

密级:\_\_\_\_\_



**中国科学院大学**  
University of Chinese Academy of Sciences

## 硕士学位论文

呼伦贝尔市森林脑炎地理环境影响因子探测及风险评价

作者姓名: \_\_\_\_\_ 李一凡 \_\_\_\_\_

指导教师: \_\_\_\_\_ 王卷乐 研究员 \_\_\_\_\_

\_\_\_\_\_ 中国科学院地理科学与资源研究所 \_\_\_\_\_

学位类别: \_\_\_\_\_ 理学硕士 \_\_\_\_\_

学科专业: \_\_\_\_\_ 地图学与地理信息系统 \_\_\_\_\_

研 究 所: \_\_\_\_\_ 中国科学院地理科学与资源研究所 \_\_\_\_\_

二〇一六年 五 月

**Assessment of Tick-borne Encephalitis Risks and Relative  
Geographical Environment Factors in Hulunbuir**

**By  
Yifan Li**

**A Dissertation/Thesis Submitted to  
The University of Chinese Academy of Sciences  
In partial fulfillment of the requirement  
For the degree of  
Master of Science**

**Institute of Geographic Sciences and Natural Resources Research  
Chinese Academy of Sciences**

**May, 2016**

# 致 谢

岁月匆匆，时光荏苒，转眼间，三年的硕士生活已接近尾声，又到了要道别的时刻。如果一定要用两个字来形容这三年，那么此刻涌上心头的便是“感恩”二字了。时光让人成长，也让人铭记。感恩时光，让我遇到了一个无比细心、勤奋的导师，几个相谈甚欢的舍友，还有一群“乐天派”的师弟们。值此论文脱稿之际，谨向所有在读研期间帮助关系过我的老师和同学致以最诚挚的感谢。

首先，我要感恩我的导师王卷乐老师。三年的学习生涯中，王老师总是用鼓励的方式来激励我一步一步走向前，一点一点去改变。因为知道我有演讲紧张的问题，老师总是鼓励我多参加学术交流活动，多与人结交。三年来，从张家界到武汉，从台北到昆明，一点一点，我渐渐改变了演讲紧张的症结。感谢老师为我提供了一个如此广阔的学习平台，为我树立了一个无比勤奋的学习榜样，为我开启了一个乐观向上的生活视角。

其次，我要感恩课题组的各位老师和师兄妹们。感谢杨飞老师和吕宁老师，感谢你们在我开题、中期以及整个学习过程中所给予的学术指导和帮助；感谢高孟绪师兄，您一丝不苟的学习和工作态度，是我未来工作态度的标榜；感谢柏中强、王末、祝俊祥以及其他已毕业的师兄师姐，感谢你们在我学习中的一次次悉心指导以及每一次出差里生活上的照应；感谢我的舍友杨芬、姜瀟月同学，因为有你们，宿舍才有温馨；感谢我的一群“乐天派”师弟师妹们，感谢永青、小强和吕鑫，感谢你们无私地帮助我分担了许多工作，让我有时间找工作、写论文，感谢刘清、奋民和苏萍，感谢你们用积极健康的运动观念，给课题组，也给我带来了健康的体魄。

最后，我要感谢我的家人。感谢爸爸妈妈和亲爱的姐姐，因为有你们在我人生道路中不断地支持和鼓励，我才能走到今天。感谢你们为我撑起了一个家，让我不管走到哪里，都感觉得到温暖，感受到爱与被爱。

李一凡

2016年5月 于北京

# 摘要

森林脑炎是一种以蜱虫为主要传播媒介的自然疫源性疾病，其分布和传播与地理环境关系密切，探索其与地理环境因子的关系对于掌握现有和潜在的森林脑炎自然疫源地空间分布，预测森林脑炎的风险区域具有重要意义。本研究以我国森林脑炎发病率最高的内蒙古呼伦贝尔市为研究区，通过统计和空间自相关方法分析了森林脑炎的流行时空特征，利用地理探测器和多元逐步线性回归等方法探测地形、气象、植被等地理环境因子与森林脑炎发病的相关关系，基于探测获取的地理因子，通过回归建模方法对其风险的空间分布进行评价。主要结论如下：

（1）呼伦贝尔市森林脑炎的发病情况具有明显的春夏季高发特性，患病人群中男性显著多于女性，且具有明显的职业特性；该市 2006~2013 年森林脑炎发病的空间分布具有显著的空间自相关特性。

（2）地理环境因子与森林脑炎发病情况的影响程度依照因子解释力大小被分为三个等级，影响程度最大的是土壤类型、降雨量和土地覆被类型。进一步探测发现土壤和土地覆被在空间上具有很大的相关性，与植被、地形、气候等要素相关程度较高，可以作为回归模型中其他要素选择的标准和借鉴。

（3）呼伦贝尔市森林脑炎发病风险的空间分布主要表现为中部出现高发病风险的三角带，东西部是低风险地带。中部三角带主要分布于额尔古纳市、根河市、鄂伦春自治旗西部以及牙克石市，三角内部发病风险表现为自南向北发病风险逐渐升高的空间格局。三角带内，相对湿度是与该区域高发病率呈现最大正相关性的要素，大部分区域的发病风险与 NDVI、降雨量和坡向呈正相关关系，与坡度呈负相关关系，北部部分区域与 DEM 呈负相关关系。

**关键词：**森林脑炎，地理环境因子，风险分布，空间分析

# Abstract

Tick-borne encephalitis (TBE) is a natural foci disease, transmitted by ticks. Its distribution and transimission is closely related with geographic and environmental factors. Getting the relationship between environmental factors and TBE risks will be of great importance to grasp China's existing and potential natural foci of TBE. Also, it will help to detect the risk areas of TBE. This paper chose Hulunbuir City as study area for it had the largest number of reported TBE cases. Firstly, statistical methods and spatial autocorrelation correlation analysis were used to detect the spatial and temporal distribution characteristics of TBE. Then, in the grid data base, Geo-Detector was used to help detecting the relationship between topographic, climatic, vegetation factors and TBE incidence. Finally, this paper has used geographical weighted regression (GWR) to regress the distribution of TBE risks with some geographic and environmental factors. The main conclusions are as follows:

1) TBE cases in Hulunbuir were mostly being infected in spring and summer time. Among these cases, male cases' number was significantly more than female cases. Most of them were doing some work related with forest operation. Besides, from 2006 to 2013, the distribution of TBE cases in Hulunbuir has significant spatial autocorrelation features.

2) Factors were divided into three classes in accordance with its explanatory power on the distribution of TBE risks. Soil type, precipitation and land cover were found to have the greatest effect on TBE incidence. Further probe found that soil type and land cover were closely related in high risk area. The formation of these two factors are influenced greatly by local vegetation, topography, climate and other factors. Thence, soil type and land cover could be a reference when selecting the other elements into the regression model.

3) In Hulunbuir, spatial distribution of TBE risks was featured with a high risk triangular belt in central part, which included Erguna, Genhe, Yakeshi county and Oroqen Autonomous Banner. The eastern and western part were both in low risk. In the triangular area, risks gradually increased from south to north. Humidity, NDVI, precipitation and slope were all positively correlated with high TBE risks, among wich humidity has the maximum correlation. Slope was negatively correlated with TBE risks in the whole triangular area. While, DEM was negatively correlated with TBE risks only in the northern part.

**Keywords:** Tick-borne encephalitis, geographic and environmental factors, spatial distribution of TBE risks, spatial analysis.

# 目 录

致 谢 .....	III
摘 要 .....	IV
Abstract .....	V
目 录 .....	VI
第一章 绪论 .....	1
1.1 选题背景 .....	1
1.2 TBE 流行病学概况 .....	1
1.2.1 TBE 流行形势 .....	1
1.2.2 TBE 流行病学过程 .....	3
1.2.3 TBE 季节和人群特性 .....	4
1.3 国内外研究综述 .....	5
1.3.1 自然疫源性疾病相关因子 .....	5
1.3.2 自然疫源性疾病风险探测 .....	7
1.4 研究目标与技术路线 .....	12
1.5 论文结构 .....	13
第二章 研究区、数据和方法 .....	15
2.1 研究区 .....	15
2.2 数据 .....	16
2.2.1 数据源 .....	16
2.2.2 数据预处理 .....	17
2.3 研究方法 .....	18
2.3.1 空间平滑 .....	18
2.3.2 空间自相关 .....	19
2.3.3 格网方法 .....	20
2.3.4 地理探测器 .....	21

2.3.5	多元线性回归 .....	23
2.3.6	地理加权回归 .....	25
第三章	森林脑炎流行时空特征 .....	27
3.1	流行病学特征 .....	27
3.1.1	时间特征 .....	27
3.1.2	人群分布 .....	27
3.2	空间特征 .....	29
3.2.1	全局空间自相关 .....	29
3.2.2	局域空间自相关 .....	30
3.3	本章小结 .....	32
第四章	森林脑炎地理环境影响因子探测 .....	33
4.1	地理因子空间分布特征 .....	33
4.2	地理因子探测 .....	39
4.2.1	数据预处理 .....	39
4.2.2	因子探测 .....	41
4.2.3	探测结果及不确定性分析 .....	44
4.3	本章小结 .....	45
第五章	森林脑炎风险评价 .....	47
5.1	多元逐步线性回归 .....	47
5.1.1	数据预处理 .....	47
5.1.2	模型建立 .....	48
5.2	地理加权回归 .....	50
5.2.1	数据预处理 .....	50
5.2.2	模型建立 .....	51
5.3	结果分析 .....	51
5.3.1	回归系数空间分布 .....	51
5.3.2	风险空间分布评价 .....	54

5.4	本章小结.....	57
第六章	结论与展望 .....	59
6.1	结论.....	59
6.2	展望.....	60
	参考文献.....	61
	作者简介.....	69



# 第一章 绪论

## 1.1 选题背景

自然疫源性疾病(Natural Focus Disease)是一种可不依赖人类,在自然界即可流行和传播的传染病。存在自然疫源性疾病病毒的地域称自然疫源地,病原体、传播媒介和宿主动物是构成自然疫源地的 3 个基本要素,人一旦进入自然疫源地,就可能会通过不同方式被感染而发病(张启恩,2003)。自然疫源性疾病的分布和传播与地理环境关系密切,探索其与地理环境因子的关系对于掌握中国现有和潜在的自然疫源地分布,预测中国自然疫源性疾病的流行趋势具有重要意义。

森林脑炎(Tick-borne Encephalitis, TBE)是一种由黄病毒属蜱传脑炎病毒所引起的中枢神经系统急性传染病,该疾病主要分布于蜱虫数量较多的地区,受地理环境的影响极大,是一类典型的自然疫源性疾病,该病在中国东北地区的病例数量较多,对当地人民的生产、生活造成较大影响。TBE 的分布与其所在的地理环境关系密切,但缺少二者相互影响的定量探测研究,因而难以根据不同地域地理环境的差异、有针对性地对我国典型地区的 TBE 发病的风险空间分布进行科学预测和判断。

中国正处于区域开发、城市化以及众多重大建设工程快速发展的阶段,“一带一路”战略、国防安全、生态文明、美丽中国等建设都需要掌握自然疫源性疾病的区域自然疫源特性及风险空间分布。本研究以中国 TBE 发病情况最为严重的内蒙古呼伦贝尔市为对象,探测 TBE 相关地理环境因子并对其风险分布进行评价。

本研究依托于国家科技基础性工作专项项目:中国自然疫源性疾病流行病学图集编研(2013FY114600)。

## 1.2 TBE 流行病学概况

### 1.2.1 TBE 流行形势

世界范围内,TBE 的主要疫源地包括西欧、俄罗斯西伯利亚地区以及远东三个地区。其中,欧洲型的宿主主要是蓖子硬蜱(*Ixodes ricinus*)(Barbour et al, 1983),西西伯利亚和远东型的宿主主要为全沟硬蜱(*Ixodes persulcatus*)(万康林等, 1998)。TBE 发病病例数最多的国家是俄罗斯,1990~2009 年间,俄罗斯平均每年新增 TBE 病例数约为 5950,占到欧洲总体的一半以上(Süss, 2008b)。2000~2010 年间,欧洲整体(除俄罗斯)的报告病例数保持在相对稳定的态势,只是在 2003、2006 以

及 2009 这三年略有增加, 其中捷克、立陶宛、拉脱维亚、德国和斯洛维尼亚的报告病例数分别占到欧洲整体的 25%、15%、11%、11%和 10%, TBE 病例的整体分布范围与该地区 TBE 的主要媒介昆虫——蓖麻蜱虫的分布大致一致, 均在北纬 40°~60°的地中海到斯堪的纳维亚半岛之间(Schwarz et al, 2009), 随着气候变化, 篦子硬蜱的分布范围已从 61° N 以南, 扩张到了 66° N 附近的河流和大的湖泊沿岸(Gray et al, 2009)。

国内方面, 根据刘育京等(1979)和尹德铭等(2000)对中国 TBE 流行病学历程的回顾, TBE 最早由日本人宫部勋和北野分别于 1943 年和 1944 年在中国东北林区发现, 当时该病被当地人称为“抬头难”, 解放后, 森林砍伐由过去的冬季作业改为常年作业, 各林区患者突增, 后经政府大力组织防治, 1953 年后病例数开始显著下降, 但由于疫源地的普遍性和不确定性, 每年仍有新增患者。蔡增林等(1996)对东北三省部分地区的 TBE 流行病学调查结果表明, 1992~1994 年该地区仍存在 TBE 自然疫源地。结合中国各地区自然环境、林型组成、媒介和宿主类型以及 TBE 的流行情况, 毕维铭等(1997)及 Lu 等(2008)将中国 TBE 疫源地划分为东北、西北和西南 3 个主要疫源区, 其中疫情最为严重的东北疫源区又可划分为长白山地 and 大小兴安岭 3 个疫源地。董景红等(2007)对内蒙古大兴安岭林区 TBE 的研究表明, 1991~2002 年间, 大兴安岭林区的 TBE 病例数不断增多, 且疫区逐年扩大, 并发现当时疫情持续多发的根本原因与人群普遍缺乏免疫力和疫苗接种覆盖面不足有关。

毕维铭等(1997)根据森林生态系原则、县域基础原则、森林分布与自然疫源地相统一的原则, 结合中国林区 TBE 疫源地的自然环境、林型组成、宿主种类以及疾病流行情况, 将中国 TBE 疫源地划分为三大疫源区和六大疫源地。东北疫源区包括长白山地、小兴安岭、大兴安岭疫源地; 西北疫源区包括天山、阿尔泰山疫源地; 西南疫源区包括位于川滇藏三省横断山地林区在内的西南可疑疫源地, 如图 1 所示。

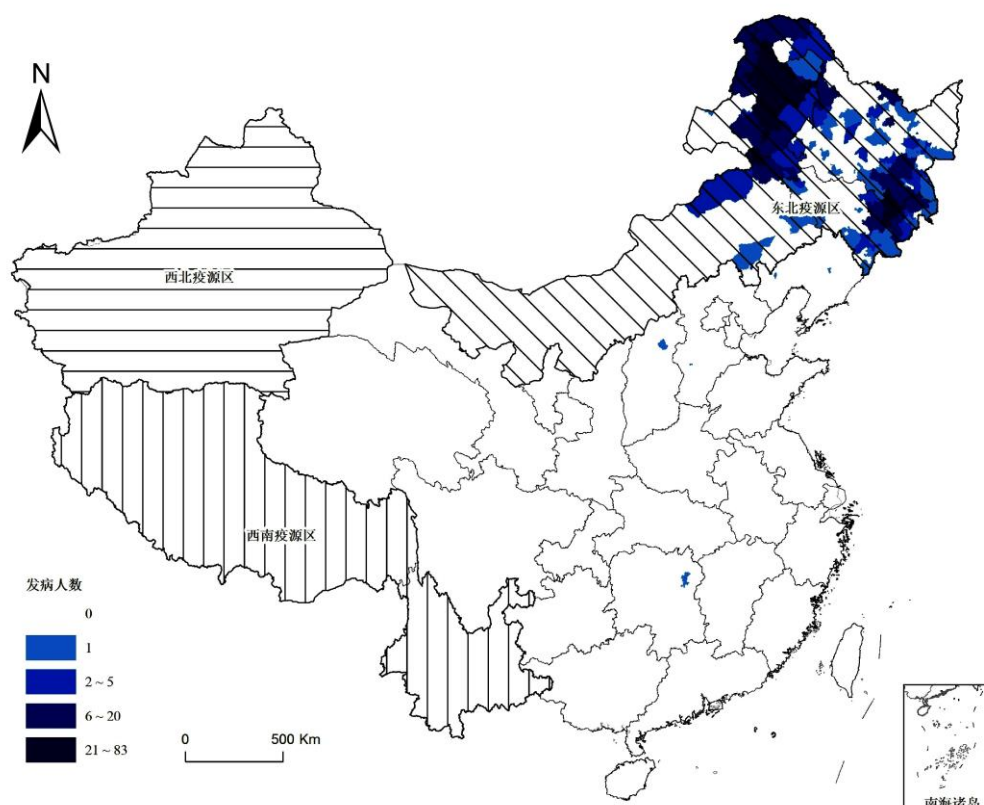


图 1 全国 TBE 三大疫源区分布

注：按 2011 年全国各县市 TBE 发病人数分布图编绘，病例统计数据来源于军事医学科学院。

### 1.2.2 TBE 流行病学过程

TBE 的发生需要具备宿主动物、传播媒介和传播途径三个重要环节。

(1)宿主动物：除蜱虫以外，TBE 的宿主动物还有啮齿动物(Knap et al, 2012, Achazi et al, 2011)、食虫类、食肉类、偶蹄类、鸟类(Geller et al, 2013)及一些家畜动物。成蜱的宿主主要是大型野生动物，例如，狗、鹿、黑熊、牛(林场饲养)等，幼、稚蜱的宿主主要是小型哺乳动物和鸟类，例如棕背鼠平、红背鼠平、花鼠、林姬鼠、黑线姬鼠、东方田鼠、中国蹶鼠、松鼠等(Kim et al, 2008, Walder et al, 2006, Takeda, 1999, Takashima et al, 1997)。

(2)传播媒介：蜱虫是 TBE 传播最主要的媒介和宿主。中国林区已发现的蜱类已有 10 多种，已从全沟硬蜱、嗜群血蜱、日本血蜱、森林革蜱和边缘草蜱中分离出病毒，其中全沟硬蜱是中国 TBE 最主要的传播媒介，广泛分布于中国北方林区和其他林区，以针阔混交林中数量最多(毕维铭, 1997)。TBE 病毒能够在蜱体内保持 3 年之久，而在脊椎动物的血液循环中最多仅能停留 10 余天，虽然两者对病毒在自然疫源地中的循环都很重要，但蜱既是传播媒介，也是长期贮存宿主，因此以蜱虫的情况作为该疫源地的强度标志更能够反映实际情况。

图 2 为蜱虫生命周期与 TBE 病毒的传递过程示意图。蜱的生活史主要有卵、幼虫、若虫和成虫四个时期(詹希美, 2002)。成虫吸血后交配落地，在第  $t-1$  年的

冬季于草根、树根、畜舍等处表层缝隙中产卵，在第  $t-1$  年的春秋季节，虫卵发育成幼虫和若虫，再经历过一个冬天，在第  $t$  年的春秋季节发育成成虫。各发育期的蜱虫都要吸血，饱血后的蜱虫可胀大几倍至几十倍，雌硬蜱甚至可以达 100 倍(Lindgren, 1998, 张思华, 2012)。

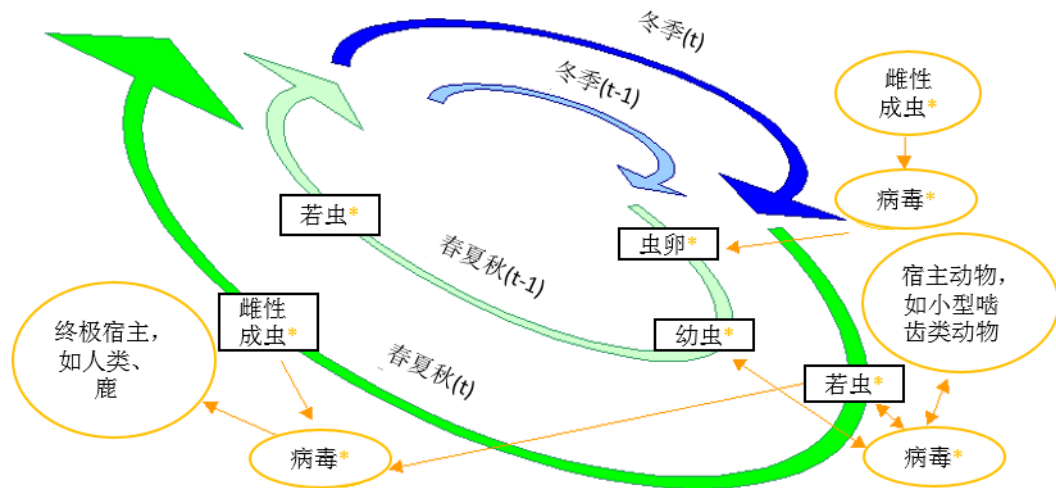


图 2 蜱虫生命周期与 TBE 病毒的传递过程示意图(Lindgren, E., 1998)

(3)传播途径: 自然带毒蜱叮咬吸血时, 病原体从涎液排出, 注入到宿主体内, 导致感染; 蜱在吸血过程中, 排出的粪便或基节液含有病原体, 可以通过蜱咬伤口或皮肤黏膜小伤口侵入机体引起感染; 带毒蜱被碾碎或压破, 其体液中病原体溢出, 经皮肤黏膜伤口感染人体(曹务春等, 1999)。

### 1.2.3 TBE 季节和人群特性

森林工作人员进入森林活动, 一旦被带有 TBE 病毒的蜱虫咬伤, 就会患上 TBE, 因而发病人群具有明显的职业特点, 几乎全为进入林区的人员, 尤其以伐木工人、地质勘探、森林调查、野生植物采摘者以及林区旅游者为多, 如, 121955~1962 年黑龙江省 292 病例中, 林业工人占 67.81% (毕维铭, 1997)。

TBE 发病有明显的季节特征, 如中国东北林区 TBE 的主要媒介——全沟硬蜱的成虫活动期在 4~8 月, 高峰在 5~6 月初, 6~7 月日均温升至  $20^{\circ}\text{C}$  以上时, 蜱虫的数量又逐渐减少, 幼虫和若虫的活动季节较长, 一般在 3 月底至 4 月初日温达到  $0^{\circ}\text{C}$  开始活动, 一直能够持续至 9~10 月间, 一般有两个高峰, 主峰常在 6~7 月, 次峰约在 8~9 月间, 9 月中下旬均温降到  $10^{\circ}\text{C}$  以下, 蜱虫的活动逐渐停止(张思华, 2012)。人群的发病时间通常较蜱虫的活动时间推后 2 周, 一般在 5~8 月份病例数最多(图 3)。

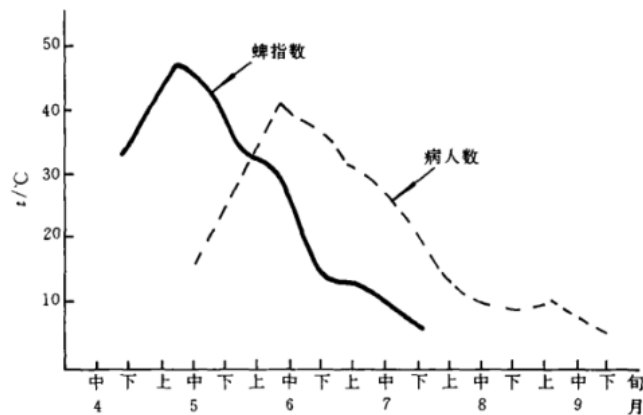


图3 蜱虫与发病人数时间关系示意图(毕维铭, 1997)

### 1.3 国内外研究综述

#### 1.3.1 自然疫源性疾病相关因子

综合 1.2.2 节中 TBE 传播的三个基本环节以及现有对蜱虫和蜱传性疾病的研究, 影响 TBE 发生的因素可以分为自然和人文两个方面, 这两类因素会对人类与蜱虫的接触机率、宿主媒介以及各自的生活环境产生交叉影响(图 2)。除 TBE 外, 常见的蜱传性疾病有莱姆病、Q 热、蜱传回归热、新疆出血热等。自然因素包括地貌、气候、土壤和植被等(Rogers et al, 2003, Hay et al, 1997, Rogers et al, 1993, Meade et al, 1988), 人文因素包括社会政策、经济活动、军事防备、人口数量变化、城市扩张、景观格局特点及其变化等(Randolph, 2008, Šumilo et al, 2008, Hoogstraal, 1979)。

##### 1) 地貌因子

地貌主要通过海拔高度和坡向影响水热的再分配, 进而对媒介蜱虫及其宿主动物的分布产生影响。Merler 等(1996)利用决策树法对意大利特伦蒂诺山区的蓖麻蜱虫分布密度进行分类, 发现该地区决定蜱虫分布的最主要因素是海拔高度和地质环境; Toomer(1997)在研究非洲地区彩饰花蜱的分布时, 结合数字高程模型(DEM)和其他变量模拟得到该类蜱虫的大致分布区域; Randolph 等(2000)在研究欧洲 TBE 分布情况时, 利用 DEM 和经过傅里叶变换后的 NOAA-AVHRR 遥感影像作为预测变量, 结果发现欧洲中心地区和波罗的海沿岸的 5 个蜱虫分布聚集点与 DEM 的分布规律具有极大的联系; Materna 等(2008)发现海拔高度通过影响林种的分布情况进而决定蜱虫密度, 并发现中欧卡尔可诺希山地区海拔小于 1000 m 的针阔混交林区最适宜于蜱虫的生长, 坡向对小范围研究区的蜱虫密度影响极高; Jacobsen (2014)对美国蒙大纳州安氏革蜱的分布进行了研究, 发现该地区南坡的地貌条件更适宜于该类蜱虫的生长, 可能与山地南坡具有更好的光照条件、较高的温度、较少的积雪量有关。

##### 2) 气候因子

气温、光照时数、降雨等气候因子在一定程度上决定了宿主和媒介的生活范围，从而影响自然疫源地的分布。

极端最高和最低温度直接影响蜱虫在各个生命周期的存活率，并对其宿主动物的生命活动产生影响。Lindgren(1998)发现冬季温度偏高可能会使得更多的蜱虫及其宿主动物存活下来，早到的春天或者晚到的秋天同样也会延长蜱类的活动时间，增加蜱虫的存活率，由此发现冬季的平均气温、最低气温以及低温天数，春季和秋季的最低气温及持续天数都可能会影响蜱虫的生长发育过程和 TBE 发病情况。Eisen(2007)在研究美国拉瑞莫郡的安氏革蜱分布时，发现蜱密度在日最高温达到 11℃时达到最高密度值的一半，但在日最高温达到 21℃时，却会低于最高密度值的一半，说明记得高低温会影响蜱虫的生长发育过程和 TBE 发病情况。Süss 等(2008)认为温度的升高可能会通过加速蜱虫发育、延长蜱虫发育周期、增加蜱虫产卵量 3 个方面，增加蜱虫密度，从而影响蜱传性疾病发病率，甚至会改变疫源地分布范围。

光照时数能够影响蜱虫的行为滞育(寻找宿主)和发育停滞(转型)(Randolph, 2004, Gray, 1998)，但由于光照时数会与其他各类气象因子对蜱虫在不同阶段的生长发育产生交叉影响，作用效果较为复杂，因而现阶段还没有确定光照时数对蜱虫生长发育影响的定量结果(Perret, 2003)。

降雨通过影响空气相对湿度(RH)而对蜱虫的生长发育及其寻求宿主活跃程度产生影响。Kahl 等(1988)认为相对湿度(RH)会影响蜱虫在无宿主期或发育转型期对半饱和空气中水汽的转化和吸收，蜱虫生长最适宜的 RH 值应该在 70%~80%，过高或过低的 RH 都会对蜱虫的生长发育产生影响。

### 3) 土壤因子

土壤类别(主要为土壤 pH 值、有机质含量以及植被类型)和土壤质地(由土壤粒度决定)决定了土壤水分的涵养能力，土壤湿度过高不利于无宿主蜱虫的生长，但却有利于真菌、线虫等对蜱虫生长产生威胁的生物的滋生。Guerra 等(2002)在研究美国莱姆病的主要媒介——肩突硬蜱的分布时发现，灰土等酸度较高或粘土比过高的土壤具有较大的湿度，不适宜于蜱虫的生长发育，淋溶土和软土等粒度较大的土壤更适宜于蜱虫的生长发育。通常可以利用地面温度指数(LST)来代替土壤温度(Randolph et al, 2000a, Randolph et al, 2000b)，用于弥补由土壤温度监测样本少所带来的缺陷。研究表明，月 LST 的平均值较其最大值具有更多的信息量，与蜱密度的相关程度更高(Randolph et al, 2000a)。

### 4) 植被因子

植被除了能够为宿主和媒介提供适宜的生活环境以外，还可以为部分宿主动物提供丰富的食物。蜱虫密度与森林类型和林型结构密切相关，落叶林和针阔混交林中蜱密度较高，针叶林和无林覆盖区中蜱密度则较小(Randolph, 2000)。

Jackson(2005)研究美国马里兰地区植被景观与当地 1996~2000 年莱姆病发病情况的相关性时,发现森林与草地的边缘对比指数与莱姆病发病情况密切相关,森林密度达到 53%时,当地莱姆病发病率最高。

植被指数作为利用遥感方法对地面光合作用活动的一种测量指标,可用来定性和定量地评价植被覆盖及其生长活力。Kitron 等(1997)利用植被指数来判定 Wisconsin 地区植被覆盖区,发现春秋季节高植被指数值和蜱虫密度与当地居民患莱姆病的发病率密切相关;Estrada-Peña(2001)在研究中指出,植被指数可以作为预测部分地区蜱虫数量的重要因子。

### 5) 土地利用

人文因子可能会通过改变疫源地的疫源性强弱程度、人与疫源地的接触频率等来影响自然疫源性疾病的发病情况。土地利用状况是人地关系影响的直接表现方式。不同土地利用类型的有机组合可以折射出宿主和媒介生境的适宜程度,从而反映不同自然疫源地的疫源性强弱程度。有关蜱传性疾病与土地利用类型之间的直接关系研究尚不多,但从相关自然疫性疾病与土地利用之间的关系研究可以得到借鉴。例如,方立群(2009)在研究中发现土地利用类型对肾综合征出血热(HFRS)传播流行的影响存在多重共线性,对 HFRS 传播和流行产生主要影响的是不同土地利用类型组合所反映的不同生境条件;疟疾、乙脑、登革热等蚊媒自然疫源性疾病的分布与水田、林地、草地、建筑用地等能够滋生蚊虫的土地利用类型的相关性极大(Vanwambeke et al, 2007, 张少森等, 2012)。此外,经济水平以及居民的文化程度等因素能够在一定程度上反映出居民对卫生健康的重视程度、接种疫苗的健康意识以及居住环境的卫生条件等,进而影响自然疫源性疾病的流行情况。以流行性乙型脑炎为例,唐咸艳等(2011)发现经济水平较好的桂南、桂东南地区因乙脑疫苗接种率高、卫生条件较好,使得这些地区乙脑的发病率较桂西北地区相对较低。

## 1.3.2 自然疫源性疾病风险探测

### 1.3.2.1 风险特征探测

自然疫源性疾病的风险特征主要是指疾病发生的相对危险区域和风险聚集情况。空间平滑、空间自相关分析和空间聚类是风险特征探测的常用技术。空间平滑能够有效解决由于人口基数不同而导致发病率方差不稳定的问题,空间聚类能够标识出相对的热点区域,空间自相关分析能够揭示自然疫源性疾病的风险分布特征,这些方法都能为风险因子的探测奠定基础。

#### (1) 空间平滑

空间平滑是用于解决由于人口基数不同而导致发病率方差不稳定的有效空间计算方法,其本质是利用其他空间单元的属性值进行加权计算的空间平滑方法。经验贝叶斯平滑(Empirical Bayes Smoothing, EBS)是基于各空间单元的原始发病

率数据,利用整个研究区某一特定加权指标(例如人口值)进行平均运算的平滑算法。Zhang 等(2015)在研究山东省聊城市 2007~2011 年手足口病的发病情况时,利用经验贝叶斯平滑纠正了由原始发病率所带来的空间分布误差;Beneš 等(2011)利用贝叶斯平滑对捷克波西米亚中部 TBE 发病风险进行了模拟;Zhang 等(2014)利用经验贝叶斯平滑对 2005~2012 年全国各县级市 HFRS 发病人数进行空间平滑,最终得到以百分数显示的 HFRS 发病人数空间分布图。

## (2) 空间聚类

空间聚类是环境健康研究中最主要的分析技术之一,它能够标识出相对的热点区域,进而引导研究者作出地理环境因子与疾病发生之间有意义的假设(申思,2008)。目前,流行病学中常用的基于同质性数据的空间聚类方法有空间扫描统计量法(Kulldorff et al, 1997)和空间分层聚类法。前者主要用来探测空间上离散分布点的聚集区,后者是依据属性值的相似程度将空间区域划分为多级别的子区域(Carvalho et al, 2009)。在空间扫描统计量的应用研究中,方立群(2004)利用扫描统计量法对北京市海淀区 1997~2002 年人群 HRFS 发病情况进行空间聚类分析,得到了 HFRS 高发病热点区域;Li 等(2012)运用改进的空间扫描分析法对山东省平度市 2009 年 5 月手足口病发病情况进行双层方向异性风险聚类分析,得到了更好的聚类效果。在空间分层聚类方法的应用研究中,Curtis 等(2002)利用最近邻生长的层次聚类算法对美国西弗吉尼亚州的浣熊狂犬病进行分级热点探测;Wang 等(2006)利用基于区域最近邻生长的层次聚类算法对北京城区 2003 年 SARS 发病病例点实现分级热点探测,发现大部分一级聚类环散落分布在三环路,二级聚类环则在西北方向聚集。

## (3) 空间自相关

空间自相关是依据事物或现象对空间位置的依赖关系,衡量全局或局域范围内,事物空间自相关性的计算方法(冯军等,2011),可分为全局和局部两种自相关分析方法。前者主要用于描述整个研究区域的空间对象某一属性取值的空间分布状态,以判断该对象在空间上是否存在聚集性;后者能够进一步揭示空间单元与其相邻近的空间单元属性特征值之间的相似性或相关性,可用于识别“热点区域”以及数据的异质性。全局自相关的常用度量指标包括 Moran's  $I$  和 Getis-Ord  $G$  等。Kitron 等(1997)以 Moran's  $I$  为指标对美国威斯康星州莱姆病三个年段的空间分布进行了研究,发现各年段中,莱姆病的发病情况均呈显著的空间正相关;范新生等(2005)运用 Moran's  $I$  对中国 SARS 的空间分布进行了分析,结果表明 SARS 疫情的空间分布在疫情发源阶段以及新暴发中心形成阶段具有很强空间自相关性;Hu 等(2011a)采用 Moran's  $I$  等指标,对澳大利亚昆士兰州历年的登革热病例资料进行了空间自相关性统计,发现该地区登革热的空间分布在某些时段呈现聚集性。局域自相关的度量指标包括局域 Moran's  $I$ (也称 LISA)、局域 Getis-Ord  $G$



等, 曹志冬等(2008)在对广州 SARS 流行的空间风险因子与空间相关性特征的研究中, 计算了广州市 SARS 发病率的 LISA, 由此得到了该研究区域的局部空间聚集特征。

### 1.3.2.2 风险因子探测

相关性分析法主要用于定量分析疾病分布与地理环境因子之间的关系, 该方法可分为经典统计相关性分析和空间相关性分析两类(图 4)。经典统计法按照是否具有时间滞后效应分为简单相关性分析法和互相关分析法, 简单相关性法按照统计量是否服从正态分布分为参数统计和非参数统计; 空间相关性分析的代表为基于空间格数据的相关性分析。

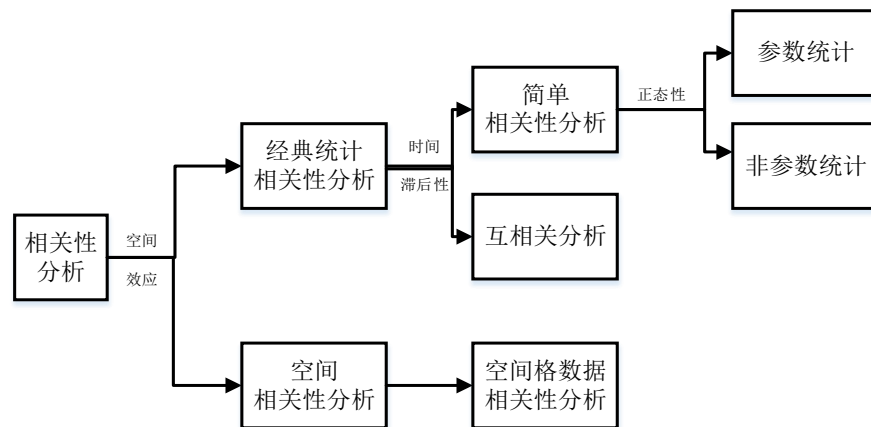


图 4 相关性分析方法

#### (1) Spearman 秩相关分析

由于大部分自然疫源性疾病是区域小概率疾病, 发病率不符合正态分布。研究中, 一般利用 Spearman 秩相关分析计算各类因子与疾病发病率之间的相关性, 并以此作为因子初选的依据。Schwarz 等(2009)利用 Spearman 秩相关分析方法对蜱虫数量与气象因素和土壤情况进行相关性分析; Lukan 等(2010)利用 Spearman 双边检验对 TBE 发病情况与年平均降雨量和温度的相关性进行分析; Ruiz-Fons 等(2012)利用 Spearman 秩相关对各环境因子与处于不同时期蓖麻蜱虫的密度进行相关性分析; Žáková 等(2013)利用 Spearman 秩相关分析对蜱虫在不同温度下的密度与杀虫剂使用剂量作了相关性分析。

#### (2) 互相关分析

气候、植被指数等因素对媒介动物的生长和发育的影响存在时间滞后性, 研究者在研究具有时间滞后效应的变量与自然疫源性疾病发病情况之间相关关系时, 一般采用互相关分析。互相关分析可以用来确定气候、植被等因子与自然疫源性疾病发病情况之间的时间滞后效应。Daniel 等(2009)利用该方法得到捷克共和国 1994~2004 年的周平均气温与 1 到 5 周滞后的 TBE 发病情况具有相关性; 张文义(2010)在研究中国东北大兴安岭林区气候因素对姬鼠型疫区 HFRS 传播的影响时, 利用互相关分析得到具有时间滞后性的相关性分析定量结果。

### (3) 空间格数据分析

基于空间格数据(Lattice data)的空间相关性探测能满足相关性分析对空间属性的依赖。格数据又称面状数据(Areal data),是指以空间格状单元存储的属性数据集。由于遥感图像和社会经济、疾病数据多以规则像元或不规则多边形的格数据形式进行存储,因而格数据空间相关性分析法更具有实际应用性(武继磊, 2004, 王劲峰等, 2010)。目前基于格数据形式的分析软件主要有地理探测器(王劲峰等, 2010)、快速查询软件(Rapid Inquiry Function, RIF)(Aylin et al, 1999)以及格网数据的空间流行病学快速挖掘软件(Fast Grid-Based data mining, FGBASE)(Baker et al, 2014)。Wang J F 等(2010a)利用地理探测器探测得到与山西省和顺县新生儿神经管畸形发病情况有关的地理环境及经济因子,并分析了因子之间的交互作用; Hu Y(2011)利用地理探测器对四川汶川地震后该地区五岁以下儿童死亡分布与该地区地震次数、倒塌房屋面积以及地形因素作了探测和分析; 通拉嘎等(2014)利用地理探测器分析了地理环境因子对 2009 年湖北省钉螺分布的影响。

#### 1.3.2.3 风险预测建模

传染病的风险预测模型按照其模型机理分为系统动力学模型和数理分析模型两类。系统动力学模型(Susceptive-infective-removal, SIR)是按照一般传染病传播机制通过数量关系描述传染病的传播过程、分析感染人数的变化规律、解释传染病的发展趋势,但由于该类模型在参数的设定过程中受主观因素和经验因素的影响较大,且参数的物理意义不明确,因而在应用时具有一定的局限性。

数理分析模型包括时间序列模型、灰色预测、神经网络、回归模型等。

##### (1) 时间序列模型

时间序列模型是指依据变量的长期监测数据及其随机扰动项所建立的模型,用以预测该变量的未来发展趋势。差分整合移动平均自回归模型(Autoregressive Integrated Moving Average, ARIMA)是应用较为广泛的时间序列模型。由于 ARIMA 等时序模型的建立需要有长时间序列的监测数据,对参与建模的时序数据要求较高,且模型中未包括影响疾病发生的因子情况,难以为特定地区特定地理环境条件下 TBE 的靶向控制提供精确的参考依据。

##### (2) 灰色预测

灰色预测是通过对系统因素进行关联度分析,以及对原始数据的生成处理来寻求系统变动的规律。通过灰色模型(Grey Model, GM)对变量的定量分析,灰色预测可以进行灰色时间序列预测、畸变预测、波形预测以及系统预测。这种预测方法的优点是:建模过程简单,模型表达式简单,求解方便,且该模型允许少数数据预测,允许对灰因果律事件进行预测,具有可检验性等。胡明等(1991)曾利用 GM 方法根据济南铁路局 1984~1988 年细菌性痢疾的发病率进行建模,并依此预测了该地区 1989 年该疾病的发病率情况,预测精度良好,具有一定可信度;刘

洁等(2013)利用 GM 模型对中国 1990~2009 年乙脑发病率数据进行拟合, 预测得到 2010~2012 年发病率情况, 与实际值基本一致。但是, GM 模型也存在一些缺陷, 如: 中指数型变化的单调性使得模型在中长期预测时可能会产生较大误差; 对随机性、波动性较大的数据, 拟合结果较差; 且不能考虑统计变量的空间属性。因而, 该方法不适宜于受地理环境因子影响较大、且具有空间异质性特点的自然疫源性疾病的发病情况的预测。

### (3) 回归模型

回归模型可以基于地理环境因子与疾病发病情况之间的定量关系, 回归并模拟疾病的风险分布情况, 为其他区域或者未来该地区的风险预测提供定量参考。依据模型是否考虑“空间效应”, 回归模型可分为经典统计回归模型和空间回归模型两类。经典统计回归模型包括普通最小二乘回归(Ordinary Least Square, OLS)、广义线性模型(Generalized Linear Model, GLM)、时间序列泊松回归模型和逻辑斯蒂(Logistic)回归模型等。空间回归模型中应用最广泛的是基于“空间相关性”和“空间异质性”的地理加权回归(Geographic Weighted Regression, GWR)(王劲峰等, 2010)。本节主要针对时间序列泊松回归、Logistic 回归以及 GWR 三种模型, 进行综述(李一凡等, 2015)。

#### a) 时间序列泊松回归模型

时间序列泊松回归模型在考虑变量的自相关性、季节性、长期趋势和“滞后效应”的同时, 能定量地研究气候变化和疾病传播之间的相关关系(Peng et al, 2008), 对气候、植被指数等具有时间滞后性的变量能够进行良好的回归建模。Peng 等(2008)结合气象因子, 利用时间序列泊松回归模型模拟, 得到澳大利亚亚热带地区考虑气候变量滞后性的弯曲杆菌感染病的发病模型; 张文义(2010)结合气象因子, 利用时间序列泊松回归模型, 模拟得到中国东北大兴安岭林区考虑气候变量滞后性的 HFRS 传播模型。

#### b) Logistic 回归模型

在小概率疾病的研究中, 将预测量和回归量按一定标准分为多级, 用分级尺度代换数字, 更易于揭示预测量和回归量之间的关系(孙振宇, 2008)。Logistic 回归模型可以根据因变量取值类别的不同, 对因变量进行回归。Glass 等(1995)利用 Logistic 回归模型模拟了美国巴尔的摩莱姆病的发病模型; Cumming(1999)基于 DEM 和气温插值数据利用 Logistic 回归模型模拟得到非洲地区精度良好的蜱虫分布结果; 张文义(2010)在研究北京市 HFRS 感染动物宿主与环境变量的关系时, 将利用单因素分析筛选出的具有统计学意义的环境变量纳入多因素 Logistic 回归分析, 并建立模型。

#### c) GWR 模型

综合“空间相关性”和“空间异质性”两种思想对变量进行空间回归模拟的 GWR 模型是一类较经典的空间回归模型, 它能更为精确地反映疾病分布在空间

上的变化(胡艺, 2012)。Wimberly 等(2008)利用 GWR 方法结合气候、土地覆被、宿主(鹿)密度对美国西南和中南部蜃传性疾病病原体分布进行了空间回归; Bayles 等(2014)利用 GWR 模型结合生态、人文因子, 对美国密苏里地区蜃传埃里希体病的发病情况进行了模拟。

## 1.4 研究目标与技术路线

通过对相关地理环境因子以及风险探测研究方法的国内外综述, 发现现阶段国内关于 TBE 风险的研究存在以下几点问题: 1) 地理环境因子会改变人类与蜃虫的接触机率, 并对宿主和媒介的生活环境产生影响, 现阶段的研究缺乏对 TBE 发病风险与地理环境因子的定量认识, 不能够为 TBE 风险的预测及建模提供良好的理论依据; 2) 对发病风险的回归和预测建模大部分是基于统计单元的, 且没有考虑到空间异质性, 所提供的空间分布结果不能够满足日益精细的防疫防控需求。

基于以上分析, 本研究的目标是以我国 TBE 发病率最高的内蒙古呼伦贝尔地区为研究区, 通过统计和空间自相关方法分析了 TBE 的流行时空特征, 利用地理探测器和多元逐步线性回归等方法探测地形、气象、植被等地理环境因子与 TBE 发病的相关关系, 基于探测获取的地理因子, 预测和评价了呼伦贝尔市 TBE 发病风险的空间分布。研究方案如图 5 所示:

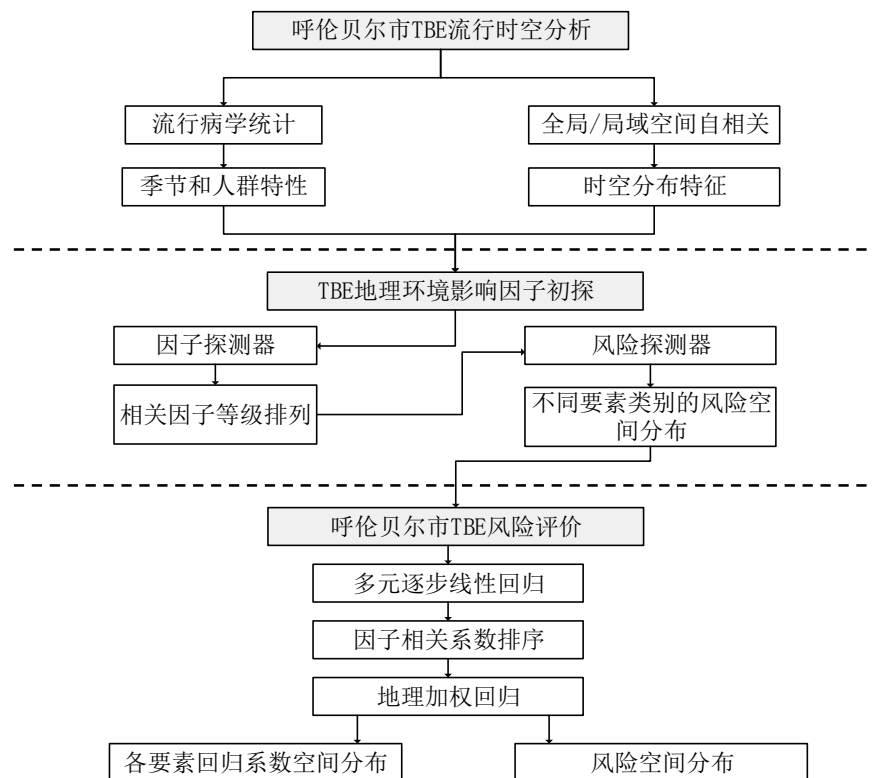


图 5 研究方案

本研究方案主要包括以下三个部分：

1) 呼伦贝尔市 TBE 流行时空特征分析。利用流行病学统计方法和全局/局域空间自相关分析方法分别得到呼伦贝尔市 TBE 发病的季节人群特征和时空分布特征。

2) TBE 地理环境影响因子初探。利用地理探测器中的因子探测器探测出相关地理环境因子对发病风险解释力的大小值，并进行等级划分，再利用风险探测器探测第一等级中部分地理环境因子空间划分情况下 TBE 发病风险的空间分布，并进一步分析这些地理环境因子之间的相关关系。

3) 呼伦贝尔市 TBE 风险评价。首先利用多元线性逐步回归对 TBE 发病风险进行回归，对各地理环境因子的相关系数大小值进行排序，在此基础上考虑各地理环境因子在物理意义和数值上的共线性，利用不同的要素组合进行地理加权回归，从中最终选择拟合效果最优的回归模型，得到各要素回归系数的空间分布以及 TBE 发病风险的空间分布。

## 1.5 论文结构

本论文分六章，如图 6 所示。

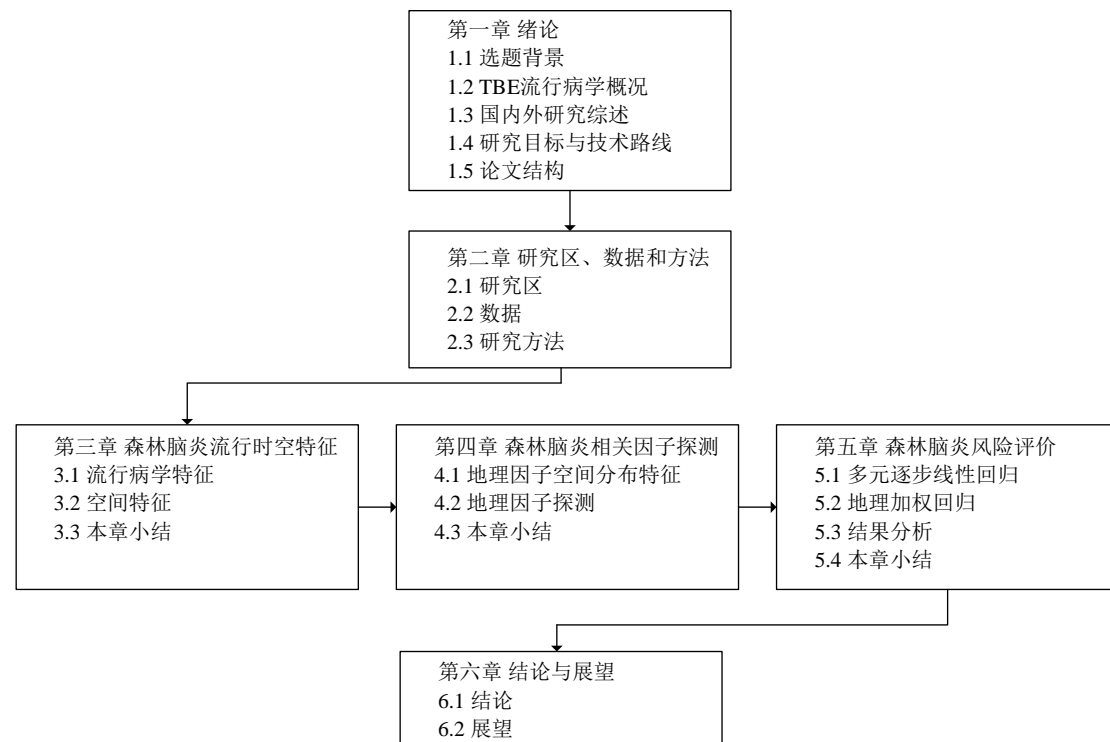


图 6 论文章节结构图

第一章是绪论，介绍了选题的背景、目的和意义，对国内外研究进展进行了归纳、总结，之后介绍了技术路线，最后是论文结构。

第二章是研究区、数据和方法，介绍了研究区、数据、数据预处理以及研究所涉及的方法。

第三章是 TBE 流行时空特征，介绍了通过流行病学统计方法和空间分析方法所探测到的研究区内 TBE 流行时空特征。

第四章是利用地理探测器中两个主要因子初步探测并分析与研究区 TBE 发病具有相关关系的地理环境因子。

第五章是利用回归模型对呼伦贝尔市 TBE 发病风险进行模拟和评估，首先利用多元逐步线性回归对各地理环境因子回归系数进行排序，再进行因子的选择和组合，最后选择最优的地理加权回归模型，分析各因子回归系数的空间分布以及发病风险空间分布的特征。

第六章是结论和展望，对研究中所得到的 TBE 流行时空特征、因子初探结果以及风险评价结果等内容进行了总结，并根据现有研究的不足对后续工作进行了展望。

## 第二章 研究区、数据和方法

### 2.1 研究区

呼伦贝尔市地处东经  $115^{\circ} 31' \sim 125^{\circ} 04'$ ，北纬  $47^{\circ} 05' \sim 53^{\circ} 20'$ ，隶属于内蒙古自治区，北与俄罗斯联邦、西与蒙古国、南与兴安盟、东与黑龙江省相邻，是蒙古高原的组成部分。该市东西 630 km，南北 700 km，总面积 25.3 万  $\text{km}^2$ ，是全国管辖面积最大的城市级行政单位，现辖 2 个市辖区、5 个县级市、4 个旗以及 3 个自治旗(其中包括 159 个乡镇、苏木)。全市总人口约为 269.7 万人，农业和非农业人口分别占人口总数的 37.4% 和 62.6%，其比例为 1: 1.67。呼伦贝尔市是一个多民族聚居城市，少数民族中以蒙古族为主体，全市汉族人口约占 82.31%，蒙古族人口约占 9.02%，其他少数民族人口占 8.66%。

南北走向的大兴安岭位于呼伦贝尔市中部，使全市形成岭东、岭中、岭西三大地理单元和经济类型区域：岭东以农区为主，包括扎兰屯、阿荣旗、莫力达瓦达斡尔族自治县，地势较低，海拔约为 200~500 m，多为低山丘陵与河谷平原，土壤以黑土、暗棕壤、黑钙土和草甸土为主，土质肥沃，自然肥力高，盛产大豆、玉米、水稻；岭中以大兴安岭林区为主，海拔约为 700~1700 m，包括牙克石市、根河市、鄂伦春自治旗，有林地面积 1.90 亿亩，约占全市土地总面积的 50%，森林覆盖率约为 49%，森林活立木总蓄积量 9.5 亿  $\text{m}^3$ ，占全国的 9.5%，林区主要树种为兴安落叶松、白桦、柃树、樟子松、黑桦、山杨、蒙古柞等；岭西以草原牧区为主，海拔约为 550~1000 m，包括陈巴尔虎旗、新巴尔虎右旗、新巴尔虎左旗、鄂温克自治旗，多年生草本植物是组成呼伦贝尔草原植物群落的基本生态特征。

呼伦贝尔市有丰富自然资源。矿产资源方面，全市探查到的各类矿产达 40 余种，矿点 370 多处，已探明的矿点有 57 处，主要有煤炭、石油、铁、铜、铅、锌、钼、金、银、铌、铍、铟、镉、硫铁矿、芒硝、萤石、重晶石、溴、水泥灰岩等。煤炭探明储量是辽宁、吉林、黑龙江三省总和的 1.8 倍。水资源方面，呼伦贝尔市水资源总量为 286.6 亿  $\text{m}^3$ 。其中，地表水资源量 272 亿  $\text{m}^3$ ，占全国地表水资源量的 1%，占全区地表水资源量的 73%；地下水资源总量 14.6 亿  $\text{m}^3$ 。全市人均占有水资源量为 1.1 万  $\text{m}^3$ ，高于世界人均占有量，是全国人均占有量的 4.66 倍。野生动植物资源方面，呼伦贝尔市共有野生植物 1400 多种，有经济价值的野生植物达 500 种以上，主要有野生药用植物、野生经济植物、野生油料植物、野生纤维植物、野生淀粉植物、野生食用植物、野生果品植物等。呼伦贝尔市野生动物品种和数量繁多。据不完全统计，全市野生动物种类占全国种类总数的 12.3%，占自治区的 70% 以上，居第一位。在这些动物中，受国家保护的一、

二、三类野生动物和受自治区保护的野生动物品种有 30 余种，其中有些是珍稀兽类和禽类，以鹿科动物最为有名。全市 313 种鸟类中，受国家保护的鸟类有 60 多种，如丹顶鹤、白头鹤、白鹤、灰鹤、大天鹅、小天鹅等。

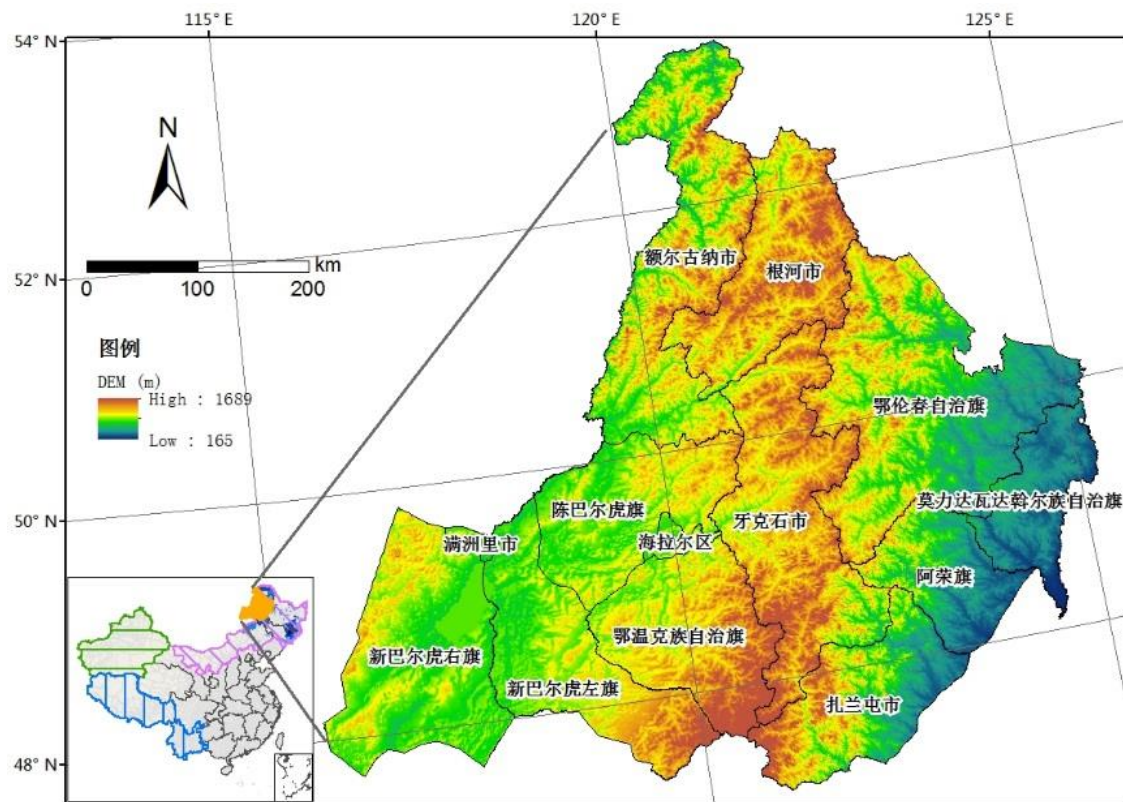


图 7 呼伦贝尔市位置分布(底图为 DEM)

## 2.2 数据

### 2.2.1 数据源

#### 1) 疾病数据

疾病数据包括呼伦贝尔市 2006~2013 年 TBE 发病病例直报数据，病例信息包含病人的出生日期、性别、发病时间、家庭和工作地址等，该数据由军事医学科学院提供。

#### 2) 地理环境要素数据

地理环境要素数据主要包括覆盖呼伦贝尔市整体范围的 DEM、坡度、坡向等地形因子，温度、降雨量、降雨频率、相对湿度、光照时数等气象因子，两种植被指数，土壤类型空间分布和土地覆被空间分布等数据。

地形数据主要包括高程(DEM)、坡度(Slope)、坡向(Aspect)三类数据，原始高程数据为 Shuttle Radar Topography Mission (SRTM) 空间分辨率为 90 m 的 DEM 数据(标准绝对高程精度  $\pm 16$  m，绝对平面精度  $\pm 20$  m)，该数据集下载自 NASA (<http://www.nasa.gov/>)。



气象数据下载自中国气象数据共享网(<http://data.cma.gov.cn/>)地面累年值月值数据集,包括气温、光照时数、相对湿度、降雨量、降雨频度共 5 个指标从 2006 年 1 月到 2013 年 12 月的月值数据,月值数据是基于实时上传的地面自动站逐小时数据文件及日值数据文件而研制。气温是空气的温度,单位是摄氏度( $^{\circ}\text{C}$ ),可以表征一个地方的热状况特征;日照时数是指太阳每天在垂直于其光线的平面上的辐射强度超过或等于  $120\text{ W/m}^2$  的时间长度,所用的光照时数是月光照时数,单位是小时(h);相对湿度是空气中实际所含水蒸汽密度和同温度下饱和水蒸汽密度的百分比值,单位是百分数(%),用来表示空气中的水汽含量多少或空气潮湿程度的物理量;降雨量是指在一定时间内降落到地面的水层深度,单位是毫米(mm);降雨频度是指日降水量大于等于  $0.1\text{ mm}$  的天数,单位是天。

植被数据选择了归一化植被指数(NDVI)和增强植被指数(EVI)这 2 类植被指数,数据来源于空间分辨率为  $1\text{ km}$ 、时间分辨率为逐月的全球植被指数数据集(MOD13A3 数据集),下载自 NASA 的 EOSDIS(<https://earthdata.nasa.gov/>)。

土壤类型和土地覆被数据取自国家地球系统科学数据共享平台(<http://www.geodata.cn/>)全国 1:400 万土壤类型分布图(1980s 土壤普查成果)和 2005 年中国 1:25 万土地覆被数据。呼伦贝尔市范围内共有包括暗棕壤、栗钙土、棕色针叶林土、沼泽土等在内的 11 个土壤一级类,以及 22 个土壤二级类。土地覆被数据包括森林、草地、农田、聚落、湿地水体、荒漠等共计 6 个土地覆被类型的一级类和 17 个二级类,其中,森林包括常绿针叶林、落叶针叶林、落叶阔叶林和灌丛共 4 个二级类,草地包括草甸草地和典型草地共 2 个二级类。

其他数据:人口数据取自于由国务院人口普查办公室于 2012 出版的《中国 2010 年人口普查分乡、镇、街道资料》,1:160 万内蒙古大兴安岭林业管理局林业区划图和森林分布图取自大兴安岭林区森林资源监测中心地图室。

### 2.2.2 数据预处理

疾病数据:依据直报病例的住址信息对病例进行空间化处理。在此过程中,由于缺少病例家庭或工作地址的精确经纬度坐标,无法将病例数据精确匹配到空间点位上,本文以乡镇、街道行政单元对病例住址进行统计。首先,对 2006~2013 年每年各乡镇/街道 TBE 发病病例数进行质量检查。此 8 年间,呼伦贝尔市共有 680 条上报 TBE 病例,其中有 5 名病例的家庭或工作地址均不在呼伦贝尔市,其余 675 名病例可按本市住址进行空间化处理。在计算原始发病率时,以 2010 年全国乡镇级别人口统计数据为基准,求得 2006~2013 年每年及 8 年平均的原始 TBE 发病率。

地形数据:坡度和坡向数据是基于原始 DEM 栅格数据利用 ArcGIS10.2 (<https://www.arcgis.com/>)中的 Slope 和 Aspect 工具计算生成的,空间分辨率均为  $90\text{ m}$ 。Slope 利用最大平均值法来计算每一个像元的坡度值,面的朝向就是待处

理像元的坡向,坡度值越小,地势越平坦,坡度值越大,地势越陡峭。Aspect 由  $0^{\circ}$  到  $359.9^{\circ}$  之间的正度数表示,以北为基准方向按顺时针进行测量。

气象数据:为了减少由插值带来的边界误差,选择呼伦贝尔市及其周边的气象站监测的月值数据作为插值基准,在 ArcGIS 内对研究区 2006~2013 年各指标逐月进行普通克里格空间插值,得到各指标逐月的空间分布栅格图层,最后通过栅格计算,得到年均温(Temperature, Temp)、年均降雨量(Precipitation, PPT)、年均降雨频率(Precipitation frequency, PF)、年均相对湿度(Relative humidity, RH)和光照时数(Sun hours, SH)共 5 个指标的空间分布。

植被数据:通过对研究区范围内 MOD13A3 数据集中 2006~2013 年 8 年月值 NDVI 和 EVI 数据进行遥感拼接、栅格均值计算等,最终获得研究区范围内 8 年 NDVI 和 EVI 的平均值。

最后一步为数据的格网化处理,主要应用于第四章“森林脑炎地理环境影响因子探测”和第五章“森林脑炎风险评价”的数据预处理过程。首先,利用 ArcGIS 的渔网(fishnet)工具生成空间尺度为  $10\text{ km} \times 10\text{ km}$  覆盖整个呼伦贝尔市的格网体系,共 2340 个网格;然后,分别针对不同要素的栅格图层进行要素数据的格网化提取,对土地覆被类型和土壤类型等类别型变量,利用 ArcGIS 区域统计中的众数(Majority)规则进行要素统计,对发病率、植被指数、气象要素以及地形要素等数值型变量采用平均值(Mean)规则进行区域统计,特别地,由于第四章中数值型变量在数据预处理过程中先进行分级分类处理,因而所有变量均利用众数规则进行统计,详见 4.2.1 节。

## 2.3 研究方法

### 2.3.1 空间平滑

对于小概率疾病,其发生率的可靠性因不同空间位置上的人口基数不同而有较大差异,需要调整到大体一致且较为稳定的水平,其后的各种统计结果才更具可靠性和可比性,这就是所谓的“小数问题”(Shi et al, 2013)。

经验贝叶斯平滑(Empirical Bayes Smoothing, EBS)是基于各空间单元的原始发病率数据,对原始发病率的估计值向整体发病率进行调整的算法。这种平滑算法之所以被称为经验贝叶斯平滑,因为它借用了“先验分布”来避免因分子(病例数)或分母(人口数)太小,这是贝叶斯思想的核心,而其“经验”之处在于这种先验分布是基于已存在观测值的整体特征的。

如果原始发病率的方差很小,那么调整后的发病率基本不变,相反的,如果原始发病率方差很大,发病率将会调整至整体均值的水平。从经验贝叶斯的角度来看,整体均值是先验分布。位置  $i$  的发病率平滑值  $\hat{\pi}_i$  的计算公式如下,

$$\hat{\pi}_i = \frac{C_i}{\omega_i P_i + (1 - \omega_i)\theta} \quad (1)$$

其中，权重  $\omega_i$  的计算公式为：

$$\omega_i = \frac{\phi}{\phi + (\theta / P_i)} \quad (2)$$

式中， $\theta$  为人口均值， $\phi$  为人口方差， $P_i$  和  $C_i$  分别为第  $i$  区的人口和 TBE 病例数。当人口数  $P_i$  很大时， $\theta / P_i$  接近于 0， $\omega_i$  接近于 1， $\hat{\pi}_i$  接近于原始人口数；当  $P_i$  值越来越小时， $\hat{\pi}_i$  受  $\theta$  的影响则越来越大。

### 2.3.2 空间自相关

空间自相关是依据事物或现象对空间位置的依赖关系，衡量全局或局域范围内事物空间自相关性的方法(冯军等，2011)。如果邻近空间位置的观察值非常接近，并且有统计学意义，提示存在正空间自相关；如果从数值排列未能得出模式的，则为独立或随机的。该类方法可分为全局和局部两种自相关分析方法。

#### (1) 全局空间自相关

全局空间自相关主要用于描述整个研究区域的空间对象某一属性取值的空间分布状态，以判断该对象在空间上是否存在聚集性，常用的度量指标包括 Moran's  $I$ 、Getis-Ord  $G$ 。选取全局 Moran's  $I$ ，通过测量邻近空间位置观察值近似程度得到整个研究区域内各斑块的正负空间自相关特性。

Moran's  $I$  指数的计算公式如下：

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^m w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{s^2 \sum_{i=1}^n \sum_{j=1}^m w_{ij}} \quad (3)$$

式中： $x_i$  表示  $i$  要素的属性值， $x_j$  表示  $j$  要素的属性值， $\bar{x}$  表示要素属性的平均值； $n$  为要素的数目， $m$  为位于点  $i$  距离为  $d$  范围内的所有邻近要素  $j$  的数目， $w_{ij}$  为要素  $j$  对要素  $i$  的权重系数； $s^2$  为  $x$  的方差。

空间自相关特性取决于随机模式下的期望值  $E(I)$ 。当点的数量很大时， $E(I)$  则接近于 0。随机模式下，Moran's  $I$  接近于  $E(I)$ 。若相邻的点趋近于具有相近的值(即空间相关)，则 Moran's  $I$  大于  $E(I)$ ；反之，若相邻的点趋近于不等的值(即空间不相关)，Moran's  $I$  小于  $E(I)$ 。Z-Score 表示空间分布区域随机分布的可能性(Hu et al, 2011b)。

#### (2) 局域空间自相关

全局空间自相关指数使用一个单一值来反映整体上的分布模式，难以探测不同位置局部区域的空间关联模式，局部空间关联指数能进一步揭示空间单元与其

相邻近的空间单元属性特征值之间的相似性或相关性，可用于识别“热点区域”以及数据的异质性。常用的局域空间自相关度量指标包括局域 Moran's  $I$  (Anselin, 1995)、局域 Getis-Ord  $G$  (Ord et al, 1995)。与全局空间自相关分析相似，为进一步区分高值聚集或是低值聚集的两种正空间自相关，选择局域 Getis-Ord  $G$  指数，其计算公式为：

$$G_i(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_j x_j} \quad (4)$$

$$U(G_i) = \frac{G_i - E(G_i)}{\sqrt{\text{var}(G_i)}} \quad (5)$$

式中， $i$  表示要求  $G$  值的要素， $j$  表示  $i$  附近的要素， $x_j$  表示  $j$  要素的属性值， $w_{ij}(d)$  为给定距离阈值下  $i$ 、 $j$  两者空间关系的权重矩阵。

当局部 Getis-Ord  $G$  的值大于数学期望，并且有统计学意义时，提示存在“热点区”；当局部 Getis-Ord  $G$  的值小于数学期望，提示存在“冷点区”。

### 2.3.3 格网方法

地理格网系统是在二维离散的网格化空间上表达复杂的地表时空分布与动态变化，即以格网为单元描述或表达地物的属性分类、统计分级以及变化参数、虚拟现实等。格网思想最早起源于农耕社会的“井田制”土地划分方法，“井”表示四周没有边界的网格，“田”代表四周封闭的网格 (陈述彭等, 2002)。现代科技中的格网已超出了传统意义上空间位置划分的范畴，它是构建在互联网上的一组新兴技术，将高速互联网、计算机、大型数据库、传感器等设备融为一体，为科技工作者和广大普通用户提供丰富的资源、功能和相关服务 (李德仁, 2005)。地理学中，格网模型被用作地理空间的多层次划分，这既是空间位置的划分方法，也是特定空间位置范围内资源、环境、社会、经济属性的信息载体和数据容器，能够为多源信息的集成提供一种良好的方法。相比于其他方式更为高效的运算速度，地理格网能够提高针对于各种空间信息的时空分析能力，深度挖掘数据中蕴藏的内在规律，提炼其中有效的信息，全面系统地认识自然、资源、环境、人口、社会经济诸多问题(王卷乐等, 2012)。

对不同空间模型和不同变量类型的要素数据进行格网化的方法也有所不同。空间数据模型被分为矢量及栅格两种模型(Congalton, 1997)。矢量数据模型用点、线、面等集合对象来表示简单的空间要素，这三种几何对象及其所表示的要素间的区别在于维度和性质。栅格数据是适用于表达连续现象的数据模型，其数据模型是用规则的格网来覆盖整个空间，格网中的各个像元值与其位置上的空间现象特征相对应，像元值的变化反映了现象的空间变异(Chang, 2010)。由于栅格数据

具有存储结构简单且运算速度快的优势,在数据格网化的时候,通常先将矢量数据转化为栅格格式,再进行栅格数据的格网化处理。以气象站点数据(矢量点)的格网化为例,首先对气象站点的各种属性数据进行插值处理,得到覆盖研究区整体的栅格数据,再利用预先设定的统一的格网系统对栅格数据按照一定的规则进行分区统计(Zonal Statistics),使得格网系统内每一个格网都对应具有一个属性值,统计规则包括均值(Mean)、最大值(Max)、最小值(Min)、众数(Majority)、值域(Range)、标准差(STD)、总和(SUM)等(ArcGIS Help 10.2)。

变量依据其属性数据的特点被划分为名义型、有序型或连续型三种变量。名义型变量是没有顺序之分的类别变量,例如土地覆被类型数据,不同的土地覆被类型(森林、草地、耕地、建设用地等)之间没有顺序之分,即使赋予了不同土地覆被类型各自的 ID 值作为其标识符,也并不意味着这些土地覆被类型之间就有了顺序;有序型变量表示一种顺序关系,而非数量关系,比如表示干旱程度的变量,其属性值分为湿润、半湿润、半干旱和干旱等不同等级,我们只知道湿润的区域比半湿润的区域要湿润,但是并不知道更加湿润的程度有多少;连续型变量可以呈现为某个范围内的任意值,并同时表示了顺序和数量,例如年龄变量就是一个连续型变量,它能够表示成 10.1 或 15.3 的值,显然,15 岁的人比 14 岁的人年长 1 岁(Kabacoff, 2011)。在进行格网化时,名义型和有序型变量可以依据面积最优法(众数规则)进行格网化属性赋值,连续型变量可以通过均值(Mean)、最大值(Max)、最小值(Min)、众数(Majority)、值域(Range)、标准差(STD)、总和(SUM)等规则进行格网化属性赋值,也可以先通过等级划分,划分为有序型变量之后再依据面积最优法进行格网化赋值,这种方法可以依据变量属性值域的等级划分区间对空间范围进行区域划分。

#### 2.3.4 地理探测器

地理探测器(Wang et al, 2010b)使得格网的概念能够应用到空间流行病学的研究中来。如图 8 所示,在数据预处理阶段,首先在预定义的格网体系中,将疾病分布图层与疑似因素图层做空间叠加,以此来计算影响要素空间类别分区内疾病流行率的均值和方差。之后,对要素的不同空间类别分区之间进行疾病流行率均值差异的显著性检验。若某种要素的类别分区之间的疾病流行率均值差异显著,且每个类别分区内部流行率的变异非常小,则该要素可以解释疾病流行率的空间变异 (Li et al, 2015)。

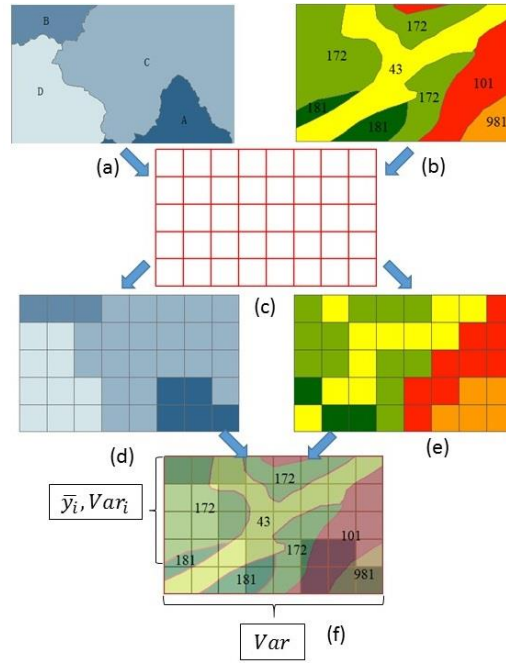


图 8 地理探测器数据预处理原理

地理探测器由四种探测器组成，即，风险探测器、因子探测器、生态探测器和交互作用探测器。风险探测器通过比较不同类别分区之间健康风险指标的平均值以搜索健康风险的区域，均值显著大的类别分区，健康风险也就越大。因子探测器用于检验某种地理因素是否是形成健康风险空间分布格局的原因。生态探测器通过比较各个要素间健康风险指标总方差的差异，来探究不同的地理要素在影响疾病的空间分布方面的作用是否有显著的差异。交互探测器可以识别危险因子之间的交互作用(王劲峰，2010)，主要利用因子探测器，通过比较同一要素类别各子区域发病率的累计离散方差与格网个数之积占整体离散方差与格网个数之积的比例，从而得到各要素对该疾病空间分布的解释力。

要素对发病率空间分布的解释力可以表示为：

$$P_{D,T} = 1 - \frac{\sum_{i=1}^{D_n} n_{D_i} Var_{D_i}}{nVar} \quad (6)$$

式中，D 代表某类地理环境要素，T 代表发病风险(本研究用发病率代替发病风险)， $P_{D,T}$  是该类要素对发病风险空间分布的解释力； $n_{D_i}$  是该类要素某一子区域  $D_i$  内格网的个数， $Var_{D_i}$  代表该子区域内发病风险的方差， $n$  是研究范围整体的格网个数， $Var$  是该区域整体方差。

如果确实存在一些影响因子可以使得子类别分区间的均值具有最大差异，子区域内方差最小，则 TBE 发病率的空间分布受该类因子的影响很大。在一些极

端情况下, 当  $\sum_{i=1}^{D_n} n_{D_i} Var_{D_i}$  等于 0, 则  $P_{D,T}$  值等于 1, 这种要素所构成的分区称为“完美分区”。

### 2.3.5 多元线性回归

回归分析能够用于探测地理环境因子与疾病发病情况之间的定量关系, 并进一步模拟得到疾病的风险分布情况。本研究选用多元线性回归(Multiple Linear Regression, MLR)分析。

MLR 模型的一般形式如下,

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (7)$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + \varepsilon_i \quad (i=1, \dots, n) \quad (8)$$

式中,  $x_i$  为自变量,  $y$  为因变量,  $\beta_0$  为截距,  $\beta_i$  为偏回归系数,  $\varepsilon$  为随机误差, 并假定  $\varepsilon_i \sim N(0, \sigma^2)$ 。回归参数估计使用普通最小二乘法(OLS), 根据残差平方和最小的原则得到最优方程。

最优回归方程的选择方法有: 全局择优法、逐步剔除法、逐步引入法、逐步回归法等。

a) 全局择优法是求出所有可能的回归模型, 共有  $(2^m - 1)$  个模型,  $m$  为自变量的个数, 再按某一准则选择最优模型, 该方法的局限性是模型的个数会随自变量个数的增加而呈指数倍增加。

b) 逐步剔除法是先将自变量全部放入方程, 然后逐步剔除的回归方法。选择偏回归平方和最小的变量, 作  $F$  检验, 得到相应的  $P$  值, 依据  $P$  值的大小决定该自变量是否被剔除, 结束该步骤以后, 利用剩余变量建立新的回归方程, 并重复以上操作。该方法的局限性是自变量高度相关时, 可能得不到正确的结果。

c) 逐步引入法是自变量从无到有、从少到多的回归方法, 因变量首先对每一个自变量作直线回归, 对回归平方和最大的自变量作  $F$  检验, 有意义 ( $P < 0.05$ ) 则引入, 在此基础上, 计算其它自变量的偏回归平方和, 选取偏回归平方和最大者作  $F$  检验, 重复以上操作, 局限性是, 后续变量的引入可能会使先进入方程的自变量变得不重要。

d) 逐步回归法, 即先将所有自变量分别与因变量建立一元线性回归方程, 按照偏回归平方和的大小将变量引入到方程中, 将剩余变量再分别与因变量以及已引入的自变量建立二元回归方程, 并检验回归方程, 剔除不显著变量。引入和剔除变量的标准是当  $F$  统计量的显著性概率小于等于 0.05 时, 变量被引入回归方程, 当显著性大于等于 0.10 时, 变量被移出回归方程。与其他方法相比, 逐步回归法有如下优点: 它所含的自变量个数较少, 便于应用; 它的剩余标准差也较

小, 方程的稳定性较好; 由于每步都作检验, 保证了方程中所有的自变量都有显著性。回归检验的方法有方差检验法、决定系数法、t 检验法等。

#### a) 方差检验法

方差检验法(F 检验)为多元线性回归模型中对回归模型整体所做的显著性检验。零假设是所有回归系数均为 0, 备择假设是所有的回归系数不全为 0。

$$SS_T = SS_R + SS_e \quad (9)$$

$$F = \frac{SS_R / p}{SS_e / (n - p - 1)} \sim F(p, n - p - 1) \quad (10)$$

上式中, 总平方和( $SS_T$ )可以被分解为回归平方和( $SS_R$ )与残差平方和( $SS_e$ )两个部分,  $n$  为  $SS_T$  的自由度,  $p$  为  $SS_R$  的自由度。

在假设检验水平为  $\alpha$  的情况下, 若  $F \leq F_{\alpha}(p, n-p-1)$ , 接受零假设, 说明所有的解释变量与回归变量之间都不存在回归关系; 若  $F > F_{\alpha}(p, n-p-1)$ , 则拒绝原假设, 意味着解释变量与回归变量之间存在回归关系。

#### b) 决定系数法

决定系数( $R^2$ )用于检验回归方程的拟合优度, 即自变量能解释因变量变异的百分比, 取值(0, 1), 越接近 1 模型拟合越好。

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_e}{SS_T} \quad (11)$$

式中,  $SS_T$  为总平方和,  $SS_R$  为回归平方和,  $SS_e$  为残差平方和。

实际中, 随着自变量个数  $p$  的增加和样本数  $n$  的变化, 必然会使得  $R^2$  不断变化, 在进行拟合优度检验的判定时, 一般采用校正决定系数( $R_c^2$ ), 以消除自变量的个数以及样本量的大小对  $R^2$  的影响。 $R_c^2$  的计算公式如下,

$$R_c^2 = 1 - (1 - R^2) \frac{n-1}{(n-1)-p} = 1 - \frac{SS_e / (n-1-p)}{SS_T / (n-1)} \quad (12)$$

#### c) t 检验法

F 检验和  $R^2$  检验都是将所有的解释变量作为一个整体来检验它们与被解释变量  $y$  的相关程度以及回归效果, 为进一步确定模型中哪些是重要的解释变量, 哪些是非重要的解释变量, 可以进行 t 检验。零假设是回归系数为 0, 备择假设是回归系数不为 0。

零假设成立的条件下,  $t = \frac{\hat{\beta}_j}{s(\hat{\beta}_j)}$ , ( $j=1, 2, \dots, k-1$ )。其中,  $s(\hat{\beta}_j)$  表示参数

估计值  $\hat{\beta}_j$  的估计标准差。在假设检验水平为  $\alpha$  的情况下, 若样本计算的  $|t| \leq$



$t\alpha/2 (T - k)$ ，则接受零假设，若样本计算的 $|t| > t\alpha/2 (T - k)$ ，则拒绝零假设。

d) 其他检验指标

AICc (校正赤池信息量准则)是衡量统计模型拟合优良性的一种标准，是由日本统计学家赤池弘次创立和发展的，可以权衡所估计模型的复杂度和此模型拟合数据的优良性，AICc 值越小，模型拟合效果最佳。

### 2.3.6 地理加权回归

地理加权回归(GWR)的实质是局部加权最小二乘法，其中的权为待估点所在的地理空间位置到其他各观测点的地理空间位置之间的距离函数。这些在各地理空间位置上估计的参数值描述了参数所研究的地理空间位置的变化情况，用以探索空间数据的非平稳性。GWR 的数学模型形式如下(Brunsdon et al,1996, Fotheringham, 2003):

$$y_i = a_0(u_i, v_i) + \sum_k a_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (13)$$

式中， $y_i$  为第  $i$  点的因变量； $x_{ik}$  为第  $k$  个自变量在第  $i$  点的值， $k$  为自变量记数； $i$  为样本点记数； $\varepsilon_i$  为残差， $(u_i, v_i)$  为第  $i$  个样本点的空间坐标； $a_k(u_i, v_i)$  为连续函数  $a_k(u, v)$  在  $i$  点的值。GWR 的估计值是：

$$a(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y \quad (14)$$

式中， $W(u_i, v_i)$  为距离权重矩阵，是一个对角矩阵，对角元素为  $(W_{i1}, W_{i2}, \dots, W_{in})$ ，非对角线元素为零， $n$  是样本量， $W_{ij}$  为第  $j$  点对  $i$  点的影响，一种定义是  $W_{ij} = \exp(-d_{ij}^2/h^2)$ ，这里  $d_{ij}$  为  $i$ 、 $j$  两点间距离， $h$  为自定义带宽(王劲峰等, 2010)。



## 第三章 森林脑炎流行时空特征

### 3.1 流行病学特征

#### 3.1.1 时间特征

综合 2006~2013 年 8 年间呼伦贝尔市各年份以及各旬(每 10 天为 1 旬)的 TBE 发病病例数(如图 9), 2006~2013 年, TBE 每年的发病病例数基本稳定在 50~80 之间, 只在 2007 和 2011 年出现了两次峰值, 2007 年 TBE 发病病例数达 95 人, 2011 年 TBE 发病病例数上升到 214 人, 为该市历年发病人数最高的一年。

TBE 病例的患病时间集中于 4~9 月份, 具有明显的春夏季高发特性。主要表现为, 每年 4 月份开始出现 TBE 患者, 是发病的起始期; 4 月下旬到 5 月上旬, 患者数逐渐增加, 到 5 月中下旬时, TBE 患者数急剧增加, 是快速发展期; 到 6 月上中旬, 患者数出现峰值, 是顶峰期; 6 月下旬开始, 患者数出现了大幅下降的趋势, 进入下降期; 到 7 月中旬以后, 每旬 TBE 病例数已经不超过 10 例, 8 月份以后, 发病人数越来越少, 几乎没有病例出现。季节性的高发特点与 TBE 流行病学过程有关, 只有在作为宿主和传播媒介的蜱虫发育为成虫, 逐渐寻找大型哺乳动物作为宿主动物时, 成虫所携带的 TBE 病毒才有可能传播给人类, 这个时间恰好是春夏季节(如 1.2.2 节图 2 所示)。TBE 发病的季节特性与毕维铭等所总结的中国 TBE 发病的季节分布特性基本一致(见 1.2.2 节)。

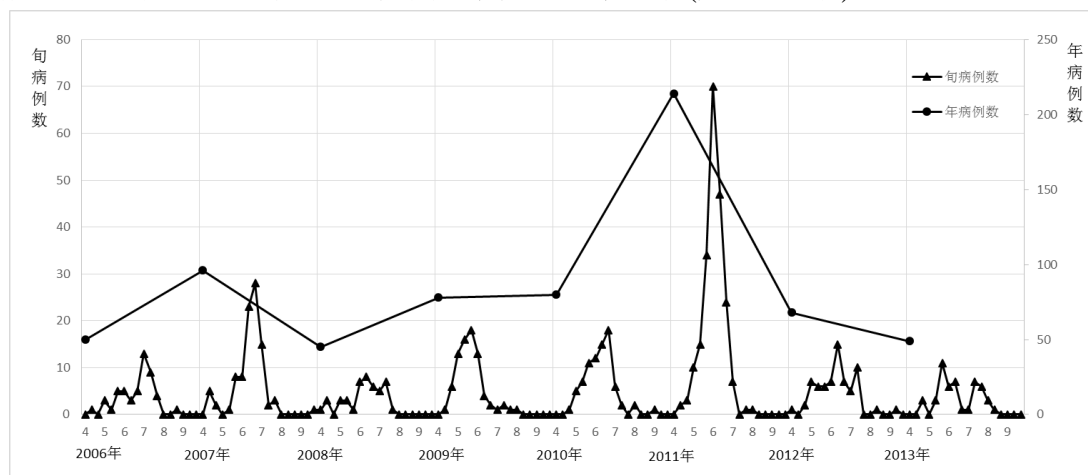


图 9 2006~2013 年 TBE 发病病例数

#### 3.1.2 人群分布

##### 1) 男女比例

如图 10 所示, 综合 2006~2013 年每月 TBE 病例数发现, 除 2013 年 8 月份女性 TBE 患病数(7)略大于男性 TBE 患病数(3)以外, 其他时间中, 男性病例数几

乎都远远超过女性病例数，8年间男性 TBE 病例数有 491 例，女性 TBE 病例数有 189 例，男性 TBE 病例数约为女性的 2.6 倍。

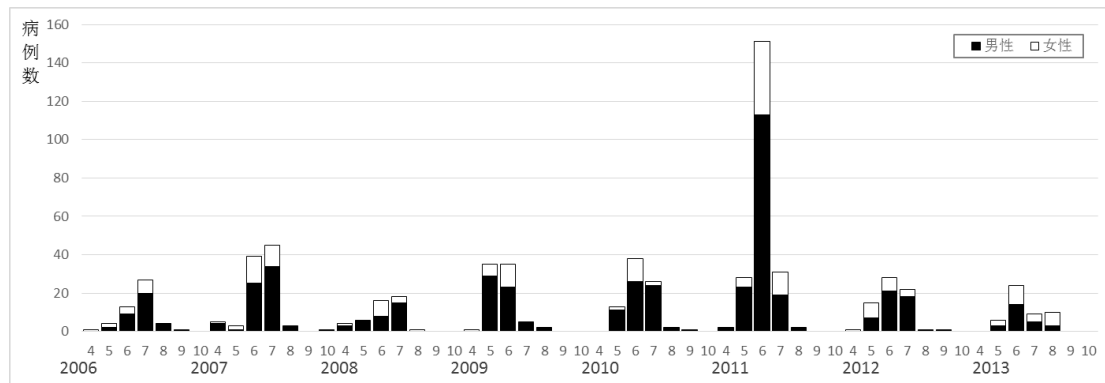


图 10 2006~2013 年各月份 TBE 男女病例数

## 2) 年龄特征

图 11 为 2006~2013 年呼伦贝尔市 TBE 发病病例的年龄构成情况，以 10 岁为年龄间隔，将 2006~2013 年 8 年间 TBE 所有病例划分为 10 个年龄段，发现处于 40~49 岁的人数最多，有 251 人，约占总人数的 1/3；30~39 岁的人数次之，有 178 人，约占总人数的 1/4；50~69 岁和 21~30 岁的人数分别为 98 和 77，平均每年约有 10 个新增病例；其他年龄段 TBE 病例数较少，共占有患病病例数的 11.18%。

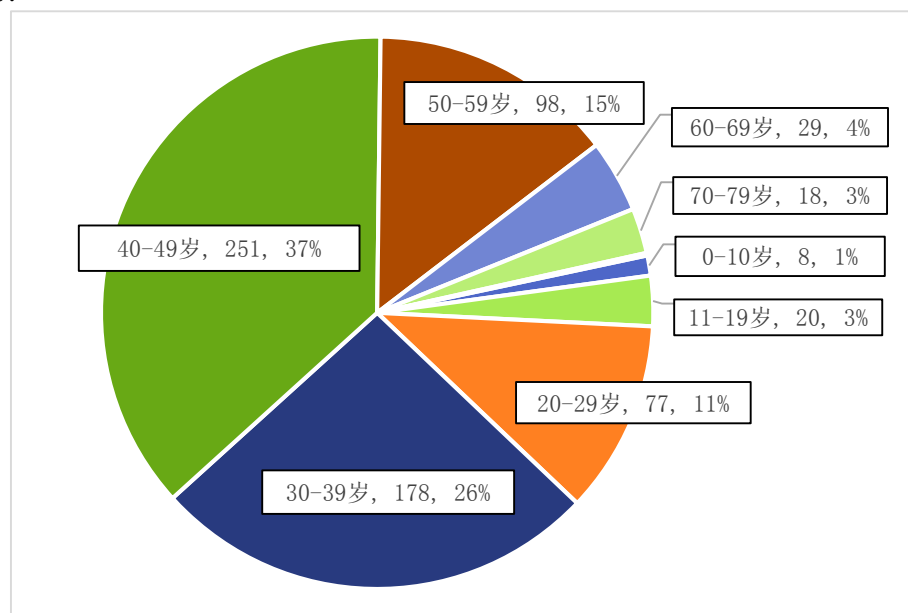


图 11 2006~2013 年呼伦贝尔市 TBE 发病病例年龄构成

## 3) 职业特征

图 12 为 2006~2013 年呼伦贝尔市 TBE 发病病例的职业构成情况，在已知职业的 481 例病例中，工人是患 TBE 病例数最多的职业，共有 205 人患病；家政、家务及待业人员的患病人数仅次于工人，有 167 人患病；农民和干部职员患 TBE 的病例数分别为 30 和 27，另有少数病例的职业为教师(4)、离退人员(16)、学生

(12)、牧民(9)。进一步对患 TBE 工人的工作单位进行分析,发现大部分患 TBE 的工人均是林业、林场和林研所等与森林作业相关的机构人员。

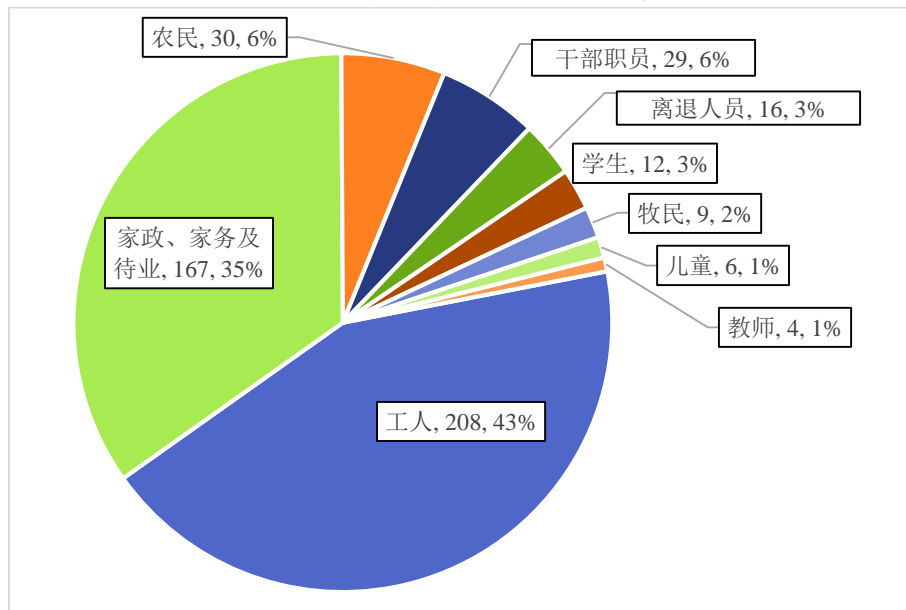


图 12 2006~2013 年呼伦贝尔市 TBE 发病病例职业构成

## 3.2 空间特征

### 3.2.1 全局空间自相关

利用增量空间自相关方法测量 2006~2013 年 8 年间平均 TBE 病例数一系列距离的空间自相关距离阈值,  $z$  得分为峰值时所对应的距离值为 106818.21 m, 表明在该距离处 2006~2013 年 8 年平均病例数的空间分布具有最大的空间自相关特性。将距离阈值设置为 106818.21 m, 利用 ArcGIS 中全局 Moran's  $I$  工具计算 2006~2013 年每年病例数以及 8 年平均病例数的全局空间自相关结果, 如表 1 所示。

表 1 2006~2013 年呼伦贝尔市各乡镇 TBE 病例数全局空间自相关计算结果

	2006 年	2007 年	2008 年	2009 年	2010 年	2011 年	2012 年	2013 年	2006~2013 平均值
Moran's $I$	0.156	0.113	0.092	0.089	0.110	0.106	0.079	0.115	0.144
期望值	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008
方差	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.004	0.003
$z$ 得分	3.205	2.085	1.791	1.707	2.169	2.003	1.485	2.081	2.655
$p$ 值	0.001	0.037	0.073	0.088	0.030	0.045	0.137	0.037	0.008

根据表 1 中各年份的计算结果, 2006 年  $p$  值小于 0.01, 该年 TBE 患者的空间分布仅有小于 1% 的可能性是随机分布的, 出现数据聚集的可能性大于随机分布的可能性, 能够显著地拒绝零假设, 这说明呼伦贝尔市在 2006 年 TBE 患者的空间分布具有显著的聚集特性, 且具有空间正相关模式; 2007、2010、2011、2013

年的  $p$  值介于 0.01~0.05 之间, 这 4 年 TBE 患者的空间分布仅有小于 5% 的可能是随机分布的, 出现数据聚集的可能性大于随机分布的可能性, 且能够显著地拒绝零假设, 说明这 4 年间呼伦贝尔市 TBE 患者的空间分布具有较为显著的聚集特性, 且具有空间正相关模式; 2008 和 2009 年  $p$  值介于 0.05~0.10 之间, 这两年 TBE 患者的空间分布仅有 10% 的可能性是随机的, 出现空间聚集的可能性大于随机分布的可能性, 但不能显著地拒绝零假设, 说明这两年呼伦贝尔市 TBE 患者的空间分布可能呈现空间聚集的特点; 2012 年  $p$  值大于 0.10, TBE 患者出现显著随机分布的空间特性, 无法拒绝零假设。

表 1 最后一列是 2006~2013 年 8 年间 TBE 病例数平均值的全局自相关计算结果,  $p$  值小于 0.01, 8 年间 TBE 患者的空间分布仅有小于 1% 的可能性是随机分布的, 出现空间聚集的可能性大于随机分布的可能性, 且能够显著地拒绝零假设, 说明 8 年间呼伦贝尔市 TBE 发病情况具有非常显著的聚集特性, 且具有空间正相关模式。

### 3.2.2 局域空间自相关

将距离阈值设置为 106818.21 m, 利用 ArcGIS 中局域 Getis-Ord  $G$  工具计算 2006~2013 年每年以及 8 年平均 TBE 病例数的 Getis-Ord  $G$  局域空间自相关结果如图 13 所示。

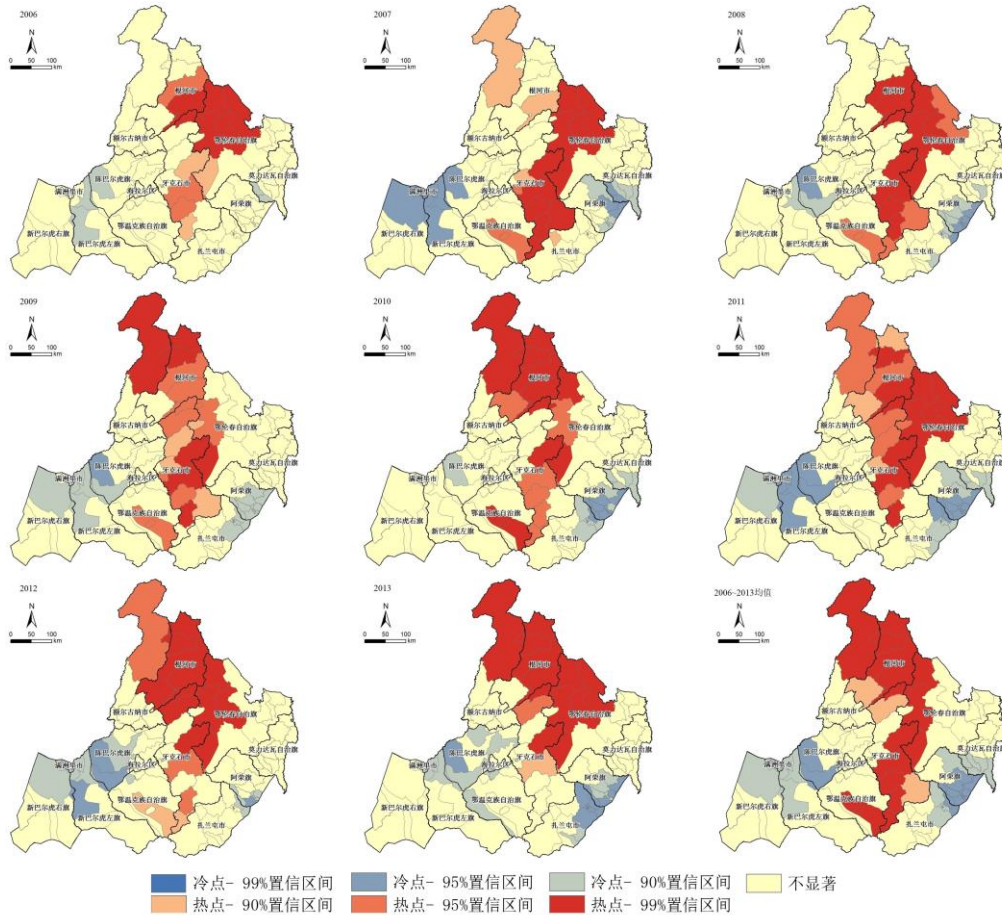


图 13 2006~2013 年呼伦贝尔市各乡镇 TBE 病例数局域自相关“冷热点”空间分布

由 2006~2013 年 TBE 病例数均值的局域空间自相关空间分布可知，发病病例数的低值分布区主要聚集于呼伦贝尔市东西两侧，其中，西侧主要是满洲里市及其附近的几个乡镇，东侧的低值聚集区是靠近黑龙江省的阿荣旗。TBE 病例数的高值聚集区分布在呼伦贝尔市岭中的大部分地区，集中偏向大兴安岭山脉的东麓，覆盖了额尔古纳市北部、根河市大部、鄂伦春自治旗西部、牙克石东南部以及鄂温克自治旗东南面小部分地区。

由 2006~2013 年各年 TBE 病例数局域空间自相关的空间分布可知，2006 年，TBE 病例数的高值区主要聚集于大兴安岭北部地区，包括根河河西街道以及鄂伦春自治旗西北部克一河、甘河、阿里河、大杨树、诺敏河、毕拉河的几个林业局；2007、2008 年，TBE 病例数的高值聚集区逐渐向南部延伸，一直到牙克石南部塔尔气镇，高值聚集区呈现带状分布；2009 年，呼伦贝尔最北部莫尔道嘎镇成为新的高值聚集区，该年的高病例数聚集区分为两个，一个是最北部莫尔道嘎镇以及根河市的满归镇和阿龙山镇，另一个是牙克石中部免渡河镇和乌尔其汉镇；2010 年，高病例数聚集区在此前基础上，扩展到鄂温克族自治旗的伊敏苏木和红花尔基镇；2011 年高病例数聚集区回归到中北部在鄂温克族自治旗的西北部各林业局以及与其相接壤的牙克石市部分乡镇；2012~2013 年高病例数聚集区由中北部扩展到最北部莫尔道嘎镇。

### 3.3 本章小结

本章主要利用统计和空间分析的方法探究呼伦贝尔市 TBE 发病的时空特征。利用传统的流行病学统计方法,可以得到:1)该市 TBE 的患病时间主要集中在 4~9 月份,具有明显的春夏季高发特性;2)男性明显多于女性,8 年间男性 TBE 病例数约为女性的 2.6 倍;3)具有明显的职业特性,与林业作业相关的职业患病人数较多。利用全局和局域空间自相关分析,可以得到:1)呼伦贝尔市整体的空间自相关特性在不同年份表现为不同的聚集性特点,其中 2006 年全市 TBE 病例数具有最高的空间聚集性特点,2012 年 TBE 病例数表现为随机分布特征;2)2006 年 TBE 发病率的高值聚集区集中于大兴安岭北部区域,2007~2008 年逐渐向南部延伸,2009~2010 年有两个高发病率聚集区,2011 年回归到大兴安岭的中北部区域,2012~2013 年高发病率聚集区继续向北扩展到莫尔道嘎镇。



## 第四章 森林脑炎地理环境影响因子探测

由第三章分析可见，TBE 病例的职业特征与森林作业相关性较大，发病率的分布具有明显聚集性，说明 TBE 与地理环境具有较强的影响关系。基于地理格网的地理探测器方法能够同时适应于栅格和矢量数据，在空间异质性的基础上得到各要素对发病率的解释能力，并进一步探测出风险的分布情况。本章首先对各地理环境因子的空间分布特征进行分析，然后基于对研究区地理环境的分析特点，利用地理探测器探测影响 TBE 的地理环境因子。

### 4.1 地理因子空间分布特征

结合 3.2.2 节呼伦贝尔市 TBE 发病率的局域空间自相关特性，对研究区内地理环境要素的特点进行分析。

图 14-16 分别为 DEM、Slope、Aspect 的空间分布。由地形要素的空间分布可得，呼伦贝尔市按地形特点可以分为岭中、岭东、岭西三个地理单元。岭中海拔较高且坡度较陡，地势较为复杂，额尔古纳市最北部原始森林区的平均海拔约为 600~700 m，沿大兴安岭逐渐向南向东，至根河境内，山体的东西跨度逐渐变大，平均海拔逐渐增大，除山间河谷地带，大部分地区海拔高于 800 m，再到中心牙克石市境内，除西南部海拉尔河谷区域地势略低，其他地区平均海拔也在 700~900 m 之间，继续向南到新巴尔虎左旗、鄂温克族自治旗以及扎兰屯市南部交界山区，山体面积达到最大，海拔高度达到最高，最南端部分地区海拔高度达到 1000 m 以上，是呼伦贝尔市地势最为复杂，坡度最为陡峭的区域。沿中部山脉向东向西，坡度逐渐减小，海拔逐渐较低，东部靠近松辽平原处，平均海拔约为 150~300 m，西部呼伦湖海拔约为 500~600 m 左右，最西部与蒙古国相邻地区地势稍高，海拔约为 650~860 m 左右。全区坡向值域为  $0^{\circ}$  ~  $237^{\circ}$ ，95% 的格网单元其坡向大致分布在  $130^{\circ}$  ~  $220^{\circ}$  左右，中值约为  $174^{\circ}$ ，即大部分地理格网的坡向是朝南的。

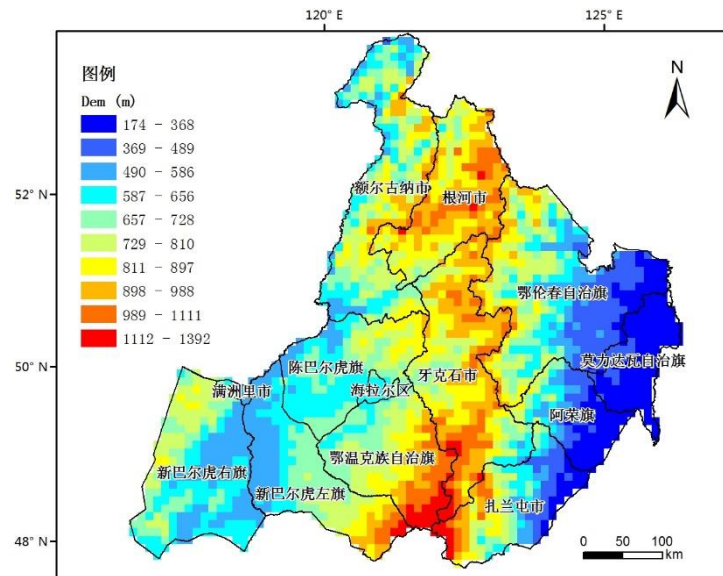


图 14 研究区 DEM 空间分布

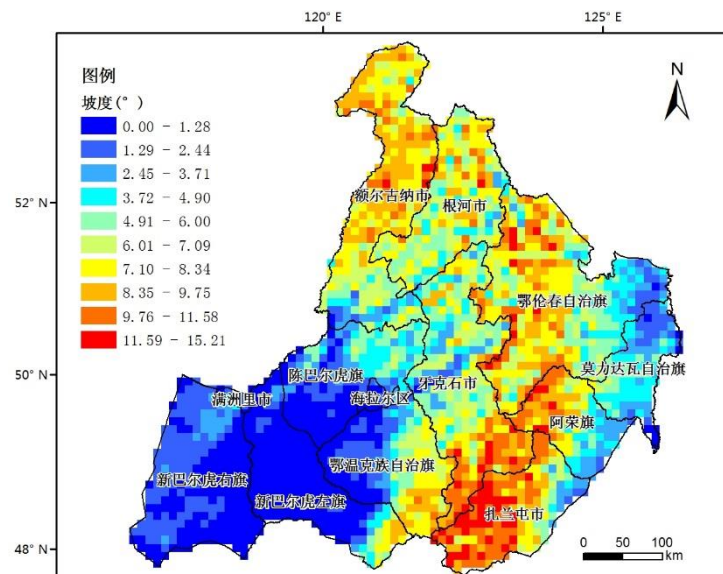


图 15 研究区坡度空间分布

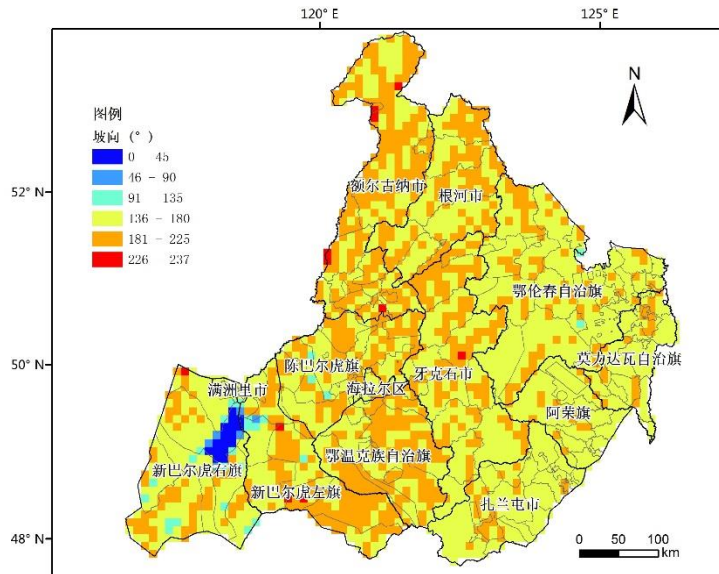


图 16 研究区坡向空间分布

图 17 至 21 分别是研究区 2006~2013 年 8 年间温度、相对湿度、降雨频度、降雨量以及日照时数平均值的空间分布，原始数据为气象监测站点的空间插值数据。由气象要素空间分布可得，研究区温度和相对湿度的空间分布具有很大的负相关特性，平均温度由北向南逐渐变大，岭中北部地区年均温最小，而相对湿度的空间分布恰好相反，岭中北部地区相对湿度最大，自南向北逐渐减小；全市日降水量大于等于 0.1 mm 的天数(降雨频度)平均每年大致有 71 到 135 天左右，这项指标的空间分布与相对湿度的空间分布相似，北部降雨频度较南部更高一些；降雨量沿大兴安岭山脉自西北向东南逐渐递增，这与大兴安岭山脉地形关系十分密切，岭东位于大兴安岭山脉的迎风坡，对偏南暖湿气流起阻挡和抬升作用，月均降雨量达到 40 mm 以上，西部靠近蒙古国的新巴尔虎右旗月均降雨量只有 2mm 左右；日照时数的空间分布呈现由西南部向东北部逐年减少的趋势，按每月有 30 天，西南部平均每天的日照时数大致为 8.5 h，北部地区大致为 7 h。

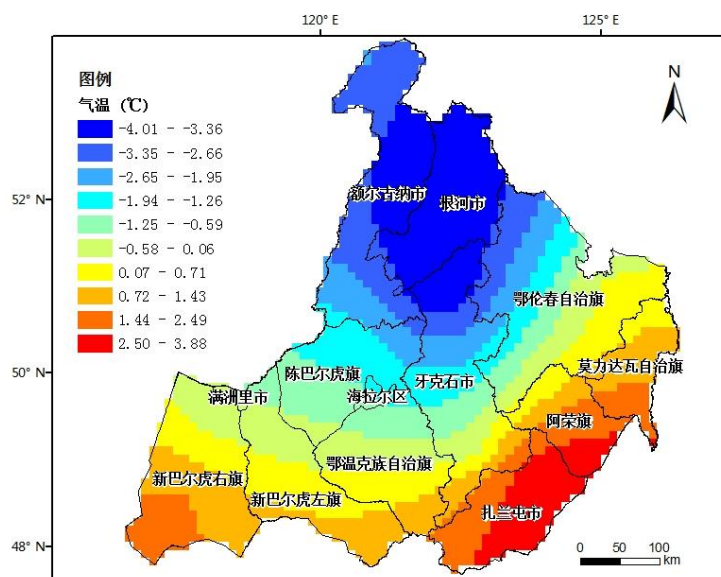


图 17 研究区气温空间分布

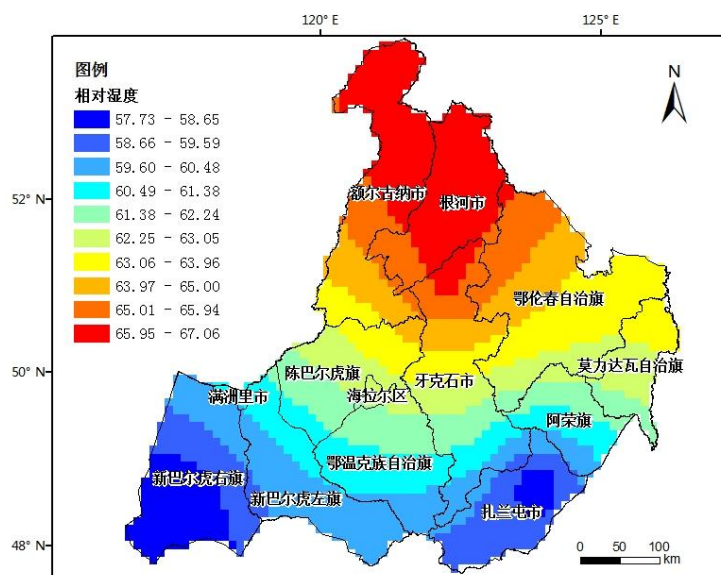


图 18 研究区相对湿度空间分布

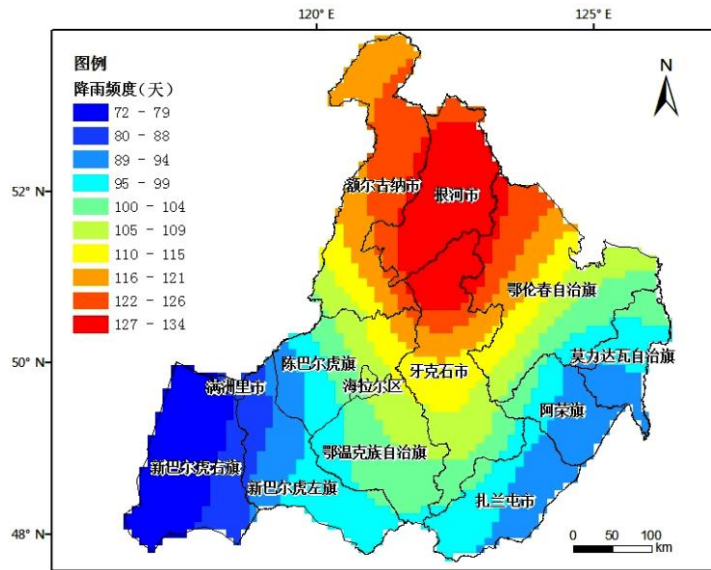


图 19 研究区降雨频度空间分布

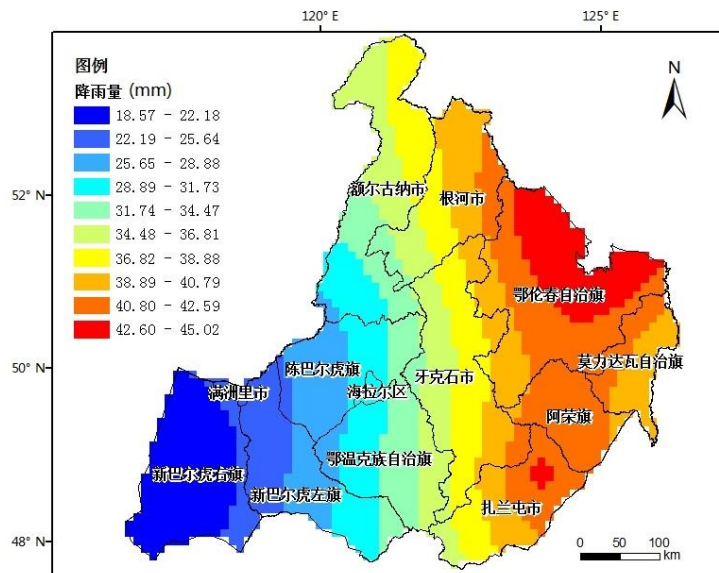


图 20 研究区降雨量空间分布



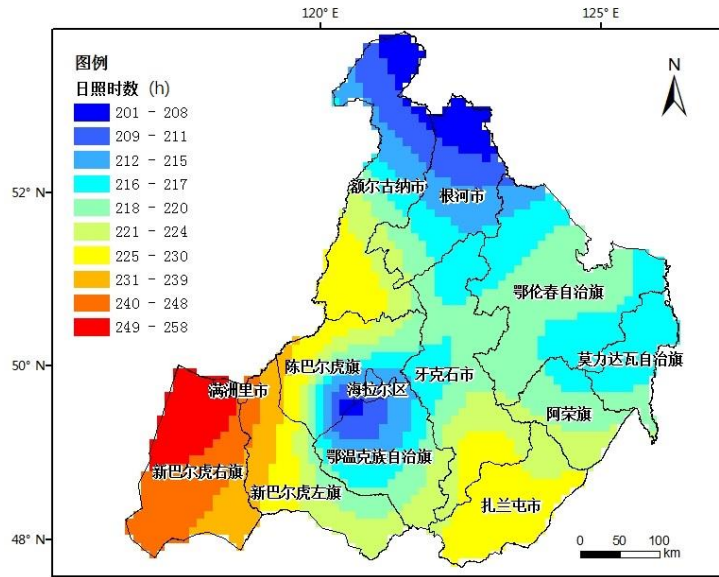


图 21 研究区日照时数空间分布

图 22、23 分别是 NDVI 和 EVI 的空间分布, 由植被指数的空间分布可得, 研究区内两种植被指数的空间分布具有很大的相关性, 高植被覆盖区主要聚集在岭中根河市、鄂伦春自治旗、牙克石市及其临近的部分县镇, 额尔古纳市北部大兴安岭原始森林区 NDVI 最高, 约在 0.53~0.57 左右, 东部部分乡镇的 NDVI 值较小, 约在 0.28~0.32 左右, 西部除呼伦湖 NDVI 值为负值以外, 其他区域 NDVI 值大致在 0.27 以下, 植被覆盖度为三个分区中最小。

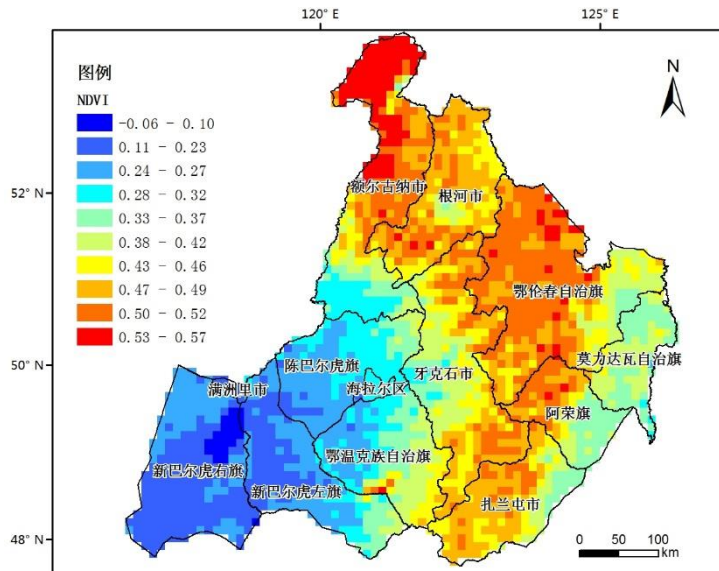


图 22 研究区 NDVI 空间分布

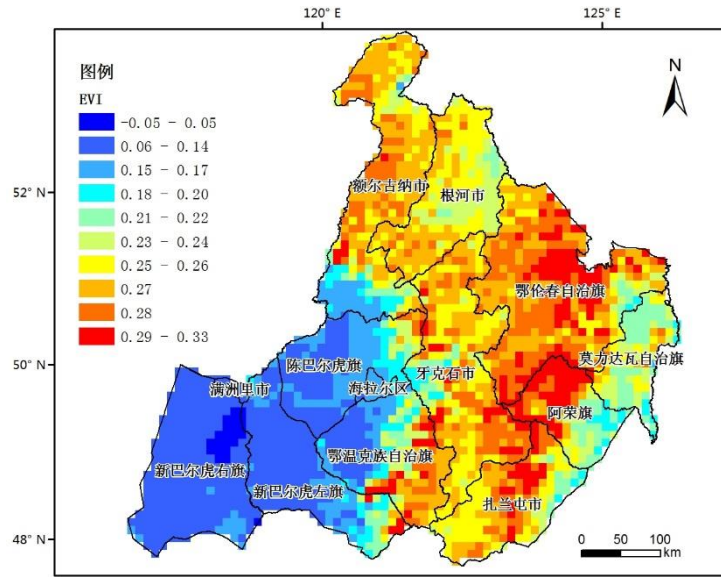


图 23 研究区 EVI 空间分布

## 4.2 地理因子探测

### 4.2.1 数据预处理

依据 3.2 和 4.1 节中分别对 TBE 发病率和各类地理环境因子的空间异质性探测结果, TBE 发病率和各类地理环境因子的分布具有空间异质性的特点。因此, 本文利用地理探测器对呼伦贝尔市地理环境影响因子对 TBE 风险分布的影响机理进行初步探测。参与计算的地理环境因子包括 DEM、坡度(Slope)、坡向(Aspect)、温度(Temp)、降雨量(PPT)、降雨频率(PF)、相对湿度(RH)、光照时数(SH)、归一化植被指数(NDVI)、增强植被指数(EVI)共 10 个数值型变量, 以及土壤类型和土地覆被类型 2 个类别型变量。数据预处理的过程包括分级分类和格网化两个步骤。

1) 分级分类。数值型变量直接进行格网化使得每个格网都有不同的属性值, 即使格网的属性值相近, 但数值上仍具有差异性, 为减少研究区内数值差异对探测结果的影响, 本节中对所有数值型地理环境要素以及 TBE 发病率本身进行分级分类。数据分级的方法包含等间距法(Equal Interval, EI)、几何分级法(Geometrical Interval, GI)、自然分割法(Jenks Natural Break, NB)等。等间距法可以将数值型变量依据值域划分为相等间隔的连续区间, 几何分级法对具有正态分布的数值型变量划分结果良好, 自然分割法对具有偏态的数值型变量划分结果良好。参考 Wu 等(2016)利用不同分级方法所得的实验结果, 本文利用自然分割法将 TBE 发病率、地形(DEM、Slope)、植被指数(NDVI、EVI)以及 5 个气候要素(Temp、PPT、PF、RH、SH)划分为 5 个等级, 依据先验知识, 坡向(Aspect)通过等间距法划分为东南、西南等 8 个方向类别, 分级结果如图 24 以及表 2、3 所示。

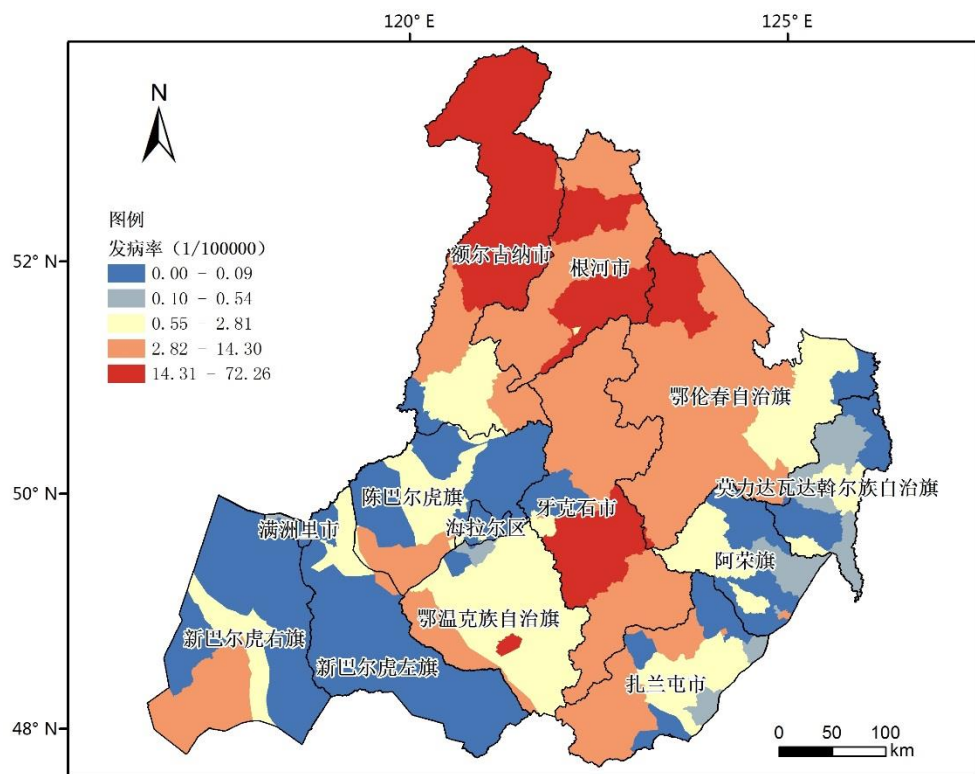


图 24 呼伦贝尔市 TBE 发病率等级空间分布

表 2 地形要素和植被指数的分级结果

DEM / m	Slope / °	NDVI	EVI	Level
165 - 466	0 - 2.86	-0.09 - 0.1	-0.08 - 0.05	1
480 - 660	2.87 - 6.86	0.11 - 0.28	0.06 - 0.17	2
670 - 820	6.87 - 11.43	0.29 - 0.38	0.18 - 0.23	3
830 - 1,000	11.44 - 17.53	0.39 - 0.47	0.24 - 0.27	4
1,100 - 1,700	17.54 - 48.58	0.48 - 0.67	0.28 - 0.41	5



表 3 气象要素分级结果

Temp / °C	PPT / mm	PF/ 天	RH / %	SH / h	Level
-4.08 - -2.57	18.48 - 24.32	72 - 85	57.69 - 59.5	201 - 213	1
-2.56 - -1	24.33 - 29.84	86 - 99	59.51 - 61.16	214 - 221	2
-0.99 - 0.38	29.85 - 34.64	100 - 109	61.17 - 62.86	222 - 231	3
0.39 - 1.92	34.65 - 39.22	110 - 121	62.87 - 64.82	232 - 245	4
1.93 - 3.92	39.23 - 45.06	122 - 135	64.83 - 67.11	246 - 258	5

2) 格网化处理。利用 ArcGIS 中区域统计工具对发病率和各要素因子的矢量和栅格图层进行格网化处理, 所采用的格网化准则为众数准则(Majority), 即以格子内等级或类别面积最大的属性等级或类别定义为该格子的属性。

#### 4.2.2 因子探测

首先, 利用因子探测器计算所有地理环境因子对 TBE 发病风险空间分布的解释力  $P_{D,T}$  值(见 2.3.4 节), 并依据解释力值对各要素进行分级分析。根据各类要素对发病率空间分布的解释力  $P_{D,T}$  的大小, 所有地理环境要素被分为 3 个等级: 第一等级的要素( $P_{D,T} \geq 20\%$ )包括土壤类型(25.83%)、降雨量(24.90%)和土地覆被类型(20.44%); 第二等级的要素( $5\% \leq P_{D,T} < 20\%$ )包括降雨频度(11.04%)、温度(6.40%)及光照时数(5.80%); 第三等级的要素( $P_{D,T} < 5\%$ )包括坡向(1.22%)、高程(1.06%)、坡度(0.56%)、相对湿度 (0.45%)、增强植被指数 (0.41%)和归一化植被指数(0.19%)。进一步利用风险探测器, 探测第一等级的各要素类型对 TBE 发病风险影响的空间差异。

1) 土壤类型: 如图 25、26 所示, 棕色针叶林土(1)、灰色森林土(4)以及草原风沙土(9)所分布的岭中山区 TBE 发病风险等级最高; 沼泽土(2)和暗棕壤(3)所分布区域的 TBE 发病风险等级处于中间位置, 沼泽土主要分布于人口稠密的大兴安岭山谷地带, 暗棕壤主要分布于呼伦贝尔东部农场区; 黑钙土(6)和栗钙土(8)所分布的西部呼伦贝尔大草原 TBE 发病风险等级最低。

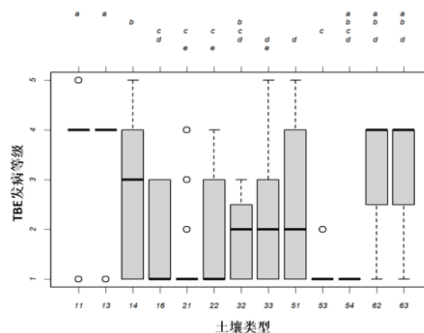


图 25 不同土壤类型间 TBE 发病风险分布差异

注：横坐标的数字对应图 25 中土壤类型，黑色横线表示各要素类型内 TBE 均值所在等级类别，箱体的上下线分别表示上下四分位等级数，上下边缘线分别是最大和最小等级数，空心小圆圈表示异常值，a 到 e 的 5 个字母是多重比较结果，凡有一个相同标记字母的即为均数差异不显著的类型，凡没有相同标记字母的即为均值差异显著的类型。

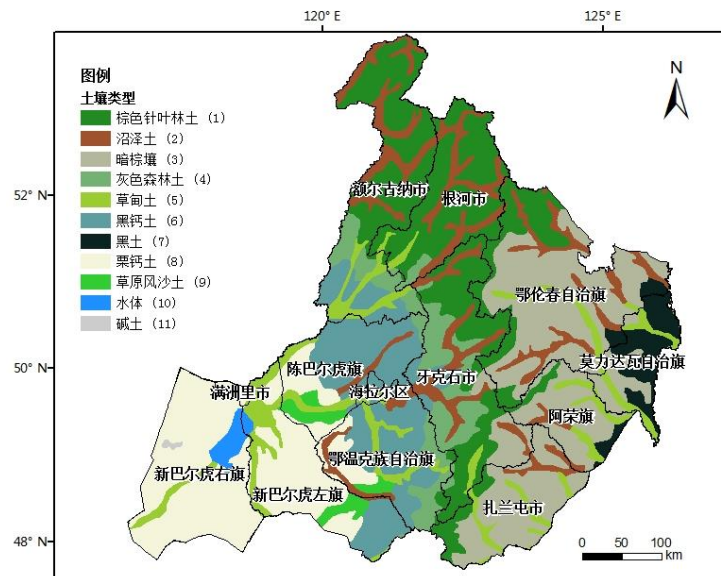


图 26 呼伦贝尔市土壤类型空间分布

2)降雨量：研究表明各类气候因子对 TBE 发病风险具有交互影响，它们会对 TBE 主要传播媒介——蜱虫的生命周期、生活环境产生影响，降雨的空间分布与 TBE 发病风险的相互作用如图 27、28 所示，第 4 降雨等级(34.65~39.22 mm)区域内 TBE 发病风险最高，第 5 降雨等级(39.23~45.06 mm)区域内 TBE 发病风险次之，其他区域降雨量都较低。

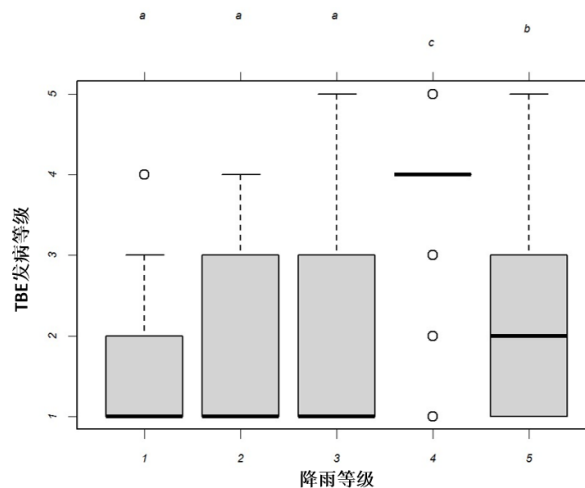


图 27 不同降雨等级 TBE 发病风险分布差异

注：横坐标的数字对应图 27 中的降雨等级，黑色横线表示各要素类型内 TBE 均值所在等级类别，箱体的上下线分别表示上下四分位等级数，上下边缘线分别是最大和最小等级数，空心小圆圈表示异常值，a 到 e 的 5 个字母是多重比较结果，凡有一个相同标记字母的即为均数差异不显著的类型，凡没有相同标记字母的即为均值差异显著的类型。

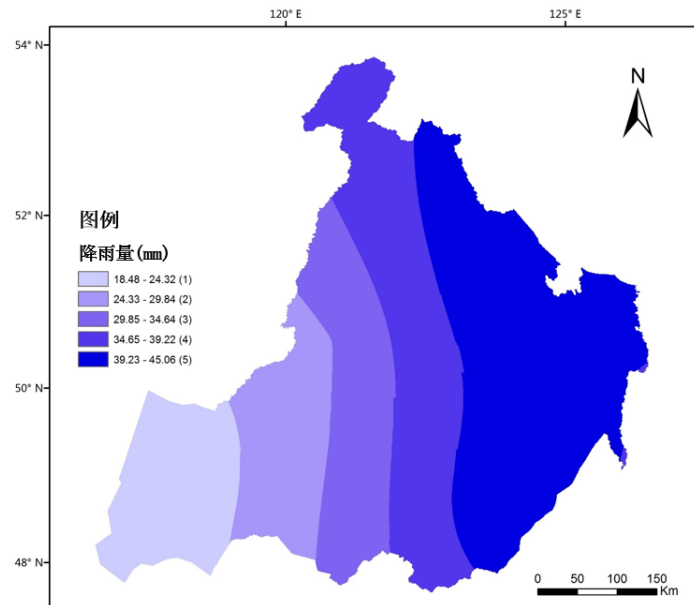


图 28 降雨等级空间分布

3)土地覆被类型: 如图 29、30 所示, 分布于呼伦贝尔中北部山区的常绿针叶林(11)和落叶针叶林(13)内 TBE 发病风险等级最高, 大兴安岭中南部阔叶针叶林(14)TBE 发病风险等级次之, 其他土地覆被类型 TBE 发病风险的等级较低。

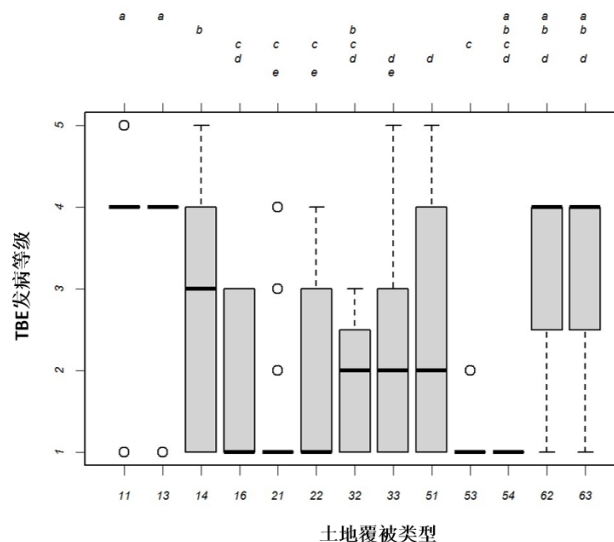


图 29 不同土地覆被类型间 TBE 发病风险分布差异

注: 横坐标的数字对应图 29 中的土地覆被类型, 黑色横线表示各要素类型内 TBE 均值所在等级类别, 箱体的上下线分别表示上下四分位等级数, 上下边缘线分别是最大和最小等级数, 空心小圆圈表示异常值, a 到 e 的 5 个字母是多重比较结果, 凡有一个相同标记字母的即为均值差异不显著的类型, 凡没有相同标记字母的即为均值差异显著的类型。

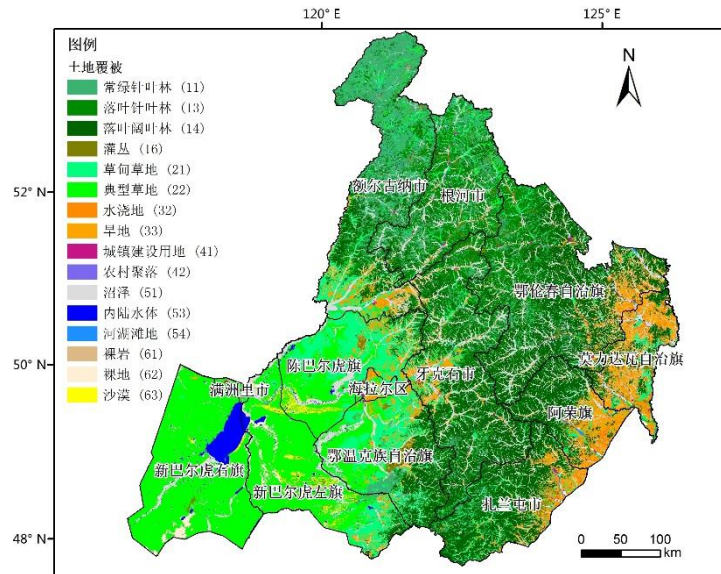


图 30 呼伦贝尔市土地覆被空间分布

#### 4.2.3 探测结果及不确定性分析

因子探测器的初步探测结果可以将各地理环境因子分为三个等级，土壤类型、降雨量以及土地覆被是影响研究区 TBE 发病风险空间分布最主要的因素，部分气候因素(降雨频度、温度和光照时数)是影响 TBE 发病风险空间分布的次要因素，地形、植被和相对湿度对 TBE 发病风险空间分布的影响最小。所得结果可以初步揭示出土壤、土地覆被类型这两类名义型变量是影响 TBE 发病率空间分布的重要因子。风险探测器所得结果表明，发病率等级最高的土壤类型——棕色针叶林土、灰色森林土与发病率等级较高的土地覆被类型——常绿针叶林、落叶针叶林以及落叶阔叶林都集中分布在岭中山区，说明决定 TBE 发病风险的土壤类型和土地覆被类型在空间分布上具有很大的相关特性。棕色针叶林土和灰色森林土主要形成于北方针叶林地带，每年大量凋落的枯枝落叶归还于土壤表层，使得土壤表层形成枯枝落叶层和腐殖质层，该土壤结构和土壤类型为蜱虫的生存提供了适宜的土壤湿度环境。针叶林是寒温带的地带性植被，主要树种为冷杉、云杉和落叶松等，林下生活着多种哺乳动物，这些动物是蜱虫寄宿的主要宿主动物。相同的植被覆盖、土壤结构以及气候特点说明土壤类型和土地覆被类型与这些要素的相关程度较高，可以作为其他数值型变量要素选择的标准和借鉴。

地理探测器所探测的相关要素的结果初步揭示了地理环境因子的影响程度，但其结果仍具有不确定性。在探测初期数据预处理的过程中，数值型变量的等级划分方法和个数的选择以及在数据格网化过程中所选用的格网大小将对探测结果产生影响。

### 4.3 本章小结

本章主要分为两个部分，第一部分内容包括对呼伦贝尔市地理环境因子空间分布特点进行分析，第二部分内容包括利用地理探测器对呼伦贝尔市 TBE 发病风险相关的地理环境因子进行探测。呼伦贝尔市地理环境因子的空间分布按地形特点可以分为岭中、岭东、岭西三个地理单元，岭中地势较为复杂且植被覆盖度较高，岭东和岭西的地形相对较为简单，且植被覆盖度相对较低；研究区内温度和相对湿度具有很大的负相关特性；东南部降雨量大于西北部，但北部降雨频度高于南部。对 TBE 发病风险相关的地理环境因子的探测结果为地理环境因子依照解释力大小分为三个等级，进一步利用风险探测器探测第一等级的三个要素对 TBE 发病风险影响的空间差异，分析可得土壤类型和土地覆被类型在空间分布上具有很大的相关特性，且与其他地理环境要素具有相关性，可以作为其他数值型变量要素选择的标准和借鉴。



## 第五章 森林脑炎风险评价

在地理环境因子初探所得结果的基础上,为进一步提高对地理环境因子的识别能力,判断 TBE 风险的空间分布,本章首先利用多元线性逐步回归方法,得到整个研究区基于空间尺度为 10 km×10 km 格网系统的 2006~2013 各年间 TBE 发病率的均值与各要素的回归关系模型;参考多元线性回归模型各要素回归系数的大小,建立发病风险与各要素的地理加权回归模型,最终得到基于空间异质性的各要素回归系数空间分布和 TBE 发病风险的空间分布。

### 5.1 多元逐步线性回归

#### 5.1.1 数据预处理

为减小由“小数问题”(Shi X et al, 2013)而带来的统计不稳定性,利用 GeoDa(<https://geodacenter.asu.edu/>)中经验贝叶斯平滑(Empirical Bayes Smoothing, EBS)方法进行空间平滑。由于平滑后的发病率分布曲线仍呈现严重的右偏态,使用多种变换方法对发病率进行变换,以消减由于因变量非正态分布而带来数据偏态的问题。表 4 是发病率经以 e 为底的对数变换(Lg(TBE\_EBS))、平方根变换(Sqrt(TBE\_EBS))、倒数变换(1/(TBE\_EBS))而得的变换结果。结果表明,对数变换所得的偏度(Skewness)和峰度(Kurtosis)最接近于其各自的标准误差(Std. Error),图 31、32 分别为原始的及经过对数变换后的 TBE 发病率分布曲线,可见变换后的分布曲线基本呈正态分布。参与回归的变量包括坡度(Slope)、坡向(Aspect)、数字高程模型(DEM)、归一化植被指数(NDVI)、增强植被指数(EVI)、相对湿度(RH)、气温(Temp)、降雨量(Prep)、降雨频率(PF)、光照时数(SH)等 10 个地理环境因子,由于土地覆被类型及土壤类型为类别型变量(哑变量),不适宜参与到其他变量均为数值型变量的高斯回归模型的计算中,因而,风险评价过程中未选择这两个变量。

表 4 原始及变换后 TBE 发病率分布曲线的峰度及偏度统计值

	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
TBE_EBS	2340	3.317	0.051	17.109	0.101
Lg(TBE_EBS)	2340	-0.198	0.051	-0.883	0.101
Sqrt(TBE_EBS)	2340	1.063	0.051	1.327	0.101
1/(TBE_EBS)	2340	3.635	0.051	16.550	0.101

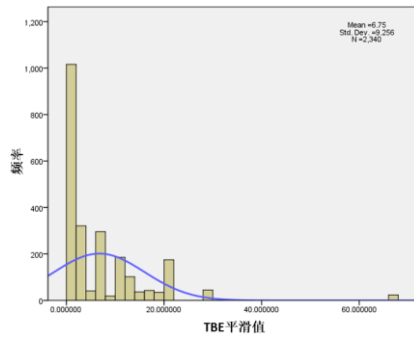


图 31 原始 TBE 发病率分布曲线

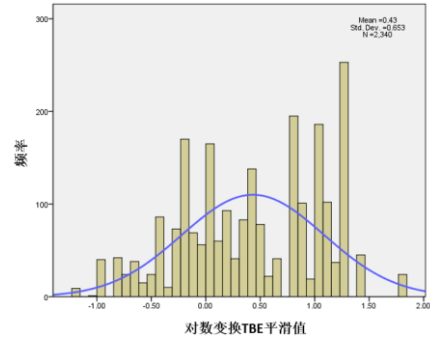


图 32 对数变换后的 EBS 平滑发病率曲线

5.1.2 模型建立

首先，利用 Spearman 秩相关性分析计算 TBE 发病率与各要素之间的相关性。结果如表 5 所示，各类要素与 TBE 发病率相关系数的 p 值均小于 0.01，因而这些要素与 TBE 发病率均呈显著性相关。

表 5 TBE 发病率与相关要素的相关系数

要素类型	坡向	坡度	DEM	EVI	NDVI	降雨量	降雨频度	光照时数	相对湿度	气温
相关系数	0.15	<b>0.56</b>	0.46	0.46	<b>0.63</b>	0.28	<b>0.68</b>	-0.40	<b>0.60</b>	<b>-0.60</b>
p 值	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01

注：加粗代表相关系数大于 0.5。

分别建立所有因子与 TBE 变换发病率的一元线性回归方程，选择偏回归平方和最大的要素引入方程；再将剩余要素分别与 TBE 原始发病率以及已引入的要素建立二元回归方程，同样选择偏回归平方和最大的要素引入方程，检验并剔除已引入变量中 F 统计量不显著的要素；剩余变量依照以上方法逐步引入到回归方程中。

多元逐步线性回归方法最终共得到 12 个回归方程，所涉及到的要素及回归优度如表 6 所示。最优模型的可决系数为 0.580，该模型包括了 Slope、DEM 两个地形要素，NDVI、EVI 两个植被指数以及 Temp、PF、Prep、RH 四个气候要素，共计 8 个要素类型。



表 6 回归模型及模型检验指标

模型编号	要素种类	F 值	显著性	可决系数(R <sup>2</sup> )	校正决定系数(R <sub>c</sub> <sup>2</sup> )
1	PF	1457	0.000	0.384	0.384
2	PF, Slope	916	0.000	0.440	0.439
3	PF, Slope, Temp	831	0.000	0.516	0.516
4	PF, Slope, Temp, NDVI	670	0.000	0.535	0.534
5	PF, Slope, Temp, NDVI, RH	610	0.000	0.566	0.566
6	PF, Slope, Temp, NDVI, RH, EVI	527	0.000	0.576	0.574
7	Slope, Temp, NDVI, RH, EVI	632	0.000	0.575	0.574
8	Slope, Temp, NDVI, RH, EVI, Aspect	530	0.000	0.577	0.576
9	Slope, Temp, NDVI, RH, EVI, Aspect, DEM	456	0.000	0.578	0.577
10	Slope, Temp, NDVI, RH, EVI, Aspect, DEM, Prep	402	0.000	0.580	0.578
11	Slope, Temp, NDVI, RH, EVI, Aspect, DEM, Prep, PF	358	0.000	0.580	0.579
12	<b>Slope, Temp, NDVI, RH, EVI, DEM, Prep, PF</b>	<b>402</b>	<b>0.000</b>	<b>0.580</b>	<b>0.579</b>

最优回归方程如下，系数参照表 4 系数列：

$$\lg(R_{TBE}) = 15.336 + 0.026X_{Slope} - 0.207X_{DEM} + 5.148X_{NDVI} - 3.978X_{EVI} - 0.449X_{Temp} - 0.263X_{RH} + 0.022X_{Prep} - 0.006X_{PF} \quad (15)$$

式中， $R_{TBE}$  代表原始 TBE 发病率， $\lg(R_{TBE})$  代表经过对数变换后的 TBE 发病率， $X_{Slope}$  代表坡度变量， $X_{DEM}$  代表高程变量， $X_{NDVI}$  代表归一化植被指数， $X_{EVI}$  代表增强植被指数， $X_{Temp}$  代表温度变量， $X_{RH}$  代表相对湿度， $X_{Prep}$  代表降雨量， $X_{PF}$  代表降雨频度。如表 7 所示，在显著性水为 0.01 的水平下，除 DEM 和降雨频度以外的回归常数和偏回归系数均显著；在显著性水为 0.05 的水平下，所有系数都显著。

标准化方程如下，系数见表 4 标准系数列：

$$\lg(R_{TBE}) = 0.133Z_{Slope} - 0.064Z_{DEM} + 0.897Z_{NDVI} - 0.379Z_{EVI} - 1.376Z_{Temp} - 1.041Z_{RH} + 0.236Z_{Prep} - 0.145Z_{PF} \quad (16)$$

式中，Z 表示各要素标准化以后的替代变量。

表 7 最优回归模型回归系数

	系数	标准系数	t 值	显著性
Constant	15.336		11.040	0.000
Slope/ °	0.026	0.133	4.548	0.000
DEM/ km	-0.207	-0.064	-2.337	0.020
EVI	-3.978	-0.379	-7.723	0.000
NDVI	5.148	0.897	14.969	0.000
PF/ days	-0.006	-0.145	-2.546	0.011
Prep/ mm	0.022	0.236	4.475	0.000
Temp/℃	-0.449	-1.376	-14.604	0.000
RH/ %	-0.263	-1.041	-11.111	0.000

按照标准化系数绝对值 $|Z_k|$ 的大小进行排列，气温( $|Z_k|=1.376$ )>相对湿度( $|Z_k|=1.041$ )>NDVI( $|Z_k|=0.897$ )>EVI( $|Z_k|=0.379$ )>降雨量( $|Z_k|=0.236$ )>降雨频率( $|Z_k|=0.145$ )>坡度( $|Z_k|=0.133$ )>DEM( $|Z_k|=0.064$ )。在不考虑空间异质性的前提下，呼伦贝尔市整体范围内地理环境因子与 TBE 发病风险的关系按照从大到小的顺序可以排列为：气温和相对湿度、植被指数、降雨和地形情况。

## 5.2 地理加权回归

### 5.2.1 数据预处理

GWR 自变量因子的选择是在多元线性逐步回归标准化系数的基础上，参考 4.3.1 节地理环境因子空间分布进行的。要素间的相关性检验结果表明，相对湿度(RH)与气温(Temp)间的相关性系数约为-0.89，降雨量(Prep)与降雨频率(PF)间的相关性系数约为 0.60，NDVI 与 EVI 间的相关性系数约为 0.95，这 3 对因子在数值上具有较大的相关性，且其物理意义相近，是造成 5.1.2 节中多元逐步线性回归模型存在共线性（最优模型在第 4 及更高维度的条件数 VIF 大于 10）的主要原因。因此，本实验在保留 3 个地形因子(坡度、坡向、DEM)的基础上，对 RH 与 Temp、Prep 与 PF、NDVI 与 EVI 这 3 对具有较大相关性的要素进行组合，最终选择拟合效果最优的因子组合作为最优 GWR 模型。

呼伦贝尔范围内相对湿度和降雨量的值域范围较小，值域分别为[57.73, 67.06]和[18.57, 45.02]，在 ArcGIS 中进行 GWR 运算时，出现了相同值的空间聚类现象，无法参与运算。结合这两个变量的数值特点，对其进行了放大处理，相对湿度利用 $(RH-57)*10000$ 进行数值变换，降雨量利用 $(Prep-18)*10000$ 进行数值变换，变换后的指标分别用相对湿度变换指标(RH\_CH)和降雨量变换指标(Prep\_CH)来替代。

5.2.2 模型建立

由于降雨频度的空间自相关程度较为显著，在进行空间自相关检验时，Z-Score 为 66.59，远大于期望值，不满足 GWR 对要素的条件要求，因而在降雨量和降雨频度这一对相关较大的要素中，选择降雨量变换指标(Prep\_CH)参与计算。最终选定 4 种要素组合参与 GWR 计算，如表 8 所示。

表 8 GWR 不同要素组合模型

模型序号	模型要素组合	R <sup>2</sup>	R <sub>c</sub> <sup>2</sup>	AICc	残差平方和
1	DEM, Slope, Aspect, Prep_CH, NDVI, RH_CH	0.98	0.99	7.55	0.24
2	DEM, Slope, Aspect, Prep_CH, EVI, RH_CH	——	——	——	——
3	DEM, Slope, Aspect, Prep_CH, NDVI, Temp_CH	0.87	0.88	56.85	11.56
4	DEM, Slope, Aspect, Prep_CH, EVI, Temp_CH	0.96	0.96	24.46	3.42

注：模型 2 不满足 GWR 模型运算的条件，因而无计算结果。

拟合效果最好的是第 1 个模型，包括 3 类地形要素、NDVI、降雨量和相对湿度变换指标的模型 1，模型整体的决定系数(R<sup>2</sup>)约为 0.980，校正决定系数(R<sub>c</sub><sup>2</sup>)约为 0.986，AICc 和残差平方和的值分别约为 7.55 和 0.237，是三个模型中最小的，具有很好的模拟效果。

5.3 结果分析

5.3.1 回归系数空间分布

图 33 到图 38 是 GWR 回归模型 1 的系数空间分布图，呼伦贝尔市 TBE 发病风险与不同地理环境因子间的相关程度具有空间异质性。

1)相对湿度：呼伦贝尔北部 3 个县市及最南部部分乡镇的 TBE 发病风险与相对湿度呈正相关，这些地区的 TBE 发病风险随该地区相对湿度的增加而增加；而位于西部的新巴尔虎右旗 TBE 发病风险却与相对湿度呈负相关关系，这些区域内相对湿度较大的区域，发病风险反而较小。

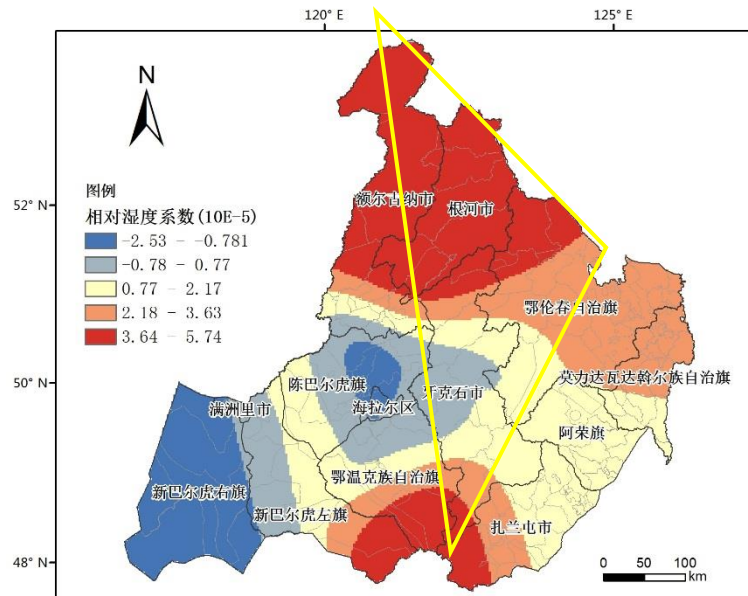


图 33 相对湿度变换值回归系数空间分布

2)植被指数: 由图 34 中 NDVI 系数的空间分布可知, 呼伦贝尔市的陈巴尔虎左旗、鄂伦春自治旗以及最南端部分乡镇 TBE 发病风险与 NDVI 呈现正相关性, 这些区域内 TBE 发病风险会随植被覆盖程度的增加而增加; 在发病风险最高的牙克石市中部乌奴耳镇和免渡河镇, 发病风险与 NDVI 呈现负相关性, 该地区 TBE 发病风险的增加与植被覆盖程度的增加并无直接关系。

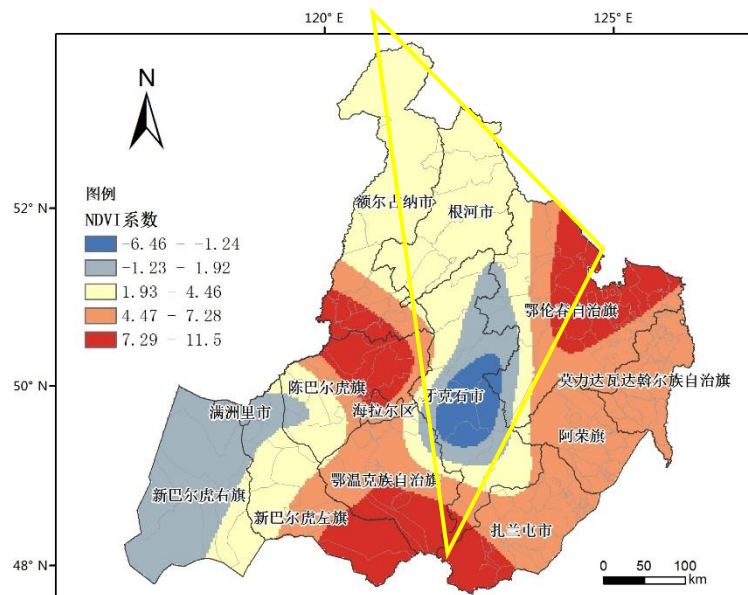


图 34 NDVI 回归系数空间分布

3)降雨量: TBE 发病风险与降雨量在呼伦贝尔市西南部为负相关关系, 该地区 TBE 的发病风险随该地区降雨量的增多而减少; 在中部牙克石市呈现正相关关系, 呼伦贝尔中部地区随 TBE 发病风险随该地区降雨量的增大而增大。

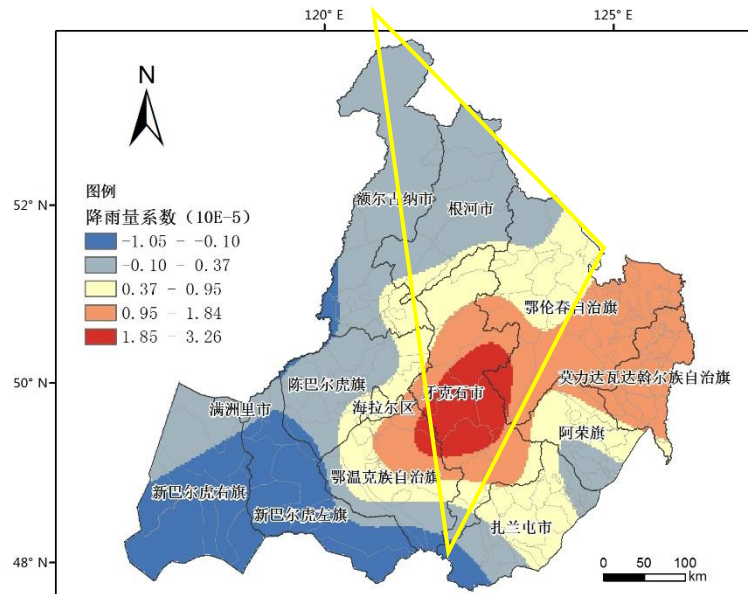


图 35 降雨量变换值回归系数空间分布

4)DEM: 呼伦贝尔 TBE 发病风险与 DEM 的相关程度自西向东呈现由负相关, 到中部相关程度较小, 再到东部正相关的特点。在呼伦贝尔西部, 高程较高的地区发病风险较高程较低的地区发病风险小; 在呼伦贝尔东部, 高程较高的地区发病风险较高程较低的地区发病风险大。

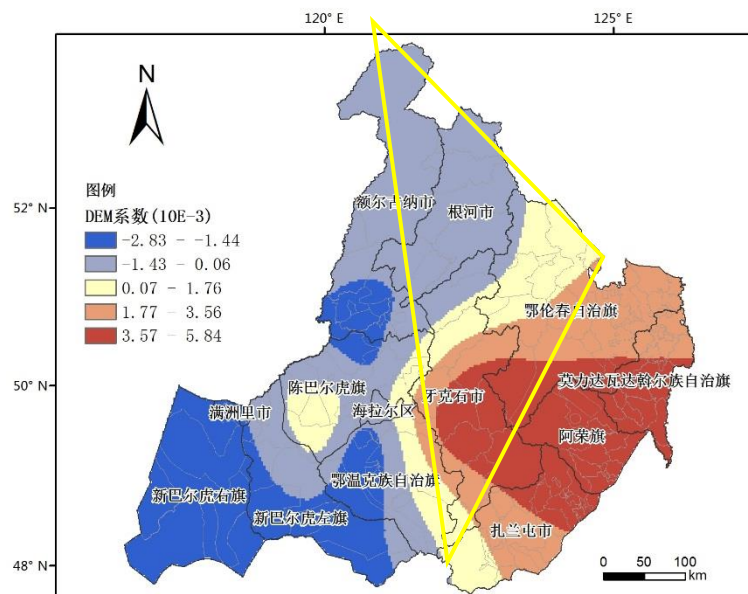


图 36 DEM 回归系数空间分布

5)坡度: 呼伦贝尔 TBE 发病风险与坡度的相关程度自西向东依次由正相关, 到中部相关程度较小, 再到东部呈现负相关的特点, 呼伦贝尔市西部草原范围内, TBE 发病风险在坡度较陡的地区比坡度平缓的地区高, 而在东部广阔的农田分布区内, TBE 发病风险在坡度较陡的地区比坡度平缓的地区低。

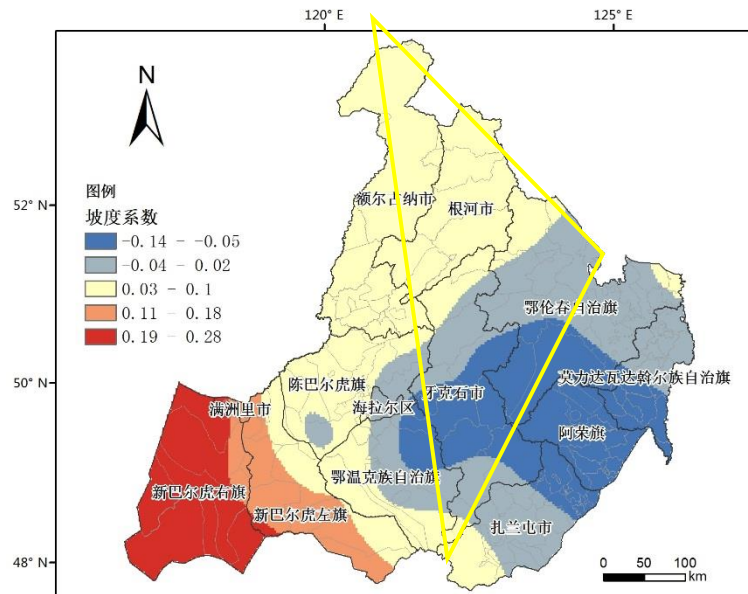


图 37 坡度回归系数空间分布

6)坡向: 坡向在整个呼伦贝尔北部地区与 TBE 发病风险的相关程度并不明显,而在南部地区则呈现出两种不同的态势,在鄂温克自治旗和新巴尔虎右旗内, TBE 发病风险与坡向呈现负相关关系,而在东南部阿荣旗和扎兰屯市, TBE 发病风险与坡向呈现正相关关系。

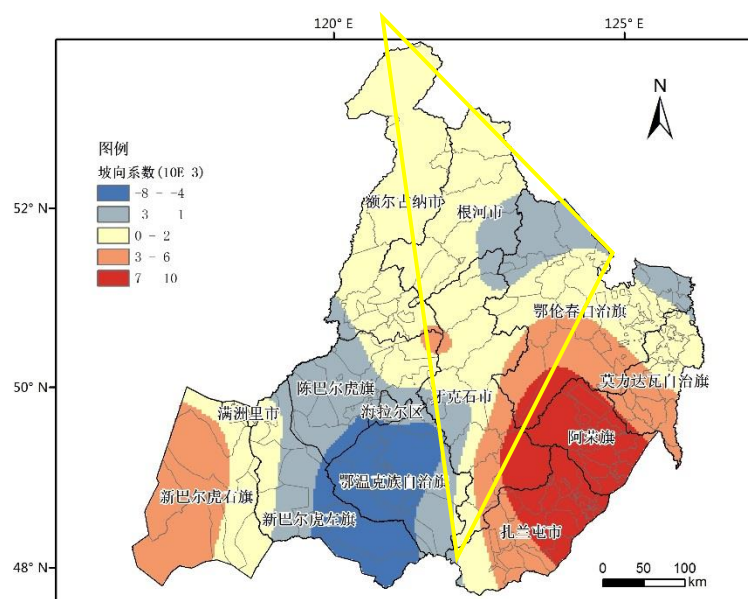


图 38 坡向回归系数空间分布

### 5.3.2 风险空间分布评价

图 39 为平滑发病率经过对数变换所得发病率的空间分布。如图所示,呼伦贝尔市中部免渡河和乌奴耳镇是 TBE 发病率最高的区域,沿大兴安岭向南向北



各乡镇 TBE 发病率较东西两侧各乡镇更高，到大兴安岭北部额尔古纳市的莫尔道嘎镇和根河市阿龙山镇、河西办事处以及鄂伦春自治旗的甘河林业局，TBE 发病率显著高于其他各乡镇；东部阿荣旗、莫力达瓦自治旗以及西部新巴尔虎左旗、新巴尔虎右旗、陈巴尔虎旗以及满洲里市 TBE 发病率显著低于中部各县市。

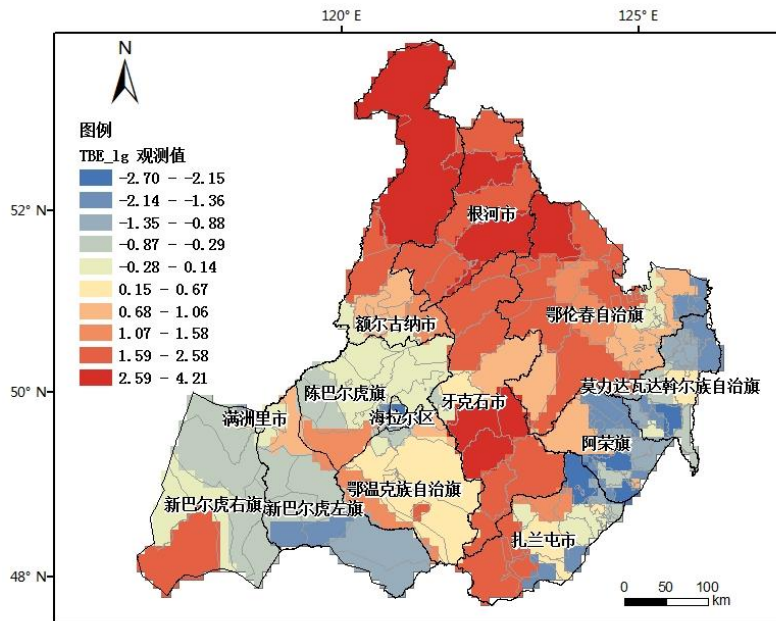


图 39 TBE 平滑发病率对数变换观测值

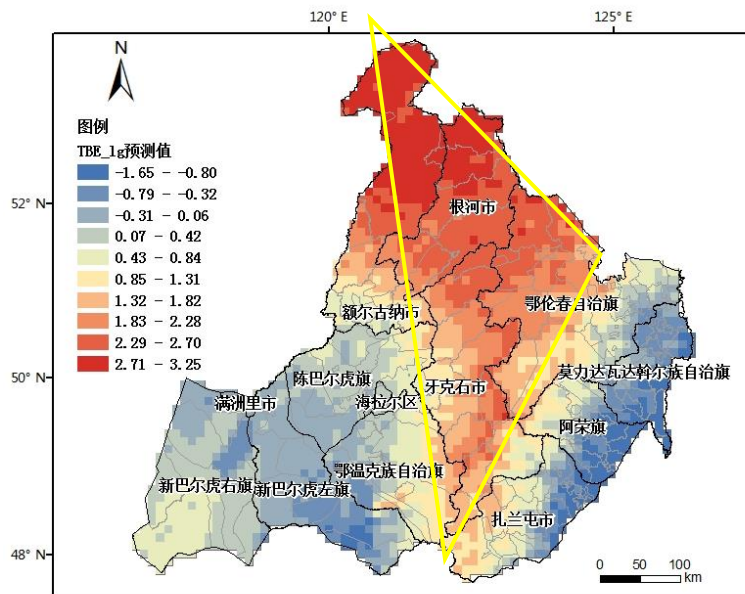


图 40 TBE 平滑发病率变换回归值

图 40 为 GWR 回归值的空间分布，以下用经过对数变换的平滑发病率回归值作为 TBE 发病风险的表达指标。格网化的空间分布打破了乡镇行政界线，能够利用更为细致的格网单元表达 TBE 发病风险的空间分布情况，具体表现在：

1) 中部高发病风险三角带

中部是 TBE 高发病风险的三角带，向东向西发病风险逐渐降低。高发病风险三角带的具体边界自额尔古纳及根河北部大兴安岭原始森林分布区，沿中俄边界一直到鄂伦春自治旗的西北部是 TBE 发病风险最高的三角边区域。顺势向南延伸到大兴安岭山区政治、经济、医疗和交通中心的牙克石市，东侧以牙克石市与鄂伦春、莫力达瓦以及阿荣旗的交界线，同时也是大兴安岭东翼与东北平原的交界线为界，西侧以牙克石市与呼伦贝尔市市政府所在的海拉尔区和鄂温克族自治旗的交界线，同时也是大兴安岭西翼与呼伦贝尔大草原的交界线为界；南部顶角汇集于牙克石市最南端的塔尔气镇落叶针叶林区。

结合 5.3.1 节中各地理环境因子回归系数的空间分布情况，三角区域不同空间位置上 TBE 发病风险与不同地理环境要素的相关性不同。三角内的大部分区域 TBE 发病风险随当地相对湿度的增大而增大，北部和南部顶角处与相对湿度的正相关性最大，相对湿度每增加 0.1，发病风险就会增加 3.64~5.74，从顶角到三角形中部，发病风险与相对湿度的相关程度越来越小，到牙克石中部由正相关变为负相关；除中部牙克石市部分乡镇，三角内大部分区域 TBE 发病风险与 NDVI 均呈现正相关性，北部额尔古纳和根河市部分乡镇内，NDVI 每增大 0.1，发病风险会增加 0.19~0.45，南部顶角和东北部鄂伦春自治旗部分县市发病风险与 NDVI 呈正相关特性，NDVI 每增大 0.1，发病风险会增加 0.73~1.15；三角内大部分区域发病风险与降雨量呈现正相关关系，中间牙克石市中部平均每月的降雨量每增大 1 mm，发病风险会增加 1.85~3.26；三角北部额尔古纳市和根河市 TBE 发病风险与高程呈负相关，高程每增高 1000 m，发病风险会减小 1.44~2.83，中南部区域 TBE 发病风险与高程呈正相关，高程每增高 1000 m，发病风险会增大 1.77~5.84；三角北部额尔古纳市和根河市 TBE 发病风险与坡度呈正相关，坡度每增加 10°，TBE 发病风险增加 0.3~1，三角南部发病风险与坡度呈负相关，坡度每减小 10°，TBE 发病风险减小 0.5~1.4。

## 2)西部低风险地带

呼伦贝尔市西部是低发病风险空间聚集地带，主要分布于大兴安岭岭西呼伦贝尔草原。东侧大兴安岭西翼和西南角新巴尔虎右旗靠近蒙古国处发病风险较其他乡镇发病风险略高，中部新巴尔虎左旗发病风险较小。

结合 5.3.1 节中各地理环境因子回归系数的空间分布情况，该区域不同空间位置上 TBE 发病风险与不同地理环境要素的相关性不同。西部新巴尔虎右旗发病风险与相对湿度呈负相关，即相对湿度每增加 0.1，发病风险就会减少 0.78~2.53，东北部与高风险三角区相接区域发病风险与相对湿度呈正相关，即相对湿度每增加 0.1，发病风险也会增加 3.64~5.74；除西部新巴尔虎右旗以外，其他区域与 NDVI 呈现不同程度的正相关特性；该区域内西南部新巴尔虎右旗和新巴尔虎左旗境内发病风险与降雨量呈负相关，即降雨量每增加 1 mm，发病风险



就会减少 0.78~2.53，鄂温克自治旗东北部与高风险三角区相接区域发病风险与降雨量呈正相关，即降雨量每增加 1 mm，发病风险也会增加 1.85~3.26；大部分区域与高程呈现负相关，高程每增加 1000 m，发病风险将会降低最多 2.83；大部分区域与坡度呈正相关，坡度没增加 1°，发病风险会增加 0.03~0.28。

### 3) 东部低风险条带

呼伦贝尔市东部呈现为低风险条带区，发病风险自大兴安岭东翼向东逐渐降低，主要包括莫力达瓦自治旗、阿荣旗和扎兰屯三市。

结合 5.3.1 节中各地理环境因子回归系数的空间分布情况，该区域大部分地区的 TBE 发病风险与相对湿度、NDVI、降雨量、DEM 和坡向均具有正相关关系，相对湿度、NDVI 和降雨量在该风险条带的最南端和最北端与其 TBE 发病风险的相关程度高于中间地带，而 DEM 和坡向在该风险条带的最南端和最北端与其 TBE 发病风险的相关程度低于中间地带。

## 5.4 本章小结

本章首先利用多元线性逐步回归得到各地理环境因子与发病风险之间标准化回归系数的大小排列，结合局域自相关特性，选择 4 种要素组合参与地理加权回归，选择回归效果最佳的地理加权回归模型，得到该模型下各地理环境因子回归系数的空间分布，以及回归风险空间分布并进行分析。格网化的 TBE 发病风险的空间分布打破了乡镇行政界线，以更为细致的格网单元体现出呼伦贝尔市 TBE 发病风险空间格局，并分析得到各风险分区域各地理环境因子的相关特性。



## 第六章 结论与展望

### 6.1 结论

本研究以中国 TBE 发病情况最为严重的内蒙古呼伦贝尔市为研究区，结合时空分析方法分析该市 TBE 流行的时空特征，利用地理探测器初步探测 TBE 相关地理环境因子，并通过回归建模方法对其风险的空间分布进行评价，得到的主要结论如下：

1) 通过对呼伦贝尔市 TBE 发病情况的流行病学调查，发现该市 TBE 的发病情况具有明显的春夏季高发特性，患病人群中男性多于女性，且具有明显的职业性特点。空间自相关的探测结果表明，2006~2013 年间，该市 TBE 发病的空间分布具有显著的空间自相关特性。8 年间，TBE 高发病率聚集区先由大兴安岭中北部区域逐步向南延伸，2011 年回归到大兴安岭中北部区域，2012~2013 年高发病率聚集区继续向北扩展到莫尔道嘎镇。

2) 通过地理探测器的因子探测器对 TBE 相关的地理环境因子进行初步探测，所有的因子依照解释力大小被分为三个等级，第一等级的要素包括土壤类型、降雨量和土地覆被类型，第二等级的要素包括降雨频度、温度和光照时数，第三等级的要素包括坡度、坡向、相对湿度、EVI 和 NDVI。进一步利用风险探测器探测不同土壤类型和土地覆被类型内发病风险的空间差异，发现这两类因子有很大的空间相关性，不同土壤和土地覆被类别的形成与植被、地形、气候等要素具有很大相关性，可将土壤和土地覆被作为回归模型中其他要素选择的标准和借鉴。

3) 地理加权回归所得的 TBE 发病风险的空间分布主要表现为中部出现高发病风险的三角带，东西部是低风险地带。中部三角带包括额尔古纳市、根河市、鄂伦春自治旗西部以及牙克石市，三角内部发病风险又表现为自南向北发病风险逐渐升高的格局。三角带内，相对湿度是与高发病风险正相关关系最大的要素，带内大部分区域的 TBE 发病风险与 NDVI、降雨量和坡向呈正相关关系，与坡度呈负相关关系，北部部分区域与 DEM 呈负相关关系。地理加权回归中各因子的回归系数表明在不同空间位置上各因子与发病风险的相关程度不同：相对湿度在呼伦贝尔北部三个县市及最南部部分乡镇的 TBE 发病风险与相对湿度呈正相关，发病风险最高的牙克石市发病风险与 NDVI 呈负相关，与降雨量呈正相关，呼伦贝尔 TBE 发病风险与高程的相关程度自西向东依次由负相关变为正相关，与坡度在空间上的相关程度与高程相反。

本文的创新性和特色主要有两点：1) 通过各类数据的格网转化，结合地理探测方法，揭示不同地理环境因子对 TBE 发病影响关系的强弱程度。初步结果表明地理环境因子对 TBE 风险的影响由强到弱可分三个等级，并定量刻画了第一

等级的土壤、土地覆被和降雨量因子的解释力。2)基于地理加权回归建模,揭示了呼伦贝尔市 TBE 风险发病风险的空间分异特征,发现 TBE 发病高风险的三角区域,并得到各地理环境因素与 TBE 发病风险相关关系的空间分异特征,可为 TBE 的预防和防控提供风险评价模型和数据支持。

## 6.2 展望

基于本文的研究和发现的问题,计划在后续研究中对以下三点继续深入研究:

1) 继续深入探测更多地理环境因子与 TBE 的影响关系,并在今后研究中加入人文要素对 TBE 的影响研究。本研究中,部分地理环境因子(如高程和坡向)在区域统计过程中损失太多信息量,为得到更多、更精细地理环境因子与 TBE 的影响关系,可结合文献综述其他蜱传性疾病以及其他研究区域有关地理环境因子及人文因子的定量研究,对数据进行逻辑分级。例如,Materna 等(2008)发现海拔高度通过影响林种的分布情况进而决定蜱虫密度,并发现中欧卡尔可诺希山地区海拔小于 1000 m 的针阔混交林区最适宜于蜱虫的生长,可以结合该地形特点,可对研究区内地形要素以 1000 m 进行分级,再参与后续研究。

2) 加强 TBE 风险评价建模方法对比研究。TBE 的风险评价应用了多元线性回归和地理加权回归两种方法对发病风险进行回归,这两种方法具有各自的优缺点,例如,多元线性回归等数理方法能够很好地探测出发病风险的时间变化特征,却忽略了地理环境因子在不同空间位置上对疾病风险产生的不同影响;GWR 方法强于探索传染病传播与风险因子的在空间上的联系,却无法在时间上和空间上进行推广。本研究中多元线性回归只被应用于 GWR 回归因子的选择中,在后续研究中应该结合这两个方法各自的优缺点,更好地探测发病风险的空间分布特征。

3) 扩大研究区域和研究的时间序列,提示更大地理单元、更长时间序列上的 TBE 风险分布。本文以呼伦贝尔市 TBE 发病情况为例,研究该地区 TBE 地理环境影响因子,并对该地区的 TBE 空间风险进行评价。在后续研究中,可以结合已得 TBE 发病风险与地理环境因子的相关关系,扩展至整个东北地区更长时间序列的 TBE 发病风险空间分布的研究,以期为 TBE 的防疫防控提供更多方法和结果支撑。

## 参考文献

- 毕维铭, 邓厚培, 卜笑妍. 1997. 森林脑炎自然疫源地地区划研究[J]. 首都师范大学学报:自然科学版, 18(2): 100-107.
- 蔡增林, 鲁志新, 胡玲美, 等. 1996. 东北三省部份地区森林脑炎疫源地流行病学调查[J]. 微生物学杂志, 16(1): 19-22.
- 曹务春, 张习坦, 许荣满. 1999. 蜱类及蜱媒疾病的公共卫生学意义[J]. 中国公共卫生, 15(3): 43-44.
- 曹志冬, 王劲峰, 高一鸽, 等. 2008. 广州SARS流行的空间风险因子与空间相关性特征[J]. 地理学报, 63(9): 981-993.
- 陈述彭, 陈秋晓, 周成虎. 2002. 网格地图与网格计算[J]. 测绘科学, 27(4):1-6.
- 董景红, 朱建华, 尹凤茹. 2007. 内蒙古大兴安岭林区森林脑炎、莱姆病流行病学调查研究[J]. 中国预防医学杂志, 8(6): 718-719.
- 范新生, 应龙根. 2005. 中国SARS疫情的探索性空间数据分析[J]. 地球科学进展, 20(3): 282-291.
- 方立群. 2004. 地理信息系统和遥感技术应用于肾综合征出血热流行病学研究[D]. 北京: 中国人民解放军军事医学科学院.
- 方立群. 2009. 肾综合征出血热时空分布及环境危险因素研究[D]. 北京: 中国人民解放军军事医学科学院.
- 冯军, 吴晓华, 李石柱, 等. 2011. 空间统计分析方法及相关软件在传染病研究中的应用[J]. 中国血吸虫病防治杂志, 23(2): 217-220.
- 国务院人口普查办公室, 国家统计局人口和就业统计司. 2012. 中国2010年人口普查分乡、镇、街道资料[M]. 北京: 中国统计出版社.
- 胡明, 赵春暖. 1991. 灰色模型在传染病预测中的应用[J]. 铁道医学, 19(3): 167-168.
- 胡艺. 2012. 基于空间模型的小概率地理健康事件生态学研究[D]. 北京:中国地质大学(北京).
- 蒋敏, 李晓松, 冯子健, 等. 2008. 四川省HIV/AIDS空间自相关分析[J]. 现代预防医学, 35(22): 4329-4331.
- 李德仁. 2005. 论广义空间信息网格和狭义空间信息网格[J]. 遥感学报, 9(5): 513-520.
- 刘洁, 何钦成. 2013. 我国流行性乙型脑炎发病的聚类分析及其未来发病趋势预测[J]. 中国媒介生物学及控制杂志, 24(2): 152-160.
- 李一凡, 王卷乐, 高孟绪. 2015. 自然疫源性疾病地理环境因子探测及风险预测研究综述[J]. 地理科学进展. 34(7): 926-935.
- 刘育京, 蒋豫图, 郭存三, 等. 1979. 吉林省森林脑炎自然疫源地的某些流行病学特点[J]. 军事医学, (1): 109-120.

- 申思. 2008. 基于文本地址的空间点数据分析研究与应用[D]. 北京:北京大学.
- 孙振宇. 2008. 多元回归分析与Logistic回归分析的应用研究[D]. 南京:南京信息工程大学.
- 唐成艳, 甘文烨, 徐斌, 等. 2011. 广西壮族自治区1989—2006年流行性乙型脑炎时空动态趋势分析[J]. 中华流行病学杂志, 32(3): 274-278.
- 通拉嘎, 徐新良, 付颖, 等. 2014. 地理环境因子对螺情影响的探测分析[J]. 地理科学进展, 33(5): 625-635.
- 万康林, 张哲夫, 窦桂兰, 等. 1998. 全沟硬蜱是我国北方地区莱姆病螺旋体的主要传播媒介[J]. 实用预防医学, 5(6): 3-6.
- 王劲峰, 廖一兰, 刘鑫. 2010. 空间数据分析教程[M]. 北京: 科学出版社.
- 王卷乐, 孙九林. 2015. 格网化资源环境综合科学调查的若干问题思考[J]. 地球信息科学学报, 17(7): 758-764.
- 武继磊. 2004. Lattice数据分析方法、模型及其应用[D]. 北京:中国科学院大学(地理科学与资源研究所).
- 尹德铭, 刘瑞璋. 2000. 对中国东北林区森林脑炎防治研究的回顾[J]. 中华流行病学杂志, 21 (5): 387-389.
- 詹希美. 2002. 人体寄生虫学(第五版)[M]. 北京: 人民卫生出版社.
- 张启恩. 2003. 我国重要自然疫源地与自然疫源性疾病[M]. 沈阳: 辽宁科学技术出版社.
- 张少森, 周水森, 汤林华, 等. 2012. 黄淮流域疟疾传播与土地利用状况变化关系的研究[J]. 中国寄生虫学与寄生虫病杂志, 30(2): 102-108.
- 张思华. 2012. 蜱与蜱传疾病最新研究进展[J]. 安徽预防医学杂志, 18(1): 45-48.
- 张文义. 2010. 基于环境因素预测肾综合征出血热和疟疾传播风险[D]. 北京: 中国人民解放军军事医学科学院.
- Achazi K, Růžek D, Donoso-Mantke O, et al. 2011. Rodents as sentinels for the prevalence of Tick-Borne encephalitis virus[J]. Vector-Borne and Zoonotic Diseases, 11(6): 641-647.
- Anselin L. 1995. Local indicators of spatial association-LISA[J]. Geographical analysis, 27(2): 93-115.
- Aylin P, Maheswaran R, Wakefield J, et al. 1999. A national facility for small area disease mapping and rapid initial assessment of apparent disease clusters around a point source: the UK Small Area Health Statistics Unit[J]. Journal of Public Health Medicine, 21 (3): 289-98.
- Baker D M, Valleron A J. 2014. An open source software for fast grid-based data-mining in spatial epidemiology (FGBASE)[J]. International Journal of Health Geographics, 13 (1): 46.
- Barbour A G, Burgdorfer W, Hayes S F. et al. 1983. Isolation of a Cultivable Spirochete from Ixodes ricinus Ticks of Switzerland[J]. Current Microbiology, (8): 123-126.
- Bayles, B R, Allan B F. 2014. Social-ecological factors determine spatial variation in human incidence of tick-borne ehrlichiosis[J]. Epidemiology and Infection, 142 (9): 1911-1924.

- Beneš V, Bodlák K, Mølleret J, et al. 2011. A case study on point process modelling in disease mapping[J]. *Image Analysis & Stereology*, 24(3): 159-168.
- Brunsdon C, Fotheringham A S, Charlton M E. 1996. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity[J]. *Geographical analysis*, 4(28): 281-298.
- Carvalho A X Y, Albuquerque P H M, Almeida J G R, et al. 2009. Spatial hierarchical clustering[J]. *Revista Brasileira de Biologia*, 27 (3): 411-442.
- Chang K T. 2006. Introduction to Geographic Information System[M]. Boston: McGraw-Hill Higher Education.
- Congalton R G. 1997. Exploring and evaluating the consequences of vector-to-raster and raster-to-vector conversion[J]. *Photogrammetric Engineering and Remote Sensing*, 63(4): 425-434.
- Cumming G S. 1999. The evolutionary ecology of African ticks[D]. Oxford: University of Oxford.
- Curtis A, Leitner M. 2002. Using hierarchical nearest neighbor analysis and animation to investigate the spatial and temporal patterns of raccoon rabies in West Virginia[M]. *Geographic Information Systems & Health Applications*, Hershey PA: Idea Group Publishing.
- Daniel M, Kriz B, Danielova V, et al. 2009. Changes of meteorological factors and tick-borne encephalitis incidence in the Czech Republic[J]. *Epidemiol Mikrobiol Imunol*, 58 (4): 179-187.
- Eisen L. 2007. Seasonal pattern of host-seeking activity by the human-biting adult life stage of *Dermacentor andersoni* (Acari: Ixodidae) [J]. *Journal of medical entomology*, 44(2): 359-366.
- Estrada-Peña A. 2001. Distribution, abundance, and habitat preferences of *Ixodes ricinus* (Acari: Ixodidae) in northern Spain[J]. *Journal of Medical Entomology*, 3(38): 361-370.
- Fotheringham A S, Brunsdon C, Charlton M. 2003. Geographically weighted regression: the analysis of spatially varying relationships[M]. John Wiley & Sons.
- Geller J, Nazarova L, Katargina O, et al. 2013. Tick-Borne pathogens in ticks feeding on migratory passerines in western part of Estonia[J]. *Vector-Borne and Zoonotic Diseases*, 13(7): 443-448.
- Glass G E, Schwartz B S, Morgan J M, et al. 1995. Environmental risk factors for Lyme disease identified with geographic information systems[J]. *American Journal of Public Health*, 85(7): 944-8.
- Gray J S, Dautel H, Estrada-Peña A, et al. 2009. Effects of climate change on ticks and Tick-Borne diseases in Europe[J]. *Interdisciplinary Perspectives on Infectious Diseases*, 2009: 1-12.
- Gray J S. 1998. The ecology of ticks transmitting Lyme borreliosis[J]. *Experimental and Applied Acarology*, 22(5): 249-258.
- Guerra M, Walker E, Jones E, et al. 2002. Predicting the risk of Lyme disease: habitat suitability for *Ixodes scapularis* in the north central United States[J]. *Emerging Infectious Disease*, 8(3): 289-97.
- Hay S I, Packer M J, Rogers D J. 1997. The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease[J]. *International Journal of Remote Sensing*,

- 18(14): 2899-2930.
- Hoogstraal H. 1979. The epidemiology of tick-borne Crimean-Congo hemorrhagic fever in Asia, Europe, and Africa[J]. *Journal of medical entomology*, 4(15): 307-417.
- Hu W, Clements A, Williams G, et al. 2011a. Spatial analysis of notified dengue fever infections[J]. *Epidemiology and Infection*, 139 (3): 391-399.
- Hu Y, Wang J F, Li X H, et al. 2011b. Geographical Detector-Based Risk Assessment of the Under-Five Mortality in the 2008 Wenchuan Earthquake, China[J]. *Plos One*, 6(6): e21427.
- Jackson L E. 2005. The relationship of land-cover pattern to Lyme disease[D]. Chapel Hill: University of North Carolina.
- Jacobsen T. 2014. Topographical variables influencing *Dermacentor andersoni* distribution in Montana[D]. Helena: Carroll College.
- Kabacoff R I. 2011. R in action: data analysis and graphics with R[M]. Manning Publications Co.
- Kahl O, Knülle, W. 1988. Wirtssuchaktivität der Schildzecke, *Ixodes ricinus* (Acari, Ixodidae) und ihre Durchseuchung mit Lyme-Spirochäten und dem Frühsommer-Meningoenzephalitis (FSME)-Virus in Berlin (West) [J]. *Mitt. Dtsch. Ges. Allg. Angew. Entomol*, (6): 223-225.
- Kim S, Yun S, Han M G, et al. 2008. Isolation of Tick-Borne encephalitis viruses from wild rodents, south korea[J]. *Vector-Borne and Zoonotic Diseases*, 8(1): 7-14.
- Kitron U, Kazmierczak J J. 1997. Spatial analysis of the distribution of Lyme disease in Wisconsin[J]. *American Journal of Epidemiology*, 145(6): 558-66.
- Knap N, Korva M, Dolinšek V, et al. 2012. Patterns of Tick-Borne encephalitis virus infection in rodents in Slovenia[J]. *Vector-Borne and Zoonotic Diseases*, 12(3): 236-242.
- Kulldorff M, Feuer E J, Miller B A, et al. 1997. Breast cancer clusters in the northeast United States: a geographic analysis[J]. *American journal of epidemiology*, 146(2): 161-70.
- Li X Z, Wang J F, Yang W Z, et al. 2012. A spatial scan statistic for nonisotropic two-level risk cluster[J]. *Statistics in Medicine*, 31(2): 177-187.
- Li Y F, Wang J L, Gao M X. 2015. Geographical detection of Tick-borne Encephalitis risks in Hulunbuir, China[C]. *Geoinformatics 2015*. 1-7
- Lindgren E. 1998. Climate and Tickborne Encephalitis[J]. *Ecology and Society*, 2 (1): 5.
- Lukan M, Bullova E, Petko B. 2010. Climate warming and Tick-borne Encephalitis, Slovakia[J]. *Emerging Infectious Diseases*, 16(3): 524-526.
- Lu Z, BRÖKER M, Liang G D. 2008. Tick-Borne Encephalitis in Mainland China[J]. *Vector-Borne and Zoonotic Diseases*, 8(5): 713-720.
- Materna J, Daniel M, Metelka L, et al. 2008. The vertical distribution, density and the development of



- the tick *Ixodes ricinus* in mountain areas influenced by climate changes (The Krkonoše Mts., Czech Republic)[J]. *International Journal of Medical Microbiology*, 298(S1): 25-37.
- Meade M S, Florin J W, Gesler W M. 1988. *Medical geography*[M]. Guilford Press.
- Merler S, Furlanello C, Chemini C, et al. 1996. Classification tree methods for analysis of mesoscale distribution of *Ixodes ricinus* (Acari: Ixodidae) in Trentino, Italian Alps[J]. *Journal of Medical Entomology*, 33(6): 888-893.
- Ord J K, Getis A. 1995. Local spatial autocorrelation statistics: Distributional issues and an application[J]. *Geographical analysis*, 27(4): 286-306.
- Peng B, Cameron A S, Zhang Y, et al. 2008. Weather and notified *Campylobacter* infections in temperate and sub-tropical regions of Australia: An ecological study[J]. *Journal of Infection*, 57 (4): 317-323.
- Perret J L, Guerin P M, Diehl P A, et al. 2003. Darkness induces mobility, and saturation deficit limits questing duration, in the tick *Ixodes ricinus*[J]. *Journal of Experimental Biology*, 206(11): 1809-1815.
- Randolph S E, Green R M, Peacey M F, et al. 2000a. Seasonal synchrony: the key to tick-borne encephalitis foci identified by satellite data[J]. *Parasitology*, 121(1): 15-23.
- Randolph S E, Rogers D J, 2000b. Fragile transmission cycles of tick-borne encephalitis virus may be disrupted by predicted climate change[J]. *Proceedings of the Royal Society B: Biological Sciences*, 267(1454): 1741-1744.
- Randolph S E. 2000. Ticks and tick-borne disease systems in space and from space[J]. *Advances in Parasitology*, 47: 217-243.
- Randolph S E. 2004. Tick ecology: processes and patterns behind the epidemiological risk posed by ixodid ticks as vectors[J]. *Parasitology*, 129 (7): S37-S65.
- Randolph S E. 2008. Tick-borne encephalitis incidence in Central and Eastern Europe: Consequences of political transition[J]. *Microbes and Infection*, 10(3): 209-216.
- Rogers D J, Packer M J. 1993. Vector-borne diseases, model, and global change[J]. *Lancet*, 8882(342): 1282-1284.
- Rogers D J, Randolph S E. 2003. Studying the global distribution of infectious diseases using GIS and RS[J]. *Nat Rev Microbiol*, 1(3): 231-237.
- Ruiz-Fons F, Fernandez-de-Mera I G, Acevedo P, et al. 2012. Factors driving the abundance of *Ixodes ricinus* ticks and the prevalence of zoonotic *I. ricinus*-borne pathogens in natural foci[J]. *Applied and Environmental Microbiology*, 78(8): 2669-2676.
- Schwarz A, Maier W A, Kistemann T, et al. 2009. Analysis of the distribution of the tick *Ixodes ricinus* L. (Acari: Ixodidae) in a nature reserve of western Germany using Geographic Information Systems[J]. *International Journal of Hygiene and Environmental Health*, 212 (1): 87-96.
- Shi X, Miller S, Mwenda K, et al. 2013. Mapping disease at an approximated individual level using

- aggregate data: a case study of mapping New Hampshire birth defects[J]. *International Journal of Environmental Research and Public Health*, 10(9): 4161-4174.
- Šumilo D, Asokliene L, Avsic-Zupanc T, et al. 2008. Behavioural responses to perceived risk of tick-borne encephalitis: Vaccination and avoidance in the Baltics and Slovenia[J]. *Vaccine*, 26(21): 2580-2588.
- Süss J, Klaus C, Gerstengarbe F W, et al. 2008a. What Makes Ticks Tick? Climate Change, Ticks, and Tick-Borne Diseases[J]. *Journal of Travel Medicine*, 15 (1): 39-45.
- Süss J. 2008b. Tick-borne encephalitis in Europe and beyond the epidemiological situation as of 2007[J]. *Euro Surveillance*, 13 (4-6):1-8.
- Takashima I, Morita K, Chiba M, et al. 1997. A case of tick-borne encephalitis in Japan and isolation of the virus[J]. *Journal of clinical microbiology*, 35(8): 1943-1947.
- Takeda T, Osada I T. 1999. Isolation of tick-borne encephalitis virus from wild rodents and a seroepizootologic survey in hokkaido, Japan[J]. *The American journal of tropical medicine and hygiene*, 2(60): 287-291.
- Toomer, J. 1997. Predicting pan-African Tick distributions using remotely sensed surrogates of meteorological and environmental conditions [R]. Undergraduate project for FHS Biological Sciences, Oxford.
- Vanwambeke S O, Lambin E F, Eichhorn M P, et al. 2007. Impact of land-use change on Dengue and Malaria in Northern Thailand[J]. *EcoHealth*, 4(1): 37-51.
- Walder G, Lkhamsuren E, Shagdar A, et al. 2006. Serological evidence for tick-borne encephalitis, borreliosis, and human granulocytic anaplasmosis in Mongolia[J]. *International Journal of Medical Microbiology*, 296: 69-75.
- Wang J F, Li X H, Christakos G, et al. 2010. 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*, 24(1): 107-127.
- Wang J F, Li X H, Christakos G, et al. 2010. 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*, 24(1): 107-127.
- Wang J F, McMichael A J, Meng B, et al. 2006. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area[J]. *Bulletin of the World Health Organization*, 84(12): 965 - 968.
- Wimberly M C, Yabsley M J, Baer A D, et al. 2008. Spatial heterogeneity of climate and land-cover constraints on distributions of tick-borne pathogens[J]. *Global Ecology and Biogeography*, 17(2): 189-202.

- Wu R, Zhang J, Bao Y, et al. 2016. Geographical Detector Model for Influencing Factors of Industrial Sector Carbon Dioxide Emissions in Inner Mongolia, China[J]. Sustainability, 8(2): 149-161.
- Zakovska A, Nejezchlebova H, Bartonkova N, et al. 2013. Activity of the tick *Ixodes ricinus* monitored in a suburban park in Brno, Czech Republic, in association with the evaluation of selected repellents[J]. Journal of Vector Ecology, 38(2): 295-300.
- Zhang S, Zhao J. 2015. Spatio-temporal epidemiology of hand, foot and mouth disease in Liaocheng City, North China[J]. Experimental and Therapeutic Medicine, (9):811-816.
- Zhang W, Wang L, Liu Y, et al. 2014. Spatiotemporal Transmission Dynamics of Hemorrhagic Fever with Renal Syndrome in China, 2005–2012[J]. PLoS Neglected Tropical Diseases, 8(11): e3344.



## 作者简介

### 一、个人简介

李一凡，女，1992年3月出生，山东潍坊人。

2013年6月毕业于中国矿业大学（北京），测绘工程专业，获工学学士学位。

2013年9月进入中国科学院地理科学与资源研究所就读硕士研究生。

### 二、论文成果

- [1] **李一凡**, 王卷乐, 高孟绪. 2015. 自然疫源性疾病的地理环境因子探测及风险预测研究综述[J]. 地理科学进展. 34(7): 926-935.
- [2] **Yifan Li**, Juanle Wang, Mengxu Gao. 2015. Geographical detection of Tick-borne Encephalitis risks in Hulunbuir, China[C]. Geoinformatics 2015, 1-7.(EI收录)
- [3] **李一凡**, 王卷乐, 祝俊祥. 基于地理分区的蒙古国景观格局分析[J].干旱区地理. (2016年待刊, 已录用)
- [4] 田静, 王卷乐, **李一凡**, 周玉洁, 郭海会, 祝俊祥. 2014. 基于决策树方法的蒙古高原土地覆盖遥感分类——以蒙古国中央省为例[J]. 地球信息科学学报. 16(3): 460-469.
- [5] 王卷乐, 高孟绪, **李一凡**. 2015. 地理大数据服务于智慧医疗与健康的思考[J]. 中国卫生信息管理. 12(6):634-637.
- [6] Juanle Wang, Yujie Zhou, Lijun Zhu, Mengxu Gao, **Yifan Li**. 2014. Cultivated Land Information Extraction and Gradient Analysis for a North-South Transect in Northeast Asia between 2000 and 2010[J]. Remote Sensing. 6: 11708-11730.
- [7] Juanle Wang, Hongsheng Li, Eryang Chen, **Yifan Li**, et al. 2014. Challenges Facing Chinese Map Libraries and Librarians: From Paper to Digital Worlds and Services[J]. Journal of Map and Geography Information. 10: 266-287.
- [8] 高孟绪, 王卷乐, 曹春香, 李群, **李一凡**. 2015. 基于地理信息系统和生态位模型的青海省喜马拉雅旱獭空间分布预测[J]. 中华地方病学杂志. 34(5):318-321.
- [9] Juanle Wang, Yongjie Zhang, Fei Yang, Xiaoming Cao, Zhongqiang Bai, Junxiang Zhu, Eryang Chen, **Yifan Li**, Yingying Ran. 2014. Spatial and temporal variations of chlorophyll-a concentration from 2009 to 2012 in Poyang Lake, China[J]. Environmental Earth Science, 1-13.

### 三、参与项目

1. 科技基础性工作专项(2013FY114600), 中国自然疫源性疾病流行病学图集编研

2. 科技基础性工作重点项目(2011FY110400), 格网化资源环境综合科学调查规范;
3. 中国科学院重点部署项目子课题(KZZD-EW-08-02-03), 蒙古资源环境关键问题及科学数据库建设