应用 LUR 模型探究空间维度上
不同城市景观对成都市空气质量的影响

唐习开，胡文哲，刘子奕，冯佰欢

摘要：城市空气质量近年来作为影响居民健康的重要因素之一，已近成为重多学课研究的热点。现行的空气质量监测办法主要为设立监测站对监测站附近的空气质量各项指标进行测量，但是这样无法获得整体高分辨率的空气质量情况。本实验以成都市为研究区，引入 LUR 模型探究不同城市景观和空气质量之间的关系，发现与 PM2.5 相关性较强的地理自变量有 1km 缓冲区内的建筑面积、1km 缓冲区内的道路长度、3km 缓冲区内的绿地面积、3km 缓冲区内的水域面积。利用 MATLAB 进行线性拟合之后进行改良得到回归模型，在需要精确插值的范围内设立了 25 个测量点，用得到的回归模型模拟各个测量点的 PM2.5 浓度，最后使用空空间插值得到空间整体分布情况。分析 PM2.5 分布情况存在空间梯度差异，整体呈城中心高周围低且西北到东南的带状分布。

关键词：成都市；LUR 模型；GIS；空气质量；PM2.5；空气污染

引言

十八大以来，生态文明建设被重视，空气质量作为环境污染的一个评估因素，十分重要。据《2017 年中国生态环境状况公报》，338 个城市中超出标准空气质量水平的城市高达 70.7%，以 PM2.5 为主要污染物的天数占重度及以上污染天数的 74.2%。迄今为止，空气污染已经成为中国城市举足轻重的一大问题，这不仅影响了人们的生活质量，同时也降低了居民生活的幸福感。

成都市是四川省省会，政治、经济等中心，同时成都市的空气污染也很严重，年平均空气质量指数排名靠后。对于空气质量的检测，传统的站点检测只能局限反映站点周围的空气质量情况，并不能从宏观的角度进行分析评估。本项目旨在利用 GIS 中的 LUR 模型，跟踪统计成都市 7 个空气检测站点数据，运用逐步回归分析的数学方法，筛选出不同月份对空气质量影响最显著的地理要素，最后建立多元线性回归方程预估成都市整体的污染情况图，从而反映污染物和影响污染因素之间的关系。在此基础上，我们将提出改善城市景观建设空气质量的建议，为生态文明城市建设提供理论参考。

细颗粒物 PM2.5 标准是美国于 1997 年提出的，主要是为了更有效地监测随着工业化的不断发展而出现的细颗粒物。2013 年 2 月，中国科学技术术语委员会将 PM2.5 命名为细颗粒物。细颗粒物的化学成分主要包括有机碳(OC)、元素碳(EC)、硝酸盐、硫酸盐、铵、钠

（Na^+）等。以 PM2.5 为代表的空气污染物不仅对人体产生伤害，它还影响了降水改变了，因此世界各国都在寻求改善空气污染的有效方法。

要想控制好空气污染，首要任务是做好空气质量检测，分析清清楚空气污染物的来源和影响因素。（LUR 模型全称为 Land Use regression, LUR）土壤污染回归模型，1997, Briggs 等在城市空气污染的 GIS 研究中，LUR 模型首次应用于大气污染状况空间化。目前，LUR 模型已经成功地应用于欧洲和北美等多地的空气污染空间分析上，并取得不错的模拟效果，国内第一次使用此方法分析空气污染，是陈莉、白志鹏、苏笛等人在 2009 年发布的论文《利用 LUR 模型模拟天津市大气污染物浓度的空间分布》一文中，此篇文章详细的叙述了利用此模型对天津市污染物的模拟实验方法。

近些年来 LUR 模型也在不断地应用中被改进，李爽、翟亮、桑会勇等学者在 2018 年发布的论文《基于改进ＬＵＲ模型的大区域ＰＭ２．５浓度空间分布模拟》中，针对传统的土地使用回归（LUR）模型于大气污染物浓度模拟时预测变量信息丢失的缺陷, 这是一种将主成分分析 PCA（Principle component analysis）与逐步多元线性回归（stepwise multiple linear regression, SMLR）相结合的一种改进 LUR（PCA + SMLR）模型模拟大面积的 PM2.5 浓度的空间分布的方法。

但是通过查阅相关文献发现，往年还没有基于 LUR 模型对成都市空气污染情况拟合的研究，所以本小组开展这次实验所获得的结果可以说是首次给成都市的空气污染治理提供空间尺度上的参考。

1 研究方法

1.1 研究区概况

成都位于四川盆地的西部，东经 102° 54′ ~104° 53′，北纬 30° 05~31° 26′。属于属亚热带季风性气候，具有早春、夏热、秋凉、冬暖的气候特征，年平均气温 16℃，年降雨量约 1000 毫米，全年多云雾日照时间短，主导风向为东北偏北风。

1.2 原始数据获取

1.2.1 空气污染数据的获取

空气污染常见类型可以分为 PM2.5、PM10、SO2、NO2、03、CO（单位μg/m3），本实验具体采集成都市金泉两河、大石西路、君平街、美国驻成都总领事馆、三瓦窑、沙河铺、十里店七个空气检测点的 PM2.5、PM10、NO2 三种污染程度数据，空气污染数据来源为

AQI 中国（可以查询到每一天的空气质量检测记录）和中国环境监测总局（通过第三方应用 AirVisual 可查询到每个小时的污染情况）。

1.2.2 城市地理相关变量的选取

城市常见的地理相关变量可总结为：交通、土地利用、人口密度、自然地理要素以及气象要素等，获取途径为中国科学院计算机网络信息中心地理空间数据云、全国地理信息资源目录服务系统网站、中国科学院资源环境科学数据中心。从这些网站下载到高精度的地理栅格矢量数据。至于降水量和风向等气候数据的获取，可以从国家气象信息中心注册查询获得。

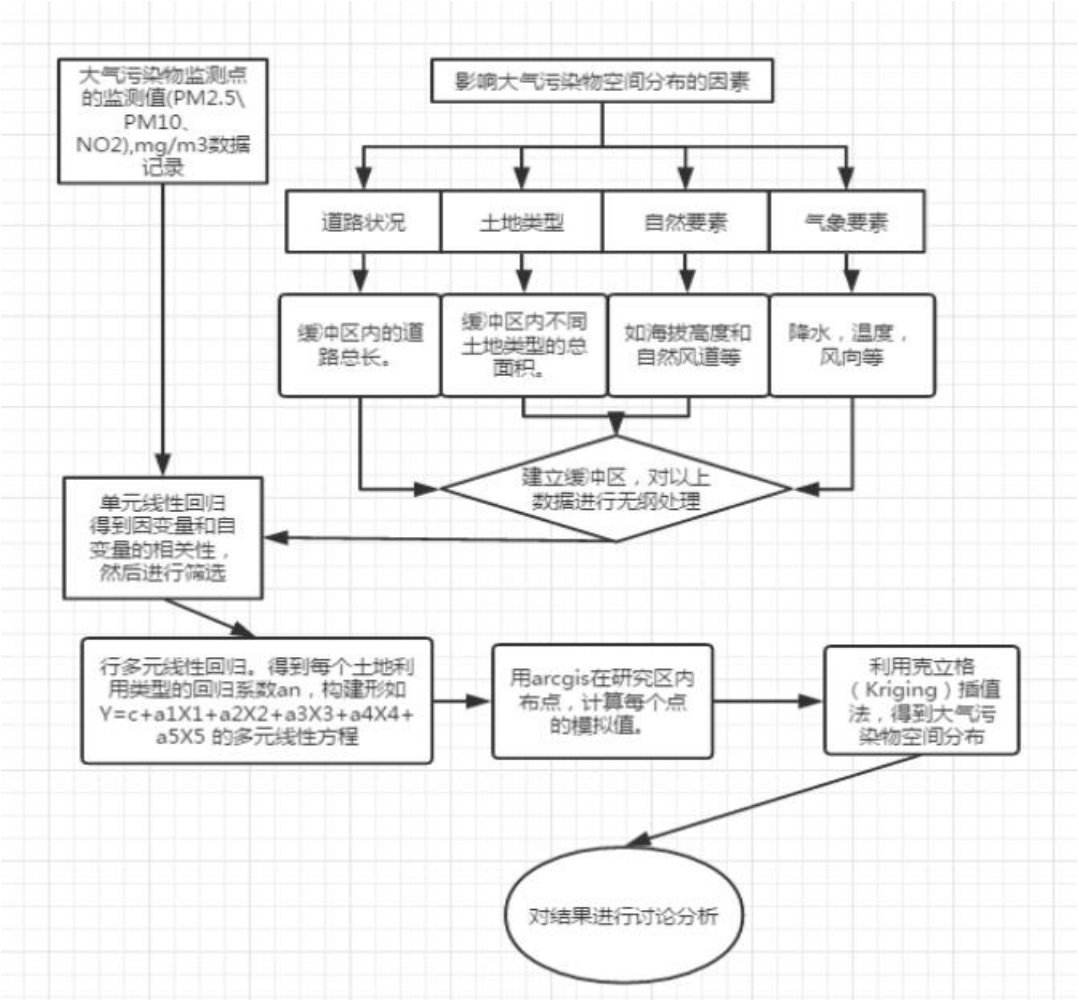


图 1 成都空气污染精细化插值研究流程

2 数据统计的处理

2.1 地理要素的统一

由于相关的地理要素单位不统一，需要通过转化变为无纲数据。故以七个检测站为中心，建立半径为 1KM、2KM、3KM 的缓冲区。通过 google earth 对获取到的成都市交通图进行

几何纠正，然后用 arcgis 中的空间叠加功能获得每个缓冲区内的道路总长度，作为交通状况的自变量。同样，土地利用类型的处理方法相似，也是先对获得地图几何纠正，然后利用 arcgis 的空间叠加功能计算出每个缓冲区内的不同利用类型面积，作为土地利用类型自变量。风向的处理较为特殊，根据成都市风向玫瑰图，得知成都市风频最大的风向是 NNE、22.5°，将风向利用公式 WindIndex=[1-cos(θ -202.5)]/2 转化为 0-1 的无纲量。

2.2 LUR 模型的构建

LUR 模型全称（land use regression, LUR）土壤污染回归模型，通过计算污染物影响因素（如道路长度，建筑密度，植被面积等等）和污染指数之间的相关系数，选取相关系数高的自变量因素构建多元回归方程，然后通过对地图插值模拟每个区域的污染情况。

首先对各污染物浓度、污染物影响因素的指标进行均一化处理，之后利用 SPSS 软件的单元线性回归功能，分别计算 1KM、2KM、3KM 缓冲区内不同土地利用类型等自变量与因变量 PM2.5 月均浓度、因变量 PM10 月均浓度、NO2 月均浓度的相关性。

再挑选出相关性最强的 4 个因素，绘制各因素与污染物平均浓度的散点图，判断其与浓度之间的函数关系类型，得到每个土地利用类型的回归系数 a_n , b_n，构建污染物浓度方程，然后以 Arcgis 为平台，以成都市中心城区为全域进行网格嵌套（5×5），取网格交点为模拟点，对成都市中心城区进行空间均匀布点，进而利用多元回归方程计算每个模拟点的污染物年均浓度值，并进行 Kriging 插值，重分类，最终得到 PM2.5、PM10 和 N02 年均浓度空间模拟分布图。

3 结果与分析

由于对于 PM2.5，PM10 以及 NO2 的统计建模过程类似，因此我们以 PM2.5 的建立回归模型过程为例展示模型。

3.1 相关性分析

将 PM2.5 的平均浓度与各地理相关变量进行双变量相关分析，结果如下表，选择各类变量中与浓度相关性最高的变量与浓度进行回归分析，在此模型中则应该选择 x_1, x_4, x_8 以及 x_{11}，但是在后文进行拟合回归时，由于水域数据以及绿地数据的缺省，出现了完美拟合的情况，因此我们选择 x_1, x_4, x_9 以及 x_{12} 作为建模的自变量。

表 1　地理相关变量与 PM2.5 浓度相关性表

地理相关变量	R	R^2	最优解
1km 道路（x_1）	0.177	0.031329	是

2km 道路（x_2）	-0.109	0.011881	
3km 道路（x_3）	-0.156	0.024336	
1km 建筑（x_4）	0.129	0.016641	是
2km 建筑（x_5）	0.087	0.007569	
3km 建筑（x_6）	0.085	0.007225	
1km 绿地（x_7）	-0.217	0.047089	
2km 绿地（x_8）	-0.402	0.161604	是
3km 绿地（x_9）	-0.347	0.120409	
1km 水域（x_{10}）	0.129	0.016641	
2km 水域（x_{11}）	-0.698	0.487204	是
3km 水域（x_{12}）	-0.269	0.072361	

表 2　地理相关变量与 PM10 浓度相关性表

地理相关变量	R	R^2	最优解
1km 道路（x_1）	-0.328	0.107584	1
2km 道路（x_2）	0.255	0.065025	
3km 道路（x_3）	0.305	0.093025	
1km 建筑（x_4）	-0.635	0.403225	1
2km 建筑（x_5）	-0.199	0.039601	
3km 建筑（x_6）	-0.116	0.013456	
1km 绿地（x_7）	0.29	0.0841	
2km 绿地（x_8）	0.931	0.866761	1
3km 绿地（x_9）	0.9	0.81	
1km 水域（x_{10}）	-0.635	0.403225	
2km 水域（x_{11}）	0.911	0.829921	1
3km 水域（x_{12}）	0.809	0.654481	

表 3　地理相关变量与 NO2 浓度相关性表

地理相关变量	R	R^2	最优解
1km 道路（x_1）	-0.384	0.147456	1
2km 道路（x_2）	0.04	0.0016	
3km 道路（x_3）	0.046	0.002116	
1km 建筑（x_4）	-0.755	0.570025	1
2km 建筑（x_5）	-0.378	0.142884	
3km 建筑（x_6）	-0.281	0.078961	
1km 绿地（x_7）	0.763	0.582169	
2km 绿地（x_8）	0.976	0.952576	1
3km 绿地（x_9）	0.843	0.710649	
1km 水域（x_{10}）	-0.755	0.570025	
2km 水域（x_{11}）	0.856	0.732736	1
3km 水域（x_{12}）	0.527	0.277729	

3.2 散点图

以 7 个统计站 3 月，4 月，5 月的平均 PM2.5 浓度为因变量，3.1 中关联性最高的因素为自变量，绘制散点图，观察各自变量与因变量之间有无明显的函数关系，散点图如下：

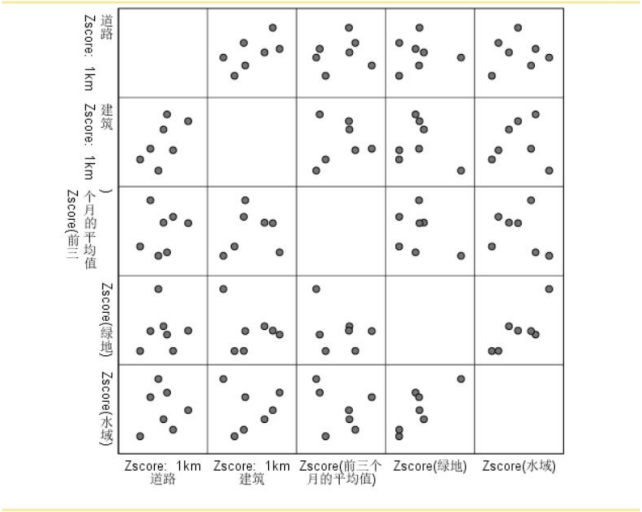


图 1 相关性明显因子与 PM2.5 浓度间散点图

3.3 回归模型

首先，我们利用 MATLAB 进行线性拟合，得到的 R 值为 0.172，调整后的 R^2 为-1.484，故拟合效果很差，得到的方程为

$$Y = 40.849 + 0.502 * x_1 - 3.25 * x_4 - 8.48 * x_9 + 0.202 * x_{12}$$

之后我们将线性拟合得到的解作为初始值，分别依次尝试四个自变量单独作为幂函数的情况，发现四种情况均能最终拟合成 R 方为 1 的情况。考虑到道路对于 PM2.5 的影响可能更大，同时该情况下非线性拟合的结果与线性拟合的结果更加接近，故选择这种情况得到的拟合方程作为最终的拟合方程，利用 MATLAB 进行线性迭代，最终函数在第 11 次之后收敛，各参数情况如表 4，所得到的方程为：

$$Y = 40.746 + 0.407 * x_1^{1.559} - 0.2 * x_4 + 0.62 * x_9 - 1.587 * x_{12}$$

而得到的线性回归参数如表 5，根据公式 R2= 1 -（残差平方和）/（修正平方和）= 1.000，由此可以确定 PM2.5 的浓度与这些地理因子之间成线性相关关系。

表 4 线性回归迭代历史记录

		参数					
迭代编号	残差平方和	a1	a2	a3	a4	a5	a6
1.0	6.766	40.000	.500	-.325	-.848	.202	1.000

1.1	20.179	40.795	.583	-.325	.825	-1.519	1.000
1.2	2.858	40.215	.645	-.266	-.745	-.274	1.619
2.0	2.858	40.215	.645	-.266	-.745	-.274	1.619
2.1	.238	40.544	.528	-.200	.123	-1.269	1.559
3.0	.238	40.544	.528	-.200	.123	-1.269	1.559
3.1	.000	40.746	.407	-.200	.620	-1.587	1.559
4.0	.000	40.746	.407	-.200	.620	-1.587	1.559
4.1	.000	40.746	.407	-.200	.620	-1.587	1.559
5.0	.000	40.746	.407	-.200	.620	-1.587	1.559
5.1	.000	40.746	.407	-.200	.620	-1.587	1.559

表 5　线性回归参数分析

源	平方和	自由度	均方
回归	6740.879	6	1123.480
残差	.000	1	.000
修正前总计	6740.879	7	
修正后总计	4.945	6	

同样的，NO2 的 R^2 为 1.00，回归方程为：

$$Y = 72.470 + 2 * x_1 - 8.19 * x_4 + 4.594 * x_{12}$$

PM10 的回归方程为：

$$Y = 72.654 - 4.349 * x_1 + 4.686 * x_9 + 5.276 * x_{12}$$

3.4 PM2.5 浓度的精细化插值

3.4.1 基准点的建立

在建立了回归模型之后，在成都市区内以绿地分布范围进行 5 行乘 5 列网格嵌套，网格交叉点为精细化插值的基准点。以交叉点为中心建立缓冲区统计缓冲区内的地理要素数量。

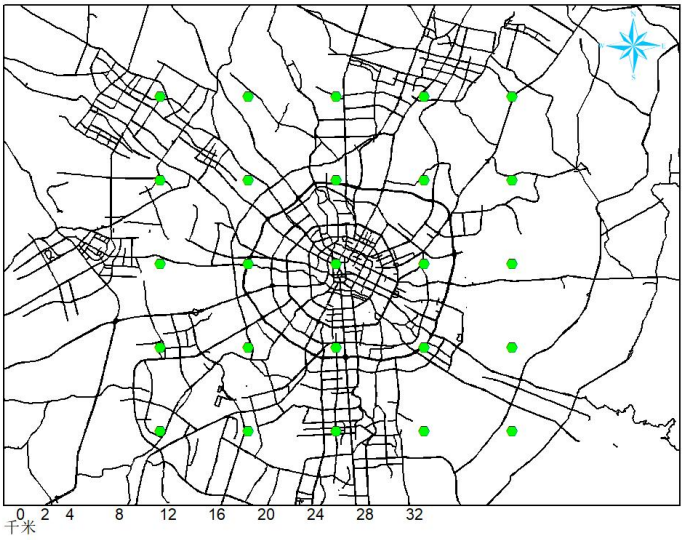


图 3　5×5 交叉点分布图

3.4.2 克里金插值法

克里格法(Kriging)又称空间局部插值，是一种基于变函数理论和结构分析的有限区域定义化变量的无偏最优估计方法。它是地质统计学的主要内容之一。这种方法最早是由南非采矿工程师 D. R.　Krrige(1951 年)在寻找黄金时使用的。

克里格法适用于区域化变量的空间相关性，即当方差图和结构分析结果表明区域化变量具有空间相关性时，可采用克里格法进行插值或外推。否则，它是不可行的。其实质是利用区域化变量的原始数据和变量函数的结构特征对未知样本点进行线性无偏和最优估计。

3.4.3 插值结果

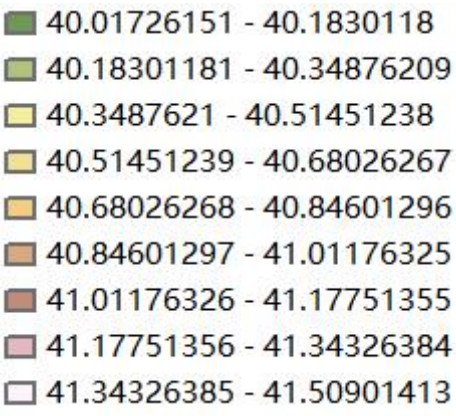


图 4 克里金插值结果分布区间



图 5 克里金插值结果

4 结论

建立 LUR 模型对成都城区的空气污染水平进行了高精度的模拟，通过相关性分析筛选出对空气质量影响较强的地理自变量，其次数学建模筛选出效果较好的回归模型，最后网格点模拟进行空间插值得到了关于 PM2.5 的浓度分布图。

研究结果表明，PM2.5 相关性较强的地理自变量有 1km 缓冲区内的建筑面积、1km 缓冲区内的道路长度、3km 缓冲区内的绿地面积、3km 缓冲区内的水域面积，精细化插值的结果显示，空气污染较为严重的区域多集中在城区，并且有西北东南向的条带状分布。

这种研究方法具有一定的创新性，它打破了传统的监测点无法得到高分辨率的空气污染浓度分布的限制。因此只要能获得一定范围内的地理因子信息，就可以对这一区域的污染物的含量进行预测。 所以我们认为这一模型可以被很好的利用于空气质量的监测、城市设计以及对于工程项目的环保审批。它可以与现有的环境监测站形成很好的互补关系，利用监测站的不断更新的数据进行定期的自我修正，以便适应不同季节，不同历史时期对于城镇的全域环境监测。

但是，我们的项目仍存在一些问题可待改善：因为成都市环境监测站数量有限而导致的样本容量少，可以将环境因子条件进一步的分类，如将道路信息细化为高速公路，主干道，

旁支道路等进行更精细的预测。而这些也正是我们的模型未来完善的目标，借此实现更精准的环境预测。

参考文献：

[1] Briggs D,Collins S,Elliot P,et al. Mapping urban air pollution using GIS: a regression-based approach [J]. Geogr. Inf. Sci., 1997, 11:699-718.

[2] 陈莉,白志鹏,苏笛,游燕,李华敏,刘全.利用 LUR 模型模拟天津市大气污染物浓度的空间分布[J].中国环境科学,2009,29(07):685-691.

[3] 焦利民,许刚,赵素丽, 等.基于 LUR 的武汉市 PM2.5 浓度空间分布模拟[J].武汉大学学报（信息科学版）,2015,40(8):1088-1094.DOI:10.13203/j.whugis20130785.

[4] Kerolyn K. Shairsingh,Cheol-Heon Jeong,Jonathan M. Wang,Jeffrey R. Brook,Greg J. Evans. Urban land use regression models: can temporal deconvolution of traffic pollution measurements extend the urban LUR to suburban areas?[J]. Atmospheric Environment,2018.

[5] rain M A, Blair R, Finkelstein N, et al. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies [J]. Atmospheric Environment, 2007,41:3453-3464.

[6] 宋万营,杨振,王平平,丁启燕,李星明.基于 LUR 模型的大气 PM_(2.5)浓度分布模拟与人口暴露研究——以湖北省为例[J].华中师范大学学报(自然科学出版),2019,53(03):451-458.

[7] 杜诗薇. 基于 LUR 模型的四川省 PM_(2.5)时空分布特征研究[D].成都理工大学,2018.

[8] 杨婉莹,刘艳芳,刘耀林,安子豪,银超慧.基于LUR模型探究城市景观格局对PM2.5浓度的影响——以长株潭城市群为例[J].长江流域资源与环境,2019,28(09):2251-2261.

[9] 鲁凤杰. 基于 LUR 模型的城市交通颗粒物污染空间分布研究[D].长安大学,2019.

[10] 常恬君. 改进 LUR 模拟 NO_2/SO_2 浓度空间分布及暴露风险评价[D].华东师范大学,2019.

[11] 李爽,翟亮,桑会勇,邹滨,方新,甄云鹏.基于改进 LUR 模型的大区域 PM_(2.5)浓度空间分布模拟[J].武汉大学学报(信息科学版),2018,43(10):1574-1579+1587.

[12] 贺佳,贺亮,张涛,霍培书,刘拓,张丹宁.基于GIS和LUR模型的西安市PM_(2.5)浓度空间分布模拟研究[J].环境科学与管理,2017,42(02):57-60.

[13] 江笑薇,任志远,孙艺杰.基于 LUR 和 GIS 的西安市 PM_(2.5)的空间分布模拟及影响因素[J].陕西师范大学学报(自然科学版),2017,45(03):80-87+106.

[14] 汉瑞英,陈健,王彬.利用 LUR 模型模拟杭州市 PM_(2.5)质量浓度空间分布[J].环境科学学报,2016,36(09):3379-3385.

[15] 刘武. 基于 LUR 模型的长沙市城区 NO_2 和 PM_(10)浓度的时空分布模拟[D].湖南大学,2016.

[16] L.F. Weissert,J.A. Salmond,G. Miskell,M. Alavi-Shoshtari,D.E. Williams. Development of a microscale land use regression model for predicting NO 2 concentrations at a heavy trafficked suburban area in Auckland, NZ[J]. Science of the Total Environment,2018,619-620.

[17] Kerolyn K. Shairsingh,Cheol-Heon Jeong,Jonathan M. Wang,Jeffrey R. Brook,Greg J. Evans. Urban land use regression models: can temporal deconvolution of traffic pollution measurements extend the urban LUR to suburban areas?[J]. Atmospheric Environment,2018.

[18] 陈林. 基于 GIS 的流域水文数据的时空分析——以格兰德河流域径流数据为例[D]. 山东科技大学, 2010.

[19] 满旺, 陈明发, 聂芹,等. 上海城区 PM2.5 精细化空间分布插值方法研究[J]. 宁夏大学学报(自然科学版), 2019, 40(01):90-95.

[20] 邵步粉. 基于 ArcGIS 的上海地区低温冰冻雨雪灾害的风险区划和极值、重现期研究[D]. 2010.
[21] 张晓娟. PM2.5 的生成来源及应对措施[J]. 考试周刊, 2014, 000(005):194-194.