

Sentinella v3 项目完善与实验结果

01 改进后的项目与实验结果

Sentinella 面向菲律宾奎松市的卫生与社区巡查人员，连接城市病例趋势、洪涝背景、持续积水观测与可核查的行动。项目已实现未来目标建模、数据追溯、时间检查、历史回放、GFD事件条件模型探索与传感器观测校验。实体设备和当地试点仍属于待验证工作。

健康模块从“同周气象解释同周病例”改为预测未来1、2、4周后开始的七天病例总数。输入增加病例历史与季节信息，按季度滚动训练；模型、逐次预测和参数选择记录均可复现。下表是2018 – 2022年历史评估，病例与气象均假设延迟7天可用。

提前量	次数	延续基线MAE	病例模型MAE	误差下降	加气象MAE
1周	244	37.10	35.24	5.0%	36.87
2周	238	45.04	38.62	14.3%	42.32
4周	226	58.75	40.66	30.8%	49.52

MAE单位为病例数，越低越好。病例模型为三个固定种子预测的平均；1周约下降5.0%，2周14.3%，4周30.8%。1、2周误差差值区间跨零，不能写成稳定显著提升。气象版本的三个提前量均未优于病例版本，不宣称气象或积水增益已获证明。

所有结果都属于单城市、既有历史数据的探索性回放；评估年份此前已经查看过，不能称全新盲测。历史数据截至2022年，未接入2026年监测，不能作为当前疫情或街区警报。

02 数据来源与日期处理

病例数据

源表为用户提供的奎松市MW/n城市周数据，572条记录，2012 – 2022年；572个病例数与原建模表逐条一致。来源说明指向Eligado等2026年PLOS论文，DOI 10.1371/journal.pntd.0014444。病例是城市聚合值，不能推断个人感染或精确街区病例。

两处五位数年份22012、22013有明确修改记录，原文件不覆盖。其他日期全部保留；22处相邻日期并非相隔7天，10对区间重叠，共20条重叠记录被隔离。没有插值病例，没有擅自把原始周重新编号成官方流行病学周。

气象来源已独立重建

从NASA POWER Daily API下载4079天数据，覆盖2011-11-01至2022-12-31；查询位置14.676 ° N、121.0437 ° E，社区AG，时间标准LST。原始响应、请求参数、产品元数据与哈希已保存。按旧表区间聚合后，六个原气象字段均为572/572条在1e-6精度内吻合。

旧temp_max/min实际是周内“每日平均气温”的最大/最小值。v3以dailymean_max/min准确命名，同时另列真正T2M_MAX/MIN的日最高/最低气温周均值。LST是NASA地方太阳时；与菲律宾民用日界的对应仍需确认。粗网格数据不能当作街道温度实测。

GFD 与 NAMRIA 补充资料

GFD的10个菲律宾事件原始包对应2012 – 2017年；两个研究框共20个GeoTIFF裁剪，保留有效像元、非永久洪水像元及来源哈希。GFD原始许可为CC BY-NC 4.0。NAMRIA官方2020年洪水图来自水深测量和访谈，但图像不等于可用的原始测点表。

尚不能从历史副本确认的条件

官方周定义、病例发布日期与回填修订、NASA历史发布版本及病例数据再发布许可仍需确认。旧outbreak_flag不再使用；旧表的洪涝和积水列全为空，没有加入主模型。新接入的GFD事件比例来自另行核验的卫星资料，不填造街道积水记录。

03 未来预测协议与时间检查

预测发行日是源表一个非重叠起始日期的00:00。h周目标的起始日必须精确等于发行日加7h天，目标终止为再加6天；找不到精确目标就不评估，不用“下一行”替代。1周表示目标窗口在7天后开始，而不是同周拟合。

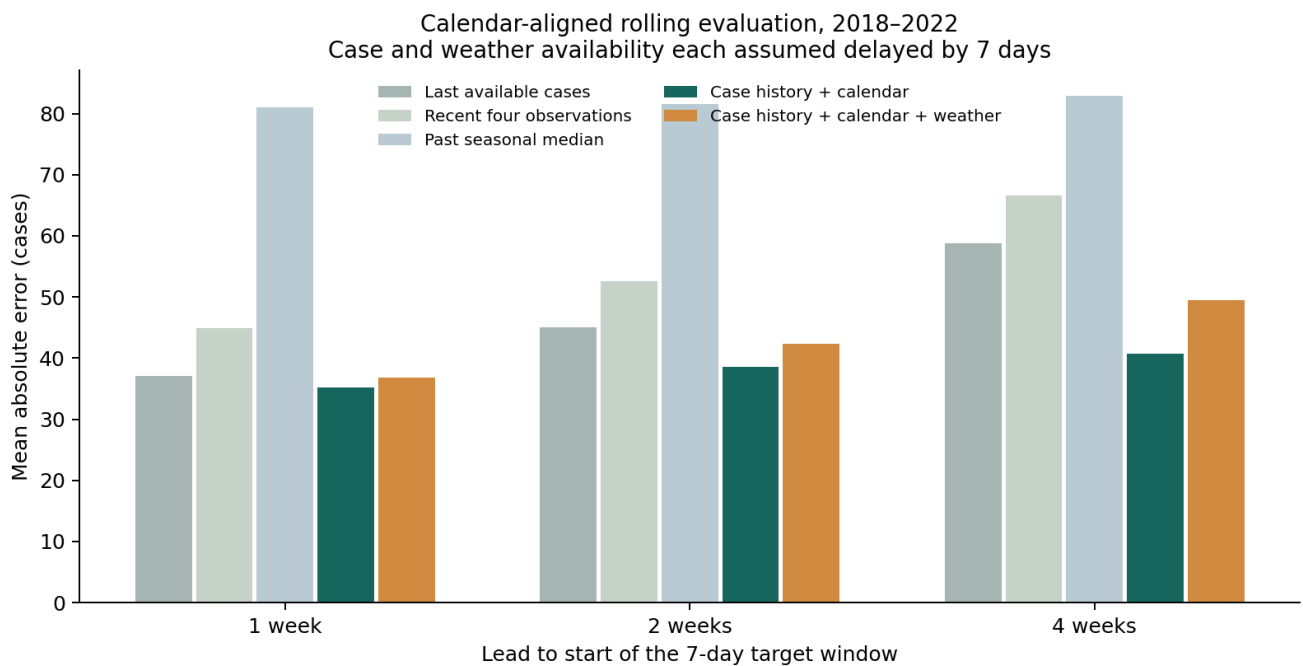
基准情景假设病例与气象各延迟7天；模型只能使用假设可用时间严格早于发行日的输入。病例输入包括最近可用病例、其实际年龄、近4条观测统计和按准确日期查询的滞后；气象用过去7/14/28天的完整日序列聚合。缺失特征只用训练集拟合的中位数填补，目标病例不填补。

初始训练仅包含2017年前可取得的标签；2017年验证用于比较Ridge、随机森林与梯度提升及预先列出的参数。参数选择在2018年开始前冻结。历史评估逐季度扩展训练窗口，训练标签的假设可用日必须早于该季度首次发行日，较远目标不会泄漏回训练。

比较最近可用病例延续、近4条平均、仅历史数据拟合的季节中位数、病例与季节模型、再加入气象模型。不同模型在相同发行日和目标样本上比较。高病例阈值由初始训练目标的75%分位数冻结；它只用于研究性排序指标，不是官方疫情暴发阈值，也不是感染概率。

随机种子42、43、44全部报告。确定性的梯度提升在本设置下三个种子结果相同，不能当作三份独立数据。当前观测有时间相关性，误差差异采用按年保留比例、8条保留观测的配对移动块重采样；所有统计推断均为条件探索，不能证明因果或跨城市效果。

04 完整对照与历史回放



图中展示相同样本上的五种方法。季节基线只使用当时已可获得的历史同季节观测中位数；由于源周日期不规则，不能把上一年最近一行暗称为同一流行病学周。新病例模型的4周误差下降更明显，加入气象并未额外改善。

figures/02_future_replay.png记录三个提前量在各时期的实际病例、病例模型与延续基线。2020年只是分段分析边界，没有排除2020 – 2022年。峰值预测仍有误差，较低平均误差不表示每周都更好。

05 稳健性与证据边界

提前量	目标时期	次数	延续基线MAE	病例模型MAE
1周	2018 – 2019	98	68.79	57.47
1周	2020 – 2022	146	15.84	20.31
2周	2018 – 2019	96	82.59	60.85
2周	2020 – 2022	142	19.65	23.59
4周	2018 – 2019	92	106.63	63.58
4周	2020 – 2022	134	25.88	24.91

2020 – 2022年中，1周和2周病例模型没有超过延续基线；4周仅有较小改善。因此不能只引用完整期平均误差，声称跨时期稳定有效。若将病例和气象延迟都设为14天，模型已另行完整重跑并保留全部结果；这说明可用性假设会影响结论。

病例模型相对延续基线的配对MAE差95%条件区间：1周[-7.04,4.38]例，2周[-14.07,0.43]例，4周[-30.76,-7.27]例。负值表示病例模型误差较低。6个探索对照的近似p值做Holm校正；4周对照约0.024。区间与p值不涵盖数据修订、真实发布延迟、重新选模或新城市差异。

不同提前量有效样本数不同，不宜简单比较AUC。旧0.93仍只属于历史同周分类口径，不作为本次未来预测证据。v3主指标是病例MAE，额外排序指标见模型指标.csv；没有把AUC或病例拟合图解释为防控后的病例减少。

06 GFD 洪涝特征模型探索

新接入GFD事件窗口的非永久水体像元比例，分别来自奎松周边和马尼拉研究框；两个研究框均非官方行政边界。项目保留原始事件覆盖表、20个裁剪栅格与事件锚点样本。NAMRIA 2020年洪水图作为相邻地区现场调查案例，不把地图颜色转换为虚构站点观测。

对10个2012 – 2017年GFD事件，按事件结束至少14天后寻找第一个能匹配全部未来目标的病例发行日，最多等待28天。其中9个事件满足条件。固定Ridge参数，前6个事件训练，按时间留出后3个事件；病例特征与加入两项GFD比例的版本在相同事件、相同未来目标比较。

未来目标	留出事件数	病例MAE	病例+GFD MAE
1周	3	43.48	40.93
2周	3	101.71	102.99
4周	3	69.51	76.75

这里只留出3个事件，1周方向稍好、2周和4周方向稍差，数字仅描述这3次事件，不能进行稳健显著性或泛化判断。部分事件中心远离马尼拉，研究框中的水体识别不能自动归因于该事件。事件目录是选择性样本，未收录周不当作无洪水。GFD全事件栅格的当年公开时间未核验，所以上述14天只是分析锚点，并非当时一定可用的输入；此实验不能替代2018 – 2022主评估或证明提前预警增益。

下一步优先取得覆盖2018 – 2022年的连续、带真实发布时点的洪水或积水资料及正式城市边界；锁定时间边界后，才在完整周序列上做病例、病例+气象、病例+洪涝的同发行日消融。对巡查优先级的价值，则另用有时间和位置的现场水深、回执与工作人员工时验证。

07 可运行产品与积水数据接口

历史预测演示.html已接入真实滚动预测。用户可以选择1/2/4周、病例模型或气象版本及年份，查看实际病例、当时预测、延续基线与发行日期。页面只回放已计算的归档结果，不在浏览器中重新训练模型，不显示当前风险。

代码提供独立的训练、评估、归档预测和传感器数据校验入口。models内保存六个模型包，每包含三种种子模型、特征和时间边界；未来目标、训练截止及逐次预测均可导出审计。持续积水观测模板.csv只有字段表头，没有伪造的观测。

传感器接口要求site_id、带时区的observed_at、水深cm、水温 ° C、质量标记、校准编号与来源编号。重复时间戳会拒收；缺校准、缺来源或无效水深会标记为不可用。连续湿状态只计算“连续湿样本的时间跨度”，缺测或超过两倍采样间隔时重置，并标出左删失，不能把断网期自动算作持续积水。

默认15分钟采样、1cm湿状态筛选均为可配置工程假设，不是蚊媒或疾病阈值。GFD事件模型与未来现场节点模型是两项不同实验；尚无真实节点数据可与病例对齐，因此不声称现场积水增益。节点仍需做台架精度、排水、防虫、维护与现场安全检查。

公众信息不展示个体病历和住户病例地址；任务最终由当地工作人员审核。模型或数据过期时停止现时预测，不把缺失信息标为低风险。旧合成巡查流程可继续作为用户任务演示，但必须与真实历史模型回放分别标注。

08 参赛表达与九十天实施

建议对外表述：Sentinella已建立可复现的奎松市城市级历史预测模块及持续积水观测接口；在明确时间假设的历史回放中，4周预测较病例延续基线降低平均误差约30.8%，气象额外增益和真实场景效果尚待验证。所有数字必须同时注明时期、样本和假设。

目标用户是地方卫生及社区巡查人员；潜在付费方为地方政府或公共卫生项目，而非已确认客户。先验证是否节省工作人员核查时间、提高可处置点命中率和回执完成率，再决定订阅服务。旧稿的大额收入和年费仍是未经付费访谈验证的假设。

第1 – 15天：请合作方核对官方流行病学周、GFD/NASA/CEMS的历史发布时间、正式城市边界和数据使用授权；获取覆盖2018 – 2022年的连续洪涝资料，并用连续新增病例登记前瞻评价协议。第16 – 30天：完成节点台架测试与校准，培训3 – 5名工作人员并试用任务清单。

第31 – 60天：经许可在一个社区3 – 5个点位影子运行，记录观测质量、维护和现场回执，不自动发布疾病或灾害警报。第61 – 90天：在相同巡查资源下比较既有工作流程与候选清单，报告可处置点命中率、工时与用户理解；若积水或病例事件不足，延长验证，不把90天当作病例减少的证明。

首个小试点仍沿用约10,320元的预算假设，须取得本地零件、通信和维护报价后重算，不含跨国差旅与全职人工。未签合作方、真实节点观测、新视频录制、公网发布和正式报名仍需团队完成；本包没有代为对外联系或提交。

来源：NASA POWER Daily API文档 <https://power.larc.nasa.gov/docs/services/api/temporal/daily/>；病例来源论文 <https://doi.org/10.1371/journal.pntd.0014444>；比赛 <https://ai4climate.world/en>。输入哈希、API响应及原始病例来源记录均随本地核验包保存。