

文章编号:1001-1498(2016)02-0256-05

基于气象因子的马尾松毛虫发生率空间格局研究

王庆, 毕猛, 杜婷, 廖怀建, 石雷*

(中国林业科学研究院资源昆虫研究所, 云南昆明 650224)

摘要: [目的] 为预测未来我国马尾松毛虫的潜在变化趋势, 以 2002—2012 年全国范围内马尾松毛虫的地级逐年平均发生率作为预测指标, [方法] 运用偏最小二乘回归方法, 获得马尾松毛虫平均发生率与相关气象因子的回归方程, 并结合地理空间数据与未来气象数据, 得到马尾松毛虫平均发生率空间格局模型。[结果] 表明: 以筛选后的 12 个气象因子建立的马尾松毛虫平均发生率空间格局模型精度达到 86.98%, 具有较强的可靠性。据此预测 2020s, 2050s, 2080s 的马尾松毛虫平均发生率空间格局, 并与 2002—2012 年的空间格局相比, 结果显示: 华东及华中地区虫害中度和重度发生面积均明显增加, 有扩散的趋势; 华东地区的轻度发生面积总体为缩减; 而华南部分地区虫害轻度发生面积扩增。[结论] 以偏最小二乘回归方法所得的空间格局模型具有实际预测意义, 可以预测我国未来马尾松毛虫平均发生率的变化趋势。

关键词: 马尾松毛虫; 气象因子; 发生率; 偏最小二乘; 空间格局

中图分类号: S791.248

文献标识码: A

Spatial Pattern of *Dendrolimus punctatus* Incidence Rate Based on Meteorological Factors

WANG Qing, BI Meng, DU Ting, LIAO Huai-jian, SHI Lei

(Research Institute of Resources Insects, Chinese Academy of Forestry, Kunming 650224, Yunnan, China)

Abstract: [Objective] Taking the average incidence rate based on China's nationwide data from 2002–2012 as indicator to predict the potential trend of *Dendrolimus punctatus* incidence rate. [Method] By means of partial least squares regression, the regression equation about average incidence rate and the related meteorological factors was obtained. Combined with the geographic spatial data and future meteorological data, the spatial pattern model of the average incidence rate of *D. punctatus* was established. [Result] The spatial pattern model of *D. punctatus*' average incidence rate built by 12 selected meteorological factors has the prediction accuracy of 86.98%. Based on this model, the spatial pattern models for 2020s, 2050s, and 2080s were established. It was predicted that compared with 2002–2012, the area of moderate and severe insect pests in East and Central China would significantly increase, and there would be a trend of spreading. The mild incidence area would decrease in East China, while the mild incidence area would have a trend of amplification in parts of Southern China. [Conclusion] The spatial pattern model obtained by partial least squares regression method can be used to predict the potential changes of the average incidence rate of *D. punctatus* in China.

Keywords: *Dendrolimus punctatus* Walker; meteorological factor; insect incidence rate; partial least squares; spatial pattern

收稿日期: 2015-09-10

基金项目: 国家林业局林业公益性行业科研专项“重大森林病虫害防控技术的关键理论基础”(201204501); 云南省技术创新人才培养计划(2012HB054)。

作者简介: 王庆(1989—), 男, 硕士研究生, 主要从事森林病虫害预测预警研究。

* 通讯作者。

马尾松毛虫(*Dendrolimus punctatus* Walker),属鳞翅目(Lepidoptera)、枯叶蛾科(Lasiocampidae),是我国历史性森林害虫,主要对马尾松造成危害,可造成大量松树死亡^[1]。气候是影响马尾松毛虫扩散、爆发的重要因素,马尾松毛虫发育需要一定的积温,平均温度越高,完成一世代数所需天数越少^[2],年积温升高,虫害发生面积也会随之增大^[3]。已有研究表明,若马尾松毛虫卵的孵化和幼虫的发育时期适逢降雨,则对其幼虫成活率十分有利^[4-5],而高温则会使幼虫的死亡率提升^[6]。光照、湿度、风向和风力等其他气候因子也对马尾松毛虫成灾有不同程度的影响:日照时数,尤其是前一年2月和当年11月的日照时数与松毛虫的暴发具有显著的正相关性^[7-8];低湿对马尾松毛虫卵与低龄幼虫的发育成活有明显的不良影响,如果空气相对湿度持续低于75%,马尾松毛虫就不能正常的生长发育^[9];大风可以使羽化的松毛虫传播更远的距离,可导致松毛虫成虫的大转移,强台风及其引起的暴雨可使虫体大量死亡^[10-11]。

为预测松毛虫的发生发展,国内外学者提出了许多预测方法,如基于经验模型、回归模型^[12-13]的线性预测,人工神经网络^[14]和logistic回归^[15]等非线性预测以及地统计学等方法^[16],但这些方法都只限于局部地区,所选的气候因子也有较大的局域性,对全国范围的大尺度分析预测少有研究。本文采用偏最小二乘(Partial Least Squares, PLS)回归方法进行预测,与其他方法(如神经网络)相比,PLS有计算简便、模型精度高、预测效果好、无需剔除变量或样本等优点,可解决许多普通线性回归无法解决的问题^[17-18]。尤其是当利用多种气象因子作因变量时,各气象因子之间往往存在较高的相关性,用PLS方法做回归建模,可避免因变量内部相关而对模型的准确性和合理性造成的影响,比对逐个因变量做多元回归更加有效,其结论更加可靠^[19-20]。本研究利用2002—2012年全国范围内马尾松毛虫的发生数据,分析其平均发生率与气象因子的相关性,并筛选出主要影响因子,建立马尾松毛虫平均发生率的预测模型。在ClimateChina气象数据模拟软件的支持下,结合地理空间数据,预测未来全国范围内马尾松毛虫的潜在变化趋势。

1 研究数据与方法

1.1 研究数据

2002—2012年全国范围内马尾松毛虫逐年发生数据来自国家林业局森林病虫害防治总站。利用监测信息完整的82个地级行政区的马尾松毛虫发生数据,计算出各地平均发生率AIR(表1)。

表1 马尾松毛虫地级平均发生率统计数据

地区名称	总发生面积/万 hm ²	总监测面积/万 hm ²	平均发生率/%
安庆市	15.109 8	874.202 7	1.728 4
百色市	1.912 3	629.559 3	0.303 8
常德市	2.379 2	339.038 7	0.701 7
巢湖市	4.049 3	139.260 0	2.907 7
潮州市	0.015 3	22.017 3	0.069 5
崇左市	5.730 2	647.158 7	0.885 4
滁州市	5.386 8	226.491 3	2.378 4
福州市	3.163 9	348.942 7	0.906 7
佛山市	0.011 3	4.632 7	0.243 9
广安市	7.436 9	111.904 7	6.645 7
.....
遵义市	0.376 7	258.654 7	0.145 6

利用我国全境航天飞机雷达地形测绘使命(SRTM)的数字高程模型(DEM)数据,选取138 465 95个地理坐标点数据(经纬度、高程)均匀分布于我国全境,各地级行政区2002—2012年的平均气象数据及未来各个时期平均气象数据均由气象数据模拟软件ClimateChina计算得出。该软件使用广泛应用于气象数值模拟的PRISM插值算法开发,并加以改进^[21-24],使其具有较高的准确性。可导出我国任意点的历史气象数据及未来2020s(2010—2039)、2050s(2040—2069)、2080s(2070—2099)的气象数据,主要包括年平均气温、降雨量,月平均气温、降雨量,年温湿系数,无霜期等气象因子数据。输出的气象数据已通过中国气象科学数据共享服务网验证。

1.2 研究方法

参考相关文献及前人研究成果^[25-26],利用偏最小二乘方法建模关键是运用舍一交叉验证法来确定主成分个数,可通过预测残差平方和(PRESS)的最小值来确定;同时,利用变量投影重要性VIP指标来判断所选自变量的重要性。VIP值越大,自变量对因变量的解释作用越明确,通常选取VIP值大于1的自变量来参与建模^[27]。

用马尾松毛虫平均发生率(*AIR*)预测值的决定系数(R^2)、平均误差(*AE*)、总预测相对误差(*PRE*)等指标来评价模型的精度^[25],相关数据的计算在SAS软件上运行。

2 结果与分析

2.1 气象因子的筛选

初步选取 ClimateChina 输出的 68 个气象因子作为相关性分析的备选变量,与经对数变化后的 2002—2012 年马尾松毛虫 *AIR* 数据($\ln AIR$)进行 Pearson 积矩相关性分析,并做双侧显著性检验。分析结果显示:2 月平均气温(*TAV02*)、5 月平均气温(*TAV05*)、10 月平均气温(*TAV10*)、9 月平均最高气温(*TMX09*)、1 月平均最低气温(*TMN01*)、9 月平均最低气温(*TMN09*)、冬季平均最低气温(*TMIN_wt*)、春季平均气温(*TAV_sp*)、秋季平均气温(*TAV_at*)、夏季降水量(*PPT_sm*)、最热月和最冷月之间温差(*TD*)、无霜期结束的儒略日(*eFFP*)等 12 个气象因子与 $\ln AIR$ 呈显著相关,将其作为建立回归模型的人选因子。

在进行 PLS 回归时,计算出参与建立模型的 12 个气象因子的 *VIP* 值,其中,2 月平均气温、秋季平均气温、春季平均气温、最热月和最冷月之间温差、5 月平均气温、10 月平均气温、1 月平均最低气温的 *VIP* 值均大于 1,说明这几个气象因子对 $\ln AIR$ 的解释作用很强。其他几个因素虽然 *VIP* 值未达到 1(0.72 ~ 0.86),但由于 9 月气温、12 月气温、夏季降水量和无霜期的长短都对马尾松毛虫的发生发展具有影响,所以,仍将其作为入选变量,参与模型的建立(表 2)。

表 2 气象因子与 $\ln AIR$ 的相关性系数以及 *VIP* 值

气象因子	<i>P</i> 值	<i>VIP</i>
2 月平均气温(<i>TAV02</i>)	0.007 5	2.368
5 月平均气温(<i>TAV05</i>)	0.020 3	1.299
10 月平均气温(<i>TAV10</i>)	0.001 3	1.121
9 月平均最高气温(<i>TMX09</i>)	0.004 7	0.865
1 月平均最低气温(<i>TMN01</i>)	0.009 0	1.003
9 月平均最低气温(<i>TMN09</i>)	0.007 8	0.772
冬季平均最低气温(<i>TMIN_wt</i>)	0.009 2	0.743
春季平均气温(<i>TAV_sp</i>)	0.012 8	1.851
秋季平均气温(<i>TAV_at</i>)	0.001 5	2.001
夏季降水量(<i>PPT_sm</i>)	0.008 5	0.720
最热月和最冷月之间温差(<i>TD</i>)	0.002 5	1.301
无霜期结束的儒略日(<i>eFFP</i>)	0.037 3	0.843

2.2 空间格局模型的建立

用马尾松毛虫 $\ln AIR$ 作为因变量,筛选出的 12 个气象因子作为自变量进行 PLS 回归运算,*PRESS* 值随主成分数的增加而不断减小,当主成分数为 10 时,*PRESS* 达到最小值 0.688 1。因此,选取 10 个主成分参与建模,得到马尾松毛虫 *AIR* 空间格局模型,计算公式如下:

$$AIR = \exp(33.799\ 6 + 4.865\ 9 \times TAV02 + 3.247\ 5 \times TAV05 + 3.124\ 5 \times TAV10 + 2.162\ 1 \times TMX09 + 3.415\ 1 \times TMN01 + 0.842\ 3 \times TMN09 - 1.497\ 7 \times TMIN_wt - 6.546\ 4 \times TAV_sp - 9.530\ 7 \times TAV_at + 0.003\ 5 \times PPT_sm + 1.375\ 1 \times TD - 0.062\ 9 \times eFFP)$$

对模型进行精度检验,结果显示:2011—2012 年马尾松毛虫 *AIR* 实测值和预测值具有线性关系(图 1),决定系数(R^2)为 0.852 2,平均误差(*AE*)为 -0.265,总预测相对误差(*PRE*)为 13.02%,整体预测精度超过 85%,达到 86.98%,说明模型精度较高,有实际预测意义。

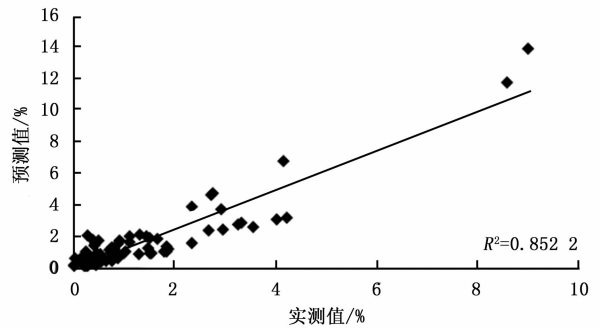


图 1 马尾松毛虫平均发生率精度检验

2.3 空间格局分析

将地理坐标点数据导入 ClimateChina 气象模拟软件,得到 2002—2012 年、2020s、2050s 和 2080s 4 个时期参与建模的 12 个气象因子数据。利用回归模型,计算出这些地理坐标点的马尾松毛虫 *AIR* 值,通过 ArcGIS 软件,得到我国马尾松毛虫 *AIR* 从 2002—2012 年到 2080s 的空间格局分布示意图(图 2)。

按照《林业有害生物发生及成灾标准》,将各地区的马尾松毛虫的发生程度分为无、轻度($AIR < 10\%$)、中度(AIR 为 $10\% \sim 20\%$)和重度($AIR > 20\%$)4 个等级,并分别统计出 4 个时期我国马尾松毛虫各等级的发生面积(表 3)。

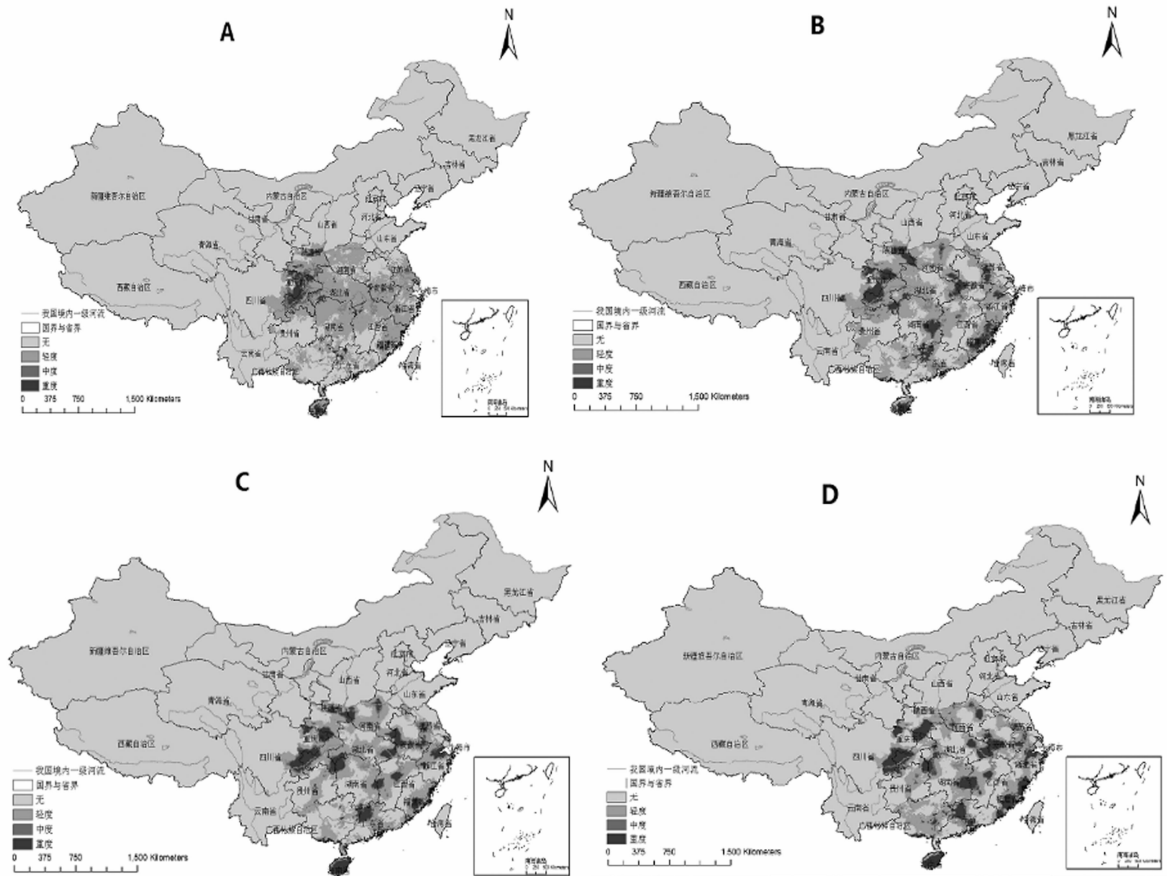


图2 4个时期马尾松毛虫平均发生率空间格局变化示意图(A、B、C、D分别为2002—2012年、2020s、2050s、2080s)

表3 马尾松毛虫各等级发生面积及比值

时间	轻度发生面积 /万 km ²	占总发生面积比值 /%	中度发生面积 /万 km ²	占总发生面积比值 /%	重度发生面积 /万 km ²	占总发生面积比值 /%	总发生面积 /万 km ²
2002—2012年	91.135	84.88	8.667	8.07	7.567	7.05	107.369
2020s	79.014	73.64	11.046	10.29	17.237	16.07	107.297
2050s	69.148	68.53	10.083	9.99	21.674	21.48	100.905
2080s	62.598	62.75	10.350	10.38	26.807	26.87	99.755

预测结果显示:2002—2012年到2080s,马尾松毛虫的总体发生面积有递减的变化趋势,轻度发生面积逐渐减少;中度发生面积略有升高,但变化不大;重度发生面积则呈明显的增加趋势。从发生区域来看,安徽西部、湖北中部、江西东北部、江苏东部地区马尾松毛虫轻度发生面积呈缩减趋势,贵州中西部、广东南部地区轻度发生面积有所扩增,陕西东南部、安徽西部、湖北东部、湖南东部、江苏西部、浙江北部、福建东部、广东北部地区虫害中度和重度发生面积均明显增加,而后2050s—2080s虫害重度发生面积基本保持稳定。

3 结论与讨论

在本研究筛选出的与马尾松毛虫危害相关的12个气象因子中,2月份平均气温、1月份平均最低气温及冬季平均最低气温直接影响马尾松毛虫幼虫的越冬死亡率,冬季的持续低温可使幼虫冻结,若出现“倒春寒”的先温暖后寒冷的气温变化,则苏醒的幼虫会大量冻死,从而影响其发生率。夏季降水带来的温度、湿度的变化也直接影响到马尾松毛虫的发育,9月正值马尾松毛虫幼虫孵化时期,温度和降水显著影响马尾松毛虫的孵化率,冬季气温和无霜期的长短则影响马尾松毛虫翌年的暴发。据未来气象模拟估计^[28-29],随着全球变暖趋势,21世纪中叶

我国南方地区的平均气温将升高 $1.2 \sim 1.6^{\circ}\text{C}$, 到 21 世纪末期平均气温仍会升高 $3.0 \sim 4.0^{\circ}\text{C}$ 左右; 这段时期秋季的平均降雨量增加, 冬季平均降雨量则减少, 霜冻日数在全国范围内也呈减小趋势, 气温年较差的变化在长江以南地区呈增加趋势^[30]。这对马尾松毛虫在未来时期的流行扩散均起有利作用。

本研究建立的模型预测精度达到 86.98%, 得到了较好的预测结果。通过气象软件的插值算法, 可对某些无气象统计数据的地区进行预测预报, 对病虫害防治有十分重要的意义。如果气候变化趋势与所预测气象数据一致, 则未来马尾松毛虫发生情况也将向模型所预测的结果方向发展。由于本研究是基于气候的稳定变化来进行模拟, 没有考虑突发的极端灾害性天气对发生率所造成的影响, 所模拟的气象数据也是某一时期的平均数据而非逐年的, 而且病虫害的流行扩散对气候变化有明显的滞后效应, 所以预测的未来某一时期的马尾松毛虫发生率空间格局只是其潜在分布而非实际分布。马尾松毛虫的发生发展与气象因子是否只是简单的线性关系也有待商榷, 若能与人工神经网络等数学方法相结合^[31], 既能解决变量之间的相关性, 又能解决普通回归模型不能很好解释的非线性问题, 将可能会获得更加精确、科学的预测预报模型。

参考文献:

- [1] 刘兴平, 温小遂. 马尾松毛虫成灾原因及其综合治理[J]. 江西植保, 2007, 30(2): 61-64.
- [2] 曾菊平, 戈峰, 苏建伟, 等. 我国林业重大害虫松毛虫的灾害研究进展[J]. 昆虫知识, 2010, 47(3): 451-459.
- [3] 刘志明, 晏明, 张旭东, 等. 气象因子对落叶松毛虫的影响分析[J]. 灾害学, 2002, 17(4): 32-37.
- [4] 李振华. 雨日对马尾松毛虫诱发作用的初探[J]. 江西林业科技, 1997(5): 32-33.
- [5] 何忠, 韩瑞东, 刘向辉, 等. 环境温度对马尾松毛虫发育与存活的影响[J]. 应用生态学报, 2006, 17(3): 483-488.
- [6] 李兆麟, 何忠, 贾凤友. 马尾松毛虫幼虫发育的光周效应[J]. 林业科学研究, 1993, 6(3): 276-281.
- [7] 侯陶谦. 中国主要害虫——松毛虫的综合防治[M]. 北京: 中国林业出版社, 1979.
- [8] 王学锋, 周德丽, 童青, 等. 思茅林区松毛虫害发生与气象因子的关系[J]. 云南林业科技, 2003(3): 77-82.
- [9] 何忠, 席瑞华, 王序英, 等. 环境湿度对马尾松毛虫卵和低龄幼虫发育与成活的影响[C]//李典谟, 康乐, 吴钜文, 等. 昆虫学创新与发展, 北京: 中国科学技术出版社, 2002: 508-513.
- [10] 孙思清, 邵守英, 刘鸿岩. 台风暴雨对赤松毛虫口密度的影响[J]. 山东林业科技, 1997(6): 22-23.
- [11] 余会康, 黄岩彬. 古田水口林区松毛虫发生的气候成因分析[J]. 四川林业科技, 2005, 26(6): 83-85.
- [12] 厉悦, 刘敏, 张合平等. 基于多元统计分析的影响马尾松毛虫灾害发生的生态环境因素研究[J]. 江苏农业科学, 2012, 40(5): 90-92.
- [13] 田万银, 徐华潮. 浙江沿海防护林马尾松毛虫的预测预报模型[J]. 环境昆虫学报, 2012, 34(4): 401-406.
- [14] 陈绘画, 朱寿燕, 崔相富, 等. 基于人工神经网络马尾松毛虫发生量预测模型的研究[J]. 林业科学研究, 2003, 16(2): 159-165.
- [15] 高倩, 安英博, 吴凤祥. 基于 Logistic 回归的落叶松毛虫预测模型研究[J]. 河北农业大学学报, 2011, 34(6): 108-110.
- [16] 向昌盛. 基于地统计学定阶的松毛虫发生面积组合预测[J]. 计算机应用研究, 2012, 29(3): 984-987.
- [17] Wold S, Albano C. Modeling Data Tables by Principal Component and PLS: class Patterns and Quantitative Predictive Relations[J]. Analysis, 1984, 12: 477-485.
- [18] Höskuldson A. PLS regression methods[J]. Journal of Chemometrics, 1988, 2: 221-228.
- [19] 高惠璇. 两个多重相关变量组的统计分析(偏最小二乘回归与 PLS 过程)[J]. 数理统计与管理, 2002, 21(2): 58-64.
- [20] 王惠文, 吴载斌, 孟浩. 偏最小二乘回归的线性与非线性方法[M]. 北京: 国防工业出版社, 2006.
- [21] Wang T, Hamann A, Spittlehouse D L, et al. Development of scale-free climate data for western Canada for use in resource management[J]. International Journal of Climatology, 2006, 26(3): 383-397.
- [22] Hamann A, Wang T L. Models of climate normal for geneecology and climate change studies in British Columbia[J]. Agricultural and Forest Meteorology, 2005, 128(3-4): 211-221.
- [23] Daly C, Gibson W P, Taylor G H, et al. A knowledge-based approach to the statistical mapping of climate[J]. Climate Research, 2002, 22(2): 99-113.
- [24] Mitchell T D, Jones P D. An improved method of constructing a database of monthly climate observations and associated high-resolution grids[J]. International Journal of Climatology, 2005, 25(6): 693-712.
- [25] 毕猛, 杜婷, 马思佳, 等. 基于气象因子的杨树溃疡病发生率空间格局研究[J]. 林业科学研究, 2014, 27(1): 24-30.
- [26] 王庆, 毕猛, 马思佳, 等. 基于气象因子的松墨天牛发生率空间格局研究[J]. 林业科学研究, 2015, 28(1): 61-66.
- [27] Chong I G, Jun C H. Performance of some variable selection methods when multicollinearity is present[J]. Chemometrics and Intelligent Laboratory Systems, 2005, 78(1-2): 103-112.
- [28] 汤剑平, 陈星, 赵鸣, 等. IPCC-A2 情景下中国区域气候变化的数值模拟[J]. 气象学报, 2008, 66(1): 13-25.
- [29] 江志红, 张霞, 王冀. IPCC-AR4 模式对中国 21 世纪气候变化的情景预估[J]. 地理研究, 2008, 27(4): 787-799.
- [30] 王冀. 中国地区极端气温变化的模拟评估及其未来情景预估[D]. 南京: 南京信息工程大学, 2008.
- [31] 陈南祥, 曹连海, 李梅, 等. 偏最小二乘回归神经网络的矿坑涌水量预测[J]. 吉林大学学报, 2005, 35(6): 765-770.

(责任编辑:张玲)