

Hyper Spectral Estimation of Soil Water Content Based on Neural Network

Minghe Wang¹, Xiaoqian Wu¹, Jindong Li¹, Chengzhi Sun²

¹School of Surveying Science and Engineering, Shandong University of Science and Technology, Qingdao Shandong

²Junan County Land and Resources Bureau, Linyi Shandong

Email: 1246562871@qq.com

Received: Oct. 4th, 2018; accepted: Oct. 18th, 2018; published: Oct. 25th, 2018

Abstract

Hyperspectral remote sensing is widely used in rock minerals, soil, vegetation and other fields due to its extremely high spectral resolution. Using the spectral reflectance data of 84 soil samples in a certain place, by analyzing the relationship between reflectance spectrum and soil water content, the BP (back propagation) neural network inversion model was established by mathematically transforming the spectral data and using the inversion factors obtained by the correlation analysis method. The results show that in the prediction of soil water content, the model established by BP neural network is effective, and the first-order differential time model with spectral transformation to square root has higher precision.

Keywords

Hyper Spectral, Soil Moisture Content, Back Propagation Neural Network

基于神经网络的土壤含水量高光谱估测

王鸣鹤¹, 吴晓倩¹, 李晋东¹, 孙成志²

¹山东科技大学测绘科学与工程学院, 山东 青岛

²临沂市莒南县国土资源局, 山东 临沂

Email: 1246562871@qq.com

收稿日期: 2018年10月4日; 录用日期: 2018年10月18日; 发布日期: 2018年10月25日

摘 要

高光谱遥感凭借其极高的光谱分辨率在岩石矿物、土壤、植被等领域的应用越来越广泛。本文使用某地84

个土壤样本的光谱反射数据，通过分析反射率光谱与土壤含水量之间的关系，对光谱数据进行一系列的数学变换，利用相关分析法获取的反演因子建立BP神经网络反演模型。结果表明，在土壤含水量预测研究中，利用BP神经网络建立的模型有效，且光谱变换为平方根的一阶微分时模型精度较高。

关键词

高光谱，土壤含水量，BP神经网络

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

土壤含水量是陆地和大气能量交换过程中的重要因子，对水的运移、陆地表面蒸发、碳循环具有很强的控制作用，是生态、气候、水文、农业等领域衡量土壤干旱水平的重要指标，也是表征土壤肥力的重要指标。传统的土壤水分测定方法，如取土烘干法、张力计法、中子水分仪法等均是以点测量为基础，虽然精度高，但工作量大，范围有限，已经难以满足区域性土壤含水量监测的实际需要；土壤水分模型方法通过建立水分平衡方程求解土壤水分，可提供适时的土壤水分信息，但实验需要大量相关参数，估测误差较大。用遥感反演手段获取土壤含水量，具有范围大、时间短、分辨率高的特点，弥补了传统方法上的不足。其中高光谱遥感是利用多波段的电磁波从物体中获取有关数据，波段划分更窄，能获取更多的窄波段的光谱信息，产生一条完整而连续的地物光谱曲线[1] [2] [3] [4] [5]。本文采用了BP神经网络模型对土壤含水量进行了反演，利用多层神经网络可以寻找最优因子，建立的土壤水分含量的预测模型精度高，且为快速准确预测土壤含水量提供理论依据。

2. 土壤含水量 BP 神经网络估测模型

人工神经网络是基于生物神经网络原理由大量简单处理单元相互连接而形成的复杂网络系统，它具有较强的学习能力、抗干扰能力、容错能力[6]。人工神经元相当于一个多输入多输出的非线性阈值器件。其神经元的输出可以描述为：

$$y_k = f \left(\sum_{j=1}^n w_{kj} x_j - \theta_k \right) \quad (1)$$

式中： x_1, x_2, \dots, x_n 为输入信号， $w_{k1}, w_{k2}, \dots, w_{kj}$ 为神经元 k 的权值， θ_k 为阈值， f 为激活函数， y_k 为神经元 k 的输出。

人工神经网络中较为典型的神经网络模型有BP神经网络、Hopfield网络、CPN网络等，应用较多的是具有非线性传递函数神经元构成的前馈网络中采用误差反向传播算法作为其学习算法的前馈网络(Back Error Propagation, BP) [7]。BP网络包含了神经网络理论中最精华的部分，由于其结构简单、可塑性强，得到了广泛的应用。特别是它的数学意义明确、步骤分明的学习算法更使其具有广泛的应用背景。BP网络由输入层、输出层和若干隐含层组成，并且一般认为网络是严格分层的，即当且仅当两相邻层的节点才有可能互联。简单的三层BP网络的结构如图1所示。

BP神经网络计算模型如下

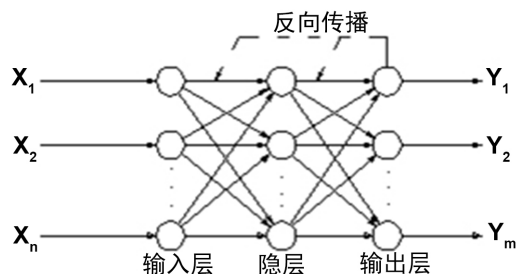


Figure 1. BP network structure diagram

图 1. BP 网络结构图

$$y = f^2 \left(\theta^2 f^1 \left(\theta^1 x + b^1 \right) + b^2 \right) \quad (2)$$

其中, y 为输出层, x 为输入层, f^1 为隐藏层传递函数, f^2 为输出层传递函数, b^1 为隐藏层偏差, b^2 为输出层偏差, θ^1 为输入层权重, θ^2 为隐藏层权重[8]。

3. 算例

以某地的土壤区作为研究对象, 其主要土地利用类型为耕地、林地、牧草地、未利用地(沙地、废弃地)、建设用地等, 选择地势较平坦、土壤裸露区作为样区, 并考虑各种土地利用类型和土壤类型, 每个样区内选择 4~5 个具有代表性的测点, 每个测点采集一个表层土(约 20 cm)土样, 共采集 84 个土样。

3.1. 光谱数据的平滑处理

光谱曲线的平滑处理是为了去除包含在信号内的少量噪声, 得到平稳的光谱波形。采用 9 点加权移动平均方法对光谱数据进行平滑处理。若给出了光谱曲线的 n 个测点的序列 $(R_i, i=1, 2, \dots, n)$, 则第 i 的值取包含前后各 9 个点的加权平均值[9], 即

$$R_i = 0.04R'_{i-4} + 0.08R'_{i-3} + 0.12R'_{i-2} + 0.16R'_{i-1} + 0.20R'_i + 0.16R'_{i+1} + 0.12R'_{i+2} + 0.08R'_{i+3} + 0.04R'_{i+4} \quad (4)$$

式中, R_i 表示 9 点平滑处理后的值。经平滑处理后形成新的光谱曲线的 n 个测点的序列 $(R_i, i=1, 2, \dots, n)$ 。

3.2. 光谱数据变换方法

为了寻找对有含水量敏感的光谱指标, 主要采用的变换方法如表 1 所示。

利用变换方法对光谱数据进行变换, 并与含水量进行相关分析, 得到反射率及其各种变换形式与含水量相关系数和包络线去除得到的各种参量与含水量的相关系数。如图 2 所示。由图可看出, 反射率 R 倒数的对数、对数的一阶微分、平方根的一阶微分与含水量的相关系数较大, 数据在峰值处大于 0.6, 表明其相关性较大, 适宜线性模型的建立, 而倒数、对数的倒数相关系数达到 0.6 的波段较多, 但是峰值较少, 不适合模型的建立。

光谱反演因子选取是挖掘光谱中蕴含的有效信息, 通过对各种变换方法认真比较, 最终选择在通过一阶微分变换、对数的一阶微分变换和平方根的一阶微分变换中选取反演因子。

采用单相关分析方法提取光谱反演因子, 顾及相邻波段间的光谱数据相关性较高, 在选择波段时尽量使其离散化。利用一阶微分变换、对数的一阶微分变换和平方根的一阶微分变换各选取了 5 个波段数据作为反演因子。

Table 1. Transformation method table
表 1. 本文使用的变换方法表

变换方法
1) 反射率的倒数 $1/R$
2) 反射率的对数 $\ln R$
3) 反射率对数的倒数 $1/\ln R$
4) 反射率的一阶微分 R'
5) 反射率对数的一阶微分 $(\ln R)'$
6) 平方根的一阶微分 $(\sqrt{R})'$

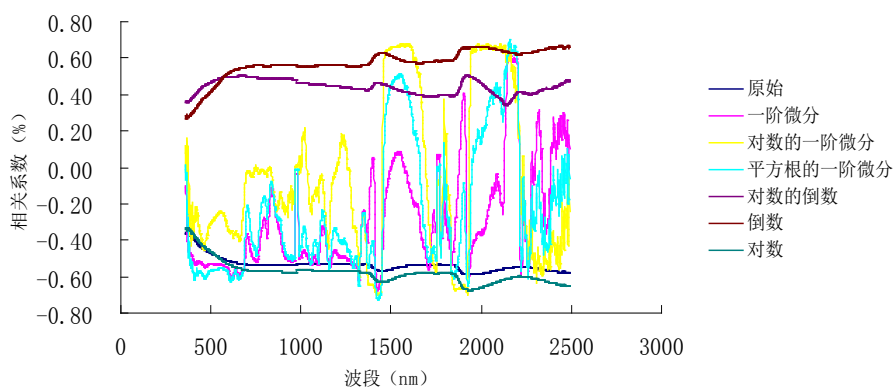


Figure 2. Transformation comprehensive map
图 2. 变换综合图

3.3. 数据处理过程

异常样本(落在总体之外的数据点)的产生是人为测量过程中带来的误差,包括土壤属性的化学测定和土壤光谱数据的测量产生的异常数据,异常值的存在可能会影响模型的准确性建模之前要先对异常样本进行删除。李希灿等提出剔除异常样本的方法是,首先对土壤含水量从小到大进行排序,对应的反演指标也进行排序,分别绘制排序后的土壤含水量和反演指标的分布曲线,然后,在分布曲线的两端寻找异常样本[10][11]。

一共剔除了 3 个异常样本,在剔除异常样本之后 R2 明显增大。把剔除异常样本后剩余的 81 个样本分成两组,一组 66 个样本为建模样本,另一组 15 个样本为检验样本。

利用变换后的光谱数据与土壤含水量相关关系曲线选取的反演因子,光谱数据与土壤含水量建立 BP 神经网络模型。将 BP 神经网络的隐含层数设为 1~3 层,隐节点数设在 4~5 个之间,隐含层传递函数采用 logsig 函数或者 tansig 函数,输出层传递函数采用 purelin 函数,网络训练函数采用 traincgh 函数或者 trainlm 函数。基于特征吸收波段的分析建模时,将 66 个将反射率变换之后的数据按照 BP 神经网络输入向量,含水量实测值作为输出向量,通过网络中隐含层节点数、隐含层数这 2 个参数的不同组合,得到一系列的模型,然后分析误差从而得出最佳模型。

4. 结果分析

将 81 个样本数据分为两组,66 个用于网络训练、15 个用于模型检验。基于 BP 神经网络法建立了多个土壤含水量的高光谱反演模型,如表 2~4 所示,结果表明,用真实含水量与预测值做差,在剔除差

Table 2. Accuracy comparison between first-order differential models of square roots**表 2.** 平方根的一阶微分模型之间的精度对比

变换方法	层数及节点数	平均相对误差	R ²
平方根的一阶微分	1-5	11.949	0.922
	1-4	13.030	0.912
	2-5-5	13.238	0.911
	2-4-5	12.991	0.924
	3-5-5-5	14.107	0.917
	3-4-4-5	14.653	0.919

Table 3. Accuracy comparison between first-order differential models**表 3.** 一阶微分模型之间的精度对比

变换方法	层数及节点数	平均相对误差	R ²
一阶微分	1-5	15.401	0.852
	1-4	58.981	0.071
	2-5-4	18.501	0.791
	2-4-5	17.107	0.804
	3-4-5-5	16.980	0.831
	3-5-4-4	18.926	0.779

Table 4. Accuracy comparison between first-order differential models of logarithms**表 4.** 对数的一阶微分模型之间的精度对比

变换方法	层数及节点数	平均相对误差	R ²
对数的一阶微分	1-5	24.137	0.819
	1-4	22.003	0.726
	2-4-5	22.556	0.734
	2-5-5	21.151	0.788
	3-4-5-5	22.421	0.747
	3-5-4-5	23.823	0.706

距较大的样本后, 取含水量与预测值差值平方和, 平方根的一阶微分神经网络模型, 平均相对误差值最小, 为 11.949, 决定系数最大, 为 0.922。

对前述的反演因子分别建立土壤含水量的多元线性回归估测模型、BP 神经网络估测模型, 精度分析如图 3, 图 4 所示, 详细数据如表 5 和表 6 所示。BP 神经网络的决定系数为 0.922, 多元线性回归模型的决定系数为 0.869, 通过比较可以得出 BP 神经网络模型预测值与实测值较为接近。

5. 结束语

研究采集了某地 0~20 cm 表层土 84 个土壤样品, 用标准化学方法测得土壤样品的水含量, 并在实验室内测定了土壤样品的反射率光谱。运用 BP 神经网络法建立了含水量的多个反演模型, 并对模型进行

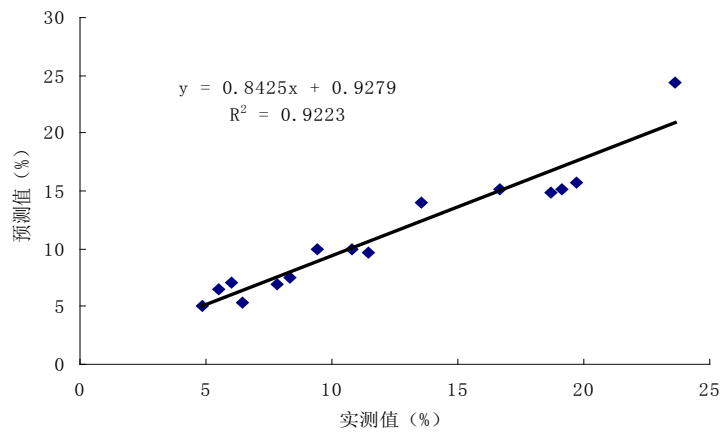


Figure 3. BP neural network test sample accuracy analysis chart
图 3. BP 神经网络检验样本精度分析图

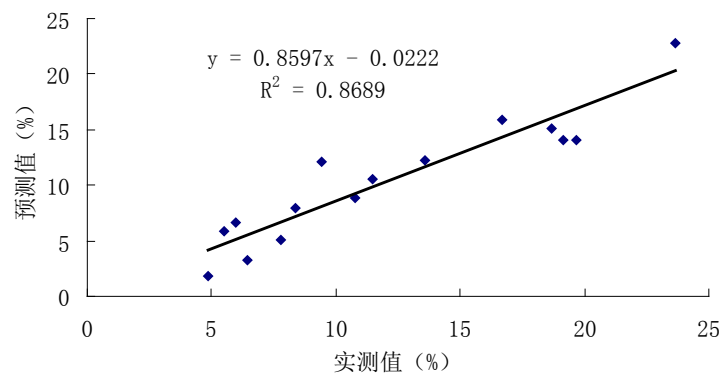


Figure 4. Multiple linear regression test samples accuracy analysis chart
图 4. 多元线性回归检验样本精度分析

Table 5. BP neural network test sample accuracy analysis table
表 5. BP 神经网络检验样本精度分析表

样本	实测值	预测值	相对误差
hs60	4.870	4.988	3.064
hs30	5.480	6.542	19.372
hs29	6.000	7.081	18.017
hs46	6.453	5.270	-18.338
hs56	7.810	6.991	-10.490
hs37	8.367	7.477	-10.636
hs16	9.400	10.000	6.378
hs17	10.790	9.980	-7.507
hs71	11.450	9.690	-15.373
hs21	13.580	13.943	2.674
hs73	16.690	15.161	-9.159
hs2	19.160	15.125	-15.170
hs74	18.680	14.914	-20.162
hs76	19.690	15.779	-19.864
hs24	23.640	24.355	3.026

Table 6. Linear regression test sample accuracy analysis table
表 6. 线性回归检验样本精度分析表

样本	实测值	预测值	相对误差
hs60	4.870	1.772	-63.620
hs30	5.480	5.796	5.775
hs29	6.000	6.610	10.163
hs46	6.453	3.272	-49.303
hs56	7.810	5.103	-34.663
hs37	8.367	7.958	-4.889
hs16	9.400	12.102	28.742
hs17	10.790	8.845	-18.029
hs71	11.450	10.510	-8.213
hs21	13.580	12.197	-10.181
hs73	16.690	15.830	-5.154
hs74	19.160	14.116	-26.328
hs2	18.680	15.164	-18.821
hs76	19.690	14.097	-28.406
hs24	23.640	22.822	-3.461

了验证。土壤水含量与土壤光谱之间的关系比较复杂，具有一定的不确定性。对样本的光谱反射率数据与土壤含水量进行分析，利用平方根的一阶微分得到的模型精度最高，其中平均相对误差值为 11.949%，决定系数为 0.922。由于土壤含水量与高光谱反射率之间并不存在严格意义上的线性关系，采用神经网络的方法进行反演建模，虽然是非线性分析方法，但是要更准确地揭示土壤含水量与高光谱反射率之间的内在联系，有待进一步实验与研究。

参考文献

- [1] 尚璇, 李西灿, 徐邮邮, 刘莎莎. 土壤水与有机质对高光谱的作用及交互作用规律[J]. 中国农业科学, 2017(8): 1465-1475.
- [2] 谢伯承, 薛绪掌, 刘伟东, 等. 基于包络线法对土壤光谱特征的提取及其分析[J]. 土壤学报, 2005, 42(1): 171-175.
- [3] 郑曼迪, 熊黑钢, 乔娟峰, 等. 基于高光谱的不同人类干扰程度下荒漠土壤有机质含量估算模型[J]. 干旱区地理, 2018, 41(2): 384-392.
- [4] 丁国香. 基于神经网络的土壤有机质及全铁含量的高光谱反演研究[D]: [硕士学位论文]. 南京: 南京信息工程大学, 2008.
- [5] 程朋根, 吴剑. 土壤有机质高光谱遥感和地统计定量预测[J]. 农业工程学报, 2009, 25(3): 142-147.
- [6] 何挺, 王静. 土壤有机质光谱特征研究[J]. 武汉大学学报·信息科学版, 2006, 11(31): 975-979.
- [7] 贺军亮, 蒋建军. 土壤有机质含量的高光谱特性及其反演[J]. 中国农业科学, 2007, 40(3): 638-643.
- [8] 唐启义, 冯明光. DPS 数据处理系统 - 实验设计统计分析 & 数据挖掘[M]. 北京: 科学出版社, 2007: 625-676.
- [9] He, J.L., Jiang, J.J., Zhou, S.L., et al. (2007) The Hyperspectral Characteristics and Retrieval of Soil Organic Matter Content. *Scientia Agricultura Sinica*.
- [10] 戴昌达. 中国主要土壤光谱反射特性分类与数据处理的初步研究[M]. 遥感文集. 北京: 北京科学出版社, 1981: 315-323.
- [11] 卢艳丽, 自由路. 基于主成分回归分析的土壤有机质高光谱预测与模型验证[J]. 植物营养与肥料学报, 2008, 14(6): 1076-1082.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2329-549X，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：gst@hanspub.org