

基于遥感图像的土壤含水量信息提取研究

崔丽霞, 王 蕾

(唐山学院 智能与信息工程学院, 河北 唐山 063000)

摘要: 土壤含水量是监测农作物生长状况的关键参数, 近年来国内外学者在利用高光谱遥感数据监测土壤含水量技术方面进行了大量的研究。文章在分析相关研究进展的基础上, 以唐山地区为研究区域, 重点选取乐亭县、滦南县为研究对象, 将 Landsat 8 遥感数据所获取的植被指数和地表温度作为反演土壤含水量的两个基本参数, 再结合实测的地表温度数据, 利用 BP 神经网络建立模型, 从而实现了对该研究区域内土壤含水量的提取。通过研究发现, 利用该方法能够提取土壤含水量, 提取结果准确性较高。

关键词: 唐山地区; 土壤含水量; 神经网络; 遥感图像

中图分类号: TP79; S152.7 **文献标志码:** A **文章编号:** 1672-349X(2017)06-0021-04

DOI: 10.16160/j.cnki.tsxyxb.2017.06.005

A Study on Determination of Soil Moisture Content Based on Remote Sensing Images

CUI Li-xia, WANG Lei

(College of Intelligence and Information Engineering, Tangshan University, Tangshan 063000, China)

Abstract: Soil moisture content is a key parameter to monitor crop growth. In recent years, a great deal of research has been done on the monitoring of soil water content by using hyperspectral remote sensing data both at home and abroad. Based on previous research on the subject, and with Tangshan as the area to be studied, Leting County, Luannan County as the research subjects, and the vegetation index and surface temperature measured by Landsat 8 as two basic parameters for the determination of soil moisture content, the authors of this paper have developed a method to determine the soil moisture content of this area after establishing a neural network model with BP and taking into consideration the real surface temperature measured. Further research shows that soil moisture content determined with this method is accurate.

Key Words: Tangshan; soil moisture content; neural network; remote sensing image

土壤水在全球水循环过程中起着十分重要的作用, 土壤水分的多少是判断农田作物是否

干旱的重要指标。根据遥感的基础理论, 可见光、近红外、热红外和微波波段都能反映裸土或

基金项目: 河北省科技厅计划项目(15227014)

作者简介: 崔丽霞(1987—), 女, 河北唐山人, 讲师, 硕士, 主要从事图像处理研究。

植被覆盖条件下土壤的水分信息。在可见光波段,土壤的反射率随波长的增加而增加,随着土壤水分的增加而减少。热红外波段可以探测土壤及植被冠层的温度,然后利用神经网络建立模型可估算土壤水分^[1]。由于干土壤与液体水的介电常数有显著区别,因此也可以利用微波波段来监测土壤水分。目前的土壤含水量监测手段还是以遥感技术中广为使用的可见光与热红外波段结合为主。本文将利用 Landsat 8 遥感数据所获得的植被指数和地表温度作为反演土壤含水量的两个基本参数,再结合地面实测的土壤含水量数据,利用 BP 神经网络建立模型,以实现土壤含水量的遥感反演。

1 研究区域及图像数据介绍

1.1 研究区域概况

唐山地区位于东经 117°31′到 119°19′之间,北纬 38°55′到 40°28′之间。地处河北省东部,华北平原东北部,地理位置较好,东面与母亲河上都河相望,西边是繁华的天津市,南边靠近渤海湾,北边与燕山相伴,和承德市之间有长城之隔。唐山东西向长约 135 km,南北宽度约 150 km,陆地总面积达到 13 472 km²。目前,唐山市辖 7 个市辖区、5 个县、2 个县级市。图 1 为唐山地区行政图。



图 1 唐山地区行政地图

1.2 气候

唐山地区在中国的北方,四季比较分明,属温带半湿润大陆性气候。冬季气候寒冷,雨雪稀少;春季降水不多,大风频繁;夏季炎热,雨量集中;秋季冷暖变化显著,少雨。全市多年平均气温 12.5℃,平均降水量为 644.4 mm。主要有滦河、冀东沿海、蓟运河三大水系。

1.3 地貌

唐山全市地势北高南低,自西北向东南倾斜,山地、丘陵和平原呈明显的地貌阶梯。土壤分布随地形由山区向平原依次变化,北部山区为山地棕色森林土及褐土,丘陵地区多为淋溶褐土,平原地区多为草甸褐土。其中研究对象滦南县无山地丘陵,属平原地貌,以洪积冲积平原、冲积平原、海积平原为主;乐亭县地处华北断块内东北部,境地内部主要为中生界、新生界沉积层,地面为燕山褶皱带南缘、渤海北岸滨海平原,其平原由滦河冲击扇和滨海平原两部分所组成。

1.4 数据源

本研究采用的数据源是 Landsat 8 影像数据,覆盖整个研究区域,影像云量较少,图像的分辨率为 30 m。获取影像日期为 2015 年 8 月 15 日。

2 研究原理及方法

2.1 归一化植被指数(NDVI)和地表温度

归一化植被指数(NDVI)对植被监测的灵敏度较高,当植被在生长过程中一直处于缺水状态时,植被的生长就受到影响,对应的植被指数将会降低。通常利用波段计算来定量确定 NDVI。

$$NDVI = (NIR - R) / (NIR + R),$$

其中, NIR , R 为遥感数据不同波段的反射率值。

NDVI 的值在 -1 到 1 之间,负值表示地物为水、雪等或有云遮盖,对可见光高反射;0 表示为岩石或裸土等, NIR 和 R 近似相等;正值,表示有植被覆盖,且随覆盖度增大而增大。

Landsat 8 数据的植被指数计算为:

$$NDVI = (B5 - B4) / (B5 + B4),$$

其中 $B4$, $B5$ 为 Landsat 8 数据中波段 4 和波段 5 的反射率值。

地表温度反演的研究大多采用地表温度劈窗算法,但是该算法对于 Landsat 8 影像精确度并不高,这是由于 TIRS 11 热红外波段的定标参数不精确,使得已提出的 2 个劈窗算法的误差都较大。因此针对 landsat 8 热红外波段地表温度

(LST)反演本研究采用了单通道算法(SC)^[2]。

2.2 实测数据

利用 Landsat 8 影像,分辨出农田地块,找出其中影像单元比较均一、面积约为 1 km² 的区域,然后在其中选择实测人员较容易到达的样点若干个。

利用 GPS 定位,以研究区域内乐亭县、滦南县的平原地区为主要对象进行采样,最后筛选出 50 个样点进行实地监测。

2.3 BP 神经网络模型

神经网络在遥感图像处理方面有着广泛的应用,如在图像分类、图像压缩、图像复原与重建、图像边缘检测、图像去噪、图像纹理分析等方面,均显示出识别率高、学习力强和高容错特性等优势。神经网络中的 BP (Back-Propagation Network) 网络是一个单向传播的多层前馈网络,该网络包括输入层、隐含层和输出层,同层节点之间不连接,每层节点的输出只影响下一层节点的输入^[3]。利用 Matlab 建立一个 BP 神经网络,需要计算输入矢量、输出矢量、误差矢量以及误差平方和,如图 2 所示。

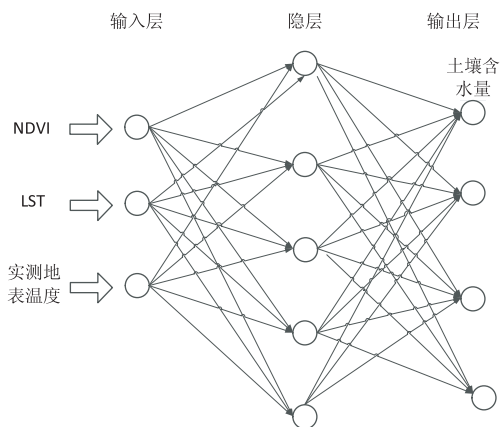


图 2 本模型中的 BP 神经网络结构

当训练矢量的误差平方和小于设定的误差目标时训练会被停止;当训练矢量的误差平方和不小于设定的误差目标时,在输出层会计算误差的变化,从而利用反向传播的方法调整权值。反复操作此过程,直到停止训练。在建立神经网络模型时,根据实际情况给出了 P 输入矢量和 T 目标矢量,并且确定了选择神经网络

的结构,例如有几层神经网络、每一层被训练的神经元数、每一层的激活函数等^[3]。

BP 神经网络的训练步骤^[4]:

(1)用小的随机数对每一层的权值 W 和偏差 B 初始化,用来保证网络不被大的加权输入饱和,并进行以下参数的设定或初始化:

I. 设定期望误差最小值 mse;

II. 设定最大循环次数;

III. 设定修正权值的学习速率 lr ,一般情况下 $lr=0.01\sim0.7$ 。

(2)用函数 newcf.m 或 newff.m 来建立一个多层的前向神经网络。

(3)用 $net=train(net,P,T)$ 来训练。

BP 神经网络算法通常由两部分组成:一部分是信息的正向传递,另一部分是误差的反向传播。在正向传播过程中,输入信息从输入层经隐含层逐层计算并传递给输出层,下一层神经元的状态只会被上一层神经元影响。如果期望值没有被输出层输出,则计算误差的变化值。在反向传播过程中,通过网络将误差信号沿原来的连接通路反传回来修正各层神经元的权值,直至达到期望目标。

在本研究中输入层一共建立了 3 个神经元,分别是 NDVI, LST 和实测地表温度,输出层建立了 4 个土壤含水量神经元。当前网络层的权值为 1,循环间隔为 100,训练目标最小误差设置为 0.05,动量因子为 0.9,训练次数为 1 000,训练最小误差为 e^{-3} ^[4]。输出结果中设置蓝色表示含水量为 80%~100%,代表此处为含水量较高地物;红色,表示含水量为 60%~80%,代表此地物含水量为中上等;墨橘黄色表示含水量为 40%~60%,代表此处地物含水量为中等;绿色表示含水量为 20%~40%,含水量为中下等,黄色表示含水量为 0~20%,代表该地物含水量较低。训练结果如图 3 所示。

从图 3 中可以清楚地看出河流湖泊、城区、山体植被、农田等被区分开来。其中蓝色区域为河流湖泊或是海洋,含水量较高;红色为农田,含水量在中上等;橘黄色为植被和农田结合处,受部分植被的遮挡,该处的含水量在 40%

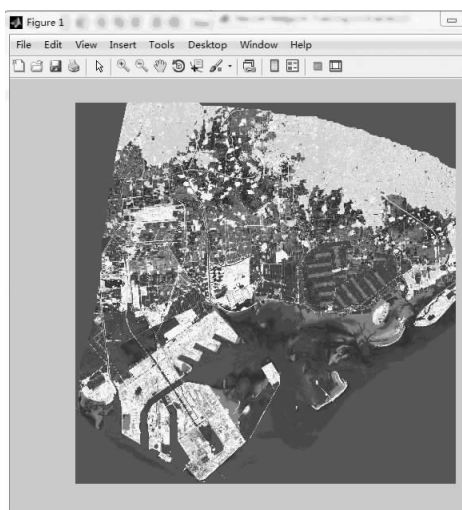


图 3 土壤含水量训练结果

~60%，为中等；绿色区域为植被茂盛地区，该处含水量在中下等；黄色为建筑区域，该区域的土壤含水量较低，有些地区甚至含水量为 0。

为了进一步验证该结果的准确性，本研究中采用 ENVI 软件对同一图像进行了土壤含水量的提取，图 4 为提取结果图。通过图 3 与图 4 对比可以看出，最终得到的土壤含水量图像基本符合实际情况，但是部分地区也存在一定的误差。由于 BP 神经网络模型在建立过程中应用的数据更多，因此相对于其他的遥感图像土壤含水量的提取结果更准确些。神经网络模型也存在一定的局限性，在建立模型时遥感图像越多研究结果越准确，本文在研究过程中受到技术的限制，只采用了一幅图像中的多个采样地进行了研究，使研究结果仍然存在一定的误差。

3 结论

(1) 利用 Landsat 8 图像可以获得植被指数，能够反演出地表温度，再结合实地测量的地表温度建立 BP 神经网络模型，从而能够估算出土壤含水量。

(2) 建立 BP 神经网络模型时，采样样本的多少与研究结果的准确性有关，在本研究中样本数据使用较少，对结果的准确性有一定的影响。



图 4 ENVI 提取土壤含水量结果

(3) 由于遥感探测地区与实际测量地表温度地区范围不大相同，需要对部分未采样地区的地表温度进行估算，而且还需要考虑采样过程中降水及部分农田灌溉对土壤含水量的影响，天气有云时对探测温度的影响，等等。这些因素都会使建立的模型产生误差，影响计算的结果，因此，本研究有待于进一步深化。

参考文献：

- [1] 赵立军. 基于 MODIS 数据的北京地区土壤含水量遥感信息模型研究[D]. 北京：中国农业大学，2004.
- [2] 崔丽霞，王蕾. 基于 Landsat 8 遥感图像的土壤含水量提取研究[J]. 唐山学院学报，2016，29(6)：48-50.
- [3] 葛超，王蕾，曹秀爽. MATLAB 技术大全[M]. 北京：人民邮电出版社，2014：782-792.
- [4] 曹兆伟. 基于人工神经网络的海岛遥感影像地物分类研究[D]. 上海：上海海洋大学，2016.

(责任编辑：夏玉玲)