

Enhancing Malaria Prediction Accuracy in Burkina Faso

A Novel Approach in Digital Health

P. Venkova, Q. Guilmard, D. Harvey, M. Hervin, C.S.B Kompaoré, D. Loureiro, G. Saral Fragkos, F. Triclin, S. Toé , W- Valkenburg



Terre des Hommes, Lausanne, Switzerland

Background

- **The Integrated e-Diagnostic Approach (leDA):** Digitizes WHO/UNICEF clinical protocols for better child healthcare in Burkina Faso.
- **Machine Learning Integration:** Leverages data to predict malaria outbreaks up to three months ahead in 11 regions and 60 districts in Burkina Faso.
- **Strategic Planning:** Includes a five-tier system for improved health planning and disease management.

Data

- Data encompassing **weekly malaria cases, tests, and consultations** from **1,089** Primary Healthcare Centers (PHCs) across **60** districts over a **2.5-year** period was analyzed. This dataset included integrated **rainfall** data from **CHIRPS** and **surface water** metrics specific to each PHC.
- A rolling window analysis with a span of **26 weeks** was conducted, yielding in excess of **100,000 time-series datasets** for detailed examination of the correlation between malaria incidence and rainfall patterns.

Methods

- Developed a library with **100,000 Gaussian Process models**. Assessed models using the latest 13 weeks data to select the top performers based on likelihood scores. Selected models to forecast up to 13 weeks into the future.
- Created a **five-tier epidemic alert system** (triggers at **5pc, 32pc, 68pc, 95pc**) (Fig 2).
- Validated epidemic triggers by **Random Forest Regressor** using **131** epidemic and **300** non-epidemic instances over **13-week** periods. Performance measured by **Precision** (indicating the accuracy of the system when predicting an epidemic) and **Recall** (reflecting the system's ability to identify all actual epidemics).

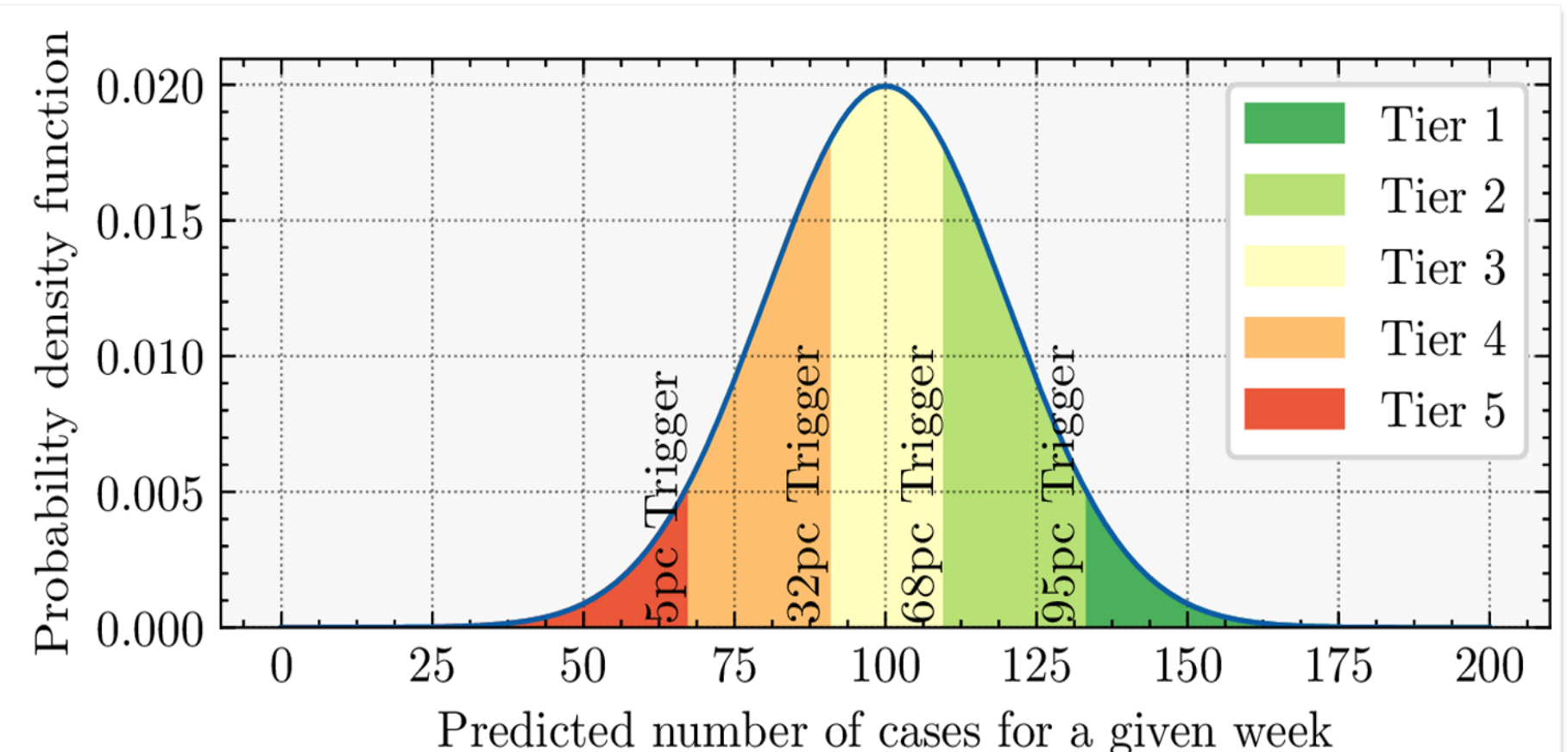


Fig. 2: Predicted Malaria Incidence and Alert Tiers at PHC
Each tier is defined by how many of those limits cross the World Health Organization (WHO) threshold for an epidemic at any point during the predicted weeks.
Tier 5, highlighted in red, indicates a critical situation where the predicted number of malaria cases surpasses the WHO threshold across all percentiles.

Results

Based on the Random Forest Regressor we define the risk for epidemic given the tier system as follows (Fig. 3)

- Tier 1: No Alert: <1%
- Tier 2: Low Alert: 30%
- Tier 3: Medium Alert: 50%
- Tier 4: High Alert: 84%
- Tier 5: Very High Alert: >99%

Percentile Trigger	Precision (%)	Recall (%)
5th	> 99	4.58
32nd	83.67	31.30
50th	67.97	66.41
68th	51.08	90.07
95th	32.27	> 99

Fig. 3: Precision and Recall for different percentile triggers

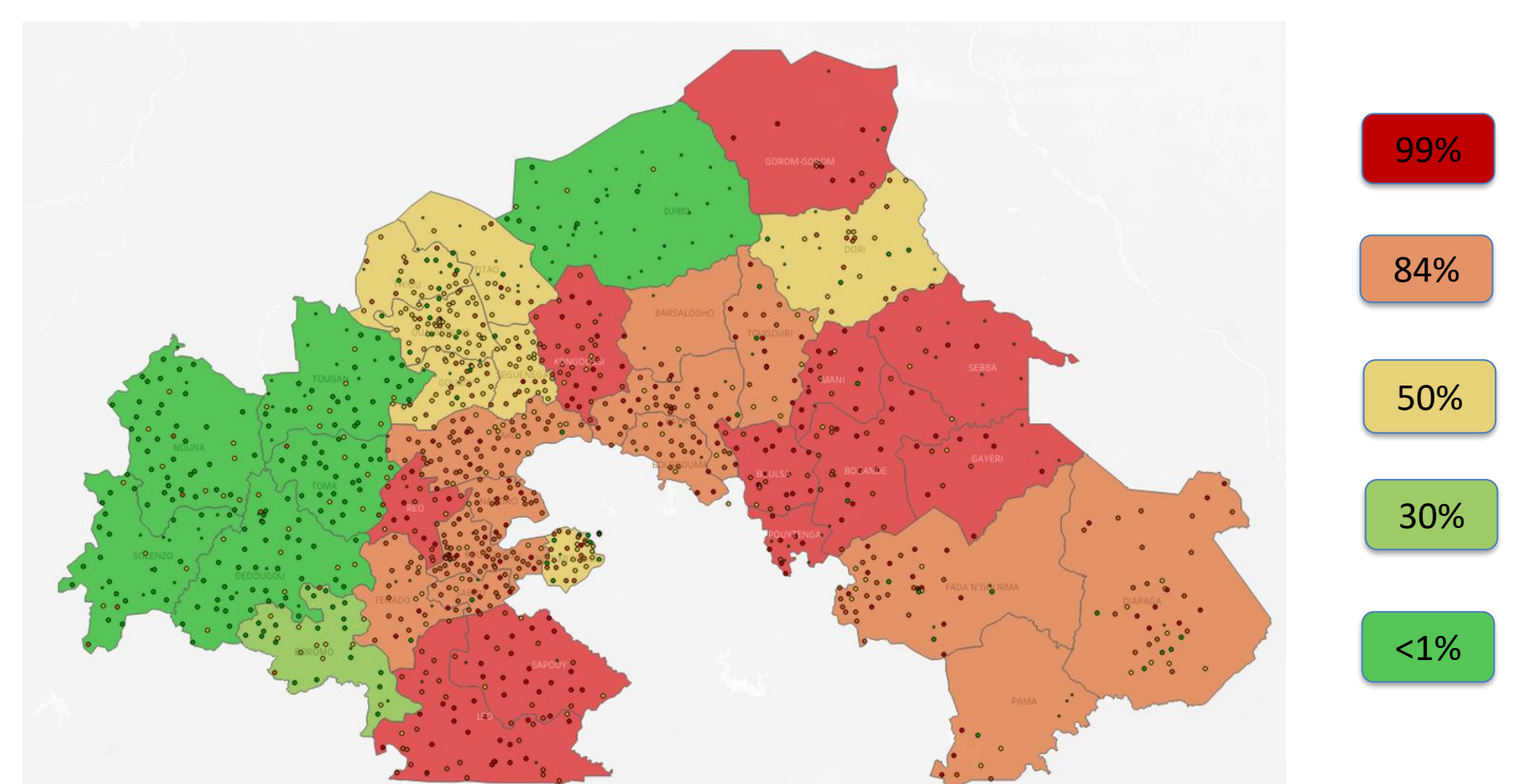


Fig. 4: Projected Epidemic Risk for Burkina Faso Districts

Conclusion

- ✓ The hybrid model, integrating Gaussian Processes with Random Forest, forecasts malaria incidence across districts with a **47.3%** validity rate at a **68%** confidence level and **65.8%** at **95%**.
- ✓ Coupled with this predictive ability is a five-tier alert system characterizing epidemic risks ranging from below **30%** to above **99%**.

Contact : gozde.saralfragkos@tdh.org

