

Spatio-Temporal Dynamics of *Aedes* Mosquito Populations under Climatic Variability in Bhagalpur, Bihar: A Hierarchical Statistical Modelling Approach

Pratyasha Tripathi

Tilka Manjhi Bhagalpur University, Bhagalpur, Bihar, India.

pratyashat25@gmail.com.

Abstract

Populations of *Aedes aegypti* and *Aedes albopictus* are shaped by far more than weather alone. Their abundance reflects the combined influence of mosquito development and survival, availability of breeding habitats, characteristics of the built environment, human-modified surroundings, and climatic conditions. These factors operate at different spatial and temporal scales and often interact with one another. As a result, analyses based only on seasonal averages, simple pairwise correlations, or independent Poisson regression may provide an incomplete or overly precise picture of climate–mosquito relationships, particularly when mosquito counts are over dispersed and the same surveillance locations are observed repeatedly. We therefore propose a pilot statistical framework for prospective *Aedes* surveillance in Bhagalpur, Bihar, in which climatic exposures may act nonlinearly or after a time lag, while residual spatial and temporal variation is explicitly represented in the model.

The proposed field design is based on weekly mosquito surveillance at geographically distributed sentinel sites using a standardized adult-trap or ovitrap protocol. Mosquito counts are recorded together with temperature, rainfall, relative humidity, sampling effort, and selected local environmental characteristics. The primary analytical approach uses a hierarchical negative-binomial model with a log link and an offset to account for variation in sampling effort. Climatic effects may be represented through quadratic or smooth functions, and lagged exposures are included where biologically plausible. Residual differences among surveillance sites are represented through a correlated Gaussian spatial effect, while temporal dependence is modelled using an autoregressive process. The statistical formulation includes the relevant probability mass functions, probability density functions, and cumulative distribution functions. A zero-inflated extension is considered only when the observed field data provide evidence for a distinct structural-zero process. The main inferential questions concern the overall contribution of climate, nonlinear temperature effects, delayed rainfall effects, residual spatial heterogeneity, and temporal persistence.

Because a standardized longitudinal *Aedes* surveillance series is not yet available for Bhagalpur, the statistical framework was evaluated using Monte Carlo simulation rather than presenting simulated observations as evidence about local mosquito ecology. Each simulated data set represented 16 sentinel sites followed weekly for 52 weeks. Counts were generated from an NB2 process incorporating extra-Poisson variation, correlated site-level spatial effects, AR(1) temporal dependence, and prespecified climatic effects. The number of simulation replications was chosen to provide adequate Monte Carlo precision. In total, 10,000 data sets were generated under non-null climatic effects and a further 10,000 under a climatic-null scenario, giving 20,000 simulated data sets. Three analytical strategies were compared: an independent Poisson generalized linear model, an independent NB2 generalized linear model, and an NB2 model including site fixed effects with site-cluster-robust inference.

In a representative simulated surveillance year, the mean mosquito count was 5.15 per site-week, while the variance was 30.86, corresponding to a variance-to-mean ratio of 5.99. Zero counts occurred in 11.8% of site-weeks. Across the 10,000 non-null simulations, all three approaches recovered the prespecified climatic coefficients with little average bias, but their estimates of uncertainty differed substantially. The Poisson model produced only 64.1% coverage for the linear temperature effect, 64.8% for the quadratic temperature effect, and 45.0% for lagged rainfall, despite the nominal 95% confidence level. Under the climatic-null simulations, the corresponding Type-I error rates were 36.9%, 33.7%, and 34.5%, respectively. These results show that ignoring overdispersion and dependence can lead to severely anti-conservative inference.

The independent NB2 model performed considerably better, with coverage ranging from 89.5% to 91.7% and Type-I error rates between 8.9% and 11.0%. This indicates that overdispersion accounted for an important part of the inferential distortion, although it did not fully resolve the problem. When persistent site-level heterogeneity was additionally controlled and site-cluster-robust inference was used, coverage improved to 94.2–95.2%, while Type-I error rates fell to 5.0–5.8%, closely matching the nominal level. With 10,000 simulation replications, the Monte Carlo standard error around a true 5% rejection probability was approximately 0.0022, indicating that these simulation estimates were obtained with good numerical precision.

The simulation results support a more focused methodological conclusion. In longitudinal *Aedes* surveillance, modelling the count variance correctly is essential, but it may not be sufficient when repeated observations from the same locations remain correlated and persistent site differences are present. A hierarchical spatio-temporal framework is therefore warranted not merely because it is more flexible, but because the principal inferential targets such as climatic effects, lag-response

relationships, and hotspot probabilities depend on having well-calibrated uncertainty estimates. Importantly, the climatic coefficients used in the simulation are design values and should not be interpreted as estimates for Bhagalpur. Prospective field surveillance is still required before any substantive ecological interpretation, local risk estimate, or early-warning claim can be made.

Keywords: *Aedes aegypti*; *Aedes albopictus*; mosquito abundance; negative binomial regression; spatio-temporal modelling; climate; rainfall; temperature; hierarchical model; Monte Carlo simulation; Bihar.

How to cite this article: Pratyasha Tripathi. Spatio-Temporal Dynamics of Aedes Mosquito Populations under Climatic Variability in Bhagalpur, Bihar: A Hierarchical Statistical Modelling Approach. Int J Drug Deliv Technol. 2026;16(85s):1989-2000. DOI: 10.25258/ijddt.16.85s.197.

Introduction

Mosquito surveillance may appear simple when reduced to a spreadsheet: traps are placed, mosquitoes are counted, and the numbers are entered for analysis. In reality, the process that produces those numbers is much more complex. The number of adult *Aedes* mosquitoes observed at a particular place and time is influenced by many biological and environmental processes, including egg laying, larval development, emergence, survival, host-seeking behaviour, the availability of breeding sites, and the likelihood that a trap actually captures mosquitoes that are present in the area. Weather can influence several of these processes, but its effects are often delayed and rarely follow a simple linear pattern. At the same time, nearby locations may share environmental conditions that are not directly measured, and repeated observations from the same site are naturally related because mosquito populations carry biological continuity from one week to the next. If these features are ignored, a statistical model may look precise while still giving misleading conclusions.

This issue is important because *Aedes aegypti* and *Aedes albopictus* are major vectors of dengue, chikungunya and Zika viruses. Their population dynamics are strongly shaped by local conditions. The same amount of rainfall, for example, may have very different effects in neighbourhoods that differ in household water-storage practices, vegetation, drainage, shade, container availability or vector-control activity. In rural southern India, Tewari et al. (2004) reported year-round breeding of *Ae. aegypti*, whereas *Ae. albopictus* showed a more seasonal pattern, with small water-holding containers acting as important persistent breeding habitats. A later longitudinal study from Bengaluru also demonstrated marked local variation in *Aedes* larval ecology across different urban settings, again highlighting the importance of artificial containers and fine-scale environmental structure (Dharmamuthuraja et al., 2023). These findings suggest that regional weather variables should not be assumed to act in the same way at every sampling location.

Time introduces another important layer of complexity. Using a 109-week *Ae. aegypti* time series, Simões et al. (2013) showed that mosquito population dynamics were influenced not only by climatic variables but also by time-lagged effects, interactions and information contained in previous mosquito abundance. Kache et al. (2020), analysing data from 338 trap sites, similarly found that estimated environmental effects on *Ae. albopictus*

depended on the time window over which precipitation and related covariates were summarized. Temperature effects are also unlikely to be linear across the full climatic range. Mechanistic work on *Aedes*-borne virus transmission has demonstrated unimodal thermal relationships rather than a simple increase in transmission potential with each degree of warming (Mordecai et al., 2017). Although transmission suitability and mosquito abundance are not identical outcomes, the broader lesson is relevant: the shape of a temperature response should be estimated from the data rather than imposed for convenience.

Spatial variation is equally important. Villela et al. (2015) used a Bayesian hierarchical model to estimate *Ae. aegypti* abundance and spatial density while carrying uncertainty through to the resulting hotspot patterns. Yang et al. (2021) extended this type of approach in a large *Aedes* surveillance data set by combining mixed effects with a zero-inflated negative-binomial model to account for occurrence, abundance, climate, host density and interspecific structure. Evidence from India also points to substantial geographical variation in mosquito dynamics. Whittaker et al. (2022) identified distinct seasonal patterns across Indian mosquito time series and related them to rainfall, temperature, standing water and land use. Although their analysis focused mainly on anopheline mosquitoes, the statistical implication is more general: mosquito populations sharing the same broad climatic region should not automatically be treated as statistically identical.

Recent modelling approaches have increasingly tried to bring these elements together by allowing climatic effects to be both nonlinear and delayed within explicitly spatial and temporal models. Ortega-Lenis et al. (2024), for example, analysed dengue risk using delayed and nonlinear effects of temperature, rainfall and humidity within a Bayesian spatio-temporal framework. Distributed lag nonlinear models, introduced by Gasparrini, Armstrong and Kenward (2010), provide a general way to describe both the shape of an exposure-response relationship and how that relationship unfolds over time. This is particularly relevant for mosquito ecology. Rainfall in one week may create breeding habitats whose influence on adult abundance becomes visible only after larval development, whereas very heavy rainfall may wash out breeding containers or reduce larval survival. The more informative question, therefore, is not simply whether rainfall is

correlated with mosquito counts, but how mosquito abundance changes across a biologically meaningful lag period after accounting for other spatial and temporal sources of variation.

Bhagalpur, in eastern Bihar, experiences a clear monsoon cycle and provides a useful setting for developing such a framework. In this study, climate is not assumed to be the single or dominant explanation for local *Aedes* variation. Rather, it is treated as one component of a broader ecological system that also includes relatively stable site characteristics, changing local environmental conditions, and unobserved spatial and temporal processes. The pilot is therefore prospective as well as analytical. The statistical framework is specified before large-scale surveillance begins so that future field data are collected at the spatial and temporal resolution required to answer the intended scientific questions.

This distinction is important because the present work is a methodological pilot rather than a completed ecological survey of Bhagalpur. The climatic coefficients used in the simulation are not estimates of effects already observed in the district and should not be interpreted as local evidence. The role of simulation is different. It allows us to ask what happens when mosquito count data truly contain overdispersion, site-to-site heterogeneity and temporal dependence. Under such conditions, how much can commonly used count models underestimate uncertainty or generate false evidence of climatic effects? Addressing this question before launching a full surveillance programme is valuable because it helps identify modelling approaches that may appear convenient but are statistically unsafe.

The study is therefore organized around four connected aims. First, it proposes a practical weekly sentinel-surveillance design for *Aedes* mosquitoes in Bhagalpur. Second, it develops a hierarchical negative-binomial model in which climatic exposures may act with delay or follow nonlinear patterns, while residual spatial and temporal dependence are modelled explicitly. Third, the probability distributions, likelihood structure and inferential targets are described in sufficient detail to make the analysis reproducible. Finally, a Monte Carlo simulation is used to examine what is lost when overdispersion or dependence is ignored and to assess whether a one-year pilot could contain enough information to recover effects of a realistic statistical magnitude.

1. Materials and Methods

The pilot is designed around 16 sentinel locations sampled once in every epidemiological week for 52 consecutive weeks, giving at most 832 site-week observations. This is large enough to reveal basic distributional and seasonal features while still being realistic for a team that must maintain consistent trapping and environmental recording. The 16 sites should not be chosen simply because they are easy to reach. A stratified design should deliberately represent contrasts in urban density, peri-urban transition,

vegetation, drainage, household water storage, proximity to markets or transport corridors and the abundance of artificial water-holding containers. Coordinates should be recorded once the sites are fixed and retained as continuous spatial data rather than immediately collapsed into administrative categories.

Adult female *Aedes* counts per standardized trap-night or trap-day are the preferred primary outcome, with *Ae. aegypti* and *Ae. albopictus* identified separately whenever laboratory capacity allows. If the project instead relies on ovitraps or larval surveys, the response definition must remain consistent over time and sampling effort must be measured explicitly. A pooled *Aedes* outcome may be useful for routine surveillance, but species-specific models are scientifically more informative because the two species need not respond to the same environmental conditions in the same way. Trap type, deployment duration, failed collections and malfunction should all be recorded. Most importantly, a week in which no trap was operated must remain missing; it must never be coded as a biological zero. Weekly mean temperature, accumulated rainfall and mean relative humidity form the core time-varying exposures. Daily observations can be drawn from an official meteorological source and, where resources permit, supplemented with site-level sensors to capture microclimate. Raw measurements should be retained alongside prespecified lagged versions. A 0–4-week window is a reasonable starting point for the pilot, although the final lag range should be justified biologically and fixed before confirmatory analysis. Site information should also describe factors that may change slowly or remain fairly stable, such as urbanization, vegetation, container density, drainage, household water storage and local control activity. These measurements will never capture the entire mosquito habitat, but they can prevent obvious site differences from being absorbed incorrectly into a climatic coefficient.

Sampling effort should enter the model explicitly. Let E_{it} denote the effective trapping effort for site i in week t . If one trap operated for two nights and another for one night, the expected counts should not be compared as if exposure were identical. For a log-link count model, effort is included as an offset, $\log(E_{it})$, so that the remaining model describes a rate per unit of sampling effort. When effort is truly identical for all observations, the offset is constant and can be omitted without changing estimated contrasts. Quality control comes before hypothesis testing. For every site, the weekly count series should be plotted and checked for gaps, abrupt changes and implausible values. The zero fraction, mean, variance and variance-to-mean ratio should be summarized, meteorological measurements should be screened for impossible or inconsistent values, and missing observations should be described by site and season. The first exploratory issue is whether an equidispersed Poisson model is even plausible. A variance much larger than the mean is a warning sign rather than a formal model-selection rule; the final judgement should come from fitted-model diagnostics and predictive checks.

The design must also distinguish technical repetition from biological replication. Several observations from the same

location are not several independent sites, and several traps placed a few metres apart during the same week should not automatically be counted as separate biological replicates. If multiple traps are used to improve detection, they can be represented as nested observations or combined into a clearly defined site-week outcome while preserving total trapping effort. Making this distinction at the design stage avoids pseudoreplication later in the analysis. The principal hypotheses are prespecified as follows. For the vector of climatic regression parameters β_C , the global climate hypothesis is

$$H_{01}: \beta_C = \mathbf{0} \quad \text{versus} \quad H_{11}: \beta_C \neq \mathbf{0}.$$

For a quadratic temperature representation with coefficients β_T and β_{T^2} , nonlinearity is assessed through

$$H_{02}: \beta_{T^2} = 0 \quad \text{versus} \quad H_{12}: \beta_{T^2} \neq 0.$$

For a prespecified lag- ℓ rainfall coefficient $\beta_{R,\ell}$,

$$H_{03}: \beta_{R,\ell} = 0 \quad \text{versus} \quad H_{13}: \beta_{R,\ell} \neq 0.$$

The spatial and temporal components are assessed through the variance and correlation parameters,

$$H_{04}: \sigma_s^2 = 0 \quad \text{versus} \quad H_{14}: \sigma_s^2 > 0,$$

and

$$H_{05}: \rho = 0 \quad \text{versus} \quad H_{15}: \rho \neq 0.$$

Because a variance parameter lies on the boundary of its parameter space under H_{04} , ordinary Wald tests are not ideal. In a Bayesian implementation, posterior mass away from zero, posterior predictive diagnostics and predictive improvement are preferable; in a likelihood implementation, restricted likelihood-ratio procedures or parametric bootstrap methods can be used. Climatic hypotheses should likewise be interpreted with confidence or credible intervals and estimated effect sizes rather than with a binary significance label alone.

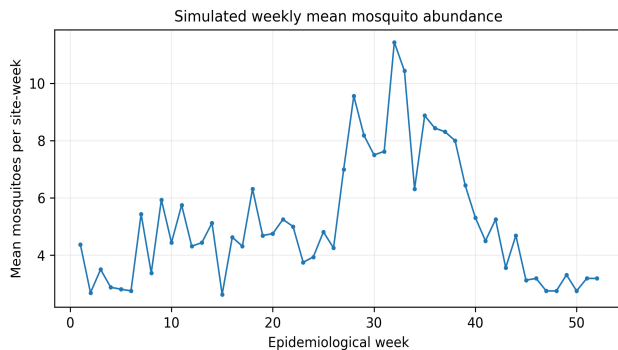


Figure 1. Simulated weekly mean mosquito abundance across the 16-site pilot network. This is an illustrative trajectory from the prespecified data-generating mechanism and is not an observed Bhagalpur time series.

Statistical Framework

Let Y_{it} be the mosquito count at site $i = 1, \dots, m$ and week $t = 1, \dots, T$. The simplest count model is Poisson. Conditional on an expected count $\mu_{it} > 0$,

$$Y_{it} | \mu_{it} \sim \text{Poisson}(\mu_{it}).$$

Its probability mass function (PMF) is

$$P(Y = y | \mu) = \frac{e^{-\mu} \mu^y}{y!}, \quad y = 0, 1, 2, \dots,$$

and its cumulative distribution function (CDF) is

$$F_Y(y | \mu) = P(Y \leq y) = e^{-\mu} \sum_{k=0}^{\lfloor y \rfloor} \frac{\mu^k}{k!}, \quad y \geq 0,$$

with $F_Y(y) = 0$ for $y < 0$. The Poisson distribution imposes

$$E(Y | \mu) = \text{Var}(Y | \mu) = \mu.$$

This equality is often too restrictive for mosquito counts. The proposed response distribution is therefore negative binomial in the NB2 parameterization. Write $Y \sim \text{NB}(r, p)$, where $r > 0$ is a size parameter and $0 < p < 1$. The PMF for the number y of failures before r successes is

$$P(Y = y | r, p) = \frac{\Gamma(y+r)}{\Gamma(r)\Gamma(y+1)} p^r (1-p)^y, \quad y = 0, 1, 2, \dots$$

The corresponding CDF is

$$F_Y(y | r, p) = I_p(r, \lfloor y \rfloor + 1),$$

where $I_p(a, b)$ is the regularized incomplete beta function. To express the distribution directly through the conditional mean μ and overdispersion parameter $\alpha > 0$, set

$$r = \frac{1}{\alpha}, \quad p = \frac{r}{r + \mu} = \frac{1}{1 + \alpha\mu}.$$

Then

$$E(Y | \mu, \alpha) = \mu,$$

and

$$\text{Var}(Y | \mu, \alpha) = \mu + \alpha\mu^2.$$

As $\alpha \rightarrow 0$, the NB2 variance approaches the Poisson variance and the negative-binomial model approaches the Poisson model. The parameter α therefore has a direct inferential role: it quantifies extra-Poisson variation after measured covariates and latent processes have been included.

The conditional mean is linked to predictors by

$$\log(\mu_{it}) = \log(E_{it}) + \eta_{it},$$

where E_{it} is sampling effort and η_{it} is the structured linear predictor. For an initial pilot model,

$$\eta_{it} = \beta_0 + \beta_1 ZT_{it} + \beta_2 Q_{it} + \beta_3 ZR_{i,t-2} + \beta_4 ZH_{i,t-1} + \mathbf{y}^T \mathbf{X}_i + b_i + u_{it}.$$

Here ZT_{it} is standardized temperature, $ZR_{i,t-2}$ is standardized rainfall lagged by two weeks, $ZH_{i,t-1}$ is standardized humidity lagged by one week and $\mathbf{X}_i$ is a vector of measured site characteristics. The quadratic temperature component is centered as

$$Q_{it} = ZT_{it}^2 - \overline{ZT^2},$$

which reduces unnecessary correlation between the intercept and quadratic term. The particular one- and two-week lags above are candidate terms used in the pilot simulation; a real analysis should compare a biologically prespecified lag set rather than assume those exact lags are true for Bhagalpur.

The exponentiated coefficient $\exp(\beta_j)$ is an incidence-rate ratio for a one-unit increase in a linear standardized predictor when interactions and nonlinear terms do not modify that contrast. With a quadratic term, the instantaneous temperature effect is not constant. The derivative on the log-mean scale is

$$\frac{\partial \log \mu_{it}}{\partial ZT_{it}} = \beta_1 + 2\beta_2 ZT_{it},$$

so the direction and magnitude of the temperature association can change across the observed range. Reporting a single percentage change per degree is therefore inappropriate when β_2 is materially different from zero.

The spatial site effect is represented as a multivariate Gaussian vector,

$$\mathbf{b} = (b_1, \dots, b_m)^T \sim N_m(\mathbf{0}, \Sigma_s),$$

with exponential covariance

$$(\Sigma_s)_{ij} = \sigma_s^2 \exp\left(-\frac{d_{ij}}{\phi}\right),$$

where d_{ij} is the distance between sites, σ_s^2 is spatial variance and $\phi > 0$ is a range parameter controlling the rate at which correlation decays with distance. For a q -dimensional Gaussian random vector $\mathbf{X} \sim N_q(\boldsymbol{\mu}, \Sigma)$, the probability density function (PDF) is

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{q/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}.$$

Its multivariate CDF is defined by

$$F_{\mathbf{X}}(\mathbf{x}) = P(X_1 \leq x_1, \dots, X_q \leq x_q) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_q} f_{\mathbf{X}}(\mathbf{z}) d\mathbf{z},$$

for which no general elementary closed form exists. Numerical integration or specialized algorithms are therefore used when multivariate Gaussian probabilities are required.

The univariate Gaussian distribution is used repeatedly for latent innovations and priors. If $X \sim N(\mu, \sigma^2)$, its PDF is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right],$$

and its CDF is

$$F_X(x) = \Phi\left(\frac{x - \mu}{\sigma}\right),$$

where

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{v^2}{2}\right) dv.$$

For completeness, the positive rainfall multiplier used in the simulation is Gamma distributed. Under the shape-scale parameterization, if $G \sim \text{Gamma}(a, \theta)$ with shape $a > 0$ and scale $\theta > 0$, its PDF is

$$f_G(g) = \frac{g^{a-1} \exp(-g/\theta)}{\Gamma(a) \theta^a}, \quad g > 0,$$

and its CDF is

$$F_G(g) = P(G \leq g) = \frac{\gamma(a, g/\theta)}{\Gamma(a)}, \quad g \geq 0,$$

where $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ is the complete gamma function and $\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$ is the lower incomplete gamma function. The simulation used $a = 2.5$ and $\theta = 0.4$ only to generate positive week-to-week rainfall variation around a monsoon-like seasonal envelope; these parameters are not climatic estimates for Bhagalpur.

Temporal dependence may be represented by a first-order autoregressive process,

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \stackrel{\text{ind}}{\sim} N(0, \sigma_\varepsilon^2), \quad |\rho| < 1.$$

Conditionally,

$$u_{it} | u_{i,t-1} \sim N(\rho u_{i,t-1}, \sigma_\varepsilon^2),$$

so its transition PDF is

$$f(u_{it} | u_{i,t-1}) = \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left[-\frac{(u_{it} - \rho u_{i,t-1})^2}{2\sigma_\varepsilon^2}\right],$$

and the corresponding conditional CDF is

$$P(U_{it} \leq u | u_{i,t-1}) = \Phi\left(\frac{u - \rho u_{i,t-1}}{\sigma_\varepsilon}\right).$$

Under stationarity, $u_{i1} \sim N\{0, \sigma_\varepsilon^2/(1 - \rho^2)\}$. Positive ρ implies persistence: an unobserved positive departure from expected abundance tends to remain positive in the immediately following week.

For observed counts $\mathbf{y}$, latent effects $\mathbf{b}$ and $\mathbf{u}$, and parameter vector $\boldsymbol{\theta}$, the conditional likelihood is

$$L(\boldsymbol{\theta}; \mathbf{y} | \mathbf{b}, \mathbf{u}) = \prod_{i=1}^m \prod_{t=1}^T P\{Y_{it} = y_{it} | \mu_{it}(\boldsymbol{\theta}, b_i, u_{it}), \alpha\},$$

where the probability term is the NB2 PMF defined above. A full hierarchical likelihood additionally multiplies this conditional likelihood by the multivariate Gaussian density of $\mathbf{b}$ and the AR(1) transition densities of $\mathbf{u}$. In a Bayesian formulation,

$$\pi(\boldsymbol{\theta}, \mathbf{b}, \mathbf{u} | \mathbf{y}) \propto L(\boldsymbol{\theta}; \mathbf{y} | \mathbf{b}, \mathbf{u}) \pi(\mathbf{b} | \boldsymbol{\theta}) \pi(\mathbf{u} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}).$$

Because predictors are standardized, weakly informative Gaussian priors can be placed on fixed effects, for example $\beta_j \sim N(0, 2.5^2)$. Positive parameters can be modelled on the log scale, such as $\log \alpha \sim N(0, 1)$, $\log \sigma_s \sim N(-1, 1)$ and $\log \sigma_\varepsilon \sim N(-1, 1)$; the spatial range can likewise be assigned a Gaussian prior on $\log \phi$. Temporal correlation can be transformed by $z_\rho = \text{atanh}(\rho)$ with a Gaussian prior. Since each prior is Gaussian on the stated transformed scale, its PDF and CDF are given by the normal expressions defined above. Sensitivity analysis should examine whether posterior conclusions depend materially on plausible alternative prior scales.

As the surveillance network grows, direct manipulation of a dense spatial covariance matrix becomes less attractive computationally. A Matérn Gaussian field represented through the stochastic partial differential equation (SPDE) formulation of Lindgren, Rue and Lindström (2011) offers an efficient alternative. Together with integrated nested Laplace approximation (INLA), introduced for latent Gaussian models by Rue, Martino and Chopin (2009), this approach can fit many non-Gaussian spatio-temporal models without relying on conventional Markov-chain Monte Carlo. Such machinery is unnecessary for a 16-site pilot, which can be fitted directly, but it becomes valuable if the project later expands to a much denser geocoded network.

If the real surveillance data contain more zeros than the fitted negative-binomial model can explain, a zero-inflated negative-binomial model can be considered. Let π_{it} be the probability that observation (i, t) belongs to a structural-zero state. Its PMF is

$$P(Y_{it} = 0) = \pi_{it} + (1 - \pi_{it})P_{NB}(0 | \mu_{it}, \alpha),$$

and, for $y > 0$,

$$P(Y_{it} = y) = (1 - \pi_{it})P_{NB}(y | \mu_{it}, \alpha).$$

The CDF is

$$F_{ZI}(y) = \begin{cases} 0, & y < 0, \\ \pi_{it} + (1 - \pi_{it})F_{NB}(y | \mu_{it}, \alpha), & y \geq 0. \end{cases}$$

A logistic model can be used for π_{it} . This extension should not be selected merely because the observed data include zeros; ordinary negative-binomial counts naturally include zeros. It becomes scientifically meaningful only if diagnostics and ecological reasoning support a distinct structural absence process.

Distributed lag nonlinear modelling can replace individual lag terms once a sufficiently long series is available. Let $w(X_{it}, \ell)$ be a cross-basis function describing exposure X at lag ℓ . Then the climatic component can be written

$$\eta_{it}^{(c)} = \sum_{\ell=0}^L w(X_{i,t-\ell}, \ell),$$

with basis functions across both exposure and lag dimensions. This avoids selecting a single lag by repeated hypothesis testing and permits cumulative climate effects to be summarized over a prespecified window. However, a 52-week pilot may be too short for highly flexible distributed-lag surfaces; quadratic terms and a small number of prespecified lags are therefore more stable for the first-stage study.

Model checking is treated as part of the analysis rather than as a final cosmetic step. Randomized quantile or simulation-based residuals should be inspected against fitted values, calendar time and site. Residual spatial pattern can be assessed through Moran-type summaries, while temporal dependence can be checked with residual autocorrelation functions. Posterior predictive simulations should reproduce the observed proportion of zeros, the upper tail of the count distribution, seasonal amplitude and between-site heterogeneity. Validation must also reflect the intended use of the model. Holding out entire sites tests spatial generalization; holding out later weeks tests forecasting. A random row-wise split would leak closely related observations into both training and test sets and would almost certainly make predictive performance look better than it would be in practice.

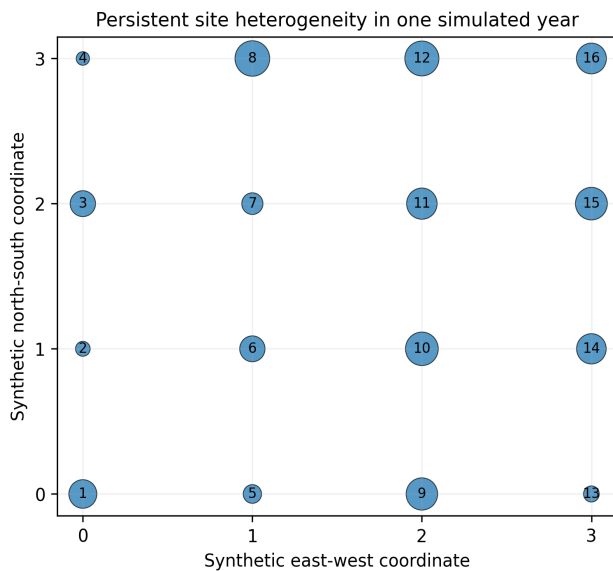


Figure 2. Bubble size represents the mean simulated count at each of 16 synthetic sentinel locations. The figure illustrates why a model may require site-level spatial structure even under a shared climatic regime.

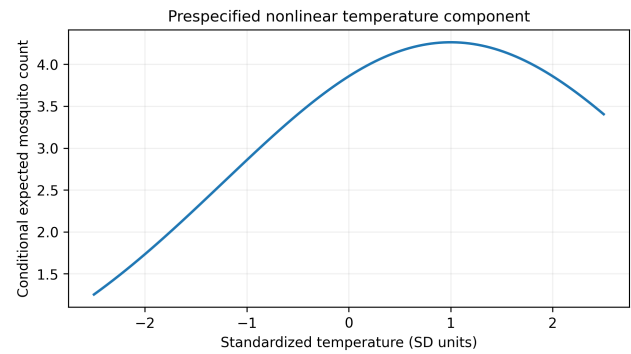


Figure 3. Conditional expected count generated by the prespecified quadratic temperature component while other covariates are held at reference values. The curve illustrates a nonlinear effect and is not an estimated Bhagalpur exposure-response function.

2. Simulation Study

The revised Monte Carlo experiment was designed around a narrower and more testable question: how much inferential distortion is introduced when a longitudinal mosquito process with overdispersion, persistent site heterogeneity and serial dependence is analysed using successively more realistic methods? The experiment does not attempt to reconstruct Bhagalpur ecology. Every climatic coefficient, covariance parameter and seasonal feature was fixed before simulation and serves only as a known benchmark. The design follows the principle that simulation studies should be planned as experiments, with prespecified data-generating mechanisms, estimands and performance measures (Burton et al., 2006; Morris et al., 2019).

Each simulated data set contained $m = 16$ sites followed for $T = 52$ weeks ($N = 832$ site-week observations). Sites were placed on a synthetic 4×4 coordinate grid. Spatial effects were generated from a zero-mean Gaussian field with standard deviation $\sigma(s) = 0.35$ and exponential correlation range $\phi = 1.4$ grid units. Each site also received a standardized environmental score U_i . Temperature followed a smooth annual cycle centred near 26 degrees C with site-week noise and a modest site-specific shift. Rainfall was generated from a broad monsoon-like seasonal envelope multiplied by Gamma (shape = 2, scale = 0.5) variation and augmented by positive random noise. Relative humidity followed a related seasonal envelope with independent perturbations. Each climatic variable was standardized over the complete simulated network-year before entering the count model. For the temporal latent process, independent site-specific AR(1) series were generated with $\rho = 0.60$ and innovation standard deviation $\sigma_t = 0.18$. The conditional log mean under the non-null data-generating mechanism was

$$\eta_{it} = 1.25 + 0.20ZT_{it} - 0.10Q_{it} + 0.25ZR_{i,t-2} + 0.12ZH_{i,t-1} + 0.30U_i + b_i + u_{it}.$$

Counts were then generated from an NB2 distribution with $\alpha = 0.50$, so that

$$\text{Var}(Y_{it} | \mu_{it}) = \mu_{it} + 0.50\mu_{it}^2.$$

The regression coefficients 0.20, -0.10, 0.25 and 0.12 were selected as simulation design values that create a mixture of strong and moderate signals; they are not prior estimates

for Bhagalpur. For example, $\beta_3 = 0.25$ corresponds to $\exp(0.25) = 1.284$, a 28.4% multiplicative change in the conditional mean per one-standard-deviation increase in the simulated two-week-lag rainfall predictor, holding the remaining components fixed. The quadratic temperature term deliberately makes the marginal temperature effect vary across the exposure range.

Three analysis strategies were compared. The first was a conventional Poisson GLM with log link, which assumes conditional equidispersion and independent observations. The second was an NB2 GLM with the generating overdispersion value $\alpha = 0.50$ supplied to the variance function. Giving the NB2 model the correct α deliberately isolates the effect of residual dependence; in field data α would be estimated. The third analysis used an NB2 regression with 15 site indicators (one site as reference) and site-cluster-robust sandwich standard errors. The site indicators absorb persistent site-level heterogeneity, while the sandwich covariance allows arbitrary residual correlation within each site's 52 repeated observations. Because only 16 independent site clusters are available, inference for the robust model used a t critical value with 15 degrees of freedom rather than a large-sample normal critical value. This third estimator is a pragmatic benchmark, not a replacement for the final continuous-space hierarchical model.

The model was specified as

$\log(\mu_{it}) = \beta_0 + \beta_1 Z^{(T)}_{it} + \beta_2 Q_{it} + \beta_3 Z^{(R)}_{it} + \beta_4 Z^{(H)}_{it} + \delta_i$,
where δ_i denotes the fixed effect associated with site i , with one site taken as the reference category to ensure model identifiability.

The cluster-robust covariance matrix of β was estimated as $\hat{V}_{CR}(\beta) = c B^{-1} [\sum_{i=1}^m S_i S_i^T] B^{-1}$,

with the finite-sample correction

$c = (m / (m - 1)) ((N - 1) / (N - p))$,

where B is the model-based information matrix and S_i represents the total NB2 score contribution from site i . To account for the limited number of site clusters, inference was based on the t -distribution with $m - 1$ degrees of freedom, t_{m-1} .

The replication count was determined from Monte Carlo precision rather than from a conventional round number. We generated $R = 10,000$ data sets under the non-null climate mechanism and another $R = 10,000$ after setting all climatic coefficients to zero, giving 20,000 independently generated data sets. For each procedure and each target coefficient, linear temperature, quadratic temperature and two-week-lag rainfall, we recorded mean estimate, bias, root mean squared error, empirical standard deviation, mean model-based standard error, nominal 95% interval coverage and rejection probability. Under the non-null mechanism, the rejection probability is power; under the null mechanism it estimates Type-I error.

The statistical performance measures were defined explicitly. If $\hat{\beta}_r$ is the estimate from replication $r = 1, \dots, R$ and β is the true generating value, Monte Carlo bias is

$$\text{Bias}(\hat{\beta}) = \frac{1}{R} \sum_{r=1}^R (\hat{\beta}_r - \beta),$$

and root mean squared error is

$$RMSE(\hat{\beta}) = \left\{ \frac{1}{R} \sum_{r=1}^R (\hat{\beta}_r - \beta)^2 \right\}^{1/2}.$$

Coverage is the proportion of replications for which

$$\hat{\beta}_r - 1.96 SE(\hat{\beta}_r) \leq \beta \leq \hat{\beta}_r + 1.96 SE(\hat{\beta}_r).$$

Monte Carlo uncertainty was quantified explicitly. For an estimated performance probability $\hat{p}$ -hat based on R independent replications,

$$MCSE(\hat{p}) = \sqrt{[\hat{p}(1 - \hat{p}) / R]}.$$

With $R = 10,000$, the maximum possible MCSE occurs at $p = 0.5$ and equals 0.005.

$$MCSE_{\max} = \sqrt{[0.5(1 - 0.5) / 10,000]} = \sqrt{[0.25 / 10,000]} = 0.005.$$

For a true Type-I error rate of 0.05 (or equivalently, coverage probability of 0.95),

$$MCSE = \sqrt{[0.05 \times 0.95 / 10,000]} \approx 0.00218.$$

, corresponding to an approximate 95% Monte Carlo half-width of 0.00427. Thus, a reported null rejection rate near 5% can be interpreted to within roughly four-tenths of a percentage point from Monte Carlo sampling alone. This degree of precision was considered appropriate because calibration, rather than only point-estimate bias, is a primary outcome of the simulation.

Independent pseudo-random streams were generated from fixed master seeds for the non-null and null scenarios so that all numerical summaries and figures are reproducible. Monte Carlo uncertainty is shown in the coverage and Type-I error figures. As throughout this paper, every graph derived from the simulation is labelled as simulated; none represents observed mosquito surveillance in Bhagalpur.

3. Results

The representative simulated network-year showed the two features that motivated the analysis: aggregation and persistence. Mean abundance was 5.15 mosquitoes per site-week, whereas the variance was 30.86, producing a variance-to-mean ratio of 5.99. Zero counts occurred in 11.8% of site-weeks. The weekly network average followed a marked seasonal wave and the synthetic sites retained visibly different annual mean counts. These features arose directly from the prespecified data-generating mechanism rather than from smoothing or post hoc modification.

Table 1. Monte Carlo performance across 10,000 non-null and 10,000 null replications.

Met hod	Param eter	T ru e be ta	Me an bet a- hat	R MS E	95 % co ve r (%)	Po we r (%)	Ty pe -I (%)
Pois son	Tempe rature	0. 20	0.2 00 6	0.0 408	64 .1	10 0.0	36. 9
Pois son	Tempe rature ²	- 0. 10	- 0.1 01 2	0.0 447	64 .8	91. 3	33. 7
Pois son	Rainfal l lag 2	0. 25	0.2 46 2	0.0 460	45 .0	10 0.0	34. 5
NB 2	Tempe rature	0. 20	0.2 00 4	0.0 387	89 .5	10 0.0	11. 0
NB 2	Tempe rature ²	- 0. 10	- 0.1 01 2	0.0 414	91 .7	76. 6	8.9
NB 2	Rainfal l lag 2	0. 25	0.2 47 9	0.0 408	89 .8	10 0.0	9.7
NB 2 + site FE + CR	Tempe rature	0. 20	0.2 00 1	0.0 365	95 .2	99. 9	5.0
NB 2 + site FE	Tempe rature ²	- 0. 10	- 0.1 00 6	0.0 401	94 .8	65. 4	5.2

Met hod	Param eter	T ru e be ta	Me an bet a- hat	R MS E	95 % co ve r (%)	Po we r (%)	Ty pe -I (%)
+							
CR							
NB 2 + site FE + CR	Rainfal l lag 2	0. 25	0.2 48 5	0.0 370	94 .2	10 0.0	5.8

Note: NB2 uses the generating $\alpha = 0.50$ to isolate the effect of dependence. FE = site fixed effects; CR = site-cluster-robust sandwich inference with $t(15)$ critical values. Monte Carlo SE for a 5% rejection probability is approximately 0.22 percentage points.

Across the 10,000 non-null replications, average coefficient estimates were close to their generating values under all three methods. For linear temperature (true $\beta = 0.20$), mean estimates were 0.2006 for Poisson, 0.2004 for NB2 and 0.2001 for NB2 with site fixed effects and cluster-robust inference. For quadratic temperature (true $\beta = -0.10$), the corresponding means were -0.1012, -0.1012 and -0.1006. For two-week-lag rainfall (true $\beta = 0.25$), they were 0.2462, 0.2479 and 0.2485. Bias was therefore small relative to the empirical sampling variability. This is an important point: approximate unbiasedness of a coefficient does not guarantee valid inference about that coefficient.

The standard-error comparison exposed the main weakness of the Poisson analysis. For linear temperature, empirical SD across replications was 0.0408, but the average Poisson SE was only 0.0193, less than half of the observed sampling variability. For quadratic temperature, empirical SD was 0.0447 versus an average Poisson SE of 0.0213; for lagged rainfall, 0.0459 versus only 0.0137. The NB2 model moved the standard errors substantially closer to empirical variability, but still tended to underestimate it. The site-fixed-effect cluster-robust analysis was much better calibrated: mean SE-to-empirical-SD ratios were approximately 0.986 for linear temperature, 0.975 for quadratic temperature and 0.956 for lagged rainfall.

These differences were reflected directly in confidence-interval coverage. Nominal 95% Poisson intervals covered the generating value in only 64.1% of replications for linear temperature, 64.8% for quadratic temperature and 45.0% for lagged rainfall. Independent NB2 regression improved

coverage to 89.5%, 91.7% and 89.8%, respectively, but remained systematically below 95%. With site fixed effects and cluster-robust inference, coverage rose to 95.2%, 94.8% and 94.2%. The Monte Carlo SE of these latter coverage estimates was approximately 0.0022–0.0023, so the remaining deviations from 95% are small relative to the failures seen under the simpler models.

Power must be read together with calibration. Poisson regression rejected the moderate quadratic-temperature null in 91.3% of non-null replications, compared with 76.6% for independent NB2 and 65.4% for the calibrated site-fixed-effect cluster-robust analysis. The higher Poisson rejection rate is not evidence of superior sensitivity: the same procedure also had grossly inflated false-positive rates and severe undercoverage. For linear temperature and lagged rainfall the prespecified signals were strong enough that all methods had essentially complete power. The quadratic term is therefore more informative for pilot design, suggesting that one year of data from 16 sites may identify a broad nonlinear tendency but may not estimate subtle curvature with high precision.

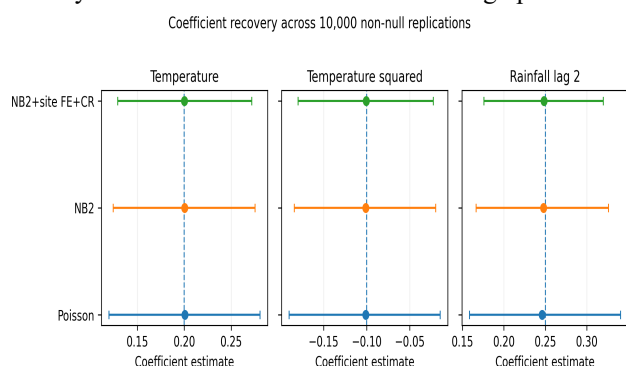


Figure 4. Monte Carlo recovery of the prespecified climatic coefficients across 10,000 non-null simulations. Points show mean estimates and horizontal intervals show the empirical 2.5th and 97.5th percentiles of the estimator distribution; dashed vertical lines indicate the generating values.

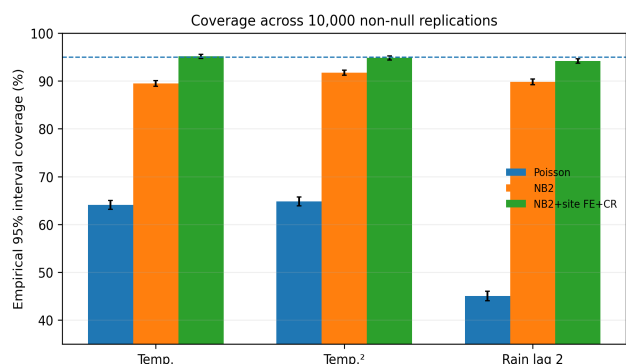


Figure 5. Empirical coverage of nominal 95% intervals across 10,000 non-null simulations. Error bars are 95% Monte Carlo intervals for the estimated coverage probability; the dashed horizontal line marks the nominal 95% target.

The climatic-null experiment provides the most direct test of calibration. When the true temperature, quadratic-temperature and rainfall coefficients were exactly zero, Poisson regression rejected at rates of 36.9%, 33.7% and

34.5%, respectively. Independent NB2 regression reduced these rates to 11.0%, 8.9% and 9.7%, demonstrating that extra-Poisson variation accounted for a substantial part of the problem but not all of it. The site-fixed-effect cluster-robust procedure produced Type-I error rates of 5.0%, 5.2% and 5.8%. With 10,000 replications, the corresponding Monte Carlo SEs were about 0.0022–0.0023, making these values compatible with near-nominal calibration and clearly distinct from the anti-conservative behaviour of the simpler models.

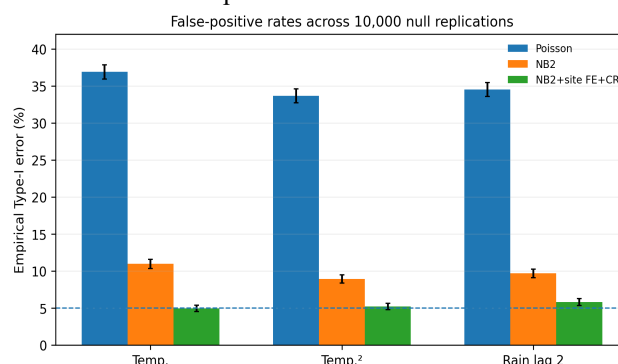


Figure 6. Empirical false-positive rates across 10,000 climatic-null simulations. Error bars are 95% Monte Carlo intervals for each estimated rejection probability; the dashed horizontal line denotes the nominal 5% significance level.

The simulation therefore separates three issues that are easily conflated in applied work. First, coefficient bias was small under all procedures. Second, allowing the variance to exceed the mean materially improved standard-error calibration. Third, residual dependence and persistent site heterogeneity still mattered after overdispersion had been addressed. A model can consequently report an approximately correct average coefficient and yet attach an unreliable confidence interval and p-value to it. For surveillance studies, where scientific conclusions often turn on uncertainty rather than only on the sign of a coefficient, this distinction is fundamental.

Taken together, the 20,000 simulation replications provide a more precise minimum standard for the field analysis. Aedes counts should be modelled using a distribution that permits overdispersion, and repeated observations from the same locations should not be treated as independent when residual dependence remains. The good calibration of the site-fixed-effect cluster-robust benchmark shows that much of the distortion is attributable to heterogeneity and within-site dependence; the planned hierarchical spatial model remains scientifically valuable because, unlike fixed site indicators, it can estimate spatial covariance, borrow strength across locations, predict at unsampled locations and quantify hotspot uncertainty.

4. Discussion

The revised simulation makes the inferential problem easier to see, mainly because the Monte Carlo error is now quite small compared with the differences observed between the methods. Across 10,000 null replications, the independent Poisson model declared climate effects significant in nearly one-third of the simulated studies, even though the true coefficients were set exactly at zero. Interestingly, this was not because the coefficient estimates were badly biased. On average, they remained close to the values used to generate the data. The real problem was the uncertainty around those estimates. The Poisson standard errors represented only about 30–48% of the actual sampling variability for the three main coefficients. As a result, the confidence intervals became too narrow and the tests rejected the null hypothesis much more often than they should have.

Allowing the variance to be larger than the mean corrected a substantial part of this problem. This is not only a statistical adjustment but also makes biological sense. Mosquito counts are naturally affected by many factors that are difficult to measure fully, such as small differences in breeding opportunities, habitat quality, trap performance, water availability and local environmental conditions. Because of this, the variability in mosquito abundance can easily be greater than what a Poisson model allows. Negative-binomial and zero-inflated negative-binomial models have also been used in previous *Aedes* studies for similar reasons. In the present simulations, the NB2 model improved 95% coverage from values as low as 45% under the Poisson model to around 90%, while false-positive rates fell from roughly one-third to about 9–11%.

However, the fact that the NB2 model still did not achieve completely correct coverage is also important. Overdispersion and dependence are not the same thing. The NB2 distribution can handle extra variability in the counts, but it does not automatically account for the fact that observations from the same surveillance site are related. Once persistent differences between sites were represented through site fixed effects, and the remaining within-site correlation was handled using cluster-robust inference, coverage increased to approximately 94–95% and Type-I error came down to about 5.0–5.8%. This benchmark model is still simpler than the full hierarchical spatial model proposed in the study. For example, it cannot estimate the spatial range, borrow information smoothly between neighbouring locations or predict a spatial effect at an unsampled site. Still, it gives an important message: the gain from the hierarchical approach cannot be explained by the negative-binomial variance alone. The dependence created by repeated surveillance also needs to be taken seriously.

The results also support the biological argument for considering lagged and nonlinear climate effects. Rainfall occurring in one week may not immediately increase the number of adult mosquitoes. Water may first collect in

containers, eggs may hatch, larvae have to develop and adults emerge only later. On the other hand, very heavy rainfall may sometimes reduce immature survival by washing larvae out of small containers. Temperature has an equally complicated role because it affects mosquito development, survival, activity and reproduction through several pathways. It is therefore difficult to justify a single constant linear effect over the whole temperature range. Previous studies have also reported lagged climate-dependent mosquito dynamics and nonlinear thermal relationships. The quadratic term used in the present pilot is deliberately simple. Its purpose is not to claim that the true relationship is exactly quadratic, but simply to avoid forcing the model to assume a straight-line response from the beginning.

For the proposed Bhagalpur study, climatic lags should therefore be selected using biological reasoning as well as statistical simplicity. One tempting approach is to fit several different lag periods, look for whichever one gives the smallest p-value, and then report that lag. This can easily produce selective results and makes later replication difficult. A more defensible strategy would be to decide on a biologically reasonable lag window before the main analysis, compare only a limited number of candidate structures, and examine both predictive performance and uncertainty. Distributed-lag approaches can eventually provide greater flexibility, but a very flexible lag-response surface may be difficult to estimate reliably from only one year of pilot data. For the initial study, relatively simple climate functions and a restricted 0–4 week lag window are likely to be more realistic. Greater flexibility can be introduced once a longer surveillance series becomes available.

The spatial component should also be interpreted carefully. If a site repeatedly shows a positive spatial effect, this does not automatically mean that the location is inherently dangerous. It simply means that, after accounting for the variables included in the model, mosquito counts at that site tend to remain above the overall network expectation. There may be several reasons for this pattern, such as unmeasured container density, vegetation, microclimate, human movement, housing conditions, vector-control activity or even differences in trap placement. The spatial effect is therefore useful in two ways. Statistically, it reduces the risk that unexplained spatial clustering will be incorrectly attributed to climate. From a field perspective, consistently high residual sites may also indicate locations that deserve closer ecological investigation.

These points have practical consequences for the field design. A surveillance network should ideally represent different environmental conditions rather than simply place as many traps as possible in locations that are easiest to access. Sixteen sites may be sufficient for a pilot, but they will provide only limited information about the exact form and range of spatial correlation. If the first year proves operationally feasible, a larger study with around

24–30 sentinel locations followed for at least two full annual cycles would provide much stronger information. Thirty sites observed weekly for 104 weeks would produce up to 3,120 site-week observations, but the benefit is not just the larger sample size. More sites improve the ability to understand spatial variation, while a second year helps determine whether a seasonal pattern is reproducible or was simply unusual during one particular monsoon season.

Model validation should also reflect the real purpose of the model. If the aim is forecasting, the model should be tested on later weeks that were not used in estimation. If the aim is prediction at new locations, entire sites or groups of nearby sites should be withheld during validation. Randomly dividing individual site-week observations into training and test sets would be less convincing because observations from the same site, or from neighbouring weeks, already share information. This type of leakage can make predictive performance look much better than it actually is when the model is finally asked to predict a genuinely new location or a future period.

Zero counts also need to be interpreted biologically rather than treated as a purely statistical nuisance. A zero may mean that mosquitoes were genuinely scarce, that the habitat was unsuitable, that mosquitoes were present but not detected by the trap, or simply that a zero occurred naturally under an over dispersed count process. For this reason, a zero-inflated model should not be added only because the data contain many zeros. Such a model assumes that there is an additional process producing structural zeros. Whether this assumption is reasonable for Bhagalpur will have to be judged from the trapping protocol, the observed zero pattern and model diagnostics. In some settings, where presence and positive abundance clearly represent two different ecological processes, a hurdle model may also be considered.

One useful feature of the hierarchical framework is that it can move the analysis from simple counts toward uncertainty-aware spatial mapping. For every site, or even for an unsampled prediction location, the model can provide a full predictive distribution rather than a single estimated number. This means that maps can show not only expected abundance, but also the probability that mosquito abundance exceeds a chosen surveillance threshold. This is more informative than simply colouring a map according to observed trap counts, because one unusually large catch may create the appearance of a hotspot even when there is considerable uncertainty around it.

At the same time, mosquito abundance should not be confused with human disease risk. A location with high *Aedes* density is not automatically a location with high dengue incidence. Human cases also depend on several other factors, including viral introduction, circulating serotypes, immunity, population movement, human–mosquito contact, case detection and public-health intervention. A later phase of the work could examine

whether mosquito surveillance adds predictive value beyond weather and seasonality, but that would require separately collected epidemiological data and a linked modelling framework. The present pilot deliberately remains focused on the mosquito population itself.

The Monte Carlo experiment also gives some guidance about study design, although it should not be treated as a complete power analysis. Under the prespecified alternative, the stronger linear temperature and rainfall effects were detected almost every time, while the moderate quadratic temperature effect had a power of only 65.4% under the best-calibrated method. This suggests that a 16-site, 52-week pilot may be sufficient to describe overdispersion, broad seasonal patterns and relatively strong climate signals, but it may still have limited ability to estimate weak curvature, interactions or highly flexible lag structures. A more complete design analysis for the larger study should therefore vary factors such as effect size, spatial correlation range, overdispersion, missing observations, zero inflation, number of sites and duration of follow-up. It would also be better to judge designs using interval width, coverage, prediction error and hotspot classification, rather than relying on statistical power alone.

The simulation also reminds us why power should never be interpreted separately from Type-I error. In these simulations, the Poisson model appeared to have high power for the quadratic temperature term, but much of this apparent advantage came from standard errors that were too small. A model that produces more p-values below 0.05 is not necessarily discovering more real effects. Sometimes it is simply more willing to interpret random variation as evidence. Simulation studies are useful here because the true values are known in advance, so both false positives and missed effects can be examined directly.

There are several possible computational approaches for fitting the final hierarchical model. A conventional Bayesian MCMC analysis is possible, but latent-Gaussian methods may also be attractive because the mosquito counts are non-Gaussian while the proposed spatial and temporal components are Gaussian. INLA therefore offers a practical alternative, and an SPDE formulation could become useful if the surveillance network later becomes much larger. For the small pilot, however, it may not be necessary to begin with the most computationally complicated method. A direct likelihood-based or Bayesian fit using a relatively simple spatial covariance structure should be adequate. The important point is to represent the structure of the data correctly; the software used to do it is secondary.

The final statistical results should also be presented in a way that remains meaningful to biological researchers. For example, if rainfall is standardized, exponentiating its regression coefficient gives the multiplicative change in expected mosquito abundance associated with a one-

standard-deviation increase in rainfall at the specified lag. When temperature enters the model nonlinearly, a plot of predicted abundance over the observed temperature range, together with uncertainty bands, is more informative than reporting one coefficient alone. Similarly, the spatial variance describes how much persistent site-to-site variation remains after measured climate and habitat variables have been considered, while the AR(1) parameter indicates how quickly unexplained changes in mosquito abundance fade from one week to the next. These quantities are usually easier to interpret biologically than a long table of regression coefficients.

The limitations of the pilot should remain visible. The simulated climate series were created to resemble seasonal behaviour and were not based on actual measured Bhagalpur weather. The coefficients used to generate the mosquito counts are hypothetical and should not be quoted as biological effects for the district. The comparison models were also deliberately simpler than the final proposed hierarchy, meaning that the simulation mainly demonstrates what happens when overdispersion and dependence are ignored rather than proving that every component of the future Bayesian model is optimal. Detection probability was not modelled separately either. In a larger study, repeated traps or other designs might help distinguish true abundance from imperfect detection, but that is beyond the scope of the initial pilot.

One strength of the proposed work is that the statistical thinking comes before, rather than after, field data collection. In many ecological studies, data are collected first and only later is a statistical method chosen according to what the data happen to look like. Here, the intended analysis helps determine what information should be collected from the beginning: exact location, week, trapping effort, trap condition, species, weather, habitat characteristics and intervention history. It also keeps the first-year objectives realistic. The pilot should primarily establish whether surveillance is feasible, assess the quality of the data, describe the count distribution and provide preliminary estimates of lag structure and variance components. Strong causal conclusions or a fully operational prediction system would belong to a later, independently validated phase.

The same statistical logic is not restricted to mosquito surveillance. Over-dispersed counts with spatial clustering and temporal persistence are common in studies of parasites, plant diseases, microbial colonies, wildlife trapping and many other biological processes. The broader lesson is quite simple: choose a probability model that resembles the way the data are generated, represent the dependence created by the sampling design, report uncertainty realistically, validate the model in the way it will eventually be used, and do not mix simulated evidence with conclusions drawn from real field observations.

5. Conclusion

This pilot develops a statistical framework for studying how *Aedes* mosquito abundance may vary across space, time and climatic conditions in Bhagalpur, Bihar. The main statistical issue is straightforward, even if the modelling is not: mosquito surveillance produces structured count data. The variance may be much larger than the mean, observations from nearby sites may share environmental influences, weekly counts from the same location may remain correlated, and weather may affect mosquito abundance after a delay or in a nonlinear manner. These are not minor technical complications. They directly affect the reliability of estimated climate effects and the confidence placed in apparent hotspots.

The proposed working model therefore combines an NB2 count distribution with a log link and an offset for sampling effort, standardized climatic variables, lagged rainfall and humidity, a nonlinear temperature component, measured site-level characteristics, a spatial Gaussian effect and AR(1) temporal dependence. If the eventual field data provide convincing evidence for a separate structural-zero mechanism, a zero-inflated extension can be considered. More elaborate distributed-lag functions or SPDE-based spatial models would be better introduced later, when a larger data set is available to support that additional complexity.

The 20,000-replication Monte Carlo experiment gives fairly clear evidence for the main methodological argument. When simulated mosquito counts contained overdispersion, persistent site differences and temporal dependence, independent Poisson regression produced coefficient estimates that were not badly biased on average, but its uncertainty estimates were seriously misleading. Coverage of nominal 95% intervals fell to 45.0–64.8%, while Type-I error increased to 33.7–36.9%. Moving to an NB2 model improved these results substantially, but did not completely solve the problem. When site heterogeneity and repeated-measure dependence were additionally addressed, coverage increased to 94.2–95.2% and Type-I error returned to approximately 5.0–5.8%. The message from the simulation is therefore not simply that the negative-binomial distribution is preferable to Poisson. It is that both the variance structure and the dependence created by the surveillance design matter for reliable inference.

None of these simulated results tells us whether temperature, rainfall or humidity actually has a particular effect in Bhagalpur. Those relationships can only be estimated from prospective field observations. The logical next step is therefore a standardized 16-site, 52-week pilot, followed, if the surveillance system works well, by a larger network of approximately 24–30 sites monitored over at least two annual cycles. The full analysis should estimate overdispersion from the field data, model spatial dependence directly, use blocked spatial and temporal

validation, and report uncertainty around both climate-response curves and hotspot probabilities.

Ultimately, the scientific question should move beyond statements such as “rainfall was significant” or “temperature was correlated with mosquito density.” More useful questions are how mosquito abundance changes across different climatic conditions and lag periods, which places remain persistently high after measured environmental differences are considered, how long unexplained population changes persist, and how well these patterns can be predicted at new locations or future times. A hierarchical spatio-temporal model provides a practical way of asking these questions while avoiding the false statistical precision that was clearly visible in the simulation.

References

1. Tewari SC, Thenmozhi V, Katholi CR, Manavalan R, Munirathinam A, Gajanana A. Dengue vector prevalence and virus infection in a rural area in south India. *Tropical Medicine & International Health*. 2004;9(4):499-507. doi:10.1111/j.1365-3156.2004.01103.x.
 2. Simões TC, Codeço CT, Nobre AA, Eiras AE. Modeling the non-stationary climate dependent temporal dynamics of *Aedes aegypti*. *PLoS ONE*. 2013;8(8):e64773. doi:10.1371/journal.pone.0064773.
 3. Villela DAM, Codeço CT, Figueiredo F, Garcia GA, Maciel-de-Freitas R, Struchiner CJ. A Bayesian hierarchical model for estimation of abundance and spatial density of *Aedes aegypti*. *PLoS ONE*. 2015;10(4):e0123794. doi:10.1371/journal.pone.0123794.
 4. Mordecai EA, Cohen JM, Evans MV, Gudapati P, Johnson LR, Lippi CA, et al. Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. *PLoS Neglected Tropical Diseases*. 2017;11(4):e0005568. doi:10.1371/journal.pntd.0005568.
 5. Kache PA, Eastwood G, Collins-Palmer K, Katz M, Falco RC, Bajwa WI, et al. Environmental determinants of *Aedes albopictus* abundance at a northern limit of its range in the United States. *American Journal of Tropical Medicine and Hygiene*. 2020;102(2):436-447. doi:10.4269/ajtmh.19-0244.
 6. Yang B, Borgert BA, Alto BW, Boohene CK, Brew J, Deutsch K, et al. Modelling distributions of *Aedes aegypti* and *Aedes albopictus* using climate, host density and interspecies competition. *PLoS Neglected Tropical Diseases*. 2021;15(3):e0009063. doi:10.1371/journal.pntd.0009063.
 7. Whittaker C, Winskill P, Sinka M, Pironon S, Massey C, Weiss DJ, et al. A novel statistical framework for exploring the population dynamics and seasonality of mosquito populations. *Proceedings of the Royal Society B: Biological Sciences*. 2022;289(1972):20220089. doi:10.1098/rspb.2022.0089.
 8. Dharmamuthuraja D, Rohini PD, Lakshmi MI, Isvaran K, Ghosh SK, Ishtiaq F. Determinants of *Aedes* mosquito larval ecology in a heterogeneous urban environment-a longitudinal study in Bengaluru, India. *PLoS Neglected Tropical Diseases*. 2023;17(11):e0011702. doi:10.1371/journal.pntd.0011702.
 9. Ortega-Lenis D, Arango-Londoño D, Hernández F, Moraga P. Effects of climate variability on the spatio-temporal distribution of dengue in Valle del Cauca, Colombia, from 2001 to 2019. *PLoS ONE*. 2024;19(10):e0311607. doi:10.1371/journal.pone.0311607.
 10. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Statistics in Medicine*. 2010;29(21):2224-2234. doi:10.1002/sim.3940.
 11. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2009;71(2):319-392. doi:10.1111/j.1467-9868.2008.00700.x.
 12. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2011;73(4):423-498. doi:10.1111/j.1467-9868.2011.00777.x.
 13. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Statistics in Medicine*. 2006;25(24):4279-4292. doi:10.1002/sim.2673.
- Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Statistics in Medicine*. 2019;38(11):2074-2102. doi:10.1002/sim.8086.