

区域土壤重金属空间分布驱动因子影响力比较案例分析

陈运帷，王文杰，师华定，王明浩，许超*

中国环境科学研究院，北京 100012

摘要：土壤重金属空间分布受到自然和人为驱动因子的共同影响，识别并评价各驱动因子对土壤重金属空间分布的影响力，对解译土壤重金属污染源、重金属迁移、探明重金属空间分布模式有重要意义。为研究不同背景下各驱动因子对土壤重金属空间分布影响的差异，并异质背景下的高影响力因子，采用地理探测器和随机森林模型，以贵州省安顺市作为土壤重金属高背景值典型地区、辽宁省葫芦岛市作为人类活动高强度典型地区，探讨市域尺度下土壤 pH、地面累年值年值气温、坡度、海拔、地面累年值年值降水量(08:00)、夜间灯光指数和县 GDP 等 7 种驱动因子的空间分异性，及其对研究区内 As、Cd、Cr、Hg、Pb 等 5 种重金属空间分布的单独及交互驱动力强度。结果表明：安顺市 5 种重金属的质量分数平均值均高于葫芦岛市，但 $w(\text{Hg})$ 差异不明显；在市域内采样点平均间距为 10 km 的尺度下，安顺市和葫芦岛市土壤重金属分布受到自然因子的影响程度强于人为因子；在安顺市和葫芦岛市，土壤 pH、海拔和县 GDP 是 7 项驱动因子中在异质环境背景下驱动能力较稳定，并且是对土壤重金属分布影响能力最大的 3 个因子，适宜在上述地区作为土壤重金属含量多元非线性回归模型的通用参数。研究显示，该研究所采用的方法可以应用于评价其他因子对土壤重金属含量空间分布影响力，以及用于土壤重金属含量回归方程自变量的选择。

关键词：土壤污染；重金属；驱动因子；地理探测器；随机森林模型

中图分类号：X144

文章编号：1001-6929(2019)00-00-00

文献标志码：A

DOI：10.13198/j.issn.1001-6929.2018.12.06

Comparative Case Study on the Influence of Spatial Distribution of Heavy Metals in Regional Area

CHEN Yunwei, WANG Wenjie, SHI Huading, WANG Minghao, XU Chao*

Chinese research academy of environmental sciences, Beijing 100012

Abstract: The spatial distribution of soil heavy metals is influenced by natural and human driving factors. Identifying and evaluating the influence of these factors is of great significance for identifying the sources, migration and spatial distribution patterns of heavy metals in soil. To research the effects of various driving factors on the spatial distribution of heavy metal in soil under different backgrounds, and to find the high-impact factors under heterogeneous backgrounds, Anshun City and Huludao City, which are respectively characterized with high background value of heavy metals and high intensity of human activity, were chosen as the study area. The driving factors including pH, annual ground temperature, annual ground precipitation (08:00), elevation, slope, night-time light index and GPD were taken into account. Methods including geodetector and random forests were adopted to analyze the individual and interactive effects of driving factors on the spatial distribution of As, Cd, Cr, Hg and Pb in the typical regions. Results showed that the mean concentration of each heavy metal in Anshun City was higher than Huludao City, but the distinction of Pb concentration was not significant. In the average sampling density of 10 km, the influence of natural driving factors was stronger than human factors in typical regions. Soil pH, elevation and GPD were the strongest driving factors that greatly affected the spatial distribution of heavy metals in soil among the referred 7 factors. Meanwhile, they were relatively stable in the heterogeneous backgrounds, and were suitable to be used as the general parameters in the multiple nonlinear regression model for calculating heavy metal concentrations in soils in the study area. The method adopted in this study can be referred to evaluate the influence of other factors on the spatial distribution of heavy metal concentrations in soil. Moreover, it is useful to select the independent variables in the regression equation to predict heavy metal concentrations

Keywords: soil contamination; heavy metals; driving factor; geodetector; random forests model

收稿日期: 2018-05-12 修订日期: 2018-11-14

作者简介: 陈运帷(1985-), 男, 北京人, beiwolf007@163.com.

*责任作者: 许超(1985-), 男, 河北石家庄人, 助理研究员, 硕士, 主要从事环境遥感监测与区域污染状况分析研究, xuchao@craes.org.cn

基金项目: 国家重点研发计划项目(No.2018YFF0213401)

Supported by the National Key Research and Development Program of China(No.2018YFF0213401)

土壤重金属污染是土壤污染的一种主要形式, 近年来常有农产品重金属含量超标的新闻报道, 使得全社会对于土壤重金属污染的关注不断增加. 不同于污染场地调查, 大范围的非建设用地土壤污染情况调查受到人力、财力限制, 不可能通过高密度采样来获取调查结果. 因此, 通过使用地理统计方法结合适当数量和分布的样本, 插值获取区域土壤污染因子的空间分布和含量是最为有效和常用的技术方法.

空间插值方法种类繁多, 较为常用的是反距离权重法、径向基函数插值法和克里金插值法等. 这些插值方法基于不同的假设条件和数据特征, 数据特征不符或插值参数设置不合适会导致插值结果产生较大误差^[1]. 土壤污染的形成原因复杂, 受到污染源分布、重金属存在形式^[2]、土壤质地和土壤母质等多种因素的影响, 不同地区、不同污染物的空间分布特征和规律并不相同. 李湘凌等^[3]在合肥义州地区对土壤重金属含量空间分布的研究表明, Cu、Pb、As 元素采用普通克里金法进行插值结果最优, 而 Zn 元素采用反距离加权法最优, 对于 Cd、Hg 元素则采用径向基函数插值法最优. 袁峰等^[4]在对安徽省铜陵市多种重金属元素进行插值方法比较的研究中也发现, 不同重金属元素的最佳插值方法(均方差最小)并不一致. 谢云峰等^[5]通过对北京市通州区土壤中 Cd 元素多种插值方法的比较研究中发现, 由于土壤重金属污染常呈现小范围内高浓度值集中出现的特征, 所以克里金插值法存在的平滑效应导致其插值结果在极值计算和高浓度值区域分布预测上表现不佳, 精度低于反距离权重和径向基函数法.

克里金插值法在土壤重金属空间分布研究中的应用由最开始的依靠地理自相关性进行线性插值, 发展为结合地理、环境信息对污染物进行线性回归以提升插值精度, 常用的方法是协同克里金插值法和回归克里金插值法^[6]. 如庞凤等^[7]以表层土壤 Cu 元素作为协同因子, 通过协同克里金插值法预测深层土壤中的 $w(\text{Cu})$; 姜勇等^[8]建立表层土壤与深层土壤之间 $w(\text{Zn})$ 的线性回归方程, 计算深层土壤 $w(\text{Zn})$ 回归值并进行普通克里金插值获取空间重金属含量及其分布模式; Hengl 等^[9]采用回归克里金插值法, 借助地理和环境因子, 对土壤有机物含量、土壤表面温度进行了空间插值. 在上述研究中, 与普通克里金插值法相比, 考虑协同变量的克里金插值法的插值精度提高了 3%~10%, 且达到相同插值精度所需的采样点数量明显减少. 可见, 合理使用协同变量辅助克里金插值能够提高空间插值的预测精度, 同时也能降低实现高精度预测所需的采样点数量, 降低调查成本.

使用协同变量辅助克里金插值的关键在于选取与插值对象相关的协同变量, 但土壤污染成因复杂, 同时受到成土过程中自然因子和人类开发利用时的人为活动因子的共同影响^[10-12]. 我国幅员辽阔, 不同地区自然因素各异, 人类活动强度分布不均. 虽然在我国西南地区土壤重金属含量背景值高于其他地区是一个客观事实^[13], 但近几十年的快速发展使得人为因子对土壤重金属含量的影响大大增加. 已有关于自然和人为因子对土壤影响力的分析多基于多元统计方法, 关注于重金属含量之间的统计关系. 如 Wong 等^[14]使用

相关系数分析和主成分分析法对珠三角地区其他农用地、水稻土和自然土壤中的 Co、Cu、Pb 和 Zn 进行统计分析, 通过主成分分析结果结合污染物来源特征, 推测珠三角地区 Pb 主要来源于汽车尾气排放, Cu 和 Zn 的富集与农业耕作活动有关; 朱求安等^[15]应用地形数据作为协同因子, 配合径向基函数对中国南部地区酸雨空间分布进行插值估计, 其精度高于单纯使用普通克里金和样条函数插值; Liénard 等^[16]使用相关系数、主成分分析和半变异函数等统计方法, 对比利时 Sclaigneaux 市某锌矿石处理厂旧址周边 3 km 区域内 7 种微量元素和 6 种主量元素进行分析; 并以距污染源距离、土地利用方式和风向为辅助变量, 对上述元素进行协方差分析, 得出土壤中 Cd、Cu、Pb、Zn 含量的空间分布主要受风向影响. 通过主成分分析等统计学方法, 分析污染物之间的数据特征并结合周边污染源的污染特征可以倒推污染源. 但这些方法对污染物空间分布驱动能力缺少定量分析, 且无法对多因子交互作用影响力进行评估.

鉴于此, 该研究从影响因子与土壤重金属含量统计关系的角度出发, 选取位于我国东北地区的辽宁省葫芦岛市和西南地区的贵州省安顺市为两个典型区域, 借助地理探测器和随机森林模型两种工具, 探讨市域尺度下两个典型区域内自然因子和人为因子对土壤重金属元素(As、Cd、Cr、Hg、Pb)空间分布的驱动能力和特征, 并试图寻找在不同地理、人文条件下均对土壤重金属含量具有较大影响的因子, 这种因子即可作为土壤重金属多元非线性回归方程的重要自变量. 研究结果既可作为土壤调查分层布点中层次划分因子, 也可以为区域土壤重金属含量非线性回归模型的建立提供参考.

1 材料与方法

1.1 研究区概况

1.1.1 葫芦岛市

葫芦岛市是辽宁省西部的海滨城市, 下辖连山区、龙港区、南票区、兴城市(县级)、绥中县和建昌县, 总面积 10400 km², 海岸线长 258 km. 年均气温为 9.0 °C, 年平均降水量为 563~642 mm^[17]. 辽宁省土壤中 5 种重金属(As、Cd、Cr、Hg、Pb)含量背景值在全国均处于较低水平, 但葫芦岛市域土壤和农产品重金属含量的调查结果显示, 葫芦岛锌厂、杨家杖子矿周边和葫芦岛市区的土壤的重金属含量远高于其含量背景值^[18,19], 重金属通过食物链传导造成二次污染的风险较大^[20]. 另有研究显示, 葫芦岛本地的蔬菜中重金属含量高于其他地方运输进入葫芦岛的蔬菜.

选取葫芦岛市作为典型地区是因为其所处区域土壤重金属含量背景值较低, 但通过上述研究研究结果可知, 在葫芦岛锌厂等人类活动强度较大的区域已出现土壤重金属含量高于周边区域的情况. 因此, 该研究以葫芦岛市体现市域尺度下人为因子对土壤重金属含量分布的影响能力.

1.1.2 安顺市

安顺市位于贵州省中西部, 下辖西秀区、平坝区、普定县、镇宁布依族苗族自治县、关岭布依族苗族自治县、紫云苗族布依族自治县等 6 个区(县), 总面积 9 267 km², 总人口 300×10⁴ 人.

安顺市是世界上典型的喀斯特地貌集中区域, 全境海拔 560~1 500 m, 平均海拔为 1 102~16 94 m. 全年雨量充沛, 年均降水量 1 360 mm, 年均气温 14 °C.

安顺市矿产资源丰富, 以煤矿开采为优势产业. 根据张玉涛等^[22]对安顺市郊区某煤矿区周边菜园土壤重金属的研究和董钧铭等^[23]对于安顺市多种类型蔬菜含 Pb 含量的综合分析, 安顺市土壤中除 Pb 含量是贵州省土壤重金属含量背景值的 3 倍外, 其他重金属含量均接近含量背景值, Pb 含量较高可能是由当地燃煤

粉尘沉降造成的。值得注意的是，王志瑞^[24]对安顺市一级水源保护区内土壤重金属含量的研究显示，在人类活动强度较低的水源保护区内的土壤仍然存在 Cd 和 As 超标现象。

贵州省土壤 5 种重金属含量背景值在全国属于较高水平。安顺市自然资源开采主要以煤矿开采和洗选为主，对土壤重金属污染比较严重的有色金属采选、冶炼等行业占比很低。因此，该研究将安顺市作为土壤重金属含量背景值高、但人类活动影响较为单一的对比区域。

1.2 数据获取

为获取研究区土壤样品，于 2012 年在葫芦岛市和安顺市市域内，按照 10 km×10 km 网格中心方式进行布点采样，分别在葫芦岛市、安顺市布设土壤采样点 95、76 个。在 20 m×20 m 内采用五点梅花采样法，采集土壤表层 0~20 cm 土壤混合制成样品。各重金属总量测试方法参考 GB 15618—1995《土壤环境质量标准》。



图 1 研究区土壤样点分布

Fig.1 Map of soil sample point in study areas

两个典型地区 2012 年县 GDP 数据分别来源于《2013 年辽宁省统计年鉴》和《2013 年安顺市统计年鉴》。

研究区内 1981—2010 年地面累年值年值气温(简称“年值气温”)和地面累年值年值降水量(08:00)(简称“年值降水量”)气象站点数据来源于中国气象网(<http://data.cma.cn/site/index.html>)，年值气温和年值降水量栅格数据根据上述数据样条插值生成。

海拔数据源于 SRTM90 m 数字高程模型，并在其基础上使用 ArcGIS 相关工具制作海拔分类和坡度数据。

夜间灯光指数是由美国国防气象卫星计划(DMSP)所发射的卫星所获取的，该卫星由于处于太阳同步低地球轨道，可以获取地球 20:30—21:30 的卫星影像^[25]，夜间灯光指数已广泛用于表征人类活动和经济发展水平^[26]。研究区 1km×1km 分辨率 2012 年 DMSP/OLS 稳定灯光影像下载自地理国情监测云平台(<http://www.dsac.cn>)。

虽然土壤样品采集时间为 6 a 之前(2012 年)，但土壤重金属空间分布模式的演变受气温、降水量等相对稳定的驱动因子所影响，变化速度较大气和水污染物扩散缓慢。为配合土壤样品采集时间，在该研究中，反映人类活动强度的 GDP 和夜间灯光指数数据也均采用了采样年(2012 年)的数据。因此所揭示的研究区土壤重金属空间分布与驱动因子之间的关系能够适用于当前及未来土壤重金属含量预测及演绎需求。

土壤重金属污染的产生原因总体上可归纳为自然因素和人为因素，自然因素主要包括气候、母质、生

物、地形等，人为因素主要包括农业开发、工业污染和人居活动等。根据土壤发生、土壤重金属生成和迁移规律并综合数据的可获取性，选取年值气温和年值降水量表征气候因子，海拔和坡度表征地形因子，夜间灯光指数表征人居活动因子，县 GDP 表征总体人类活动强度。

1.3 研究方法

1.3.1 地理探测器

地理探测器是中科院地理科学与资源研究所王劲峰等^[27, 28]提出的一种评价样本空间分异性和自变量分异性对因变量分异性影响力的统计学方法，其核心思想是基于这样的假设：如果某个自变量对某个因变量有重要影响，那么自变量和因变量的空间分布应该具有相似性。

地理探测器包含 4 个探测器：分异及因子探测器、交互作用探测器、风险区探测器、生态探测器。

分异及因子探测器：探测变量 Y 的空间分异性，以及探测某变量 X 在多大程度上解释了变量 Y 的空间分异性。分异性强度用 q 值度量，表达式^[29]为

$$q = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2} = 1 - \frac{SSW}{SST} \quad (1)$$

$$SSW = \sum_{h=1}^L N_h \sigma_h^2 \quad (2)$$

$$SST = N \sigma^2 \quad (3)$$

式中： L 为变量 Y 或变量 X 的分层(strata)数，即分类或分区； N_h 和 N 分别为层 h 和全区的单元数； σ_h^2 和 σ^2 分别为层 h 和全区的 Y 值的方差；SSW、SST 分别为层内方差之和(within sum of squares)、全区总方差(total sum of squares)。

q 的值域为[0, 1]，值越大说明 Y 的空间分异性越明显；如果分层是由自变量 X 生成的，则 q 值越大表示自变量 X 对变量 Y 的解释力越强，反之则越弱。极端情况下， q 值为 1 表明变量 X 完全控制了 Y 的空间分布， q 值为 0 则表明变量 X 与 Y 没有任何关系， q 值表示 X 对 Y 的解释率为 $100 \times q\%$ 。分异探测器可以指示区域内变量空间异质性的强度^[30]，高强度的空间分异性是空间插值的前提。

以年值气温、年值降水量、海拔、坡度、夜间灯光指数和市内各县 GDP 作为探测变量 X ，分别对 5 种重金属含量 Y 进行分异性探测，探测某单一变量对重金属空间分布格局的影响力。

交互作用探测器：识别不同风险变量 X_i 之间的交互作用，即评估变量 X_1 和 X_2 共同作用时是否会增加或减弱对因变量 Y 的解释力，或这些变量对 Y 的影响是否为相互独立的。评估方法：首先分别计算两个变量 X_1 和 X_2 对 Y 的 q 值—— $q(X_1)$ 和 $q(X_2)$ ，再计算二者交互时的 q 值—— $q(X_1 \cap X_2)$ ，最后对 $q(X_1)$ 、 $q(X_2)$ 与 $q(X_1 \cap X_2)$ 进行比较。两个变量之间的关系可分为 5 类(见图 2)。

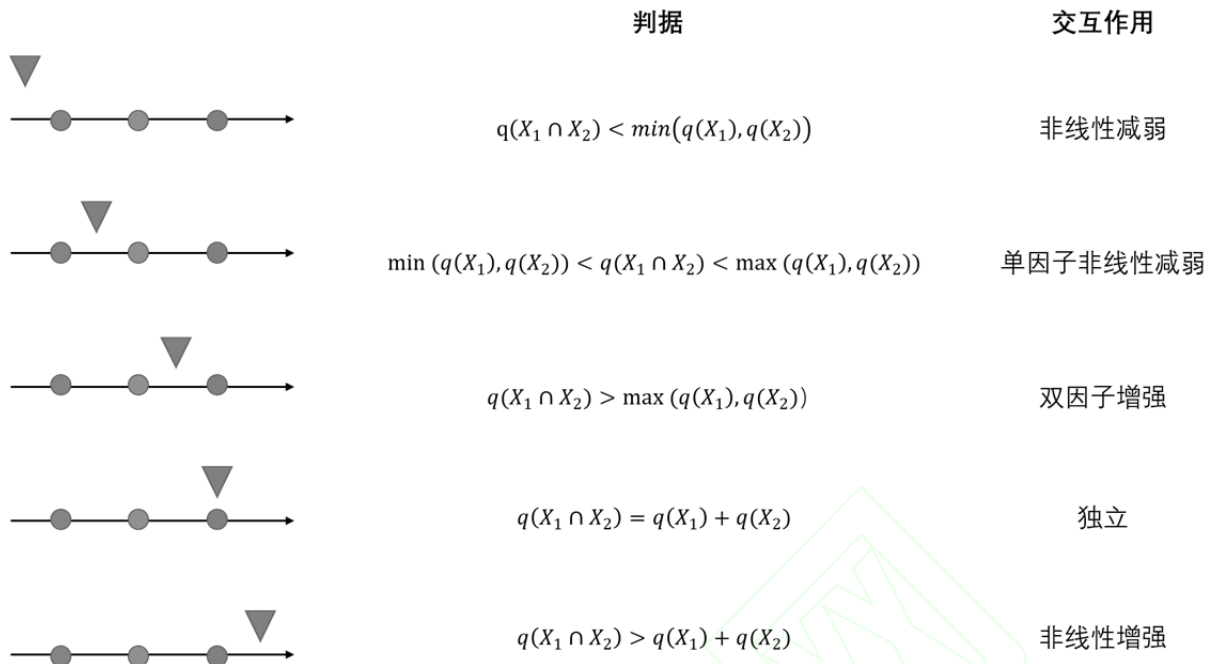


图2 两个自变量对因变量交互作用的类型

Fig.2 Types of interaction between two covariates

地理探测器对变量分异性、变量间交互作用分析基于数据统计特征所形成的层次，这种层次的划分既可以通过专业知识人工识别，也可以根据数据聚类分析获得。这种特性使地理探测器的使用领域跳出了地理学的范畴，而被许多学者应用于公共医疗、环境保护等领域^[31-34]。以年值气温、年值降水量、海拔、坡度、夜间灯光指数和各县 GPD 作为探测变量 X ，分别对 5 种重金属含量 Y 进行分异性探测。探测在两个变量的交互作用下，各“变量对”对研究区土壤重金属空间分布格局的影响力。

该研究未使用风险区探测器和生态探测器，不予赘述。

1.3.2 随机森林模型

随机森林模型(RF)是由 Breiman^[35]探索出来的一种机器学习算法，其本质是以决策树为基础的模型，多用于数据分类。为了保证分类的完整性，在特征采样中，以随机的形式从 M 个特征中选择 m 个 ($m \ll M$) 组成一个决策树，将数百个随机决策树组合为随机森林。剪枝作为随机森林实现过程中的一个重要步骤，是将期望错误率最小和最小描述长度原则相结合，防止最终的分析结果出现过分拟合的现象^[36]。通过标记好的训练数据对完成剪枝的随机森林进行训练，确定每一棵决策树决策节点的阈值。完成模型训练后，每个决策树对待分类数据进行分类判断，随机森林根据模型中每棵树分类的“投票”结果，选择票数最多的类别为该条数据的分类结果。

随机森林模型在进行决策树节点变量选择和阈值确定时，经常采用的决策方式是基于置换检验和基于最小 Gini 系数这两种方式，该研究选取最小 Gini 系数法。Gini 系数是一种对系统熵的有效近似，对于决策树这种二分结构，在决策树节点 τ 的 Gini 系数 ($i(\tau)$) 计算方法为

$$i(\tau) = 1 - p_1^2 - p_0^2 \quad (4)$$

式中： p_1 、 p_0 为某一变量在任意阈值 k 下二分后，两个子类样本数占原样本总数的比例。

变量 x 在决策树节点 τ 处根据阈值 k 分为 l 和 r 两个部分后，整个系统 Gini 值的变化量 ($\Delta i(\tau)$) 为

$$\Delta i(\tau) = i(\tau) - p_l i(\tau_l) - p_r i(\tau_r) \quad (5)$$

式中： P_l 、 P_r 为两个子类样本数 l 和 r 分别占原样本总数的比例。

在节点 τ 中对变量 x 进行穷举，并分别计算 x 的所有阈值选取所能得出的 Gini 值变化量。选取 Gini 值下降最多时所对应的“变量-阈值”对，就可以得到主导节点 τ 分支的变量及划分阈值。作为这种方法的副产物，对随机森林节点各变量 Gini 值变化量进行统计，即可明确随机森林模型在进行模型训练的过程中主要参考了哪些变量，反映了每个变量对数据分布驱动力的大小。

随机森林模型以其对异常值不敏感和对空值的容忍度高的优势，特别适用于环境领域的污染物分析或生态领域的生态生产力等，在数据获取时离群值和空值较为常见的研究领域。随机森林模型可用于分类也可以用于回归，体现自变量与因变量之间的非线性拟合关系，并给出各自变量对因变量的影响程度，这也十分适用于解释环境、土壤等影响因子间的复杂的非线性交互作用关系^[37]。

该研究使用 Python 语言的 scikit-learn 工具包的随机森林模块，以重金属含量为因变量、各驱动因子为自变量，设计决策树数量为 1000，进行随机森林回归训练，输出各驱动因子在稳定回归关系中的权重。随机森林变量权重值总和为 1，每个因子权重数值越大，表示该因子对研究区形成目前的土壤重金属空间分布状态的影响力越大。

2 结果与分析

2.1 描述性统计

统计结果(见图 2)显示，安顺市域内 5 种重金属质量分数平均值均高于葫芦岛市，其中， $w(\text{As})$ 、 $w(\text{Cd})$ 、 $w(\text{Cr})$ 、 $w(\text{Pb})$ 显著较高， $w(\text{Hg})$ 未通过显著性检验。葫芦岛市的杨家杖子矿、锦西炼油厂和葫芦岛硫酸厂等工矿企业均已开工(采)近百年，而安顺市占优势的煤矿产业大多是 1949 年以后陆续开始开采的。相较而言，葫芦岛市的人类活动强度和历史累积性影响明显高于安顺市。但土壤重金属质量分数描述性统计显示，安顺市 4 种重金属含量明显高于葫芦岛市。

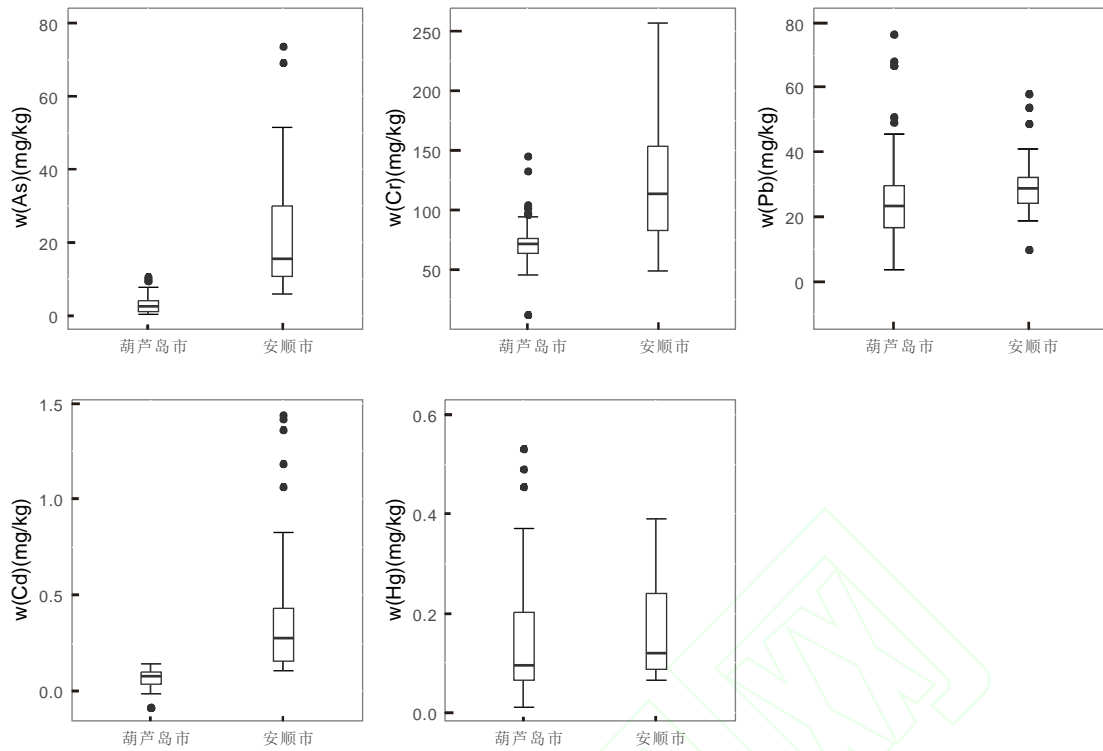


图3 研究区5种重金属质量分数箱型图

Fig.3 The box-whisker plot of five heavy metals concentration in research area

2.2 重金属含量分布驱动力地理探测器分析

采用地理探测器进行驱动因子影响力分析要求驱动因子为类型数据，需对连续型驱动因子数据进行离散化处理，所采用离散化方法应能够体现出指标间的最大差异化水平，以使地理探测器分析结果更加精确，不同指标间阈值不存在相互影响关系。根据 Cao 等^[38]的研究，不同的离散化方法和分类数量均会显著影响地理探测器的分析精度。参考其研究结果，该研究对年值气温和年值降水量数据采用等分法，将二者分别划分为7个类型。对海拔、坡度、夜间灯光指数和土壤 pH 采用 Jenks Natural Breaks 法各分为7类，其中，坡度和安顺市夜间灯光指数分类中因第6、7类的阈值相同，故合并为6类。各因子分类阈值详见表1。

表1 驱动因子离散化阈值

Table 1 The threshold value of driving factors discretization

城市		类别						
		1	2	3	4	5	6	7
安顺市	海拔	787	900	1090	1200	1300	1400	1500
	坡度	1.65	6.46	11.15	16.29	21.28	35.78	
	夜间灯光指数	6.0	7.0	10.0	11.0	14.0	16.0	
	土壤 pH	4.7	5.0	5.4	5.7	6.0	6.2	6.4
葫芦岛	海拔	20	89.0	175.0	288.0	399.0	550.0	676.0
	坡度	1.02	3.06	5.83	9.70	20.13	27.88	
	夜间灯光指数	5.0	8.0	11.0	18.0	33.0	56.0	62.0
	土壤 pH	4.59	5.2	5.8	6.5	7.10	7.5	8.0

将各采样点检测的 5 种重金属质量分数与采样点位置 6 种驱动因子分类值分别导入地理探测器进行计算，结果如表 2~7 所示。

表 2 研究区域土壤重金属含量驱动力强度

Table 2 Driving force strength of soil heavy metal concentration in typical areas

重金属	城市	统计量	年降水量	年气温	海拔	夜间灯光指数	坡度	县 GDP	土壤 pH
As	安顺市	<i>q</i>	0.116	0.190	0.215	0.176	0.093	0.067	0.145
		<i>P</i>	0.767	0.480	0.309	0.843	0.938	0.871	0.640
	葫芦岛市	<i>q</i>	0.148	0.099	0.137	0.086	0.087	0.074	0.120
		<i>P</i>	0.011*	0.086	0.026*	0.131	0.371	0.284	0.130
Cd	安顺市	<i>q</i>	0.179	0.227	0.235	0.039	0.294	0.179	0.185
		<i>P</i>	0.532	0.377	0.237	0.975	0.390	0.420	0.511
	葫芦岛市	<i>q</i>	0.125	0.116	0.114	0.086	0.067	0.086	0.333
		<i>P</i>	0.029*	0.043*	0.047*	0.692	0.521	0.308	0.000**
Cr	安顺市	<i>q</i>	0.199	0.102	0.167	0.080	0.349	0.143	0.113
		<i>P</i>	0.473	0.812	0.474	0.995	0.452	0.627	0.783
	葫芦岛市	<i>q</i>	0.048	0.046	0.122	0.058	0.041	0.050	0.051
		<i>P</i>	0.487	0.505	0.107	1.000	1.000	1.000	0.806
Hg	安顺市	<i>q</i>	0.155	0.174	0.277	0.223	0.109	0.172	0.098
		<i>P</i>	0.615	0.508	0.184	0.817	0.941	0.480	0.840
	葫芦岛市	<i>q</i>	0.070	0.062	0.060	0.070	0.017	0.016	0.038
		<i>P</i>	0.244	0.318	0.348	0.559	0.982	0.928	0.679
Pb	安顺市	<i>q</i>	0.103	0.340	0.234	0.257	0.031	0.273	0.098
		<i>P</i>	0.825	0.103	0.362	0.983	1.000	0.393	0.842
	葫芦岛市	<i>q</i>	0.034	0.039	0.036	0.097	0.160	0.027	0.060
		<i>P</i>	0.680	0.612	0.699	0.612	0.686	0.924	0.343

注：*表示 $P < 0.05$ ；**表示 $P < 0.01$ 。下同。

表3 典型地区土壤w(As)驱动力交叉影响强度

Table 3 The cross-influence strength of driving force of soil As concentration in typical areas

城市	项目	年降水量	年气温	海拔	夜间灯光指数	坡度	县 GDP	土壤 pH
安顺市	年降水量	0.116						
	年气温	0.930	0.190					
	海拔	0.894	0.659	0.215				
	夜间灯光指数	0.568	0.412	0.551	0.176			
	坡度	0.307	0.389	0.630	0.358	0.093		
	县 GDP	0.571	0.619	0.539	0.531	0.523	0.067	
	土壤 pH	0.785	0.525	0.807	0.743	0.358	0.500	0.145
	葫芦岛市	年降水量	0.148					
	年气温	0.258	0.099					
	海拔	0.292	0.302	0.137				
夜间灯光指数	0.283	0.233	0.251	0.086				
	坡度	0.426	0.317	0.322	0.210	0.087		
	县 GDP	0.256	0.170	0.340	0.225	0.281	0.074	
	土壤 pH	0.447	0.378	0.494	0.328	0.261	0.301	0.120

表4 典型地区土壤w(Cd)驱动力交叉影响强度

Table 4 The cross-influence strength of driving force of soil Cd concentration in typical areas

城市	项目	年降水量	年气温	海拔	夜间灯光指数	坡度	县 GDP	土壤 pH
安顺市	年降水量	0.179						
	年气温	0.968	0.227					
	海拔	0.490	0.460	0.235				
	夜间灯光指数	0.293	0.262	0.268	0.039			
	坡度	0.490	0.977	0.488	0.326	0.294		
	县 GDP	0.405	0.967	0.478	0.230	0.497	0.179	
	土壤 pH	0.511	0.964	0.985	0.575	0.979	0.960	0.185
	葫芦岛市	年降水量	0.125					
	年气温	0.204	0.116					
	海拔	0.289	0.264	0.114				
夜间灯光指数	0.200	0.226	0.365	0.086				
	坡度	0.407	0.370	0.352	0.218	0.067		
	县 GDP	0.255	0.195	0.289	0.200	0.311	0.086	
	土壤 pH	0.525	0.535	0.564	0.456	0.502	0.438	0.333

表5 典型地区土壤w(Cr)驱动力交叉影响强度

Table 5 The cross-influence strength of driving force of soil Cr concentration in typical areas

城市	项目	年降水量	年气温	海拔	夜间灯光指数	坡度	县 GDP	土壤 pH
安顺市	年降水量	0.199						
	年气温	0.899	0.102					
	海拔	0.693	0.346	0.167				
	夜间灯光指数	0.434	0.253	0.309	0.080			
	坡度	0.582	0.827	0.655	0.515	0.349		
	县 GDP	0.577	0.895	0.568	0.299	0.533	0.143	
	土壤 pH	0.639	0.934	0.952	0.512	0.839	0.826	0.113
	葫芦岛市	年降水量	0.048					
	年气温	0.111	0.046					
	海拔	0.336	0.310	0.122				
	夜间灯光指数	0.240	0.221	0.279	0.058			
	坡度	0.208	0.163	0.253	0.169	0.041		
	县 GDP	0.245	0.289	0.361	0.372	0.303	0.050	
	土壤 pH	0.394	0.303	0.496	0.358	0.173	0.386	0.051

表6 典型地区土壤w(Hg)驱动力交叉影响强度

Table 6 The cross-influence strength of driving force of soil Hg concentration in typical areas

城市	项目	年降水量	年气温	海拔	夜间灯光指数	坡度	县 GDP	土壤 pH
安顺市	年降水量	0.155						
	年气温	0.800	0.174					
	海拔	0.739	0.430	0.277				
	夜间灯光指数	0.635	0.403	0.555	0.223			
	坡度	0.458	0.563	0.594	0.393	0.109		
	县 GDP	0.605	0.510	0.573	0.392	0.436	0.172	
	土壤 pH	0.869	0.705	0.805	0.593	0.460	0.780	0.098
	葫芦岛市	年降水量	0.070					
	年气温	0.113	0.062					
	海拔	0.261	0.217	0.060				
	夜间灯光指数	0.261	0.241	0.285	0.070			
	坡度	0.235	0.210	0.188	0.153	0.017		
	县 GDP	0.137	0.118	0.171	0.229	0.137	0.016	
	土壤 pH	0.332	0.299	0.307	0.326	0.328	0.226	0.038

表7 典型地区土壤w(Pb)驱动力交叉影响强度

Table 7 The cross-influence strength of driving force of soil Pb concentration in typical areas

城市	项目	年降水量	年气温	海拔	夜间灯光指数	坡度	县 GDP	土壤 pH
安顺市	年降水量	0.103						
	年气温	0.924	0.340					
	海拔	0.674	0.596	0.234				
	夜间灯光指数	0.629	0.465	0.545	0.257			
	坡度	0.278	0.702	0.485	0.441	0.031		
	县 GDP	0.627	0.677	0.604	0.474	0.387	0.273	
	土壤 pH	0.899	0.876	0.659	0.598	0.601	0.711	0.098
	年降水量	0.034						
葫芦岛市	年降水量	0.034						
	年气温	0.194	0.039					
	海拔	0.396	0.237	0.036				
	夜间灯光指数	0.283	0.307	0.254	0.097			
	坡度	0.371	0.354	0.371	0.380	0.160		
	县 GDP	0.191	0.347	0.350	0.277	0.437	0.027	
	土壤 pH	0.347	0.477	0.457	0.366	0.464	0.275	0.060

地理探测器因子驱动力分析结果(见表 2)显示,多数驱动因子的地理探测器分析结果—— q 统计量的 P 值大于 0.05,未能通过显著性检验,这可能与采样精度设置有关.根据 Lu 等^[20]对于葫芦岛锌厂周边土壤重金属污染的研究,企业对于周边土壤重金属含量影响范围不超过 10 km.该研究中土壤采样点平均间距约为 10 km,该尺度避免了分析结果受到人为活动影响,但同时损失了部分自然因子的差异信息,造成分析结果的显著性不明显.

值得注意的是,地理探测器驱动力分析结果中,通过显著性检验的几个“污染物-因子”对都出现在葫芦岛市,分别是“As-年降水量”、“As-海拔”、“Cd-年降水量”、“Cd-海拔”、“Cd-年气温”、“Cd-土壤 pH”.这表明相较于 Cr、Hg 和 Pb,自然因子对于 As 和 Cd 在土壤中分布格局的影响更为明显和确定.与安顺市相比,葫芦岛市年降水量、海拔、坡度等自然因子的地域差异性较小,相似自然条件的地域成片分布,这也是葫芦岛市土壤 As、Cd 空间分异特征更为明显的原因之一.

从两个典型地区 5 种土壤重金属空间分布驱动因子交互作用分析结果中分别提取自然因子、人为因子中 q 统计量最高的因子对.结果(见表 8)显示,与葫芦岛市相比,安顺市土壤重金属分布驱动因子交互作用影响力较高,自然因子影响力与人为因子影响力之间的差异也比较明显.这表明在市域尺度下以安顺市土壤重金属空间分布更多受到自然因子的影响,而葫芦岛市人为和自然因子都对土壤重金属分布影响造成一定影响.

对最高影响力因子对进行分析发现,土壤 pH、海拔和地区 GDP 是出现频率为前三的驱动因子,且在安顺市和葫芦岛市均有出现.这表明虽然两个典型地区在自然条件、人类活动强度、土壤重金属来源等方面有较大差异,但土壤 pH、海拔和地区 GDP 值对两个典型区内土壤重金属空间分布较强的影响力并未因上述条件的变化而发生变化.

表 8 不同重金属驱动因子最强交叉影响力

Table 8 The strongest cross influence of 5 heavy metal driving factors in typical areas

重金属	城市	最高影响力的自然因子	自然因子交叉影响力	最高影响力的人为因子	人为因子交叉影响力
As	安顺市	年降水量-年气温	0.930	夜间灯光指数-土壤 pH	0.743
	葫芦岛市	海拔-土壤 pH	0.494	夜间灯光指数-土壤 pH	0.328
Cd	安顺市	海拔-土壤 pH	0.985	县 GDP-土壤 pH	0.960
	葫芦岛市	海拔-土壤 pH	0.564	夜间灯光指数-土壤 pH	0.456
Cr	安顺市	海拔-土壤 pH	0.952	县 GDP-土壤 pH	0.826
	葫芦岛市	海拔-土壤 pH	0.496	县 GDP-土壤 pH	0.386
Hg	安顺市	年降水量-土壤 pH	0.869	县 GDP-土壤 pH	0.780
	葫芦岛市	年降水量-土壤 pH	0.332	夜间灯光指数-土壤 pH	0.326
Pb	安顺市	年气温-年降水量	0.924	县 GDP-土壤 pH	0.711
	葫芦岛市	年气温-土壤 pH	0.477	坡度-夜间灯光指数	0.380

2.3 随机森林模型驱动力分析

随机森林模型中各因子权重如表 9 所示. 在两个研究内影响土壤重金属空间分布的主要因子是土壤 pH、年气温、海拔等自然因子. 个别自然因子对土壤重金属空间分布的影响力接近或超过 0.5(安顺市海拔对 Cd 和 Cr 的影响力、葫芦岛土壤 pH 对 Cd 的影响力), 表明这些自然因子对土壤重金属的空间分布有较大影响, 这与地理探测器所得结论相同. 但相较于地理探测器, 随机森林模型对人为因子影响力的评价结果较低, 鲜有影响力超过 0.1 的因子, 这与随机森林模型对因子连续数据离散化的方式有关. 随机森林模型因子离散化参考的是 Gini 不纯度值, 最终训练结果取决于数据特征和随机采样形成的样本, 离散化级别数和阈值无法人为确定. 这种离散化方式的差异有可能造成了两个模型在人为因子对土壤重金属空间分布影响力评价结果上的区别.

表 9 随机森林模型中土壤重金属驱动因子影响力

Table 9 Influence of soil heavy metal driving factors in random forest model

城市	重金属	土壤 pH	年气温	坡度	海拔	年降水量	夜间灯光指数	县 GDP
安顺市	As	0.21	0.24	0.09	0.13	0.15	0.09	0.10
	Cd	0.05	0.09	0.17	0.52	0.07	0.01	0.08
	Cr	0.16	0.10	0.04	0.43	0.15	0.03	0.08
	Hg	0.18	0.28	0.07	0.23	0.13	0.03	0.08
	Pb	0.10	0.35	0.03	0.35	0.13	0.02	0.01
葫芦岛市	As	0.19	0.15	0.11	0.27	0.18	0.05	0.05
	Cd	0.48	0.08	0.08	0.14	0.13	0.05	0.04
	Cr	0.18	0.13	0.10	0.25	0.15	0.08	0.10
	Hg	0.17	0.25	0.10	0.14	0.25	0.05	0.03
	Pb	0.17	0.10	0.25	0.09	0.13	0.17	0.09

2.4 研究不足及后续研究方向

受数据可获得性的影响, 该研究中人为因子仅选取了县级 GDP、第二产业 GDP 和夜间灯光指数. 若能够在进行驱动力分析时加入对人类活动刻画更为直接的因子, 则人为因子影响力分析将更为精确.

土壤污染物经迁移转化呈现的空间分布格局受到多个因子的共同影响, 由于数据可获得性等原因, 笔者仅选取了 7 项因子进行驱动力研究, 并非表明土壤污染物空间分布仅受或主要受该 7 项因子的影响. 该

研究所关注的重点在于寻找在不同自然及人类活动背景下均能够对土壤污染物空间分布模式具有较强影响的因子,从而提高土壤污染物多元非线性回归方程的泛化能力.如能获取土壤基质、理化性质、有机碳等其他因子,后续研究均可依照笔者提出的研究思路进一步拓展高影响驱动因子或受自然、人类活动干扰轻微的驱动因子清单.

虽然在大尺度下一个地区的土壤 pH 是有已完成的调查成果,可以查阅获取.但严格来说,一个地区的土壤 pH 仍然只能通过采样后测定获取,该研究所用土壤 pH 数据即来源于样品测定结果.但考虑到该研究显示土壤 pH 对于重金属空间分布影响能力较强和测定土壤 pH 的成本低廉,在中小尺度下采用土壤 pH 作为协同因子参与土壤重金属分布插值计算是提高插值精度较为经济、有效的手段.

市域尺度下探讨土壤重金属空间分布驱动力相关研究较少,可供参考的相关研究结论不足.李雨等^[39]在使用地理探测器对湖南省湘潭县土壤重金属影响因子分析中认为,GPD、平均气温和平均相对湿度的影响力最大.这与笔者在人为因子的分析结论一致,与自然因子分析结论相似.

影响土壤重金属空间分布的驱动因子较多且作用机理复杂.在与该研究中典型区相似的区域,以该研究所揭示的高影响力“通用”因子为基础,构建土壤重金属含量的多元回归方程,生成辅助数据,则能够更好地提高土壤重金属含量制图精度和降低调查成本.

3 结论

a) 与葫芦岛市相比,安顺市内 5 种重金属质量分数平均值均较高,其中, $w(\text{As})$ 、 $w(\text{Cd})$ 、 $w(\text{Cr})$ 、 $w(\text{Pb})$ 显著较高, $w(\text{Hg})$ 未通过显著性检验.

b) 在采样点间距为 10 km 的尺度下,安顺市和葫芦岛市土壤重金属分布受到自然因子的影响强于人为因子.其中,安顺市土壤重金属分布受海拔影响较强,葫芦岛市土壤重金属分布受土壤 pH 影响较强.

c) 土壤 pH、海拔和县 GPD 值是对土壤重金属分布影响能力最大的 3 个因子,且其影响力在两个典型区和 5 种重金属中均较高.三项因子在进行两两分析时,对研究区土壤重金属含量空间分布的解释能力具有非线性增强效果.因此,该三项因子适宜作为研究区土壤重金属含量计算模型的通用参数,因子结合使用能够进一步提高计算模型的拟合能力.

参考文献(References):

- [1] ROBINSON T P, METTERNICHT G. Testing the performance of spatial interpolation techniques for mapping soil properties[J]. *Computers and Electronics in Agriculture*, 2006, 50(2): 97-108.
- [2] 陈志凡,赵焯,郭廷忠,等. 污灌条件下重金属在耕作土壤中的积累与形态分布特征:以北京市通州区凤港减河污灌区农用地为例[J]. *地理科学*, 2013(8): 1014-1021.
CHEN Zhifan, ZHAO Ye, GUO Tingzhong, *et al* Impacts of sewage irrigation on heavy metal distribution and chemical fractions in arable soils: a case study about sewage-irrigated farmlands of the fenggangjian river in tongzhou district of Beijing, China[J]. *Scientia Geographica Sinica*, 2013(8): 1014-1021.
- [3] 李湘凌,张颖慧,杨善谋,等. 合肥义城地区土壤重金属污染评价中典型插值方法的对比[J]. *吉林大学学报(地球科学版)*, 2011(01): 222-227.
LI Xiangling, ZHANG Yinghui, YANG Shanmou, *et al*. Comparison of typical interpolation methods for pollution evaluation of soil heavy metals in Yicheng district, HeFei[J]. *Journal of Jilin University(Earth Science Edition)*, 2011(01): 222-227.
- [4] 袁峰,白晓宇,周涛发,等. 元素空间分布插值方法的对比研究:以铜陵地区土壤中的重金属元素为例[J]. *地学前缘*, 2008(5): 103-109.

- YUAN Feng,BAI Xiaoyu,ZHOU Taofa ,*et al.* Comparison between methods for interpolation of studying spatial distribution of elements:a case study of soil heavy metals in Tongling area, South China[J]. *Earth Science Frontiers*, 2008(5): 103-109.
- [5] 谢云峰,陈同斌,雷梅,等. 空间插值模型对土壤 Cd 污染评价结果的影响[J]. *环境科学学报*, 2010(4): 847-854.
XIE Yunfeng,CHEN Tongbin,LEI Mei ,*et al.* Impact of spatial interpolation methods on the estimation of regional soil Cd[J]. *Acta Scientiae Circumstantiae*, 2010(04): 847-854.
- [6] HENGL T, HEUVELINK G B M, STEIN A. A generic framework for spatial prediction of soil variables based on regression-kriging[J]. *Geoderma*, 2004, 120(1): 75-93.
- [7] 庞凤,李廷轩,王永东,等. 县域农田土壤铜含量的协同克里格插值及采样数量优化[J]. *中国农业科学*, 2009(08): 2828-2836.
PANG Su,LI Tingxuan,WANG Yongdong ,*et al.* Spatial interpolation and sampling numbers of the concentration of copper in cropland soil on county scale using cokriging[J]. *Scientia Agricultura Sinica*, 2009(08): 2828-2836.
- [8] 姜勇,李琪,张晓珂,等. 利用辅助变量对污染土壤锌分布的克里格估值[J]. *应用生态学报*, 2006(01): 97-101.
JIANG Yong,LI Qi,ZHANG Xiaokei ,*et al.* Kriging prediction of soil zinc in contaminated field by using an auxillary variable[J]. *Chinese Journal of Applied Ecology*, 2006(01): 97-101.
- [9] HENGL T, HEUVELINK G B M, ROSSITER D G. About regression-kriging: From equations to case studies[J]. *Computers & Geosciences*, 2007, 33(10): 1301-1315.
- [10] MANTA D S, ANGELONE M, BELLANCA A, *et al.* Heavy metals in urban soils: a case study from the city of Palermo (Sicily), Italy[J]. *Science of the Total Environment*, 2002, 300(1-3): 229-243.
- [11] FACCHINELLI A, SACCHI E, MALLEN L. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils[J]. *Environmental Pollution*, 2001, 114(3): 313-324.
- [12] SUN C, LIU J, WANG Y, *et al.* Multivariate and geostatistical analyses of the spatial distribution and sources of heavy metals in agricultural soil in Dehui, Northeast China[J]. *Chemosphere*, 2013, 92(5): 517-523.
- [13] 国家环境保护局,中国环境监测总站. 中国土壤元素背景值[M]. 北京: 中国环境科学出版社, 1990.
- [14] WONG S C, LI X D, ZHANG G, *et al.* Heavy metals in agricultural soils of the Pearl River Delta, South China[J]. *Environmental Pollution*, 2002, 119(1): 33-44.
- [15] 朱求安,江洪,宋晓东. 基于空间插值方法的中国南方酸雨时空分布格局模拟及分析[J]. *环境科学研究*, 2009, 22(11): 1237-1244.
QIUAN Zhu,HONG Jiang,XIAODONG Song. Simulation and Analysis of Spatial-temporal Patterns of Acid Rain in Southern China Based on Spatial Interpolation[J]. *Research of Environmental Sciences*, 2009, 22(11): 1237-1244.
- [16] LIÉNARD A, BROSTAUX Y, COLINET G. Soil contamination near a former Zn–Pb ore-treatment plant: Evaluation of deterministic factors and spatial structures at the landscape scale[J]. *Journal of Geochemical Exploration*, 2014, 147: 107-116.
- [17] 葫芦岛市地方志工作办公室. 葫芦岛市志(综合卷)[M]. 辽宁省葫芦岛市: 海天出版社, 2009: 538.
- [18] 李亮亮,依艳丽,王延松,等. 葫芦岛市连山区、龙港区土壤重金属空间分布及污染评价[J]. *土壤通报*, 2006, 37(3): 495-499.
LI Liangliang,YI Yanli,WANG Yansong ,*et al.* Spatial distribution of soil heavy metals and population evaluation in Lianshan country and Longgang country ,Huludao city[J]. *Chinese Journal of Soil Science*, 2006, 37(3): 495-499.
- [19] 李亮亮,王延松,张大庚,等. 葫芦岛市土壤铅空间分布及污染评价[J]. *土壤*, 2006(04): 465-469.
LI Liangliang,WANG Yansong,ZHANG Dageng ,*et al.* Spatial distribution of soil Pb and its pollution evaluation of Huludao city[J]. *Soils*, 2006(04): 465-469.
- [20] LU C A, ZHANG J F, JIANG H M, *et al.* Assessment of soil contamination with Cd, Pb and Zn and source identification in the area around the Huludao Zinc Plant[J]. *Journal of Hazardous Materials*, 2010, 182(1-3): 743-748.
- [21] 郑娜,王起超,刘景双,等. 葫芦岛市土壤、蔬菜重金属污染空间变化规律[J]. *环境科学*, 2009(07): 2071-2076.
ZHENG Na,WANG Qichao,LIU Jingshuang ,*et al.* Spatial variation of heavy metals contamination in the soil and vegetables of Huludao city[J]. *Environmental Science*, 2009(07): 2071-2076.
- [22] 张玉涛,程劲松,李琳,等. 安顺市煤矿区菜园土壤重金属含量及其赋存特征研究[J]. *蔬菜*, 2014(07): 5-9.
- [23] 董钧铭,王立秋,徐林,等. 安顺市 27 种当地种植蔬菜中铅含量分析[J]. *环境与职业医学*, 2015(10): 960-962.
DONG Junming,WANG Lingqiu,XU Lin ,*et al.* Lead concentrations of 27 local grown vegetables in Anshun city[J]. *Journal of*

- Environmental and Occupational Medicine, 2015(10): 960-962.
- [24] 王志瑞. 不同土地利用方式下农村土壤重金属污染研究[J]. 科技展望, 2014(21): 266-267.
- [25] ELVIDGE C D, IMHOFF M L, BAUGH K E, *et al.* Night-time lights of the world: 1994–1995[J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2001, 56(2): 81-99.
- [26] DOLL C N H, MULLER J, MORLEY J G. Mapping regional economic activity from night-time light satellite imagery[J]. *Ecological Economics*, 2006, 57(1): 75-92.
- [27] WANG J F, HU Y. Environmental health risk detection with GeogDetector[J]. *Environmental Modelling & Software*, 2005, 20(10): 114-115.
- [28] WANG J F, LI X H, CHRISTAKOS G, *et al.* Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China[J]. *International Journal of Geographical Information Science*, 2010, 24(1): 107-127.
- [29] 王劲峰,徐成东. 地理探测器:原理与展望[J]. 地理学报, 2017, 72(1): 116-134.
WANG Jinfeng,XU Chengdong. Geodetector: Principle and prospective[J]. *Acta Geographica Sinica*, 2017, 72(1): 116-134.
- [30] 董玉祥,徐茜,杨忍,等. 基于地理探测器的中国陆地热带北界探讨[J]. 地理学报, 2017, 72(01): 135-147.
DONG Yuxiang,XU Qian,YANG Ren, *et al.*. Delineation of the northern border of the tropical zone of China's mainland using Geodetector[J]. *Scientia Geographica Sinica*, 2017, 72(01): 135-147.
- [31] SHRESTHA A, LUO W. Analysis of groundwater nitrate contamination in the central valley: comparison of the geodetector method, principal component analysis and geographically weighted regression[J]. *ISPRS International Journal of Geo-Information*, 2017, 6(10): 297.
- [32] XU Q, DONG Y, WANG Y, *et al.* Determinants and identification of the northern boundary of China's tropical zone[J]. *Journal of Geographical Sciences*, 2018, 28(1): 31-45.
- [33] ZHOU C, CHEN J, WANG S. Examining the effects of socioeconomic development on fine particulate matter (PM_{2.5}) in China's cities using spatial regression and the geographical detector technique.[J]. *Science of the Total Environment*, 2018, 619–620: 436-445.
- [34] XU C D, XIAO G X. Spatiotemporal risk mapping of hand, foot and mouth disease and its association with meteorological variables in children under 5 years[J]. *Epidemiology & Infection*, 2017, 145(14): 1-9.
- [35] BREIMAN L. Random Forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [36] OLIVEIRA S, OEHLER F, SAN-MIGUEL-AYANZ J, *et al.* Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest[J]. *Forest Ecology and Management*, 2012, 275: 117-129.
- [37] CUTLER D R, JR E T, BEARD K H, *et al.* Random forests for classification in ecology.[J]. *Ecology*, 2007, 88(11): 2783-2792.
- [38] CAO F, YONG G, WANG J. Optimal discretization for geographical detectors-based risk assessment[J]. *Mapping Sciences & Remote Sensing*, 2013, 50(1): 78-92.
- [39] 李雨,韩平,任东,等. 基于地理探测器的农田土壤重金属影响因子分析[J]. 中国农业科学, 2017(21): 4138-4148.
LI Yu,HAN Ping,REN Dong, *et al.*. Influence factor analysis of farmland soil heavy metal based on the geographical detector[J]. *Scientia Agricultura Sinica*, 2017(21): 4138-4148.