

基于 BP 神经网络的马尾松人工林胸径-树高模型预测

卯光宪^{1,2}, 谭 伟^{1,2}, 柴宗政^{1,2}, 赵 杨¹, 杨深钧^{1,2}

(1. 贵州大学 林学院, 贵州 贵阳 550025; 2. 贵州大学 林业信息工程研究中心, 贵州 贵阳 550025)

摘要: 【目的】马尾松 *Pinus massoniana* 是中国南方主要用材树种, 建立高效的马尾松人工林胸径-树高预测模型, 可为马尾松人工林经营提供理论指导。【方法】以贵州省黔中地区马尾松人工林为研究对象, 基于 82 块样地 (25 m×25 m) 的 4 284 株马尾松单木数据, 选取 6 个常用的广义非线性模型进行拟合, 从中筛选出拟合效果最好的模型。使用相同的数据确定最佳隐层节点数量后, 经过反复训练建立基于 BP 神经网络的马尾松胸径-树高预测模型。【结果】在 6 个广义非线性模型中, 拟合效果最佳为 Korf 模型 ($R^2=0.650$); 马尾松适宜的隐藏层节点数为 2, 适宜的模型结构 (输入层节点数: 隐藏层节点数: 输出层节点数) 为 1:2:1, 模型预测精度达 0.717。【结论】广义非线性模型能较好地拟合马尾松人工林胸径-树高关系, 但与 BP 神经网络模型相比, BP 神经网络不需要依赖经验模型, 也不用模型筛选, 而且 BP 神经网络模型具有较高的决定系数和较低的均方根误差, 拟合精度优于广义非线性模型。图 5 表 5 参 35

关键词: 森林经理学; 胸径; 树高; 马尾松; BP 神经网络

中图分类号: S757.9

文献标志码: A

文章编号: 2095-0756(2020)04-0752-09

Diameter-height model for *Pinus massoniana* plantations based on BP neural network

MAO Guangxian^{1,2}, TAN Wei^{1,2}, CHAI Zongzheng^{1,2}, ZHAO Yang¹, YANG Shenjun^{1,2}

(1. College of Forestry, Guizhou University, Guiyang 550025, Guizhou, China; 2. Research Center of Forestry Information Engineering, Guizhou University, Guiyang 550025, Guizhou, China)

Abstract: [Objective] To provide theoretical guidance for the management of the plantation of *Pinus massoniana*, a wood widely used in the south, this study is aimed at the establishment of a high-efficiency diameter-height Model. [Method] With 4 284 *P. massoniana* plants from 82 sample plots in central Guizhou Province as the study subjects, 6 generalized nonlinear models were selected to fit the diameter-height relationship. After the determination of the optimal number of hidden layer nodes, repeated training was conducted to establish the BP neural network-based model of diameter-height relationship for *P. massoniana* plantation. [Result] Among the 6 generalized nonlinear models, the model that fits the best is Korf function ($R^2=0.650$) while the two hidden layers for BP neural network model were suitable for *P. massoniana* plantation with 1:2:1 as the appropriate model structure (number of input layer nodes: number of hidden layer nodes: number of output layer nodes) was and the model prediction accuracy reaching 0.717. [Conclusion] The generalized nonlinear model can fit the diameter-height relationship of *P. massoniana* plantation pretty well, yet in comparison, the BP neural network works better in that it's not reliant on empirical models and has better

收稿日期: 2019-08-21; 修回日期: 2020-01-16

基金项目: 贵州省林业厅资助项目 (黔林资复〔2012〕331 号); 贵州省科技支撑计划项目 (黔科合支撑〔2017〕2520-1 号); 贵州省科技计划项目 (黔科合基础〔2019〕1076 号); “十三五”国家重点研发计划项目 (2017YFD0601201)

作者简介: 卯光宪, 从事森林经理研究。E-mail: maoguangxian@163.com。通信作者: 谭伟, 教授, 博士, 从事森林可持续经营、林业信息与“3S”技术研究。E-mail: wtan@gzu.edu.cn

fitting accuracy for the high R^2 yet lower root mean squared error. [Ch, 5 fig. 5 tab. 35 ref.]

Key words: forest management; diameter at breast height; tree height; *Pinus massoniana*; BP neural network

胸径和树高是林业调查的重要因子, 胸径通常易测且结果精准, 但树高的测量往往低效费时, 结果偏差相对较大^[1]。因此, 为了更加方便准确地获得林木树高, 李春明等^[2]通过非线性混合效应模型很好地模拟了栓皮栎 *Quercus variabilis* 树高与胸径的关系; 张连金等^[3]通过幂函数模拟了 12 种不同林分的胸径与树高关系; CHAI 等^[4]选取常用的 16 个林业模型对贵州省平坝区柳杉 *Cryptomeria fortunei* 人工林树高-胸径进行拟合, 选出了拟合精度相对较高的 3 个模型。此类研究通常需要从常用的生长模型或其扩展模型中选出很多模型作为候选模型, 再对它们分别进行拟合, 通过一系列的模型评价指标对其进行筛选, 再结合树种生物学特性及模型的适用性选择出表现最好的模型。然而候选模型的种类和数量常常因人而异, 少则几种^[5-6], 多则有几十种^[7], 过少则精度不佳, 过多则计算量大, 费时费力。但是, 林木的生长过程复杂, 导致传统的模型难以达到较高的预测精度。人工神经网络不依赖现存模型, 可以最大限度地逼近任意非线性变化趋势。BP(back propagation) 神经网络也称误差反向传播网络, 在很多领域均有应用, 是当前应用最广的神经网络模型之一, 也是林业研究中常用的模型。在林业调查中, 通过建立树高-胸径的关系模型来预测树高往往具有事半功倍的效果, 同时在林木生长预测、生物量的估计, 森林资源的监测与评价中具有较高的使用价值^[4]。刘鑫等^[8]、董云飞等^[9]运用 BP 神经网络分别拟合了天然云冷杉 *Abies fabri* 针阔混交林和杉木 *Cunninghamia lanceolata* 的树高曲线, 表明 BP 神经网络有操作简单、拟合精度高、泛化能力强等特点。目前以 BP 神经网络预测树高的研究主要集中于长白山落叶松 *Larix olgensis*^[10]、油松 *Pinus tabulaeformis*^[11]、兴安落叶松 *Larix gmelinii*^[12] 等树种构成的林分。马尾松 *Pinus massoniana* 作为南方主要用材树种, 由于经营不及时、地理环境特殊等原因, 大面积马尾松人工林林分结构不合理等^[13-16], 并且对其树高模型预测方面的研究还较少。基于此, 本研究以黔中马尾松人工林为对象, 通过胸径-树高建立 BP 神经网络, 并与 6 个候选模型中筛选出的最佳模型进行比较, 探究模型在马尾松人工林树高预测中的表现和适用性, 以期对马尾松人工林树高估测及森林资源的监测与评价提供理论依据。

1 研究地区与研究方法

1.1 研究区概况

研究区 (26°10'~27°21'N, 106°59'~107°16'E) 位于黔中息烽县、开阳县、乌当区、龙里县、平坝区 5 个县 (区), 该区地貌类型复杂多样, 按形态划分为山地、丘陵和坝地三大类型; 地势起伏较大, 大致呈东西向延展, 西部、南部较高, 东部、北部较低, 地形坡度大多都在 25°以下; 平均海拔 1 200 m 左右, 属于亚热带湿润温和型气候, 年均气温 15.3 ℃, 年均降水量 1 200 mm; 森林土壤主要有黄壤、石灰土和紫色土。黄壤为地带性土壤, 一般土层较深厚。常见乔木主要有马尾松、枫香 *Liquidambar formosana*、华山松 *Pinus armandii*、柳杉、楸树 *Catalpa bungei*、梓树 *Catalpa ovata*、亮叶桦 *Betula luminifera*、白栎 *Quercus fabri*、麻栎 *Quercus acutissima* 等。

1.2 研究方法

1.2.1 数据来源 数据来源于 5 个黔中县 (区) 马尾松可持续经营试点区域经营对照样地, 起源为人工林, 未实施过任何经营措施, 且人为干扰相对较小。其中, 样地数为息烽县 4 个, 开阳县 8 个, 乌当区 22 个, 龙里县 16 个, 平坝区 32 个, 样地大小为 25 m×25 m, 调查林分基本情况 (坡向、坡位、海拔、郁闭度等), 对样地中胸径大于 5 cm 的林木进行每木检尺并通过网格计算其相对位置, 记录因子有树高、胸径、冠幅、第一活枝高、林木质量、相对位置, 在样地的 4 个角及中间设置 5 个 5 m×5 m 的样方, 记录主要灌木的平均高度、地径、株数、冠幅、盖度, 同时设置 5 个 1 m×1 m 的的样方, 记录主要草本的平均高度、盖度。

1.2.2 传统胸径-树高模型的建立 基于胸径-树高散点图 (图 1) 进行模型拟合, 其中, 训练数据 3 430 株, 验证数据 854 株。基于模型的数学特性及模型参数的生物学意义, 选择常用的 6 个非线性生长模型 (表 1)。

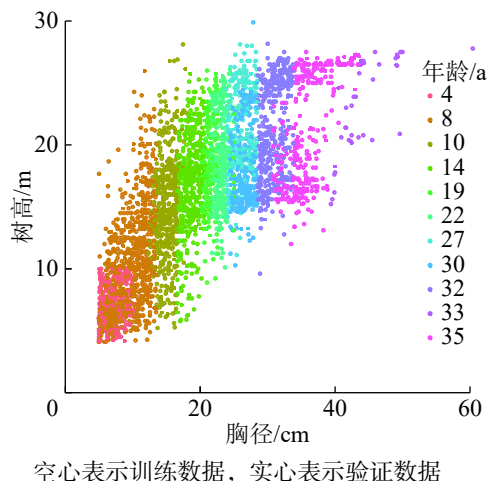


图1 马尾松模型拟合胸径-树高散点图

Figure 1 Diameter-height scatter diagram for model fitting data for

P. massoniana

表1 非线性胸径-树高候选模型

Table 1 Nonlinear diameter-height growth models selected for use			
模型名称	参数/个	表达式	参考文献
Näslund	2	$H(D) = 1.3 + \frac{D^2}{(aD+b)^2}$	[17]
Curtis	2	$H(D) = 1.3 + \frac{aD}{(1+D)^b}$	[18]
Logistic	3	$H(D) = 1.3 + \frac{a}{1 + b\exp(-cD)}$	[19]
Weibull	3	$H(D) = 1.3 + a[1 - \exp(-bD^c)]$	[20]
Gomperz	3	$H(D) = 1.3 + a\exp[-b\exp(-cD)]$	[21]
Korf	3	$H(D) = 1.3 + a\exp(-bD^{-c})$	[22]

说明: H 表示树高; D 表示胸径; a , b , c 为模型参数

1.2.3 BP神经网络 这是多层前馈神经网络的一种,其特点为:信号向前传播,误差反向传播(图2)。过程主要分为2个阶段:第1阶段是信号向前传播,由输入层经过一到多个的隐藏层,最后到达输出层;第2阶段是误差反向传播的过程,从输出层到隐藏层,再到输入层,根据误差调节隐藏层到输出层的权重和偏置,直到达到目标误差限或预设训练次数为止。神经网络的基本组成单元见图3, $x_1 \sim x_n$ (1, 2, 3, ..., n) 为神经元的输入信号, $\omega_1 \sim \omega_n$ (1, 2, 3, ..., n) 为权重, b 为偏置, Σ 为求和节点, σ 为激活函数。根据 Kolmogorov 定理,单隐层的 BP 神经网络就可以无限逼近任意连续的非线性曲线^[16]。但过分拟合会导致模型的泛化能力差,所以在实际的应用中还要结合样本本身的特性来控制拟合程度。

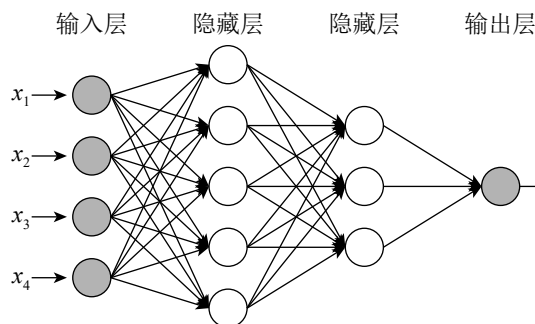


图2 BP神经网络结构图

Figure 2 BP neural network structure

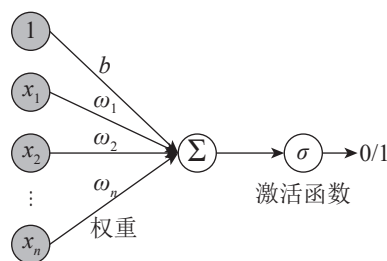


图3 神经元结构图

Figure 3 BP neural structure

1.2.4 BP神经网络模型的建立 基于 Matlab(2016a) 软件自带工具箱 nntool, 建立单隐层的 BP 神经网络以估测树高, 以输入层为胸径, 以输出层为树高。根据 BP 神经网络的隐层节点数可用 $N_h = \sqrt{N_{in} + N_{out}} + h$, 其中: N_h 为隐层节点数, N_{in} 为输入层节点数, N_{out} 为输出层节点数, h 为 1~10 的整数^[23]。本研究通过在范围内逐个测试, 对比其估测精确度及泛化能力以确定最佳隐层节点数。建模时, 设置目标精度为 0.001, 最大迭代次数为 1 000, 学习率为 0.01, 因胸径-树高关系曲线形似“高关型曲线”, 所以以 sigmoid(logsig) 函数为隐层传递函数, 以 purelin 函数输出层传递函数, Levenberg-Marquardt 算法相对其他算法有缩短训练时间的特点, 因此以 Levenberg-Marquardt 算法为模型训练算法。

1.2.5 模型拟合效果评价 对胸径-树高模型的选择通过决定系数 (R^2)、相对平均绝对误差 (R_{MA})、均方根误差 (R_{MSE}) 和 Akaike 信息准则 (A_{IC}) 评价模型的拟合效果(表2), R^2 越大、 R_{MA} 越小、 R_{MSE} 越小、 A_{IC} 越小, 模型的拟合精度越高。

1.2.6 数据处理 数据统计使用 Excel 2016, BP 神经网络的构建及拟合使用 Matlab(2016a) 软件自带工具箱 nntool, 70% 为训练数据, 15% 为验证数据, 15% 为测试数据; 传统的胸径-树高模型的拟合主要使

用 R 语言 `sampling` 程序包进行分层抽样，抽取 80% 为训练数据，20% 为检测数据，借助 `lmfor` 程序包进行模型拟合，`ggplot2` 程序包作图。

2 结果与分析

2.1 样地马尾松基本情况

所选样地马尾松平均年龄为 18 a，胸径为 5.0~60.4 cm，树高为 4.1~29.9 m(表 3)。训练数据包括 3 430 株林木，胸径为 5.0~50.0 cm，树高为 4.1~29.9 m，覆盖了整个径阶和树高范围。测试数据包括 854 株林木，胸径为 5.0~60.4 cm，树高为 4.1~27.8 m，基本覆盖所有径阶和树高范围，说明分层抽样的可靠性。

表 2 模型评价指标

Table 2 Model performance criteria selected for this study	
评价指标	表达式
决定系数(R^2)	$R^2 = 1 - \frac{\sum_{i=1}^n (H_i - \widehat{H}_i)^2}{\sum_{i=1}^n (H_i - \overline{H}_i)^2} \frac{n-1}{n-p-1}$
相对平均绝对误差(R_{MA})	$R_{MA} = \sum_{i=1}^n \left \frac{H_i - \widehat{H}_i}{\widehat{H}_i} \right / n \times 100\%$
均方根误差(R_{MSE})	$R_{MSE} = \sqrt{\frac{\sum_{i=1}^n (H_i - \widehat{H}_i)^2}{n-p}}$
Akaike信息准则(A_{IC})	$A_{IC} = n \times \ln R_{MSE} + 2p$
说明： H_i 表示树高实测值； $\widehat{H}_i$ 表示树高预测值； $\overline{H}_i$ 表示树高平均值； n 表示样本量； p 表示模型参数数量	

表 3 建模数据和检测数据统计

Table 3 Summary statis for all sampled trees and trees used for model calibration and validation									
径阶/cm	株数/株	胸径/cm				树高/m			
		平均值	标准差	最大值	最小值	平均值	标准差	最大值	最小值
6	370	5.9	0.6	6.9	5.0	7.1	2.2	18.7	4.0
8	345	7.9	0.6	8.9	7.0	8.4	2.5	19.6	4.2
10	300	9.9	0.6	10.9	9.0	9.8	3.1	24.3	4.2
12	294	11.9	0.6	12.9	11.0	12.4	4.0	26.0	4.3
14	297	14.0	0.6	14.9	13.0	14.4	3.9	24.3	5.2
16	275	15.9	0.6	16.9	15.0	15.3	4.0	26.9	5.3
18	289	18.0	0.6	18.9	17.0	16.8	4.0	28.1	5.8
20	257	20.0	0.6	20.9	19.0	18.0	3.3	25.3	6.7
22	312	22.0	0.6	22.9	21.0	18.7	3.4	26.0	8.9
24	314	23.9	0.6	24.9	23.0	19.2	3.6	26.3	10.1
26	287	26.0	0.6	26.9	25.0	19.7	3.7	28.1	10.3
28	228	27.9	0.6	28.9	27.0	20.0	3.8	29.9	9.6
30	205	29.9	0.6	30.9	29.0	20.3	3.5	28.2	13.3
32	181	31.9	0.6	32.9	31.0	21.1	3.9	27.5	12.8
34	132	33.9	0.6	34.9	33.0	20.6	4.0	27.0	12.0
36	90	35.8	0.6	36.9	35.0	21.5	4.4	27.5	13.6
38	43	37.8	0.6	38.9	37.0	22.2	4.3	27.0	13.1
40	33	40.0	0.6	40.9	39.0	24.0	3.8	27.0	15.2
42	30	42.0	0.6	42.9	41.0	25.1	2.5	27.3	20.2
44	15	43.6	0.4	44.3	43.0	25.6	2.2	27.3	20.6
46	4	45.9	0.5	46.5	45.5	24.9	3.2	27.5	20.4
50	8	49.4	0.4	50.0	49.0	26.3	2.2	27.5	20.9
60	1	60.4	—	60.4	60.4	27.8	—	27.8	27.8
训练数据	3 430	19.5	9.3	50.0	5.0	15.9	6.0	29.9	4.1
验证数据	854	19.3	9.3	60.4	5.0	15.7	6.0	27.8	4.1

说明：“—”表示空值

2.2 基于广义非线性模型预测胸径-树高关系

2.2.1 模型训练 所有的模型拟合都是基于 3 430 株林木的胸径，拟合结果如图 4 所示。整体来看，Logistic 模型拟合效果最差，对于胸径较小和中等 (5.0~25.0 cm) 的林木，除 Logistic 模型外，其他模型

的拟合效果都比较好；对于胸径较大的林木，由于林木树高分化严重，所以6个模型拟合效果都不是很好。

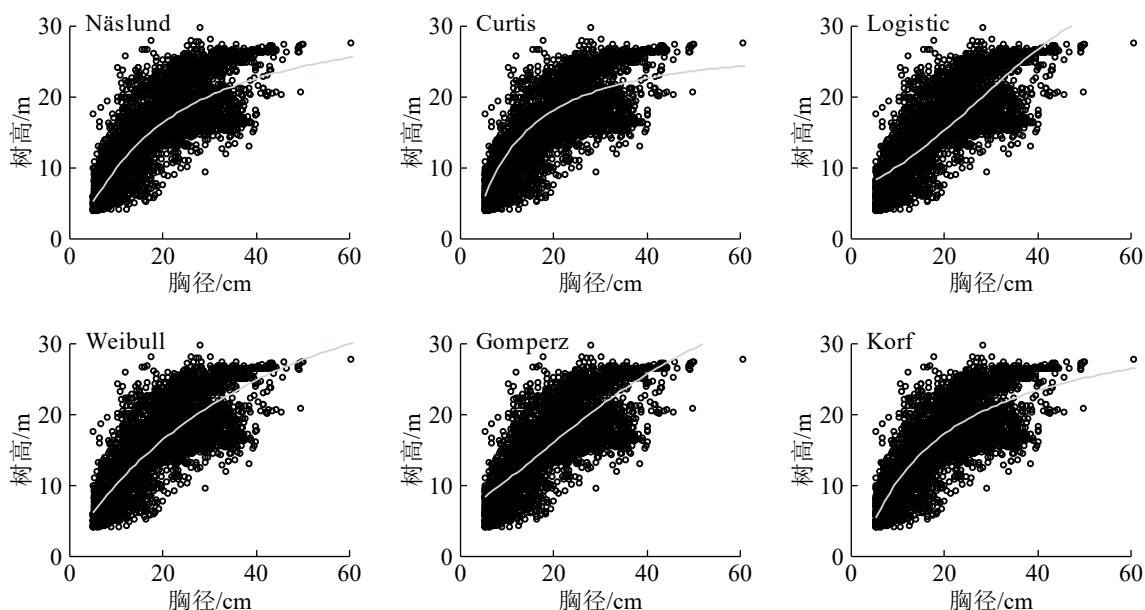


图4 马尾松人工林6个非线性模型胸径-树高曲线

Figure 4 Six nonlinear diameter-height curves for a *P. massoniana* plantation

从6个非线性模型中选择出对马尾松胸径-树高关系拟合效果最佳的模型，拟合结果如表4所示。由表4可知：除了Logistic模型外，其他模型拟合效果都较好， R^2 最小值为0.568， R^2 从小到大依次为Logistic模型、Gompertz模型、Curtis模型、Weibull模型、Näsrlund模型、Korf模型， R_{MSE} 从大到小依次为Logistic模型、Gompertz模型、Curtis模型、Weibull模型、Korf模型； R_{MSE} 最大值仅为4.218。综合考虑4个评价标准，Korf模型拟合效果最佳。

2.2.2 模型测试 用于模型测试的林木854株，胸径(5.0~60.4 cm)。由表4可知：表现最佳的是Korf模型，检测样本的模型拟合结果与训练样本一致，Korf模型和Näsrlund模型能够较好地预测林木树高，可选取这2个模型作为研究区马尾松人工林胸径-树高预测模型。此外，由于Korf模型为3参数模型，较2参数模型复杂，但是当研究区大、样本较复杂时有较好的拟合效果，而2参数模型较适合于小面积、数据量较小的情况。因此，选用Korf模型为最佳的胸径-树高模型。

表4 马尾松人工林6个候选模型的参数估计及性能评价

Table 4 Parameter estimates and performance criteria of 6 nonlinear height-diameter models for a *P. massoniana* plantation

模型	a	b	c	训练样本				验证样本			
				R^2	$R_{MA}/\%$	R_{MSE}	A_{IC}	R^2	$R_{MA}/\%$	R_{MSE}	A_{IC}
Näsrlund	1.614	0.176	—	0.635	25.018	4.044	23 176.40	0.623	20.089	3.630	4 626.377
Curtis	27.15	9.538	—	0.626	23.808	4.094	23 282.51	0.622	18.645	3.634	4 628.369
Logistic	37.18	5.804	0.064	0.568	27.210	4.402	23 903.70	0.551	21.652	3.963	4 777.537
Weibull	37.18	0.031	0.941	0.629	24.640	4.079	23 252.32	0.615	19.600	3.671	4 646.573
Gompertz	37.18	2.016	0.039	0.603	25.797	4.218	23 538.00	0.589	20.493	3.790	4 701.199
Korf	37.18	6.796	0.698	0.650	23.585	3.963	23 003.99	0.641	18.877	3.545	4 587.180

说明：“—”表示空值

2.3 基于BP神经网络预测胸径-树高关系

2.3.1 隐层节点数的确定 以胸径为输入变量，以树高为输出变量，建立隐层为 N_h 的BP神经网络。根据经验公式求得隐层数为2.414~11.414，由于神经网络训练结果有一定的波动，为增加模型的容错率，

一次将隐层数设置为 2~11 的整数，每个节点数训练 10 次并计算 R^2 和 R_{MSE} 的平均值 (表 5)，随着隐层节点数的增加，所有数据 R_{MSE} 最大仅为 3.742，大部分 R^2 都大于 0.700，但在测试数据中存在 R^2 小于 0.700 的情况。对于训练数据和验证数据，随着隐层节点的增大， R^2 也逐渐增大， R_{MSE} 对于训练数据和验证数据的变化情况与 R^2 相同，随着隐层节点数的增加逐渐减小；对于测试数据，因为是独立验证样本，未参与建模，所以测试数据的 R^2 和 R_{MSE} 上下波动，最大和最小分别为 0.717 和 3.513。

对于测试数据，随着隐层数的增加，预测精度未得到明显提升。为进一步确定隐层节点数量，将神经网络的预测树高与对应的胸径建立散点图，当数据出现过拟合时说明该隐层节点数已不可取。从图 5 可见：从节点数为 3 开始就出现了分化及变形，所以确定此时已经出现过拟合的情况。考虑到模型的实用性及马尾松的生物学特性，故选择图像出现过拟合时隐层节点数的前一个节点数作为最佳隐层节点数，即最佳隐层节点数为 2。

2.3.2 适宜模型 确定最佳隐层节点数后，在同一个隐层节点数的基础上不断训练模型，最终选择合适的结构 [输入层节点数 (N_{in}):隐层节点数 (N_h):输出层节点数 (N_{out})=1:2:1] 为最后的模型，经训练得到的传递函数模型： $h_1=\text{logsig}(2.580\ 6+0.510\ 32D)$ ； $h_2=\text{logsig}(1.614\ 1+1.512\ 8D)$ ； $H=\text{purelin}(1.111+2.155\ 1h_1+1.403\ 5h_2)$ 。其中： h_i 为隐层神经元的输出， $i=1, 2$ ； D 为胸径；purelin 为线性函数；logsig 为对数 S 型函数； H 为树高。

2.4 模型的比较及验证

运用 6 个广义非线性模型 (表 4) 与 BP 神经网络 (表 5) 对 82 块样地的抽样数据对比发现：广义非线性模型的 R^2 均为 0.500~0.700， R_{MSE} 最小为 3.963；而 BP 神经网络训练数据、验证数据、测试数据的 R^2 均大于 0.700， R_{MSE} 最大仅为 3.742，优于广义非线性模型的最大值。因此，在马尾松胸径-树高建模方面，BP 神经网络模型比广义非线性模型效果好。

3 讨论

在研究胸径-树高的关系时，模型选择不但要考虑模型的性能及拟合精度，还需考虑其生物学意义^[24]。尽管所有模型都具有较好的拟合效果，但是 Weibull、Näslund、Korf、Gomperz、Curtis 这 5 个模型对马尾松人工林胸径-树高关系拟合的效果相近，均具有较高的 R^2 和较低的 R_{MA} 、 R_{MSE} 、 A_{IC} 。与此前的大部分研究结果相似^[25-27]，CHAI 等^[4] 表明：Weibull 和 Näslund 模型的拟合精度最高。SHARMA 等^[28] 通过对挪威云杉 *Picea abies* 胸径-树高关系的拟合也发现：Weibull 模型有较好的拟合精度。这可能是因为这几个模型具有较好的灵活性，能够通过参数来调节曲线形状以适应各个树种的生长情况。此外，树高的生长随着胸径的生长逐渐变大后趋于平缓，生长曲线呈 S 型，而这 5 个函数都是 S 型模型，因而拟合效果较好。

本研究显示：最佳模型结构 (输入节点数:隐藏层节点数:输出层节点数) 为 1:2:1，训练数据、验证数据、测试数据的 R^2 分别达 0.712、0.717、0.717， R_{MSE} 分别达 3.546、3.655、3.513。与传统模型相

表 5 不同隐层节点数的 10 次拟合统计量平均值

Table 5 Average statistics of fitting with different hidden layers for 10 times

节点数	决定系数 R^2			均方根误差 R_{MSE}		
	训练数据	验证数据	测试数据	训练数据	验证数据	测试数据
2	0.712	0.717	0.717	3.546	3.655	3.513
3	0.713	0.713	0.709	3.535	3.643	3.554
4	0.714	0.714	0.695	3.527	3.564	3.554
5	0.715	0.714	0.695	3.346	3.512	3.645
6	0.722	0.720	0.703	3.335	3.412	3.616
7	0.724	0.721	0.710	3.310	3.345	3.598
8	0.726	0.726	0.708	3.282	3.312	3.612
9	0.726	0.726	0.705	3.256	3.245	3.721
10	0.728	0.728	0.709	3.238	3.231	3.645
11	0.729	0.729	0.701	3.234	3.203	3.742

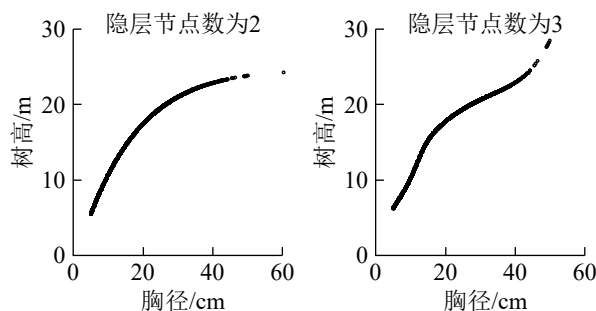


图 5 过拟合前后胸径-树高散点图

Figure 5 Diameter-height scatter diagram for over-fitting

比, BP神经网络模型的 R^2 均明显大于传统模型, R_{MSE} 均小于传统模型, 这与黄旭光等^[29]、徐志扬^[30]的结果相似, 说明BP神经网络在建模方面比传统模型更具有优越性。杜志等^[31]通过对杉木和马尾松的树高曲线拟合得到相同结论。王轶夫等^[32]在对马尾松生物量模型的研究中也发现: BP神经网络优于传统模型。由于无限逼近原理, 所以在拟合过程中容易出现过拟合的情况, 本研究通过输入变量与输出变量所形成的散点图判断得出: 当隐层节点数为3时出现过拟合。本研究结果可为研究区马尾松人工林的经营节约调查成本, 提高经营效率。充分考虑地理因素所训练出的模型具有较广的适用性^[33], 后续研究应充分调查不同立地条件下的该类林分, 增加样本量和调查因子, 增多神经网络的输入变量, 以提高模型的泛化能力。

BP神经网络在拟合效果上虽然较传统模型好, 但是其内部结构并不清楚, 无法展示其详细的内部计算过程, 这是BP神经网络的不足之处。同时, 为了预防BP神经网络过早的产生过拟合, 后续研究可使用改进的自适应遗传算法^[34]、改进粒子群算法^[35]等优化BP神经网络以达到更好的预测效果。

4 结论

本研究选用6个林业常用的传统模型, 对胸径-树高关系进行拟合, 结果发现: 除了Logistic模型外, 其他5个候选模型均表现较好。考虑研究区较大, 所以选择具3个参数的Korf模型作为马尾松胸径-树高模型。而用BP神经网络对树高进行估测显示: 最佳模型结构(输入节点数:隐藏层节点数:输出层节点数)为1:2:1, BP神经网络模型的 R^2 均明显大于传统模型, R_{MSE} 均小于传统模型, 说明BP神经网络在建模方面比传统模型更具有优越性, 更适合黔中地区马尾松人工林树高预测。

5 参考文献

- [1] MEHTÄTALO L, de MIGUEL S, GREGOIRE T G. Modeling height-diameter curves for prediction [J]. *Can J For Res*, 2015, **45**(7): 826–837.
- [2] 李春明, 李利学. 基于非线性混合模型的栓皮栎树高与胸径关系研究[J]. *北京林业大学学报*, 2009, **31**(4): 7–12.
LI Chunming, LI Lixue. Height-diameter relationship for *Quercus variabilis* Blume plantations based on nonlinear mixed model [J]. *J Beijing For Univ*, 2009, **31**(4): 7–12.
- [3] 张连金, 孙长忠, 辛学兵, 等. 北京九龙山不同林分树高与胸径相关生长关系分析[J]. *中南林业科技大学学报*, 2014, **34**(12): 66–70.
ZHANG Lianjin, SUN Changzhong, XIN Xuebing, et al. Allometric relationship between height and diameter at breast height of different stand in Beijing Jiulong Mountain [J]. *J Cent South Univ For Technol*, 2014, **34**(12): 66–70.
- [4] CHAI Zongzheng, TAN Wei, LI Yuanyuan, et al. Generalized nonlinear height-diameter models for a *Cryptomeria fortunei* plantation in the Pingba region of Guizhou Province, China [J]. *Web Ecol*, 2018, **18**(1): 29–35.
- [5] 童洁, 石玉立. 加格达奇3种森林类型树高-胸径的曲线拟合[J]. *东北林业大学学报*, 2017, **45**(2): 6–11, 21.
DONG Jie, SHI Yuli. Tree height-diameter model in Jiagedaqi [J]. *J Northeast For Univ*, 2017, **45**(2): 6–11, 21.
- [6] 王冬至, 张冬燕, 张志东, 等. 基于非线性混合模型的针阔混交林树高与胸径关系[J]. *林业科学*, 2016, **52**(1): 30–36.
WANG Dongzhi, ZHANG Dongyan, ZHANG Zhidong, et al. Height-diameter relationship for conifer mixed forest based on nonlinear mixed-effects model [J]. *Sci Silv Sin*, 2016, **52**(1): 30–36.
- [7] 卢军, 张会儒, 雷相东, 等. 长白山云冷杉针阔混交林幼树树高-胸径模型[J]. *北京林业大学学报*, 2015, **37**(11): 10–25.
LU Jun, ZHANG Huiru, LEI Xiangdong, et al. Height-diameter models for saplings in a spruce-fir mixed forest in Changbai Mountains [J]. *J Beijing For Univ*, 2015, **37**(11): 10–25.
- [8] 刘鑫, 王海燕, 雷相东, 等. 基于BP神经网络的天然云冷杉针阔混交林标准树高-胸径模型[J]. *林业科学研究*, 2017, **30**(3): 368–375.
LIU Xin, WANG Haiyan, LEI Xiangdong, et al. Generalized height-diameter model for natural mixed spruce-fir coniferous and broadleaf forests based on BP neural network [J]. *For Res*, 2017, **30**(3): 368–375.
- [9] 董云飞, 孙玉军, 王轶夫, 等. 基于BP神经网络的杉木标准树高曲线[J]. *东北林业大学学报*, 2014, **42**(7): 154–156, 165.
DONG Yunfei, SUN Yujun, WANG Yifu, et al. Generalized height-diameter model for Chinese fir based on BP neural

- network [J]. *J Northeast For Univ*, 2014, **42**(7): 154 – 156, 165.
- [10] 沈剑波, 雷相东, 李玉堂, 等. 基于 BP 神经网络的长白落叶松人工林林分平均高预测[J]. 南京林业大学学报 (自然科学版), 2018, **42**(2): 147 – 154.
- SHEN Jianbo, LEI Xiangdong, LI Yutang, *et al.* Prediction mean height for *Larix olgensis* plantation based on Bayesian-regularization BP neural network [J]. *J Nanjing For Univ Nat Sci Ed*, 2018, **42**(2): 147 – 154.
- [11] 徐步强, 张秋良, 弥宏卓, 等. 基于 BP 神经网络的油松人工林生长模型[J]. 东北林业大学学报, 2011, **39**(12): 33 – 35.
- XU Buqiang, ZHANG Qinliang, MI Hongzhuo, *et al.* Growth model of *Pinus tabulaeformis* plantation based on BP neural network [J]. *J Northeast For Univ*, 2011, **39**(12): 33 – 35.
- [12] 金星姬, 贾炜玮, 李凤日. 基于 BP 人工神经网络的兴安落叶松天然林全林分生长模型的研究[J]. 植物研究, 2008, **28**(3): 370 – 374, 384.
- JIN Xingji, JIA Weiwei, LI Fengri. Whole stand growth model for natural *Dahurian larch* forests based on BP ANN [J]. *Bull Bot Res*, 2008, **28**(3): 370 – 374, 384.
- [13] 刘建忠, 余娜. 贵州省马尾松林地质量空间评价与低效林防控措施研究[J]. 湖北农业科学, 2016, **55**(20): 5202 – 5206.
- LIU Jianzhong, YU Na. Quality spatial evaluation of *Pinus massoniana* and control measures of low-efficiency forest in Guizhou Province [J]. *Hubei Agric Sci*, 2016, **55**(20): 5202 – 5206.
- [14] 何婧, 韦小丽, 徐海. 贵州省马尾松林分树种多样性分析[J]. 西部林业科学, 2016, **45**(3): 26 – 29, 37.
- HE Jing, WEI Xiaoli, XU Hai. Tree species diversity of *Pinus massoniana* forest in Guizhou Province [J]. *J West China For Sci*, 2016, **45**(3): 26 – 29, 37.
- [15] 王科, 谭伟, 戚玉娇. 近自然经营间伐对黔中马尾松天然次生纯林生长的初期效应[J]. 浙江农林大学学报, 2019, **36**(5): 886 – 893.
- WANG Ke, TAN Wei, QI Yujiao. Initial effects of close-to-nature thinning on a natural secondary *Pinus massoniana* pure forest in central Guizhou [J]. *J Zhejiang A&F Univ*, 2019, **36**(5): 886 – 893.
- [16] 浦瑞良, 宫鹏, YANG R. 应用神经网络和多元回归技术预测森林产量[J]. 应用生态学报, 1999, **10**(2): 129 – 134.
- PU Ruiliang, GONG Peng, YANG R. Forest yield prediction with an artificial neural network and multiple regression [J]. *Chin J Appl Ecol*, 1999, **10**(2): 129 – 134.
- [17] PESCHEL W. Mathematical methods for growth studies of trees and forest stands and the results of their application [J]. *Tharandter Forstl Jahrb*, 1938, **89**: 169 – 247.
- [18] CURTIS R O. Height-diameter and height-diameter age equations for second-growth Douglas fir [J]. *For Sci*, 1967, **13**(4): 365 – 375.
- [19] PEARL R, REED L J. On the rate of growth of the population of the United States since 1790 and its mathematical representation [J]. *Proc Nati Acad Sci*, 1920, **6**(6): 275 – 288.
- [20] WEIBULL W. A statistical distribution function of wide applicability [J]. *J Appl Mech*, 1951, **18**(3): 293 – 297.
- [21] GOMPERZ B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies [J]. *Phil Trans Roy Soc London*, 1832, **115**: 513 – 585.
- [22] LUNDQVIST B. On the height growth in cultivated stands of pine and spruce in Northern Sweden [J]. *Medd Fran Statens Skogsforskningsinst*, 1957, **47**(2): 1 – 64.
- [23] 张立明. 人工神经网络的模型及其应用[M]. 上海: 复旦大学出版社, 1993: 43 – 46.
- [24] 罗恒春, 张超, 魏安超, 等. 云南松林分平均胸径生长模型及模型参数环境解释[J]. 浙江农林大学学报, 2018, **35**(6): 1079 – 1087.
- LUO Hengchun, ZHANG Chao, WEI Anchao, *et al.* Average DBH growth model of a stand with environmental parameters for *Pinus yunnanensis* in central Yunnan, China [J]. *J Zhejiang A&F Univ*, 2018, **35**(6): 1079 – 1087.
- [25] HUANG S S, TITUS S J, WIENS D P. Comparison of nonlinear height-diameter functions for major Alberta tree species [J]. *Can J For Res*, 1992, **22**(9): 1297 – 1304.
- [26] GUPTA A, SRIVASTAVA R, JABEEN N, *et al.* Height estimation model for *Eucalyptus tereticornis* grown under semiarid conditions of India [J]. *Indian J Agrofor*, 2008, **10**(2): 34 – 39.
- [27] KALBI S, FALLAH A, BETTINGER P, *et al.* Mixed-effects modeling for tree height prediction models of Oriental beech in the Hyrcanian forests [J]. *J For Res*, 2018, **29**(5): 1195 – 1204.

- [28] SHARMA R P, VACEK Z, VACEK S, *et al.* Modelling individual tree height-diameter relationships for multi-layered and multi-species forests in central Europe [J]. *Trees*, 2019, **33**(1): 103 – 119.
- [29] 黄旭光, 周俊朝, 黄柏华, 等. 基于人工神经网络对栎树天然林地位指数模拟系统的研究[J]. 河南农业大学学报, 2015, **49**(2): 190 – 194.
HUANG Xuguang, ZHOU Junchao, HUANG Bohua, *et al.* Study of oak growth dynamic simulation system based on artificial neural network [J]. *J Henan Agric Univ*, 2015, **49**(2): 190 – 194.
- [30] 徐志扬. 基于 BP 神经网络的马尾松树高曲线模型[J]. 林业调查规划, 2015, **40**(2): 6 – 8, 73.
XU Zhiyang. Height-diameter model for *Pinus massoniana* based on BP neural network [J]. *For Inventory Plann*, 2015, **40**(2): 6 – 8, 73.
- [31] 杜志, 甘世书. 基于 BP 神经网络的杉木和马尾松树高曲线模型研究[J]. 中南林业调查规划, 2017, **36**(4): 36 – 39.
DU Zhi, GAN Shishu. Height-diameter models for *Cunninghamia lanceolata* and *Pinus massoniana* based on BP neural network [J]. *Cent South For Inventory Plann*, 2017, **36**(4): 36 – 39.
- [32] 王轶夫, 孙玉军, 郭孝玉. 基于 BP 神经网络的马尾松立木生物量模型研究[J]. 北京林业大学学报, 2013, **35**(2): 17 – 21.
WANG Yifu, SUN Yujun, GUO Xiaoyu. Single-tree biomass modeling of *Pinus massoniana* based on BP neural network [J]. *J Beijing For Univ*, 2013, **35**(2): 17 – 21.
- [33] AHMADI K, ALAVI S J, TABARI M, *et al.* Non-linear height-diameter models for oriental beech (*Fagus orientalis* Lipsky) in the Hyrcanian forests, Iran [J]. *Biotechnol Agron Soc Environ*, 2013, **17**(3): 431 – 40.
- [34] 苏崇宇, 汪毓铎. 基于改进的自适应遗传算法优化 BP 神经网络[J]. 工业控制计算机, 2019, **32**(1): 67 – 69.
SU Chongyu, WANG Yuduo. BP neural network optimized by improved adaptive genetic algorithm computer engineering and applications [J]. *Ind Control Comput*, 2019, **32**(1): 67 – 69.
- [35] 李松, 刘力军, 翟曼. 改进粒子群算法优化 BP 神经网络的短时交通流预测[J]. 系统工程理论与实践, 2012, **32**(9): 2045 – 2049.
LI Song, LIU Lijun, ZHAI Man. Prediction for short-term traffic flow based on modified PSO optimized BP neural network [J]. *Syst Eng-Theory Pract*, 2012, **32**(9): 2045 – 2049.