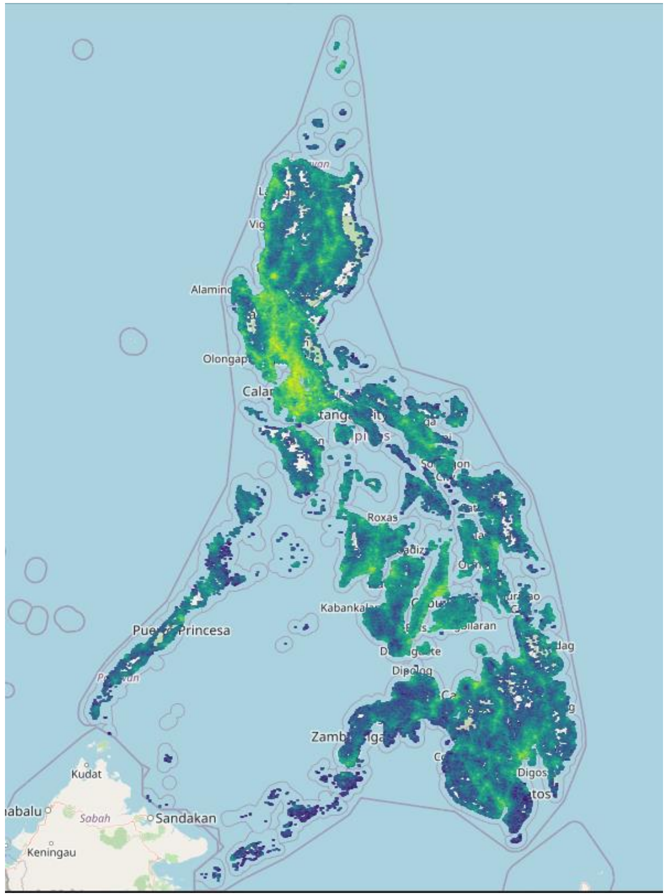
Poverty Mapping
for Public Policy Impact Evaluation

By: Teo Fwu Chyi

Sep 2024

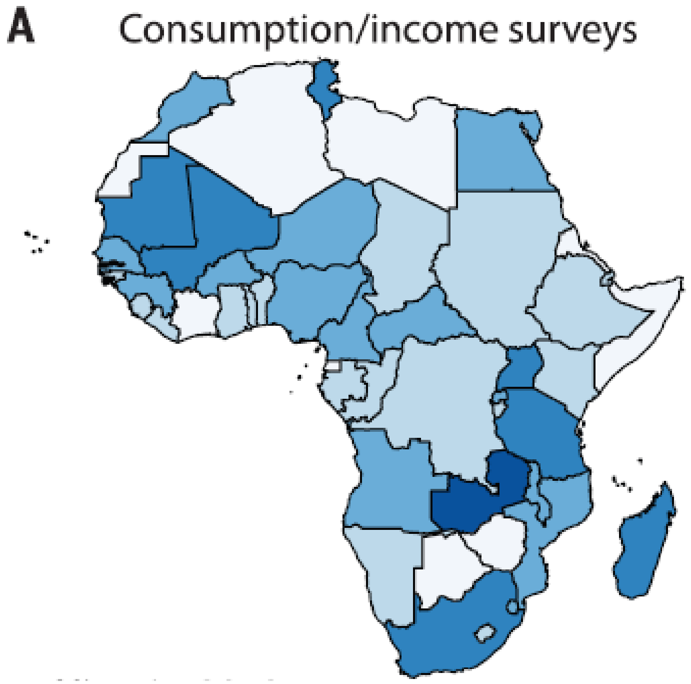
Overview

- Open-source geo-data to map poverty levels
  - Satellite imagery
  - Geographic Information System (GIS) data
- Computer Vision + Machine Learning
- Compare different approaches → feasibility and practicality
- Illustrative example: Evaluate effectiveness of PPPP program in the Philippines



Poverty Mapping

- Understand geographic distribution of people living in poverty



Income distribution map (Source: Jean et al, 2016)

Demographic and Health Survey (DHS)

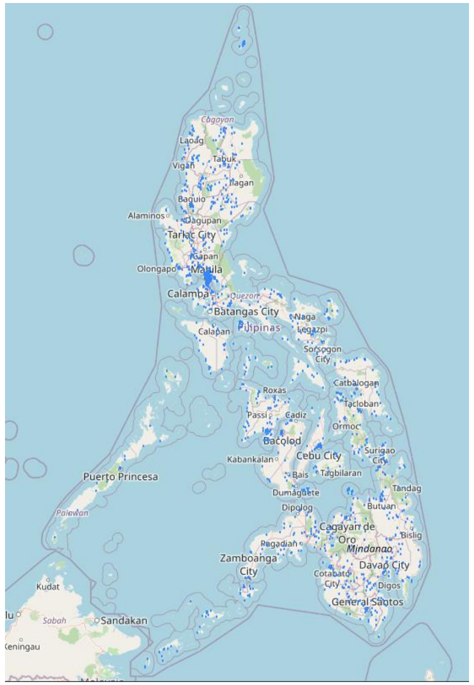
- Costly: 1.5M USD
- Low frequency: 3~5 years
- Low spatial coverage

(Source: Tingzon et al, 2019)

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790-794.
Tingzon, I., Orden, A., Go, K. T., Sy, S., Sekara, V., Weber, I., ... & Kim, D. (2019). Mapping poverty in the Philippines using machine learning, satellite imagery, and crowd-sourced geospatial information. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 425-431.

Poverty Mapping

- Understand geographic distribution of people living in poverty



Philippines 2022 DHS

Demographic and Health Survey (DHS)

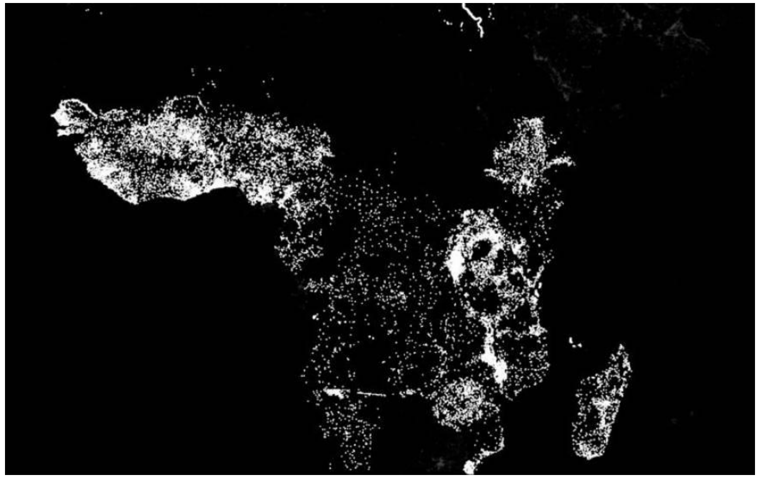
- Costly: 1.5M USD
- Low frequency: 3~5 years
- Low spatial coverage

(Source: Tingzon et al, 2019)

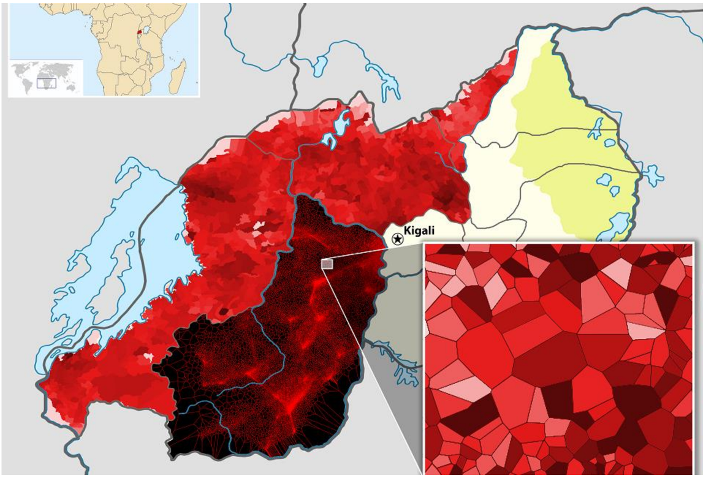
Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790-794.
Tingzon, I., Orden, A., Go, K. T., Sy, S., Sekara, V., Weber, I., ... & Kim, D. (2019). Mapping poverty in the Philippines using machine learning, satellite imagery, and crowd-sourced geospatial information. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 425-431.

Poverty Mapping

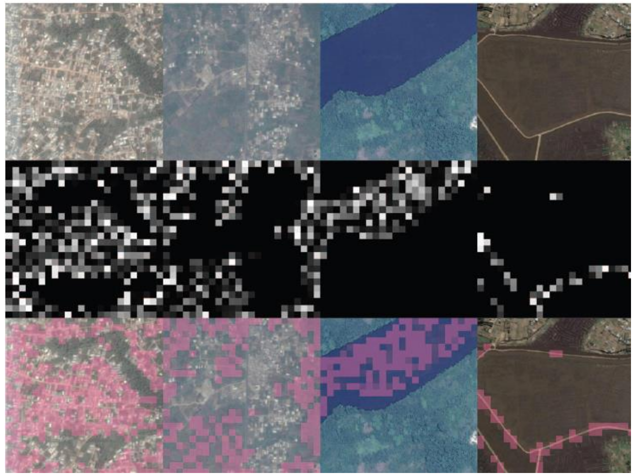
- Machine learning on geodata?



Nightlight (Source: Stanford School of Earth 2016)



Mobile data (Source: Blumenstock et al, 2015)



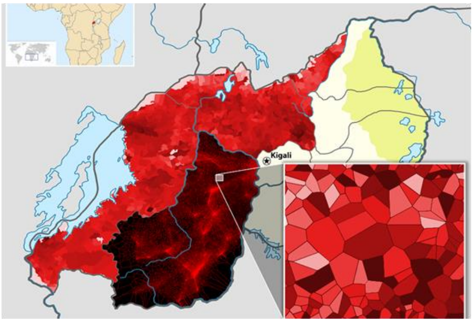
Daytime satellite images trained on nightlight as feature extractor
(Source: Jean et al, 2016)

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, *350*(6264), 1073-1076.
Stanford researchers use dark of night and machine learning to shed light on global poverty. Stanford School of Earth, Energy & Environmental Sciences. (2016, March 17). https://pangea.stanford.edu/d7-archive/sesd7/news/stanford-researchers-use-dark-night-and-machine-learning-shed-light-global-poverty/index.html
Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790-794.

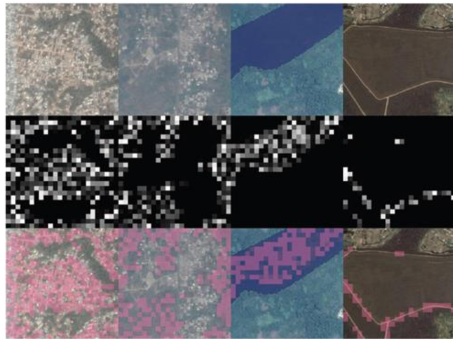
Poverty Mapping

- Feasibility?
  - Nightlight → freely available
  - Mobile data → proprietary and limited
  - Daytime satellite images → high resolution can be costly
    - Jean et al, 2016 → 2.5m per pixel
    - Tingzon, 2019 → 1m per pixel (Google Maps API)
- Organizations/governments with limited funds
  - Prohibitive costs - does approach still work with free, lower quality data?
  - Data science expertise?

Nightlight (Source: Stanford School of Earth 2016)



Mobile data (Source: Blumenstock et al, 2015)

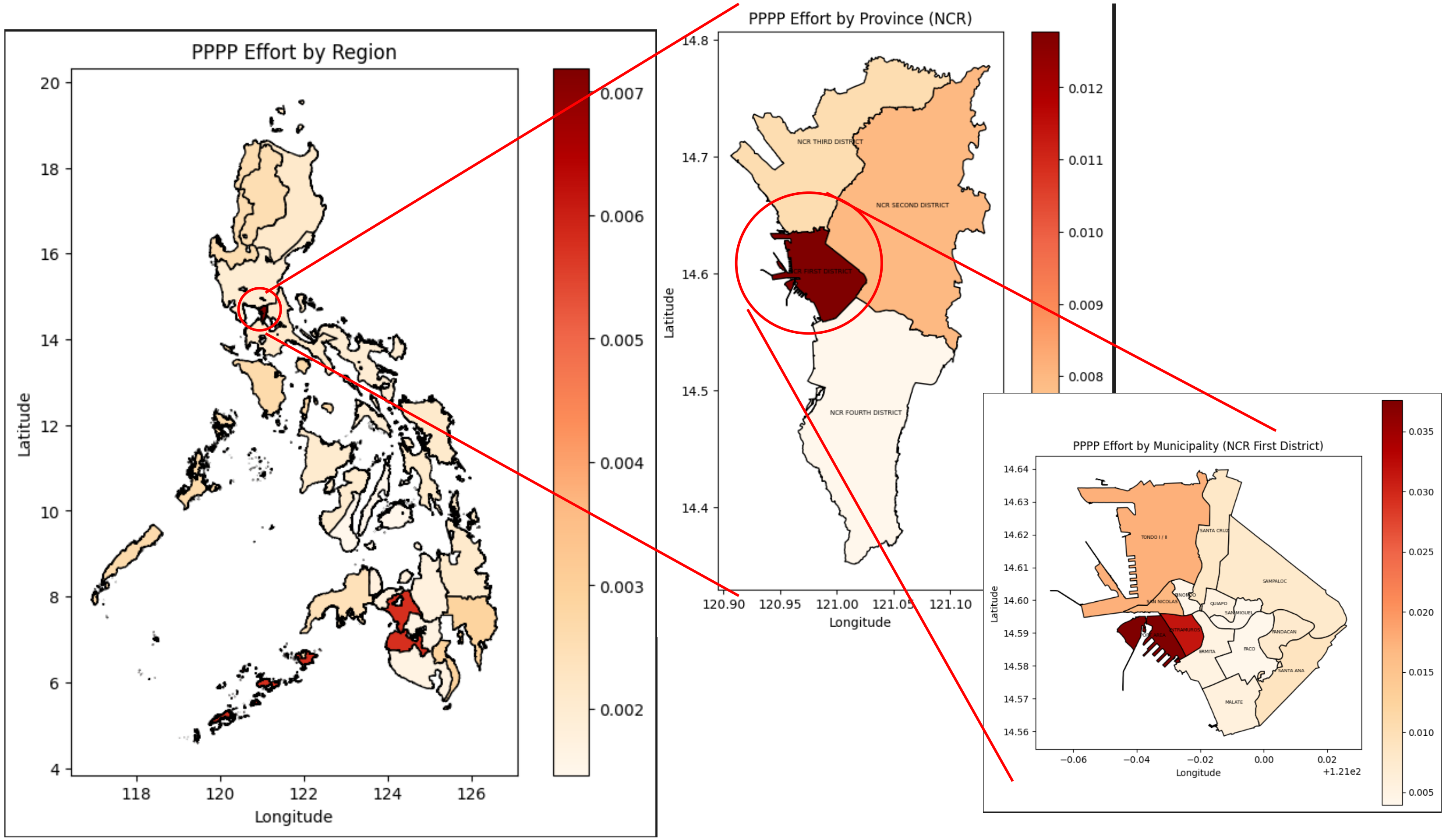


Daytime satellite images trained on nightlight as feature extractor
(Source: Jean et al, 2016)

Objective

Investigate feasibility of applying machine learning based poverty mapping approaches

Approach – Geo-temporal Data



Hierarchical geography (2020-2023)

Approach – Geo-temporal Data

Open Source Data

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Approach – Geo-temporal Data

<u>Open Source Data</u>

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Lower bit images
8-bit instead of 16-bit images for paid versions

Lower resolution
10m/pixel instead of 1~2.5m/pixel in original studies

Approach – Open Source Material

Open Source Data

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Open Source Software

Python
geopandas, rasterio, ThinkingMachines codes, camelot, timm

QGIS

Approach – Overall Strategy

Open Source Software

Python
geopandas, rasterio, ThinkingMachines codes, camelot, timm

QGIS

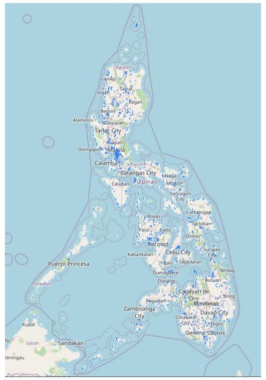
Open Source Data

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)



Approach – Overall Strategy

Open Source Software

Python
geopandas, rasterio, ThinkingMachines codes, camelot, timm

QGIS

Open Source Data

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

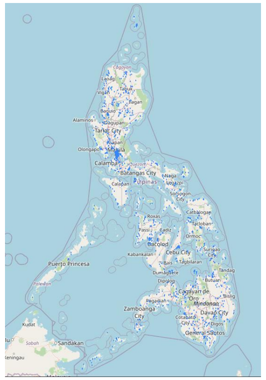
Daytime satellite images
via WMS on QGIS (EOX)

Train

Model

Wealth Index (DHS) at limited locations

Approach – Overall Strategy



Open Source Software

Python
geopandas, rasterio,
ThinkingMachines codes,
camelot, timm

QGIS

Open Source Data

OOKLA
Broadband and mobile
data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby
schools, roads, markets

Nightlight
Earth Observation Group
(EOG)

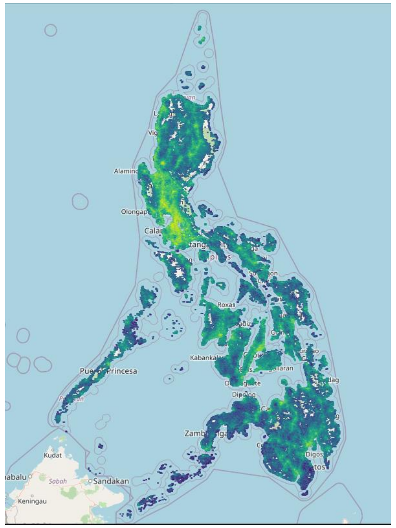
Daytime satellite images
via WMS on QGIS (EOX)

Train

Model

Wealth Index (DHS) at limited locations

Predict



Map poverty levels of entire country

Approach – Overall Strategy

Open Source Software

Python
geopandas, rasterio,
ThinkingMachines codes,
camelot, timm

QGIS

Open Source Data

OOKLA
Broadband and mobile
data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby
schools, roads, markets

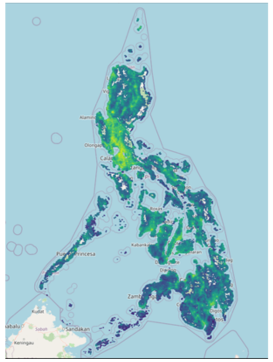
Nightlight
Earth Observation Group
(EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Train Model

Wealth Index (DHS) at limited locations

Predict



Map poverty levels of entire country

Sample Evaluation

Pantawid
Pamilyang
Pilipino
Program

Conditional Cash Transfer (CCT)
Poverty alleviation strategy

Data Acquisition

Data Acquisition

Pantawid Pamilyang Pilipino Program

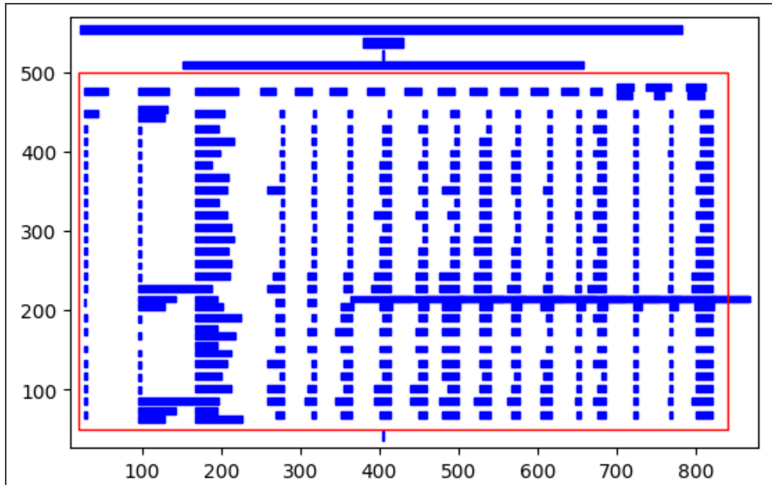
MONTHLY REPORT ON PANTAWID PAMILYA COVERAGE 26

Appendix

Appendix 1. Number of Active Pantawid Pamilya Beneficiary Households By Region, Province, City/Municipality, and Set

REGION	PROVINCE	MUNICIPALITY	SET 1	SET 2	SET 3	SET 4	SET 5	SET 6	SET 7	SET 8	SET 9	FNSP	HSF	IPs in GIDA	Itinerant IPs	Grand Total
NCR	NCR FIRST DISTRICT	BINONDO	2	-	1	-	-	56	3	-	-	-	39	-	-	101
		ERMITA	-	1	1	51	12	2	7	-	-	-	114	-	-	188
		INTRAMUROS	-	1	-	83	-	2	122	-	-	-	47	-	-	255
		MALATE	3	-	2	86	5	17	668	24	-	-	158	-	1	964
		PACO	-	1	3	614	79	60	219	14	-	-	43	-	-	1,033
		PANDACAN	2	-	-	354	8	18	464	27	-	-	33	-	-	906
		PORT AREA	1,251	7	8	241	19	2,278	212	19	63	-	113	-	-	4,211
		QUIAPO	-	3	1	14	5	3	359	-	-	-	34	-	-	419
		SAMPALOC	7	3	7	2,681	677	177	249	7	-	1	217	-	-	4,026
		SAN MIGUEL	-	-	-	55	5	7	122	-	-	-	10	-	-	199
		SAN NICOLAS	4	1	2	153	5	16	1,027	7	1	-	159	-	-	1,375
		SANTA ANA	5	3	3	324	20	45	2,053	19	-	-	90	-	-	2,562
		SANTA CRUZ	5	1	8	479	16	70	936	22	-	-	306	-	-	1,843
		TONDO I / II	555	18	21	5,038	128	16,953	1,282	71	3	-	346	-	-	24,415
	NCR FIRST DISTRICT Total		1,834	39	57	10,173	979	19,704	7,723	210	67	1	1,709	-	1	42,497

camelot



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	REGIONPROVINCEMUNICIPALITY	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	FNSP	HSF	IPsinGIDA	ItinerantIPs	GrandTotal
0	NCR FIRST			NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	NCR \nBINONDO	2	-	1	-	-	56	3	-	-	-	39	-	-	101
2	DISTRICT			NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	ERMITA	-	1	1	51	12	2	7	-	-	-	114	-	-	188
4	INTRAMUROS	-	1	-	83	-	2	122	-	-	-	47	-	-	255
5	MALATE	3	-	2	86	5	17	668	24	-	-	158	-	1	964
6	PACO	-	1	3	614	79	60	219	14	-	-	43	-	-	1,033
7	PANDACAN	2	-	-	354	8	18	464	27	-	-	33	-	-	906
8	PORT AREA	1,251	7	8	241	19	2,278	212	19	63	-	113	-	-	4,211

Custom cleaning

	REGION	PROVINCE	MUNICIPALITY	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	FNSP	HSF	IPsinGIDA	ItinerantIPs	GrandTotal
1	NCR	NCR FIRST DISTRICT	BINONDO	2	0	1	0	0	56	3	0	0	0	39	0	0	101
3	NCR	NCR FIRST DISTRICT	ERMITA	0	1	1	51	12	2	7	0	0	0	114	0	0	188
4	NCR	NCR FIRST DISTRICT	INTRAMUROS	0	1	0	83	0	2	122	0	0	0	47	0	0	255
5	NCR	NCR FIRST DISTRICT	MALATE	3	0	2	86	5	17	668	24	0	0	158	0	1	964
6	NCR	NCR FIRST DISTRICT	PACO	0	1	3	614	79	60	219	14	0	0	43	0	0	1033

2020-2023 reports

- **14 reports, 50~60 pages of tables**
- **Number of enrolled households by region, province, municipality**

Data Acquisition

<u>Open Source Data</u>

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Wealth Index (DHS) at limited locations

ThinkingMachines code

Data Acquisition

<u>Open Source Data</u>

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Wealth Index (DHS) at limited locations

Adapt ThinkingMachines code

+

Locating dump files for past data

Data Acquisition

<u>**Open Source Data**</u>

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

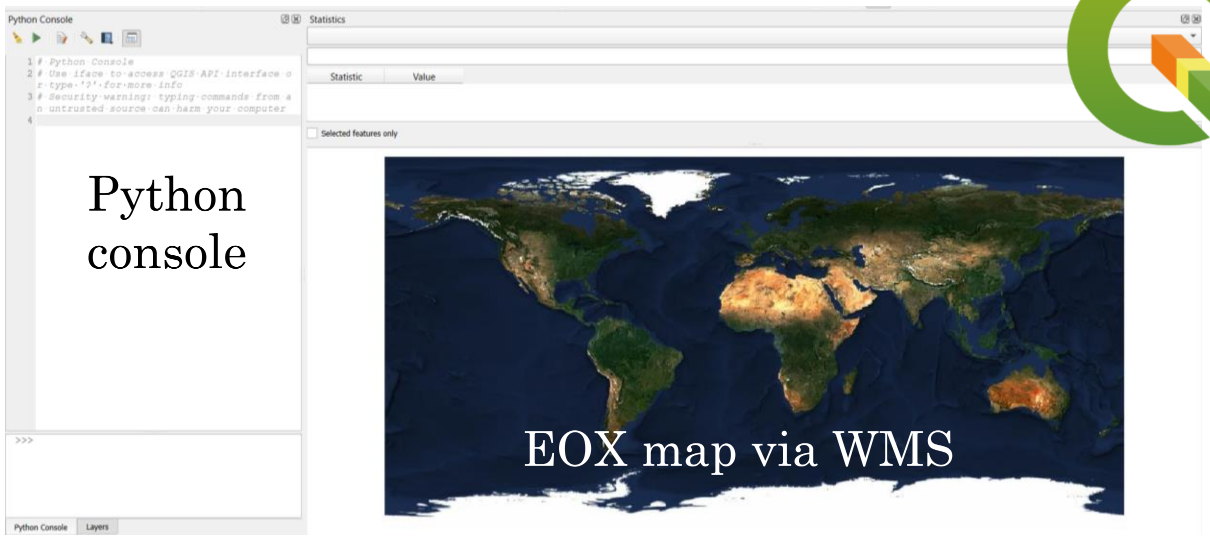
Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Wealth Index (DHS) at limited locations

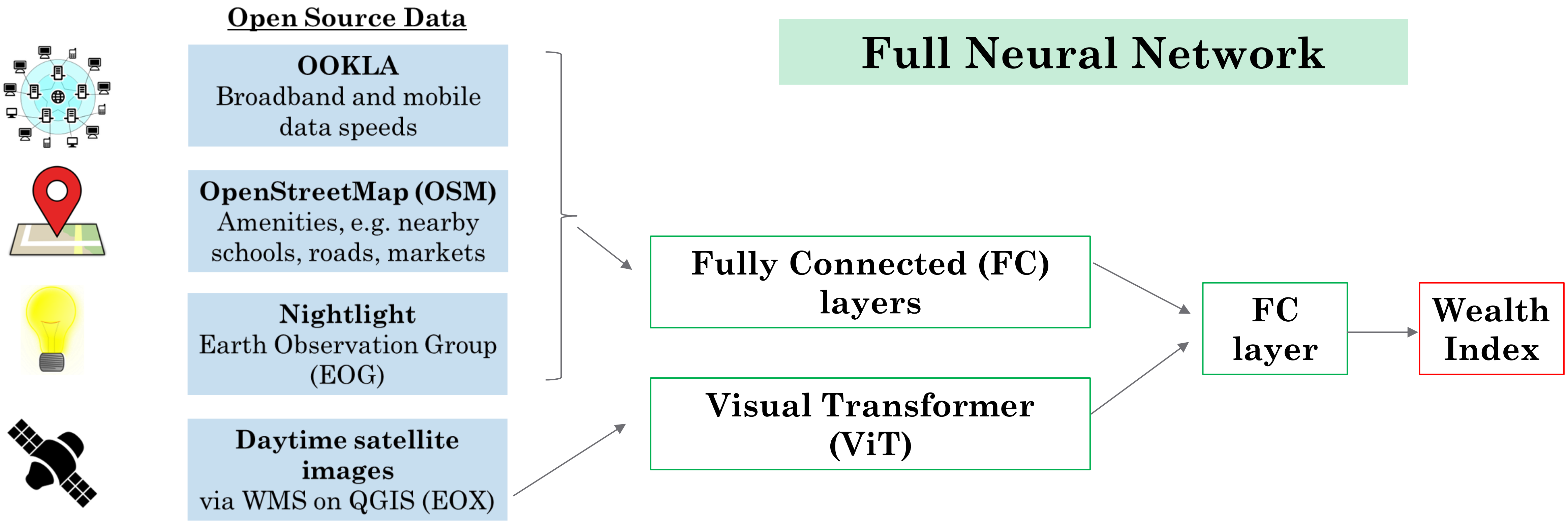
WMS via QGIS for EOX cloudless Sentinel-2 images + Python script

~45,000 image tiles per year

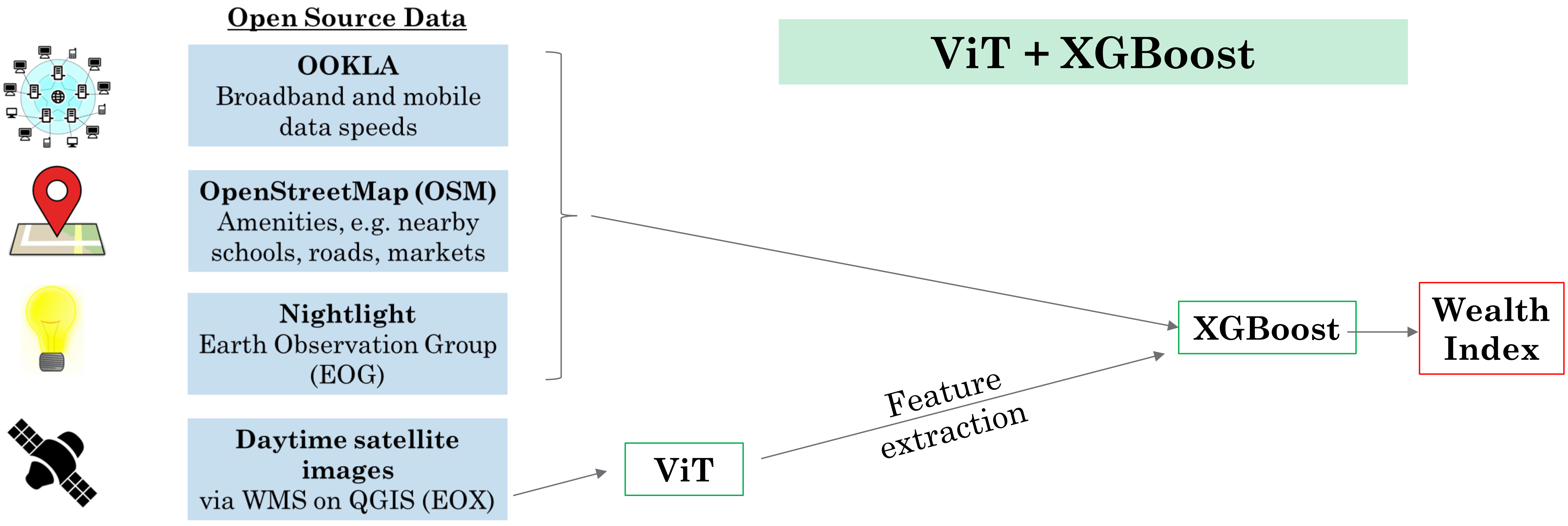


Model Training

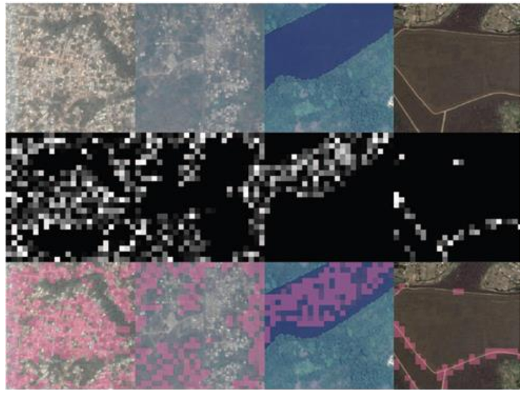
Models Considered



Models Considered



Models Considered



Daytime satellite images trained on nightlight as feature extractor
(Source: Jean et al, 2016)

ViT (on nightlight) + XGBoost

<u>Open Source Data</u>

OOKLA
Broadband and mobile data speeds

OpenStreetMap (OSM)
Amenities, e.g. nearby schools, roads, markets

Nightlight
Earth Observation Group (EOG)

Daytime satellite images
via WMS on QGIS (EOX)

Pre-trained

ViT

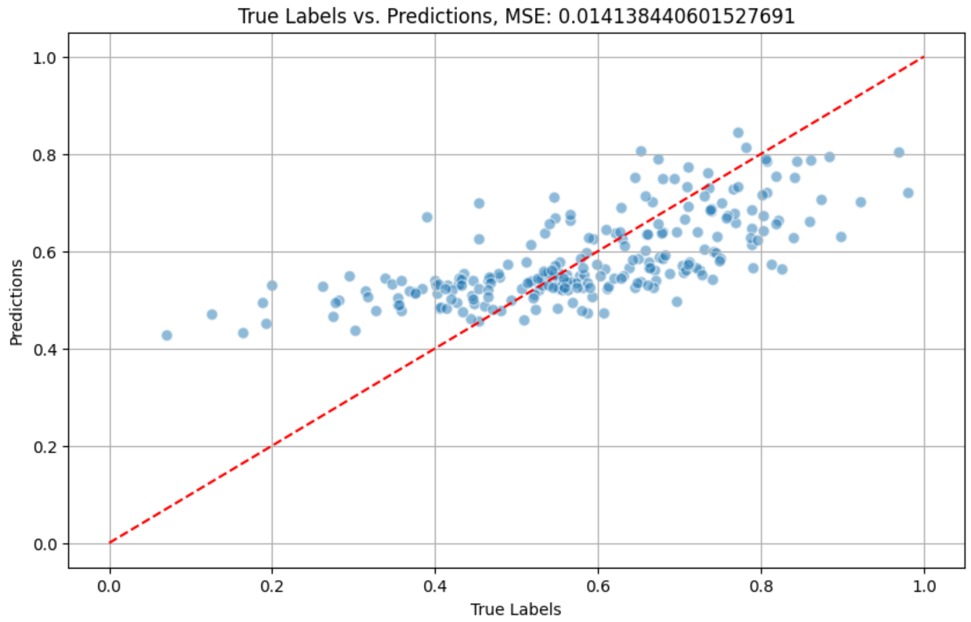
Feature extraction

XGBoost

Wealth Index

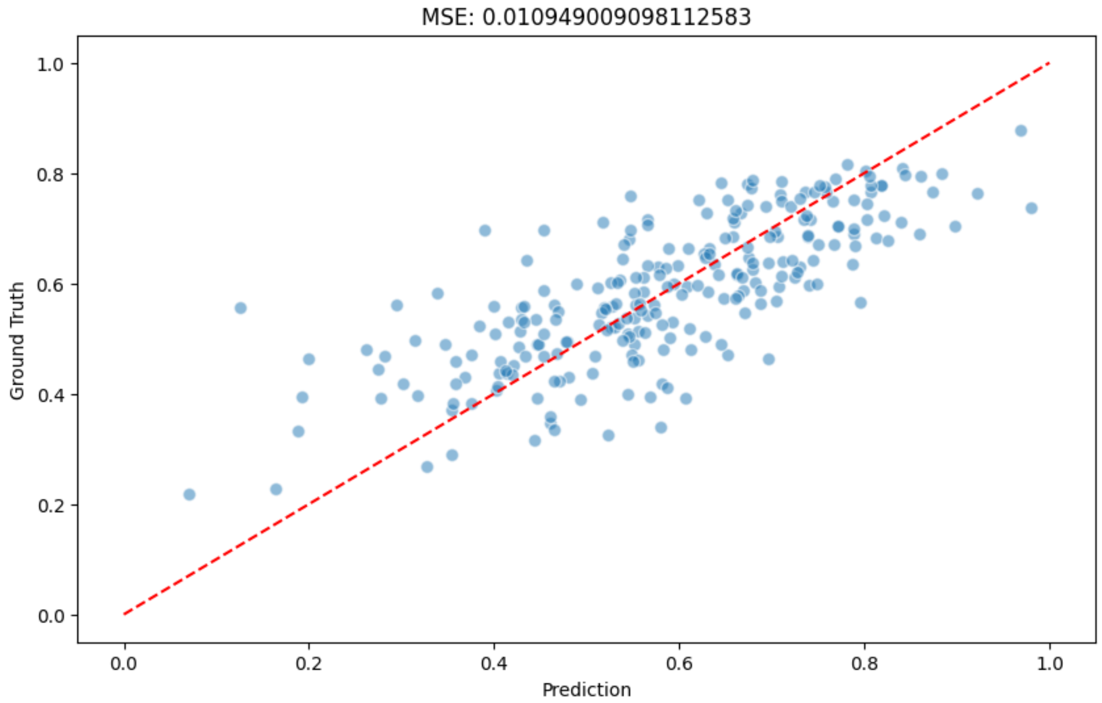
Model Performance – Pred vs True

Full Neural Network



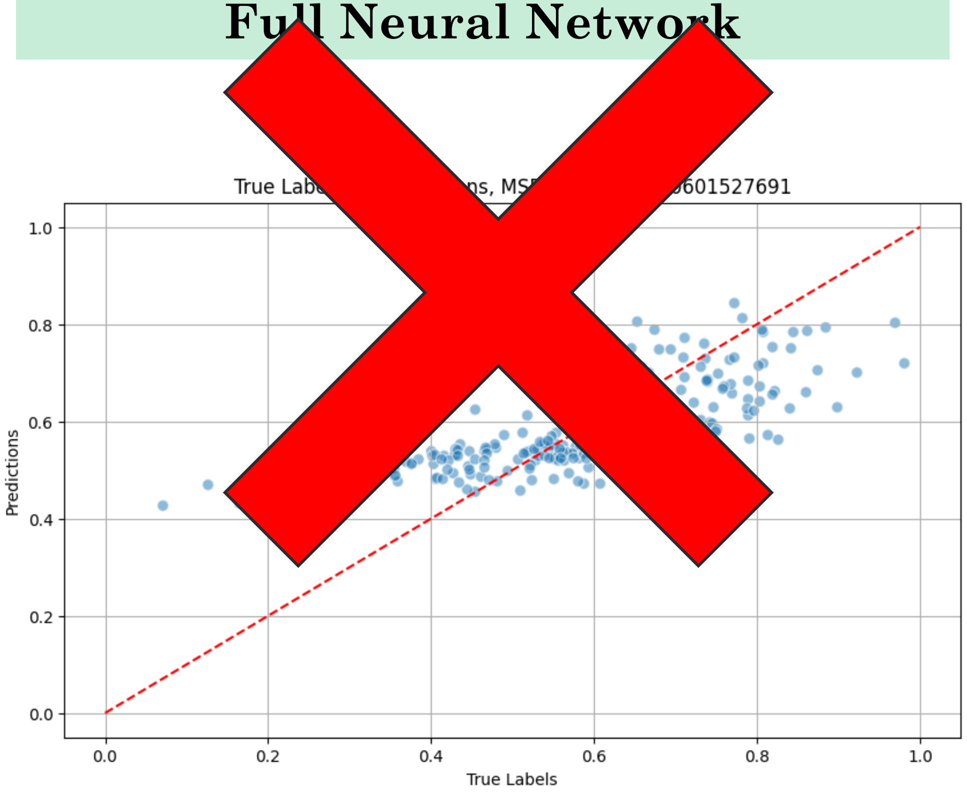
ViT (on nightlight) + XGBoost

ViT + XGBoost



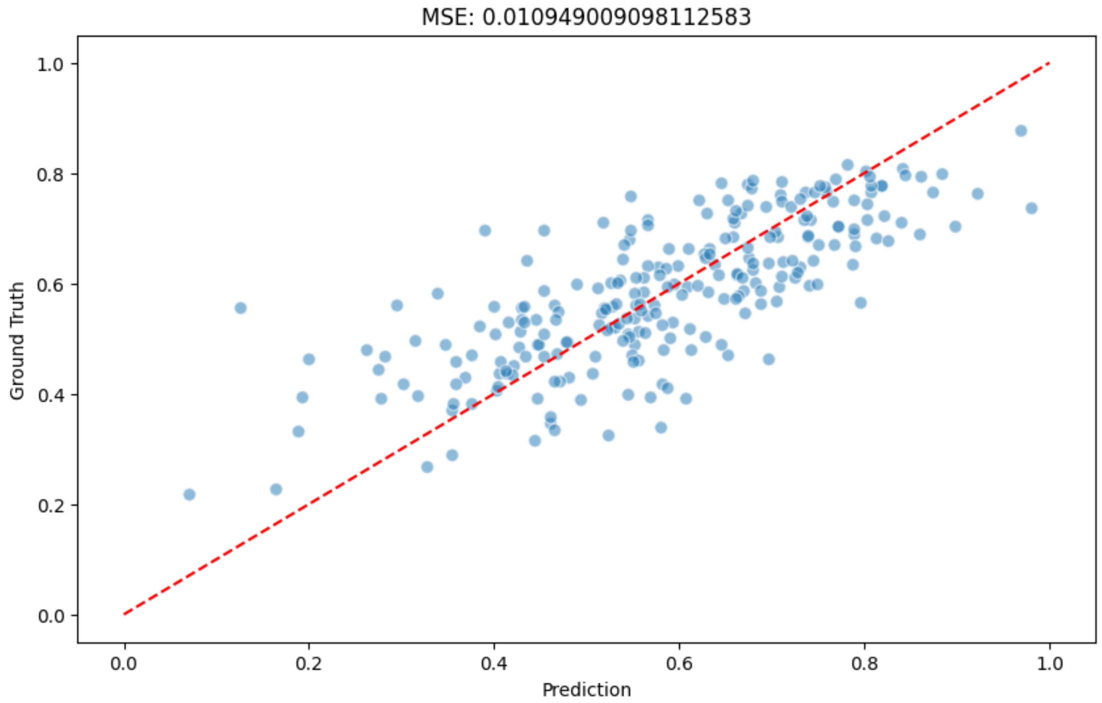
Model Performance – Pred vs True

Full Neural Network



ViT (on nightlight) + XGBoost

ViT + XGBoost

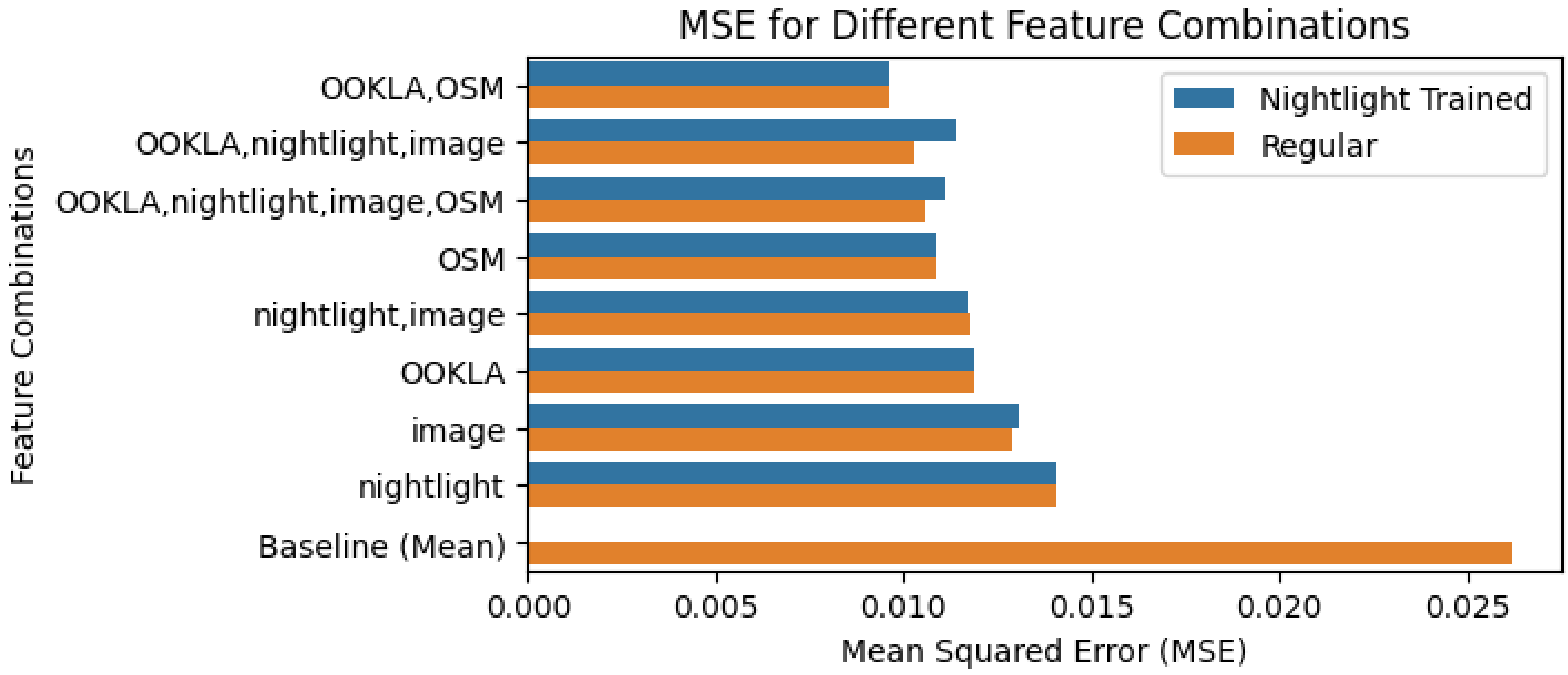


Model Performance

Feature/Model Selection

ViT + XGBoost

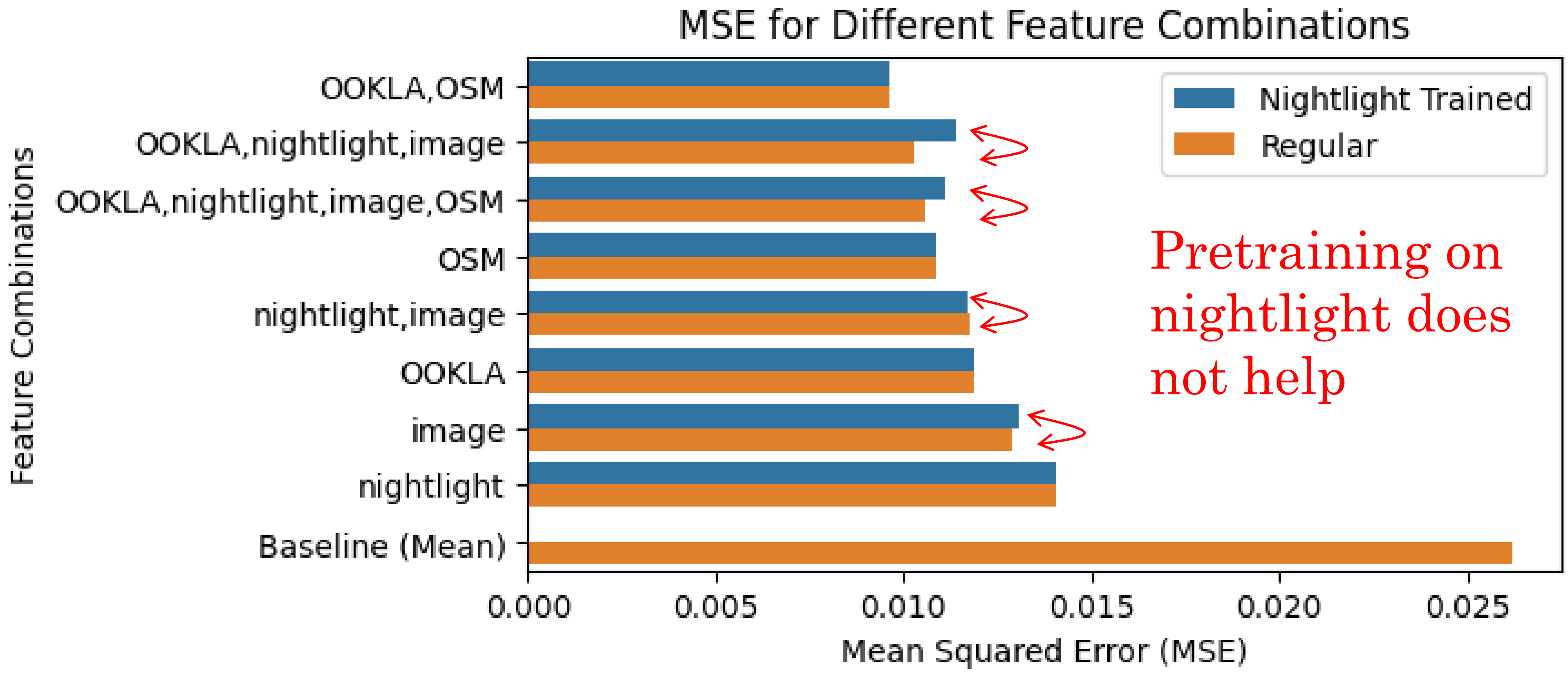
ViT (on nightlight) + XGBoost



Model Performance
Feature/Model Selection

ViT + XGBoost

ViT (on nightlight) + XGBoost



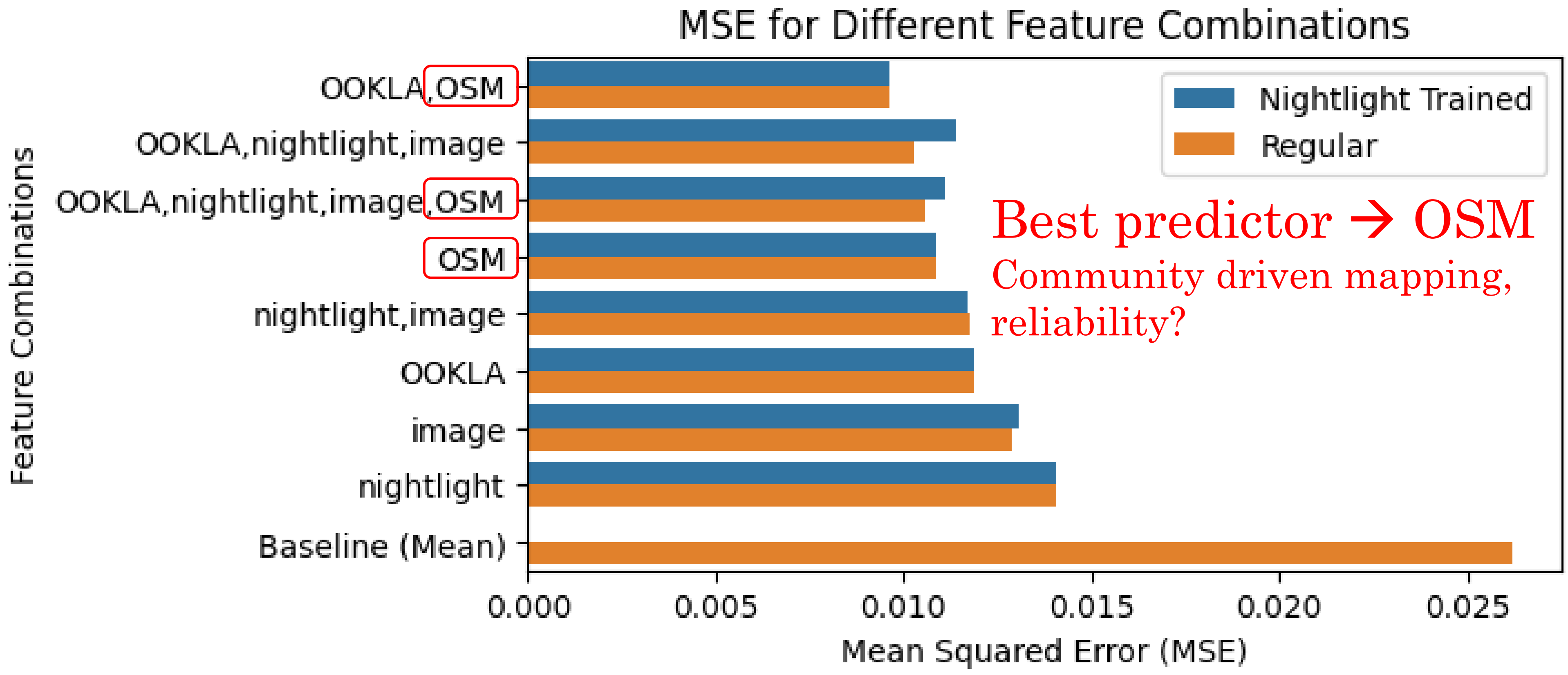
Pretraining on nightlight does not help

Model Performance

Feature/Model Selection

ViT + XGBoost

ViT (on nightlight) + XGBoost



Best predictor → OSM

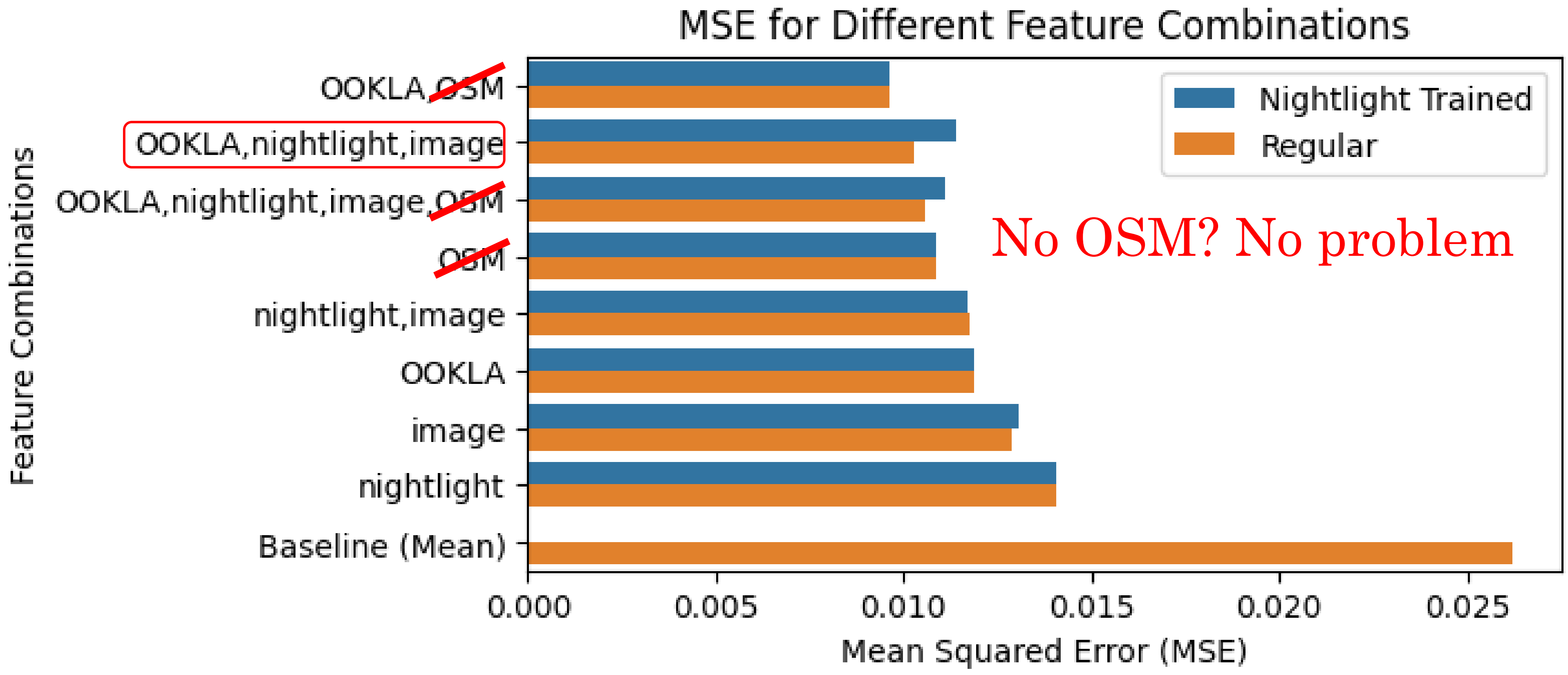
Community driven mapping, reliability?

Model Performance

Feature/Model Selection

ViT + XGBoost

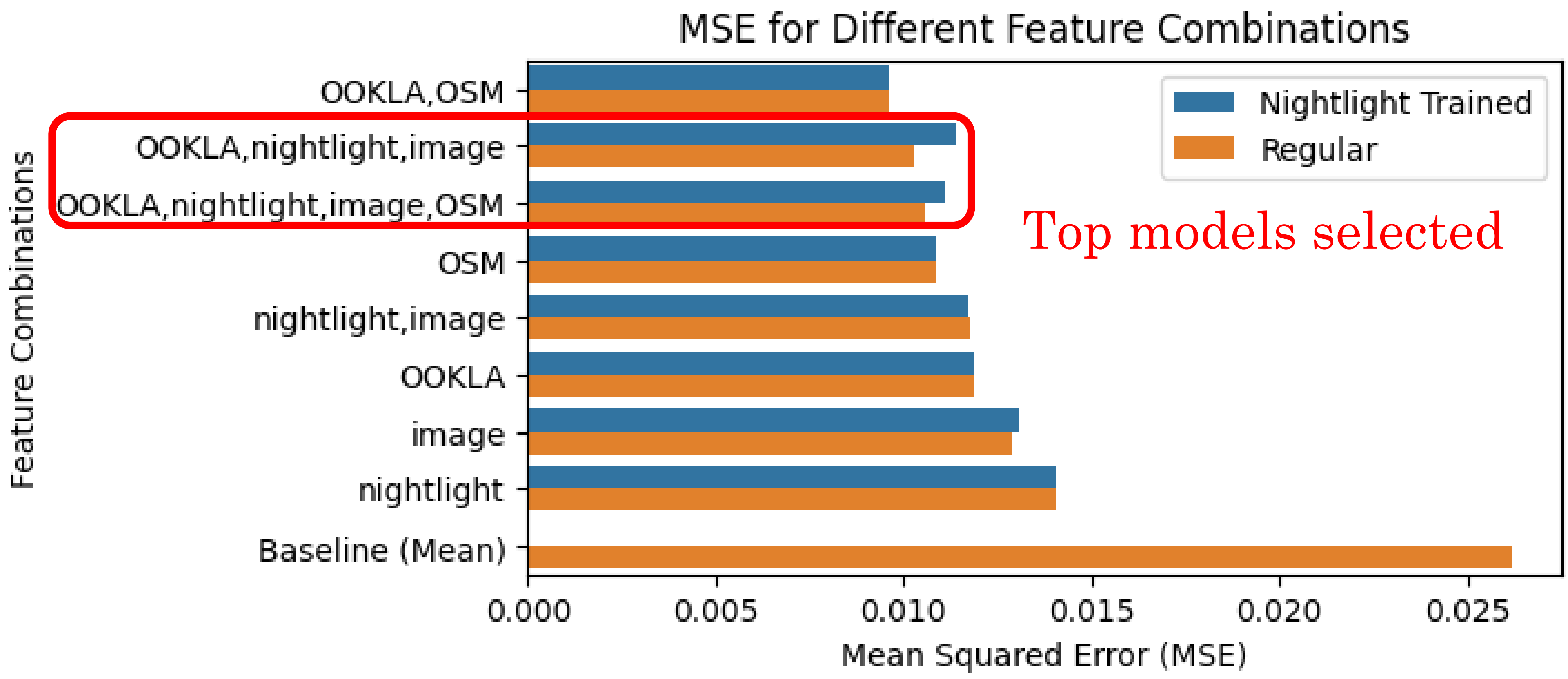
ViT (on nightlight) + XGBoost



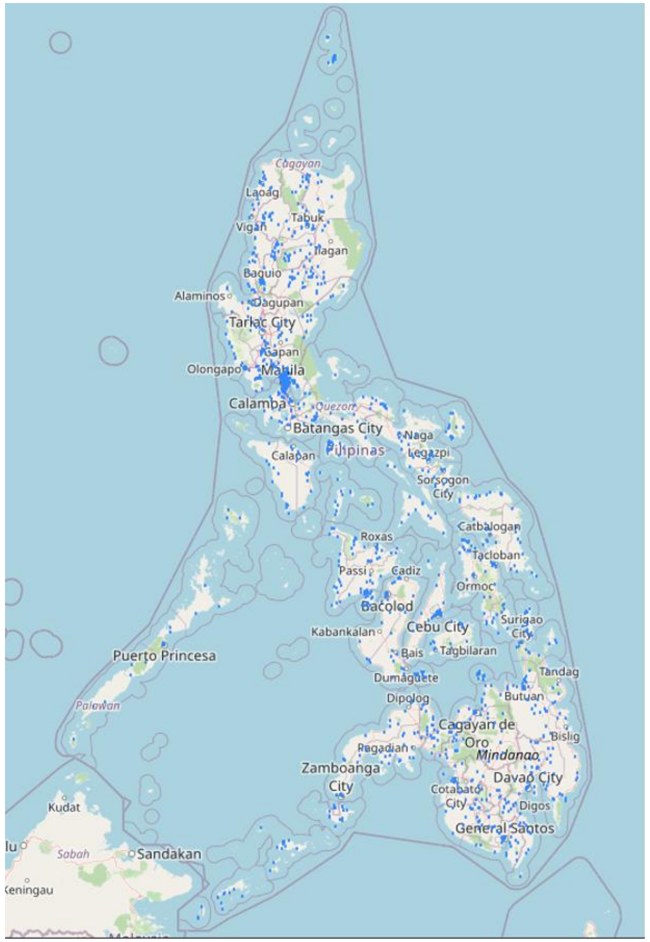
Model Performance
Feature/Model Selection

ViT + XGBoost

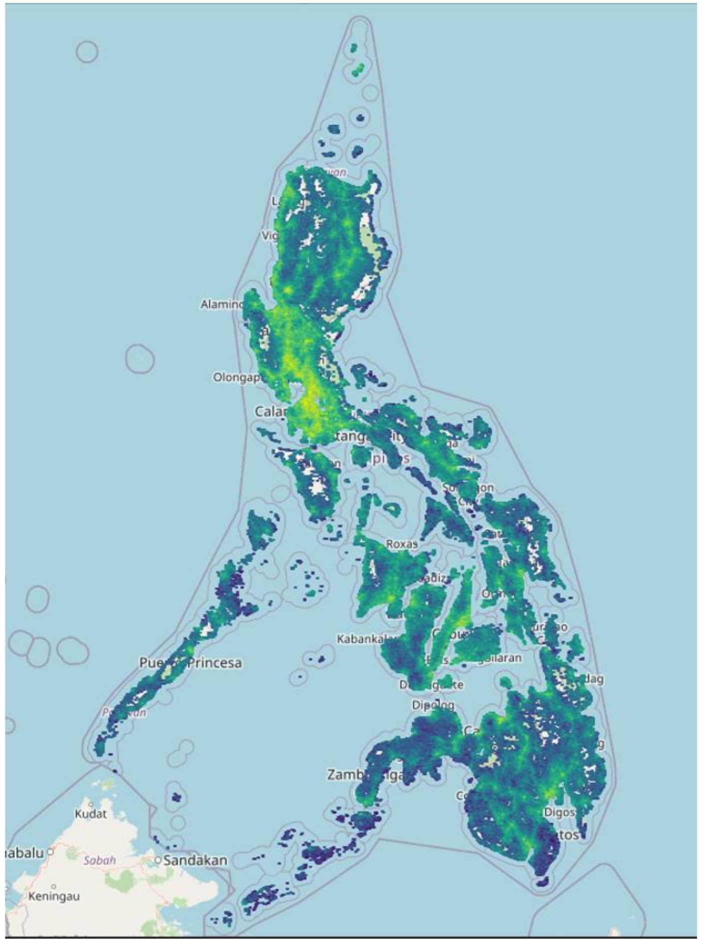
ViT (on nightlight) + XGBoost



Mapping coverage



DHS coverage



Satellite mapping coverage

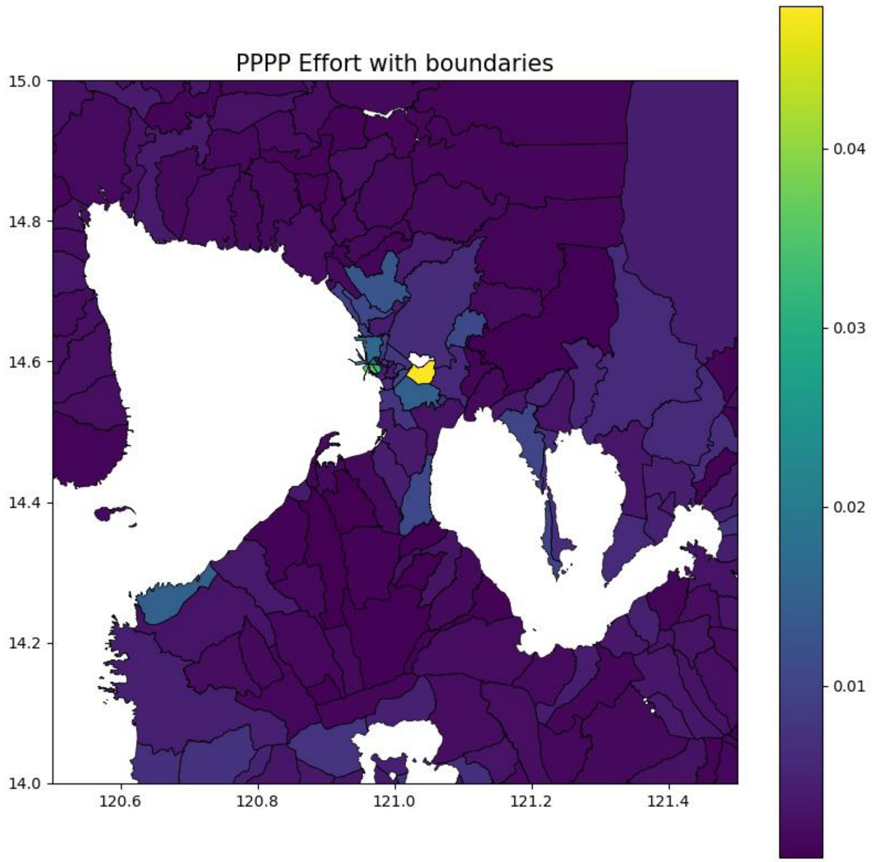
Impact Evaluation

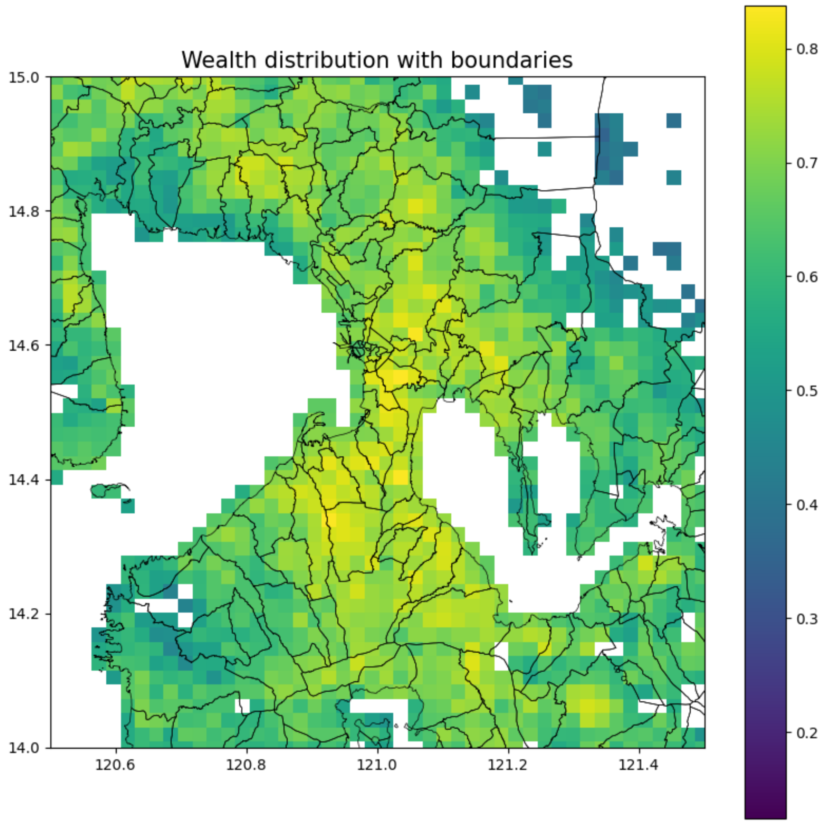
Effectiveness

Pantawid Pamilyang Pilipino Program

PPPP Effectiveness

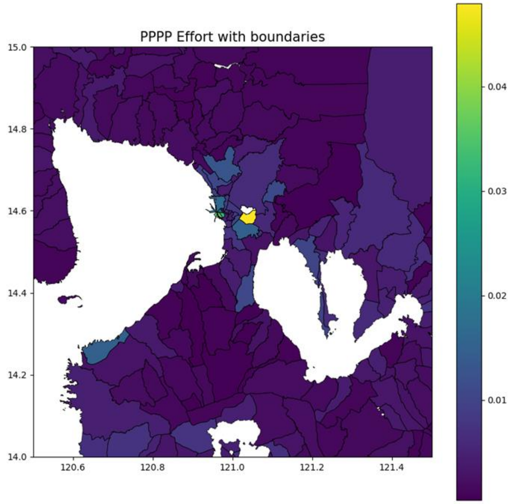
PPPP activity → Wealth index change between years (2020→2023)

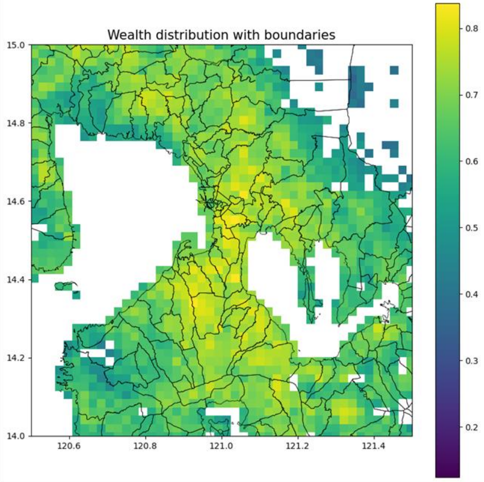




PPPP Effectiveness

PPPP activity → Wealth index change between years (2020→2023)





Conclusion → Increased PPPP activity slightly <u>reduces</u> wealth

(*p*-value=0.019)

- Simple example **for illustration only → only PPPP considered here**
- Effect of pandemic? Political uncertainty? Natural disasters?
- Increased PPPP spending in response to anticipated rising poverty levels?

Conclusion

Conclusion

- Machine learning techniques can be applied → poverty mapping
- Lower quality satellite imagery with computer vision
  - Decent predictive power
  - No advantage to pretraining with nightlight
- **Open source material → adequate for poverty mapping needs**