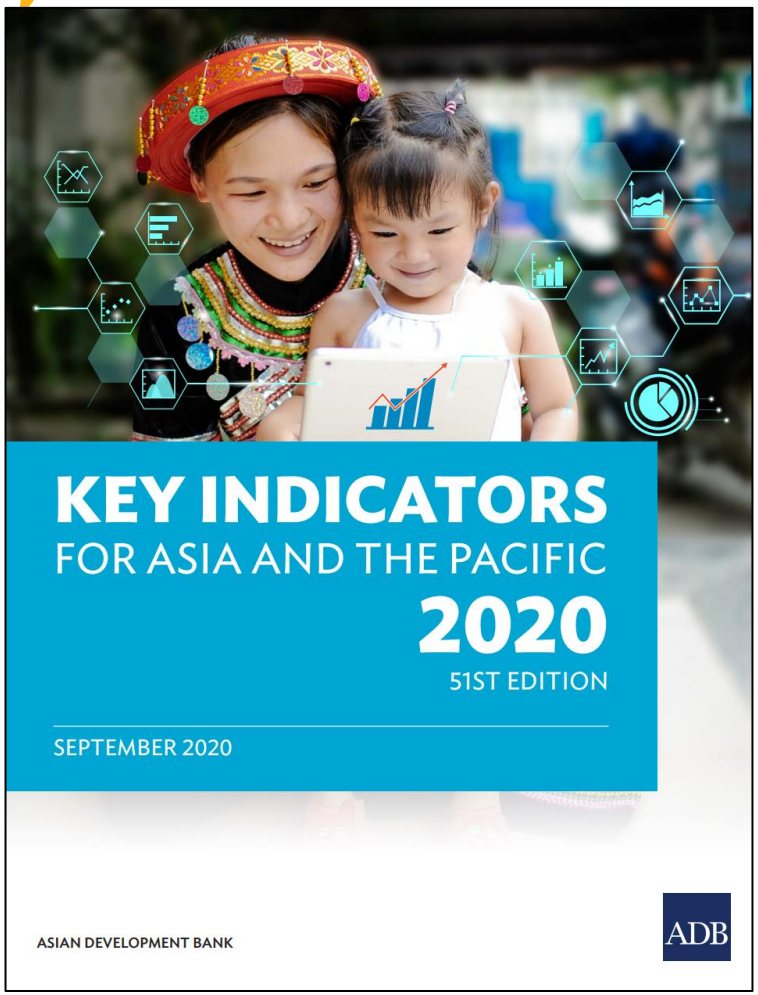


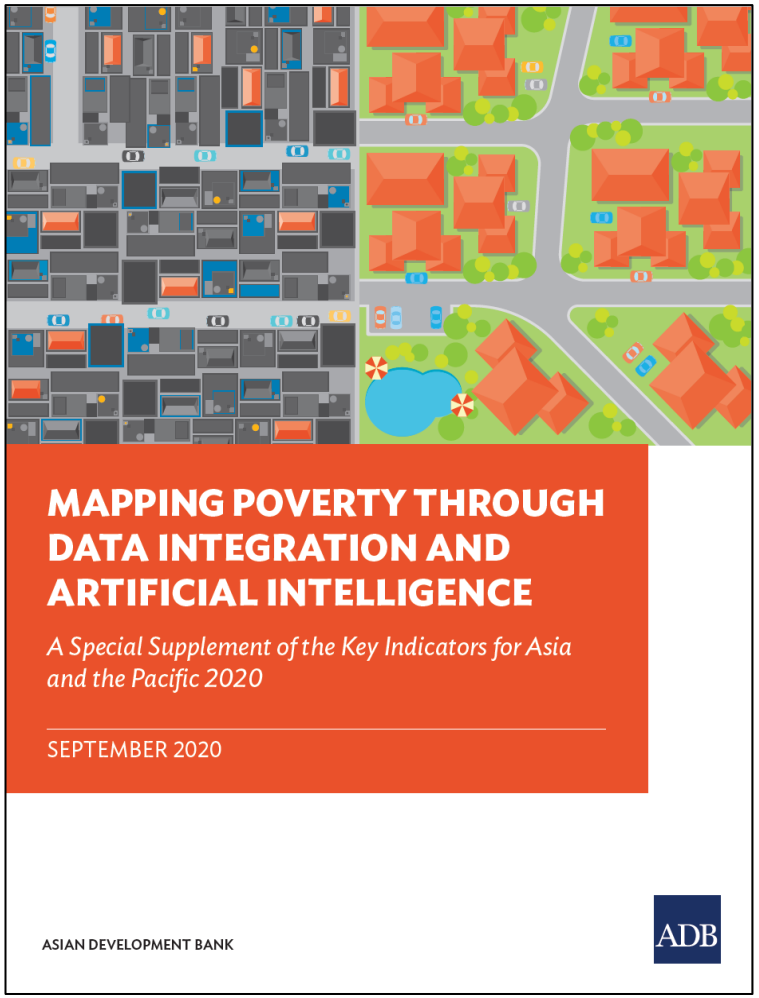
Mapping Poverty through Data Integration and Artificial Intelligence

Arturo M. Martinez Jr.
Statistics and Data Innovation Unit
Economic Research and Regional Cooperation Department
Asian Development Bank

ADB



https://www.adb.org/publications/key-indicators-asia-and-pacific-2020



https://www.adb.org/publications/mapping-poverty-data-integration-ai

ADB

Applying Artificial Intelligence to Poverty Prediction

The 2030 Sustainable Development Agenda espouses the ‘leave no one behind’ principle to ensure that improved living standards can be enjoyed by everyone.

The LNOB principle calls for availability of data disaggregated by **location**, age, gender, ethnicity, income status, disability status, etc., to speak on behalf of people who may be left behind by inequitable socioeconomic progress.

From the perspective of data producers, particularly national statistics offices which are considered as major compilers of SDG data, data disaggregation entails redesigning existing data collection systems to collect more information.

However, **conventional data sources are already costly** to begin with (e.g., hhld income and expenditure survey is roughly US$1.7 million to conduct)

Applying Artificial Intelligence to Poverty Prediction

How can we reconcile the need for granular data and need to contain costs at manageable levels?

Could we capitalize on availability of alternative data sources to complement conventional data sources to meet the disaggregated data requirements of the SDGs?

ADB's **Data for Development** aims to strengthen the capacity of NSOs to meet the disaggregated data requirements of the 2030 Sustainable Development Agenda by capitalizing on both conventional and **innovative data** sources.

Applying Artificial Intelligence to Poverty Prediction



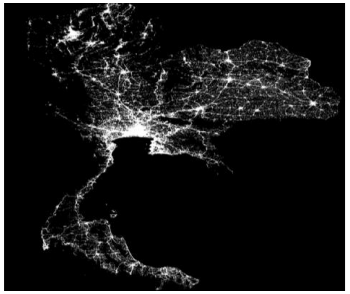




Image Source: Solar Quotation



Image Source: Earth Imaging Journal

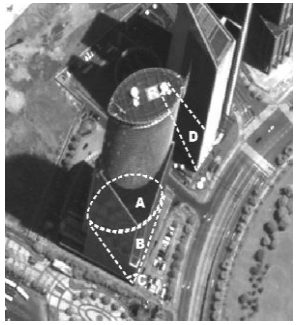


Image Source: Pan et al., 2008



Image Source: AgriLand



Image Source: Wang et al., 2016

ADB

Applying Artificial Intelligence to Poverty Prediction

CNN requires volumes of poverty-labelled images as input data for training which we do not have.



Proportion of area X's population living below poverty line: **25%**

Poverty statistics are typically derived from hhld surveys which are designed to provide reliable estimates at national, regional, or provincial-levels only.

These do not provide sufficient data for CNN training

Applying Artificial Intelligence to Poverty Prediction

Researchers from Stanford University proposed a transfer learning technique wherein instead of training the computer vision algorithm to directly predict poverty, the algorithm predicts intensity of night light.



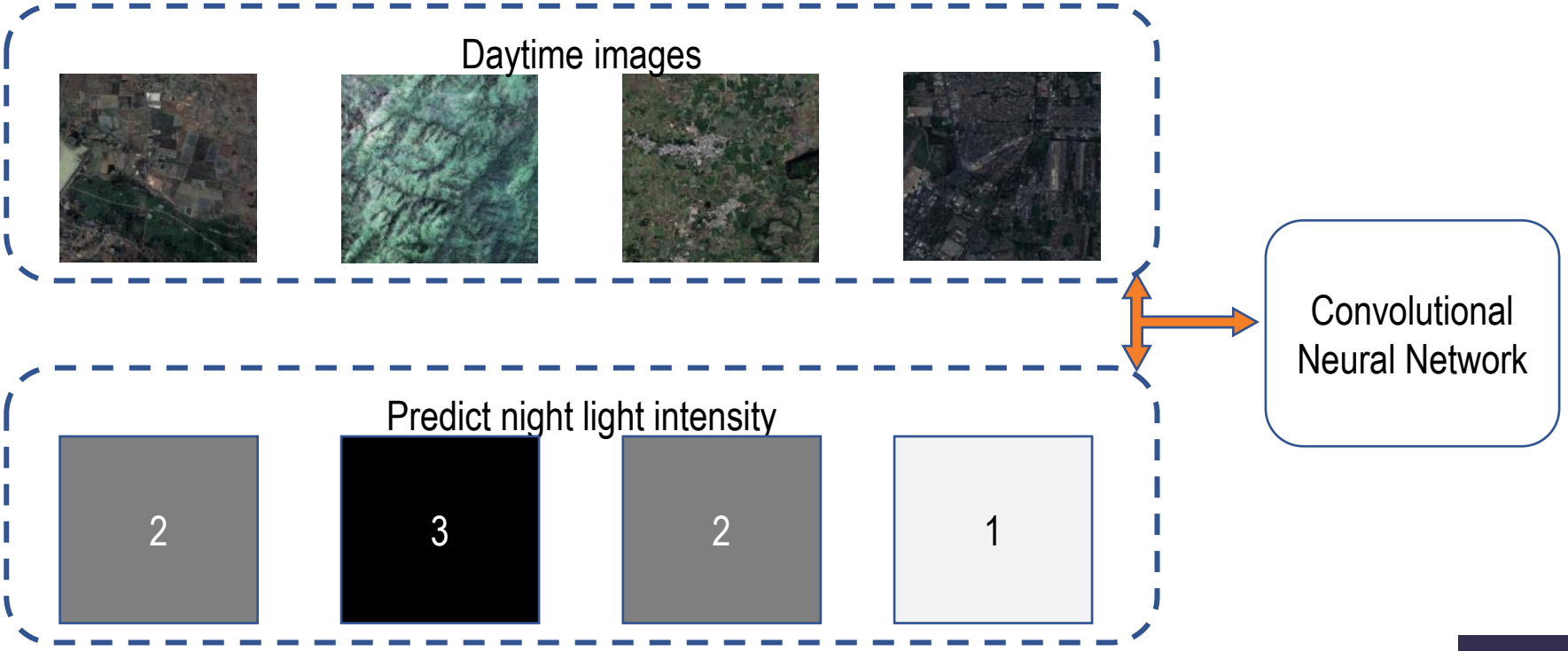
Proportion of area X's population living below poverty line:
25%



Intensity of night lights is correlated with economic activity -> poverty.

Data on night lights are more granular -> provides more data for training a computer vision algorith,

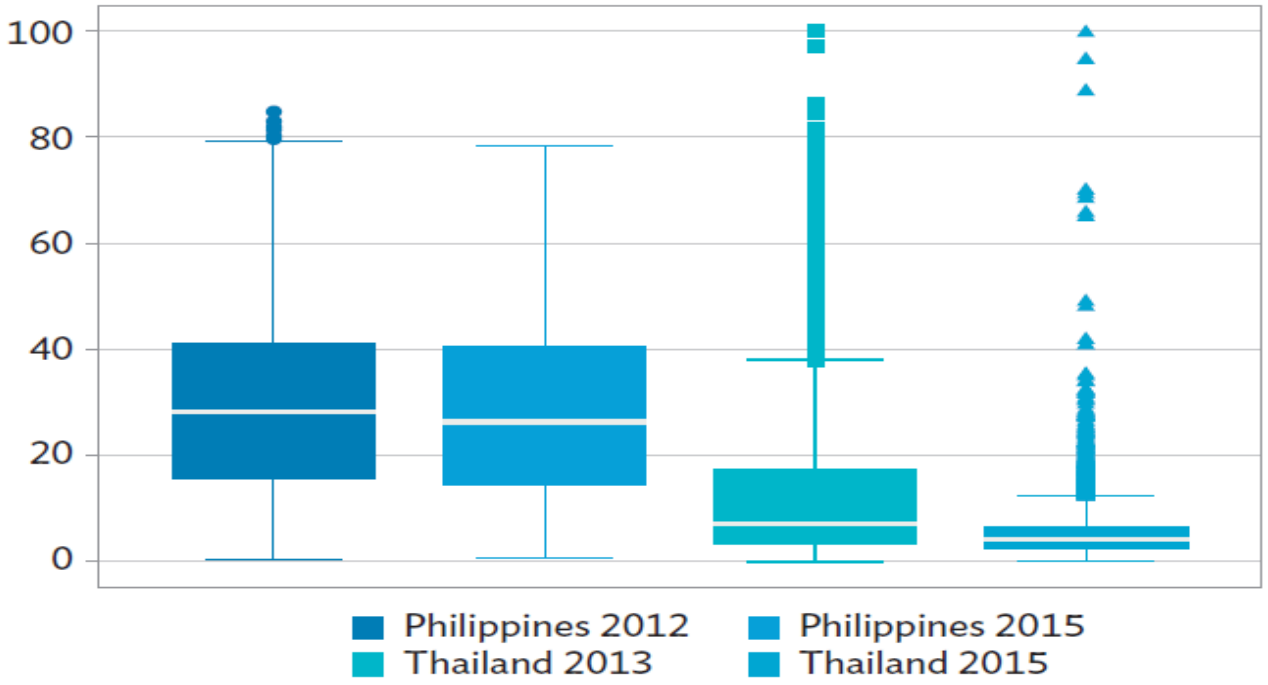
Methods: Using AI for Poverty Mapping



Applying on PHI and THA data

We use the estimated proportion of population living below the national poverty line as compiled by the Philippine Statistics Authority and National Statistical Office of Thailand through SAE techniques: approx. 1600 *Municipal / city-level in the Philippines* ; 7000 *Tambon-level in Thailand*

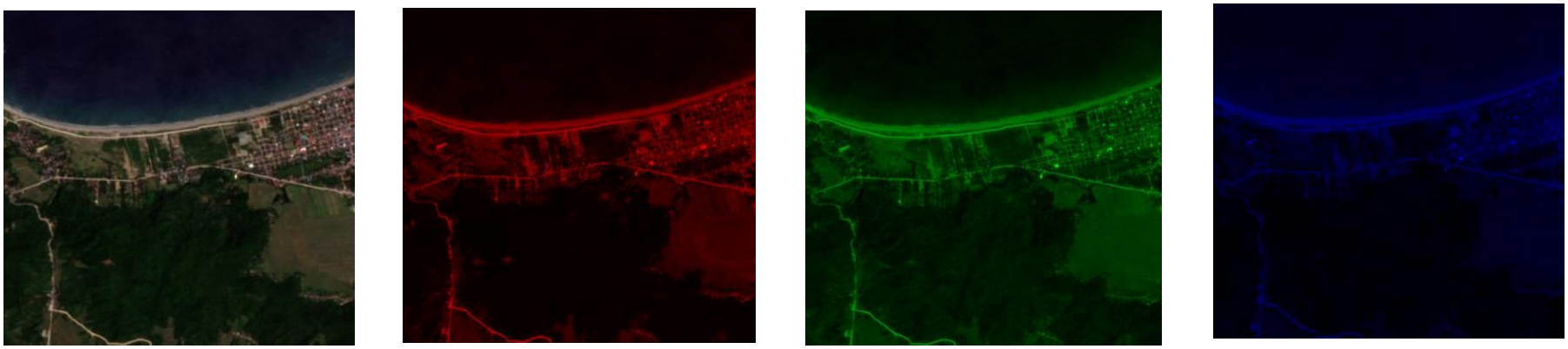
Figure: Distribution of Poverty Rates in the Philippines and Thailand



Applying on PHI and THA data

The input data were obtained using georeferenced and tagged image files. These image files are stored in three-dimensional arrays, with each pixel represented in red, green, and blue color bands.

Figure. Image Color Bands within a Georeferenced Image File

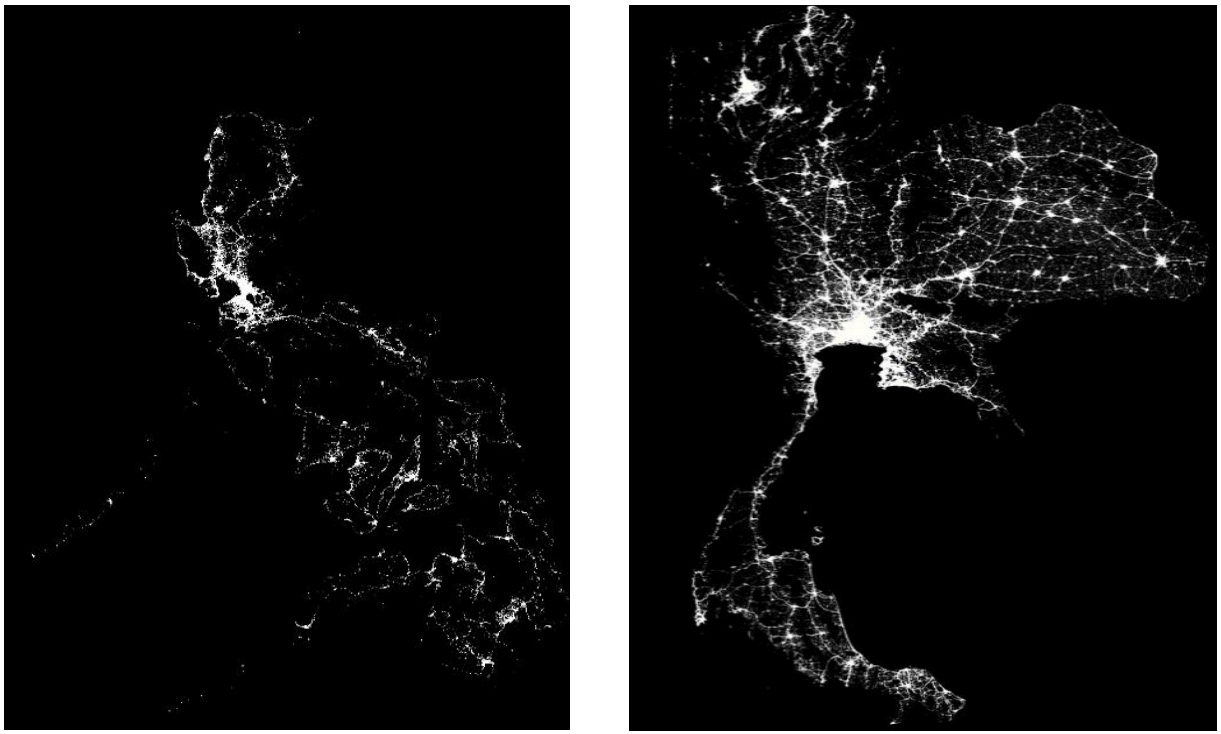


Source: Sentinel Images

Applying on PHI and THA data

We use data on night lights compiled by the Visible Infrared Imaging Radiometer Suite (VIIRS). The intensity levels were categorized into discrete groups using combination of Gaussian Mixed Models and heuristic methods.

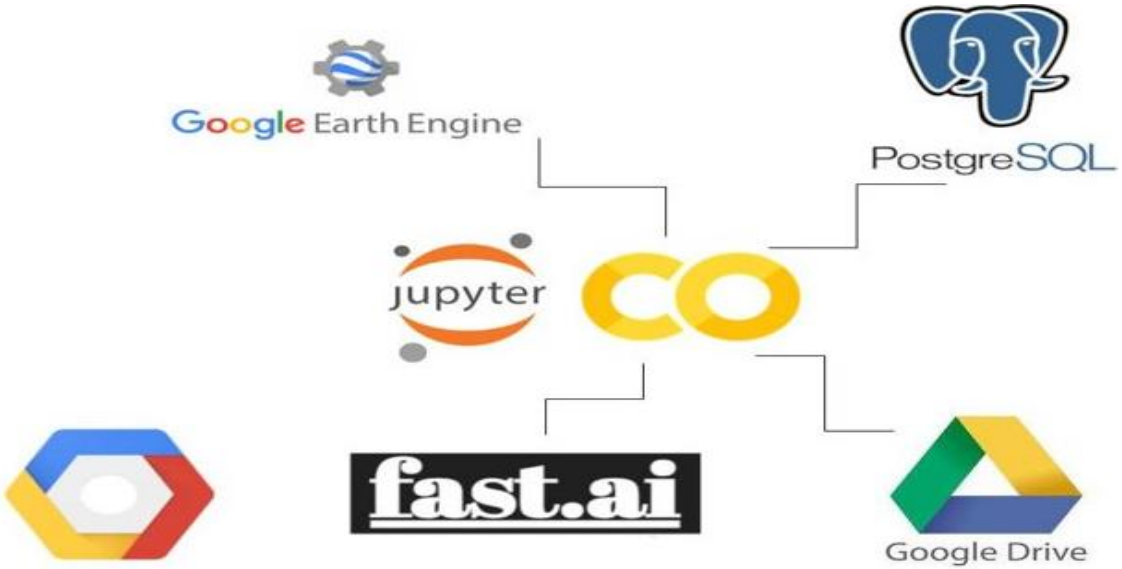
Figure. Intensity of Night Lights



Applying on PHI and THA data

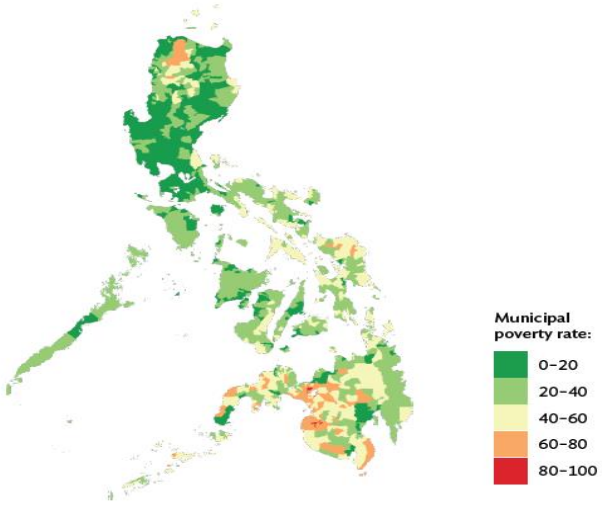
To avoid starting from scratch, we use an off-the-shelf CNN called ResNet34. This algorithm has been pretrained using the ImageNet database to ensure that it is capable of identifying simple features. ImageNet is regarded as a solid benchmark performer in computer vision predictions.

We also leveraged on readily available analytical platforms:

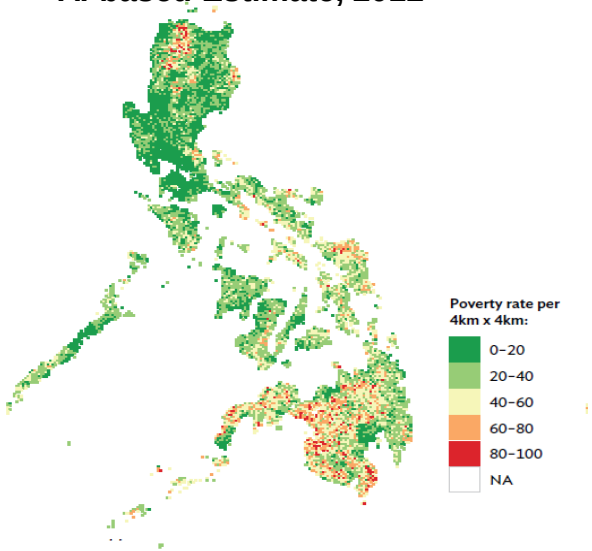


Outlining the Key Findings of the Feasibility Study

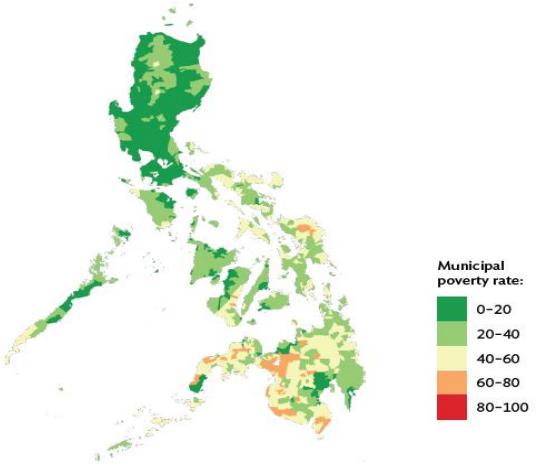
Published, 2012



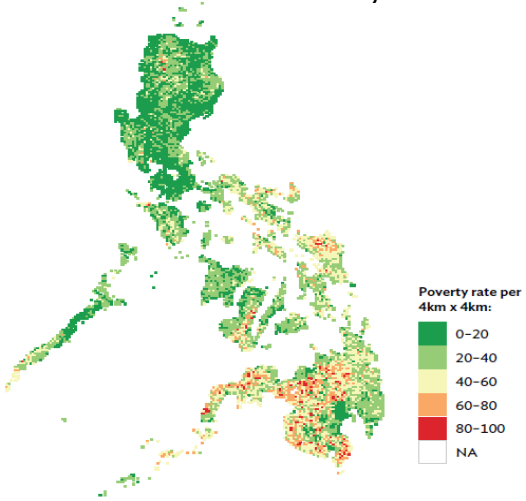
AI-based Estimate, 2012



Published, 2015

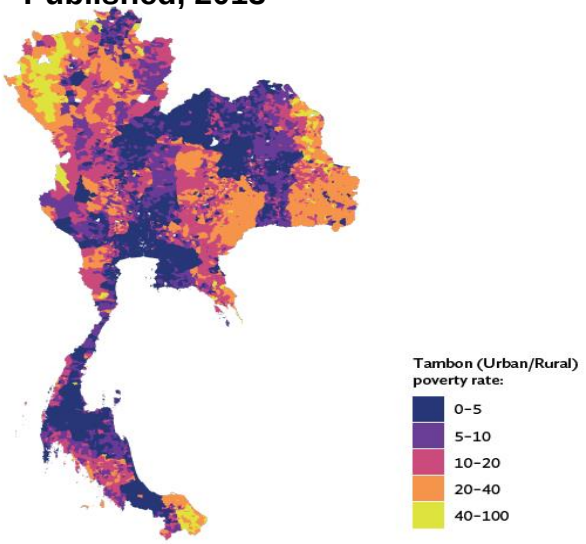


AI-based Estimate, 2015

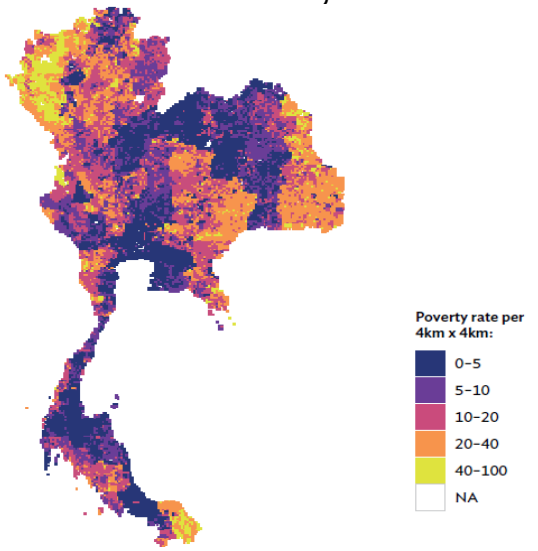


Outlining the Key Findings of the Feasibility Study

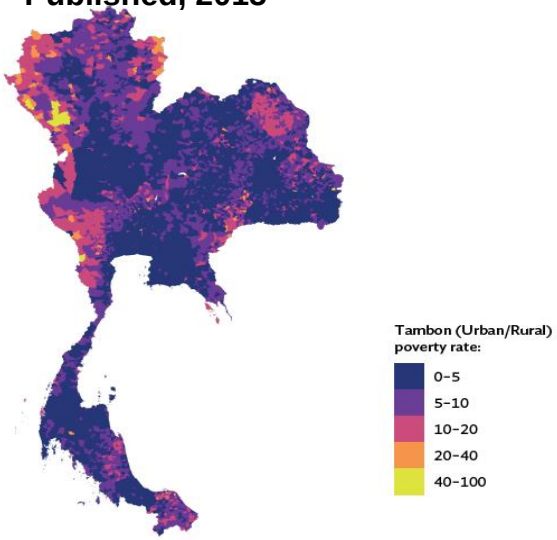
Published, 2013



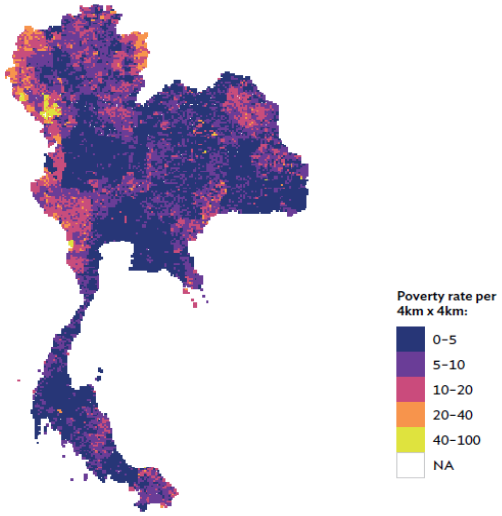
AI-based Estimate, 2012



Published, 2015

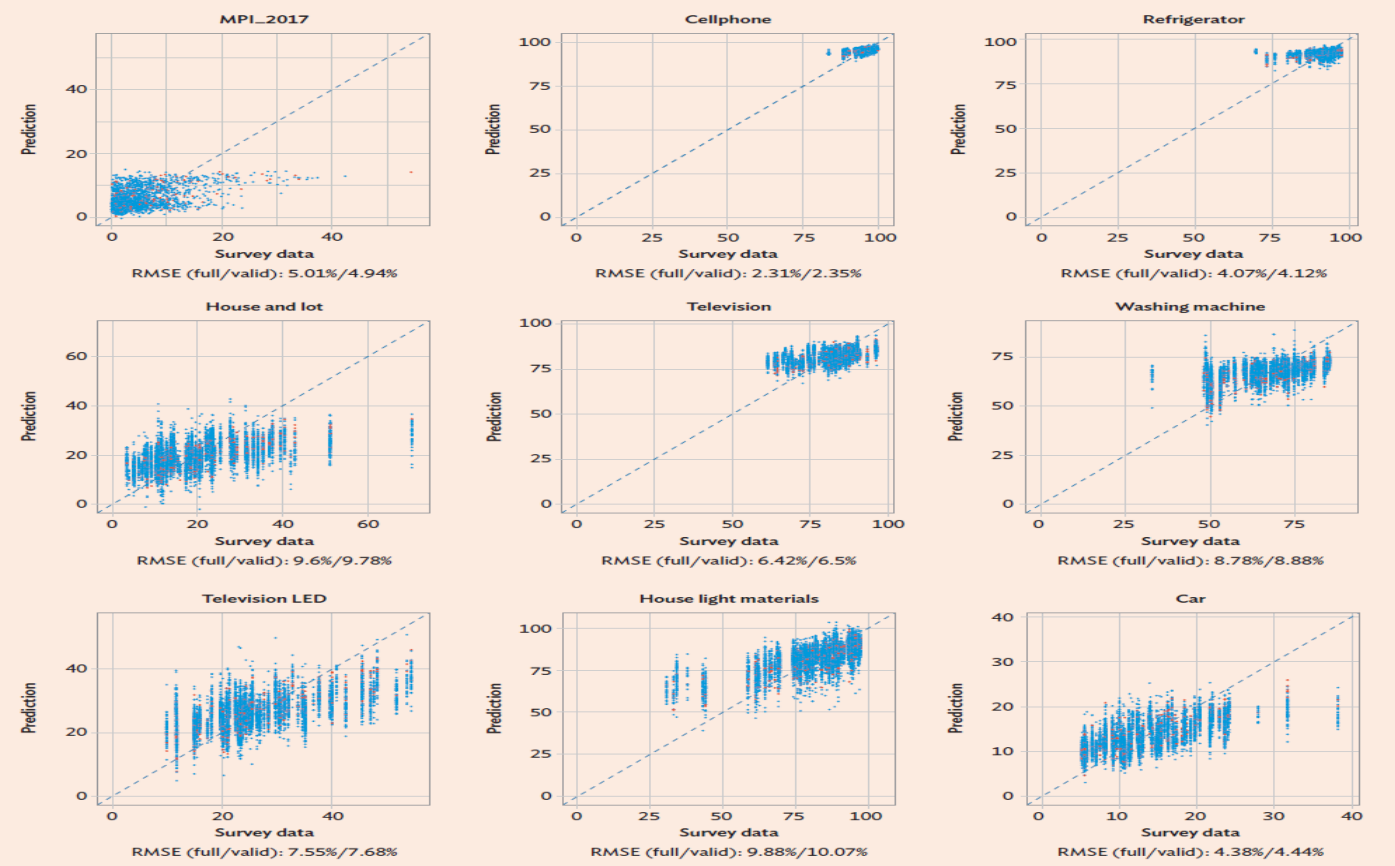


AI-based Estimate, 2015



Outlining the Key Findings of the Feasibility Study

We could potentially use a similar methodology to predict other non-income measures of poverty (e.g., multidimensional poverty index, ownership of durable assets, etc).



full = full data set; valid = validation set.
Source: Calculations and graphics generated by the study team.

Summary and Conclusion

The results of applying artificial intelligence on specific datasets from the Philippines and Thailand are encouraging. Even using publicly accessible satellite imagery, in which the resolutions are not as fine as those in commercially sourced images, the study produced predictions that aligned with the government-published poverty estimates.

For the purposes of small-scale poverty studies using nontraditional data, NSOs can streamline their resources and avoid substantial upfront costs by capitalizing on publicly accessible satellite imagery, affordable cloud services, and computational tools.

Nevertheless, further gains in the granularity of statistics potentially can be made if higher resolution imagery is used in future studies. NSOs also need to make investments on faster local computation equipment, and other IT infrastructure, and more importantly, on developing their human capital, such as appropriately skilled data scientists, to collect, process, and analyze geospatial and other nontraditional data, and to integrate new data collection methodologies into their work programs.

Summary and Conclusion

To fully capitalize on innovative data sources, forging partnerships and cooperation between multiple stakeholders is critical.

Ongoing engagement between data stakeholders and policymakers can help ensure that the statistics being compiled by NSOs are used to enhance society and improve living standards, particularly for the poor and disadvantaged.



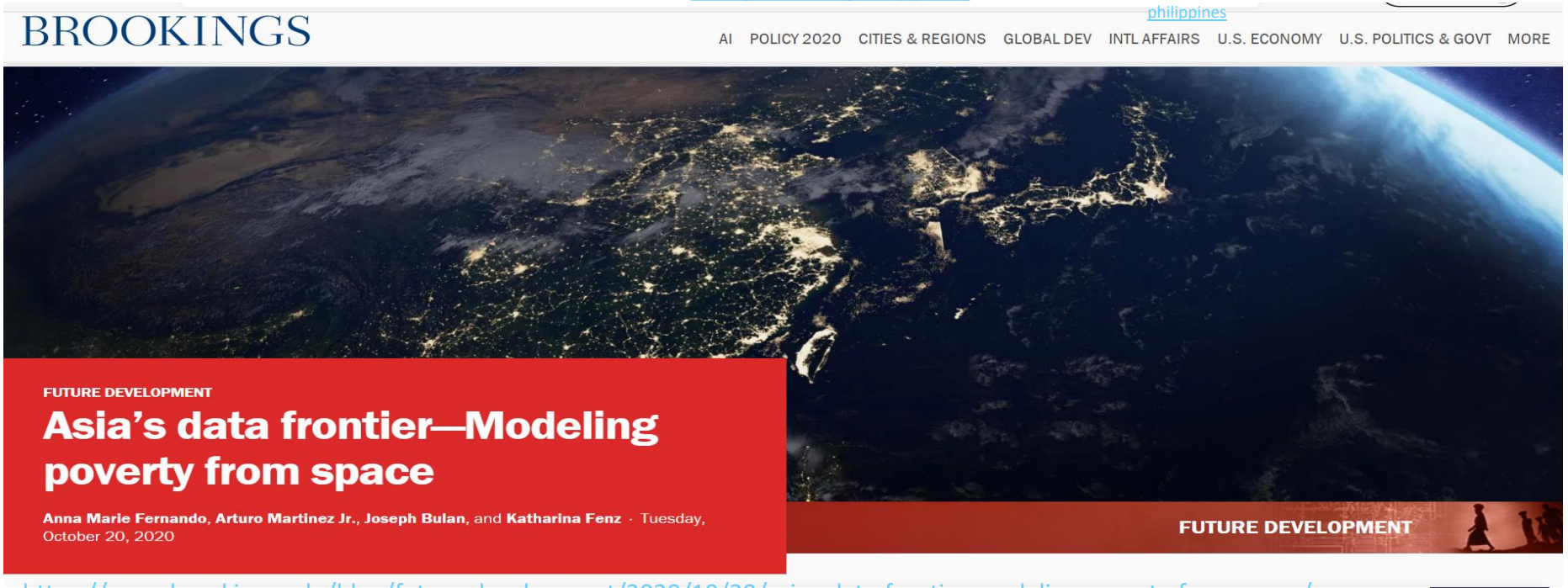
https://www.adb.org/news/events/webinar/asian-impact



https://development.asia/insight/using-machine-learning-satellite-images-map-poverty



https://blogs.adb.org/blog/here-s-how-we-used-satellite-data-to-map-poverty-in-thailand-and-philippines



https://www.brookings.edu/blog/future-development/2020/10/20/asias-data-frontier-modeling-poverty-from-space/

Thank you!

email: amartinezjr@adb.org