



(12) 发明专利

(10) 授权公告号 CN 103529189 B

(45) 授权公告日 2016.01.06

(21) 申请号 201310268497.0

(22) 申请日 2013.06.28

(73) 专利权人 四川农业大学

地址 625014 四川省成都市温江区惠民路
211号

(72) 发明人 李启权 王昌全 李冰 张新
高雪松 张毅 袁大刚

(74) 专利代理机构 北京众合诚成知识产权代理
有限公司 11246

代理人 龚燮英

(51) Int. Cl.

G01N 33/24(2006.01)

G06F 19/00(2011.01)

(56) 对比文件

JP H11 - 316221 A, 1999.11.16, 全文.

CN 101788553 A, 2010.07.28, 全文.

CN 102680413 A, 2012.09.19, 全文.

JP H01311248 A, 1989.12.15, 全文.

李启权等. 基于神经网络模型和地统计学方法的土壤养分空间分布预测. 《应用生态学报》. 2013, 第24卷(第2期), 第459-466页.

李启权等. 丘陵区土壤有机质空间分布预测的神经网络方法. 《农业环境科学学报》. 2012, 第31卷(第12期), 第2451-2458页.

审查员 潘艺茗

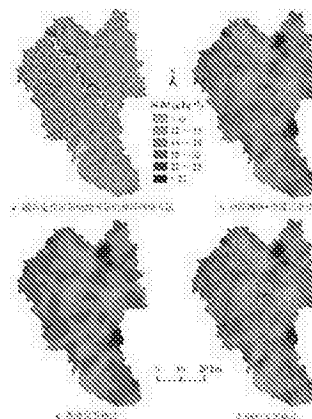
权利要求书2页 说明书8页 附图2页

(54) 发明名称

一种基于定性和定量辅助变量的土壤有机质空间分布预测方法

(57) 摘要

本发明公开了一种基于定性和定量辅助变量的土壤有机质空间分布预测方法,其特征在于:步骤如下:1)、数据来源;2)、辅助数据来源与处理;3)、研究方法;4)、对照方法。由于采用了上述技术方案,与现有技术相比,本发明运用人工神经网络模型,在融合土壤类型、地形因子以及植被指数等定性和定量辅助环境变量的基础上,开展土壤有机质含量的空间分布预测,以期为区域高精度土壤性质空间分布预测提供方法参考。



1. 一种基于定性和定量辅助变量的土壤有机质空间分布预测方法,其特征在于:步骤如下:

1)、数据来源:以选取的研究区地形图、土地利用现状图和土壤类型分布图为辅助信息,在兼顾代表性与均匀性原则的基础上综合考虑地形、土壤类型这些信息进行土壤采样点的布设,在每个土壤采样点处采用多点混合的方法采集表层土壤样品;每个采样点均以GPS记录其地理坐标和海拔高度,同时详细记录样点所在的环境信息;采集到的样品带回实验室经自然风干后磨碎过1mm筛,采用重铬酸钾容量法测定每个土样有机质含量;

2)、辅助数据来源与处理:

研究中定量因子主要选择了地形参数和植被指数;地形参数是土壤性质空间分布预测中最常用的辅助环境变量,包括高程、坡度、平面曲率、剖面曲率及地形湿度指数;基于研究区1:5万地形图在ArcGIS9.3中生成30m分辨率的数据高程模型,进一步利用ArcGIS9.3的空间分析和栅格计算功能获取研究区30m分辨率的坡度、平面曲率、剖面曲率及地形湿度指数分布图;

基于遥感影像获取的植被指数能较好地反映区域地表植被的生长状态和植被覆盖度信息,是常用于土壤性质空间分布预测的又一辅助变量;在研究区采用MODIS遥感数据产品中的增强型植被指数作为预测研究区土壤有机质含量空间分布的定量植被因子;

3)、研究方法:

研究方法采用的是融合定性和定量辅助变量的径向基函数神经网络方法;

将各采样点土壤有机质的实测值和环境因素之间的关系表示为:

$$Z(x_{i,k}, y_{j,k}) = m(k) + r(x_{i,k}, y_{j,k}) \quad (1)$$

$$r(x_{i,k}, y_{j,k}) = f(t(x_i, y_j), v(x_i, y_j)) \quad (2)$$

式中: $Z(x_{i,k}, y_{j,k})$ 表示属于第k种土壤类型的土样在 (x_i, y_j) 处的有机质含量值, (x_i, y_j) 为采样点坐标,其行列号分别为i和j; $m(k)$ 为第k种土壤类型有机质含量平均值; $r(x_{i,k}, y_{j,k})$ 为采样点 (x_i, y_j) 有机质含量减去该点所属土壤类型有机质含量平均值后的残差; $t(x_i, y_j)$ 为点 (x_i, y_j) 的地形参数值, $v(x_i, y_j)$ 为点 (x_i, y_j) 的植被指数值;即假定特定点位上土壤有机质的变异首先由该点所处的土壤类型决定,而其残差值则由该点所处的地形植被条件这些局地环境因素确定;平均值 $m(k)$ 和残差 $r(x_{i,k}, y_{j,k})$ 相互独立;

各点位残差值与局地环境因素间的关系采用径向基函数人工神经网络模型表达;该模型是一种具有单隐层的三层前馈型神经网络模型,预测方法过程如下:

(1)、依据采样点土壤有机质含量值,分别统计研究区中各个土壤类型土壤有机质含量的平均值及每个土壤采样点对应的残差值;

(2)、依据计算出的各个土壤类型有机质含量平均值和研究区土壤类型分布图,制作研究区各土壤类型有机质含量平均值分布图;

(3)、以各土壤采样点的地形参数值及植被指数值作为网络输入,以对应点位归一化后的土壤有机质残差值作为网络输出,构建神经网络模型,在MATLAB中模拟得到研究区土壤有机质的残差分布图;

(4)、最后将研究区各土壤类型有机质含量平均值分布图加上神经网络模型预测得到的残差分布图,即得到研究区土壤有机质含量空间分布图;

4)、对照方法:

选用神经网络模型与普通克里格结合的方法、普通克里格法及回归克里格法作为对照方法；回归克里格法首先采用多元逐步回归进行土壤有机质与环境因子间的回归预测，再以普通克里格法对回归预测结果残差进行插值，最后将回归预测结果和普通克里格法对残差的估计值相加，即得到研究区土壤有机质的空间分布图；神经网络模型与普通克里格结合的方法则是以神经网络模型替代回归克里格法中的回归模型完成对研究区土壤有机质的空间分布预测；

5)、预测精度评价方法

采用独立验证样点集对各方法预测结果进行评价，即随机从 2346 个土壤样点提取 20% 作为验证点，其余 80% 作为建模点，以平均绝对误差、均方根误差和平均相对误差对验证点的预测值和实际观测值进行对比分析，得出精度评价结果。

一种基于定性和定量辅助变量的土壤有机质空间分布预测方法

技术领域

[0001] 本发明涉及一种基于定性和定量辅助变量的土壤有机质空间分布预测方法。

背景技术

[0002] 土壤有机质是土壤特性的重要指标,也是植物矿质营养和有机营养的源泉,在区域土壤肥力、农业生产和环境保护等方面均具有重要的作用和意义。由于多种成土因素和生态过程的影响,土壤有机质含量在空间上呈现非均匀分布。通常基于采样获得的样点土壤有机质含量信息,很难满足区域农业生产和环境保护中对面域土壤有机质含量分布特征的实际需求。因此,精确获取土壤有机质含量的空间分布信息,是准确掌握区域土壤肥力状况、科学的土壤养分管理和区域环境保护的实际需要。

[0003] 土壤性质的空间分布格局是在各种地表环境因素共同作用下形成的。引入相关地表环境因子作为辅助变量的土壤性质空间分布预测方法,在不同程度上考虑了相关环境因子对土壤性质空间分布的影响,其预测精度较仅基于样点数据进行空间内插的方法有明显的提高。在各类辅助变量中,定量因子如气象、地形特征参数等因子相对易于操作,被广泛用于土壤性质空间分布预测。定性因素如土地利用方式、土壤类型和母质类型等对土壤性质空间变异也具有显著的影响,该类因素甚至会弱化定量因子对土壤性质空间变异的作用。因此,在预测方法中同时引入定性和定量辅助变量是进一步提高土壤性质空间分布预测精度的必然要求和最有效的途径之一。然而,在土壤性质空间分布预测中,定性因素的使用明显偏少。

[0004] 在引入环境变量的预测方法中,多元回归模型、协同克里格法以及回归克里格法等方法被广泛用于刻画土壤性质与环境因子间的关系,进而预测土壤性质的空间分布。然而,土壤性质与环境因素间是一种极其复杂的非线性关系,准确表达土壤性质与环境因素间的复杂关系是提高预测精度的关键。神经网络模型能较准确地揭示复杂系统中多元环境变量与目标变量间的非线性映射关系,被成功应用于描述土壤与环境因子间的非线性关系,在土壤性质的空间分布预测研究中受到越来越多的关注。

发明内容

[0005] 本发明要解决的技术问题是克服上述缺陷,提供一种运用神经网络模型,在融合土壤类型、地形因子以及植被指数等定性和定量辅助环境变量的基础上,开展土壤有机质含量的空间分布预测的基于定性和定量辅助变量的土壤有机质空间分布预测方法。

[0006] 为解决上述问题,本发明所采用的技术方案是:

[0007] 一种基于定性和定量辅助变量的土壤有机质空间分布预测方法,其特征在于:步骤如下:

[0008] 1)、数据来源:以选取的研究区地形图、土地利用现状图和土壤类型分布图为辅

点的布设,在每个土壤采样点处采用多点混合的方法采集表层土壤样品;每个采样点均以 GPS 记录其地理坐标和海拔高度,同时详细记录样点所在的环境信息;采集到的样品带回实验室经自然风干后磨碎过 1mm 筛,采用重铬酸钾容量法测定每个土样有机质含量;

[0009] 2)、辅助数据来源与处理:

[0010] 研究中定量因子主要选择了地形因子和植被指数;地形因子是土壤性质空间分布预测中最常用的辅助环境变量,包括高程(H)、坡度(S)、平面曲率(Ct)、剖面曲率(Cp)及地形湿度指数(TI);基于研究区 1:5 万地形图在 ArcGIS9.3 中生成 30m 分辨率的数据高程模型,进一步利用 ArcGIS9.3 的空间分析和栅格计算功能获取研究区 30m 分辨率的坡度、平面曲率、剖面曲率及地形湿度指数分布图;

[0011] 基于遥感影像获取的植被指数能较好地反映区域地表植被的生长状态和植被覆盖度信息,是常用于土壤性质空间分布预测的又一辅助变量;在研究区采用 MODIS 遥感数据产品中的增强型植被指数作为预测研究区土壤有机质含量空间分布的定量植被因子;

[0012] 3)、研究方法:

[0013] 研究方法采用的是融合定性和定量辅助变量的径向基函数神经网络方法;

[0014] 将各采样点土壤有机质的实测值和环境因素之间的关系表示为:

$$[0015] \quad Z(x_{i,k}, y_{j,k}) = m(k) + r(x_{i,k}, y_{j,k}) \quad (1)$$

$$[0016] \quad r(x_{i,k}, y_{j,k}) = f(t(x_i, y_j), v(x_i, y_j), \dots) \quad (2)$$

[0017] 式中: $Z(x_{i,k}, y_{j,k})$ 表示属于第 k 种土壤类型的土样在 (x_i, y_j) 处的有机质含量值, (x_i, y_j) 为采样点坐标,其行列号分别为 i 和 j; $m(k)$ 为第 k 种土壤类型有机质含量平均值; $r(x_{i,k}, y_{j,k})$ 为采样点 (x_i, y_j) 减去该点所属土壤类型有机质含量平均值后的残差。 $t(x_i, y_j)$ 为点 (x_i, y_j) 的地形参数值, $v(x_i, y_j)$ 为点 (x_i, y_j) 的植被指数值。即假定特定点位上土壤有机质的变异首先由该点所处的土壤类型决定,而其残差值则由该点所处的地形植被条件等局地环境因素确定;平均值 $m(k)$ 和残差 $r(x_{i,k}, y_{j,k})$ 相互独立。

[0018] 各点位残差值与局地环境因素间的关系采用径向基函数人工神经网络模型表达;该模型是一种具有单隐层的三层前馈型神经网络模型,预测方法过程如下:

[0019] (1)、依据采样点土壤有机质含量值,分别统计研究区中各个土壤类型土壤有机质含量的平均值及每个土壤采样点对应的残差值;

[0020] (2)、依据计算出的各个土壤类型有机质含量平均值和研究区土壤类型分布图,制作研究区各土壤类型有机质含量平均值分布图;

[0021] (3)、以各土壤采样点的地形及植被特征参数值作为网络输入,以对应点位归一化后的土壤有机质残差值作为网络输出,构建神经网络模型,在 MATLAB 中模拟得到研究区土壤有机质的残差分布图;

[0022] (4)、最后将研究区各土壤类型有机质含量平均值分布图加上神经网络模型预测得到的残差分布图,即得到研究区土壤有机质含量空间分布图;

[0023] 4)、对照方法:

[0024] 选用普通克里格法、回归克里格法及神经网络模型与普通克里格结合的方法作为对照方法;回归克里格法首先采用多元逐步回归进行土壤有机质与环境因子间的回归预测,再以普通克里格法对回归预测结果残差进行插值,最后将回归预测结果和普通克里格法对残差的估计值相加,即得到研究区土壤有机质的空间分布图;神经网络模型与普通克里

格结合的方法则是以神经网络模型替代回归克里格法中的回归模型完成对研究区土壤有机质的空间分布预测。

[0025] 由于采用了上述技术方案,与现有技术相比,本发明以四川盆地丘陵区的三台县为研究区,运用人工神经网络模型,在融合土壤类型、地形因子以及植被指数等定性和定量辅助环境变量的基础上,开展土壤有机质含量的空间分布预测,以期为区域高精度土壤性质空间分布预测提供方法参考。

附图说明

[0026] 图 1 是研究区土壤有机质含量及其对数转换值的频率分布图;

[0027] 图 2 是不同预测方法对土壤有机质含量空间分布的预测结果图。

具体实施方式

[0028] 实施例:

[0029] 一种基于定性和定量辅助变量的土壤有机质空间分布预测方法,步骤如下:

[0030] 1)、数据来源:以选取的研究区地形图、土地利用现状图和土壤类型分布图为辅信息,在兼顾代表性与均匀性原则的基础上综合考虑地形、土壤类型等信息进行土壤采样点的布设,在每个土壤采样点处采用多点混合的方法采集表层土壤样品。全区共采集土样 2346 个采样点,每个采样点均以 GPS 记录其地理坐标和海拔高度,同时详细记录样点所在的环境信息如土壤类型、耕种作物和农地轮作方式等信息。采集到的样品带回实验室经自然风干后磨碎过 1mm 筛,采用重铬酸钾容量法测定每个土样有机质含量。

[0031] 2)、辅助数据来源与处理:

[0032] 研究中定量因子主要选择了地形因子和植被指数;地形因子是土壤性质空间分布预测中最常用的辅助环境变量,包括高程(H)、坡度(S)、平面曲率(Ct)、剖面曲率(Cp)及地形湿度指数(TI);基于研究区 1:5 万地形图在 ArcGIS9.3 中生成 30m 分辨率的数据高程模型,进一步利用 ArcGIS9.3 的空间分析和栅格计算功能获取研究区 30m 分辨率的坡度、平面曲率、剖面曲率及地形湿度指数分布图。

[0033] 基于遥感影像获取的植被指数能较好地反映区域地表植被的生长状态和植被覆盖度信息,是常用于土壤性质空间分布预测的又一辅助变量;在研究区采用 MODIS 遥感数据产品中的增强型植被指数作为预测研究区土壤有机质含量空间分布的定量植被因子。考虑到土壤有机质的积累特点,本发明中选用了 2001 至 2006 年每年 7 月和 8 月 16d 合成的 250m 分辨率的 EVI 数据,在对所下载的 EVI 数据进行质量检查和处理后取平均值并重采样至 30m 分辨率,即用 2001 至 2006 年生物量最大时的平均植被指数作为定量辅助变量之一。为方便计算,所有定量因子均进行归一化处理。

[0034] 研究表明,丘陵区土壤类型较土地利用类型对土壤有机质含量空间分布的影响大,因此研究中选择研究区土壤类型作为本研究的定性辅助变量。研究区土壤类型分布图来源于第二次土壤普查。在对研究区土壤类型图扫描后利用 ArcGIS9.3 进行矢量化,得到研究区数字土壤类型分布图。

[0035] 3)、研究方法:

[0036] 研究方法采用的是融合定性和定量辅助变量的径向基函数神经网络方法;

[0037] 将各采样点土壤有机质的实测值和环境因素之间的关系表示为：

$$[0038] \quad Z(x_{i,k}, y_{j,k}) = m(k) + r(x_{i,k}, y_{j,k}) \quad (1)$$

$$[0039] \quad r(x_{i,k}, y_{j,k}) = f(t(x_i, y_j), v(x_i, y_j), \dots) \quad (2)$$

[0040] 式中： $Z(x_{i,k}, y_{j,k})$ 表示属于第 k 种土壤类型的土样在 (x_i, y_j) 处的有机质含量值， (x_i, y_j) 为采样点坐标，其行列号分别为 i 和 j ； $m(k)$ 为第 k 种土壤类型有机质含量平均值； $r(x_{i,k}, y_{j,k})$ 为采样点 (x_i, y_j) 减去该点所属土壤类型有机质含量平均值后的残差。 $t(x_i, y_j)$ 为点 (x_i, y_j) 的地形参数值， $v(x_i, y_j)$ 为点 (x_i, y_j) 的植被指数值。即假定特定点位上土壤有机质的变异首先由该点所处的土壤类型决定，而其残差值则由该点所处的地形植被条件等局地环境因素确定；平均值 $m(k)$ 和残差 $r(x_{i,k}, y_{j,k})$ 相互独立。

[0041] 各点位残差值与局地环境因素间的关系采用径向基函数人工神经网络模型表达；该模型是一种具有单隐层的三层前馈型神经网络模型，预测方法过程如下：

[0042] (1)、依据采样点土壤有机质含量值，分别统计研究区中各个土壤类型土壤有机质含量的平均值及每个土壤采样点对应的残差值；

[0043] (2)、依据计算出的各个土壤类型有机质含量平均值和研究区土壤类型分布图，制作研究区各土壤类型有机质含量平均值分布图；

[0044] (3)、以各土壤采样点的地形及植被特征参数值作为网络输入，以对应点位归一化后的土壤有机质残差值作为网络输出，构建神经网络模型，在 MATLAB 中模拟得到研究区土壤有机质的残差分布图；

[0045] (4)、最后将研究区各土壤类型有机质含量平均值分布图加上神经网络模型预测得到的残差分布图，即得到研究区土壤有机质含量空间分布图；

[0046] 4)、对照方法：

[0047] 选用普通克里格法、回归克里格法及神经网络模型与普通克里格结合的方法作为对照方法；回归克里格法首先采用多元逐步回归进行土壤有机质与环境因子间的回归预测，再以普通克里格法对回归预测结果残差进行插值，最后将回归预测结果和普通克里格法对残差的估计值相加，即得到研究区土壤有机质的空间分布图；神经网络模型与普通克里格结合的方法则是以神经网络模型替代回归克里格法中的回归模型完成对研究区土壤有机质的空间分布预测。

[0048] 针对上述基于定性和定量辅助变量的土壤有机质空间分布预测方法预测精度评价方法采用独立验证样点集对各方法预测结果进行评价，即随机从 2346 个土壤样点提取 20% (469 个) 作为验证点，其余 80% (1877 个) 作为建模点，以平均绝对误差 (MAE)、均方根误差 (RMSE) 和平均相对误差 (MRE) 对验证点的预测值和实际观测值进行对比分析，得出精度评价结果。

[0049] 上述基于定性和定量辅助变量的土壤有机质空间分布预测方法结果与分析：

[0050] 1、描述性统计分析。

[0051] 根据建模点统计分析表明，如表 1 所示，研究区土壤有机质质量分数的平均值为 17.97g/kg；变异系数处于 10 ~ 100% 之间，为中等程度的变异性。不同土壤类型土壤有机质含量差异较大，水稻土有机质质量分数平均值达到 22.49g/kg，远高于其它 3 个土壤类型；其次是潮土和紫色土，平均值分别为 15.87 和 14.09g.kg⁻¹；黄壤土最小，平均值为 13.66g.kg⁻¹。从变异程度来看，水稻土和紫色土有机质含量变异程度略高于另 2 个土类，

但均属于中等程度变异。

[0052] 从数据的分布形态来看,原始数据呈明显的偏态分布,通过对数转换后其偏度值和峰度值以及分布频率图明显更接近正态分布,如表 1、图 1 所示。因此,对数转换后的数值用于研究中各种预测方法的计算。

[0053] 表 1 研究区不同土壤类型有机质含量描述性统计特征

[0054]

土壤类型	样本数	最小值 (g·kg ⁻¹)	最大值 (g·kg ⁻¹)	平均值 (g·kg ⁻¹)	标准差 (g·kg ⁻¹)	变系系 数 (%)	原始数据		对数转换值	
							偏度	峰度	偏度	峰度
潮土	54	10.6 0	25.2 0	15.8 7	3.4 5	21.7 4	0.1 7	-0.8 4	-0.1 5	-0.9 5
黄壤土	90	6.70	22.3 0	13.6 6	3.0 3	22.1 9	0.5 8	0.16	-0.0 8	0.26
紫色土	113 7	4.20	36.9 0	14.0 9	3.8 5	27.3 2	1.2 2	3.46	0.06	0.96
水稻土	106 5	9.00	47.6 0	22.4 9	6.3 2	28.0 9	0.8 1	1.14	-0.0 8	0.03
全部建模 点	187 7	4.20	47.6 0	17.9 7	6.6 3	36.8 9	1.0 8	1.42	0.15	-0.2 4

[0055] 2、土壤有机质与环境变量的关系分析。

[0056] 利用方差分析对研究区不同土壤类型有机质含量差异进行检验,结果表明,如表 2 所示,研究区 4 种土壤类型有机质含量存在极显著的差异 ($P < 0.001$),即土壤类型的空间分布格局对研究区土壤有机质含量的空间分布具有重要的影响。

[0057] 表 2 不同土壤类型间有机质含量的方差分析

[0058]

方差来源	自由度	方差和	平均方差	F 值	P 值
组间方差	3	32819.39	10939.80	411.75	<0.001
组内方差	1873	49764.18	26.57		
总方差	1876	82583.56			

[0060] 土壤有机质含量与 6 个定量因子间的相关分析结果表明,研究区土壤有机质含量明显受到地形和植被因素的影响,如表 3 所示。其中,海拔高度和坡度与土壤有机质质量分数存在极显著的负相关关系,相关系数分别为 -0.06 ($P = 0.0067 < 0.01$) 和 -0.22 ($P = 4.17E-21 < 0.01$),这说明地势越高、坡度越大的点位上,土壤有机质含量越低。地形湿度指数和植被指数与土壤有机质质量分数间的相关系数分别为 0.24 和 0.17,均达到极显著的

正相关关系 ($P = 6.87E-27$ 和 $5.67E-13$), 这表明地势越低洼、植被盖度越大的位置上, 土壤有机质含量越高。此外, 地形剖面曲率与土壤有机质含量也有一定程度的相关性 (相关系数 0.06 , $P = 0.02$)。

[0061] 表 3 不同土壤类型有机含量及各土壤类型均值残差与环境因子间的相关性系数 ($n = 1877$)

[0062]

数据项	海拔	坡度	平面曲率	剖面曲率	地形湿度指数	植被指数
原始值	-0.06**	-0.22**	-0.03	0.06*	0.24**	0.17**
均值残差	0.02	-0.09**	-0.02	0.02	0.10**	0.13**

[0063] 将各样点有机质含量去除所属土类平均值后得到的残差与各定量因子进行相关分析, 结果表明, 如表 3 所示, 均值残差与坡度的相关系数为 -0.09 , 达到极显著的负相关关系 ($P = 9.35E-05$); 与地形湿度指数和植被指数的相关系数分别为 0.10 和 0.13 , 均表现为极显著的正相关关系 ($P = 9.43E-06$ 和 $2.81E-08$)。这表明, 去除各土壤类型平均值后, 研究区土壤有机质含量的差异主要与地表坡度、地形湿度指数以及植被覆盖状况的差异有关。

[0064] 3、半方差分析

[0065] 将建模点土壤有机质含量值、多元逐步回归预测的残差值及神经网络预测结果的残差进行对数转换, 使其更接近正态分布, 在此基础上完成对上述 3 个数据项的半方差分析。结果表明, 如表 4, 3 个数据项均符合球状模型, 模型的确定系数均在 0.85 以上, 即所选的理论半方差模型的拟合程度较高, 能较好地反映各指标的空间结构特征。从模型的参数来看, 3 个数据项的块金效应值在 $0.742 \sim 0.765$ 之间, 变程在 $6.00 \sim 7.00\text{km}$ 之间, 这表明, 研究区土壤有机质的空间相关程度较弱, 空间自相关的范围不大; 回归方程及神经网络模型预测残差的半方差模型参数与原变量变化较小, 基本保留了原变量的空间结构特征。

[0066] 表 4 研究区土壤有机质及其回归和神经网络预测残差值的半方差模型参数

[0067]

数据项	模型	块金值	基台值	块金效应	变程	决定系数	残差
ln SOM	球状	0.092	0.124	0.742	7.00	0.91	1.91E-04
ln MLR residuals	球状	0.130	0.170	0.765	6.00	0.87	3.58E-04
Ln RBFNN residuals	球状	0.112	0.147	0.762	6.50	0.85	3.28E-04

[0068] \ln SOM:土壤有机质的对数转换值; \ln MLR residuals 土壤有机质回归残差的对数转换值; \ln RBFNN residuals: 土壤有机质神经网络预测残差的对数转换值。

[0069] 4、土壤有机质含量空间分布预测结果

[0070] 图 2 为不同预测方法得到的研究区土壤有机质含量空间分布图。从预测结果图来看,研究区土壤有机质含量空间分布趋势并不明显,这与研究区破碎的地形有关。不同方法预测结果差异明显。普通克里格法 (OK) 预测结果较平滑,高值区和低值区呈明显的块状分布 (图 2c)。回归克里格法 (RK) 和神经网络模型与普通克里格法相结合的方法 (RBFNN+OK) 预测结果相似,能在一定程度上体现土壤有机质含量随地形变化的细节信息 (图 2b 和 d); 但均在研究区东北部和东部出现 2 块明显的呈块状分布的高值区域。融合定性和定量辅助变量的神经网络方法预测结果图中没有明显成块的高值区域分布,如图 2a 图中白色区域为水域,高值区主要出现在地势低洼的冲沟中。

[0071] 在川中丘陵区,地势较低洼的冲沟中主要分布土壤类型为水稻土,统计结果已表明其有机质含量明显高于其它土壤类型,如表 1 所示。相关分析表明地形低洼处土壤有机质含量较高,如表 3 所示。其它相关研究的统计分析结果表明,川中丘陵区微地形下有机质含量差异为:丘脚>丘腰>丘顶;这种微地形下土壤有机质含量的分布特征决定了成块状分布的高值区和低值区均不符合地形破碎的研究区的实际情况。因此,融合定性和定量辅助变量的神经网络方法对研究区土壤有机质空间分布格局的预测结果与研究区实际情况更为一致。

[0072] 2.5 预测精度评价

[0073] 表 5 是不同预测方法对 469 个验证样点预测结果的误差统计结果。从误差分析可以看出,融合定性和定量辅助变量的神经网络方法 (ST+RBFNN) 预测误差明显小于其它 3 种方法。其中,ST+RBFNN 对 469 个验证样点预测结果的平均绝对误差较普通克里格法、回归克里格法及神经网络与普通克里格法结合的方法分别降低了 30.78%、27.43%和 25.70%;平均相对误差分别降低了 35.27%、31.86%和 30.06%,均方根误差分别降低了 22.15%、19.35 和 17.94%,误差降低幅度明显。

[0074] 从不同土壤类型来看,融合定性和定量辅助变量的神经网络方法对紫色土验证点预测精度提高幅度最大,各项误差较其它 3 种方法的降低幅度在 25.93 ~ 49.27%之间,平均降低 38.24%;其次是黄壤土,各项误差的降低幅度在 17.92 ~ 46.19%之间,平均降低 30.04%;再次是潮土,各项误差的降低幅度在 12.50 ~ 25.70%之间,平均降低 19.42%;而降低幅度最小的是水稻土,各项误差的降低幅度在 0.59 ~ 18.55%之间,平均降低 10.08%。这主要是由于所选择的地形因素相对而言能更好地揭示前 3 种土壤类型的发育特点,而水稻土分布的区域各地形因子差异相对较小。

[0075] 表 5 不同预测方法预测误差分析

[0076]

误差参数	预测方法	全部验证点	潮土	黄壤土	紫色土	水稻土
MAE	OK	4.97	3.94	3.36	4.91	5.23
	RK	4.74	3.87	2.55	4.70	5.01
	RBFNN+OK	4.63	3.76	2.88	4.61	4.87
	ST+RBFNN	3.44	3.29	2.06	2.82	4.26
MRE	OK	31.02	27.65	28.47	39.80	21.35
	RK	29.47	26.39	21.13	37.56	21.02
	RBFNN+OK	28.71	25.99	23.40	36.84	20.42
	ST+RBFNN	20.08	20.69	15.32	20.19	20.30
RMSE	OK	6.05	4.50	3.87	5.64	6.68
	RK	5.84	5.02	3.07	5.48	6.43
	RBFNN+OK	5.74	4.71	3.27	5.40	6.31
	ST+RBFNN	4.71	3.73	2.52	4.00	5.56

[0077] 注:OK:普通克里格法;RK:回归克里格法;RBFNN+OK:径向基函数神经网络与普通克里格相结合的方法;ST+RBFNN:融合定性定量辅助变量的神经网络方法;MAE:平均绝对误差;MRE:平均相对误差;RMSE:均方根误差。

[0078] 根据本发明的实施例,可以得出如下结论:

[0079] (1) 研究区土壤有机质质量分数在 4.20 ~ 47.60g. kg⁻¹ 之间,平均为 17.97g. kg⁻¹;变异系数值 36.89%,为中等程度的变异性。不同土壤类型中,水稻土有机质含量最大,其平均质量分数为 22.49g. kg⁻¹;其次是潮土和紫色土,平均值分别为 15.87 和 14.09g. kg⁻¹;黄壤土最小,平均值为 13.66g. kg⁻¹。方差分析结果表明,不同土壤类型间土壤有机质的含量差异达到极显著水平 ($F = 411.75, P < 0.001$),即土壤类型因素对该区域土壤有机质含量的空间分布具有显著的影响。

[0080] (2) 相关分析表明,土壤有机质含量与研究区海拔高度、坡度存在极显著的负相关关系 ($P < 0.01$),与地形湿度指数以及植被指数存在极显著的正相关关系 ($P < 0.01$),与剖面曲率也有一定的相关性 ($P < 0.05$)。在去除土壤类型因素后,坡度、地形湿度及植被因子是引起研究区土壤有机质空间变异的主要因子。

[0081] (3) 半方差分析表明,研究区土壤有机质的块金值与基台值之比在 0.742 ~ 0.765 之间,空间自相关性较弱;变程在 6.00 ~ 7.00km 之间,空间自相关的范围不大。

[0082] (4) 融合土壤类型因素以及地形植被因素的神经网络模型较普通克里格法、回归克里格法以及神经网络与普通克里格相结合的方法精度有明显提高;该方法对 469 个验证点预测结果的平均绝对误差较其它 3 种方法分别降低了 30.78%、27.43% 和 25.70%;平均相对误差分别降低了 35.27%、31.86% 和 30.06%,均方根误差分别降低了 22.15%、19.35 和 17.94%,误差降低幅度明显。该方法可为复杂环境区域土壤性质的空间变异研究提供研究思路。

[0083] 本发明不局限于上述的优选实施方式,任何人应该得知在本发明的启示下做出的结构变化,凡是与本发明具有相同或者相近似的技术方案,均属于本发明的保护范围。

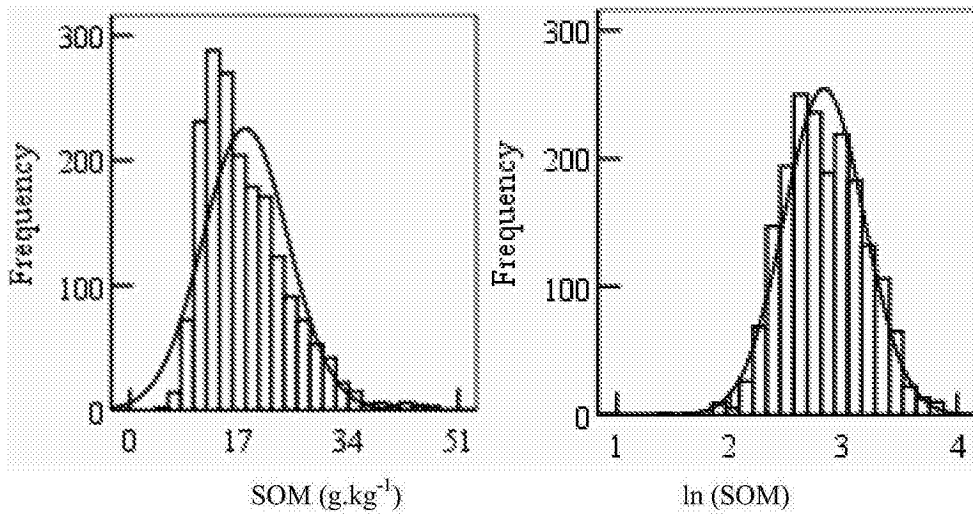


图 1

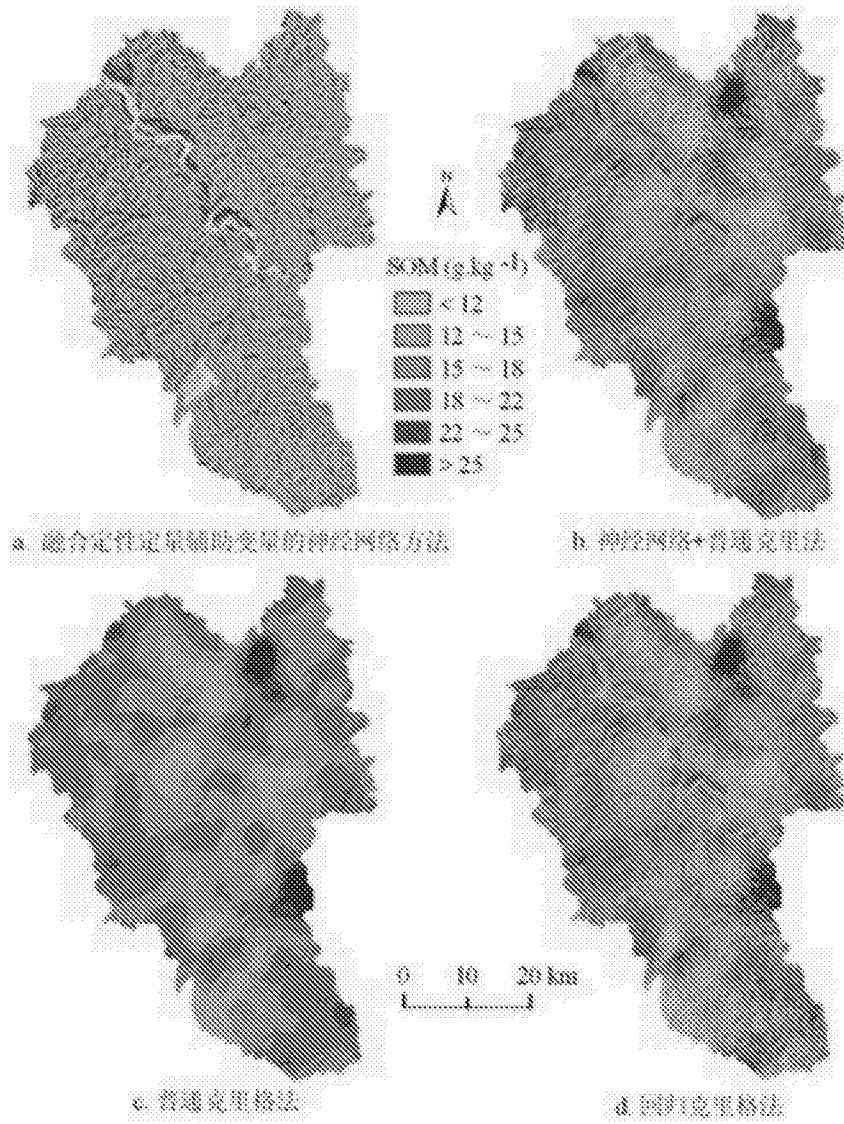


图 2