

# 基于遗传优化神经网络的多源遥感数据反演土壤水分

关韵桐<sup>1</sup>, 李金平<sup>1,2</sup>

(1. 云南师范大学 旅游与地理科学学院, 云南 昆明 650500;  
2. 西部资源环境地理信息技术教育部工程研究中心, 云南 昆明 650500)

**摘要:** 为快速反演较高精度土壤水分, 提出用遗传算法优化后的神经网络辅以多源遥感数据的方法进行地表土壤水分反演。首先建立4层神经网络并用遗传算法优化此网络, 之后以雷达数据不同极化(VV、VH、VH/VV)的后向散射系数、雷达入射角、光学数据的归一化植被指数(NDVI)、以及高程数据作为网络的输入, 土壤水分数据为输出, 对网络进行训练与仿真, 再运用地表实际测量数据与反演数据做对比验证。结果表明: 反演结果与实际测量数据相关性良好,  $R^2$  可达0.79。采用遗传算法对神经网络优化的土壤水分反演方法可行, 且添加光学数据等辅助数据后土壤水分反演效果更优, 为多源遥感土壤水分的协同反演研究提供新思路。

**关键词:** 土壤水分; 多源遥感数据; 遗传优化算法; 神经网络; 反演土壤水分

中图分类号: S152.7; P237 文献标识码: A 文章编号: 1672-643X(2019)02-0252-05

## Soil moisture inversion based on genetic optimization neural network and multi-source remote sensing data

GUAN Yuntong<sup>1</sup>, LI Jinping<sup>1,2</sup>

(1. School of Tourism and Geographic Science, Yunnan Normal University, Kunming 650500, China; 2. GIS Technology Engineering Research Center for West-China Resources and Environment of Educational Ministry, Kunming 650500, China)

**Abstract:** In order to quickly invert higher precision soil moisture, this paper used the genetic algorithm optimized neural network and multi-source remote sensing data to invert the surface soil moisture. First, a four-layer neural network was established and the network was optimized by genetic algorithm. Then, we used the backscattering coefficient of radar data with different polarizations (VV, VH, VH/VV), radar incident angle, normalized vegetation index (NDVI) of optical data, and elevation data as input to neural network, the soil moisture data as the output to train and simulate the network. Finally, the inversion data was verified by the actual measured data of the surface. The results showed that the correlation between the inversion data and the measured data is high ( $R^2 = 0.79$ ). The method of genetic algorithm optimized neural network is feasible to calculate soil moisture content, and the soil water inversion effect is better after adding auxiliary data such as optical data. This study provides a new idea for the collaborative inversion of soil moisture in the multi-source remote sensing.

**Key words:** soil moisture; multi-source remote sensing data; genetic optimization algorithm; neural network; inversion

## 1 研究背景

土壤含水量影响着与地球系统有效运行有关的各种过程, 它对土壤物理性质有很强的控制作用, 影响土壤的生长、侵蚀、径流和渗透, 同时土壤生物地

球化学影响微生物的活动和  $\text{CO}_2$  的产生<sup>[1]</sup>。土壤含水量监测在气象学、水文学和农业等各种学科中起着不可或缺的作用<sup>[2]</sup>。因此对土壤含水量的研究具有重要意义。测量土壤含水量的方法大致分为3类: 取样测量、模型模拟、遥感监测<sup>[3]</sup>。取样测量

收稿日期: 2018-11-12; 修回日期: 2018-12-28

基金项目: 国家自然科学基金项目(41461087)

作者简介: 关韵桐(1994-), 女, 辽宁辽阳人, 硕士研究生, 主要研究方向为遥感信息提取与分析。

通讯作者: 李金平(1964-), 男, 河北秦皇岛人, 教授, 博士, 主要从事地图制图学与地理信息工程教学科研工作。

方法操作简单,测量误差较小,但无法大规模测量且效率不高;模型模拟方法在土壤水分监测中具有一定研究意义,但土壤水分变化是复杂的动态循环过程,影响土壤水分变化的因素较多,无法完全用纯理论公式进行反演;遥感检测方法具备全天候,可大范围观测的优势,为土壤水分监测提供新思路。在主动微波遥感测量中,雷达后向散射得强度很大程度上取决于目标物的介电属性,同时土壤水分与土壤介电属性之间具有很强的相关性,因此 SAR 数据特别适用于土壤水分估算。主要方法是建立雷达后向散射与土壤含水量间的半经验/经验模型,通过实验数据确定函数关系式<sup>[4]</sup>,孔金玲等<sup>[5]</sup>提出在稀疏植被条件下,利用 C 波段雷达数据基于高级积分方程模型(AIEM)构建  $S^3/l$  的组合粗糙度参数建立土壤水分反演模型;余凡等<sup>[6]</sup>利用多极化 ASAR 数据构建粗糙度参数  $S^3/l^2$ ,在不测量地表粗糙度的条件下获取土壤水分。然而半经验/经验模型存在一定限制,因此非线性的理论模型应用较广,在非线性问题的处理中,人工神经网络(ANN)、粒子群算法等智能优化算法效果更好。韩玲等<sup>[7]</sup>利用粒子群算法求解雷达后向散射与土壤水分的非线性方程组,提高反演精度。Baghdadi 等<sup>[8]</sup>使用多层神经网络(MLP)在 IEM 模型的基础上进行土壤地表参数反演,得到了较好的反演结果;Santi 等<sup>[9]</sup>使用 ANN 技术基于 ASAR 数据反演土壤水分。

雷达后向散射强度易受地表粗糙度与植被的影响,为减少此类影响因素的干扰,本文协同利用 C 波段双极化 Envisat-ASAR 雷达数据及 Landsat TM 光学数据、DEM 数据探讨遗传算法优化神经网络的土壤水分反演方法,即构建一个神经网络,运用遗传算法对该网络的阈值权值进行优化,将遥感数据及土壤水分数据分别作为输入输出带入优化后的神经网络中进行网络训练,通过与实地测量数据的对比验证该方法在土壤水分反演中的可行性。传统微波遥感土壤水分反演方法多为建立经验/半经验模型,然而模型适用区域及范围会存在一定限制,利用遗传优化神经网络的方法解决后向散射系数与土壤水分之间的非线性关系,可在一定程度上提高反演的精度。

## 2 研究区概况及数据处理

### 2.1 研究区概况

研究区位于甘肃省张掖市临泽县,中心地理位置为北纬 39.261 9°、东经 100.291 9°,海拔 1 385

m,研究区为平原,地势平坦,属大陆性温带干旱气候,年均降水量 121.5 mm。研究区属盐渍化草甸地区,土地类型以裸土为主,其中沙土含量占 16.7%,黏土含量占 8.5%。实测数据由“黑河综合遥感联合试验”于 2008 年 7 月 11 日临泽草地加密观测区测得,测量期间无降雨、无人工灌溉,因此可忽略降水及人为活动对土壤水分的影响。本文选用试验区内大小为 120 m × 120 m 的 D、E 两个样方,其中 D 样方植被覆盖类型为大麦,E 样方植被覆盖类型为苜蓿,采集两个样方内共计 98 个实测点的表层 0~5 cm 土壤体积含水量数据以及经纬度坐标、土壤容重、土壤温度等数据。

### 2.2 遥感数据

2.2.1 Envisat-ASAR 数据 根据土壤水分实测数据选取同日期(2008 年 7 月 11 日)过境的 Alternating Polarization 工作模式、IS4 入射角下的 Level 1B 级 Envisat-ASAR 数据,具有 VV、VH 两种极化方式,空间分辨率为 30 m。雷达数据预处理采用欧空局 SNAP 软件,对 ASAR 数据进行辐射定标、滤波、地形校正等处理。辐射定标处理之后,将雷达数据的 DN 值转化为后向散射系数值(dB),采用 5 × 5 窗口的 Lee low-pass filter 滤波器进行滤波处理去除噪声干扰,使边缘特征得以保留同时平均图像<sup>[10]</sup>。结合 30 m 的 SRTM DEM 数据进行地形校正,根据各样点经纬度坐标读出 ASAR 数据的后向散射系数及入射角度。

2.2.2 Landsat TM 数据 微波与植被和土壤之间存在相互作用,导致植被覆盖对土壤水分反演造成影响。植被量,介电特性和分布模式影响微波后向散射对土壤水分的敏感性。选取相近时间(2008 年 7 月 23 日)无云 Landsat TM 数据,在 ENVI5.1 软件下对数据进行辐射定标、大气校正等预处理操作,选取 B3(Red)、B4(NIR)波段进行归一化植被指数(NDVI)计算:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

根据各样点经纬度坐标提取各点对应的 NDVI 值。

## 3 研究方法

大部分基于 SAR 数据的土壤水分反演是假设雷达后向散射系数与土壤含水量之间存在一种线性关系,也有一些研究探索两者之间的非线性关系<sup>[11]</sup>。本文基于 BP 神经网络探索两者间的非线性关系。神经网络因其具有较强的自学能力而广泛应

用于各个领域,但其本身存在易陷入局部最小值、依赖设计结构等缺陷使得有时无法寻找到全局最优值。遗传算法虽不具备自学能力,但其具备可收敛至寻找到全局最优解的能力,因此,运用遗传算法优化神经网络能够改善神经网络本身存在的缺点,提高学习能力<sup>[12]</sup>。

运用 GA 对 BP 网络进行优化的详细步骤如下:

(1) 确定网络结构:建立一个 4 层的 BP 网络,确定各层神经元个数以及传递函数、BP 网络的具体参数、训练次数等。图 1 为神经网络模型结构图。

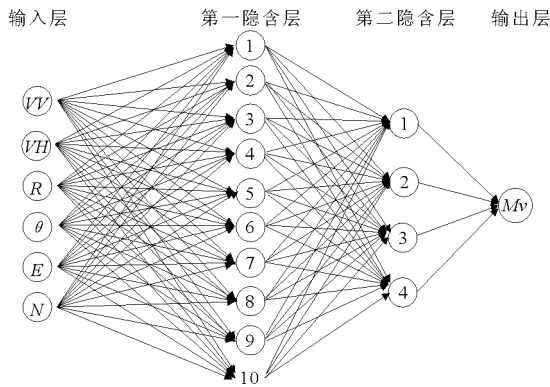


图 1 神经网络模型结构

(2) 编码:确定染色体长度即编码长度  $S$ :

$S = R \cdot S_1 + S_1 \cdot S_2 + S_2 \cdot S_3 + S_1 + S_2 + S_3$  (2)  
式中:  $R$ 、 $S_1$ 、 $S_2$ 、 $S_3$  分别为输入层、第一隐含层、第二隐含层、输出层的神经元个数。

(3) 初始种群选取:选取较大的种群数量更易于找到全局最优解,但同时也会增加迭代次数进而增加训练时间,故一般取 20 ~ 100 为宜<sup>[13]</sup>。经过多次试验,本文种群数量确定为 70。

(4) 选择:在当前群体中选择出最优异的个体,把最优个体作为父体直接遗传或由交叉配对发展出新的最优个体并遗传到下一代。

(5) 交叉运算:交叉运算提高了遗传算法的搜索能力,是遗传算法区别于其他进化算法的重要运算过程,具有重要作用。采用实数交叉法,  $a_k$  为第  $k$  个染色体,  $a_l$  为第  $l$  个染色体,在  $j$  位的交叉过程:

$$\begin{cases} a_{kj} = (1 - b)a_{kj} + ba_{lj} \\ a_{lj} = (1 - b)a_{lj} + ba_{kj} \end{cases} \quad (3)$$

(6) 变异操作:种群中随机选取个体通过变异产生新个体<sup>[14]</sup>,类似于生物情况,在遗传算法中极少发生变异,故通常选取较小的变异概率。

(7) 达到终止条件时,停止优化,否则继续重复以上步骤,得到神经网络的最优权值和阈值。

(8) 训练网络,当网络精度达到要求后进行仿

真预测,得到预测结果。

GA - BP 的流程如图 2:

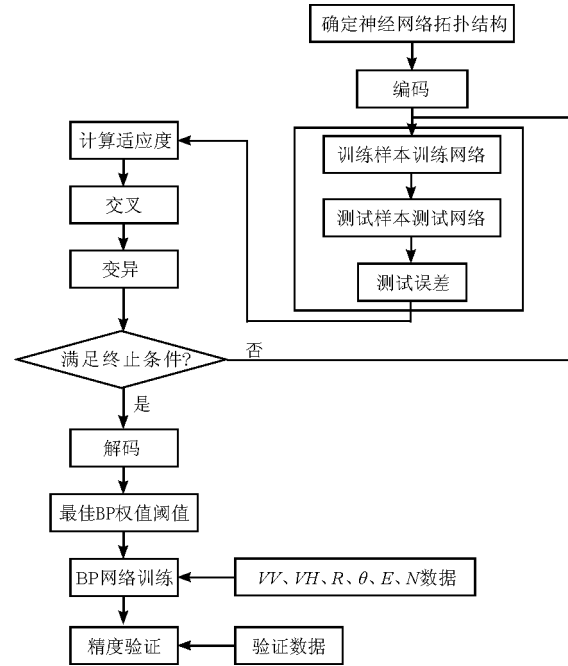


图 2 GA - BP 流程

## 4 土壤水分反演

### 4.1 反演试验

采用 ANN 方法分析 ASAR 及 Landsat TM 数据协同反演土壤水分,借助 MATLAB 2014a 软件,以 ASAR 数据的  $VV$ 、 $VH$ 、 $VH/VV$  极化下后向散射系数、入射角度、SRTM DEM (30 m) 数据的高程信息、Landsat TM 数据的 NDVI 值作为网络的输入,土壤体积含水量数据则为输出,设计编写 BP 神经网络。其中  $VH/VV$  的后向散射系数在角度一定的情况下仅于地表粗糙度有关,DEM 数据可表示入射角的变化,NDVI 数据表示植被覆盖情况。根据文献[15]的方法,在 98 个实测样点的土壤体积含水量数据中选取 2/3 作为训练数据,余下 1/3 作为验证数据,建立结构为 4 层的 BP 神经网络,1 个输入层 6 个神经元,第 1 隐含层 8 个神经元,第 2 隐含层 6 个神经元,输出层 1 个神经元,训练函数为 Levenberg - Marquardt 数值优化算法 trainlm,因其能使网络快速收敛,适用于小型神经网络<sup>[16]</sup>。采用 GA 优化设计好的 BP 的网络权值、阈值,优化后的神经网络用来训练及土壤水分反演。GA 优化 BP 神经网络的具体参数设置如表 1,图 3 为优化神经网络土壤水分反演的结果图,由图 3 可知,土壤水分含量范围为 0.14 ~ 0.76  $m^3/m^3$ ,裸地部分土壤水分含量较低,当

植被覆盖度增大时,土壤水分含量相对增高,该反演结果与实际情况大抵一致。

表 1 GA-BP 参数设置

参数	数值	参数	数值
第 1 隐含层神经元	3 ~ 14	种群数	70
第 1 隐含层激活函数	tansig	遗传代数	100
第 2 隐含层神经元	1 ~ 5	染色体最大长度	109
第 2 隐含层激活函数	logsig	变异概率	0.08
输出层激活函数	logsig	交叉概率	0.6

### 4.2 反演结果验证

为确定遗传优化神经网络土壤水分反演的有效性,设计 4 种组合方案进行试验,图 4 为不同方案的

土壤水分反演值与土壤水分实测值的对比,表 2 为相关模型的相关系数( $R^2$ )值。

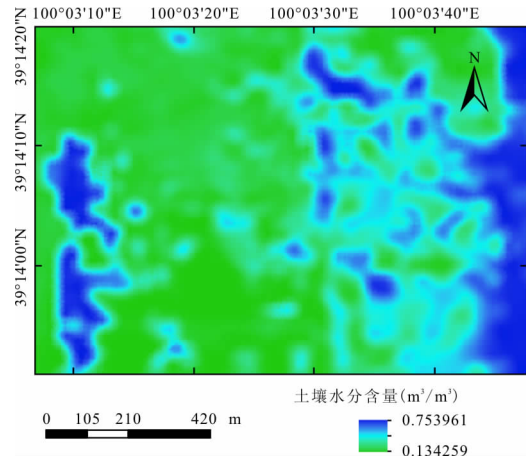


图 3 GA-BP 反演土壤水分分布图

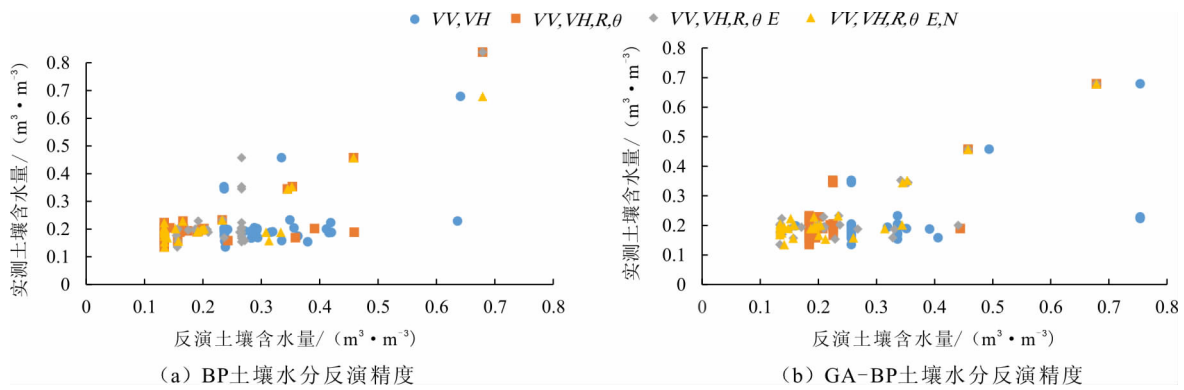


图 4 土壤水分反演值与土壤水分实测值对比图

表 2 不同组合方式下 BP 神经网络与 GA-BP(优化后)神经网络土壤水分反演精度对比

输入变量	BP ( $R^2$ )	GA-BP ( $R^2$ )
$VV, VH$	0.21	0.26
$VV, VH, R, \theta$	0.62	0.69
$VV, VH, R, \theta, E$	0.62	0.67
$VV, VH, R, \theta, E, N$	0.76	0.79

其中  $VV, VH$  为 ASAR 数据  $VV, VH$  极化下的后向散射系数;  $R$  为  $VH$  与  $VV$  后向散射系数比值;  $\theta$  为雷达入射角度;  $E$  为高程;  $N$  为  $NDVI$ 。

由表 2 可知,对比不同组合方案,采用遗传算法优化后的神经网络反演精度普遍高于未优化的神经网络,随着输入变量的增多,反演结果与实际测量值的相关性系数也随之增大,对以上 4 种方案的具体分析如下:

(1)当输入参数仅为 ASAR 数据不同极化的后向散射系数时,相关系数最低。

(2)土壤水分除受到雷达后向散射系数影响以外亦受雷达入射角度的影响<sup>[17]</sup>,在神经网络输入变量中添加雷达入射角度数据后,相关系数有所提升,  $R^2$  可达 0.69。

(3)输入参数中添加高程数据后相关系数无明显变化,或因研究区地势较为平坦,无明显高低起伏变化,后向散射系数随高程变化不明显,故未影响土壤水分的反演结果。

(4)由于研究区样方内存在植被,ASAR 数据后向散射系数为地表和植被散射之和,因此引入  $NDVI$  数据减弱植被散射的影响,提高反演精度,  $R^2$  为 0.79。

相较于单一雷达数据,雷达数据协同光学数据等辅助数据的反演精度更高,同时经遗传算法优化后的神经网络进一步提升土壤水分反演的精度,由此可见,基于遗传优化神经网络的多源数据土壤水分协同反演方法可行。

## 5 结 论

通过遗传算法优化神经网络进行土壤水分反演实验,在优化后的神经网络中输入雷达后向散射系数( $VV$ 、 $VH$ 、 $VH/VV$ )、雷达入射角度、高程、归一化植被指数数据进行土壤水分反演,通过反演结果与实际测量数据的对比验证,结果如下:

(1)通过GA优化BP网络,在缩短网络训练时间的基础上寻找到全局最优解,并提高收敛速度,弥补了神经网络及遗传算法的缺陷。

(2)使用遗传优化神经网络方法代替传统的模型方法处理后向散射系数与土壤水分之间的非线性关系。

(3)分别将单一雷达数据、协同雷达与光学数据等辅助数据作为遗传优化神经网络的输入数据,通过对比分析发现后者反演精度高于前者,体现出主动与被动遥感土壤水分协同反演具备一定的潜能。证明该方法可有效预测土壤水分,对农业生产、灾害监测、地下水补给情况等提供支持。

由于实测数据有限,未能考虑其他影响雷达后向散射系数和土壤水分的因素,使得在神经网络反演过程中可能存在一定偏差,可以考虑通过以下因素进一步改进所提出的反演方法:(1)增加实测数据种类及数量,在优化神经网络的输入参数中添加如土壤表面粗糙度参数、土壤类型、土壤成分等数据;(2)结合光学数据,除 $NDVI$ 以外还需考虑温度植被干旱指数( $TVDI$ );(3)研究GA-BP最佳参数的设定方法,优化神经网络训练速率等。

### 参考文献:

- [1] BARRETT B, WHELAN P, DWYER E. Detecting changes in surface soil moisture content using differential SAR interferometry[J]. International Journal for Remote Sensing, 2013, 34(20):7091-7112.
- [2] 姜良美. 基于微波遥感农田土壤水分反演研究[D]. 湘潭:湖南科技大学, 2012.
- [3] 王大龙,舒英格. 土壤含水量测定方法研究进展[J]. 山地农业生物学报, 2017, 36(2):61-65.
- [4] 高峰,王介民,孙成权,等. 微波遥感土壤湿度研究进展[J]. 遥感技术与应用, 2001, 16(2):97-102.
- [5] 孔金玲,甄珮珮,李菁菁,等. 基于新的组合粗糙度参数的土壤水分微波遥感反演[J]. 地理与地理信息科学, 2016, 32(3):34-38.
- [6] 余凡,赵英时. 合成孔径雷达反演裸露地表土壤水分的新方法[J]. 武汉大学学报(信息科学版), 2010, 35(3):317-321.
- [7] 韩玲,秦小宝,陈鲁皖. 双极化SAR数据反演裸露地表土壤水分[J]. 测绘工程, 2018, 27(2):7-12.
- [8] BAGHDADI N, CRESSON R, HAJJ M E, et al. Estimation of soil parameters over bare agriculture areas from C-band polarimetric SAR data using neural networks[J]. Hydrology & Earth System Sciences Discussions, 2012, 9(3):1607-1621.
- [9] SANTI E, PALOSCIA S, PETTINATO S, et al. Comparison between SAR soil moisture estimates and hydrological model simulations over the scriveria test site[J]. Remote Sensing, 2013, 5(10):4961-4976.
- [10] WANG Xin, GE Lingling, LI Xiaojing. Evaluation of filters for envisat asar speckle suppression in pasture area[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2012, 1-7:341-346.
- [11] HOSSAIN A K M, EASSON G. Soil moisture estimation in South-Eastern New Mexico using high resolution synthetic aperture radar (SAR) data[J]. Geosciences, 2016, 6(1):1.
- [12] 余凡,赵英时,李海涛. 基于遗传BP神经网络的主被动遥感协同反演土壤水分[J]. 红外与毫米波学报, 2012, 31(3):283-288.
- [13] 谢小山. 基于遗传算法和BP神经网络的铁路客运量预测研究[D]. 成都:西南交通大学, 2010.
- [14] 刘浩然,赵翠香,李轩,等. 一种基于改进遗传算法的神经网络优化算法研究[J]. 仪器仪表学报, 2016, 37(7):1573-1580.
- [15] 哈斯巴干,马建文,周自江,等. 基于气象数据与AVHRR热红外数据的人工神经网络分类方法[J]. 中国科学院大学学报, 2003, 20(3):328-333.
- [16] 刘丽娜. 基于遗传优化BP神经网络算法的土壤含水量反演研究[D]. 成都:电子科技大学, 2011.
- [17] 刘伟,施建成,余琴,等. 地表土壤水分与雷达后向散射系数及入射角之间关系研究[J]. 国土资源遥感, 2004(3):14-17.