Using Mobile Phone Data For Measuring SDGs

Over the past decade, national statistical offices have been tasked with producing accurate, high quality statistics faster, more frequently and with a reduced burden on respondents. Yet they are expected to achieve this within the same or even a smaller budget. To face this challenge, national statistical offices can turn to innovative data sources like passive mobile positioning data (MPD).

Even though the private sector was the first to harness big data, the practice has now expanded to the global statistical community. The United Nations Statistical Commission and national statistical offices are looking into ways of using big data sources to complement official statistics and better meet their objectives for providing timely and accurate evidence for policy-making.

A project, led by ITU, was conducted in 2020 to demonstrate how big data can be used to produce internationally agreed ICT SDG indicators 9.c.1 (Proportion of population covered by a mobile network) and 17.8.1 (Individuals using the Internet). The feasibility of using MPD for both indicators was tested in both Brazil and Indonesia, one of each presented in the case studies. The collaboration showed that public and private sector organisations can work together for societal interest to leave no one behind.

Any national statistical office (or other organisation) who wants to calculate SDG indicators from MPD should follow certain steps

1 DATA ACCESS

Before the project begins, NSOs should ensure the availability of necessary data processing infrastructure, data science skills and access to the following data:

- mobile positioning data, i.e. metadata on calls, messages and internet usage;
- mobile network cell locations, i.e. from the mobile network operator, the telecom regulator or crowdsourced database like OpenCelliD;
- GIS files of local administrative units, i.e. shapefiles;
- population grid, i.e. from WorldPop database;
- any necessary spatial data, i.e. digital elevation model.

2 INPUT QUALITY ASSURANCE (QA)

Check input data to ensure quality before calculations:

- prepare and run QA on local administrative unit layers;
- run QA of mobile positioning data to ensure it is reliable for the required task;
- download and validate cell location data;
- download and validate any external data (i.e. population grid, digital elevation model), compare with other available data, if necessary.

3 PROCESSING

Implement all processing steps:

- calculate coverage areas;
- calculate and validate home locations;
- calculate indicators.

4 OUTPUT QUALITY ASSURANCE

Implement all processing steps:

- NSO validates results of the calculated indicators;
- calculated indicators are visualised to analyse regional differences.

INDONESIA

USING MPD TO MEASURE SDG 9.c.1

The project was implemented by BPS Statistics Indonesia, together with Bappenas (Ministry of Planning), ITU and Positium, and was submitted for Voluntary National Review. The project partners have previously been collaborating to calculate official tourism statistics with MPD since 2016. Now, mobile positioning data was used to estimate the SDG indicators 9.c.1 and 17.8.1. Here we present the results of one of the two indicators included in the project, indicator 9.c.1, proportion of population covered by a mobile network. While the indicator has already been calculated by BPS in Indonesia every year with administrative data, there were no geographic breakdowns and it took time to compose these, resulting in delays in publication.

OBJECTIVES

- Monitor SDG indicator 9.c.1 and include it in Indonesia's Voluntary National Review annually, while increasing the timeliness and disaggregation of both indicators;
- answer policy needs for SDG indicators and meet the 2030 SDG targets set by the Ministry of Development Planning;
- explore new data sources to reduce cost from and necessity of conducting an annual household survey;
- compare MPD results with the household survey to check for accuracy.

CHALLENGES

- Access to data – operators are not very keen on sharing the necessary data;
- knowledge required for the new methodology for measuring ICT indicators;
- infrastructure challenges – using the mobile network operator sandbox, electricity breakouts, long data processing time given the data volumes necessitated investments in infrastructure;
- assessing the differences between population estimates from BPS and WorldPop.

SOLUTION

BPS Statistics Indonesia itself gained access to the data and conducted verification of the results. The ITU Handbook with its defined methodology made data processing smoother at every step. To improve processing power, investments were made into infrastructure and software. Based on the digital elevation model, a coverage model was built, and assumptions for signal propagation were made. Crowdsourced data (from OpenCellID database) was used to determine cell locations. A population grid was harnessed to validate population projection from MPD.

Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation

Target 9.c: Significantly increase access to information and communications technology and strive to provide universal and affordable access to the Internet in least developed countries by 2020

Indicator 9.c.1: Proportion of population covered by a mobile network, by technology

9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

KEY RESULTS

For the first time, regions in Indonesia had data for mobile network coverage. Thanks to big data, the coverage could be calculated for even 1 km grid level. The calculated coverage areas were realistic and, therefore, calculated indicator values were assessed to have good quality. The mobile network coverage indicator from big data allows deep inspection of areas where coverage is still missing compared to population figures, whereas previously this analysis was only done at municipality border level in Indonesia.

It was proven that additional data sources – OpenCelliD and WorldPop – can be used to calculate the proportion of population covered with mobile networks without requiring any partnerships. OpenCelliD data is similar to data from mobile operators, and use of the digital elevation model provided a way to include the topography of the country in the coverage calculation. The mobile network coverage indicator from big data allows deep inspection of areas where coverage is still missing compared to population figures, whereas previously this analysis was only done at municipality border level in Indonesia. The initial exercise developed into a method which, combined with gained experience, will be used to replicate the calculations regularly and submit timely results to ITU.
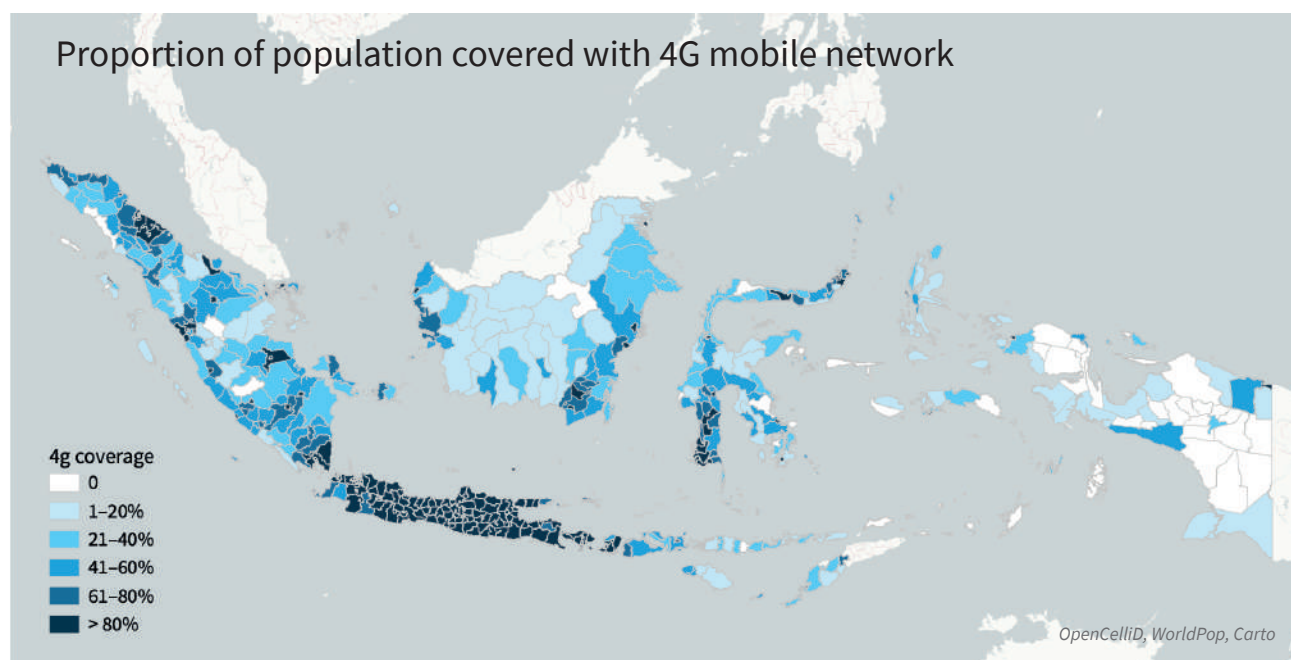

Figure 1. Proportion of population covered by 4G

USING MPD TO MEASURE SDG 17.8.1

In Brazil, an innovative data source – mobile positioning data (MPD) – was used to obtain trusted statistics for SDG indicators for the first time. The project was led by the Brazilian Institute of Geography and Statistics (IBGE), who is responsible for finding out how many individuals use the Internet (17.8.1). Although indicator 9.c.1 is not a focus for IBGE, it was also calculated, but will not be covered within this case study. IBGE collaborated with the Regional Center for Studies on the Development of the Information Society (cetic.br), ITU and Positium. The results of the analysis were verified for accuracy and robustness by comparing with similar data from traditional data sources (surveys). Through this process, IBGE gained experience in handling and processing MPD.

OBJECTIVES

- Have results on a more disaggregated level than the household survey provides;
- compare results with the household survey to determine whether mobile phone data can complement it;
- get timely data and statistical results.

CHALLENGES

- Data access – not all mobile network operators were willing to provide data for a pilot project;
- not knowing what methodology or skills are needed to work with this type of data;
- reliability of the data received from the mobile network operator is unknown as well as the method for assessing data quality;
- infrastructure to handle volumes of big data.

SOLUTION

IBGE reached out to experts on mobile phone big data and to other countries who had previously done a similar analysis. With the support of the ITU Handbook and its defined methodology, every step of the data handling was smooth. Working together with national stakeholders, they were able to facilitate access to the data and verify the results. This included doing a quality check on raw data to assess if it was fit-for-use. Investments were made to infrastructure, particularly to improve software and hardware for data processing, and to train data scientists and staff in this area of MPD for statistics.

Goal 17: Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development

Target 17.8: Fully operationalize the technology bank and science, technology and innovation capacity-building mechanism for least developed countries by 2017 and enhance the use of enabling technology, in particular information and communications technology

Indicator 17.8.1: Proportion of individuals using the Internet

17 PARTNERSHIPS FOR THE GOALS

KEY RESULTS

Internet use statistics were very close to the household survey results when comparing access to the internet using a mobile phone. IBGE gained new experience and knowledge in working with MPD. Now, a protocol has been made for the future incorporation of MPD in the IBGE statistical production pipeline. Results are surprisingly similar for the total area of study, with a difference of up to 1.04% in the city of Rio de Janeiro. Therefore, the results suggest the method is robust and can be used to produce official statistics. Despite the definition difference between the SDG indicator and the measure using MPD, the last can be a good proxy for the first one.
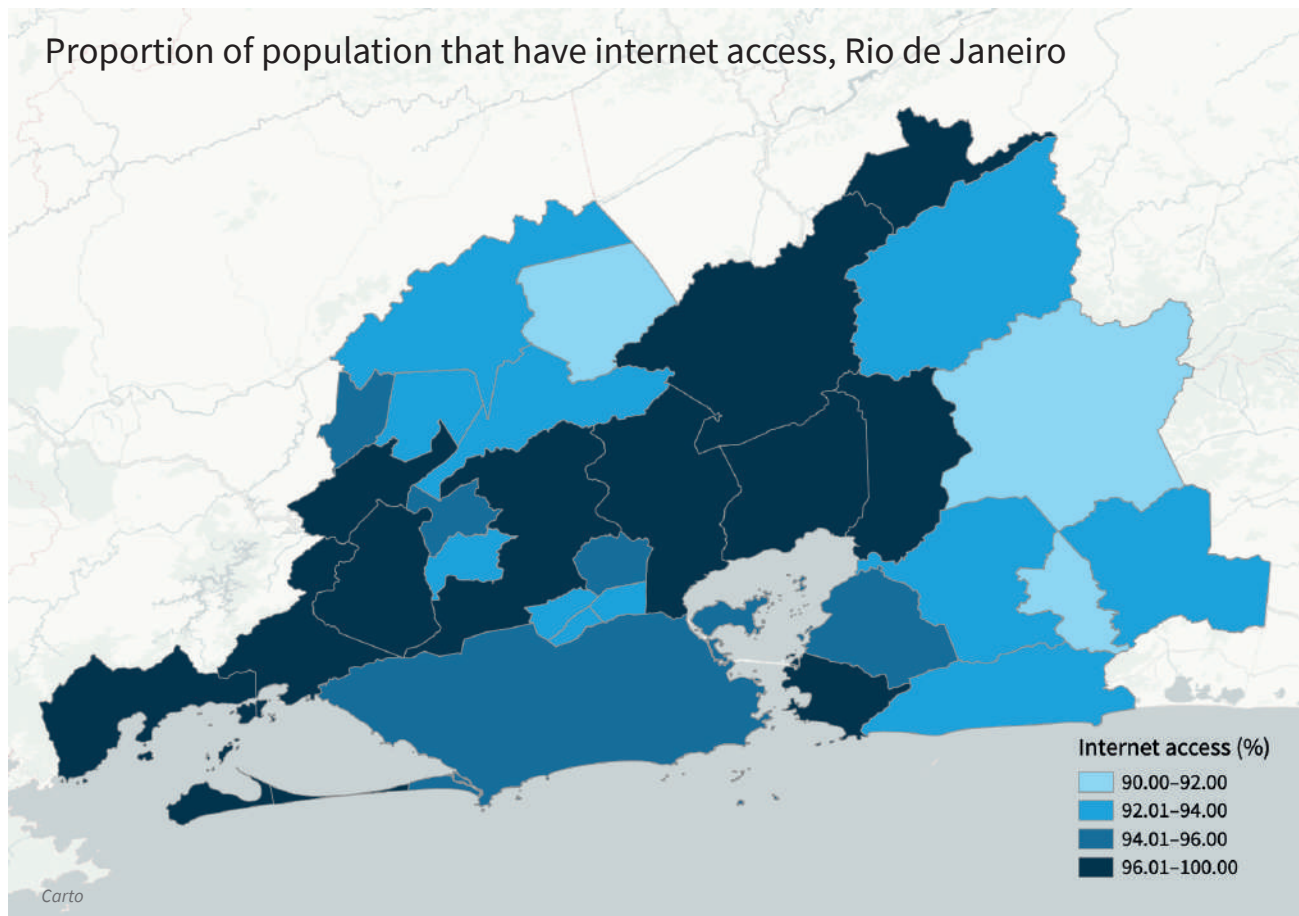
Proportion of population that have internet access, Rio de Janeiro

Internet access (%)
90.00–92.00
92.01–94.00
94.01–96.00
96.01–100.00

Figure 2. Proportion of population that have internet access, Rio de Janeiro

CONCLUSIONS

The projects in Indonesia and Brazil had a positive outcome, showing good results in terms of quality of MPD and how well it compares to reference data. This suggests that MPD, together with other datasets, can be used to calculate ICT SDG indicators in other countries as well. MPD helps to produce statistics in a timely and transparent manner with increased spatial resolution. The digital nature of MPD means it is generated constantly and is therefore very well suited for monitoring SDG trends over time. What is more, indicator 9.c.1 can be calculated even without access to mobile network operator data, simply by following the steps presented in the ITU Handbook created within this project.