2024 Infodengue Sprint: Dengue Fever Predictive Modeling in Brazil

Dear participants,

As we approach the forecast submission deadline, we prepared this document describing important information regarding the forecast submission and evaluation processes and the next steps of the Sprint.

Important Dates:
1. Test 1 and Test 2 forecast submission deadline - August 16th, 2024
2. Presentation of participant's models - August 16th, 2024
3. 2025 dengue forecast submission deadline - August 25th, 2024
4. Presentation of the ensemble forecast model- August 30th, 2024

1. Forecast submission - 16 Aug

By 16 Aug, all teams should have uploaded their predictions in the Mosqlimate platform. Instructions can be found here.

2. The Aug 16th workshop

In this workshop, we will present the motivations for the Sprint and the expected outcomes. Each team will have 10-15 minutes to present their results for the testing phase. The following topics are requested to be included in the presentation:

- Variables used;
- Description of the forecast model;
- How the uncertainty is computed.
- Results
- Discussion of main challenges for the forecast

The workshop ends with a presentation of the next steps:

1. In the following week, the Mosqlimate group will assess the performance of each model using a set of scores, as described below. Individual scores will be calculated for each state and each year that correspond to test 1 and test 2. Based on these scores, we will rank the models according to different challenges, also described below. The results of this analysis will be shared with the teams.

2. Using the most up-to-date dengue data, we invite all teams to provide forecasts for 2025 using their models by Aug 25th.

3. The 2025 Forecast submission deadline - August 25th

- The forecast of the 2025 dengue season should be conducted for all 27 federal units. If your best model varies between regions, we recommend using the same model for units located within the same region.
- Upload the forecast to the platform as before.

4. The August 30th workshop

This is the final workshop where the Mosqlimate team will present the results of the testing phase, which includes evaluations of the individual models and ensembles. Based on the highest performing ensembles, the forecast for 2025 will be shared. All teams will have the opportunity to discuss the final results.
Additionally, we will provide a draft of the technical report for review and feedback.

Methods

Scoring

The logarithmic score, CRPS[1] and the interval score[2] will be computed using the `scoringrules[3]` Python package.

The **CRPS** will be computed using the equation below:

$$CRPS(N(\mu_i, \sigma_i^2), y_i) = \sigma_i\{\omega_i[\Phi(\omega_i) - 1] + 2\phi(\omega_i) - \frac{1}{\sqrt{\pi}}\}, \quad i = 1, \ldots, W_f. \quad (1)$$

where $\Phi(\omega_i)$ and $\phi(\omega_i)$ is the cumulative distribution function (CDF) and the probability density function (PDF) of the standard normal distribution, respectively, evaluated at the normalized prediction error $\omega_i = \frac{y_i - \mu_i}{\sigma_i}$. Additionally, y_i represents the cases observed in week i, μ_i is the mean forecasted value in week i and σ_i is the standard deviation of the forecast on week i. From the 90% prediction interval, $\sigma_i = \frac{u_i - l_i}{3.29}$, where u_i represents the *upper* limit and l the *lower* limit of the predictions. W_f is the final week considered, 52 or 26.

The **logarithmic score** is computed using the equation below:

$$LogS(N(\mu_i, \sigma_i^2), y_i) = log(\frac{\phi(\omega_i)}{\sigma_i}), \qquad i = 1, \ldots, W_f. \qquad (2)$$

Following [4], the **interval score** is defined as

$$S_\alpha^{int}(l_i, u_i; y_i) = (u_i - l_i) + \frac{2}{\alpha}(l_i - y_i)I\{y_i < l_i\} + \frac{2}{\alpha}(y_i - u_i)I\{y_i > u_i\},$$

$$i = 1, \ldots, W_f, \quad (3)$$

where I is the indicator function, and $\alpha = 0.1$. W_f is the final week considered, 52 or 26.

INFO DENGUE

Other metrics

The following metrics will be calculated as additional feedback to teams but will not affect the ranking of the models:

- Average scores at these regions of interest of the prediction window:
    - Epidemic onset: Weeks between growth start and peak
    - Epidemic peak: 3 week window centered on the peak
- The time lag, maximizing cross-correlation between predictions and data.

Ranking

For each year and state, the models will be assessed according to the six scores listed in the table below.

Average Score S = $\frac{1}{W_f}\sum_i^{W_f} S_i$	**Score (S) used**	**Final Week (W_f)**
S_1	CRPS	52
S_2	CRPS	26
S_3	Log Score	52
S_4	Log Score	26
S_5	Interval Score	52
S_6	Interval Score	26

The models will be ranked according to each score, that is, each model will receive a Rank R1, R2, …, R6, for each year and state.
Finally, the final ranking $R_{Y,S}$ of the models will be calculated with the following formula, for each **year** and **state**:

$$R_{Y,S} = \sum_{i=1}^{6} 1/R_i$$

A global ranking will be calculated using a similar method.

Ensemble

Two methods were selected for building the ensemble models, chosen for their ability to combine the strengths of different types of models. Models will be added to the ensemble incrementally, following the ranking order described above, until no further performance improvement is observed.

Voting regressor

A voting regressor is a meta estimator that takes the average or weighted averages of the forecasts produced by all the models in the ensemble. When using a weighted average, various weighting schemes will be explored to optimize the ensemble's performance.

Stacking regressor

The Stacking regressor is similar to the Voting regressor, with the difference that it combines the individual forecasts using a regression model specified. This final estimator is trained using cross-validation.

References

1. Gneiting, T., Raftery, A. E., Westveld, A. H. & Goldman, T. Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. (2005) doi:10.1175/MWR2904.1.

2. Bracher, J., Ray, E. L., Gneiting, T. & Reich, N. G. Evaluating epidemic forecasts in an interval format. *PLOS Comput. Biol.* **17**, e1008618 (2021).

3. Zanetta, Francesco, and Sam Allen. *Scoringrules: A Python Library for Probabilistic Forecast Evaluation*. 2024, GitHub, https://github.com/frazane/scoringrules. Accessed 8 Aug. 2024.

4. Gneiting, Tilmann, and Adrian E. Raftery. "Strictly proper scoring rules, prediction, and estimation." *Journal of the American statistical Association* 102.477 (2007): 359-378.

Financial Support:

Project Mosqlimate (grant 226088/Z/22/Z)