https://doi.org/10.56093/JISAS.V80I3.2
Available online at https://isas.live/newsite
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 80(3) 2026 411–418

Classical vs Bayesian Approaches to Modeling Overdispersed Count Data: Case Study of Tomato Leaf Curl

Briefully Kharnongrum[1], Kimthianngak[1], Nidahamteinsar Papang[1], Priti Pradhan[1], Biwash Gurung[2] and Ramnunpuia[1]

[1]North Eastern Hill University, Shillong
[2]G.D. Goenka University, Gurugram

Received 18 June 2026; Revised 01 July 2026; Accepted 29 July 2026

SUMMARY

Count data consisting of non-negative integer values arising from counting events require specific statistical methods for analysis. This study compares Classical and Bayesian approaches for estimating count regression models using Tomato Leaf Curl (TLC) data consisting of 809 observations. It is found that the simple Poisson regression model performs poorly due to over dispersion, where the variance greatly exceeds the mean. The Negative Binomial model, which handles over dispersion, provides a good fit to the data under both approaches. Both frameworks produced nearly the same parameter estimates confirming the consistency of the results. However, the Bayesian framework provided a full posterior distributions enables posterior predictive checks supports Leave-One-Out Cross-Validation (LOO-CV) model comparison, providing a richer inferential tool for uncertainty. Zero-inflated Poisson model is also used for count data modelling but it is not applicable here due to the absence of zero counts in the data. Overall, the Bayesian framework provides a broader view of uncertainty around the estimates.

Keywords: Count data models; Poisson model; Negative Binomial model; Classical estimation; Bayesian estimation; Markov Chain Monte Carlo (MCMC); Over dispersion; Tomato Leaf Curl (TLC); Leave-One-Out Cross-Validation (LOO-CV).

1. INTRODUCTION

In statistics, count data refers to data that record countable quantities. These are non-negative integers — 0, 1, 2, 3, and so on — that come from counting events, not from ranking or scoring them [Cameron & Trivedi,2013]. Analysing count data is different from analysing binary data, where each observation is only 0 or 1, and from ordinal data, which uses integers on an arbitrary scale where only the order matters [McCullagh & Nelder,1989]. Count data usually pile up around a few small integer values. Count outcomes show up often in real-world applications. Fields like epidemiology, public health, insurance and social sciences frequently use count variables to represent how many times an event occurs within a set time period [Cameron & Trivedi,2013].

Count data not optimally modelled with a normal distribution as the assumptions of the ordinary least-squares regression are violated (homoscedasticity, normality and linearity), the use of statistical techniques generally causes biased and inefficient results [Fox, 2016]. In such cases, the Poisson and Negative Binomial (NB) regression model are commonly used for modeling outcome and explanatory variable[Dobson & Barnett, 2018]. The Poisson regression model assumes equal mean and variance [McCullagh & Nelder, 1989]. However, this is unrealistic in real world application. The Negative Binomial extends Poisson regression, provides greater flexibility to model the relationship between the conditional mean and variance, unlike the more restrictive Poisson model [Cameron & Trivedi, 2013]. When data shows excess zeros and over dispersion – where the sample variance exceeds

Corresponding author: Briefully Kharnongrum
E-mail address: dragokharnongrum@gmail.com

the mean- zero-inflated and hurdle models are more appropriate [Lambert, 1992]. The Zero-Inflated Poisson (ZIP) model, introduced by Lambert (1992) for quality control data, was found to outperform both the Negative Binomial and Poisson models.

In recent years, Bayesian methods have gained considerable attention for parameter estimation in count data models [Gelman *et al.* , 2013]. Bayesian estimation applies Bayes' theorem to updates prior beliefs with observed data, combining prior information and the likelihood function to obtain posterior distributions [Gelman *et al.* , 2013]. Bayesian estimation is flexible, allowing incorporation of external information or expert knowledge into the estimation process but requires complex computations like Markov Chain Monte Carlo (MCMC), especially for large datasets [Albert, 2009]. It is particularly useful for complex count data with over dispersion, excess zeros or time dependence [Lunn *et al.* ,2013]. Software like R enables Bayesian parameters estimation even when analytical solution are unavailable. Tomato Leaf Curl virus (ToLCV), transmitted by the whitefly *Bemisiatabaci*, is is the most economically significant biotic stresses affecting tomato cultivation in India and across South Asia. In West Bengal, ToLCV causes extensive crop losses and poses a persistent challenge to growers [Gurung *et al.* ,2022]. Quantitative modelling of disease incidence counts provide actionable insights for disease monitoring and management. Although biologically meaningful predictors such as temperature, humidity and vector population density are known drivers of ToLCV incidence, structured longitudinal datasets linking these variables with incidence counts at sufficient scale remain limited in the published literature.

While the Negative Binomial model outperforms the Poisson for over dispersion count data, this study offers three novel contributions. First, this study analyzes real agricultural data – weekly TLC disease counts from West Bengal –to demonstrate the influence of six weather factors on disease occurrence. Such studies using agricultural are rare. Second, it compares classical and Bayesian approaches for Poisson model and Negative Binomial model. This shows the extra benefits of Bayesian methods by providing complete probability distribution, better prediction checks and model comparison using LOO-CV. Lastly, it demonstrates that applying the Poisson model to over dispersion count data can produce misleadingly narrow confidence intervals, making results appear more precise than they are. However, the Bayesian Negative Binomial model handles uncertainty more realistically making it useful for disease prediction and pest management decision.

2. LITERATURE REVIEW

The formal study of count data models started in 1972 when John Nelder and Robert Wedderburn introduced the Generalized Linear Model (GLM) framework, establishing Poisson regression as a standard for counts events. The GLM structure uses link functions and provides a unified way to fit different distributions to count data. To handle over dispersion — where variance exceeds the mean —Lawless introduced the Negative Binomial regression model in 1987. Cameron Trivedi's 1998 book provided a comprehensive overview of Poisson, Negative Binomial, Hurdle, and zero-inflated models, focusing on practical issues like over dispersion and excess zeros. Bayesian methods for count data gained traction as computational tools advanced. Gelfand and Smith (1990) demonstrated the use of Markov Chain Monte Carlo (MCMC) for complex models, with Gelman *et al.* (2013) further developing Bayesian regression techniques. Congdon (2005) and Minka (2003) contributed to Bayesian applications and prior selection for count models. The Negative Binomial model remains central for over dispersed counts and is preferred in practice for its flexibility.

When data exhibit excess zeros, as in many health and epidemiological studies, standard models struggle. Lambert (1992) introduced the Zero-Inflated Poisson (ZIP) model to address this, and Hall (2000) added random effects for hierarchical data. MCMC methods became essential for Bayesian estimation, with tools like JAGS (Plummer, 2003) facilitating practical application.

Bayesian count models are widely used in health sciences and epidemiology for disease mapping and spatial analysis. Researchers like Best *et al.* (2005), Lawson (2013), and Banerjee *et al.* (2014) have advanced hierarchical Bayesian approaches for spatial and grouped count data, enabling robust modeling of clustering and spatial dependence.

3. DATA DESCRIPTION AND EXPLORATORY ANALYSIS

The dataset consists of 809 weekly observations of Tomato Leaf Curl (TLC) disease cases with six other weather-related factors that may affect the disease such as Maximum Temperature (MaxTemp), Minimum Temperature (MinTemp), Mean Relative Humidity (MeanRH), Rainfall, Sunshine hours and Wind speed. The values are essentially non-negative integers which makes them appropriate for count regression modelling.

4. MATERIALS AND METHODS

4.1 Data Source

Field experiments were conducted at experimental plots in research station farm of Bidhan Chandra Krishi Viswavidyalaya (BCKV), Kalyani, Nadia (Elevation: 16m, Latitude: 22° 59' 15.2'' and Longitude: 88° 27' 26.6''). Tomato crops were raised for eight years i.e. 2011-2012 to 2018-2019 and tomato leaf curl disease incidence was recorded periodically at 7 days interval and the number of leaf curl infected plants was noted by visual observation and accordingly percent disease incidence was calculated.

4.2 Classical (Maximum Likelihood) Estimation

Using R software, two classical count regression models were estimated by maximum likelihood. The classical Poisson model was first used to study the TLC count data and it specifies

$$\log(\mu) = \alpha + \beta_1 MaxTemp_std + \ldots + \beta_6 Wind_std$$

assuming mean = var (Y) = μ.

The classical Negative Binomial model handles over dispersion by adding a dispersion parameter θ, yielding

$$Var(Y) = \mu + \frac{\mu^2}{\theta}$$

A Zero-Inflated Poisson model was also considered but it is not applicable due to the absence of zeros.

Over dispersion was tested using Cameron- Trivedi test and the likelihood ratio test, which compared Poisson and Negative Binomial model. And the model adequacy was assessed using Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and log-likelihood.

4.3 Bayesian (MCMC) Estimation

Both the classical models were estimated using Hamiltonian Monte Carlo through the rstanarm package. Here, a weakly informative Normal (0, 10) priors were assigned to all regression coefficients and intercept. The models were run using four chains each with 5,000 iterations and the first 1,000 iterations were used for warmup, producing 16,000 posterior samples for analysis. Convergence was assessed using the $\hat{R}$-hat statistic (threshold < 1.05) and effective sample size (n_eff). And no thinning was applied. Posterior predictive checks (PPC) were also used to assess model adequacy. Lastly, the Bayesian framework used Leave-One-Out Cross-Validation (LOO-CV) and Pareto-k diagnostics to check which model gives better prediction.

5. CLASSICAL (MLE) RESULTS

5.1 Classical Poisson Regression

The Classical Poisson model showed that Minimum Temperature, Sunshine and Wind had an impact on disease incidence [MinTemp_std (b = -0.344), Sunshine_std (b = 0.079) and Wind_std (b = -0.045)].

However, Maximum Temperature and Mean Relative Humidity had some effect but not strong enough to be considered [MaxTemp_std (p = 0.074) and MeanRH_std (p = 0.087)].

Lastly, Rainfall had no effect.

The Cameron–Trivedi over dispersion test produces a dispersion parameter of 9.18(z = 18.45, p < 0.001), strongly violating the equi dispersion assumption. Also, the residual deviance of 7,516.8 on 802 degrees of freedom and a chi-squared goodness-of-fit p-value ≈ 0 with AIC value of 11,680.62 indicates that the model does not fit the data well.

5.2 Classical Negative Binomial Regression

The Classical NB model showed moderate over dispersion in the data, estimating theta (θ) = 3.360 (SE = 0.186). Here, Minimum Temperature, Sunshine and Wind had a clear effect on the disease incidence [MinTemp_std (b = -0.351, p < 0.001), Sunshine_std (b = 0.075, p = 0.001) and Wind_std (b = -0.078, p = 0.006)].

However, maximum Temperature, Mean Relative Humidity and Rainfall does not have any effect on the disease incidence.

The Negative Binomial model with a likelihood ratio (LR) test yielding LR = 4812.42 ($p < 0.001$) and AIC value dropping from 11,680.62 to 6,870.20 performs better than the Poisson model. Hence, the Negative Binomial model is a better fit for the observed data.

Table 2. Classical Model Comparison: Log-Likelihood, AIC, and BIC

Model	Log-Likelihood	AIC	BIC	df
Classical Poisson	-5833.31	11680.62	11713.49	7
Classical Neg. Binomial	-3427.10	6870.20	6907.77	8

5.3 Classical Diagnostic Plots

Fig. 2 presents residual diagnostics for both classical models. The Poisson model displays marked heteroscedasticity in Pearson residuals (sum of squared Pearson residuals = 7,426.77). The NB model residuals are substantially better behaved (sum = 754.12), confirming adequate accommodation of over dispersion.

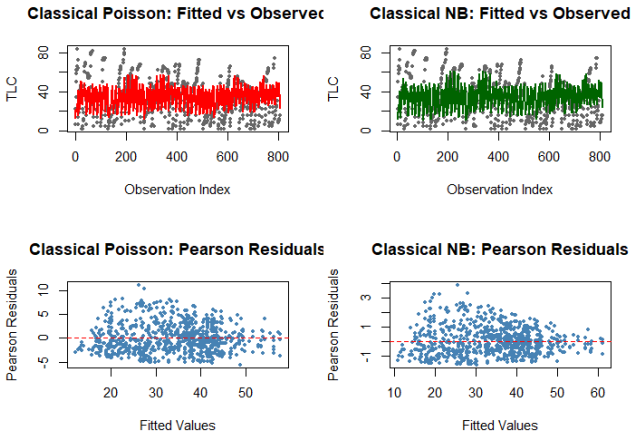


Fig. 2. Classical model diagnostics. Top: fitted values (lines) versus observed TLC counts (grey points) for Poisson (red) and NB (green). Bottom: Pearson residual plots

6. BAYESIAN (MCMC VIA STAN) RESULTS

6.1 Prior Specification and MCMC Convergence

From Bayesian model checks, we get $\hat{R}$-hat values which is equal to 1 or 1.001 (threshold < 1.05) and the effective sample sizes were very large (ranging from 4,212 to 20,907). Also, the Monte Carlo errors were very small to the point that they are negligible (<0.003). This shows excellent convergence and sampling efficiency.

6.2 Bayesian Poisson Regression

The Bayesian Poisson model yielded a results similar to the Classical Poisson model. It shows that the Minimum Temperature, Sunshine and Wind had clear impact on TLC disease incidence since their credible ranges does not include zero, making them an important predictors [MinTemp_std had a posterior mean of -0.344 (95% CI: -0.379 to -0.309), Sunshine_std of 0.079 (95% CI: 0.064 to 0.093) and Wind_std of -0.045 (95% CI: -0.062 to -0.027)] . However, the Maximum Temperature, Mean Relative Humidity and Rainfall impacts were uncertain since their credible intervals included zero.

6.3 Bayesian Negative Binomial Regression

The Bayesian Negative Binomial also gave results similar to Classical model. It estimated a reciprocal parameter of 3.310 (SD = 0.185; 95% CI: 2.959–3.685) which remained consistent with the classical theta estimate of 3.360. Here, the Minimum Temperature, Sunshine and Wind again had clear impact on the TLC disease count with all three credible intervals excluding zero [MinTemp_std had a posterior mean of -0.353 (95% CI: -0.476 to -0.235), Sunshine_std of 0.076 (95% CI: 0.030 to 0.123) and Wind_std of -0.079 (95% CI: -0.137 to -0.020].

Table 3. Classical Regression Coefficient Estimates (Poisson and Negative Binomial)

Model	Predictor	Estimate	SE	p-value	IRR (95% CI)	Model	Estimate	SE	p-value	IRR (95% CI)
Classical Poisson	MaxTemp_std	0.0328	0.0183	0.074	0.997–1.071	Classical Neg. Binomial	0.0236	0.0629	0.707	0.905–1.159
	MinTemp_std	-0.344	0.0179	< 0.001***	0.684–0.734		-0.3513	0.0607	< 0.001***	0.625–0.792
	MeanRH_std	-0.02	0.0119	0.087	0.957–1.003		-0.0311	0.0405	0.442	0.894–1.051
	Rainfall_std	0.0077	0.0157	0.623	0.977–1.039		-0.0172	0.053	0.746	0.885–1.092
	Sunshine_std	0.0788	0.0072	< 0.001***	1.067–1.097		0.0748	0.0231	0.001**	1.030–1.129
	Wind_std	-0.045	0.0088	< 0.001***	0.940–0.973		-0.0784	0.0288	0.006**	0.872–0.980

*Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. IRR = Incidence Rate Ratio with 95% profile likelihood confidence interval. Predictors are standardised (mean = 0, SD = 1). Shaded rows: Classical Negative Binomial.*

Table 4. Bayesian Posterior Summaries: Poisson and Negative Binomial Models

Model	Predictor	Post. Mean	Post. SD	95% Credible Interval	Model	Post. Mean	Post. SD	95% Credible Interval
Bayesian Poisson	MaxTemp_std	**0.032**	0.0181	-0.003 – 0.068	Bayesian Neg. Binomial	0.0251	0.0635	-0.097 – 0.152
	MinTemp_std	**-0.344**	0.0178	-0.379 – -0.309*		-0.3531	0.0614	-0.476 – -0.235*
	MeanRH_std	**-0.02**	0.0119	-0.044 – 0.003		-0.0314	0.0407	-0.112 – 0.049
	Rainfall_std	**0.008**	0.0156	-0.023 – 0.038		-0.0162	0.054	-0.122 – 0.090
	Sunshine_std	**0.079**	0.0072	0.064 – 0.093*		0.0762	0.0237	0.030 – 0.123*
	Wind_std	**-0.045**	0.0089	-0.062 – -0.027*		-0.0785	0.0299	-0.137 – -0.020*

*Note: * 95% credible interval excludes zero. Post. Mean = posterior mean; Post. SD = posterior standard deviation. Predictors are standardised. Shaded rows: Bayesian Negative Binomial.*

7. DIAGNOSTIC AND POSTERIOR PREDICTIVE CHECKS

7.1 Posterior Predictive Checks

Fig. 3 assesses whether replicated data generated from each model's posterior resembles the observed TLC distribution (dashed black line). As seen below, the Bayesian Negative Binomial model performs well since the posterior draws are very close to the observed TLC curve. This captured both the location and large variability in the data. On the other hand, the Bayesian Poisson posterior draws are too narrow which severely underestimates the spread of the data.

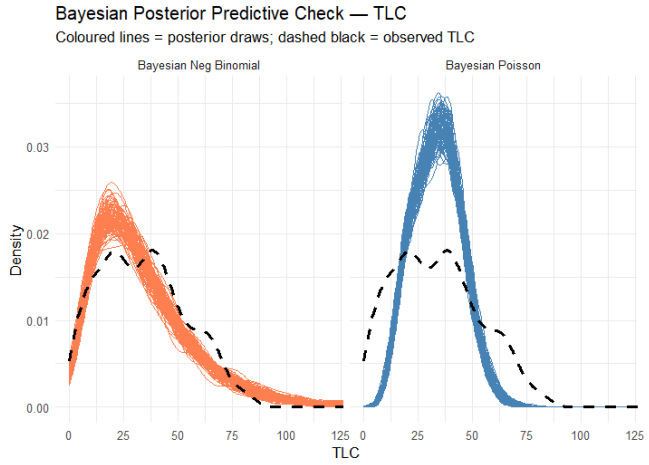


Fig. 3. Bayesian Posterior Predictive Check. Coloured lines show 100 posterior predictive draws; dashed black line shows the observed TLC density. The NB model (left) substantially outperforms the Poisson model (right) in capturing observed variability

7.2 Observed vs Posterior Predicted TLC

Fig. 4 shows observed vs Posterior Predicted TLC values. The Negative Binomial shows prediction that are scatter around the diagonal line, giving a more realistic picture of uncertainty. The Poisson model gives predictions that are too compressed, failing to capture the full spread of the data.

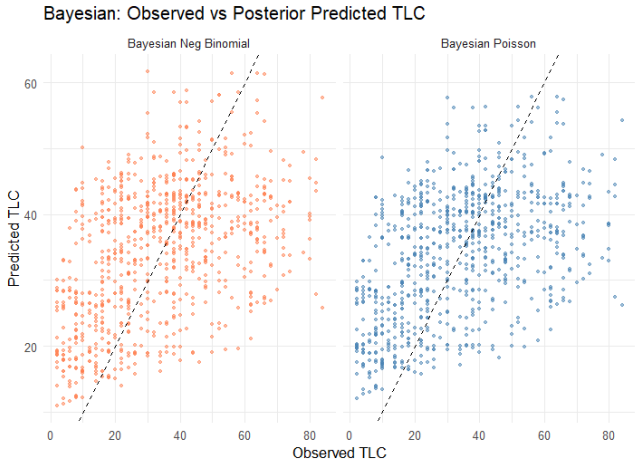


Fig. 4. Observed vs posterior mean predicted TLC for the Bayesian Negative Binomial (left) and Bayesian Poisson (right) models. The dashed diagonal represents perfect prediction

7.3 Posterior Distributions of Predictor Coefficients

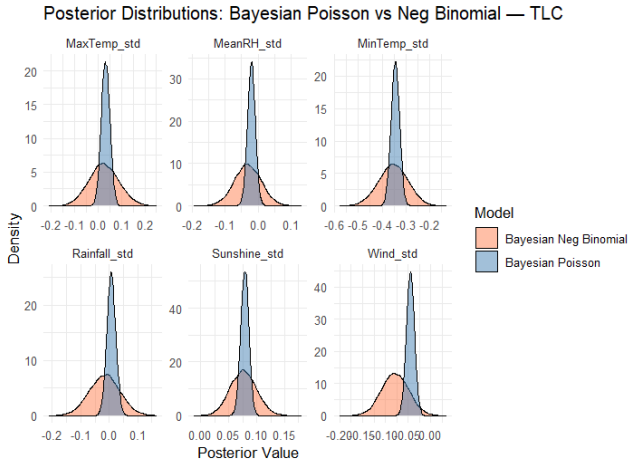


Fig. 5. Posterior distributions of standardised predictor coefficients for Bayesian Poisson (blue) and Bayesian Negative Binomial (coral) models

Fig. 5 overlays the posterior distributions for all six standardised predictors. The Bayesian NB posteriors (coral) are substantially wider than the Bayesian Poisson posteriors (blue) for all predictors, reflecting the Poisson model's artificially inflated precision. This is especially clear for Minimum Temperature (MinTemp_std), Sunshine (Sunshine_std) and Wind (Wind_std).

7.4 Forest Plot: Posterior Means and 95% Credible Intervals

The Minimum Temperature has the strongest negative effect since the credible intervals excluded zero in both Bayesian models. Sunshine shows a moderate positive effect and Wind showing a moderate negative effect. The Bayesian Negative Binomial has a wider credible interval which shows more uncertainty, correctly reflecting the variation in the data.

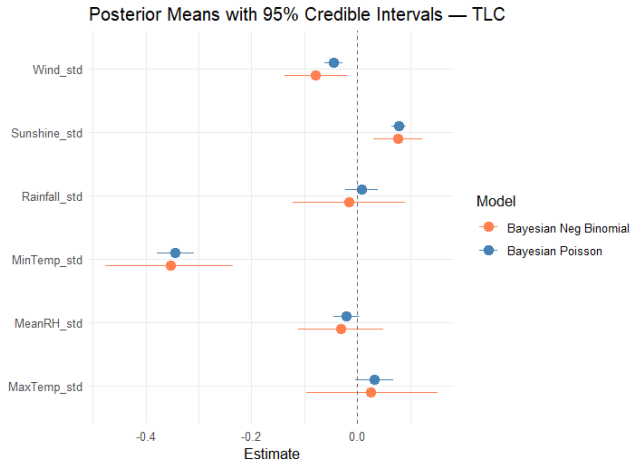


Fig. 6. Forest plot of posterior means and 95% credible intervals for all standardised predictors under Bayesian Poisson (blue) and Bayesian Negative Binomial (coral) models

8. BAYESIAN MODEL COMPARISON: LOO-CV

Leave-One-Out Cross-Validation (LOO-CV) estimates each model's expected log predictive density (ELPD) for new observations. Both models had reliable prediction results since all Pareto-k diagnostic values were below 0.7.

Table 5. Bayesian LOO-CV Model Comparison

Model	ELPD	SE (ELPD)	LOOIC	SE (LOOIC)
Bayesian Poisson	-5865.8	158.6	11731.6	317.2
Bayesian Neg. Binomial	-3434.7	19.8	6869.4	39.5
Difference (NB − Poisson)	+2431.1	143.9	—	—

Note: ELPD = Expected Log Predictive Density (higher is better); LOOIC = LOO Information Criterion (lower is better). Difference row reports Bayesian NB minus Bayesian Poisson.

The Bayesian Negative Binomial model has a higher ELPD value (closer to 0) than the Poisson model (-3,434.7 vs. -5,865.8) with a difference of +2,431.1 ELPD units. Also, the LOOIC value is much lower. Hence, the Bayesian Negative Binomial model is a better choice for the observed TLC data.

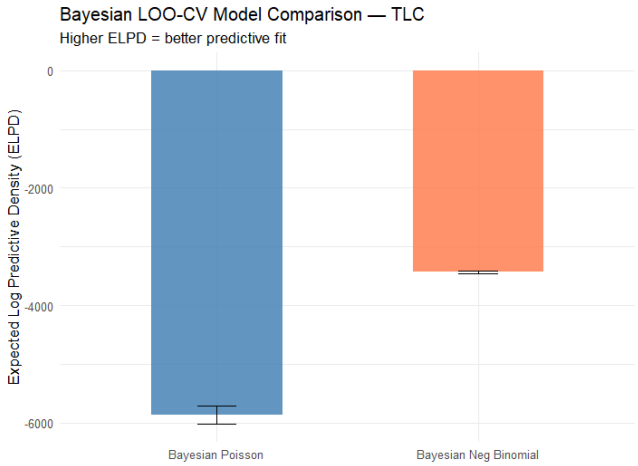


Fig. 7. Bayesian LOO-CV comparison. Higher (less negative) ELPD indicates better predictive performance. Error bars show ±1 SE. The Bayesian Negative Binomial model substantially outperforms Bayesian Poisson

Minimum temperature was the most influential factor, with higher values reducing Tomato Leaf Curl (TLC) incidence by about 29-30% ($p < 0.001$) across all models. Moderate increases in sunshine hours led to a slight rise in TLC cases (7-8%), while higher wind speed consistently reduced incidence by 4-8%. Maximum temperature, mean relative humidity, and rainfall had no significant effects, with small, uncertain impacts. Additionally, high correlation between humidity and rainfall made it difficult to distinguish their individual effects.

9. CONCLUSION

This study underscores the importance of choosing suitable statistical models for over dispersed count data. Using 809 Tomato Leaf Curl (TLC) disease observations from West Bengal, both classical and Bayesian analyses showed that the Negative Binomial model significantly outperformed the Poisson model, as indicated by lower AIC and higher ELPD values. The consistency between classical and Bayesian estimates reinforces the results' reliability. Notably, the Bayesian framework offered richer inference by highlighting the Poisson model's tendency to underestimate uncertainty, whereas the Bayesian Negative Binomial model provided more realistic credible intervals and improved predictive accuracy.

Among weather variables, minimum temperature had the strongest negative association with TLC incidence, while sunshine hours had a moderate positive effect and wind speed a moderate negative effect. Maximum temperature, humidity, and rainfall showed no significant effects, likely due to multicollinearity. These results are valuable for TLC disease forecasting. The Bayesian Negative Binomial model's ability to deliver both point predictions and uncertainty estimates makes it well-suited for early warning systems and agricultural planning. Future research should incorporate temporal dependencies, vector population data, and multi-site observations to further enhance model accuracy and generalizability.

ACKNOWLEDGEMENTS

The authors wish to express their sincere thanks to ICAR–National Centre for Integrated Pest Management and the National Innovations in Climate Resilient Agriculture (NICRA) project for making the dataset available for this research.

DECLARATIONS

- **Conflict of Interest**

  The author declares that he has no conflict of interest.

- **Data Availability**

  The dataset analysed during the current study is publicly available at 10.54386/jam.v24i4.1818.

- **Code Availability**

  The R code used in this study was developed by the authors and is available at GitHub repository. **https://github.com/BraveDi/Count-time-series-Bayesian-**

- **Ethical Approval**

  Ethical approval was not required for this study because it did not involve human participants, animals, or collection of identifiable personal information.

- **Clinical Trial Number**

  Not Applicable. This study did not involve a clinical trial.

- **Consent To Publish**

  Not applicable. This study did not involve the recruitment or participation of human subjects.

- **Consent To Participate**

  This study did not involve the recruitment or participation of human subjects.

- **Funding**

  The author received no specific funding or financial support for this study.

  The authors confirm that the paper is not under review or in publication in full or partial in any other Journal.

REFERENCES

Albert, J. (2009).*Bayesian Computation with R.* Springer.

Banerjee, S., Gelfand, A.E., and Carlin, B.P. (2025). *Hierarchical modeling and analysis for spatial data*. Chapman and Hall/CRC.

Best, N., Richardson, S., and Thomson, A. (2005). A comparison of Bayesian spatial models for disease mapping. *Statistical methods in medical research*, **14**(1), 35-59.

Brooks, S., Gelman, A., Jones, G., and Meng, X.L. (Eds.). (2011). *Handbook of markov chain monte carlo*. CRC press.

Cameron, A.C., and Trivedi, P.K. (1998).*Regression Analysis of Count Data.* Cambridge University Press.

Cameron, A.C., andTrivedi, P.K. (2013). *Regression analysis of count data* (No. 53). Cambridge university press.

Congdon, P. (2005). *Bayesian models for categorical data*. John Wiley & Sons.

Dobson, A.J., andBarnett, A.G. (2018). *An introduction to generalized linear models*. Chapman and Hall/CRC.

Fox, J. (2016). *Applied regression analysis and generalized linear models*. Sage publications..

Gelfand, A.E., and Smith, A.F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, **85**(410), 398-409.

Gelman, A., Carlin J.B., Stern, H.S., Dunson, D.B., Vehtari, A and Rubin, D.B. (2013).*Bayesian Data Analysis* (3rd ed.). CRC Press, Chapman & Hall.

Greenwood, M., and Yule, G.U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal statistical society*, **83**(2), 255-279..

Gurung, B., Dutta, S., Singh, K.N., Lama, A., Vennila, S. and Gurung, B. (2022). Development of weather-based forewarning model for tomato leaf curl infestation. *Journal of Agrometeorology*, **24**(4), 424-426.

Hall, D.B. (2000). Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics*, **56**(4), 1030-1039.

Hilbe, J.M. (2011). *Negative binomial regression*. Cambridge University Press.

John Nelder, J.A., and Robert Wedderburn, R.W.M. (1972). *Generalized Linear Models*. Journal of the Royal Statistical Society: Series A (General), **135**(3), 370–384.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, **34**(1), 1-14.

Lawless, J.F. (1987). Negative binomial and mixed Poisson regression. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 209-225.

Lawson, A.B. (2018). *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. Chapman and Hall/CRC.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. (2013). The BUGS book. *A practical introduction to Bayesian analysis, Chapman Hall, London*.

McCullagh, P., and Nelder, J.A. (1989).*Generalized Linear Models* (2nd ed.). Chapman & Hall.

Minka, T.P. (2003). Bayesian inference, entropy, and the multinomial distribution. *Online tutorial*..

Nelder, J.A., and Wedderburn, R.W. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *135*(3), 370-384.

Plummer, M. (2003, March). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, No. 125.10, pp. 1-10).