

The View from Above: Applications of Satellite Data in Economics

Donaldson D. & Storeygard A.

Journal of Economic Perspectives, 2016

1 Introduction

这篇论文主要介绍了如何在经济学研究中使用遥感数据，以及这种新型数据源在未来的研究潜力。遥感数据主要有三个优势：

- 能得到其他方法难以获取的信息
遥感数据可以低成本地重复收集之前难以获得的面板数据，比如灯光、降雨降雪、风速、洪水、设施分布、植被覆盖、作物选择、农业产出、乡村发展、建筑类型、道路、污染、水产质量等数据。
 - Burgess et al.(2012) 研究过印度尼西亚的森林退化，印度尼西亚限制森林砍伐，但地方官员却常遭到贿赂以放松管制。因此官方数据并不可靠，而 Burgess et al.(2012) 使用卫星对森林的观测数据作为更客观的数据来源，发现省内增加一个监管区域，会使伐木率增加 8.2%。随着数据精度和算法的进步，未来也许会出现全球范围的研究。
 - Jayachandran(2009) 通过 1997 年印度尼西亚森林大火每日卫星观测数据，估计森林大火的颗粒物污染造成了 16400 儿童及胎儿死亡。这种问题在以前难以研究，然而现在遥感数据还能提供甲烷、温室气体等数据。
- 空间分辨率超乎寻常
以前的遥感数据分辨率只有 $30\text{m} \times 30\text{m}$ ，土地使用、建筑类型、农作物类型等信息只能靠估算。而 1999 年后，出现了精确到 1m 的数据，且 2014 年后，美国允许私人公司出售精确到 0.5m 以下的遥感数据。因此也衍生出了很多针对高分辨率数据的研究方法 (Turner, Haughwout, and van der Klaauw 2014)
 - Marx, Stoker, and Suri(2015) 利用 0.5m 精度图像中铁皮屋顶的反射率差异（新屋顶反射率高，旧屋顶由于生锈反射率低，且观测的内罗毕地区都喜欢用铁皮屋顶），估计当地居民房屋投资，进而研究统治者种族偏好对居民生活条件影响。
 - 所谓的高精度也是相对的，利用精度为 1km 的夜间照明数据就能测算更精细的居住分布和财富分布情况 (Elvidge et al. 1997; Henderson, Storeygard, and Weil, 2012)。而 Bleakley and Lin (2012) 通过照明数据证实了美国地理上的水运分界线是现在的大城市的起源。
- 覆盖范围广
卫星数据的收集不受政治动荡和自然灾害的影响，且不同地区数据口径一直，而且有的卫星会定期（每周甚至每天）对同一地区扫描。
 - Costinot, Donaldson, and Smith(2016) 利用全球卫星数据，研究了气候变化对农业经济的影响
 - 所谓的高精度也是相对的，利用精度为 1km 的夜间照明数据就能测算更精细的居住分布和财富分布情况 (Elvidge et al. 1997; Henderson, Storeygard, and Weil, 2012)。而 Bleakley and Lin(2012) 通过照明数据证实了美国地理上的水运分界线是现在的大城市的起源。

2 A Primer on Remote Sensing for Economics

- 卫星轨道
卫星轨道分为地球同步静止轨道 (Geostationary) 和太阳同步轨道 (Sun-synchronous)。地球同步静止轨道卫星的优点是可以对同一地区观测到所有时段的数据, 但缺点是由于卫星距离较远 (36000km), 图像分辨率较低。太阳同步轨道卫星的优点是距离较近 (6000km), 图像分辨率高, 但存在分辨率与采集频率的取舍, 对同一地区的采集频率越高, 分辨率越低。每周更新的图像精度为 1m, 而每天更新的精度为 1km。
- 传感器和光谱
一个卫星有一个到多个传感器, 每个传感器能提供一到多种数据。不同的光谱数据有不同的用途, 比如植物生长的不同阶段, 其光谱不同, 因此可以利用这种变化监控植物生长 (Normalized Difference Vegetation Index (NDVI))。红外线数据也可以用来观测温度和云层 (Novella and Thiaw, 2012)。而 LiDAR 可以测量建筑物高度。因此, 根据不同研究需求, 要选择合适的传感器和光谱数据。
- 影像处理
社会科学研究中从不使用卫星的原始数据, 他们使用的数据都是经过矫正处理的。比如:
 - 正射校正: 卫星观测地球的相邻地区存在角度差异, 需要矫正原始图像经纬度。
 - 多重叠加: 将同一地区的多张图像叠加, 减少随机误差, 比如去除云雾等等。
- 影像解析
除了单个像素点包含的信息, 像素点的组合也是关注的重点 (比如捕捉汽车、建筑物、农作物等)。这些图像的解析依靠非监督聚类方法和有监督的机器学习方法 (Vogelmann et al., 2001; Burchfield et al. 2006)。而监督学习需要大量人工标注好的数据。
- 主要数据来源
卫星传感器信息数据库可以在世界气象组织 OSCAR 数据库找到 (<http://www.wmo-sat.info/oscar/>), 主要数据库原文有列举, 不再赘述。

3 Applications of Remote Sensing Data in Economics

主要研究主题如下:

- 夜间灯光
用夜间灯光数据估计经济活动, 与传统 GDP 指标比较
- 气候
研究短期天气变化和长期气候变化对人类活动的影响, 比如温度升高引起的干旱怎样导致非洲冲突的可能性上升 (Harari and La Ferrara, 2015), 气候变化对儿童死亡率的影响 (Kudamatsu, Persson and Stromberg, 2016)
- 农业用地和作物选择
农业政策如何影响农业用地规划, 估计某一地区的农业产出
- 城市用地
监测城市土地 (居民区、商业区、工业区和交通网络) 和非城市土地 (水域、岩石、沙地、森林、草地、农田、湿地) 的用途, 捕捉城市发展过程。另一方面也可以估计土地政策的实际影响。
- 建筑类型
之前提到的用铁皮屋顶反射率衡量居民房屋建设, 另外也可以估计建筑体量, 考察建筑用地。

- 自然资源
根据植被覆盖，考察林业发展或者非法伐木的经济影响。还有研究旅游资源（沙滩质量）对经济发展的影响。
- 污染检测
测算森林大火的污染影响、空气污染对婴儿死亡率的影响。以及讨论污染的产生原因（政策因素、国际贸易因素等）。
- 多种数据来源整合利用
根据多种卫星数据来源估算家庭收入、贫困率等指标。

4 Potential Pitfalls in the Use of Remote Sensing Data

- 数据体量
卫星数据体量巨大，需要针对性的处理方法和模型。
- 空间相关性
相邻的数据点存在相关性，不加处理可能会产生有偏的结果。
- 测量误差
卫星传感器数据不一定准确，比如对同一地区在不同时间观测的角度、气象条件都有波动，会造成数据的误差；还有一些由于测量原理造成的误差，比如森林繁密的地方地表高度数据更可能不准确。因此，知道数据产生的原理和所依据的假设很重要。
- 隐私问题
高分辨率影像数据将会在一定程度上侵犯个人隐私，研究者需要予以重视。

5 Further Reading

- 原文: <http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.30.4.171>
- 其余资源在原文 pdf 中均可直接点击浏览