On-demand Model Validation Built into Infectious Disease Early Warning Systems: Malaria Forecasts in Ethiopia with R package epidemiar

Dawn Nekorchuk[1], Justin Davis[1], Teklehaimanot Gebrehiwot[2,3,3,4,5], and Michael Wimberly[6]

[1]University of Oklahoma Norman Campus
[2]Bahir Dar University
[3]Health, Development, and Anti-Malaria Association
[4]Amhara Public Health Institute
[5]Amhara National Regional State Health Bureau
[6]University of Oklahoma

November 24, 2022

Abstract

Developing and implementing a malaria early warning system that integrates public health surveillance with monitoring of related environmental factors is the goal of the Epidemic Prognosis Incorporating Disease and Environmental Monitoring for Integrated Assessment (EPIDEMIA) project. Collaborating with our Ethiopian partners on requirements, we developed the R package epidemiar to provide a generalized set of functions for disease forecasting, plus customized code including a Google Earth Engine script for environmental data and formatting scripts for distributable reports with maps and graphs. Since 2019, a local team at Bahir Dar University in Ethiopia has been using EPIDEMIA to produce weekly malaria forecasting reports. Intensive anti-malarial efforts in the Amhara region of Ethiopia have resulted in declining malaria incidence, with a 75% decrease in cases between 2013 and 2018 (561,101 to 137,445 cases). In this context of potentially changing malaria transmission patterns, continual model evaluation past the initial model development is warranted. We built model validation and assessment tools into the epidemiar R package for on-demand evaluation for any historical period. Validation statistics included Mean Error (ME), Mean Absolute Error (MAE), and proportion of observations that fell inside the prediction intervals. Evaluation can be made for one through n-week ahead predictions, and include comparisons with two naïve models: persistence of last known value, and average cases from that week of the year. Building validation into the early warning system provides more opportunities to learn about the model via the validation results. We can identify locations where the models perform best with district-level results. With on-demand implementation and time-range flexibility, we can also investigate how accuracy changes over time, which is of particular interest in places like Ethiopia with changing patterns and declining trends due to anti-malarial programs.

On-demand Model Validation in Infectious Disease Early Warning Systems
Malaria Forecasts in Ethiopia Using R Package epidemiar

Dawn M. Nekorchuk[1], Justin K. Davis[1], Worku Awoke[2], Abere Mihretie[3], Aklilu Getinet[3], Mastewal Worku[4], Teklehaimanot Gebrehiwot[4], Michael C. Wimberly[1]

[1]Department of Geography & Environmental Sustainability, University of Oklahoma, USA; [2]School of Public Health, Bahir Dar University, Ethiopia; [3]Health, Development, and Anti-Malaria Association, Ethiopia; [4]Amhara Public Health Institute, Ethiopia

EPIDEMIA research project

Epidemic Prognosis Incorporating Disease and Environmental Monitoring for Integrated Assessment

- Create early warning malaria forecast reports integrating epidemiological & environmental data

R package `epidemiar`

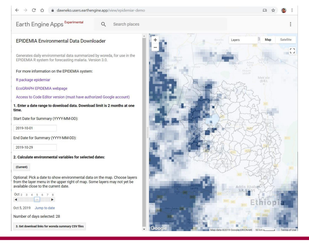
- Modeling, forecasting & validation functions
- Flexible – supports various environmentally-mediated diseases, locations, environmental variables
- Event detection: Farrington improved algorithm
- https://github.com/EcoGRAPH/epidemiar

Custom R project

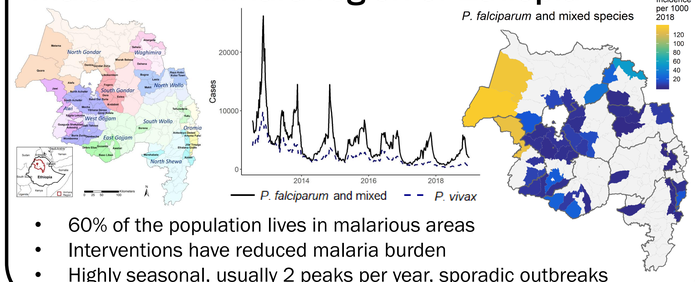
- Region & disease-specific settings for local data input, modeling, event detection, and report formatting
- https://github.com/EcoGRAPH/epidemiar-demo

Google Earth Engine App

- Environmental data summaries
- Fast, on-demand
- Small download



Malaria in Amhara region of Ethiopia



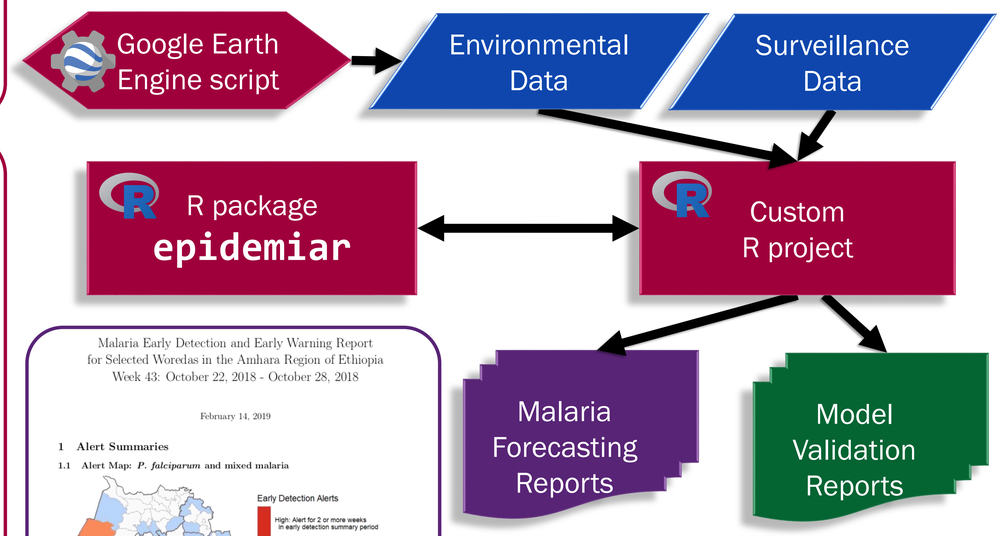
- 60% of the population lives in malarious areas
- Interventions have reduced malaria burden
- Highly seasonal, usually 2 peaks per year, sporadic outbreaks

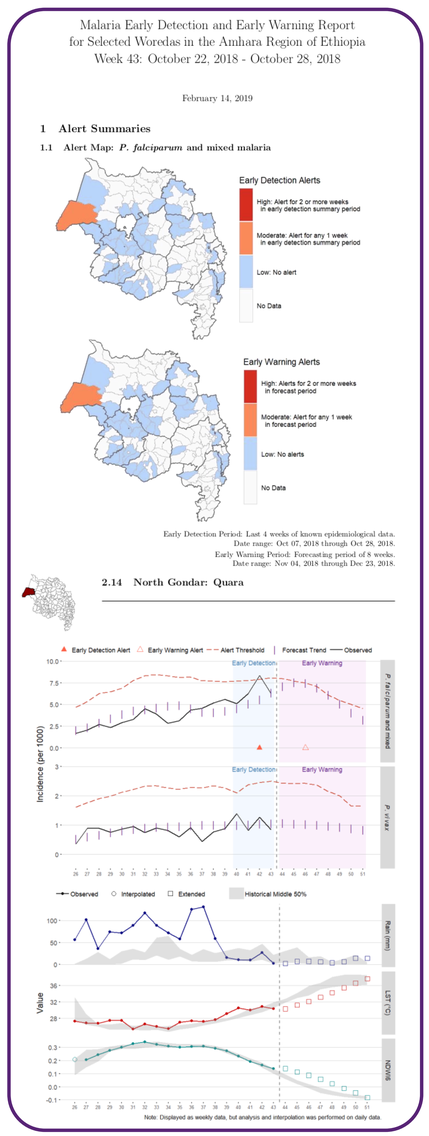
Malaria Forecast Reports

- Developed with researcher and public health feedback
- Currently forecasts 8 weeks into the future
- Early detection and early warning alerts
- Can be run from a desktop/laptop
- Map summaries and per district reports
- Includes environmental factors for context

COLLEGE OF ATMOSPHERIC AND GEOGRAPHIC SCIENCES
DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL SUSTAINABILITY
The UNIVERSITY of OKLAHOMA

EPIDEMIA Forecasting System




Malaria Early Detection and Early Warning Report
for Selected Woredas in the Amhara Region of Ethiopia
Week 43: October 22, 2018 - October 28, 2018

February 14, 2019

1 Alert Summaries

1.1 Alert Map: *P. falciparum* and mixed malaria

Early Detection Alerts
High: Alert for 2 or more weeks in early detection summary period
Moderate: Alert for any 1 week in early detection summary period
Low: No alert
No Data

Early Warning Alerts
High: Alerts for 2 or more weeks in forecast period
Moderate: Alert for any 1 week in forecast period
Low: No alerts
No Data

Early Detection Period: Last 4 weeks of known epidemiological data.
Date range: Oct 07, 2018 through Oct 28, 2018.
Early Warning Period: Forecasting period of 8 weeks.
Date range: Nov 04, 2018 through Dec 23, 2018.

2.14 North Gondar: Quara

Early Detection Alert | Early Warning Alert | Alert Threshold | Forecast Trend | Observed
Incidence (per 1000)
P. falciparum and mixed
P. vivax
Observed | Interpolated | Extended | Historical Middle 50%
Rain (mm)
LST (°C)
NDMI
Value
Note: Displayed as weekly data, but analysis and interpolation was performed on daily data.

Model Assessment & Validation

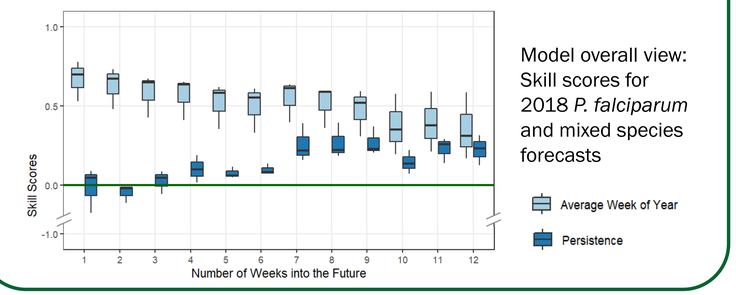
- Built into forecasting model system
- On demand, user-specified time range
- User-specified future forecast period

Skill Scores

- Accuracy statistics
  - Mean Absolute Error (MAE),
  - Root Mean Squared Error (RMSE)
  - R^2 (variance explained),
- Compare against naïve models
  - Average week of the year: based on historical epidemiological data
  - Persistence: last known value carried forward *n* weeks
- Skill score shows relative improvement of forecast model over naïve model, calculated per accuracy statistic

Validation statistics for 2018 *P. falciparum* & mixed species

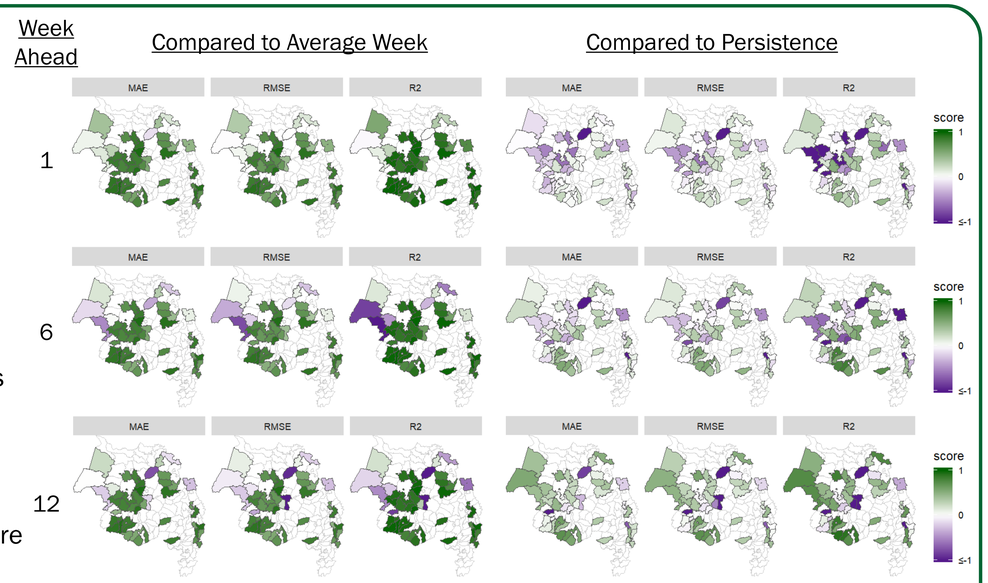
Week Ahead	MAE	RMSE	R^2
1	19.5	54.0	0.76
2	21.1	59.7	0.71
4	23.5	68.0	0.62
6	25.3	76.9	0.51
8	26.0	73.6	0.55
10	27.5	92.6	0.29
12	26.8	95.5	0.25



Model overall view: Skill scores for 2018 *P. falciparum* and mixed species forecasts

Geographic View of Validation

- Geographic insights at district level
- Maps help visualize geographic patterns
- Low skill could indicate presence of malaria drivers other than environmental variables
- Increases the transparency of the modeling and forecasts
- Identifies locations where model works well and where it does not



Skill scores for 2018 *P. falciparum* & mixed species forecasts

Contact

Dawn.Nekorchuk@ou.edu or Twitter: @nekodawn
EcoGRAPH Research Group: http://ecograph.net

Acknowledgements

This work is supported by the National Institute of Allergy and Infectious Diseases (R01-AI079411) and the USAID Adaptation Thought Leadership and Assessments (ATLAS) project (AID-OAA-TO-14-00044)