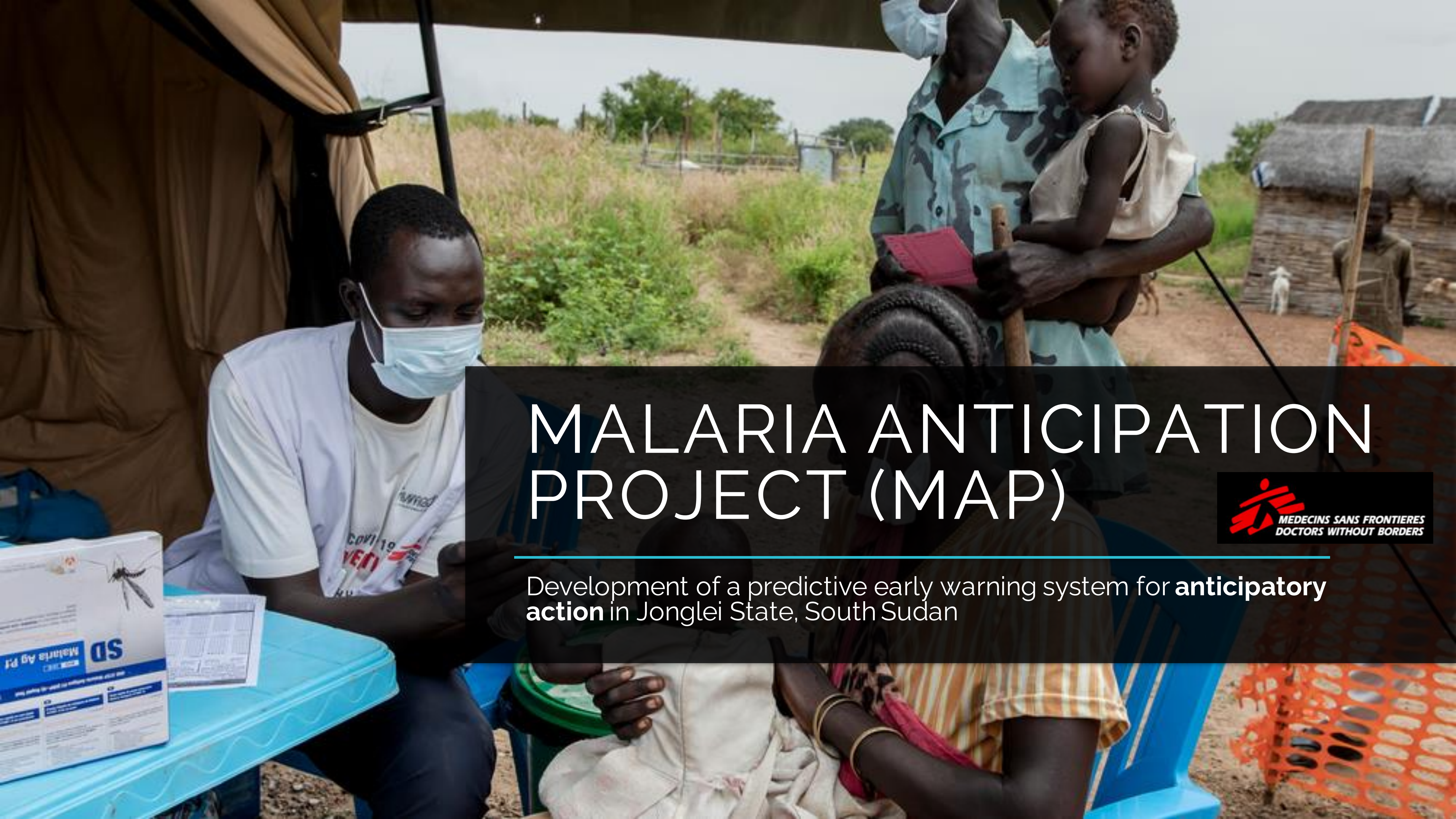


MALARIA ANTICIPATION PROJECT (MAP)

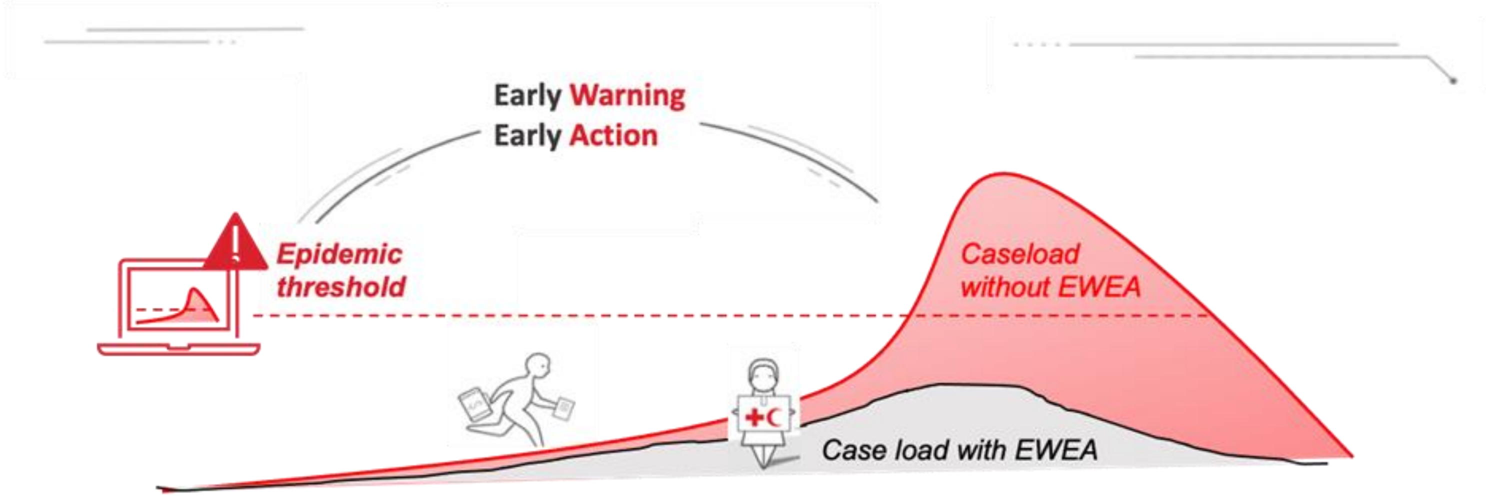
MEDECINS SANS FRONTIERES
DOCTORS WITHOUT BORDERS

Development of a predictive early warning system for **anticipatory action** in Jonglei State, South Sudan

PROBLEM STATEMENT

Urgent need to develop and implement strategies to anticipate malaria disease burden, given:

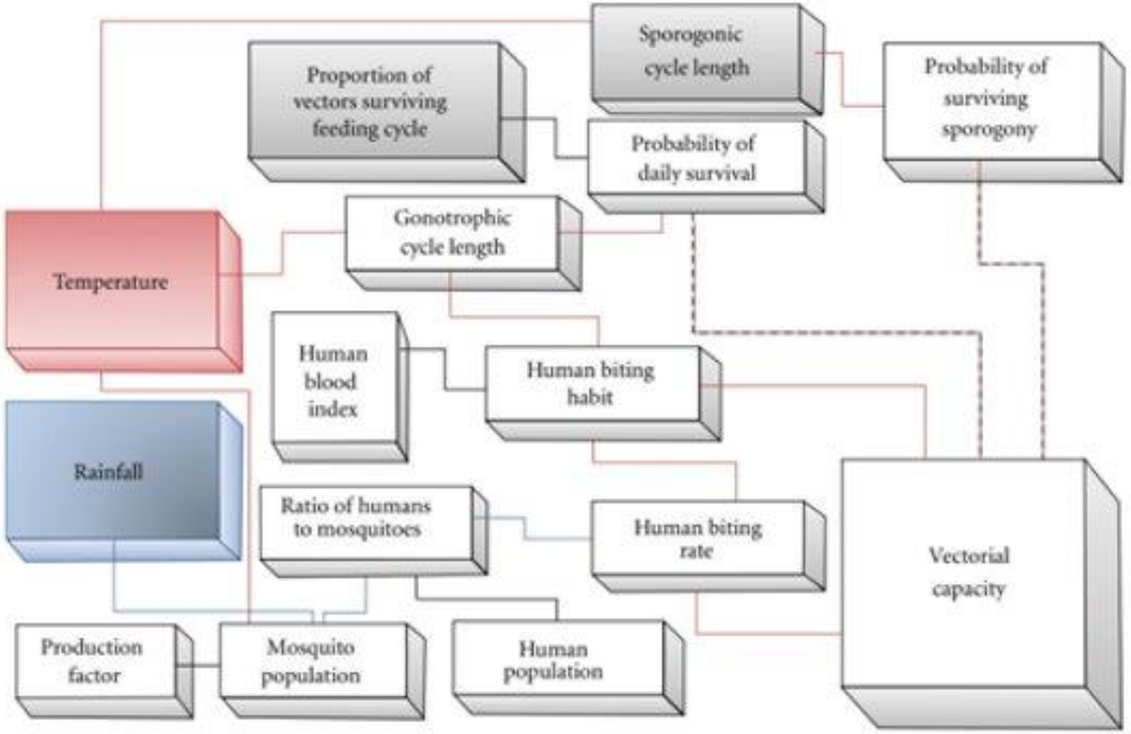
- evolving climate and environmental conditions
- focus on prioritisation/resource allocation given multitude of needs
- novel preventative interventions may have a limited window of implementation



Source: *Red Cross Climate Center*

WHY ENVIRONMENTAL VARIABLES?

- There is a wealth of scientific literature linking environmental variables with malaria dynamics in a variety of locations.
- Data easy to access retrospectively; continuous datasets; no need to invest in situ data collection.
- May be able to add other variables in future



Malaria vector capacity model (from Ceccato et al, 2012).

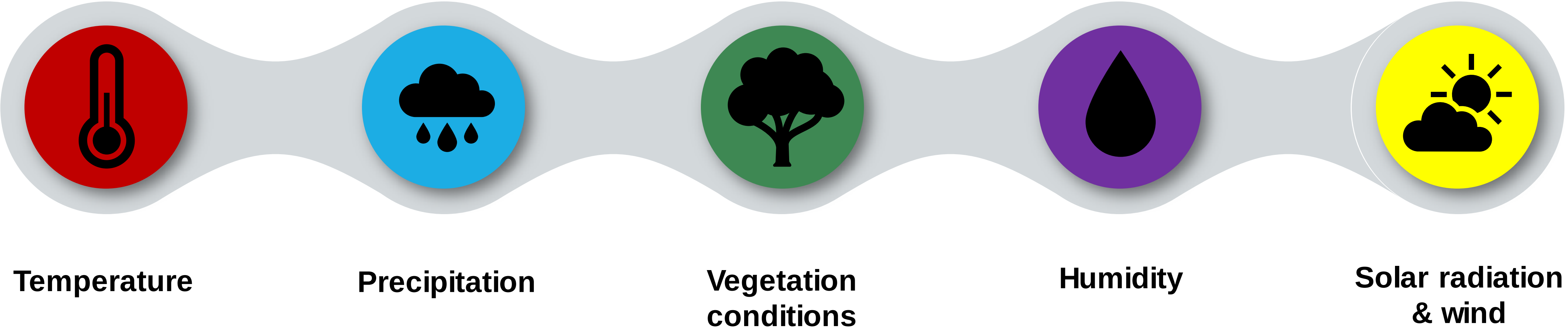

SOUTH SUDAN
SUDAN
ETHIOPIA
CENTRAL AFRICAN REPUBLIC
N
0 100 200 km
Renk
NORTHERN UPPER NILE
Kodok
FASHODA
Malakal
CENTRAL UPPER NILE
MAIWUT
LATJOR
Nasser
Maiwut
BIEH
Waat
Akobo
AKOBO
Panriang
RUWENG
NORTHERN LIECH
Bentiu
Abyei
AWEIL EAST
Mayen Abun
TWIC
Wanyjok
Aweil
GOGRIAL
AWEIL
LOL
Raja
Kuajok
Leer
FANGAK
Ayod
TONJ
WAU
Wau
Tonj
JONGLEI
Cueibet
GOK
Rumbek
Yirol
EASTERN LAKE
Bor
Pibor Town
BOMA
TAMBURA
WESTERN LAKE
Tambura
TEREKEKA
Terekeka
AMADI
Mundri
MARIDI
Maridi
GBUDWE
Yambio
Juba
KAPOETA
Kapoeta
GIS

MODEL DEVELOPMENT

Aim: Forecast malaria burden using machine learning models

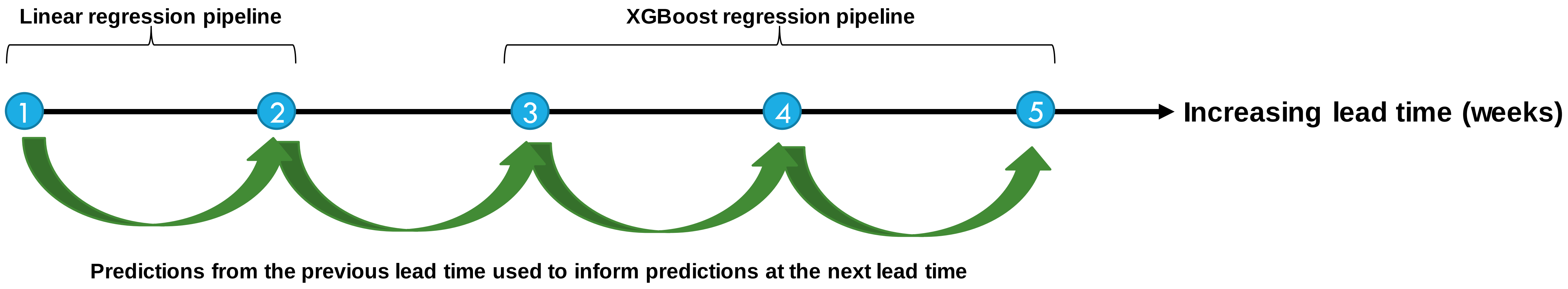
Key differentiators in modelling approach:

- Use over 20 meteorological and environmental variables
- Intensive feature engineering to develop proxies for antecedent conditions
- Use tailored model performance criteria



MODEL DEVELOPMENT

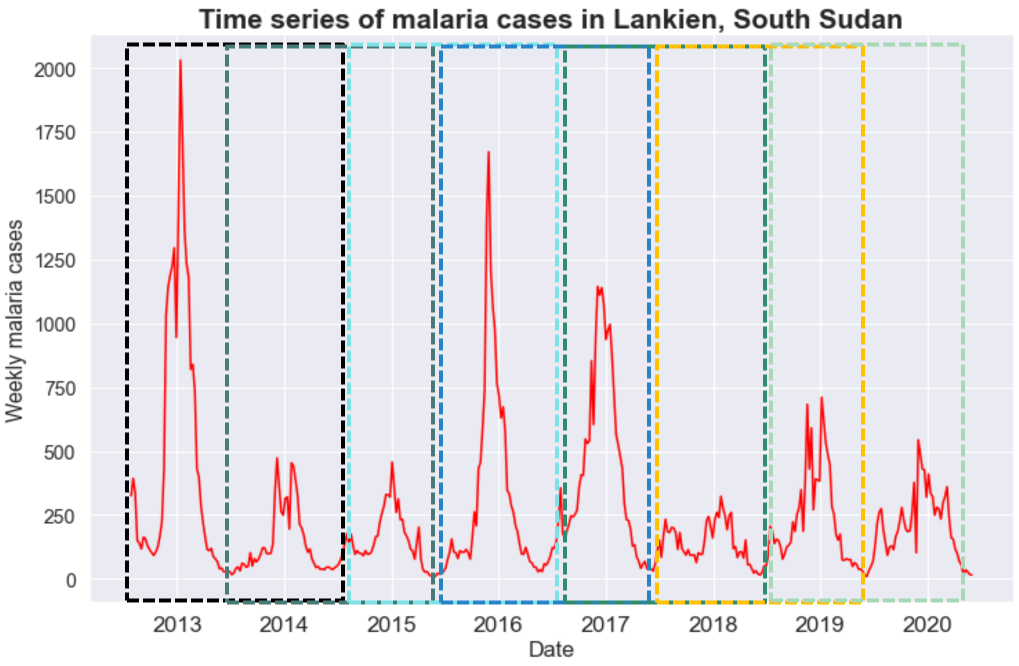
A recursive modelling approach combining linear regression and extreme gradient boosting (XGBoost).

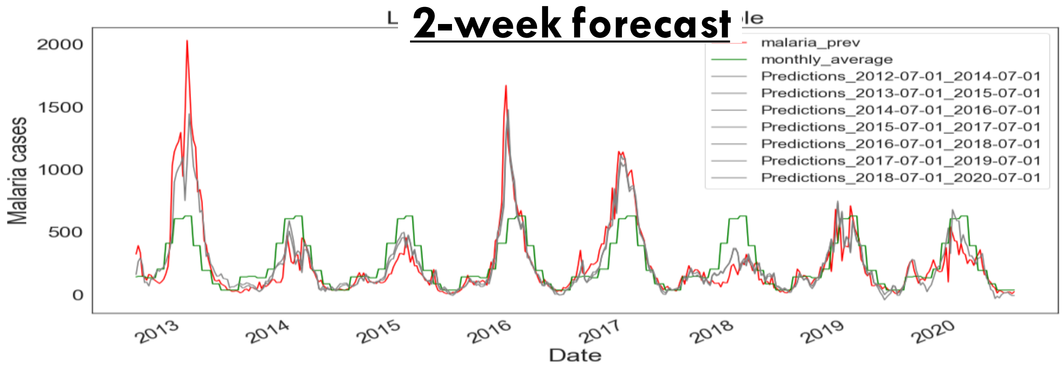


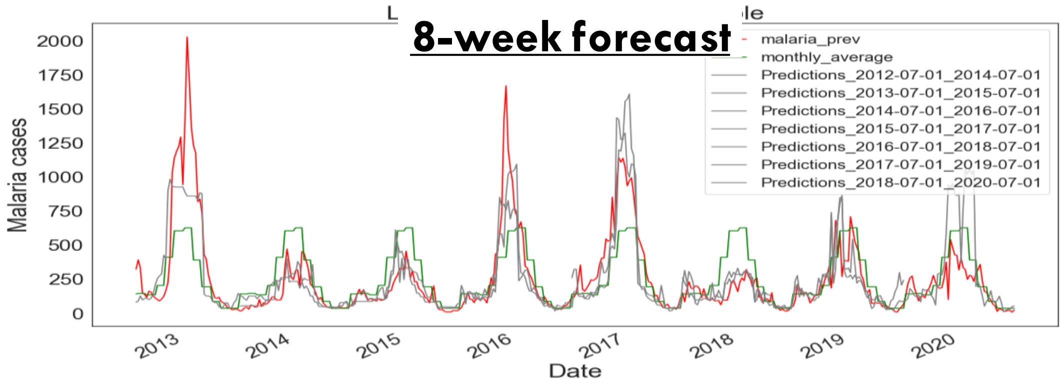
MODEL RESULTS

CROSS VALIDATION (2012-2020)

Promising results up to an 8-week forecast, particularly for predicting risk relative to the monthly average.





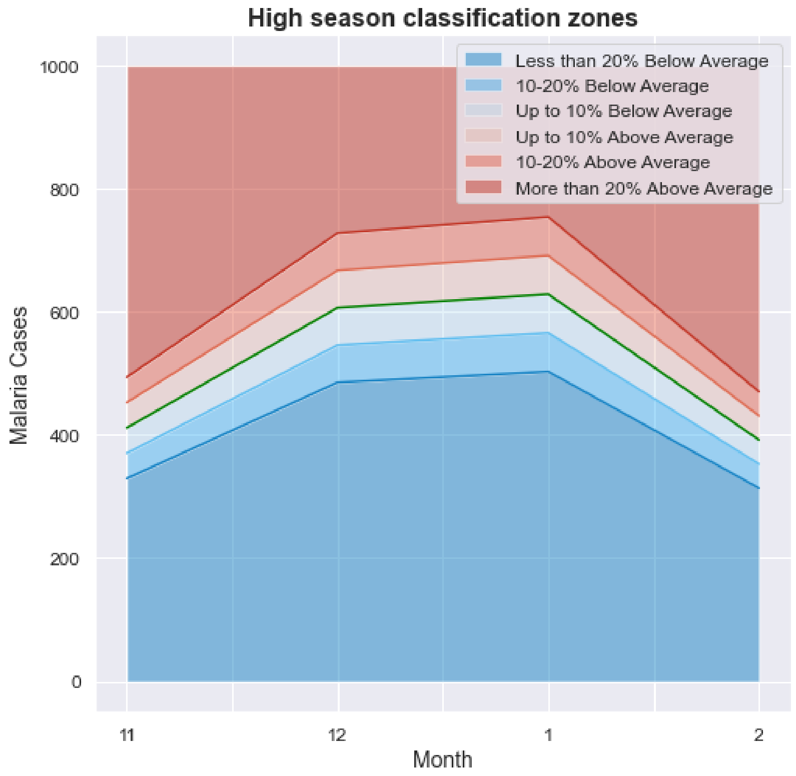


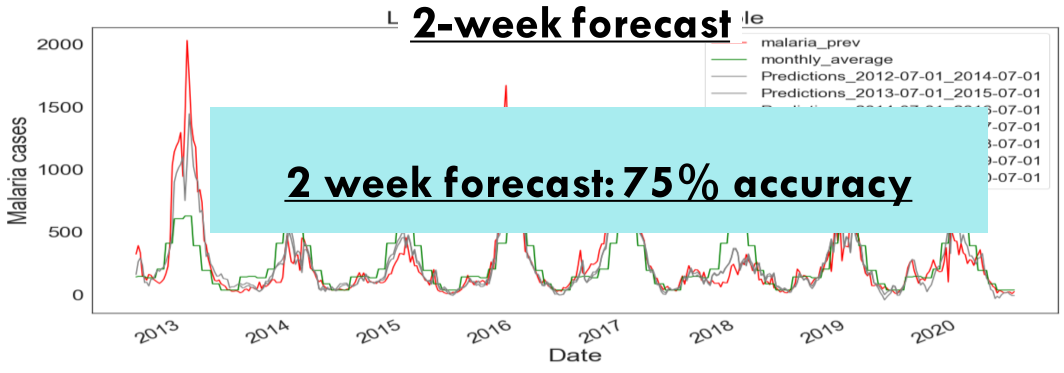
MEDECINS SANS FRONTIERES ARUP

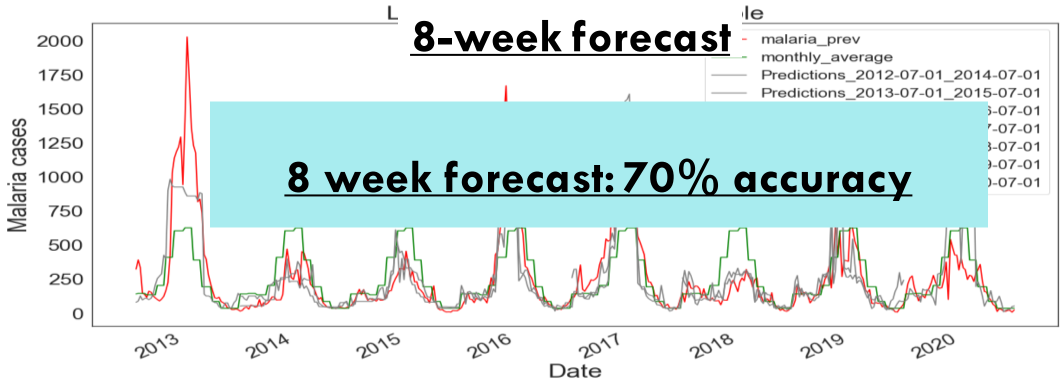
MODEL RESULTS

CROSS VALIDATION (2012-2020)

Promising results up to an 8-week forecast, particularly for predicting risk relative to the monthly average.



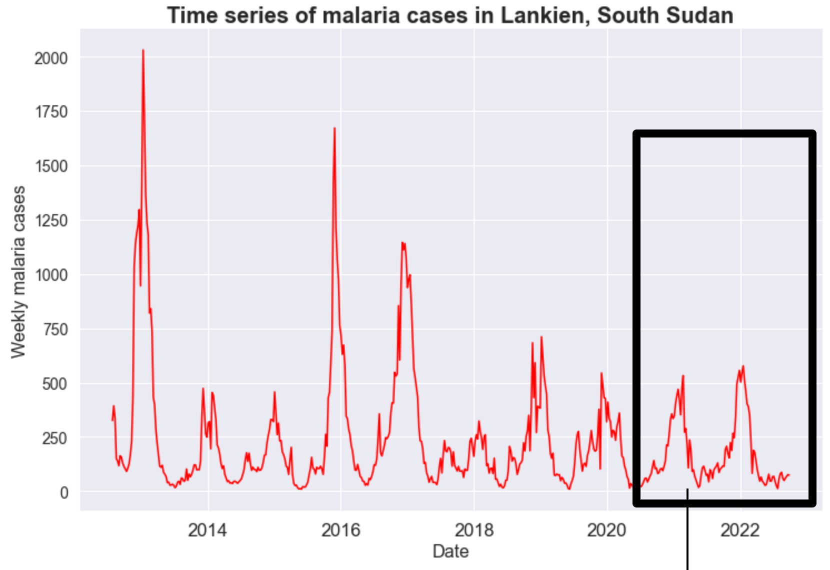




MODEL RESULTS

VALIDATION DATASET (2020-2022)

Overpredicting malaria risk for forecasts over two weeks.



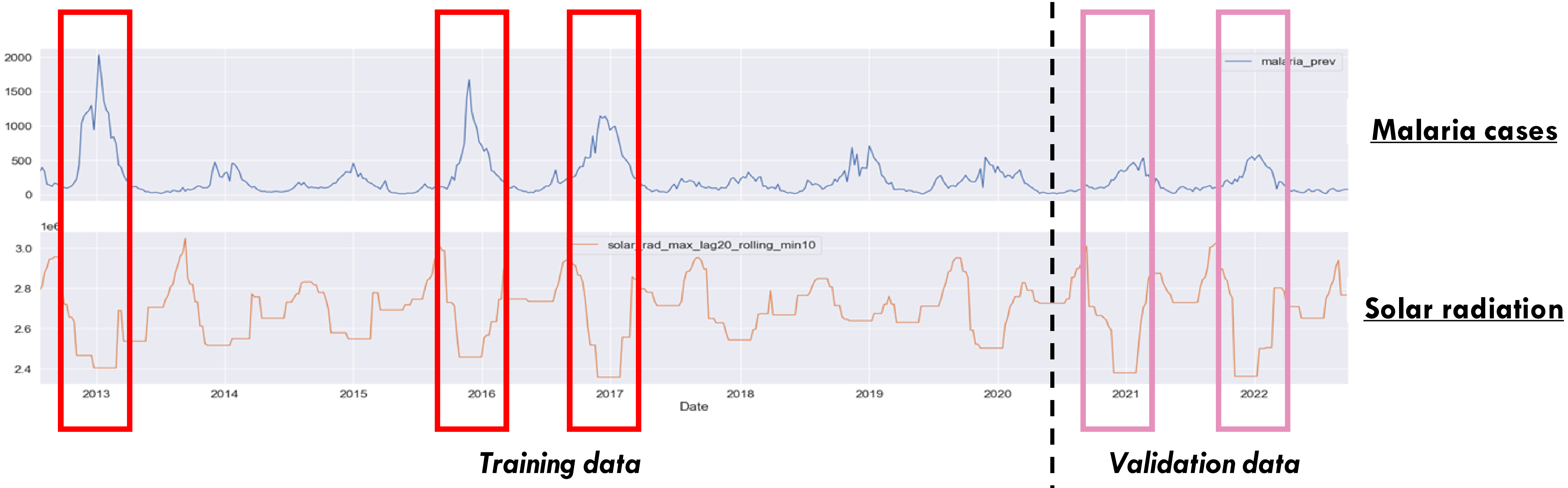
Validation data (2020-22)

MODEL RESULTS

VALIDATION DATASET (2020-2022)

Driven by the relationship between malaria and solar radiation in the training dataset.

This shows the **limitations in the length of the training dataset** and/or **model vulnerability to changing conditions**



COMMUNITY UNDERSTANDING OF ANTICIPATORY ACTION, CLIMATE & MALARIA

Rationale

- To interrogate/corroborate the models we have developed and challenge biases and assumptions around community malaria anticipation .
- To operationalise the predictions by understanding community care/seeking in relation to malaria and any perspectives they have on anticipatory action.
- **Methodology**

	Participant group	Analysis
Life histories (n=8)	Community elders Community leaders Agro-pastoralists	• Code-based narrative construction • Comparison with environmental variables picked up in models
Key Informant Interviews (n=6)	Health workers (biomedical & traditional) with experience of malaria control interventions	Code-based inductive and deductive analysis
Focus group discussions (n=6)	Community members (n=58)	Code-based inductive and deductive analysis

THEME 1: KNOWLEDGE OF MALARIA

"In the rainy season [mosquitoes] are huge and dark in colour. After the rainy season they are white and thin… it is the dark one I think that causes too much malaria."
KII4 – (biomedical health staff)

© Nicola Flamigni/MSF

THEME 2: PREDICTION AND DISSEMINATION

"We normally know if the annual coming of rain changes, especially if it comes from the west - we know that water will be much more than usual and that will lead to an outbreak of diseases. The good rain for the crops is from either the north or north west." FGD2 (women)

© Paul Odongo/MSF

THEME 3: CHANGES TO LIFE OVER TIME AND IMPACT ON HEALTH AND WELLBEING

"It was in the year 2017 that the climate change and that made all these disasters, such as flooding, drought, poverty and disease outbreak. They have become too much. 15 years back when WFP dropped food down, we wondered what is this food dropped for?" FGD 4 (men)

© Nicola Flamigni/MSF

MEDECINS SANS FRONTIERES ARUP

KEY FINDINGS

The modelling approach shows promise but also has limitations

- **Modelling has the potential to inform anticipatory action**
- **There are limitations in forecasting approaches**
- **Opportunity to work with communities to leverage existing anticipatory practices**
- **Difficult to compare two knowledge systems**

BENEFITS AND OPPORTUNITIES

More timely emergency response aligned with communities' needs

Improved efficiency and resource allocation

Climate & community adaptation of MSF projects

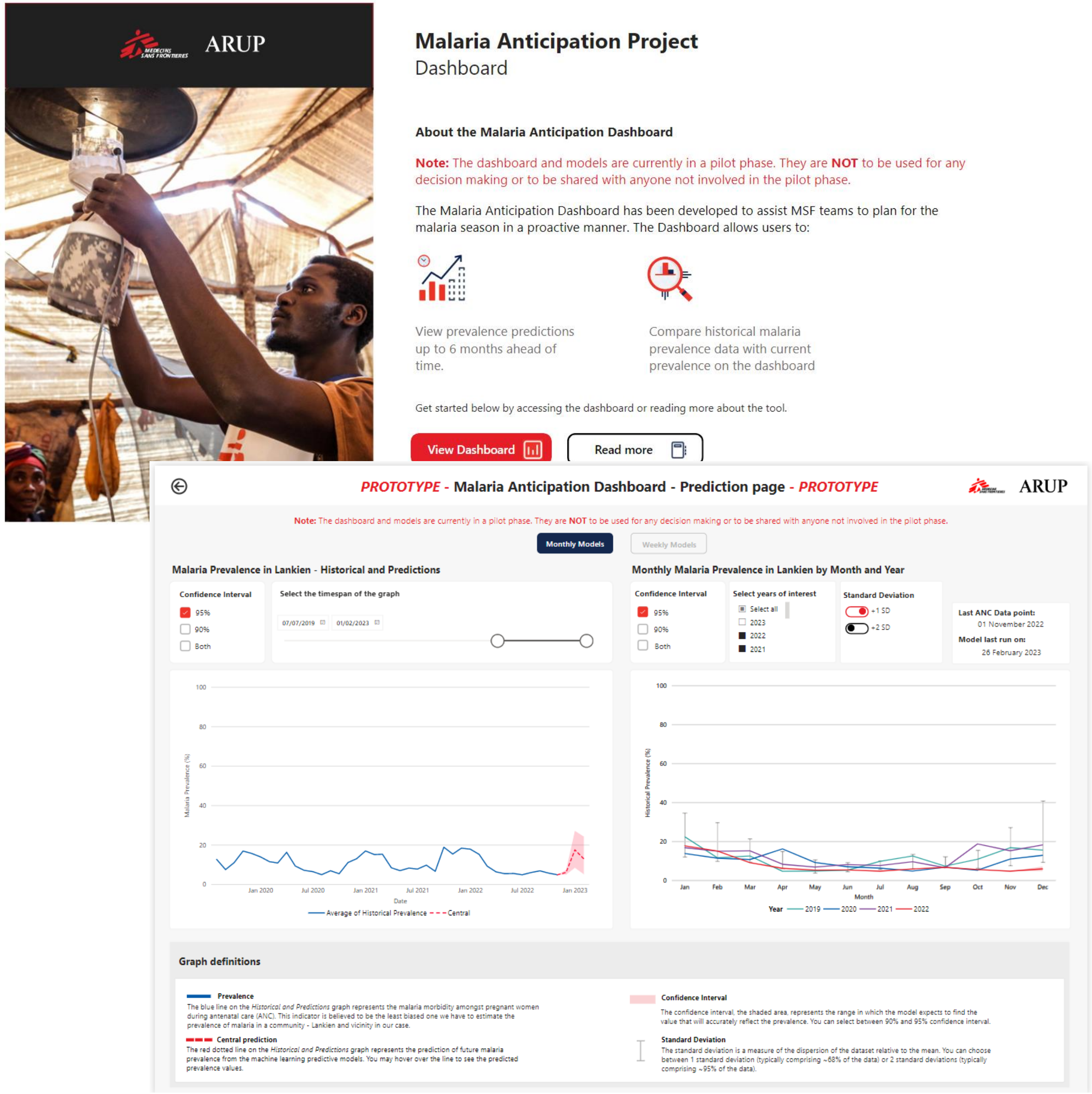
Maximising use of data

Potential application to other climate sensitive diseases

Shifting power and supporting field-level data-driven decision-making

Growing anticipatory action for health ecosystem

THANK YOU



ARUP
Malaria Anticipation Project
Dashboard
About the Malaria Anticipation Dashboard
Note: The dashboard and models are currently in a pilot phase. They are NOT to be used for any decision making or to be shared with anyone not involved in the pilot phase.
The Malaria Anticipation Dashboard has been developed to assist MSF teams to plan for the malaria season in a proactive manner. The Dashboard allows users to:
View prevalence predictions up to 6 months ahead of time.
Compare historical malaria prevalence data with current prevalence on the dashboard
Get started below by accessing the dashboard or reading more about the tool.
View Dashboard
Read more
PROTOTYPE - Malaria Anticipation Dashboard - Prediction page - PROTOTYPE
ARUP
Monthly Models
Weekly Models
Malaria Prevalence in Lankien - Historical and Predictions
Monthly Malaria Prevalence in Lankien by Month and Year
Graph definitions

ADDITIONAL SLIDES

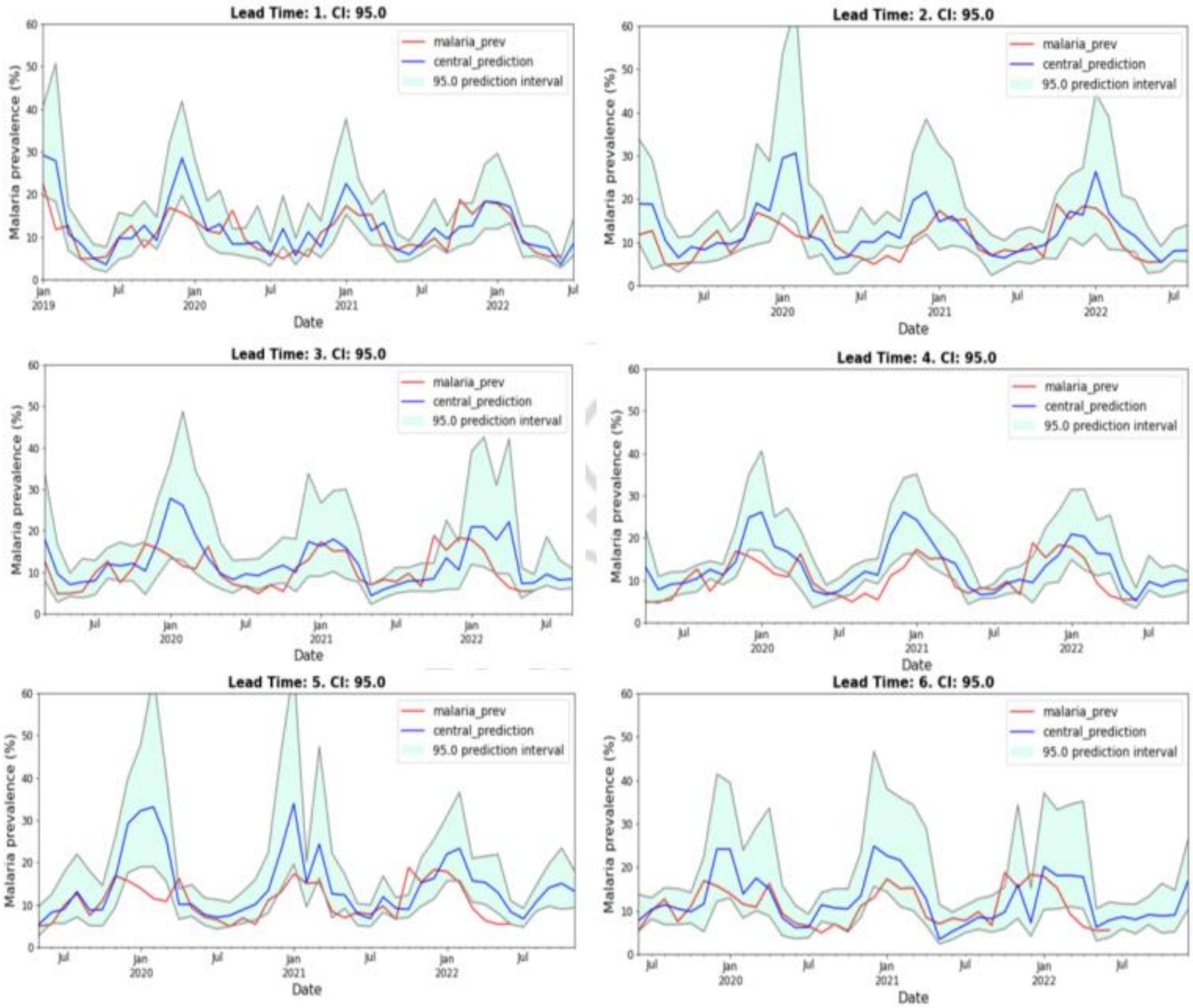
KEY FINDINGS SO FAR – MODEL PERFORMANCE

Results from model development (2012-2022)

- ANC models show a high predictive performance at one month lead time (79% observed malaria prevalence fall within 95% CI) This decreases at 2-3 months (70%), but remains promising. Drops off at 3 months with decreased performance 4-6 months
- OPD models show very high predictive performance at 2 weeks lead time (classification accuracy of 75%). This steadily decreases with each increasing lead time but remains high until 8 weeks (70%). Under/over predictions are more likely with longer lead times

Results from 2020/21 and 2021/22 malaria seasons

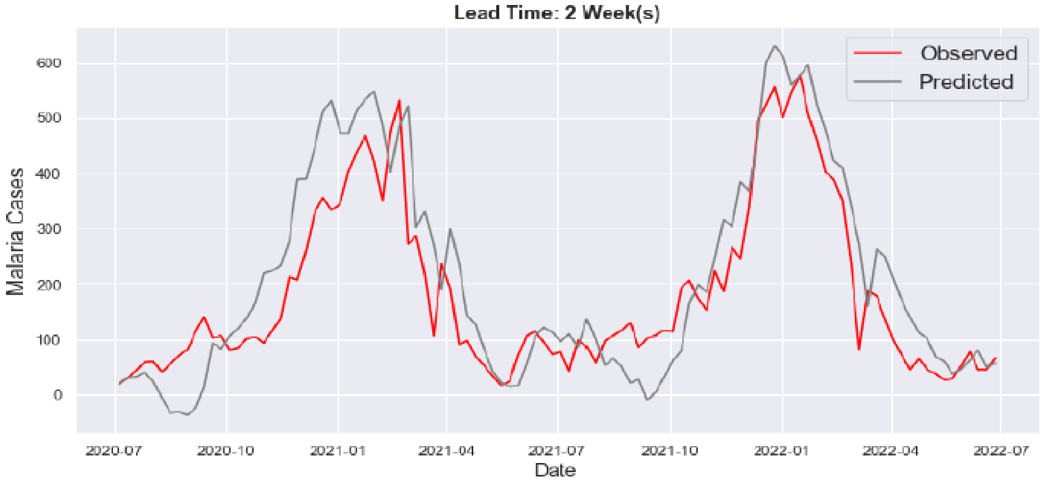
- **Background:** 2020-2022 were years were Lankien was impacted by COVID-19 and flooding and there was an unexpected earlier peak in October 2021 as well as a later peak in December 2021
- **ANC**: The models with lead times of 1-3 months provide reasonably good predictions for the magnitude of the malaria peak for 2020-21. For 2021-22, none of the models were able to accurately predict the earlier peak in October 2021. However, all models do a reasonably good job of predicting the magnitude and timing of the later peak
- **OPD:** there is an overprediction for both seasons at 4-8 weeks lead time thought to be associated with solar radiation (20 week lag) lead time.
- **Hypothesis:** lower solar radiation results in higher rainfall and flooding and therefore more malaria cases, until the flood extent reaches a tipping point where mosquito habitats are washed out and malaria cases are lower than normal. Another is that the 2020 expansion of the Sudd Marshes has disrupted environment-malaria relationship.

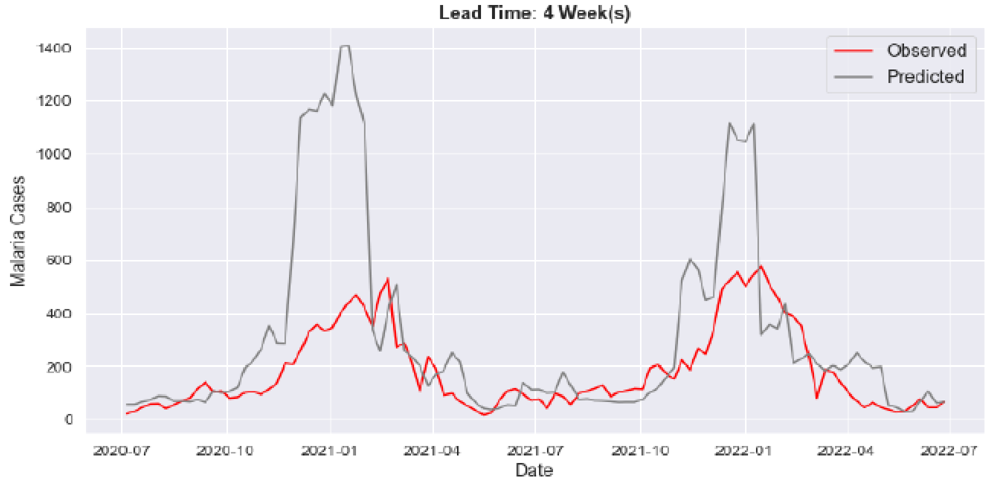


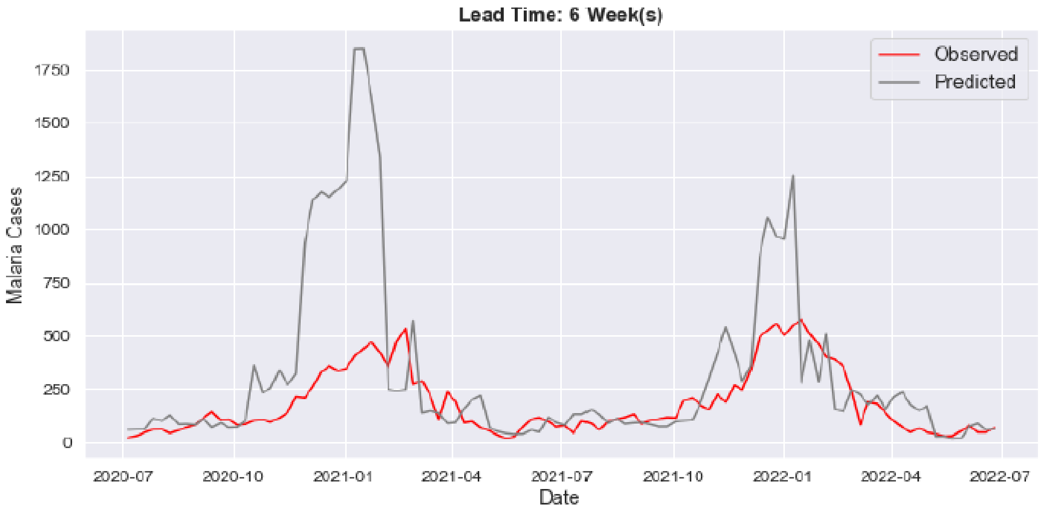
ANC MODEL PERFORMANCE ON VALIDATION DATASET (JANUARY 2019 - JULY 2022)

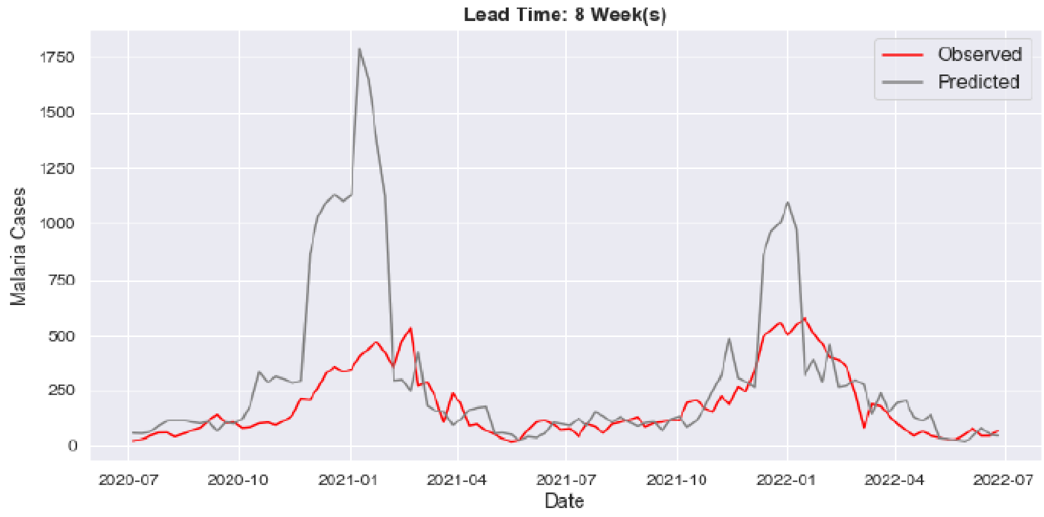
LEAD TIME IN MONTHS

OPD MODEL PERFORMANCE ON VALIDATION DATASET (JULY2020-JULY 2022) LEAD TIME WKS



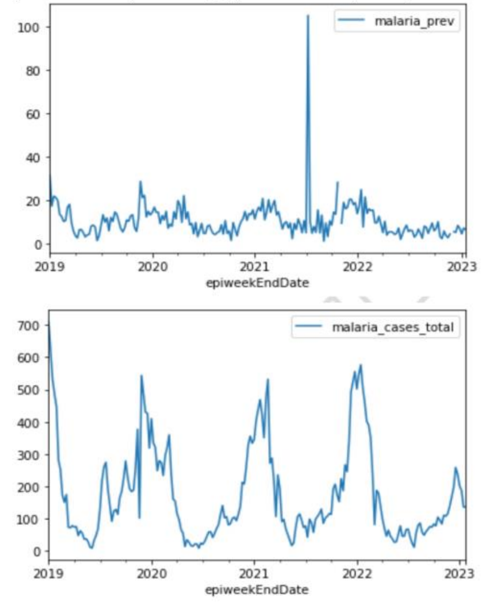






KEY FINDINGS SO FAR – MODEL PERFORMANCE

Figure 15: Malaria data from January 2019 to January 2023, provided by MSF for ANC prevalence (top) and OPD cases (bottom).



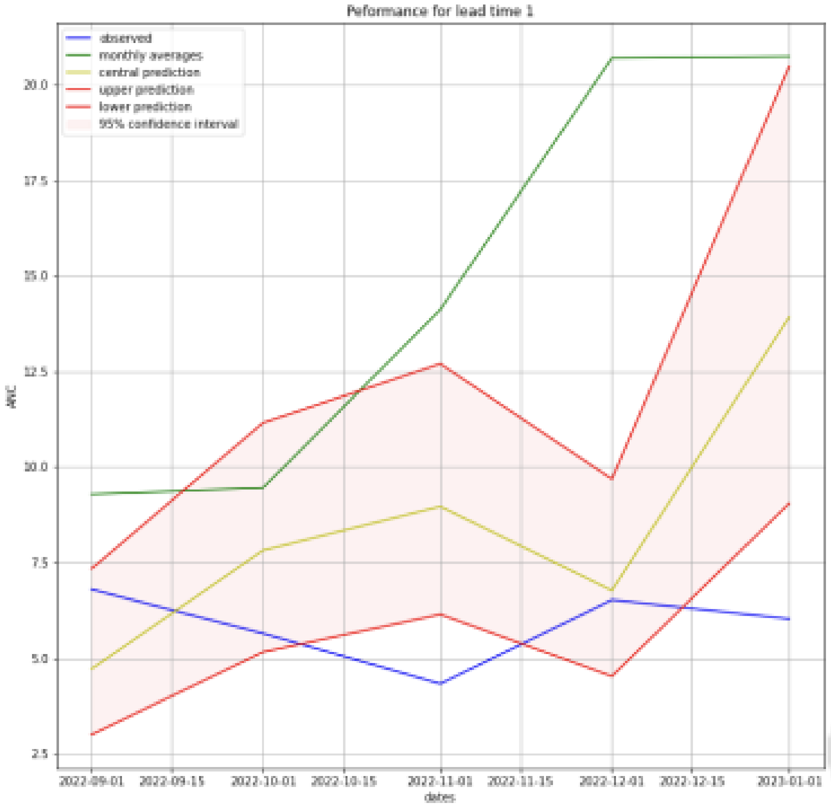
Results from 2022-2023 years

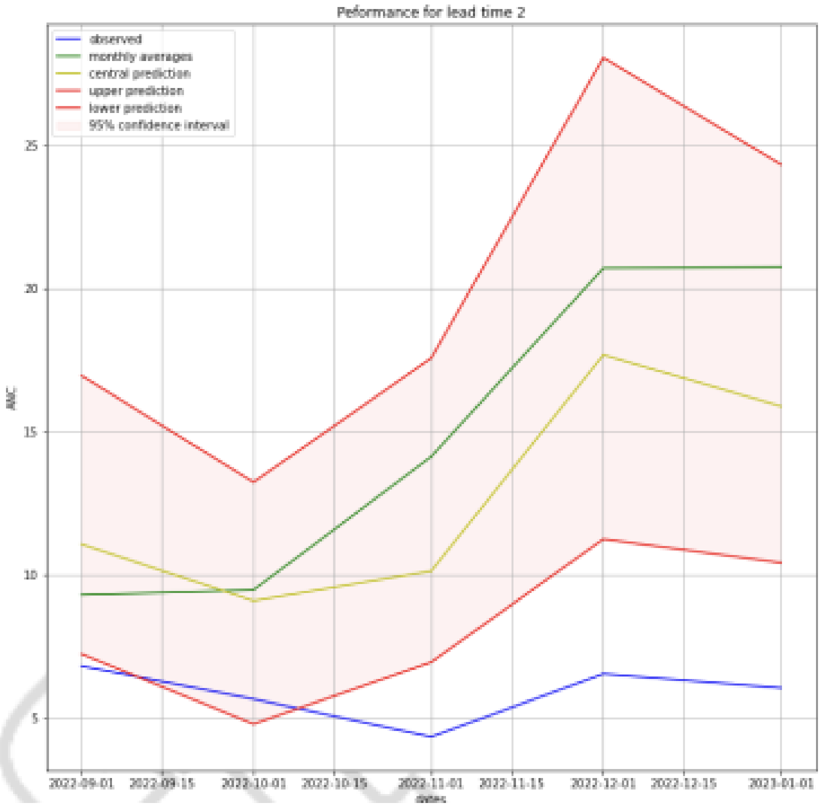
Background: 2022/23 malaria season was lower than usual – Observed peak in both ANC malaria prevalence and OPD malaria cases was the lowest peak seen across entire time series (2012-23).

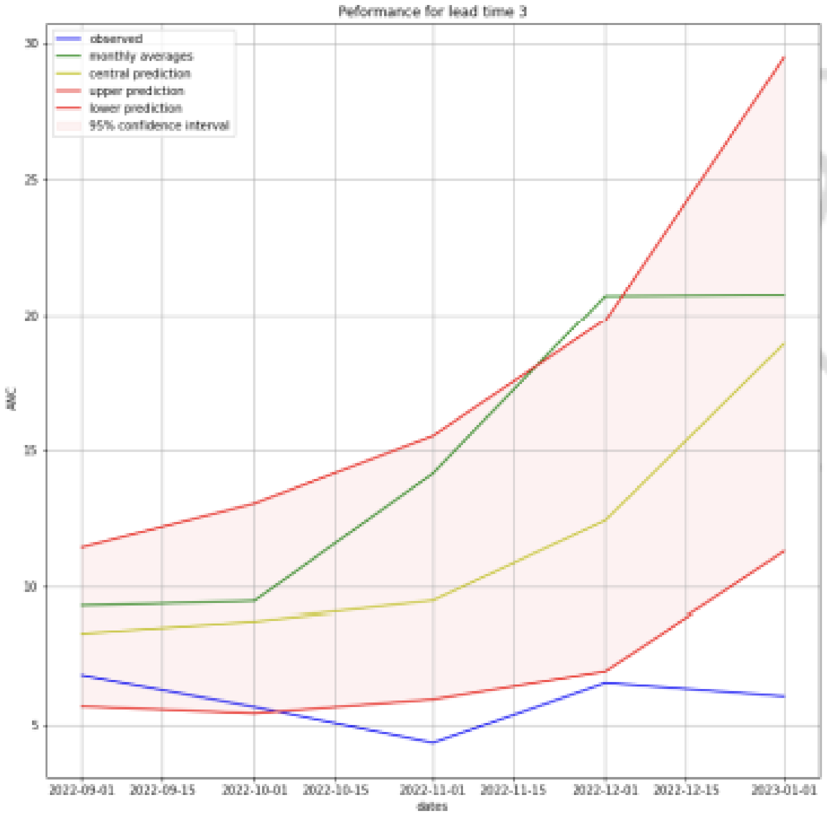
- **ANC:** At all lead times from one to three months, malaria prevalence is overpredicted for the peak malaria season between November and January & size of overprediction increases with lead time (although all lower than monthly average)

- **OPD:** For 2022/2023 season both models predicted that the malaria prevalence/number of cases would be below the long term average for Lankien – seen as successful outcome as models could anticipate lower malaria burden

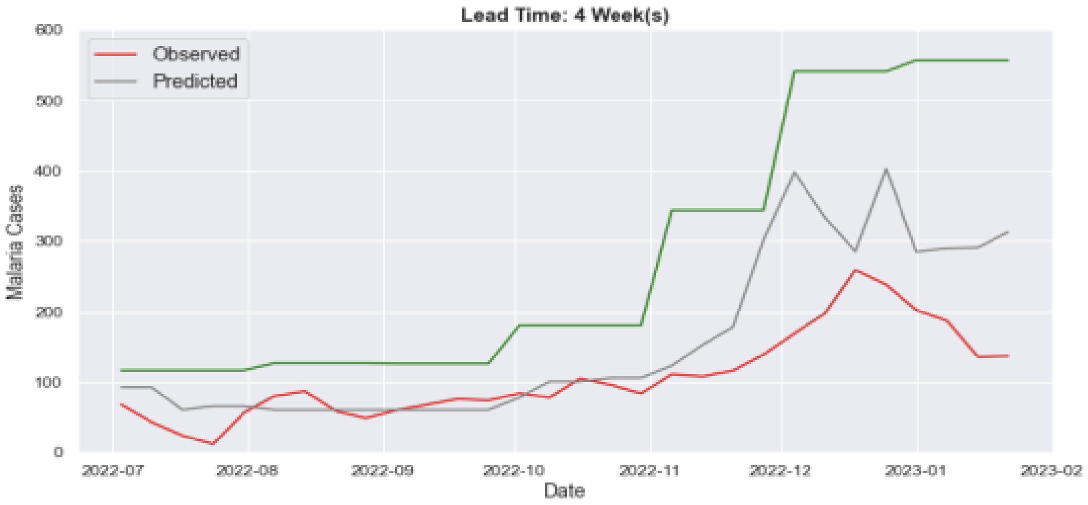
ANC MODEL PERFORMANCE (2022-2023)

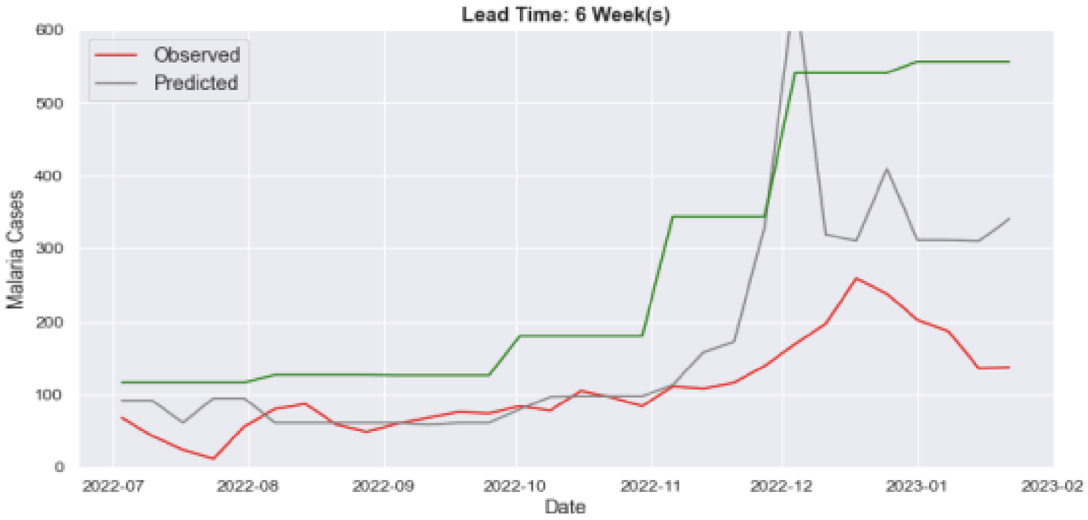


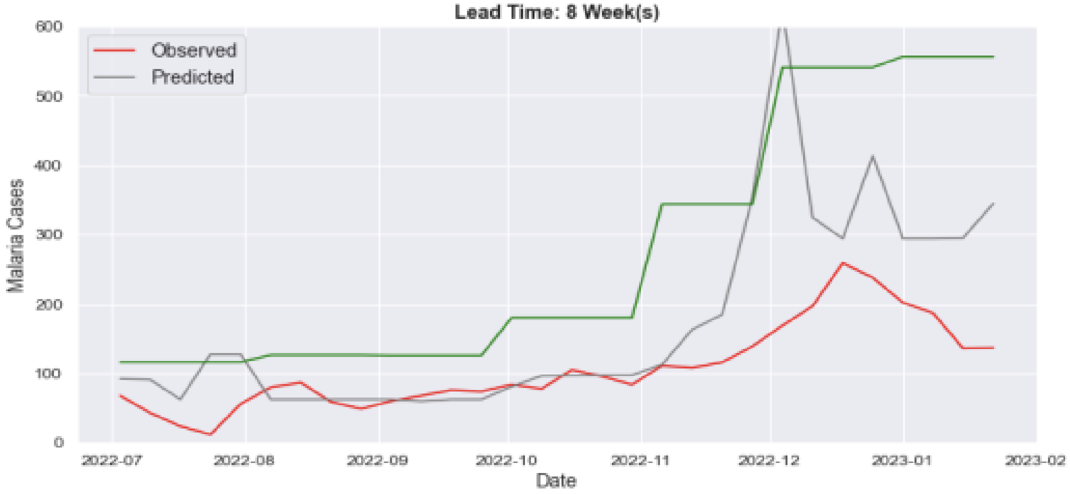




OPD MODEL PERFORMANCE (2022-2023)







MODEL PERFORMANCE SUMMARY

- **The shorter the lead time, the more accurate the prediction** for both ANC and OPD models, however longer lead times give more time for planning anticipatory action
- ANC models may be of use to in-practice responders for the 1 month lead time, possibly 2 months. The OPD models show potential at the 2-5 week lead times
- The **OPD models may be more reliable than the ANC models** due to larger number of data points
- **Model performance is limited by the size of the training dataset**, particularly for ANC
- Most important environmental predictors across different lead times, which could be further characterised in future, especially solar radiation:
  - **ANC:** Normalised Difference Vegetation Index (NVDI) (2 month lag), relative humidity (3-5 month lag), Normalised Difference Water Index (NDWI) (5 months lag), solar radiation (10-11 month lag)
  - **OPD:** Normalised Difference Water Index (NDWI) (16 week lag) , wind speed mean (13 week lag) and solar radiation (20 week lag)

OTHER KEY FINDINGS

- **Enthusiastic involvement of health teams:** there is appetite to use predictive models for anticipatory action based on field-level workshops user engagement feedback from pilot phase and feedback from in MoH. South Sudan teams has a **preference for OPD models** over ANC models for operational planning
- **Models can be successfully run on OCA cloud server,** but code review by internal and experts found code is complex and could potentially be simplified
- **Local/indigenous knowledge essential** to contextualise the predictions – communities should be involved in the co-design of anticipatory interventions in future phases

LOOKING FORWARD: OPPORTUNITIES

- **More timely emergency response aligned with communities' needs:** potential paradigm shift in malaria EPREP and response that allows more time for implementing preventive and curative measures and collaboration with other actors and communities
- **Improved efficiency and resource allocation:** through the development of thresholds for different anticipatory actions this helps to direct resources to where they are most needed and reduce need for emergency supply orders
- **Improved understanding of climate linkages to malaria:** puts a climate lens on our operations putting
- Approach can potentially be applied to **other climate-sensitive infectious diseases**
- **Growing anticipatory action for health ecosystem:** very few anticipatory in health projects globally and this project is a first for MSF. MSF are coordinating Interagency Working Group on Anticipatory Action for Health with OCHA and Red Cross Climate Centre to share knowledge and best practice,

CHALLENGES & LESSONS LEARNED

- Machine learning models need a **long epidemiological time series** to train the models – in terms of scalability, this excludes a number of MSF projects
- MSF has **limited data science capacity in house** - MAP reliant on external actors for model development
- Models **not fully automated** –DHiS2 data is currently manually uploaded but hoping this can be automated in future phases
- EWS currently only developed using MSF data. Need to explore **how to incorporate data from MoH/other actors** to ensure sustainability of EWS should MSF change/handover project
- **Need to put in contingencies** should specific environmental data sources stop being published open access – vulnerabilities of the models currently being explored with support from internal and external software experts.