Hybrid neural–mechanistic modeling of leptospirosis transmission with environmental drivers: Evidence from Thailand

*Sumet Khumphairan, Sudarat Chadsuthi, Peter Fransson, Yichao Liu, **Charin Modchang**, Joacim Rocklöv, and Ekaterina Kostina*

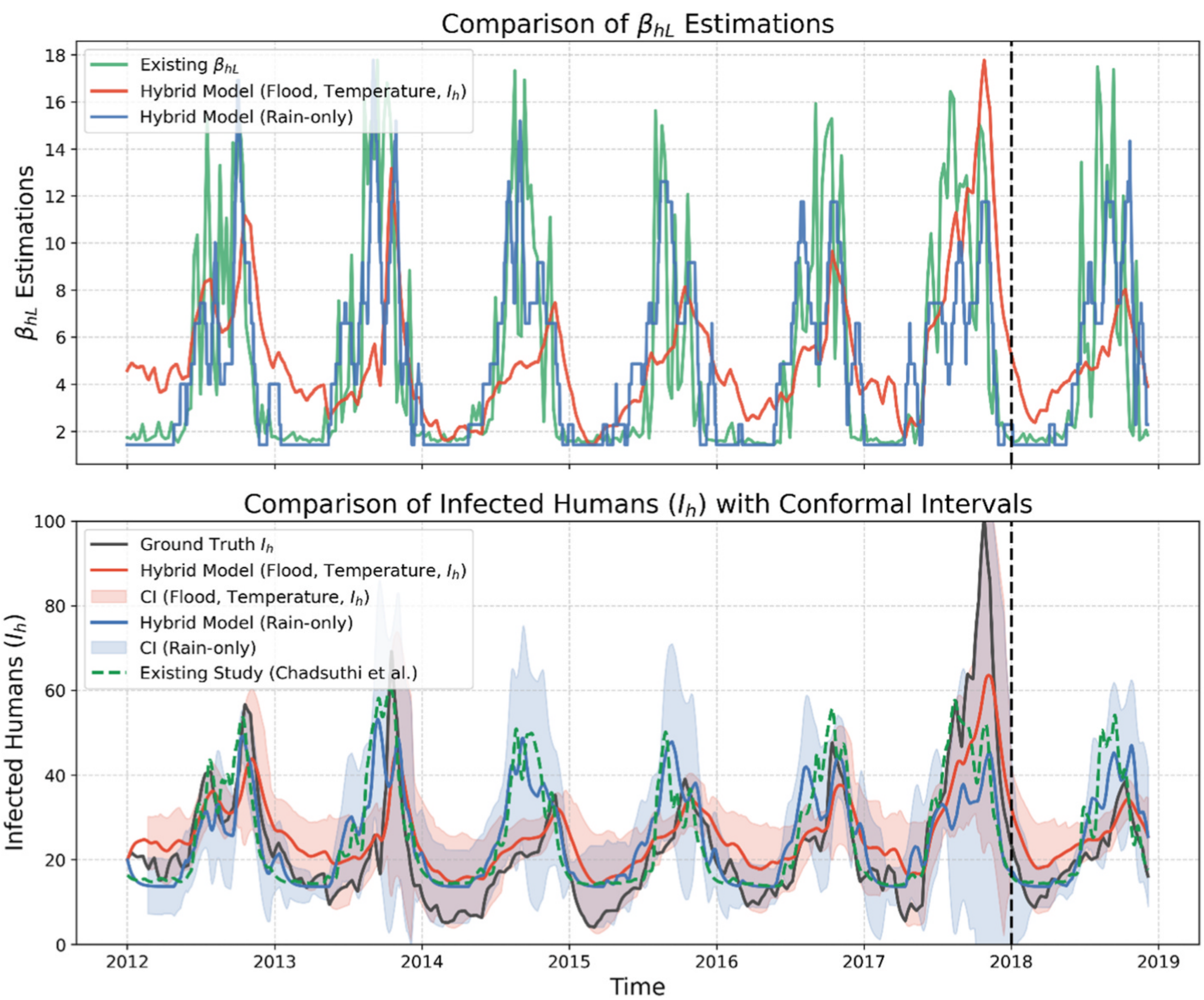
Rationale and objectives: Leptospirosis is a serious bacterial disease spread through water and soil contaminated by infected animals, causing an estimated one million cases and 60,000 deaths worldwide every year. In Thailand, outbreaks follow a strong seasonal rhythm, peaking during the heavy rains and flooding of the monsoon season, especially among rural and farming communities. Predicting these outbreaks is difficult because the rate at which the disease spreads keeps changing with the weather. Traditional mathematical models assume this transmission rate stays fixed, while purely data-driven AI models can track the changes but reveal little about the underlying biology. The researchers therefore set out to build a hybrid model that combines the strengths of both approaches—using artificial intelligence to learn how the transmission rate shifts over time, while keeping the biological structure of a classical disease model—to better estimate and forecast leptospirosis transmission in Thailand.

Summary: The team developed a hybrid framework that pairs a Long Short-Term Memory (LSTM) neural network—a type of AI designed to find patterns in time-series data—with a compartmental disease model that tracks how infection moves between people, animals, and the contaminated environment. The neural network learns the key transmission rate from the environment to humans, drawing on satellite measurements of rainfall, flooding, and temperature together with reported human cases. Using daily surveillance data from 2012 to 2018, the team trained and tested ten versions of the model built from different combinations of these inputs, then benchmarked each one against actual case counts and against an earlier modeling study.

Outcome: The best-performing model, which combined flooding, temperature, and human case data, predicted infections far more accurately than the previous modeling study, cutting one key error measure by about 61%. Importantly, models built on environmental data alone—particularly rainfall or flooding—still forecast outbreaks well without needing case reports, making them valuable when surveillance data are delayed or incomplete. The study also found that adding more variables did not always help: simpler input combinations often generalized better than more complex ones. The approach stayed robust even when case reporting was artificially delayed or undercounted, and because it relies on routinely available climate data, it is well suited to early warning in data-limited settings. While demonstrated here for leptospirosis, the same framework could be adapted to other environmentally driven diseases such as dengue, cholera, and malaria.

Research grant and acknowledgement: National Research (NU) and the National Science, Research and Innovation Fund (NSRF) (Grant Agreement No. R2568B013), and the Pandasia project funded by the European Union's Horizon Europe research and innovation programme (Grant Agreement No. 101095444).

Related SDGs goal: 3. Good health and well-being.



Graphical summary: Comparison of transmission-rate (β_{hL}) estimations and infected human cases (I_h) between the proposed hybrid model and an existing study. **(Top)** Estimated transmission rate $\beta_{hL}(t)$ over time: the orange line is the hybrid model using flooding, temperature, and human cases (I_h); the blue line is the rain-only model; and the green line shows the existing study by Chadsuthi et al. **(Bottom)** Predicted infected human cases $I_h(t)$: the black line is the ground truth, the orange line is the hybrid model with flooding, temperature, and I_h, and the blue line is the rain-only hybrid model, while the green line is simulated using the existing study. Shaded bands show 95% conformal prediction intervals for each hybrid model (orange and blue), computed with a 50-day rolling calibration window. The vertical dashed line marks the train–test split (end of 2017).

Related publication:

Khumphairan S, Chadsuthi S, Fransson P, Liu Y, **Modchang C**, Rocklöv J, and Kostina E. Hybrid neural–mechanistic modeling of leptospirosis transmission with environmental drivers: Evidence from Thailand. *Computers in Biology and Medicine*. https://doi.org/10.1016/j.compbiomed.2026.111632