Epidemic Prognosis Incorporating Disease and Environmental Monitoring for Integrated Assessment (EPIDEMIA)

Prospects for Scaling up Malaria Early Warning in Ethiopia

Michael C. Wimberly, Dawn M. Nekorchuk, Justin K. Davis

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

USAID
FROM THE AMERICAN PEOPLE

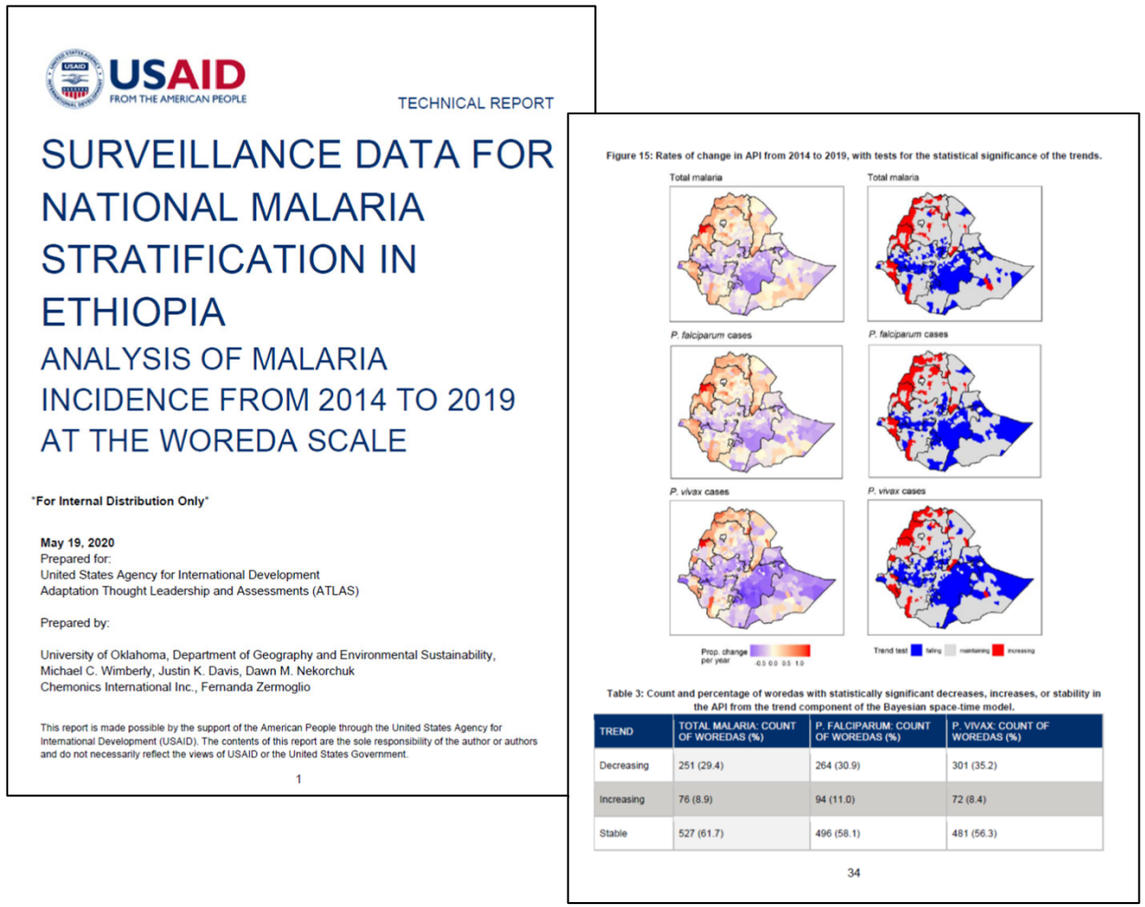
The objective of this project is **to develop a roadmap for scaling-up and implementing malaria early warning at a national level** in Ethiopia and other malaria-impacted countries.

- Supported by USAID through the Adaptation Thought Leadership and Assessment (ATLAS) project.
- To accomplish this objective we have:
  - Evaluated the suitability of national-level malaria surveillance data for malaria stratification and forecasting
  - Updated the EPIDEMIA software to support malaria forecasting for multiple regions in Ethiopia

Today we will present the current status of EPIDEMIA and hope to get your feedback on what is needed for scale-up.

We produced two USAID reports, in collaboration with EPHI, on data quality and usability for malaria stratification and early warning.

- Concluded that PHEM data are suitable for malaria stratification and early warning
- Described methods for harmonization, screening, and imputation
- Presented preliminary estimates of malaria stratification and trends at a national level



USAID
FROM THE AMERICAN PEOPLE

TECHNICAL REPORT

SURVEILLANCE DATA FOR NATIONAL MALARIA STRATIFICATION IN ETHIOPIA

ANALYSIS OF MALARIA INCIDENCE FROM 2014 TO 2019 AT THE WOREDA SCALE

"For Internal Distribution Only"

May 19, 2020
Prepared for:
United States Agency for International Development
Adaptation Thought Leadership and Assessments (ATLAS)

Prepared by:

University of Oklahoma, Department of Geography and Environmental Sustainability,
Michael C. Wimberly, Justin K. Davis, Dawn M. Nekorchuk
Chemonics International Inc., Fernanda Zermoglio

This report is made possible by the support of the American People through the United States Agency for International Development (USAID). The contents of this report are the sole responsibility of the author or authors and do not necessarily reflect the views of USAID or the United States Government.

1

Figure 15: Rates of change in API from 2014 to 2019, with tests for the statistical significance of the trends.

Total malaria | Total malaria
P. falciparum cases | *P. falciparum* cases
P. vivax cases | *P. vivax* cases

Prop. change per year
Trend test: falling, maintaining, increasing

Table 3: Count and percentage of woredas with statistically significant decreases, increases, or stability in the API from the trend component of the Bayesian space-time model.

TREND	TOTAL MALARIA: COUNT OF WOREDAS (%)	P. FALCIPARUM: COUNT OF WOREDAS (%)	P. VIVAX: COUNT OF WOREDAS (%)
Decreasing	251 (29.4)	264 (30.9)	301 (35.2)
Increasing	76 (8.9)	94 (11.0)	72 (8.4)
Stable	527 (61.7)	496 (58.1)	481 (56.3)

34

The UNIVERSITY of OKLAHOMA

USAID
FROM THE AMERICAN PEOPLE

Structure of this presentation

- **BACKGROUND** – rationale behind the design of EPIDEMIA
  - Break for questions and discussion
- **DATA AND METHODS** – environmental and epidemiological data used to calibrate predictive models
  - Break for questions and discussion
- **APPLICATION** – software tools that integrate methods and data to generate and evaluate malaria forecasts
  - Final questions and discussion

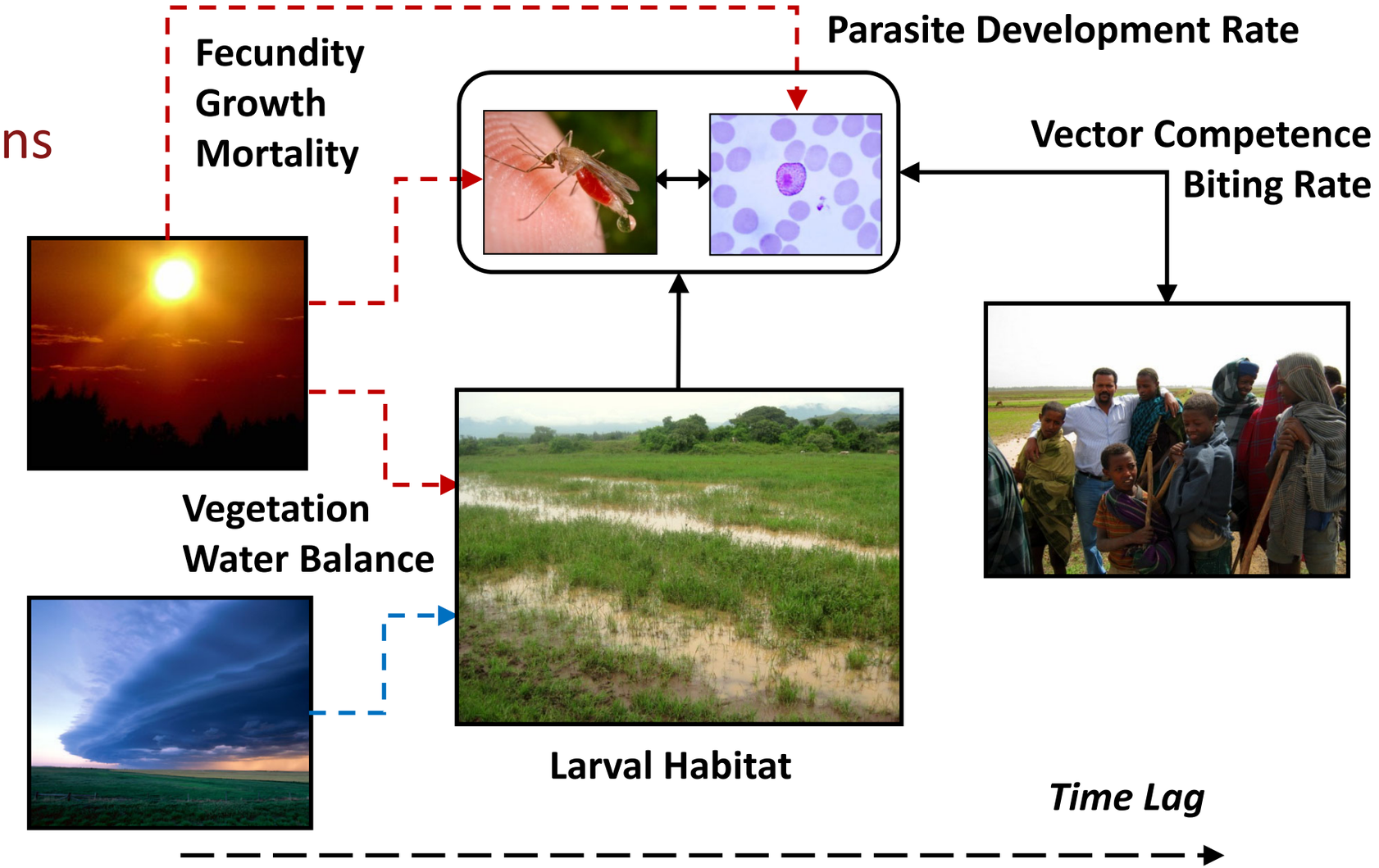
BACKGROUND

- Scientific basis for predicting malaria with climate data
- Limitations of climate-driven forecasts
- Design and development of EPIDEMIA

Climate factors such as temperature and precipitation affect mosquito populations and malaria transmission.

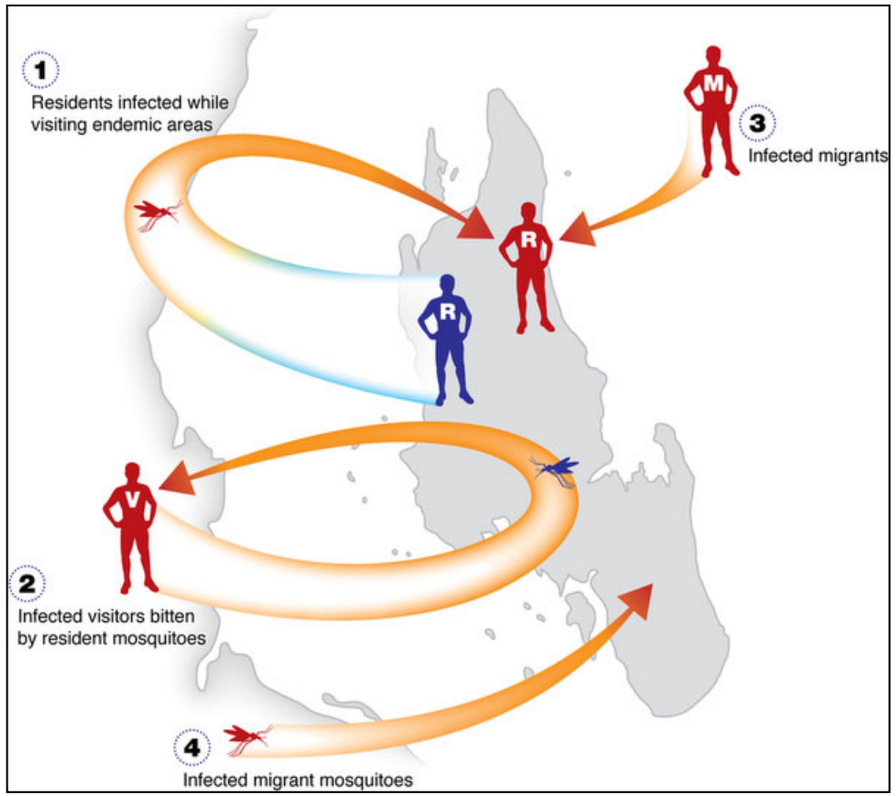
Temperature

Rainfall



However, numerous other non-climatic factors also influence malaria transmission.

- Movement of people and parasites
- Accessibility and quality of health care
- Interventions for malaria prevention and control



Le Menach et al. 2011, Nature Scientific Reports

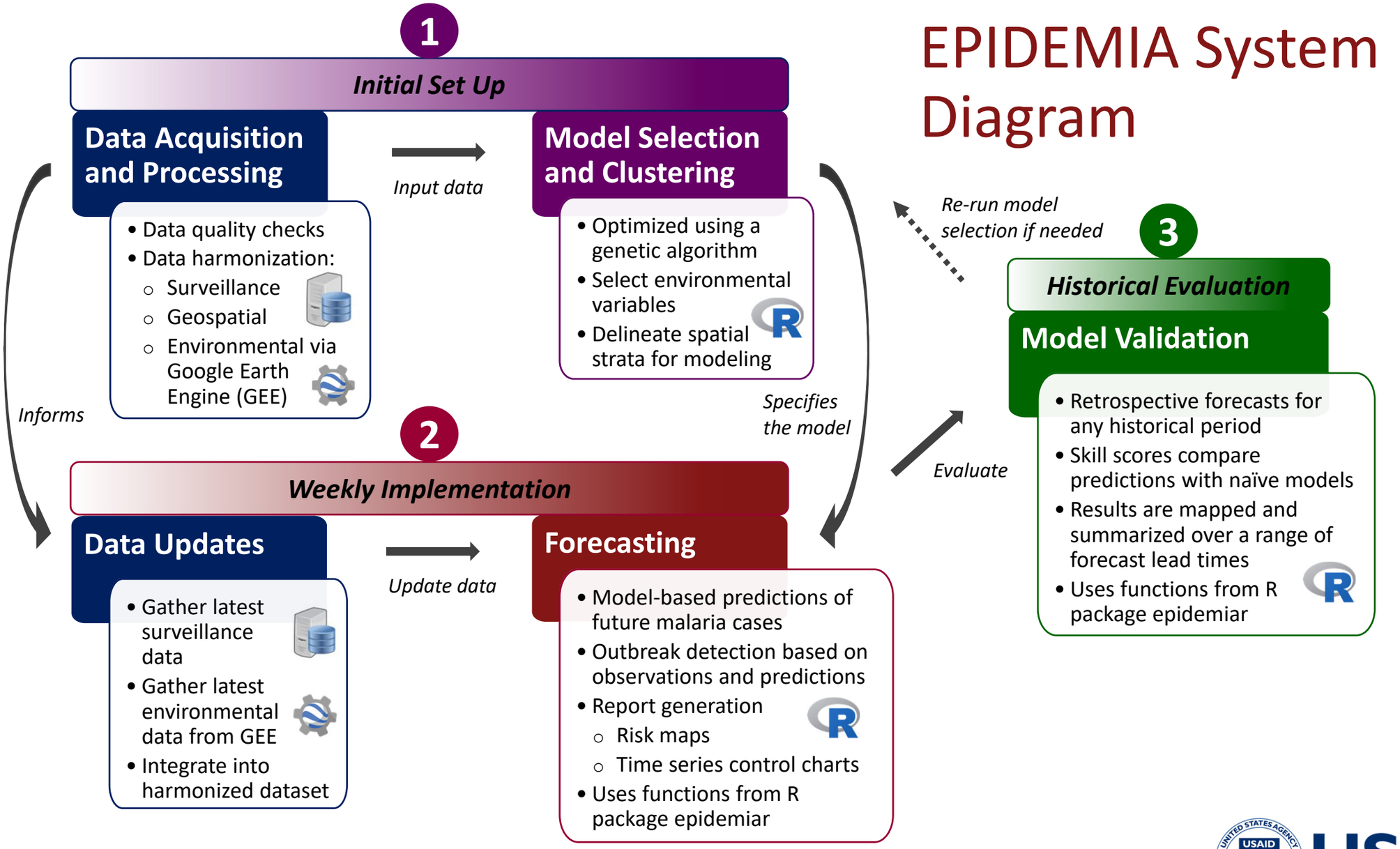
To address the challenge of malaria forecasting, we developed the Epidemic Prognosis Incorporating Disease and Environmental Monitoring for Integrated Assessment (EPIDEMIA) system.

- Co-developed by the University of Oklahoma (formerly South Dakota State University); Health, Development, and Anti-Malaria Association, Amhara Public Health Institute, and Bahir Dar University
- Piloted in the Amhara region
- Data driven, combining malaria surveillance data with environmental observations from Earth-observing satellites
- Multiple approaches for outbreak prediction
  - Early warning forecasts based on environmental models
  - Early detection of outbreaks based on recent malaria case fluctuations

EPIDEMIA has features that make it particularly well-suited for scaling up malaria early warning.

- Free and open-source software using freely available data
  - R language and environment for statistical computing
  - NASA remote sensing data products accessed through Google Earth Engine
- Automates time-consuming steps in the forecasting workflow
  - Acquires, processes, and formats data
  - Fits predictive models and generates forecasts for future time steps
  - Produces detailed formatted reports with maps and charts
- Can be applied in any locations where the appropriate malaria and environmental data are available
- Requires trained analysts to operate the software and interpret the outputs

EPIDEMIA System Diagram



Break for questions

Points to consider:

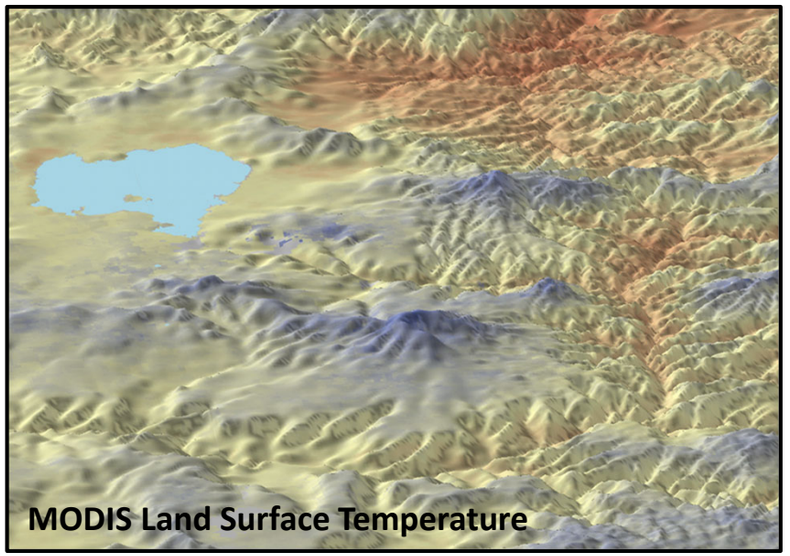
- What are the current gaps in our understanding of the effects of climate on malaria outbreaks?
- What are the benefits and challenges of implementing malaria early warning at a national level in Ethiopia?

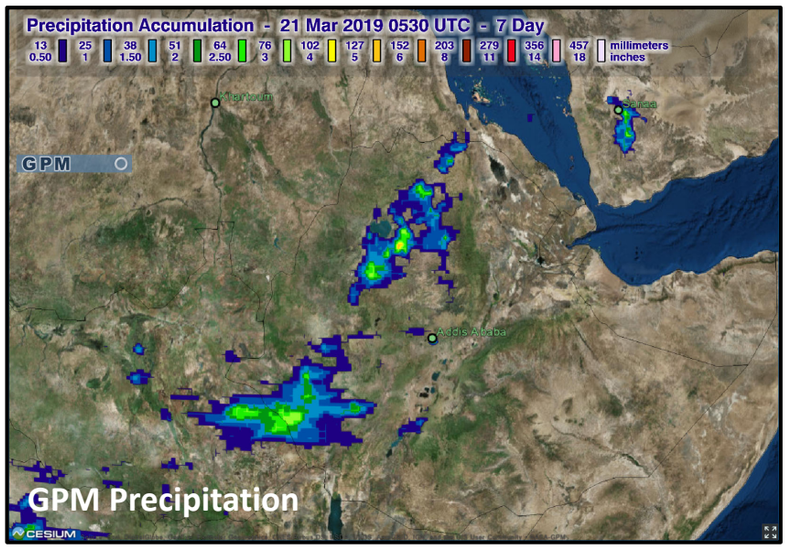
DATA AND METHODS

- Data sources
- Harmonization
- Screening and imputation
- Predictive modeling

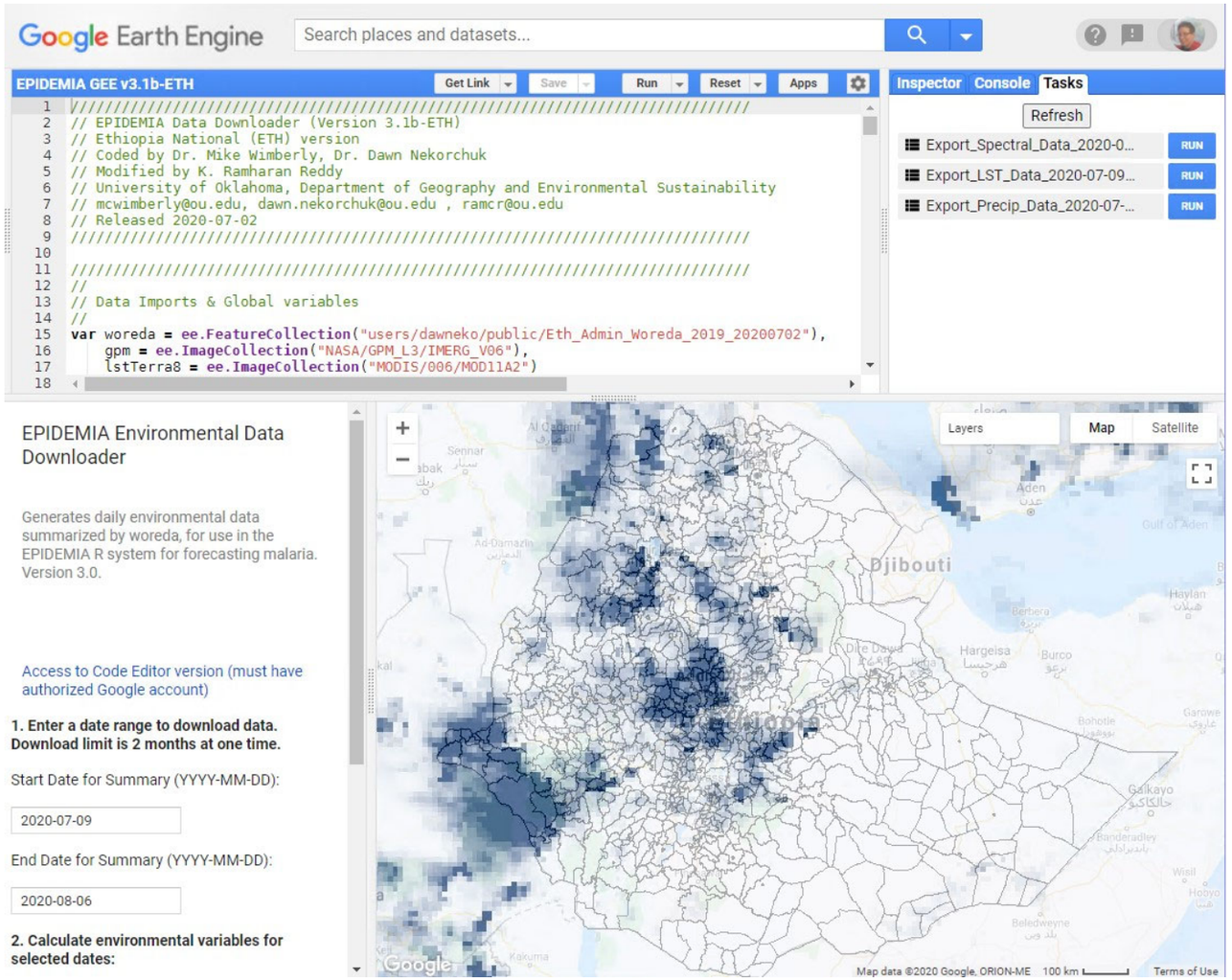
We use environmental data collected by NASA Earth-observing satellites.

- High-quality, spatially contiguous observations with frequent measurement intervals
- Available FREELY for almost any location in the world
- Variety of datasets available:
  - Land surface temperature (MODIS LST)
  - Precipitation (Global Precipitation Measurement Mission)
  - Vegetation greenness and moisture indices (MODIS optical-IR reflectance)

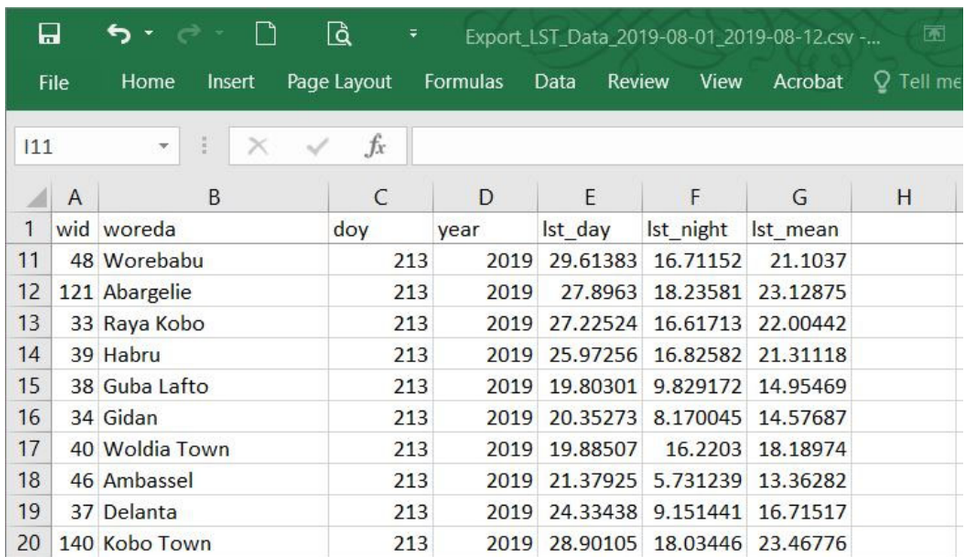




These environmental data are accessed and processed in the cloud using Google Earth Engine (GEE).



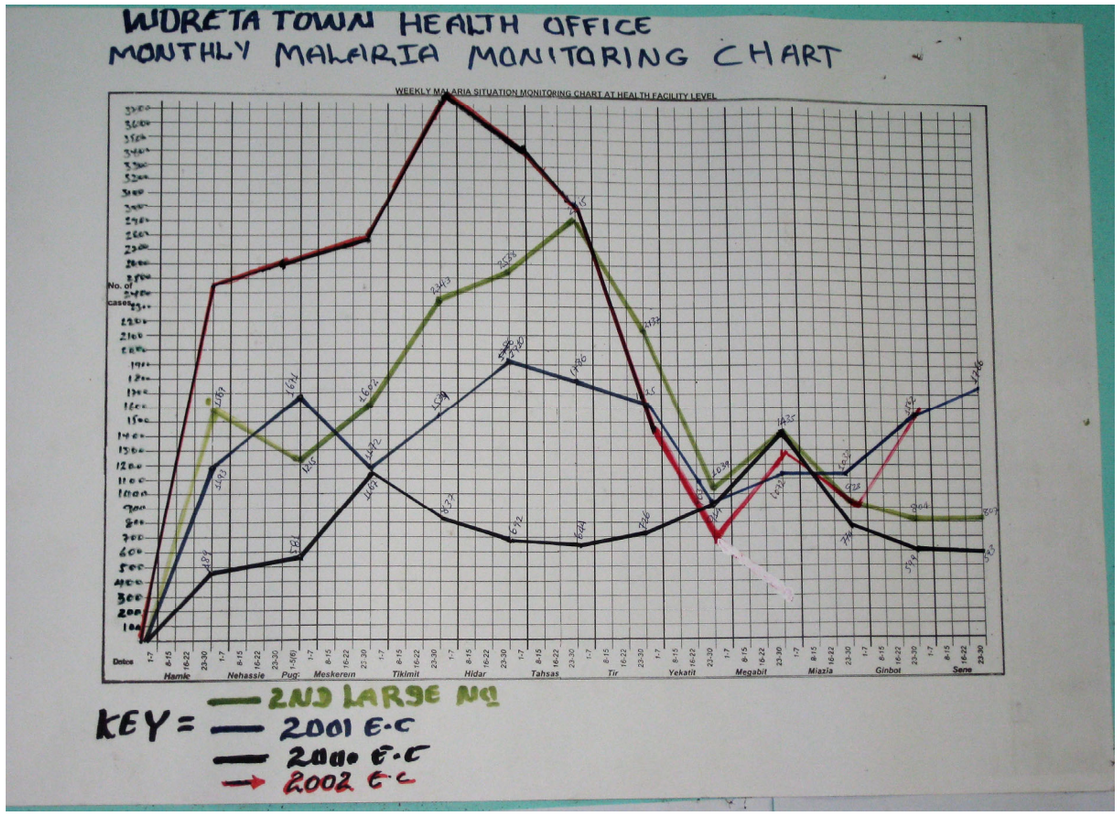
- Web-based GEE tool for data access
- Small, summarized data files for download



	A	B	C	D	E	F	G	H
1	wid	woreda	doy	year	lst_day	lst_night	lst_mean	
11	48	Worebabu	213	2019	29.61383	16.71152	21.1037	
12	121	Abargelie	213	2019	27.8963	18.23581	23.12875	
13	33	Raya Kobo	213	2019	27.22524	16.61713	22.00442	
14	39	Habru	213	2019	25.97256	16.82582	21.31118	
15	38	Guba Lafto	213	2019	19.80301	9.829172	14.95469	
16	34	Gidan	213	2019	20.35273	8.170045	14.57687	
17	40	Woldia Town	213	2019	19.88507	16.2203	18.18974	
18	46	Ambassel	213	2019	21.37925	5.731239	13.36282	
19	37	Delanta	213	2019	24.33438	9.151441	16.71517	
20	140	Kobo Town	213	2019	28.90105	18.03446	23.46776	

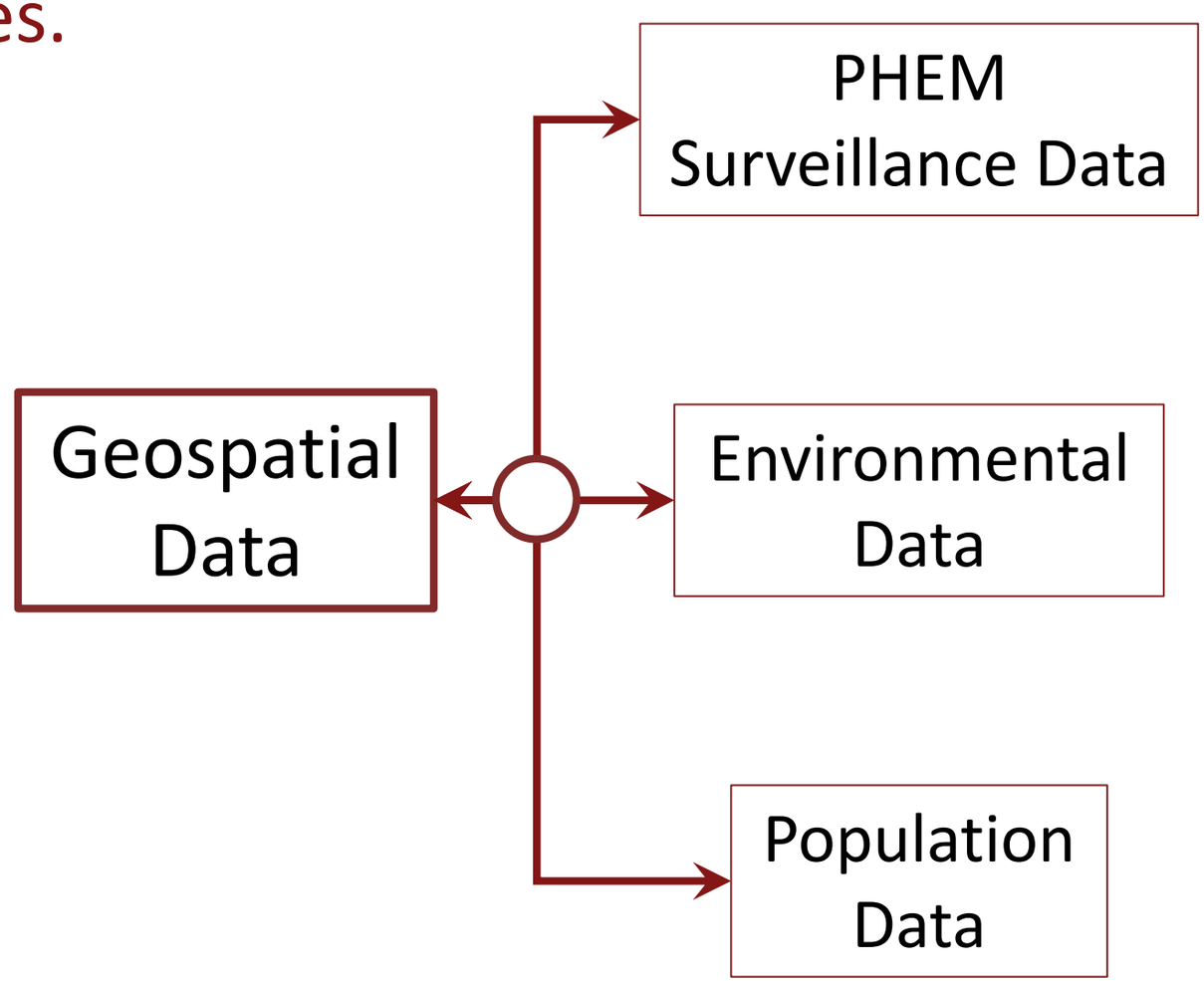
We use **weekly malaria surveillance data** from Ethiopia's Public Health Emergency Management (PHEM) system.

- Woreda-level counts of confirmed malaria cases by species
  - *Plasmodium falciparum*/mixed
  - *Plasmodium vivax*
- Require multiple levels of processing
  - Harmonization with other data sources
  - Screening for suspect values
  - Imputation of missing and suspect data



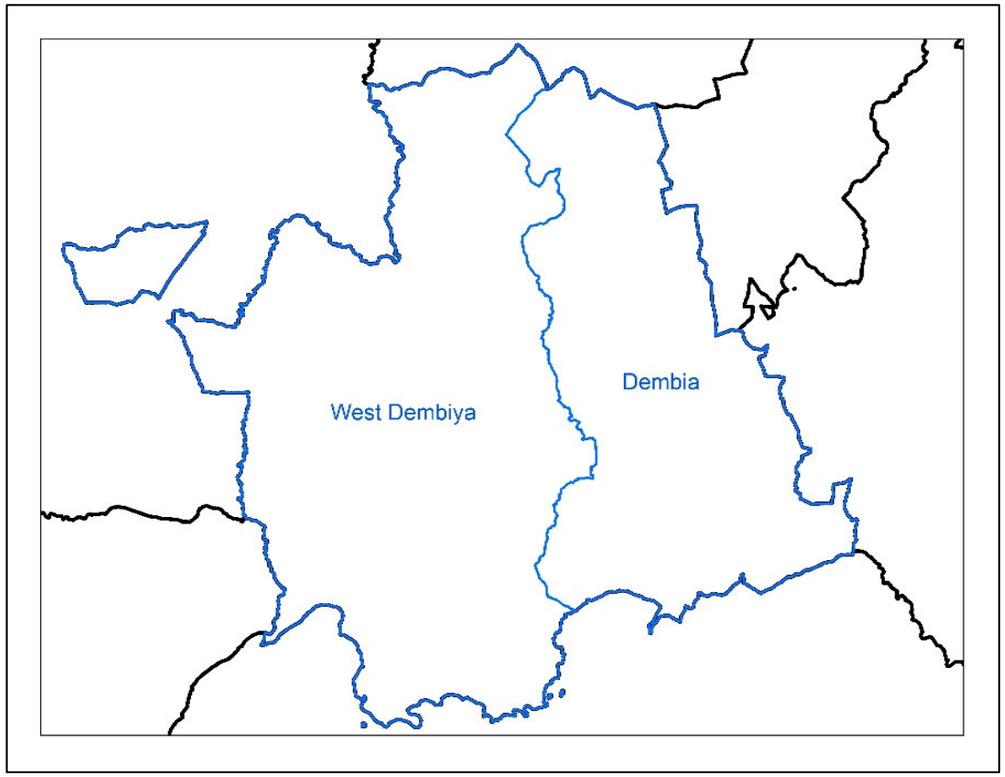
Spatial and temporal harmonization requires crosswalks between the main data sources.

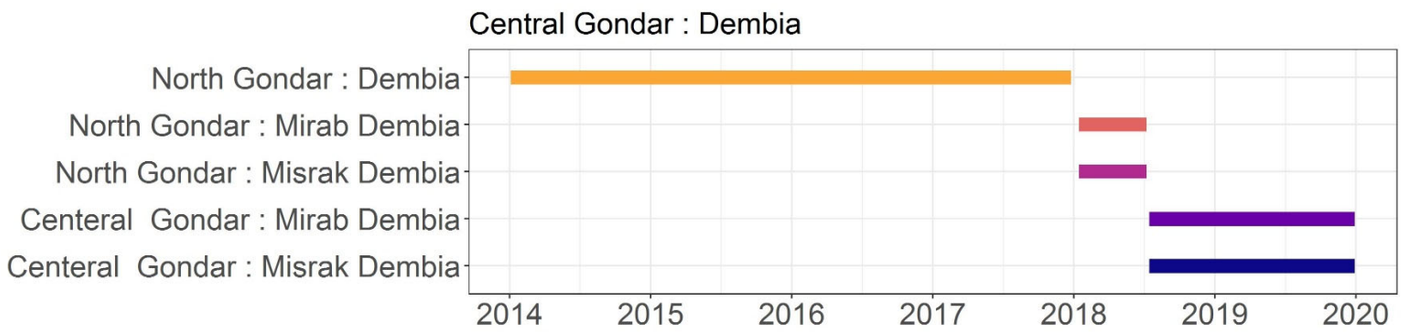
- Goal is to produce consistent time-series for modeling
- Check for missing data
- Reconcile variations in spellings
- Handle geographic changes
  - Woreda boundary changes
  - Zone boundary changes
  - Woreda splits
  - Separation of towns and woredas

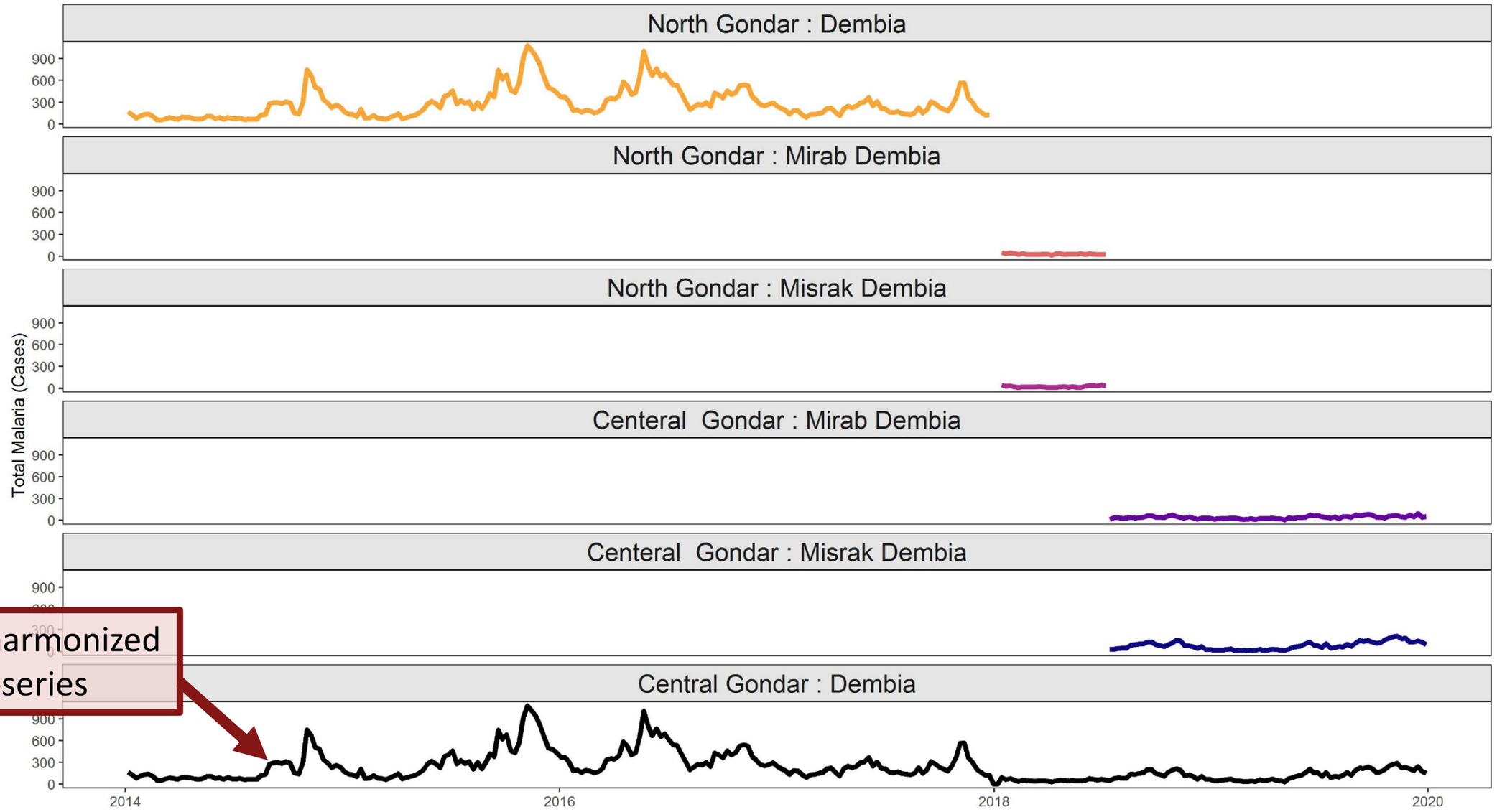


Harmonization Example: Dembia

- Geographic changes
  - Split from “Dembia” into “Misrak Dembia” & “Mirab Dembia”
  - Change of zone from “North Gondar” to “Centeral Gondar”
- Spelling differences
  - “West Dembiya” vs. “West Dembia” vs. “Mirab Dembia”
  - “Central Gondar” vs. “Centeral Gondar”
- All merged into “Dembia” to create a standardized time series for modeling

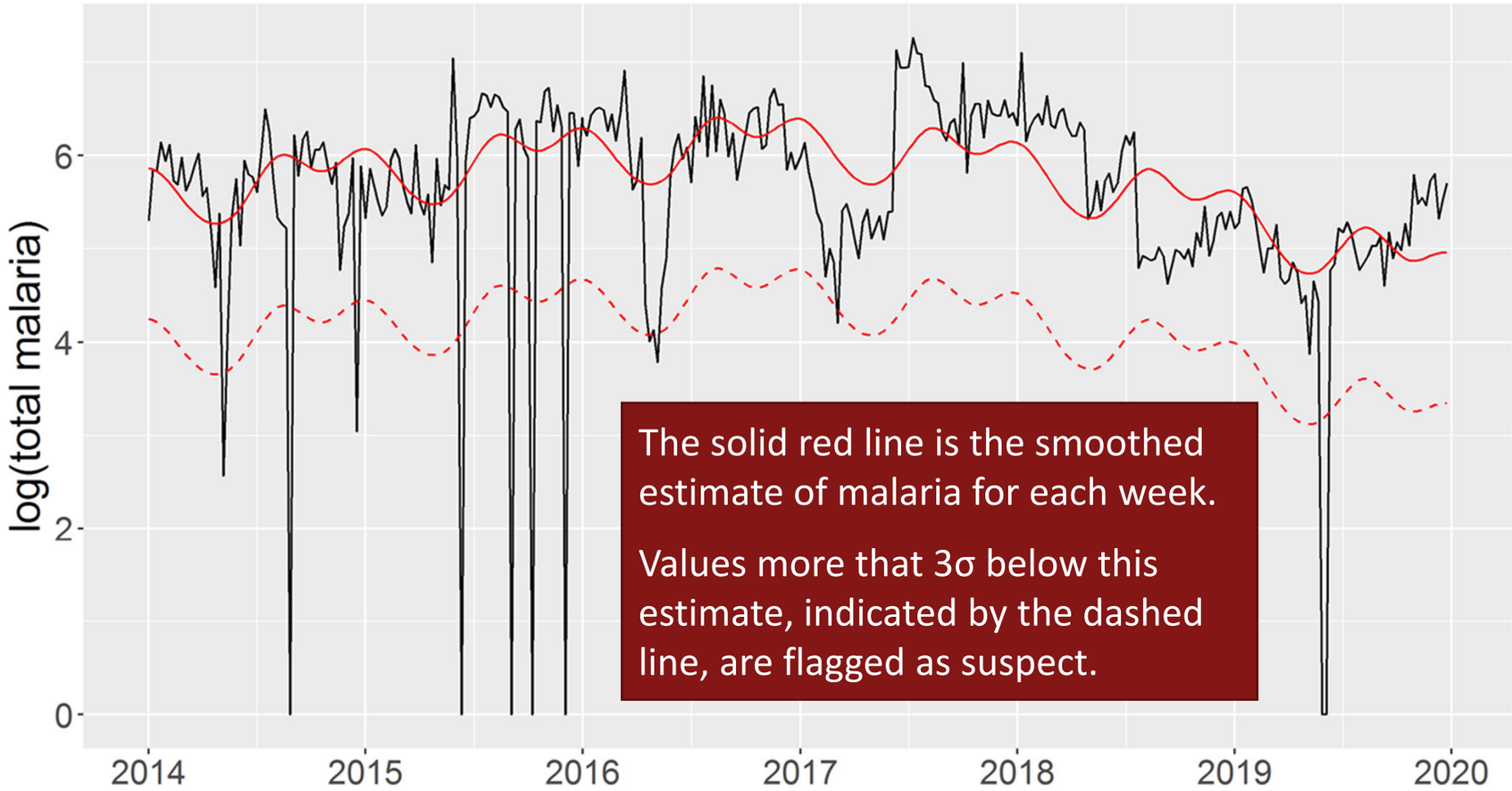




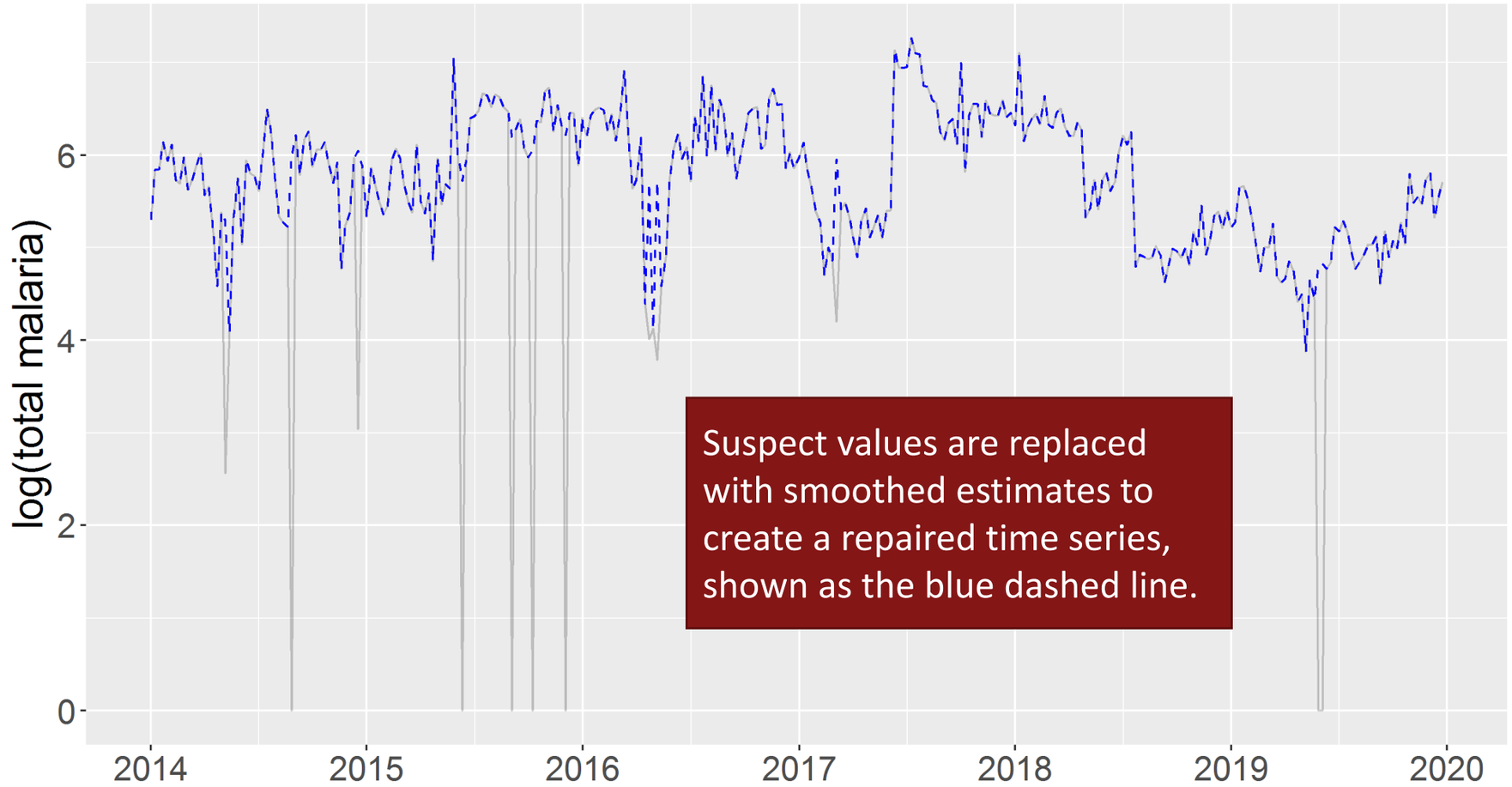


Full harmonized time-series

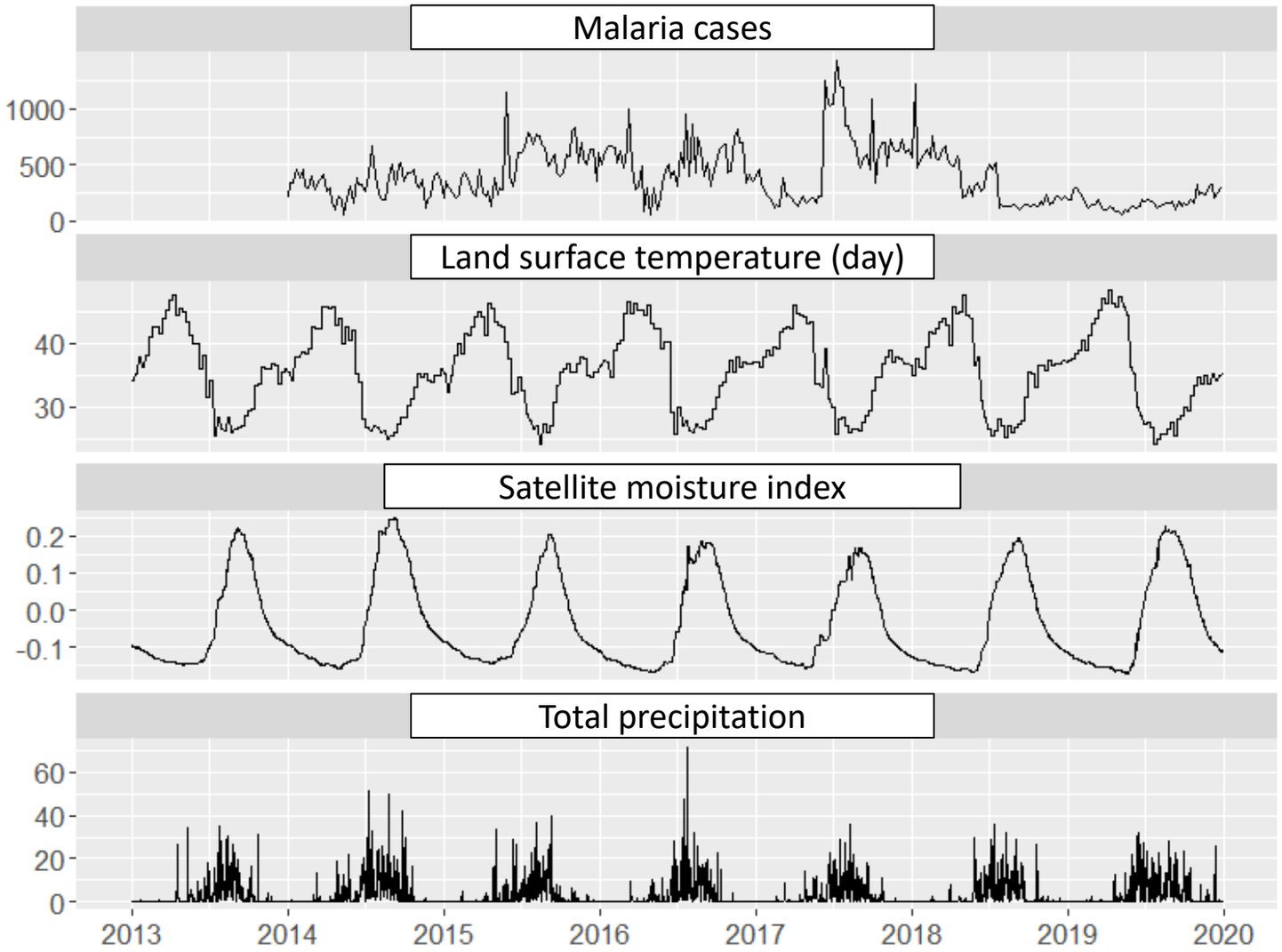
We used a robust linear model to screen out weeks where there were unusually low case counts.



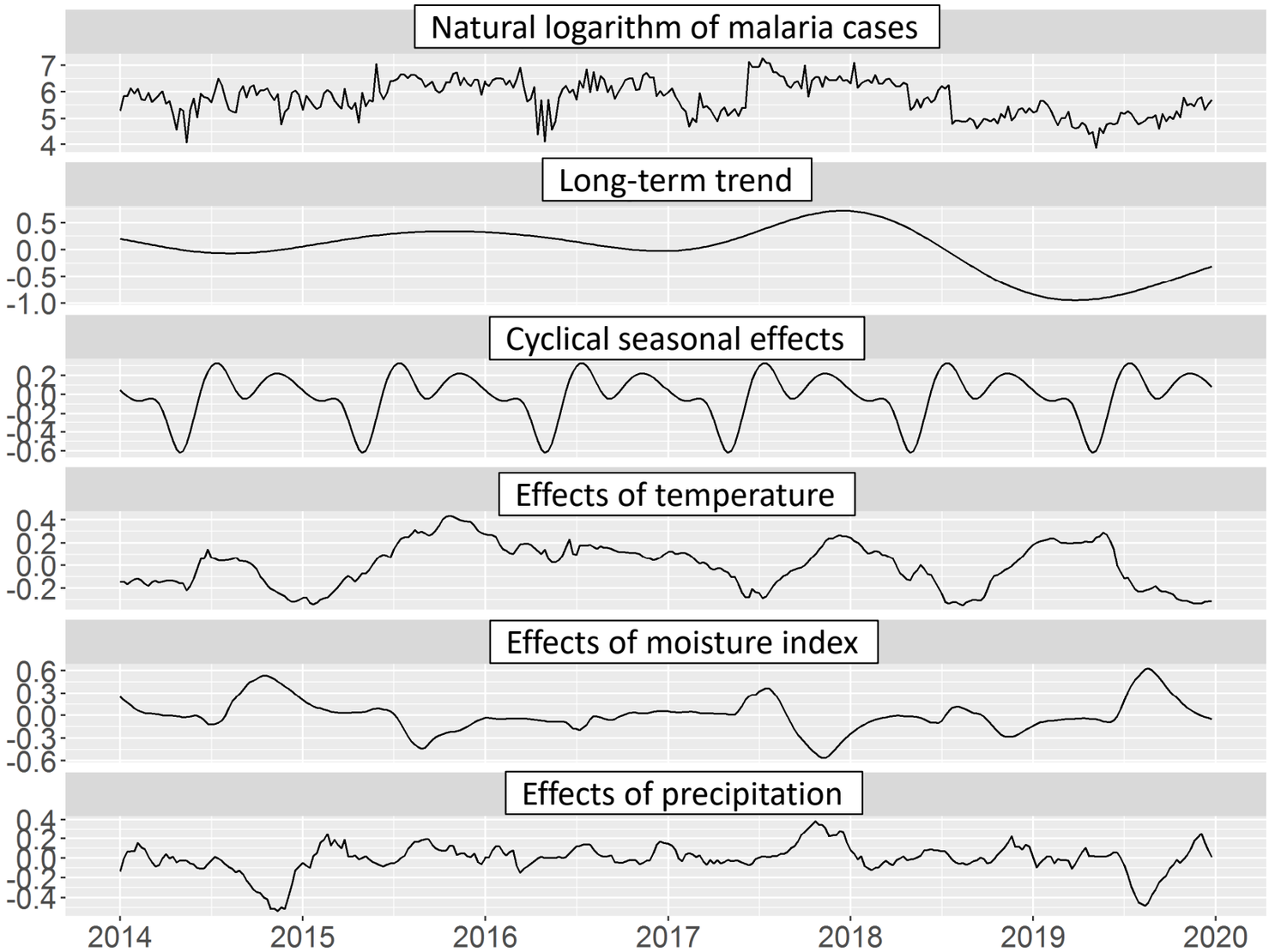
We then replaced suspect and missing data with the smoothed estimates from the robust model.



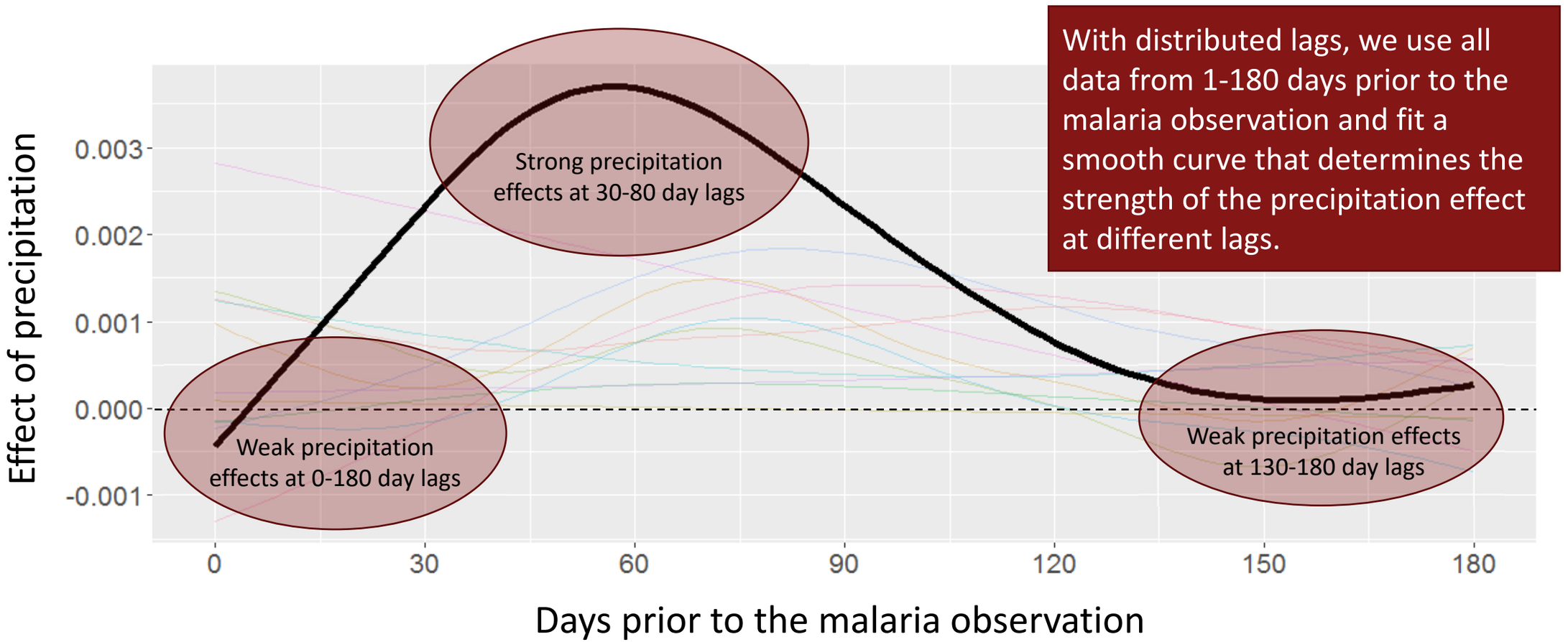
After data harmonization, screening, and imputation, we have a complete time series of data for each woreda.



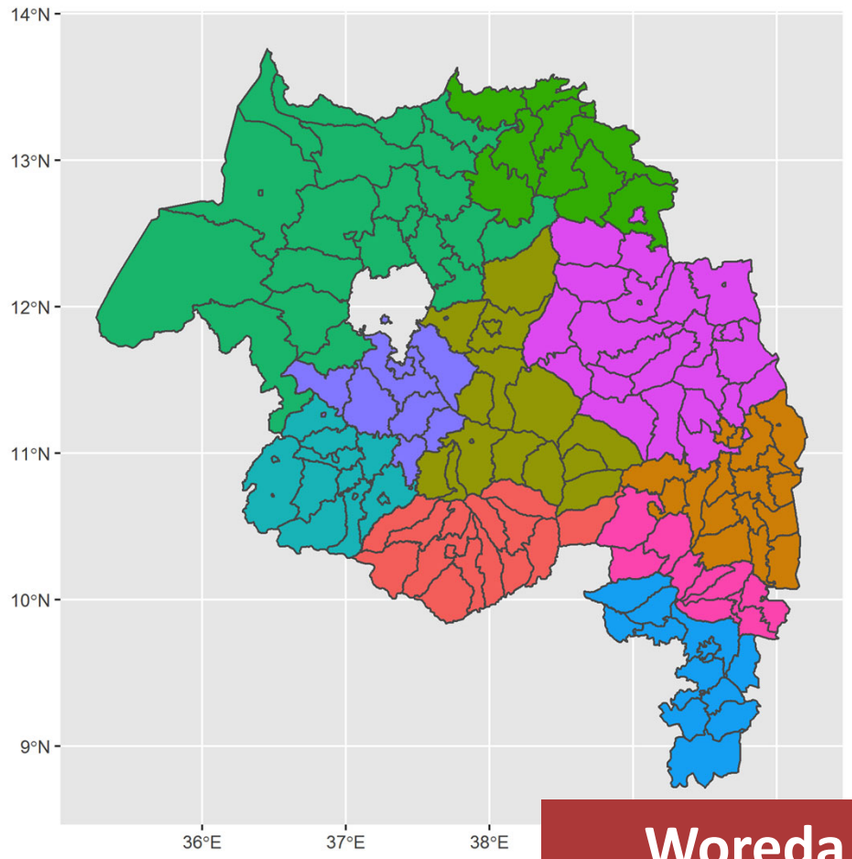
Time series models are used to decompose the malaria data into a set of predictable components.

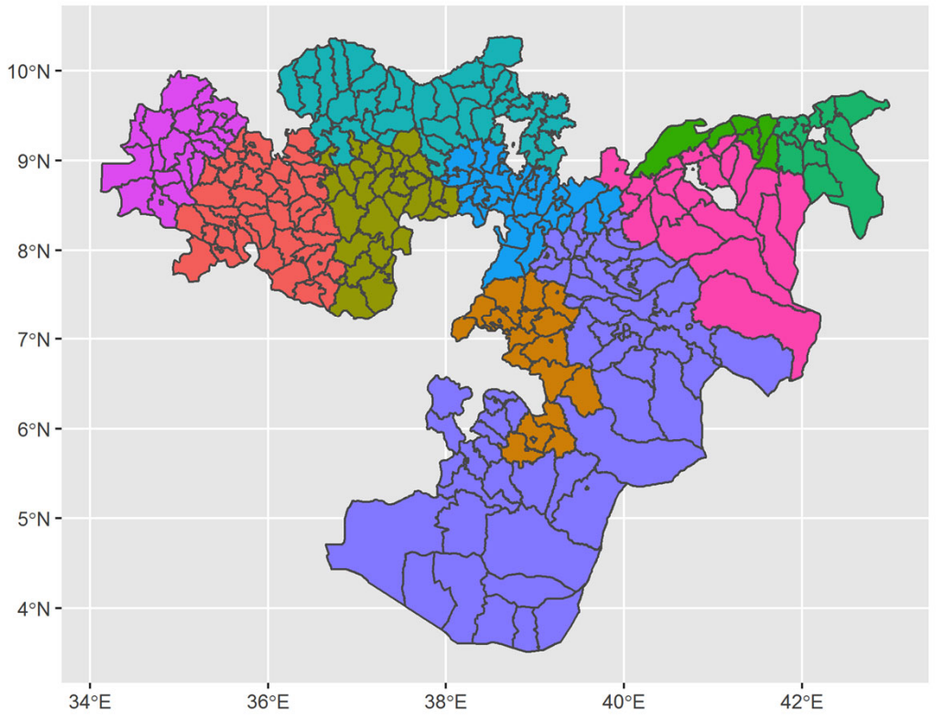


We use *distributed lags* to model the delayed effects of variation in environmental variables on malaria



We have developed novel **optimization methods** based on evolutionary algorithms to identify cluster of woredas for modeling





Woreda clusters are based on their environmental sensitivities.

The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

Break for Questions

Points to Consider:

- What additional data could be used to improve predictions of malaria outbreaks?

APPLICATION

- Forecasts and report generation

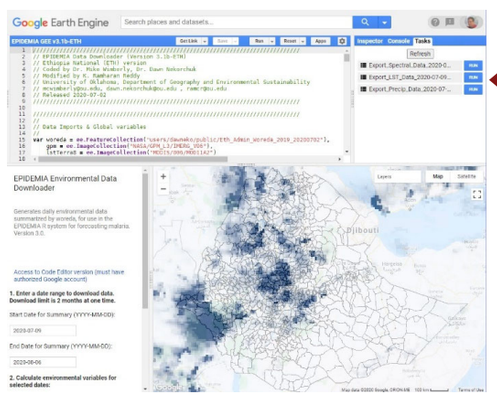
- Validation and report generation

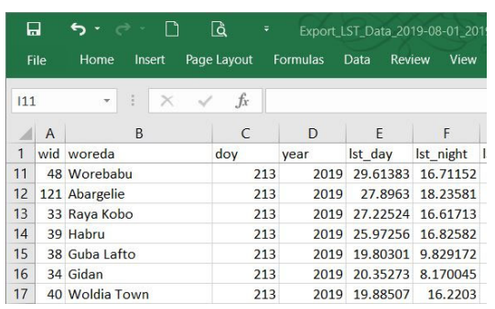
You can generate an EPIDEMIA forecast by obtaining the necessary data and running a customized script for the area of interest.

Local Surveillance Data

1. Obtain epidemiological data

Google Earth Engine (GEE)





	A	B	C	D	E	F
1	wid	woreda	doy	year	lst_day	lst_night
11	48	Worebabu	213	2019	29.61383	16.71152
12	121	Abargelie	213	2019	27.8963	18.23581
13	33	Raya Kobo	213	2019	27.22524	16.61713
14	39	Habru	213	2019	25.97256	16.82582
15	38	Guba Lafto	213	2019	19.80301	9.829172
16	34	Gidan	213	2019	20.35273	8.170045
17	40	Woldia Town	213	2019	19.88507	16.2203

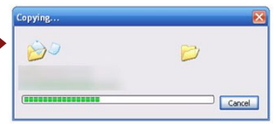
2a. Run GEE script

2b. Download environmental data

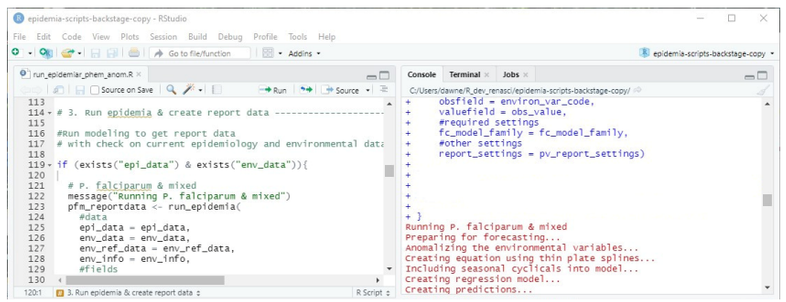
User

Customized R Project

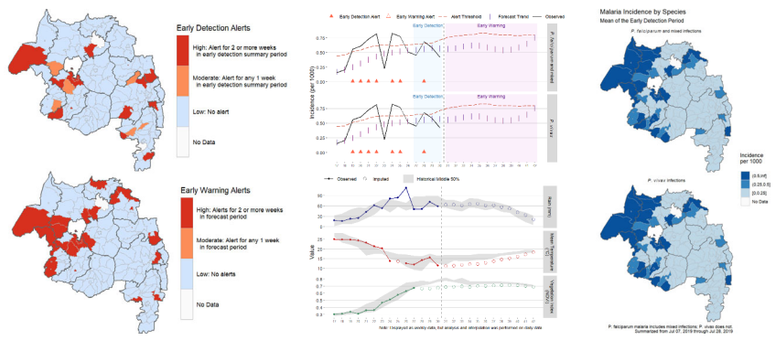
3a. Add data files



3b. Run R script



3c. Forecast report is generated and returned



R Packages epidemiar & clusterapply

Script calls the functions in epidemiar and clusterapply to perform the modeling and forecasting based on the user's data and specified settings

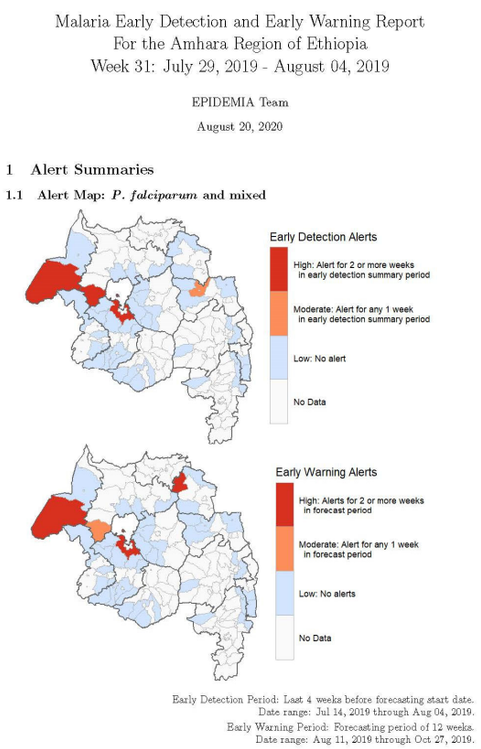
Returns report data to custom scripts for formatting and presentation of results

The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

REPORTS are automatically generated with visualizations of the data and forecasts.

Summary **Alert Maps**



Malaria Early Detection and Early Warning Report
For the Amhara Region of Ethiopia
Week 31: July 29, 2019 - August 04, 2019

EPIDEMIA Team

August 20, 2020

1 Alert Summaries

1.1 Alert Map: *P. falciparum* and mixed

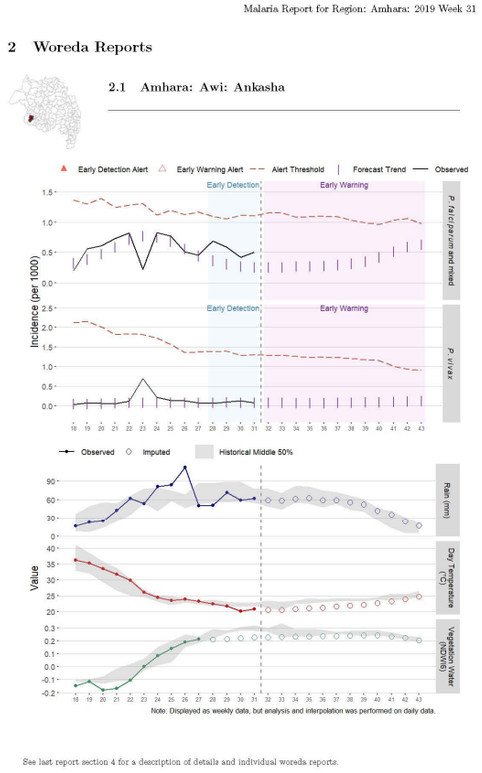
Early Detection Alerts
High: Alert for 2 or more weeks in early detection summary period
Moderate: Alert for any 1 week in early detection summary period
Low: No alert
No Data

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 before forecasting start date.
Date range: Jul 14, 2019 through Aug 04, 2019.
Early Warning Period: Forecasting period of 12 weeks.
Date range: Aug 11, 2019 through Oct 27, 2019.

1

Time series charts for each woreda



Malaria Report for Region: Amhara: 2019 Week 31

2 Woreda Reports

2.1 Amhara: Awi: Ankasha

Early Detection Alert | Early Warning Alert | Alert Threshold | Forecast Trend | Observed

Early Detection | Early Warning

Incidence (per 1000)

P. falciparum and mixed

P. vivax

Observed | Imputed | Historical Middle 50%

Value

Rain (mm)

Day Temperature (°C)

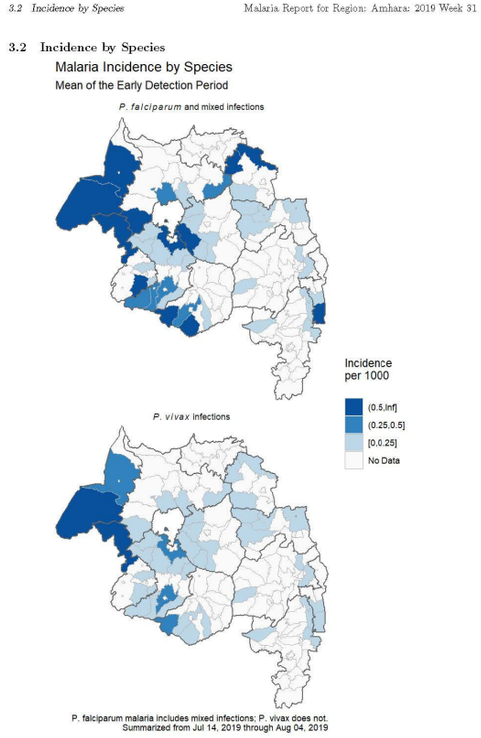
Vegetation Water (NDWI6)

Note: Displayed as weekly data, but analysis and interpolation was performed on daily data.

See last report section 4 for a description of details and individual woreda reports.

6 Report compiled on 20 August 2020

Maps of recent malaria incidence



3.2 *Incidence by Species* Malaria Report for Region: Amhara: 2019 Week 31

3.2 Incidence by Species

Malaria Incidence by Species
Mean of the Early Detection Period

P. falciparum and mixed infections

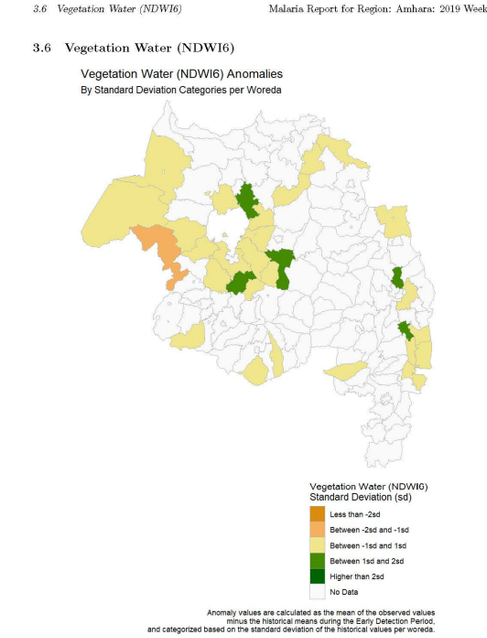
P. vivax infections

Incidence per 1000
(0.5,Inf]
(0.25,0.5]
[0.0,25]
No Data

P. falciparum malaria includes mixed infections; P. vivax does not.
Summarized from Jul 14, 2019 through Aug 04, 2019

54 Report compiled on 20 August 2020

Maps of recent climate anomalies



3.6 *Vegetation Water (NDWI6)* Malaria Report for Region: Amhara: 2019 Week 31

3.6 Vegetation Water (NDWI6)

Vegetation Water (NDWI6) Anomalies
By Standard Deviation Categories per Woreda

Vegetation Water (NDWI6) Standard Deviation (sd)
Less than -2sd
Between -2sd and -1sd
Between -1sd and 1sd
Between 1sd and 2sd
Higher than 2sd
No Data

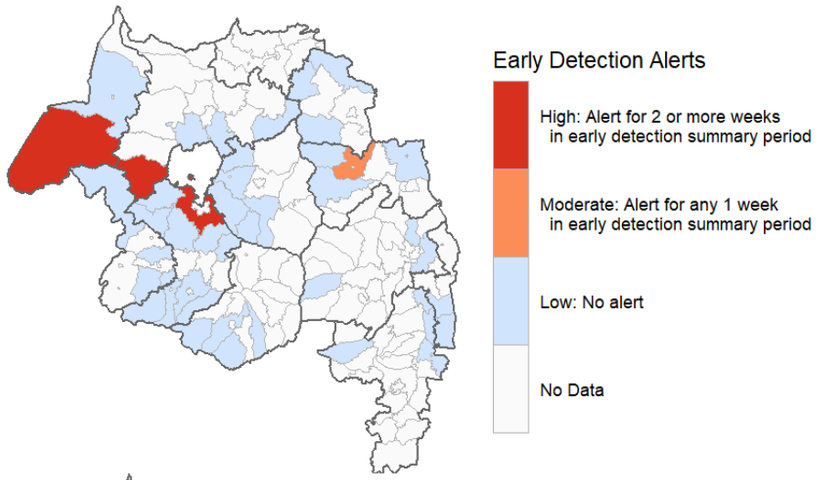
Anomaly values are calculated as the mean of the observed values minus the historical means during the Early Detection Period, and categorized based on the standard deviation of the historical values per woreda.

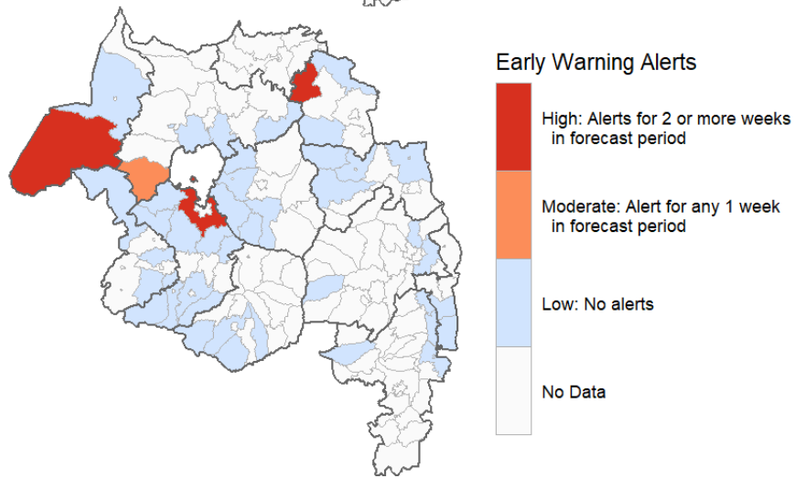
The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

Alert Maps highlight woredas where there is evidence of an outbreak.

- Early Detection: observed cases in recent weeks
- Early Warning: environmental predictions of cases in upcoming weeks
- Separate alerts for *falciparum* and *vivax* malaria





1.2 Alert Listing: P. falciparum and mixed Malaria Report for Region: Amhara: 2019 Week 31

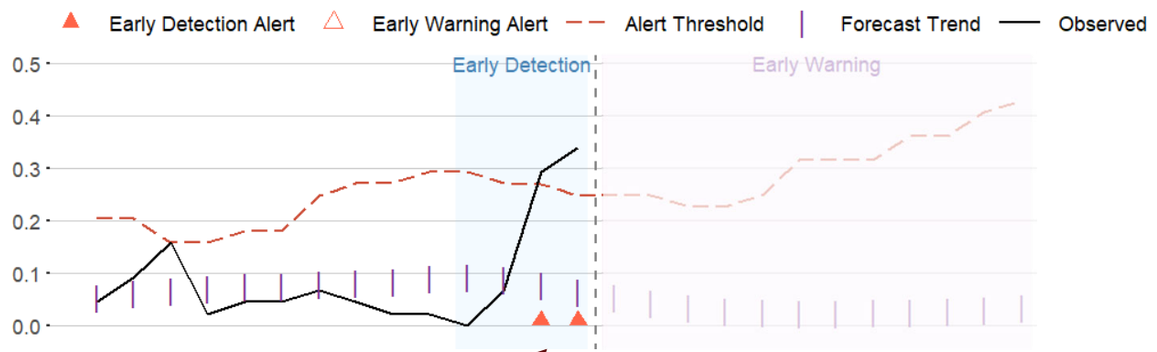
1.2 Alert Listing: *P. falciparum* and mixed

1.2.1 Amhara

	Zone	Woreda	Early Detection	Early Warning	Both
1	Central Gondar	Alfa	High	Medium	Yes
2	North Wello	Lasta	Medium	Low	-
3	Wag Hamra	Sahila	Low	High	-
4	West Gojam	Bahirdar Zuria	High	High	Yes
5	West Gondar	Quara	High	High	Yes

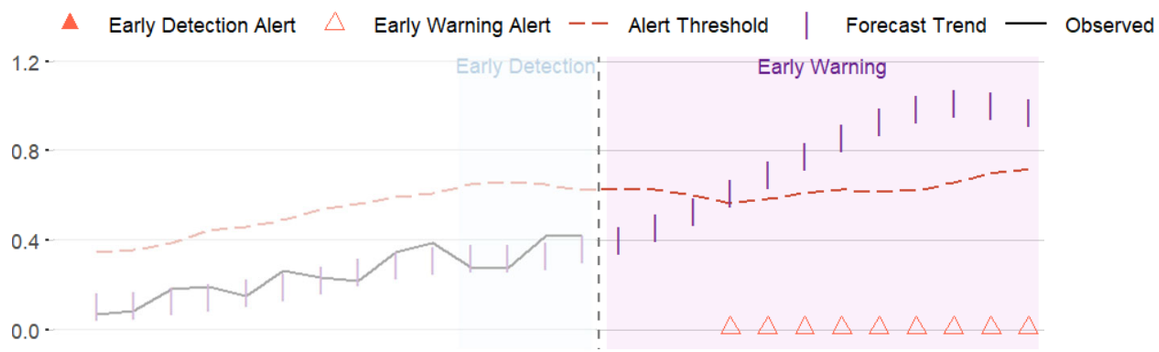
Alert Thresholds can detect emerging outbreaks and **early warning of future outbreaks.**

Early Detection



Early Detection Alerts triggered when the observed value is higher than the threshold

Early Warning

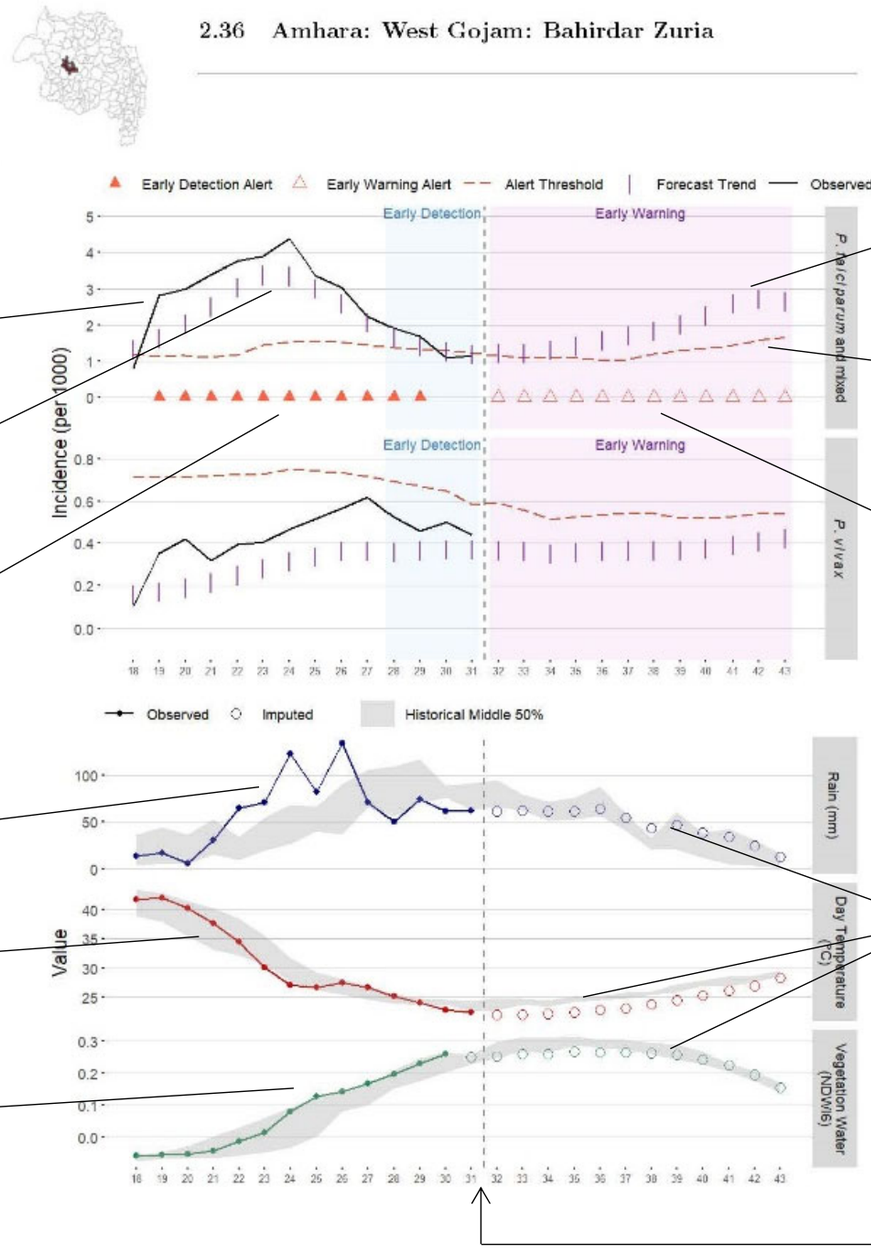


Early Warning Alerts triggered when the forecasted values are higher than the threshold

Thresholds are calculated using Farrington Improved algorithm:

- Based on historical case counts and patterns
- Allows adjustments for trend, seasonality, prior outbreaks and more

Each woreda report provides **observations over the 14 preceding weeks and forecasts for the 12 upcoming weeks**.



Observed malaria incidence (solid lines)

Previous 1-week ahead forecasts (vertical bars)

Early detection alerts (solid triangles)

Future malaria forecasts (vertical bars)

Outbreak threshold (dashed line)

Early warning alerts (hollow triangles)

Precipitation (blue line)

Temperature (red line)

Satellite moisture index (green line)

Extrapolations of recent climate trends

Date of forecast

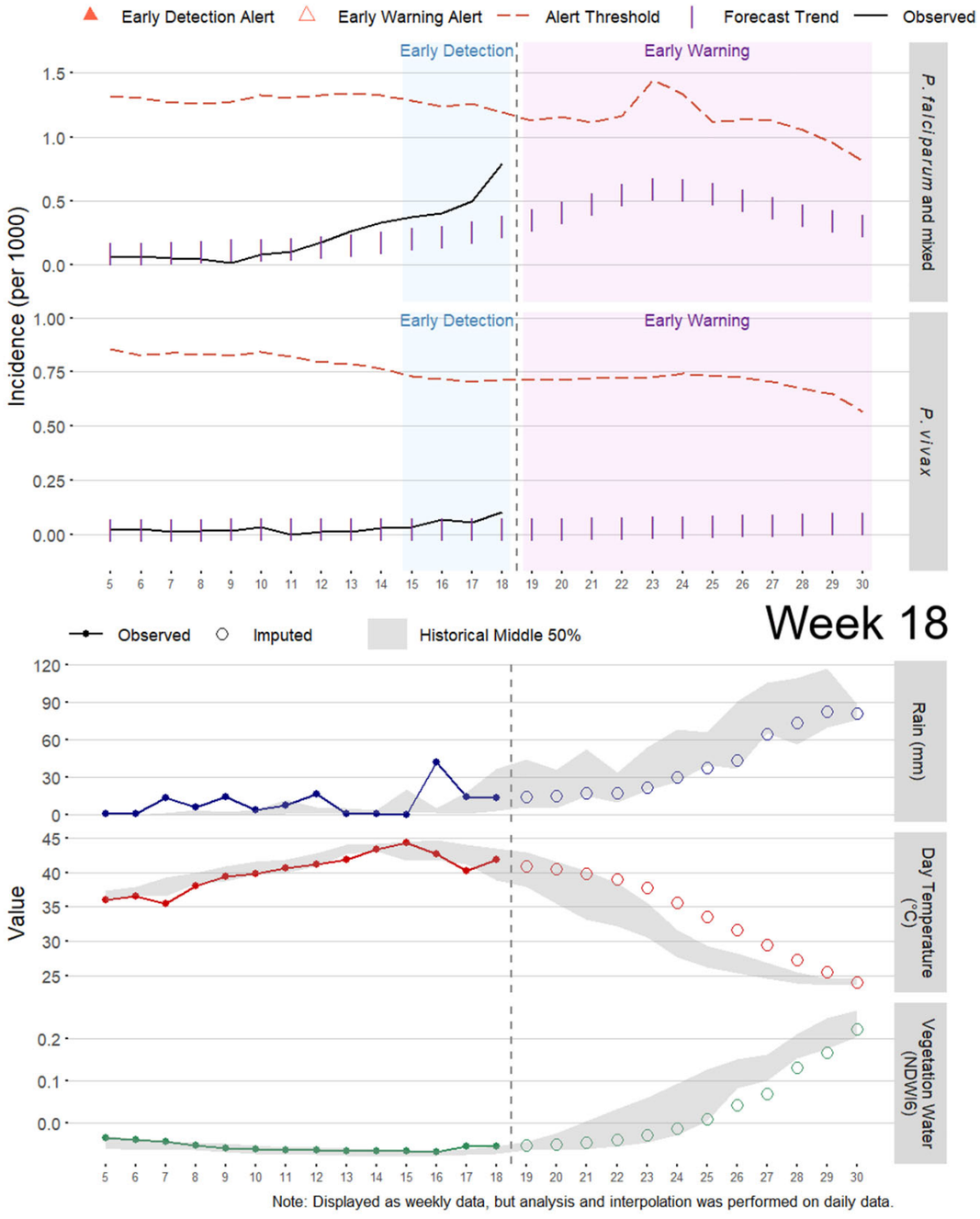
The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

This video shows how forecasts changed from weeks 18-52 of 2019 in Bahir Dar Zuria.

Early detection of an outbreak in weeks 19-28 and

Early warning of an outbreak in weeks 40-49.

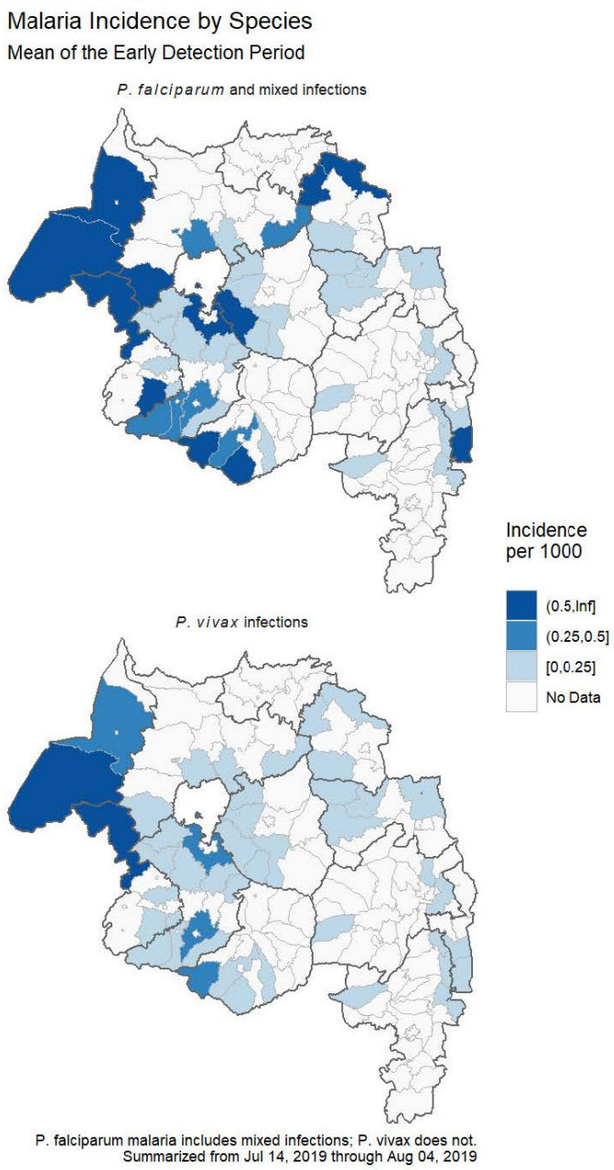


The UNIVERSITY of OKLAHOMA

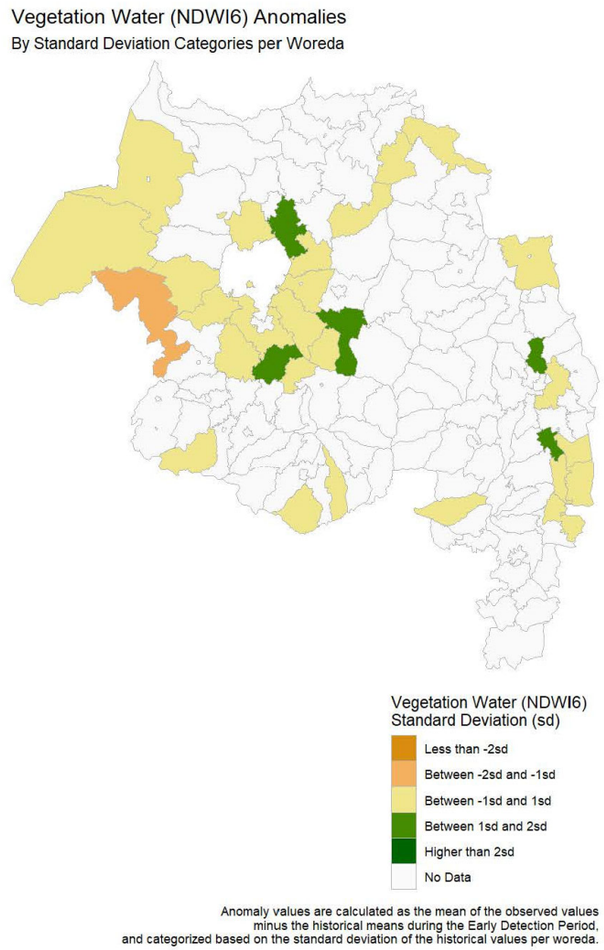
USAID FROM THE AMERICAN PEOPLE

Forecasting reports also include a **variety of maps that describe recent patterns of malaria incidence and climate anomalies**.

3.2 Incidence by Species



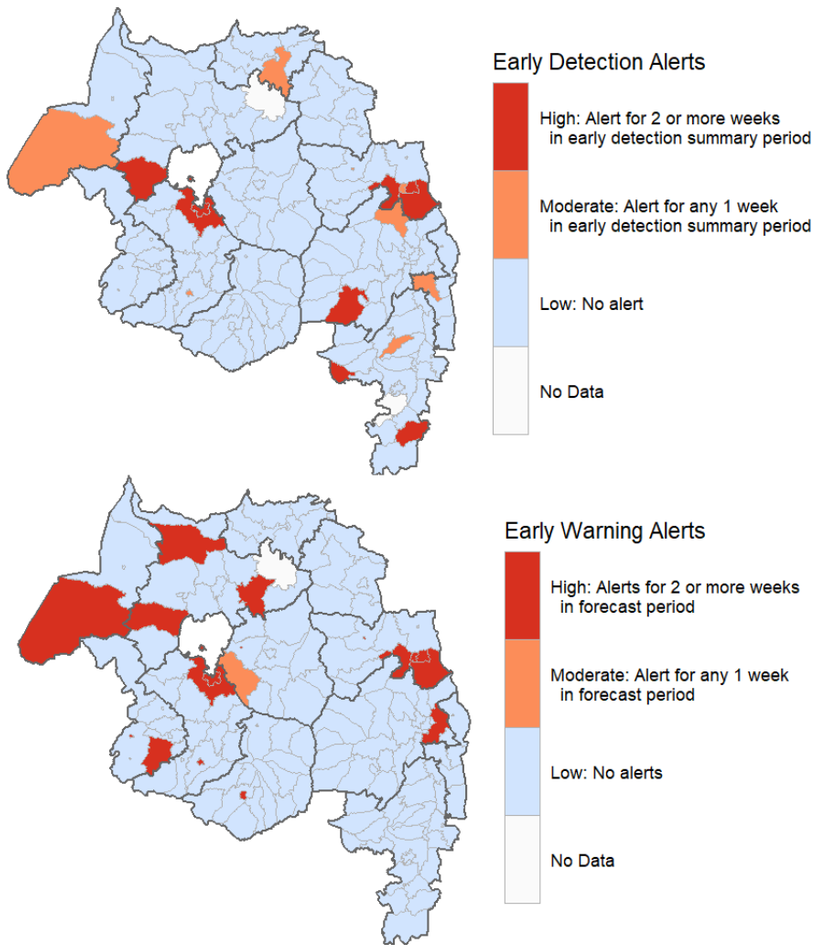
3.6 Vegetation Water (NDWI6)

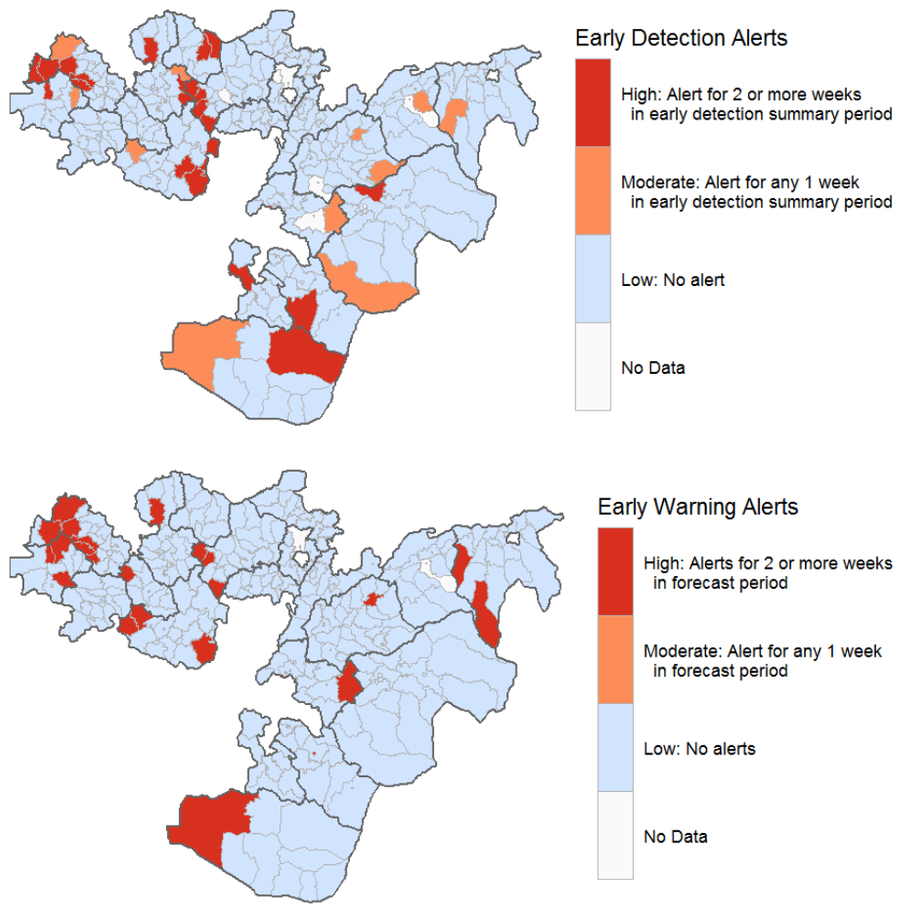


The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

The current version of EPIDEMIA can use the PHEM dataset to generate **region-wide forecasts anywhere in Ethiopia**.

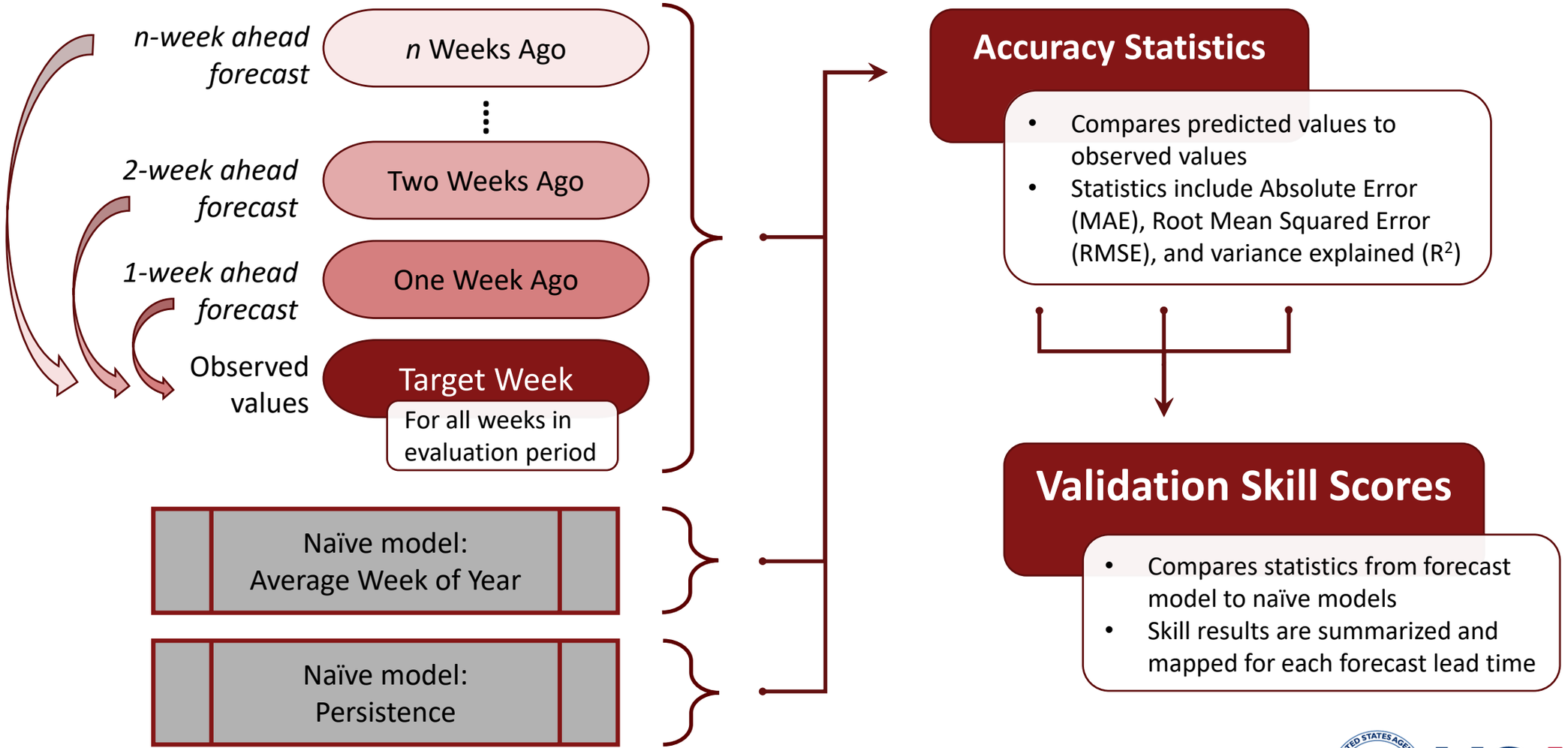




The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

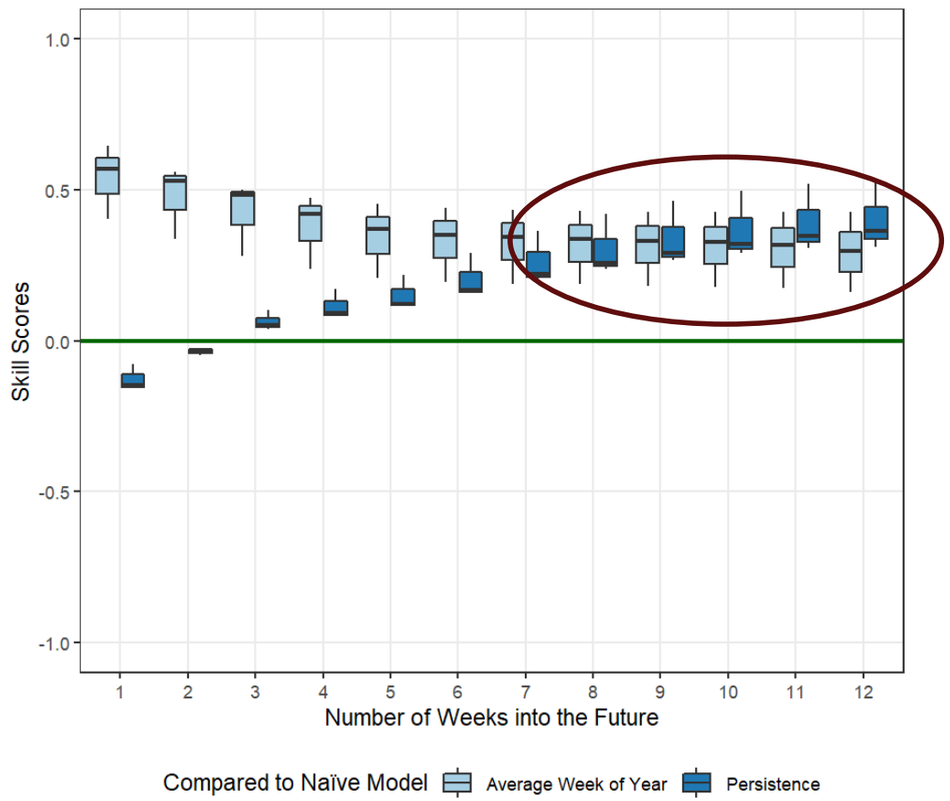
Validation is based on historical data and compares forecasts over a range of historical dates to subsequent observations.



Skill scores show whether the model improves accuracy compared to naïve baseline predictions.

- Summary charts display the distribution of skill scores for forecasts made over a range of lead times.
- Positive skill scores mean that the environmental data are improving the predictions of malaria.
- EPIDEMIA forecasts generally have high predictive skill out to 12 weeks in the future.

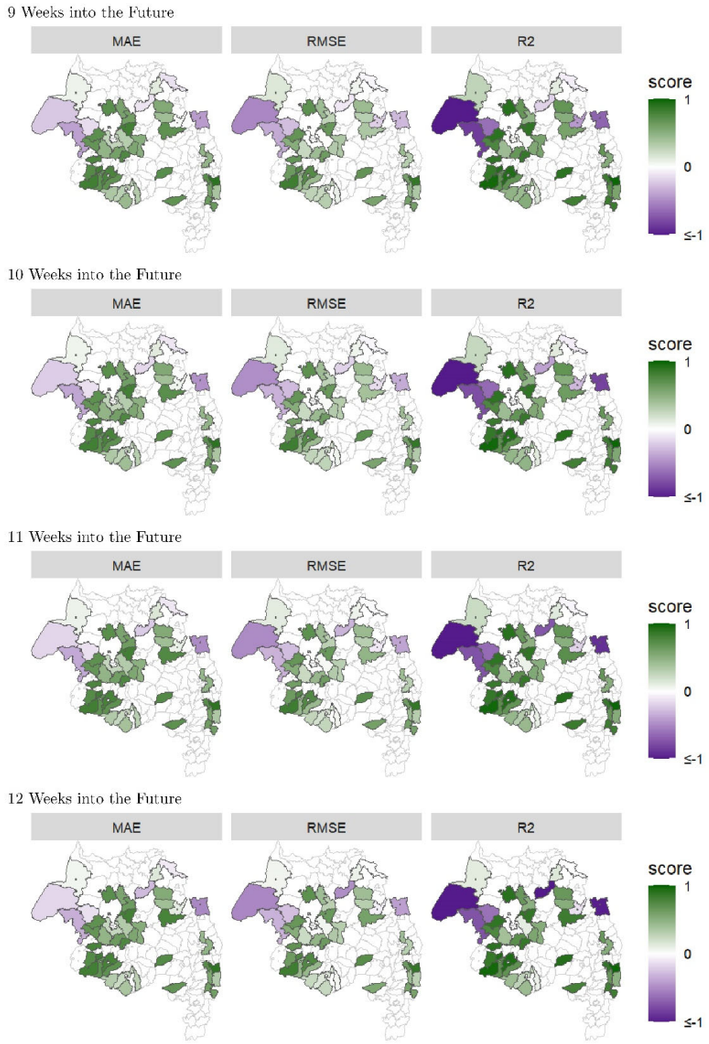
Skill score summary chart for Amhara (47 pilot woredas) from Jan. 1 2018-Dec. 31 2019



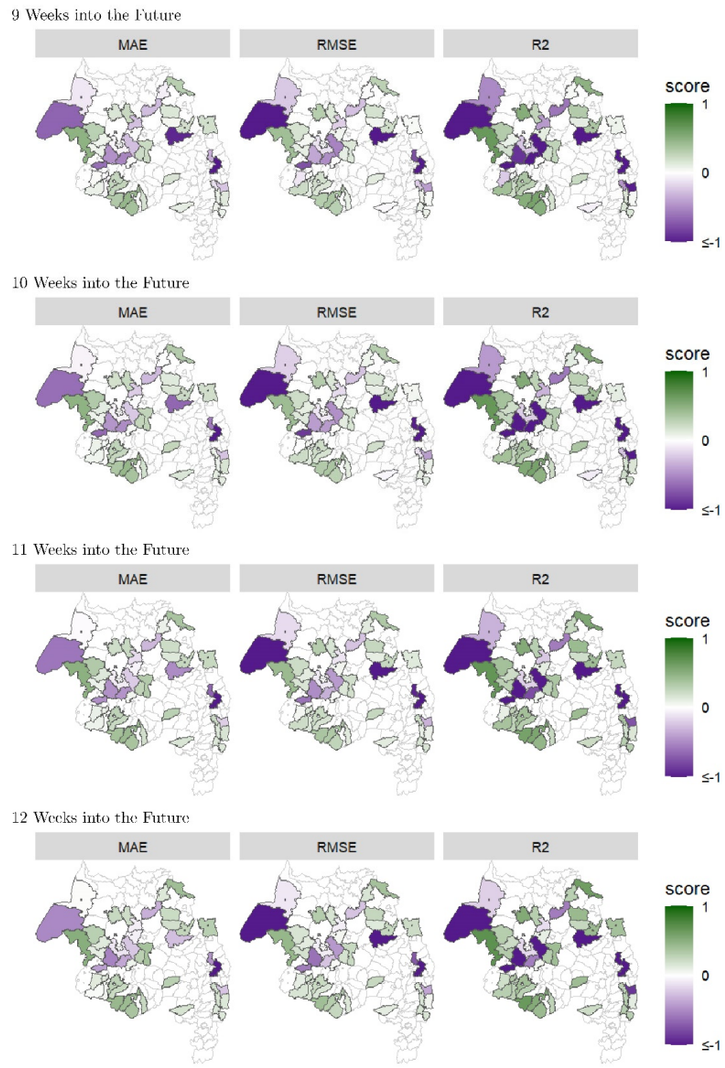
Maps of the validation results **identify specific woredas where the forecasts consistently have high or low skill.**

SKILL = Model Accuracy

Skill versus persistence



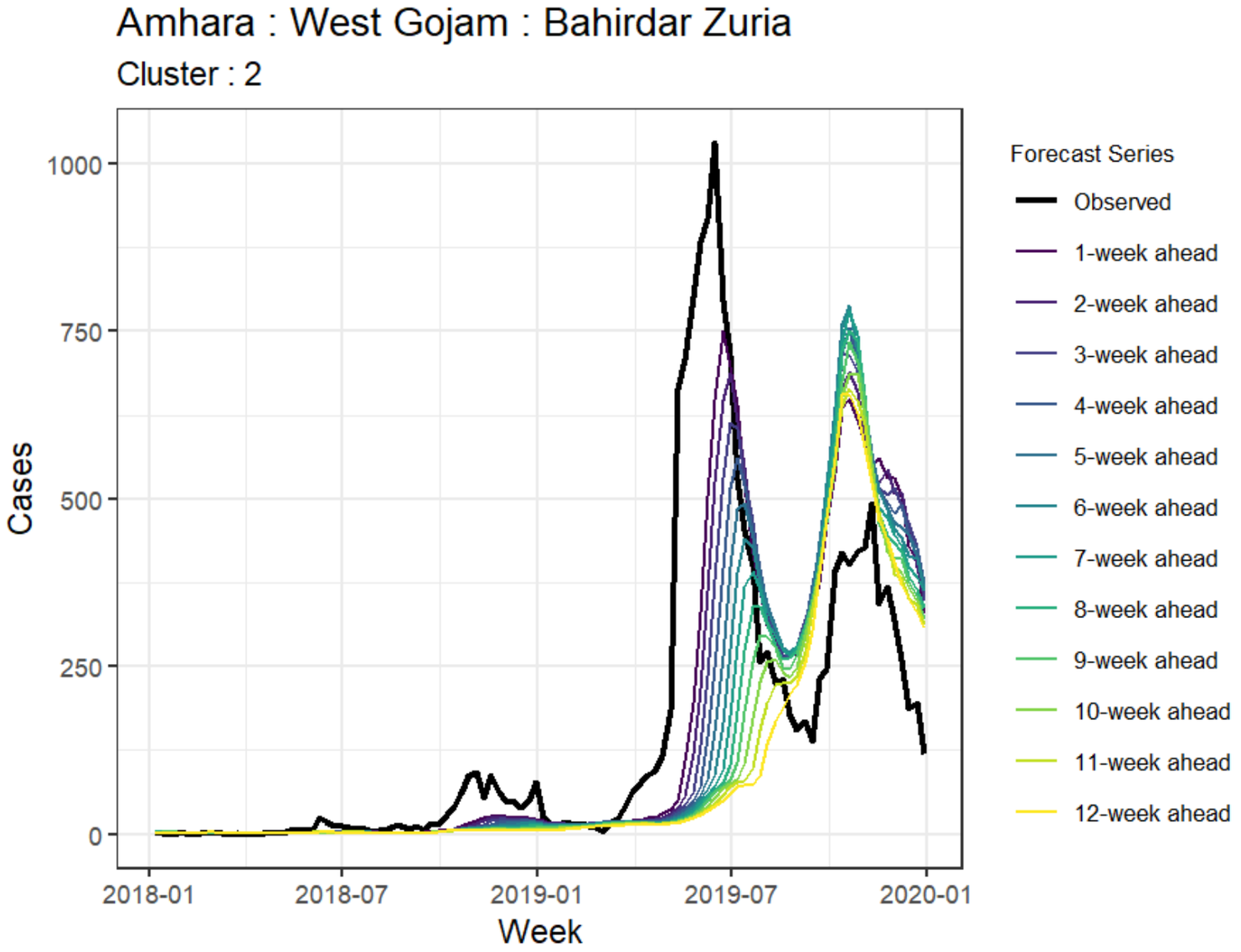
Skill versus average week of year



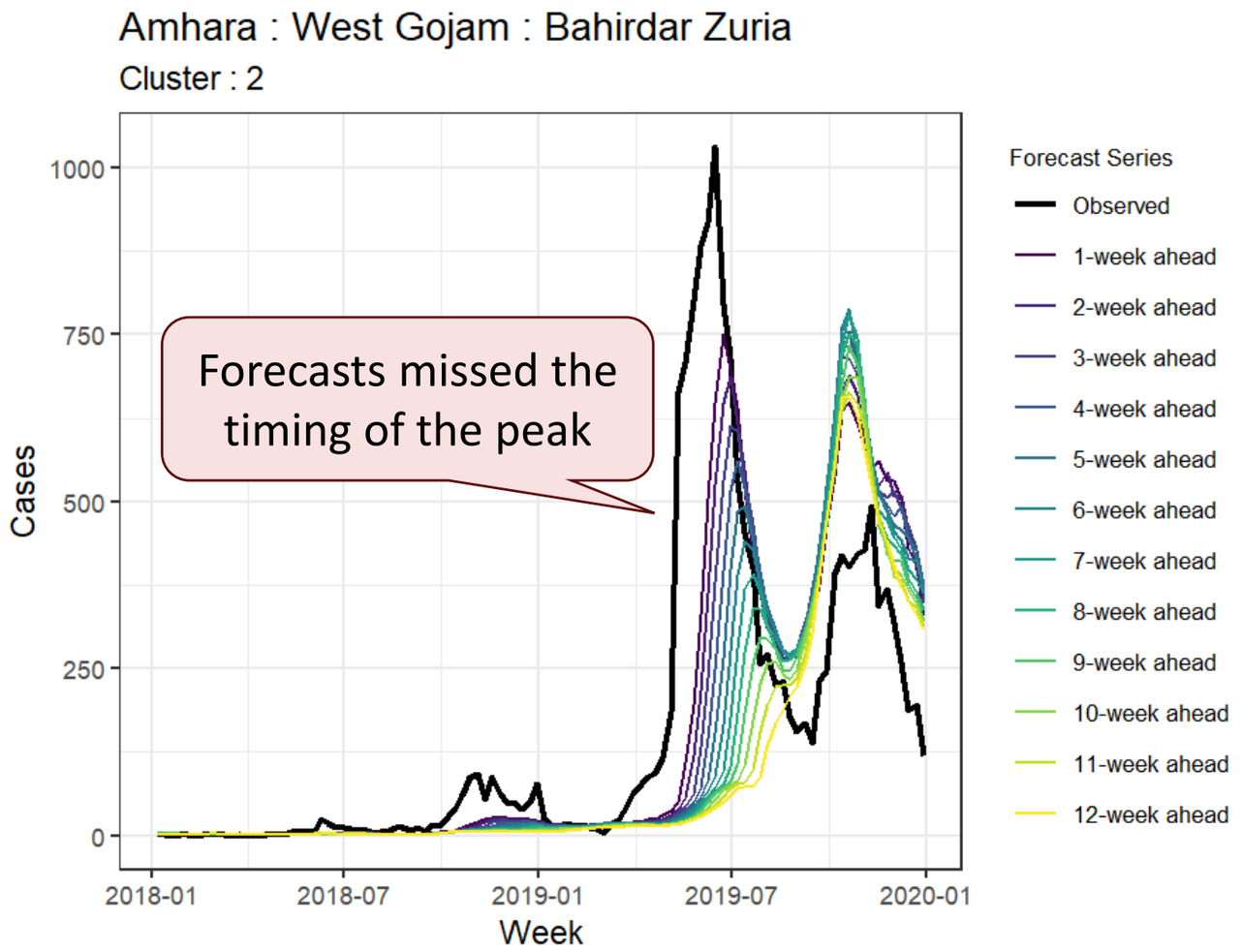
The UNIVERSITY of OKLAHOMA

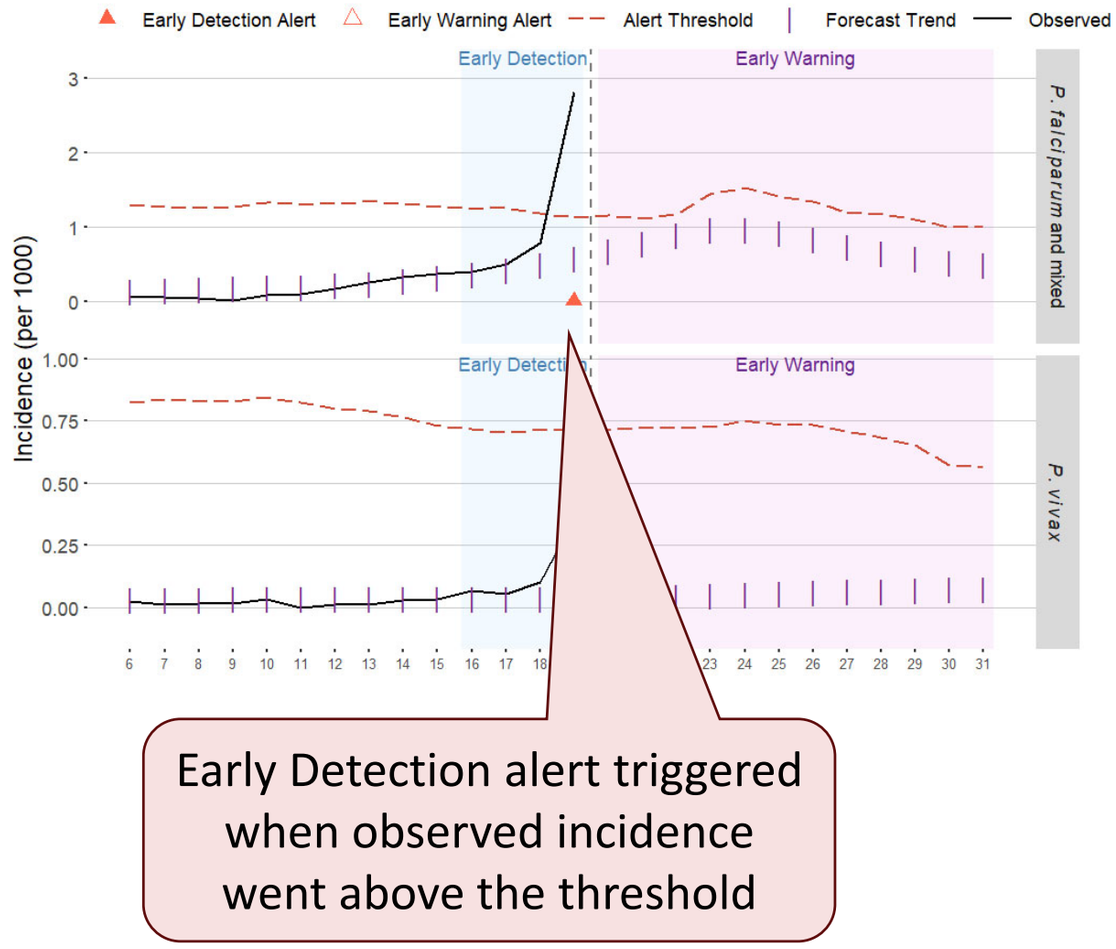
USAID FROM THE AMERICAN PEOPLE

Week-ahead forecast charts show predictions made over a range of weeks before each observation, providing more insights into the timing of the predictions.



Early warning forecasts missed the timing of this outbreak, but an early detection alert was triggered at the beginning of the outbreak.

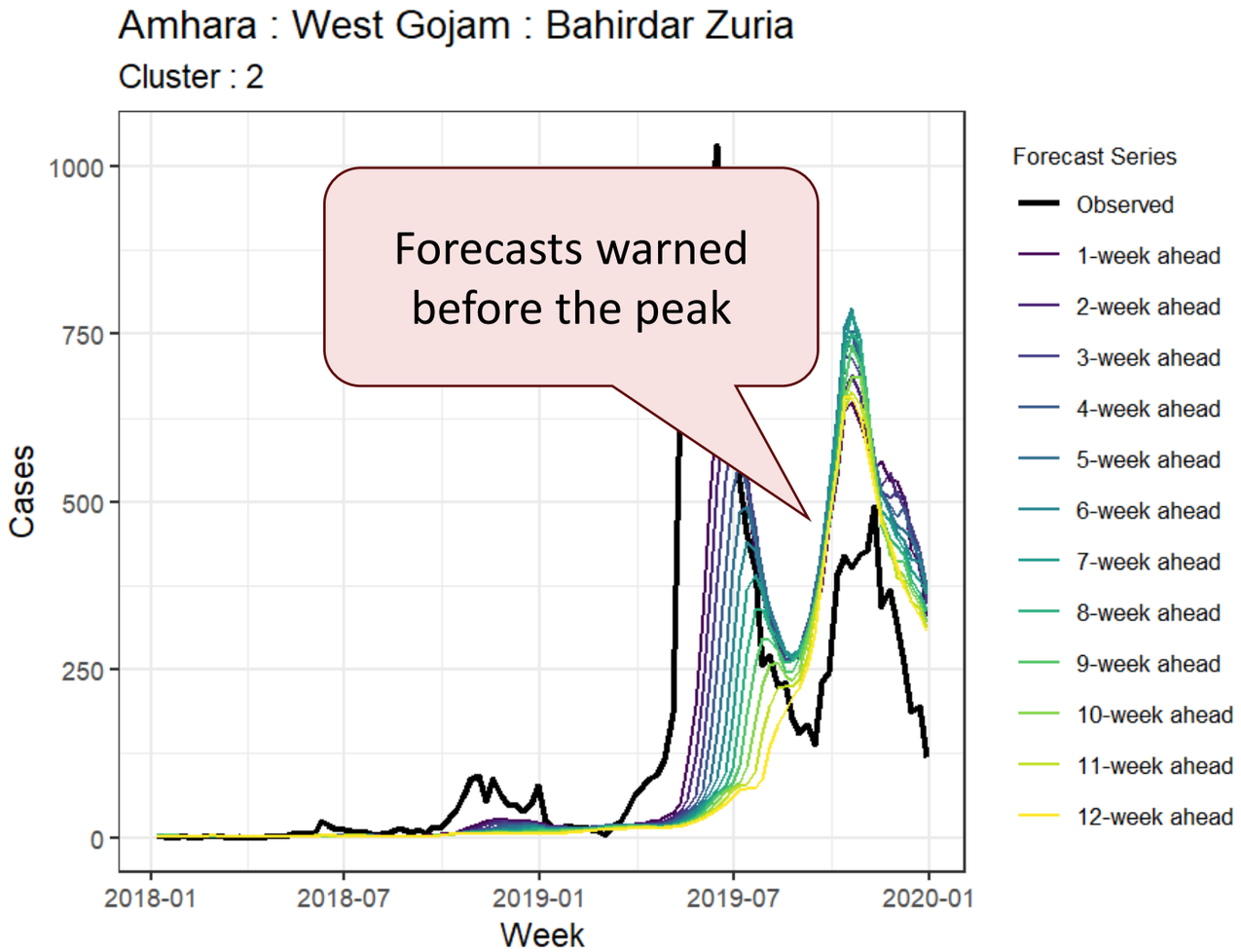


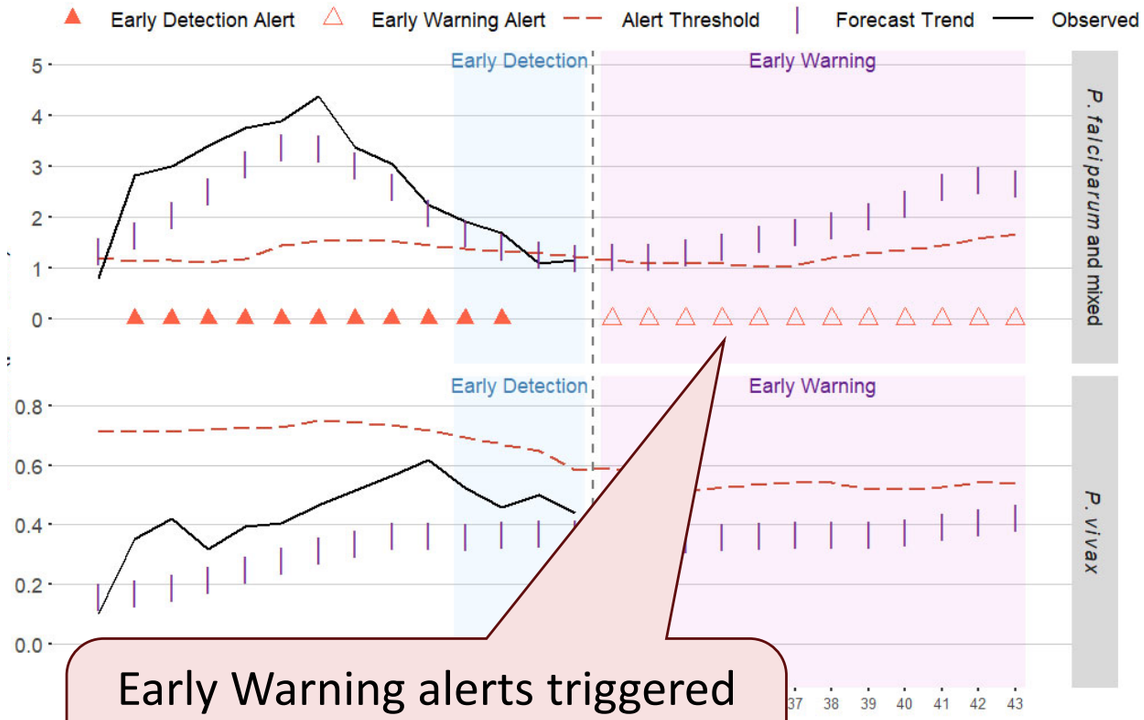


The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

Early warning forecasts predicted this outbreak almost three months before the peak.





The UNIVERSITY of OKLAHOMA

USAID FROM THE AMERICAN PEOPLE

In summary, the EPIDEMIA system currently offers a mature software environment for generating malaria forecasts anywhere in Ethiopia.

- Free and open source software can be applied over multiple regions.
- Automates time-consuming data processing tasks.
- Uses freely available environmental data.
- Incorporates cutting-edge predictive analytics that extract climate-malaria relationships from big datasets.
- Generates detailed forecasting reports with charts and maps.
- Facilitates accuracy assessment of the forecasts.
- Implementation will require more engagement with Ethiopian stakeholders as users and co-developers.

End of talk – Questions?

Points to consider:

- What institutions should be involved in the use and continued development of malaria early warning systems in Ethiopia?
- What training and capacity building activities are required to facilitate implementation?
- What further updates should be made to EPIDEMIA to improve its suitability for implementing malaria early warning in Ethiopia?