

实验实训方案

案例 02 · 降水异常与气象干旱的时空聚集分析

案例编号	难度	学时	交付格式
02	基础—进阶	6	可编辑 DOCX

使用说明：本文件为实验实训方案的可编辑 Word 版本；Notebook、数据说明和代码仍在平台对应案例页提供下载。

实训问题

三个月累计降水形成的 SPI-3 如何识别干旱频率？末年 SPI-3 的局部高低值聚集哪些通过显著性检验？

实训目标

完成数据状态核对、Notebook 操作、案例核心算法、图件检查和结果边界说明；能将结论逐项对应到代码输出与数据证据。

数据与前置条件

- **数据/产品**：CHIRPS v2.0
- **官方入口**：<https://data.chc.ucsb.edu/products/CHIRPS-2.0/>
- **产品说明**：<https://www.chc.ucsb.edu/data/chirps>
- **当前随包运行模式**：上面列出的是真实产品入口；Notebook 只使用固定随机种子的合成教学输入，不会自动下载或读取这些真实产品。合成输出只用于演示算法和图件链路，不是实测生态环境结论。
- **真实数据准备**：先阅读本案 [data/README.md](#)、[data/字段字典.md](#)、[references/DATASET_MANIFEST.csv](#) 和官方产品文档；按版本核对字段/波段、时间、单位、CRS/transform、质量标记、NoData、AOI 与许可。实际产品输入尚未接入 Notebook 时，不得声称已完成真实产品分析。

本案核心方法

三个月累计降水与 SPI-3、Gamma 分布拟合与零降水概率、短样本回退方法标记、 $SPI \leq -1$ 干旱频率、Global Moran I 与 Local Moran 置换检验、BH-FDR 校正

核心 Python 接口：[src/core_model.py](#) 中的 `spi3`, `drought_frequency`, `moran_i`, `lisa_quadrants`, `lisa_statistics`。学生不需要在命令行手动调用函数；按学生指导书运行 Notebook 即可。若图表或文件与方案写法不一致，请先暂停并反馈教师，不要自行猜测或伪造结果。

实训流程

第一次先打开本案例的 Notebook，点一次‘重启内核并全部运行（Restart Kernel → Run All）’，等所有代码格跑完。下面五步是按顺序检查这一次的运行结果，不用每一步都重新运行。Notebook 每个代码格前都有标题，照着本表看对应结果即可，不需要在命令行里手动输入函数名。以后改了参数，再找到对应标题的代码格，从那里运行到 Notebook 结束。

步骤	学生操作	本步产物	检查点
1. 确认月份和数据模式	Notebook 一次性运行完成后，先看输入摘要。随包输入是 10 年合成月降水网格，不是 CHIRPS 实测数据；月份应连续，网格大小和缺测值会在摘要中显示。CHIRPS 是可下载的真实数据入口，但 Notebook 不会自动联网下载。	Notebook 输入摘要；课后手工填写的数据模式、月份数、网格形状和单位	月份从起始月连续排列；记录网格形状、单位和 NoData 约定，并注明当前是合成数据。
2. 计算 SPI-3	SPI-3 是先把连续 3 个月的降水量加起来，再换成标准化的干湿程度。程序按日历月份拟合 Gamma 分布，并处理零降水；样本太少时会按代码说明退回简化 Z 分数。这个教学序列只有 10 年，每个月份大约只有 10 个样本，时间太短，不能当成规范的长基准期 SPI。打开结果表的 <code>spi_method</code> 看本次采用的方法；表里没有每个网格的备用方法计数。	<code>outputs/student_run/summary.csv</code> 的 <code>spi_method</code> 字段； <code>outputs/student_run/figures/01_spatiotemporal_overview.png</code>	知道 SPI-3 用连续 3 个月累计降水；记录表中实际方法说明；不把备用方法结果称为标准长基准期 SPI。
3. 计算干旱频率和 Global Moran I	Notebook 一次性运行完成后，查看 <code>summary.csv</code> 和图 02。程序把 SPI 小于或等于 -1.0 ($SPI \leq -1$) 的有效月份比例作为干旱频率，再用网格相邻权重计算 Global Moran I。它回答的是“高低值是否倾向于聚在一起”，不是趋势检验。	<code>outputs/student_run/summary.csv</code> 的 <code>spi_threshold</code> 、 <code>mean_drought_frequency</code> 、 <code>global_moran_I_drought_frequency</code> ； <code>outputs/student_run/figures/02_trend_significance.png</code>	默认阈值为 $SPI \leq -1.0$ ；干旱频率只用有效月份作分母；能用一句话解释 Moran I 是空间聚集指标。
4. 查看 Local Moran 结果	打开局部诊断表和图 03。这里用的是最后一年（10 年序列的第 10 年）12 个月平均的 SPI-3 场，不是全 10 年的干旱频率图。HH 表示 SPI 偏高且周边也偏高（相对偏湿），LL 表示 SPI 偏低且周边也偏低（相对偏干），HL/LH 表示高低值相邻；只有通过置换检验并经 BH-FDR 多重检验校正后 $q \leq 0.05$ 的单元才标为显著。	<code>outputs/student_run/lisa_cell_diagnostics.csv</code> （其中 <code>analysis_field</code> 标明分析的是末年平均 SPI-3）； <code>outputs/student_run/figures/03_spatial_diagnostic.png</code>	表中能找到 <code>analysis_field</code> 、 <code>quadrant</code> 、 <code>p_value</code> 、 <code>q_value_BH</code> ；当前使用 4999 次置换、随机种子 102、q 阈值 0.05；象限标签本身不代表显著。

步骤	学生操作	本步产物	检查点
5. 只做默认阈值的敏感性比较	查看图 04, 比较 SPI 阈值 -0.5、-1、-1.5、-2 对总体干旱频率的影响。当前代码没有对每个阈值重新计算 LISA 显著格数; 多阈值 LISA 是选做扩展。	outputs/student_run/figures/ 04_sensitivity.png; outputs/student_run/summary.csv	报告写清这是总体频率敏感性, LISA 只对应默认阈值和末年场, 并说明真实研究需使用足够长的基准期。

学生交付物

运行 Notebook 后, 提交以下实际生成的文件:

- outputs/student_run/figures/01_spatiotemporal_overview.png
- outputs/student_run/figures/02_trend_significance.png
- outputs/student_run/figures/03_spatial_diagnostic.png
- outputs/student_run/figures/04_sensitivity.png
- outputs/student_run/lisa_cell_diagnostics.csv
- outputs/student_run/summary.csv

另外, 请自己新建 outputs/student_run/结果说明.md, 用一页写清输入是什么、关键参数、图表发现、敏感性结果和结论限制。Notebook 不会自动生成这份说明。

故障处理

- **Notebook 找不到输入:** 先按本案运行手册和 Notebook 相对路径启动; 不要使用作者电脑绝对路径。
- **图件空白或出现异常值:** 检查维度、有效掩膜、NaN/NoData、单位和阈值; 先停止解释, 再修复输入。
- **真实产品与示例输入不匹配:** 核对官方字段字典和产品版本; 完成显式映射、单位/CRS 检查后再运行, 不能直接把真实栅格塞入模拟数组流程。
- **显著性或趋势过度解释:** 记录空间/时间依赖、权重或检验条件; 图上分组、象限或 Z 分数不自动代表统计显著。
- **结果无法复现:** 记录依赖版本、随机种子、数据版本/哈希、处理步骤和输出文件名。

评分量规 (100 分)

维度	分值	可核验证据
数据来源、状态与元数据	20	官方入口、版本/模拟状态、空间时间支持、CRS/单位/NoData 记录

维度	分值	可核验证据
数据质检与预处理	20	输入形状、有效样本、掩膜和数据限制说明
核心算法与参数解释	25	spi3, drought_frequency, moran_i, lisa_quadrants, lisa_statistics 对应代码、参数表和结果表
GIS 图件、验证与敏感性	20	图件追溯表、图例/单位/范围、至少一项敏感性分析
可复现性与解释边界	15	环境/种子/来源记录、合成与真实数据区分、限制说明
合计	100	

论文与解释边界

核心论文和公开全文入口见 [references/论文与案例对应.md](#)。论文只支撑方法或数据产品背景，不代表本案例复现论文的研究区、数据版本或数值结果。