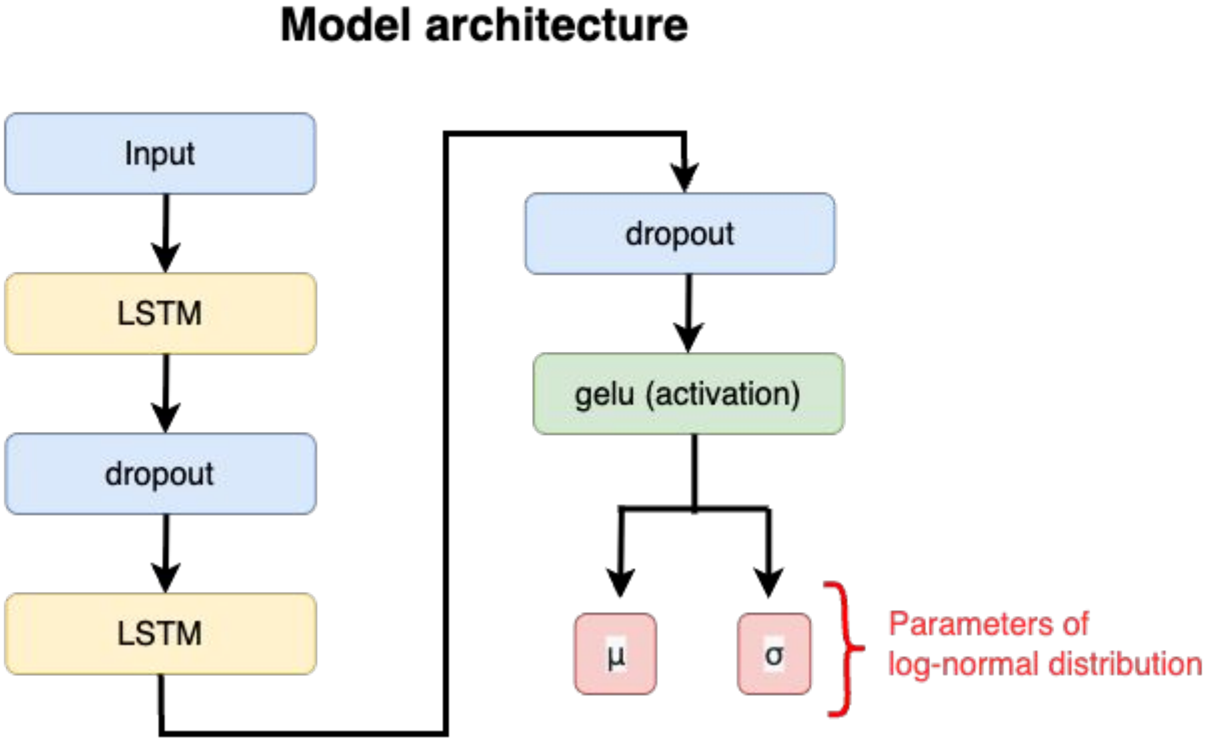
Dengue oracle (Model ID 155 e 156)



Dengue oracle (Model ID 155 e 156)

Model 1 (M1) - base:

Temporal predictors:

- cases;
- epiweeks.

Static predictors:

- pop_norm.

Model 2 (M2) - with covariates:

Temporal predictors:

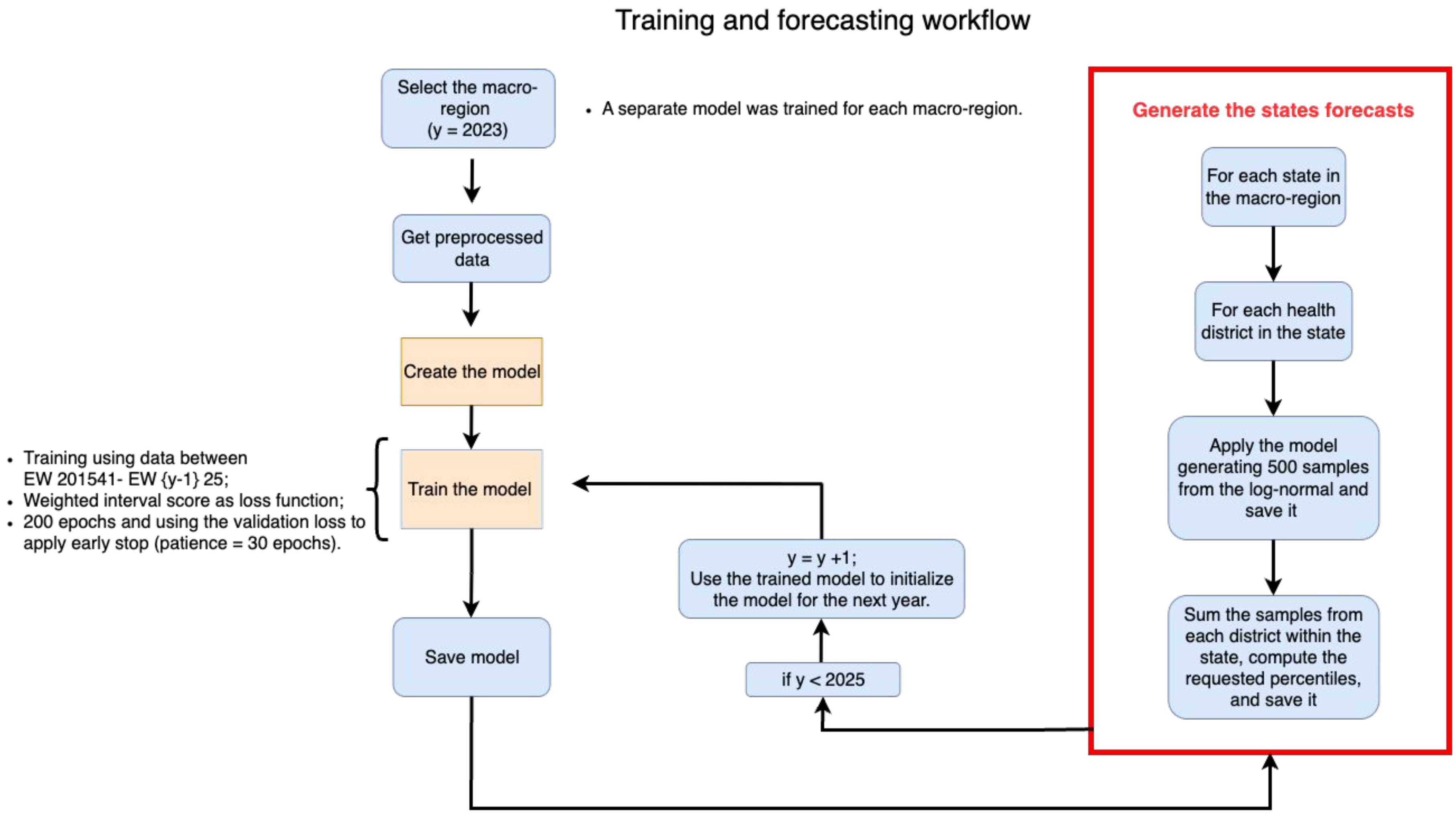
- cases;
- epiweeks;
- enso value.

Static predictors:

- pop_norm;
- biome predominant.

Independent of the set of features the train and test samples are generated using the data (after normalization by max values) from **all the regional health of the macro-region.**

The static features were encoded as time series with constant values.


Training and forecasting workflow
Select the macro-region (y = 2023)
• A separate model was trained for each macro-region.
Get preprocessed data
Create the model
• Training using data between EW 201541- EW {y-1} 25;
• Weighted interval score as loss function;
• 200 epochs and using the validation loss to apply early stop (patience = 30 epochs).
Train the model
Save model
y = y +1; Use the trained model to initialize the model for the next year.
if y < 2025
Generate the states forecasts
For each state in the macro-region
For each health district in the state
Apply the model generating 500 samples from the log-normal and save it
Sum the samples from each district within the state, compute the requested percentiles, and save it

Team: JBD - Mosqlimate

Members:

- Beatriz Laiate, Ph.D.
- Davi Sales Barreira, Ph.D.
- Julie Souza, Ph.D.

Institution: School of Applied Mathematics - Fundação Getúlio Vargas (FGV/EMAp)

Forecasting Methodology

Data: Dengue cases by municipalities and week, ENSO index, and Brazilian states coordinates.

Model: Chronos, a language model framework for time series forecasting. The model is pre-trained on a large dataset of time series data, and was fine-tuned on the dengue cases data.

Tools: ‘Autogluon‘, a library that provides a high-level interface for training and evaluating time-series models.

Post-Processing:

- Sort quantiles to ensure monotonicity.
- Set negative values to zero.
- Interpolate and extrapolate quantiles to obtain the required intervals.

References

Ansari, Abdul Fatir; Stella, Lorenzo; Turkmen, Caner; Zhang, Xiyuan; Mercado, Pedro; Shen, Huibin; Shchur, Oleksandr; Rangapuram, Syama Sundar; Arango, Sebastian Pineda; Kapoor, Shubham; et al.
Chronos: Learning the language of time series.
arXiv preprint arXiv:2403.07815, 2024.

Shchur, Oleksandr; Turkmen, Ali Caner; Erickson, Nick; Shen, Huibin; Shirkov, Alexander; Hu, Tony; Wang, Bernie.
AutoGluon–TimeSeries: AutoML for probabilistic time series forecasting.
International Conference on Automated Machine Learning, PMLR, pp. 9–1, 2023.

Hierarchical Bayesian mixed model with covariate interactions

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

t	=	temporal index
a(t)	=	annual index
w(t)	=	week index
s	=	spatial index
U(s)	=	state index
A	=	absolute temperature
T	=	temperature anomaly
L	=	long-lag drought index
S	=	short-lag drought index
N	=	niño index
C	=	pre-2018 index (binary)

$$y_{s,t} \mid \mu_{s,t}, K \sim \text{NegBin}(\mu_{s,t}, K)$$

$$\log(\mu_{s,t}) = \log(\rho_{s,t}) + \log(p_{s,a(t)})$$

population offset

$$\underbrace{\log(\rho_{s,t})}_{\text{dengue incidence rate}} = \alpha + \underbrace{(\beta_T X_T + \beta_L X_L + \beta_S X_S + \beta_{T,L} X_T X_L + \beta_{T,S} X_T X_S + \beta_{L,S} X_L X_S + \beta_{T,L,S} X_T X_L X_S + \beta_A X_A)_K + \beta_N X_N + \beta_{C,U(s)} X_{C,U(s)}}_{\text{climate and oceanic covariates}} + \underbrace{\delta_{w(t),U(s)} + \gamma_{a(t)} + u + v}_{\text{spatio-temporal random effects}}$$

dengue incidence rate

climate and oceanic covariates

spatio-temporal random effects

Source: Fletcher & Mila et al. (in prep)

Hierarchical Bayesian mixed model with covariate interactions

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

t	=	temporal index
a(t)	=	annual index
w(t)	=	week index
s	=	spatial index
U(s)	=	state index
A	=	absolute temperature
T	=	temperature anomaly
L	=	long-lag drought index
S	=	short-lag drought index
N	=	niño index
C	=	pre-2018 index (binary)

$$y_{s,t} \mid \mu_{s,t}, K \sim \mathrm{NegBin}(\mu_{s,t}, K)$$

$$\log(\mu_{s,t}) = \log(\rho_{s,t}) + \log(p_{s,a(t)})$$

population offset

$$\log(\rho_{s,t}) = \alpha + (\beta_T X_T + \beta_L X_L + \beta_S X_S + \beta_{T,L} X_T X_L + \beta_{T,S} X_T X_S + \beta_{L,S} X_L X_S + \beta_{T,L,S} X_T X_L X_S + \beta_A X_A)_K + \beta_N X_N + \beta_{C,U(s)} X_{C,U(s)} + \delta_{w(t),U(s)} + + u + v$$

dengue incidence rate

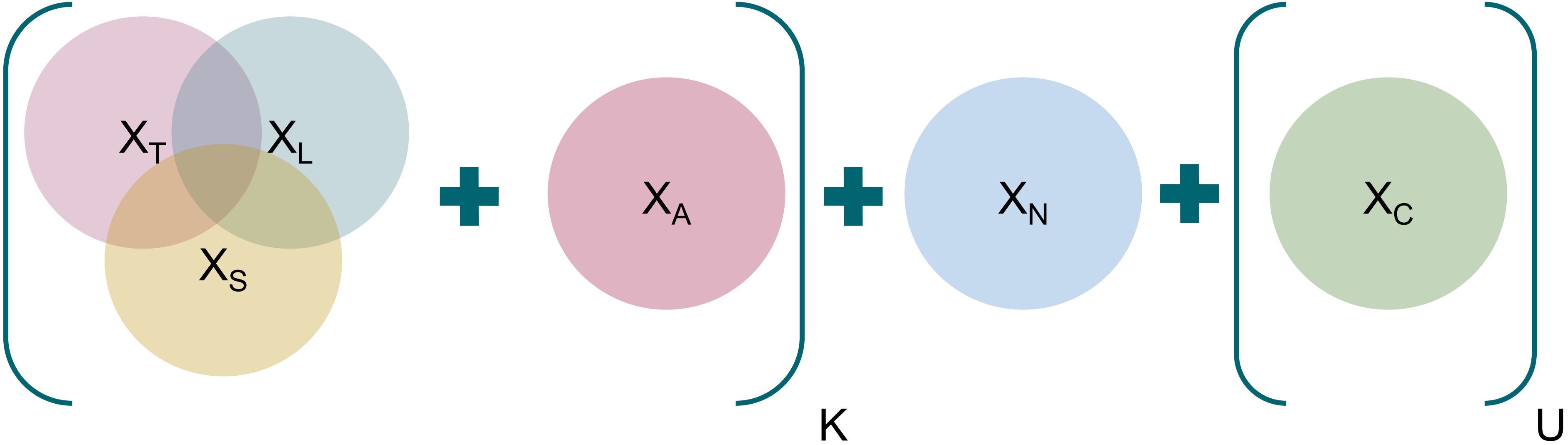
climate and oceanic covariates

spatio-temporal random effects

Source: Fletcher & Mila et al. (in prep)

Best model for predicting the upcoming dengue season in Brazil

Barcelona Supercomputing Center
Centro Nacional de Supercomputación



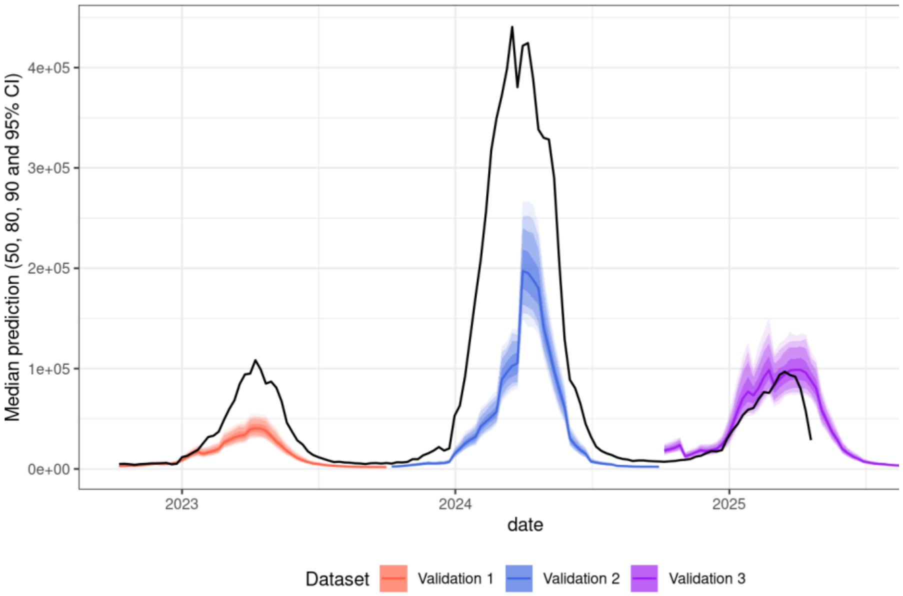
T	=	6-month mean temp anomaly	(lagged 1 month)
L	=	12-month SPEI	(lagged 3 months)
S	=	3-month SPEI	(lagged 1 month)
A	=	6-month mean abs temp	(lagged 1 month)
N	=	3-month Niño 3.4 index (ONI)	(lagged 7 months)
C	=	Pre-2018 binary index	-
K	=	Köppen classification	-
U	=	State	-

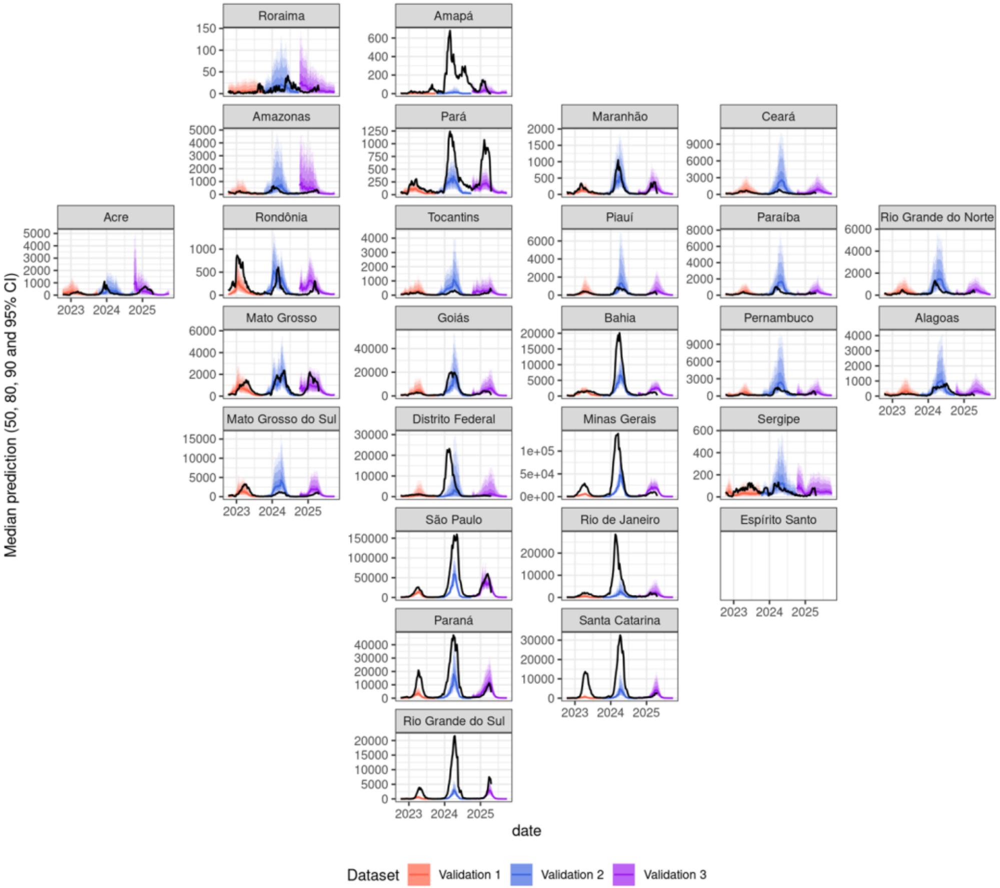
+ $(\text{Weekly RE})_U$ and Spatial RE

Source: Fletcher & Mila et al. (in prep)

Weekly national and state-level validation results for 3 dengue seasons

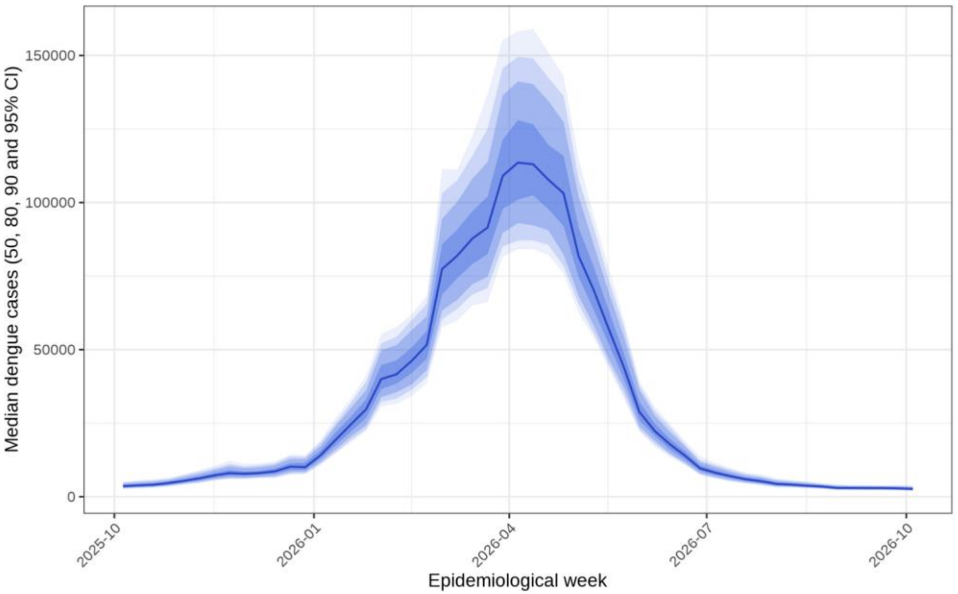
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

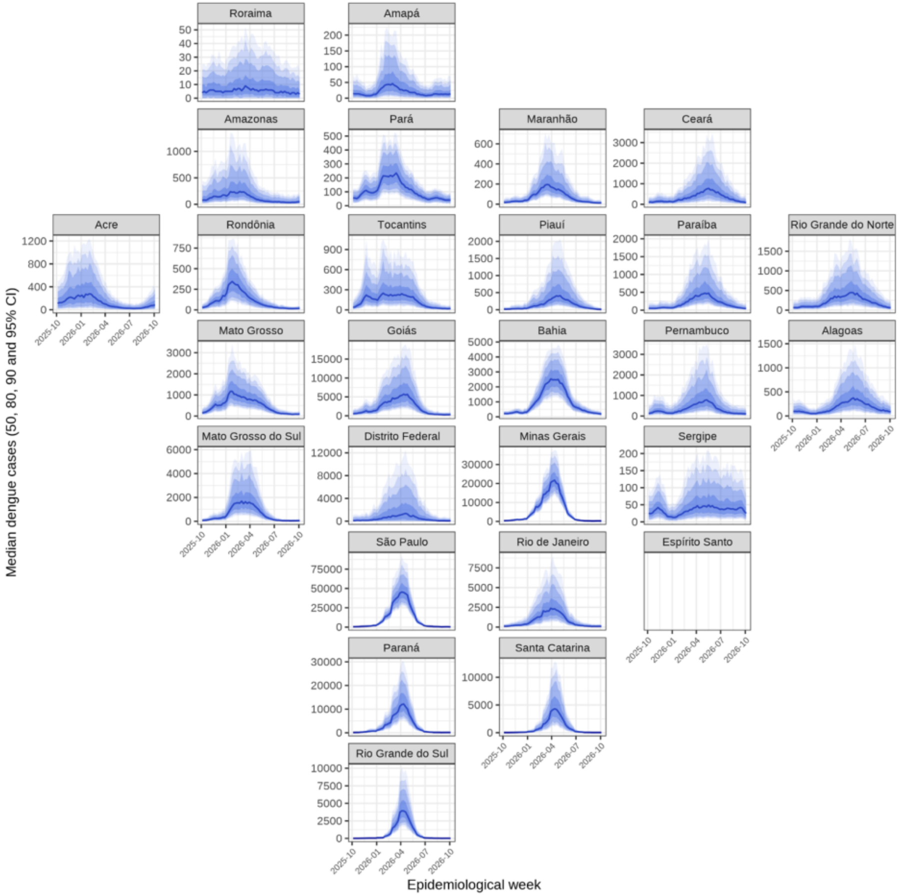




Source: Fletcher & Mila et al. (in prep)

Weekly national and state-level dengue forecasts for 2025/26 season





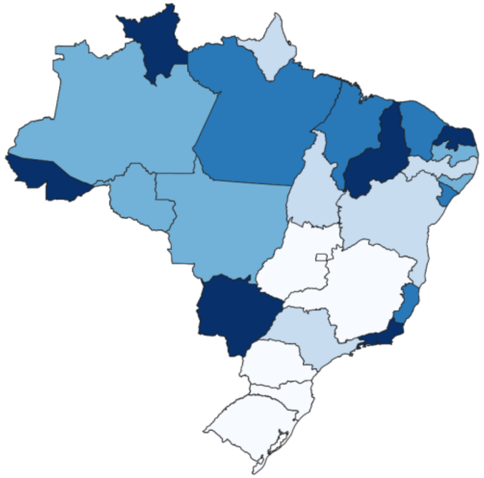
Source: Fletcher & Mila et al. (in prep)

Data

1 Target: Weekly number of dengue cases per state in Brazil

2 Time-varying Covariates

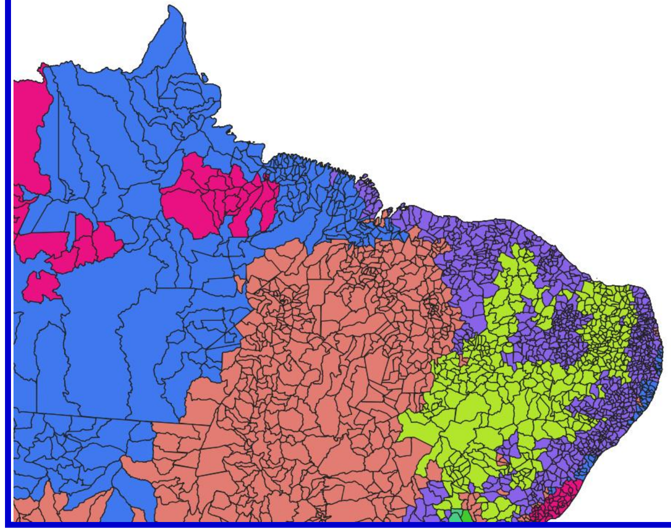
Population-weighted aggregation of municipality-level climate values to state-level (min, med, max)



- Temperature
- Precipitation
- Relative Humidity
- Pressure

3 Static Covariates

Used **proportion of state population** living in each **Köppen** climate class and **Brazilian biome** (aggregating from municipalities)

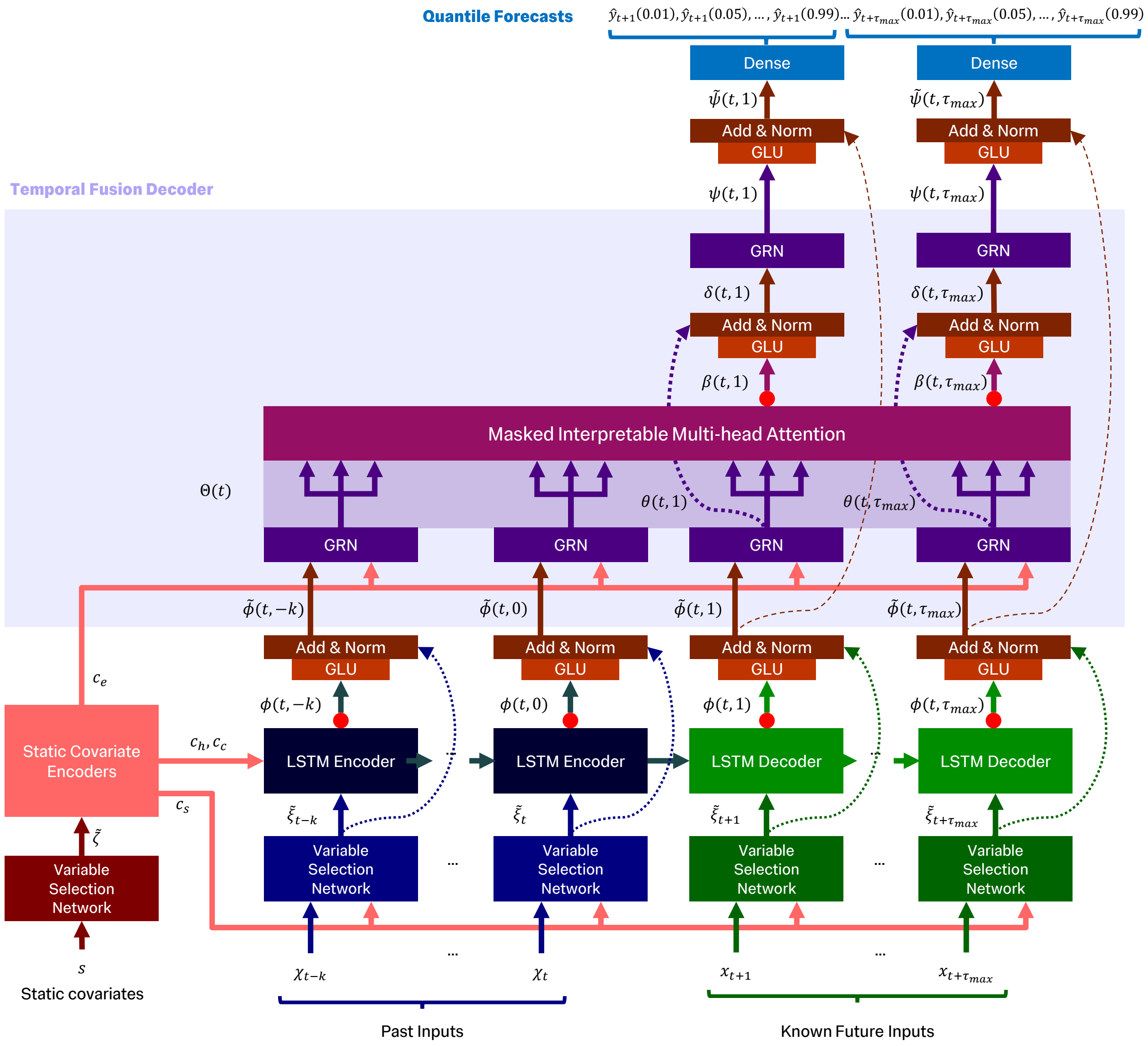


Model

Temporal fusion transformer (TFT): deep-learning-based forecasting method

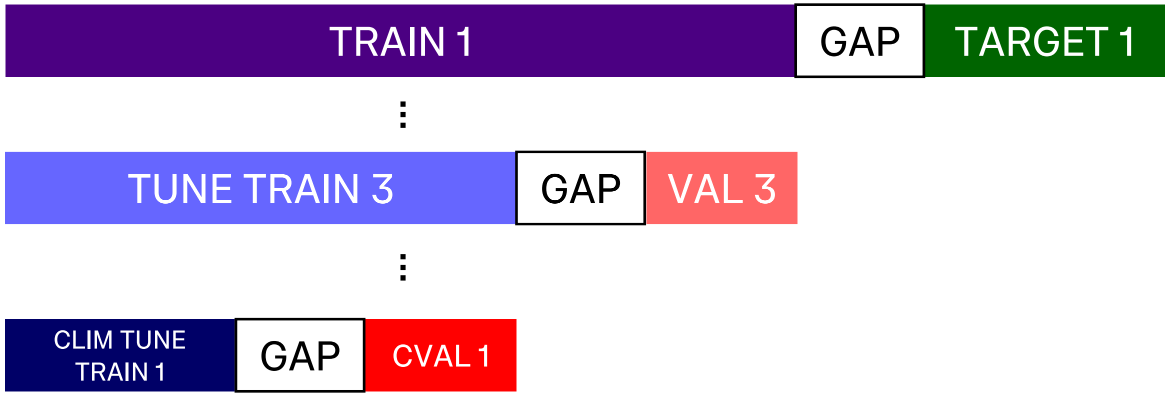
Main features:

1. **Automated variable selection** and importance weighting
2. **Scalable complexity** through use of skip connections
3. **Flexibility** in use of covariates (static, past, or future covariates all allowed)
4. Multi-head **attention** for learning longer range temporal patterns



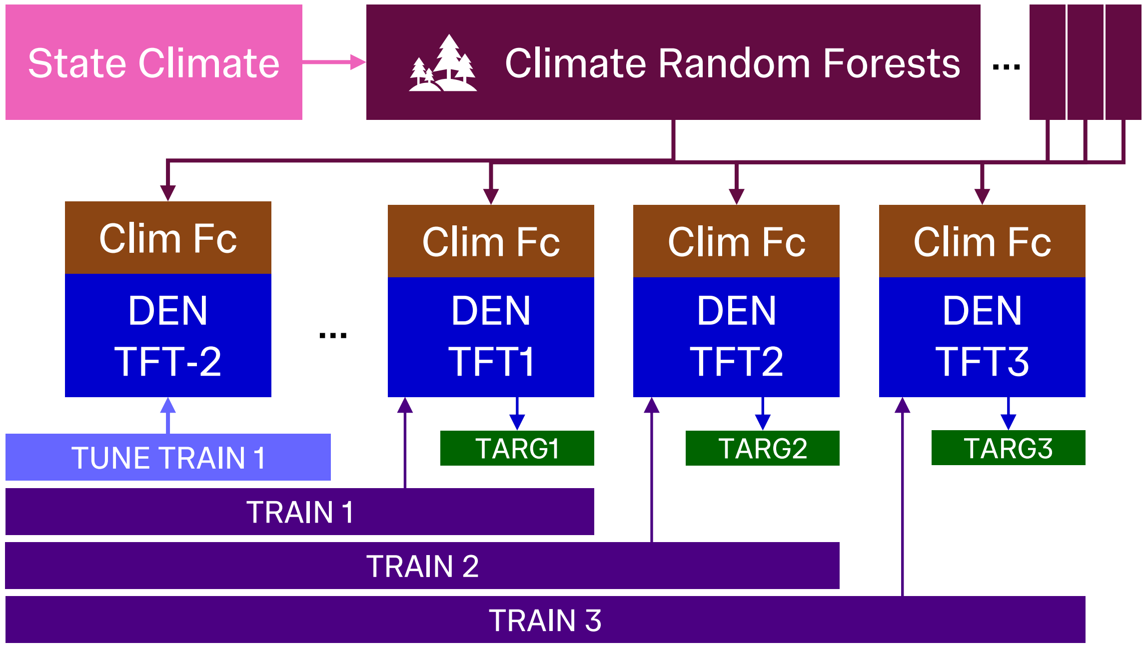
1. Lim B, Arık SÖ, Loeff N, Pfister T. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *Int J Forecast* 2021; **37**(4): 1748-64.
2. Herzen J, Lässig F, Piazzetta SG, et al. Darts: user-friendly modern machine learning for time series. *J Mach Learn Res* 2022; **23**(1).

Pipeline



- Initial training set split into **multiple train-validation folds**
- Hyperparameters tuned using **Optuna** optimising on **mean quantile loss (MQL)** across three validation sets

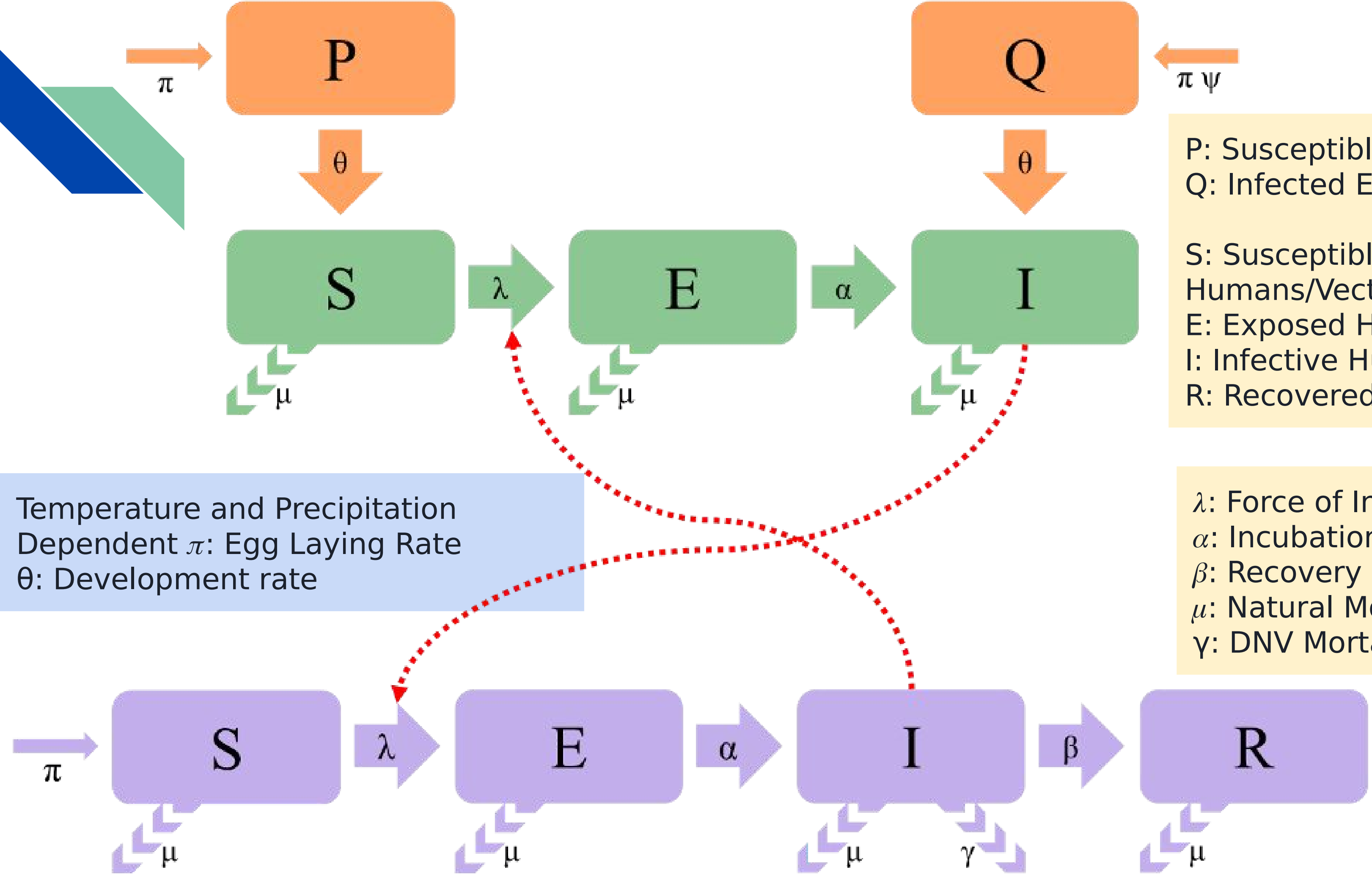
Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York, NY, USA: Association for Computing Machinery: 2623–31.



- Random forests trained to generate **climate forecasts** used in dengue model as future covariates
- **Incremental fine-tuning** of previous models as new data “arrives” then forecasts are generated

ISI Dengue Model

ISI Foundation, Turin (Italy)
Davide Nicola



Temperature and Precipitation
Dependent π: Egg Laying Rate
θ: Development rate

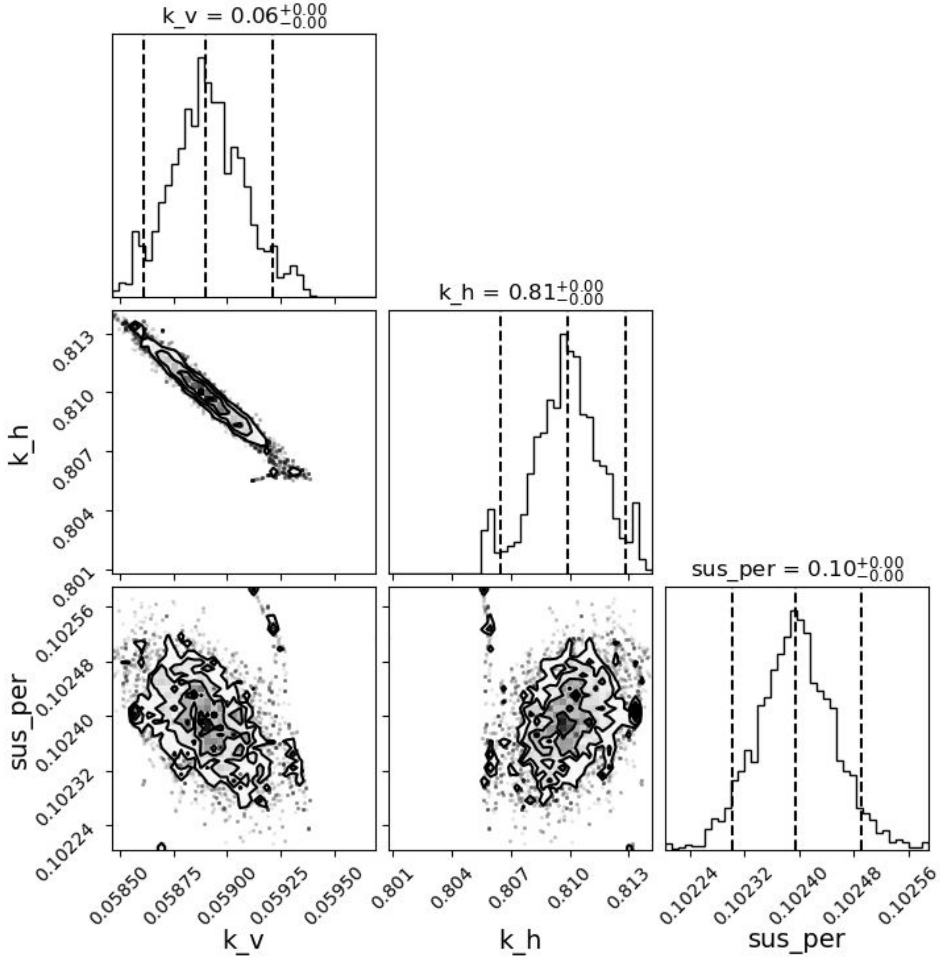
P: Susceptible Eggs
Q: Infected Eggs

S: Susceptible Humans/Vectors
E: Exposed Humans/Vectors
I: Infective Humans/Vectors
R: Recovered Humans

λ: Force of Infection
α: Incubation Rate
β: Recovery Rate
μ: Natural Mortality Rate
γ: DNV Mortality Rate

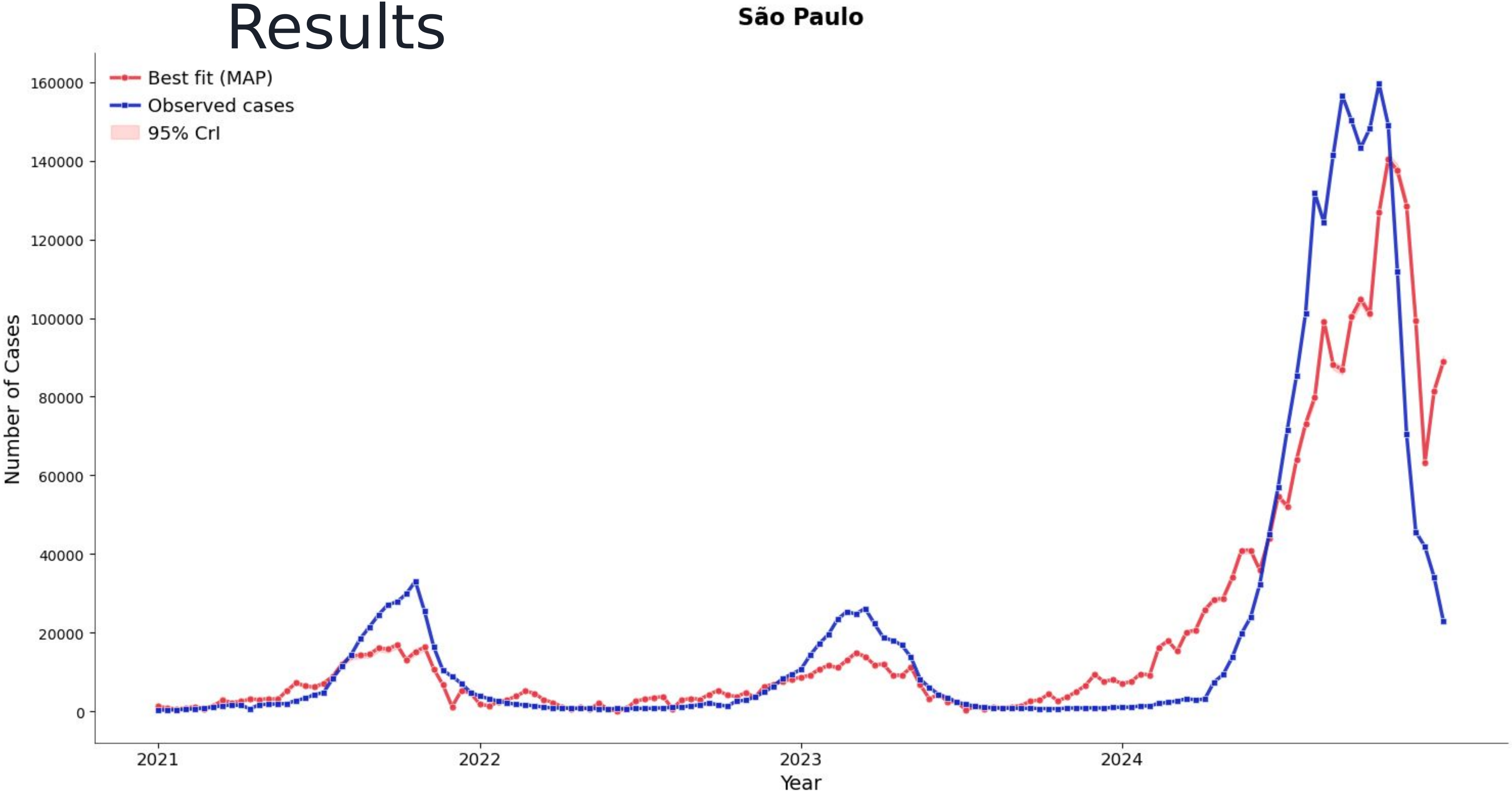
Model Calibration

- Parameters:
  - Human to Vector Transmission Probability -> Beta Distribution (α = 2.0 - β = 20.0)
  - Vector to Human Transmission Probability -> Beta Distribution (α = 8.0 - β = 5.5)
  - Percentage of Susceptible Population -> Beta Distribution (α = 3.0 - β = 20.0)
- Methods:
  - Markov Chain Monte Carlo (PyMC)
    - 6 chains
    - 5000 samples
    - 5000 tuning samples



Corner Plot and Posterior of the MCMC

Results



São Paulo
Best fit (MAP)
Observed cases
95% CrI
160000
140000
120000
100000
80000
60000
40000
20000
0
Number of Cases
2021
2022
2023
2024
Year

Dengue Forecast Model with Seasonal and Gravity Components

Marcio Maciel Bastos

October 14, 2025

Outline

1. Introduction

2. Content

This model combines seasonal dynamics, representing the **yearly dengue cycle**, with **spatial interaction**, where transmission strength depends on population and distance between regions.

Model

- **Sazonality:** $g_s(t) = \sum_{n=0}^{N} a_n(\sigma_s)\cos\left(\frac{2\pi n(t-\mu_s)}{T}\right)$, $a_n(\sigma_s) = \exp\left(-\frac{2\pi^2\sigma_s^2 n^2}{T^2}\right)$, $\max_t g_s(t) = 1$.
- **Gravity:** $\mathrm{grav}_{i,t} = \kappa(\mathrm{POP}_{i,t})^{\alpha_S}\sum_j(\mathrm{POP}_{j,t})^{\beta_I} d_{ij}^{-\gamma}$.
- **Logit:** $\eta_{s,t} = \mathrm{amp}_s\, g_s(t) + \mathrm{grav}_{s,t}$.

Model

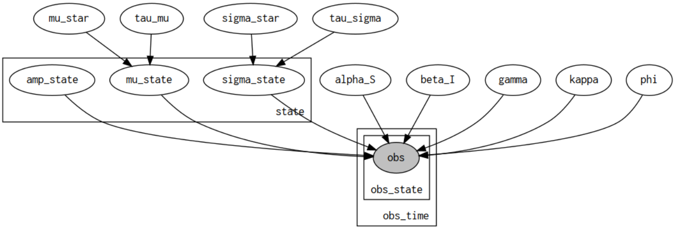


Figure 1: Bayesian Model