

实验实训方案

案例 03 · ERA5-Land 热浪识别与空间分布

案例编号	难度	学时	交付格式
03	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

在合成日最高温序列中，热浪日数和最长连续热浪如何按网格分布？改变阈值后结果是否稳定？

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**：ERA5-Land
- **官方入口**：<https://cds.climate.copernicus.eu/>
- **产品说明**：<https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land>
- **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

日最高温合成序列、前 365 天固定百分位阈值、连续至少 3 天热浪识别、最后 365 天热浪日数与最长持续期、85/90/95 百分位敏感性

核心 Python 接口：[src/core_model.py](#) 中的 `percentile_threshold`, `heatwave_mask`, `longest_run`。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. 确认温度时间轴和单位	Notebook 一次性运行完成后，查看输入摘要。当前输入是 5 年合成日最高温度；先看时间长度、网格形状和温度单位。真实 ERA5-Land 数据接入时还要检查 UTC 时间、K 转 °C、网格和缺测。	Notebook 输入摘要	记录日数、网格形状和单位，并明确合成状态。
2. 用前 365 天计算固定 90 分位阈值	查看阈值单元的运行结果。程序只用前 365 天作为基准，为每个网格算一个固定的 90 分位温度阈值；它不是按每个日历日滑动计算的阈值。	<code>outputs/student_run/summary.csv</code> 的 <code>mean_threshold_C</code> ; <code>outputs/student_run/figures/01_spatiotemporal_overview.png</code>	能指出基准期是前 365 天、百分位是 90，并在图 01 中看到阈值线。
3. 识别连续热浪	查看热浪单元的运行结果。温度超过各网格阈值且连续至少 3 天才算一次热浪；缺测日不能被拼成连续事件。	<code>outputs/student_run/summary.csv</code> 的 <code>minimum_duration_days</code> 、 <code>max_run_days</code> ; <code>outputs/student_run/figures/03_spatial_diagnostic.png</code>	能解释‘热日’和‘连续热浪’的区别，并确认最短持续日数为 3 天。
4. 看最后 365 天的热浪日数	打开图 02。图中是最后 365 天每个网格的热浪日数空间聚合，不是跨多年趋势显著性图。	<code>outputs/student_run/figures/02_trend_significance.png</code>	图题、颜色条和统计时段一致；不把这张图写成长期趋势或健康风险结论。
5. 比较百分位阈值并写边界	打开图 04，比较 85、90、95 分位阈值下的标记比例。报告中写明再分析温度与站点观测的区别，合成结果不能直接用于真实热风险判断。	<code>outputs/student_run/figures/04_sensitivity.png</code> ；一页结果说明	至少记录一次阈值比较，并同时写出基准期、最短持续日数和数据状态。

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/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
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	<code>percentile_threshold</code> , <code>heatwave_mask</code> , <code>longest_run</code> 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。