

基于遥感影像以及宏病毒组学的病毒生态分布识别方法构建与应用

选题背景和意义

- 背景
- 意义

在新冠全球大流行的背景下，我国开始重视野生动物病原监测工作，病原监测的基本方法就是通过抽样调查的方法来估计病原在野外的分布情况，从而防患未然。但由于实际情况中无法对所有的野生动物进行普查，所以现在迫切需要一种根据有限样本来估计采样点附近尽可能大的范围内病原分布情况的方法

本课题结合了宏病毒组学以及遥感卫星影像等技术，提供了一种针对识别病毒生态分布的非参数估计方法，填补了病原监测从抽样调查到分布估计的鸿沟

国内外研究现状

- 病毒检测技术升级
- 生态环境决定病毒分布
- 遥感影像能监测决定病毒分布的生态特征
- 利用遥感影像监测病原分布的案例

PCR、血清学到宏病毒组

生态环境决定宿主分布（生态位理论）

宿主分布决定病毒分布，目前大部分宿主动物分布已知（IUCN）

利用生态特征可以对病毒分布进行建模（Olival 2017）

描述遥感影像的基本原理和GEE中要用到的数据集

遥感影像协助预测疟疾高发地区

研究内容

- 病毒生态分布识别方法描述
- 技术路线

基本假设：生态环境决定病毒分布，所以生态环境条件相似的区域，病毒组成也相似；

输入：采集到的样本对应的GPS、病毒检出矩阵；输出：采样点附近n米范围内每个GPS点对应的病毒矩阵的估计；

子问题拆解：1. 病毒分布是由哪些生态因素决定的？2. 如何建立生态因素和病毒分布之间的关系模型？3. 如何评估估计方法的可靠性？4. 如何在实际病原监测工作中应用？

文献调研，回答哪些生态因素会对病毒分布产生影响

利用遥感卫星影像获取对应的生态特征数据

根据生态环境条件相似的区域，病毒组成也相似的基本假设制作模拟样本集合，并从中随机抽样，搭建生态因素和病毒分布之间的关系模型，通过交叉验证评估模型可靠性

用东非蝙蝠、中国蚊虫监测抽样调查结果估计东非、中国的病毒分布情况

数据收集

- 生态数据收集
- 病毒传播方式分类
- 模拟病毒分布及模拟采样

哪些因素可能会如何影响到病毒传播，需要重点考虑

这些因素如何从公共数据库中获取，缓冲距离如何设置

可以结合东非蝙蝠或中国蚊虫实际生态特征和病毒矩阵的聚类结果验证

气溶胶、粪口、接触、垂直等

根据生态因素的前三个主成分以及病毒的传播方式建立函数，加入随机扰动

建模评估

- 数据预处理方法
- 模型结构
- 模型性能评估

如何将各种类型的生态数据转为向量，分类型、0-1型、数值型

模型有几层？每一层有几个神经元？是否采用drop out？激励函数如何设置？

交叉验证

GEE可视化预测结果，对比与实际模拟分布的差异

案例研究

- 东非蝙蝠
- 国内蚊虫

结果讨论

- 应用场景和意义
- 局限性