Welcome!

NETHOPE SOLUTIONS CENTER

Webinar Series

We will get started at the top of the hour!

>>Please activate the Zoom Chat window for panelist and attendee interactions<<

YOU WILL EXPERIENCE SILENCE UNTIL WE START

WELCOME!

NetHope Solutions Center

Webinar:

Lessons Learned From Practical Implementations of AI/ML in the International Development Sector

March 31, 2020

CRS CATHOLIC RELIEF SERVICES

Releasing children from poverty Compassion in Jesus' name

NETHOPE

Housekeeping

- Let's keep this interactive:
  Post questions in Zoom Chat window
  for the Q&A and discussion session

- Look for a follow-up email
  with link to recording and collateral
  on NetHope Solutions Center

SOLUTIONS CENTER

- Please respond to webinar satisfaction poll
  presented after the webinar

NETHOPE

Host:

Leila Toplic, Lead for Emerging Technologies Initiative, NetHope

Speakers:

James Campbell, Regional Technical Advisor for Monitoring, Evaluation, Accountability and Learning for Southern Africa,
Catholic Relief Services

Dr. Filip Ponulak, Director of Data Science, Compassion International

NETHOPE

Lessons Learned from Practical Implementations of AI /ML in the International Development Sector

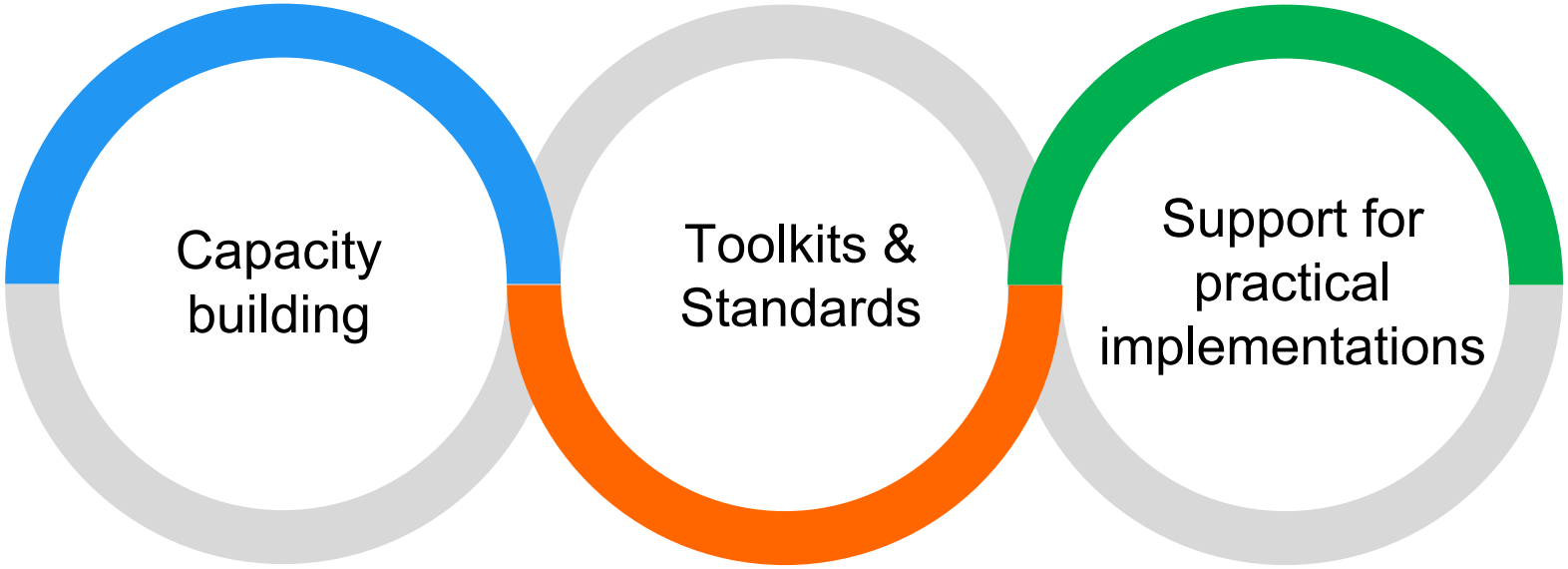
Host: Leila Toplic, Lead for Emerging Technologies, NetHope

Presenters:
Dr. Filip Ponulak, Director of Data Science, Compassion International

James Campbell, Regional Technical Advisor for Monitoring, Evaluation, Accountability and Learning for Southern Africa, Catholic Relief Services

NetHope's

AI Working Group



Focus: *Grow understanding of AI capabilities, tools & services. Share lessons learned from current practical implementations*

Focus: *Practical toolkits for international development sector.*

Focus: *Support for collective impact, multi-agency practical implementations solving real problems.*

Capacity Building

5 focus areas based on member needs:

- AI/ML – Capabilities, Benefits, Use Cases
- Lessons learned from practical implementations
- Ethical, responsible development and use of AI
- AI/ML tools and services that don't require specialize expertise
- Evaluating / using AI/ML for field programs and internal operations

Learn more here: https://solutionscenter.nethope.org/et

Today: Lessons learned from 2 practical implementations

Compassion International is exploring the use of ML to track the progress and effectiveness of their anti-poverty programs and interventions.

Catholic Relief Services is using Machine Learning to gain a better understanding of resilience in rural areas in Malawi.

As you begin evaluating AI for your work, consider using questions from AI Suitability Framework to guide the process.

AI Suitability Framework is a set of 32 questions designed to guide nonprofits and partners through the evaluation process to determine suitability of AI for international development and to plan for sustainability.

SAMPLE QUESTIONS.

1-What problem are you trying to solve?

2-How is the problem being addressed today?

3-Why is AI better than the current solution?

4-What is the solution?

5-What kinds of data does your solution need?

6-What resources do you need to support the development, implementation, and maintenance of the solution?

7-What are the potential biases that AI may introduce or amplify in your context?

8-What is your approach to maintaining the solution?

Host:

Leila Toplic, Lead for Emerging Technologies Initiative, NetHope
leila.toplic@nethope.org

Speakers:

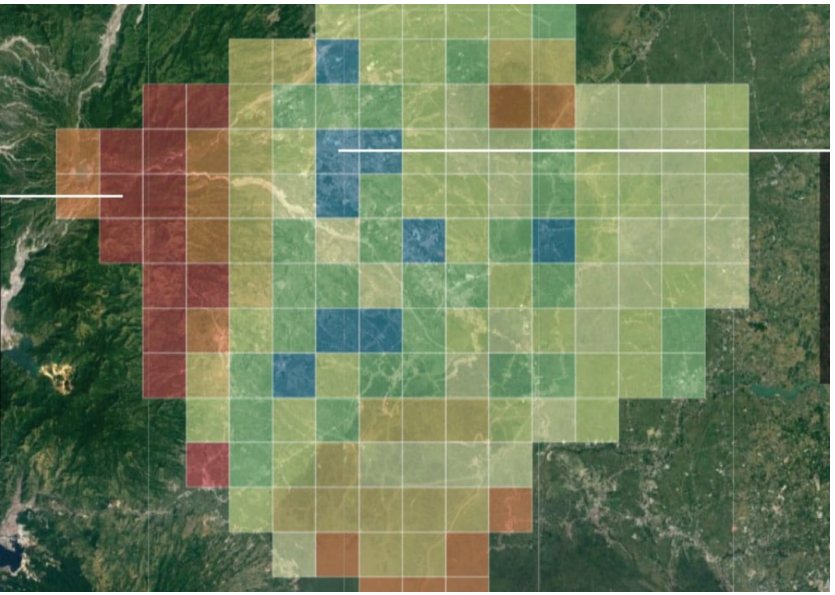
James Campbell, Regional Technical Advisor for Monitoring, Evaluation, Accountability and Learning for Southern Africa,
Catholic Relief Services
james.campbell@crs.org

Dr. Filip Ponulak, Director of Data Science,
Compassion International
fponulak@us.ci.org

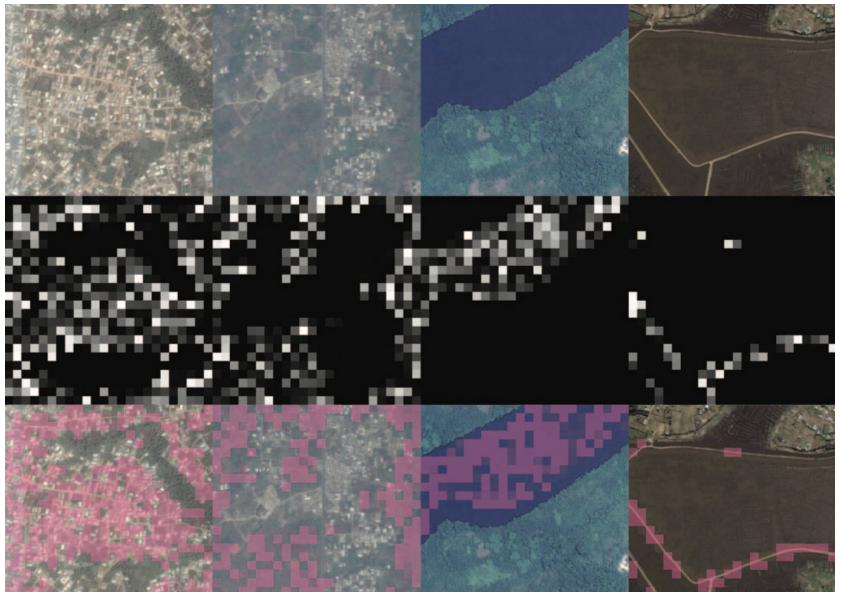
NETHOPE

Poverty Mapping based on Satellite Imagery

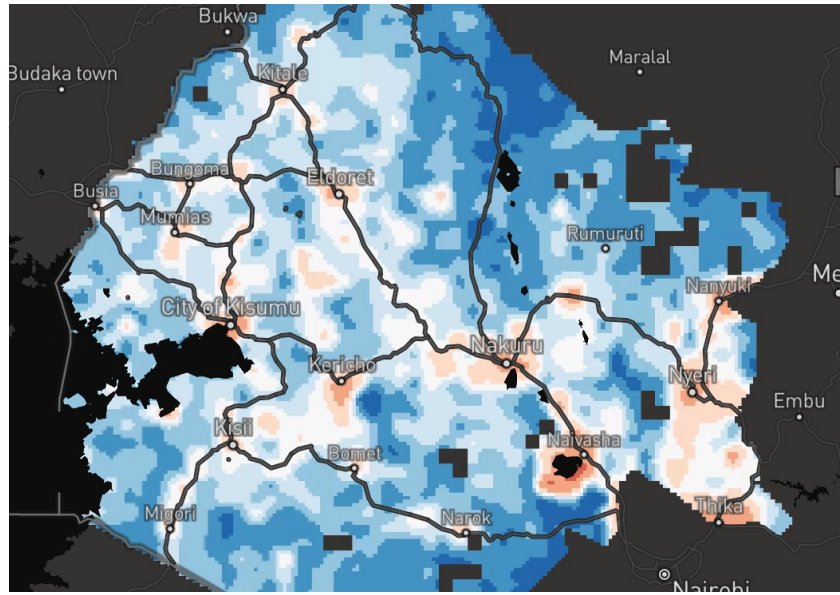
Dr. Filip Ponulak
Director of Data Science
Compassion International



Source: www.thinkingmachin.es



Source: Jean et al., Science 2016



Source: www.demo.atlas.ai.co

NETHOPE

What problem we are trying to solve

Find accurate, scalable methods of poverty mapping in order to identify and serve the most children in the greatest need of poverty alleviation.

COMPASSION INTERNATIONAL

2,000,000
SPONSORED CHILDREN

330,000,000
CHILDREN GLOBALLY IN EXTREME POVERTY

E

How is the problem being addressed today?

- ***Compassion's current poverty mapping is based on field surveys (paper) and on the knowledge of local staff in the areas of poverty.***

- ***While this approach gives us highly reliable information, it has its limitations:***
  - ***high cost of information gathering;***
  - ***limited frequency of updates;***
  - ***limited scalability capabilities;***

NETHOPE

New approach: poverty mapping using satellite images & ML

- ***We are currently running a pilot study of poverty mapping using machine learning to identify features in the satellite imagery indicative of economic activity.***
- ***Our work is based on the approach proposed by Neal Jean, et al. (Science, 2016 & AAAI 2016)***
- ***Advantages of the new approach:***
  - ***high spatial and temporal resolution,***
  - ***better scalability***
  - ***low implementation costs***

Myanmar (source: Digital Globe)

- *Jean et al. in "Combining satellite imagery and machine learning to predict poverty", Science (2016)*
- *Xie, et al., "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping", Proc. AAAI (2016)*

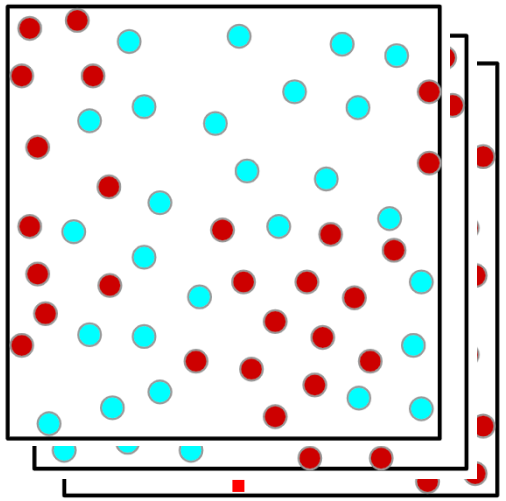
Method overview

At their very core ML algorithms are designed to:

(1) learn from examples

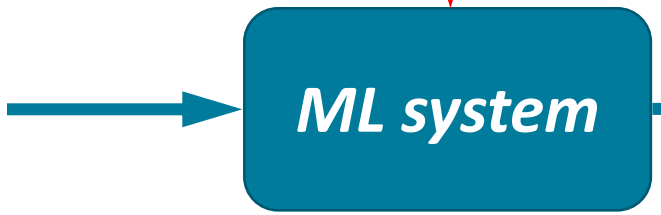
(2) find patterns in data, and

(3) generalize acquired knowledge

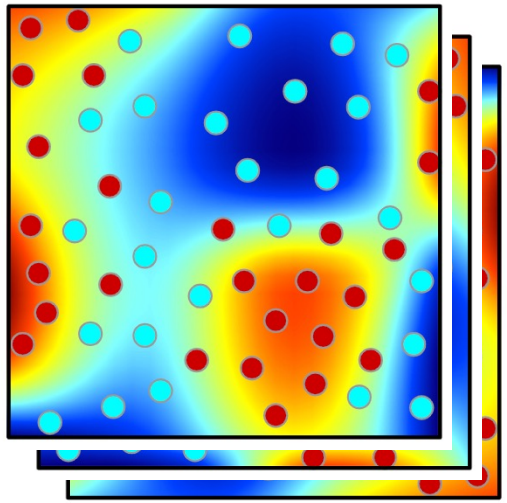


training signal: *"ground truth" poverty values at selected locations (e.g. survey-based)*

input: *sat. images*



output: *poverty risk maps*



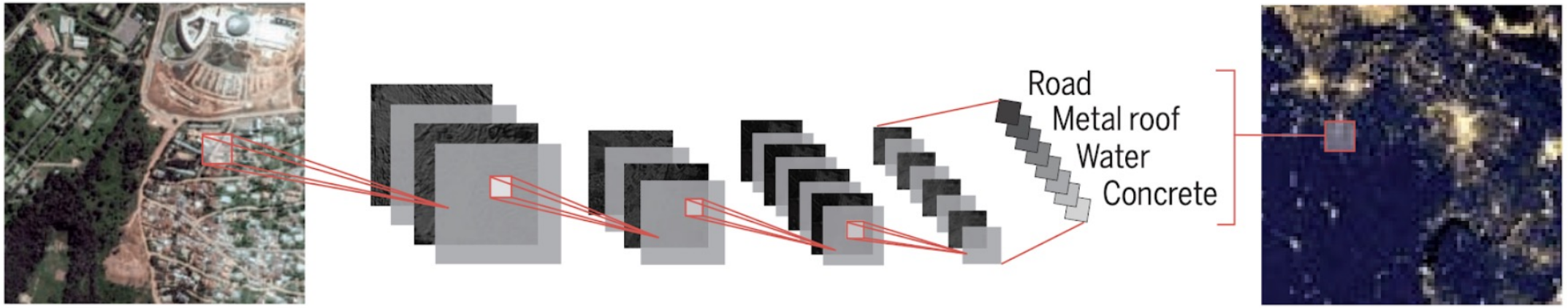
Method overview

Transfer learning: intermediate training step

Daytime satellite photos capture details of the landscape

Convolutional Neural Network (CNN) associates features from daytime photos with nightlight intensity

Satellite nightlights are a proxy for economic activity



Sources: Fighting poverty with data, Blumenstock, Science (2016),
Combining satellite imagery and machine learning to predict poverty, Jean et al., Science (2016)

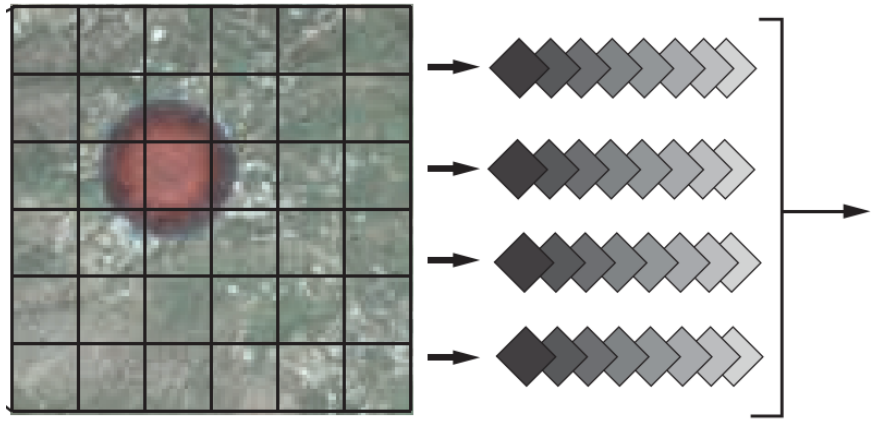
NETHOPE

Method overview

Regression model

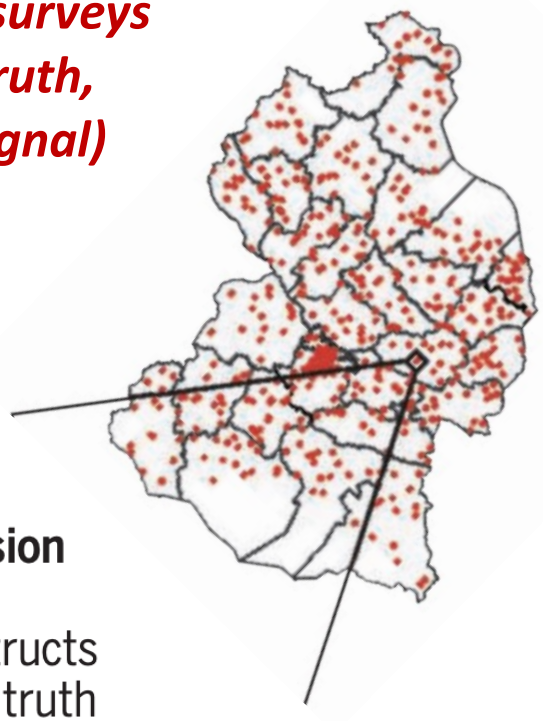
Household surveys (ground truth, training signal)

CNN processes satellite photos



Features from multiple photos are averaged

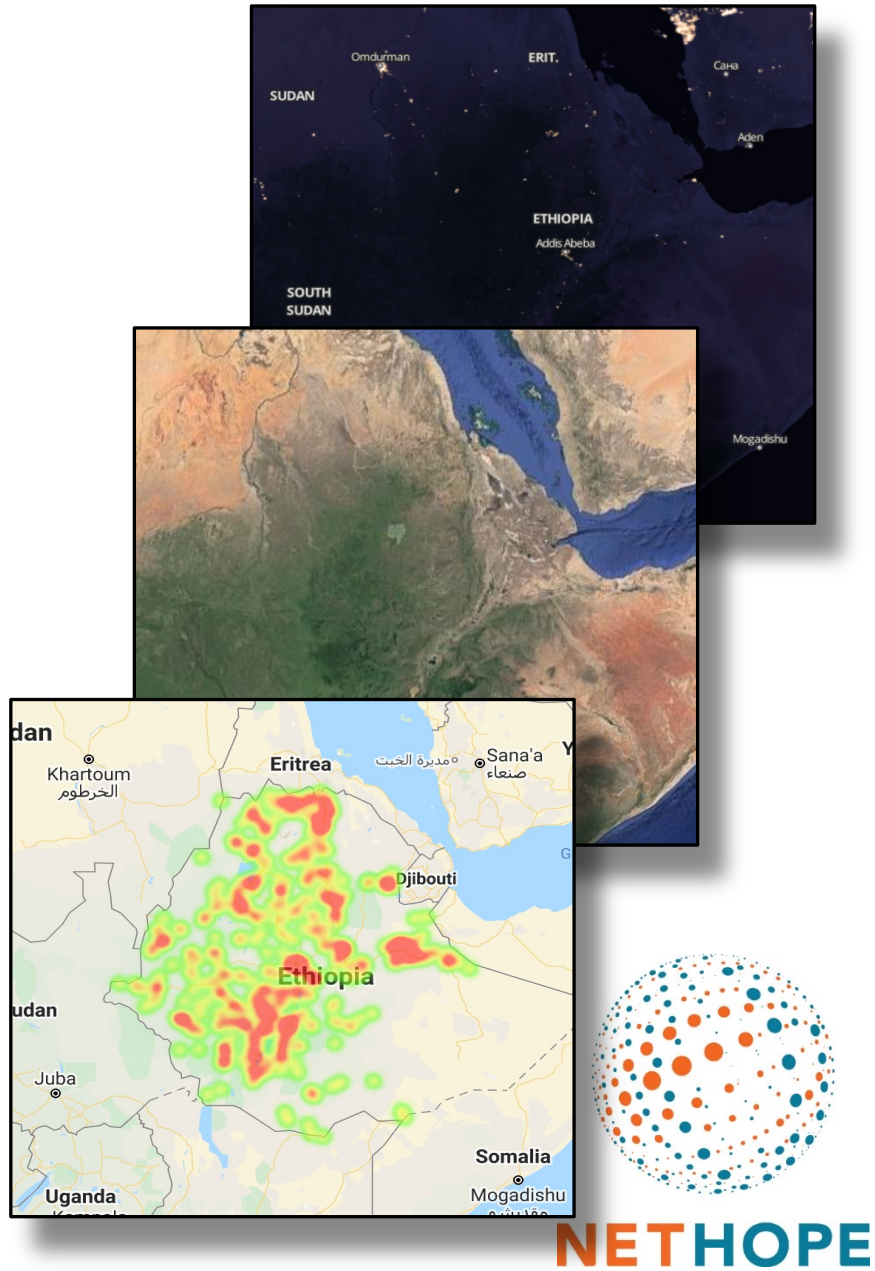
Ridge regression model reconstructs ground truth estimates of poverty



NETHOPE

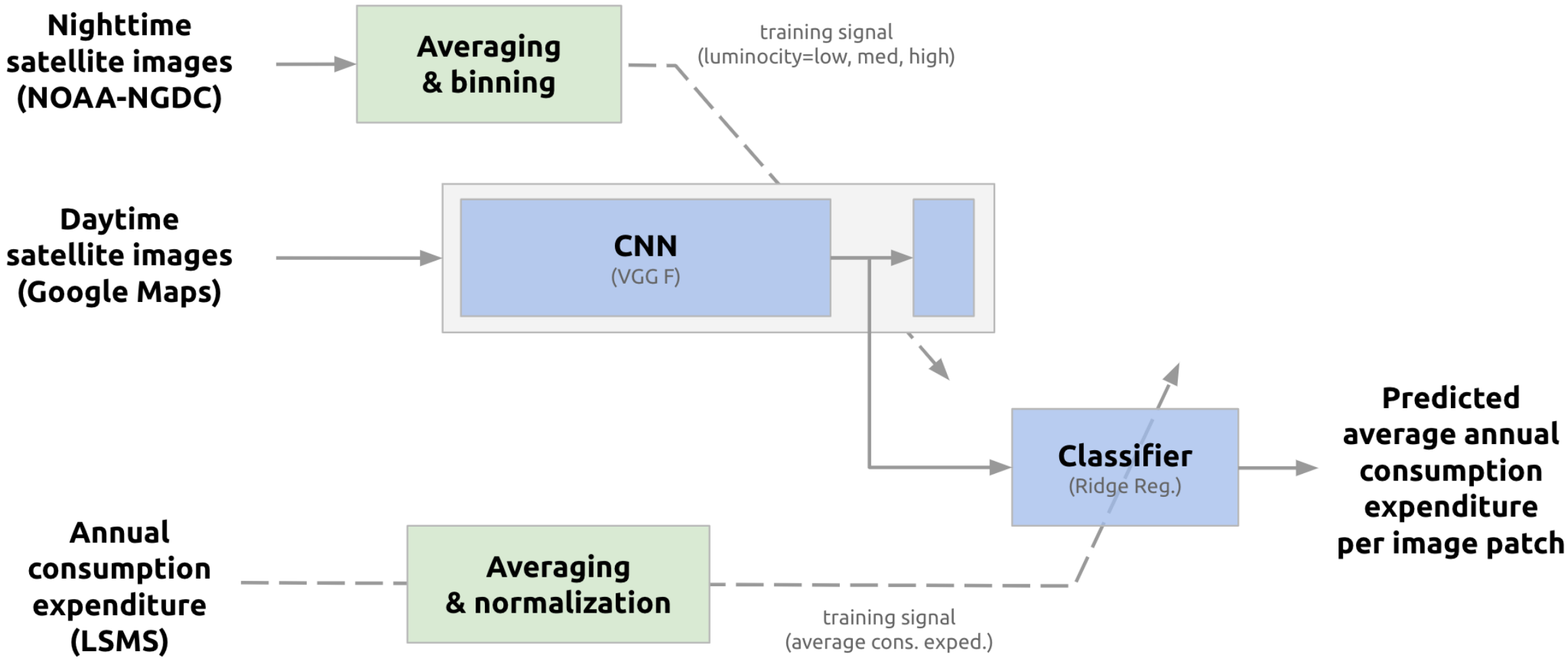
Data and data sources used in the approach:

1. ***Nighttime sat. images: National Oceanic and Atmospheric Administration's National Geophysical Data Center (NOAA-NGDC)***
2. ***Daytime satellite images (Google Static Maps API)***
3. ***"Ground truth" – local data (e.g. survey-based):***
   - ***World Bank's: Living Standards Measurement Surveys (LSMS) (publicly available)***
   - ***Compassion's proprietary poverty indicators***



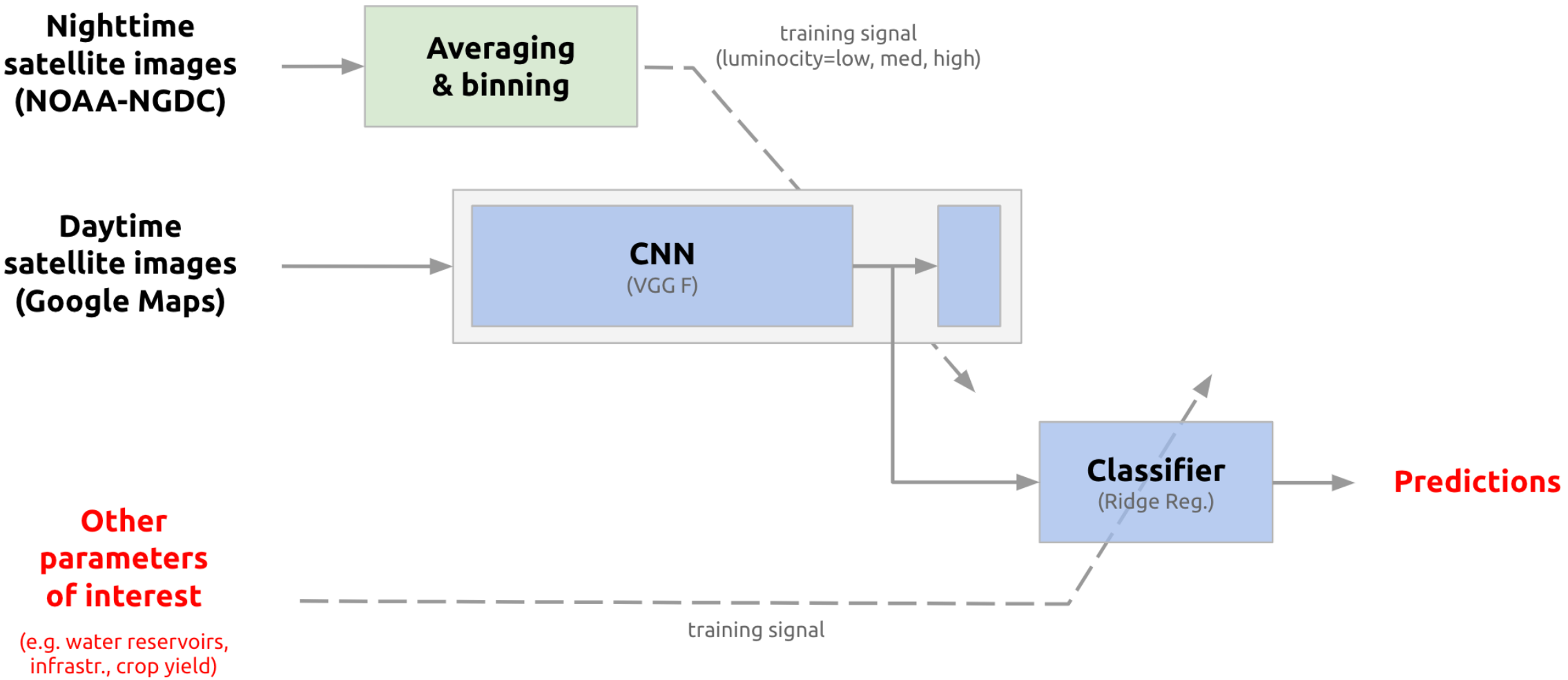
NETHOPE

System diagram (current stage)



NETHOPE

System diagram (other opportunities)



NETHOPE

Pilot study

Tanzania

99,332
CHILDREN SERVED

459
CHURCHES SERVING

Ethiopia

120,638
CHILDREN SERVED

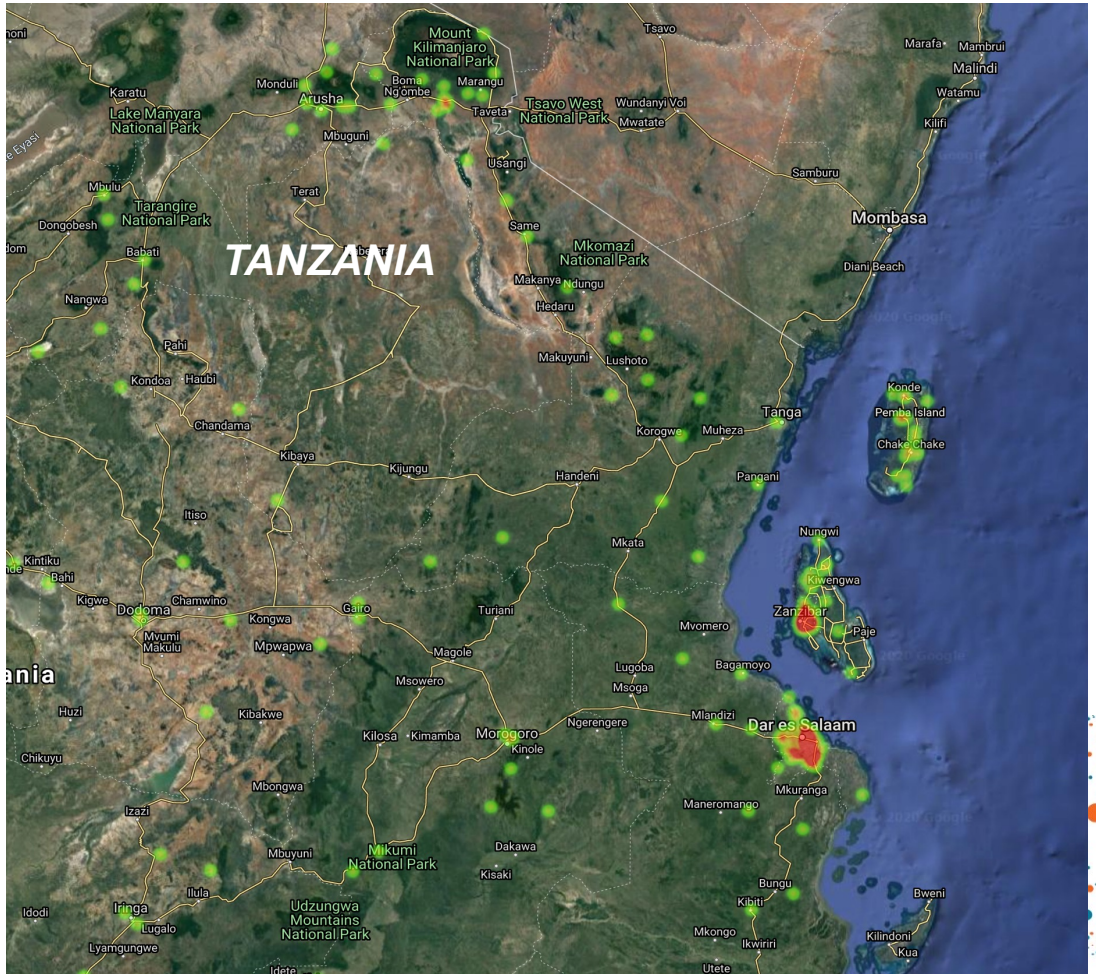
514
CHURCHES SERVING

NETHOPE

Pilot study: Tanzania

Nighttime & daytime sat. images (overlay)

"Ground truth" data: LSMS clusters



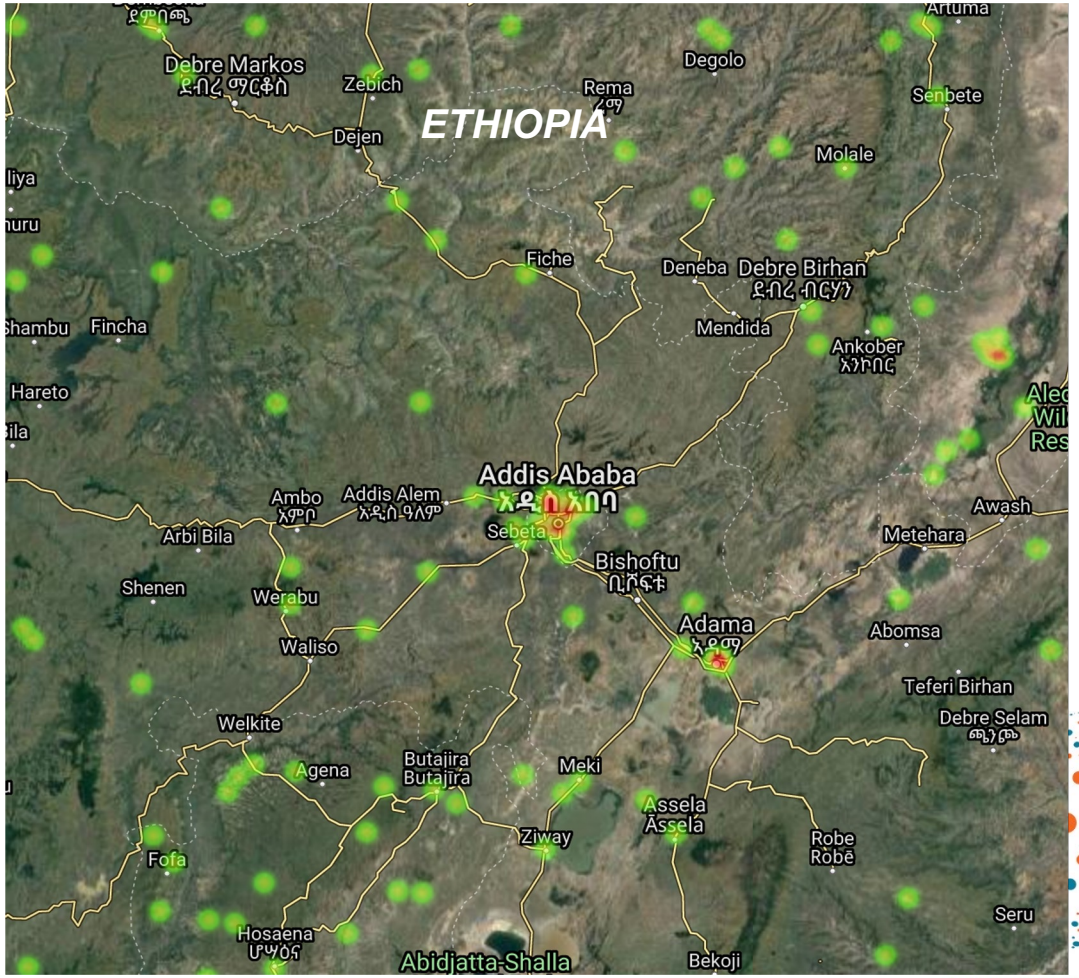
(source: Google Maps & NASA Worldview Earthdata; World Bank – LSMS/2015)

NETHOPE

Pilot study: Ethiopia

Nighttime & daytime sat. images (overlay)

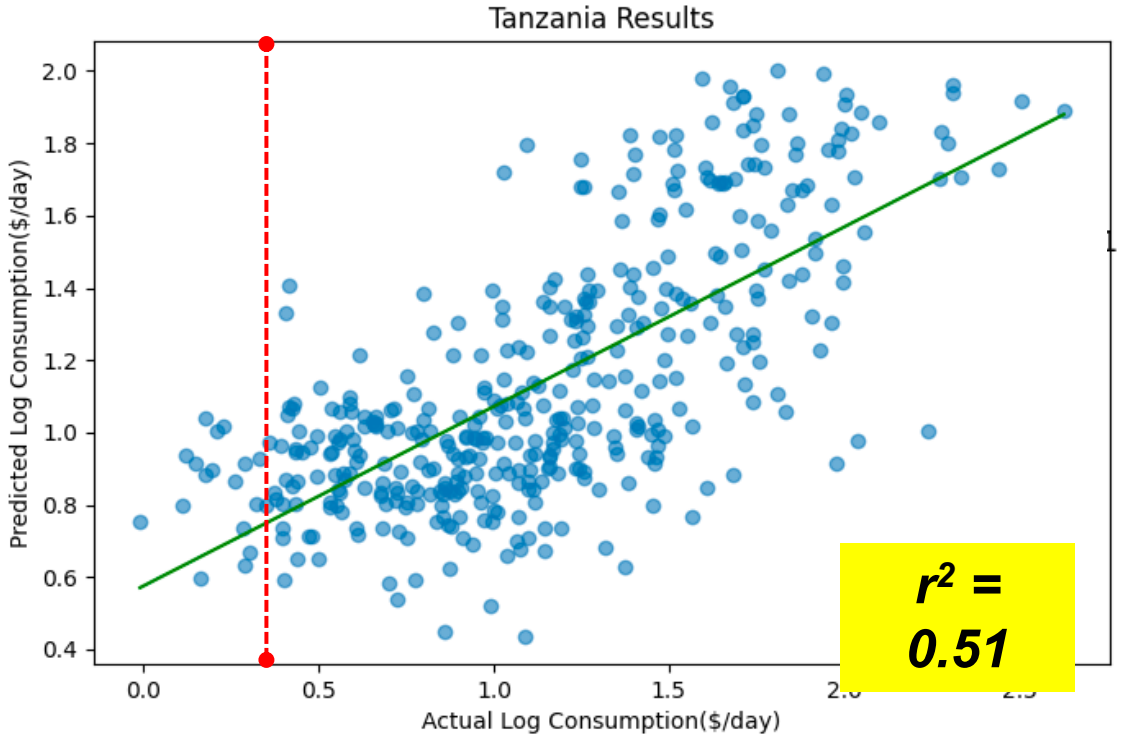
"Ground truth" data: LSMS clusters

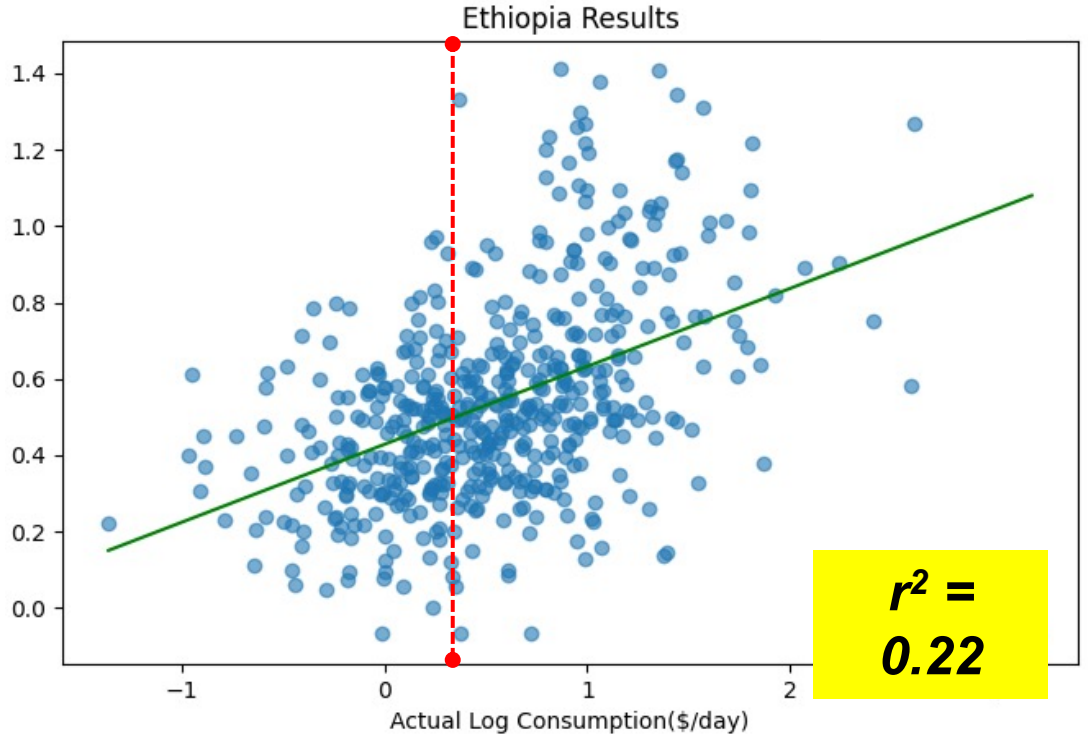


(source: Google Maps & NASA Worldview Earthdata; World Bank – LSMS/2015)

NETHOPE

Pilot study: results





Legend:

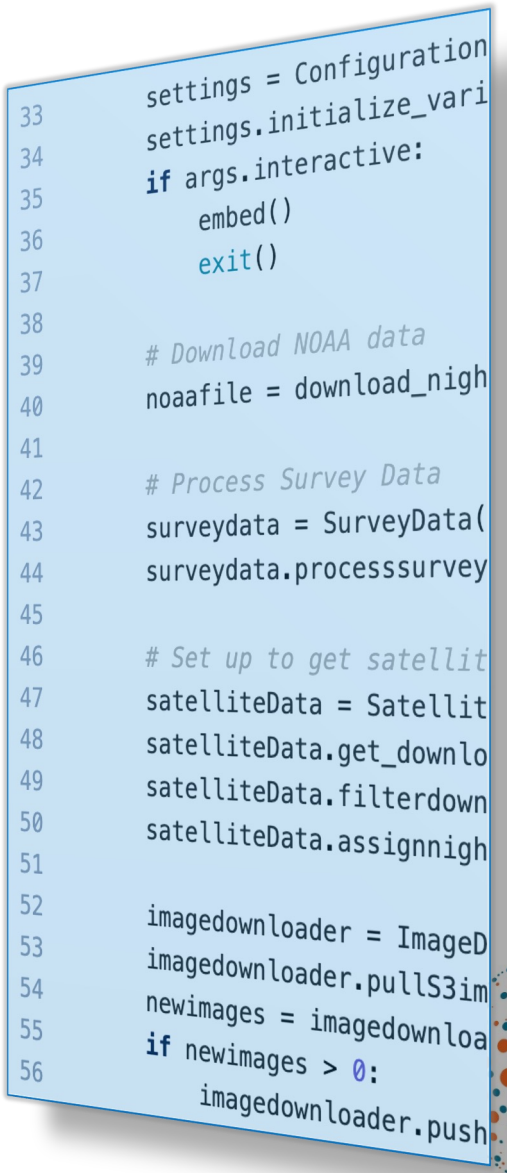
- international poverty line ($1.90/day),

r^2 *- coefficient of determination - informs us what % of the variation in avg. household consumption is explained by the model*

NETHOPE

Resources needed for the development, implementation, and maintenance of the solution.

- *Compute environment:*
  - *Python / AWS infrastructure (SageMaker + S3)*
- *Data sources:*
  - *Public data:*
    - *daytime- & –nighttime sat. images*
    - *training data: poverty indicators (e.g. LSMS)*
  - *Proprietary data:*
    - *training data: poverty indicators (e.g. PPI)*
- *Teams:*
  - *Data Science / Innovations Team*
  - *Global Program / Frontline Church Partners*

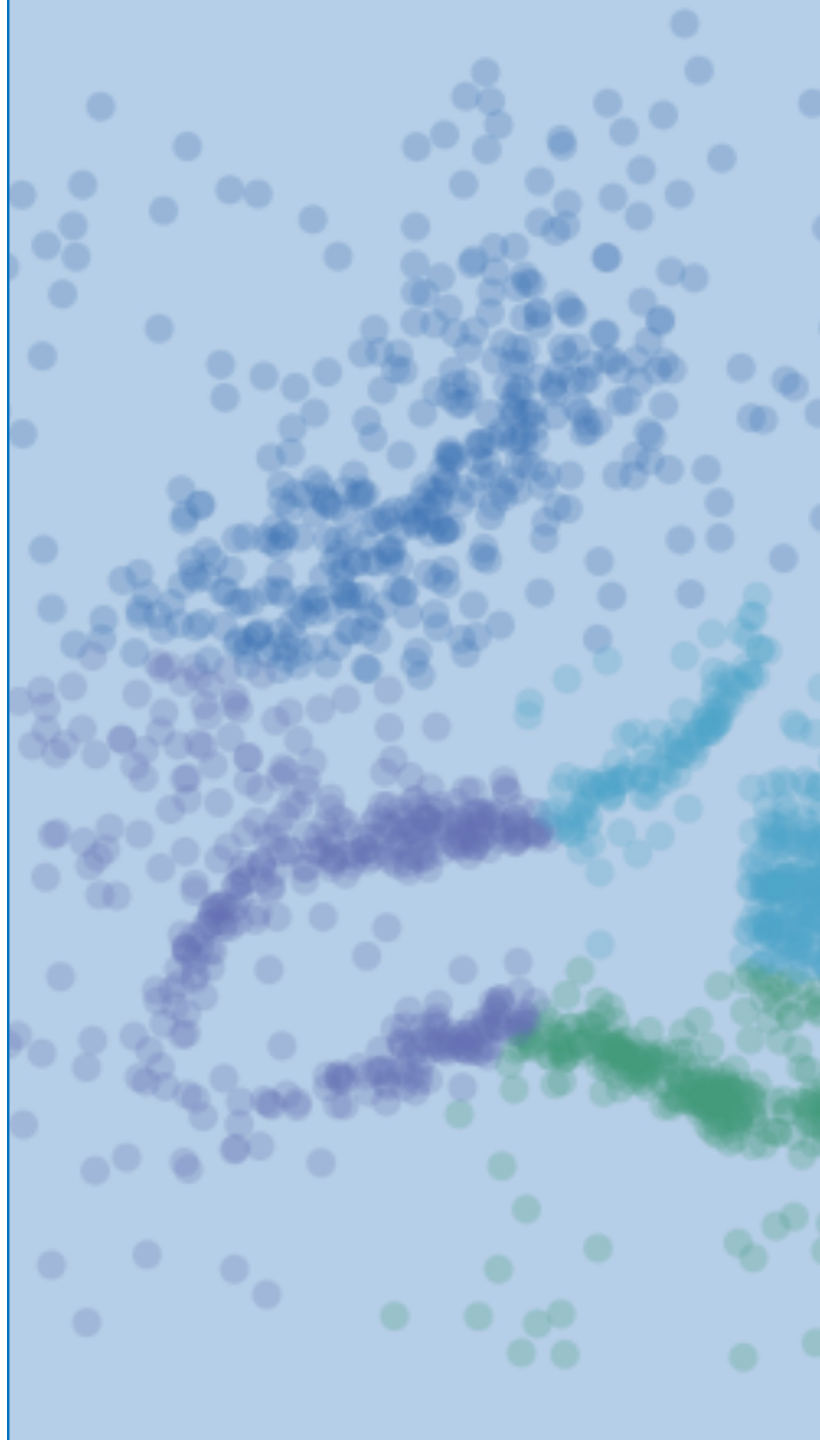


NETHOPE

Challenges & potential biases that AI may introduce:

Algorithm's predictive power is limited by:

1. *limited and/or outdated survey data – resulting in incorrect & low-fidelity predictions*
2. *nighttime images – poor differentiation for areas of extreme poverty*
3. ***features extracted from images not indicative of poverty (or of other parameters of interest)***
4. ***non-uniform geographical distribution of survey / reference data (e.g. overrepresentation of urban areas)***
5. ***quality & resolution of daytime-/nighttime- images***



Opportunities

1. *"Ground-truth" data combined from multiple sources*
2. *New data sources for transfer learning (e.g. infrared sat. images, aerial images, cellular data usage)*
3. ***ImageNet → "SatNet"***
4. ***Model training for better generalization across countries / regions***
5. ***Other improvements of the original algorithm (transfer training on full nighttime images, etc.)***

Our approach to maintaining the solution

1. ***Ensuring continues alignment with our stakeholder, that is frontline churches and national office partners, to understand which further insights they are looking for from this solution and direct feature enhancements.***

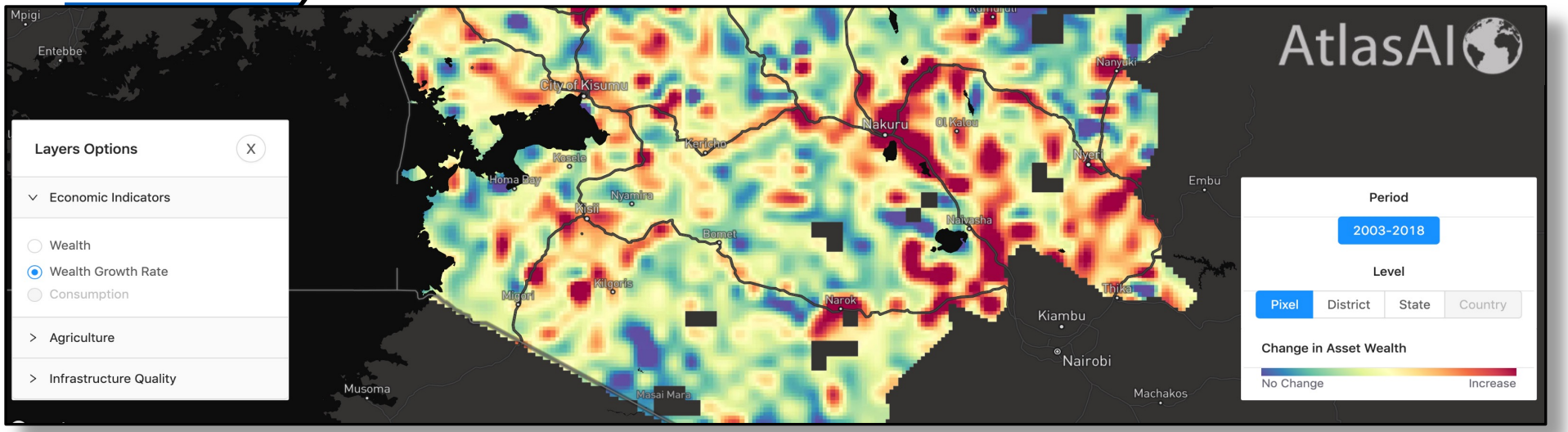
2. ***Streamline data acquisition (programmatic approach with access to API)***

3. ***Feedback loop mechanism with the broader community to incorporate both technical/algorithmic improvements and data-source updates.***

Partnership Opportunities

- *Opportunities to partner on additional data sources, such as:*
  - *HD satellite images at specific dates/times,*
  - *alternative data for transfer learning (e.g. infrared images)*
  - *alternative field-data*
- *(example: partnership with [Baylor University](#))*

- *Technology partners – remote sensing solutions (example: [AtlasAI.co](#))*



NETHOPE

Credits

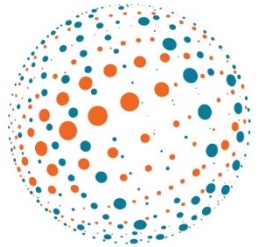
- *Nick Otto (Compassion International)*
- *Kate Williams / Mario Carrillo / Paul D'Angelo (Compassion International)*
- *Neal Jean – for helpful discussions & making the reference code repository publicly available*
- *Jatin Mathur – for making the reference code repository publicly available*

References

- ***Jean et al., "Combining satellite imagery and machine learning to predict poverty", Science (2016)***
- ***Xie, et al., "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping", Proc. AAAI (2016)***
- ***World Bank, Living Standards Measurement Study, www.microdata.worldbank.org***
- ***Jatin Mathur source code: github.com/jmather625/predicting-poverty-replication***
- ***NOAA-NGDC – nighttime satellite images, www.ngdc.noaa.gov***
- ***Google Static Maps API - daytime sat. images, cloud.google.com/maps-platform***

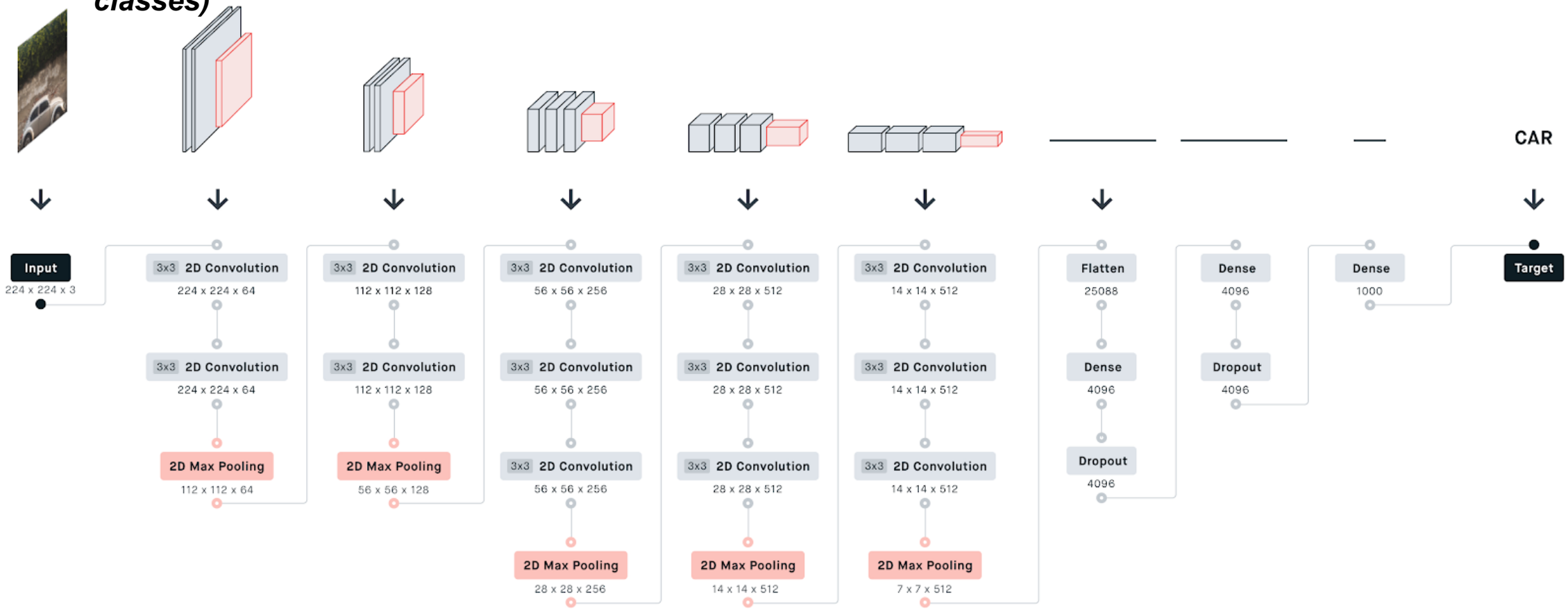
NETHOPE

Appendix



Convolutional Neural Network (VGG)

- *VGG (Visual Geometry Group) is one of the most popular CNN architectures*
- *Introduced by Simonyan & Zisserman in 2014 (“Very Deep Convolutional Networks for Large-Scale Image Recognition”)*
- *Achieves 92.7% top-5 test accuracy in ImageNet (a dataset of over 14 million images belonging to 1000 classes)*



Measuring Indicators for Resilience Analysis (MIRA

Helping Communities Identify and Plan for Risk

James Campbell

Catholic Relief Services

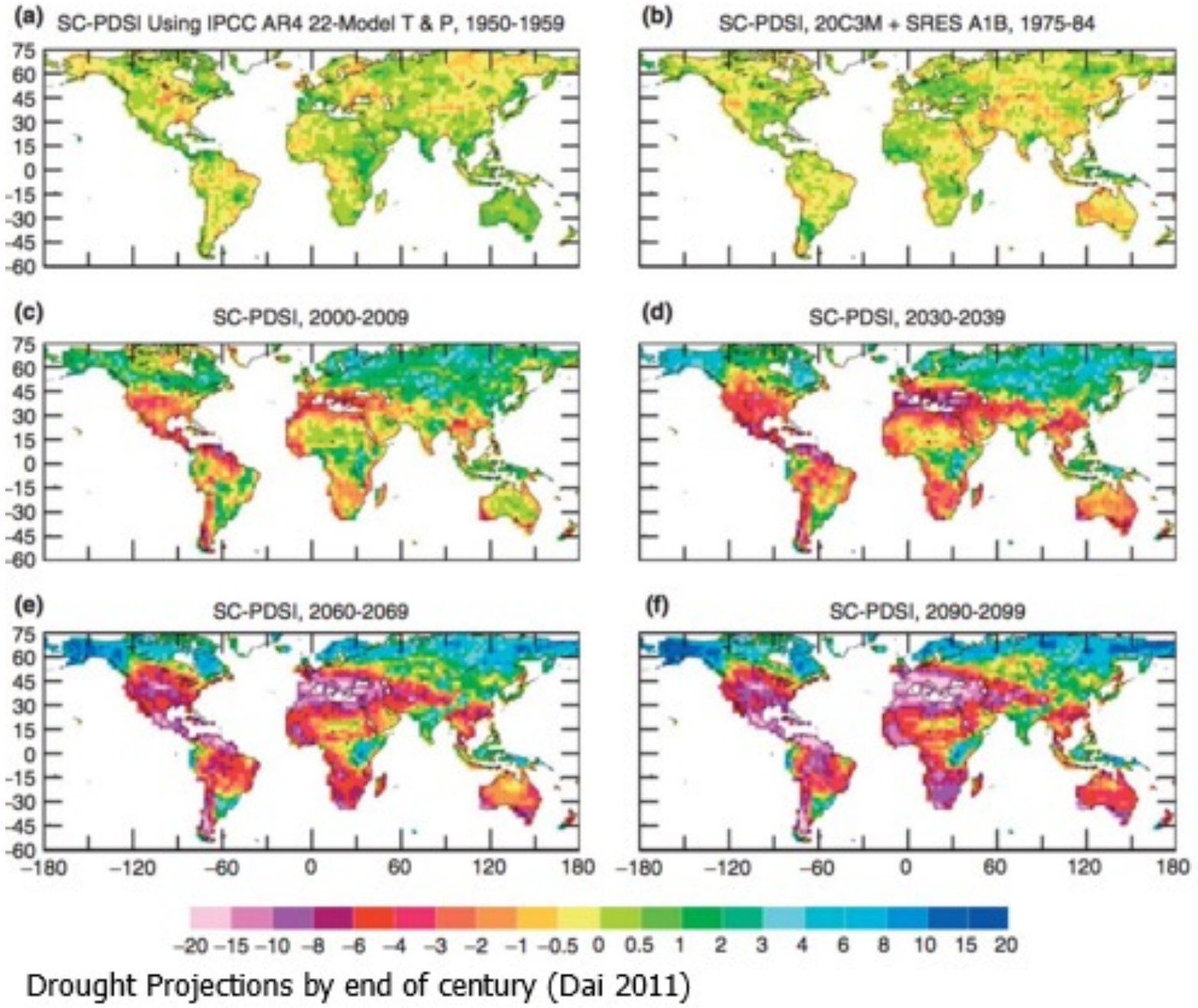
M&E Regional Technical Advisory for Southern Africa

y enumerator collecting monthly data. Photo - CRS

What problem are you trying to solve?

The goal of the Measurement Indicators for Resilience Analysis (MIRA) study was to develop a resilience measurement platform to measure how effective our programs are at increasing the resilience capacities of participants to shocks and stresses.

**Resilience concept - "the capacity that ensures adverse stressors and shocks do not have long-lasting adverse development consequences" (FSIN, RM-TWG, 2014, p. 6).*



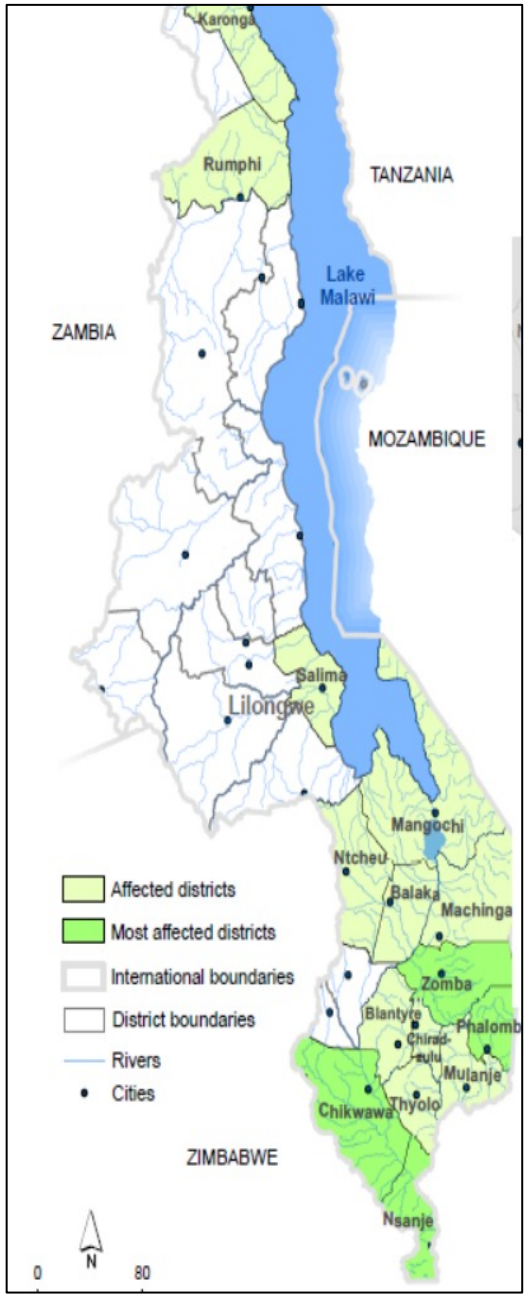
How is the problem being addressed today?

- *In the first few months of 2015, flooding displaced an estimated 230,000 people, damaged about 64,000 hectares of land and destroyed the asset wealth of many.*

- *Proof of concept by CRS and Cornell University with 580 households in Chikwawa District. Expanded to 2200 household in two adjacent districts in phase II (entire DFAP project area).*

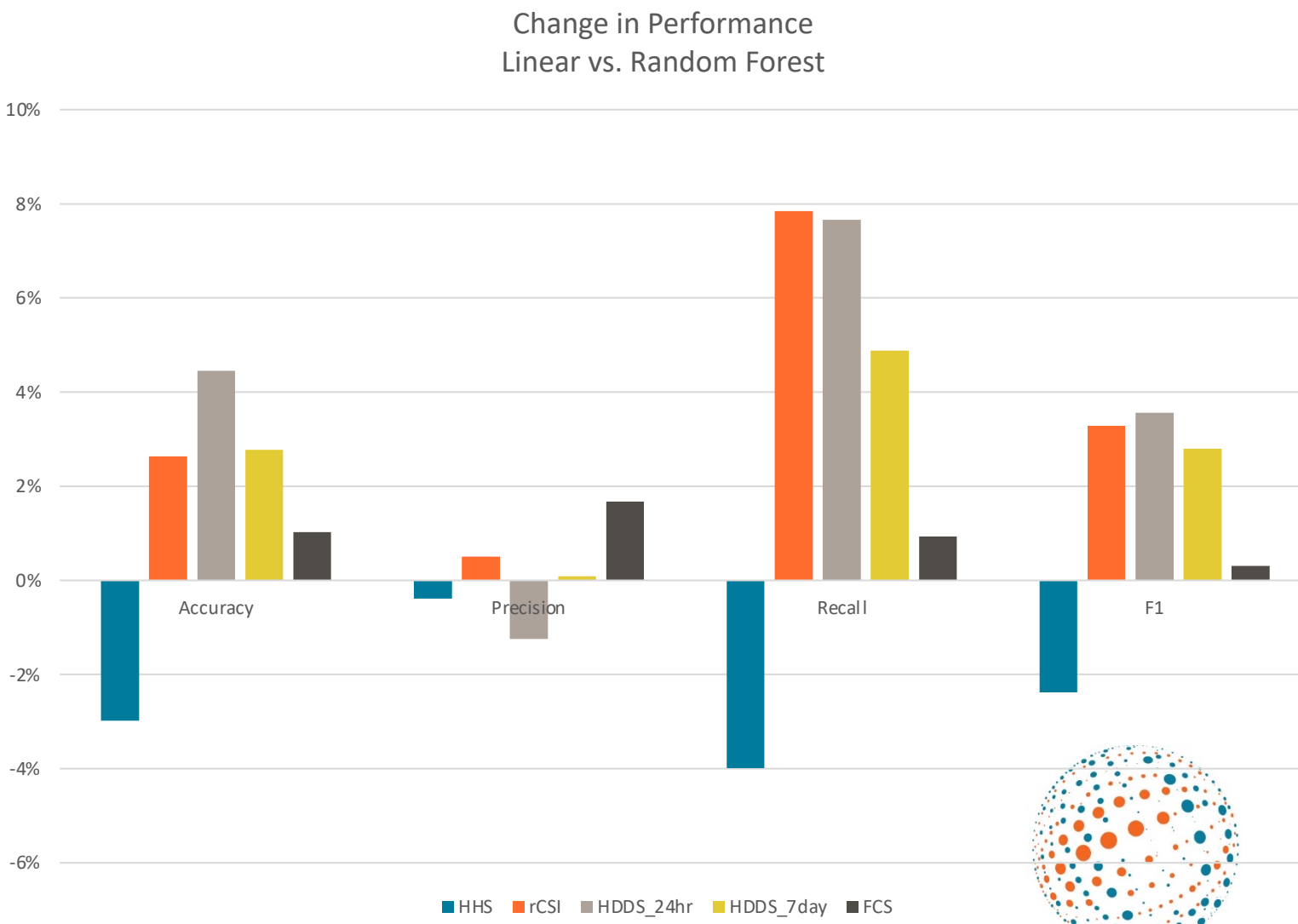
- *Approach was developed to focus specifically on the resilience building aims of the UBALE project and explain variations in recovery from shocks.*

2015 Flood



Why is AI better than the current solution?

- ***In terms of analytics for the study we used several different approaches such as regression analysis & machine learning;***
- ***Initially, we did not include any aspects of ML or AI , this occurred after we had implemented the protocol and realized that we could employ ML techniques***
- ***A comparison of predictive performance between using a standard linear model vs. the random forest model***



NETHOPE

What is the solution?

Evaluated two machine learning algorithms (LASSO and Random Forest) to predict responses to shocks & inform targeting

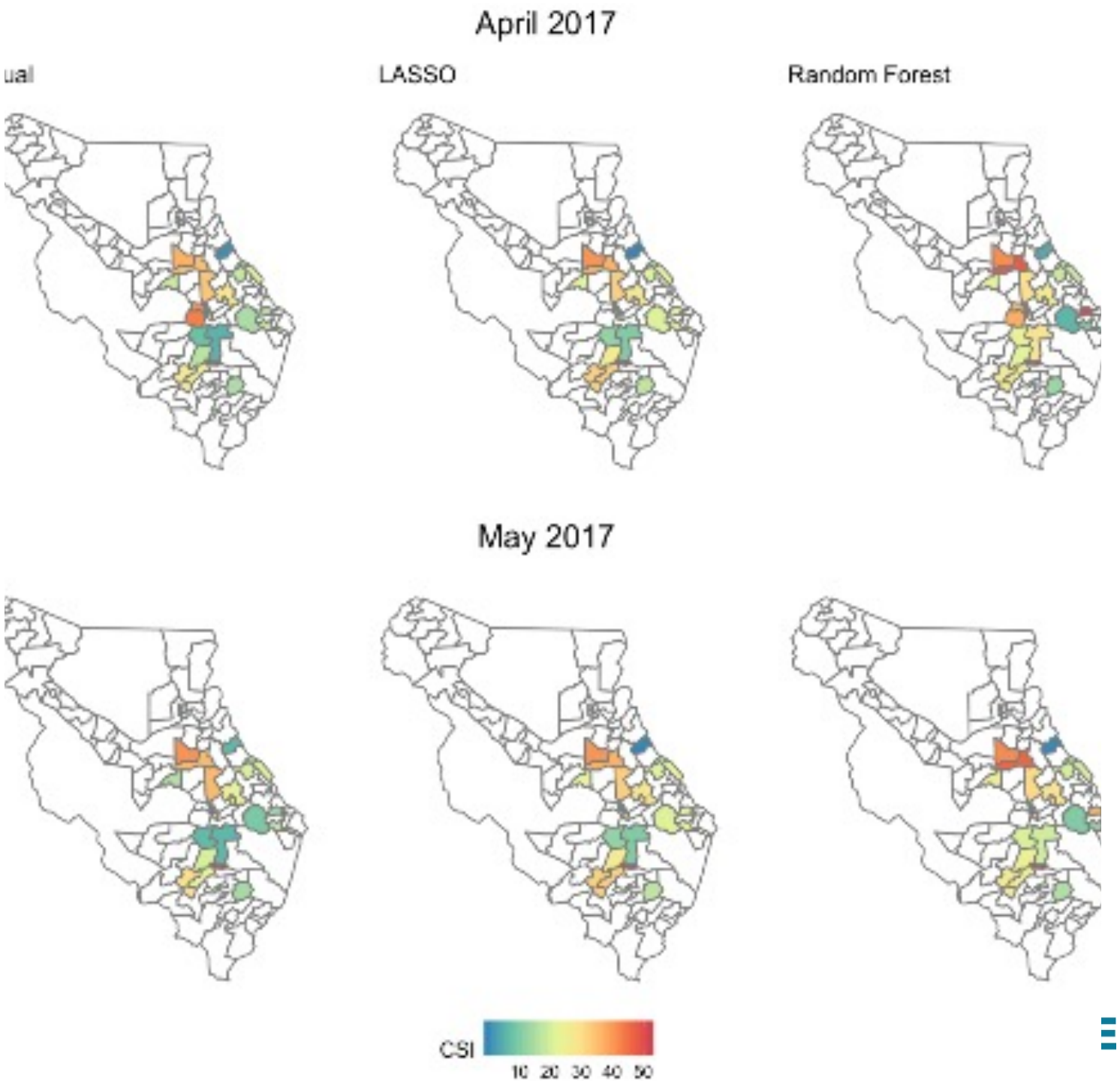
	LASSO	Random Forest
	CSI Last Month	CSI Last Month
	Group Village Head	Group Village Head
	Received Assistance from Government Last Month	Age
	Received Assistance as Food Last Month	Education
	Distance to Drinking Water	Distance to Drinking Water
	Quality of Floor	Quality of Roof
	Sold Assets Last Month	Dietary Diversity Score
	Purchased Assets Last Month	Land Farmed (ha)
	Lives in Flood Plain	Lives in Flood Plain
	Experienced Drought Last Month	Pregnant or Nursing Household Member
N	4308	4308
R^2	56.4 %	55.6 %

Sample restricted to training set, (June-March) and tested on April-May

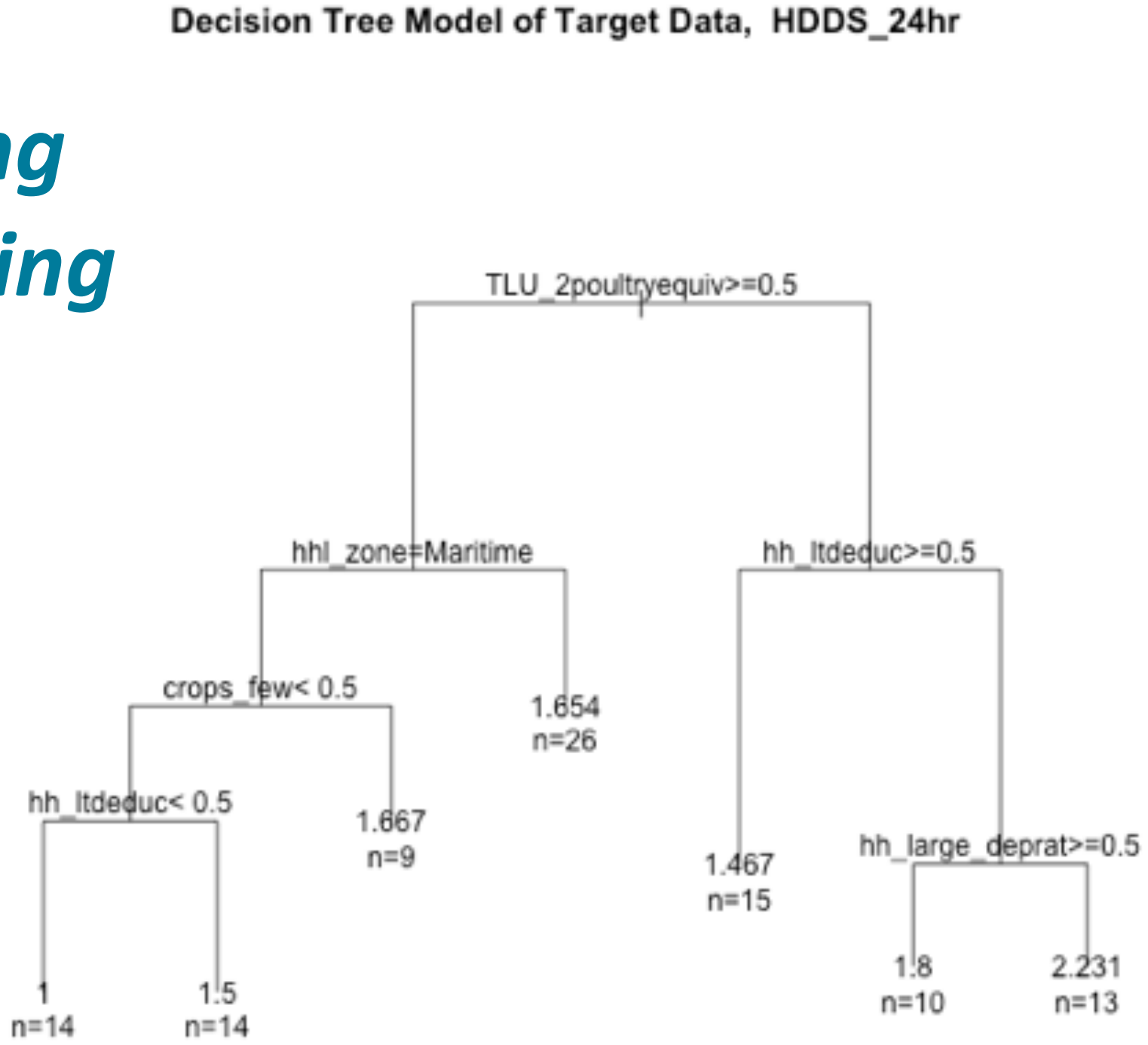
Predictive Capacity

Good predictors include

- ***location of fields***
- ***proximity to flood plains***
- ***sex of household head***
- ***previous shocks experienced***



Targeting using Machine Learning



What kinds of data does your solution need?

Focus of variables

1. *Well-being outcomes (e.g., food security indicators)*
2. *Shocks experienced that have adverse effects*
3. *Capacities and characteristics that may prevent adverse effects and/or enable recovery*

Data Collection Process

1. *MIRA involves first a baseline survey and then monthly data collection using a Computer Assisted Personal Interviewing (CAPI) platform*
2. *Trained enumerators embedded in the communities*
3. *Panel data collected from same households over time (2100 HHs across 3 districts in Southern Malawi)*

Empirical challenge

- *Resilience focus*
- *Low burden*
- *Generalizable*
- *Accessible and useful*

NETHOPE

What resources do you need to support the development, implementation, and maintenance of the solution?

Technical

- ***Cornell University: Supports current research and analytics, quantifying households' resilience for the purpose of impact evaluation, targeting and forecasting***

- ***University of Illinois and University of Texas: test the reliability of these solutions and how this data can be used to make out of sample predictions***

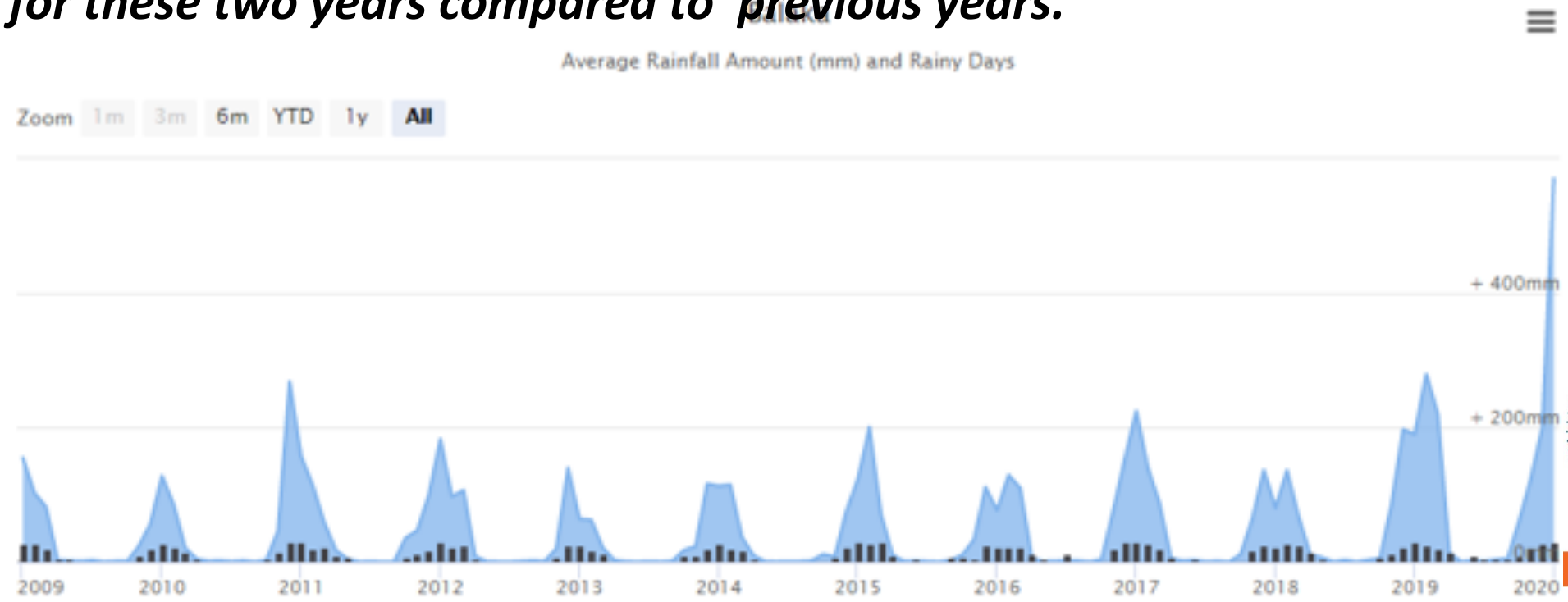
Non-technical

- ***Field Supervisor***
- ***Community Engagement - Community leaders, Village Development Committee & Village Civil Protection Committee Members***
- ***Embedded Enumerators***

NETHOPE

What potential biases could AI introduce or amplify in your context?

- ***The amount of rainfall received from 2016 and 2018 was far less than the amount of rain received in 2017, 2019, and the first month of 2020 (possibly an outlier).***
- ***Results may be biased if we use 2016 – 2018 data to train the model and predicting outcomes for the 2019 – 2020 seasons considering that the district has experienced max rainfall for these two years compared to previous years.***



IETHOPE

What is your approach to maintaining the solution?

- ***CRS has been awarded a Microsoft AI for Humanitarian Action grant.***

- ***Currently working with Microsoft to develop a plug and play machine learning solution that any project which implements the MIRA protocol can utilize to identify household characteristics driving food insecurity and predict future shock outcomes using a comprehensive set of algorithms.***

PARTNERSHIPS

- ***Microsoft is providing technical support (e.g. data scientists)***
- ***Will also maintain relationships with the aforementioned research institutions to guide the associated learning***

NETHOPE

Next Steps

Opportunities

- ***Expanded and established a cohort in the Grand Sud in Madagascar in 2018, implementing the MIRA protocol***

- ***Initiating partnership with USAID, the World Bank, and Concern Worldwide to scale up MIRA in 9 additional districts in Southern Malawi***

- ***GoM has expressed an interest in adopting the MIRA protocol to track progress of its National Resilience Strategy***

NETHOPE