

# 基于 GIS 与神经网络的黄土丘陵区土壤水分模型研究\*

赫晓慧<sup>1,2</sup> 温仲明<sup>1\*\*</sup> 焦 锋<sup>1</sup> 杨勤科<sup>1</sup> 李 锐<sup>1</sup>

(11 中国科学院水利部水土保持研究所 杨陵 712100; 21 郑州大学环境与水利学院 郑州 450001)

**摘 要** 土壤水分不足是黄土高原丘陵区植被建设的主要限制因子,土壤水分的空间分布受外界气象因子、土地利用与复杂地形等的影响,关系比较复杂。本研究利用黄土丘陵区纸坊沟流域的土壤水分试验资料,建立了基于 GIS 的 BP 神经网络模型,模型中同时考虑了多个因子对土壤水分空间分布的影响,利用实测资料对网络进行训练后对整个流域进行了预测,预测结果与实际情况较为一致,表明应用 GIS 与 BP 神经网络研究区域复杂地形下的土壤水分分布规律是可行的。

**关键词** 土壤水分模型 BP 神经网络 GIS 黄土丘陵区

Application feasibility of GIS based neural network model in soil water analysis in the hilly Loess Plateau. HE Xiaohui<sup>1,2</sup>, WEN Zhongming<sup>1</sup>, JIAO Feng<sup>1</sup>, YANG Qinko<sup>1</sup>, LI Rui<sup>1</sup> (1. Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling 712100, China; 2. School of Environment and Water Conservancy, Zhengzhou University, Zhengzhou 450001, China), CJE, 2007, 15(6): 40~43

**Abstract** The main limiting factor of vegetation construction in hilly areas of Loess Plateau is soil moisture. The spatial distribution of soil moisture is influenced by meteorological factors, land uses and complex topographies, making soil moisture-vegetation relationship pretty complex. This paper utilized data from soil moisture tests in the hilly Loess Plateau of Zhifangou catchment to build a GIS based BP neural network model. The model technique includes multiple factor influences on spatial distribution of soil moisture. Field data was used in training the net which then used to forecast soil moisture of the catchment. The forecast result is in good agreement with practical situations, showing the feasibility of GIS based BP neural network in analyzing soil moisture distribution regulation in highly complex regional topographic terrains.

**Key words** Soil moisture model, BP neural network, GIS, Hilly Loess Plateau

(Received Feb. 17, 2006; revised May 28, 2006)

限制黄土高原植被与生态环境建设的关键因素是土壤水分不足<sup>[1]</sup>,当前对土壤水分模型的研究主要集中在经验性统计模型和物理性理论模型等确定性模型及随机性模型<sup>[2,10]</sup>,所得模型的复杂性及适用性各异。经验性统计模型一般应用比较简单<sup>[3,4,11~13]</sup>,如 Todd 利用地统计学方法建立了小流域地形、植被与土壤性质等的逐步回归模型,但该模型仅能反映某些统计特性,不能反映坡面土壤水分入渗与蒸发的复杂非线性特性,且其经验参数的适用范围有限,难以准确确定<sup>[11,12]</sup>。物理性理论模型具有一定的通用性,但参数较多,应用不便<sup>[5~7]</sup>。黄秉维于 1963 年提出一个比较完全的土壤水分平衡公式,该公式考虑了包括降水和坡面径流在内的 13 个参数,但其中一些参数的获取比较困难<sup>[5]</sup>;傅伯杰<sup>[6]</sup>根据地表特征(植被覆盖和土地利用)和坡度影响下的地表水扩散特征,建立了一个黄土丘陵区坡地的分布式数学模型,但该模型仅适用于土壤特征均一的坡地。

黄土高原地区土壤水分分布受地形、土地利用等众多因素的影响,而现有的土壤水分模型大多针对局部的点或一个坡面的土壤水分以及农田水分进行研究,难以对复杂地形下的区域土壤水分空间分布进行预测,因而不能为植被的空间布局提供科学依据<sup>[2,8,14,15]</sup>。GIS 和人工神经网络(ANN)中的 BP 网络模型为分析土壤水分空间分布的不确定性及其复杂性提供了有效处理手段<sup>[9,16]</sup>。人工神经网络模型可模拟各因素之间的非线性关系,并可在 GIS 支持下将模型推广到区域空间范围。本文着重对黄土丘陵沟壑区不同植被与地形条件下的土壤水分空间演变模型进行研究,以达到模拟黄土高原土壤水分时空分布的目的。

\* 国家自然科学基金项目(40301029)、中国科学院/西部之光0人才培养计划项目(B22012900)和中国科学院水土保持研究所领域前沿项目(SW04107)资助

\*\* 通讯作者

收稿日期:20060217 改回日期:20060528

# 1 研究区域概况与研究方法

## 1.1 研究区域概况

研究区域为陕西省安塞县纸坊沟流域(北纬 36°51'30", 东经 109°19'30"), 属黄土丘陵沟壑区, 植被区划上属于森林草原区, 流域面积 8127km<sup>2</sup>, 年均气温 8.8℃, 年均降水量 549.1mm, 降水年际变率大, 枯水年只有 300mm 左右, 丰水年可达 700mm 以上, 且年内分配不均, 7~ 9 月份降水占全年降水的 61.1%, 年蒸发量大于 1463mm。土壤水分物理特征为凋萎湿度 41.5%, 田间最大持水量 181.4%, 毛管断裂水含量 10%。

## 1.1.2 野外试验设计

影响土壤渗水量分布和蒸发的因素有降水、径流、温度、地形、土壤特征、植被覆盖和土地利用等。因此有必要根据黄土丘陵区实际情况设定一些前提条件, 忽略部分影响因子。基本的假设条件包括: (1) 地块, 地块划分时利用地理信息系统软件结合野外调查的结果进行划分, 且在研究中假定地块内部的性质基本一致。(2) 灌溉, 研究区域基本上以雨养农业和林业为主, 本研究未考虑灌溉的影响。(3) 活动土壤层厚度, 活动土壤层主要指植物根系可以利用水分的土壤厚度, 研究假定植被的活动土壤层厚度约为 1m。

采用 BP 神经网络方法建模的首要前提条件是有足够多典型性好和精度高的样本。为建立良好的数据基础, 在上述假设前提下, 于 2004 年 4~ 10 月在纸坊沟选取主要群落做典型抽样调查, 选取 3 个典型断面, 结合小流域断面地形部位和利用条件确定了 31 个测点进行动态监测, 其中乔木样地 8 个, 灌木样地 8 个, 草本样地 5 个, 人工果园 4 个, 农地 3 个, 共测得 101 组数据。利用 GPS 定点监测土壤水分含量, 对每一样地的立地条件和植物特征进行描述记录, 并人工监测同期降雨和温度。

## 1.1.3 神经网络简介

人工神经网络(Artificial neural networks)是 19 世纪 80 年代中后期迅速兴起的一门非线性科学, 由于其大规模并行处理、分布式储存、高度非线性关系、学习能力强等优良特征, 应用已渗透到各个领域。在目前应用较广的神经网络模型中, BP 神经网络是一种应用最为广泛的前馈性网络, 它具有良好的非线性映射能力, 结构简单, 性能良好, 与其他传统模型相比具有较好的持久性和预测性。BP 网络模型为前向多层模型, 由输入层、输出层和若干个隐含层组成。每层都包含若干个神经元, 同一层各神经元之间没有相互的连接, 相邻层的神经元之间通过权连接, 而隐含层的节点数由具体问题的复杂程度、误差减小情况等来确定。

BP 算法实际上是前向式多层网络的实际输出与要求输出之间最小均方差的一种迭代递推算法, 即用网络的实际输出与目标之间的误差来修正其权值, 使目标输出与目标期望值尽可能的接近, 就是使网络输出的误差平方和达到最小。BP 算法由两部分组成, 即信息的正向传递和误差的反向传播。在正向传递过程中, 输入信息从输入层经隐含层逐层计算传向输出层, 每一层神经元的状态只影响下一层神经元的状态。如果在输出层没有得到期望的输出, 则计算输出层的误差变化值, 然后转向反向传播, 通过网络将误差信号沿原来的连接通路反向传播回来, 修改各层神经元的权值, 直到达到期望目标。

设输入为  $p$ , 输入神经元有  $r$  个, 隐含层内有  $s_1$  个神经元, 输出层内有  $s_2$  个神经元, 对应的激活函数为  $f(x)$ , 输出为  $a$ , 目标矢量为  $T$ 。则隐含层中第  $i$  个神经元的输出为:

$$a_{1i} = f_1 \left( \sum_{j=1}^r w_{1ij} + b_{1i} \right) \quad (i = 1, 2, \dots, s_1) \tag{1}$$

激活函数  $f(x)$  一般采用 Sigmoid 模型, 即:

$$f(x) = \frac{1}{1 + \exp^{-x}} \tag{2}$$

输出层第  $k$  个神经元的输出为:

$$a_{2k} = f_2 \left( \sum_{j=1}^{s_1} w_{2kj} + b_{2k} \right) \quad (k = 1, 2, \dots, s_2) \tag{3}$$

在满足一定精度要求下一般取较小的数值, 以改善网络的概括推论能力, 在训练中网络的收敛采用输出值  $a_{2k}$  与实测值  $t_p$  的平均绝对误差进行控制, 其误差  $E(W, B)$  和训练样本集误差  $E$  可分别定义为:

$$E(W, B) = \frac{1}{2} \sum_{k=1}^{s_2} (t_p - a_{2k})^2 \tag{4}$$

$$E = \frac{1}{s_2} \sum_{k=1}^{s_2} E(W, B) \tag{5}$$

以上为网络学习的信息正向传播过程,另一个过程为误差的反向传播过程,当误差大于期望值时,则将误差反向传播,利用梯度下降法调节权值的变化:

$$W(i+1) = -G \frac{\delta E(W, B)}{\delta W} + AW(i) \tag{6}$$

式中,  $W(i)$  为  $i$  次训练时权植和域值的修正;  $G, A$  分别为比例系数和动量系数。

### 114 地理信息系统(GIS)在土壤水分模型中的应用

地理信息系统(GIS)技术具有强大的对空间数据进行储存、处理、分析和可视化能力,有着传统方法不可比拟的优越性,以土地利用图和 DEM 图为基础数据源,把基于神经网络过程的土壤水分模型结合进来,研究中尺度上环境因子与土壤水分的关系,是土壤水分和水文模型及其尺度外推的重要手段。利用 GIS 解决土壤水分模型普适性不强的问题,将有利于对土壤水分的时间和空间变化进行分析,以建立能够应用于其他区域的土壤水分模型<sup>[16]</sup>。本研究利用 GIS 软件(ARCGIS),通过 5m 栅格精度的 DEM 得到纸坊沟小流域坡度与坡向图,并与植被图、地形地貌图叠加,经过分类、交叉、归组等工作得到 149 个地块单元图,在此基础上结合野外调查建立小流域的空间数据库和属性数据库,从中提取出下文小流域 6 个输入变量的地块文件,将各变量输入 BP 神经网络进行各影响因子的计算,可得出不同地块的土壤水分含量。

## 2 结果与分析

### 211 土壤水分 BP 网络模型输入与输出变量的确定

影响土壤水分空间分布的因素很多且相互作用,而且土壤水分本身又是一个随时间而变化的动态过程,因此各因素与土壤水分的关系很难精确定量,而 BP 神经网络所建立的非线性模型则可以反映各输入变量与输出变量之间关系的复杂性与不对应性。黄土高原丘陵沟壑区沟深坡陡,坡面较长,不同断面部位、不同坡位、坡向等立地因素是黄土高原降水和热量的再分配因子,地形直接影响降雨入渗量和蒸发量,从而带来土壤水分的差异,坡向对热量影响较大,进而导致土壤储水能力不同。根据黄土丘陵区的实际情况,应用 BP 网络采用了下列 6 个变量作为网络输入:一是采样前 30 日内的降雨,降雨量对土壤水分起着决定性作用,为与采样时间对应并保持相对一致性,决定采用 30 日内的总降雨量作为输入变量;二是采样前 30 日内的平均温度,主要反映植被的蒸发量;三是植被,植被对于土壤水分的空间变化有着显著影响;四是坡位;五是坡向;六是坡度。网络的输出值为 30 日降雨时段后 1m 土壤内平均含水量,即预测值。

### 212 网络模型的建立及预测

根据以上分析,网络输入变量数  $n$  为 6,输出变量数  $m$  为 1,采用 1 个隐含层的前馈性神经网络 BP(6,

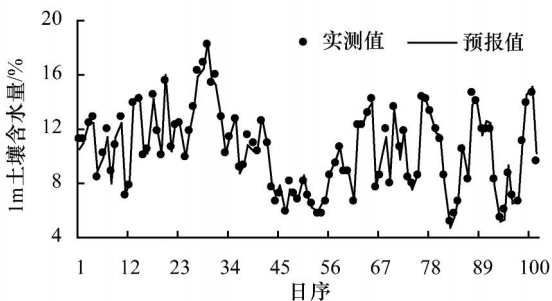


图 1 纸坊沟 1m 土壤含水量实测值与预报值比较  
Fig1 Comparing of measured soil moisture value and forecast value of 1m in Zhifanggou cathment

拟和程度较好,其中实测值与预报值的相关系数达 0.99 以上,相对误差大部分控制在 10% 内,说明该模型训练精度及预测结果可信度均较高。

### 213 水分空间分布预测

采用以上训练好的网络模型对利用 GIS 划分后的纸坊沟 149 个地块单元(居民点无预测值)进行 9 月份土壤

$q, 1)$ , 隐含层单元数  $q$  与所研究的具体问题有关,目前尚无统一的确定方法,根据网络的收敛速度及稳定性等情况,在此选用隐含层节点数  $q_5$  进行训练。为监控训练过程使之不发生/过拟合和评价建立的网络模型的性能和泛化能力,将收集到的数据随机分成训练样本和试报样本(20% 以上)2 部分。此外,数据分组时还应尽可能考虑样本模式间的平衡。因此,随机抽取野外调查试验所得 101 组实测数据中的 79 组作为训练样本,其余 22 组作为试报样本,网络允许误差设为 0.001,将数据归一化后利用 Matlab 编写程序,应用功能强大的神经网络工具箱,程序简单,易于求解。从图 1 和表 1 可知,纸坊沟土壤水分含量实测值与预报值

表 1 BP 神经网络的训练与试报结果

Tab. 1 The results of BP neural network training and test forecasting

样本类型	样本数量	平均绝对 误差 / %	实测值与预报值 相关系数(R)	相对误差 < 5% 的比例 / %	相对误差 < 10% 的比例 / %
Sample type	Sample number	Average absolute error	Correlation index R between true value and forecast value	Proportion of relatively error < 5%	Proportion of relatively error < 10%
训练	79	0.13003	0.9916	81.10	98.17
试报	22	0.13993	0.9901	63.16	95.15

水分预测,即根据以上建立的非线性关系预测同一时段内的小流域各地块单元的土壤水分含量,预测结果见图 2。可以看出,黄土丘陵区小流域整体土壤水分含量不高,水分含量低于 8% 的面积占 15%,而土壤水分高于 15% 的面积仅为 0.3%。这与本区降雨量不高有关,同时也是大面积种植乔木的必然结果。

### 3 结 论

利用 BP 神经网络所生成的能反映土壤水分与其影响因素之间的非线性模型使离散不规则分布的原始土壤水分数据的功能得到扩展。同时,利用 GIS 的支持,使该模型得到良好的数据源基础和直观的图形化显示。单纯的野外实测数据很难满足土壤学、水文学等行业的专业需求,且野外定点测试需耗费大量资源。BP 神经网络模型考虑到了时间和地形、植被等对土壤水分空间分布的影响,结合精确的 DEM,利用非线性模型而不是固定的经验值表达土壤水分的时空间分布,使得寻求区域土壤水分时空间分布规律成为可能。土壤水分建模技术仍在发展之中,各种方法的选用带有一定的主观性。BP 神经网络模型不能像地统计方法那样定量描述土壤性质的空间自相关性,也不能像传统土壤水量平衡模型、土壤水力模型等那样具有明确的物理意义,但该模型能更好地描述土壤水分分布的复杂非线性特性,不需要建立具体的数学模型,计算相对简单,不失为一种建立能反映土壤水分变化与其主要影响因子之间非线性关系模型的好方法。本研究所建立的土壤水分预报 BP 网络模型预测效果较好,模型的输入变量中降雨量、气温、植被与地形等均为较易监测或以地理信息手段获得,因此模型比较容易建立与运用。模型参数是根据一定的自然条件和试验条件下建立的,具有一定的适用范围,影响土壤水分分布的因素众多,如降雨强度、最高最低温度、日照时数、坡长、土壤容重与人为耕作措施等,在进一步的研究中仍需深入探讨。

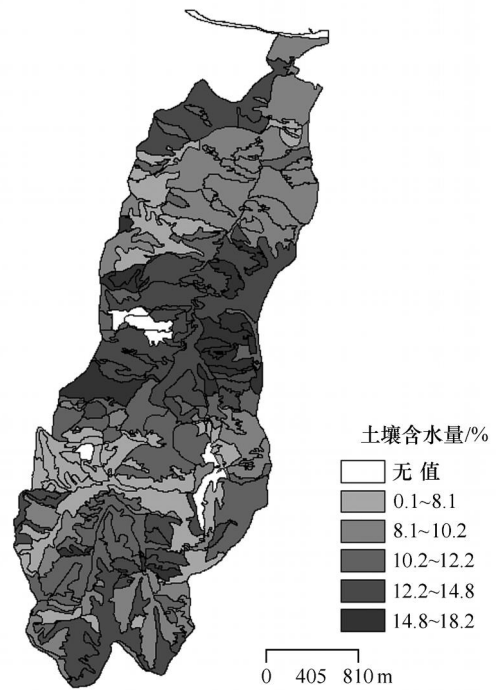


图 2 基于 BP 神经网络的小流域土壤含水量预测结果

Fig. 2 The forecasting result of soil moisture in catchment based on BP

### 参 考 文 献

- 1 吴钦孝,杨文治.黄土高原植被建设与持续发展.北京:科学出版社,1998.42~68
- 2 杨文治,邵明安.黄土高原土壤水分研究.北京:科学出版社,2002
- 3 傅伯杰,王军,马克明.黄土丘陵区土地利用对土壤水分的影响.中国科学基金,1999,13(4):225~227
- 4 邱扬,傅伯杰,王军,等.黄土丘陵小流域土壤水分时空分异与环境关系的数量分析.生态学报,2000,20(5):741~747
- 5 李忠武,蔡强国,曾光明,等.基于 GIS 的黄土丘陵沟壑区土壤水分模型研究.水利学报,2004(3):123~128
- 6 傅伯杰,杨志坚,王仰麟,等.黄土丘陵坡地土壤水分空间分布数学模型.中国科学(D辑),2001,31(3):185~191
- 7 冯国章.通用土壤水分平衡模型及其应用.西北水资源与水工程,1992,3(4):46~48
- 8 赵晓光,吴发启,刘秉正,等.黄土高原坡耕地土壤水分主要受控因子研究.水土保持通报,1999,19(1):10~14,32
- 9 尚松浩,毛晓敏,雷志栋,等.冬小麦田间墒情预报的 BP 神经网络模型.水利学报,2002(4):60~63,68
- 10 McBratney A. B., Santos M. L. M., Minasny B. On digital soil mapping. Geoderma, 2003, 117(112):3~52
- 11 Lookingbill T., Urban D. An empirical approach towards improved spatial estimates of soil moisture for vegetation analysis. Landscape Ecology, 2004, 19(4):417~433
- 12 Andrew W. W., Gunter B., Rodger B. G. Geostatistical characterization of soil moisture patterns in the Tarrawarra catchment. Journal of Hydrology, 1998, 205(112):20~37
- 13 Andrew W. W., Rodger B. G. The Tarrawarra data set: soil moisture patterns, soil characteristics, and hydrological flux measurements. Water Resources Research, 1998, 34(10):2765~2768
- 14 Svetlitchnyi A. A., Plotnitskiy S. V., Stepovaya O. Y. Spatial distribution of soil moisture content within catchments and its modeling on the basis of topographic data. Journal of Hydrology, 2003, 277(1/2):50~60
- 15 Gomez-Plaza A., Martinez-Mena M., Albaladejo J., et al. Factors regulating spatial of soil water content in small semiarid catchments. Journal of Hydrology, 2001, 253(1/4):211~226
- 16 Zhu A. X. Mapping soil landscape as spatial continua: the neural network approach. Water Resources Research, 2000, 36(3):663~667