

A Hybrid Statistical and Probabilistic Model for Forecasting Uncertain Events

Ayotunde Oluwale Ajala^{1*}, Dave Onyemaechi Ekeh²

¹ Sports Management Studies, National Institute for Sports, Lagos, Nigeria;

² Educational Management and Foundational Studies, Alex Ekwueme Federal University, Ndufu-Alike, Ebonyi State, Nigeria;

* **Corresponding author:** ayotundeajala15@gmail.com.

ABSTRACT: Forecasting uncertain events requires models that provide more than a single expected value, particularly when the target is a variable event count affected by seasonality, environmental conditions, and irregular outbreak peaks. This study proposes and evaluates a hybrid statistical-probabilistic framework for forecasting uncertain events, using weekly dengue incidence as an empirical case study. The framework integrates time-series structuring, lagged statistical features, environmental regression predictors, ensemble point forecasting, negative-binomial distribution fitting, and uncertainty interval estimation. A public dengue forecasting dataset containing 1,456 city-week observations from San Juan and Iquitos was used for chronological model evaluation. The proposed model was compared with seasonal naive, moving-average, log-ridge regression, random-forest, and gradient-boosting baselines using mean absolute error (MAE), root mean squared error (RMSE), symmetric mean absolute percentage error, prediction-interval coverage, and interval width. The hybrid model achieved the lowest validation error, with an MAE of 5.908 and an RMSE of 9.874, outperforming the seasonal naive baseline by a large margin and modestly improving over the individual ensemble learners. The probabilistic layer produced empirically conservative uncertainty intervals, with 87.67% coverage for the nominal 80% interval and 93.84% coverage for the nominal 90% interval. Feature-importance results showed that recent case history was the strongest predictor, while precipitation, temperature, humidity, and time-index variables contributed additional contextual information. The study contributes an interpretable forecasting framework that connects statistical temporal structure with probabilistic event-risk estimation. The main limitation is that the empirical evidence is based on two locations, so future research should test the framework across broader domains of outbreak, climate, financial, and operational events.

Keywords: probabilistic forecasting; uncertainty estimation; time-series modeling; outbreak prediction; hybrid statistical model.

I. INTRODUCTION

Forecasting uncertain events is a central problem in contemporary decision science because many important outcomes do not unfold as stable, deterministic trends. Disease outbreaks, weather-related risks, market instability, resource demand, operational disruptions, and public-health emergencies are shaped by historical patterns, seasonal cycles, random variation, incomplete measurement, and sudden changes in system behavior. In these settings, a forecast that reports only one expected future value can be useful, but it is rarely sufficient for decision-making. Decision-makers usually need to know not only what is likely to happen, but also how uncertain the estimate is, how wide the range of plausible outcomes may be, and whether the probability of a high-risk event exceeds a practical threshold.

Traditional statistical forecasting has strong advantages for uncertain-event prediction because it provides transparent tools for modeling temporal structure, trend, seasonality, regression effects, and residual variability. However, purely deterministic or point-based statistical models can understate risk when the event process is overdispersed, nonlinear, or influenced by local contextual variables. Probabilistic methods address this weakness by estimating distributions, confidence ranges, or prediction intervals around future outcomes. A probabilistic forecast is therefore more informative than a single point estimate because it conveys the expected outcome along with uncertainty and tail-risk information. This distinction is particularly important for count outcomes, such as weekly disease cases, where peaks are usually rare but consequential.

The present study develops a hybrid statistical-probabilistic model for forecasting uncertain events. The core idea is that historical time-series structure and contextual predictors should be used together to estimate the expected event level, while a fitted probability distribution should be used to quantify the uncertainty around that expectation. The model is hybrid in two senses. First, it combines statistical feature construction with regression and ensemble forecasting. Second, it combines point forecasts with a probabilistic layer that transforms predicted means into event distributions, prediction intervals, and exceedance probabilities. This design allows the framework to remain interpretable while still capturing nonlinear relationships and overdispersed event variation.

Dengue incidence is used as the empirical case because it is a clear example of an uncertain event process. Dengue outbreaks are influenced by climatic conditions, mosquito ecology, seasonal patterns, population exposure, and prior case dynamics. Weekly dengue case counts therefore provide a suitable test setting for a model that must capture both temporal regularity and irregular outbreak peaks. The empirical case also allows the model to be evaluated in two distinct locations with different outbreak intensities: San Juan and Iquitos. This two-location structure is valuable because a model that performs well in only one outbreak regime may not generalize to another.

The study is guided by three objectives. The first objective is to construct an interpretable hybrid framework that links time-series features, environmental predictors, and probabilistic uncertainty estimation. The second objective is to compare the proposed model with standard forecasting baselines under a chronological validation design. The third objective is to evaluate whether the model can provide meaningful prediction intervals and outbreak-risk probabilities, rather than only point forecasts.

The research questions are as follows. First, can a hybrid statistical-probabilistic model improve forecasting accuracy for uncertain event counts compared with conventional baselines? Second, can the probabilistic layer produce empirically useful uncertainty intervals for future event counts? Third, which historical and environmental variables contribute most strongly to the forecasting process? These questions are addressed using an empirical dengue forecasting dataset and a transparent modeling pipeline.

II. LITERATURE REVIEW

Recent evidence on dengue forecasting shows that meteorological variables can support practical prediction pipelines when combined with transparent model evaluation and operational interpretation [1]. Systematic early-warning evidence further shows that outbreak tools depend strongly on data quality, validation design, and the clarity of model outputs [2]. Distributional and ensemble-based dengue-risk modeling also supports the value of combining statistical distribution logic with flexible predictors for risk-sensitive forecasting [3].

Hybrid forecasting research supports the use of engineered temporal variables, regression structure, and nonlinear learners in the same forecasting pipeline [4]. Multi-location weekly dengue forecasting evidence shows that San Juan and Iquitos provide a suitable empirical setting for evaluating outbreak forecasts under different local regimes [5]. High-incidence dengue prediction studies indicate that early-warning signals are useful, but they require probabilistic interpretation to support decision-making under uncertainty [6]. Local dengue forecasting research also emphasizes that city-level and neighborhood-level structures influence predictive accuracy and should not be ignored in model design [7].

Climate-fluctuation studies show that precipitation, temperature, and regional weather dynamics contribute to dengue outbreak behavior [8]. Additional climate-focused evidence shows that environmental

variables may improve dengue forecasts differently across local contexts, which supports the city-aware design used in the present study [9]. Probabilistic forecasting theory provides the foundation for evaluating forecasts as distributions rather than isolated point estimates [10]. The calibration-sharpness principle further clarifies that a useful predictive distribution should include observed values at appropriate rates while remaining sufficiently informative for decision-making [11].

Automated time-series forecasting research supports systematic baseline construction and chronological model selection when temporal dependence is present [12]. Infectious-disease forecast evaluation requires comparison with simple seasonal and climate-driven reference models, because apparent model improvement can be misleading without clear baselines [13]. Prediction-interval research shows that uncertainty ranges should be formally assessed rather than treated as a secondary addition to point accuracy [14]. Systematic reviews of dengue prediction models confirm the need for transparent reporting, comparable evaluation standards, and interpretable performance measures [15].

Environmental-risk forecasting studies show that satellite rainfall and related climate information can strengthen early-warning systems when local surveillance data are limited or delayed [16]. Nonlinear climate-risk modeling indicates that dengue incidence may respond to delayed and nonlinear environmental effects, supporting the inclusion of lagged and contextual predictors [17]. Large-scale forecasting comparisons demonstrate that model performance depends on data structure, forecast horizon, and validation design, which justifies the comparative evaluation strategy used in this manuscript [18]. Web-based and surveillance-assisted dengue estimation also show that alternative signals can support outbreak forecasting while keeping the final output interpretable for practical risk assessment [19].

Contemporary forecasting principles stress transparent validation, benchmark comparison, and practical interpretability in applied forecasting studies [20]. Reviews of epidemiological early-warning systems show that climate-informed models are promising but methodologically heterogeneous, reinforcing the need for a reproducible framework [21]. Prediction intervals for ensemble learners provide an important route for converting flexible point forecasts into uncertainty-aware outputs [22]. Meteorological and epidemiological dengue-forecasting evidence supports the operational use of recent case counts and weather variables in applied prediction systems [23].

Probabilistic time-series forecasting research demonstrates the value of modeling full predictive distributions rather than only conditional means [24]. Recent systematic evidence on data-driven dengue surveillance shows that real-world and large-scale data approaches are expanding, but comparability and uncertainty reporting remain persistent weaknesses [25]. Transparent trend and seasonality modeling remains valuable when forecasts must be communicated to applied decision-makers rather than used only as technical outputs [26]. Reviews of prediction-interval methods support the use of parametric and distribution-based uncertainty estimation as a formal layer of model assessment [27].

Spatiotemporal statistical modeling provides a broader framework for representing dynamic risk processes across space and time [28]. Forecast-error uncertainty research shows that validation errors themselves can be uncertain under temporal dependence, which requires caution when interpreting model performance [29]. Robust probabilistic time-series forecasting further highlights the importance of reliability when predictive systems are exposed to unstable or perturbed event patterns [30]. These studies support the central position of the present manuscript: uncertain-event forecasts should report expected counts, prediction intervals, and high-risk probabilities in a single interpretable framework.

Table 1 summarizes studies that directly inform the present manuscript. The table is deliberately focused on studies and methodological sources relevant to probabilistic forecasting, statistical time-series modeling, dengue prediction, uncertainty intervals, and hybrid forecasting. It is retained as a comparative literature matrix and does not include the research gap or proposed method, because those components are treated in separate sections.

The reviewed literature shows significant progress in statistical forecasting, probabilistic forecasting, and dengue risk prediction. However, the combined evidence also shows that many studies still emphasize point accuracy over decision-oriented uncertainty outputs. The present study therefore extends prior work by combining city-aware temporal structure, environmental predictors, ensemble point forecasting, negative binomial distribution fitting, and high-risk exceedance probabilities into a single transparent framework.

Table 1. Literature review comparison table.

Study	Focus	Method / context	Relevance to this study
[1] Al Mobin et al. (2024)	Meteorological dengue forecasting	Machine-learning pipeline with meteorological predictors	Supports practical use of climate-informed dengue forecasting
[2] Baharom, Ahmad, and Hod (2022)	Dengue early-warning systematic review	Evaluation of outbreak prediction tools	Shows that forecast accuracy depends heavily on data quality and model choice
[3] Barboza et al. (2022)	Dengue risk and climate learning	GAMLSS and random forest modeling with uncertainty	Supports combining statistical distribution models and ensemble predictors
[4] Ben Taieb and Hyndman (2014)	Gradient boosting for load forecasting	Regression-based forecasting with engineered temporal features	Supports hybrid statistical and learning-based feature construction
[5] Benedum et al. (2020)	Weekly dengue forecasts	Multi-location dengue forecast comparison including Iquitos and San Juan	Strongly connected to the empirical cities and forecast horizon
[6] Buczak et al. (2014)	High-incidence dengue prediction	Classification-oriented outbreak prediction in the Philippines	Shows the usefulness of early-warning labels, but not full probabilistic counts
[7] Chen et al. (2018)	Neighborhood dengue forecasting	Real-time dengue forecasting in tropical urban Singapore	Shows that local spatiotemporal structure affects dengue predictability
[8] Chuang, Chaves, and Chen (2017)	Climatic fluctuation and dengue outbreaks	Time-series modeling of local and regional climatic effects	Supports precipitation, temperature, and regional climate variables
[9] Da Silva et al. (2024)	Climate variables and dengue forecasting	Random-forest dengue forecasting across multiple cities	Shows climate data may help differently across local contexts
[10] Gneiting and Raftery (2007)	Probabilistic forecasting and proper scoring rules	Foundational framework for evaluating probabilistic predictions	Supports the use of distribution-aware forecasts rather than point forecasts only
[11] Gneiting, Balabdaoui, and Raftery (2007)	Calibration and sharpness of probabilistic forecasts	Theoretical and diagnostic assessment of predictive distributions	Provides the calibration-sharpness logic used for prediction intervals
[12] Hyndman and Khandakar (2008)	Automatic ARIMA modeling	Statistical time-series model selection	Justifies statistical baseline comparison and chronological validation
[13] Johansson et al. (2016)	Climate-driven and seasonal dengue forecasts	Out-of-sample forecast evaluation for dengue in Mexico	Shows the importance of reference baselines in infectious-disease forecasting
[14] Khosravi et al. (2011)	Prediction intervals with neural models	Lower-upper bound estimation for uncertainty intervals	Shows why interval quality matters in forecasting evaluation
[15] Leung et al. (2023)	Dengue outbreak prediction systematic review	Review of quantitative outbreak prediction models	Shows need for methodological transparency and standard performance reporting
[16] Lippi et al. (2019)	Mosquito-borne climate suitability	Climate-sensitive transmission suitability analysis	Supports environmental variables as indirect outbreak risk indicators
[17] Lowe et al. (2018)	Climate-dengue modeling	Delayed and nonlinear climate effects on dengue risk	Supports lagged environmental variables in dengue forecasting
[18] Makridakis, Spiliotis, and Assimakopoulos (2020)	M4 forecasting competition	Large-scale comparison of statistical and computational models	Supports using multiple baselines rather than a single comparator
[19] Marques-Toledo et al. (2017)	Web-based dengue prediction	Social and surveillance signals for country/city forecasting	Highlights the value of alternative predictors but also the need for interpretable risk estimates

Study	Focus	Method / context	Relevance to this study
[20] Petropoulos et al. (2022)	Forecasting theory and practice	Synthesis of forecasting principles and validation needs	Supports transparent forecast design and robust evaluation
[21] Racloz et al. (2012)	Dengue early warning review	Climate and surveillance-based early-warning systems	Identifies early-warning potential but methodological heterogeneity
[22] Ramosaj, Schultheis, and Pauly (2021)	Random-forest prediction intervals	Out-of-bag and non-parametric interval construction	Supports uncertainty estimation around ensemble point forecasts
[23] Roster, Connaughton, and Rodrigues (2022)	Machine-learning dengue forecasting	Monthly dengue forecasting in Brazilian cities	Supports meteorological and epidemiologic predictors in operational forecasting
[24] Salinas et al. (2020)	DeepAR probabilistic forecasting	Autoregressive recurrent networks for probabilistic time series	Demonstrates modern probabilistic learning for count/time-series prediction
[25] Sylvestre et al. (2022)	Dengue data-driven surveillance review	Real-world and big-data methods for dengue prediction	Shows the growth of data-driven dengue surveillance and gaps in comparability
[26] Taylor and Letham (2018)	Forecasting at scale	Additive trend and seasonality framework	Supports transparent modeling of trend, seasonality, and holiday/time effects
[27] Tian, Nordman, and Meeker (2020)	Prediction interval methods	Review of parametric and distribution-based intervals	Supports the distribution-fitting layer of the proposed framework
[28] Wikle, Zammit-Mangion, and Cressie (2019)	Spatiotemporal statistics	Hierarchical and dynamic statistical modeling	Supports structured temporal and location-aware uncertainty modeling
[29] Xu et al. (2023)	Uncertainty intervals for prediction errors	Time-series forecast-error uncertainty intervals	Supports caution when interpreting validation errors under temporal dependence
[30] Yoon et al. (2022)	Robust probabilistic time-series forecasting	Probabilistic forecasting under perturbations	Supports robustness concerns in uncertainty-aware prediction

III. RESEARCH GAP

The literature indicates four important gaps. First, many dengue and uncertain-event forecasting studies emphasize point accuracy but provide limited probabilistic interpretation. This reduces the usefulness of forecasts when decision-makers need to know whether an event is merely expected to rise or likely to exceed a risk threshold. Second, several studies use climate and environmental variables, but their contributions vary by location, which means a city-aware design is needed. Third, hybrid models often improve predictive performance, yet their uncertainty layers are not always reported clearly. Fourth, interval coverage and interval width are rarely discussed together, even though a wide interval can be statistically safe but practically less informative.

The present study addresses these gaps by developing a framework that first estimates point forecasts and then fits a probabilistic count distribution around them. The framework is evaluated with baseline models and reports accuracy metrics, interval coverage, interval width, feature importance, and outbreak exceedance probabilities. This makes the study more useful for risk-sensitive forecasting than a deterministic model alone.

IV. CONCEPTUAL FRAMEWORK AND HYPOTHESES DEVELOPMENT

The conceptual framework assumes that uncertain event counts are generated by three interacting components: temporal memory, contextual exposure, and probabilistic variability. Temporal memory refers to the influence of recent observations and recurring seasonal cycles. Contextual exposure refers to environmental conditions such as precipitation, temperature, vegetation, and humidity. Probabilistic variability refers to the residual uncertainty that remains after temporal and contextual information has been modeled.

Based on this framework, three hypotheses are formulated.

- H1: A hybrid statistical-probabilistic model will reduce forecasting error compared with naive and single-model baselines.
- H2: A probabilistic count layer will produce prediction intervals whose empirical coverage is close to or higher than nominal coverage.
- H3: Recent case history will be the strongest predictor, while environmental variables will provide additional contextual forecasting value.

V. PROPOSED METHOD

The proposed method is a five-stage hybrid forecasting framework. First, the data are arranged as city-week time series, and missing values are handled using median-based imputation within the modeling pipeline. Second, statistical temporal features are constructed, including lagged case counts, rolling means, rolling standard deviations, seasonal sine/cosine terms, and a city-specific time index. Third, regression and ensemble forecasting models estimate the conditional expected number of weekly cases from historical and environmental predictors. Fourth, a negative binomial probabilistic layer is fitted around the hybrid point forecast to capture overdispersed event-count uncertainty. Fifth, the model reports a decision output consisting of a point forecast, prediction interval, and probability that the event count exceeds a city-specific high-risk threshold.

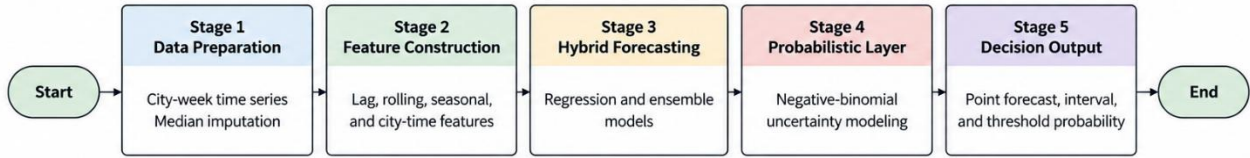


FIGURE 1. Proposed hybrid statistical-probabilistic forecasting framework.

FIGURE 1. presents the full workflow from data preparation to decision-oriented probabilistic outputs. The design separates point forecasting from uncertainty estimation so that the final output includes both expected values and risk ranges.

The general forecasting function is expressed as:

$$\hat{y}_{c,t} = F(X_{c,t}, L_{c,t}, S_t, C_c) \quad (1)$$

where $\hat{y}_{c,t}$ is the predicted event count for city c at week t , $X_{c,t}$ contains environmental predictors, $L_{c,t}$ contains lagged and rolling case-history features, S_t represents seasonality, and C_c represents city effects.

The ensemble point forecast combines the outputs of three complementary models:

$$\hat{\mu}_{c,t} = w_1 \hat{y}_{c,t}^{(1)} + w_2 \hat{y}_{c,t}^{(2)} + w_3 \hat{y}_{c,t}^{(3)}, \quad \sum_{m=1}^3 w_m = 1 \quad (2)$$

where $\hat{\mu}_{c,t}$ is the hybrid expected count and w_m denotes model weights. In this empirical implementation, log-ridge regression, random forest, and gradient boosting are combined because they represent linear-regularized, bagging-based, and boosting-based forecasting perspectives.

The probabilistic layer assumes an overdispersed count distribution around the hybrid mean:

$$Y_{c,t} \sim NB(\mu_{c,t}, \alpha_c), \quad Var(Y_{c,t}) = \mu_{c,t} + \alpha_c \mu_{c,t}^2 \quad (3)$$

where α_c is a city-specific dispersion parameter. Prediction intervals are obtained from the fitted predictive distribution:

$$PI_{q_1, q_2}(Y_{c,t}) = [Q_{NB}(q_1 | \mu_{c,t}, \alpha_c), Q_{NB}(q_2 | \mu_{c,t}, \alpha_c)] \quad (4)$$

The probability of a high-risk week is calculated as:

$$P_{risk} = P(Y_{c,t} > \tau_c | \mu_{c,t}, \alpha_c) = 1 - F_{NB}(\tau_c | \mu_{c,t}, \alpha_c) \quad (5)$$

where τ_c is the city-specific upper-quartile threshold from the historical training period. This probability output is central to the study because it transforms an ordinary forecast into a decision-support measure.

VI. STUDY CONTRIBUTION

The main contribution of this study is methodological and empirical. Methodologically, the study presents a forecasting framework that is simple enough to be reproduced, yet flexible enough to incorporate trend, seasonality, regression, ensemble learning, distribution fitting, and uncertainty intervals. Empirically, the study demonstrates the framework using real outbreak time-series data and reports both accuracy and uncertainty-calibration outcomes. In practice, the study shows how uncertain-event forecasting can support risk-sensitive decisions by reporting predicted case counts, plausible ranges, and the probabilities of high-risk weeks.

VII. METHODOLOGY

1. RESEARCH DESIGN

This study uses a secondary quantitative empirical design. The design is predictive, comparative, and probabilistic. It is predictive because it estimates future weekly event counts from historical observations. It is comparative because the proposed model is evaluated against baseline forecasting models. It is probabilistic because the final model produces prediction intervals and exceedance probabilities in addition to point forecasts.

2. DATA SOURCE AND EMPIRICAL CASE

The empirical data come from a public dengue forecasting dataset that predicts weekly dengue cases using environmental and temporal variables [31]. The observations represent weekly city-level records from San Juan and Iquitos. The target variable is the number of weekly dengue cases. The predictor variables include vegetation indices, precipitation measures, reanalysis temperature variables, station temperature variables, humidity measures, and calendar variables. The future test records were not used for validation because they do not contain observed case outcomes; they are only useful for formatting external prediction submissions.

Table 2. Dataset summary.

City	Observations	Period		Mean weekly cases	Median weekly cases	Standard deviation	Maximum weekly cases	Zero-case weeks
Iquitos	520	2000-07-01 to 2010-06-25		7.57	5	10.77	116	96
San Juan	936	1990-04-30 to 2008-04-22		34.18	19	51.38	461	4

Table 2. shows that the two cities differ substantially in outbreak intensity. San Juan has a higher mean and maximum weekly case count, while Iquitos has lower average counts but still shows clear outbreak variability.

3. SAMPLE SELECTION AND INCLUSION CRITERIA

All training records with matched city, year, and week identifiers were retained. The feature and label records matched exactly, producing 1,456 valid city-week observations. The inclusion criteria were: a valid city identifier, a valid epidemiological week, a matched observed case count, and a chronological position within the training period. The evaluation used an 80/20 chronological split within each city to avoid information leakage from future weeks into model training. This means that earlier weeks were used for model fitting, while later weeks were reserved for validation.

4. VARIABLES AND MEASURES

The dependent variable was the weekly total number of dengue cases. The independent variables included rainfall and precipitation indicators, reanalysis temperature indicators, station temperature

variables, vegetation indices, humidity indicators, week-of-year seasonality, city membership, and engineered lagged case-history variables. Lagged features included one-, two-, four-, eight-, twelve-, and fifty-two-week prior case counts. Rolling features included four-week, eight-week, and twelve-week rolling means and standard deviations. These variables were selected because uncertain event counts are usually affected by both immediate historical memory and broader seasonal/environmental conditions.

Table 3. Missing-value summary for predictors.

Variable	Missing percent
ndvi_ne	13.32
ndvi_nw	3.57
station_avg_temp_c	2.95
station_diur_temp_rng_c	2.95
ndvi_sw	1.51
station_precip_mm	1.51
ndvi_se	1.51
station_max_temp_c	1.37
station_min_temp_c	0.96
precipitation_amt_mm	0.89

Table 3 indicates that missingness is moderate and concentrated mainly in vegetation and selected station variables. The missing-value pattern is manageable and does not invalidate the dataset.

5. ANALYTICAL PROCEDURE

The analytical procedure followed seven steps. First, features and labels were merged by city, year, and week. Second, records were ordered chronologically within each city. Third, time-series features were generated to represent seasonality and case-history memory. Fourth, missing predictors were imputed inside the training pipeline. Fifth, baseline models and the proposed hybrid model were fitted using the training period. Sixth, point forecasts were evaluated on the chronological validation period. Seventh, the hybrid forecast was converted into a negative-binomial predictive distribution to estimate prediction intervals and outbreak probabilities.

The accuracy metrics were mean absolute error, root mean squared error, and symmetric mean absolute percentage error:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

$$sMAPE = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \quad (8)$$

Prediction interval coverage was calculated as the percentage of observed validation values falling inside the estimated interval:

$$Coverage = \frac{1}{n} \sum_{i=1}^n I(L_i \leq y_i \leq U_i) \times 100 \quad (9)$$

where L_i and U_i are lower and upper prediction bounds.

6. VALIDITY, RELIABILITY, AND ETHICS

Internal validity was strengthened by using chronological validation rather than random splitting. This is important because random splitting can leak future temporal information into the training process. Reliability was supported through a reproducible modeling pipeline, clearly defined feature construction,

and comparison with multiple baselines. Measurement validity was supported by the use of an established public forecasting dataset with clearly defined weekly case counts and environmental predictors. No human participants were contacted, and no personally identifiable information was used. Therefore, the study did not require direct participant consent or institutional human-subjects approval.

VIII. RESULTS

1. DESCRIPTIVE OUTBREAK PATTERN

The weekly dengue case series shows clear differences between the two empirical settings. San Juan has stronger outbreak peaks and greater variance, while Iquitos has lower counts but still shows unstable event behavior. This supports the decision to use a city-aware forecasting design rather than assuming a single common event process across both cities.

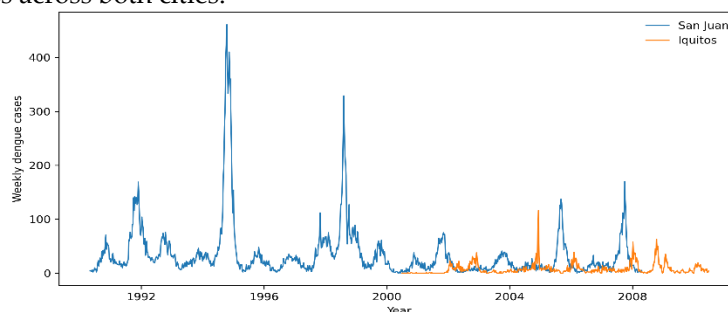


FIGURE 2. Observed weekly dengue cases by city.

FIGURE 3. shows that the target variable is not a smooth deterministic sequence. The case pattern includes seasonal variation, irregular peaks, and city-specific differences in intensity.

2. FORECASTING ACCURACY

Table 4. presents the validation performance of the proposed model and comparator models. The hybrid statistical-probabilistic model achieved the lowest mean absolute error and root mean squared error. The seasonal naive model performed the worst, indicating that simply repeating past seasonal values is insufficient for this dataset. The moving-average and log-ridge models improved on the naive baseline but did not match the nonlinear ensemble models. The hybrid model slightly improved over gradient boosting and random forest, suggesting that combining complementary forecast signals can improve stability.

Table 4. Model performance on chronological validation data.

Model	MAE	RMSE	sMAPE (%)
Hybrid statistical-probabilistic	5.908	9.874	47.12
Gradient boosting	6.335	10.36	46.71
Random forest	6.535	10.357	48.96
Log-ridge regression	10.018	18.858	64.47
Moving average	10.783	18.169	61.05
Seasonal naïve	19.277	33.335	90.97

Table 4. confirms H1: the proposed hybrid model produces the lowest validation error among all evaluated models. The improvement is strongest against simple baselines and moderate against individual ensemble learners.

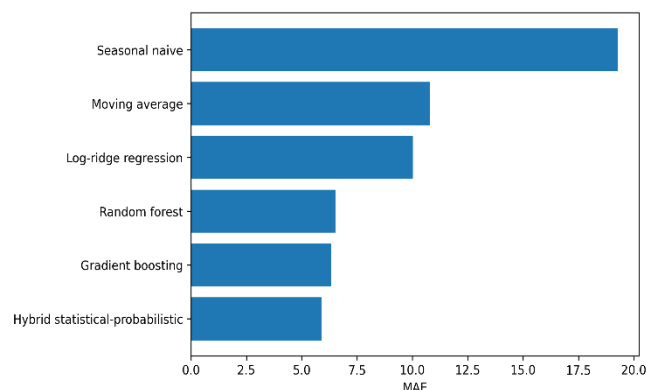


FIGURE 3. Validation forecasting error by model.

FIGURE 4. visualizes the same result using MAE. The hybrid model has the smallest error bar, while the seasonal naive model is the least reliable in this empirical case.

3. CITY-SPECIFIC VALIDATION PERFORMANCE

Because the two cities have different case distributions, the hybrid model was also evaluated separately by city. Table 5 shows that the model achieved lower absolute error in Iquitos because case counts were generally smaller, while San Juan remained more difficult due to larger outbreaks and sharper peaks.

Table 5. City-specific validation performance of the hybrid model.

City	Validation weeks	MAE	RMSE	sMAPE (%)	Mean observed cases	Mean predicted cases
Iquitos	104	4.273	6.908	62.77	9.52	8.43
San Juan	188	6.812	11.181	38.46	24.69	24.67

Table 5. shows that the model behaves differently across local outbreak regimes. This supports the city-aware interpretation of the forecasting task.

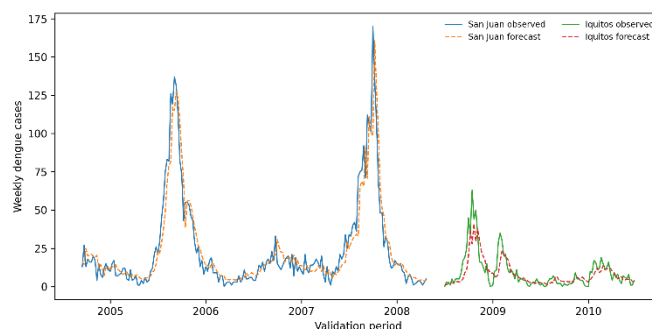


FIGURE 4. Hybrid model forecasts versus observed cases.

FIGURE 5. indicates that the hybrid model follows the general validation trajectory but still underestimates some outbreak spikes. This is expected in irregular count-event forecasting and motivates the probabilistic interval layer.

4. PROBABILISTIC FORECASTING AND UNCERTAINTY INTERVALS

The probabilistic layer converted the hybrid mean forecast into negative-binomial predictive intervals. Table 6 shows that the empirical coverage values exceeded the nominal levels. The 80% interval covered 87.67% of observed validation values, while the 90% interval covered 93.84%. This confirms H2, although the intervals are slightly conservative. In practical terms, conservative intervals may be acceptable for public-health preparedness because missing outbreak peaks can be more costly than issuing wider risk ranges.

Table 6. Prediction interval performance.

Nominal interval	Empirical coverage (%)	Mean interval width
80%	87.67	23.87
90%	93.84	30.83

Table 6 shows that the probabilistic layer provides reliable, slightly conservative prediction intervals. The model therefore provides more decision-relevant information than point forecasts alone.

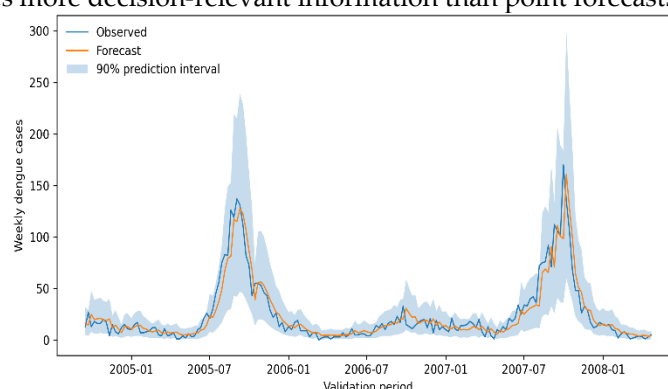


FIGURE 5. Probabilistic forecast interval for San Juan.

FIGURE 6. shows how the 90% interval surrounds the point forecast during the validation period. The wider interval during high-risk periods reflects uncertainty caused by outbreak volatility.

5. FEATURE IMPORTANCE AND OUTBREAK-RISK OUTPUT

Table 7. and Figure 7 summarize the strongest predictors. Recent case history was the dominant predictor, particularly the one-week lag and short rolling mean. Environmental variables such as precipitation and temperature contributed smaller but meaningful contextual information. This confirms H3: the recent disease trajectory carries the strongest predictive signal, while environmental indicators provide additional risk context.

Table 7. Most influential forecasting variables.

Feature	Importance
cases_lag_1	0.8799
cases_rol_mean_4	0.0357
cases_lag_2	0.0276
time_index	0.0094
cases_rol_std_12	0.0056
reanalysis_precip_amt_kg_per_m2	0.0044
station_diur_temp_rng_c	0.0025
cases_lag_12	0.0021
cases_rol_std_4	0.0019
reanalysis_min_air_temp_k	0.0018

Table 7 indicates that temporal memory is the primary forecasting signal. Environmental variables are still useful, but their effect is secondary in this empirical setup.

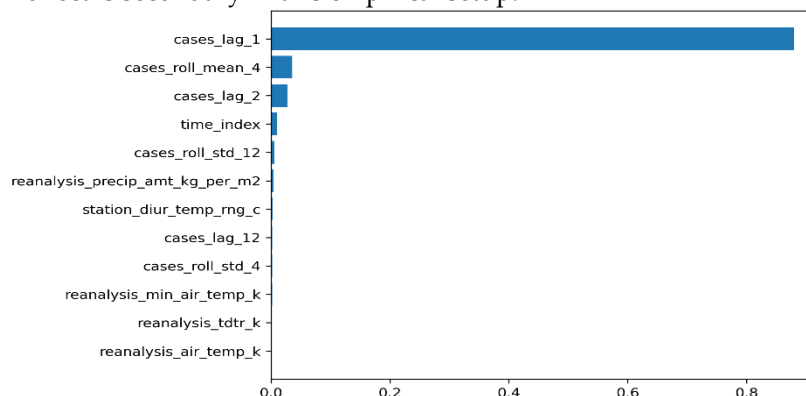


FIGURE 6. Most influential forecasting variables.

FIGURE 7. visually reinforces the dominance of recent case history. This does not mean climate variables are unimportant; rather, their predictive value appears conditional on the recent outbreak state.

The framework also produced probability-based high-risk outputs. Table 8 gives examples from validation weeks with high estimated probability of exceeding the city-specific upper-quartile threshold.

Table 8. Examples of high-risk probabilistic outputs.

city	year	Weekofyear	total_cases	hybrid_pred	prob_outbreak	p05	p95
Iquitos	2008	43	44	40.99	0.99	15	77
San Juan	2007	41	135	160.54	0.988	60	300
Iquitos	2008	45	35	37.46	0.986	13	71
Iquitos	2008	41	34	35.99	0.984	13	68
Iquitos	2008	46	16	30.78	0.973	11	59
San Juan	2007	42	106	128.3	0.973	48	240
San Juan	2005	36	131	127.57	0.972	47	239
Iquitos	2008	44	50	29.57	0.969	10	56

Table 8. demonstrates the decision-support value of the proposed method. Each forecast can be expressed as a probability of a high-risk occurrence rather than a single predicted count.

IX. DISCUSSION

The results show that the hybrid statistical-probabilistic model provides a useful balance between accuracy, interpretability, and uncertainty awareness. The hybrid model achieved the lowest validation MAE and RMSE, which supports the argument that uncertain events should be modeled using both statistical temporal structure and flexible regression components. The strong performance of lagged case variables is consistent with the logic of outbreak momentum: recent cases reflect current transmission conditions, reporting behavior, exposure to susceptibles, and short-term ecological dynamics. However, the presence of precipitation and temperature variables among the influential predictors suggests that environmental context still adds explanatory value.

The main empirical finding is not only that the hybrid model forecasts better than the naive baseline, but that it converts the forecast into a probabilistic decision output. This is important because a point estimate can hide uncertainty. For example, a predicted value of 40 cases may mean different things depending on whether the plausible range is 35 to 45 or 10 to 120. The probabilistic layer addresses this problem by estimating intervals and exceedance probabilities. In outbreak preparedness, interval and probability outputs can help decision-makers allocate surveillance, vector-control, clinical, and communication resources more cautiously.

The interval results were slightly conservative. The 80% interval covered 87.67% of observations, and the 90% interval covered 93.84%. This indicates that the fitted uncertainty layer did not underestimate risk. A conservative interval may be preferable in public-health contexts, but it also means that future refinements should attempt to improve sharpness without sacrificing coverage. Possible refinements include conformal calibration, quantile regression, Bayesian count models, dynamic dispersion parameters, and location-specific interval tuning.

The comparison among models also provides methodological insight. The seasonal naive baseline performed poorly, indicating that the previous year's seasonal value alone does not capture the irregularity of dengue events. The moving-average baseline was more stable but still unable to capture complex changes. The log-ridge model provided a transparent regression baseline but was less effective for nonlinear relationships and outbreak spikes. Random forest and gradient boosting performed well, confirming the value of nonlinear ensemble forecasting. The hybrid model outperformed these individual learners slightly, suggesting that combining multiple model families can reduce dependence on one modeling assumption.

The city-specific results demonstrate that local outbreak regimes matter. San Juan had higher outbreak intensity and larger peaks, making it more difficult to forecast in terms of absolute error. Iquitos had lower average counts, which produced lower absolute errors but does not necessarily mean the forecasting problem was easier in relative terms. This difference reinforces the importance of city-aware modeling. A pooled model without city effects may hide important heterogeneity, while fully separate models may lose the benefit of shared learning. The hybrid approach used here adopts a middle position by including city indicators and city-specific dispersion estimates.

Overall, the findings support the central claim of the manuscript: uncertain-event forecasting should combine statistical pattern recognition with probabilistic uncertainty estimation. Accuracy metrics remain necessary, but they are not sufficient. A model designed for uncertain events must also demonstrate that it can represent uncertainty in a way that is useful for decision-making. The proposed framework meets this requirement by producing point forecasts, prediction intervals, and risk probabilities from the same modeling pipeline.

X. PRACTICAL AND POLICY IMPLICATIONS

For public-health planners, the proposed framework can support early warning by identifying not only expected case counts but also the probability of high-risk weeks. This is useful for planning vector-control campaigns, assessing hospital readiness, evaluating laboratory capacity, informing public communication, and guiding surveillance intensity. For analysts, the framework offers a reproducible pipeline that can be adapted to other uncertain events such as climate extremes, market risk, equipment failures, emergency demand, and operational disruptions. For institutions, the results underscore the importance of maintaining consistent time-series records, environmental covariates, and historical event outcomes, as probabilistic forecasting depends heavily on data quality.

The framework should not be used as a standalone decision rule. Instead, it should complement expert judgment, epidemiological surveillance, and local operational knowledge. In practice, a high forecast probability should trigger closer review and preparedness action, while a wide interval should signal uncertainty that requires caution rather than overconfidence.

XI. LIMITATIONS AND FUTURE RESEARCH

This study has several limitations. First, the empirical evidence is based on two cities, which limits geographical generalization. Second, the target variable is weekly case count, but the dataset does not fully observe behavioral, intervention, vector-control, population-mobility, or surveillance-system changes that may influence outbreak dynamics. Third, the validation design evaluates one chronological holdout period rather than multiple rolling-origin validation windows. Fourth, the probabilistic layer uses a city-specific negative-binomial dispersion parameter, which is interpretable but may not capture time-varying uncertainty during unusual outbreak periods.

Future research should extend the framework to additional cities, longer time spans, and different domains of uncertain events. Methodological extensions may include conformal prediction, Bayesian hierarchical count models, zero-inflated models, dynamic dispersion, multi-step forecasting, and rolling-origin cross-validation. Future studies should also compare decision-oriented metrics such as cost-weighted errors, outbreak recall, risk-threshold calibration, and resource-allocation utility.

XII. CONCLUSION

This study developed and empirically evaluated a hybrid statistical-probabilistic model for forecasting uncertain events using weekly dengue incidence as an applied event-count case. The proposed framework combined city-week time-series structuring, lagged and rolling temporal features, environmental predictors, ensemble point forecasting, and a negative-binomial uncertainty layer. The validation evidence showed that the hybrid model achieved the strongest overall forecasting performance, with MAE = 5.908 and RMSE = 9.874. The model clearly outperformed simple seasonal and moving-average baselines and produced modest but meaningful gains over individual ensemble learners, indicating that a combined modeling strategy can improve stability when forecasting irregular outbreak counts.

Beyond point accuracy, the framework's central value lies in its probabilistic output. The negative-binomial layer produced reliable and slightly conservative prediction intervals, with empirical coverage of 87.67% for the nominal 80% interval and 93.84% for the nominal 90% interval. These results demonstrate that the model can express forecast uncertainty in a form that is directly useful for risk-sensitive planning. The feature-importance findings further showed that recent case history was the dominant forecasting signal, while precipitation, temperature, humidity, and time-index variables added contextual information. This confirms that uncertain-event forecasting benefits from combining short-term temporal memory with environmental and seasonal structure.

The study contributes to forecasting research by presenting an interpretable framework that integrates accuracy evaluation, interval reliability, and exceedance-risk estimation into a single reproducible pipeline. It also contributes to applied outbreak analytics by showing how weekly case forecasts can be translated into prediction intervals and probabilities of crossing city-specific high-risk thresholds. Nevertheless, the findings should be interpreted within the limits of a two-city empirical case study and a single chronological validation design. Future research should evaluate the framework across broader geographic areas, longer time horizons, rolling-origin validation schemes, and alternative event domains. Overall, the results support the use of hybrid statistical-probabilistic models when forecasts must guide decisions under uncertainty rather than merely report expected future values.

Author Contributions

The author conducted the conceptualization, methodology, data analysis, investigation, writing, review, editing, and final approval of the manuscript.

Funding

This research received no external funding.

Data Availability

The dataset will be available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

REFERENCES

1. Al Mobin, M., et al. (2024). Forecasting dengue in Bangladesh using meteorological parameters and machine learning models. *Scientific Reports*, 14, Article 31136. doi: 10.1038/s41598-024-83770-0
2. Baharom, M., Ahmad, N., & Hod, R. (2022). Dengue early warning system as outbreak prediction tool: A systematic review. *Risk Management and Healthcare Policy*, 15, 871-886. doi: 10.2147/RMHP.S361106

3. Barboza, L. A., et al. (2022). Assessing dengue fever risk in Costa Rica by using climate variables and machine learning techniques. arXiv. doi: 10.48550/arXiv.2204.01483
4. Ben Taieb, S., & Hyndman, R. J. (2014). A gradient boosting approach to the Kaggle load forecasting competition. *International Journal of Forecasting*, 30(2), 382-394. doi: 10.1016/j.ijforecast.2013.07.005
5. Benedum, C. M., et al. (2020). Weekly dengue forecasts in Iquitos, Peru; San Juan, Puerto Rico; and Singapore. *PLOS Neglected Tropical Diseases*, 14(10), Article e0008710. doi: 10.1371/journal.pntd.0008710
6. Buczak, A. L., et al. (2014). Prediction of high incidence of dengue in the Philippines. *PLOS Neglected Tropical Diseases*, 8(4), Article e2771. doi: 10.1371/journal.pntd.0002771
7. Chen, Y., et al. (2018). Neighbourhood level real-time forecasting of dengue cases in tropical urban Singapore. *BMC Medicine*, 16, Article 129. doi: 10.1186/s12916-018-1103-6
8. Chuang, T.-W., Chaves, L. F., & Chen, P.-J. (2017). Effects of local and regional climatic fluctuations on dengue outbreaks in southern Taiwan. *PLOS ONE*, 12(6), Article e0178698. doi: 10.1371/journal.pone.0178698
9. da Silva, S. T., et al. (2024). When climate variables improve dengue forecasting: A machine learning approach. *European Physical Journal Special Topics*. doi: 10.1140/epjs/s11734-024-01201-7
10. Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359-378. doi: 10.1198/016214506000001437
11. Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2), 243-268. doi: 10.1111/j.1467-9868.2007.00587.x
12. Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3), 1-22. doi: 10.18637/jss.v027.i03
13. Johansson, M. A., Reich, N. G., Hota, A., Brownstein, J. S., & Santillana, M. (2016). Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico. *Scientific Reports*, 6, Article 33707. doi: 10.1038/srep33707
14. Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011). Lower upper bound estimation method for construction of neural network-based prediction intervals. *IEEE Transactions on Neural Networks*, 22(3), 337-346. doi: 10.1109/TNN.2010.2096824
15. Leung, X. Y., et al. (2023). A systematic review of dengue outbreak prediction models: Current scenario and future directions. *PLOS Neglected Tropical Diseases*, 17(2), Article e0010631. doi: 10.1371/journal.pntd.0010631
16. Lippi, C. A., et al. (2019). Investigating the utility of satellite-based rainfall estimates for dengue early warning in Mexico. *PLOS Neglected Tropical Diseases*, 13(11), Article e0007564. doi: 10.1371/journal.pntd.0007564
17. Lowe, R., et al. (2018). Nonlinear and delayed impacts of climate on dengue risk in Barbados: A modelling study. *PLOS Medicine*, 15(7), Article e1002613. doi: 10.1371/journal.pmed.1002613
18. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1), 54-74. doi: 10.1016/j.ijforecast.2019.04.014
19. Marques-Toledo, C. de A., et al. (2017). Dengue prediction via the web: Tweets are a useful tool for estimating and forecasting dengue at the country and city levels. *PLOS Neglected Tropical Diseases*, 11(7), Article e0005729. doi: 10.1371/journal.pntd.0005729
20. Petropoulos, F., et al. (2022). Forecasting: Theory and practice. *International Journal of Forecasting*, 38(3), 705-871. doi: 10.1016/j.ijforecast.2021.11.001
21. Racloz, V., Ramsey, R., Tong, S., & Hu, W. (2012). Surveillance of dengue fever virus: A review of epidemiological models and early warning systems. *PLOS Neglected Tropical Diseases*, 6(5), Article e1648. doi: 10.1371/journal.pntd.0001648
22. Ramosaj, B., Schultheis, J. S., & Pauly, M. (2021). Constructing valid prediction intervals with random forests. arXiv. doi: 10.48550/arXiv.2103.05766
23. Roster, K., Connaughton, C., & Rodrigues, F. A. (2022). Machine-learning-based forecasting of dengue fever in Brazilian cities using epidemiologic and meteorological variables. *American Journal of Epidemiology*, 191(10), 1803-1812. doi: 10.1093/aje/kwac090
24. Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181-1191. doi: 10.1016/j.ijforecast.2019.07.001
25. Sylvestre, E., et al. (2022). Data-driven methods for dengue prediction and surveillance using real-world and Big Data: A systematic review. *PLOS Neglected Tropical Diseases*, 16(1), Article e0010056. doi: 10.1371/journal.pntd.0010056
26. Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37-45. doi: 10.1080/00031305.2017.1380080
27. Tian, Q., Nordman, D. J., & Meeker, W. Q. (2022). Methods to compute prediction intervals: A review and new results. *Statistical Science*, 37(4), 580-597. doi: 10.1214/21-STS842
28. Wikle, C. K., Zammit-Mangion, A., & Cressie, N. (2019). *Spatio-temporal statistics with R*. CRC Press.

-
29. Xu, H., Mei, S., Bates, S., Taylor, J., & Tibshirani, R. (2023). Uncertainty intervals for prediction errors in time series forecasting. arXiv. doi: 10.48550/arXiv.2309.07435
30. Yoon, T., Park, Y., Ryu, E. K., & Wang, Y. (2022). Robust probabilistic time series forecasting. arXiv. doi: 10.48550/arXiv.2202.11910
31. DrivenData. (2016). DengAI: Predicting disease spread [Data set]. DrivenData.