MEASUREMENT ERROR IN BAYESIAN SPATIO-TEMPORAL POISSON MODELING OF DENGUE USING REMOTE SENSING COVARIATES

SUKARNA

STATISTICS AND DATA SCIENCE STUDY PROGRAM
SCHOOL OF DATA SCIENCE, MATHEMATICS, AND INFORMATICS
IPB UNIVERSITY
BOGOR

2026

STATEMENT REGARDING THE DISSERTATION AND INFORMATION SOURCES AND COPYRIGHTS

I hereby declare that the dissertation entitled “Measurement Error in Bayesian Spatio-Temporal Poisson Modeling of Dengue using Remote Sensing Covariates” is my own original work, carried out under the guidance of my academic supervisors, and has not been submitted in any form to any other university. All sources of information from published and unpublished works by other authors have been properly cited in the text and listed in the References section at the end of this dissertation.

I hereby assign the copyright of my writing to the IPB University, Indonesia.

Bogor,   Agustus 2026

Sukarna
G1601221006

SUMMARY

SUKARNA. Measurement Error in Bayesian Spatio-Temporal Poisson Modeling of Dengue Using Remote Sensing Covariates. Supervised by HARI WIJAYANTO, ANANG KURNIA, and YENNI ANGRAINI.

Statistical modeling has evolved from the classical Linear Model (LM) to more flexible frameworks, including the Generalized Linear Model (GLM), Linear Mixed Model (LMM), and Generalized Linear Mixed Model (GLMM), enabling the analysis of non-Gaussian, hierarchical, and spatio-temporal data. These methodological advances have become increasingly important in epidemiology, where infectious diseases such as dengue exhibit substantial spatial and temporal heterogeneity. Recent developments in remote sensing provide continuous environmental information for disease surveillance; however, satellite-derived environmental indices are indirect measurements affected by sensor limitations, atmospheric disturbances, cloud contamination, and image-processing procedures. Ignoring this uncertainty may bias parameter estimation and reduce predictive performance. Although measurement-error methods have been extensively investigated in classical statistical models, their integration into Bayesian spatio-temporal disease-mapping remains relatively limited. Therefore, this dissertation develops a unified Bayesian spatio-temporal Poisson framework that explicitly accounts for measurement error in remote-sensing covariates for dengue modeling.

The proposed framework extends the Bayesian hierarchical GLMM by jointly modeling fixed effects, localized spatial dependence, temporal dependence, spatio-temporal interaction, and covariate measurement error within a single coherent inferential framework. Spatial autocorrelation is represented by Localized Conditional Autoregressive (LCAR) priors, while temporal dependence and spatio-temporal interactions are incorporated through hierarchical random effects. This hierarchical Bayesian formulation enables simultaneous estimation of all model components while propagating uncertainty across multiple levels of the model, providing a flexible and statistically rigorous approach for analyzing complex spatio-temporal epidemiological data.

The first study demonstrated that Sentinel-2-derived environmental indices provide reliable environmental proxies for explaining dengue incidence. The study compared four statistical models, including conventional Poisson regression, clustered Poisson regression, localized spatial Poisson regression, and Hierarchical Bayesian Spatio-Temporal Localized Conditional Autoregressive (HBSTLCAR) model. The results showed that explicitly accounting for spatial dependence substantially improved predictive performance and more effectively captured spatial heterogeneity than conventional regression models.

The second study incorporated measurement-error correction into a clustered Poisson regression model and demonstrated that explicitly accounting for uncertainty in remotely sensed environmental covariates improved both parameter estimation and predictive accuracy. These findings indicate that satellite-derived

environmental variables should not be regarded as error-free measurements when used in epidemiological modeling.

The third study developed the proposed HBSTLCAR Model with Measurement Error (HBSTLCAR–ME) and evaluated its performance through both simulation and empirical studies. The simulation considered nine scenarios generated from combinations of three levels of spatial dependence (none, moderate, and high) with three levels of temporal dependence (none, moderate, and high). Across nearly all scenarios, HBSTLCAR-ME consistently outperformed the corresponding model without measurement error correction, yielding lower prediction errors and improved Bayesian model-fit criteria. The simulation further demonstrated that the effectiveness of measurement-error correction depended on the underlying spatio-temporal structure, with greatest improvements observed under moderate temporal dependence. Under strong spatial dependence, the performance gap between the two models become smaller because localized spatial random effects explained much of the residual variation. In the empirical application, the proposed HBSTLCAR-ME model achieved the best overall predictive performance while preserving localized spatial and temporal risk structures. Measurement-error correction primarily improved estimation of environmental covariate effects without materially altering the underlying spatial dependence, indicating that localized spatial heterogeneity remained the primary source of variation, whereas measurement error constituted an important secondary source of uncertainty.

The principal contribution of this dissertation is the development and comprehensive validation of a unified Bayesian spatio-temporal Poisson framework that integrates localized spatial dependence, temporal correlation, spatio-temporal interaction, and measurement error in remotely sensed environmental covariates. Validation through both simulation studies and empirical epidemiological applications demonstrates that explicitly accounting for measurement uncertainty consistently improves predictive performance, enhances parameter estimation, and strengthens Bayesian inference while preserving the underlying spatio-temporal dependence structure. Furthermore, the proposed framework establishes that remotely sensed environmental indices can serve as reliable proxy variables for dengue surveillance, provided that their inherent measurement uncertainty is explicitly accommodated within a Bayesian spatio-temporal modeling framework.

Keywords: Bayesian spatio-temporal modeling; dengue; Hamiltonian Monte Carlo; measurement error; Poisson regression; remote sensing.

RINGKASAN

SUKARNA. Kesalahan Pengukuran pada Pemodelan Poisson Bayesian Spasio-Temporal Dengue Menggunakan Kovariat Penginderaan Jauh. Dibimbing oleh HARI WIJAYANTO, ANANG KURNIA, dan YENNI ANGRAINI.

Pemodelan statistik telah berkembang dari Linear Model (LM) klasik menuju kerangka pemodelan yang lebih fleksibel, yaitu Generalized Linear Model (GLM), Linear Mixed Model (LMM), dan Generalized Linear Mixed Model (GLMM), sehingga memungkinkan analisis data non-Gaussian, hierarkis, serta spasial-temporal. Perkembangan metodologi ini menjadi semakin penting dalam bidang epidemiologi karena penyakit menular, seperti demam berdarah dengue (DBD), memiliki heterogenitas spasial dan temporal yang tinggi. Di sisi lain, perkembangan teknologi penginderaan jauh telah menyediakan informasi lingkungan yang bersifat kontinu secara spasial untuk mendukung kegiatan surveilans penyakit. Namun demikian, indeks lingkungan yang diturunkan dari citra satelit merupakan hasil pengukuran tidak langsung yang dipengaruhi oleh keterbatasan sensor, gangguan atmosfer, kontaminasi awan, serta proses pengolahan citra. Pengabaian terhadap ketidakpastian tersebut dapat menyebabkan bias dalam pendugaan parameter dan menurunkan kemampuan prediksi model. Meskipun metode penanganan kesalahan pengukuran (*measurement error*) telah banyak dikembangkan pada model statistik klasik, penerapannya pada pemodelan penyakit spasial-temporal berbasis Bayesian masih relatif terbatas. Oleh karena itu, disertasi ini mengembangkan suatu kerangka pemodelan Poisson spasial-temporal Bayesian yang terpadu dengan mengakomodasi kesalahan pengukuran secara eksplisit pada peubah kovariat yang diperoleh dari penginderaan jauh untuk pemodelan DBD.

Kerangka pemodelan yang diusulkan merupakan pengembangan dari GLMM hierarkis Bayesian dengan memodelkan secara simultan efek tetap, ketergantungan spasial terlokalisasi, ketergantungan temporal, interaksi spasial-temporal, serta kesalahan pengukuran pada peubah kovariat dalam satu kerangka inferensi yang terpadu. Autokorelasi spasial direpresentasikan menggunakan prior *Localized Conditional Autoregressive* (LCAR), sedangkan ketergantungan temporal dan interaksi spasial-temporal dimodelkan melalui efek acak hierarkis. Formulasi Bayesian hierarkis ini memungkinkan seluruh komponen model diestimasi secara bersamaan sekaligus mempropagasikan ketidakpastian pada setiap tingkat hierarki model, sehingga menghasilkan pendekatan yang fleksibel dan memiliki landasan statistik yang kuat untuk menganalisis data epidemiologi spasial-temporal yang kompleks.

Penelitian pertama menunjukkan bahwa indeks lingkungan yang diturunkan dari citra Sentinel-2 dapat digunakan sebagai peubah proksi lingkungan yang andal untuk menjelaskan kejadian DBD. Penelitian ini membandingkan empat model statistik, yaitu regresi Poisson konvensional, regresi Poisson dengan efek klaster, regresi Poisson spasial terlokalisasi, serta model Hierarchical Bayesian Spatio-Temporal Localized Conditional Autoregressive (HBSTLCAR). Hasil penelitian menunjukkan bahwa pemodelan yang secara eksplisit memperhitungkan ketergantungan spasial mampu meningkatkan kinerja prediksi secara signifikan

serta menggambarkan heterogenitas spasial dengan lebih baik dibandingkan model regresi konvensional.

Penelitian kedua mengintegrasikan koreksi kesalahan pengukuran ke dalam model regresi Poisson berklaster dan menunjukkan bahwa pengakomodasian ketidakpastian pada peubah kovariat lingkungan hasil penginderaan jauh mampu meningkatkan kualitas pendugaan parameter sekaligus akurasi prediksi. Temuan ini menunjukkan bahwa peubah lingkungan yang diperoleh dari citra satelit tidak seharusnya diperlakukan sebagai hasil pengukuran yang bebas dari kesalahan dalam pemodelan epidemiologi.

Penelitian ketiga mengembangkan model Hierarchical Bayesian Spatio-Temporal Localized Conditional Autoregressive dengan Measurement Error (HBSTLCAR–ME) dan mengevaluasi kinerjanya melalui studi simulasi serta aplikasi pada data empiris. Studi simulasi dilakukan menggunakan sembilan skenario yang dibentuk dari kombinasi tiga tingkat ketergantungan spasial (tidak ada, sedang, dan tinggi) dengan tiga tingkat ketergantungan temporal (tidak ada, sedang, dan tinggi). Pada hampir seluruh skenario, model HBSTLCAR–ME secara konsisten menunjukkan kinerja yang lebih baik dibandingkan model yang tidak mengakomodasi kesalahan pengukuran, yang ditunjukkan oleh kesalahan prediksi yang lebih rendah serta nilai kriteria kecocokan model Bayesian yang lebih baik. Hasil simulasi juga menunjukkan bahwa efektivitas koreksi kesalahan pengukuran bergantung pada struktur spasial-temporal yang mendasarinya. Peningkatan kinerja terbesar diperoleh pada kondisi ketergantungan temporal sedang, sedangkan pada kondisi ketergantungan spasial yang tinggi, perbedaan kinerja kedua model menjadi lebih kecil karena efek acak spasial terlokalisasi telah mampu menjelaskan sebagian besar variasi residual. Pada aplikasi empiris, model HBSTLCAR–ME memberikan kinerja prediksi terbaik secara keseluruhan dengan tetap mempertahankan struktur risiko spasial dan temporal yang terlokalisasi. Koreksi kesalahan pengukuran terutama meningkatkan ketepatan pendugaan pengaruh peubah kovariat lingkungan tanpa mengubah secara berarti struktur ketergantungan spasial yang mendasarinya. Hasil tersebut menunjukkan bahwa heterogenitas spasial terlokalisasi tetap merupakan sumber utama variasi, sedangkan kesalahan pengukuran menjadi sumber ketidakpastian tambahan yang penting untuk dipertimbangkan.

Kontribusi utama disertasi ini adalah pengembangan dan validasi secara komprehensif terhadap suatu kerangka pemodelan Poisson spasial-temporal Bayesian yang terpadu, yang mengintegrasikan ketergantungan spasial terlokalisasi, korelasi temporal, interaksi spasial-temporal, serta kesalahan pengukuran pada peubah kovariat hasil penginderaan jauh. Validasi melalui studi simulasi dan aplikasi epidemiologi empiris menunjukkan bahwa pengakomodasian ketidakpastian pengukuran secara eksplisit mampu meningkatkan kinerja prediksi, memperbaiki pendugaan parameter, serta memperkuat inferensi Bayesian tanpa menghilangkan struktur ketergantungan spasial-temporal yang mendasarinya. Selain itu, kerangka pemodelan yang diusulkan membuktikan bahwa indeks lingkungan hasil penginderaan jauh dapat dimanfaatkan sebagai peubah proksi yang andal untuk mendukung surveilans DBD, sepanjang ketidakpastian pengukuran yang melekat pada data tersebut diakomodasi secara eksplisit dalam kerangka pemodelan Bayesian spasial-temporal.

Kata kunci: Bayesian spasio-temporal, demam berdarah dengue, Hamiltonian Monte Carlo, kesalahan pengukuran, penginderaan jauh, regresi Poisson.

MEASUREMENT ERROR IN BAYESIAN SPATIO-TEMPORAL POISSON MODELING OF DENGUE USING REMOTE SENSING COVARIATES

SUKARNA

Dissertation
as one of the requirements for obtaining a degree
Doctorate
in Study Program on Statistics and Data Science

STUDY PROGRAM ON STATISTICS AND DATA SCIENCE
SCHOOL OF DATA SCIENCE, MATHEMATICS, AND INFORMATICS
IPB UNIVERSITY
BOGOR
2026

External Examiners of the Supervisory Commission on Dissertation Closed Examinations:

1. Dr. Kusman Sadik, M.Si. ....
2. Dr. Bagus Sartono, M.Si. ....

Promoter Outside the Supervisory Commission at the Dissertation Open Promotion Session:

1. Dr. Bagus Sartono, M.Si. ....
2. Prof. Drs. Suwardi Annas, M.Si., Ph.D. ....

APPROVAL SHEET

Dissertation Title : Measurement Error in Bayesian Spatio-Temporal Poisson Modeling of Dengue using Remote Sensing Covariates
Name : Sukarna
Student ID : G1601221006

Approved by the Dissertation Committee

Supervisor:
Prof. Dr. Ir. Hari Wijayanto, M.Si. (..............................)

Co-Supervisor:
Prof. Dr. Anang Kurnia, M.Si. (..............................)

Co-Supervisor:
Dr. Yenni Angraini, M.Si. (......................)

Acknowledged by

Head of the Study Program:
Dr. Kusman Sadik, S.Si., M.Si. (......................)
NIP 19690912 199702 1 001

Dean of School of Data Science, Mathematics, and Informatics:
Prof. Dr. Ir. Agus Buono, M.Si., M.Kom. (......................)
NIP 19660702 199301 1 001

Date of Closed Examination: 20 July 2026
Date of Promotion Examination: 05 August 2026

Date of Approval:

ACKNOWLEDGEMENTS

All praise and gratitude are due to Allah for His blessings and mercy, which have enabled the completion of this dissertation entitled ***"Measurement Error in Bayesian Spatio-Temporal Poisson Modeling of Dengue Using Remote Sensing Covariates"***. Conducted from August 2024 to April 2026, this research focuses on incorporating measurement error into Bayesian spatio-temporal models for dengue incidence using environmental covariates derived from remote sensing data. The completion of this dissertation would not have been possible without the support, guidance, and encouragement of many individuals and institutions. May Allah Subhanahu wa Ta'ala reward all those who have contributed to this work and grant them blessings in this life and the hereafter.

First and foremost, I express my deepest gratitude to my beloved parents, parents-in-law, and extended family for their prayers, unconditional love, encouragement, and unwavering support throughout my academic journey. I would also like to extend my sincere appreciation and gratitude to the following individuals and institutions:

1. Prof. Dr. Hari Wijayanto, Prof. Dr. Anang Kurnia, and Dr. Yenni Anggraini, members of my dissertation supervisory committee, for their invaluable guidance, insightful suggestions, constructive feedback, and continuous encouragement throughout this research.
2. The Head and Secretary of the Statistics and Data Science Study Program for the undergraduate, master's, and doctoral programs, for their academic support and assistance.
3. All lecturers and administrative staff of the Statistics and Data Science Study Program, the School of Data Science, Mathematics, and Informatics, and the Graduate School of IPB University, for their invaluable academic guidance, administrative assistance, and unwavering support.
4. The Government of the Republic of Indonesia, through PPAPT, LPDP, the Ministry of Economic Affairs, and BPI, for providing financial and institutional support throughout my doctoral education.
5. All doctoral students in Statistics and Data Science at IPB University, particularly the 2022 cohort, as well as colleagues from the 2017, 2018, 2019, 2020, 2021, 2023, 2024, and 2025 cohorts, for their friendship, collaboration, and mutual support.
6. My friends, colleagues, and companions in various study groups at Cespo Baranangsiang and LSI Dramaga, especially the "Outlier," "Geng Cucurak," and "Hero IPB" groups, whose camaraderie, encouragement, and shared experiences made this journey both meaningful and enjoyable.
7. All individuals and institutions whose names cannot be mentioned individually, but whose contributions and support are sincerely appreciated.

Finally, it is my sincere hope that this dissertation will contribute to the advancement of scientific knowledge, methodology, and technology, and provide meaningful benefits to the academic community and society at large.

Bogor, August 2026

Sukarna

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF GLOSSARIUM

Abbreviation	Expanding of Abbreviation
AC	Administrative Cluster
AIC	Akaike Information Criterion
BI	Breteau Index
BPJS	Social Security Agency (*Badan Penyelenggara Jaminan Sosial*)
BSLCAR	Bayesian Spatial Localized Conditional Autoregressive
BYM	Besag–York–Mollié
CI	Confidence Interval
CMP	Conway-Maxwell-Poisson
CSU	Clustered Spatial Unit
DHF	Dengue Hemorrhagic Fever
DHF	Dengue Hemorrhagic Fever
DIC	Deviance Information Criterion
ENSO	El Niño-Southern Oscillation
GLM	Generalized Linear Model
GLMM	Generalized Linear Mixed Model
Green	Green Band (Band-3, B3)
GSMaP	Global Satellite Mapping of Precipitation
HBSLCAR	Hierarchical Bayesian Spatial Localized Conditional Autoregressive
HBSTLCAR	Hierarchical Bayesian Spatial-temporal Localized Conditional Autoregressive
HBSTLCAR-ME	Hierarchical Bayesian Spatio-temporal Localized Conditional Autoregressive model incorporating Measurement Error
HI	House Index
HMC	Hamiltonian Monte Carlo

Abbreviation	Expanding of Abbreviation
ICAR	Intrinsic Gaussian Conditional Autoregressive
IG	Inverse Gamma
IOD	Indian Ocean Dipole
IRS	Indoor Residual Spraying
JOIV	International Journal on Informatics Visualization
LCAR	Localized Conditional Autoregressive
LM	Linear Model
M1	Model-1
M2	Model-2
M3	Model-3
M4	Model-4
MAE	Mean Absolute Error
MC-EM	Monte Carlo Expectation–Maximization
MCMC	Markov Chain Monte Carlo
ME	Measurement Error
MJO	Madden-Julian Oscillation
MLE	Maximum Likelihood Estimation
NDBI	Normalized Difference Built-up Index
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near-Infra Red band (Band-8, B8)
NIRp	The Narrow NIR Band (Band-8A, B8A)
NUTS	No-U-Turn Sampler
OLS	Ordinary Least Squares
OR	Odd Ratio
PKM	Community health centers (*Pusat Kesehatan Masyarakat*)

Abbreviation	Expanding of Abbreviation
Puskesmas	Community health centers (*Pusat Kesehatan Masyarakat*)
Red	Red Band (Band-4, B4)
RMSLE	Root Mean Squared Logarithmic Error
SIR	Standardized Incidence Ratio
SWIR1	Shortwave Infra-Red Band (Band 11, B11)
TCP	Tweedie Compound Poisson
USGS	United States Geological Survey
WAIC	Widely Applicable Information Criterion
WHO	World Health Organization