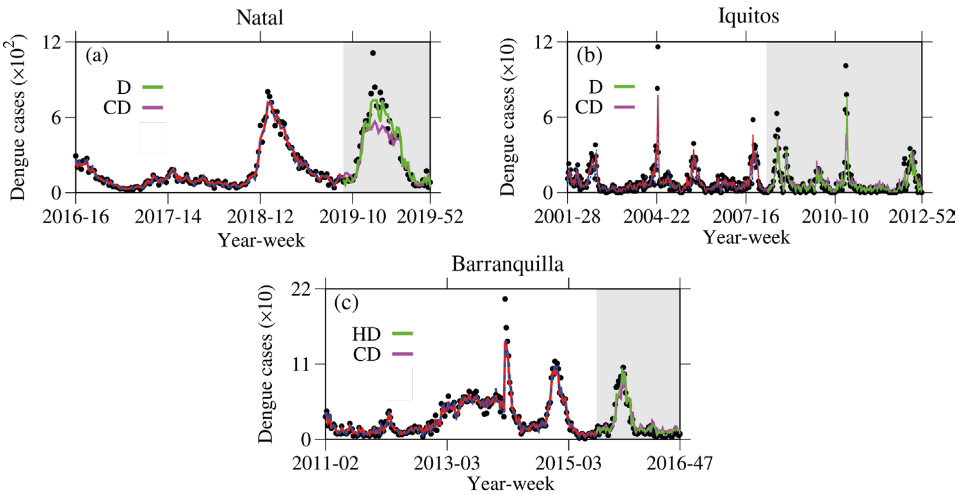
Assessing the impact of climate on dengue outbreaks

– DOI: https://doi.org/10.1051/2025103

Machine learning techniques reveal that researchers can improve their forecasts for the spread of dengue by incorporating climate data – but not in all cases



▲ Principal results obtained in the paper. The black points are real data for (a) Natal (Brazil), (b) Iquitos (Peru), and (c) Barranquilla (Colombia). The green and magenta curves are predicted by the ML technique, using dengue cases (D), climate plus dengue cases (CD), and humidity and dengue cases (HD) as input. The forecasting range is highlighted by the gray background.

Dengue is a mosquito-borne disease which infects around 390 million people globally each year. Case numbers have grown steadily in recent years, with the most severe outbreaks occurring in tropical regions of South America. To better predict how the disease spreads, it will be vital for researchers to fully understand how dengue case numbers are linked to different aspects of tropical climates.

Through new analysis published in *EPJ Special Topics*, a team led by Enrique Gabrick at the Potsdam Institute for Climate Impact, Germany, highlight how dengue forecasts can become more accurate by incorporating climate data – but also show that the success of this approach can vary between different regions. The team's results could enable researchers to develop more accurate forecasts for the spread of dengue, which may ultimately help to save thousands of lives.

Globally, the number of people infected with dengue has shown a worrying trend over the past 20 years: increasing from around 500,000 cases in 2000, to 5.2 million in 2019. "Most of these cases have been reported in tropical countries, particularly in the Americas," Gabrick explains. "This geographic preference is driven by environmental conditions and climate factors, such as temperature, humidity, and precipitation, which are critical for mosquito life cycles."

In South America, dengue infections have accelerated even further over the past year: with over 670,000 cases reported in the first five weeks of 2024. In order for public health interventions to slow down this rate of infection, it will be vital for researchers to build more accurate predictive models, which incorporate all variables with a measurable impact on future case numbers.

To address the challenge, Gabrick's team applied a machine learning technique based on 'random forest' algorithms. This method works by building many, slightly different 'decision trees': each essentially a flowchart which makes independent interpretations of new data, based on the algorithm's training data.

"We chose the random forest algorithm due to its robustness and predictive capacity," Gabrick explains. "It is based on an ensemble learning method and consists of multiple decision trees, enabling more accurate predictions than individual models. Moreover, the algorithm evaluates the importance of input variables, providing valuable insights into the factors that influence forecasts."

To test their approach, the team trained their algorithm using historical dengue case numbers from three different cities in Brazil, Peru, and Colombia – each with a tropical climate. They then used it to forecast the number of dengue cases one week ahead based on three separate considerations: current dengue cases alone; dengue cases combined with climate data (including temperature, precipitation, and humidity); and dengue cases combined with humidity alone.

Surprisingly, each of these considerations yielded the best prediction in one of each of the cities studied by the team. Altogether, the result revealed how the incorporation of climate variables leads to mixed success when predicting future cases numbers, and doesn't always help to improve forecasts. But it also shows that by carefully considering whether or not to combine case numbers with climate data, only humidity data, or neither, researchers could ultimately improve the accuracy of their predictions.

"Additionally, we underscore the importance of machine learning techniques in combining meteorological and epidemiological data to enhance forecasts," Gabrick says. "We expect that our results have a strong potential to improve the dengue forecasting and provide valuable insights for public health interventions." ■

References

[1] S.T. da Silva, E.C. Gabrick, P.R. Protachevicz *et al.*, Eur. Phys. J. Spec. Top. (2024). https://doi.org/10.1140/epjs/s11734-024-01201-7