

Bayesian Spatio-Temporal Count Models with Underreporting: A Practical Identifiability Approach for Dengue Surveillance in Costa Rica

Jimena Murillo-Montero, Boston Scientific

Underreporting is a major challenge in epidemiological surveillance because reported cases do not necessarily reflect the true incidence of disease. This issue is particularly relevant for dengue in Costa Rica, where surveillance relies on passive reporting and laboratory confirmation, potentially biasing spatio-temporal analyses and public health decision-making. This study proposes Bayesian spatio-temporal count models with explicit underreporting to jointly estimate latent disease incidence and reporting probability. The primary objective is to evaluate the practical identifiability of these models, that is, their ability to reliably distinguish the disease process from the reporting process using finite observational data. Poisson-logit and Negative Binomial-logit specifications incorporating spatial and temporal random effects will be compared through simulation studies designed to assess parameter recovery, estimation stability, and the contribution of model structure to practical identifiability. The selected model will subsequently be applied to weekly dengue surveillance data from Costa Rica (2017–2025). Simulation and application results are expected to illustrate the performance of the proposed framework and its potential to improve estimation of dengue burden under underreporting, providing more reliable evidence for epidemiological surveillance and public health decision-making.

The gender gap in biotechnology awareness: a ridge regression analysis from Costa Rica

Laura Obando Esquivel, Student

Biotechnology awareness varies considerably across populations, yet the factors driving this variation remain poorly characterized in Costa Rica. This study examines sociodemographic and attitudinal predictors of self-reported biotechnology awareness among Costa Rican adults using data from the 2022 Encuesta de Actualidades (Current Affairs Survey, University of Costa Rica, $n = 1,744$). Ridge regression (penalized logistic regression, $\alpha = 0.0014$) was applied to address multicollinearity among predictors while retaining all variables in the model. Willingness to consume genetically modified foods (OR = 1.78), educational attainment (OR = 1.64), and reading capacity (OR = 1.44) showed the strongest positive associations with awareness. Being a woman was associated with lower odds of biotechnology awareness (OR = 0.68). Perceived biotechnology knowledge (OR = 1.35) and household income (OR = 1.08) showed smaller positive associations with awareness. These findings suggest that science communication strategies targeting women and populations with lower educational attainment may be central to reducing biotechnology awareness gaps in Costa Rica.

Forecasting Weekly Dengue Relative Risk Across Costa Rican Municipalities Using Machine Learning and Digital Surveillance Data

Marianne Peña Wüst, Student

Dengue fever is a disease that poses a significant health risk in tropical and subtropical regions. Therefore, this study aims to predict weekly dengue relative risk for 82 municipalities in Costa Rica from 2020 to 2025, by using meteorological and social variables and Google Trends data. The modeling approach included machine learning techniques such as Random Forest, XGBoost, and their hurdle variants to account for excess zeros. Each of the four models was fitted independently for each municipality, in addition to a global model fitted using data from all municipalities. Performance was evaluated using Normalized Root Mean Squared Error, and variable importance and selection were computed. Random Forest models exhibited the best performance. The most important variables included relative risk lags, meteorological variables, and Google Trends topics relating to employment, mobility and finance, and dengue spread.

Modeling Milk Yield in a Tropical Dairy Herd with Generalized Additive Models

Shirley Rojas-Salazar, University of Costa Rica

Milk yield is influenced by biological factors that often exhibit nonlinear relationships, making conventional linear models inadequate to describe lactation dynamics. This study applied Generalized Additive Models (GAMs) to evaluate the effects of days in milk, parity, and breed composition on milk yield in a tropical dairy herd. Data were collected at the Diamante animal production farm in Guápiles, Costa Rica. Smooth functions were used to model nonlinear effects, and cow was included as a random effect to account for repeated measurements. The analysis identified nonlinear effects of the explanatory variables and enabled the comparison of alternative model specifications. The results demonstrate the usefulness of GAMs for modeling complex biological relationships and providing an interpretable framework for analyzing longitudinal milk production data.

