

基于 BP 神经网络模型的川中丘陵区 坡耕地土壤侵蚀预测研究

朱雪梅¹, 冯海峰¹, 林立金², 杨远祥¹, 邵继荣³, 朱波⁴, 冯济敏⁵

(1. 四川农业大学 资源环境学院, 四川 雅安 625014; 2. 雅安水土保持生态环境监测分站, 四川 雅安 625000; 3. 四川农业大学 生命科学与理学院, 四川 雅安 625014;
4. 中国科学院 水利部 成都山地灾害与环境研究所, 成都 610041; 5. 名山县水务局, 四川 名山 625100)

摘 要:利用地处川中丘陵区的遂宁水土保持试验站 6 个坡耕地径流小区 1991—2000 年的径流观测数据, 采用基于 Matlab 的 BP 神经网络模型, 将 5 个输入因子(降雨量、降雨历时、降雨强度、植被覆盖率、坡度)、9 个输入因子(在 5 个输入因子基础上增加前期降雨量、前期降雨历时、前期降雨强度及前后降雨间隔时间)、10 个输入因子(在 9 个输入因子基础上增加水土保持措施因子)作为输入因子, 冲刷量作为输出因子进行土壤侵蚀建模及预测。结果表明: 5 个输入因子对输出因子的决定系数(R^2)为 0.52, 模型效率系数(NE)为 0.48, 预测效果不理想; 9 个输入因子对输出因子的 R^2 为 0.53, NE 为 0.48, 预测效果也不理想; 10 个输入因子对输出因子的 R^2 为 0.57, NE 为 0.70, 预测效果较好, 可以达到业务预报水平。

关键词:川中丘陵区; 坡耕地; 土壤侵蚀; BP 神经网络模型

中图分类号: S157 文献标识码: A 文章编号: 1005-3409(2011)02-0001-05

Research of Soil Erosion Prediction Based on Back Propagation Neural Network Model in the Hilly Region of Central Sichuan

ZHU Xue-mei¹, FENG Hai-feng¹, LIN Li-jin², YANG Yuan-xiang¹, SHAO Ji-rong³, Zhu Bo⁴, FENG Ji-min⁵
(1. College of Resources and Environment, Sichuan Agricultural University, Ya'an, Sichuan 625014, China;
2. Ya'an Soil and Water Conservation Monitoring Substation, Ya'an, Sichuan 625000, China; 3. College of Life Science, Sichuan Agricultural University, Ya'an, Sichuan 625014, China; 4. Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China; 5. Mingshan Water Conservancy Bureau, Ya'an, Sichuan 625000, China)

Abstract: Based on back propagation (BP) neural network, the soil erosion was predicted according to the observational data from 1991 to 2000 of the six runoff plots at Suining Soil and Water Conservation Experiment Station in the central Sichuan Hilly Region. Five factors (rainfall, rainfall duration, rainfall intensity, vegetation cover, slope), nine factors (five factors plus early rainfall, early rainfall duration, pre-rainfall intensity and time interval before and after the rain), and ten factors (nine factors added to water conservation measures) are input respectively, and erosion is output in the model. The results showed as follows: the determination coefficient of 5 factors on soil erosion was 0.52, and the coefficient of efficiency of those was 0.48, which were not satisfactory, while those of 9 factors on soil erosion were 0.53 and 0.48, which were not good either, but those of 10 factors on soil erosion were 0.57 and 0.70, which achieved a satisfactory prediction results and business forecasting.

Key words: Central Sichuan hilly region; slop farmland; soil erosion; back propagation neural network model

土壤侵蚀已成为我国最严重的生态环境问题之一,开展土壤侵蚀评价与预测是有效进行水土流失监测、开展水保措施效果评估的重要基础,而土壤侵蚀预测模型正是实施这一工作的有效手段^[1]。关于土壤侵蚀的预测,国内外已研究出多种土壤侵蚀评价与预测模型,其中尤以美国的通用土壤流失方程

(USLE)和 WEPP 为代表,随着社会的发展人工神经网络被广泛应用各领域的分类、模式识别、回归预测等方面,取得了丰硕的成果,而其中使用最广泛的是BP 神经网络模型。近年来,许多学者采用 BP 等人工神经网络模型做了相关研究^[2-7],显示出其在土壤侵蚀产沙模拟研究中具有较好的预测精度。

川中丘陵区位于四川盆地中部,为四川省水土流失最严重的地区^[8],目前关于该区土壤侵蚀的研究主要集中在土壤侵蚀成因^[9]、特征^[10-13]、尺度^[14-15]及治理^[16-17]方面,在模型预测方面,仅有少量报道^[18-20],但未有采用 BP 神经网络模型建模预测的报道。本研究利用川中丘陵区的遂宁水土保持试验站坡耕地径流小区 1991—2000 年观测数据,尝试建立该区的基于 BP 神经网络这种人工智能技术的、结构简单的坡面侵蚀产沙预测模型,研究其土壤侵蚀规律,这对开发、保护川中丘陵区坡耕地资源和建设生态环境有重要的意义。

1 研究区域概况

本研究区为地处川中丘陵区的遂宁水土保持试验站(以下简称“遂宁站”)内的坡耕地农耕措施径流观测场。遂宁站位于四川省遂宁市安居镇解元村,地处四川盆地中部嘉陵江中下游丘陵区,琼江流域上游,平均海拔约 420 m,多年平均气温 17.4℃,多年平

均降雨量 885.31 mm。土壤为侏罗纪遂宁组紫色沙泥岩母质发育的红棕紫色土,其结构差,持水量小,保水能力差,pH 呈中性到微碱性,有机质含量低,抗冲刷和抗蚀能力均弱。

2 试验场地及处理

2.1 试验场地设计

坡耕地农耕措施径流观测场地分设 10°、15°两种不同坡级,其中:10°坡级设 4 个径流小区、15°坡级设 2 个径流小区,共 6 个小区。各小区坡面南向临近布置,土壤为红棕紫色土,质地属中壤土,土层厚度 60—80 cm,各小区长、宽水平投影均分别为 20 m、5 m,间隔 2 m 作保护带,四周修建截水墙,下侧修建集水槽,沉沙池和径流池。整个试验场上、下及两侧均有排水沟,场地下方修建观测房,各小区均设径流观测池及三角堰于观测房内,并安装 SW40 型日记水位计 1 台观测观测池池中的水位变化,用 φ20 cm 塑管连接集流槽与观测池。

2.2 试验处理

研究区采用顺坡耕作和横坡垄作等多种保土耕作方法,作物主要为大春红苕(或红苕间作玉米)、小春小麦,不同小区间同一年度至少设置两种不同的耕作方法,研究针对的 1991—2000 年间共实施了 11 种不同的耕作方法,具体见表 1。

表 1 各小区分年度耕作方法安排

年份	10°				15°	
	1 小区	2 小区	3 小区	4 小区	1 小区	2 小区
1991	HDLJ	HDL30	HDL20	HDLW	HDLJ	HDLW
1992	HDLJ	HDL30	HDL20	HDLW	HDLJ	HDLW
1993	HDLJ	HDL30	HDL20	HDLW	HDLJ	HDLW
1994	HDLW	HDLW	SK	SK	SK	HDLW
1995	HDLJG	HDLJ	HDLJ	HDLJG	HDLJG	HDLJ
1996	HDL30	HDL25	HDLJ	SL	HDLJ	HDL25
1998	DDL	HDLW	HDLW	SL	HDLW	SL
1999	HDLW	SL	SL	HDLW	XZ	XZ
2000	HDLW	HDLW	HDLW	HDLW	SF	SF
样本数	24	53	49	38	30	47

注:(1)1997 年因降雨计量仪器故障无观测数据。(2)HDLJ——横坡垄作(加档);HDLW——横坡垄作(未加档);HDL30——斜坡 30°垄作(即横坡垄作与等高线夹角 30°);HDL20——斜坡 20°垄作;HDL25——斜坡 25°垄作;HDLJG——横坡垄作(加档且中部开一深 0.45 m,宽 0.4 m 的小沟);SL——顺坡垄作;DDL——等高带状垄作;XZ——闲置;SK——顺坡垄作(中部开沟);SF——(深翻平整后)闲置。下同。

2.3 数据来源

本研究数据为 6 个坡耕地径流小区 1991—2000 年的径流观测数据,其它数据由遂宁水土保持试验站观测所得。

(1)降雨量观测。汛期按 4 段制(2:00,8:00,14:00,20:00)观测,非汛期按一段制(8:00)观测,并采用人工雨量器与虹吸式雨量计同步观测校正。

(2)径流量观测。一次降雨结束,待小区停止产流后,即通过安装在径流池内的水标尺直接测读水位,然后根据前后径流池不同水位计算浑水径流量。用浑水径流量减去泥沙体积即得到清水径流量。

(3)泥沙观测。把径流池内的水搅匀,在 0.4 m 水深处放置容积为 0.5 L 的平底采样器,从径流池四角及中央分别取水样 5 组置入样瓶中,在实验室内经

过滤、烘干、称重后,即可得出泥沙重量和体积。

(4)植被覆盖率。在观测径流及泥沙的同时观测对应小区内的植被覆盖率

3 模型的建立

3.1 模型输入及输出因子的选择

影响坡面土壤侵蚀的因素众多,主要有降雨、地形、土壤、植被和人为活动(包括水土保持农业措施等)5 个方面。本研究采用 BP 神经网络模型分别对 5 个输入因子、9 个输入因子和 10 个输入因子进行模拟、预测,其中选用降雨量(mm)、降雨历时(h)、雨强(mm/h)、植被覆盖率(%)、坡度(°)5 个因子作为输入因子,冲刷量为输出变量(t/km^2);增加前期降雨雨量(mm)、前期降雨历时(h)、前期降雨雨强(mm/h)、前后降雨相隔时间(h)4 个前期降雨因子,组成 9 个输入因子,冲刷量为输出变量(t/km^2);再增加水土保持措施因子(P)作为输入因子,组成 10 个输入因子,冲刷量为输出变量(t/km^2)。

3.2 建模平台

Matlab 是 MathWorks 公司推出的一款高性能的可视化数值计算软件^[21-23],目前,在各领域得到了广泛的应用。作为一种面向科学与工程计算的高级语言,它为数据分析和数据可视化、算法和使用程序开发提供了最为核心的数学和高级图形工具。它集科学计算、自动控制、信号处理、神经网络、图像处理等于一体,构成了一个方便的、界面友好的用户环境,具有极高的编程效率和运算可靠性,可以大大缩短算法软件开发周期。

3.3 模型参数及预测精度的确定

本研究分别将 5 个因子、9 个因子、10 个因子作为输入变量,输出因子为冲刷量(t/km^2),用其代表土壤侵蚀量,导入到 Matlab 中并存储为 .mat 文件,将各输入的数值都转换到 0~1 之间,即进行归一化处理,选择[0.05,0.95]为归一化的极值范围。本研究采用比例压缩法^[24]。

$$T = T_{min} + \frac{T_{max} - T_{min}}{X_{max} - X_{min}}(X - X_{min}) \quad (1)$$

式中: X——原始数据; X_{max} , X_{min} ——原始数据的最大值和最小值; T——变换后的数据,也称之为目标数据; T_{max} , T_{min} ——目标数据的最大值和最小值。

本文采用 Nash—Suttcliffe^[25]模型效率系数(NE)来衡量模型模拟值与实测值之间的拟合度。NE 值介于 0 与 1 之间,其值越接近于 1,表明模拟效果越好。

$$NE = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})^2} \quad (2)$$

式中: NE——模型效率系数; n——样本数据系列长度; $Q_{obs,i}$ ——第 i 组样本实测值(观测值); $Q_{sim,i}$ ——第 i 组样本模拟值(预测值); \bar{Q}_{obs} ——实测值平均值。

4 结果与分析

4.1 5 个输入因子建模效果

从 BP 神经网络模型训练期拟合输出值、预测期预测输出值与实测值间的对比图(图 1、图 2)可以看出,此时 BP 神经网络模型仅能大体反映原始观测值的变化趋势,但是输出值与原始观测值很多存在较大差距,即出现较多陷入局部极值的情况。对于全部的 6 个径流小区,由于其全部样本中的 5 个输入因子对输出因子的决定系数(R^2)仅为 0.52,而不同小区间存在其他重要影响因子(如前期降雨因素的影响^[26]、各小区耕作方式的不同等)未获提取并用于建模,所以导致在样本数较少的情况下,无法用其建立一个适用于全部小区的 BP 神经网络模型。

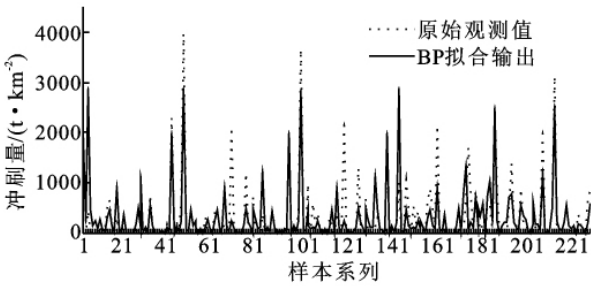


图 1 5 个输入因子 BP 神经网络模型训练期拟合输出与实测值对比(NE=0.51)

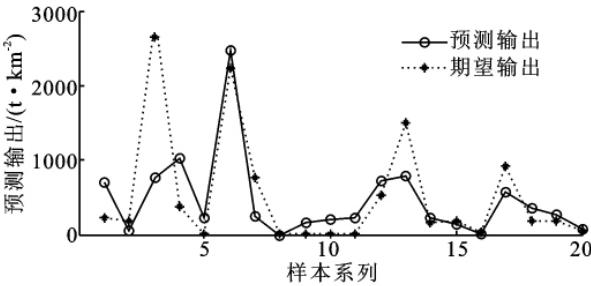


图 2 5 个输入因子 BP 神经网络模型预测输出与实测值对比(NE=0.48)

4.2 9 个输入因子建模效果

从图 3、图 4 可以看出,BP 神经网络模型仍存在较多的陷入极值的情况,据在 Matlab 中编写 BP 神经网络模型训练及预测代码程序并运行的结果显示,隐含层节点数为 15 时,预测效果最好,训练 NE 值为 0.51,预测 NE 值为 0.48,并未比 5 个输入因子建模时的预测效果有所提高,预测效率仍很低,达不到业务应用的要求。但由于 9 个输入因子对输出因子的决定系数(R^2)仅从 0.52 提高到 0.53,BP 神经网络模型建模效率仅在局部稍有改善,但总体上并无提高。

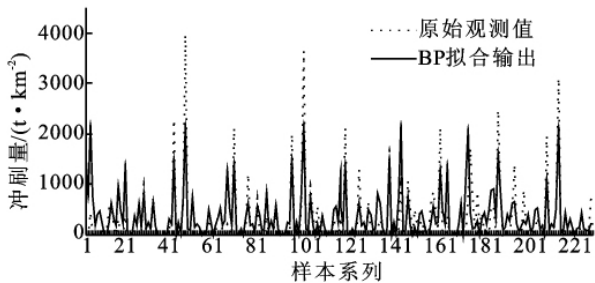


图3 9个输入因子BP神经网络模型训练期拟合输出与实测值对比(NE=0.51)

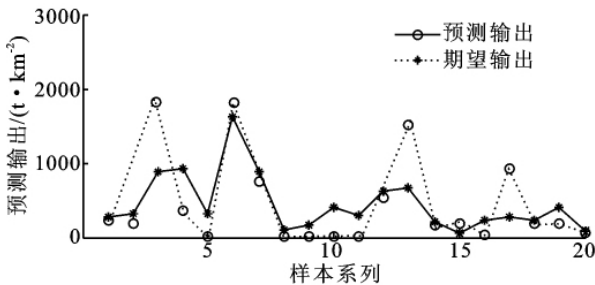


图4 9个输入因子BP神经网络模型预测输出与实测值对比(NE=0.48)

4.3 10 因子建模效果

从图 5、图 6 中可以看出,加入 P 值后,10 个输入因子对输出因子的决定系数(R^2)提高到 0.57,BP 神经网络模型建模时训练期拟合、预测期预测效果均较 9 个输入因子时有较大幅度提高,且预测 BP 神经网络模型在训练期总体已基本符合样本变化趋势,陷入极值的情况较少,模拟效果总体上较 9 个输入因子时

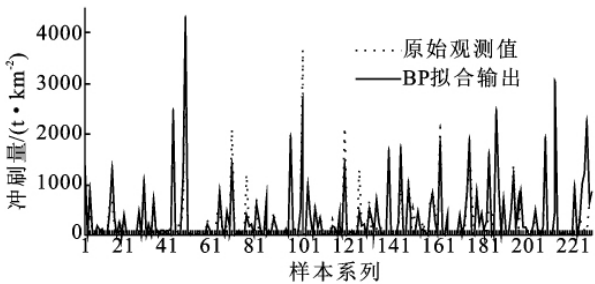


图5 10个输入因子BP神经网络模型训练期拟合输出与实测值对比图(NE=0.75)

要好。在预测期,大部分样本的预测输出值与原始观测值已十分接近甚至与之一致,预测效果也较 9 个输入因子时要好。此时,模型的预测效果为乙等,已可以应用于实际的业务预报中。

4.4 特定耕作方法单独建模

在 Matlab 中对样本数据进行归一化处理,建立 BP 神经网络模型,训练、拟合并选择测试效果最好的模型,其结果汇总如表 2。

从表 2 可以看出,BP 神经网络模型建模的总体情况为,训练期的拟合效果有 6/9 为甲等,2/9 为乙等,1/9 为丙等,无丁等;预测期的预测效果有 3/9 为甲等,2/9 为乙等,4/9 为丙等,无丁等。由此可见,BP 神经网络模型存在拟合效果好而预测效果不一定好这一波动性。建模之前输入因子对输出因子的决定系数及训练样本数对建模的效果有较大影响,一般输入因子决定系数高建立的 BP 神经网络模型的预测效果较好,但输入因子决定系数高而训练样本数较少时,预测效果仍然不理想。当训练样本数足够且输入因子对输出因子的决定系数较高时,BP 神经网络模型建模的训练拟合及预测效果较好。

因此, P 因子可对坡面土壤侵蚀产生重要影响,因而对某种特定的耕作方法用降雨、地形、植被及前期降雨因子进行 BP 神经网络建模预测时,只要具备一定数量的训练样本,且输入因子对输出因子的决定系数较高,可达到较好的建模效果。

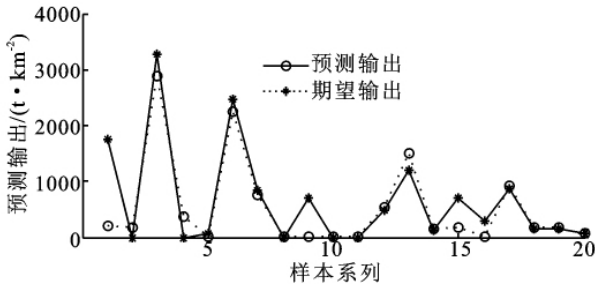


图6 10个输入因子BP神经网络模型预期输出与实测值对比图(NE=0.70)

表 2 不同耕作方法下最佳 BP 神经网络模型主要参数表

耕作方式	样本总数	R^2	测试样本	隐含层神经元数	训练 NE 值	预测 NE 值	预测等级
HDLW	78	0.608	10	9	0.70	0.62	丙等
HDLJ	26	0.741	5	49	0.90	0.61	丙等
HDL20	20	0.966	5	32	0.81	0.61	丙等
HDL25	12	0.992	4	25	0.68	0.52	丙等
HDL30	28	0.980	6	34	0.94	0.86	乙等
SF	16	0.989	4	48	0.96	0.90	甲等
XZ	18	0.975	6	42	0.99	0.99	甲等
SK	18	0.923	6	21	0.97	0.70	乙等
SL	26	0.934	5	50	0.94	0.90	甲等

注:由于 HDLJG、DDL 方式下的样本数小于 5,故对其不开展建模。

5 结 论

研究发现,对所有小区,仅使用降雨量、历时、雨强、植被覆盖率、坡度 5 个因子作为输入因子建模时,预测效果很差;在 5 个输入因子基础上增加前期降雨量、前期降雨历时、前期降雨强度及前后降雨间隔时间,效果增加不是很显著,只有增加了水土保持措施因子后,模型效果增加比较明显,说明水土保持措施因子比较关键,但是在 5 个输入因子上增加 P 后,训练样本数达不到一定的量输入因子对输出因子的决定系数较小且在样本数较少的情况下,无法用其建立一个适用于全部小区的 BP 神经网络模型。由此推测建模时,训练样本数及输入因子对输出因子的决定系数大小对建模效果有较大影响,训练样本数越多、输入因子决定系数越大,建模效果越好。而增加前期降雨雨量、前期降雨历时、前期降雨雨强、前后降雨相隔时间 4 个前期降雨因子,组成 9 个因子为输入因子后,预测效果稍有改善,但是预测精度仍然未达到业务应用水平;当加入经过量化的水土保持措施因子作为输入因子之一时,建模效果大幅提高,所建模型可以应用于业务预报。

参考文献:

[1] 汪东川,卢玉东.国外土壤侵蚀模型发展概述[J].中国水土保持科学,2004,2(2):35-40.

[2] 冯利华.基于 ANN 的土壤侵蚀研究[J].土壤侵蚀与水土保持学报,1999,5(6):105-109.

[3] 王协康,方铎.土壤侵蚀产沙量的人工神经网络模拟[J].成都理工学院学报,2000,27(2):197-201.

[4] 赵西宁,吴普特,冯浩,等.坡面土壤侵蚀产沙的神经网络模拟[J].土壤学报,2006,43(2):324-327.

[5] 李云开,杨培岭,田英杰,等.强烈侵蚀产沙区小流域土壤侵蚀强度的支持向量机预报模型研究[J].北京林业大学学报,2007,29(3):93-98.

[6] 李斌兵,郑粉莉,龙栋材,等.基于支持向量机回归的次降雨小流域侵蚀产沙预报研究:以晋西王家沟为例[J].水土保持通报,2007,27(6):120-125.

[7] 毛典辉,曾致远,王乘,等.基于支持向量机的土壤侵蚀评价预测模型研究[J].人民长江,2007,38(8):82-84.

[8] 尹忠东,苟江涛,李永慈.川中紫色土区农作型小流域水土保持措施设计策略及减蚀效益[J].农业系统科学与综合研究,2009,25(3):370-372.

[9] 莫斌,朱波,高美荣,等.紫色泥页岩的侵蚀产沙特点及影响因素分析[J].水土保持研究,2005,12(1):129-131.

[10] 郭跃.川中丘陵区土壤侵蚀的特征与行为方式[J].水土保持学报,2002,16(6):9-13.

[11] 何丙辉,缪驰远,吴咏,等.遂宁组紫色土坡耕地土壤侵蚀规律研究[J].水土保持学报,2004,18(3):9-12.

[12] 王先拓,王玉宽,傅斌,等.川中丘陵区紫色土坡耕地产流特征试验研究[J].水土保持学报,2006,20(5):9-12.

[13] 杨远祥,朱雪梅,邵继荣,等.遂宁组母质封禁后的土壤侵蚀规律分析[J].水土保持通报,2007,27(1):24-28.

[14] 贺秀斌,张信宝,文安邦.川中丘陵区侵蚀产沙的尺度单元及其研究方法[J].水土保持通报,2004,24(3):18-20.

[15] 高旭彪,孙厚才,赵永军,等.长江上游川中紫色土丘陵区小流域泥沙输移比空间尺度效应研究[J].水土保持通报,2007,27(6):130-134.

[16] 陈泽渊.川中丘陵区遂宁组母质侵蚀与治理技术研究[J].中国水土保持,2001(6):37-38.

[17] 张建华,赵夔兰,林超文,等.川中丘陵坡耕地水土保持与农业生产的发展[J].水土保持学报,2001,15(1):81-83.

[18] 文安邦,张信宝,王玉宽,等.长江上游紫色土坡耕地土壤侵蚀¹³⁷Cs 示踪法研究[J].山地学报,2001,19(增刊):56-59.

[19] 郑进军,张信宝,贺秀斌.川中丘陵区坡耕地侵蚀空间分布的 WEPP 模型和¹³⁷Cs 法研究[J].水土保持学报,2007,21(2):19-23.

[20] 严冬春,文安邦,张忠启,等.坡面版 WEPP 模型在川中丘陵区的应用研究[J].水土保持学报,2007,21(5):42-46.

[21] 张志涌.掌握和精通 Matlab[M].北京:航空航天大学出版社,1997:21-35.

[22] 葛哲学,孙志强.神经网络理论与 MATLAB R2007 实现[M].北京:电子工业出版社,2007:8-12.

[23] 蔡旭辉,刘卫国,蔡立燕. MATLAB 基础与应用教程[M].北京:人民邮电出版社,2009:24-38.

[24] 贺秀斌,张信宝,文安邦.川中丘陵区侵蚀产沙的尺度单元及其研究方法[J].水土保持通报,2004,24(3):18-20.

[25] Nash J E, Sutcliffe I V. River flow forecasting through conceptual models [J]. Journal of Hydrology,1970,273(10):282-290.

[26] 彭清娥,曹叔尤,刘兴年,等.坡面产沙 BP 神经网络模型研究[J].水土保持学报,2002,16(3):79-82.