Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Published online: 09.03.2026
DOI : 10.15672/hujms.1849711

Research Article

Hybrid conditional autoregressive generalized random forest for overdispersed spatial counts: Evidence from Kalimantan wildfire hotspots

Irfani Azis*[1], Anik Djuraidah[2], Muhammad Nur Aidi[3], Indahwati[4], Ardhasena Sopaheluwakan[5]

[1,2,3,4] *School of Data Science, Mathematics, and Informatics, IPB University, Bogor, Indonesia*
[5] *Meteorological, Climatological, and Geophysical Agency, BMKG, Jakarta, Indonesia*

Abstract

This paper proposes a negative binomial conditional autoregressive generalized random forest model for overdispersed spatial count data. The framework combines a flexible nonparametric mean function estimated via generalized random forests with conditional autoregressive spatial random effects under a negative binomial likelihood. We evaluate the proposed method through a simulation study and an empirical application to satellite-derived wildfire hotspot counts in Kalimantan, Indonesia (September 2024), using rainfall, air temperature, relative humidity, wind speed, and the number of rainless days as covariates. Across simulation scenarios, negative binomial conditional autoregressive generalized random forest yields lower prediction errors than the benchmark negative binomial conditional autoregressive random forest. In the Kalimantan application, the model achieves a root mean squared error of 3.64 and a mean absolute error of 2.11 hotspots per grid cell. Variable importance analysis indicates that air temperature is the most influential predictor of the spatial distribution of hotspots.

Mathematics Subject Classification (2020). 62M30, 62P12, 62H11

Keywords. Conditional autoregressive, generalized random forest, overdispersed count data, negative binomial, spatial modelling

1. Introduction

Modelling overdispersed count data, where the variance exceeds the mean, often requires distributional choices and inferential procedures beyond the classical Poisson framework. In many applied contexts, including public health, insurance, ecology, and disaster risk, the negative binomial (NB) distribution is widely adopted because it introduces a dispersion parameter and provides a more realistic likelihood for overdispersed outcomes. Recent empirical evaluations report that NB models and their extensions, including zero-inflated variants, frequently outperform Poisson models when overdispersion and zero inflation are

*Corresponding Author.
Email addresses: irfaniazis@apps.ipb.ac.id (I. Azis), anikdjuraidah@apps.ipb.ac.id (A. Djuraidah), muhammadai@apps.ipb.ac.id (M. N. Aidi), indahwati@apps.ipb.ac.id (Indahwati), ardhasena@bmkg.go.id (A. Sopaheluwakan)
Received: 30.12.2025; Accepted: 21.02.2026

pronounced, in both simulations and real-world studies. These findings underscore that selecting a likelihood consistent with the underlying data-generating process is fundamental for estimation stability and predictive accuracy in count-data modelling [2, 16, 29, 30]. Such characteristics are common in satellite-derived wildfire hotspot counts used for early warning and risk mapping.

Many count processes are also spatially structured, in the sense that observations in one location tend to be influenced by neighbouring areas through shared environmental, social, or physical mechanisms. Ignoring spatial dependence can bias inference and degrade prediction, particularly when residual correlation remains after controlling for observed covariates. Conditional autoregressive (CAR) models offer a principled approach for representing structured areal random effects, and modern formulations based on Gaussian Markov random fields are widely used for areal and spatio-temporal settings [21]. Recent methodological developments have extended CAR models to NB outcomes, including formulations with constant or spatially varying dispersion, enabling more realistic modeling of local nonstationarity in spatial count data [3, 10, 16]. Reviews in spatial epidemiology and related fields likewise emphasize that spatial components are essential for mapping, cluster detection, and prediction in spatial count models [25, 32].

Hierarchical CAR-based models are effective in capturing residual spatial correlation. In contrast, machine-learning methods such as random forests (RF) [6] excel at learning nonlinear relationships and complex interactions between covariates. However, standard RF assumes conditionally independent observations and does not explicitly account for spatial autocorrelation. Consequently, conventional random cross-validation can yield overly optimistic performance estimates when spatial dependence is present, motivating spatially aware validation strategies and hybrid algorithms that combine flexible learners with explicit spatial structure [11, 17, 24, 26]. In this line of work, iterative frameworks have been proposed that alternate between a forest-based mean function and a CAR-type spatial effect to absorb residual dependence while preserving nonlinearity in the covariate–response relationship [20].

In the specific context of overdispersed spatial count data, Azis et al. [5] proposed the negative binomial conditional autoregressive random forest (NBCAR-Forest) model to predict wildfire hotspot counts in Kalimantan. This framework combines an NB likelihood to accommodate overdispersion, a CAR component to represent spatial dependence, and an RF engine to capture nonlinearities and interactions. Although NBCAR-Forest demonstrated promising performance, the study also reported practical challenges, including a tendency toward overestimation under specific configurations and sensitivity to the choice of spatial validation scheme. These limitations motivate further methodological development to enhance robustness, mitigate systematic prediction bias, and strengthen inferential reliability for the spatial prediction of overdispersed counts.

A complementary strand of research addresses spatial heterogeneity through localized learning, such as geographically weighted random forest approaches. However, these methods typically do not represent residual spatial autocorrelation through CAR-type priors and are less developed for NB count likelihoods [11, 18, 27]. More broadly, existing NB–CAR–forest hybrids rely on standard RF, which is primarily predictive and can be sensitive to validation design and tuning. This leaves a gap for a likelihood-consistent forest component with a clearer inferential target within an NB–CAR framework.

Motivated by this gap, we use the generalized random forest (GRF) framework [4]. GRF generalizes standard forests by targeting parameters defined through local moment conditions and provides a formal basis for asymptotic inference under suitable regularity conditions [4]. This inferential structure makes GRF a natural candidate for strengthening the learning component within an NB–CAR hybrid model.

Consequently, we propose the negative binomial conditional autoregressive generalized random forest (NBCAR-GRF) model. The proposed approach integrates three elements:

an NB likelihood for overdispersed counts, CAR spatial random effects to account for residual areal dependence, and a GRF-based mean function to capture nonlinear predictor–response relationships and complex interactions. Relative to NBCAR-Forest, the key methodological change is replacing the RF component with GRF to enhance the flexibility of the estimated mean structure and to provide a clearer inferential foundation for the forest-based component [4–6]. Unlike standard RF, GRF targets parameters defined through local moment conditions, making the forest component explicitly parameter-seeking rather than purely predictive. This aligns naturally with the likelihood-based NB–CAR layer because the mean structure can be updated through well-defined estimating equations, while the CAR component retains residual spatial dependence [4–6].

To our knowledge, this is the first study to integrate GRF with a likelihood-based NB–CAR model for overdispersed spatial count data within an iterative hybrid estimation framework. Integrating a flexible forest learner with a spatial random-effects model is non-trivial: joint estimation can be sensitive to tuning and stopping criteria, and care is needed to prevent the forest component from absorbing structured spatial signal that should be captured by the CAR. We therefore design the algorithm and validation strategy to balance nonlinear learning with explicit spatial dependence modelling.

The specific objectives of this study are:

- to develop the NBCAR-GRF framework and an estimation procedure that couples a GRF mean function with NB–CAR spatial random effects for overdispersed areal counts;
- to evaluate predictive performance and numerical stability against NBCAR-Forest under controlled simulation scenarios with varying spatial autocorrelation and distributional settings; and
- to demonstrate the proposed model on Kalimantan wildfire hotspot counts using ERA5 meteorological covariates and to assess key predictors of spatial variation in predicted hotspots.

To evaluate NBCAR-GRF, the study comprises two components. First, a simulation study systematically varies the sample size, strength of spatial autocorrelation, feature variability, distributional settings, and the magnitude of spatial random effects. This design enables a direct comparison between NBCAR-GRF and the benchmark NBCAR-Forest under controlled scenarios that reflect realistic data conditions. Second, an empirical application focuses on satellite-derived wildfire hotspot counts in Kalimantan, Indonesia, for September 2024. Covariates include rainfall, air temperature, relative humidity, wind speed, and the number of rainless days derived from the ERA5 reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 is widely used in atmospheric and hydrometeorological applications and provides near-real-time updates [13]. The remote-sensing literature on Kalimantan indicates that wildfire hotspot dynamics are strongly linked to climate anomalies and land use and land cover patterns, supporting the use of meteorological covariates to model spatial variation in wildfire hotspot counts [8, 12, 28]. These meteorological drivers also exhibit spatial structure and jointly influence fuel moisture and fire-spread potential, further motivating their use in a spatial modelling framework. By improving spatial prediction of wildfire hotspot counts for monitoring and early warning, this work also supports Sustainable Development Goal 13 on climate action through strengthened climate-related risk information.

Figure 1 provides an overview of the NBCAR-GRF framework, highlighting the novel components and main outputs. The remainder of this article is organized as follows. Section 2 presents the NBCAR-GRF methodology, including estimation and validation procedures. Sections 3 and 4 report simulation results and empirical findings for Kalimantan wildfire hotspots. Section 5 provides discussion and concluding remarks.

2. Methods

Figure 1 summarizes the overall NBCAR-GRF workflow. The model integrates three components within an iterative scheme: (i) a negative binomial likelihood for overdispersed counts, (ii) CAR-based spatial random effects (estimated via INLA), and (iii) a GRF mean function to capture nonlinear covariate effects. The highlighted block in the figure indicates the novel GRF–NB–CAR coupling that differentiates NBCAR-GRF from NBCAR-Forest. Model performance is evaluated using RMSE and MAE.

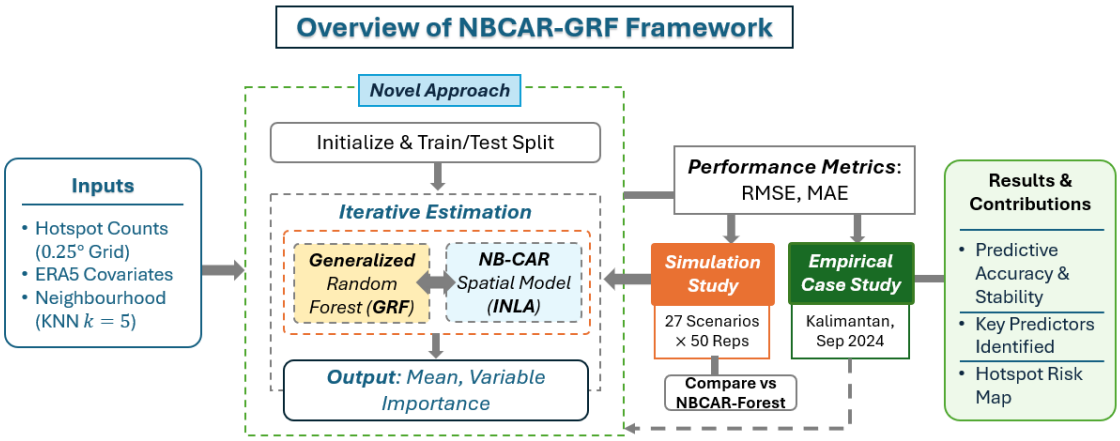


Figure 1. Overview of the NBCAR-GRF framework. Dashed boxes highlight the novel components, while the right panel summarises the main outputs and contributions.

2.1. Model formulation

Following [5], we model the count response Y_k observed in areal unit k, for $k = 1, 2, \ldots, K$. Let $x_k \in \mathbb{R}^p$ denote the covariate vector at location k. We assume a negative binomial model with mean $\mu_k > 0$ and dispersion (size) parameter $\theta > 0$, written as $Y_k \sim \text{NegBin}(\mu_k, \theta)$. Under this parameterization,

$$\mathbb{E}(Y_k) = \mu_k, \qquad \text{Var}(Y_k) = \mu_k\left(1 + \frac{\mu_k}{\theta}\right),$$

which explicitly accommodates overdispersion.

Using a log link, the linear predictor is

$$\log(\mu_k) = \eta_k = x_k^\top \beta + \phi_k, \tag{2.1}$$

where β is a regression parameter vector and ϕ_k is a CAR-type spatial random effect capturing residual dependence among neighbouring areal units. In [5], a decorrelated response is constructed to remove ϕ_k from the learning target. In this study, we adopt an IRLS-based working-response construction tailored to the negative binomial likelihood and perform decorrelation by subtracting the spatial-effect estimate from the previous iteration.

2.2. Negative binomial likelihood and IRLS-based working response

Let y_k be a realisation of Y_k. Using $\text{NB}(\mu_k, \theta)$ parameterisation with

$$p_k = \frac{\theta}{\theta + \mu_k} \in (0, 1),$$

the probability mass function can be written as

$$\Pr(Y_k = y_k \mid \mu_k, \theta) = \frac{\Gamma(y_k + \theta)}{\Gamma(\theta)\, y_k!} \left(\frac{\mu_k}{\theta + \mu_k}\right)^{y_k} \left(\frac{\theta}{\theta + \mu_k}\right)^{\theta}, \qquad y_k = 0, 1, 2, \ldots. \tag{2.2}$$

The log-likelihood contribution from location k is

$$\ell_k(\mu_k) = \log \frac{\Gamma(y_k + \theta)}{\Gamma(\theta)\, y_k!} + y_k \log \mu_k - (y_k + \theta) \log(\theta + \mu_k) + \theta \log \theta, \tag{2.3}$$

where terms not depending on μ_k may be absorbed into a constant.

Because $\eta_k = \log(\mu_k)$ and $\mu_k = \exp(\eta_k)$, we have

$$\frac{\partial \mu_k}{\partial \eta_k} = \mu_k.$$

Differentiating (2.3) yields the score with respect to μ_k,

$$\frac{\partial \ell_k}{\partial \mu_k} = \frac{y_k}{\mu_k} - \frac{y_k + \theta}{\theta + \mu_k},$$

and, by the chain rule, the score with respect to η_k becomes

$$\frac{\partial \ell_k}{\partial \eta_k} = \frac{\partial \ell_k}{\partial \mu_k} \frac{\partial \mu_k}{\partial \eta_k} = \left(\frac{y_k}{\mu_k} - \frac{y_k + \theta}{\theta + \mu_k}\right) \mu_k = \frac{\theta\,(y_k - \mu_k)}{\theta + \mu_k}. \tag{2.4}$$

The second derivative with respect to η_k is

$$\frac{\partial^2 \ell_k}{\partial \eta_k^2} = -\frac{\theta\, \mu_k\, (y_k + \theta)}{(\theta + \mu_k)^2}. \tag{2.5}$$

Using $\mathbb{E}(Y_k) = \mu_k$, the Fisher information for η_k is

$$\mathcal{I}_k(\eta_k) = -\,\mathbb{E}\left[\frac{\partial^2 \ell_k}{\partial \eta_k^2}\right] = \frac{\theta\, \mu_k}{\theta + \mu_k}. \tag{2.6}$$

2.3. IRLS update, working response, and weights

Following the IRLS (Newton–Raphson/Fisher scoring) scheme [1, 23], the update at iteration r is

$$\eta_k^{(r)} = \eta_k^{(r-1)} + \frac{\left.\dfrac{\partial \ell_k}{\partial \eta_k}\right|_{(r-1)}}{\mathcal{I}_k(\eta_k)\big|_{(r-1)}}. \tag{2.7}$$

Substituting (2.4) and (2.6) evaluated at $(r-1)$ yields the working-response form

$$\eta_k^{(r)} = \eta_k^{(r-1)} + \frac{y_k - \mu_k^{(r-1)}}{\mu_k^{(r-1)}}. \tag{2.8}$$

Equivalently, the IRLS working response can be defined as

$$z_k^{(r)} = \eta_k^{(r-1)} + \frac{y_k - \mu_k^{(r-1)}}{\mu_k^{(r-1)}}, \tag{2.9}$$

so that $z_k^{(r)} = \eta_k^{(r)}$ in (2.8). The corresponding IRLS weight is

$$w_k^{(r)} = \mathcal{I}_k(\eta_k)\big|_{(r-1)} = \frac{\theta\, \mu_k^{(r-1)}}{\theta + \mu_k^{(r-1)}} = \frac{\mu_k^{(r-1)}}{1 + \dfrac{\mu_k^{(r-1)}}{\theta}}. \tag{2.10}$$

2.4. Decorrelation of the working response

To learn the non-spatial component of the mean structure, we remove the spatial-effect estimate from the previous iteration. Specifically, we define the decorrelated working response as

$$Z_k^{(r)} = z_k^{(r)} - \phi_k^{(r-1)}. \tag{2.11}$$

This construction follows the spirit of Azis et al. [5]—namely, training the learner on a target that is as free as possible from the spatial random effect—but here $Z_k^{(r)}$ is obtained from the NB–IRLS working response rather than from the original count scale.

In this study, the GRF component is trained on $\{(x_k, Z_k^{(r)})\}_{k=1}^K$ without applying the IRLS weights in (2.10). Distributional calibration and residual spatial dependence are subsequently handled through the NB–CAR update step in the NBCAR-GRF iteration.

2.5. NBCAR-GRF algorithm

The proposed NBCAR-GRF procedure modifies NBCAR-Forest [5] by replacing the random forest learner with a generalized random forest (GRF) [4]. The algorithm proceeds as follows:

Algorithm 1. NBCAR-GRF algorithm

Step 0 (Initialization). Set tuning parameters for GRF (e.g., N_{tree}, m_{try}, and $\min_{\text{node}}$), the neighbourhood specification (D), and the maximum number of iterations (R). Initialise $\hat{\phi}_k^{(0)} = 0$ for $k = 1, 2, \ldots, K$.

Step 1 (For $r = 1, \ldots, R$).

Step 1A (NB–GLM update and decorrelation). Fit an NB–GLM with a log link using $\hat{\phi}_k^{(r-1)}$ as an offset:

$$Y_k \sim \text{NegBin}(\mu_k, \theta), \qquad \log(\mu_k) = \eta_k = x_k^\top \beta + \hat{\phi}_k^{(r-1)}. \tag{2.12}$$

Obtain $\mu_k^{(r-1)}$ and $\eta_k^{(r-1)} = \log(\mu_k^{(r-1)})$. Compute the working response $z_k^{(r)}$ using (2.9) and the decorrelated response $Z_k^{(r)}$ using (2.11).

Step 1B (GRF fit on decorrelated responses). Fit a GRF using predictors x_k and response $Z_k^{(r)}$. Obtain out-of-bag predictions for training units and predictions for test units. Denote the fitted non-spatial mean component by $\hat{m}_k^{(r)}$.

Step 1C (NB–CAR update via INLA). Treat $\hat{m}_k^{(r)}$ as fixed and fit an NB–CAR model via INLA:

$$Y_k \sim \text{NegBin}(\mu_k, \theta), \qquad \log(\mu_k) = \hat{m}_k^{(r)} + \phi_k, \tag{2.13}$$

where $\phi = (\phi_1, \ldots, \phi_K)$ follows a CAR prior defined on the adjacency graph induced by the areal units. INLA provides posterior summaries for ϕ_k and θ. Update the spatial effect using the posterior mean $\hat{\phi}_k^{(r)} = \mathbb{E}(\phi_k \mid Y)$ and pass $\hat{\phi}_k^{(r)}$ to Step 1A as the offset for the next iteration.

Stopping rule. Iterate until $r = R$ or $\max_k |\hat{\phi}_k^{(r)} - \hat{\phi}_k^{(r-1)}| < \varepsilon$, with ε typically set to 10^{-4}.

Step 2 (Prediction). At the final iteration, obtain predicted means

$$\hat{\mu}_k = \exp(\hat{m}_k^{(R)} + \hat{\phi}_k^{(R)}), \tag{2.14}$$

and report predicted wildfire hotspot counts based on $\hat{\mu}_k$.

2.6. Simulation design

This study consists of (i) a simulation study to evaluate robustness under controlled data-generating conditions and (ii) an empirical application using satellite-derived wildfire hotspot counts over Kalimantan, Indonesia, for September 2024, with meteorological covariates. NBCAR-Forest is used as the benchmark model throughout. For both simulation and empirical analyses, the data are randomly split into training and testing sets using an 80:20 proportion, and predictive performance is evaluated on the held-out testing set. The simulation workflow is summarized as follows:

1. **Sample size.** Specify the number of areal units as $n \in \{700, 1200, 2500\}$.
2. **Spatial lattice and neighbourhood structure.** Construct a regular spatial grid with a resolution of $0.25^\circ \times 0.25^\circ$. Define the spatial weight matrix W using a k-nearest neighbour (KNN) graph with $k = 5$, and row-standardise W where applicable.
3. **Spatial dependence strength.** Set the spatial autocorrelation coefficient ρ to represent three levels: low (0–0.40), medium (0.41–0.80), and high (0.81–0.99).
4. **Forest hyperparameters.** Fix GRF hyperparameters at $m_{\text{try}} = 1$, $\min_{\text{node}} = 5$, and $N_{\text{tree}} = 1000$.
5. **Covariate generation and collinearity check.** Generate X_1 and X_2 from a bivariate uniform distribution on $(0, 2\pi)$. Assess multicollinearity using the variance inflation factor (VIF).
6. **Nonlinear mean structure.** Specify a nonlinear predictor–response relationship as $m(x) = A \cdot \sin(X_1) + \log(X_2 + 0.5)$, where $A \in \mathbb{R}$ controls the magnitude of the nonlinear component. Calibrate A to achieve the desired relative contribution of covariate effects and spatial random effects. Verify nonlinearity using a likelihood ratio test that compares the nonlinear specification against a nested linear alternative.
7. **Spatial random effects and response generation.** Generate spatial random effects ϕ from a Gaussian model with covariance structure $\phi \sim N(0, (I - C)^{-1}M)$, where $C = B + B^T - BB^T$, $B = \rho W$, and $M \equiv \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$. Generate the count response from a negative binomial distribution $Y_k \sim \text{NegBin}(\mu_k, \theta)$ with $\log(\mu_k) = m(x_k) + \phi_k$.
8. **Train–test split.** Randomly split the simulated dataset into training (80%) and testing (20%) sets.
9. **Model fitting and prediction.** Fit NBCAR-GRF and NBCAR-Forest using the training set. For each model, obtain the final predicted means on the testing set using Step 2 of the corresponding algorithm, denoted by $\hat{\mu}_k$.
10. **Repetition.** Repeat Steps 1–9 for 50 replications per scenario to assess stability and consistency.
11. **Performance evaluation.** Compute RMSE and MAE by comparing the observed counts Y_k with the final algorithmic predictions $\hat{\mu}_k$:

$$\text{RMSE} = \sqrt{\frac{1}{n_{\text{test}}} \sum_{k \in T} (Y_k - \hat{\mu}_k)^2}, \qquad \text{MAE} = \frac{1}{n_{\text{test}}} \sum_{k \in T} |Y_k - \hat{\mu}_k| .$$

Here, T denotes the index set of testing units and $n_{\text{test}} = |T|$. Lower RMSE and MAE indicate better predictive accuracy. Because the response exhibits spatial dependence, random holdout evaluation may be optimistic relative to spatially blocked validation; the reported results should therefore be interpreted as baseline predictive performance under random splitting.

Overall, 27 simulation scenarios are considered. For each scenario, the model with smaller RMSE and MAE is regarded as providing better predictive performance. The simulation results are then used as supporting evidence for the empirical application.

2.7. Empirical analysis

This empirical application is a single-month (September 2024) case study intended as a proof-of-concept demonstration of NBCAR-GRF in an operational hotspot-mapping setting. The response variable is the number of hotspots within each $0.25^\circ \times 0.25^\circ$ grid cell over Kalimantan.

Predictor variables include air temperature, total precipitation, the number of rainless days, wind speed, and relative humidity, extracted from the ECMWF–ERA5 reanalysis for the same period. These meteorological drivers exhibit coherent spatial gradients over Kalimantan and jointly govern fuel moisture, ignition potential, and fire spread, which motivates their inclusion within a spatial NB–CAR modelling framework rather than a conditionally independent count model [13, 19, 22, 33].

Higher precipitation generally suppresses hotspots by increasing surface wetness, while prolonged dry spells (rainless days) intensify fuel dryness and elevate hotspot occurrence [22]. Warmer air temperatures can further exacerbate drying through enhanced evapotranspiration, increasing fire susceptibility. Wind speed facilitates fire propagation and ember transport, and lower relative humidity is commonly associated with more flammable conditions [33].

The empirical analysis proceeds as follows:

1. Compile wildfire hotspot counts for all grid cells over Kalimantan for September 2024.
2. Extract meteorological covariates from ECMWF–ERA5 at matching spatial resolution.
3. Split the empirical dataset into training (80%) and testing (20%) sets.
4. Fit the proposed NBCAR-GRF model using the training set and obtain testing-set predictions via Step 2.
5. Evaluate predictive performance using RMSE and MAE computed on the testing set.
6. Interpret results in terms of predictive performance, influential covariates, and implications for spatial prediction of wildfire hotspot risk.

3. Results of simulation study

This section reports simulation results for the proposed NBCAR–GRF model and compares its predictive performance with the benchmark NBCAR–Forest under the scenarios described in Section 2.6. Tables 1–2 summarise out-of-sample accuracy on the held-out 20% testing set using RMSE and MAE. To standardise reporting of extreme failures, both metrics were capped at 500 on the testing set. Accordingly, entries shown as "> 500" indicate configurations in which the testing error exceeded this threshold, which is consistent with severe numerical instability or divergence of the iterative fitting procedure.

Table 1. RMSE values of simulation for the NBCAR–Forest and NBCAR–GRF models.

Scenario	n	RMSE NBCAR–Forest			RMSE NBCAR–GRF		
		ρ low	ρ medium	ρ high	ρ low	ρ medium	ρ high
$\text{var}(x) < \text{var}(\phi)$	700	12.71	17.55	20.87	10.61	12.21	16.95
	1200	11.86	14.63	19.01	10.25	12.81	17.03
	2500	13.24	440.68	31.51	11.35	13.11	18.69
$\text{var}(x) \approx \text{var}(\phi)$	700	152.36	117.13	122.54	22.51	21.35	26.16
	1200	75.15	>500	204.05	20.00	24.63	30.40
	2500	>500	>500	>500	22.06	27.08	35.23
$\text{var}(x) > \text{var}(\phi)$	700	>500	>500	>500	193.84	189.24	259.59
	1200	>500	>500	>500	201.01	249.41	293.71
	2500	>500	>500	>500	214.18	276.70	399.07

Table 2. MAE values of simulation for the NBCAR–Forest and NBCAR–GRF models.

Scenario	n	MAE NBCAR–Forest			MAE NBCAR–GRF		
		ρ low	ρ medium	ρ high	ρ low	ρ medium	ρ high
$\text{var}(x) < \text{var}(\phi)$	700	6.68	7.47	8.73	5.71	6.23	7.54
	1200	6.44	7.06	8.43	5.71	6.36	7.50
	2500	6.53	26.48	9.26	5.82	6.41	7.72
$\text{var}(x) \approx \text{var}(\phi)$	700	23.94	21.47	23.13	9.07	9.55	11.16
	1200	15.03	71.25	26.26	8.91	10.36	11.58
	2500	182.04	64.15	15.00	9.35	10.58	12.61
$\text{var}(x) > \text{var}(\phi)$	700	>500	>500	>500	64.66	67.47	83.10
	1200	>500	>500	>500	69.59	77.73	88.05
	2500	>500	>500	>500	74.86	88.77	111.15

The results in Tables 1–2 cover 27 configurations (3 variance regimes $\times 3$ sample sizes $\times 3$ spatial dependence levels). Using the capped reporting rule, NBCAR–Forest exceeds the RMSE threshold in 13 of the 27 configurations (Table 1) and exceeds the MAE threshold in 9 of the 27 configurations (Table 2). In contrast, NBCAR–GRF returns finite RMSE and MAE values in all configurations reported in both tables, indicating a substantially lower frequency of extreme failures under the same simulation and evaluation protocol.

In the regime where spatial variation dominates covariate variation, $\text{var}(x) < \text{var}(\phi)$, both approaches generally yield small, finite errors, but NBCAR–GRF consistently improves accuracy. For example, at $n = 700$, NBCAR–Forest RMSE ranges from 12.71 to 20.87 as ρ increases from low to high, whereas NBCAR–GRF reduces this range to 10.61–16.95 (Table 1). A similar pattern holds for MAE: 6.68–8.73 under NBCAR–Forest versus 5.71–7.54 under NBCAR–GRF (Table 2). Across sample sizes, NBCAR–GRF remains stable in magnitude (RMSE approximately 10–19; MAE approximately 5.7–7.7), suggesting consistent out-of-sample behaviour in this regime.

A notable exception in this regime occurs for NBCAR–Forest at $n = 2500$ under medium ρ, where RMSE increases sharply to 440.68 (Table 1) and MAE rises to 26.48 (Table 2), despite much smaller errors in adjacent dependence levels. NBCAR–GRF does not show a comparable spike for the same configuration (RMSE 13.11; MAE 6.41). Because the data-generating mechanism, split proportion, and evaluation protocol are fixed within each configuration, such abrupt departures are consistent with occasional instability of the iterative fitting procedure under certain parameter combinations, rather than a systematic increase in difficulty across the entire regime.

When covariate-driven and spatial-effect-driven variations are comparable, $\text{var}(x) \approx \text{var}(\phi)$, the performance gap becomes more pronounced and instability is more frequent for NBCAR–Forest. At $n = 700$, NBCAR–Forest RMSE remains large (117.13–152.36 across ρ levels), whereas NBCAR–GRF maintains RMSE between 21.35 and 26.16 (Table 1). MAE shows a similar reduction, from 21.47–23.94 (NBCAR–Forest) to 9.07–11.16 (NBCAR–GRF) (Table 2). At $n = 1200$ with medium ρ, NBCAR–Forest exceeds the cap (RMSE > 500), while NBCAR–GRF remains finite (RMSE 24.63; MAE 10.36). Most notably, at $n = 2500$ NBCAR–Forest exceeds the RMSE threshold across all dependence levels ("> 500"), whereas NBCAR–GRF remains finite and relatively stable (RMSE 22.06–35.23; MAE 9.35–12.61). The consistent ranking under both RMSE and MAE suggests that NBCAR–GRF reduces both extreme errors (captured by RMSE) and typical absolute errors (captured by MAE) on the testing set.

The most challenging setting is $\text{var}(x) > \text{var}(\phi)$, where NBCAR–Forest exceeds the reporting threshold in every configuration for both RMSE and MAE (Tables 1–2). NBCAR–GRF, while still producing finite predictions, exhibits substantially higher errors compared with the previous regimes: RMSE ranges approximately from 189 to 399, and MAE ranges from about 65 to 111 across n and ρ. For example, at $n = 700$ RMSE increases from 193.84 (low ρ) to 259.59 (high ρ), and MAE increases from 64.66 to 83.10; at $n = 2500$, RMSE increases further to 214.18–399.07 and MAE to 74.86–111.15. These results indicate that when the covariate signal dominates and the spatial random effect is comparatively weaker under the simulation design, the prediction task becomes difficult for both approaches; however, NBCAR–GRF remains numerically stable under the capped reporting rule, whereas NBCAR–Forest frequently produces extreme errors.

Stronger spatial dependence (from low to high ρ) is generally associated with increased error whenever predictions remain finite. This is visible, for instance, in $\text{var}(x) < \text{var}(\phi)$ at $n = 700$, where NBCAR–Forest RMSE increases from 12.71 to 20.87 and NBCAR–GRF RMSE increases from 10.61 to 16.95 as ρ rises (Table 1); a similar monotone increase is observed in MAE (Table 2). The effect of sample size is more nuanced and should not be overstated as uniformly beneficial: in $\text{var}(x) \approx \text{var}(\phi)$, increasing n coincides with a marked deterioration in NBCAR–Forest stability (e.g., RMSE "> 500" at $n = 2500$), while NBCAR–GRF remains finite with only moderate changes in error.

3.1. Stability across replications

Beyond average testing-set errors, we assess numerical stability across the 50 replications by reporting the standard deviation (SD) and interquartile range (IQR) of RMSE and MAE within each configuration. Larger SD/IQR indicate greater run-to-run variability, which is consistent with less stable out-of-sample behaviour under repeated simulation draws. Entries shown as "> 500" indicate that the dispersion measure exceeds the reporting cap, consistent with repeated extreme errors or divergence under the corresponding configuration.

Table 3. Dispersion of RMSE across 50 replications (KNN neighbourhood). SD and IQR are computed from testing-set RMSE within each configuration.

Scenario	n	SD RMSE Forest			SD RMSE GRF			IQR RMSE Forest			IQR RMSE GRF		
		low	medium	high	low	medium	high	low	medium	high	low	medium	high
	700	3.40	16.94	9.43	2.77	5.77	8.28	4.32	5.31	12.38	2.70	4.06	7.07
$\text{var}(x) < \text{var}(\phi)$	1200	2.91	4.71	7.40	2.64	4.14	6.87	3.84	4.54	8.15	2.36	4.59	7.83
	2500	3.50	> 500	55.41	2.42	3.76	7.18	4.17	6.83	8.65	2.84	3.24	5.77
	700	> 500	349.32	351.11	13.91	8.60	12.36	20.02	26.83	29.73	7.76	10.61	15.34
$\text{var}(x) \approx \text{var}(\phi)$	1200	134.77	> 500	472.07	5.74	8.36	14.68	30.82	60.90	34.97	5.83	10.99	13.27
	2500	> 500	> 500	> 500	7.25	10.88	12.76	53.18	123.05	77.53	3.91	9.79	17.12
	700	> 500	> 500	> 500	107.44	72.24	138.78	> 500	> 500	> 500	115.26	86.19	143.66
$\text{var}(x) > \text{var}(\phi)$	1200	> 500	> 500	> 500	89.97	104.16	155.03	> 500	> 500	> 500	76.92	97.47	180.04
	2500	> 500	> 500	> 500	48.12	144.97	167.33	> 500	> 500	> 500	57.15	104.89	177.44

Table 4. Dispersion of MAE across 50 replications (KNN neighbourhood). SD and IQR are computed from testing-set MAE within each configuration.

Scenario	n	SD MAE Forest			SD MAE GRF			IQR MAE Forest			IQR MAE GRF		
		low	medium	high	low	medium	high	low	medium	high	low	medium	high
	700	1.10	2.41	2.03	0.73	1.05	1.64	1.77	1.41	2.65	0.99	1.36	2.19
$\text{var}(x) < \text{var}(\phi)$	1200	0.78	0.98	1.85	0.70	0.83	1.49	1.15	1.30	2.06	0.85	1.10	1.75
	2500	0.47	129.33	3.17	0.45	0.63	1.10	0.71	1.24	2.04	0.53	0.66	1.42
	700	83.47	33.54	32.82	2.22	1.94	3.17	4.90	6.58	8.17	2.71	2.90	4.57
$\text{var}(x) \approx \text{var}(\phi)$	1200	10.61	241.22	33.93	1.18	1.78	2.53	3.93	8.77	6.06	1.69	2.17	2.97
	2500	> 500	188.33	> 500	1.30	1.79	2.23	4.48	11.35	7.36	1.17	2.10	2.67
	700	> 500	> 500	> 500	18.82	17.21	27.75	463.00	> 500	> 500	24.68	19.16	36.77
$\text{var}(x) > \text{var}(\phi)$	1200	> 500	> 500	> 500	15.78	16.71	22.04	> 500	> 500	> 500	17.74	21.46	30.74
	2500	> 500	> 500	> 500	10.76	21.82	25.32	> 500	> 500	> 500	14.45	23.99	34.67

Tables 3–4 show that NBCAR–GRF generally yields smaller SD and IQR than NBCAR–Forest, indicating more stable predictive behaviour across repeated runs. The contrast is particularly pronounced in the $\text{var}(x) \approx \text{var}(\phi)$ regime, where NBCAR–Forest exhibits very large dispersion (often exceeding the reporting cap), whereas NBCAR–GRF remains finite with substantially smaller SD/IQR across spatial dependence levels. Even in the challenging $\text{var}(x) > \text{var}(\phi)$ regime, NBCAR–GRF retains finite dispersion measures, while NBCAR–Forest frequently shows extreme dispersion consistent with repeated divergence. For instance, under $\text{var}(x) \approx \text{var}(\phi)$ with $n = 700$ and medium ρ, the SD of RMSE is 349.32 for NBCAR–Forest versus 8.60 for NBCAR–GRF (Table 3), highlighting a substantial reduction in run-to-run variability.

Overall, the simulation evidence indicates that NBCAR–GRF improves predictive accuracy and exhibits substantially stronger numerical stability than NBCAR–Forest across the considered scenarios. The advantage is clearest when spatial variation is non-negligible relative to covariate variation ($\text{var}(x) \leq \text{var}(\phi)$), where NBCAR–GRF yields consistently smaller RMSE/MAE and avoids the severe spikes and threshold exceedances observed for NBCAR–Forest. When $\text{var}(x) > \text{var}(\phi)$, prediction remains difficult for both methods; however, NBCAR–GRF continues to provide finite predictions under the reporting rule, whereas NBCAR–Forest frequently exceeds the error threshold. Taken together, the consistent ranking under both RMSE and MAE supports the conclusion that NBCAR–GRF improves both typical error behaviour and robustness to extreme failures on the testing set within the scope of the simulation design.

4. Results of the empirical study

The empirical analysis examines wildfire hotspot counts over Kalimantan, Indonesia, in September 2024, represented on a regular $0.25^\circ \times 0.25^\circ$ grid. September was selected because hotspot activity typically peaks during the regional dry season, making this period informative for evaluating predictive performance under elevated fire risk. Hotspot counts were obtained from the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG). Meteorological predictors include air temperature, total precipitation, number of rainless days, wind speed, and relative humidity, all derived from the ECMWF–ERA5 reanalysis for the same period. Figure 2 visualises the spatial distribution of hotspot counts across the study region.

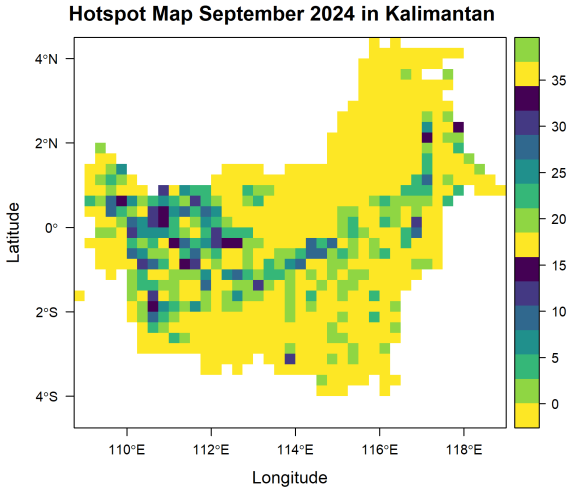


Figure 2. Wildfire hotspot counts in Kalimantan, September 2024.

Figure 2 shows substantial spatial heterogeneity in hotspot counts, characterised by pronounced clustering rather than a spatially uniform pattern. Higher hotspot intensities are primarily observed in a west-to-central corridor around 110–113°E near the equator, and in a north-eastern cluster extending toward coastal areas around 116–118°E. In contrast, several southern and interior areas exhibit consistently low counts. This clustering is consistent with residual spatial dependence in wildfire-related areal count data and motivates the inclusion of a CAR-type spatial random effect to account for dependence among neighbouring grid cells. Moreover, the sharp local contrasts suggest that hotspot intensity may respond nonlinearly to meteorological conditions and their interactions, supporting the use of flexible learners for the mean structure [15].

Beyond the presence of localised high-intensity patches, the map also indicates a large proportion of low-count grid cells, including broad areas with near-zero or consistently small values. This spatial sparsity, together with a small number of concentrated hotspots, is consistent with an overdispersed count process in which the variance exceeds the mean. This pattern provides empirical support for adopting a Negative Binomial likelihood rather than a Poisson specification, as the former can better accommodate heavy-tailed behaviour and heterogeneous variability across space.

Taken together, the observed spatial configuration suggests that a suitable model should capture both (i) residual spatial dependence among neighbouring cells (via a CAR-type random effect) and (ii) potentially complex, nonlinear relationships between hotspot intensity and covariates. Accordingly, this application provides an appropriate setting for the proposed hybrid formulation, which combines CAR-based dependence control with a GRF-based mean component for nonlinear covariate effects and interactions.

4.1. Predictive performance and variable importance

On the held-out testing set, NBCAR–GRF attains RMSE = 3.64 and MAE = 2.11 hotspots per grid cell. Variable-importance results are reported in Table 5. Air temperature is the most influential predictor (27.74%), followed by wind speed (20.34%) and the number of rainless days (20.17%). Precipitation (16.65%) and relative humidity (15.11%) also contribute to predictive performance, although with smaller marginal importance than the top three predictors.

Table 5. Variable importance from NBCAR–GRF for the September 2024 Kalimantan application.

Variable	**Importance (%)**
Air temperature	27.74
Wind speed	20.34
Number of rainless days	20.17
Precipitation	16.65
Relative humidity	15.11

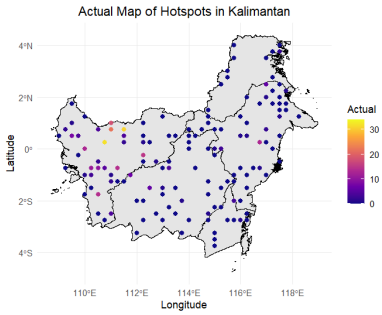
Table 5 indicates that predictive skill is not driven by a single meteorological factor. The top three variables (air temperature, wind speed, and number of rainless days) account for 68.25% of total importance, suggesting that the fitted model relies primarily on indicators of drying conditions and fire-spread potential. Nevertheless, precipitation and relative humidity together contribute a non-trivial 31.76%, indicating that short-term moisture constraints provide complementary information rather than being negligible. This balance aligns with the expectation that hotspot activity reflects both the build-up of flammability and the suppression of ignition and spread through rainfall and atmospheric moisture.

The similar importance values for wind speed and rainless days (20.34% versus 20.17%) suggest that these predictors play comparably sized roles in the fitted model for September 2024. The moderate contribution of precipitation (16.65%) relative to rainless days (20.17%) may reflect the distinction between immediate wetting effects and accumulated dry-spell conditions captured by the latter. While variable importance does not imply causality, the ranking provides a transparent summary of which meteorological inputs are most strongly used for out-of-sample prediction in this application and supports a meteorologically interpretable reading of the fitted model.

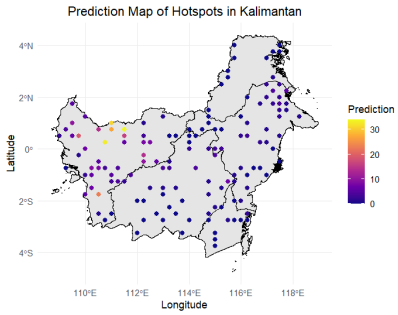
Substantively, the ranking is consistent with established wildfire mechanisms: higher temperatures accelerate fuel drying, extended dry spells increase effective drought duration, and stronger winds facilitate fire spread and ember transport. Precipitation and relative humidity act as moisture constraints that reduce ignition probability and limit fire propagation. This interpretation is consistent with prior evidence highlighting the roles of temperature, precipitation, humidity, and wind in wildfire hazard and hotspot variability, including studies relevant to Kalimantan [7, 14, 28, 31].

4.2. Spatial correspondence between observed and predicted hotspots

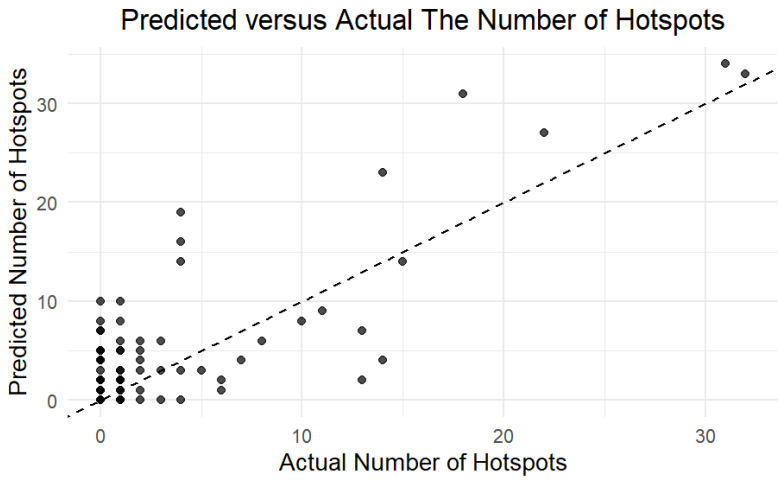
Figure 3 summarises empirical performance in terms of spatial agreement and pointwise accuracy. Panels (a) and (b) show that the predicted hotspot surface broadly reproduces the observed spatial structure: near-zero to low counts dominate much of Kalimantan, while a smaller number of grid cells exhibit moderate-to-high hotspot activity. Importantly, the predictions preserve the main spatial contrast between low-activity zones and concentrated pockets of higher activity, rather than producing an overly uniform surface. This is desirable for spatial risk screening, where identifying higher-risk clusters is often more important than fitting isolated extremes exactly.



(a) Observed hotspot counts.



(b) Predicted hotspot counts (NBCAR–GRF).



(c) Predicted versus observed hotspot counts on the testing set.

Figure 3. Observed and predicted hotspot patterns for Kalimantan (September 2024).

A closer comparison of panels (a) and (b) suggests that the model captures the *locations* of major activity corridors reasonably well, whereas matching the *intensity* of the most extreme cells is more challenging. This behaviour is expected in overdispersed areal count data where a small number of grid cells can attain unusually high values due to localised and transient conditions. In addition, CAR-type spatial random effects borrow information across neighbours and can smooth sharp peaks or redistribute part of an extreme signal across adjacent cells. As a result, predicted high-count areas may appear as small clusters rather than isolated spikes, improving spatial coherence while potentially attenuating the highest observed maxima.

Panel (c) provides a complementary pointwise diagnostic through the predicted-versus-observed scatter plot on the testing set. The dense concentration of points at low observed values reflects the prevalence of low-count cells, for which predictions remain close to the 45° reference line on average. This behaviour is consistent with the reported RMSE and MAE, indicating good accuracy in the low-to-moderate range that constitutes most observations. At higher observed counts, dispersion around the reference line increases, consistent with heteroscedastic prediction uncertainty typical of overdispersed counts with a heavy right tail.

Larger deviations from the 45° line occur primarily for grid cells with very high hotspot counts. In these extremes, the model shows occasional overprediction in some locations and underprediction in others. Such tail behaviour is common in wildfire applications because rare high-intensity events can reflect local combinations of prolonged dry spells, elevated

temperature, wind conditions, and unobserved site characteristics (e.g., land cover and human activity). Moreover, because the model incorporates spatial dependence through the CAR component, some extreme values may be partially smoothed by the neighbourhood structure. Despite these tail deviations, the fitted model generally preserves relative risk ordering: locations with higher observed counts tend to receive higher predicted values than those with lower observed counts.

Overall, Figure 3 indicates that NBCAR–GRF produces a coherent spatial reconstruction of hotspot patterns while achieving reasonable pointwise agreement on held-out cells. The combined evidence from the maps and the scatter plot supports the practical value of integrating (i) a Negative Binomial likelihood for overdispersed counts, (ii) CAR spatial random effects for residual areal dependence, and (iii) a GRF-based mean component for nonlinear covariate effects. From an operational perspective, the ability to spatially localise higher-risk clusters—even when uncertainty increases at the upper tail—is relevant for early warning screening and prioritisation of monitoring resources.

Finally, the alignment between the spatial patterns in Figure 3 and the predictor ranking in Table 5 provides an internally consistent descriptive interpretation: higher predicted values tend to occur in areas where meteorological covariates jointly indicate conditions favourable for ignition and spread. While this does not establish causality, it supports the descriptive validity of the fitted model for predictive hotspot mapping in Kalimantan.

5. Conclusion and discussion

This study proposes and evaluates NBCAR–GRF, a hybrid framework for overdispersed spatial count data that integrates (i) a Negative Binomial likelihood to accommodate overdispersion, (ii) CAR-type spatial random effects to represent structured areal dependence, and (iii) a Generalized Random Forest to capture nonlinear covariate effects and interactions. Across the simulation study, which varies sample size, spatial dependence strength, and the relative contribution of covariate signal and spatial random effects, NBCAR–GRF consistently yields lower prediction errors and improved numerical stability relative to the benchmark NBCAR–Forest.

Quantitatively, NBCAR–Forest exhibits extreme failures under a substantial fraction of configurations: across 27 settings, testing-set RMSE exceeds the reporting cap in 13/27 cases and MAE exceeds the cap in 9/27 cases (Tables 1–2). In contrast, NBCAR–GRF returns finite RMSE and MAE in all configurations. This stability advantage is also reflected in the dispersion across 50 replications. NBCAR–GRF generally produces markedly smaller SD and IQR of RMSE/MAE than NBCAR–Forest, particularly in regimes where covariate and spatial effects are of comparable magnitude, $\mathrm{Var}(X) \approx \mathrm{Var}(\phi)$ (Tables 3–4). For example, when $n = 700$ and ρ is medium in this regime, the SD of RMSE is 349.32 for NBCAR–Forest versus 8.60 for NBCAR–GRF (Table 3). Overall, performance gains are most pronounced when spatial effects are non-negligible relative to covariate effects ($\mathrm{Var}(X) \leq \mathrm{Var}(\phi)$), whereas prediction remains challenging when $\mathrm{Var}(X) > \mathrm{Var}(\phi)$; nonetheless, NBCAR–GRF retains finite errors and dispersion measures, while NBCAR–Forest frequently diverges.

In the empirical application to satellite-derived wildfire hotspot counts over Kalimantan in September 2024, using meteorological predictors on a $0.25^\circ \times 0.25^\circ$ grid, NBCAR–GRF achieves RMSE = 3.64 and MAE = 2.11 hotspots per grid cell on the held-out testing set. Variable-importance analysis identifies air temperature as the most influential predictor, followed by wind speed and the number of rainless days, consistent with established fire-weather mechanisms. These findings support the practical relevance of NBCAR–GRF for spatial hotspot mapping and early-warning-oriented risk screening.

Overall, this study makes two main contributions. First, it introduces NBCAR–GRF as an explicit integration of an overdispersed count likelihood, CAR-based spatial dependence, and GRF-based nonlinear learning within a single iterative modelling procedure. Second, it provides simulation and empirical evidence that replacing the forest component with GRF improves predictive accuracy and, crucially, reduces numerical instability relative to NBCAR–Forest under challenging configurations.

From an operational perspective, improved spatial prediction of hotspot counts can be translated into actionable early warning products. For instance, predicted mean hotspots can be converted into risk-ranked grids or threshold-based alerts that prioritise surveillance, patrol deployment, and rapid-response allocation under constrained resources. More reliable identification of emerging high-risk clusters—together with reduced unstable extreme predictions—can support earlier targeting of monitoring and mitigation efforts (e.g., directing ground checks, water resources, or firefighting readiness to predicted hotspots). In this way, NBCAR–GRF can strengthen climate-related risk information and decision support aligned with Sustainable Development Goal 13 on climate action.

5.1. Limitations and future work

The current evidence has clear limitations. The iterative procedure uses an a priori iteration budget R (with no fully data-driven stopping rule), uncertainty quantification is limited, the likelihood does not explicitly accommodate zero inflation, and computational demands may increase with finer spatial resolution or longer time windows. In addition, the empirical demonstration is restricted to a single-month case study, limiting temporal generalisability.

First, the current implementation fixes the iteration count R a priori, so efficiency and reproducibility may depend on this choice. Second, the empirical analysis covers a single month (September 2024), which restricts generalisability across seasons and interannual variability. Third, although GRF is partly motivated by its inferential properties, the present evaluation focuses on predictive accuracy and numerical stability; inferential outputs (e.g., local effect summaries with uncertainty) are deferred to future work. Fourth, the model assumes a NB count process without an explicit zero-inflation mechanism, and calibrated predictive uncertainty (e.g., prediction intervals) is not yet fully developed. Finally, computational cost may become substantial for finer grids or extended spatio-temporal windows.

Several extensions follow naturally. First, a data-driven stopping rule could be developed based on convergence of spatial effects, stabilisation of out-of-sample error, or changes in the GRF mean component. Second, the framework could be extended to zero-inflated or hurdle variants to separately model hotspot occurrence and intensity. Third, future work can more directly leverage GRF's inferential capabilities by reporting local effect summaries and associated uncertainty. Fourth, calibrated prediction intervals could be added via posterior predictive distributions in the NB–CAR stage, conformal prediction adapted to counts, or other calibrated approaches. Finally, broadening the empirical scope to multi-month or multi-year settings would enable a more rigorous assessment of temporal robustness and operational value for early warning screening.

Acknowledgements

The authors gratefully acknowledge the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG) for providing the dataset utilized in this research.

Data availability. The data that support the findings of this study are available upon reasonable request from the corresponding author.

Author contributions. All the co-authors have contributed equally in all aspects of the preparation of this submission.

Conflict of interest statement. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This study was financially supported by the Indonesian Education Scholarship [BPI ID: 202209090500], Center for Higher Education Funding and Assessment, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, and Endowment Fund for Education Agency, Ministry of Finance of the Republic of Indonesia.

References

[1] H.A. Abood, S. Khatan Ismail and A.F.T. Sabah, Re-weighted least squares: the best negative binomial regression methods in determining the congenital anomalies' risk factors, *Syst. Rev. Pharm.* **11** (12), 1956-1963, 2020.

[2] G. Alomair, Predictive performance of count regression models versus machine learning techniques: a comparative analysis using an automobile insurance claims frequency dataset, *PLoS One* **19** (12), 1-12, 2024.

[3] J. Argota Sánchez-Vaquerizo and D. Helbing, Spatial econometrics to estimate traffic reduction by transforming office space into housing and other land uses: the case for Barcelona, *Cities* **162**, 1-21, 2025.

[4] S. Athey, J. Tibshirani and S. Wager, Generalized random forests, *Ann. Stat.* **47** (2), 1148-1178, 2019.

[5] I. Azis, A. Djuraidah, M. Aidi and A. Sopaheluwakan, Spatial prediction of the number of wildfire hotspots using negative binomial conditional autoregressive random forest, *Glob. J. Environ. Sci. Manag.* **11** (4), 1-18, 2025.

[6] L. Breiman, Random forests, *Mach. Learn.* **45**, 5-32, 2001.

[7] M. Chen, J. Liu, B. Chu, D. Zhao, R. Li, T. Chen, Q. Ma, Y. Wang, P. Zhang, H. Li and H. He, Research on the influencing factors of PM2.5 in China at different spatial scales based on machine learning algorithm, *Environ. Int.* **200**, 1-10, 2025.

[8] K.A. Coskuner, Assessing the performance of MODIS and VIIRS active fire products in the monitoring of wildfires: a case study in Turkey, *iForest* **15** (2), 85-94, 2022.

[9] İ. Daşdemir, F. Aydın and M. Ertuğrul, Factors affecting the behavior of large forest fires in Turkey, *Environ. Manage.* **67** (1), 162-175, 2021.

[10] D.J. Fitri, A. Djuraidah and H. Wijayanto, Bayesian conditional negative binomial autoregressive model: a case study of stunting on Java Island in 2021, *Commun. Math. Biol. Neurosci.* **2024**, 1-17, 2024.

[11] J.K. Frank, T. Suesse and A. Brenning, An assessment of spatial random forests for environmental mapping: the case of groundwater nitrate concentration, *Environ. Model. Softw.* **193**, 1-14, 2025.

[12] L. Giglio, J. Descloitres, C.O. Justice and Y.J. Kaufman, An enhanced contextual fire detection algorithm for MODIS, *Remote Sens. Environ.* **87** (2), 273-282, 2003.

[13] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita and J.-N. Thépaut, The ERA5 global reanalysis, *Q. J. R. Meteorol. Soc.* **146** (730), 1999-2049, 2020.

[14] M.W. Jones, J.T. Abatzoglou, S. Veraverbeke, N. Andela, G. Lasslop, M. Forkel, A.J.P. Smith, C. Burton, R.A. Betts, G.R. van der Werf, S. Sitch, J.G. Canadell, C. Santín, C. Kolden, S.H. Doerr and C. Le Quéré, Global and regional trends and drivers of fire under climate change, *Rev. Geophys.* **60** (3), 1-76, 2022.

[15] W.S. Karurung, K. Lee and W. Lee, Assessment of forest fire vulnerability prediction in Indonesia: seasonal variability analysis using machine learning techniques, *Int. J. Appl. Earth Obs. Geoinf.* **138**, 1-16, 2025.
[16] J. Kim, A.B. Lawson, B. Neelon, J.E. Korte, J.M. Eberth and G. Chowell, Evaluation of Bayesian spatiotemporal infectious disease models for prospective surveillance analysis, *BMC Med. Res. Methodol.* **23** (1), 1-13, 2023.
[17] A. Kmoch, C.T. Harrison, J. Choi and E. Uuemaa, Spatial autocorrelation in machine learning for modelling soil organic carbon, *Ecol. Inform.* **86**, 1-21, 2025.
[18] F. Lu, G. Zhang, T. Wang, Y. Ye and Q. Zhao, Geographically weighted random forest based on spatial factor optimization for the assessment of landslide susceptibility, *Remote Sens.* **17** (9), 1-28, 2025.
[19] L. Luo, D. Apps, S. Arcand, H. Xu, M. Pan and M. Hoerling, Contribution of temperature and precipitation anomalies to the California drought during 2012–2015, *Geophys. Res. Lett.* **44** (7), 3184-3192, 2017.
[20] C. MacBride, V. Davies and D. Lee, A spatial autoregressive random forest algorithm for small-area spatial prediction, *Ann. Appl. Stat.* **19** (1), 485-504, 2025.
[21] Y.C. MacNab, Adaptive Gaussian Markov random field spatiotemporal models for infectious disease mapping and forecasting, *Spat. Stat.* **53**, 1-23, 2023.
[22] T.W. Mas'oed, S. Nurdiati, A. Sopaheluwakan, M.K. Najib and A. Salsabila, Modeling fire hotspots in Kalimantan, Indonesia using nested 3-copula regression based on precipitation and dry days during different ENSO phases, *Geographia Technica* **19** (2), 264-281, 2024.
[23] P. McCullagh and J.A. Nelder, *Generalized Linear Models*, 2nd ed., Chapman & Hall/CRC, 1989.
[24] C.A. Mushagalusa, A.B. Fandohan and R. Glèlè Kakaï, Random forest and spatial cross-validation performance in predicting species abundance distributions, *Environ. Syst. Res.* **13** (1), 1-24, 2024.
[25] F. Mutiso, J.L. Pearce, S.E. Benjamin-Neelon, N.T. Mueller, H. Li and B. Neelon, Bayesian negative binomial regression with spatially varying dispersion: modelling COVID-19 incidence in Georgia, *Spat. Stat.* **52**, 1-19, 2022.
[26] M. Pascolini-Campbell, J.B. Fisher, K. Cawse-Nicholson, C.M. Lee and N. Stavros, Assessment of spatial autocorrelation and scalability in fine-scale wildfire random forest prediction models, *Sci. Rep.* **15** (1), 1-15, 2025.
[27] Z. Qin, Q. Peng, C. Jin, J. Xu, S. Xing, P. Zhu and Y. Yang, Geographically weighted random forest fusing multi-source environmental covariates for spatial prediction of soil heavy metals, *Environ. Pollut.* **385**, 2025.
[28] A. Schmidt, L.M. Ellsworth, G.A. Boisen, N. Novita, A. Malik, A. Gangga, I. Albar, A.D. Nurhayati, R.P. Ritonga, A. Asyhari and J.B. Kauffman, Fire frequency, intensity, and burn severity in Kalimantan's threatened peatland areas over two decades, *Front. For. Glob. Change* **7**, 1-16, 2024.
[29] J. Stoklosa, R.V. Blakey and F.K.C. Hui, An overview of modern applications of negative binomial modelling in ecology and biodiversity, *Diversity* **14** (5), 1-25, 2022.
[30] S. Ullah, M.A.K. Barakzai and T. Xie, Performance of a negative binomial-GLM in spatial scan statistic: a case study of low-birth weights in Pakistan, *Geospat. Health* **19** (2), 2024.
[31] A. Usup and H. Hayasaka, Peatland fire weather conditions in central Kalimantan, Indonesia, *Fire* **6** (5), 1-12, 2023.
[32] Y. Wang, X. Chen and F. Xue, A review of Bayesian spatiotemporal models in spatial epidemiology, *ISPRS Int. J. Geo-Inf.* **13** (3), 1-29, 2024.
[33] Z. Xu, D. Liu and L. Yan, Temperature-based fire frequency analysis using machine learning: A case of Changsha, China, *Clim. Risk Manag.* **31**, 100276, 2021.