



(12)发明专利申请

(10)申请公布号 CN 110909467 A

(43)申请公布日 2020.03.24

(21)申请号 201911148393.X

(22)申请日 2019.11.21

(71)申请人 沈阳农业大学

地址 110866 辽宁省沈阳市沈河区东陵路
120号

(72)发明人 王帅 金鑫鑫 杨子娇 虞娜
王妍 齐丽

(74)专利代理机构 北京高沃律师事务所 11569
代理人 杨媛媛

(51)Int.Cl.
G06F 30/20(2020.01)

权利要求书2页 说明书11页 附图2页

(54)发明名称

一种三维土壤田间持水量预测方法及系统

(57)摘要

本发明公开一种三维土壤田间持水量预测方法及系统。该方法包括：获取多个采样点的土壤样本和多个环境数据；根据各所述采样点的土壤样本，确定多个土壤田间持水量；根据各所述土壤田间持水量，建立土壤田间持水量深度分布函数模型；根据所述土壤田间持水量深度分布函数模型，确定采样点的土壤田间持水量深度分布函数模型参数；根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据，预测非采样点的土壤田间持水量深度分布函数模型参数；根据非采样点的土壤田间持水量深度分布函数模型参数和土壤田间持水量深度分布函数模型，确定非采样点土壤田间持水量预测值。本发明能够提高土壤田间持水量预测的精确度和真实性。



1. 一种三维土壤田间持水量预测方法,其特征在于,包括:
 - 获取多个采样点的土壤样本和多个环境数据;
 - 根据各所述采样点的土壤样本,确定多个土壤田间持水量;
 - 根据各所述土壤田间持水量,建立土壤田间持水量深度分布函数模型;
 - 根据所述土壤田间持水量深度分布函数模型,确定采样点的土壤田间持水量深度分布函数模型参数;
 - 根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据,预测非采样点的土壤田间持水量深度分布函数模型参数;
 - 根据所述非采样点的土壤田间持水量深度分布函数模型参数和所述土壤田间持水量深度分布函数模型,确定非采样点土壤田间持水量预测值。
2. 根据权利要求1所述的三维土壤田间持水量预测方法,其特征在于,所述环境数据包括地形数据、气候数据和生物数据。
3. 根据权利要求1所述的三维土壤田间持水量预测方法,其特征在于,所述根据各所述采样点的土壤数据,确定多个土壤田间持水量,具体包括:
 - 称量原状环刀和土壤样本重量;
 - 用200目尼龙纱网包裹原状环刀刀口,下沿放入平底容器中;
 - 向所述平底容器中逐渐加水,直至水面距原状环刀上沿1mm;
 - 静置使原状环刀样在水中饱和48h,称重饱和后环刀样品;
 - 将环刀样品移至压力膜仪测定0.33bar负压含水量,所述0.33bar负压下的土壤含水量为土壤田间持水量。
4. 根据权利要求1所述的三维土壤田间持水量预测方法,其特征在于,所述根据各所述土壤田间持水量,建立土壤田间持水量深度分布函数模型,具体包括:
 - 根据分类指数函数和所述土壤田间持水量,建立土壤田间持水量深度分布函数模型。
5. 根据权利要求1所述的三维土壤田间持水量预测方法,其特征在于,所述根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据,预测非采样点的土壤田间持水量深度分布函数模型参数,具体包括:
 - 采用随机森林模型建立所述采样点土壤田间持水量深度分布函数模型参数和所述环境数据的定量关系模型;
 - 根据所述定量关系模型,预测非采样点的土壤田间持水量深度分布函数模型参数。
6. 根据权利要求1所述的三维土壤田间持水量预测方法,其特征在于,还包括:
 - 对所述非采样点土壤田间持水量预测值的预测精度进行评价。
7. 一种三维土壤田间持水量预测系统,其特征在于,包括:
 - 获取模块,用于获取多个采样点的土壤样本和多个环境数据;
 - 土壤田间持水量确定模块,用于根据各所述采样点的土壤样本,确定多个土壤田间持水量;
 - 土壤田间持水量深度分布函数模型建立模块,用于根据各所述土壤田间持水量,建立土壤田间持水量深度分布函数模型;
 - 采样点的深度分布函数模型参数确定模块,用于根据所述土壤田间持水量深度分布函数模型,确定采样点的土壤田间持水量深度分布函数模型参数;

非采样点的深度分布函数模型参数确定模块,用于根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据,预测非采样点的土壤田间持水量深度分布函数模型参数;

预测模块,用于根据所述非采样点的土壤田间持水量深度分布函数模型参数和所述土壤田间持水量深度分布函数模型,确定非采样点土壤田间持水量预测值。

8. 根据权利要求7所述的三维土壤田间持水量预测系统,其特征在于,所述土壤田间持水量深度分布函数模型建立模块,具体包括:

土壤田间持水量深度分布函数模型建立单元,用于根据分类指数函数和所述土壤田间持水量,建立土壤田间持水量深度分布函数模型。

9. 根据权利要求7所述的三维土壤田间持水量预测系统,其特征在于,所述非采样点的深度分布函数模型参数确定模块,具体包括:

定量关系模型建立单元,用于采用随机森林模型建立所述采样点土壤田间持水量深度分布函数模型参数和所述环境数据的定量关系模型;

非采样点的深度分布函数模型参数确定单元,用于根据所述定量关系模型,预测非采样点的土壤田间持水量深度分布函数模型参数。

一种三维土壤田间持水量预测方法及系统

技术领域

[0001] 本发明涉及土壤田间持水量预测领域,特别是涉及一种三维土壤田间持水量预测方法及系统。

背景技术

[0002] 土壤属性深度分布模拟的关键在于深度分布函数是否可以准确的描述土壤属性的深度分布特征 (Jobbágy and Jackson, 2001)。而深度分布函数对土壤属性的深度分布特征通常是通过两个方面进行表征:模拟的精度和结果的真实性。深度函数的模拟精度是指模型的模拟值与土壤属性值之间的数值吻合程度;而结果的真实性则是指深度函数所模拟的土壤属性所表现的随深度的变化特征与实际的吻合程度 (Kempen et al, 2011)。因此,兼顾准确性和真实性的深度分布函数是表征土壤属性随深度分布特征的基础。

[0003] 针对土壤属性深度分布特征的研究,前人已经进行了大量的研究工作并建立了不同的深度分布函数 (Meersmans et al, 2009; Mendonca-Santos, 2014; Wang et al, 2016)。目前,学者关于数字土壤制图三维尺度的研究主要集中在土壤有机碳中,缺乏对田间持水量三维制图方面的研究。土壤田间持水量是指土壤中所能保持的土壤悬着水的最大量,也是对作物有效的最大土壤含水量,被认为是一个常数。因此,土壤田间持水量三维制图研究的一个重要应用是可以用于灌溉,确定灌溉上限和计算灌水定额。前人对土壤田间持水量的研究已经取得了一些进展,估算的方法多为基于土壤转换函数方法,利用较容易测定的土壤属性(颗粒含量、有机质和容重等)来对土壤的田间持水量进行预测,当拓展到制图上,这种方法往往并不是很适用。往往需要研究人员将转换函数所涉及土壤属性进行空间预测以得到土壤属性的空间数据集,并在此基础上利用转换函数进行空间预测。此外,由于区域内各个土壤属性的预测过程中存在各种不确定性,这种方法可能将会导致预测的不确定性增加。为了达到制图的目的,另一种基于土壤类型和景观类型的多边形方法也被科研人员所提出,然而这种方法通常对数据的要求较高(完备的土壤数据库等)。并且这种方法的最大问题是并不能描述区域内土壤属性的空间变异情况。为了解决这一科学问题,一些研究人员提出了基于栅格数据集的土壤田间持水量的空间预测方法 (Hong et al, 2013; Gooley et al, 2014; Padarian et al, 2014; Tidwell et al, 2014)。因此,基于三维连续空间分布的土壤田间持水量的研究具有重要的意义及应用前景。

发明内容

[0004] 本发明的目的是提供一种三维土壤田间持水量预测方法及系统,能够提高土壤田间持水量预测的精确度和真实性。

[0005] 为实现上述目的,本发明提供了如下方案:

[0006] 一种三维土壤田间持水量预测方法,包括:

[0007] 获取多个采样点的土壤样本和多个环境数据;

[0008] 根据各所述采样点的土壤样本,确定多个土壤田间持水量;

- [0009] 根据各所述土壤田间持水量,建立土壤田间持水量深度分布函数模型;
- [0010] 根据所述土壤田间持水量深度分布函数模型,确定采样点的土壤田间持水量深度分布函数模型参数;
- [0011] 根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据,预测非采样点的土壤田间持水量深度分布函数模型参数;
- [0012] 根据所述非采样点的土壤田间持水量深度分布函数模型参数和所述土壤田间持水量深度分布函数模型,确定非采样点土壤田间持水量预测值。
- [0013] 可选的,所述环境数据包括地形数据、气候数据和生物数据。
- [0014] 可选的,所述根据各所述采样点的土壤数据,确定多个土壤田间持水量,具体包括:
- [0015] 称量原状环刀和土壤样本重量;
- [0016] 用200目尼龙纱网包裹原状环刀刀口,下沿放入平底容器中;
- [0017] 向所述平底容器中逐渐加水,直至水面距原状环刀上沿1mm;
- [0018] 静置使原状环刀样在水中饱和48h,称重饱和后环刀样品;
- [0019] 将环刀样品移至压力膜仪测定0.33bar负压含水量,所述0.33bar负压下的土壤含水量为土壤田间持水量。
- [0020] 可选的,所述根据各所述土壤田间持水量,建立土壤田间持水量深度分布函数模型,具体包括:
- [0021] 根据分类指数函数和所述土壤田间持水量,建立土壤田间持水量深度分布函数模型。
- [0022] 可选的,所述根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据,预测非采样点的土壤田间持水量深度分布函数模型参数,具体包括:
- [0023] 采用随机森林模型建立所述采样点土壤田间持水量深度分布函数模型参数和所述环境数据的定量关系模型;
- [0024] 根据所述定量关系模型,预测非采样点的土壤田间持水量深度分布函数模型参数。
- [0025] 可选的,还包括:
- [0026] 对所述非采样点土壤田间持水量预测值的预测精度进行评价。
- [0027] 一种三维土壤田间持水量预测系统,包括:
- [0028] 获取模块,用于获取多个采样点的土壤样本和多个环境数据;
- [0029] 土壤田间持水量确定模块,用于根据各所述采样点的土壤样本,确定多个土壤田间持水量;
- [0030] 土壤田间持水量深度分布函数模型建立模块,用于根据各所述土壤田间持水量,建立土壤田间持水量深度分布函数模型;
- [0031] 采样点的深度分布函数模型参数确定模块,用于根据所述土壤田间持水量深度分布函数模型,确定采样点的土壤田间持水量深度分布函数模型参数;
- [0032] 非采样点的深度分布函数模型参数确定模块,用于根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据,预测非采样点的土壤田间持水量深度分布函数模型参数;

[0033] 预测模块,用于根据所述非采样点的土壤田间持水量深度分布函数模型参数和所述土壤田间持水量深度分布函数模型,确定非采样点土壤田间持水量预测值。

[0034] 可选的,所述土壤田间持水量深度分布函数模型建立模块,具体包括:

[0035] 土壤田间持水量深度分布函数模型建立单元,用于根据分类指数函数和所述土壤田间持水量,建立土壤田间持水量深度分布函数模型。

[0036] 可选的,所述非采样点的深度分布函数模型参数确定模块,具体包括:

[0037] 定量关系模型建立单元,用于采用随机森林模型建立所述采样点土壤田间持水量深度分布函数模型参数和所述环境数据的定量关系模型;

[0038] 非采样点的深度分布函数模型参数确定单元,用于根据所述定量关系模型,预测非采样点的土壤田间持水量深度分布函数模型参数。

[0039] 根据本发明提供的具体实施例,本发明公开了以下技术效果:

[0040] 本发明提供一种三维土壤田间持水量预测方法。通过获取多个采样点的土壤样本和多个环境数据;根据各采样点的土壤样本,确定多个土壤田间持水量;根据各土壤田间持水量,建立土壤田间持水量深度分布函数模型;根据土壤田间持水量深度分布函数模型,确定采样点的土壤田间持水量深度分布函数模型参数;根据采样点的土壤田间持水量深度分布函数模型参数和环境数据,预测非采样点的土壤田间持水量深度分布函数模型参数;根据非采样点的土壤田间持水量深度分布函数模型参数和土壤田间持水量深度分布函数模型,确定非采样点土壤田间持水量预测值。采用本发明的方法能够提高土壤田间持水量预测的精确度和真实性。

附图说明

[0041] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0042] 图1为本发明三维土壤田间持水量预测方法流程图;

[0043] 图2为本发明田间持水量在1m土体内的3种分布模式图;

[0044] 图3为本发明三维土壤田间持水量预测系统结构图。

具体实施方式

[0045] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0046] 本发明的目的是提供一种三维土壤田间持水量预测方法及系统,能够提高土壤田间持水量预测的精确度和真实性。

[0047] 为使本发明的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0048] 图1为本发明三维土壤田间持水量预测方法流程图。如图1所示,一种三维土壤田

间持水量预测方法包括:

[0049] 步骤101:获取多个采样点的土壤样本和多个环境数据,所述环境数据包括地形数据、气候数据和生物数据。

[0050] 对于土壤样本的采集包括:

[0051] 1) 采样方法

[0052] 以土壤-景观模型为基础的的目的性采样方法,该方法根据土壤属性与成土环境之间的对应关系,假设相似环境下形成的土壤具有较高的相似性,在成土环境因素组合的典型位置布设采样点。

[0053] 首先,利用模糊c均值聚类方法对海拔、坡度、年均温、年降雨等主要环境因素进行聚类。

[0054] 其次,在聚类结果中,每个点由一组模糊隶属度向量构成,模糊隶属度值的高低代表了该点隶属于相应聚类组合的程度。对于每一类环境因素的组合类型,模糊隶属度的高值区通常可以代表该组合类型的中心位置。

[0055] 最终,根据环境因素组合中每个点的模糊隶属度值的大小,以及样点的可达性程度,最终选取具有较高代表性的点作为设计样点。

[0056] 2) 土壤样品采集

[0057] 基于上节设计的样点,挖掘剖面,每个剖面1.2m深或者至(准)石质接触面。根据土壤发生学知识对剖面进行层次划分。土壤的物理属性如土壤水分特征、容重、孔隙等采用环刀法按土壤发生层采集,利用原状环刀(直径5cm,体积100cm³)在每个层次采集三个重复。

[0058] 3) 土壤田间持水量(SoilField Capacity)的测定

[0059] 具体步骤如下:

[0060] a) 称量原状环刀和土样重量;用200目尼龙纱网包裹原状环刀刀口,下沿放入平底容器中;向容器中逐渐加水,在此过程中每升高1cm间隔4h,以驱除土壤中的空气,直至水面距环刀上沿1mm。

[0061] b) 静置使原状环刀样在水中饱和48h,称重饱和后环刀样品;将环刀样品移至压力膜仪测定0.33bar负压含水量(0.33bar负压含水量测定时间约为10天)。最终,0.33bar负压下的土壤含水量即为土壤田间持水量。

[0062] 对于环境数据的采集:

[0063] 所选取的环境数据包括地形因素、气候因素、生物因素、空间位置等3大类12个环境变量,变量的详细信息见表1。表1为研究中所涉及的环境变量。为了达到制图的目的,在ArcGIS 10.2 (ESRI Inc., USA) 中将所有环境变量重采样为90m分辨率的栅格数据,并转换为相同的投影系统(krasovsky_1940_albers)以便进行后续的建模与分析。

[0064] 1) 地形数据

[0065] 用于计算地形变量的源数据为90m分辨率的Shuttle Radar Topography Mission 数字高程模型(SRTM DEM)。SRTM由美国太空总署(NASA)和美国国防部国家测绘局(NIMA)以及德国与意大利航天机构共同合作完成联合测量,由美国发射的“奋进”号航天飞机搭载SRTM系统完成。其数据具有覆盖面积广、采集数据量大、精度高等特点。该数据可从美国地质调查局(United States Geological Survey, USGS) 地球资源观测与科学数据中心(<http://srtm.usgs.gov/>) 免费获取。在SAGA GIS (Olaya, 2004) 软件中,利用SRTM DEM数据

提取了6个地形变量,它们分别为:高程、坡度、坡向、平面曲率、剖面曲率和汇水面积。坡向的范围由0-360°变换为0-180°,代表着由正北向正南的连续变化,0°表示正北,180°表示正南。

[0066] 2) 气候数据

[0067] 气候数据包括年均气温和年均降水量。数据来自中国农业科学院农业资源与农业区划研究所中国生态环境背景层面建设项目完成的栅格数据(1000m分辨率),为1980-1999年的逐月平均值计算生成。

[0068] 3) 生物数据

[0069] 生物数据由遥感卫星数据组成,包括美国陆地卫星第五颗星(Landsat 5)中波段3(B3,0.63-0.69 μm)、波段4(B4,0.76-0.96 μm)、波段5(B5,1.55-1.75 μm)三个波段的反射率及其衍生的归一化植被指数(Normalized Difference Vegetation Index,NDVI)。Landsat 5数据集来自于中国科学院计算机网络信息中心的地理空间数据云平台(<http://www.gscloud.cn/>),该数据集收集了USGS免费发布的2010年的遥感影像(分辨率为30m,数据最大云量不超过10%),利用多项式几何校正方法对影像进行了几何配准,保留了Landsat 5中三个植被信息最丰富的三个波段(波段3、波段4、波段5)。归一化植被指数由公式 $(B4-B3)/(B4+B3)$ 计算得来,它用来表征植被长势。

[0070] 表1研究中所涉及的环境变量

类别	变量	单位	描述	分辨率
Class	Variable	Unit	Description	Resolution
[0071]	地形 Topography	海拔 Altitude	m 到大地水准面的绝对垂直距	90m
	地形 Topography	坡度 Slope	° 地面点法线与铅垂线之夹角	90m
		坡向 Aspect	° 坡面法线在水平面上的投影	90m
	地形 Topography	平面曲率 Surface curvature	(100 m) ⁻¹ 过任一点的水平面沿水平方向切地形表面所得的曲率值	90m
		剖面曲率 Profile curvature	(100 m) ⁻¹ 对地面坡度的沿最大坡降方向地面海拔变化率的度量	90m
	地形 Topography	汇水面积 Catchment area	m ² m ⁻¹ 单位等高线上的汇流面积	90m
		气候 Climate	年均气温 Average annual temperature	°C 1980-1999 年年平均气温
	气候 Climate	年均降水	mm 1980-1999 年年平均降水量	1000m

量			
	Average annual precipitation		
	n		
[0072]	生物 归一化植被指数 Biology NDVI	近红外波段反射值与红光波段反射值 之差除以两者之和	30m
	Landsat5 Band 3	Landsat5 第三波段	30m
	Landsat5 Band 4	Landsat5 第四波段	30m
	Landsat5 Band 5	Landsat5 第五波段	30m

[0073] 步骤102:根据各所述采样点的土壤样本,确定多个土壤田间持水量,具体包括:

[0074] 称量原状环刀和土壤样本重量。

[0075] 用200目尼龙纱网包裹原状环刀刀口,下沿放入平底容器中。

[0076] 向所述平底容器中逐渐加水,直至水面距原状环刀上沿1mm。

[0077] 静置使原状环刀样在水中饱和48h,称重饱和后环刀样品。

[0078] 将环刀样品移至压力膜仪测定0.33bar负压含水量,所述0.33bar负压下的土壤含水量为土壤田间持水量。

[0079] 步骤103:根据各所述土壤田间持水量,建立土壤田间持水量深度分布函数模型,具体包括:

[0080] 根据分类指数函数和所述土壤田间持水量,建立土壤田间持水量深度分布函数模型。

[0081] 基于所采集的80个土壤剖面点数据分析项目区的土壤田间持水量的分布模式,可将其归并为3种分布模式:类型a,随深度递减型;类型b,随深度递增型;类型c,随深度先递减再递增型,如图2所示。

[0082] 图2为本发明田间持水量在1m土体内的3种分布模式图。类型a为随深度递减型;类型b为随深度递增型;类型c为随深度先递减再递增型。

[0083] 为了模拟出研究区土壤田间持水量的深度分布特征,本发明定义了一个分类指数函数,本发明的深度分布函数模型结合了基本形式的指数函数以及土壤田间持水量的深度分布模式。其结构形式如下:

[0084] 递增: $SFC(d) = SFC_a \exp(k_a d)$

[0085] 递减: $SFC(d) = SFC_b \exp(k_b d)$ (1)

[0086] 其中, d 为土壤距地表深度(m), 表层深度为0; $SFC(d)$ 为深度 d 模拟出的土壤田间持水量(%); k 为控制土壤田间持水量随深度的增加的递减或递增速率; SFC_a 和 SFC_b 则分别对应着表层土壤田间持水量(%). 该指数函数包含一个递减型函数和一个递增型函数。图2中的类型a和类型b分别用对应的形式进行模拟。针对类型c, 可以将其划分为上部为递减型和下部为递增型, 然后分别对其进行模拟。

[0087] 步骤104: 根据所述土壤田间持水量深度分布函数模型, 确定采样点的土壤田间持水量深度分布函数模型参数。利用分类指数函数对所有样点进行模拟之后, 在每个样点位置得到了一组田间持水量深度分布函数的参数(SFC_a 、 SFC_b 、 k_a 和 k_b) (式1), 这些参数可以用来描述该位置土壤田间持水量深度分布的特点。

[0088] 步骤105: 根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据, 预测非采样点的土壤田间持水量深度分布函数模型参数, 具体包括:

[0089] 采用随机森林模型建立所述采样点土壤田间持水量深度分布函数模型参数和所述环境数据的定量关系模型。

[0090] 根据所述定量关系模型, 预测非采样点的土壤田间持水量深度分布函数模型参数。

[0091] 为了得到非采样位置的土壤田间持水量深度分布, 本发明首先利用随机森林模型建立了土壤田间持水量深度分布函数中的参数(SFC_a 、 SFC_b 、 k_a 和 k_b)与环境变量之间的定量关系模型, 环境变量如表1所示。

[0092] 再将建立的空间预测模型运用在非采样点位置, 根据非采样点的环境变量预测该点的土壤田间持水量深度分布函数参数。考虑到本发明中样点数量有限, 以及随机森林模型在精度评价方面具有较好地可靠性, 所以利用所有土壤样点参与建模。

[0093] 表2为模型预测精度统计表。表2列出了随机森林模型预测分类指数函数中参数精度。识别土壤田间持水量的深度分布类型的错误率为0.48, 表明随机森林模型对土壤田间持水量的深度分布模型的正确率接近50%。模型分布参数 SFC_a 、 SFC_b 、 k_a 和 k_b 的 R^2 分别为0.59、0.61、0.53和0.51。

[0094] 表2模型预测精度统计表

参数 Parameter	指标 Index				
	Error rate	MA E	RM SE	R ²	LUCC
深度分布类 型	0.48	-	-	-	-
[0095] <i>SFC_a</i>	-	2.16	4.25	0.59	0.62
<i>SFC_b</i>	-	1.34	3.57	0.61	0.68
<i>k_a</i>	-	0.32	0.96	0.53	0.57
<i>k_b</i>	-	0.47	1.27	0.51	0.55

[0096] 随机森林是由美国科学家Leo Breiman于2001年提出的一种基于分类回归树(Classification and Regression Tree, CART)的机器学习算法,该方法结合了多次重复采样(Bootstrapping)技术和特征随机选取技术(Breiman, 2001)。随机森林的优点体现在建模过程的随机性特征,表现在两个方面:(1)训练每棵树时,从全部训练样本中随机选取一个子集进行训练,用未被抽取的数据(out-of-bag, OOB)进行误差评价;(2)在每个节点,随机选取所有特征的一个子集,用来计算最佳分割方式。由于随机森林是基于随机方式建立的模型,包含了多个分类回归树,从而保证了模型的多样性和稳定性,可以用来解决分类和回归的相关问题。

[0097] 随机森林的训练过程可概括为三个主要步骤:

[0098] (1) 采用多次重复采样技术进行取样:从原始训练集中有放回地选取n个样本。

[0099] (2) 特征选取与模型建立:对每个训练样本,随机选取出m个属性作为分裂节点的候选属性,再对n个样本分别建立n个分类回归模型,得到n种结果。

[0100] (3) 结果输出:对于分类问题,根据n种分类结果进行简单投票表决确定最终分类结果,对于回归问题,对n个回归结果进行平均确定终值。

[0101] 本发明中,随机森林建模在R语言(R Development Core Team, 2013)中利用“randomForest”包(Liaw and Wiener, 2002)完成。在建模过程中,需完成三个参数的设置:数的总数(ntree),叶节点包含最少样本的个数(nodesize),随机选取变量的个数(mtry)。ntree的默认值是500,然而,为了得到更稳定的预测结果往往需要更多的数。本发明中设定ntree等于1000。nodesize使用默认值5,分类和回归问题中mtry分别设为4和6。

[0102] 步骤106:根据所述非采样点的土壤田间持水量深度分布函数模型参数和所述土壤田间持水量深度分布函数模型,确定非采样点土壤田间持水量预测值。

[0103] 利用上一步对研究区所模拟出来的四个模型分布参数SFC_a、SFC_b、k_a和k_b和公式1求取任何深度下的覆盖整个研究区的土壤田间持水量空间预测图。

[0104] 上述方法还包括:

[0105] 对所述非采样点土壤田间持水量预测值的预测精度进行评价。

[0106] 对预测精度进行评价时,可选取平均绝对误差(Mean Absolute Error, MAE)、均方

根误差 (Root Mean of Squared Error, RMSE)、决定系数 (Coefficient of Determination, R^2) 和林的一致性相关系数 (Lin's Concordance Correlation Coefficient, LCCC) 等四个参数:

$$[0107] \quad MAE = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i| \quad (2)$$

$$[0108] \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2} \quad (3)$$

$$[0109] \quad LCCC = \frac{2r\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\bar{Y} - \bar{X})^2} \quad (4)$$

$$[0110] \quad R^2 = \frac{\sum_{i=1}^n (X_i - \bar{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2} \quad (5)$$

[0111] 式中, X和Y分别为预测值和观测值, n为样点个数, \bar{X} 和 \bar{Y} 分别为预测值和观测值的均值, r为预测值和观测值之间的皮尔逊相关系数, σ_x 和 σ_y 分别为预测集和观测集的方差。

[0112] 其中: MAE用来评价预测值的偏差程度, 越接近0, 表示预测结果越好。RMSE用来对预测结果的总体准确性进行评价, 其值越小, 则表明模型预测精度越高。 R^2 用来评价模型的拟合优度, 其值越接近1, 表示模型的参考值越高。LCCC用来度量预测值和实测值在1比1线分布的程度, 越接近1, 预测值与观测值的吻合程度越高, 则表明模型预测能力越强。

[0113] 本发明中将深度分布函数模型与空间预测模型相结合的土壤有机碳和田间持水量的三维预测方法能够获得较高的模拟精度。精度评价结果表示模型方法可以较准确地模拟区域内土壤属性的三维分布, 可以为区域内土壤有机碳库的估算和水资源管理提供依据。环境景观复杂多变、土壤属性空间变异性较大的情况下, 引入土壤发生学知识所建立的三维土壤属性预测方法, 可以提高三维土壤属性预测精度。

[0114] 图3为本发明三维土壤田间持水量预测系统结构图。如图3所示, 一种三维土壤田间持水量预测系统, 包括:

[0115] 获取模块201, 用于获取多个采样点的土壤样本和多个环境数据。

[0116] 土壤田间持水量确定模块202, 用于根据各所述采样点的土壤样本, 确定多个土壤田间持水量。

[0117] 土壤田间持水量深度分布函数模型建立模块203, 用于根据各所述土壤田间持水量, 建立土壤田间持水量深度分布函数模型。

[0118] 采样点的深度分布函数模型参数确定模块204, 用于根据所述土壤田间持水量深度分布函数模型, 确定采样点的土壤田间持水量深度分布函数模型参数。

[0119] 非采样点的深度分布函数模型参数确定模块205, 用于根据所述采样点的土壤田间持水量深度分布函数模型参数和所述环境数据, 预测非采样点的土壤田间持水量深度分布函数模型参数。

[0120] 预测模块206, 用于根据所述非采样点的土壤田间持水量深度分布函数模型参数和所述土壤田间持水量深度分布函数模型, 确定非采样点土壤田间持水量预测值。

[0121] 所述土壤田间持水量深度分布函数模型建立模块203, 具体包括:

[0122] 土壤田间持水量深度分布函数模型建立单元,用于根据分类指数函数和所述土壤田间持水量,建立土壤田间持水量深度分布函数模型。

[0123] 所述非采样点的深度分布函数模型参数确定模块205,具体包括:

[0124] 定量关系模型建立单元,用于采用随机森林模型建立所述采样点土壤田间持水量深度分布函数模型参数和所述环境数据的定量关系模型。

[0125] 非采样点的深度分布函数模型参数确定单元,用于根据所述定量关系模型,预测非采样点的土壤田间持水量深度分布函数模型参数。

[0126] 本说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的系统而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。

[0127] 本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处。综上所述,本说明书内容不应理解为对本发明的限制。



图1

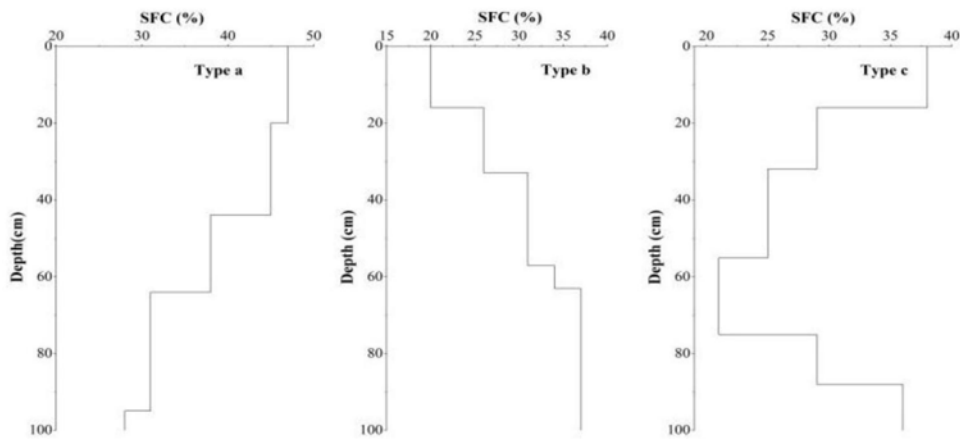


图2



图3