

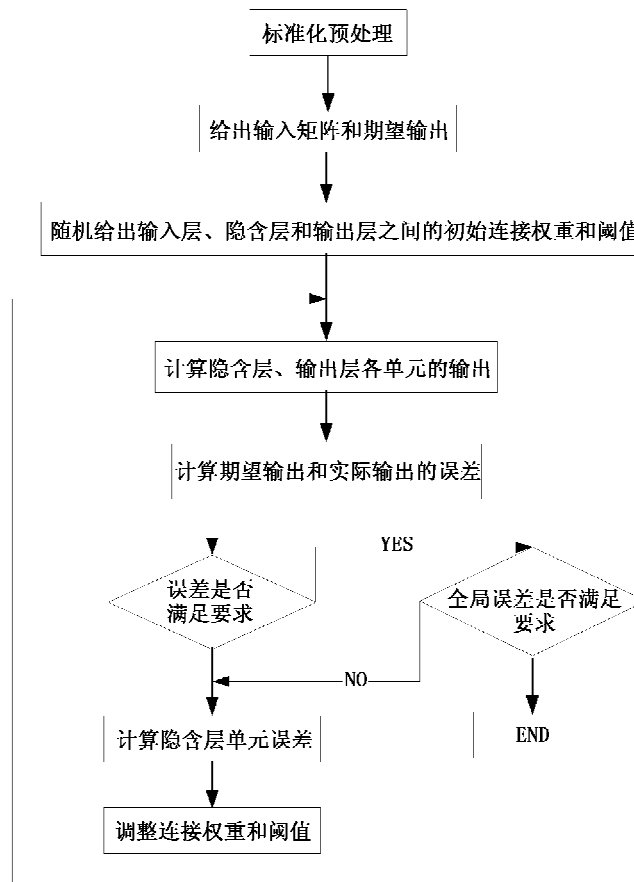


图2 3层BP网络模型计算流程图

Fig.2 The algorithm flowchart of BP neural network model with 3 layers

表1 水质因子、扇贝免疫学指标与扇贝死亡率的相关分析

Tab.1 Correlation analysis between the parameters of both seawater quality and scallop immunology and scallop mortality

项目 index	相关系数 correlation coefficient	P	偏相关系数 partial correlation coefficient	P
水温 temperature	0.646 *	0	0.682 *	0
pH	-0.733 *	0	-0.723 *	0
盐度 salinity	-0.376	0.064	-0.405	0.5
氨氮 NH ₃ -N	0.305	0.138	0.312	0.138
亚硝氮 NO ₂ -N	0.06	0.744	0.072	0.74
ALP 活力 ALP activity	0.317	0.123	0.252	0.235
ACP 活力 ACP activity	0.055	0.79	0.022	0.917
SOD 活力 SOD activity	0.056	0.792	0.092	0.669
CAT 活力 CAT activity	-0.601 *	0.001	-0.648 *	0.001
血清蛋白浓度 protein content	0.578 *	0.002	0.643 *	0.001

注: * 显著相关(P<0.01)

Notes: * Significant correlation (P<0.01)

表 2 样本归一化后数值和模型预测结果

Tab.2 Normalized values of factors and predicting results of the BP network model

样本 sample	密度 ^a density	水温 temperature	pH	CAT 活力 CAT activity	血清蛋白浓度 protein content	实际死亡等级 factual death level	预测死亡等级 ^b predicted death level
1	0	0.813	0.507	0.212	0.362	1	1(1.001)
2	0	0.915	0.366	0.149	1.000	3	3(2.9949)
3	0	0.160	0.901	0.792	0.083	1	1(1.001)
4	0	0.162	0.549	0.850	0.354	1	1(1.001)
5	0	0.915	0.127	0.116	0.644	3	3(2.9949)
6	0	1.000	0.507	0.063	0.169	3	3(3.0114)
7	0	0.162	0.493	0.771	0.352	1	1(0.9949)
8	0	0.813	0.268	0.268	0.362	1	1(0.9949)
9	0	0.915	0.254	0.069	0.689	3	3(2.9949)
10	0	1.000	0.380	0.010	0.427	3	3(3.0114)
11	0	0.160	1.000	0.515	0.056	1	1(0.9495)
12	0.333	0.000	0.648	0.373	0.181	1	1(0.9495)
13	0.333	0.502	0.423	0.220	0.563	2	2(2.0201)
14	0.333	0.157	0.845	0.231	0.000	1	1(0.9495)
15	1	0.162	0.521	0.804	0.068	2	2(3.0114)
16	1	0.915	0.000	0.000	0.732	4	4(4.0013)
17	1	1.000	0.493	0.046	0.448	2	2(2.0201)
18*	0	0.162	0.507	0.688	0.233	1	1(0.9495)
19*	0	1.000	0.465	0.003	0.536	3	3(2.9949)
20*	0	0.813	0.465	0.175	0.446	1	1(1.001)
21*	0	0.160	0.915	0.364	0.100	2	2(1.9576)
22*	0.333	0.630	0.394	0.129	0.297	3	3(3.0114)
23*	0.333	0.509	0.507	0.163	0.103	2	3(2.9949)
24*	1	0.813	0.056	0.139	0.287	1	1(1.001)
25*	1	0.160	0.873	1.000	0.100	1	1(1.001)

注: * 未参与模型构建的样本; a. 养殖笼内中每层的扇贝数; b. 括号中为模型输出值

Notes: * Samples not used to create the model; a. Scallop numbers of each layer; b. Values in the bracket represent the output values of the model

扇贝死亡等级的划分 由于模型样本由历史实测资料组成,因此不能完全代替总体未来的扇贝死亡样本。为使建立的预测模型具有实际应用效果,应对样本的输出进行分级。根据各种防护措施对应的扇贝死亡情况,将扇贝死亡程度划分为 4 个等级:周死亡率≤1%为 1 级,1%~10%为 2 级,10%~30%为 3 级,≥30%为 4 级。

2.2 模型编程

采用 MATLAB 软件编程,利用 MATLAB 中的神经网络工具箱建立 BP 神经网络模型。

输入样本矩阵 P 表示所建立的样本矩阵,其中 A:水温;B: pH;C:血清过氧化氢酶活力;D:血清蛋白浓度;E:扇贝养殖密度;n:输入变量个数。

```
P = [A1 A2 A3 ..... An;
      B1 B2 B3 ..... Bn;
      C1 C2 C3 ..... Cn;
      D1 D2 D3 ..... Dn;
      E1 E2 E3 ..... En];
```

输出目标向量 T 表示输出的目标向量,

即与输入样本对应的扇贝周死亡率等级。

$$T = [T1 \ T2 \ T3 \ \dots \ Tn];$$

建立网络模型

```
P = [0.813 0.915 0.160 ..... 1.000;
      0.507 0.366 0.901 ..... 0.493;
      .....
      0 0 0 ..... 1]; %输入数据向量
```

```
T = [1 3 1 ..... 2]; %输出目标向量
```

```
S = 20; %隐含层神经元个数
```

```
df = 10; %学习过程显示频率
```

```
[w1, b1, w2, b2] = initff (P, S, 'tansig', T, 'purelin'); %网络初始化
```

```
me = 1000; %最大训练次数
```

```
eg = 0.01; %期望误差
```

```
lr = 0.001; %初始学习速率
```

```
tp = [df me eg lr];
```

```
[w1, b1, w2, b2, ep, lr] = trainbpx (w1, b1, 'tansig', w2, b2, 'purelin', p, t, tp) %采用快速BP算法训练网络
```

```
P = [0.162 1.000 0.813 ..... 0.160;
```

```

0.507 0.465 0.465 ..... 0.873;
.....
0 0 0.....1]; % 预测样本输入
T = simuff ( P, w1, b1, 'tansig', w2, b2,
'purelin') % 预测结果
Echo off

```

隐含层神经元的数量直接影响网络的非线性性能,它与所解决问题的复杂性有关。根据经验公式(公式4)确定隐含层神经元个数为20个^[21,31]。

$$S \geq N \cdot p \cdot (p + t) \quad (4)$$

式中, n 为训练样本数, p 为输入层结点数, t 为输出层结点数。

隐含层的激活函数采用'tansig'函数(S型正切激活函数),输出层的激活函数采用'purelin'函数(线性激活函数)。快速BP算法将自适应的学习速率和附加动量法结合起来,避免BP网络落入局部极小值和收敛速率过慢,并提高训练速度。BP网络训练结束后,使用'simuff'函数即可实现对前向网络的仿真。

3 模型应用

在模型构建过程中,对模型的隐含层结点数、输入层结点数、训练样本数、训练精度以及训练步长等关键参数进行多次调整优化,使模型的稳定性和准确率达到较优。图3是所建立的BP神经网络模型训练过程的误差递减曲线,该模型经300次的学习训练,其平方误差和从67.46下降至0.0091,小于模型的期望误差 eg ,因此训练中止,即获得养殖栉孔扇贝夏季大规模死亡的BP神经网络预测模型。利用该模型进行预测的结果见表2,其中,对建模样本预测的拟合率为100%;对非建模样本预测的准确率为87.5%。

4 讨论

神经网络工具箱是MATLAB软件中重要的一个工具箱。在实际应用中,BP神经网络应用最为广泛,尤其适用于对具有非正态和非线性特征的病害发生进行预测。许多学者的相关研究结果表明,人工神经网络预测模型相对于其他的统计预测方法具有误差小、拟合率和准确率高以及操作方法简单等优点^[32-34]。网络初学者即使不了

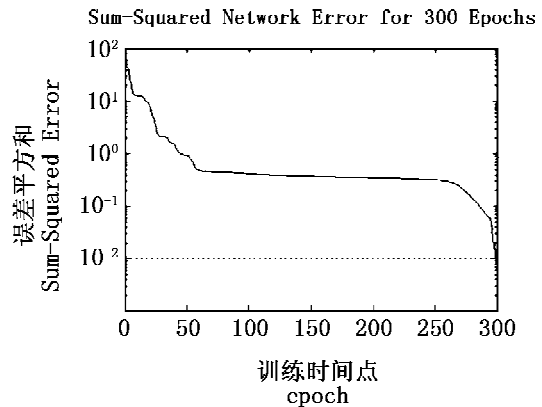


图3 模型训练时网络误差平方和的降低过程
Fig.3 The descending process of the sum-squared network error through training model

解其算法本质,也可以直接应用功能丰富的函数来实现研究目的^[19,35-36]。本文利用MATLAB软件,根据两年的野外监测数据,初步建立栉孔扇贝夏季大规模暴发性死亡的BP神经网络预测模型。结果表明,该模型的预测结果较好且稳定,预测结果与实际调查结果的拟合率为100%、准确率达到87.5%。因此运用本文方法建立环境因子、免疫因子与扇贝大规模死亡的预测模型是可行的。

在研究中我们体会到以下几点是建立人工神经网络模型的关键:(1)应尽可能地选取系统的、连续的资料,最好包括病害高发年和低发年的数据,收集影响因子时要抓住主要相关因子;(2)要结合专业知识、病情危害程度和采取防救措施所需的时间将扇贝死亡率进行适当分级,并根据网络学习训练的效果进行调整,使网络学习达到最佳效果;(3)调查资料的特征较大程度地影响网络学习的稳定性,因此数据的应用、纳入、排出问题有待进一步研究;(4)模型构建时各参数的确定没有现成的公式和答案,需要通过不断摸索来优化和解答。

将BP人工神经网络应用于养殖扇贝死亡的预测是一个新的研究方向,将其在实践中推广,仍需进一步深入研究。随着生态学、生物学、数学和计算机软件等相关学科的飞速发展,神经网络理论必将在水产动物病害的预测领域中拥有广阔的应用前景。

参考文献:

- [1] 王如才,王昭萍,张建中.海水贝类养殖学[M].青岛:青岛海洋大学出版社,1993:164-166.
- [2] 邵明凯,王海仁,张永刚,等.海阳县养殖栉孔扇贝大面积死亡原因分析[J].齐鲁渔业,1996,13(2):19.
- [3] 姚振刚.栉孔扇贝换季养殖技术[J].科学养鱼,2002,12:24.
- [4] 王远隆,杨晓岩.环境参数与扇贝的生长研究综述[J].齐鲁渔业,1991,35(4):29-31.
- [5] 孙惠玲,匡世煥,方建光,等.桑沟湾栉孔扇贝不同养殖方式及适宜养殖水层研究[J].中国水产科学,1996,4(1):60-65.
- [6] 于瑞海,王如才,田传远,等.栉孔扇贝大面积死亡原因分析及预防的探讨[J].海洋湖沼通报,1998,3:69-72.
- [7] Xiao J, Ford S E, Yang H S. Studies on mass summer mortality of cultured zhikong scallops (*Chlamys farreri* Jones et Preston) in China [J]. Aquaculture, 2005, 250: 602-615.
- [8] 王运涛,相建海.栉孔扇贝大规模死亡的原因探讨[J].海洋与湖沼,1999,30(6):770-774.
- [9] 周瑞莲,吕军仪,吴进锋,等.雷州半岛北部湾海域华贵栉孔扇贝的暴死原因调查情况[J].水产科技,2006,2:22-23.
- [10] 张福绥,杨红生.栉孔扇贝大规模死亡问题的对策与应急措施[J].海洋科学,1999,23(1):44-48.
- [11] 相建海.海水养殖生物病害发生与控制[M].北京:海洋出版社,2001:1, 29.
- [12] Weigend A, Gershenfeld N. Predicting the future and understanding the past [M]. Redwood City: Addison-Wesley, 1993: 3-11.
- [13] Haykin S. Neural networks [M]. New York: Macmillan, 1994.
- [14] 高春玉,熊鸿燕,易 东,等.智能神经网络模型用于气象因素对疟疾发病影响的初步研究[J].中华流行病学杂志,2003,24(9):831-834.
- [15] Kleinschmidt I, Sharp B L, Clarke G P, et al. Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa [J]. Am J Epidemiol, 2001, 153: 1213-1221.
- [16] Mukher A, Deshpande J M. Application of artificial neural networks in structural design expert systems [J]. Computer and Structures, 1995, 54: 367-370.
- [17] Abidi S S R, Goh A. Neural network based forecasting of bacteria-Antibiotic interactions for infectious disease control [C]. Ninth World Congress on Medical Informatics (Seoul). Amsterdam: IOS Press, 1998.
- [18] Gosukonda R M, Naghedolfeizi M, Carter J. Artificial neural networks to predict daylily hybrids [C]//Proceedings of the Fifth Computational Systems Bioinformatics Conference Workshops, Stanford: IEEE, 2005, 187-188.
- [19] 李祚泳,彭荔红.基于人工神经网络的农业病虫害预测模型及其效果检验[J].生态学报,1999,19(5):759-762.
- [20] 关 鹏,何 苗,周宝森.基于BP人工神经网络的病毒性肝炎发病率自回归预测模型[J].中国卫生统计,2003,20(5):269-271
- [21] 高 宁,邵陆寿.基于MATLAB的BP神经网络的农作物虫情测预报中的应用[J].计算机与农业,2003,7:16-16.
- [22] 冯天瑾.神经网络技术[M].青岛:青岛海洋大学出版社,1994:14.
- [23] Kohonen T. State of the art in neural computing [C]. First International Conference on Neural Networks, San Diego: IEEE, 1987, 1. 81
- [24] Rumelhart D, Hinton G, Williams R. Learning representation by backpropagating errors [J]. Nature, 1986, 323: 533.
- [25] Stornetta W, Huberman B. An improved three layer back propagation algorithm [C]//First International Conference on Neural Networks, San Diego: IEEE, 1987, 1: 637.
- [26] Yu S W, Zhu K J, Diao F Q. A dynamic all parameters adaptive BP neural networks model and its application on oil reservoir prediction [J]. Applied Mathematics and Computation, 2008, 95(1): 66-75.
- [27] Li P C, Li S Y. Learning algorithm and application of quantum BP neural networks based on universal quantum gates [J]. Journal of Systems Engineering and Electronics, 2008, 19(1): 167-174.
- [28] Toshio F, Takanori G. Theory and applications of neural networks for industrial control systems [J]. Transactions on Industrial Electronics, 1992, 39(6): 472.
- [29] 阚平凡,黄端旭.人工神经网络——模型·分析与应用[M].合肥:安徽教育出版社,1993:35.
- [30] Sofuoglu S C. Application of artificial neural networks to predict prevalence of building-related symptoms in office buildings [J]. Building and

- Environment, 2008, 43(6): 1121 - 1126.
- [31] 袁曾任. 人工神经网络及其应用[M]. 北京: 清华大学出版社, 1999: 118.
- [32] Bakhary N, Hao H, Deeks A J. Damage detection using artificial neural network with consideration of uncertainties [J]. Engineering Structures, 2007, 29(11): 2806 - 2815.
- [33] 马玉霞, 郑有飞. 模型参数对疾病发病率人工神经网络模型精度的影响[J]. 气象科学, 2003, 23(2): 153 - 160.
- [34] 陈绘画, 朱寿燕. 基于人工神经网络的马尾松毛虫发生量预测模型的研究[J]. 林业科学研究, 2003, 16(2): 159 - 165.
- [35] 从爽. 面向 MATLAB 工具箱的神经网络理论与应用[M]. 合肥: 中国科学技术大学出版社, 1999.
- [36] 匡胤. 基于人工神经网络的系统建模及 MATLAB 实现[J]. 四川理工学院学报(自然科学版), 2007, 20(5): 6 - 9.

Prediction model (ANN) for massive death of scallop *Chlamys farreri* in summer

DENG Deng¹, MAI Kang-sen¹, LI Xiao-ming², MA Hong-ming¹, TAN Bei-ping¹

(1. Marine Culture Key Laboratory, Ministry of Education, Ocean University of China, Qingdao 266003, China;
2. College of Communication Science and Engineering, Ocean University of China, Qingdao 266003, China)

Abstract: The mass mortality of cultured scallops *Chlamys farreri* often occurs in summer and brings huge loss to farmers. However, the loss could be reduced greatly by transferring and renewing scallop cages or harvesting scallops before the occurrence of the massive death of scallops. A model for predicting the death of scallops by using the principle of Artificial Neural Network (ANN), the method of Back Propagation (BP) network and MATLAB software has been developed. The data to build the prediction model were acquired from a two-year (2002 & 2003) investigation on temperature, salinity, pH, NH₃-N, NO₂-N of seawater and protein concentration, acid phosphatase activity, alkaline phosphatase activity, superoxide dismutase activity, catalase activity of scallops serum in Sanggou Bay, Rongcheng, Shandong Province. We debugged the model repeatedly by changing the key parameters, input layer node number, hidden layer node number, sample number and epoch number. After 300 times of studies and training, the sum-squared error of the prediction model decreased from 67.46 to 0.009 1. The model was tested, and the prediction accuracy was 87.5%. It is the first time that ANN was used in the aquaculture disease prediction. This model has many strong points, such as data adapting well, study momentarily and predicting accurately. The present study presents a new way for disease prediction and control of aquaculture animals.

Key words: *Chlamys farreri*; massive death; artificial neural network; prediction model