Spatio-temporal analysis of dengue fever in Brazil and its relationship to climate stressors

02364467

07/04/2023

Abstract

Dengue fever is a frequent epidemic in Brazil and has proven to be a large burden on the healthcare system of Brazil. The increase in temperatures from global warming has and will increase the risk upon areas of Brazil that where once unaffected as mosquito population densities increase . In this paper we evaluate the spatio-temporal patterns of dengue using Bayesian models and Integrated Nested Laplace Approximation across a 5 year time period (2015-2019) to understand the way in which dengue fever is spread across Brazil and the effects of climatic stressors like temperature and drought have on the spread. We find that the build-up of the virus begins in the most northern regions of Brazil and the high number of dengue fever cases are perpetuated by factors in the south eastern region containing the big cities like Rio and Sao Paulo.Climatic stressor seem to have a small effect on dengue cases however there is a chance these covariate effects are being clouded by spatial effect confounding and may need further analysis. The results highlight the importance of focusing health resources on the most northern regions of Brazil to mitigate the spread of the disease.

Introduction

Dengue is a globally expanding virus that can cause symptoms that range from a fever to more extreme cases of severe hemorrhaging (Zerfu, Kassa, and Legesse 2023) . As global warming increases global temperature , a common vector for the Dengue virus the aedes aqgypti mosquito has found more environments hospitable than before , exposing the virus to a much larger global population with putting near to a billion new people at risk within this century (Ryan et al. 2019). Brazil has been no stranger to the expansion of this virus as it has been shown to reach areas at higher latitudes and has effected 481 municipalities that have reports of dengue for the first time putting an additional 8.1 million people at risk. With the northern and southern regions being effected more due to their links with urbanization and global warming(Codeco et al. 2022).

To combat this Brazil has created a formidable surveillance system to look at the spread in dengue fever in an attempt to mitigate these cyclical dengue epidemics. However while the surveillance is robust , the healthcare system suffers from a lack of health inclusivity due to no universal healthcare availability and therefore the resources to deal with dengue fever are minimal relative to the virus outbreak numbers(Angelo et al. 2020). Being more specific in identifying these high risk areas and areas where dengue cases build up can help identify areas where these minimal health care resources can be allocated

Aims

The aim of this study is to analyze the spatio-temporal patterns of dengue fever outbreaks in Brazil over a 5-year period, using high-resolution micro region data. We also aim to evaluate the impact of climatic stressors on the spread of the disease. By identifying the factors that contribute to the cyclical nature of dengue fever outbreaks and the potential impact of climatic stressors, we hope to inform the development of targeted

public health interventions aimed at reducing the spread of the disease in Brazil.

Methods

Pre-processing

The study used data provided by the Brazilian Institute of Geography and Statistics to construct a map of Brazil's micro regions, with spatial weights between neighbors specified as snap =1000. Monthly dengue case data were obtained from the Ministry of Health Information Department, while Palmer's Drought Severity Index (PDSI) was obtained from the Climatic Research Unit at the University of East Anglia. Minimum temperature data were obtained from the NCAR Climate Data Guide, and accessibility to water networks data were obtained from the 2010 Brazilian census. Missing data for Fernando de Noronha island was removed, while no other data was removed due to no missingness being present in the dataset. The PDSI was rebanded based on parameters provided by NCAR climate data guide, while minimum temperature and percentage of people with access to water were kept continuous. A table (Table 1) was created, stratified by major region, using values from each micro region result from all 60 months, analyzing climatic stressors and case counts.

Initial SMR analysis

First a simple SMR analysis of the relative risk was conducted. In which $SMR_i = y_i/e_i$ where y_i is the sum of observed cases per microregion over 5 years (60 months) E_i is the sum of expected cases of microregion where $E_i = \frac{1}{10^5}\sum_{j=1}^{n} p_{ij}$ and i = microregion , n = total number of microregions and j = index of each microregion . Dividing by 10^5 allows to standardize the expected cases by converting them to an incidence rate per 100,000 people. The SMR of each microregion was then plotted (supplementary 2). This was used as a baseline to compare the model to more complex models that incorporate spatial and temporal effects.

Spatial analysis

A graph based on a hierarchical poisson log linear model is plotted. This is model is defined as :

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

Where O_i represents the dengue cases (O) in reigon i and assumes a poisson distribution with a mean of $\rho_i E_i$ with E_i = expected cases and ρ_i representin the relative risk in micro region i compared with the overall relative risk. A log linear function is used to model ρ_i with b_0 = fixed effect and b_1 = random effects. $\boldsymbol{b}$ represents the vector which will represent the random effects with the spatial structure being incorporated captured ICAR model in which $\boldsymbol{v}*$ is the standardized spatial structured adjacency between regions and $\boldsymbol{u}*$ = the spatially unstructured component.

The formula uses ID to identify each micro region , the adjacency matrix is held in the shape.adj object . To incorporate the ICAR structured model and the spatially unstructured effect a BYM2 model is used in which hyperparamter pc.prec and phi are used with their corresponding priors mentioned in the param argument to define the spatially structured and unstructured effects respectively (supplementary 4). No climatic stressor covaraites where used due to the inability of the model to converge. The waic is extracted to compare the two models also using the general formula used for each model (supplementary 3)

Spatio-temporal analysis

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

Where the priors from the previous spatial analysis are maintained a temporal component is added to the equation represented by t to incorporate the temporal effect. In additon to this γ_t is added which is a temporal random effect added with a rank 1 random walk in which we assume the observations and the previous observation before it are normally distributed with mean 0 and variance 1. Relating the temporal effect t to the previous temporal random effect indexed by $t-1$ with a degree of randomness. ψ_t is a temporal autocorrelation effect with an assumed normal distribution and a correlation between time point ψ_t and ψ_{t-1} where time that are further apart are less likely to be similar. Both γ_t and ψ_t have been incorporated in the f(ID_time,model"rw1") in the formula(supplementary 5).The Relative Risk and Posterior probabilities are then extracted using the extraction general extraction algorithm used for each model. (supplementary 6).

Spatio-temporal interaction (type 1 interaction)

The spatio temporal interaction can be described as :

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

While maintaining previous priors and parameters mentioned in spatial and spatio-temporal analyis wihout an interaction , a normal distibuted space time interaction paramter δ_{it} . This will account for any space time interaction that was not seen before and any in addition accounts for any additional covariate interaction across space and time. This is in incorporated f(ID.space.time,model = “iid”) in the formula (supplementary 7)

Spatio-temporal interaction type IV

Here instead of assuming the spatio-temporal interaction is linear , it is assumed that the spatio temporal interaction can shift in direction and strength. To do this the formula used and the ilna method are completely changed to implement this complex model per instruction from Goicoa et al. (2017). First a temporal structure was defined using RW1 model priors. Then a uniform distribution is implements on the standard deviation of random effects. We then define the spatial neighborhood matrix using the .graph model containing the neighbors and weights between the neighbors . Kronecker priors are formed to help define the type IV interaction using the RW1 in the formula (supplementary 8).Here “generic1” is used as the model for ID_space to incorporate an ICAR prior, “rw1” was used to define the temporal element and “generic0” was used to define a generic Gaussian prior to be used all in which the “sdunif” prior is responsible to apply the uniform distribution upon the random effects .

Results

Exploratory analysis

Initial temporal analysis

	Cases	
year	observed	expected
2015	1700074	24537.54
2016	1514569	24733.33
2017	243238	24918.95
2018	265408	25019.03
2019	1558311	25217.29

Table 1: Table showing the sum of observed and expected dengue cases in Brazil from 2015-2019

Table 1 shows that dengue cases in Brazil fluctuated greatly between 2015 and 2016, and then decreased to expected levels before rising again in 2019. This suggests that the dengue epidemic has a cyclical pattern over time.

Initial regional analysis

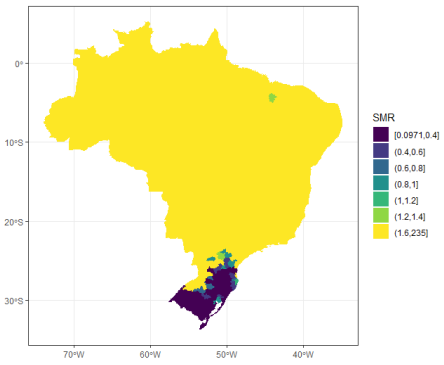


Figure 1: plot of SMR across microregion in Brazil between 2015-2019

	Centre-West (N=3120)	North (N=3840)	Northeast (N=11220)	South (N=5640)	Southeast (N=9600)	Overall (N=33420)
rebanded_pdsi						
normal	1701 (54.5%)	2116 (55.1%)	5798 (51.7%)	2684 (47.6%)	4760 (49.6%)	17059 (51.0%)
severe_extreme_drought	597 (19.1%)	654 (17.0%)	2130 (19.0%)	1248 (22.1%)	2084 (21.7%)	6713 (20.1%)
moderate_drought	504 (16.2%)	634 (16.5%)	1904 (17.0%)	1111 (19.7%)	1718 (17.9%)	5871 (17.6%)
unusually_moist	154 (4.9%)	193 (5.0%)	581 (5.2%)	294 (5.2%)	487 (5.1%)	1709 (5.1%)
very_extreme_moist	164 (5.3%)	243 (6.3%)	807 (7.2%)	303 (5.4%)	551 (5.7%)	2068 (6.2%)
tmin						
Mean (SD)	19.9 (2.30)	22.2 (1.46)	21.7 (1.68)	15.3 (3.52)	17.5 (3.05)	19.3 (3.60)
Median [Min, Max]	20.6 [10.0, 24.2]	22.4 [15.5, 25.8]	21.9 [14.7, 25.8]	15.6 [5.44, 22.6]	18.0 [7.02, 24.4]	20.2 [5.44, 25.8]
water_network (percentage)						
Mean (SD)	74.1 (12.7)	48.3 (18.4)	68.3 (13.0)	78.8 (12.5)	82.0 (12.7)	72.2 (17.0)
Median [Min, Max]	76.3 [35.9, 94.8]	46.0 [16.7, 90.5]	69.2 [24.3, 97.8]	80.7 [27.1, 95.7]	85.8 [39.7, 98.8]	74.9 [16.7, 98.8]
observed_cases (per microregion result in region)						
Mean (SD)	277 (1220)	38.4 (109)	91.1 (412)	31.4 (190)	320 (2300)	158 (1320)
Median [Min, Max]	25.0 [0, 21500]	9.00 [0, 2310]	8.00 [0, 13400]	1.00 [0, 4490]	15.0 [0, 88800]	8.00 [0, 88800]
expected_cases (per microregion result in region)						
Mean (SD)	3.05 (5.43)	2.81 (4.02)	3.04 (4.91)	3.15 (5.30)	5.44 (15.7)	3.72 (9.43)
Median [Min, Max]	1.22 [0.302, 30.4]	1.85 [0.210, 24.7]	1.73 [0.260, 38.8]	1.72 [0.250, 38.9]	2.28 [0.275, 150]	1.86 [0.210, 150]

Table 2 : Table showing stats for the covariates and observed and expected cases for each major region in Brazil based on averages taken from 60 results per micro region over the 60 month time period

Looking at the SMR plot we can see that in most of Brazil has a very high SMR due to a very large number of observed cases in some cases the observed cases being 265 times greater than the number of expected cases. However the more southern regions seem to be able to maintain a number of observed cases similar or even lower that what we would expect, this makes the southern areas a point of interest to see how the models deal with these regions given the general pattern of high counts of observed cases in the rest of Brazil. The table1 also shows some interesting insights. The northern regions have less access to the piped water than the other regions , furthermore all the regions seem to be suffering from a relatively high amount of drought with all regions experiencing above 30% of moderate to extreme drought. This provides a basis for including these climatic stressors in the model to see if they can help explain the patterns in the amount of dengue cases.

Spatial and Spatio-temporal analysis

Comparing plots

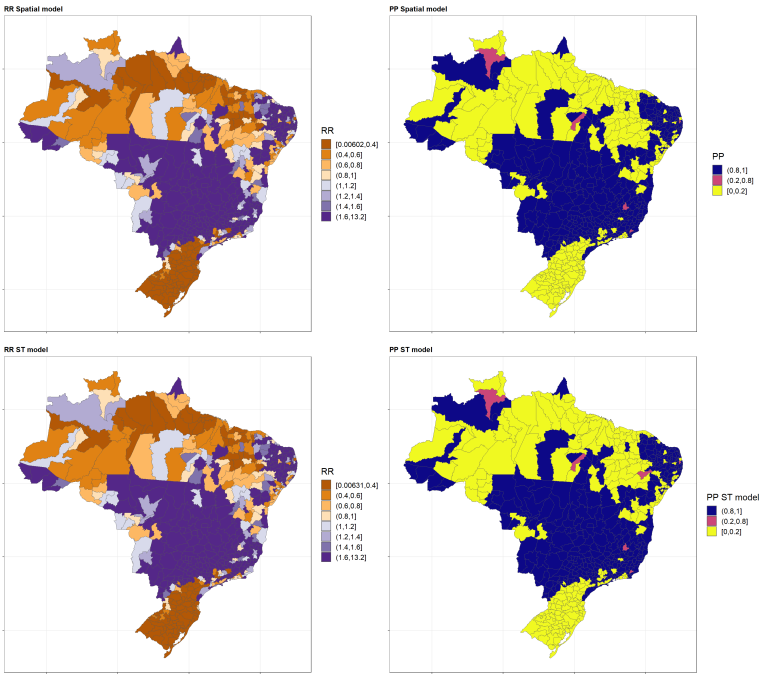


Figure 2: This figure plots the Relative Risk (RR) and Posterior probability (PP) of a spatial model and Spatio-temporal (ST) model of Brazil.

The initial spacial model and and spatio-temporal model are very similar and show clustering of high Relative Risk toward central Brazil. Based upon the posterior probability the degree of certainty within this central cluster is very high however both models show a low degree of certainty near the southern region where the observed cases are low in some areas, the models also struggle with the northern regions . It is clear that just adding a temporal effect and covariates to the model is not enough to show the underlying patterns within the data.

Adding Spatio-temporal interaction terms

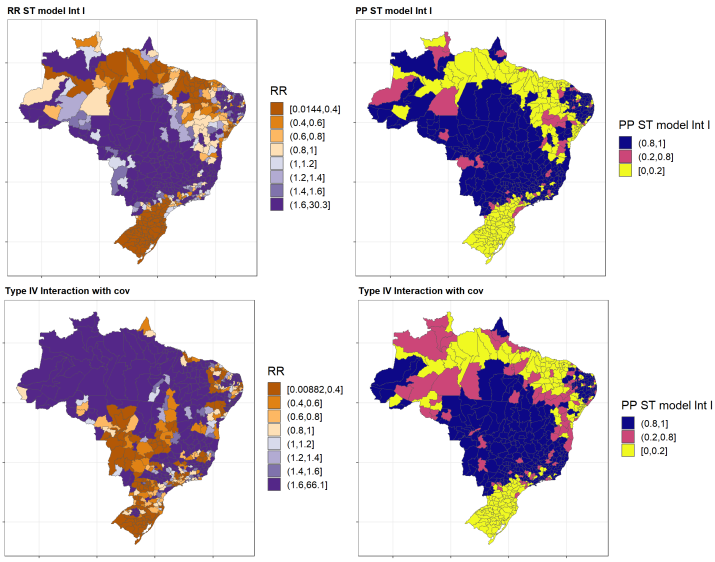


Figure 3: This shows the RR and PP plots of spatio temporal models with type I interaction (top row) and Type IV interaction (bottom row)

When we add a type I interaction the maximum relative risk increases by nearly double as more northern areas are incorporated into the model suggesting that there is in fact a spatio-temporal interaction that is changing the model drastically (figure 3). While type 1 interaction assumes a fixed spatial pattern , type 4 interaction model do not (supplemetary 9). Type 4 models where explored to see if the model can better interpret the spatial - temporal interaction over time. Looking at the type 4 interaction plot it is clear that there does seem to be changing pattern as now the

high Relative Risk cluster moves further north and again a higher relative risk is incorporated at 66.1 as more northern regions are considered high risk in the model. When looking at the underlying interaction across the 60 months , a pattern arises , in which when dengue cases are low the interaction is high in the more northern regions then the interaction moves to the south and south eastern regions when the dengue cases rise (supplementary 9).

Comparing covariate cilmatic stressors

Watanabe-Akaike information criterion (WAIC)

Table 3 : Table showing the WAIC of each model with and without covariates

Model	Waic (with cov)	WAIC (without cov)
Spatio-temporal model no interaction	21919663	21708939
Spatio-temporal model type I interaction	195923.1	216278.2
Spatio-temporal model type IV interaction	196229.48	196777.8

While especially in the final type 4 interaction model the WAIC is lower in the models with covariates the overall difference between the model is very small relative to the scale of the WAIC values for each model. Adding the covariates as fixed effects to the model is not as effective suggesting that these three covariates where not as effective in describing the model .

Analyzing the fixed effects

Table 4 : Table showing the multiplicative factor e^(mean) of each fixed effects' relationship to response variable (dengue case count)

Fixed effects:

	Multiplicative factor	mean	sd
(Intercept)		-2.734	0.235
tmin	1.17	0.156	0.005
rebanded_pdsisevere_extreme_drought	0.98	-0.02	0.029
rebanded_pdsimoderate_drought	1.06	0.055	0.024
rebanded_pdsiunusually_moist	1.00	-0.001	0.041
rebanded_pdsivery_extreme_moist	0.97	-0.034	0.045
water_network	1.02	0.016	0.003

By analyzing the fixed effect parameters of the cilmatic stressor covaraites we see that minimum temperature and moderate drought seem to have an effect on increasing dengue cases. However , given that temperature varies by 10 degrees celcius , a multiplicative factor of 1.17 is still small relative to large amount of cases.

Discussion and conclusion

The results do overall show a spatio-temporal interaction when looking at dengue fever. We see that when we consider that the interaction between the micro regions is modeled a a non - fixed effect that can change direction and magnitude of the spatio-temporal interaction between micro regions. Incorporating these interaction within this model we are able to identify the northern regions that seem to be most at risk when looking at these interaction. When looking at interaction across the 60 months the pattern in the northern regions (2017 - 2018) suggest that the build up of dengue cases occurs in the most northern regions and spreads towards the south eastern regions over time. Moving into 2019, when dengue cases are much higher, the interaction term is highest in the south eastern regions suggesting that the big cities like Rio and Sao Paulo become are responsible for spreading the disease at it peaked. Based on this we propose that dealing with dengue fever at the build up in the North may ease the burden during the a dengue fever outbreak in the larger cities later. However as we can see from the posterior proabilities , no model is able to completely explain the spatio-temporal patterns especially in the most northern and southern regions. This could be due to the low number of observed cases in those regions . We suggest a sub analysis in the future to look more specifically at the southern and northern regions or to incorporate parameters within the models that may be able to explain drastic changes in cases between regions.

Looking at the covriates involved in these models , the evidence show small multiplicative effect upon the number of dengue cases. This may be due to confounding between the spatial random effects , which has been known to effect result of the fixed effect covariates. New models , involving restricted regression can be used to alleviate this confounding making results clearer Adin et al. (2021).

Given that the northern regions of Brazil show high relative risk for dengue fever and dengue epidemics build up in the north , the effects of global warming may cause more drastic effects as mosquitoes start to find the north more habitable . Already papers have shown that the density of mosquito populations will increase as temperature rise , increasing the amount of dengue vector available Bonnin et al. (2022). Future analysis is required to further understand the relationship between global warming , dengue vectors and dengue fever in Brazil.

Supplementary material

Include here the supplementary material, such as the code (this is mandatory) or additional exploratory analyses or maps/plots. The Supplementary material is an extra session, additional to the 5-pages of actual scientific mini-project report. To include the code, you could include it into the code chunk, setting `eval=FALSE` and `echo=TRUE`. Here an example:

Supplementary 1

```
#Loading package ----
library(dplyr)        # A package for data manipulation
library(sf)           # Simple feature for R
library(spdep)        # Functions and tests for evaluating spatial patterns
library(tidyr)        # Tools to create tidy data
library(INLA)         # Integrated Nested Laplace Approximation package
library(ggplot2)      # A package for creating maps and graphs
library(viridis)      # A package providing color palettes
library(patchwork)    # A package to compose plots

# For tables in RMarkdown
library(knitr)
library(kableExtra)
#Looking at the data
#first we will set the working directory
setwd("~/OneDrive/Documents/HDA/Advanced analytics/AA_project/DATA_SCIENTIFIC_PROJECTS/
DS2_DengueBrazil")
#Load in the csv
data <- read.csv("data_2015_2019.csv")
#dealing with the missing microregion ------
#The PDSI was missing for the island FERNANDO DE NORONHA , microregion code 26019
##thus excluding microregion leaving only 557 microregions left..

###pre processing for map ###
#Install sf package
install.packages("sf")
#using the sf packages to read shapefile
library(sf)
#sf package used to read in brazil layers
shape <- read_sf(dsn = ".", layer = "shape_brazil")
#removing microregion code 26019
shape <- subset(shape, code!= 26019)
#now we hae removed that code as that microregion is unavaliable
#viewing map
library(ggplot2)
shape %>%
  ggplot() +
  geom_sf()
#creating neighbours and weights between the region
shape_nb = poly2nb(shape, snap=1000, queen=TRUE)
summary(shape_nb)
#adding the neighbour
nb2INLA("shape.graph",shape_nb)
shape.adj = paste(getwd(),"/shape.graph",sep="")
#Looking at the regions
ggplot() +
  geom_sf(data = shape, color = "blue", fill = "white") +
  coord_sf() +    #axis limits and CRS
  theme_bw() +    # dark-on-light theme
  theme(axis.title = element_text(size = 14),
        axis.text = element_text(size = 12))
#we can confirm that the arrangment of the microegion fits the data presented


###pre processing of data###
```

```
#Looking for missingness
sapply(data,is.na)
colSums(is.na(data))
#there is no missingess (good)
## we need to include expected cases
data$E  <- data$population/10^5
#use offset=log(E) within the inla function

#rebanding PDSI to categorical varaible for analysis
data$rebanded_pdsi <- cut(data$pdsi,
                                breaks = c(-Inf, -3, -2, 1.9, 2.9, Inf),
                                labels = c("severe_extreme_drought", "moderate_drough
t",
                                           "normal", "unusually_moist", "very_extreme
_moist"))
data$rebanded_pdsi <- relevel(RESP_DATA_ST$rebanded_pdsi, ref = "normal")

### Making table 1 ###

# Subset the data to include only the relevant columns
sub_data <- data1 %>%
  select(year, region_name, dengue_cases, E, water_network, tmin, pdsi, rebanded_pdsi)
%>%
  rename(observed_cases = dengue_cases, expected_cases = E)

# Create a table with counts and standard deviations
my_table <- table1(~ rebanded_pdsi + tmin + water_network + observed_cases + expected_c
ases | region_name, data=sub_data)

# Print the table
my_table
```

Supplementary 2

```
###Analysing SMR###

#Aggregating data by microregion to get total cases peer region across the 60 ##months
data_agg = data%>% group_by(code) %>%
  summarize(observed = sum(dengue_cases),
            expected = sum(E)) %>%
  dplyr::rename(O = observed, E = expected)
#producing SMR using expected and outcome values
data_agg = data_agg %>% mutate(SMR = O/E)
#producing spatial map of aggregeated SMR only looking at SMR right now
data_agg$SMRcat = cut(data_agg$SMR,
                      breaks=c(min(data_agg$SMR),
                               0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6,
                               max(data_agg$SMR)), include.lowest = T)

map_SMR = left_join(shape, data_agg, by = c("code" = "code"))
#plotting map
SMR_plot<-ggplot() + geom_sf(data = map_SMR, col = NA) + aes(fill = SMRcat) +
  theme_bw() + scale_fill_viridis_d() +
  guides(fill=guide_legend(title="SMR"))
```

Supplementary 3

```
# formula
formula_BYM2 = O ~ avg_tmin + avg_pdsi +
                     f(ID,model="bym2",
                       graph=shape.adj,
                       hyper=list(prec = list(
                       prior = "pc.prec",
                       param = c(0.5 / 0.31, 0.01)),
                       phi = list(
                         prior = "pc",
                         param = c(0.5, 2 / 3))))

<model without covariates> = inla(formula=formula_BYM2, family="poisson", data=data_ag
g,
                  offset=log(E), control.compute=list(waic=TRUE))
formula_BYM2_cov = O ~ avg_tmin + avg_pdsi +
  f(ID,model="bym2",
    graph=shape.adj,
    hyper=list(prec = list(
      prior = "pc.prec",
      param = c(0.5 / 0.31, 0.01)),
      phi = list(
        prior = "pc",
        param = c(0.5, 2 / 3))))
<model with covariates>= inla(formula=formula_BYM2_cov, family="poisson", data=data_ag
g,
                      offset=log(E), control.compute=list(waic=TRUE))
#comparing waic of two models
<model without covariates>$waic[1]
<model with covariates>$waic[1]
#ver little difference however covaraites are used for model
```

Supplementary 4

```
#formula for spatial model with covariates
formula_BYM2_cov = O ~
  f(ID,model="bym2",
    graph=shape.adj,
    hyper=list(prec = list(
      prior = "pc.prec",
      param = c(0.5 / 0.31, 0.01)),
      phi = list(
        prior = "pc",
        param = c(0.5, 2 / 3))))
sBYM.model= inla(formula=formula_BYM2_cov, family="poisson", data=data_agg,
                      offset=log(E), control.compute=list(waic=TRUE))
```

Supplementary 5

```
#formula for spatio temporal model no interaction
formula_ST_noint = O ~ f(ID_space, model="bym2", graph=shape.adj,
                         hyper=list(prec = list(
                           prior = "pc.prec",
                           param = c(0.5 / 0.31, 0.01)),
                           phi = list(
                             prior = "pc",
                             param = c(0.5, 2 / 3)))) + f(ID_time,model="rw1",
                                                          hyper=list(prec = list
                                                                     (
                             prior = "pc.prec",
                             param = c(0.5 / 0.31, 0.01))))
```

Supplementary 6 (General plot extraction algorithm)

```
#Spatial Relative risks extracted from model
RR_stBYM = c()

for(i in 1:557){
  RR_stBYM[i] = inla.emarginal(function(x) exp(x),
                               <model>$marginals.random$ID_space[[i]])
}

#Posterior probabilities (for spatial RR)
RR_stBYM_marg = <model>$marginals.random$ID_space[1:557]
PP_stBYM = lapply(RR_stBYM_marg, function(x) {1-inla.pmarginal(0,x)})

#Temporal Relative risks and CI95
RR_stRW_RR = c()
RR_stRW_lo = c()
RR_stRW_hi = c()

for(i in 1:60){
  #Posterior mean
  RR_stRW_RR[i] = inla.emarginal(function(x) exp(x),
                                 <model>$marginals.random$ID_time[[i]])
  #2.5% quantile
  RR_stRW_lo[i] = inla.qmarginal(0.025,inla.tmarginal(function(x) exp(x), <model>$margi
nals.random$ID_time[[i]]))
  #97.5% quantile
  RR_stRW_hi[i] = inla.qmarginal(0.975, inla.tmarginal(function(x) exp(x), <model>$marg
inals.random$ID_time[[i]]))
}

RR_stRW = data.frame(RR=RR_stRW_RR,low=RR_stRW_lo,high=RR_stRW_hi)
#Plot the temporal residual RRs (RR_stWR)
Temp1 = ggplot(RR_stRW, aes(seq(1,60), RR)) + geom_line() +
  ggtitle("ST model No Int") + geom_ribbon(aes(ymin=low,ymax=high), alpha=0.2) + labs(x
="month")

Temp1
#plotting porper cycle
ggplot(RR_stRW, aes(seq(1,60), RR)) +
  geom_line() +
  ggtitle("ST model No Int") +
  geom_ribbon(aes(ymin=low,ymax=high), alpha=0.2) +
  labs(x="month/year") +
  scale_x_continuous(breaks = seq(1,60,12),
                     labels = c("Month 1\n2015", "Month 13\n2016", "Month 25\n2017", "M
onth 37\n2018", "Month 49\n2019"))

#using code to plot on the map
resRR_PP_st = data.frame(resRR=RR_stBYM,
                         PP=unlist(PP_stBYM),
                         code=data_agg[,2])
# breakpoints
resRR_PP_st$resRRcat = cut(resRR_PP_st$resRR, breaks=c(min(resRR_PP_st$resRR),
                                                       0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6,
                                                       max(resRR_PP_st$resRR)),include.
lowest = T)
```

```
resRR_PP_st$PPcat = cut(resRR_PP_st$PP, c(0, 0.2, 0.8, 1.00), include.lowest = TRUE)

map_RR_ST = left_join(shape, resRR_PP_st, by = c("code" = "code"))
#plotting map
p3 = ggplot() + geom_sf(data = map_RR_ST) + aes(fill = resRRcat) +
  theme_bw() + scale_fill_brewer(palette = "PuOr") +
  guides(fill=guide_legend(title="RR")) +  ggtitle("RR Spatio-temporal model") +
  theme(text = element_text(size=15),
        axis.text.x = element_blank(),
        axis.text.y = element_blank(), plot.title = element_text(size = 12, face = "bol
d"))

p4 = ggplot() + geom_sf(data = map_RR_ST) + aes(fill = PPcat) +
  theme_bw() +
  scale_fill_viridis(
    option = "plasma",
    name = "PP Spatio-temporal model",
    discrete = T,
    direction = -1,
    guide = guide_legend(
      title.position = 'top',
      reverse = T
    )) +  ggtitle("PP ST model") + theme(text = element_text(size=15),
                                         axis.text.x = element_blank(),
                                         axis.text.y = element_blank(), plot.title = el
ement_text(size = 12, face = "bold"))

(p1|p2) / (p3|p4)
```

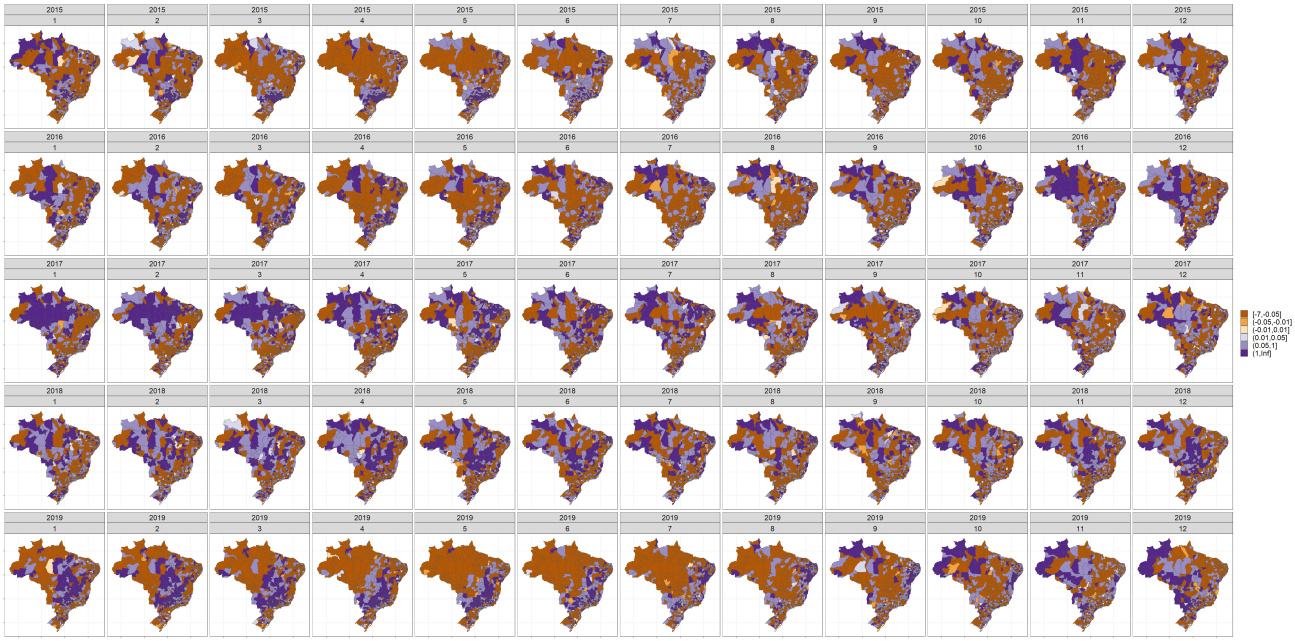
Supplementary 7

```
#formula for model with type 1 interaction
formula_ST_intI = O ~ f(ID_space, model="bym2", graph=shape.adj,
                        hyper=list(prec = list(
                          prior = "pc.prec",
                          param = c(0.5 / 0.31, 0.01)),
                          phi = list(
                            prior = "pc",
                            param = c(0.5, 2 / 3)))) +
  f(ID_time,model="rw1", hyper=list(prec = list(
    prior = "pc.prec",
    param = c(0.5 / 0.31, 0.01))))+
  f(ID.space.time,model="iid", hyper=list(prec = list(
    prior = "pc.prec",
    param = c(0.5 / 0.31, 0.01)))) + tmin + rebanded_pdsi + water_network
```

Supplmentary 8

```
#formula for building the spatio-temporal interaction model type 4 interaction
##defining temporal strcure of matrix
D1 <- diff(diag(60),differences=1)
Q.gammaRW1 <- t(D1)%*%D1

D2 <- diff(diag(60),differences=2)
Q.gammaRW2 <- t(D2)%*%D2

#definining spatial neighbourhood matrix
g <- inla.read.graph("shape.graph")
Q.xi <- matrix(0, g$n, g$n)
for (i in 1:g$n) {
  Q.xi[i,i] = g$nnbs[[i]]
  Q.xi[i,g$nbs[[i]]] = -1
}
Q.Leroux <- diag(557)-Q.xi
#knocker product
R <- kronecker(Q.gammaRW1,Q.xi)
r.def <- 557 + 60-1
A1 <- kronecker(matrix(1,1,60),diag(557))
A2 <- kronecker(diag(60),matrix(1,1,557))
A.constr <- rbind(A1,A2)
# controlling standard deviation of covariates
sdunif = "expression:
logdens = -log_precision/2;
return(logdens)"
#spatial smoothin of covaraites
lunif = "expression:
a = 1;
b = 1;
beta = exp(theta)/(1+exp(theta));
logdens = lgamma(a+b)-lgamma(a)-lgamma(b)+(a-1)*beta+(b-1)*(1-beta);
log_jacobian = log(beta*(1-beta));
return(logdens+log_jacobian)"
#creating data frame
RESP_DATA_ST<-as.data.frame(RESP_DATA_ST)

formula <- O ~ f(ID_space, model="generic1", Cmatrix=Q.Leroux, constr=TRUE,
                 hyper=list(prec=list(prior=sdunif),beta=list(prior=lunif))) +
  f(ID_time, model="rw1", constr=TRUE,
    hyper=list(prec=list(prior=sdunif))) +
  f(ID.space.time, model="generic0", Cmatrix=R, rankdef=r.def,
    constr=TRUE, hyper=list(prec=list(prior=sdunif)),
    extraconstr=list(A=A.constr, e=rep(0,557+60)))


stIntI.BYM.model = inla(formula=formula, family="poisson", data=RESP_DATA_ST, offset =
log(E),
                        control.predictor=list(compute=TRUE),
                        control.compute=list(dic=TRUE, waic=TRUE),
                        verbose=TRUE
```

Supplementary 9



References

- You should update the attached file `biblio.bib` with your references.
- The references cited in the mini-project report will be automatically inserted after this header.

Adin, Aritz, Tomás Goicoa, James S Hodges, Patrick M Schnell, and María D Ugarte. 2021. “Alleviating Confounding in Spatio-Temporal Areal Models with an Application on Crimes Against Women in India.” *Statistical Modelling* 23 (1): 9–30. https://doi.org/10.1177/1471082x211015452 (https://doi.org/10.1177/1471082x211015452).

Angelo, Marco, Walter Massa Ramalho, Helen Gurgel, Nayara Belle, and Eva Pilot. 2020. “Dengue Surveillance System in Brazil: A Qualitative Study in the Federal District.” *International Journal of Environmental Research and Public Health* 17 (6): 2062. https://doi.org/10.3390/ijerph17062062 (https://doi.org/10.3390/ijerph17062062).

Bonnin, Lucas, Annelise Tran, Vincent Herbreteau, Sébastien Marcombe, Sébastien Boyer, Morgan Mangeas, and Christophe Menkes. 2022. “Predicting the Effects of Climate Change on Dengue Vector Densities in Southeast Asia Through Process-Based Modeling.” *Environmental Health Perspectives* 130 (12). https://doi.org/10.1289/ehp11068 (https://doi.org/10.1289/ehp11068).

Codeco, Claudia T., Sara S. Oliveira, Danielle A. C. Ferreira, Thais I. S. Riback, Leonardo S. Bastos, Raquel M. Lana, Iasmim F. Almeida, Vinicius B. Godinho, Oswaldo G. Cruz, and Flavio C. Coelho. 2022. “Fast Expansion of Dengue in Brazil.” *The Lancet Regional Health - Americas* 12 (August): 100274. https://doi.org/10.1016/j.lana.2022.100274 (https://doi.org/10.1016/j.lana.2022.100274).

Goicoa, T., A. Adin, M. D. Ugarte, and J. S. Hodges. 2017. “In Spatio-Temporal Disease Mapping Models, Identifiability Constraints Affect PQL and INLA Results.” *Stochastic Environmental Research and Risk Assessment* 32 (3): 749–70. https://doi.org/10.1007/s00477-017-1405-0 (https://doi.org/10.1007/s00477-017-1405-0).

Ryan, Sadie J., Colin J. Carlson, Erin A. Mordecai, and Leah R. Johnson. 2019. “Global Expansion and Redistribution of Aedes-Borne Virus Transmission Risk with Climate Change.” Edited by Barbara A. Han. *PLOS Neglected Tropical Diseases* 13 (3): e0007213. https://doi.org/10.1371/journal.pntd.0007213 (https://doi.org/10.1371/journal.pntd.0007213).

Zerfu, Biruk, Tesfu Kassa, and Mengistu Legesse. 2023. “Epidemiology, Biology, Pathogenesis, Clinical Manifestations, and Diagnosis of Dengue Virus Infection, and Its Trend in Ethiopia: A Comprehensive Literature Review.” *Tropical Medicine and Health* 51 (1). https://doi.org/10.1186/s41182-023-00504-0 (https://doi.org/10.1186/s41182-023-00504-0).