

DOI:10.13275/j.cnki.lykxyj.2022.03.022

基于 BP 神经网络的沙柳地上生物量预测模型

程冀文¹, 王树森^{1*}, 罗于洋¹, 张 岑^{1,2}

(1. 内蒙古农业大学沙漠治理学院, 荒漠生态系统保护与修复国家林业和草原局重点实验室, 内蒙古 呼和浩特 010018;

2. 清水河县老牛湾镇人民政府 内蒙古 呼和浩特 011614)

摘要: [目的] 以库布其沙漠沙柳为研究对象, 建立基于 BP 神经网络的沙柳生物量模型, 探究不同建模因子下的沙柳生物量估算模型变化, 以期探究沙柳生物量估算模型的最优形式。[方法] 选取 6 种沙柳生长因子, 并根据与生物量相关性大小加入输入变量, 从而组成 6 组不同输入变量, 输入变量包含因子数量逐步增加 (1 ~ 6 种)。对比 BP 神经网络沙柳生物量模型不同输入变量所拟合模型的性能, 确定最佳输入变量, 并在最优输入变量的基础上, 确定 BP 神经网络隐层数量, 经过反复训练, 建立基于 BP 神经网络的沙柳生物量估算模型。[结果] 基于 BP 神经网络的沙柳生物量模型最优结构, 即输入层节点数 (N_{in}): 隐层节点数 (N_h): 输出层节点数 (N_{out}) 为: 4 : 9 : 1。其中训练数据 $R^2=0.97$, $RMSE=0.67$, $MAE=0.50$; 测试数据 $R^2=0.96$, $RMSE=1.10$, $MAE=0.77$ 。[结论] 基于 BP 神经网络的沙柳生物量, 随着输入变量中输入因子的数量不断增加, 发现其 R^2 、 $RMSE$ 、 MAE 所表现出的模型性能逐渐变好, 但是输入变量每增加 1 种后, 当输入因子数量为 5 时, 模型精度相比输入因子数量为 4 时提升幅度较小, 考虑模型使用时的精度和野外工作的便利性, 输入层最优输入因子数为 4 种, 当隐层数为 9 时模型性能表现为最优。

关键词: 沙柳; 生物量估算模型; BP 神经网络; 库布其沙漠

中图分类号: S718.5

文献标志码: A

文章编号: 1001-1498(2022)03-0193-06

沙柳 (*Salix psammophila* C.Wang et Ch.Y. Yang) 又名北沙柳, 在库布齐沙漠有广泛分布, 常作固沙造林树种^[1-2], 也是我国北部地区防风固沙的优良树种和当地典型的能源树种^[2-3]。通常情况下, 沙柳在沙丘中下部有少量分布, 在滩地沙区大量密集分布^[4]。沙柳具备生长速率较快的特点, 并且人工栽培沙柳在短时间内就能产生体积很大的灌丛, 因此能够在水土保持与荒漠化防治、水源涵养、气候调节等生态效益方面起到重要的作用^[5]。由于沙柳木质优良, 还可以广泛用于造纸、人造板等工业用途^[6]。

在干旱半干旱风沙区, 沙生灌木有着防风固沙、水土保持等作用, 通常野外调查使用的直接收

割法会对地表造成破坏从而发生土壤侵蚀现象^[7]。因此探究灌木生物量与易测因子间的关系, 建立沙生灌木生物量模型对沙生灌木生物量研究有很大的助力。目前, 国内外学者在生物量模型方面有大量的研究^[8-10], 通常选取林木因子作为自变量构建灌木生物量模型, 但是由于各区域之间自然环境条件的巨大差别, 以及不同气候区灌木植被的形态特性等具有很大不同, 因此很多地区灌木的生物量预测模型仅在当地有较好的解释力。因此对于灌木构建适应其生境的生物量预测模型是十分必要的^[11]。

近年来, 许多学者将 BP 神经网络应用在林业上^[12-14]。研究发现, BP 神经网络模型精度相较于传统更好^[15-19]。但是目前的研究大多集中在高大乔

收稿日期: 2021-09-28 修回日期: 2021-12-28

基金项目: 内蒙古自治区应用技术与开发项目“内蒙古地区多功能树种选育及扩繁关键技术与示范”(2019GG004)

* 通讯作者: 王树森, 博士, 博士生导师, 研究方向为水土保持与荒漠化防治。E-mail: wsswt@126.com

木的生物量估算方面,对于运用 BP 神经网络进行灌木,尤其是沙生灌木的生物量估算却鲜有报道。因此,利用沙柳的不同林木因子建立 BP 神经网络生物量估算模型,探究 BP 神经网络在沙生灌木生物量方面的应用是必要的。本研究以库布其沙漠沙柳为研究对象,建立了沙柳 BP 神经网络的生物量预测模型,探究不同林木因子对沙柳 BP 神经网络生物量预测模型的影响,以期对沙柳以及其他沙生灌木生物量建模与估测工作提供参考依据。

1 材料与方法

1.1 研究区概况

研究区域地处库布齐沙漠,处于鄂尔多斯市杭锦旗独贵塔拉镇境内,地理坐标为 $108^{\circ}15' \sim 108^{\circ}24' E$, $40^{\circ}48' \sim 40^{\circ}39' N$,属于典型温带大陆性干旱、半干旱季风气候。年均降水量约 400 mm^[20]。其范围内木本野生植物主要有小叶杨

(*Populus simonii* Carr),沙枣(*Elaeagnus angustifolia* Linn),沙柳(*Salix psammophila* C.Wang et Ch.Y. Yang)等^[6]。

1.2 研究方法

1.2.1 数据采集 外业数据采集时间为 2017 年 4 月。在研究地选择迎风坡风蚀、丘间低地沙埋和未沙埋 3 种立地条件,共选择 91 株沙柳。根据沙柳的形态学特性,选取用于 BP 神经网络生物量预测模型建模的因子有丛高(H)、地径(D)、冠幅(C)、基部分支数(N_t)、单枝最大分支数(N_b)、单枝最大分枝次数(N_{bt})、枝条粗度(D_b)、芽眼数(N_e)。采用皮尺和游标卡尺对以上林木因子进行测量和记录(表 1)。记录工作结束后,将所选取的沙柳齐地面刈割,使用电子天平现场称取鲜质量。将现场刈割的沙柳枝条带回实验室使用烘箱 $70^{\circ}C$ 烘干 24 h 至恒质量,使用天平称取干质量,计算沙柳地上部分生物量。

表 1 建模数据统计

Table 1 Summary statistics for model modeling and validation

	丛高(H) Height/m	地径(D) Ground diameter/m	冠幅(C) Crown width/m	基部分支数 (N_t) Number of basal branches	枝条粗度 (D_b) Branch diameter/cm	芽眼数(N_e) Number of eyes	单枝最大分 支数(N_b) Maximum number of branches	单枝最大分 枝次数(N_{bt}) Maximum branch level	生物量(B) Biomass/kg
平均值	2.97	0.54	2.92	45.62	1.33	67.07	16.91	4.16	4.65
最大值	4.45	1.89	4.93	173.00	2.90	115.00	51.00	6.00	33.92
最小值	0.96	0.04	1.43	3.00	0.61	21.00	4.00	2.00	0.05
标准差	0.53	0.37	0.89	36.34	0.42	20.67	6.92	0.91	5.69

1.2.2 BP 神经网络 BP 神经网络由包括输入层、隐层和输出层共 3 部分组成(图 1),BP 神经网络中神经元结构如图 2 所示, $x_1 \sim x_n$ (1, 2, 3, \dots , n) 为输入信号, $\omega_1 \sim \omega_n$ (1, 2, 3, \dots , n) 为权重, b 为偏置, Σ 为求和节点, σ 为激活函数。按

照 Kolmog-orov 定理,单隐层的 BP 神经网络将能够无限接近任意连续的非线性曲线^[21]。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。

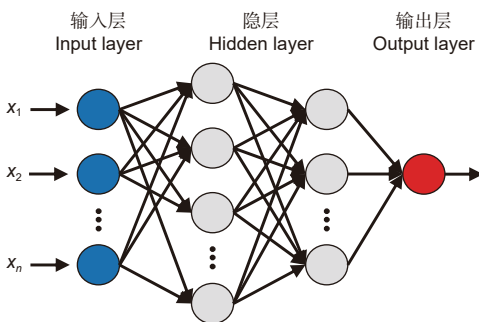


图 1 BP 神经网络结构

Fig. 1 Structure of BP neural network

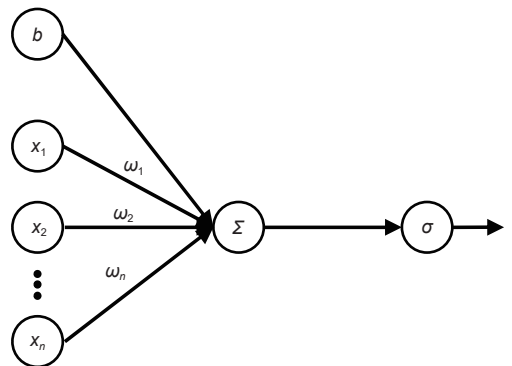


图 2 神经元结构

Fig. 2 Neuron structure

1.2.3 BP神经网络模型的建立 构建沙柳BP神经网络生物量预测模型使用Matlab软件中的nntool工具箱,以沙柳林木因子作为输入层,沙柳生物量实测值为训练目标。 $N_h = \sqrt{N_{in} + N_{out}} + h$ 为确定BP神经网络的隐层节点数的经验公式,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,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,隐层传递函数为logsig,输出层传递函数为purelin,使用后的模型训练算法为L-M算法。

1.2.4 归一化处理与反归一 对于参与建模的数据采取归一化处理,可以让神经网络的训练效率有所提升,并且在拟合后还必须对输出数据采取反归一处理,使输出数据重新转换回沙柳的生物量的输出值。据此,本研究通过下列公式对输入因子数据和输入目标数据进行归一化处理:

$$Y = Y_{\min} + \frac{Y_{\max} - Y_{\min}}{X_{\max} - X_{\min}}(X - X_{\min})$$

式中: X 为未进行处理的值; X_{\max} 、 X_{\min} 为代表未进行处理数据组中的最大和最小值; Y 为进行处理后的值; Y_{\max} 和 Y_{\min} 分别为进行处理后的数据组中的最大和最小值。

当BP神经网络拟合结束后,对所得数据进行反归一处理,也就是将程序运行得到的数据映射到实际预测值。

1.2.5 模型评价与检验 采用决定系数 R^2 ,均方根误差(RMSE)和平均绝对误差(MAE)对模型性能进行评价。

2 结果与分析

2.1 模型输入因子的筛选

前期可供作为输入因子的备选指标有丛高(H)、地径(D)、冠幅(C)、基部分支数(N_t)、单枝最大分支数(N_b)、单枝最大分枝次数(N_{bt})、枝条粗度(D_b)、芽眼数(N_e),对各因子与生物量(B)之间进行相关性分析,结果如图3所示。

其中各因子与生物量相关性表现为:冠幅(0.751) > 基部分支数(0.677) > 地径(0.631) > 枝条粗度(0.589) > 芽眼数(0.555) > 丛高(0.550) > 单枝最大分枝次数(0.179) > 单枝最大分枝数(0.132),其中丛

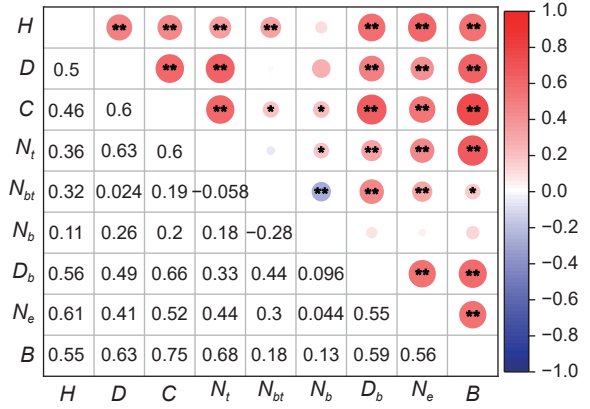


图3 沙柳各因子相关性分析

Fig. 3 Relationship of factors in *Salix psammophila*

高、地径、冠幅、基部分支数、芽眼数、枝条粗度与生物量间表现为极显著相关($P < 0.01$);单枝最大分枝次数与生物量间表现为显著相关($P < 0.05$);单枝最大分枝数与生物量之间显著相关。由此选取丛高、地径、冠幅、基部分支数、芽眼数、枝条粗度为模型最终输入因子。对于传统生物量模型而言,将丛高、地径、冠幅、基部分支数、芽眼数、枝条粗度分别作为自变量拟合生物量模型。对于BP神经网络模型而言,根据其相关性高低逐个加入模型输入层,确定输入层为输入因子数为1~6。

2.2 BP神经网络模型的建立

根据相关性大小,共设置6组输入变量 N_{in} 分别为1~6。输出变量 N_{out} 为生物量。由于输出值都为1,输入变量为1~6,则根据经验公式 $N_h = \sqrt{N_{in} + N_{out}} + h$ 可得,当输入因子数为1时,BP神经网络隐层数 N_h 的取值范围为2.41~11.41之间;当输入因子数为2~5时,BP神经网络隐层数 N_h 的取值范围最小值为2.73~3.45,隐层数 N_h 的取值范围最大值为11.73~12.45;当输入因子数为6时,BP神经网络隐层数 N_h 的取值范围为3.65~12.65。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,则将隐层数 N_h 选取为整数,当输入因子数为1时,BP神经网络隐层数 N_h 的取值范围为2~11的整数;当输入因子数为2~5时,BP神经网络隐层数 N_h 的取值范围为3~12的整数;当输入因子数为6时,BP神经网络隐层数 N_h 的取值范围为4~13的整数。对不同输入因子数各10隐层的模型分别训练10次,不同输入变量拟合情况见表2。

表 2 BP 神经网络模型不同输入变量拟合优势度评价
Table 2 Evaluation of fitting dominance of different input variables in BP neural network model

输入因子 Input factor	决定系数均值 Coefficient of determination R^2	均方根误差均值 Root mean square error $RMSE$	平均绝对误差均值 Mean absolute error MAE
C	0.70	2.97	1.90
C、 N_t	0.75	2.10	1.36
C、 N_t 、D	0.86	1.99	1.30
C、 N_t 、D、 D_b	0.89	1.37	0.90
C、 N_t 、D、 D_b 、 N_e	0.90	1.31	0.82
C、 N_t 、D、 D_b 、 N_e 、H	0.91	1.24	0.78

说明：H：丛高（Height）；D：地径（Ground diameter）；C：冠幅（Crown width）； N_t ：基部分枝数（Number of basal branches）； D_b ：枝条粗度（Branch diameter）； N_e ：芽眼数（Number of eyes）； N_b ：单枝最大分支数（Maximum number of branches）； N_{bt} ：单枝最大分枝次数（Maximum branch level）；B：生物量（Biomass）。

可以看出，模型的精度随着输入因子个数的增加而增加（图 4），当输入因子数量为 5 时，模型精度相比输入因子数量为 4 时提升不明显，为了平衡模型的易用性和精度，则选择 4 输入因子为模型的最优输入变量数。

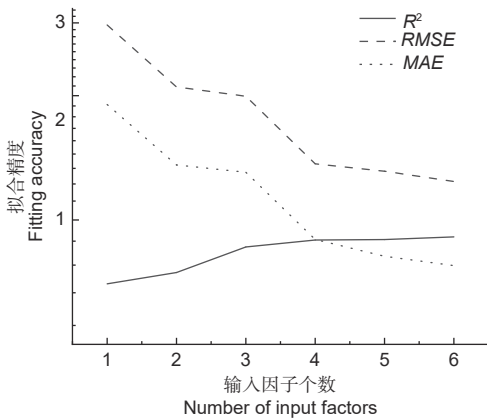


图 4 随着输入因子增加模型性能变化

Fig. 4 Fitting accuracy with different number of input factors

2.3 BP 神经网络最优模型的确定

当输入因子数为 4 时 BP 神经网络隐层数 N_h 的取值范围为 3~12 的整数，不同节点数的模型拟合情况（见表 3），其中当隐层数为 9 时，其 R^2 在各隐层数中表现为最大， $RMSE$ 和 MAE 表现为最小，由此得知，当隐层数为 9 时，模型拟合效果最好。

由此，最佳输入变量和最佳隐层数确定后，在

表 3 输入因子数为 4 时 BP 神经网络模型拟合优势度评价
Table 3 Evaluation of fitting dominance of BP neural network model when the N_{in} factors is 4

节点数 Nodes	训练数据 Training			测试数据 Test		
	R^2	$RMSE$	MAE	R^2	$RMSE$	MAE
3	0.802	2.035	1.272	0.920	1.431	0.853
4	0.891	1.534	1.090	0.879	1.506	0.897
5	0.893	1.322	0.870	0.701	2.646	1.776
6	0.894	1.518	0.882	0.784	2.579	1.315
7	0.951	0.859	0.600	0.831	2.068	1.014
8	0.872	1.639	1.101	0.800	2.488	1.192
9	0.968	0.666	0.503	0.961	1.102	0.768
10	0.833	1.933	1.185	0.858	1.848	1.001
11	0.900	1.288	0.858	0.929	1.385	0.812
12	0.938	0.883	0.613	0.937	1.304	0.809

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为最后的模型，其中：最优结构 [输入层节点数 (N_{in})：隐层节点数 (N_h)：输出层节点数 (N_{out})] = 4 : 9 : 1。经过训练所得传递函数模型为：

$$H = \text{purelin} (1.506 1h_1 + 1.151 6h_2 + 0.226 53h_3 - 1.447 2h_4 - 0.234 79h_5 + 1.225 4h_6 - 1.263 5h_7 - 0.155 27h_8 - 2.062 6h_9 - 0.853 28);$$

$$h_1 = \text{logsig} (0.844 89C + 1.865 2N_t + 3.207 6D - 4.063 2D_b + 6.167 7);$$

$$h_2 = \text{logsig} (6.942C + 0.381 08N_t + 4.981 5D - 1.046 5D_b - 5.647 2);$$

$$h_3 = \text{logsig} (1.635C - 2.413 6N_t - 0.163 18D - 4.040 7D_b - 2.106 9);$$

$$h_4 = \text{logsig} (0.940 89C + 2.839 6N_t - 6.912 5D - 1.823 8D_b - 1.757 9);$$

$$h_5 = \text{logsig} (1.224C - 2.715 7N_t - 2.839 4D - 4.562 5D_b - 2.169);$$

$$h_6 = \text{logsig} (1.193 2C + 3.811 7N_t - 7.989 1D - 0.911 14D_b - 0.853 18);$$

$$h_7 = \text{logsig} (1.955 4C - 2.838 1N_t + 2.935 2D - 1.89D_b + 3.127 6);$$

$$h_8 = \text{logsig} (-3.880 3C + 0.570 5N_t - 2.167 3D - 2.634 9D_b - 3.617 8);$$

$$h_9 = \text{logsig} (-3.781 5C + 2.025 1N_t + 2.536 2D + 0.408 18D_b - 4.022 4);$$

3 讨论

灌木生物量模型应选择所研究灌木的易测因子作为模型构建的自变量,往往不同形态特征的灌木,宜采用不同的变量构建生物量模型,所以应根据所研究灌木的形态特征来选取最优灌木易测因子作为模型构建的自变量^[11]。由于BP神经网络模型的自身特性,增加输入变量的数量,可以使得模型的泛化能力提高,增加模型精度^[22-23]。对于本研究来讲,模型输入变量尽可能的囊括沙柳所有可以测量的形态指标,并以沙柳各因子与沙柳生物量的相关关系为基础进行筛选,在最优模型中输入因子包括冠幅 C 、基部分支数 N_t 、地径 D 、枝条粗度 D_b 共4种。输入变量中输入因子数量的增加可以使得模型精度不断上升。

在风沙区,由于风大沙多,灌木基部容易产生沙埋的现象,使得基径的测量比较困难,考虑到模型在日常生产中的易用性和实用性,在构建生物量模型时要结合实际工作考量不同的建模因子^[24]。本研究中地径与生物量相关性为0.631。如果使用相关性不如地径但更为简单易测的两个或多个因子进行建模,是否在模型性能上可以替代相对不易测的地径参与建模的模型呢?在未来的沙区灌木生物量模型研究中,应将实际工作中灌木因子的测量难易程度加入建模输入因子筛选的条件中,对比低相关但相对易测因子与高相关性但测量困难的因子对生物量估算模型性能的影响。

本研究选择6种BP神经网络的输入因子均为沙柳的测树因子,并未包括环境因子,将林分因子和气候因子等变量共同参与生物量预测模型的拟合,对于模型的性能有较大的提高^[17]。坡向、坡位、立地条件、海拔和林分密度等环境因素对林木的生长发育也有很大的影响,将以上因子加入到模型的拟合会有效提高模型精度^[25]。本研究中尚未将丘间低地未沙埋、丘间低地沙埋、迎风坡风蚀3种立地条件考虑到模型的建模当中,接下来的研究中要探究不同立地条件对的沙柳BP神经网络生物量预测模型的影响。

4 结论

本研究以库布其沙漠沙柳为研究对象,为其分别建立传统沙柳生物量模型和基于BP神经网络的沙柳生物量模型,探究不同建模因子下的沙柳生物

量估算模型变化。根据沙柳生长因子相关性大小,输入变量所包含的输入因子由1种逐渐增加至最大6种,拟合发现,随着输入变量中输入因子的数量不断增加,输入变量达到6种时 R^2 (0.91)、 $RMSE$ (1.24)、 MAE (0.78)所表现出的模型性能最好,输入变量每增加1种后,模型性能提升幅度不尽相同,当输入因子数量为5时,模型精度相比输入因子数量为4时提升幅度较小,考虑到在模型使用时要同时兼顾其使用精度和野外实地调查时的工作强度和调查量大小,故选择输入变量中,输入因子为4种时为最佳,则其输入层最优输入因子数为4种。在确定模型结构输入为4,输出为1的前提下,对模型各隐层数不断训练,当隐层数为9时,模型性能表现为最优,则基于BP神经网络的沙柳生物量估算模型最优结构[输入层节点数(N_{in}):隐层节点数(N_h):输出层节点数(N_{out})] = 4:9:1。

参考文献:

- [1] 米志英,高永,白存德,等.库布齐沙漠沙柳无性快速培育技术对比研究[J].中国沙漠,2008(2):318-321.
- [2] 陈国鹏,赵文智.毛乌素沙地南缘沙柳(*Salix psammophila*)丛生枝年龄结构与动态特征[J].中国沙漠,2015,35(6):1520-1526.
- [3] Li X, Huang Y, Gong J, et al. A study of the development of bio-energy resources and the status of eco-society in China[J]. Energy, 2010, 35(11): 4451-4456.
- [4] 米志英.库布齐沙漠沙柳培育关键技术研究[D].呼和浩特:内蒙古农业大学,2008.
- [5] 刘志刚.典型沙生灌木平茬技术与机具研究[D].北京:中国农业大学,2013.
- [6] 张岑.人工平茬沙柳固碳能力的研究[D].呼和浩特:内蒙古农业大学,2018.
- [7] 杨昊天,李新荣,王增如,等.腾格里沙漠东南缘4种灌木的生物量预测模型[J].中国沙漠,2013,33(6):1699-1704.
- [8] 赵蓓,郭泉水,牛树奎,等.大岗山林区几种常见灌木生物量估算与分析[J].东北林业大学学报,2012,40(9):28-33.
- [9] 季蕾,亢新刚,张青,等.吉林金沟岭林场不同密度天然云冷杉林下主要灌木生物量模型[J].浙江农林大学学报,2016,33(03):394-402.
- [10] 王文栋,白志强,阿里木·买买提,等.天山林区6种优势种灌木林生物量比较及估测模型[J].生态学报,2016,36(9):2695-2704.
- [11] 童新风,杨红玲,宁志英,等.科尔沁沙地优势固沙灌木的生物量预测模型[J].中国沙漠,2018,38(3):553-559.
- [12] 王立海,邢艳秋.基于人工神经网络的天然林生物量遥感估测[J].应用生态学报,2008,19(2):261-266.

- [13] 黄家荣, 孟宪宇, 关毓秀. 马尾松人工林直径分布神经网络模型研究[J]. 北京林业大学学报, 2006, 28 (1): 28-31.
- [14] 车少辉, 张建国, 段爱国, 等. 杉木人工林胸径生长神经网络建模研究[J]. 西北农林科技大学学报:自然科学版, 2012, 40 (3): 84-92.
- [15] 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.
- [16] 卯光宪, 谭伟, 柴宗政, 等. 基于BP神经网络的马尾松人工林胸径-树高模型预测[J]. 浙江农林大学学报, 2020, 37 (4): 752-760.
- [17] 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.
- [18] 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.
- [19] 陈佳琦, 赵鹏祥, 祁宁, 等. 基于BP神经网络的油松人工林树高模型研究[J]. 西北林学院学报, 2020, 35 (1): 212-217.
- [20] 唐宽燕. 库布齐沙漠东缘植被动态与环境的关系[D]. 北京: 中国农业科学院, 2017.
- [21] Tianyun, Jiazhiqing, Yangxiaohui. Improving shrub biomass estimations in the Qinghai-Tibet Plateau: Age-based Caragana intermedia allometric models[J]. NRC Research Press, 2014, 90(2): 154-160.
- [22] 曾伟生, 唐守正. 利用度量误差模型方法建立相容性立木生物量方程系统[J]. 林业科学研究, 2010, 23 (6): 797-803.
- [23] 梁瑞婷, 王轶夫, 邱思玉, 等. 人工神经网络与相容性生物量模型预测单木地上生物量的比较[J/OL]. 应用生态学报: 1-10. [2021-10-01]. <https://doi.org/10.13287/j.1001-9332.202201.001>.

Aboveground Biomass Model of *Salix psammophila* Based on BP Neural Network

CHENG Ji-wen¹, WANG Shu-sen¹, LUO Yu-yang¹, ZHANG Cen^{1,2}

(1. Desert Control Science and Engineering, Key Laboratory of National Forestry and Grassland Administration on Desert Ecosystem Conservation and Restoration, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia, China;
2. People's Government of Laoniawan Town, Qingshuihe County, Hohhot 011614, Inner Mongolia, China)

Abstract: [Objective] To accurately predict the aboveground biomass of *Salix psammophila* in Kubuqi Desert, the aboveground biomass models including different predictors were developed using BP neural network. [Methods] Six independent variables were selected and added to the biomass model according to their correlation coefficients. The number of input variables gradually increased from 1 to 6. The best BP biomass model with the optimal input variables was selected based on the model performance. [Results] The best structure of *S. psammophila* BP biomass model was that the number of input layer nodes (N_{in}), hidden layer nodes (N_h) and output layer nodes (N_{out}) were 4 : 9 : 1, respectively. R^2 , RMSE, and MAE output from training data was 0.97, 0.68, 0.50, respectively; and 0.96, 1.16, 0.78, respectively from test data. [Conclusion] The performance of *S. psammophila* aboveground biomass model based on BP neural network became better with the increasing number of input variables. However, when the number of input variables was 5, the improvement of model performance was slightly better than the model with 4 input variables. Considering the model accuracy and model application, the optimal number of input variables in the input layer is 4, and the model performance is the best when the number of hidden layers is 9.

Keywords: *Salix psammophila*; biomass model; BP neural network; Kubuqi desert

(责任编辑: 彭南轩)