

# Spatial-temporal analysis of dengue fever in Brazil during 2015-2019

By 01911899

12/04/2024

## Abstract

This ecological study employs Bayesian hierarchical models and spatial generalized linear mixed-effect models (GLMMs) to analyse the spatial-temporal dynamics of dengue fever in Brazil from 2015 to 2019. Using data from 557 microregions over 60 months, we aim to investigate the relationships between climatic variables (monthly maximum temperature and Palmer's drought severity index) and dengue incidence. Despite identifying high-risk areas primarily in Southeast and Centre-West regions, the periodic outbreak pattern observed did not exhibit a clear yearly trend. The chosen spatial-temporal model with type I interaction and two covariates accurately predicted dengue cases. Our findings provide valuable insights into dengue transmission patterns and highlight the importance of comprehensive modelling approaches in informing public health interventions and policy decisions.

## Introduction

Dengue fever is a viral infection caused by the dengue virus, which is primarily transmitted to humans through the bite of female Aedes mosquito infected with dengue virus serotypes. (Hasan 2016)

In Brazil, dengue fever has been a significant public health concern, with the country experiencing large outbreaks and high numbers of cases in recent years. For example, in 2019, Brazil reported over 1.5 million cases of dengue fever, with hundreds of deaths attributed to the disease. The prevalence and severity of dengue fever in Brazil are influenced by various factors, including climate conditions. The situation regarding dengue fever in Brazil can be severe, particularly during outbreaks when healthcare systems may become overwhelmed, and resources are strained. (Azevedo, Lorenz, and Chiaravalloti-Neto 2020) Therefore, by implementing comprehensive prevention strategies, public health authorities can effectively reduce the impact of dengue fever on individuals, communities, and societies.

Dengue fever is considered a climate-sensitive disease because its transmission dynamics are strongly influenced by climatic factors. (Lowe et al. 2011) (Lowe et al. 2013) The primary vectors responsible for transmitting the dengue virus have specific temperature and humidity requirements for breeding and survival. Dengue fever often exhibits seasonal patterns, with higher transmission rates occurring during periods of warmer temperatures and increased rainfall.

The aim of this research is to conduct a spatial-temporal analysis to investigate relationships between climatic variables and dengue fever incidence over the period of 2015-2019 across different micro regions in Brazil. Many existed studies investigate the spatial-temporal relationship using variety of covariates including climatic, demographic, socio-economic, entomologic and geographic predictors.(Aswi et al. 2019) (Lowe et al. 2011) Two climatic variables, temperature, and precipitation, are the most commonly used predictors. However, to our best knowledge, there is no other spatial-temporal analysis focused on the same regions during the same period that used both temperature and precipitation predictors. Therefore, this study will build on top of previous work by including: 1) Comparing some other work using yearly and out-of-date data, we use more up-to-date data with higher granularity (in both spatial and temporal dimensions) 2) Incorporating both monthly maximum temperature and the Palmer's drought severity index (PDSI), a measurement for precipitation, as the climatic variables, we intend to unravel the intricate relationship between environmental factors and the incidence of dengue fever.

Through this investigation, the research aims to address a series of research questions in two main folds. Research questions:

1. What are the spatial and temporal patterns of dengue fever incidence across different microregions in Brazil from 2015 to 2019, and which microregions exhibit consistently higher-than-average incidence rates? Can we build a predictive model to indicate when and where should the authorities implement relevant strategies to prevent dengue fever?
2. How variations in monthly maximum temperature and PDSI influence the dengue fever incidence, controlling the spatial and temporal variability?

By connecting these research questions to the overarching aim, this research endeavours to contribute to a comprehensive understanding of the interplay between climatic factors and dengue fever dynamics, providing insights that may inform targeted interventions and public health strategies in Brazil.

## Methods

For this spatial-temporal ecological study, we utilized Brazilian regional dengue cases and covariate data spanning 557 micro regions over a period of 60 months (12 months for each of five years from 2015 to 2019). The dengue case data were sourced from the Notifiable Diseases Information System, accessible through the Ministry of Health Information Department (DATASUS). These monthly dengue case records provide comprehensive information on reported dengue fever cases across different regions of Brazil. The geographical

delineation of the micro regions was established based on a predefined spatial structure provided by the Brazilian Institute of Geography and Statistics (IBGE). All 557 micro regions can be categorised into five larger regions: North, Northeast, Centre West, South and Southeast. Additionally, both maximum temperature and self-calibrated PDSI data were obtained from the Climatic Research Unit. These climatic variables were chosen for their relevance to dengue fever transmission dynamics and their availability for the study period, allowing for a comprehensive assessment of the relationship between climate and dengue incidence in Brazil.

There are five spatial and spatial-temporal model are conducted, which are described in the following equations. For all models, we employed spatial generalized linear mixed-effect models (GLMM) with a Poisson family and use Bayesian hierarchical framework (Banerjee, Carlin, and Gelfand 2014) to make inference of the posterior distribution of latent parameters by Integrated Nested Laplace Approximation(INLA) (Rue, Martino, and Chopin 2009).

Model-1 is the spatial-only model which is only an exploratory model to understand the spatial variation of dengue cases and serves as a reference comparison for interpretation purpose. Then we developed spatial-temporal models with and without interactions. Model-2.1 and Model-2.2 are the spatial-temporal model without interaction, where the difference between these two models is that Model-2.2 include the two covariates of interets, temperature and PDSI; while Model-3.1 and Model-3.2 are the spatial-temporal model with type I interaction, with the later including covariates as well. Both covariates are used directly as continuous number. Since the expected cases in not available, we replace the variable  $E$  in the following equation by a scaled population, i.e.  $E = Pop/10^5$ , which subsequently make  $\rho$  being the incident rate of dengue fever per 100,000 population.

### Model – 1

$$\begin{aligned} O_i &\sim \text{Poisson}(\rho_i E_i) \\ \log \rho_i &= b_0 + b_i \\ \mathbf{b} &= \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi} \mathbf{v}_* + \sqrt{\phi} \mathbf{u}_*) \end{aligned}$$

### Model – 2.2

$$\begin{aligned} O_{it} &\sim \text{Poisson}(\rho_{it} E_{it}) \\ \log \rho_{it} &= b_0 + b_i + \gamma_t + \beta_T \times Temp_{it} + \beta_P \times PDSI_{it} \\ \mathbf{b} &= \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi} \mathbf{v}_* + \sqrt{\phi} \mathbf{u}_*) \\ \gamma_t &\sim RW(1) \end{aligned}$$

### Model – 2.1

$$\begin{aligned} O_{it} &\sim \text{Poisson}(\rho_{it} E_{it}) \\ \log \rho_{it} &= b_0 + b_i + \gamma_t \\ \mathbf{b} &= \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi} \mathbf{v}_* + \sqrt{\phi} \mathbf{u}_*) \\ \gamma_t &\sim RW(1) \end{aligned}$$

### Model – 3.1

$$\begin{aligned} O_{it} &\sim \text{Poisson}(\rho_{it} E_{it}) \\ \log \rho_{it} &= b_0 + b_i + \gamma_t + \delta_t \\ \mathbf{b} &= \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi} \mathbf{v}_* + \sqrt{\phi} \mathbf{u}_*) \\ \gamma_t &\sim RW(1) \\ \delta_{it} &\sim \text{Normal}(0, \sigma_\delta^2) \end{aligned}$$

### Model – 3.2

$$\begin{aligned} O_{it} &\sim \text{Poisson}(\rho_{it} E_{it}) \\ \log \rho_{it} &= b_0 + b_i + \gamma_t + \delta_t + \beta_T \times Temp_{it} + \beta_P \times PDSI_{it} \\ \mathbf{b} &= \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi} \mathbf{v}_* + \sqrt{\phi} \mathbf{u}_*) \\ \gamma_t &\sim RW(1) \\ \delta_{it} &\sim \text{Normal}(0, \sigma_\delta^2) \end{aligned}$$

For area  $i$  at time  $t$ ,

- $O$  stands for the number of observed cases , and  $E$  is defined as the population scaled by 100,000,
- $\rho$  is the incident rate of dengue fever for 100,000,
- $Temp$  is the month maximum temperature, and  $PDSI$  is the Palmer's drought severity index,
- $\mathbf{b}$ , the spatial random effect, is specified by a Besag-York-Molli'e (BYM) prior(Besag, York, and Mollié 1991) composed by  $u_i$  and  $v_i$ , where  $u_i$  is the spatially structured component defined by an intrinsic CAR prior (IARC 2016):  $\mathbf{u} \sim ICAR(\mathbf{W}, \sigma_u^2)$ , and  $v_i$  is the unstructured component defined with prior:  $v_s \stackrel{iid}{\sim} \text{Normal}(0, \sigma_v^2)$  ,
- $\tau_b$  is the precision parameter controlling the marginal variance of the random effect with its prior defined as  $P(\sigma_{\tau_b} > 0.5/0.31) = 0.01$ , and  $\phi$  is the mixing parameter measuring the proportion of the marginal variance with its prior defined as  $P(\phi < 0/5) = 2/3$  (Riebler A 2016).

## Results

### Exploratory Data Analysis

Exploratory data analysis is conducted before the statistical models. By aggregating all micro regions data over time, we plotted the variation in the number of dengue cases over the five years. Moreover, in order to understand how temperature and PDSI vary over time, especially to investigate if the years with more severe outbreaks have unusual temperatures and PDSI, we plot the spatial and temporal variation of these two covariates in Figure 1. These plots are interpreted collaboratively with the regional incidence rate plot (shown in Figure 2) to summarise the temperature and PDSI features of regions with higher risk.

### Generalised Linear Mixed Models

For the spatial and spatial-temporal models, we present two types of plots to identify micro regions at higher risk in Figure 3: 1) the plot of posterior mean of incidence rate ratio (IRR) and 2) the plot of posterior probabilities (PP) that the IRR  $> 1$ .

## Covariates, Hyperparameters and Model comparison

Then we extracted the posterior median with 95% CI of the IRR of two covariates (in Table 1) and all hyperparameters (in Table 2) respectively. The Watanabe–Akaike information criterion (WAIC) of all five models is shown in Table 3 to provide a comparable measurement evaluating model fit and model complexity. We further visualise the predictive values of each micro regions at each time point to compare the predictive performance of different models.

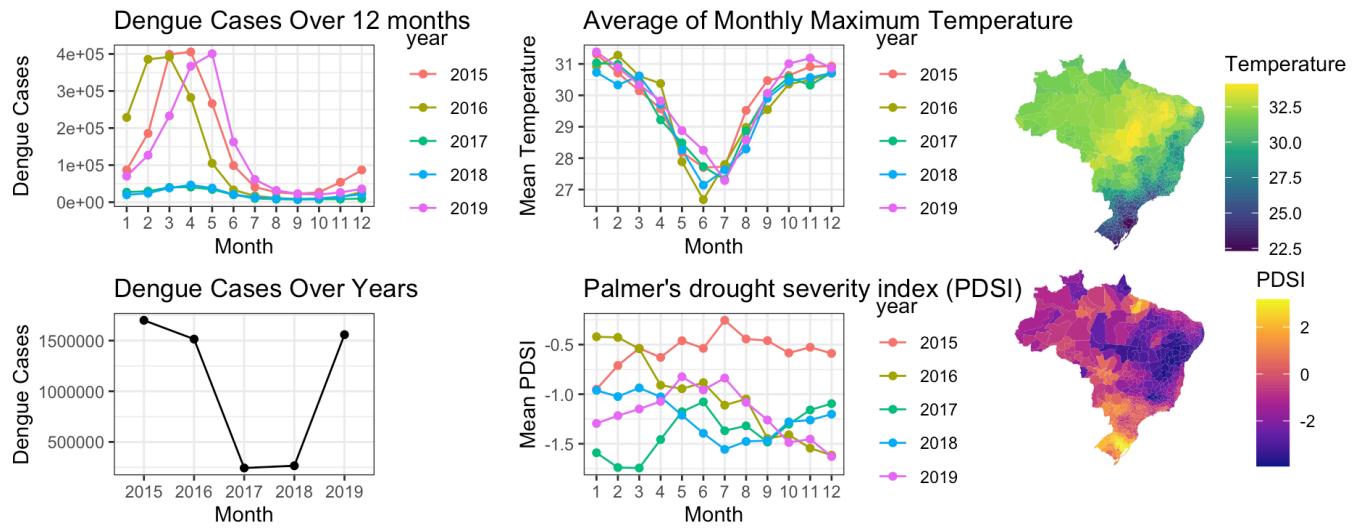


Figure 1. (Left-up) Plot of monthly variation of the number cases between 2015 and 2019. (Left-bottom) Plot of yearly variation of the total number of cases. (Middle-up) Plot of monthly average of monthly maximum temperature for the five years. (Middle-bottom) Plot of the monthly average of PDSI for the five years. (Right-up) Map of average of monthly maximum temperature. (Right-bottom) Map of average PDSI

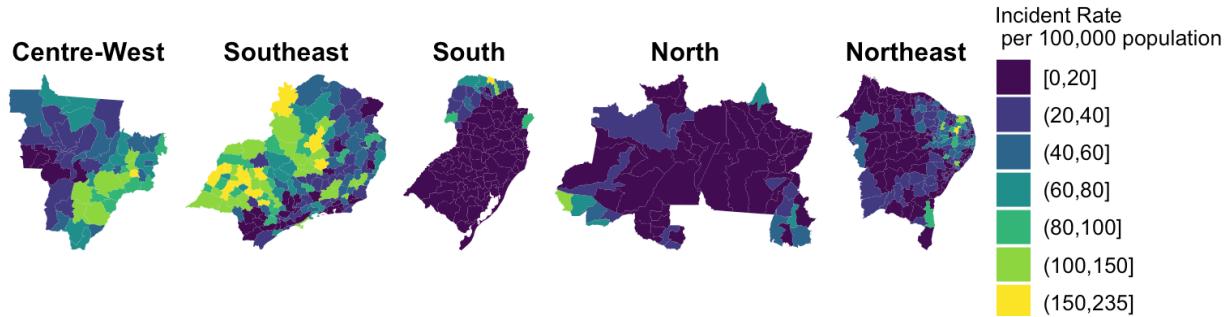
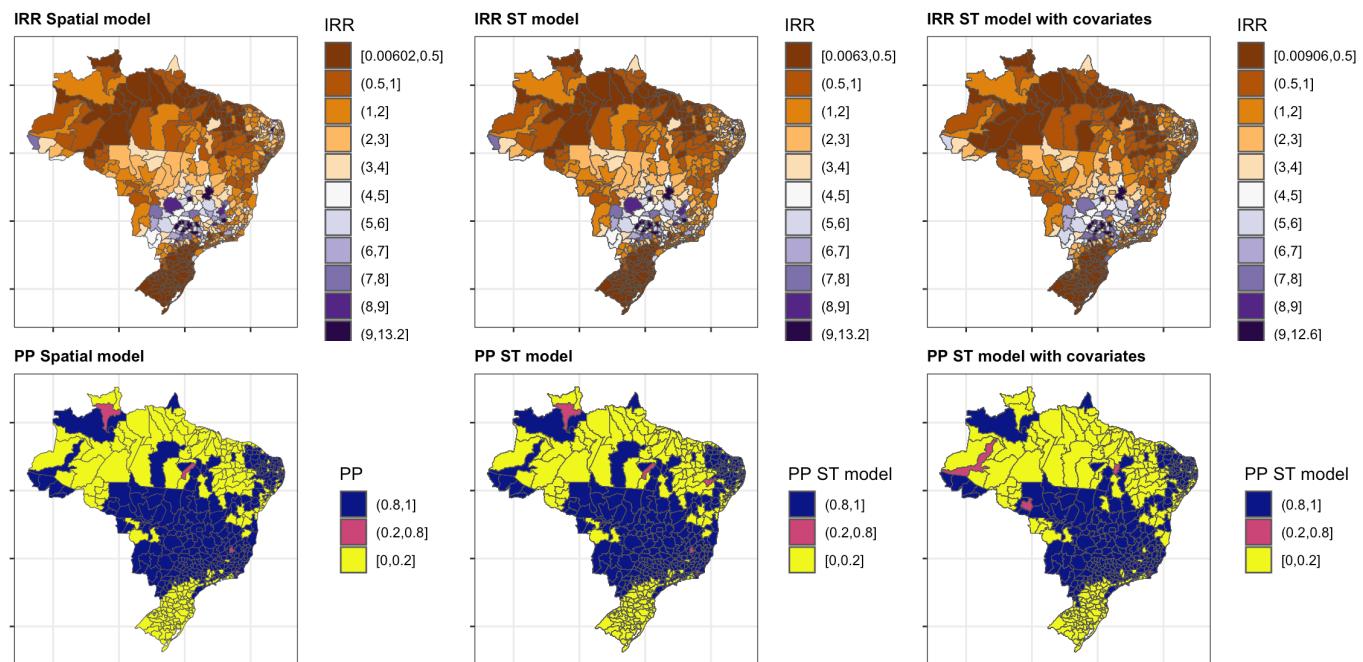


Figure 2. Brazil dengue disease incidence rate per 100,000 population.



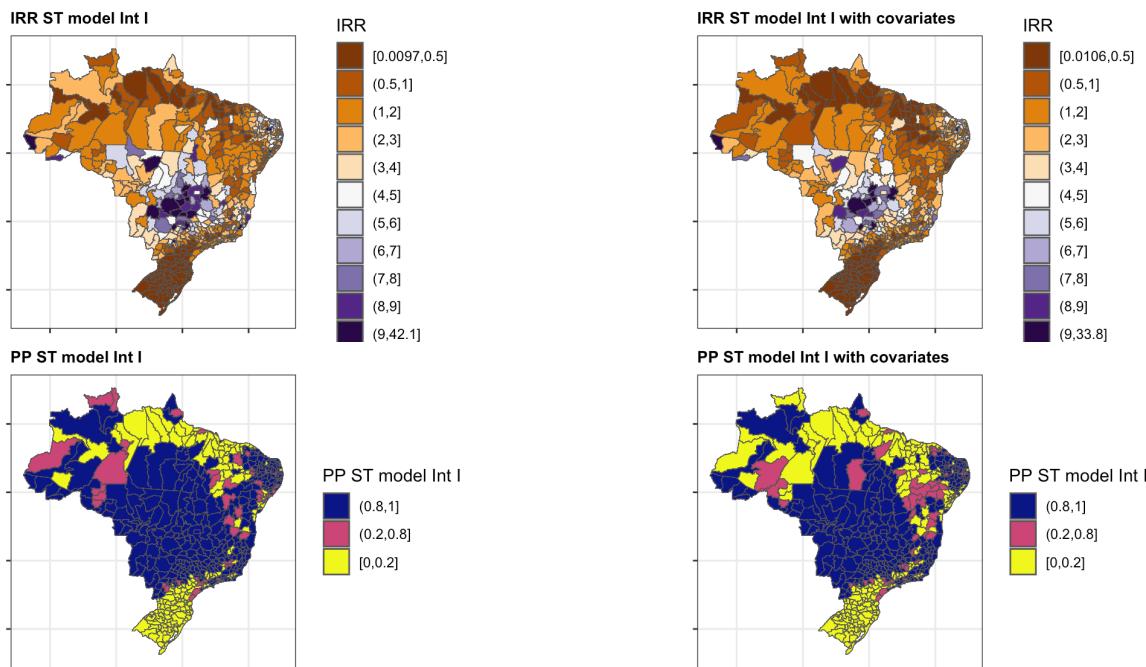


Figure 3. (Left) Map of posterior mean of incidence rate ratio of each model. (Right) map of posterior probabilities (PP) that the IRR > 1

Table 1. Posterior median with 95% CI of the IRR of two covariates

	SpatTemp no int with covariates	SpatTemp typel with covariates
Temperature	1.06 (1.059, 1.061)	1.006 (0.995, 1.017)
PDSI	0.946 (0.946, 0.947)	0.964 (0.954, 0.974)

Table 2. Posterior median with 95% CI of all hyperparameters

	SpatTemp no int	SpatTemp no int with covariates	SpatTemp typel	SpatTemp typel with covariates
Precision for ID_space	0.655 (0.557, 0.754)	0.674 (0.568, 0.779)	0.805 (0.698, 0.928)	0.941 (0.798, 1.079)
Phi for ID_space	0.97 (0.903, 0.995)	0.966 (0.894, 0.994)	0.925 (0.893, 0.949)	0.791 (0.757, 0.827)
Precision for ID_time	3.531 (2.444, 5.016)	3.528 (2.44, 5.002)	3.351 (2.301, 5.181)	2.27 (1.841, 2.768)
Precision for ID_st			0.61 (0.596, 0.622)	0.615 (0.605, 0.632)

Table 3. WAIC of the different models

model	WAIC
Spatial	6237
SpatTemp no Int	1639486
SpatTemp no Int with covariates	1640271
SpatTemp typel	629160
SpatTemp typel with covariates	196468

## Predictive Values

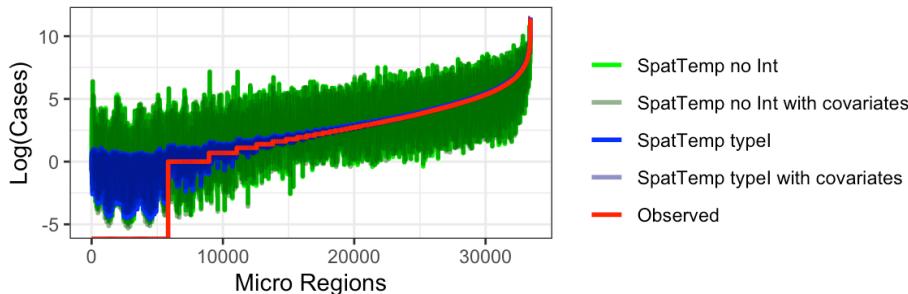


Figure 4. Plot of the predictive number cases of each spatial-temporal model and the observed cases.

## Discussion and conclusion

As shown in Figure 1, there were relatively more severe dengue fever outbreaks in three years (2015, 2016 and 2019), with over 1.5 million cases diagnosed each year. The outbreak of dengue fever occurred periodically, peaking time ranging from February to May. Even in 2017 and 2018 when Brazil had a significantly smaller outbreak, the relative periodic pattern can still be observed.

However, in both spatial and temporal dimensions, the mean value of monthly maximum temperature varies smoothly without a specific unusual pattern and there is no clear pattern for the PDSI as well. By calculating the incidence rate for each micro regions, as shown in Figure 2, we can observe that the most severe region is Southeast followed by Centre-West. The eastern side of the Northeast region and the most northern side of the South region also present some outbreaks of dengue fever. By combining these severe regions with the maps of temperature and PDSI in Figure 1, these areas actually all present have about 30 degree in temperatures and -2 in PDSI.

From a series of plots in Figure 3, we can observe that the spatial and spatial-temporal model without interaction identifies similar regions, but by including an interaction term, more areas in Centre-West and North are identified with PP > 0.8. In addition, these two interaction models also present a larger IRR for those regions at the highest risk. The maximum posterior IRR calculated by the interaction models (41.5 and 33.8 for models with and without covariate respectively) are about three times larger than the corresponding maximum IRR calculated by the models without interaction term, and they are all much lower than the true observed IRR. This means that the spatial-temporal models without interaction perform worse in capturing the long-tail most risky regions properly.

On the other hand, the pattern between models with covariates and models without covariates is not significantly different. These results are also supported by the results shown in Table 2. We can see that by adding the interaction term, there is no large variation in the IRR of both covariates. However, the positive effect of temperature becomes insignificant after adding the interaction, while the negative effect of PDSI remains statistically significant.

Combining the results from Table 2 and Table 3, we can find that the spatial-temporal model with Type I interaction and covariates has the highest  $\tau$  and smallest  $\phi$ , which means it balances better between the structured random effect and the unstructured random noise. The results of WAIC also show that the model with type I interaction and covariates is the best model among all four spatial-temporal models. On top of that, Figure 4 further confirms this model's performance by the predictive number of cases which has a smaller variance and much more accurate prediction compared to the model without interaction.

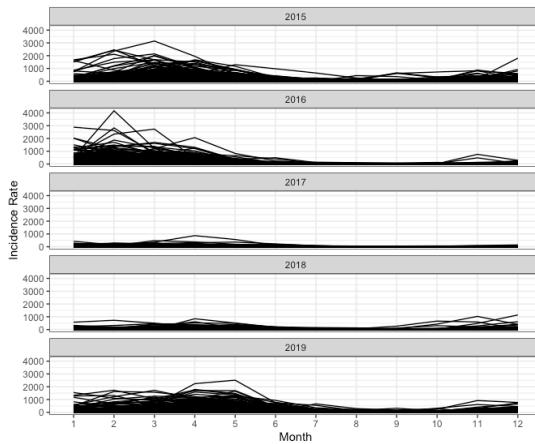
One significant limitation of our study is the assumption of type I interaction, which assumes unstructured spatial and temporal effects. This assumption does not fully align with the observed patterns in our data, as our exploratory data analysis revealed periodic patterns and spatial concentrations of dengue fever. However, due to memory limitations caused by the large dataset size, we were unable to compute the Kronecker product required for type IV interaction. As a result, we opted to include only type I interaction in our analysis, potentially limiting the capture of complex spatial and temporal dynamics in our model.

In conclusion, our study provides valuable insights into the spatial and temporal dynamics of dengue fever incidence in Brazil, as well as the influence of climatic variables on disease transmission. Through our analysis, we identified regions with higher dengue risk, particularly in the Southeast, Centre-West, eastern side of the Northeast region, and the most northern side of the South region. While periodic outbreaks of dengue fever were observed, our analysis revealed no clear trend of increasing or decreasing incidence over time. Our spatial-temporal modelling approach, incorporating type I interaction and two covariates (monthly maximum temperature and Palmer's drought severity index), emerged as the most effective in accurately predicting the number of dengue cases.

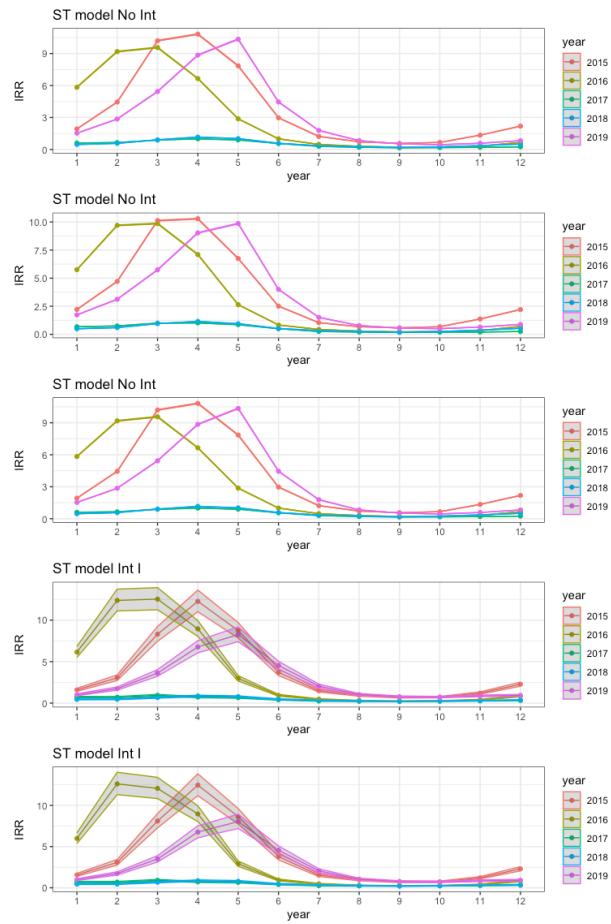
Moving forward, our findings highlight the importance of targeted interventions in high-risk areas identified by our spatial analysis. By focusing resources and efforts on these regions, public health authorities can implement more effective vector control measures, campaigns, and interventions to reduce dengue transmission and mitigate the impact of outbreaks. Furthermore, our analysis can enable timely detection of outbreaks and facilitate proactive responses to mitigate their spread.

Additionally, while our study found no significant effect of monthly maximum temperature on dengue incidence after controlling for spatial and temporal factors, the negative association between Palmer's drought severity index (PDSI) and dengue fever incidence warrants further investigation.

## Supplementary material

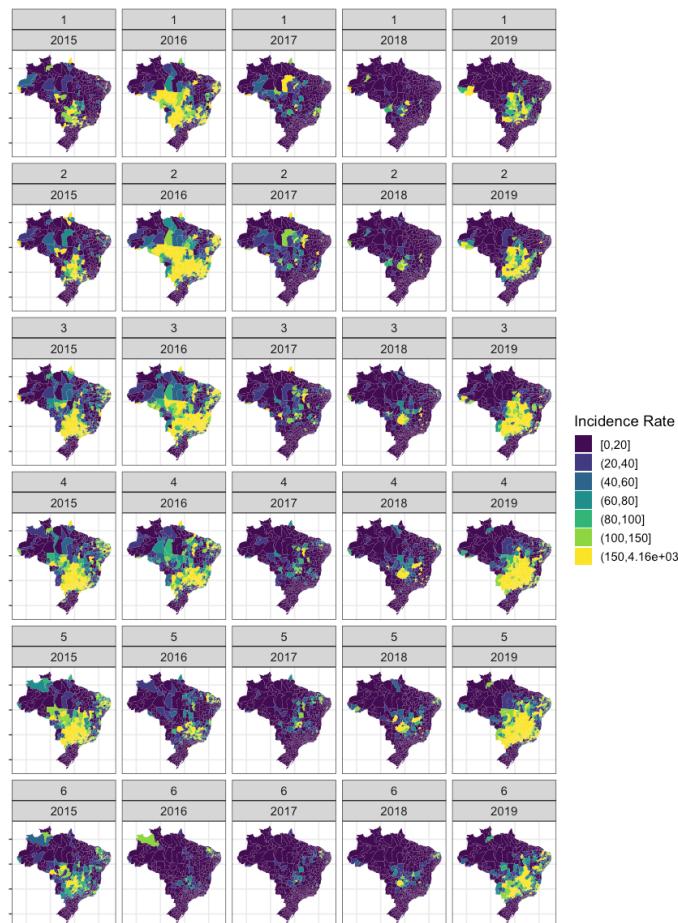


Supplementary Figure 1. Incidence rate over five years (no aggregation)

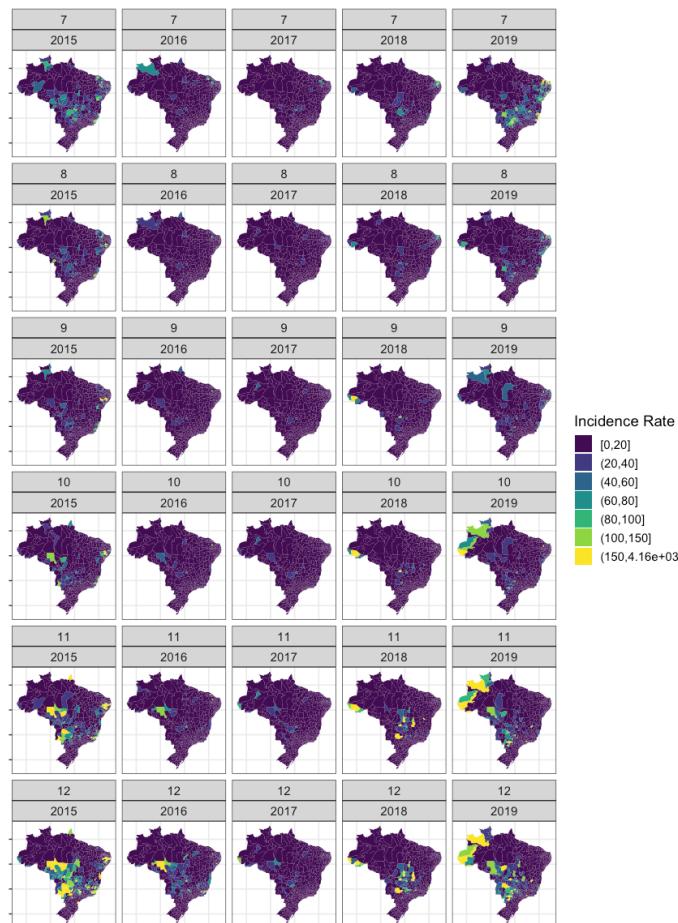


Supplementary Figure 2. Posterior mean and 95% CI of the temporal random effect for all five models

## Spatial-temporal analysis of dengue fever in Brazil during 2015-2019

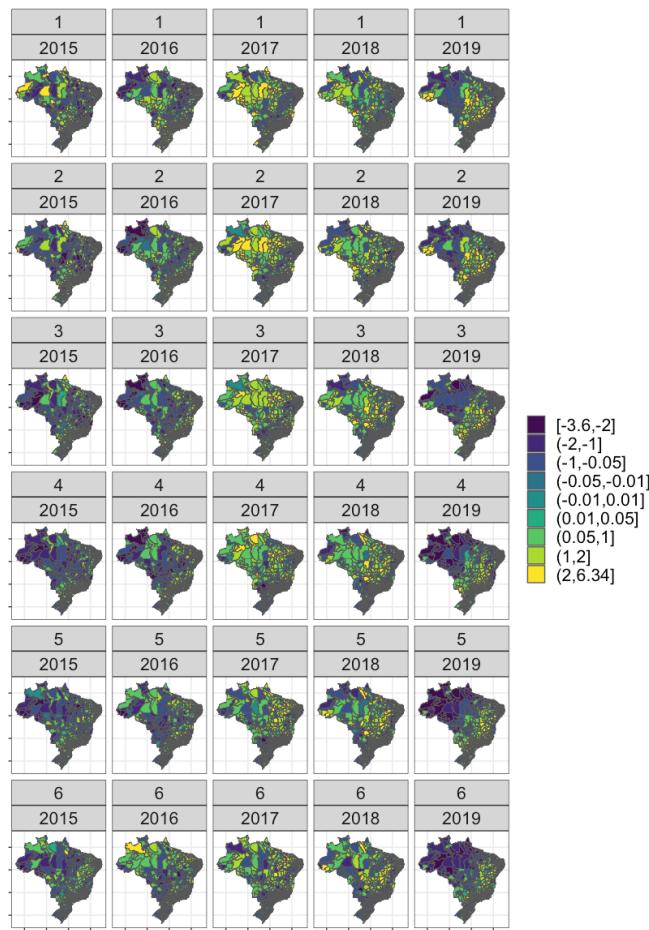


Supplementary Figure 3. Plots of spatial and temporal variation over five years (no aggregation)

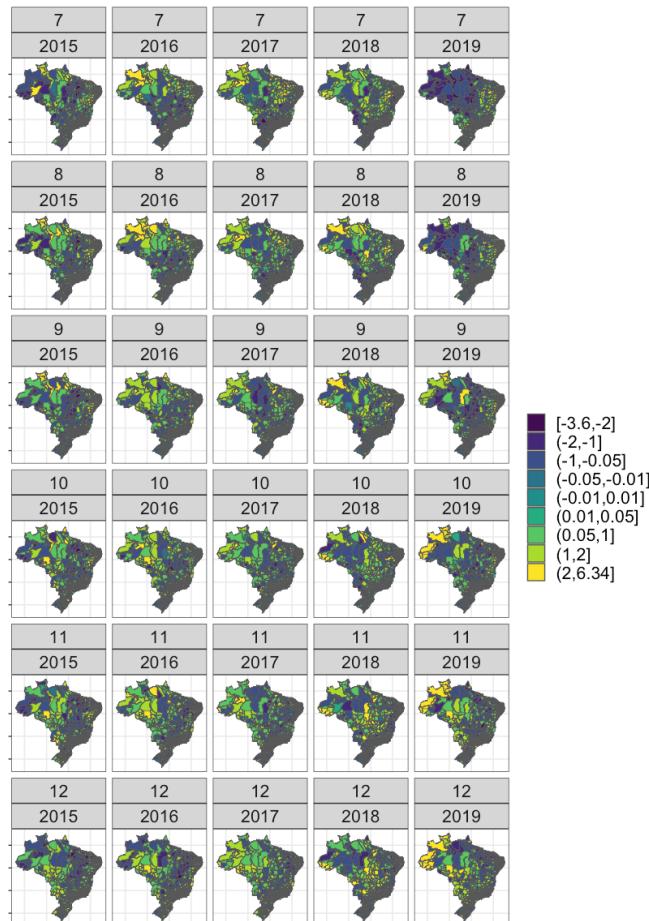


Supplementary Figure 3. Plots of spatial and temporal variation over five years (no aggregation)

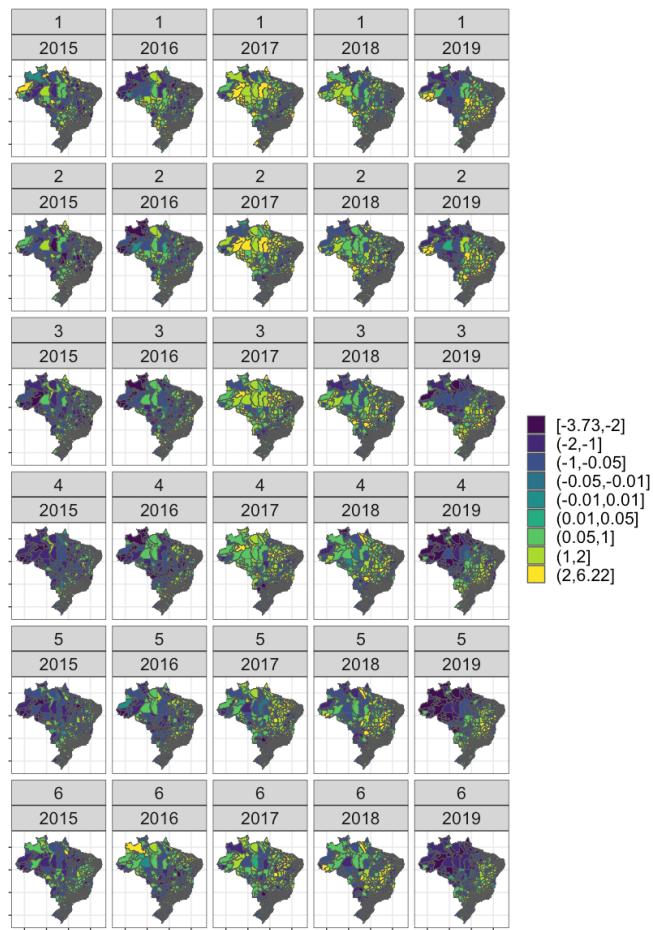
## Spatial-temporal analysis of dengue fever in Brazil during 2015-2019



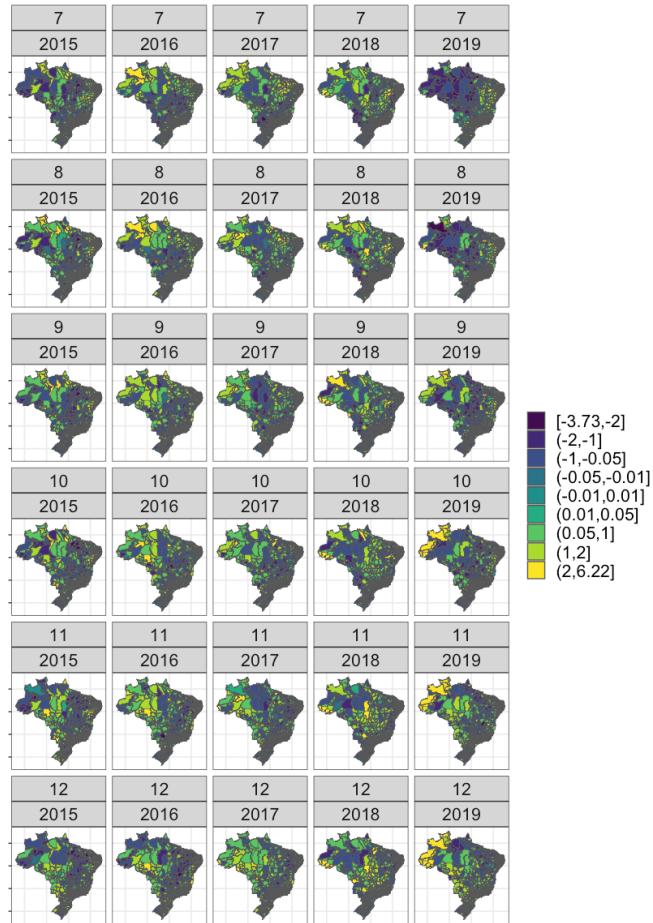
Supplementary Figure 4. Map of the random effect of the spatical-temporal model with interaction term.



Supplementary Figure 4. Map of the random effect of the spatical-temporal model with interaction term.



Supplementary Figure 5. Map of the random effect of the spatial-temporal model with interaction term and covariates.



Supplementary Figure 5. Map of the random effect of the spatial-temporal model with interaction term and covariates.

## References

- Aswi, A., S. M. Cramb, P. Moraga, and K. Mengersen. 2019. "Bayesian Spatial and Spatio-Temporal Approaches to Modelling Dengue Fever: A Systematic Review." *Epidemiology and Infection* 147: e33. <https://doi.org/10.1017/S0950268818002807>.
- Azevedo, Thiago S de, Camila Lorenz, and Francisco Chiaravalloti-Neto. 2020. "Spatiotemporal evolution of dengue outbreaks in Brazil." *Transactions of The Royal Society of Tropical Medicine and Hygiene* 114 (8): 593–602. <https://doi.org/10.1093/trstmh/traa030>.
- Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2014. *Hierarchical Modeling and Analysis for Spatial Data*. 2nd ed. Chapman & Hall/CRC.
- Besag, J., J. C. York, and A. Mollié. 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics (with Discussion)." *Annals of the Institute of Statistical Mathematics* 43: 1–59.
- Hasan, Sami Faisal1; Alalowi, Shamimul; Jamdar. 2016. "Dengue Virus: A Global Human Threat." *Journal of International Society of Preventive and Community Dentistry* 6(1): 1–6.
- IARC. 2016. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 109. Outdoor Air Pollution*. Lyon, France: International Agency for Research on Cancer.
- Lowe, Rachel, Trevor C. Bailey, David B. Stephenson, Richard J. Graham, Caio A. S. Coelho, Marilia Sá Carvalho, and Christovam Barcellos. 2011. "Spatio-Temporal Modelling of Climate-Sensitive Disease Risk: Towards an Early Warning System for Dengue in Brazil." *Computers & Geosciences* 37 (3): 371–81. <https://doi.org/https://doi.org/10.1016/j.cageo.2010.01.008> (<https://doi.org/10.1016/j.cageo.2010.01.008>).
- Lowe, Rachel, Trevor C. Bailey, David B. Stephenson, Tim E. Jupp, Richard J. Graham, Christovam Barcellos, and Marilia Sá Carvalho. 2013. "The Development of an Early Warning System for Climate-Sensitive Disease Risk with a Focus on Dengue Epidemics in Southeast Brazil." *Statistics in Medicine* 32 (5): 864–83. <https://doi.org/https://doi.org/10.1002/sim.5549> (<https://doi.org/10.1002/sim.5549>).
- Riebler A, Simpson D, Sørbye SH. 2016. "An Intuitive Bayesian Spatial Model for Disease Mapping That Accounts for Scaling." *Statistical Methods in Medical Research* 25(4): 1145–65. <https://doi.org/doi:10.1177/0962280216660421> (<https://doi.org/doi:10.1177/0962280216660421>).
- Rue, H., S. Martino, and N. Chopin. 2009. "Approximate Bayesian Inference for Latent Gaussian Model by Using Integrated Nested Laplace Approximations (with Discussion)." *J. R. Statist. Soc. B* 71: 319–92.