

The Study on Estimating Grassland NPP based on BP Neural Network Algorithm

Zhijun Tong*, Xiangqian Li, Xingpeng Liu, Jiquan Zhang

School of Environment, Northeast Normal University; Institute of Natural Disaster Research,
Northeast Normal University, Changchun 130024, China

基于 BP 神经网络算法的草原 NPP 估算研究 ——以锡林郭勒草原为例

佟志军*, 李向前, 刘兴朋, 张继权

东北师范大学环境学院, 东北师范大学自然灾害研究所, 长春 130024, 中国

Abstract

Grassland NPP is an important component of the carbon cycle of terrestrial ecosystems, which can directly measure the carrying capacity of grassland ecosystem. Considering the disadvantages of the existing NPP estimated models, this study adopt BP neural network model to estimate the NPP for XilinGol grassland based on the meteorological, field sample, DEM and remote sensing data. Our results showed that: (1) The trained BPNN not only has excellent performance for estimating NPP for the trained dataset, but also has good generation and portability, which can be applied for other pixels; (2) The spatial pattern of the grassland NPP has obviously horizontal zonal regularity, which rises from 0 to $>300\text{gC/m}^2$ from west to east.

Keywords: Grassland; Estimate; NPP; BPNN; XilinGol

*通讯作者: 佟志军, 副教授, 硕士生导师, 主要从事生态安全评价与管理、草原火灾风险预警等方向研究。
Email: gis@nenu.edu.cn

摘要

草原 NPP 是陆地生态系统碳循环的重要组成部分, 是草原生态系统承载力水平的直接衡量指标。本文考虑到现有三类 NPP 估算模型的缺陷, 决定采用 BP 神经网络算法从数据挖掘的角度, 结合气象、遥感、野外等多源遥感数据, 以锡林郭勒盟草原为研究区, 精确估算其 NPP 的数值。研究结果表明: (1) 训练后的 BPNN 模型具备优秀的 NPP 模拟性能, 且其泛化性和可移植性良好, 可用来估算研究区其它像元的 NPP 数值; (2) 锡林郭勒盟草原 NPP 空间格局存在较为明显的水平地带性规律, 其值自西向东由 0 逐渐上升至 300gC/m^2 以上。

关键词: 草原; 估算; NPP; BPNN; 锡林郭勒

1. 引言

作为陆地生态系统的重要组成部分, 草原约占陆地面积的 24%, 是地球陆地表面仅次于森林的第二个绿色覆被层, 与森林、海洋并列地球的三大碳库[1], 在全球碳循环过程中起着举足轻重的作用。草原净初级生产力 (Net Primary Productivity, NPP) 是指草原植被光合作用同化碳量与暗呼吸消耗碳量之差, 是草原生态系统碳循环过程的重要分量[2-4]。它直接表征草原生态系统在自然条件下的生产能力

和质量状况,是判定草原生态系统的碳源/汇功能和评价其生态承载力的核心指标[5-6]。因此,草原 NPP 的精准估算有助于清楚地认知草原生态承载力水平,对合理利用和开发草场资源,指导区域未来生产和规划具有重要的现实意义。

目前,气候生产潜力模型、生态过程模型和光能利用率模型是估算区域尺度陆地生态系统 NPP 的主要方法。其中,气候生产潜力模型的机理是根据实测点的数据建立回归方程,进而实现某点植被 NPP 的估算,这些模型虽然已经在世界各地得到了不同程度的验证,但模型考虑的植被和环境因子相对简单,忽略了许多影响 NPP 估算结果的植被生态生理反应和过程,其实质是植被 NPP 与环境因子的简单回归,且没有与特定的植被种类相关联,理论基础不足,模型估算结果误差较大;生态过程模型充分考虑了植被的生理反应和生态过程,然而所需参数众多,且这些参数在大尺度范围内很难获得;基于遥感数据驱动的光能利用率模型虽然能实现实时的陆地生态系统 NPP 的监测,然而模型的关键参数——最大光能利用率尚未有统一公认的取值标准,为 NPP 的估算带来较大的不确定性,备受学者争议。

随着人工智能的快速发展,利用人工神经网络(Artificial Neural Network,ANN)模型模拟植被 NPP 不失为解决现有问题的一种有效尝试。目前,ANN 模型已在风险、气象、水文、地质等诸多科学领域取得了较大的成就,

尤其是应用广泛的 BP 神经网络(Back Propagation Neural Network,BPNN)模型,它具有自组织、自适应和自学能力,强调从实际中总结经验,避免了复杂的数据分析和建模过程,其自适应的信息处理方式可完成复杂的输入-输出的非线性映射,特别适用于处理需要同时考虑多种因素和不精确条件的模糊信息问题[7-8]。NPP 变化成因复杂,与众多环境因子存在着复杂的耦合关系。有鉴于此,本文采用 BPNN 模型来实现草原 NPP 的精准估算。

2. 材料和方法

2.1 研究区域概况

锡林郭勒盟(简称锡盟)位于中国正北方,其地理坐标介于东经 $111^{\circ}09' \sim 119^{\circ}58'$,北纬 $41^{\circ}35' \sim 46^{\circ}46'$ 之间,土地总面积 202580km²,拥有 104.69 万人口(2016 年),下辖二市九旗一县(锡林浩特市、二连浩特市、东乌珠穆沁旗、西乌珠穆沁旗、阿巴嘎旗、苏尼特左旗、苏尼特右旗、正蓝旗、正镶白旗、镶黄旗、太仆寺旗、多伦县)(图 1)。就地形而言,该地区以高平原为主,多种地貌并存,其海拔介于 760~1900m 之间,地势总体呈现出南高北低的态势;就气候而言,该地区属于北温带干旱、半干旱的大陆性季风气候,其主要气候特点是风大、少雨、寒冷、干燥;就植被而言,该地区地处欧亚大陆草原带的中部,其中草原面积 192515km²,占全盟总面积的 95.03%,是我国北方温带草原的典型区域,其植被类型主要有草甸草原、典型草原和荒漠草原。

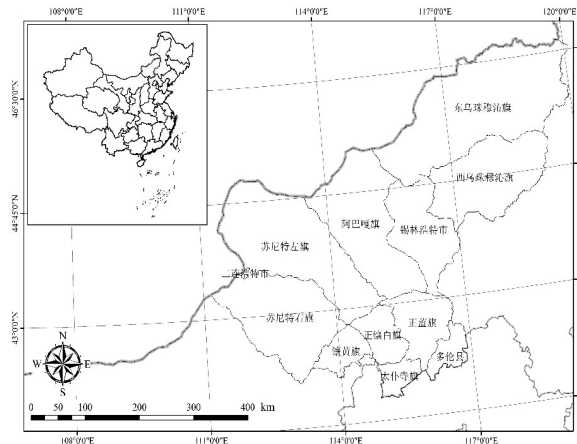


图 1 研究区地理位置

2.2 数据来源及预处理

本文所用到的数据可划分为四大类：NPP 实测数据、气象观测数据、DEM 数字高程数据和卫星遥感数据。其中，NPP 实测数据来源于 2013 年课题组野外采集的样方数据，共计 510 个，采集时间为草原植被生长茂盛的 8 月份。气象数据来自中国气象数据网 (<http://data.cma.cn>)，资料包括 2013 年 8 月份的逐日气温和降水气象要素。为保证空间插值精度，本文根据站点提供各气象要素在研究时段内（2013 年 8 月）的连续性，共选取锡盟及周边地区 36 个常规气象观测站点。DEM 数字高程数据来自地理空间数据云平台 (<http://www.gscloud.cn/>)。卫星遥感数据来自于美国国家航空航天局（National Aeronautics and Space Administration, NASA）的 MODIS 产品平台 (<https://ladsweb.nascom.nasa.gov/data>)，本文主要选取两类遥感产品：MOD13A3 和 MOD15A2H。其中，MOD13A3 主要负责提供锡盟草原的 NDVI 数据，MOD15A2H 主要负责提供锡盟草原的 LAI 和 FPAR 数据，时间选定为 2013 年 8 月份。值得注意的是，本文研

究区锡盟草原的地理区域涉及到 MODIS 的 h25v04 和 h26v04 景号，需经过 MRT(MODIS Reprojection Tool)的拼接、投影等步骤，最终才能获得完整覆盖研究区域的遥感影像。此外，本文将各影像的无效像元（主要为水体、裸地和居民点）均统一为 0 值。

2.3 研究方法—BP 神经网络

人工神经网络 (Artificial Neural Networks, ANN) 是一种模仿人脑神经网络结构和功能的数字信息处理模型。BP 神经网络，即误差反向传播神经网络，是神经网络模型中使用最广泛的一类。它因其经典的反向传播学习算法而命名，是一个典型的多层网络。Hecht Nielsen 证明一个三层的 BP 网络足以实现任意 n 维到 m 维空间的复杂映射[90]。因此 BPNN 经常采用 3 层结构，即输入层、隐含层、输出层。层与层之间全连接，同一层神经元之间不连接（图 2-4）。BPNN 算法的核心思想是通过反向传播模型学习后的误差来调整学习过程中的权重，最终使得整个网络的总误差达到最小，从而实现自学习、自组织、自适应的目的。

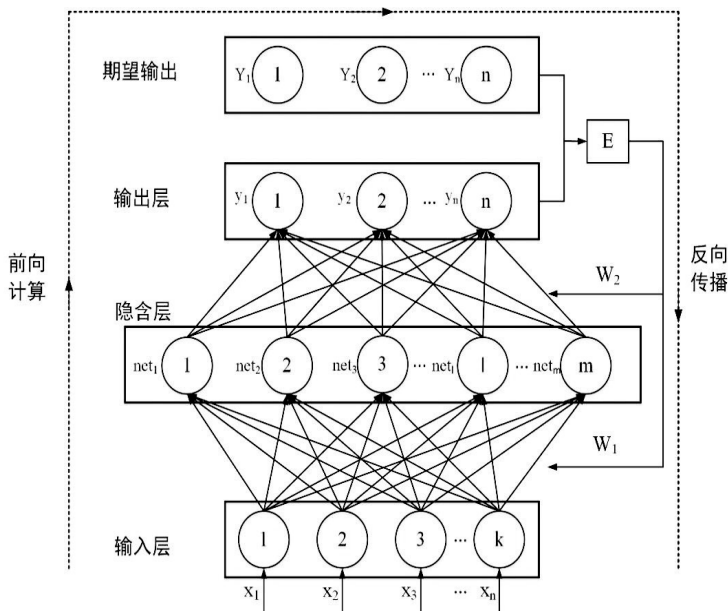


图 2 三层 BP 神经网络

具体算法如下:

$$net_j = \sum_i w_{ij} x_i + \theta_j \quad (1)$$

$$y_j = f(net_j) \quad (2)$$

其中, x_i 为前一层节点 i 的输出, w_{ij} 为连接权值, θ_j 为改变单元 j 活性的阈值, 范围一般设定为 $[-1, 1]$, f 采用 Sigmoid 激活函数, 即

$$f(a) = \frac{1}{1 + e^{-a}} \quad (3)$$

对某个训练样本而言, 其实际输出与期望输出的误差定义为:

$$E = \frac{1}{2} \sum_{k=1}^l (t_k - o_k)^2 \quad (4)$$

其中, l 为输出层单元数, t_k 为输出层第 k 个单元的期望输出, o_k 为输出层第 k 个单元的

实际输出。处理所有训练样本一次所花费的时间称为一个周期, 经过若干周期的训练后, 可以使误差 E 小于设定阈值。此时, 认为 BPNN 模型已达到收敛状态, 结束迭代学习过程。

3. 结果与分析

3.1 模型的构建

本文基于 BPNN 算法原理, 依托野外样方数据、DEM 数据、气象观测数据和卫星遥感数据, 利用 ArcGIS 平台强大的空间分析技术和 MATLAB 平台快速的矩阵运算能力, 进而实现锡林郭勒盟草原 NPP 的估算, 具体流程如下:

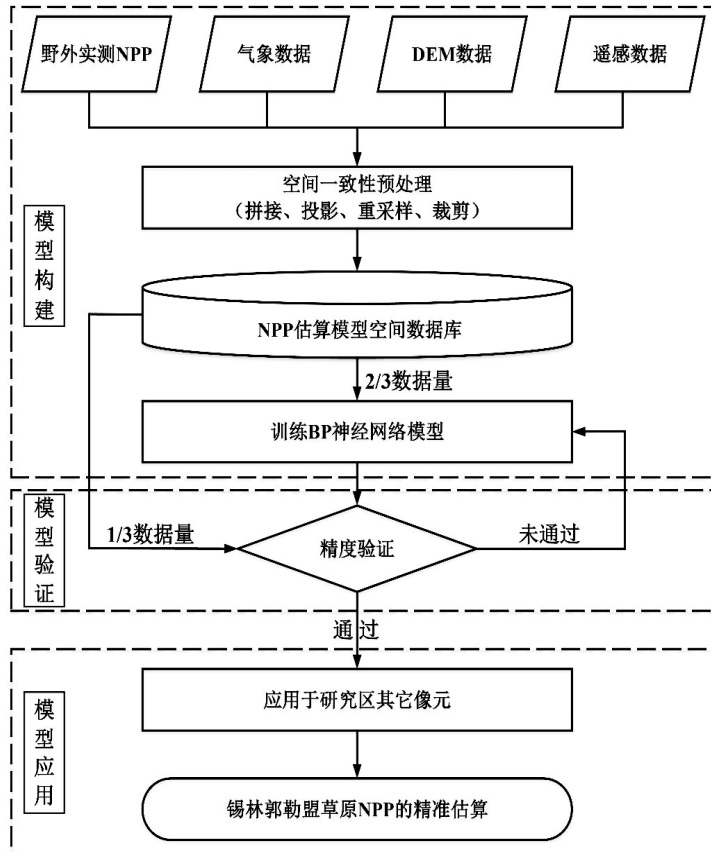


图3 基于 BPNN 算法的 NPP 估算模型构建流程

值得注意的是, 本文总计获得 510 个野外实测 NPP 数值, 随机选取其中的 2/3 作为 BPNN 训练样本数据集的输出, 利用 ArcGIS 软件选取像元对应位置上的 DEM、NDVI、月

平均气温、月降水量、LAI 与 FPAR 等要素作为训练网络模型的输入; 剩余的 1/3 用作训练后网络的测试, 检验基于 BPNN 算法所构建的 NPP 估算模型的泛化性和可移植性。

3.2 模型的验证

本文利用 MATLAB 软件作为计算平台，基于 340 个 NPP 实测值和相应环境因子值构建 BPNN 模型。在此基础上，本文以这 340 个参与建模的 NPP 模拟值和实测值的拟合度 (R^2) 和均方根误差 (Root Mean Squared Error, RMSE) 来对模型性能进行评定。

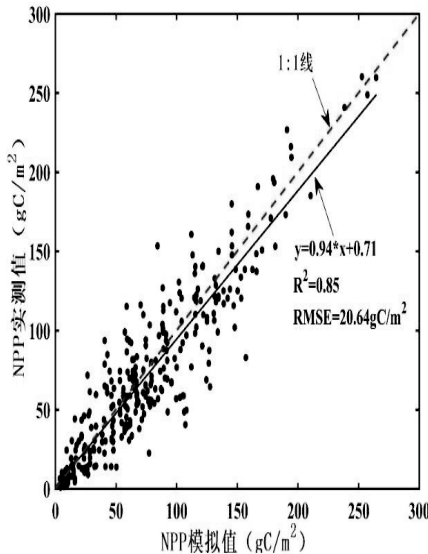


图 4 BPNN 模拟草原 NPP 性能

图 4 显示，训练后的 BPNN 模型具有优秀的 NPP 估算能力，其输入训练集的模拟值和实测值之间的拟合优度达 0.85，RMSE 为 20.64gC/m²。尽管如此，在将其应用到研究区其它像元之前，仍需进一步验证它的泛化性和可移植性。也就是说，需要用预留的测试样本数据集来对训练后 BPNN 的进一步进行精度检验。若通过，则将其应用于研究区剩余像元，否则，则应调整模型相应参数，继续训练，直至通过检验。本文依托预留的测试数据集，采用 RMSE 和决定系数作为诊断指标，对所构建 BPNN 模型的泛化性进行评价分析。如图 5 显示，模型的 RMSE 仅为 21.45gC/m²，拟合优度达 0.85，意味着训练后的模型可解释 NPP 变异的 85% 的信息量，估算的平均值略低于实测值，推测这可能是由于样方和像元尺寸不一引起的。总体而言，所训练的 BPNN 模型具备良好的可移植性和泛化性，可被用来估算研究区剩余像元的 NPP 值。

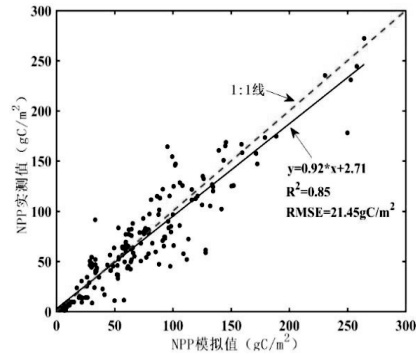


图 5 预留测试数据集的验证

3.3 模型的应用

本文借助 MATLAB 平台，利用训练后的 BPNN 模型估算了研究区剩余像元的 NPP 数值，实现了锡林郭勒盟草原 NPP 的精准估算，结果如图所示。

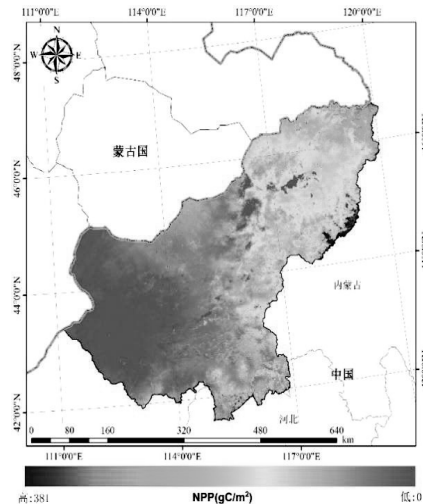


图 6 模型结果：锡林郭勒草原 NPP 的空间格局 (2013 年 8 月)

由图 6 可以看出，锡林郭勒盟草原于 2013 年 8 月长势旺盛期间的 NPP 空间分布存在着明显的地域性差异，其值自西向东由 0 逐渐上升至 300gC/m² 以上，有着较为明显的水平地带性规律。这种地带性规律符合研究区实际状况，进一步印证了所构建模型的合理性。研究区东部和东北部是大兴安岭西麓的延伸余脉，地形主要以低山丘陵为主，水热条件配合较好，草原植被类型以草甸草原为主，这些区域属于 NPP 高值区。研究区中部水热条件相对较好，

植被类型主要以典型草原为主, NPP 数值较高。研究区西部和西南部常年蒸发量大于降水量, 水分分配不均, 植被类型多以荒漠草原为主, 属于 NPP 低值区。

4. 结论

本文考虑到现有 NPP 估算模型的固有缺陷, 尝试从人工智能算法入手, 基于 BP 神经网络模型, 结合野外实测、气象观测、卫星遥感等多源数据, 进而实现了锡盟草原 NPP 的估算, 是对现有生态系统 NPP 估算模型的有力补充和有效尝试, 主要得到以下结论:

(1) 训练后的 BPNN 模型具有优秀的 NPP 估算能力, 其输入训练集的模拟值和实测值之间的拟合优度达 0.85, RMSE 为 20.64gC/m²。

(2) 训练后的 BPNN 模型具备良好的泛化性和可移植性。经预留的测试数据集验证, 模型的 RMSE 仅为 21.45gC/m², 拟合优度达 0.85, 可用来估算研究区剩余像元的 NPP。

(3) 模拟出的锡盟草地 NPP 空间格局和现实相符, 其值自西向东由 0 逐渐上升至 300gC/m² 以上, 有着较为明显的水平地带性规律。

Acknowledgements

This study is supported by the National Science Foundation of China (41571491 and 41371495); The China Special Fund for Meteorological Research in the Public Interest (GYHY201506001-6); The National Key Technology R&D Program of China under Grant (2013BAK05B01); The Fundamental Research Funds for the Central Universities of China (2412016KJ046).

致谢

本研究得到了国家自然科学基金项目 (41571491 和 41371495)、公益性行业 (气象) 科研重大专项 (GYHY201506001-6)、国家重点技术研发项目 (2013BAK05B01)、中央高校基本科研业务费专项资金资助 (2412016KJ046)

参考文献

- [1] 陈佐忠, 汪诗平, 王艳芬. 内蒙古典型草原生态系统定位研究最新进展. 植物学报, 2003, 20(4): 423-429.
- [2] Matsushita B, Tamura M. Integrating remotely sensed data with an ecosystem model to estimate net primary productivity in East Asia. Remote Sensing of Environment, 2001, 81(1): 58-66.
- [3] Field C B, Behrenfeld M J, Randerson J T, et al. Primary production of the biosphere: Integrating terrestrial and oceanic components. Science, 1998, 281(5374): 237-240.
- [4] Imhoff M L, Bounoua L, Defries R, et al. The consequences of urban land transformation on net primary productivity in the United States. Remote Sensing of Environment, 2004, 89(4): 434-443.
- [5] 朱文泉, 陈云浩, 徐丹, 等. 陆地植被净初级生产力计算模型研究进展. 生态学杂志, 2005, 24(3): 296-300.
- [6] 赵俊芳, 延晓冬, 朱玉洁. 陆地植被净初级生产力研究进展. 中国沙漠, 2007, 27(5): 780-786.
- [7] Marcek D. Risk scenes of managerial decision-making with incomplete information: an assessment in forecasting models based on statistical and neural networks approach. Journal of Risk Analysis and Crisis Response, 2013, 3(1): 13-21.
- [8] Yang G, Gu C, Huang Y, et al. BP neural network integration model research for hydraulic metal structure health diagnosing. International Journal of Computational Intelligence Systems, 2014, 7(6): 1148-1158.