

# 荒漠草原植物群落现存量的 BP 神经网络模拟与预测研究

张 爽<sup>1</sup>, 卫智军<sup>1\*</sup>, 吕世杰<sup>2</sup>, 刘文静<sup>1</sup>, 白玉婷<sup>1</sup>

(1. 内蒙古农业大学生态环境学院, 内蒙古 呼和浩特 010019;

2. 内蒙古农业大学理学院, 内蒙古 呼和浩特 010018)

[摘要] 本研究以内蒙古锡林郭勒盟苏尼特右旗短花针茅荒漠草原为研究对象, 采用分层取样技术, 分别对碱韭和短花针茅为主的草地植物群落进行调查(测定指标为高度、盖度、密度和地上现存量), 获得禾本科为主、碱韭为主以及由二者构成的复合样本数据。通过 MATLAB 软件平台, 把植物群落的数量特征(高度、盖度和密度)作为输入因子, 以群落现存量作为输出因子来建立 BP 神经网络模型, 并对预测结果进行检验。结果显示, 禾本科为主、碱韭为主以及由二者构成的复合样本, 采用 BP 神经网络平均预测准确率分别为 94.1%, 92.98%, 91.01%。因此, BP 神经网络可作为草地植物群落地上现存量模拟与预测的有效工具之一。采用 BP 神经网络对不同草地植物类群进行模拟和预测可能会存在差异, 但这种差异会随着样本容量的增大或训练精度的增加而弱化。

[关键词] 荒漠草原; 现存量; BP 神经网络; MATLAB; 模拟与预测

[中图分类号] Q145+.2; Q332 [文献标识码] A [文章编号] 2095—5952(2016)01—0053—07

## The research of simulation and prediction on plant community biomass in desert steppe based on BP neural network

ZHANG Shuang<sup>1</sup>, Lü Shijie<sup>2\*</sup>, WEI Zhijun<sup>1</sup>, LIU Wenting<sup>1</sup>, BAI Yuting<sup>1</sup>

(1. College of Ecology and Environmental Science, Inner Mongolia Agricultural University  
Huhhot 010019, China,

2. College of Science and Environmental Science, Inner Mongolia Agricultural University  
Huhhot 010019, China,)

**Abstract:** The grassland biomass of desert steppe is a very important research field of prairie research, it is not only resolve accumulation of the dry matter of desert steppe ecosystem, but also resolve its biomass so that effect the function of ecology. This paper take plant community of poaceae and allium polyrhizum and its compound plant community in desert steppes as the main research object and the data was come from an investigation which held in Inner Mongolia Xilinguole. According to the quantitative features (natural height, relative coverage degree, relative density) of plant community in desert steppe and the practical data of biomass, we use MATLAB established a BP neural network model. We took quantitative features of plant community as input factors, and took corresponding practical biomass as output factor, so that we constructed a three layers BP neural network and simulated the biomass of plant community. The result shows that the standing crop of poaceae, allium polyrhizum and mixed plant communities dominated the average prediction accuracy rate belong to a good level. This model could predict plant community biomass and it has higher accuracy.

**Key words:** desert steppe; standing crop; BP neural network; MATLAB; simulation and prediction

随着近代科学技术的发展, 人类对于草地的  
传统价值观念正在经历着非常深刻的变化, 有些国

家和地区已经开始把单纯的草地收获利用转向为  
最大限度的发挥草地高效性能和维持生态平衡的

发展方向<sup>[1-3]</sup>。只有当草地的计量以重量或能量来表示时,才有可能对草地生态系统的综合效益和可更新资源的潜力做出非常客观的定量估计和对草地生态系统的稳定与平衡做出正确的定量判断。因此,草地植物现存量的研究是草地生态学研究的一个必不可少的内容,也对阐述草地生态学研究中的问题具有重要的意义。传统的草地植被地上生物量的测定方法有直接收割法和水热法等,其中直接收割法是比较有效且准确的测定方法。但采用直接收割法测定地上生物量会耗费大量的人力、物力和财力并且在地域上有一定的局限性。故此,关于草地现存量模拟与预测的研究则显得尤为重要。

目前 BP 神经网络被大量地应用于工学、天文学、生物学、生态学等领域。其优势在于具有很强的学习能力和记忆力并且具有较高的构建模型能力和良好的数据拟合能力,在建模时不需要给出具体的数学函数,可以一次性引入多个解释变量,并同时输出多个估测量。黄家荣<sup>[8]</sup>等将 BP 神经网络建模技术应用于多种林分生长模型。王立海<sup>[9]</sup>等应用 BP 神经网络建立了森林生物量非线性遥感模型系统,绘制了吉林省汪清林业局经营区森林生物量分布图。在生物量的估测上,BP 神经网络模型表现出了其独特的优越性能。

因此,本文构建 BP 神经网络<sup>[10]</sup>模型,对荒漠草原<sup>[11,12]</sup>草地地上现存量进行模拟与预测研究,拟探讨其在草地现存量预测上的可行性,解决草地描述样方数据(高度、盖度和密度)与产量样方(地上现存量)数据不能逐一对应问题,为草地现存量建模与预测工作提供参考依据,同时也为草地动态监测提供节省物力和人力提供数据支撑。

## 1 材料与方法

### 1.1 试验区概况

本研究在 1999 年建立的内蒙古农业大学苏尼特右旗教学科研基地上进行,行政区划属内蒙古锡林郭勒盟苏尼特右旗朱日和镇,地理位置为 E112°47'11.2",N42°15'48.7"。试验区植被以亚洲中部荒漠草原种占主导地位,植被草层低矮,一般高度为 10~35cm;植被稀疏,盖度为 10~50%。短花

针茅(*Stipabreviflora*)为建群种,它在群落外貌和群落环境方面起到决定作用。优势种为无芒隐子草(*Cleistogenessongorica*)和碱韭(*Allium polyrhizum*)。主要伴生种有细叶韭(*Allium tenuissimum*)、银灰旋花(*Convolvulus ammannii*)等。该地区地形平坦,有明显的钙积层,主要分布在 10~35cm。土壤为淡栗钙土,腐殖质层厚 5~10cm。年平均降水量 177.2mm,且 60%~80%的降水集中在牧草生长旺季的 7 月到 9 月。

### 1.2 数据采集及预处理

#### 1.2.1 数据的采集

分别在禾本科为主和碱韭为主的植物群落试验区内,随机选取 10 个 1m×1m 的样方,观测样方内植物群落数量特征(盖度、高度和密度)及其地上现存量。

在群落样方内,地上现存量采用其地面剪割法收集,然后将野外得到的样品带回实验室洗净,并将其置于 60℃烘箱内烘干至恒重(约 48h),获得植物群落的地上现存量<sup>[16]</sup>。

因此,数据可分为 3 组,第一组为禾本科为主的草地植物群落样本数据,第二组为碱韭为主的草地植物群落样本数据,第三组为复合样本草地植物群落样本数据(即第一组和第二组整合在一起)。

#### 1.2.2 数据预处理<sup>[17]</sup>

因为所收集的样本数据各指标的单位不同,为了消除量纲的影响,本文利用以下公式对样本各指标进行归一化处理,即将输入、输出数据映射到[0, 1]范围内,训练结束后再反映射到原始数据范围。

$$x'_{ij} = \frac{x_{ij} - \{x_{ij}\}_{\min}}{\{x_{ij}\}_{\max} - \{x_{ij}\}_{\min}}$$

式中 $\{x_{ij}\}_{\min}$ 表示第 j 个实测指标中最小值;

$\{x_{ij}\}_{\max}$ 表示第 j 个实测指标中最大值;

### 1.3 研究方法

#### 1.3.1 模型的选择

根据植物群落数量特征(高度、盖度、密度)和现存量的实测数据,在随机抽取的情况下,选取总体的 70%为神经网络学习的训练样本,取余下的 30%作为神经网络学习的检验样本。基于 MATLAB(R2012b)软件平台,把植物群落数量

特征(高度、盖度、密度)作为输入因子,把相应实测群落的现存量作为输出因子,建立 BP 神经网络模型<sup>[14,15]</sup>对荒漠草原植物群落现存量进行拟合训练。

### 1.3.2 模型模拟结果检验

本文采用决定系数( $R^2$ )和均方误差(Mean Squared Error, MSE)进行模型训练结果的检验和模型精度评价。其中  $R^2$  和 MSE 的计算式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

式中  $y_i$  为现存量实测值,  $\hat{y}_i$  为模型预测值,  $\bar{y}$  为现存量实测值的均值,  $n$  为现存量实测值自由度。

### 1.3.3 模型预测结果检验

神经网络模型建立后,为检验模型的预测效果。取余下 30% 的样本数据对模型进行检验,采用相对误差(E)进行模型预测结果的检验,其计算式为:

$$E = \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\%$$

式中  $y_i$  为现存量实测值,  $\hat{y}_i$  为模型预测值。

## 2 结果与分析

### 2.1 禾本科为主的实验区

#### 2.1.1 网络的训练与预测

BP 神经网络的训练初始参数期望误差选取为 0.001,最大训练次数选为 10000,初始学习速率为 0.05。神经网络经随机初始化,利用函数 traingdm 对网络进行训练,当迭代次数达到 5146 步时,均方误差就达到了目标误差的要求。训练变化曲线如图 4 所示。

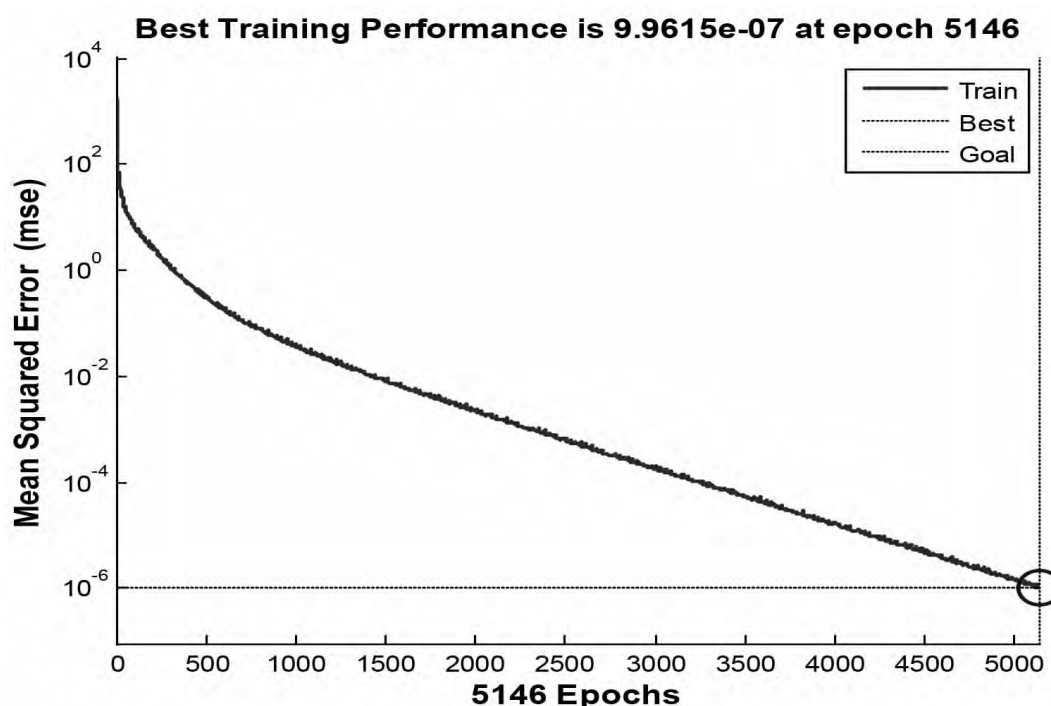


图 4 实验区植物群落现存量的误差训练曲线示意图

Fig. 4 Sketch map of the error training curve of the plant community in experimentation area

模型训练结束后将训练的输出数据还原,以目标向量对比模型训练后的输出数据,其实测值与预

测值拟合结果见表 1。

表 1 实验区植物群落现存量的 BP 神经网络训练样本结果

Table 1 The results of BP neural network training sample of the plant community in the experimentation area

序号	实测值 $y_i$	预测值 $\hat{y}$
1	62.55	62.63
2	18.55	18.78
3	30.25	30.75
4	30.48	29.89
5	51.11	51.11
6	29.38	29.11
7	61.26	61.18

由表 1 可以看到,实测值与预测值非常接近,经分析其相关系数达到 0.99,平均相对误差为 0.86%,可见 3 个输入因子与输出因子之间关系密切,可用禾本科为主群落的高度、盖度和密度对草地植物群落地上现存量进行模拟。这也说明采用 BP 神经网络能够精确的揭示草地植物群落高度、盖度、密度与地上现存量之间的数量关系,获得 BP 神经网络模型十分理想。

2.1.2 网络的检验

取禾本科为主实验区余下的 3 组数据对模型进行检验,BP 神经网络测试 results 和野外实测结果

见表 2。

表 2 实验区植物群落现存量的 BP 神经网络测试样本结果

Table 2 The results of the BP neural network test sample of the plant community in the experimentation area

序号	实测值 $y_i$	预测值 $\hat{y}$	绝对误差 E	相对误差 % E
1	45.09	42.93	2.16	4.79
2	29.11	30.61	1.50	5.15
3	24.56	22.68	1.88	7.66

在表 2 可以看到,该模型的实测值与预测值的误差很小,平均预测准确率为 94.10%,属于较好水平,因此 BP 神经网络模型可用于禾本科为主植物群落现存量的模拟和预测。

2.2 碱韭为主实验区草地样本的 BP 神经网络拟合训练

2.2.1 网络的训练与预测

同禾本科为主试验区草地样本 BP 神经网络训练一样,期望误差选取为 0.001,最大训练次数选为 2000,初始学习速率为 0.05。神经网络经初始化,利用函数 traingdm 对网络进行训练迭代到 489 步时,均方误差就达到了目标误差的要求。训练变化曲线如图 5 所示。

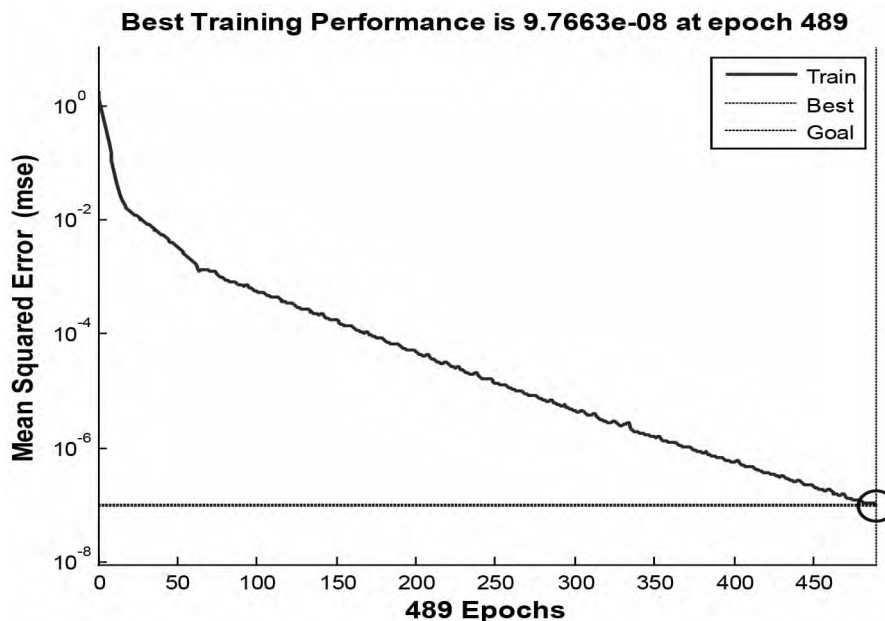


图 5 碱韭为主草地样本植物群落现存量的误差训练曲线示意图

Fig. 5 Schematic diagram of the error curve allium polyrhizum based training samples of the grassland standing crop of plant communities

根据建立的 BP 神经网络模型,获得其拟合值,将植物群落地上现存量野外实测值和 BP 神经

网络拟合值整理在一起,结果见表 2。

表 3 碱韭为主草地样本植物群落现存量的

BP 神经网络训练样本结果

Table 3 The results of BP neural network training samples of the plant community samples allium polyrhizum grassland standing crop composite sample

序号	实测值 $y_i$	预测值 $\hat{y}$
1	22.77	22.73
2	9.94	10.05
3	20.45	20.56
4	18.80	19.16
5	17.71	17.56
6	20.96	20.68
7	21.80	21.68

由表 3 可以看到, 实测值与预测值非常接近, 经分析其相关系数达到 0.99, 平均相对误差为 0.93%, 可见 3 个输入因子与输出因子之间关系密切, 可用碱韭为主群落高度、盖度和密度对草地植物群落地上现存量进行模拟。这也说明采用 BP 神经网络能够精确的揭示草地植物群落高度、盖度、密度与地上现存量之间的数量关系, 获得 BP 神经网络模型十分理想。

## 2.2.2 网络的检验

取碱韭为主实验区余下的 3 组数据对模型进

行检验, BP 神经网络测试结果和野外实测结果见表 4。该模型的实测值与预测值的误差很小, 平均预测准确率为 92.98%, 属于较好水平, 因此 BP 神经网络模型可用于碱韭为主植物群落现存量的模拟和预测。

表 4 碱韭为主草地样本植物群落现存量的

BP 神经网络测试样本结果

Table 4 plant community samples allium polyrhizum grassland standing cropBP neural network test sample

序号	实测值 $y_i$	预测值 $\hat{y}$	绝对误差 $e$	相对误差 % $E$
1	20.60	18.45	2.15	10.437
2	22.15	21.23	0.92	4.153
3	16.35	17.41	1.066	4.83

## 2.3 复合草地样本的 BP 神经网络拟合训练

### 2.3.1 网络的训练与预测

同禾本科为主试验区草地样本 BP 神经网络训练一样, 复合样本 BP 神经网络模型训练时期望误差选取为 0.001, 最大训练次数选为 10000, 初始学习速率为 0.05。神经网络经初始化, 利用函数 traingdm 对网络进行训练迭代到 7179 步时, 均方误差就达到了目标误差的要求。训练变化曲线如图 6 所示。

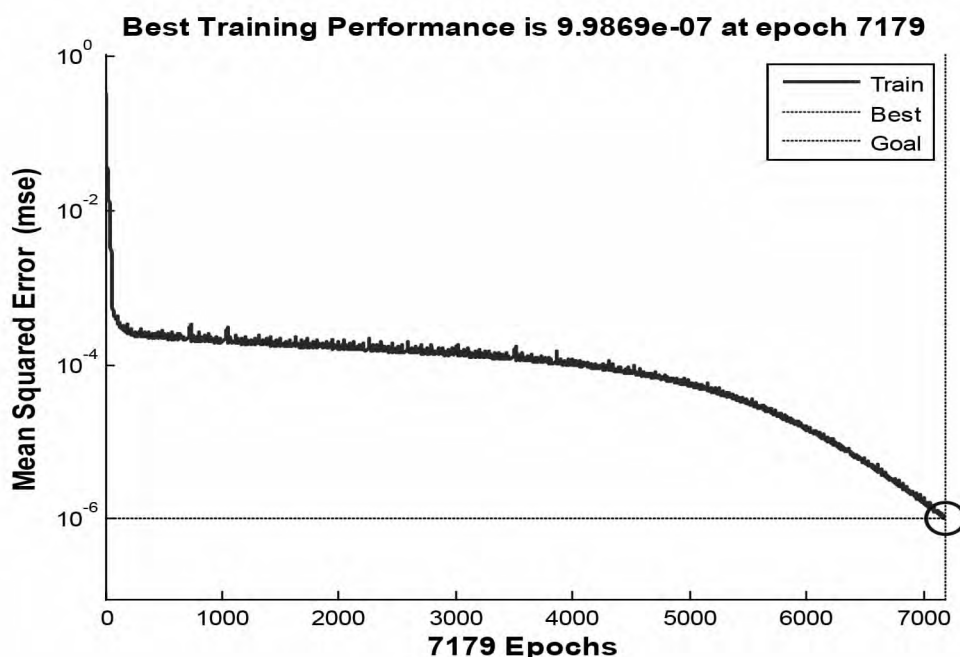


图 6 植物群落复合样本现存量的误差训练曲线示意图

Fig. 6 Error training curves of composite samples of plant community

根据建立的 BP 神经网络模型, 获得其拟合值, 将植物群落地上现存量野外实测值和 BP 神经网络拟合值整理在一起, 结果见表 5。由表 5 可以看到, 实测值与预测值非常接近, 经分析其相关系数达到 0.99, 平均相对误差为 4.25%, 可见 3 个输入因子与输出因子之间关系密切。所以, 不管采用层内样本还是复合样本, 草地植物群落高度、盖度和密度与草地植物群落地上现存量之间可建立 BP 神经网络关系。这也进一步说明采用 BP 神经网络能够精确的揭示草地植物群落高度、盖度、密度与地上现存量之间的数量关系。

表 5 植物群落复合样本现存量的

BP 神经网络训练样本结果

Table 5 The results of BP neural network training samples of the plant community composite sample

序号	实测值 $y_i$	预测值 $\hat{y}$
1	61.26	63.21
2	62.55	62.55
3	45.09	43.47
4	29.11	30.47
5	24.56	23.80
6	30.25	30.11
7	30.48	32.65
8	51.11	48.73
9	18.55	19.10
10	29.38	27.44
11	20.96	21.12
12	22.77	22.71
13	21.80	22.46
14	20.60	21.03

### 2.3.2 网络的检验

取复合样本余下的 6 组数据对模型进行检验, BP 神经网络测试结果和野外实测结果见表 6。该模型的实测值与预测值的误差很小, 平均预测准确率为 91.01%, 属于较好水平, 因此 BP 神经网络模

型可用于复合样本植物群落现存量的模拟和预测。

表 6 植物群落复合样本现存量的

BP 神经网络测试样本结果

Table 6 Test sample results of BP neural network in plant community composite sample

序号	实测值 $y_i$	预测值 $\hat{y}$	绝对误差 $e$	相对误差 % $E$
1	22.15	20.87	1.28	5.779
2	20.45	21.02	0.57	2.787
3	18.80	16.90	1.90	10.106
4	17.71	19.21	1.50	8.470
5	9.94	11.36	1.42	14.286
6	16.35	14.30	2.05	12.538

## 3 讨论与结论

### 3.1 结论

(1) 采用 BP 神经网络进行草地植物群落地上现存量的模拟与预测能够获得较高的拟合率和预测精度(预测精度可达 90% 以上)。

(2) 采用 BP 神经网络对不同草地植物类群进行模拟和预测可能会存在差异, 但这种差异会随着样本容量的增大或训练精度的增加而弱化。

### 3.2 讨论

本文基于 BP 神经网络模型对荒漠草原植物群落地上现存量的模拟与预测研究取得了比较好的结果, 但荒漠草原植物群落现存量的评价与预测研究仍是一个逐渐完善的过程, 还有一些问题需要进一步的讨论及研究。

(1) BP 神经网络模型的评价精度与训练样本的数量和质量高度相关。但由于时间和精力限制, 本文所选取的训练样本有限。希望在进一步的研究中, 能够选取更多不同地区和年份的指标作为神经网络的训练样本, 以便更好的反映荒漠草原植物群落地上现存量的发展规律, 从而印证 BP 神经网络在植物群落地上现存量评价中的应用价值。

(2) 本文中, 仅仅基于 BP 神经网络对荒漠草原植物群落现存量的模拟与预测研究进行了最基

本的尝试,目前 BP 神经网络还存在一些不足,它的计算过程受到网络学习参数、网络层数、隐含层节点数、学习步长等因素的影响,在计算过程中会产生一定的波动。因此,关于 BP 神经网络模型参数的设定及训练样本、检验样本的选取等问题还需进一步研究。

(3)影响荒漠草原植物群落地上现存量的因素除了植物群落的数量特征(高度、盖度、密度)外还有其他的一些可控与不可控的因素,因此应进一步增加影响因子的个数,以提高预测的精度。

## 参考文献:

- [1] Debra M. Cassels, Robert L. , et al Effects of grazing management on standing crop dynamics in tallgrass-prairie[J]. J. Range Management 1995,48:81—84.
- [2] Ralphs M H. 短时间高强度放牧对草地植被的影响[J]. 草原与牧草,1991,2:34—41.
- [3] Raymond F. Angell, Crested wheatgrass and shrub response to continuous or rotational grazing[J]. J. Range Management 1997,50:160—164.
- [4] 王得利,吕新龙,罗卫东等. 不同放牧密度对草原植被特征的影响分析[J]. 草业学报,1996,5(3):28—33.
- [5] Robert L. Gillen F. Ted Mccollum III, et al. Tallgrass prairie response to grazing system and stocking rate[J]. J. Range Management 1998,51:139—146.
- [6] 汪诗平,李永宏,陈佐忠. 内蒙古典型草原载畜适宜放牧率的研究. II. 以牧草地上现存量与净初级生产力为管理目标[J]. 草地学报,1999,7(3):198—203.
- [7] 李永宏,陈佐忠,汪诗平,等. 草原放牧系统持续管理试验研究:试验设计及放牧率对草地系统影响分析[J]. 草地学报,1999,7(3):174—182.
- [8] 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.
- [9] 王立海,邢艳秋. 基于 BP 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.
- [10] 董长虹, MATLAB 神经网络与应用[M]. 北京:国防工业出版社,2007
- [11] 杨持. 生态学[M]. 北京:高等教育出版社,2008
- [12] 卫智军,韩国栋,赵钢,李德新. 中国荒漠草原生态系统研究[M]. 北京:科学出版社,2013
- [13] 王鑫厅,梁存柱,王炜. 尺度与密度:测定不同尺度下的种群密度[J]. 植物生态学报,2013,37(2):104—110.
- [14] 叶世伟,神经网络原理[M]. 北京:机械工业出版社,2004
- [15] 孙亮, MATLAB 语言与控制系统仿真,北京:北京工业大学出版社,2012
- [16] 王敏,苏永中,杨荣,杨晓. 黑河中游荒漠草地上和地下生物量的分配格局[J]. 植物生态学报,2013,37(3):209—219.
- [17] 张金屯. 数量生态学(第二版)[M]. 北京:科学出版社,2011.

[收稿日期] 2016—01—30

[基金项目] 内蒙古自然科学基金项目(2015MS0349);国家自然科学基金项目(31460126);国家重点基础研究计划项目(2014CB138800)。

[作者简介] 张 爽(1994—),女,硕士研究生,主要研究方向为草地生态与管理, e-mail: zs123shuang@163.com。

[通讯作者] 卫智军 E-mail: nmndwzj@163.com。