Modeling the Association Between Climatic Conditions and Dengue Incidence Using Monthly and Weekly Surveillance Data in Bauru, 2010–2018.

Camila Batista Rahal

Introduction

This study builds upon a previous project conducted in R (GitHub Repository: https://github.com/camilarahal/uni-capstone-project-dengue) and aims to model the association between climatic variables and dengue incidence in Bauru, Brazil using python and regression analysis. The goal is to explore possible association between temperature, humidity, precipitation, and dengue transmission.

Data collection

Dengue cases data originate from official health surveillance databases, where notifiable diseases such as dengue, Zika, and chikungunya are recorded. Cases are reported weekly, with most classified based on clinical and epidemiological criteria. Climatic variables used as explanatory factors include: Temperature (tempmean): Average daily temperature per week; humidity (humidmean): Average daily relative humidity per week; precipitation: Monthly accumulated precipitation.

Sources

- Precipitation data: https://www.ipmetradar.com.br/2estHist.php
- Cases, temperature and humidity data: https://info.dengue.mat.br/services/api

Data processing

Weekly dengue case and climate data were aggregated to a monthly level, and monthly precipitation data was merged with the dataset. Missing climatic values were imputed using linear interpolation, a common approach for time-series data. Due to overdispersion in the data (variance: 3,064,512.39; mean: 556.66), Negative Binomial Regression was selected as it accommodates this issue, unlike Poisson regression, which assumes equal mean and variance. To account for the delayed effects of environmental conditions on dengue transmission, one-month lagged variables for temperature, humidity, and precipitation were introduced.

Regression Analysis and Results

Monthly model

Dep. Variable:	monthly_cases	**No. Observations:**	131
Model:	GLM	**Df Residuals:**	127
Model Family:	NegativeBinomial	**Df Model:**	3
Link Function:	Log	**Scale:**	1.0000
Method:	IRLS	**Log-Likelihood:**	-856.19
Date:	Fri, 28 Mar 2025	**Deviance:**	240.13
Time:	22:29:20	**Pearson chi2:**	371.
No. Iterations:	11	**Pseudo R-squ. (CS):**	0.7956
Covariance Type:	nonrobust		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
Intercept	-11.9252	1.448	-8.233	0.000	-14.764	-9.086
temp_mean_lag1	0.3719	0.039	9.451	0.000	0.295	0.449
humidity_mean_lag1	0.1198	0.012	10.005	0.000	0.096	0.143
precipitation_lag1	-0.0015	0.001	-1.230	0.219	-0.004	0.001

The Negative Binomial regression model at the monthly level suggests that temperature and humidity (both lagged by one month) are significant predictors of dengue incidence. Higher temperatures (coef = 0.3719, p ¡ 0.001) are associated with an increase in dengue cases, while humidity (coef = 0.1198, p ¡ 0.001) also shows a positive correlation with dengue incidence. However, precipitation (coef = -0.0015, p = 0.219) is not statistically significant, indicating that it does not have a clear effect on monthly dengue cases. The model demonstrates good performance, with a Pseudo R^2 of 0.7956 and a log-likelihood of -856.19. Given the lack of significance of precipitation, this variable was removed, and the dataset was restructured to assess weekly trends.

Weekly model

Dep. Variable:	casos	**No. Observations:**	573
Model:	GLM	**Df Residuals:**	569
Model Family:	NegativeBinomial	**Df Model:**	3
Link Function:	Log	**Scale:**	1.0000
Method:	IRLS	**Log-Likelihood:**	-2691.5
Date:	Fri, 28 Mar 2025	**Deviance:**	682.50
Time:	22:29:20	**Pearson chi2:**	663.
No. Iterations:	58	**Pseudo R-squ. (CS):**	0.9018
Covariance Type:	nonrobust		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
Intercept	1.0579	0.519	2.040	0.041	0.041	2.074
tempmed_lag1	0.0062	0.015	0.427	0.669	-0.022	0.035
umidmed_lag1	0.0245	0.004	6.143	0.000	0.017	0.032
casos_lag1	0.0047	0.000	43.148	0.000	0.005	0.005

The Negative Binomial regression model at the weekly level suggests that temperature is not a significant predictor of dengue cases (coef = 0.0062, p = 0.669), indicating that short-term fluctuations in temperature do not strongly impact dengue transmission. However, humidity (coef = 0.0245, p ¡ 0.001) is a significant predictor of weekly dengue fluctuations, suggesting that humidity plays a more immediate role in dengue incidence. Additionally, the number of dengue cases from the previous week (coef = 0.0047, p ¡ 0.001) is also a strong predictor, highlighting the persistence of dengue cases over time. The model exhibits strong performance, with a Pseudo R^2 of 0.9018 and a log-likelihood of -2691.5.

Interpretation and Conclusions

The findings indicate that temperature influences dengue cases over a longer monthly timeframe, while humidity has a more immediate effect on weekly fluctuations. Additionally, dengue cases exhibit persistence over time, suggesting an autoregressive effect. The difference in model results suggests that mosquito life cycles (7-10 days) might align better with weekly modeling, but due to data constraints, a one-month lag was used instead.

These insights emphasize the importance of considering both short-term and long-term climatic effects when predicting dengue outbreaks. Future studies should integrate finer-scale precipitation data and explore additional time-series approaches for improved modeling accuracy.