

马尾松自疏规律的人工神经网络模型研究

吴承祯 洪伟

(福建林学院资环系, 南平 353001)

摘要 森林自然稀疏规律的研究已经有了很大发展, 并提出了许多经验的或理论的表达式。本研究介绍了人工神经网络方法, 首次建立了马尾松人工林自然稀疏规律的三层前馈反向传播神经网络模型。仿真结果表明, 人工神经网络模型能很好地符合实际的观测资料, 具有良好的使用价值, 从而丰富了该领域的研究方法。

关键词 人工神经网络; 马尾松; 密度变化

中图分类号 Q948.121

STUDY ON SELF-THINNING LAW OF *PINUS MASSONIANA* BASED ON ARTIFICIAL NEURAL NETWORK

Wu Chengzhen Hong Wei

(Dept. of Resources & Environment, Fujian Forestry College, Nanping 353001)

Abstract The research on forest self-thinning has been developed greatly, and a lot of theoretical or empirical formulations were proposed. In this paper a method of artificial neural network is introduced, and a model of three-level feedforward back-propagation neural network to simulate self-thinning law of *Pinus massoniana* is established for the first time. The result of simulation shows that artificial neural network model can fit the observed data very well, and it is very useful in practice, which will enrich the research method of forest self-thinning.

Key words Artificial neural network; *Pinus massoniana*; Density change

植物种内竞争的结果—自然稀疏现象, 是植物种群生态的一大特征。自然稀疏又可叫作密度制约死亡率(density-dependent mortality)。植物种群所特有的自然稀疏现象引起了生态学家和林业工作者的极大兴趣, 最突出的成果就是反映同龄纯林存活密度和平均个体重量之间关系的负3/2自疏法则^[1], 这是植物种群生态学中颇为成熟的理论, 被人们称之为生态学中心法则^[2]和生态学基本原则^[3]。前人在研究森林自疏过程时较多地应用负3/2法则^[4-6]。然而, 不能否认这些法则的局限性, 因为它们要么不能拟合种群生活史的全过程, 要么缺乏与种群生长的连结。为此, 一些科学家试图利用密度与生长的理论, 提出能描述密度变化过程的理论或经验模型^[7-15], 但由于这些模型的复杂性在一定范围内又限制了它们的应用。由于森林自疏过程反映着不同年

龄阶段的林分株数的结构与动态规律, 即反映着自然状况下林分密度的变化规律, 因此, 研究探讨实用的森林自疏规律预报与模拟新方法, 提高预测预报水平, 对于实现森林密度的合理配置、提高林木产量、实现林业的可持续发展具有十分重要的理论价值和实际生产意义。为此, 本文首次提出应用人工神经网络的多层误差反传网络结构模型和一维时间序列拓展的方法, 发展了一种新的马尾松人工林自然稀疏规律预测与模拟模式。

1 人工神经网络的原理和实现

1.1 人工神经网络

人工神经网络(Artificial neural network, 简称 ANN) 是 80 年代以来国际上迅速兴起的一种非线性科学, 尤其近 20 年来形成了第二次研究新热潮, 并且在人工智能、信号处理、自动控制 and 模式识别等研究领域取得了令人瞩目的研究成果^[16-18]。但目前 ANN 在生态领域的应用尚未见报道。

人工神经网络是一种模拟人脑基本结构和功能来处理非线性与高阶性知识信息问题的新方法, 其特点是具有很强的自组织、自适应学习和记忆联想能力^[16]。在各种人工神经网络的模型中, 目前应用最为广泛的多层误差反传网络模型是一种非线性映射网络^[18], 是采用最小均方差学习方式, 在进行输入样本和期望输出的网络学习时是一种有教师的训练。学习阶段是由输出层误差反向逐层传播到输入层, 然后把所得结果与所需结果进行比较, 如果两者相差很大, 则通过调节连接权重再反传通过网络。如此反复, 只有当实际输出与期望输出之间差距足够小时, 才能说明网络学会了输入信息。人工神经网络的这种自适应学习能力和记忆联想能力, 使得它为解决复杂的非线性高维问题提供了有力的工具并且展示了广阔的前景^[17]。

1.2 三层误差反传网络方法的实现

我们在研究马尾松人工林自然稀疏规律模拟模型时, 从众多的人工神经网络模型中选用误差反传 Back propagation (B-P) 网络模型, 主要是因为 B-P 算法不仅理论依据坚实, 物理概念清晰, 通用性好; 而且网络在学习训练时, 除了学习因子和动量因子事先选定外, 网络的大量参数是网络对输入的原始数据进行不断学习后获得的。即 B-P 网络是通过学习, 从原始数据中“提取”信息逼近规律, 而不是人为赋予某种规律的。

以三层 B-P 网络为例, 具体证明给出网络的训练方法和过程。该网络包括一个输入层, 一个输出层和一个隐含层(如果是多层网络则有多个隐含层)。设网络的学习样本输入和期望输出为 $A_k, C_k(k=1,2,\dots,m)$ 。随机地给出网络从输入层到隐含层的初始连接权值为一组随机小量 V_{hi} , 同时也随机地给出一组隐含层到输出层的连接权函数 W_{ij} 以及隐含层单元的阈值 θ_i 和输出层单元的阈值 r_j 。然后对学习样本输入和期望输出 $A_k, C_k(k=1,2,\dots,m)$ 进行下列计算:

(1) 根据连接权矩阵(初始时刻为给定的一组随机小量)和学习样本的输入计算隐含层新的

$$\text{激活值: } b_i = f\left(\sum_{h=1}^n A_h V_{hi} + \theta_i\right) \quad (1)$$

其中 $i=1, 2, \dots, p$, 激活值函数为 Sigmoid 函数: $f(x) = 2/(1 + e^{-x}) - 1$ (2)

$$(2) \text{ 计算输出层单元的激活值: } C_j = f\left(\sum_{i=1}^p W_{ij} b_i + r_j\right) \quad (3)$$

其中 $j=1, 2, \dots, q$, (初始时刻 W_{ij} 为一组给定的随机小量)。

$$(3) \text{ 计算输出层单元的一般化误差: } d_j = C_j(1 - C_j)(C_j^k - C_j) \quad (4)$$

其中 $j=1, 2, \dots, q$, C_j^k 为输出层单元 j 的期望输出。

$$(4) \text{ 计算隐含层单元相对于每一个 } d_j \text{ 的误差: } e_i = b_i(1 - b_i)(\sum_{j=1}^q W_{ij}d_j) \quad (5)$$

其中 $i=1, 2, \dots, p$ 。

$$(5) \text{ 调整隐含层单元到输出层单元的连接权值: } \Delta W_{ij} = \alpha b_i d_j \quad (6)$$

其中 $i=1, 2, \dots, p$ 和 $j=1, 2, \dots, q$, α 为学习因子 ($0 < \alpha < 1$)。

$$(6) \text{ 调整输出层的阈值: } \Delta r_i = \alpha d_j \quad (7)$$

其中 $j=1, 2, \dots, q$ 。

$$(7) \text{ 调整输入层单元到隐含层单元的连接权值: } \Delta V_{hi} = \beta A_h e_i \quad (8)$$

其中 $h=1, 2, \dots, n$ 和 $i=1, 2, \dots, p$, β 为动量因子 ($0 < \beta < 1$)。

$$(8) \text{ 调整隐含层单元的阈值: } \Delta \theta_i = \beta e_i \quad (9)$$

(9) 重复第 (1) 到 (8) 式的计算步骤, 直到 $j=q$ 和 $k=m$, 计算实际输出与期望输出的误差, 当全部样本的输出误差小于设定的收敛误差时, 训练结束。在训练过程中, B-P 网络的权值是被逐步调整的, 通过训练, 网络从实际输入样本中获得了知识信息。并且可以看到, 存贮神经网络知识的正是由隐含层连接起来的输入层和输出层的连接权值和相应的阈值。当网络学习结束后, 这些连接权值和阈值都确定了, 根据这些参数和实际资料输入, 只要进行一些简单的加法和乘法运算即可进行模拟和预测。

2 马尾松自疏规律的人工神经网络 B-P 模型的建立

2.1 研究区概况与资料的收集

闽北地区, 地处福建省北部。其西北面以武夷山脉与江西交界, 北面以仙霞岭山脉与浙江毗邻, 东西面则有鹫峰、戴云山脉。全区山岭耸立, 丘陵起伏, 沟谷交错, 一般海拔在 500—1000 m。气候属中亚热带气候区域, 气候温暖, 雨量充沛, 湿度大, 阴天多, 年平均气温 14—19℃, 极端最高气温 39.7℃, 在高海拔处的极端最低温度 -11.5℃, 年平均降水量在 1500—1900 mm, 雨量分布不均, 最高在 7—8 月, 最低在 1—2 月, 年蒸发量 1200—1500 mm, 相对湿度为 80%—85%。土壤主要为发育在花岗岩、砂岩、片麻岩、页岩等成土母质上的红壤、黄红壤, 土层深厚, 质地为轻壤至中粘壤, 大部分 A+AB 层大于 25 cm, 土壤肥沃。该区主要植被为温暖湿润性针叶林, 群落组成种有杉木 (*Cunninghamia lanceolata*)、马尾松 (*Pinus massoniana*)、柳杉 (*Cryptomeria fortunei*)、湿地松 (*P. elliotii*) 及火炬松 (*P. taeda*) 等。马尾松一般分布于海拔 1000 m 以下, 多为人工林单优群落。

马尾松是强阳性树种, 当其林分生长到 5—6 年时, 树与树间的树冠开始出现重叠, 个体间对资源的竞争也就开始了; 随着时间的推移, 竞争加剧, 产生分化, 生长力弱的林木逐渐沦为被压木而最终死亡。我们在闽北马尾松人工林林区收集了 153 块标准地资料, 各标准地面积为 0.067 hm², 林木年龄都在 5—20 年之间, 郁闭度均接近 0.8 且林分结构较完整, 未受人为干

扰。据此, 应用“空间差异代替时间变化”^[12,15]来研究马尾松人工林自然稀疏规律, 整理得到马尾松人工林 5-20 年的密度动态。

2.2 马尾松自疏规律的人工神经网络模型的建立

用 B-P 网络模型对马尾松人工林自疏规律进行模拟, 取 5-18 年的马尾松人工林密度观测资料作为神经网络学习的基本原始数据, 并且将 19、20 年的马尾松人工林密度资料作为独立样本预报检验资料。

为了满足 B-P 网络 Sigmoid 函数的条件, 对一维马尾松人工林密度观测序列 (X_i) 作如下交换: $E_i = (X_i - S_{min} + \delta) / S_{max}$ (10)

式中: S_{min} 和 S_{max} 为样本序列中最小值和最大值, δ 取一小量, 以保证变换后的 E_i 序列最大值略小于 1 和最小值略大于 0。再根据后延时间 τ 将变换后的马尾松人工林密度一维时间序列 (E_i) 作飘移, 拓展成多维序列。其具体计算公式为, 对于一维时间序列: $E(t_1), E(t_2), \dots, E(t_n)$ (11)

按以下公式将其拓展成多维序列: $E(t_1), E(t_2), \dots, E(t_m)$ (12)

$$\begin{aligned}
 &E(t_1 + \tau), E(t_2 + \tau), \dots, E(t_m + \tau) \\
 &\dots \dots \dots \\
 &E(t_1 + k\tau), E(t_2 + k\tau), \dots, E(t_m + k\tau) \quad \text{其中 } m < n.
 \end{aligned}$$

取 $E(t_m), E(t_m + \tau), \dots, E(t_m + k\tau)$ 序列作为学习网络的期望输出, 其余序列作为网络学习的样本输入信息。本文在实际计算时是将 1 维观测序列, 延拓为 4 维, 因此, 网络的实际学习样本长度为 11。表 1 给出了由 (10) 和 (12) 式变换得出的马尾松人工林 5-18 年密度变化人工神经网络学习输入矩阵和期望输出。

表 1 标准化 B-P 网络训练集
Table 1 Standardized back propagation training set of the network

学习样本输入				期望输出			
Learning samples input				Expected output			
0.7282	0.7178	0.6342	0.5872	0.4430	0.3826	0.3490	0.2819
0.7178	0.6342	0.5872	0.5436	0.3826	0.3490	0.2819	0.2483
0.6342	0.5872	0.5436	0.5034	0.3490	0.2819	0.2483	0.1946
0.5872	0.5436	0.5034	0.4430	0.2819	0.2483	0.1946	0.1443
0.5436	0.5034	0.4430	0.3826	0.2483	0.1946	0.1443	0.0839
0.5034	0.4430	0.3826	0.3490				

采用三层误差反传 (B-P) 网络输出层为一个节点, 输入层为三个节点, 隐含层包括五个节点。在学习过程中学习因子和动量因子分别取 0.65 和 0.45, 神经网络中权系数和阈值的初值取 (0, 0.3) 之间的随机数。网络输出以 $H = \sum_{i=1}^n (A_i - C_i)^2$ 来考核网络学习状况, 并不断迭代使 H 趋于最小。根据公式 (1) - (9), 经过 12 850 次运算, 收敛误差 $H = 0.0102$ 趋于最小, 学习结束。B-P 模型列于表 2。

网络学习的拟合值和实测值变化见表 3。可以看出, 网络经过学习后拟合效果是令人满意的, 说明人工神经网络应用于马尾松人工林自疏规律研究是可行的。

表2 马尾松自疏规律的B-P模型

Table 2 Back propagation model of self-thinning law of *Pinus massoniana*

隐含层阈值 Threshold value in concealing layer	-0.03864	-0.005308	0.1163	-0.1143	-0.05594
输入层与隐含层连接权值 Linking value between input and concealing layer	0.3677	0.3059	0.1722	0.4352	0.3687
	0.3740	0.4547	0.2604	0.3464	0.4094
	0.2960	0.2169	0.2024	0.3623	0.2827
隐含层与输出层连接权值 Linking value between output and concealing layer	-0.2834	1.2018	1.1223	0.5809	1.5292
输出层阈值 Threshold value in output layer			1.1989		

表3 马尾松自疏过程观察值与模拟值 (Individuals hm^{-2}) 比较Table 3 Comparison of observed value and simulated value (Individuals hm^{-2}) of *Pinus massoniana* population density change during self-thinning

年龄	观察值	模拟值	年龄	观察值	模拟值	年龄	观察值	模拟值
Age (a)	Observed value	Simulated value	Age (a)	Observed value	Simulated value	Age (a)	Observed value	Simulated value
8	3840	3664	12	2925	3066	16	2085	2152
9	3645	3535	13	2775	2853	17	1860	1888
10	3465	3385	14	2475	2626	18	1590	1638
11	3195	3238	15	2325	2381			

3 B-P模型与其它模型比较

用建立的B-P网络模型计算训练样本,得到对应的模拟值,依此计算出相关指数和剩余标准差及平均模拟精度(表4)。前人在研究森林自然稀疏问题时多采用幂函数^[9],兰斌等^[10]曾针对前人所提出的密度自疏规律均为离散型数学模型的缺陷,提出应用连续型概率密度函数数学模型来研究马尾松人工林自疏规律。将B-P模型模拟结果与常用的幂函数及兰斌等^[10]用指数分布模型和Pearson-III分布模型拟合结果进行比较(表4)。可见,B-P网络模型的相关指数和模拟精度结果皆比幂函数、指数分布模型和Pearson-III型分布模型高,而剩余离差平方和皆比幂函数、指数分布和Pearson-III型分布模型小。而且在建立幂函数、指数分布和模型Pearson-III型分布时,常需对模型参数进行优化方能达到满意效果^[20]。而B-P模型自身具有学习训练功能,可以建立各种非线性映射,一般只要具备486以上的计算机条件即可进行工作,使用方便,而且精度可以人为控制,值得在生态学领域应用。

表4 B-P模型与其它模型的模拟结果比较

Table 4 Comparison back propagation model with other models for simulation

项目	B-P模型	幂函数	指数分布模型	Pearson-III分布模型
Item	B-P model	Power model	Index model	Pearson-III model
相关指数 Correlative index	0.99	0.97	0.98	0.94
剩余离差平方和 Surplus square	110804	365874.6	216325	679275
精度 Precision (%)	96.8	94.28	95.00	95.01

4 B-P模型预测

根据训练好的网络参数,对未参加学习训练的19、20年的马尾松人工林密度样本进行预测

预报, 并对这2年的独立样本预测预报结果进行检验, 确定其预报精度。从表5可见, 预报精度是令人满意的, 19、20年龄的预测预报平均相对误差仅为1.32%。由上述进行的预测步骤, 我们可以清楚地看到, 对于19、20年龄的马尾松人工林密度预测预报具有真正意义上的试验预报。因为在对人工神经网络学习时, 未来年龄的样本没有参加网络学习。另外, 森林自疏规律除与林分生长特性有关外, 还与前期林分密度有相关性, 本文采用将一维时间序列拓展和人工神经网络相结合的方法, 充分利用了这一相关性信息, 从而提高了预报精度。

表5 马尾松密度 (Individuals hm^{-2}) 变化预测结果检验

Table 5 Test of forecast results of density (Individuals hm^{-2}) change of *Pinus massoniana*

年龄 Age (a)	实测值 Observed value	预测值 Prediction value	绝对误差 Absolute error	相对误差 Relative error
19	1350	1334	16	1.19
20	1035	1020	15	1.45
平均 Mean			15.5	1.32

5 讨论

林木分化和剧烈竞争的结果必然使林分出现自然稀疏现象。马尾松人工林自然稀疏大致可分为3个阶段: 一是郁闭初期(5-12年), 此时林木高径生长缓慢, 林木分化较弱; 二是强烈的自然稀疏阶段(13-20年), 此时为高径速生期, 林木分化很明显, 被压木大量枯死, 两极分化明显; 三是自然稀疏微弱阶段(20年后), 进入材积生长盛期, 林木分化定型, 树冠疏开, 林分株数相对稳定。本文将马尾松人工林密度一维时间序列作飘移, 拓展成为多维序列再结合三层B-P网络模型建立了马尾松人工林自然稀疏规律模型, 研究结果表明所建立的B-P模型能切实反映马尾松人工林林分自然稀疏过程, 其模拟与预测精度能满足林业生产预测预报的要求。同时, 由于网络在学习训练后, 只要进行简单的运算即可作出未来时刻的马尾松人工林自疏密度预测, 十分适合林业生产预报推广。且采用人工神经网络和一维时间序列拓展的方法, 从马尾松人工林自疏实测序列中发掘有用的信息是一种新的尝试。其结果确实包含有以前很少利用的信息, 从而为森林自疏规律的研究提供了一种新的方法。

人工神经网络是一个高度的非线性映射, 它可以模拟人脑的若干基本特征, 是处理非线性复杂问题的有力工具。其在生态学领域的广泛应用还有待于广大生态学家共同努力, 本文仅作了一些尝试。在网络训练时, 有时会在极小点附近产生振荡导致不收敛以及根据不同的输入样本矩阵, 怎样选取更合适的模型结构, 学习参数和节点数等有关问题, 均有待于进一步深入探讨。

马尾松人工林自疏研究和成功预测预报, 对于马尾松资源的密度优化管理具有理论上和实际上的重大意义, 它可以指导我们选择抚育时间、抚育次数和抚育标准。随着马尾松人工林经营管理的深入, 可以重新收集其自疏过程密度资料建立新的B-P模型, 以维持其生命力和应用价值。

参考文献

- 1 Yoda K, Kira T, Qgawa H et al. Self-thinning in overcrowded pure stands under cultivated and natural

- conditions. *J Biol*, 1963, (14):107-129
- 2 White J. Demographic factors in populations of plants. In: Solbrig O T ed. *Demography and Evolution in Plant Population*. Oxford:Blackwell, 1980, 21-48
 - 3 Hutchings M L. Ecological laws in search of a theory. *New Scientist*, 1983, 98:765-767
 - 4 Silvertown J W. *Introduction of Plant Population Ecology*. Longman, London and New York, 1982, 3-296
 - 5 O'Neill R V, De D L. Angelis comparative productivity and biomass relations of forest ecosystem. In: Reichie D E ed. *Dynamic Properties of Forest Ecosystem*. Cambridge England: Cambridge University Press, 1981, 411-419
 - 6 赵学农. 哀牢山大果石栎林种群自我调节与竞争的初步研究. *植物生态学与地植物生态学学报*, 1991, 15(2):183-189
 - 7 Hozumi K. Ecological and mathematical considerations on self-thinning in even-aged pure stands. I. Mean plant weight-density trajectory during the course of self-thinning. *Bot Mag Tokyo*, 1977, 90:165-179
 - 8 Hozumi K. Ecological and mathematical considerations on self-thinning in even-aged pure stands. III. Effect of the linear growth factor on self-thinning. *Bot Mag Tokyo*, 1983, 96:171-191
 - 9 方精云. 一种描述植物种群自疏过程的经验模型. *林业科学*, 1995, 31(3):247-252
 - 10 兰斌, 洪伟, 苏宝川等. 马尾松幼龄林自然稀疏规律的研究. *林业勘察设计*, 1996, (1):10-13
 - 11 张大勇, 赵松龄. 森林自疏过程中密度变化规律的研究. *林业科学*, 1985, 21(4):369-374
 - 12 张利权. 浙江省松阳县黄山松种群的年龄结构与分布格局. *植物生态学与地植物学学报*, 1988, 14(4):328-334
 - 13 杨永祥. 云南松林分自然稀疏规律的研究. *云南林业科技通讯*, 1981, (4):419-425
 - 14 唐守正. 同龄纯林自然稀疏规律的研究. *林业科学*, 1993, 29(3):234-241
 - 15 董鸣. 缙云山马尾松种群年龄结构的初步研究. *植物生态学与地植物学学报*, 1987, 13(1):28-35
 - 16 庄镇泉. *神经网络与神经计算机*. 北京: 科学出版社, 1992, 7-51
 - 17 焦李成. *神经网络的应用与实现*. 西安: 西安电子科技大学出版社, 1993, 87-125
 - 18 Hechi-Nielsen R. Theory of back propagation neural network. *Int J Conf on Neural Network*, Washington D C, 1989, (1):593-605
 - 19 石培礼, 杨修, 钟章成. 桉柏混交林种群生物量动态与密度调节. *应用生态学报*, 1997, 8(4):341-346
 - 20 洪伟, 潘辉. 用三次设计法最优拟合 Taylor 幂法则模型. *福建林学院学报*, 1991, 11(3):253-259