

基于人工神经网络的土壤颗粒组成制图^①

孙艳俊^{1,2}, 张甘霖^{1*}, 杨金玲^{1,2}, 赵玉国¹

(1 土壤与农业可持续发展国家重点实验室(中国科学院南京土壤研究所), 南京 210008;

2 中国科学院研究生院, 北京 100049)

摘要: 以浙江西苕溪流域为研究区, 综合考虑地形和土壤类型等信息, 采集典型土壤样本, 测定土壤颗粒组成, 并基于土壤颗粒组成与景观位置和特征之间的关系, 利用径向基函数(RBF)神经网络建立了高程、坡度、平面曲率、剖面曲率、径流强度系数和地形湿度指数 6 个地形因子与土壤颗粒组成之间的非线性映射关系, 预测土壤颗粒组成的空间分布。验证结果表明, RBF 神经网络方法能够挖掘出地形因子信息与土壤颗粒组成之间的非线性映射关系, 其预测精度较高, 模型稳定性较好, 是一种低成本、高效率的制图方法。

关键词: 土壤颗粒组成; 地形因子; 神经网络; 数字土壤制图

中图分类号: P934

土壤颗粒组成是土壤物理性质之一, 指土壤中不同大小直径的矿物颗粒的组合状况。作为较稳定的土壤自然属性, 土壤颗粒组成是影响土壤其他一系列物理与化学性质的重要因子, 土壤颗粒组成与土壤通气、保肥、保水状况及耕作的难易有密切关系; 土壤颗粒组成状况是拟定土壤利用、管理和改良措施的重要依据。土壤颗粒组成也是进行陆面过程及大气过程模拟必需的一个重要参数^[1-2]。长期以来, 土壤颗粒组成信息主要依赖于土壤物理分析方法, 要获取区域的土壤颗粒组成分布, 就要进行大量的采样和分析, 速度慢、效率低。

近年来, 数字土壤制图方法在全世界范围内得到了非常广泛的研究, 覆盖了不同地区、不同环境下各种土壤属性及土壤类型的预测制图。目前土壤属性制图主要基于两类方法, 一类是基于模型的方法; 另一类是基于数据驱动的, 如空间内插。目前空间内插在土壤属性制图领域应用最为广泛, 主要有移动平均法^[3]、反比距离加权法^[4-5]、样条函数法^[6]、趋势面分析法^[7]以及克里格法^[8-10]等。这些方法主要是基于几何特征或线性加权方式, 都没能利用母质、气候、地形信息等现有的土壤成土因素信息。现在随着 3S 技术的发展, 人们已经可以很方便地获取地形、气候等方面信

息, 合理利用这些信息就能提高制图精度或降低制图成本, 如协同克里格方法^[11]、回归克里格方法^[12]等。

神经网络是由大量的神经元连接而成的非线性系统, 具有并行分布处理、自组织、自适应、自学习的特点, 适于处理动态的、随机性、非线性数据。径向基函数(RBF)神经网络是目前应用很广泛的一种具有单隐层的三层前馈网络, 能以任意精度逼近任意连续函数。雷能忠等^[13]利用神经网络对土壤全氮进行空间预测, 丁国香^[14]利用神经网络进行有机质和全铁的高光谱反演, RBF 神经网络应用于土壤属性制图的优点在于当样本数据较少时预测能力优于其他空间预测方法^[15], 其实质仍然是建立成土环境与土壤属性之间的关系, 这种关系可能是非线性的, 适合通过神经网络方法来实现。

本研究以浙江西苕溪流域为研究区, 利用 RBF 神经网络方法结合地形因子信息^[16]对土壤颗粒组成进行空间预测, 并对预测结果进行验证分析, 为该流域的水文模拟、营养盐输出等研究提供基础土壤数据支持。

1 研究区域与方法

1.1 研究区概况

研究区域西苕溪流域位于浙江省西北部, 包括安

^①基金项目: 中国科学院知识创新工程重要方向项目(KZCX2-YW-Q10-3)、江苏省自然科学基金项目(BK2008058)和国家自然科学基金项目(40771092)资助。

* 通讯作者(glzhang@issas.ac.cn)

作者简介: 孙艳俊(1986—), 男, 江苏镇江人, 硕士研究生, 主要从事资源演变分异规律与预测制图等方面的研究。

吉县大部和长兴县的一部分，属太湖流域中部丘陵区，位于 30°22'58"~30°53'42"N，119°13'52"~119°57'12"E，属于中亚热带向北亚热带的过渡性区域，气候温和、雨量充沛、四季分明，有明显的垂直气候差异。流域内地形变化较大，高程范围为 2~1 564 m，周长 340 km，面积 2 071 km²。北部地势变化稍缓，南部多高山地形变化较复杂，山丘盆谷相间，具有山、丘、岗、谷、沟、盆地和平原多种地貌类型。土壤类型具有明显的地带性和多样性，北部和中部平原地区以水稻土（水耕人为土）为主，南部和流域边界的高山丘陵地带主要分布着红壤（富铁土）、黑色石灰土（淋溶土或锥形土）、粗骨土（锥形土）和黄壤（淋溶土）。复杂的地形、地貌对该区域土壤发生、发育具有重要影响作用。

1.2 数据准备

基础数据包括土地利用图、土壤类型图、基于 SRTM 的 90 m × 90 m 数字高程模型 (DEM) 和 43 个野外土壤采样点 (图 1)。在样点布设时，尽量确保覆盖不同的土壤类型和地形组合，采样时记录每个样点的地形景观特征、植被、利用方式和母质等，采集土壤表层的土壤样品，实验室利用 LS230 激光粒度仪测定了土壤颗粒组成 (体积百分比)，按照土壤质地划分的卡庆斯基制^[17]表示为土壤砂粒百分比、黏粒百分比和粉粒百分比。

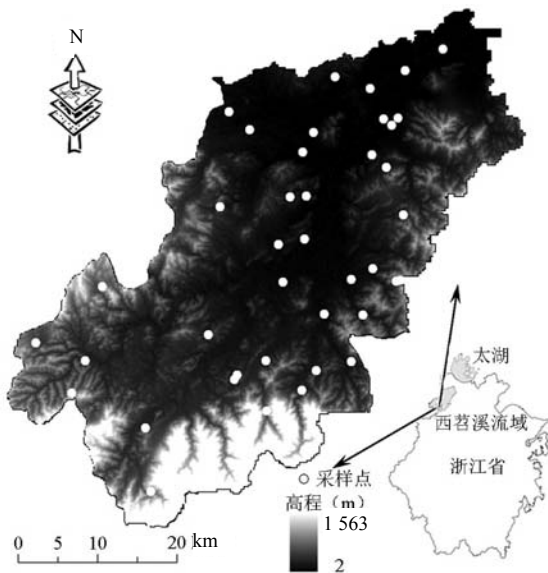


图 1 采样点在流域 DEM 上的分布

Fig. 1 Distribution of sampling sites on basin-DEM

从研究区的数字高程模型提取坡度、坡向、平面曲率、剖面曲率，在 SAGA GIS 软件中对高程、坡度、

坡向等这些简单地形因子进行综合计算提取复合地形指数 (composite terrain index, CTI) 和径流强度指数 (surface slops intensity, SPI)^[18]，其中复合地形指数 (CTI) 和径流强度指数 (SPI) 分别反映了土壤中水分状况和水流侵蚀强度^[19]。最后建立包含高程、坡度、坡向、平面曲率、剖面曲率、复合地形指数、径流强度指数的地形因子数据库。

1.3 研究方法

1.3.1 RBF 神经网络 RBF 神经网络具有优秀的离散数据内插特性，可以提供完备的、最优逼近功能；同时具有结构简单、训练速度快、函数逼近能力强、不存在局部最优问题等特点，由该网络构成的系统是有界、稳定的。RBF 神经网络包含输入层、隐层和输出层，输入层节点传递信号到隐层，隐层节点由作用函数组成，输出层通常是简单的线性函数。结构如图 2 所示。

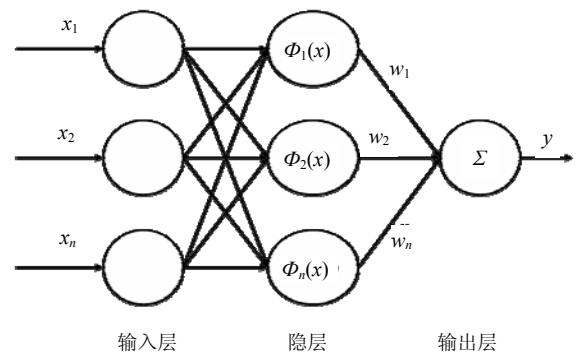


图 2 RBF 神经网络结构

Fig. 2 Structure of RBF network

隐层节点作用函数采用高斯核函数,即:

$$u_j = \left[\frac{(X - c_j)^T (X - c_j)}{2\sigma_j^2} \right] \quad j = 1, 2, 3, \dots, m$$

式中, u_j 是第 j 个隐层节点的输出, $X = (x_1, x_2, \dots, x_n)$ T 是输入样本, c_j 是高斯函数的中心值, σ_j^2 是标准化参数, m 是隐层节点个数。

RBF 网络的输出为隐层节点输出的线性组合,即:

$$y = \sum_{j=1}^m w_j u_j \quad i = 1, 2, 3, \dots, n$$

RBF 神经网络的映射过程是特征输入经过非监督聚类形成空间分布,然后再通过线性对应形成映射关系。由图 1 可见,输入层到隐层实现从模式向量 X 到 u_j 的非线性映射, X 通过非线性径向基函数获得与各个聚类中心竞争归属。若 X 在 c_j 的一定范围内,当 X 越接近于 c_j , u_j 也越接近于 1,从而获得竞争胜利。而隐层到输出层实现由 u_j 到 y 的线性映射。

1.3.2 网络结构 网络结构包括输入层神经元、输出层神经元。本文以高程、坡度、坡向、平面曲率、剖面曲率、复合地形指数和径流强度指数这 7 个预测自变量作为网络输入层。以砂粒、黏粒和粉粒百分比分别作为网络输出层，建立了 3 个神经网络：砂粒 RBF 网络、黏粒 RBF 网络和粉粒 RBF 网络。

1.3.3 制图过程 利用神经网络进行预测制图主要分为网络训练和网络模拟两步。网络训练过程就是建立输入变量和输出变量之间的非线性对应函数关系的过程，初始化网络权重，然后在计算过程中不断增加隐含层神经元个数并调整网络权重直到误差达到预期目标，训练好的网络包含确定的网络结构和连接权重。网络拟合过程就是用训练好的网络模型对研究区地形因子数据库进行模拟计算，输出的结果就是相应的土壤颗粒体积百分比。

本研究在 MATLAB 中编程实现了神经网络过程^[20]。为了减少网络训练误差，首先要对网络训练数据进行归一化处理，准备好训练数据后，调用 `newrb` 函数建立进行网络训练，其调用格式为：

`net = newrb(input, output, goal, spread, mn, df)`

式中，`input` 为标准化后的高程、坡度、坡向、平面曲率、剖面曲率、复合地形指数和径流强度指数这 7 个地形因子组成的向量，`output` 为标准化后的土壤颗粒体积百分比值，`goal` 为训练目标，`spread` 为网络扩展常数，`mn` 为最大隐层神经元个数，`df` 为显示频数。其中 `goal` 反映网络训练结果对训练集的逼近程度，`spread` 和 `mn` 是两个关键参数，对模型训练结果起决定作用^[21]，`spread` 为径向基函数伸展常数，`spread` 越大网络输出越平滑，拟合性越好；`mn` 是网络隐层神经元个数，`mn` 越大网络过程越复杂。

调用网络模拟命令进行空间预测，网络模拟调用格式如下：

`RESULT = sim(net, DATA)`

其中，`net` 是训练好的网络，`DATA` 为区域上的土壤预测自变量，`RESULT` 对应的网络输出为区域上的预测结果，最后在 ArcGIS 9.3 中进行成图输出。

1.3.4 精度评价方法 精度评价时采用的指标有：相关系数 (related coefficient, RC)、平均误差 (mean error, ME)、平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE)、相对误差 (relative error, RE) 和一致性系数 (agreement coefficient, AC)^[22]。其中，平均误差 (ME) 和平均绝对误差 (MAE) 是对模型的总体准确度的度量，其

值越接近于 0，说明模型的总体偏差越小；均方根误差 (RMSE) 是对模型精度及稳定性的度量，值越小说明模型的精度和稳定性越高；相对误差 (RE) 能够客观地反映预测结果的准确度，通常用百分比表示，其值越小说明预测值与观测值越接近；一致性系数 (AC) 描述模型预测值与实际值的吻合程度，其值在 0 到 1 内变化，越靠近 1，表明预测值与实际值之间吻合程度越高，模型越好，反之则差。

2 结果与讨论

2.1 网络优化

网络训练过程中网络扩展常数 `spread` 和隐层神经元个数 `mn` 是两个关键参数 (下面就用 `spread` 和 `mn` 表示)，因此首先要分析确定这 2 个网络训练参数，使得网络泛化性能最强，以得到最优的网络预测结果^[23]。

随着隐含层神经元个数的增加，神经网络模型训练结果的均方根误差 (RMSE) 总体上都呈下降趋势，而验证误差却逐渐增大，即随着训练能力的提高，预测能力反而下降，出现所谓“过拟合”现象^[24]，因而存在着最佳隐含层神经元个数确定的问题。同时，在一般情况下隐含层神经元个数应不小于输入因子的数量，且合理的隐含层神经元个数与模型的层数、输入因子数都有密切关系^[25-28]。兼顾训练误差和检验、预测精度，本文设定隐含层神经元个数为 10。

根据多次网络训练与验证结果，作出隐含层神经元个数 `mn=10` 情况下，网络扩展常数 `spread` 在 0~1 之间的训练集合和验证集合的误差见图 3，根据图 3a 所反映的验证误差随网络扩展常数 `spread` 的变化关系，我们发现当 `spread = 0.04` 时 (图中虚线) 模型在验证点上误差达到最小，说明网络的泛化性能最强，因此将砂粒 RBF 网络的 `spread` 设定为 0.04。对于黏粒 RBF 网络，当 `spread=0.51` 时，网络泛化性能最强；对于粉粒 RBF 网络，当 `spread=0.42` 时，网络泛化性能最强，因此最后确定黏粒 RBF 和粉粒 RBF 的 `spread` 参数分别为 0.51 和 0.42。

2.2 预测图分析

利用最优网络训练参数，在 Matlab 中调用网络训练函数进行训练，然后对砂粒 RBF 网络、黏粒 RBF 网络和粉粒 RBF 网络分别调用地形因子数据库进行网络模拟，模拟结果为二进制 ASCII 码格式，为文件赋上坐标信息后就能在 ArcGIS 中成图输出。

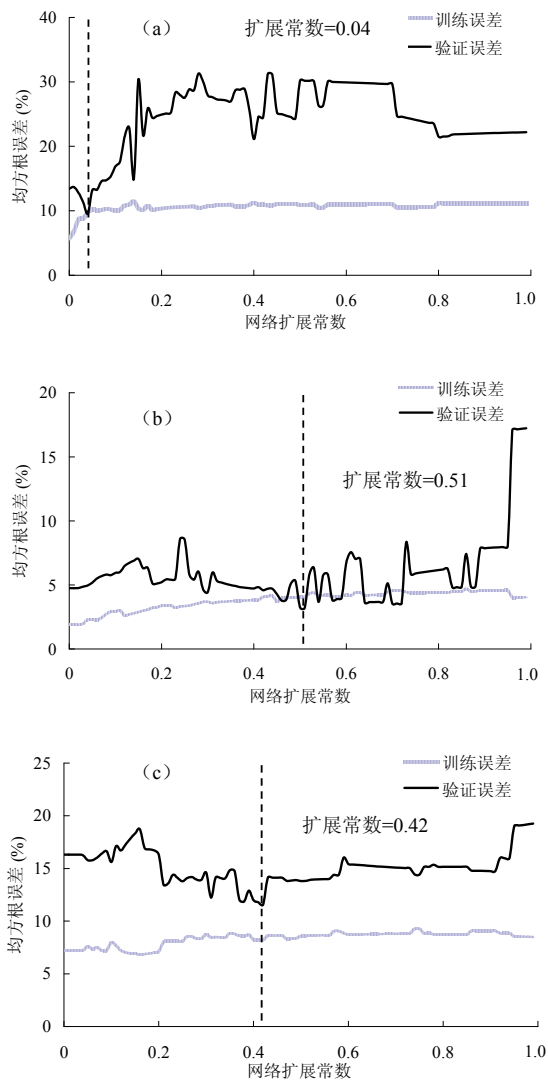


图 3 网络预测结果分析 (a 砂粒; b 黏粒; c 粉粒)
 Fig. 3 Analysis of network predicted results

由图 4 可以看出, 研究区土壤砂粒、黏粒、粉粒百分比预测结果的范围在 2.3%~57.6%、9.1%~58.4%、33.2%~86.8%。流域北部和中部土壤砂粒百分比比较高, 东部和西部黏粒百分比比较高, 流域南部土壤质地情况变化较复杂。根据《安吉土壤》中二次土壤普查调查, 流域北部为潜育型水稻土, 主要有黄泥砂田、泥砂田两土属, 其中黄泥砂田含砂大于 0.1 mm 的砂粒百分比在 30% 以上, 泥砂田大多由溪滩改田过来, 含大量粗砂; 流域中部为渗育型水稻土, 主要有山地黄泥土、砂性黄泥田、白砂田三土属, 土壤含砂砾, 且质地砂黏比>1; 流域东部和西部分布着红壤亚类的黄筋泥土和红黏土两土属, 土体发育好, 黏粒百分比高; 流域南部土壤类型复杂, 质地变化也较复杂。本研究所获得的土壤颗粒组成分布图与第二次土壤普查的结果是相吻合的, 且所得到的连续性土壤颗粒组成分布图, 与传统的多边形图斑图相比能更好地反映土壤颗粒组成的空间变异, 且可获取任意点的土壤颗粒组成值, 便于应用于环境模拟等模型。

从区域上来看, 可以将砂粒、黏粒和粉粒百分比预测结果进行叠加来对预测结果进行初步分析。叠加后的实际值应该为 100%, 对实际值与预测值的相对误差进行分析统计, 发现在绝大多数栅格上的相对误差在 25% 以内, 其中相对误差在 20% 以内的栅格数占 88%, 相对误差在 15% 以内的栅格数占 79%, 相对误差在 10% 以内的栅格数占 62%, 相对误差在 5% 以内的栅格数占 40%, 对于砂粒、黏粒和粉粒百分比叠加分析表明预测结果整体上有较高精度。

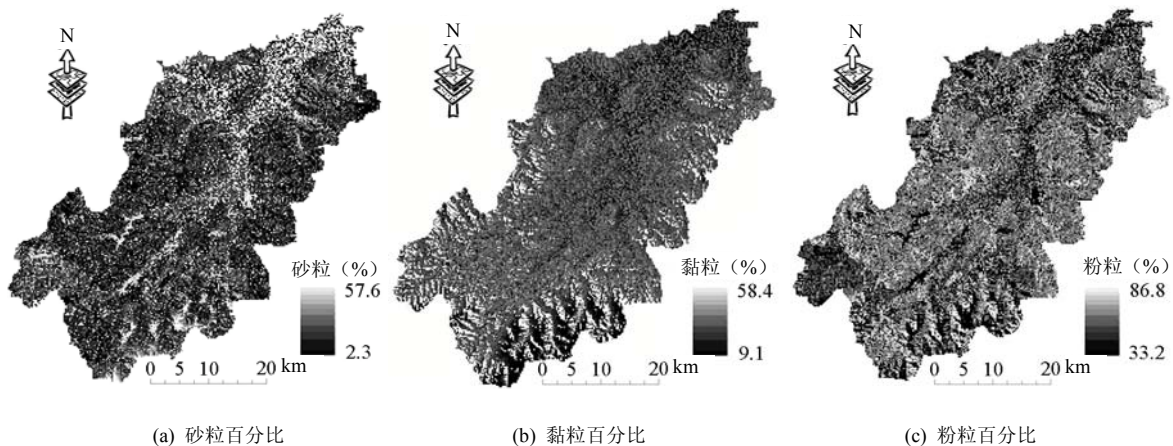


图 4 土壤颗粒组成网络预测结果

Fig. 4 Predication of soil particle composition by RBF neural net

2.3 精度评价

采用 16 个验证样点对砂粒、粉粒、黏粒百分比分布图进行验证。预测值与实测值之间有很好的相关性，

都达到了极显著的水平，其相关系数为砂粒 0.87、黏粒 0.63，粉粒 0.71 ($P < 0.01$, $n = 16$)，可以看出本研究有较高的预测精度 (表 1)。

表 1 验证点相关信息

Table 1 Information of validation points

验证点	土壤类型	砂粒百分比 (%)		黏粒百分比 (%)		粉粒百分比 (%)	
		实测值	预测值	实测值	预测值	实测值	预测值
AJ01	粗骨土	10.27	8.10	20.30	21.37	69.44	70.53
AJ11	粗骨土	11.32	13.54	18.50	17.90	70.19	58.56
AJ12	粗骨土	7.36	10.08	19.95	20.46	72.70	69.46
AJ04	红壤	42.33	31.60	15.65	18.91	42.03	49.49
AJ10	红壤	8.25	9.70	20.90	23.06	70.86	57.24
AJ13	红壤	18.87	17.97	18.70	18.86	62.43	70.17
AJ08	黄壤	15.75	17.38	19.25	20.44	65.01	62.18
AJ09	黄壤	16.41	15.16	19.55	18.56	64.04	57.28
AJ03	潞育水稻土	55.39	41.38	9.12	18.57	35.50	40.05
AJ14	潞育水稻土	10.09	11.49	17.20	17.97	72.72	65.54
AJ15	潞育水稻土	7.60	8.13	15.90	15.41	76.50	65.46
AJ16	渗育水稻土	15.70	17.66	14.55	16.88	69.76	65.46
AJ02	渗育水稻土	13.47	11.53	17.50	18.49	69.03	74.98
AJ05	黑色石灰土	15.11	14.46	23.55	22.06	61.35	71.47
AJ06	黑色石灰土	10.42	12.16	28.45	21.68	61.13	52.16
AJ07	紫色土	15.76	21.54	20.20	21.01	64.04	57.44

对验证点集上的预测结果进行误差分析 (表 2)，砂粒、黏粒、粉粒百分比的平均绝对误差为 3.19%、2.07%、7.07%，均方根误差为 4.89%、3.21% 和 7.81%。作为区域预测的结果，这样的预测精度基本与前人研究相似或略好，如在 Qi 等^[29]的类似研究中，预测结果砂粒百分比和粉粒百分比的平均绝对误差为 7.75%、6.49%，均方根误差为 12.65%、11.11%，预测效果不如

本研究结果，说明预测结果有较高的精度。在 Zhu 等^[30]的研究中，土壤属性预测值的均方根误差是观测平均值的 26%，低于传统土壤图的 38%。本文对于黏粒百分比和粉粒百分比均方根误差是观测平均值的 17% 和 12%，误差变异较小，结果可信度较高，而砂粒百分比均方根误差是观测平均值的 28%，这是由于砂粒 RBF 网络的 spread 常数较小，网络输出不平滑的原因。

表 2 预测图验证结果

Table 2 Validation results of prediction map

项目	平均误差 (%)	平均绝对误差 (%)	均方根误差 (%)	一致性系数	均方误差/均值
砂粒百分比	-0.76	3.19	4.89	0.91	0.28
黏粒百分比	0.77	2.07	3.21	0.92	0.17
粉粒百分比	-2.45	7.07	7.81	0.99	0.12

3 结论

本研究利用 RBF 网络模型建立典型采样点上的土壤砂粒、黏粒和粉粒百分比与地形因子信息之间的非线性对应关系，并进行土壤属性的预测。预测结果与

传统土壤调查结果相吻合，同时对 3 个模型预测结果叠加分析与理论值 100% 做相对误差分析，表明 3 种模型预测结果整体上有较高精度。验证点上实测值与预测值分析结果模型有较高的精度和较好的可靠性。因此使用神经网络模型能够很好挖掘出地形因子等土

壤辅助变量与土壤颗粒之间的非线性函数对应关系, 用于土壤属性预测制图, 是一种低成本高效率的制图方法, 并且预测结果具有较高精度和可靠性。所获得的连续性土壤颗粒组成分布图, 与传统的多边形图斑图相比能更好地反映土壤颗粒组成的空间变异, 且可获取任意点的土壤颗粒组成值, 便于应用于环境模拟等模型。

致谢: 感谢中国科学院南京地理与湖泊研究所李恒鹏博士为本研究提供了研究区土壤类型分布图。

参考文献:

- [1] Duan Q, Schaake J, Koren V. A priori estimation of land surface model parameters // Lakshmi V, Albertson J, Schaake J. Land Surface Hydrology, Meteorology and Climate: Observations and Modeling, Water Science and Applications. vol. 3. Washington DC: AGU, 2001: 77-94
- [2] Dai YJ, Zeng XB, Dickinson RE, Muhammad S, Ranga M. The common land model (CLM). Bull. Amer. Meteor. Soc., 2003, 84(8): 1013-1023
- [3] 王珂, 许红卫, 史舟, Bailey JS, Jordan C. 土壤钾素空间变异性和空间插值方法的比较研究. 植物营养与肥料学报, 2000, 6(3): 318-322
- [4] 肖玉, 谢高地, 安凯. 土壤速效磷含量空间插值方法比较研究. 中国生态农业学报, 2003, 11(1): 56-58
- [5] Brodsky L, Vanek V, Bazalova M, Balik J. The differences in the interpolation methods for mapping spatial variability of soil property. Rostlinna-Vyroba, 2001, 47(12): 529-535
- [6] 石小华, 杨联安, 张蕾. 土壤速效钾养分含量空间插值方法比较研究. 水土保持学报, 2006, 20(2): 68-72
- [7] 李雅琦, 田均良, 刘普灵. 黄土高原土壤元素地域分异规律. 西北农业学报, 2000, 9(3): 63-66
- [8] 霍霄妮, 李红, 张微微, 孙丹峰, 周连第, 李保国. 北京耕作土壤重金属多尺度空间结构. 农业工程学报, 2009, 25(3): 223-229
- [9] 师荣光, 赵玉杰, 周启星, 李野, 刘凤枝, 孙丽. 苏北优势农业区土壤磷空间变异性研究. 农业工程学报, 2008, 24(1): 80-84
- [10] Kerry R, Oliver MA. Forest soil acidification assessment using principal component analysis and geostatistics. Geoderma, 2007, 140(4): 374-382
- [11] 李艳, 史舟, 王人潮, 黄明祥. 海涂土壤剖面电导率的协同克立格法估值及不同取样数目的比较研究. 土壤学报, 2004, 41(3): 434-443
- [12] 鲁程鹏, 束龙仓, 张颖, 王勇. 稀疏数据插值问题的回归克里格方法. 水电能源科学, 2009, 27(1): 81-84
- [13] 雷能忠, 王心源, 蒋锦刚, 黄大鹏. 基于BP神经网络插值的土壤全氮空间变异. 农业工程学报, 2008, 24(11): 130-134
- [14] 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 (硕士学位论文). 南京: 南京信息工程大学, 2008
- [15] 李启权, 王昌全, 岳天祥, 李冰, 杨娟. 基于RBF神经网络的土壤有机质空间变异研究方法. 农业工程学报, 2010, 26(1): 87-93
- [16] Lagacherie P. Digital soil mapping: A state of the art // Hartemink AE, McBratney A, Mendonca-Santos ML. Digital Soil Mapping with Limited Data. Netherlands: Springer, 2008: 3-14
- [17] 胡宏祥, 马友华. 卡庆斯基制与美国制间土粒分析结果的转换. 安徽农业科学, 2004, 32(6): 1156-1157
- [18] Florinsky IV, Eilers RG, Manning GR, Fuller LG. Prediction of soil properties by digital terrain modeling. Environmental Modeling & Software, 2002, 17: 295-311
- [19] 赵量, 赵玉国, 李德成, 孙孝林, 张甘霖. 基于模糊集理论提取土壤-地形定量关系及制图应用. 土壤学报. 2007, 44(6): 961-967
- [20] 闻新, 周露, 李翔. MATLAB 神经网络仿真与应用. 北京: 科学出版社, 2003: 258-284
- [21] 孙福振, 李艳, 李业刚. 基于 RBF 神经网络的土壤分类设计. 安徽农业科学, 2009, 37(13): 6293-6294
- [22] 周启鸣, 刘学军. 数字地形分析. 北京: 科学出版社, 2006
- [23] 唐南奇, 谭明军. 基于人工神经网络农用地分等研究分等模型与精度检测. 福建农林大学学报(自然科学版), 2004, 33(2): 241-244
- [24] 姚荣江, 杨劲松, 邹平, 刘广明. 区域土壤水盐空间分布信息的BP神经网络模型研究. 土壤学报, 2009, 46(5): 788-794
- [25] 屈忠义, 陈亚新, 史海滨, 魏占民. 地下水文预测中BP网络的模型结构及算法探讨. 水利学报, 2004(2): 88-93
- [26] Sahoo GB, Raya C, Wadeb HF. Pesticide prediction in groundwater in North Carolina domestic wells using artificial neural network. Ecological Modeling, 2005, 183: 29-46
- [27] Shen ZQ, Shi JB, Wang K, Kong FS, Bailey JS. Neural network ensemble residual Kriging application for spatial variability of soil properties. Pedosphere, 2004, 14(3): 289-296
- [28] Panlin C, Francois A, Ranlon A, Bernard B. Artificial neural network modeling of water table depth fluctuations. Water Resource Research, 2001, 37(4): 885-896
- [29] Qi F, Zhu AX, Harrower M, Burt JE. Fuzzy soil mapping based on prototype category theory. Geoderma, 2006, 136: 774-787
- [30] Zhu AX, Band LE. A knowledge-based approach to data integration for soil mapping. Canadian Journal of Remote Sensing, 1994, 20(4): 408-418

Mapping of Soil Particle Composition Based on Artificial Neural Network

SUN Yan-jun^{1,2}, ZHANG Gan-lin¹, YANG Jin-ling^{1,2}, ZHAO Yu-guo¹

(1 *State Key Laboratory of Soil and Sustainable Agriculture (Institute of Soil Science, Chinese Academy of Sciences), Nanjing 210008, China;*

2 *Graduate University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Taking Xitiaoqi basin, Zhejiang Province as the study area, a digital soil mapping technique was introduced in this paper based on artificial neural network method. The study integrated comprehensively regional topography and soil types and collected 43 soil profiles for model development and validation. Using the radial basis function (RBF) neural network, the study established a nonlinear relation between six terrain factors, i.e., elevation, slope, plan curvature, profile curvature, surface slope intensity, composite terrain index and soil particle composition and then simulated the space distribution of soil particle composition. Validation results showed that the RBF neural network method could provide a reliable spatial prediction for soil mapping. The model performance of RBF neural network as shown by prediction accuracy and stability is acceptable, which shows that the method is a mapping tool of low cost and high efficiency in the hilly area.

Key words: Soil particle composition, Terrain factor, Neural network, Digital soil mapping