Statistical modelling for infectious diseases

Part 2: Epidemic thresholds and delay correction

Leo Bastos
PROCC/Fiocruz

Leonardo.bastos@fiocruz.br
@leosbastos

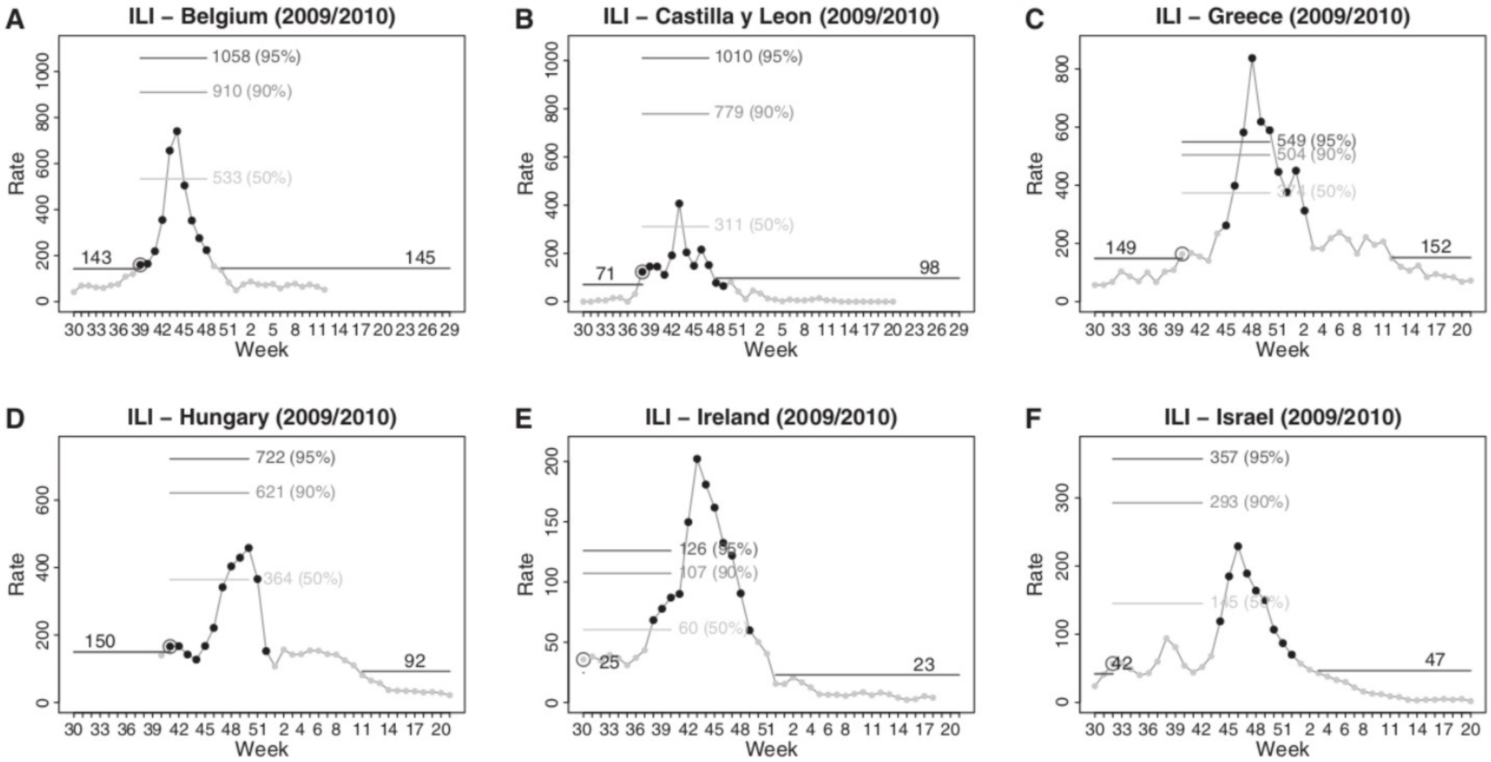
Summary

- Defining epidemic thresholds
  - MEM
  - Predictive distribution based
- Delay correction
  - Background
  - Nowcasting
  - Extensions

Moving Epidemic Method (MEM)

- Vega et al (2012) proposed an algorithm that based on historical data of ILI
  1) Estimates start, duration and end of an epidemic
  2) Calculates thresholds pre- and post-epidemic
  3) Provide different intensity levels for an epidemic period

MEM



A
ILI – Belgium (2009/2010)
1058 (95%)
910 (90%)
533 (50%)
143
145
Rate
Week
B
ILI – Castilla y Leon (2009/2010)
1010 (95%)
779 (90%)
311 (50%)
71
98
C
ILI – Greece (2009/2010)
549 (95%)
504 (90%)
152
149
D
ILI – Hungary (2009/2010)
722 (95%)
621 (90%)
364 (50%)
150
92
E
ILI – Ireland (2009/2010)
126 (95%)
107 (90%)
60 (50%)
25
23
F
ILI – Israel (2009/2010)
357 (95%)
293 (90%)
42
47

MEM: R package

The Moving Epidemic Method

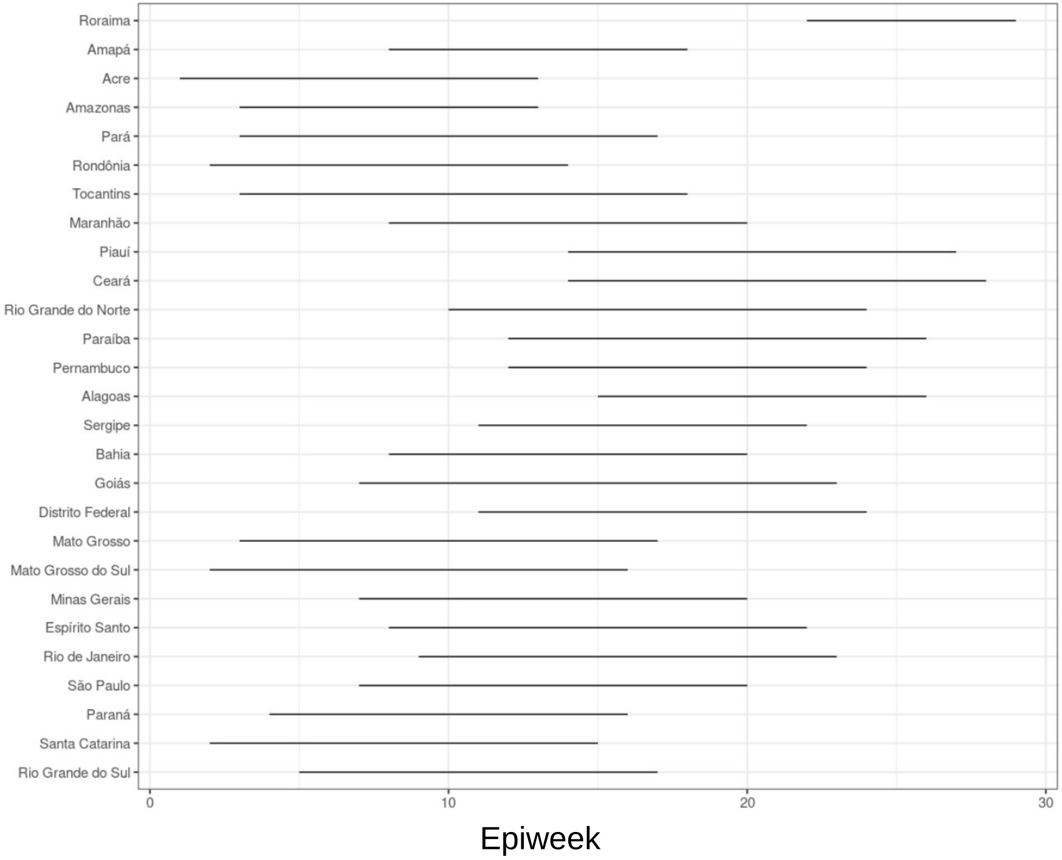
Documentation for package ‘mem’ version 2.16

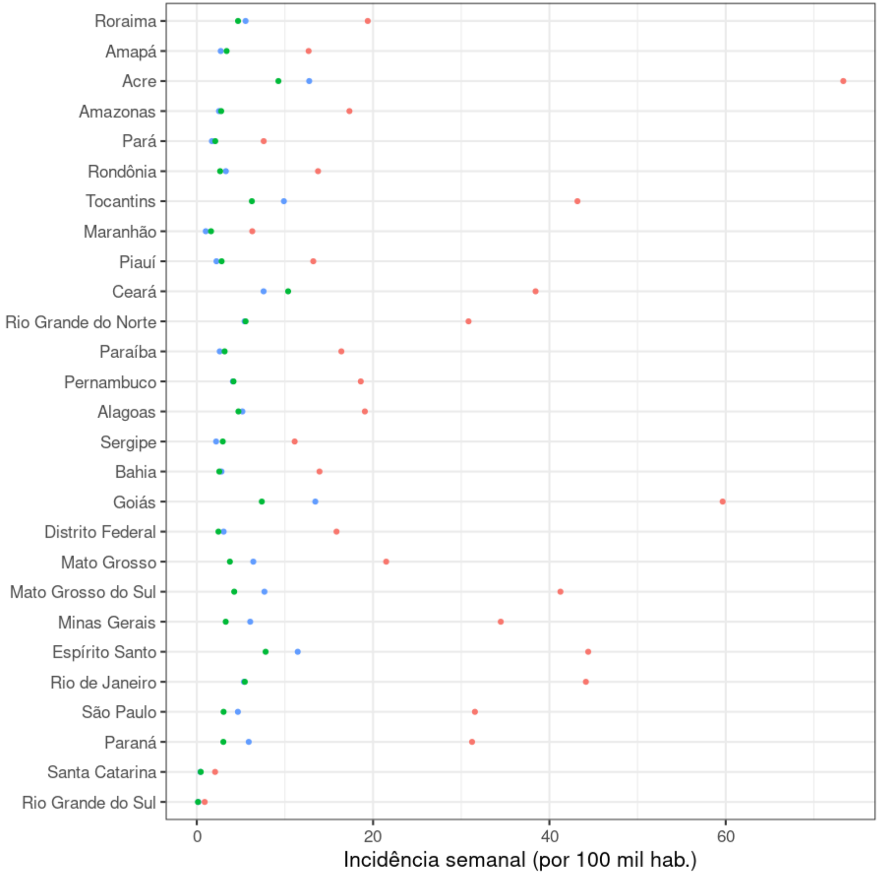
- DESCRIPTION file.

Help Pages

epimem	Deprecated function(s) in the mem package
epitiming	Deprecated function(s) in the mem package
flucyl	Castilla y Leon influenza crude rates
flucylraw	Castilla y Leon influenza standarised rates
full.series.graph	Creates the historical series graph of the datasets
memevolution	Evolution of estimators
memgoodness	Goodness of fit of the mem
memintensity	Thresholds for influenza intensity
memmodel	Methods for influenza modelization
memstability	Stability of indicators
memsurveillance	Creates the surveillance graph of the current season
memsurveillance.animated	Creates the animated graph of the surveillance of the current season
memtiming	Influenza Epidemic Timing
memtrend	Methods for influenza trend calculation
optimum.by.inspection	Inspection calcultation of the optimum
processPlots	Full process plots for mem
roc.analysis	Analysis of different indicators to find the optimum value of the window parameter
summary.epidemic,plot.epidemic,print.epidemic	Influenza Epidemic Timing
summary.flu,plot.flu,print.flu	Methods for influenza modelization
transformdata	Data transformation
transformdata.back	Data transformation
transformseries	Transformation of series of data

MEM: Dengue in Brazilian states






WE BROUGHT HER IN TO HELP WITH OUR ~~SALES~~ PREDICTIONS.
Disease
N.FAKES

Predicting infectious diseases

$$Y_t \sim NegBin(Pop_{Ano[t]}\lambda_t, \phi)$$

$$log(\lambda_t) = \alpha + \beta_{Ano[t]} + \gamma_{Semana[t]}$$

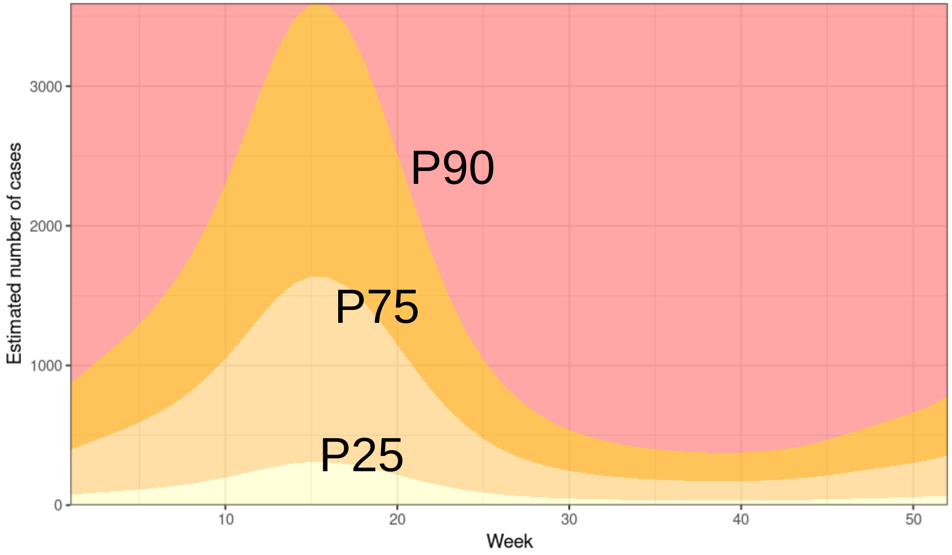
$$\beta_{Ano} \sim \text{“iid”}$$

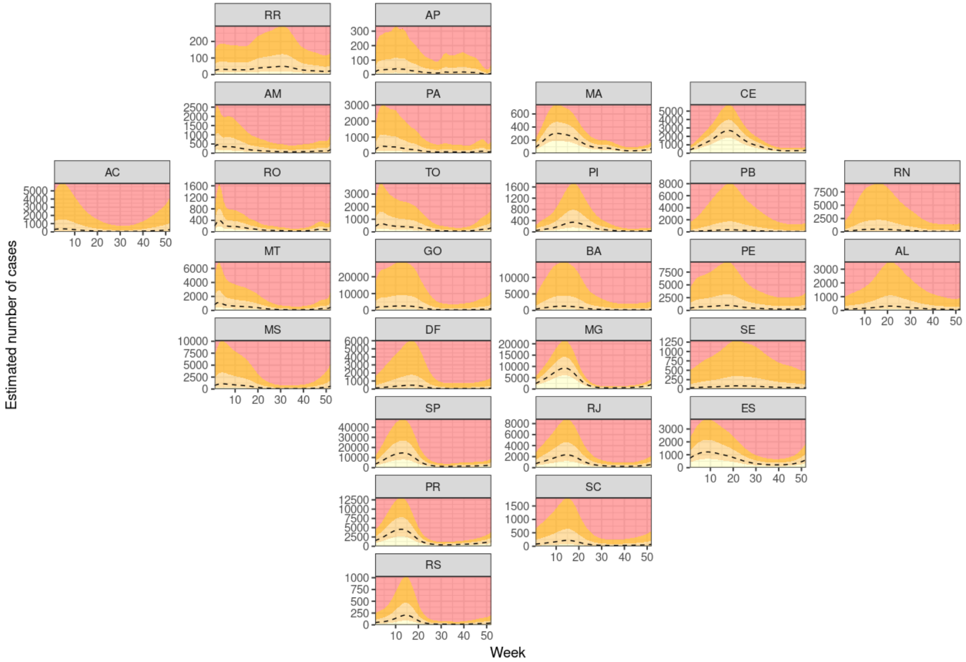
$$\gamma_{Semana} \sim \text{“cyclic rw2”}$$

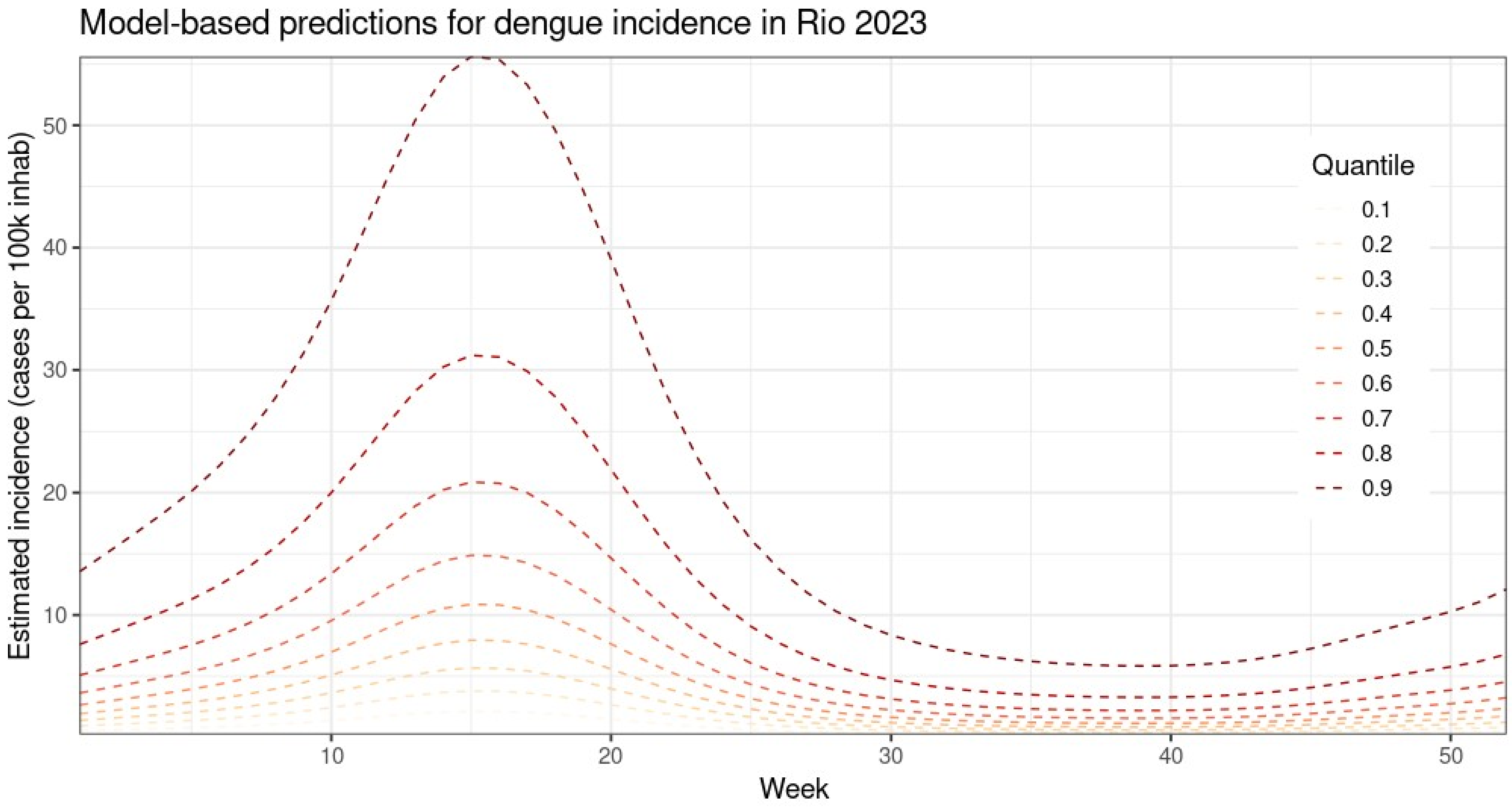
$$p(y_{Sem:1:52,Ano:2023} | y_{Sem:1:52*,Ano:2010:2022}) = ?$$

Dengue predictions for 2023

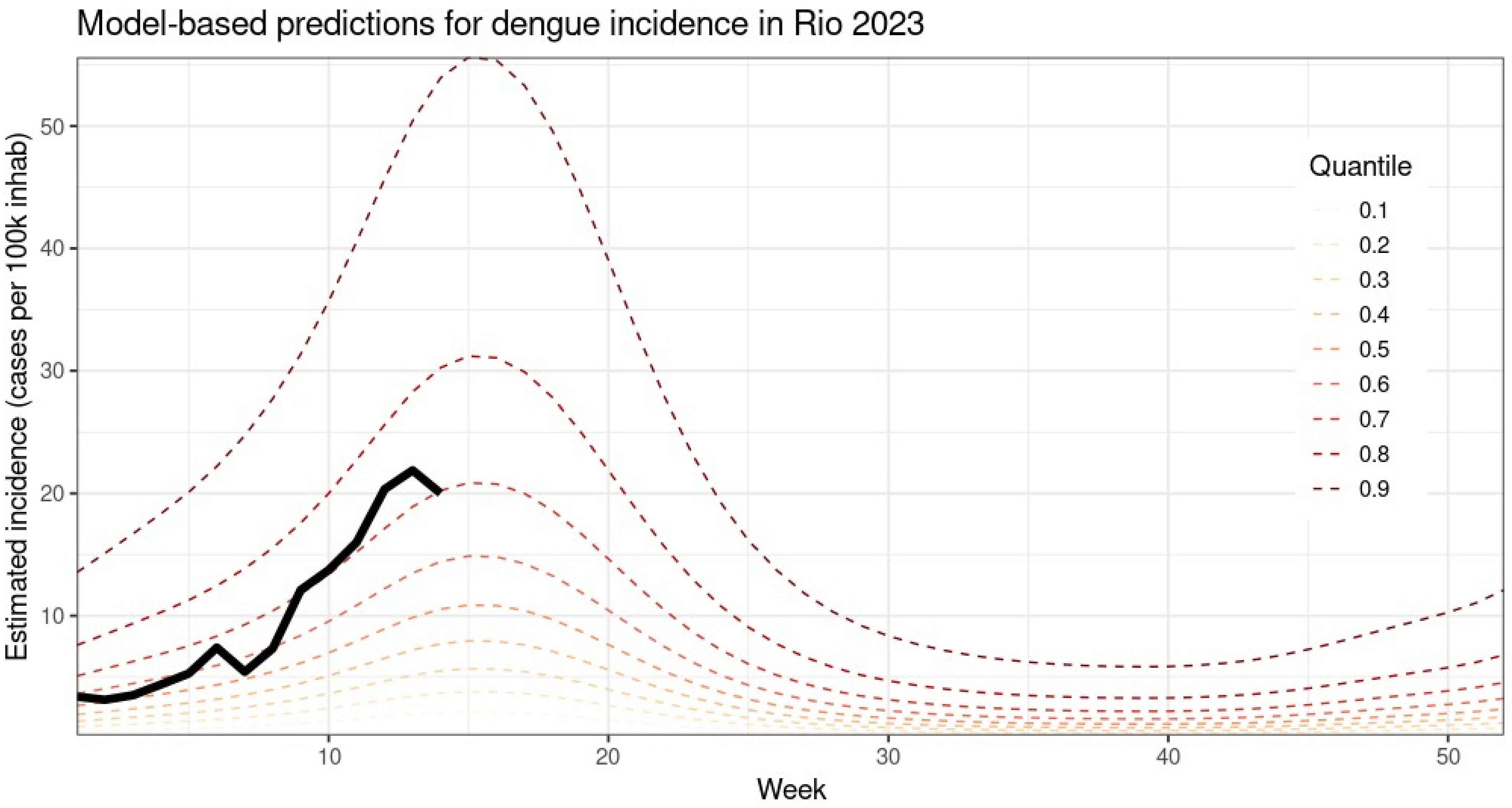
Rio de Janeiro







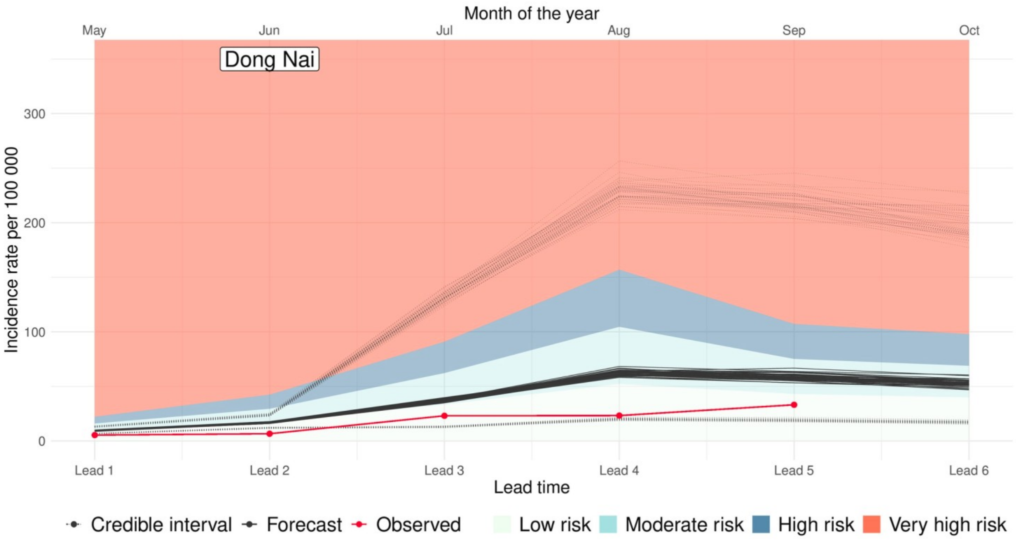
Model-based predictions for dengue incidence in Rio 2023
Estimated incidence (cases per 100k inhab)
50
40
30
20
10
10
20
30
40
50
Week
Quantile
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

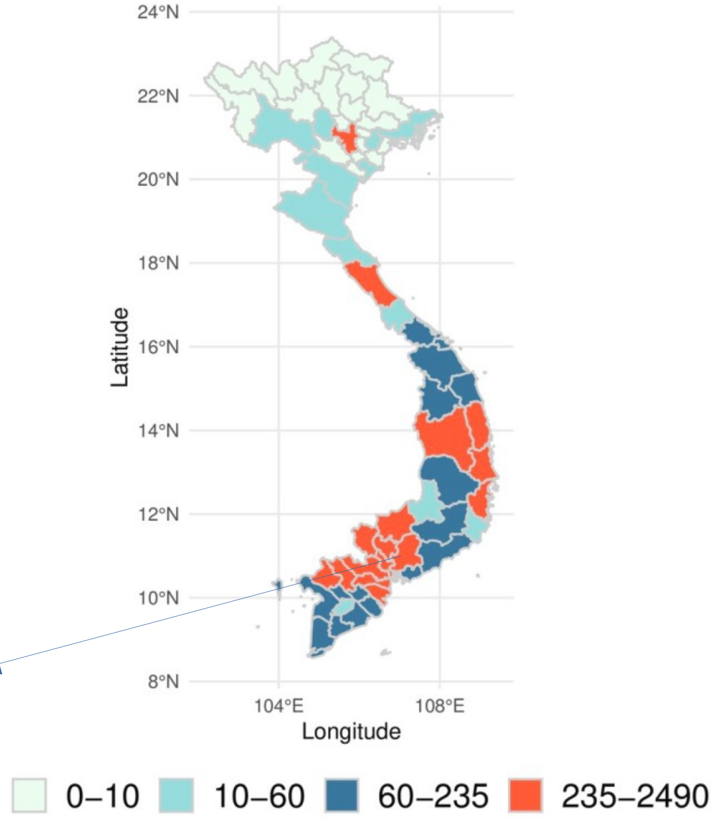


Model-based predictions for dengue incidence in Rio 2023
Estimated incidence (cases per 100k inhab)
Week
Quantile
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
10
20
30
40
50

Predicting dengue in Vietnam

- Forecasts made once a year in November
- Climate variables (GloSea5*)
  - Hindcast from GloSea5 (Copernicus)

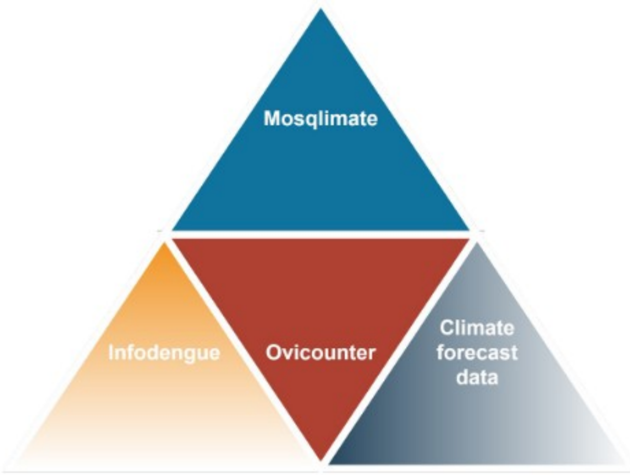


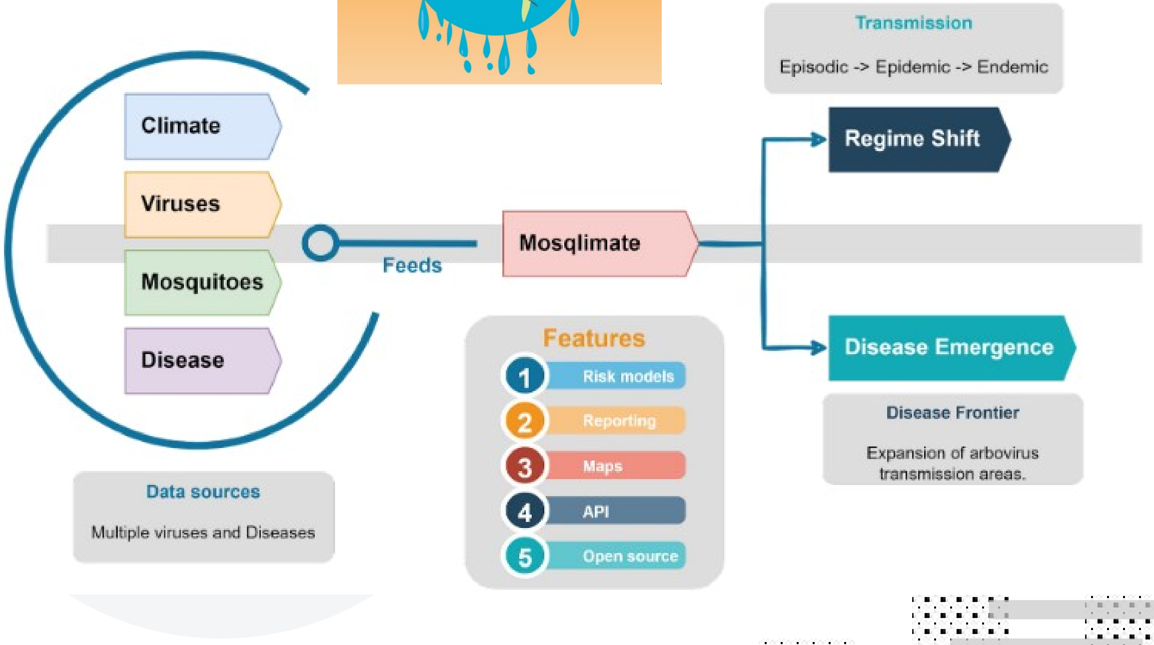


* GloSea5: Met Office Global Seasonal forecasting system version 5

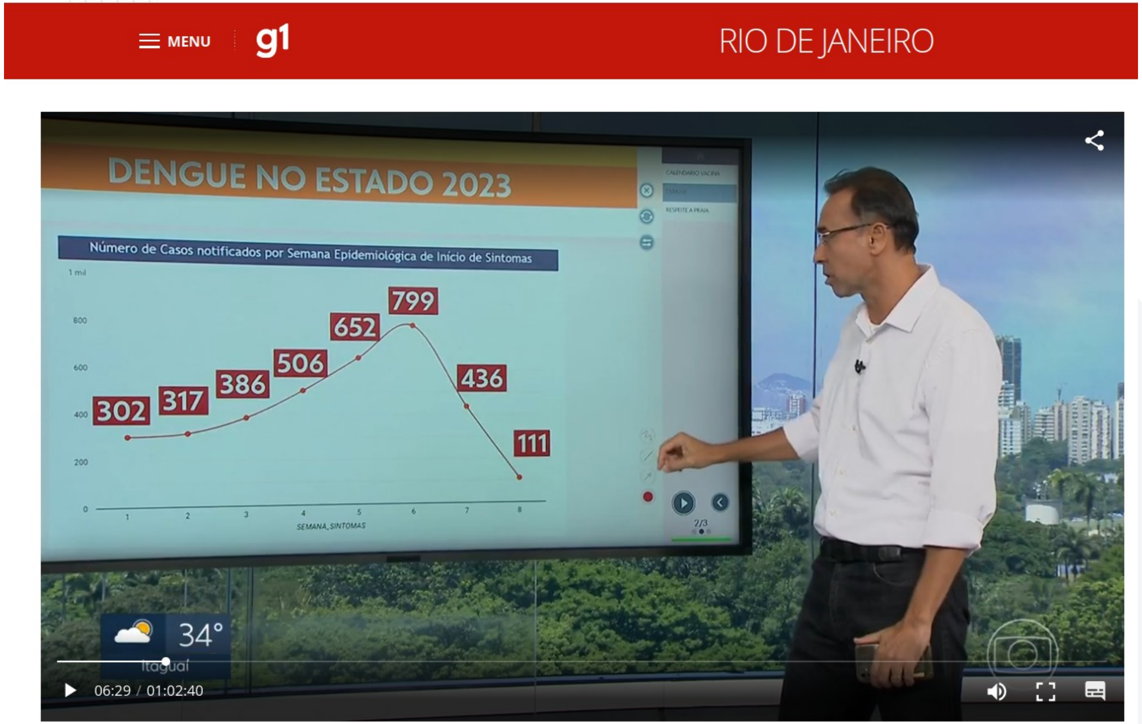
Predicting dengue in Brazil

Mosqlimate Project





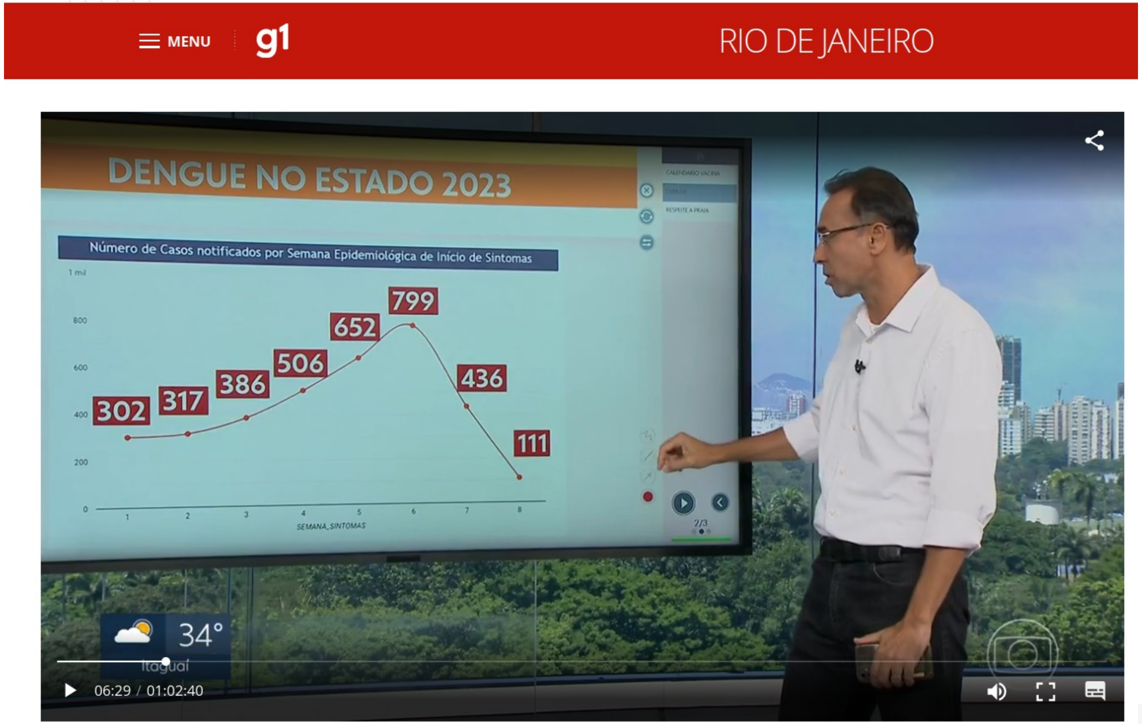
Notification delay problem

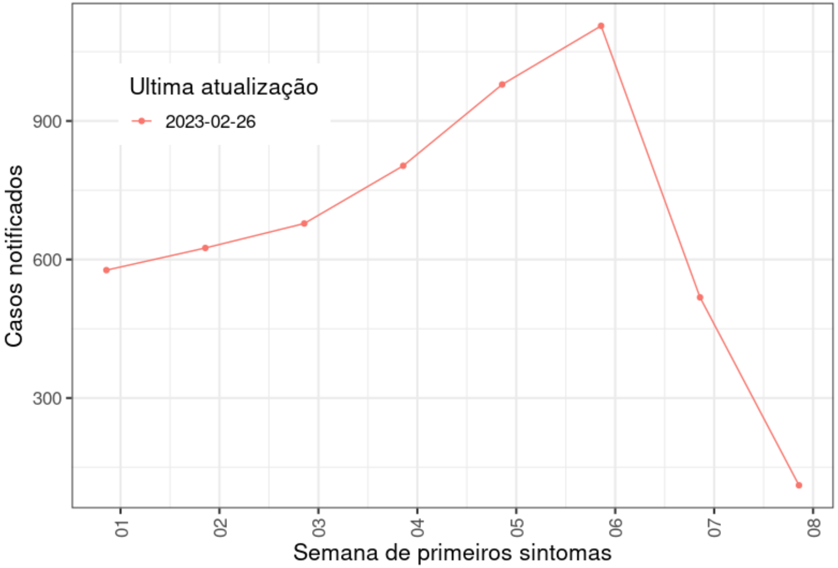


Rio, 2023-03-04

- End of week 9
- The TV presenter emphasizes that
  - the peak occur on week 6
  - “good news” cases are dropping

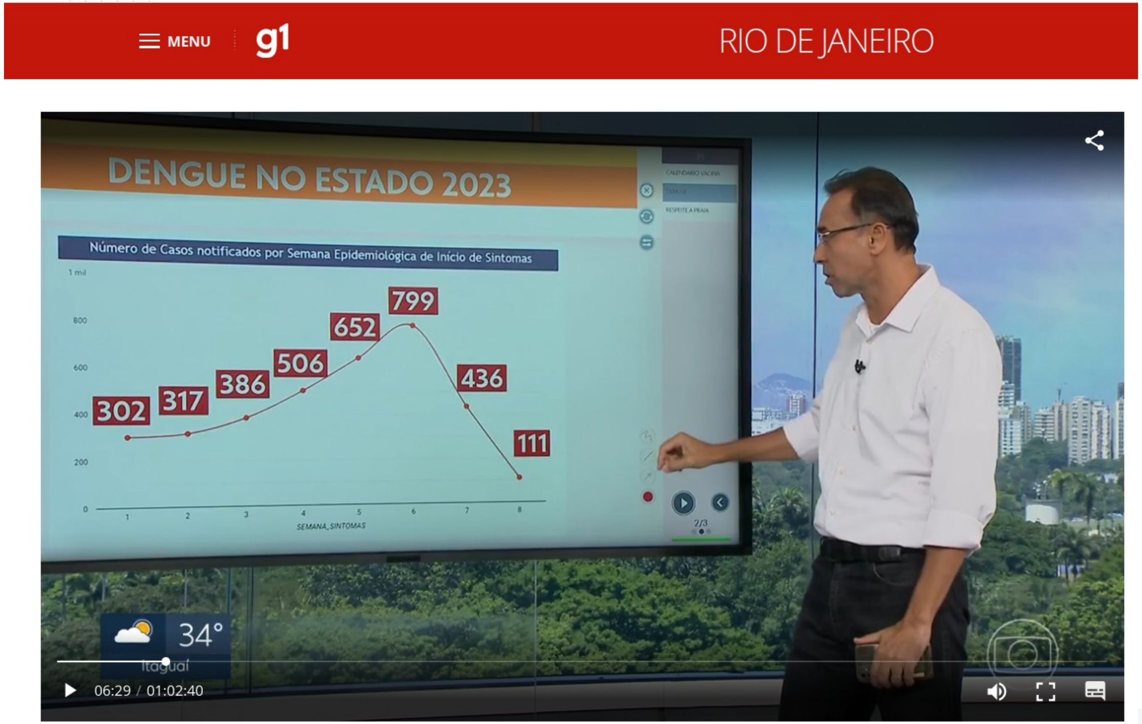
Notification delay problem



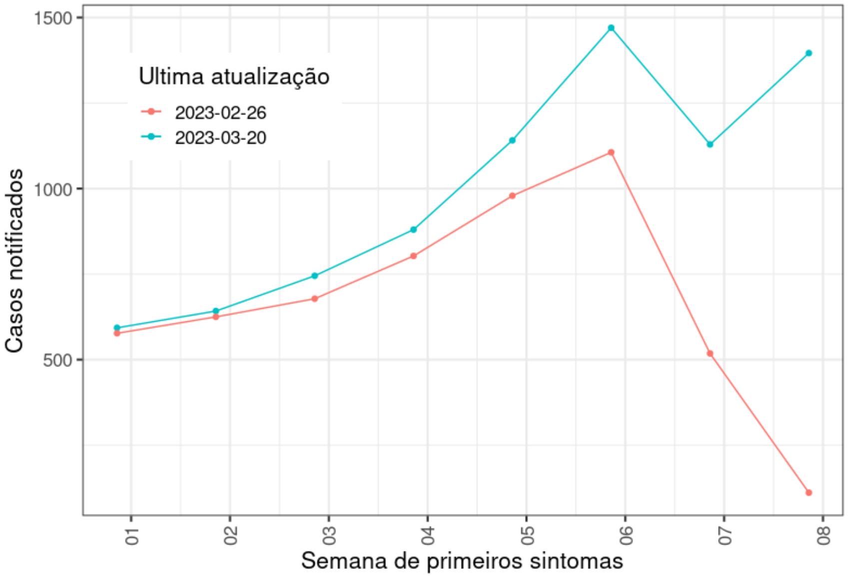


Rio, 2023-03-04

Notification delay problem

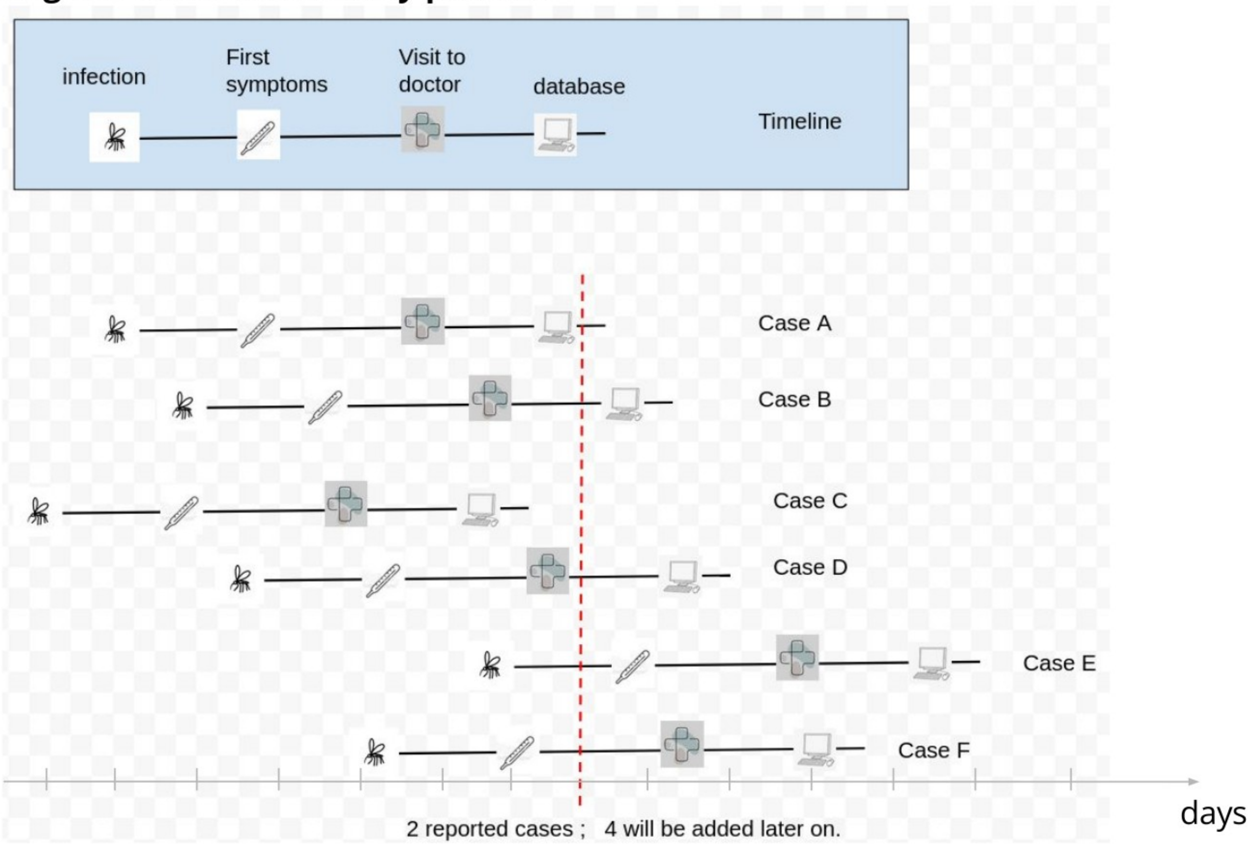


Rio, 2023-03-04



Why does the delay occur?

Modeling the observation delay process

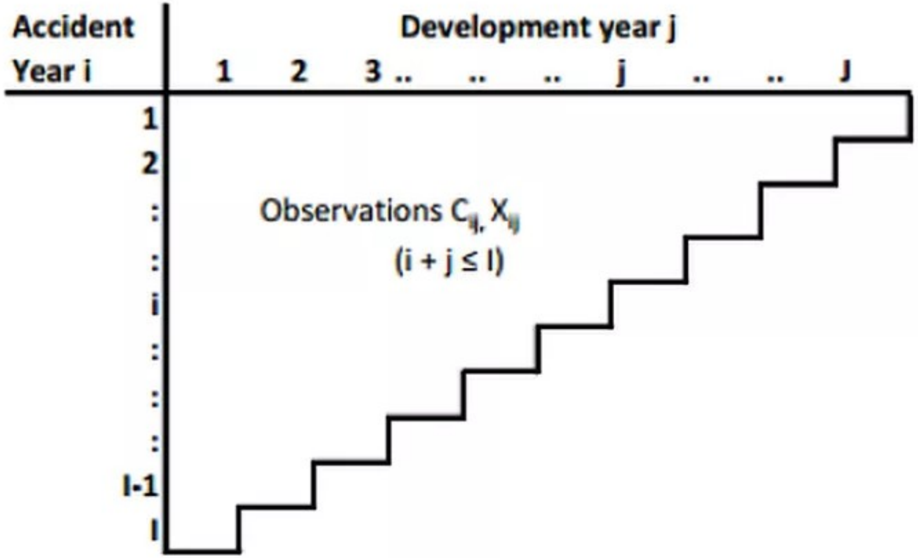


Delay correction models

- Actuarial sciences 1980/90
  - Chain ladder model

Kremer (1982), Two-way ANOVA

$$C_{ij} = E_i S_j R_{ij}$$



Renshaw (1989), Modelo linear misto
Verhall (1990), Modelo de Espaço de estados + Linear Bayes

$$Y_{ij} = \log(C_{i,j}) = \mu + \alpha_i + \beta_j + e_{ij}$$

Delay on HIV/AIDS in 1980s

- Brookmeier and Damiano (1989)
- Zeger, See and Diggle (1989)

$$Y_{t,u} \sim Poisson(\mu_{t,u})$$

$$\log(\mu_{t,u}) = s(t, \beta) + d(u, \theta)$$

- Brookmeier and Liao (1990), EUA
- Barbosa and Struchiner (1998), Brasil

MONITORING THE EPIDEMIC

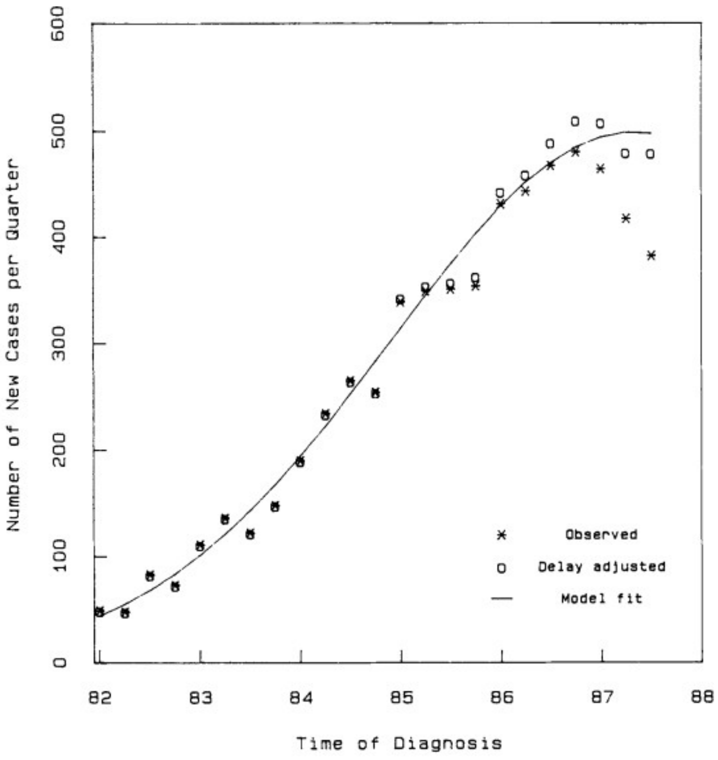


Figure 2. AIDS incidence: homosexuals in north-east

Bayesian Nowcasting

- Dengue and Influenza*

$$n_{t,d} \sim NegBin(\lambda_{t,d}, \phi)$$

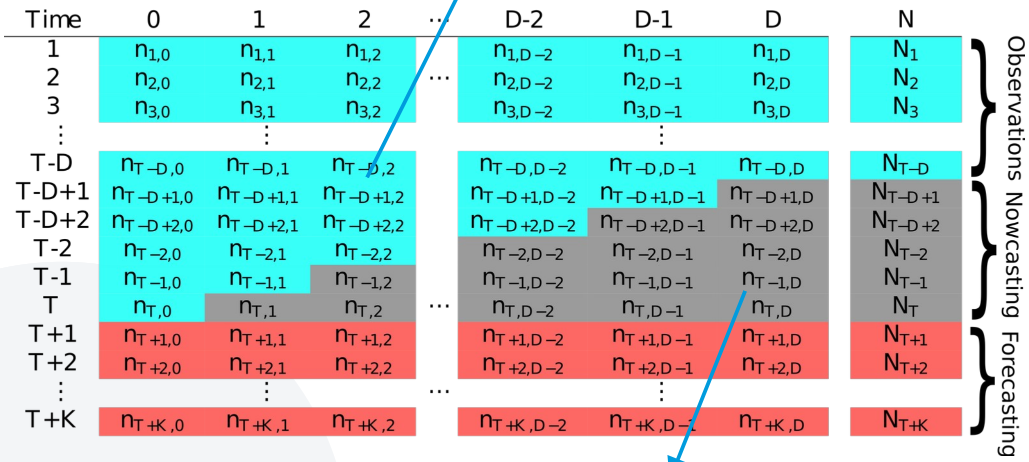
  - Bastos et al. (2019) (INLA)

$$\log(\lambda_{t,d}) = \alpha + \beta_t + \gamma_d$$

  - McGough et al. (2020) (MCMC)

$$\lambda_{t,d} = \delta_d e^{\beta_t}$$

$$\{n_{t,d}\} = \{n_{t,d} : T \leq t + d\}$$



$$\{n^*_{t,d}\} = \{n_{t,d} : T > t + d\}$$

$$P(\{n^*_{t,d}\} \mid \{n_{t,d}\}) = \int_{\theta \in \Theta} p(\theta, \{n^*_{t,d}\} \mid \{n_{t,d}\}) d\theta$$

$$= \int_{\theta \in \Theta} p(\theta \mid \{n_{t,d}\}) p(\{n^*_{t,d}\} \mid \theta) d\theta$$

Bayesian Nowcasting

- Bastos et al. (2019) foi implemented in:
  - InfoGripe
  - InfoDengue
  - R package: *nowcaster*

INFO GRIPE

INFO DENGUE

Nowcaster

`nowcaster` is a R package for "nowcasting" epidemiological time-series. Every single system of notification has an intrinsic delay, `nowcaster` can estimate how many counts of any epidemiological data of interest (*i.e.*, daily cases and deaths counts) by fitting a negative binomial model to the time steps of delay between onset date of the event, (*i.e.*, date of first symptoms for cases or date of occurrence of death) and the date of report (*i.e.*, date of notification of the case or death).

`nowcaster` is based on the `R-INLA` and `INLA` packages for "**I**ntegrated **N**ested **L**aplace **A**pproximation" algorithm to Bayesian inference. `INLA` is a fast alternative to others methods for Bayesian inference like **MCMC**. An introduction to `INLA` can be found here.

`nowcaster` is build for epidemiological emergency use, it was constructed for the Brazilian Severe Acute Respiratory Illness (SARI) surveillance database (SIVEP-Gripe).

Installing

Before installing the package certify you have an active installation of `INLA`, to do so you can run the follwing code:

```
install.packages("INLA",
                 repos=c(getOption("repos"),
                 INLA="https://inla.r-inla-download.org/R/stable"),
                 dep=TRUE)
```

Links

Browse source code

Report a bug

License

Full license

GPL (>= 3)

Citation

Citing nowcaster

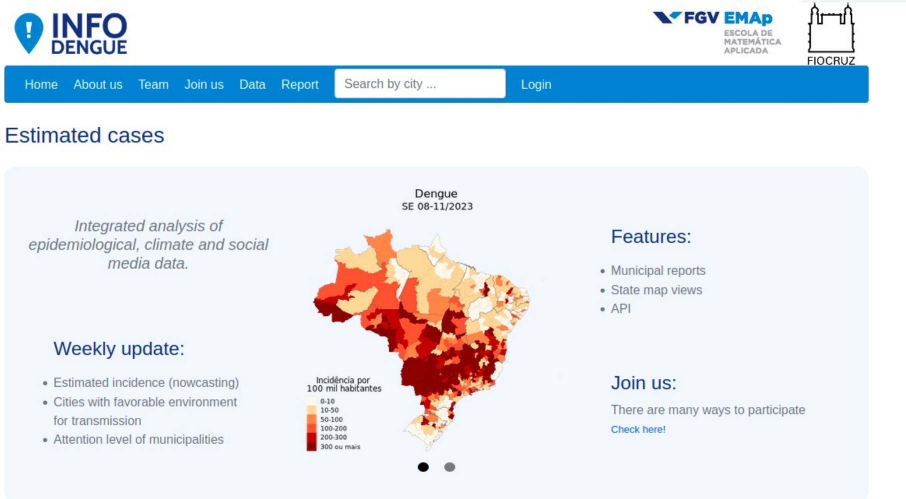
Developers

Rafael Lopes
Author, maintainer

Leonardo Bastos
Author

https://covid19br.github.io/nowcaster/

(Nearly) Real-time monitoring systems



Model performance

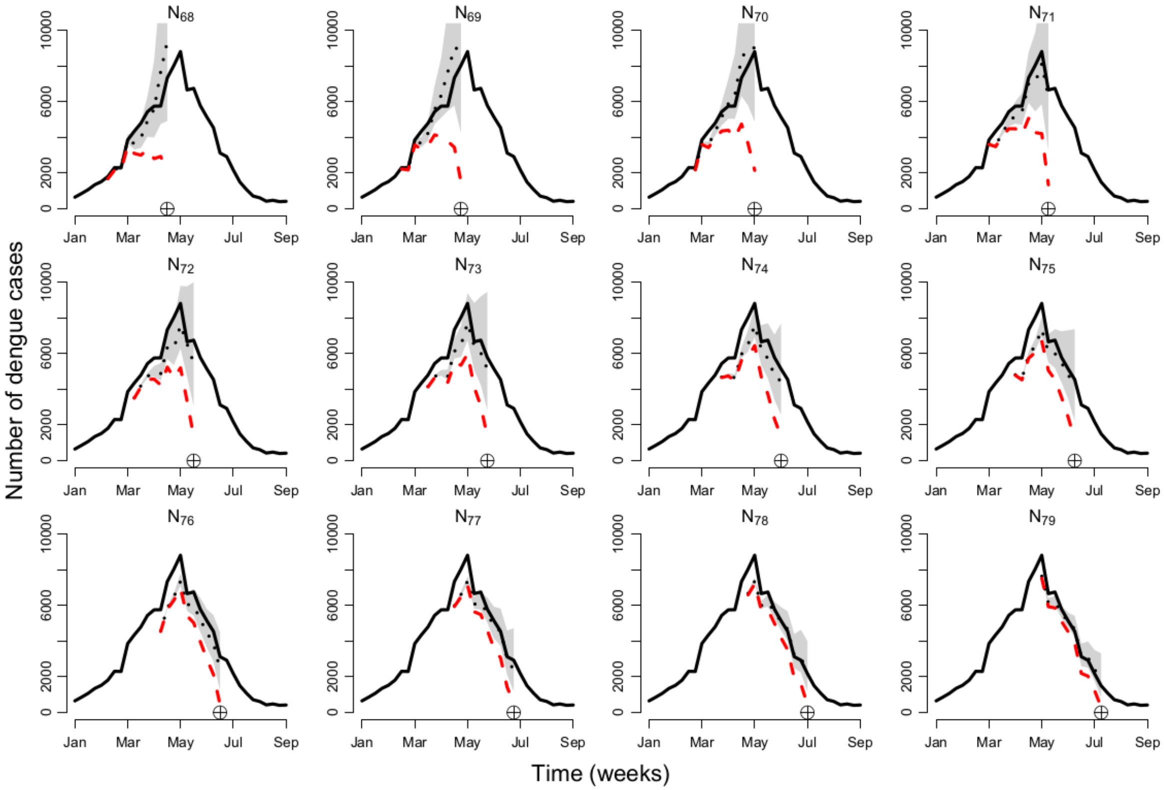


FIGURE 8 Time series of dengue cases in Rio de Janeiro for 12 epidemic weeks starting from the 15th epidemic week of 2012 on the top left ($T = 68$). The black line shows the eventually reported number of cases, the red dashed line shows the number of currently reported cases, and the black dotted line shows model predictions (of the eventually reported number of cases) along with 95% prediction intervals. The circled cross symbol indicates the epidemic week $T = 68, 69, \ldots, 79$ [Colour figure can be viewed at wileyonlinelibrary.com]

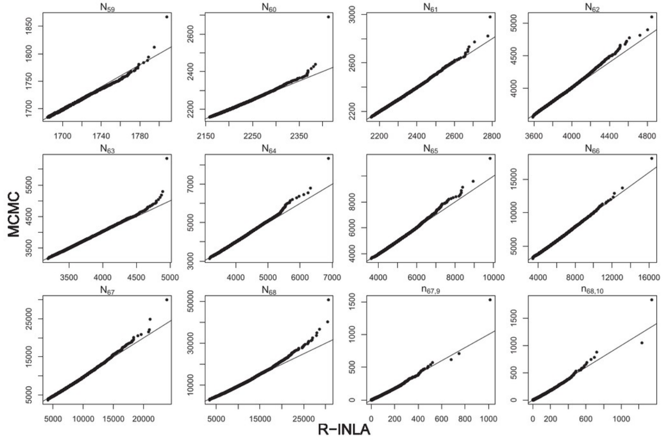
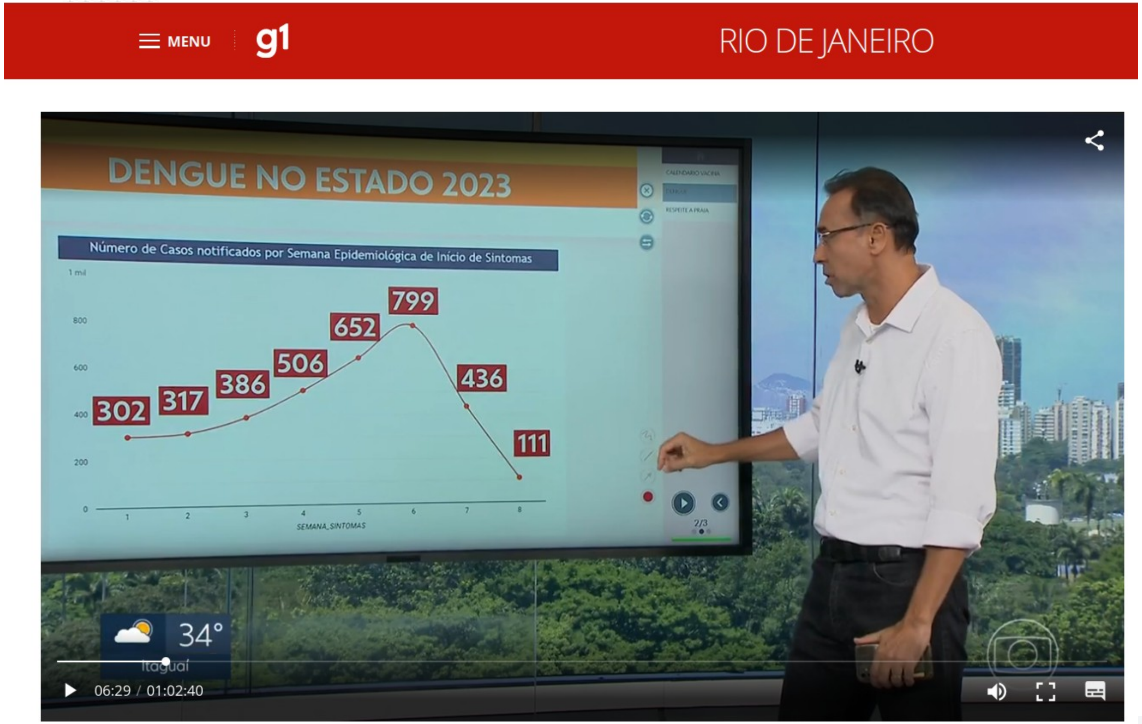


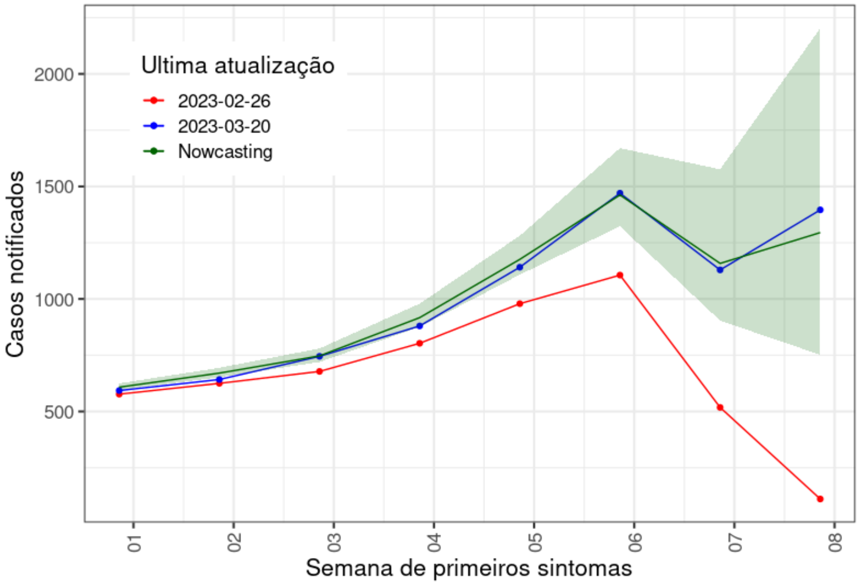
FIGURE 6 Q-Q plots comparing R-INLA and MCMC samples from the predictive distribution of the total counts N_t for $t = 59, \ldots, 68$, where $T = 68$ is the 15th epidemic week of 2012. INLA, integrated nested Laplace approximation; MCMC, Markov chain Monte Carlo

INLA versus MCMC

Notification delay problem

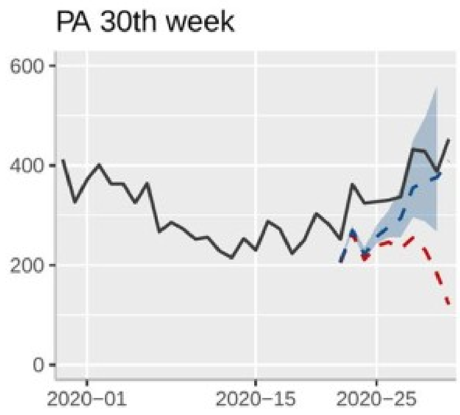


Rio, 2023-03-04



Notification delay (Malaria in the Amazon)

- Malaria is a mosquito-borne disease endemic in the Amazon region (99% of the notified cases)
- There is notfication delay, then the nowcasting is needed!



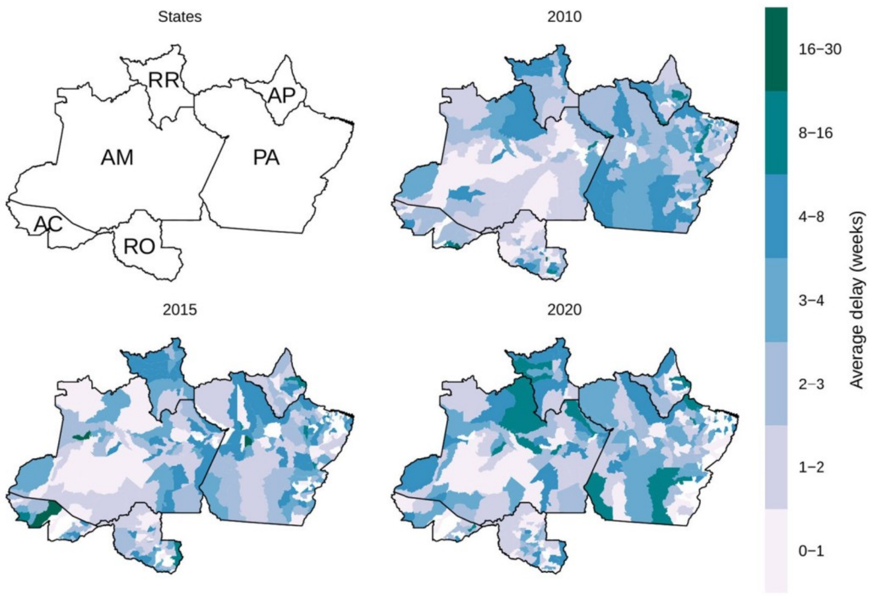


Fig. 2 State units of the Brazilian Legal Amazon and average delays in weeks per municipality in years 2010, 2015 and 2020

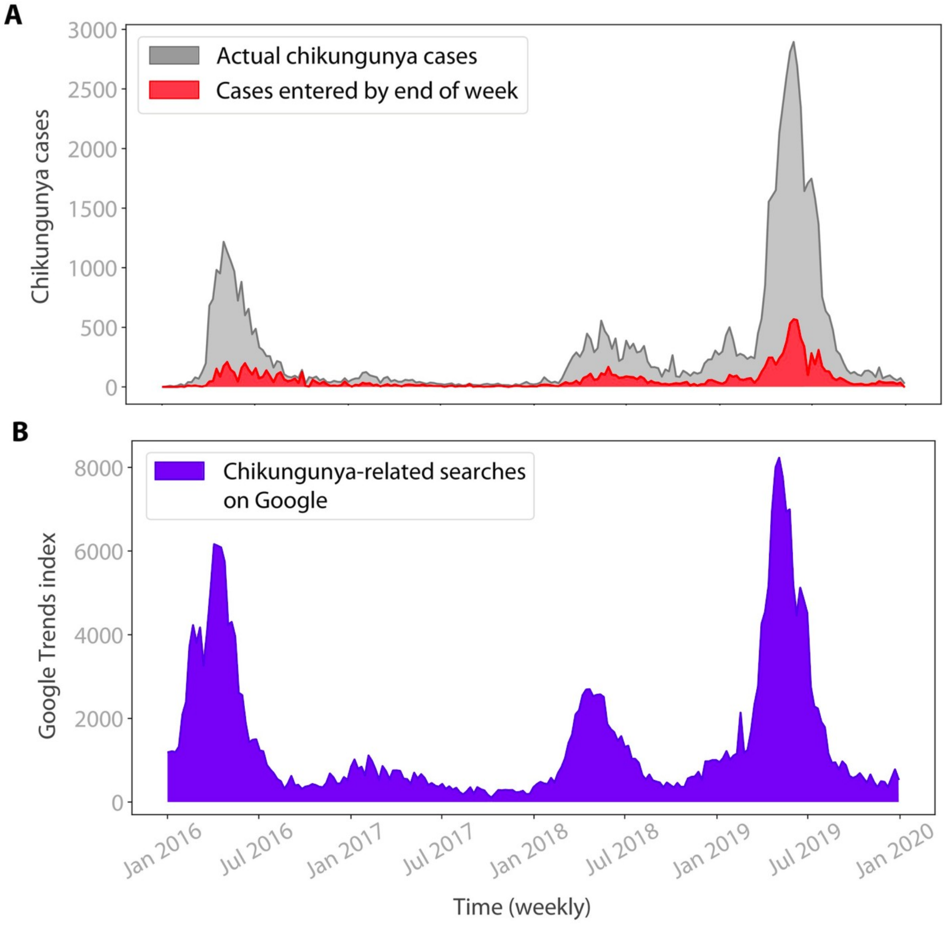
Ayala et al. (2023, Malaria Journal)

Improving the delay correction

- Covariates may improve the corrections

$$n_{t,d} \sim NegBin(\lambda_{t,d}, \phi)$$

$$\log(\lambda_{t,d}) = \alpha + \beta_t + \gamma_d + \mathbf{x}_{t,d}^T \delta$$



Bastos et al. (2019)

Mizzi (2019, PhD Warwick), Miller et al. (2022)

Spatial nowcasting

- Disease dynamics may vary on space
  - Population immunity, different ways to fight a possible epidemic event, vector adaptation, etc
- Delay may also vary in space
  - Structural problems, lack of training etc

- In Bastos et al. (2019), we added a spatial structure

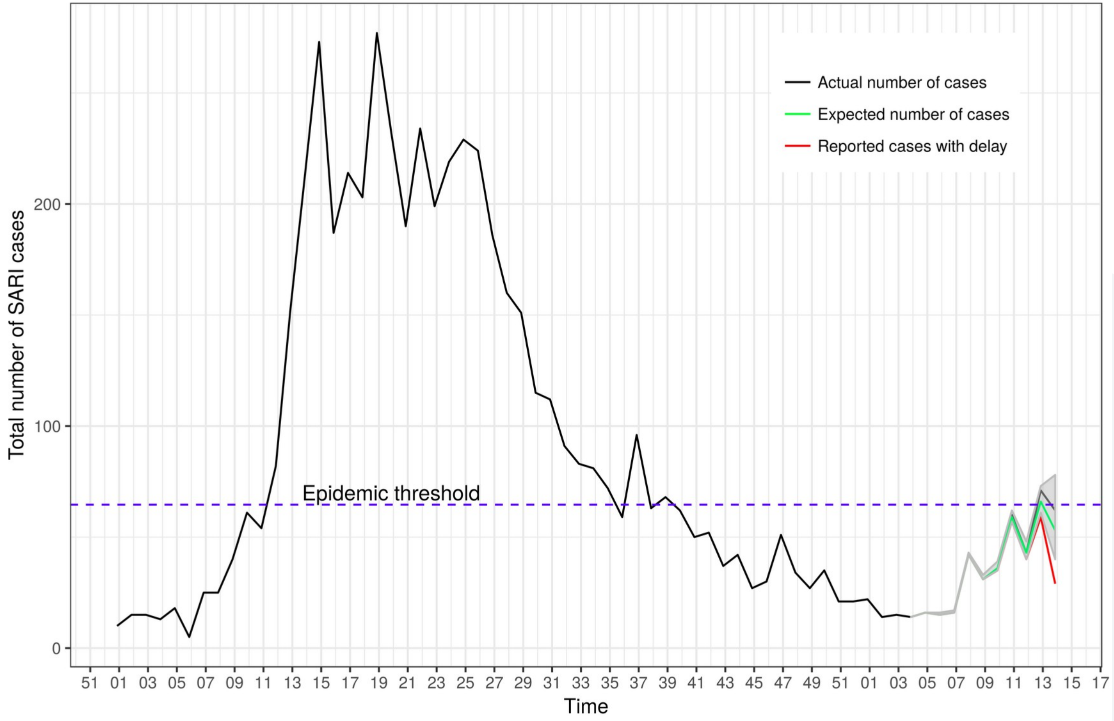
$$n_{t,d,r} \sim NegBin(\lambda_{t,d,r}, \phi)$$

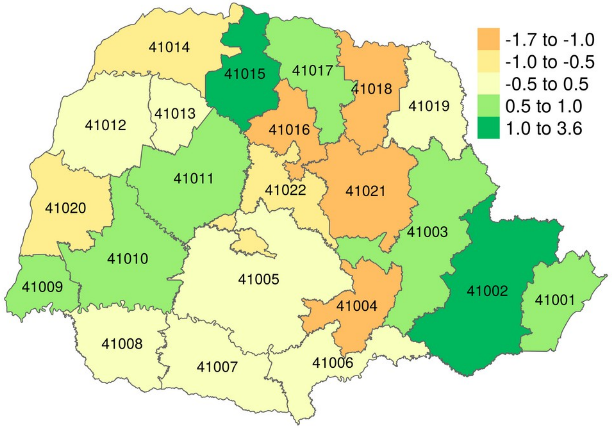
$$\log(\lambda_{t,d,r}) = \alpha_r + \beta_{t,r} + \gamma_{d,r}$$

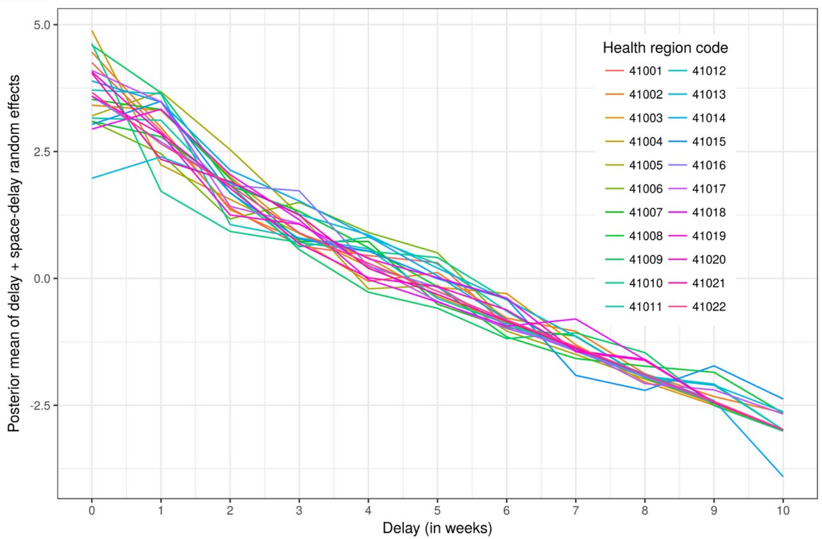
$$\Sigma_{\beta_{t,r}} = \Sigma_t \otimes \Sigma_r$$
$$\Sigma_{\gamma_{d,r}} = \Sigma_d \otimes \Sigma_r$$

Spatial random effects is modelled using CAR

SARI in PR 2017





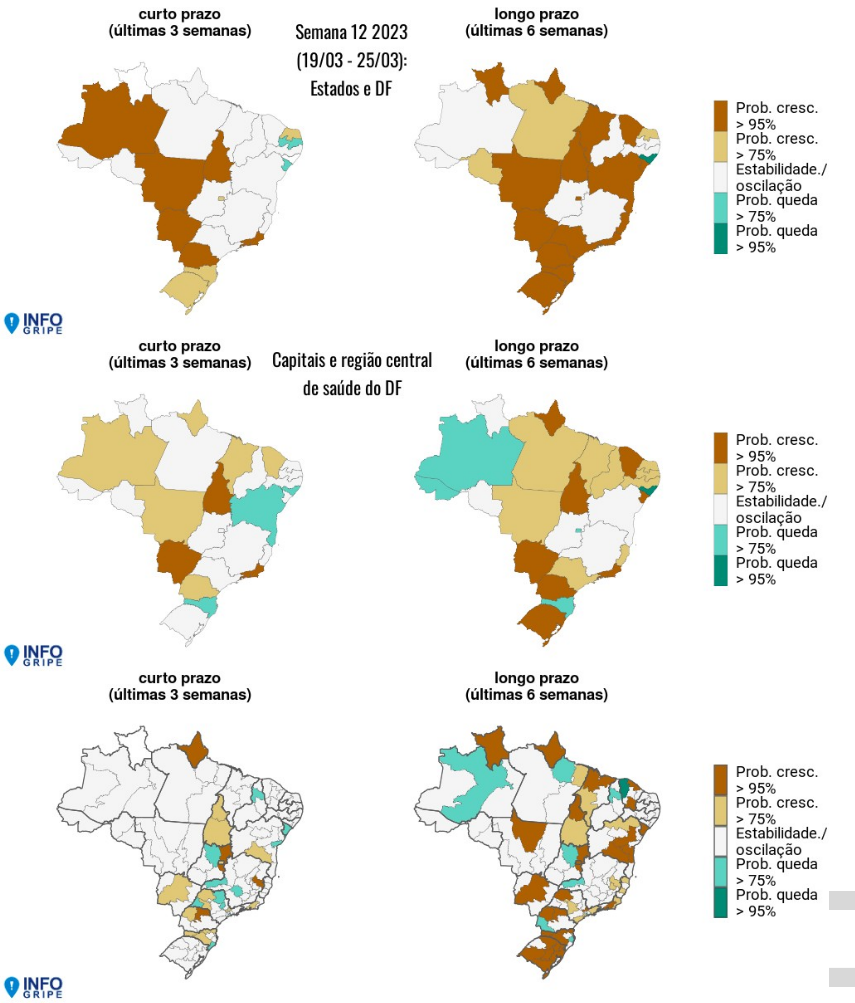


SARI during the pandemic

- We needed a simple and fast measure to show trend
- A linear model coefficient applied on each sample sample trajectory
- Monte Carlo probabilities can de calculated

$$N_t = A + Bt + e_t, \quad t = 1, 2, .., W$$

$$P(B > 1) \qquad P(B < 1)$$



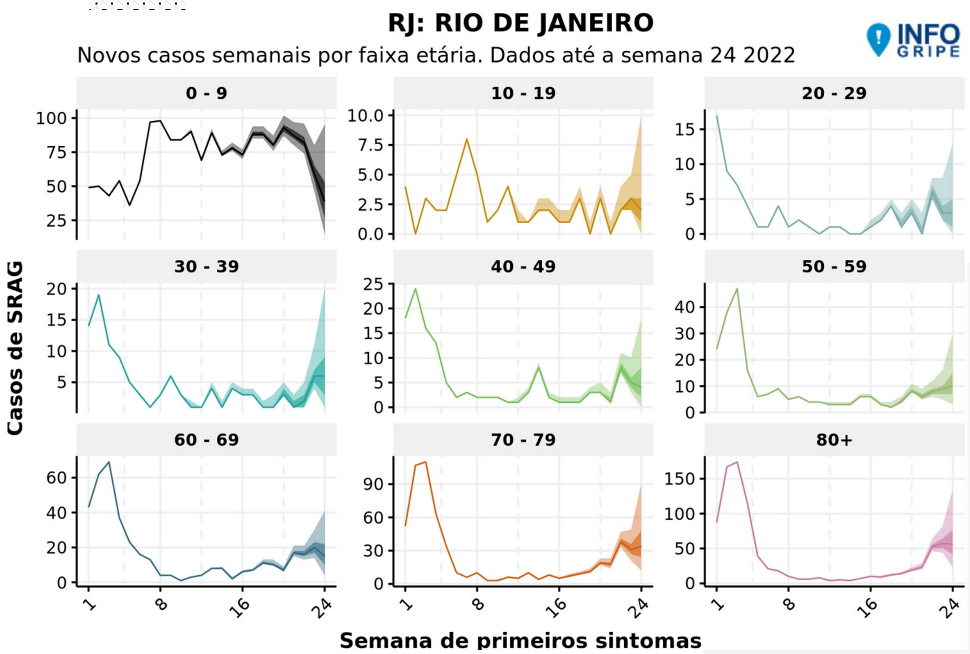
Age-dependent nowcasting

- The dynamic of COVID-19 hospitalizations varies on age groups
- Hence, the nowcasting model should consider age

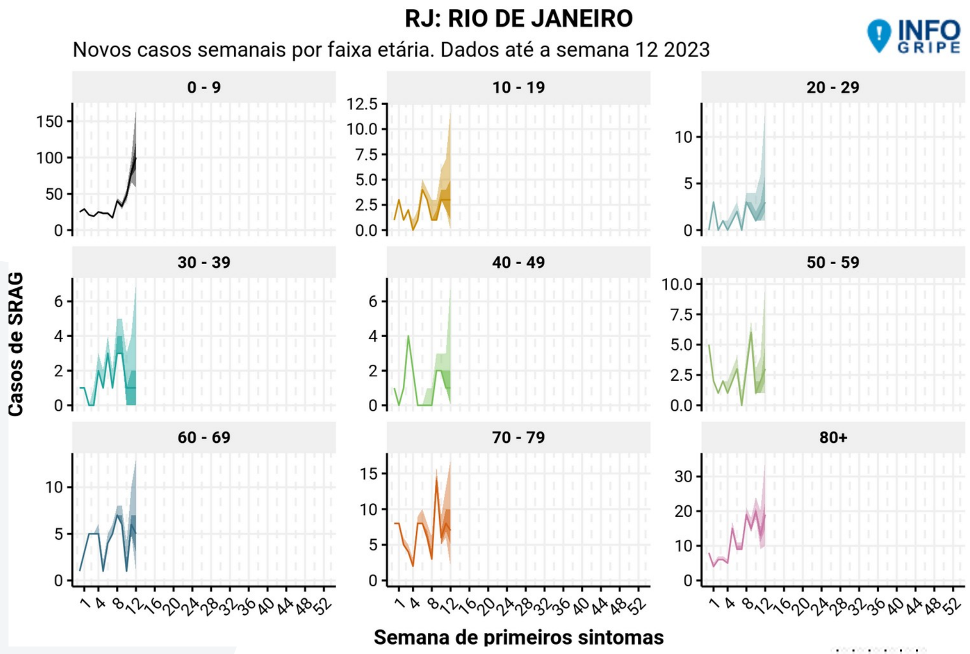
$$n_{t,d,g} \sim NegBin(\lambda_{t,d,g}, \phi)$$

$$\log(\lambda_{t,d,g}) = \alpha_r + \beta_{t,g} + \gamma_{d,g}$$

Age-dependent nowcasting



SARI due to COVID-19



SARI due to RSV

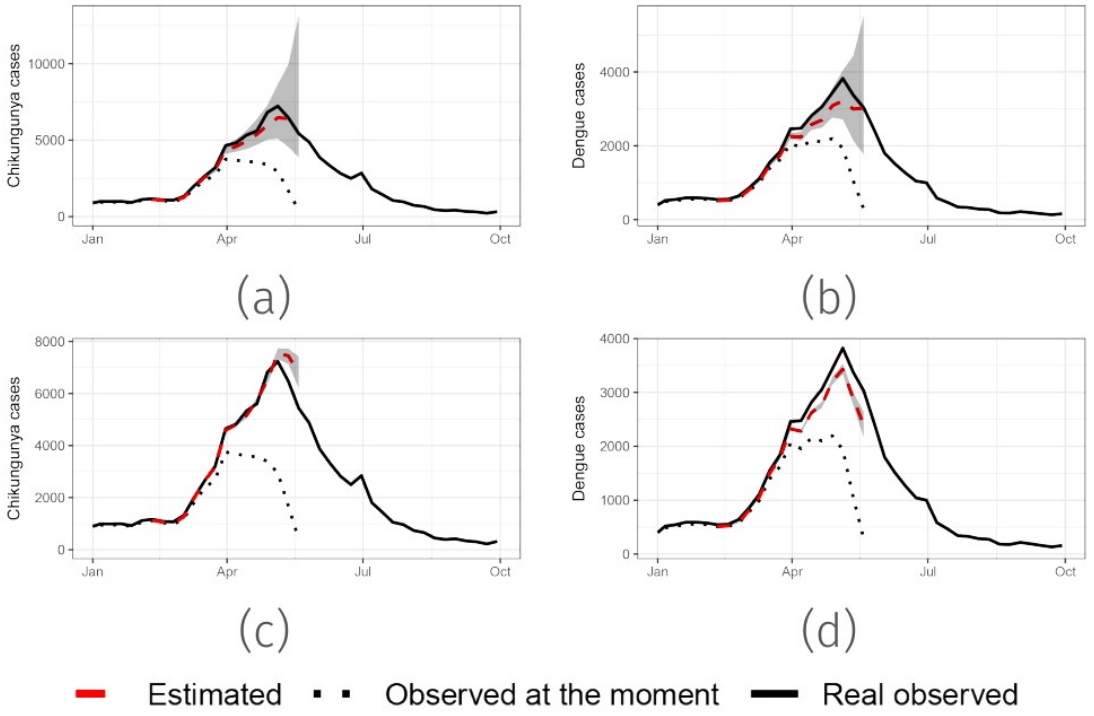
Multivariate nowcasting

Dengue and Chikungunya joint model

Shared component model

$$n_{t,d}^{(i)} \sim NegBin(\lambda_{t,d}^{(i)}, \phi^{(i)})$$

$$\log(\lambda_{t,d}^{(i)}) = \alpha + \beta_t + \gamma_d + \beta_t^{(i)} + \gamma_d^{(i)}$$



Dos Santos (2023, Msc, UFRJ)

Thank you!

Leo Bastos
PROCC/Fiocruz

Leonardo.bastos@fiocruz.br

@leosbastos