

基于 ANN 的栎类天然次生林林分直径分布研究

闫东锋, 李纪亮*, 李国勇, 黄家荣, 郭欢欢, 李纪军 (1. 河南农业大学林学院园艺学院, 河南郑州 450002; 2. 中国科学院成都生物研究所, 四川成都 610041; 3. 河南平顶山市农业科学研究所, 河南平顶山 467000)

摘要 利用宝天曼自然保护区栎类天然次生林 32 块标准地的调查数据, 采用人工神经网络建模方法, 以相对直径作输入向量, 以株数累计频率作输出向量, 在 MATLAB 软件中进行模型训练, 建立人工神经网络模型。对主要树种建立的 7 个模型进行的精度分析表明, 该模型能够很好地描述栎类天然次生林主要树种的直径分布。

关键词 直径分布; 人工神经网络模型; 栎类天然次生林; 宝天曼

中图分类号 S75 文献标识码 A 文章编号 0517-6611(2007)34-11046-02

Study on Stand Diameter Distribution of Oak Natural Secondary Forest Based on ANN

YAN Dong-feng et al (College of Forest and Horticulture, Henan Agricultural University, Zhengzhou, Henan 450002)

Abstract Based on the data of the investigation on 32 standard plots of oak natural secondary forest in Baotianman Natural Reserve, with relative diameter as input vector and accumulated frequency of plant number as output vector, the artificial neural network modeling method was used to conduct model training in MATLAB software and establish the artificial neural network model. The precision analysis on 7 models established by main tree species showed that this model could better describe the diameter distribution of main tree species in oak natural secondary forest.

Key words Diameter distribution; Artificial neural network model; Oak natural secondary forest; Baotianman

长期以来, 林分直径分布规律的数学描述多借助于概率论中的各种概率分布函数^[1], 但这些传统方法运算繁琐、模型过于依赖数学方程式^[2]。人工神经网络(简称 ANN)是由大量简单神经元广泛连接而成, 用以模拟人脑思维方式的复杂网络系统, 以其具有良好的非线性映射能力、自学习适应能力和并行信息处理能力^[3]等独特的优点引起了广泛的关注。ANN 有无限逼近任意非线性系统的能力, 且不依赖于现存的数学函数, 在非线形系统建模方面具有明显的优势^[4]。目前, 已有利用人工神经网络模型方法来研究林分直径分布的研究报道, 但利用该方法研究栎类天然次生林还未见报道。笔者以河南宝天曼国家级自然保护区栎类天然次生林为研究对象, 用人工神经网络建模技术对林分主要树种直径规律进行研究, 以便更好地指导森林保护和经营。

1 研究区概况

宝天曼国家自然保护区位于河南省内乡县境内, 总面积达 9 304 hm²。该区地处暖温带向北亚热带过渡地区, 属季风型大陆性气候, 年平均气温 15.1℃, 年降水变率为 17.9%, 年平均降水量 855.6 mm, 年蒸发量 991.6 mm, 年平均相对湿度为 68%。区内土壤划分为山地棕壤、山地黄棕壤和山地褐土 3 个土类, 土层一般为 30~40 cm。是我国中部地区植物群落类型最复杂、植物种类最丰富的地区, 共有高等植物 223 科、1 002 属、2 711 种及变种。

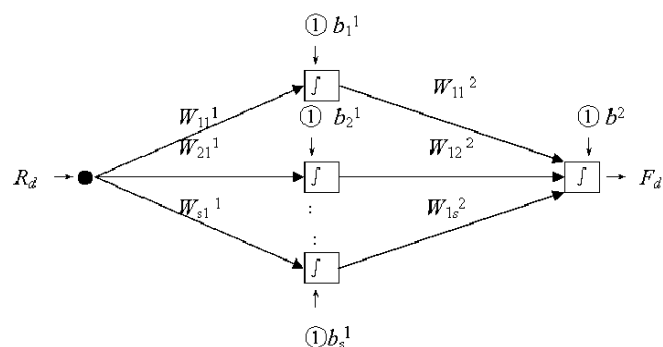
2 材料与方法

2.1 研究材料 在对保护区内栎类天然次生林的分布、生长、生境条件研究的基础上, 选取不同树种组成、不同生境条件、林相完整、有代表性的 32 块样地作为调查的标准地。对标准地内所有在起测径阶以上的林木进行每木检尺, 调查每株乔木树种的种类、高度、胸径、冠幅等诸项因子。

2.2 研究方法

2.2.1 模型建立。以相对直径作输入向量, 以株数累计频

率作输出向量, 隐层节点数 s 个(隐层节点数经过对比分析后确定), 相应的 BP 神经网络模型, 如图 1 所示, MATLAB 数学模型^[5]公式: $F_d = \text{logsig}(\sum_{i=1}^s W_{i1}^2 H_i + b^2)$; $H_i = \text{logsig}(W_{i1}^1 R_d + b_i^1)$ ($i = 1, 2, 3, \dots, s$)。式中, F_d 为株数累计频率; W_{i1}^2 为从隐层 i 节点至输出层 1 节点的连接权; H_i 为隐层 i 节点的输出值; b^2 为输出层 1 节点的阈值; W_{i1}^1 为从输入层的 1 节点至隐层的 i 节点的连接权; R_d 为径阶上限相对直径; b_i^1 为隐层 i 节点的阈值; s 为隐层神经元数, 可对其由少至多取不同的数值进行模拟。



输入值 1 的节点 输入层节点 表示作用函数为 logsig

图 1 栎类天然次生林林分直径分布神经网络模型结构

2.2.2 模型训练。从 32 块标准地中随机抽取 2 块作模型检验, 将剩下的 30 块标准地主要树种调查数据统计整理, 以相对直径作输入向量, 以株数累计频率作输出向量建立模型。隐层节点的确定是先给定一个较小的值, 根据训练的结果逐渐增加隐层节点数, 从而找出最佳节点数, 然后再输入适当的训练步数。在 MATLAB 软件中进行模型训练, 最终得到拟合效果最佳的模型。

2.2.3 模型检验。为检验所建立的神经网络模型的精度和应用推广能力, 用未参加模型训练的标准地主要树种数据对该模型进行精度检验, 并把检验拟合结果和实际结果进行一致性检验。

3 结果与分析

3.1 BP 神经网络模型拟合 通过对栎类天然次生林林分 7 个主要树种 BP 神经网络模型进行训练, 经过对比分析得到拟合各个树种直径分布规律效果最佳的神经网络模型, 结果见表 1。

基金项目 河南省教育厅自然科学基金项目(2007220002) 和河南农业大学林学院园艺学院青年教师科研基金项目。

作者简介 闫东锋(1979-), 男, 河南平舆人, 助教, 从事森林资源可持续经营管理方面的研究。* 通讯作者。

收稿日期 2007-07-06

表1 主要树种BP神经网络模型训练结果

树种	网络对象	网络结构	权值个数	训练步数	误差	拟合精度
锐齿栎	mynet1	1 2 1	7	1 000	5.94E 06	0.997 7
短柄柞树	mynet2	1 3 1	10	500	9.34E 06	0.996 5
化香	mynet3	1 3 1	10	500	7.73E 21	0.992 5
栓皮栎	mynet4	1 3 1	10	1 500	4.01E 05	0.992 2
椴树	mynet5	1 3 1	10	200	3.67E 06	0.998 5
鹅耳枥	mynet6	1 2 1	7	500	7.14E 04	0.969 2
三桠乌药	mynet7	1 2 1	7	200	7.42E 18	1.000 0

各模型的权值矩阵和阈值矩阵为:

$$\text{mynet1} : Iw\{1,1\} = [1.656 0 \quad -122.702 9]$$

$$Lw\{2,1\} = [11.276 1 \quad -0.381 1]$$

$$b\{1\} = [-1.535 7 \quad 60.983 0]$$

$$b\{2\} = [-5.373 7]$$

$$\text{mynet2} : Iw\{1,1\} = [-3.740 0 \quad -1.748 2 \quad 5.782 2]$$

$$Lw\{2,1\} = [112.216 8 \quad -486.402 6 \quad 255.831 4]$$

$$b\{1\} = [26.835 9 \quad 7.205 3 \quad 2.798 7]$$

$$b\{2\} = [116.801 4]$$

$$\text{mynet3} : Iw\{1,1\} = [-0.633 3 \quad 25.619 6 \quad 37.658 0]$$

$$Lw\{2,1\} = [-139.525 7 \quad 69.111 2 \quad -0.691 3]$$

$$b\{1\} = [3.611 1 \quad -6.243 6 \quad -24.561 3]$$

$$b\{2\} = [64.813 2]$$

$$\text{mynet4} : Iw\{1,1\} = [-431.886 9 \quad -23.011 0 \quad 2.337 5]$$

$$Lw\{2,1\} = [-15.769 7 \quad -1.007 0 \quad 10.650 3]$$

$$b\{1\} = [698.713 6 \quad 25.454 8 \quad -1.575 9]$$

$$b\{2\} = [9.804 3]$$

$$\text{mynet5} : Iw\{1,1\} = [46.802 0 \quad -29.914 7 \quad -2.853 0]$$

$$Lw\{2,1\} = [-1.609 7 \quad -12.298 8 \quad -3.938 1]$$

$$b\{1\} = [-101.640 2 \quad 74.7748 \quad 2.524 5]$$

$$b\{2\} = [16.209 7]$$

$$\text{mynet6} : Iw\{1,1\} = [9.162 5 \quad -9.162 5]$$

$$Lw\{2,1\} = [-5.171 1 \quad 2.149 4]$$

$$b\{1\} = [-173.919 9 \quad 6.719 9]$$

$$b\{2\} = [1.510 8]$$

$$\text{mynet7} : Iw\{1,1\} = [24.689 3 \quad 12.184 3 \quad -22.295 6]$$

$$Lw\{2,1\} = [24.546 0 \quad 2.719 3 \quad 0.500 8]$$

$$b\{1\} = [-35.558 7 \quad -9.538 2 \quad 2.383 1]$$

$$b\{2\} = [-1.873 1]$$

3.2 各径阶株数累计频率拟合精度 由于相对直径在理论上没有上限,在实际数据标定的最大径阶上限处,理论累计频率接近1但不等于1,从而使转算后的径阶频率合计不等于1。采用“截尾累入法”(BAILEY,1973),即截下最高径阶上限以后的尾部频率,将其累入最高径阶;截下最低径阶下限以前的频率将其累入最低径阶。按此处理后,将各径阶相对直径代入MATLAB数学模型公式,求出各径阶株数累计频率,然后再求出各径阶的频率,并与实测值比较。由实测值与拟合结果可知,7个主要树种种群的拟合总株数与实测总株数相等,各径阶株数基本吻合,为了进一步说明模型的拟合精度,可采用下式进行精度计算: $P_c = 1 - \frac{\sum(\text{abs}(\text{mynet1_errors}))}{\sum(f(d))}$ 。式中, mynet1_errors 为神经网络模型的拟合株数误差, $f(d)$ 为实测株数, \sum 为求和符号, abs 为求绝对值符号。所建模型的拟合精度计算结果,如表1所示。

3.3 模型的检验 用未参加模型训练的标准地数据作为检验样本,用来检验所建立的神经网络模型的优劣,检验对比结果,如表2、图2。

表2 栎类天然次生林林分直径分布检验

径阶	相对直径 cm	实测			检验		
		株数	频率	累计频率	株数	频率	累计频率
6	0.30	5	0.078 125	0.078 125	5	0.078 125	0.078 125
10	0.50	7	0.109 375	0.187 500	7	0.109 375	0.187 500
14	0.69	11	0.171 875	0.359 375	11	0.164 510	0.352 010
18	0.89	10	0.156 250	0.515 625	12	0.183 240	0.535 250
22	1.09	14	0.218 750	0.734 375	11	0.177 560	0.712 810
26	1.29	6	0.093 750	0.828 125	8	0.124 300	0.837 110
30	1.49	6	0.093 750	0.921 875	4	0.070 100	0.907 210
34	1.69	0	0.000 000	0.921 875	2	0.036 180	0.943 390
38	1.89	1	0.015 625	0.937 500	1	0.018 670	0.962 060
42	2.08	3	0.046 875	0.984 375	1	0.010 010	0.972 070
46	2.28	1	0.015 625	1.000 000	2	0.027 930	1.000 000

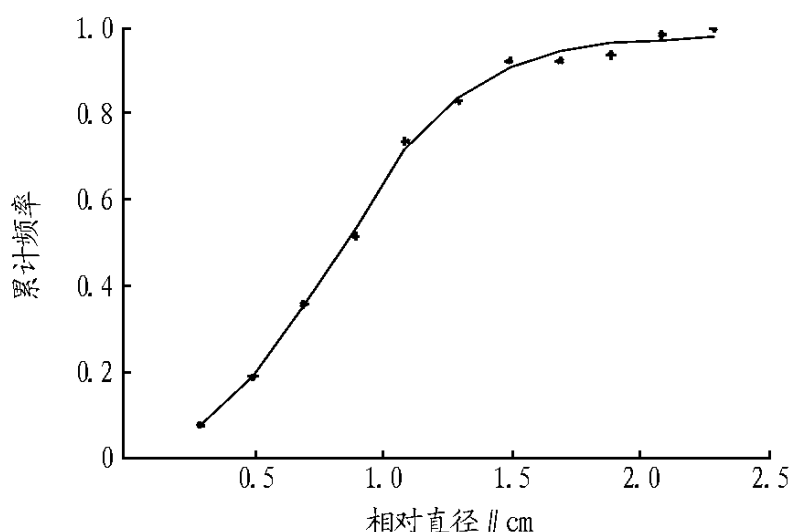


图2 模型检验拟合曲线

由表2可知,检验总株数与实测总株数相等,根据精度计算公式求出检验精度为97.95%,对林分直径的拟合达到了比较满意的效果。同时对模型进行了 χ^2 检验,分别取 $\alpha = 0.05$ 和 $\alpha = 0.01$,自由度 $f = 11 - 1 - 3 = 7$,查表得: $\chi^2_{0.05}(7) = 14.067$, $\chi^2_{0.01}(7) = 16.013$ 。经计算,所建神经网络模型的 χ^2 值为9.152,低于临界值,检验结果显著,进一步说明了所建立的BP神经网络模型对试验材料的直径分布拟合比较理想,具有很强的适应性和应用推广能力。

4 讨论

(1) 研究所建立起来的BP人工神经网络模型达到了理想的拟合效果,在对锐齿栎等7个主要树种的拟合中,拟合精度最高的为100%,最低的为96.92%。对所建立的BP人工神经网络模型检验结果也表明:检验精度达到97.95%,一致性检验达到显著水平。

(2) 在拟合时并不一定精度最高的模型就是最好的模型,还要结合林分的直径分布规律,采用定性定量相结合的方法,选择合适的神经网络模型,使其具有更强的泛化能力。

参考文献

- [1] 闫东锋,侯金芳,张忠义,等.宝天曼自然保护区天然次生林林分直径分布规律研究[J].河南科学,2006,24(3):364-367.
- [2] 陈晨,郭芳,黄家荣,等.杉木人工林直径分布BP模型的研究[J].河南农业大学学报,2005,39(4):390-393.
- [3] 黄家荣,孟宪宇,关毓秀.马尾松人工林单木生长神经网络模型研究[J].山地农业生物学报,2004,23(5):386-391.
- [4] 浦瑞良,宫鹏.应用神经网络和多元分析回归技术预测森林产量[J].应用生态学报,1999,10(2):129-134.
- [5] 黄家荣,任谊群.森林生长的人工神经网络建模[M].北京:中国农业出版社,2006.