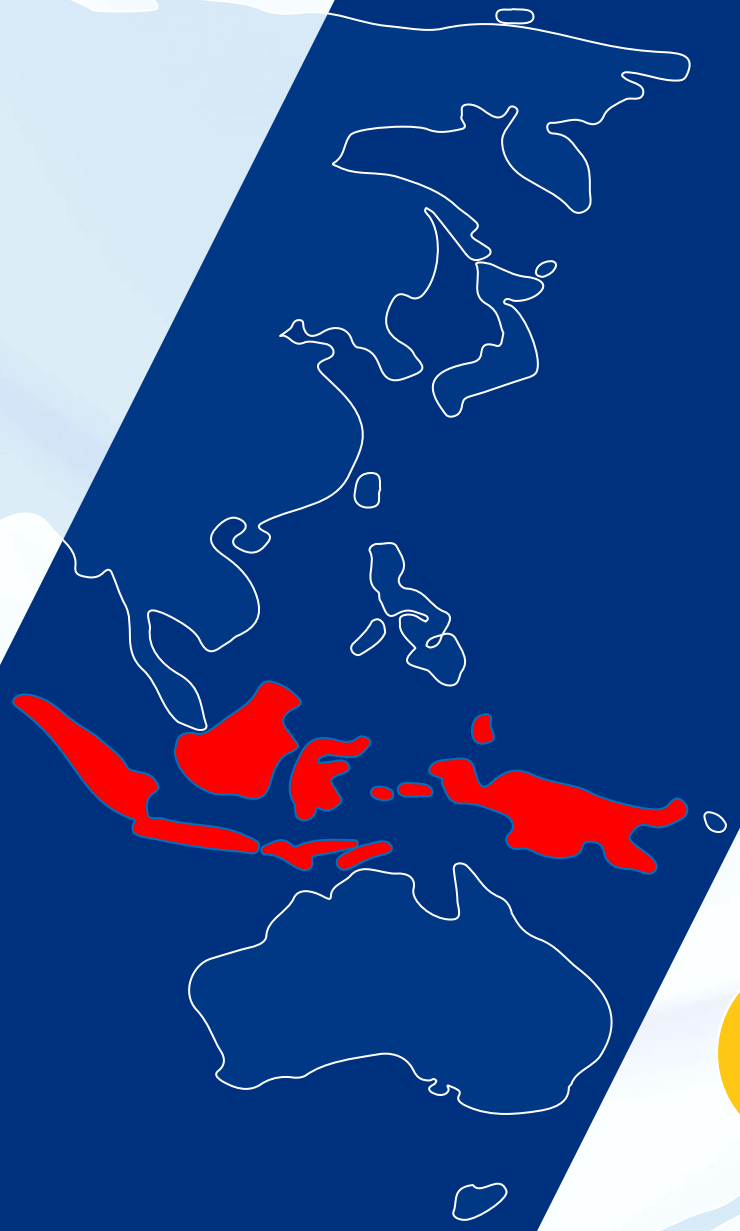
STATISTICS INDONESIA

FEASIBILITY STUDY OF SATELLITE IMAGERY ANALYSIS FOR WEALTH INDEX DEVELOPMENT IN INDONESIA

Arie Wahyu Wijayanto

UNECE Machine Learning Group
28 June 2021

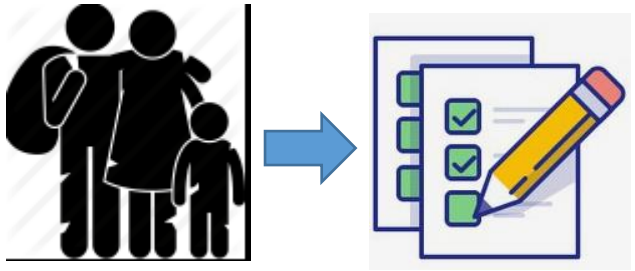


OUTLINE

- MOTIVATION
- METHODOLOGY
- PRELIMINARY RESULTS
- CONCLUDING REMARKS

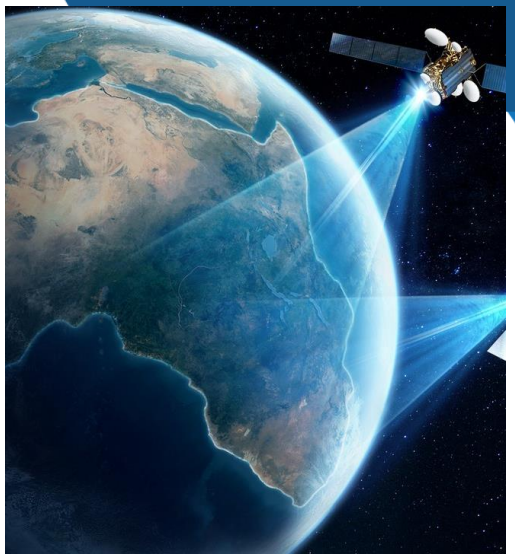
MOTIVATION

- Eliminating poverty is Indonesia's main target for sustainable development by 2030



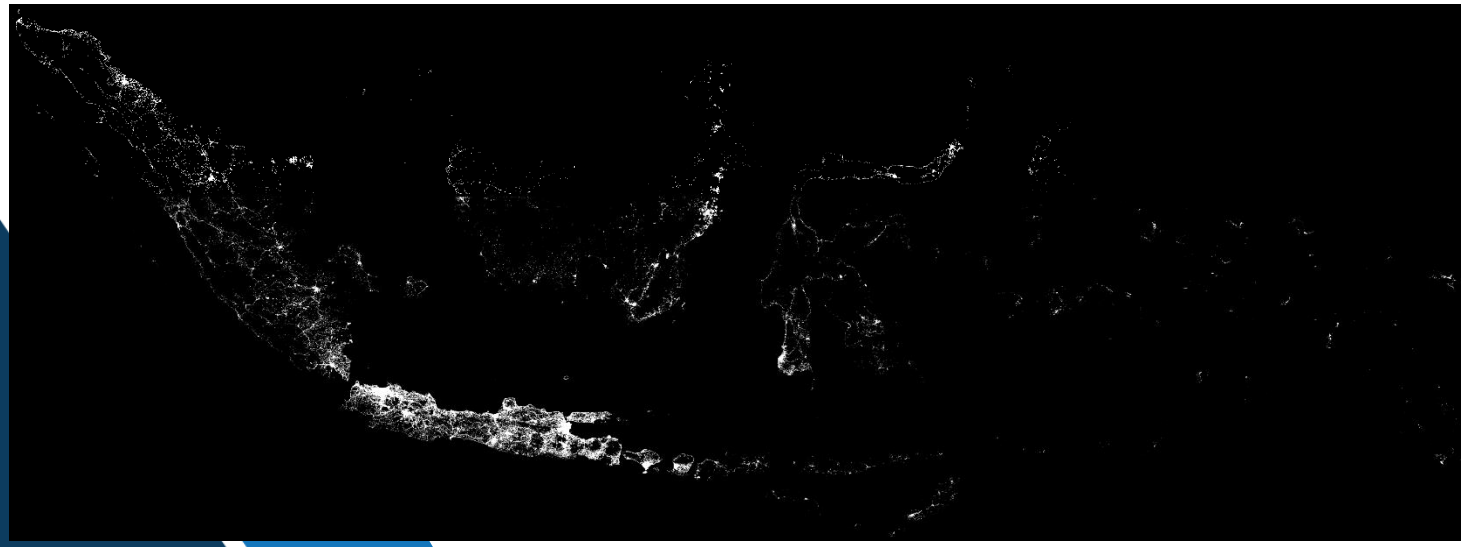
- Establishing a complete poverty database at **national scale** is costly
  - Currently available data: PSE 2005, PPLS 2008, PPLS 2011, PBDT 2015
- Poverty data estimation through biannual Households Socio-Economic Surveys (SUSENAS) are only available up to the **district level**

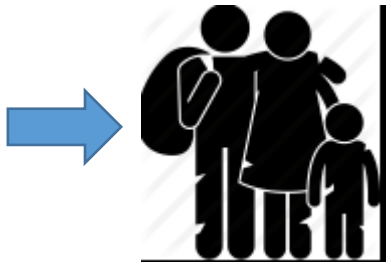
MOTIVATION



- Estimation of regional poverty using satellite imagery is a new alternative to support poverty alleviation (Chen & Nordhaus, 2011; Henderson et al., 2012; Ivan et al., 2020).

- We aim to evaluate the feasibility of estimating the **poverty spatial distribution** and **wealth index development** using satellite imagery and geospatial data to enhance the **cost effectiveness**, **granularity**, and **accuracy** of poverty statistics.





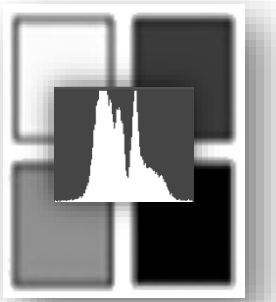
Estimating Poverty Spatial Distribution

DATA

Data

Day time satellite images

Landsat 8
Sentinel 2



Night time light intensities

NPP-VIIRS

National Polar-orbiting Partnership–Visible Infrared Imaging Radiometer Suite

Poverty Database

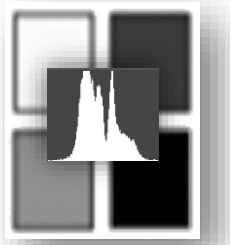
PBDT 2015

Indonesia Integrated Poverty Database

METHODOLOGY

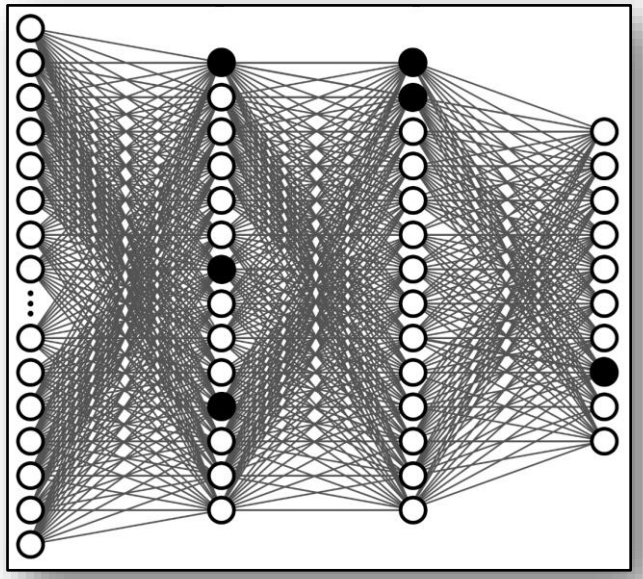
Input image

Day time satellite images



Night time light intensities

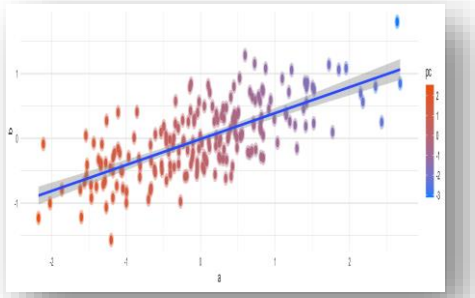
Extract features using trained machine learning algorithm



Convolutional Neural Networks (ResNet34)

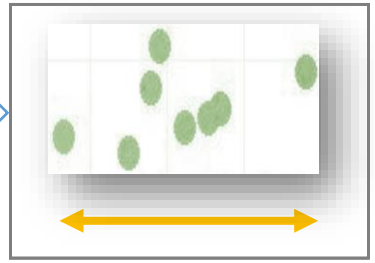
Extracted features

Trained regression model



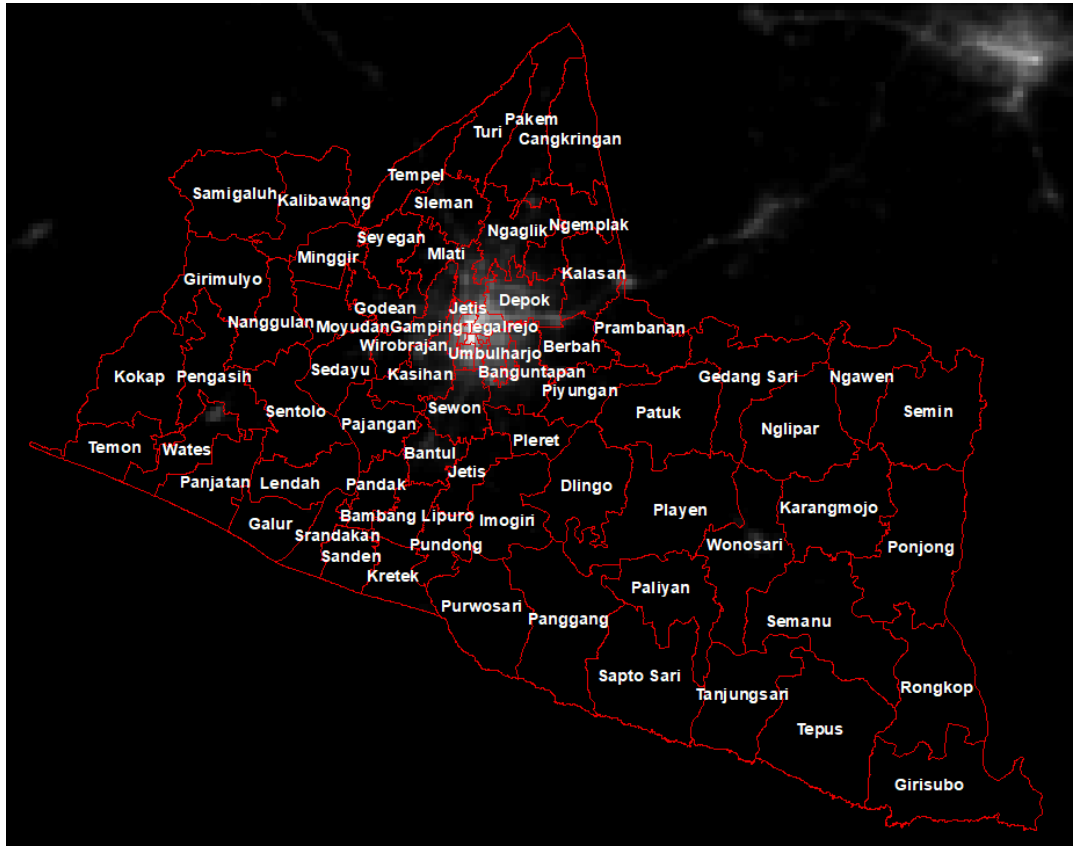
Ridge Regression
Support Vector Regression

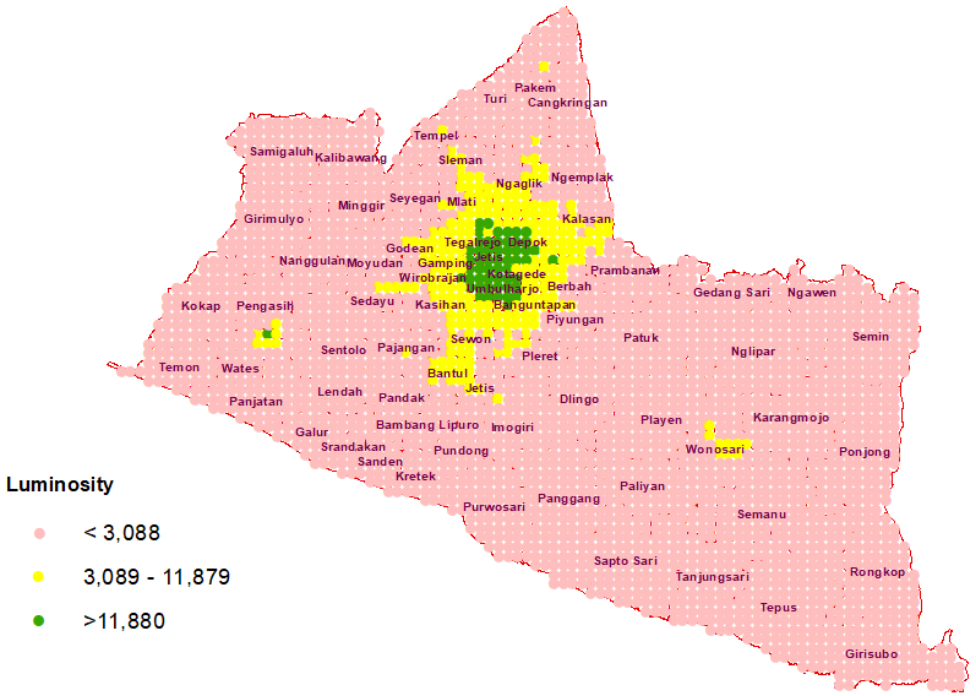
Poverty statistics indicators



PRELIMINARY RESULTS: YOGYAKARTA PROVINCE

Night-Time Lights Luminosity

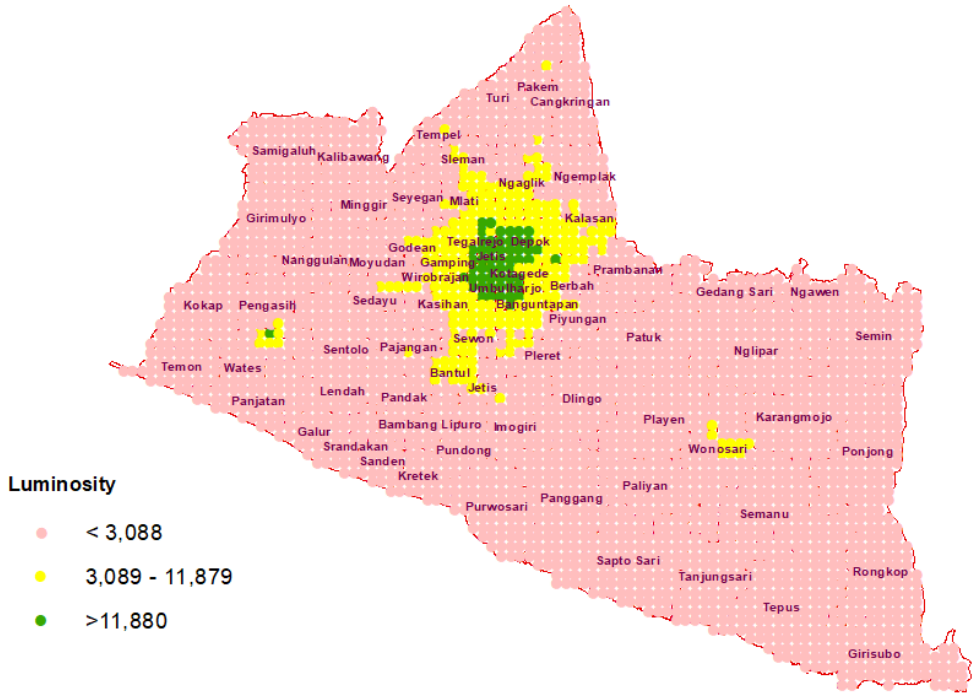




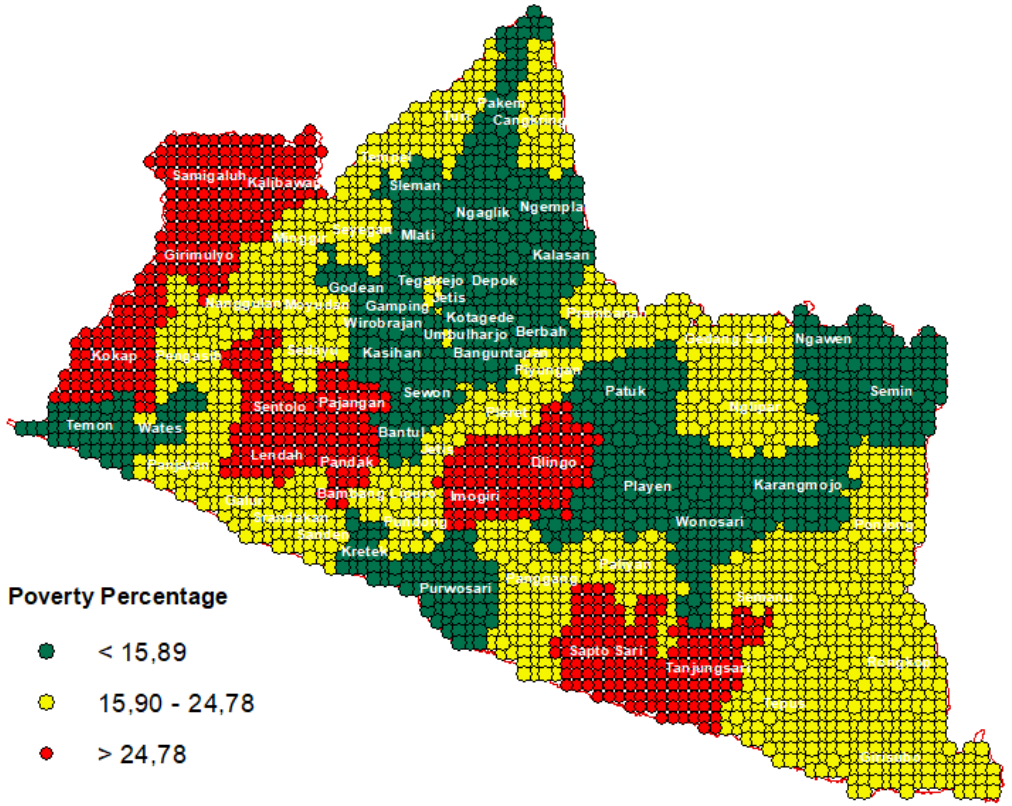
The center of Yogyakarta City has a greater luminosity intensity than rural areas and areas outside the city.

PRELIMINARY RESULTS: YOGYAKARTA PROVINCE

Night-Time Lights Luminosity



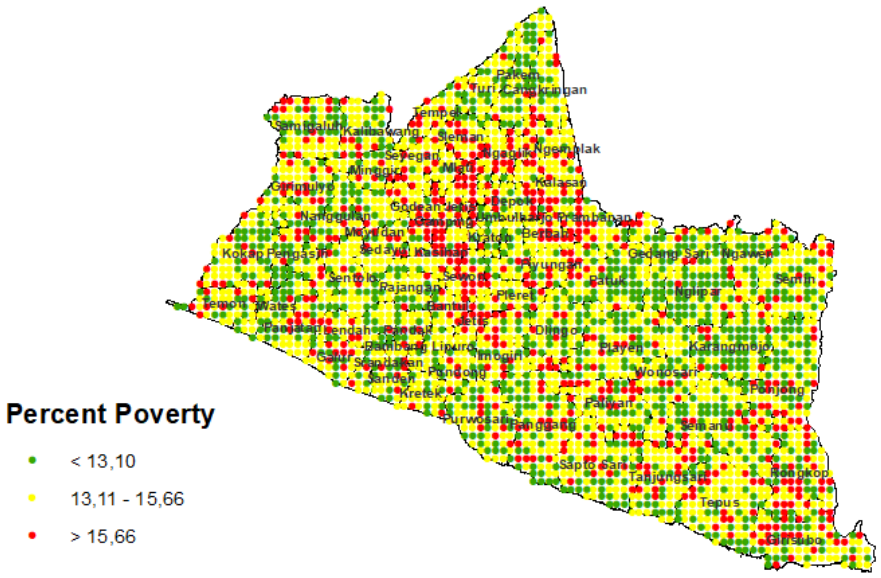
Official Poverty Distribution (PBDT 2015)



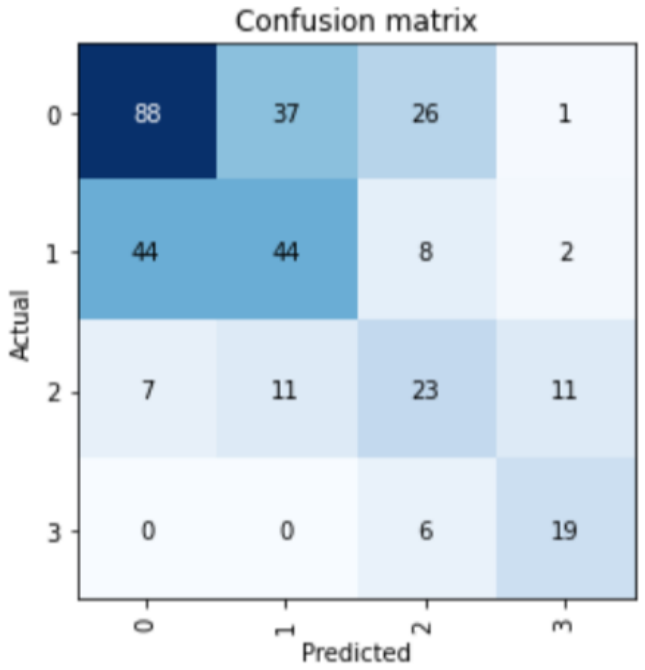
The city center of Yogyakarta province has a greater intensity of night-time light and a lower poverty rate than other areas.

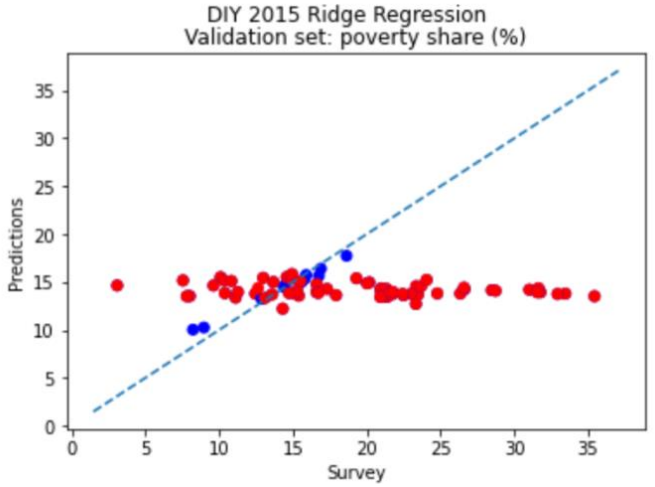
RESULTS: PREDICTED POVERTY DISTRIBUTION (DAY TIME - NIGHT TIME LIGHT DERIVED ESTIMATION)

Poverty Percentage by prediction model with RES34



CNN Model Testing and Evaluation



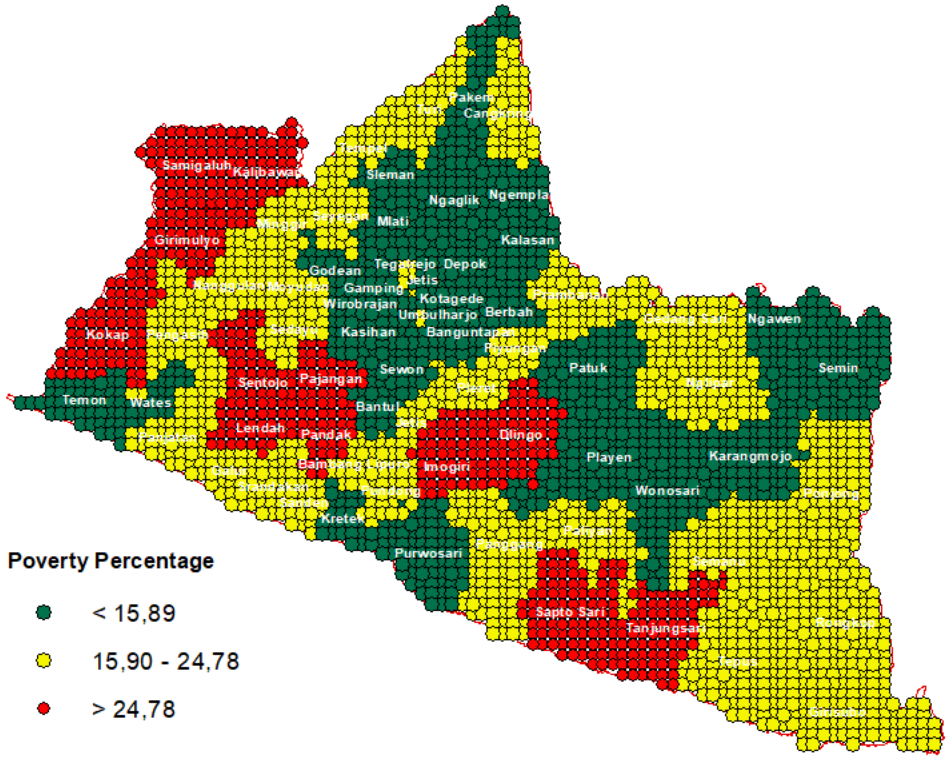


RMSE_valid	0.0896
RMSE_full	0.0861
R2_valid	-0.5537
R2_full	-0.4796
R2_train	0.9247

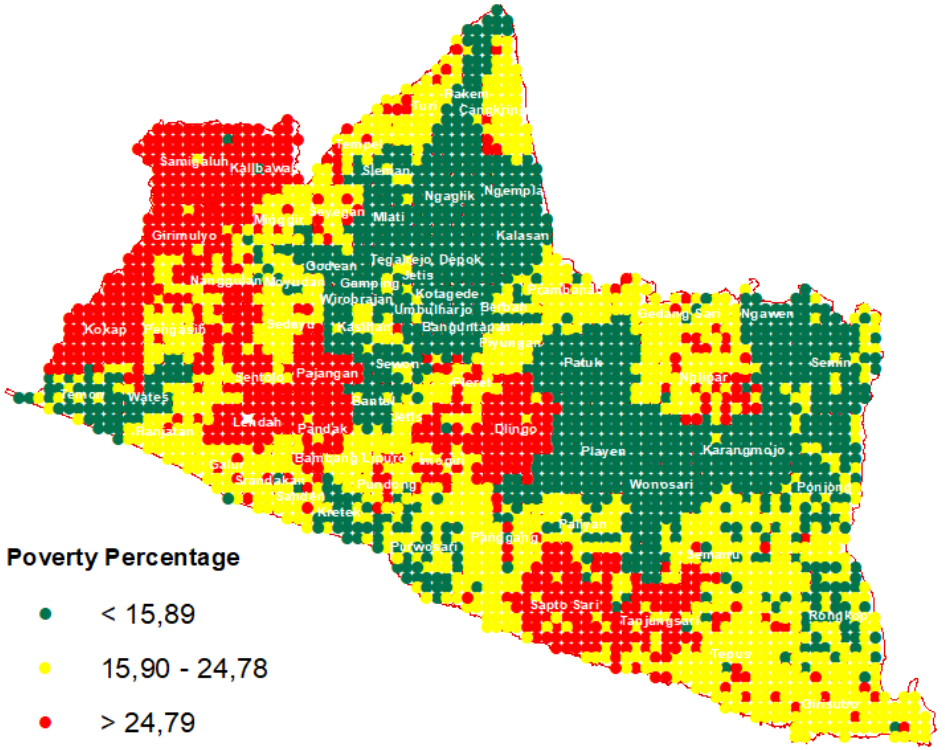
The resulted model predictions when compared with the Official Poverty Distribution (PBDT 2015)

RESULTS: PREDICTED POVERTY DISTRIBUTION (DAY TIME - NIGHT TIME LIGHT DERIVED ESTIMATION)

Distribution poverty percentage by PBDT 2015

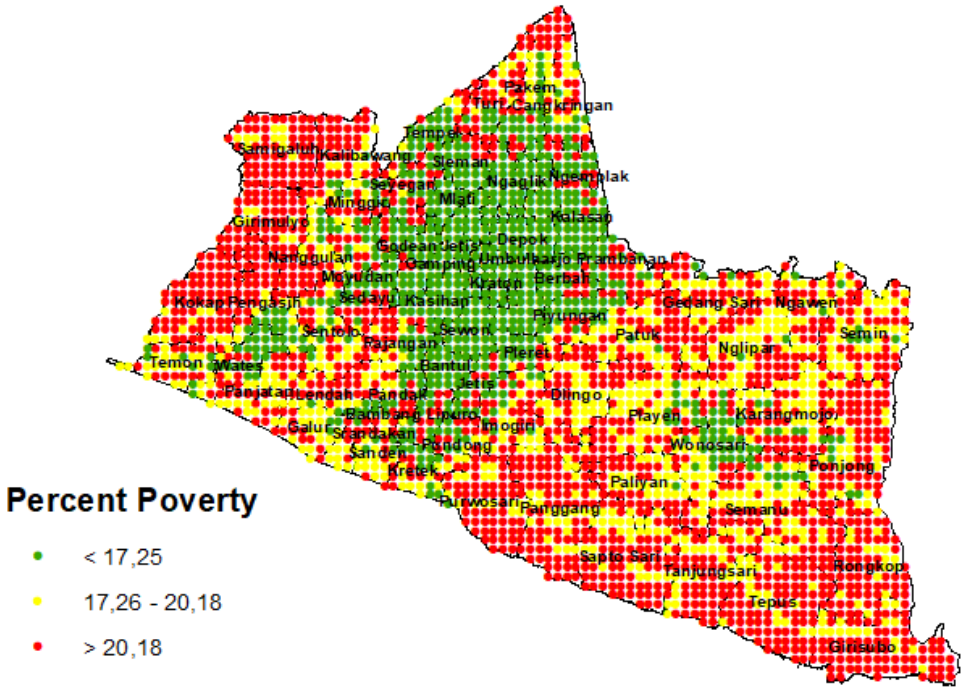


Poverty Percentage by prediction model with RES34 after it is rescaled by population grid

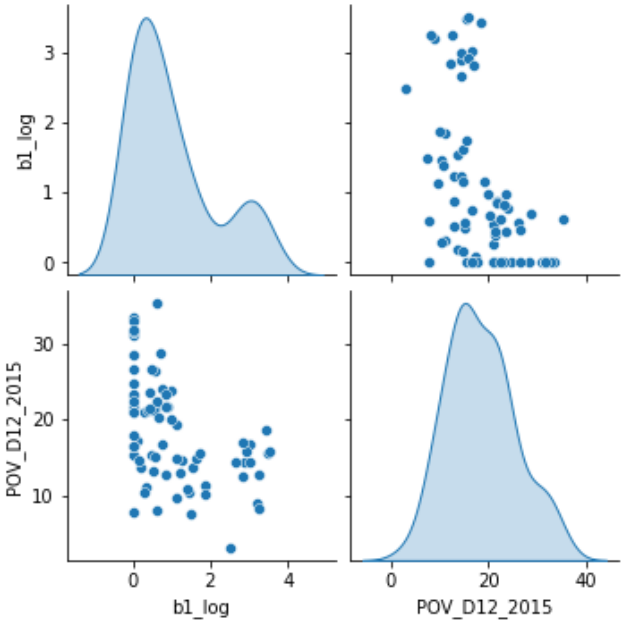


The results of the model predictions after rescaling are quite good in estimating regional poverty with an RMSE value of 8 percent

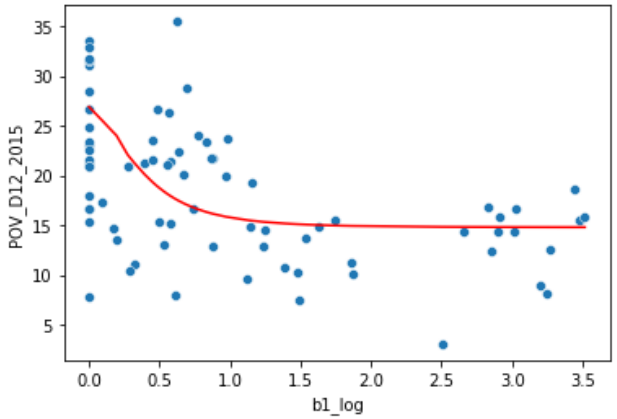
RESULTS: PREDICTED POVERTY DISTRIBUTION (NIGHT TIME LIGHT DERIVED ESTIMATION)



Using log transformation for transformation luminosity intensity to normal value distribution



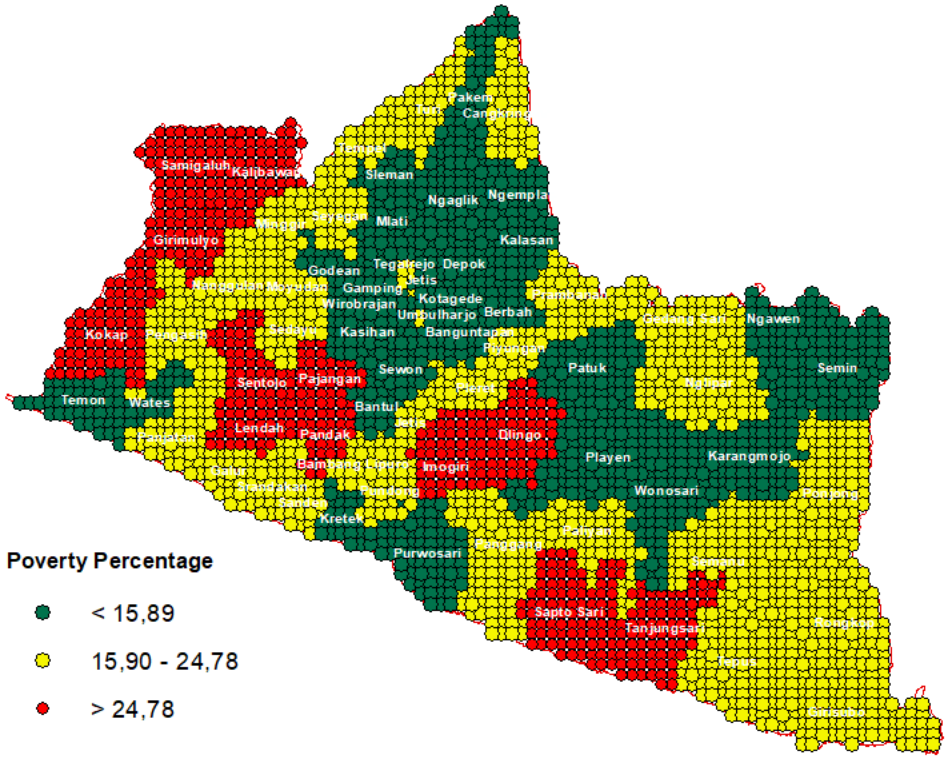
Using Support Vector Regression



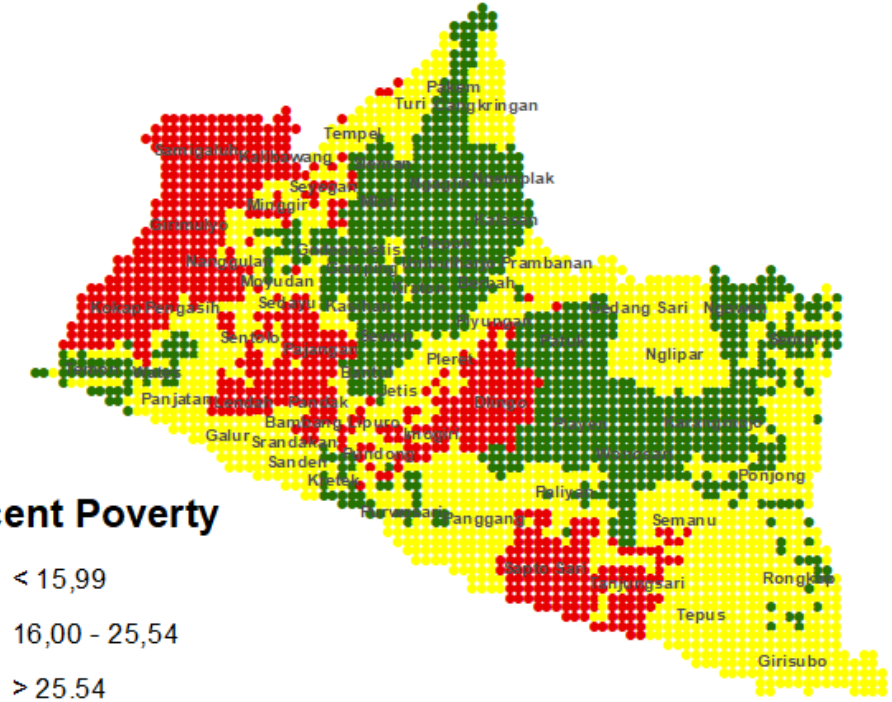
The results of the model predictions with only NTL Imagery are better than use day light imagery.

RESULTS: PREDICTED POVERTY DISTRIBUTION (NIGHT TIME LIGHT DERIVED ESTIMATION)

Official Poverty Distribution (PBDT 2015)



Poverty Percentage by prediction model with SVR after it is rescaled by population grid



Using NTL Imagery, we can predict regional poverty distribution more properly.
After rescaling using the population grid, we better spatial distribution poverty compared with the Official Poverty Database (PBDT 2015).

CONCLUDING REMARKS

KEY FINDINGS:

- The model has been able to predict poverty relatively well.
- Using only night time light derived data for prediction is better than its combination with raw day time imagery.
- Ground checks need to be done to ensure that the night light data represents the economic activities in some area.

ON GOING WORKS:

- Investigating the use of additional geospatial features in accordance with relevant studies:
  - Urban Heat Island (Buyantuyev, 2009 and Dissanayake, 2018)
  - Air Pollution from CO Emissions (Tariq, 2017)
  - Built-up Area Distribution (Faisal, 2016)
- Incorporating small area estimation

STATISTICS INDONESIA

THANK YOU