

Aedes aegypti breeds in this much water. A bottle cap.

No specific treatment exists. Prevention means covering, emptying and cleaning water containers every week.

Photo: CDC Public Health Image Library #26033 (Bishop, 2022). WHO (2025).

In the Philippines, dengue is common.

in a single year

441,902

dengue cases reported in 2019,
the year a national epidemic was declared



8 in 10

Quezon City residents who died of dengue
in early 2025 were minors

But common shouldn't mean acceptable.

DOH Epidemiology Bureau via HDX, 86 provinces. Andrade (2019). Sampang (2025).

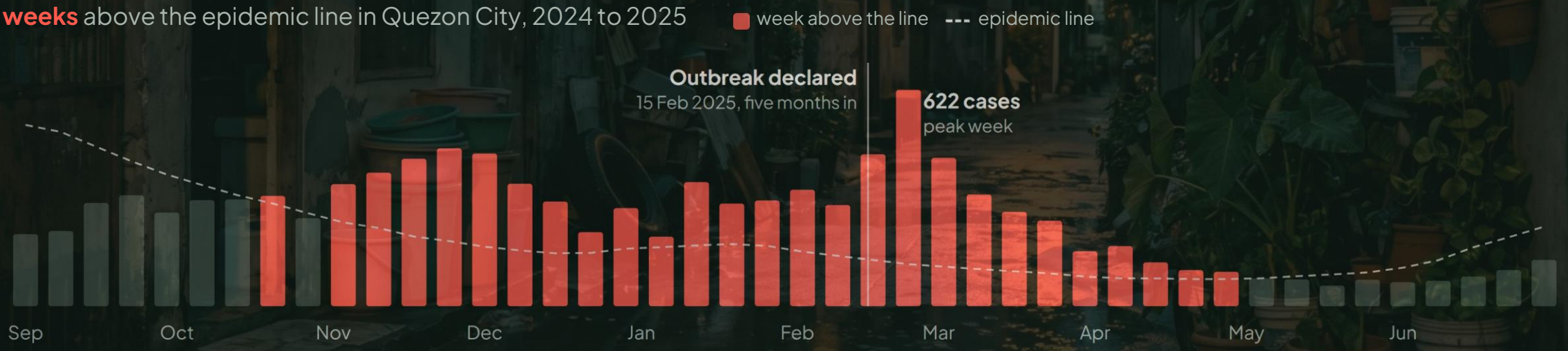
The Problem

By the time a case is counted, the outbreak has already begun.

Rain → Breeding → Bites → Sickness → Counted

weeks of warning, unused

27 weeks above the epidemic line in Quezon City, 2024 to 2025



Quezon City weekly reported cases (Pelitro, 2026). Epidemic line: mean + 2 SD of the same weeks in earlier years. Declaration: Sampang (2025).

EpiLag-PH

Who we are, and why we built it.

TEAM DIAGNODES

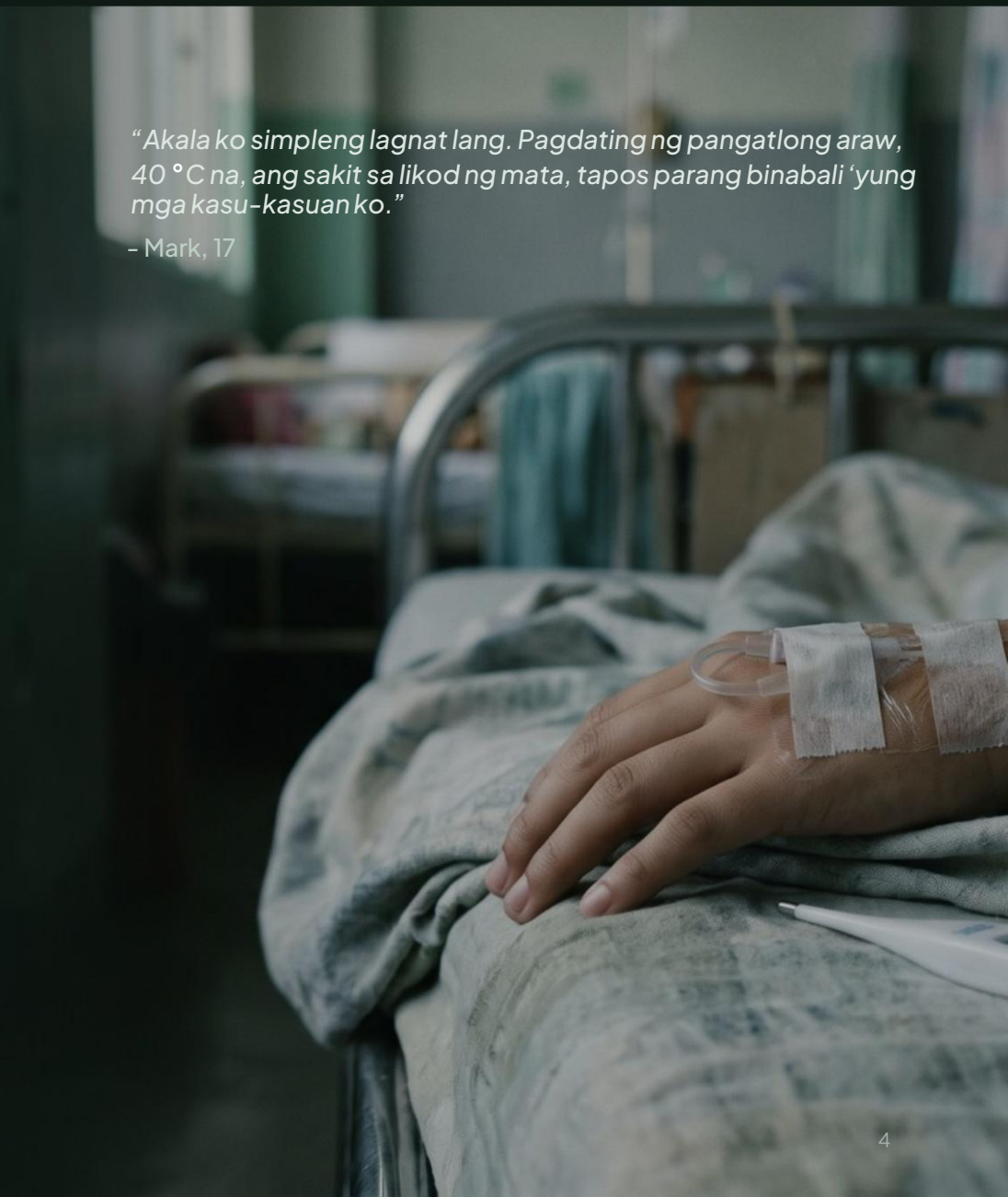
Jhezra Tolentino **Paul Dacalan** **Kishia Gaytano**

We built EpiLag-PH so a health office sees the next outbreak coming.

01
Forecast
dengue cases 2 to 8 weeks ahead

02
Alert
flag the areas heading for an outbreak

03
Act
tell health offices what to do, and where to spend



"Akala ko simpleng lagnat lang. Pagdating ng pangatlong araw, 40 °C na, ang sakit sa likod ng mata, tapos parang binabali 'yung mga kasu-kasuan ko."

- Mark, 17

Covers 86 provinces, 17 regions and Quezon City. Symptoms match those listed by WHO (2025).

Related Literature

What 19 studies say about weather and dengue, 9 of them in the Philippines.



Single-site studies, pinned.
Nationwide: Seposo 2023, Buczak 2014, Xu 2020.

Humidity
about 4 weeks, Carvajal 2018, Xu 2020

humidity acts fastest

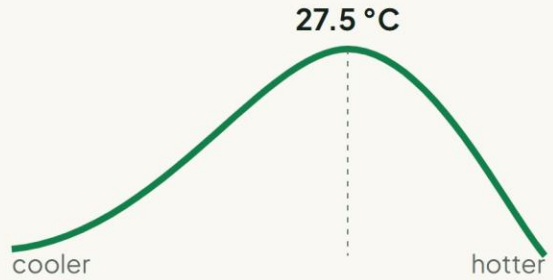
Rain
5 to 7 weeks, Sumi 2017, Carvajal 2018

Temperature
10 to 18 weeks, Seposo 2023, Mordecai 2017

temperature acts slowest: up to 4 months



HEAT BENDS



Risk rises with temperature, peaks near 27.5 °C, then falls (61 provinces, Seposo 2023).

extreme heat is bad for mosquitoes too

El Niño shifts dengue over 2 to 5 years (van Panhuis 2015)

Elevation tracks infection, not reported cases (Sy 2026)

Full review: research/lit/rrl_matrix.csv and feature_evidence.csv. Lags as reported by each study.

Related Literature

No Philippine study does all four. That is the gap EpiLag-PH fills.

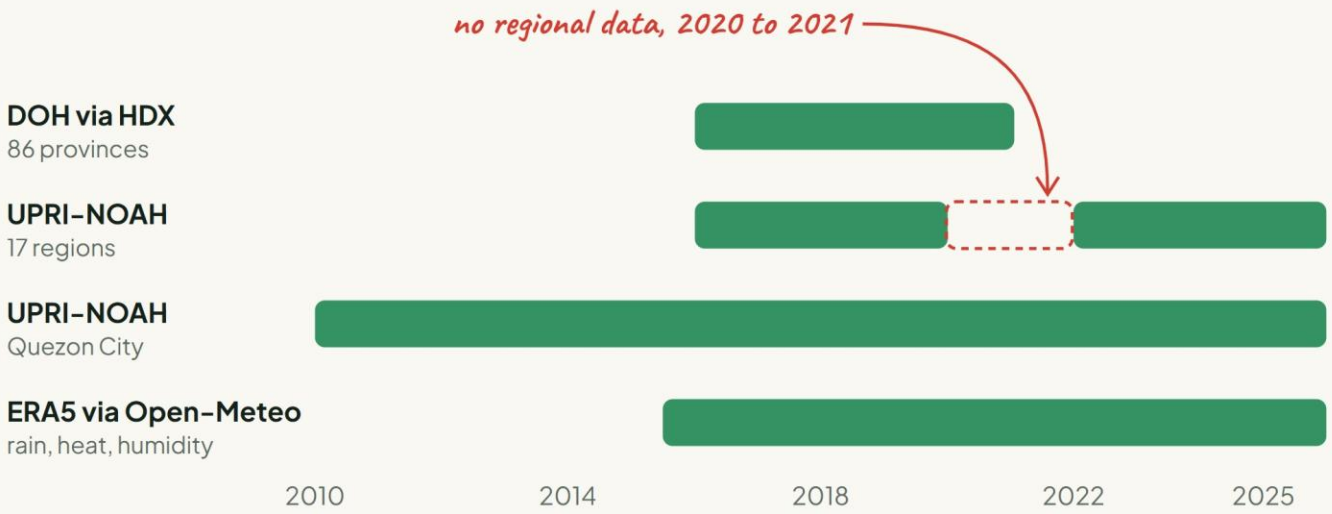
	Weekly, for many areas	Gives a likely range	Alerts scored for hits and false alarms	Tested on several unseen years
Olana et al. (2025) national, monthly		✓		
Carvajal et al. (2018) Metro Manila, weekly				
Eligado et al. (2026) Quezon City, weekly				
Buczak et al. (2014) 40 provinces, high or low	✓		✓	✓
EpiLag-PH 104 areas, weekly counts	✓	✓	✓	✓

EpiLag-PH fills the gap

Olana and Carvajal test one held-out year; Olana's ranges have no coverage check; Eligado compares alert thresholds without scoring hits; Buczak predicts classes, not counts. research/lit/positioning.md

Data Sources and Cleaning

Four public sources, cleaned in eight steps into one weekly table.



RAW FILE

REGION	YR	WN	DC_HDX	RF_HDX
REGION VII	2019	23	153	17.76

MODEL TABLE

unit_id	week	cases	rain_mm	y_lag0
R_REGION_VII	2019-06-09	153	17.76	5.04

1. **Match names to official codes**
   location names to PSGC-based area codes, fixing typos such as “Basilig” for Bislig
2. **Fold cities into their area**
   cities summed into their province or NCR district, giving 86 province units
3. **Merge split areas**
   Maguindanao, split in 2022, kept as one unit with Cotabato City for 2016 to 2020
4. **Keep only complete weeks**
   a week counts only if every reporting location sent data; otherwise it stays missing, never zero
5. **Epidemiological week to date**
   “2019, week 23” becomes the week ending 9 Jun 2019
6. **Gap-free weekly calendar**
   every area gets every week, so “4 weeks ago” always means 4 weeks
7. **Weather to weeks**
   daily ERA5 at each area's centre: rain summed, temperature and humidity averaged
8. **Scale the counts**
   log(cases + 1), so big provinces don't drown out small ones

DOH via HDX; Pelitro (2026); ERA5 (Hersbach et al., 2020); PSA and OCHA. Example row: Region VII, week ending 9 Jun 2019 (research/results/extracts/weekly_series.csv.gz). Code: src/clean_dengue.py, src/fetch_weather.py, src/features.py.

Feature Engineering

The literature's lags, turned into model inputs.

Recent cases 8 inputs
this week, 1 to 3 weeks back, 4 and 12 week averages, trend, same week last year

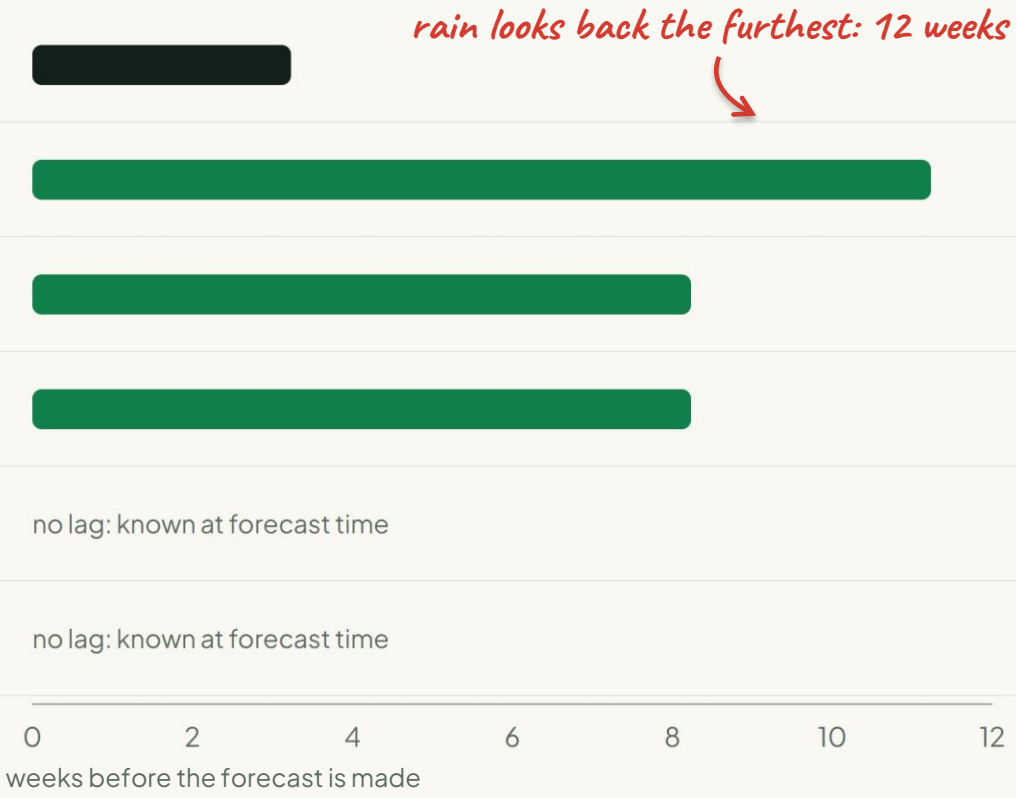
Rain 16 inputs
each of the last 12 weeks, plus 4, 8 and 12 week totals, and heavy-rain days (50 mm or more) in 8 weeks

Humidity 5 inputs
0, 2, 4, 6 and 8 weeks back

Temperature 6 inputs
0, 2, 4, 6 and 8 weeks back, plus the 4 week maximum

Time of year 2 inputs
sine and cosine, so December sits next to January

Population 1 input
log of population



38 *from 6 groups*

inputs per province forecast. Regions and Quezon City use 26: rain, but not heavy-rain days, temperature or humidity.

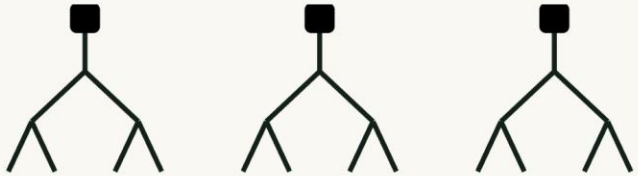
> **Every input is known on the day the forecast is made. No future data, ever.**

src/features.py, feature_columns(). Temperature and humidity were tested for regions in the research (slide 14).

Model

Quantile LightGBM with conformal calibration: a forecast, a likely range, and an alert.

GRADIENT-BOOSTED TREES, THREE PER FORECAST DISTANCE



10th
low estimate

50th
middle

90th
high estimate

Each quantile is trained with **pinball (quantile) loss**, which penalises misses on one side more than the other. 3 quantiles × 4 distances = **12 models** per area set.

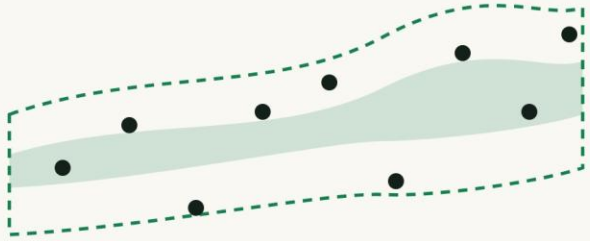
3 models: low, middle, high

HYPERPARAMETERS

`n_estimators` trees added one by one	**400**
`learning_rate` small steps, less overfitting	**0.03**
`num_leaves` how detailed each tree can be	**31**
`min_child_samples` at least 30 weeks per leaf	**30**
`subsample, colsample` 80% of rows and inputs per tree	**0.8, 0.8**

set once, the same for every test year

CONFORMALIZED QUANTILE REGRESSION



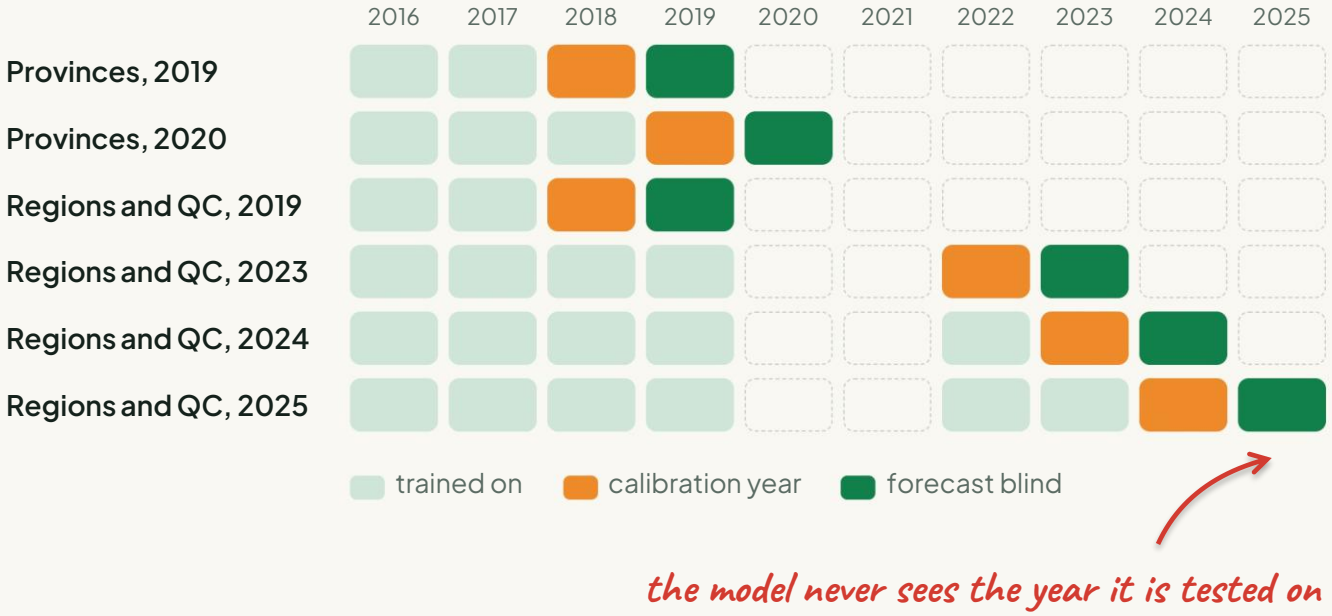
Hold out the last training year, measure how far real weeks fall outside the band, and widen it until **80%** fall inside.

filled = raw range
dashed = calibrated

src/backtest.py (PARAMS, fit_quantiles, conformal_margin). Ke et al. (2017); Romano et al. (2019).

Validation

Expanding-window time split: every test year is forecast blind.



24

checks: 6 test years × 4 forecast distances

No shuffling

rows are never mixed across time, unlike a random split

2 baselines

“same as this week” and “same week last year”

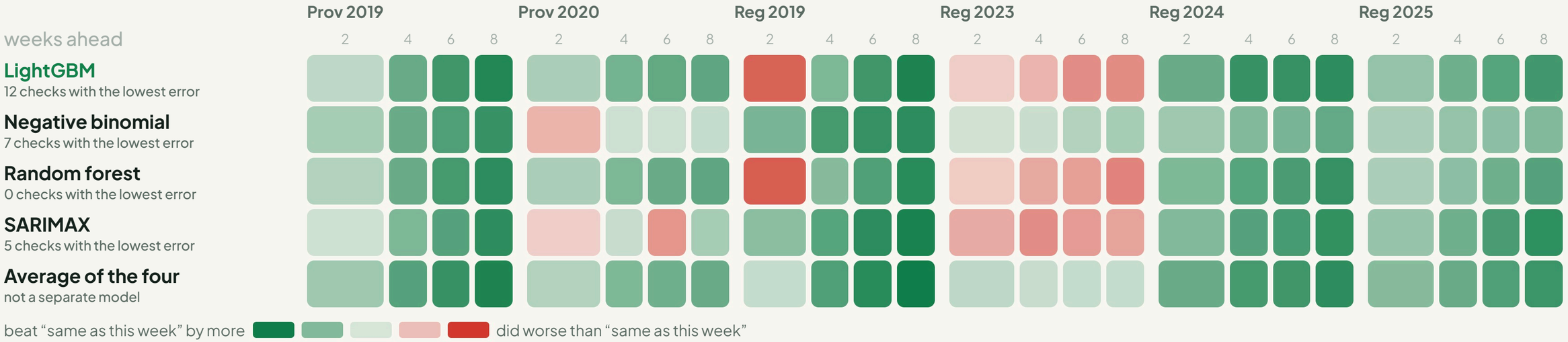
4 scores

error (MAE), range coverage, outbreaks caught, false alarms

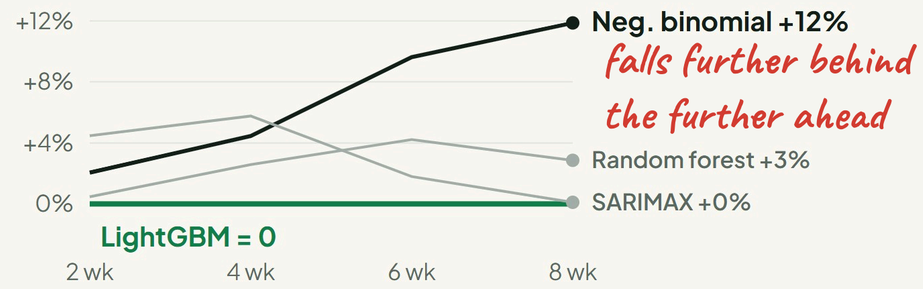
src/backtest.py, src/alerts.py; research/results/METHODS.md. An outbreak start is the first of 2 or more weeks above the epidemic line.

Model Comparison and Trade-offs

Four methods on the same blind test, plus their average. Each square is one check.



HOW MUCH LESS ACCURATE THAN LIGHTGBM, BY WEEKS AHEAD



WHY WE KEPT LIGHTGBM

12 of 24

checks with the lowest error, the most of any method

36% vs 30%

outbreaks caught early, vs the slightly more accurate average of the four

WHERE IT FALLS SHORT

Record surges: 2019 rose past anything in training

Quiet years: 2023's range came out too wide (92%)

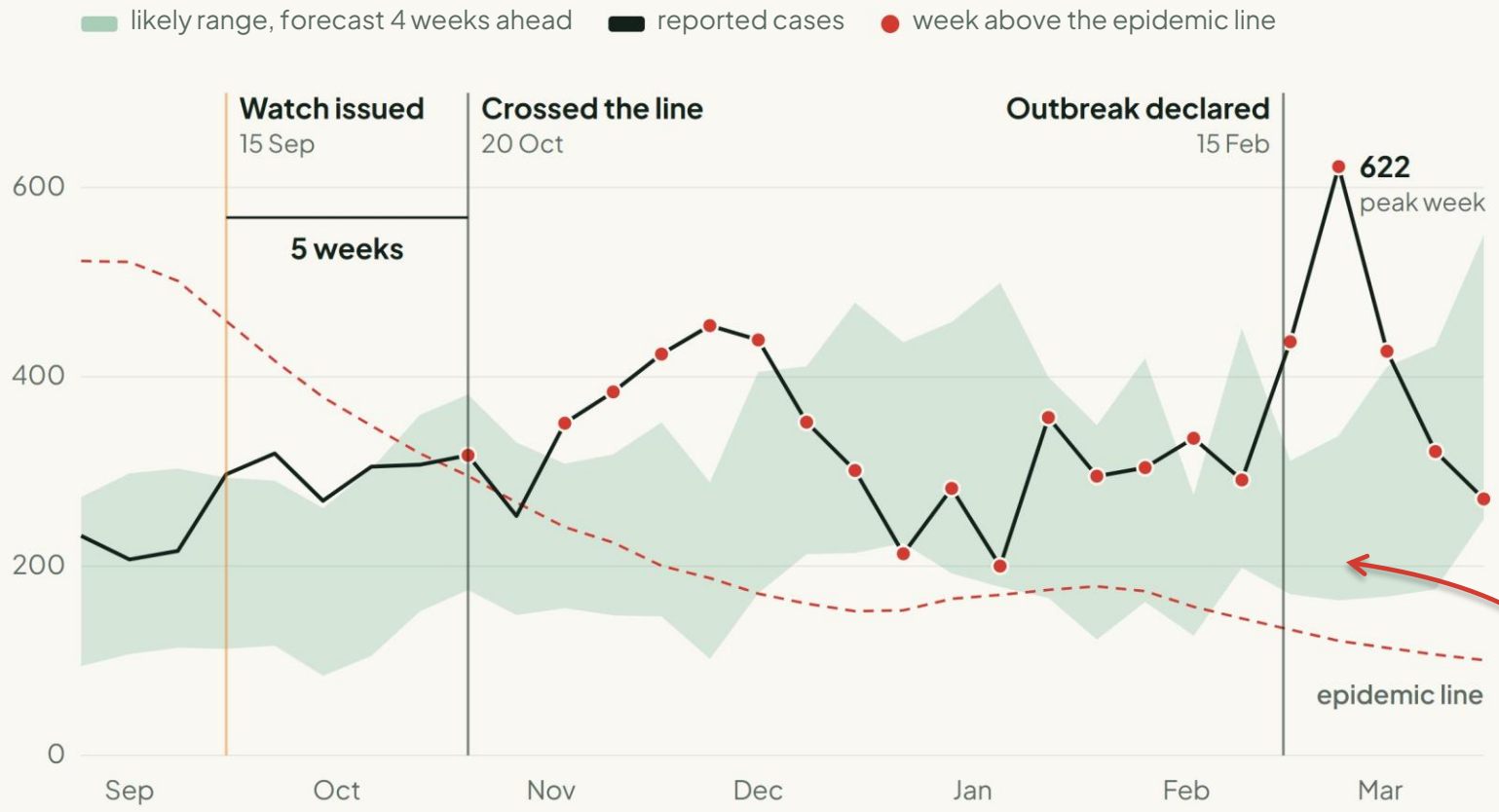
False alarms: about half of Watch alerts, on purpose

Better inputs: not live yet: chosen after seeing the test years

Color = error divided by the error of "same as this week". Prov = provinces; Reg = regions and QC. research/results/model_comparison.csv; research/analysis/causes.md.

Results

It saw Quezon City's outbreak coming five weeks early.



5 weeks

of warning before Quezon City crossed its epidemic line in 2024

236

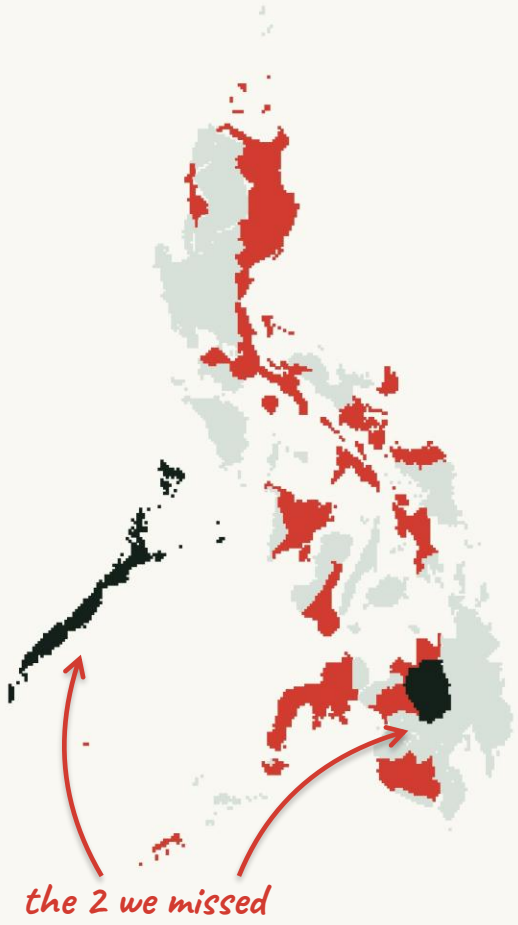
forecast for the peak week, against 622 reported. The alert was right; the size was not.

called the outbreak five weeks early, not its full height

Quezon City, weekly reported cases, Sep 2024 to Mar 2025. Forecasts use only data available at the time. Sources: Pelitro (2026); EpiLag-PH backtests.

Results

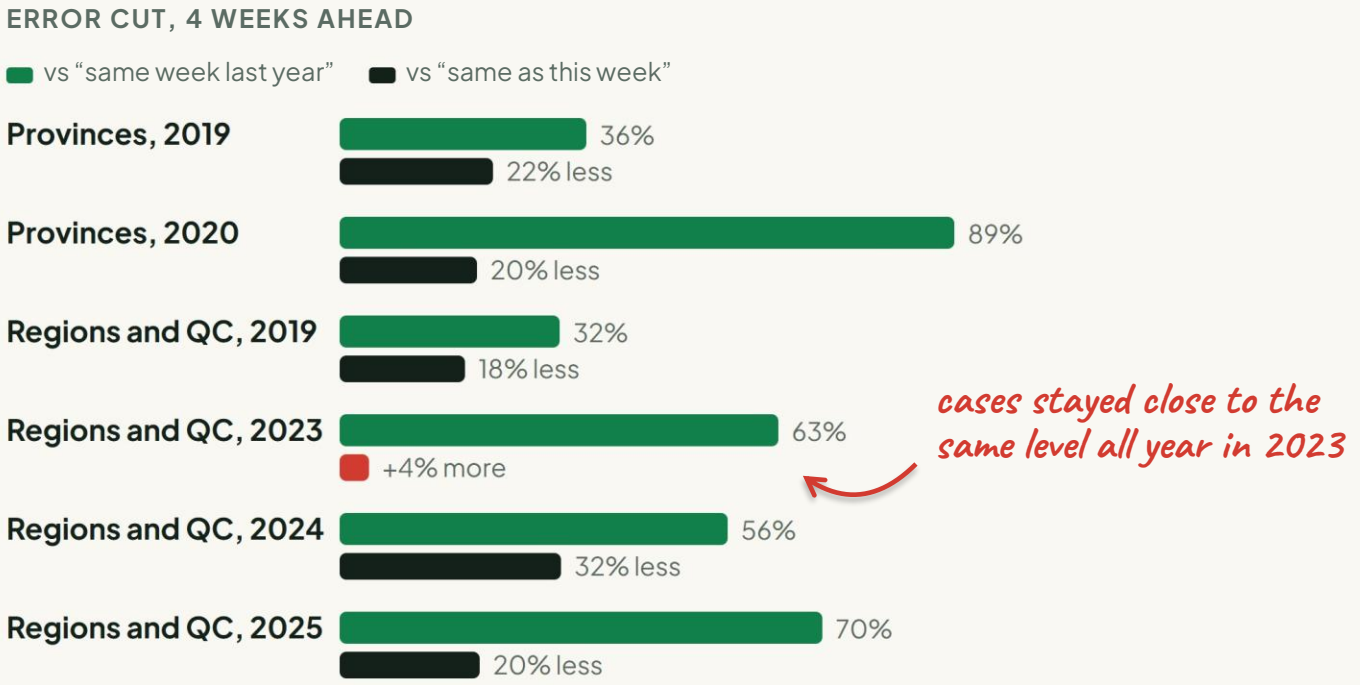
Nationwide, the 2019 replay flagged 30 of 32 provinces before they crossed the line.



30
crossed the line, flagged 4 weeks ahead

2
crossed, missed

did not cross



Ranges held 70 to 92% of real weeks across test years (target 80%).

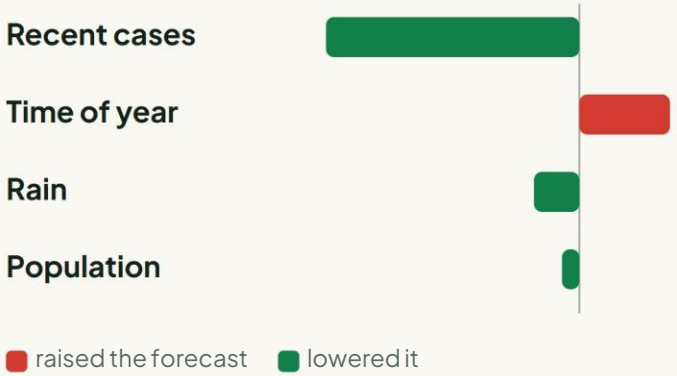
Forecasts made 2 Jun 2019 for 30 Jun 2019 (docs/data/replay_2019-06-02.json). Error: reports/backtest_province.csv, reports/backtest_region_qc.csv (mean absolute error).

Key Drivers

What the forecast relies on, and which factors earn their place.

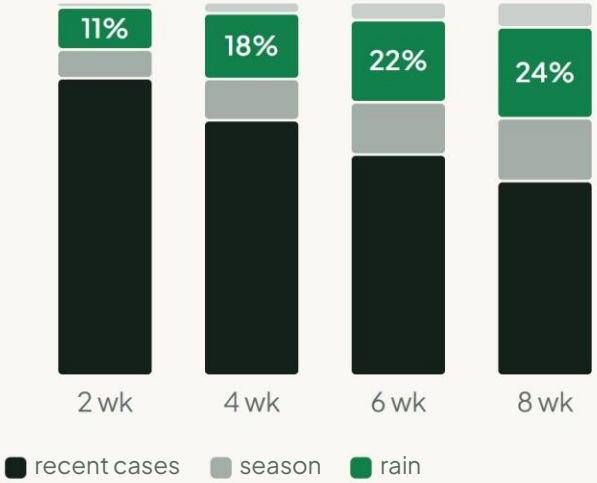
1. SHAP: ONE FORECAST, EXPLAINED

Recent cases pulled Quezon City's forecast down; the season pushed it up.



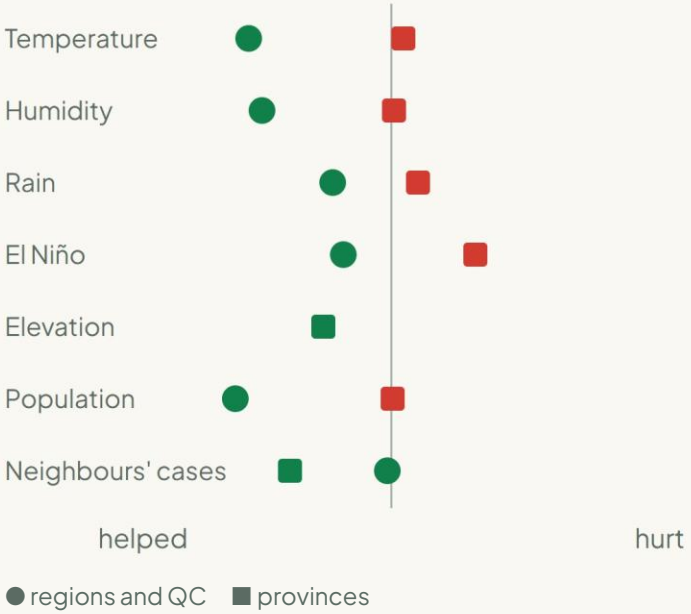
2. FURTHER AHEAD, WEATHER MATTERS MORE

Rain's share of attention: 11% at 2 weeks, 24% at 8.



3. EACH FACTOR, TESTED

Added to recent cases and season: did the error drop?

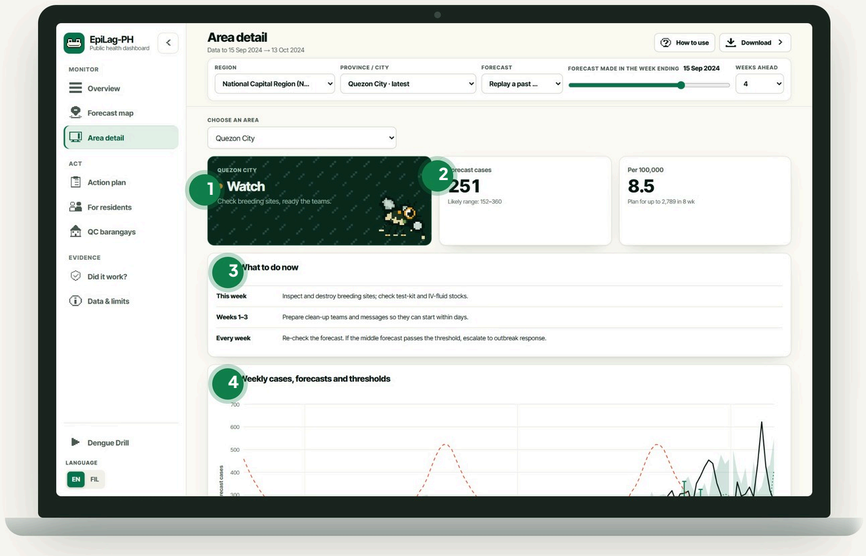


SHAP: Lundberg and Lee (2017). Quezon City latest forecast (region_qc_drivers.json); shares: shap_groups.csv; tests: feature_groups.csv (add), proposed_features.csv (neighbours).

The System

One app: an action plan for health offices, plain words and SMS for residents, a game for training.

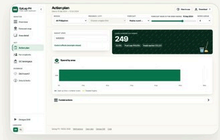
FOR HEALTH OFFICES



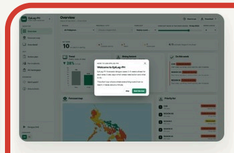
1 Alert level **2** Likely range **3** What to do **4** Against the line



Outbreak map



Budget plan



Onboarding guide

first visit? a guided tour in English and Filipino

FOR RESIDENTS

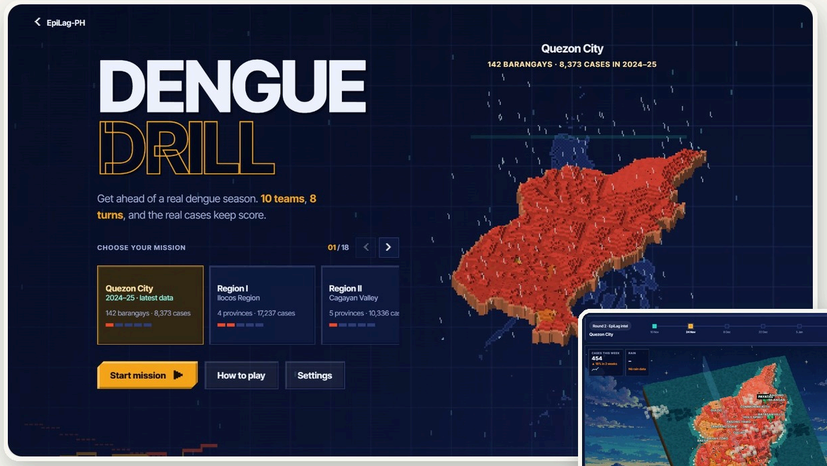


Residents' page in Filipino

SMS, NO DATA NEEDED

BANTAY DENGUE Quezon City: maaaring tumaas sa 4 linggo. Takpan ang tubig, itapon ang naipon. Lagnat 2+ araw? Magpatingin.

FOR TRAINING: DENGUE DRILL



Dengue Drill: pick a real outbreak



In play: with EpiLag-PH intel

WHEN THE COMPUTER PLAYER FOLLOWED EPILAG-PH, 2019 MISSIONS

12 regions, more cases averted | **5** fewer

TRY IT LIVE

epilag.vercel.app

Live app and Dengue Drill (epilag.vercel.app). Health office view: Quezon City, forecast of 15 Sep 2024. SMS: residents' page, 4 weeks ahead. Drill: 2019 missions; scripts/drill_data.py.

EpiLag-PH

How it works: one forecast, followed through the whole system.

Public data → LightGBM → Alert → Dashboard → Dengue Drill

2:00

Expert Consultation

Initial feedback on our pitch and live app, from two practitioners.

WHAT WE CONFIRMED

The LGU's current disease surveillance system for dengue has no forecasting. Monitoring is real-time.



Drake Jordan
Data Engineer · online consultation, 19 Nov 2026

Name withheld at his request

Epidemiology and Surveillance Officer
a Quezon City LGU · consultation, 20 Nov 2026

Initial consultation and feedback sessions, 19 and 20 Nov 2026. No partnership or endorsement is claimed. Photo: online session with Drake Jordan and our mentor, 19 Nov 2026.

Impact and Scaling Plan

Start in Batasan Hills, Quezon City's barangay with the most dengue cases. Prove it. Then grow.

1 Batasan Hills, Quezon City

the most dengue cases of the city's 142 barangays: 327 a year since 2010, 968 in 2025; about 169,000 people

Setup	Shadow run	Live: rainy season	Early-year tail	Evaluate
data link	forecasts scored, no action	a Watch triggers clean-up within a week	2025 peaked in Feb	go or no-go

Jan 2027 — Jul — Jan 2028 — Jun 2028

EXPECTED RESULTS (ESTIMATES), COMPARED WITH PAYATAS, COMMONWEALTH AND HOLY SPIRIT

JUN 2027
Forecasts scored live
first clean-up round before the rains

DEC 2027
18 fewer cases
range 6 to 31

JUN 2028
Lead time, response and false alarms scored
28 fewer cases expected (9 to 47); 57 in a 2025-sized season

Pass mark: a Watch 4+ weeks before cases cross the epidemic line, clean-up within a week, few false alarms.

₱0	Open source	No install	No patient data
to run EpiLag-PH; about 4 staff hours a week	code and open forecast data	runs in any browser, on any phone	counts only, no names (Data Privacy Act)

THEN GROW, ONE STEP AT A TIME

2 District 2
Jul 2028

102
fewer cases a season
range 34 to 170

+ Payatas, Commonwealth, Holy Spirit, Bagong Silangan: 23% of the city's cases

3 All of Quezon City
Jul 2029

445
fewer cases a season
range 148 to 742

142 barangays, 3.08 million people; money goes first where it averts the most cases per peso

4 NCR, then national
from 2030

regional epidemiology units, then the DOH Epidemiology Bureau; province forecasts already run weekly

full retrain: 69 s for 1,600 areas on a laptop CPU, no GPU
a retrain goes live only if it beats the simple baselines

each step must prove itself first

Estimates. Cases averted = season cases × 65% in the 24 busiest weeks × 30% of residents reached × effect 10/30/50% (Bowman 2016). Population: PSA 2024 census. Retrain time: reports/scale_benchmark.csv; gate: scripts/model_gate.py. No partnership is claimed.

EPILAG-PH, QUEZON CITY

Watch

5 weeks ahead

Next rainy season, we don't have to wait for the count.

Dengue is common.
But common shouldn't mean acceptable.

epilag.vercel.app



References

ABS-CBN News. (2024, July 1). *Dengue cases up as DOH asks public to search, destroy mosquito breeding grounds*. https://news.abs-cbn.com/news/2024/7/1/dengue-cases-up-as-doh-asks-public-to-search-destroy-mosquito-breeding-grounds-1435

Andrade, J. I. (2019, August 7). DOH declares national dengue epidemic. *Philippine Daily Inquirer*. https://newsinfo.inquirer.net/1151103/doh-declares-national-dengue-epidemic

Bishop, L. (2022). *Aedes aegypti larvae and pupae in the water of a plastic bottle cap* [Photograph] (ID 26033). Public Health Image Library, Centers for Disease Control and Prevention. https://phil.cdc.gov/details.aspx?pid=26033

Buczak, A. L., Baugher, B., Babin, S. M., Ramac-Thomas, L. C., Guven, E., Elbert, Y., Koshute, P. T., Velasco, J. M., Roque, V. G., Tayag, E. A., Yoon, I.-K., & Lewis, S. H. (2014). Prediction of high incidence of dengue in the Philippines. *PLOS Neglected Tropical Diseases, 8*(4), Article e2771. https://doi.org/10.1371/journal.pntd.0002771

Carvajal, T. M., Viacrusis, K. M., Hernandez, L. F. T., Ho, H. T., Amalin, D. M., & Watanabe, K. (2018). Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines. *BMC Infectious Diseases, 18*, Article 183. https://doi.org/10.1186/s12879-018-3066-0

Department of Health Epidemiology Bureau. (n.d.). *Philippine dengue cases and deaths* [Data set]. Humanitarian Data Exchange. https://data.humdata.org/dataset/philippine-dengue-cases-and-deaths

Eligado, A. A., Hasegawa, T., Hattori, Y., & Nakamura, K. (2026). Advancing outbreak detection: Hybridizing machine learning with wavelets for weekly dengue case forecasting. *PLOS Neglected Tropical Diseases, 20*(7), Article e0014444. https://doi.org/10.1371/journal.pntd.0014444

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., . . . Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society, 146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems, 30*.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*.

Mordecai, E. A., Cohen, J. M., Evans, M. V., Gudapati, P., Johnson, L. R., Lippi, C. A., Miazgowicz, K., Murdock, C. C., Rohr, J. R., Ryan, S. J., Savage, V., Shocket, M. S., Stewart Ibarra, A., Thomas, M. B., & Weikel, D. P. (2017). Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. *PLOS Neglected Tropical Diseases, 11*(4), Article e0005568. https://doi.org/10.1371/journal.pntd.0005568

Olana, K. O. A., Poprom, N., Siewchaisakul, P., Punyapornwithaya, V., & Thongprachum, A. (2025). Spatial distribution analysis and comparative forecasting of dengue resurgence in the Philippines (2025–2027): A nationwide study. *Transboundary and Emerging Diseases, 2025*, Article 7480710. https://doi.org/10.1155/tbed/7480710

Pelitro, K. J. (2026). *Weekly dengue incidence linked to satellite rainfall at city, regional, and country scales* (Version 1.0.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.19448695

Romano, Y., Patterson, E., & Candès, E. (2019). Conformalized quantile regression. *Advances in Neural Information Processing Systems, 32*.

Sampang, D. (2025, February 15). QC declares dengue outbreak as cases surge 200% this year. *INQUIRER.net*. https://newsinfo.inquirer.net/2034957/qc-declares-dengue-outbreak-as-cases-surge-200-this-year

Seposo, X., Valenzuela, S., & Apostol, G. L. (2023). Socio-economic factors and its influence on the association between temperature and dengue incidence in 61 provinces of the Philippines, 2010–2019. *PLOS Neglected Tropical Diseases, 17*(10), Article e0011700. https://doi.org/10.1371/journal.pntd.0011700

Sumi, A., Telan, E. F. O., Chagan-Yasutan, H., Piolo, M. B., Hattori, T., & Kobayashi, N. (2017). Effect of temperature, relative humidity and rainfall on dengue fever and leptospirosis infections in Manila, the Philippines. *Epidemiology and Infection, 145*(1), 78–86. https://doi.org/10.1017/S095026881600203X

Sy, A. K., Biggs, J., Chowdhury, S. K., Quinones, M. A., Salazar, F. V., Jones-Warner, W., Ashall, J., Sucaldito, M. N. L., Cutiongco-de la Paz, E., Capeding, M. R. Z., Padilla, C. D., Villanueva, S. Y. A. M., Hibberd, M. L., & Hafalla, J. C. R. (2026). The impact of elevation and population density on dengue incidence and force of infection across the Philippines. *PLOS Neglected Tropical Diseases, 20*(5), Article e0014356. https://doi.org/10.1371/journal.pntd.0014356

van Panhuis, W. G., Choisy, M., Xiong, X., Chok, N. S., Akarasewi, P., Iamsirithaworn, S., Lam, S. K., Chong, C. K., Lam, F. C., Phommasak, B., Vongphrachanh, P., Bouaphanh, K., Rekol, H., Hien, N. T., Thai, P. Q., Duong, T. N., Chuang, J. H., Liu, Y. L., Ng, L. C., . . . Cummings, D. A. (2015). Region-wide synchrony and traveling waves of dengue across eight countries in Southeast Asia. *Proceedings of the National Academy of Sciences, 112*(42), 13069–13074. https://doi.org/10.1073/pnas.1501375112

World Health Organization. (2025, August 21). *Dengue and severe dengue* [Fact sheet]. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue

Xu, Z., Bambrick, H., Yakob, L., Devine, G., Frentiu, F. D., Villanueva Salazar, F., Bonsato, R., & Hu, W. (2020). High relative humidity might trigger the occurrence of the second seasonal peak of dengue in the Philippines. *Science of the Total Environment, 708*, Article 134849. https://doi.org/10.1016/j.scitotenv.2019.134849