Opportunities and Challenges of Remote Sensing, Geospatial Data, and Machine Learning in Obtaining Accessibility and Location Information for Sustainable Development in Indonesia

T Devara[1,*]

[1]BPS-Statistics of Bima Regency, Nusa Tenggara Barat, Indonesia

*Corresponding author's e-mail: terry.devara@bps.go.id

Abstract. With the advancement of technologies so does the data collection method which creates a large, rapid, and diverse stream of data. Statistic Indonesia (BPS) has also encouraged to utilize this by starting to collect geospatial information on respondents and public facilities. To keep up with this a change needs to be made in processing methods to accommodate massive, high-dimensional, and multiform data collected in different forms such as machine learning. This progression also opens up a new opportunity for tackling various statistical data problems such as accessibility and location data. Remote sensing is one of the big data sources that undergoes a lot of changes shown in the high spatial and temporal resolution satellite imagery availability, together with the BPS geotagging data shows great promise in classifying land use and geospatial analysis. Even so, there are still some challenges in remote sensing as well as other geospatial data utilization. The goals of this review paper are to study the opportunities and challenges in utilizing remote sensing, geospatial data, and machine learning for accessibility and location information. In this paper, we explore the possibilities and limitations in its implementation into SDGs indicators that involve accessibility and location such as indicators 9.1.1, 11.1.1, 11.2.1, 11.3.1, and 11.7.1 including other variables needed for the calculation like access to public facilities. Moreover, our experiment using geotagging data shows potential in improving proportion estimation when compared to using a simple ratio. Our DEGURBA following the UN definition using machine learning LULC for dasymetric mapping also provides more insight compared to the existing data. We can conclude that there are great opportunities in applying remote sensing and other geospatial data to monitor the accessibility and location to further sustainable development in Indonesia.

1. Introduction

To ensure prosperity while considering sustainability for both developed and developing nations, the United Nations proposed a program by the name of Sustainable Development Goals (SDGs) for the good of human beings and the sustainability of the planet by 2030. Even then, Indonesia still lacks the necessary way to monitor the much-needed SDGs indicator. In 2021, BPS could only provide 92 out of 289 SDGs second-edition indicators by either itself or with stakeholders with only 50 that meet the global metadata [1]. This means that since the global indicator framework for Sustainable Development Goals was developed and agreed upon in March 2017, one-third of the allocated time has passed but

BPS can't even provide one-fifth of the indicators that meet the global metadata essential for reaching these goals.

This is due to a vast number of these indicators either can't be calculated using the conventional method or the cost of using the conventional method is too massive. The census in Indonesia was done once every 10 years and it applies to population, agriculture, and economic census. While on the year in between, data was collected using estimates from surveys which only available in villages, districts, regencies, or other levels above. Meanwhile, if the metadata were to be followed, data on a smaller unit was necessary and the data would need to be disaggregated. This limitation and unavailability of the required data added to the lack of resources both human and monetary leads to the indicator released in 2022 not being up to global parameters [1].

At the same time, technology has advanced so much that other means of data collection are available. Remote sensing data has become widely used along with machine learning as a new means of data processing and analysis. One of the remote sensing data is satellite imagery which provides spatial information on surface reflection daily and is publicly available which has proven to be applicable and helped in all kinds of fields such as crop monitoring [2], [3], or disaster relief [4]–[6]. This advancement in technology has also provided us with other forms of data other than traditional tables such as road network maps or public facilities geotagging. This data could also be used to monitor SDGs indicators such as monitoring the ratio of land consumption rate to population growth rate [7], [8] or even monitor the accessibility of public facilities such as school [9] or healthcare [10].

BPS (Statistic Indonesia) has also implemented this technology by starting to collect respondent geotagging data which provides the geolocation of respondents' residents. While it was meant to provide more accurate data such as by name and address, and an effort to provide geospatial data of the corresponding census or survey, it can also be utilized for other means such as mapping Degree of Urbanization (DEGURBA), analyzing disaster/hazard risk, or accessibility. Several other sources that provide geospatial data in Indonesia are OpenStreetMap (OSM) which offers data such as public facilities, road maps, or railway track maps, and Ina-geospatial which acts as Indonesia's national geoportal.

With the availability of these new data sources and the usage of machine learning so are the opportunities to apply it to analyze location and accessibilities to better support sustainable development in Indonesia. To that end, DEGURBA data is vital since several SDGs indicators that focus on accessibility were divided into rural or urban areas. While DEGURBA data was not yet available in Indonesia, other data that could be used to obtain it was readily available. Other than that, disaster risk analysis can only be done by using spatial data using locations and past or possible disaster maps.

Despite all the opportunities provided by remote sensing and other spatial data with earth observation data and machine learning, there is still some hurdle that needs to be tackled. The availability and quality of the collected spatial data such as public facilities location, road maps, or disaster maps still leave some room for improvement. The diversity in Indonesia's land cover also makes feature selection to improve accuracy to be quite challenging. To fully utilize remote sensing, spatial data, and machine learning, we need to identify all the challenges. Therefore, in this paper, we would like to study and explore the opportunities and challenges of remote sensing, spatial data, and machine learning in obtaining accessibility and location information for sustainable development in Indonesia by studying past research and doing a comparison with the current state in Indonesia.

2. Remote Sensing, Machine Learning, and Geospatial Data

As the city and population grow, so does the need for infrastructure and facilities to keep up with the public needs. While there are already surveys by Statistics Indonesia such as the socioeconomic surveys (SUSENAS) that monitor access to public facilities, it was still crude at the very best. This was due to the question in SUSENAS not considering accessibilities in the term that a facility was within a reachable distance and doesn't need an extra effort to access it but consider that a population has access to a facility if the population used that facility before. Thus, a better way to monitor accessibility and even the location of residents and infrastructure is of the utmost importance.

2.1. Remote Sensing
Remote sensing technologies have experienced a massive advancement in recent years. With each passing year, a new satellite was launched with improvements in sensor, resolution, bands, and revisit period. Different technologies were also equipped to better fulfill the different purposes of each satellite like collecting surface reflectance, monitoring sea surface, night light, air quality, etc. A higher resolution with more bands will provide more information which is useful in analysis, A higher image resolution also means a bigger data size and more bands mean more dimensionality in the data which in turn makes conventional analysis quite challenging.

2.2. Spatial Data
In recent years, Statistics Indonesia developed an application by the name of Wilkerstat (Statistical Working Area) to simplify the mapping process and keep the statistical working area up to date [11]. As time went on, the app evolved by providing geotagging capabilities which are already implemented in the recent census and survey, including the 2022 Social Economic Registration (REGSOSEK) and 2023 Agricultural Census. On top of that, the statistical working area was also changed from census blocks to SLS (lowest local administrative area). Not only does it make data consistencies between institutions possible, but it also helps to provide insight into the local government.

2.3. Machine Learning
Machine Learning has become a staple in analyzing remote sensing and spatial data because of its ability to process a large amount of data in a short period and its ability to take on many variables. Machine learning algorithms are also versatile in their usability which can be applied to remote sensing data to create a Land Use Land Cover (LULC) map [2], [3], [12]–[16], land change detection [17], and dasymetric mapping of population density [18]–[21] which is integral to classifying the degree of urbanization (DEGURBA) [22]–[24]. It can also be used to detect areas of interest such as built-up areas [25]–[27] and oil palm detection [28], or used to map poverty [29]–[31].

3. Opportunities and Challenges
The opportunities and challenges were made per UN SDGs indicator metadata by reviewing past research and guidelines made by the corresponding International Organization responsible for global monitoring. Some experiment was also conducted to explore the applicability of each indicator, data, and method in Indonesia due to difference in concept and definition in several variables.

3.1. The Proportion of Population with Access to Public Facilities
While there are all different kinds of public services, quality education, and health care are an integral part of a country's development. Therefore, an inclusive learning environment and affordable healthcare should be accessible to every population. The school enrollment rate in Indonesia in 2022 was 98.08 for the population aged 7-15 years old and 91.92 for the population aged 7-18 years old [32]. Whereas the proportion of the population with access to healthcare was 79.33 [33]. Since these numbers were based on SUSENAS, there are some weaknesses in which the school enrollment rate was not a good representation of an inclusive learning environment. While the school enrollment rate was high, cases where students had to travel a great distance or even had to stay in relatives' house to have access to education, was still common situation.

Based on regulation no 1 in the year 2021 by the Ministry of Education, Culture, Research, and Technology [34], a major proportion of school enrollment had to be based on the regional zonation of the school which is regulated by the local government. These regulations though, mainly can only be applied in urban areas. In West Java, a province with the largest population in Indonesia numbering 41,150 thousand [35], the education accessibility, especially for a higher level of education itself was moderate for middle-high school and moderate to low for high schools [9].

The same thing applies to healthcare access. The information collected through SUSENAS was limited to health insurance, diseases, and whether the population used a health facility such as a hospital or not without taking into consideration the hardship of accessing those health facilities. Additional spatial data such as maps of residential areas, public facilities, and roads could be utilized to estimate the distances between schools and their students' residents or between hospitals and the patients' homes. This way, we could get the accessibility information at least from the distance perspective. Accessibility was crucial especially for healthcare facilities since it was closely related to risk measurement, particularly during an emergency such as labor or accident. Studies have been done on using road network data to calculate the driving distance between residential areas and public facilities. While straight line distance remains valid, driving distance and time which could be done using the SAS FILENAME URL method in SAS version 9.1 from Google Maps were still preferred especially in the area of emergency response where the result was sensitive to the smallest difference since the cost has become negligible [36]. The problem with straight-line distance though was its inability to put uncrossable areas such as lakes, rivers, mountains, or cliffs into consideration.

Another alternative was to use floating catchment areas (FCA) methods which assume that the population will only use services within their catchment areas. This method has also been modified using spatial decomposition which is then called a two-step floating catchment area method (2SFCAM), which not only accounts for the supply but also the demand [37], [38]. The application of 2SFCAM itself was calculated using the maximum travel time or catchment size. Further improvement was made by introducing distance decay because when the catchment size was too big, the method would start measuring ‘choice’ rather than accessibility [39]. Whereas in urban areas with heterogenous transportation methods, a multi-mode 2SFCAM was introduced since a single-mode 2SFCAM tends to over-estimate while under-estimate in rural areas [10].

3.2. Rural Area with Road Access

SDGs indicator 9.1.1 Proportion of the rural population who live within 2 km of an all-season road which is also commonly known as the Rural Access Index (RAI) is the measurement of the rural population living within 2 km of an all-season road [40]. To make this measurement, population data, DEGURBA data, and road data were essential [41]–[43]. This could be achieved by combining land cover maps from remote sensing and machine learning with population census data which resulted in the urban-rural map, then combining the result with road data that could be collected from RBI (Rupa Bumi Indonesia) map or OpenStreetMap.

The proportion of the rural population with access could be roughly estimated using the proportion of rural areas within 2 km of an all-season road to the overall rural area itself. Our experiments show that better estimation could be performed by utilizing geotagging data of each household (Figures 1 and 2). This was due to the geotagging location being exclusive to residents and excluding other areas such as fields, parking lots, markets, etc. However, This method was limited by the data availability whereas the geotagging data may be unavailable and needs to be updated regularly which puts a lot of financial burdens.

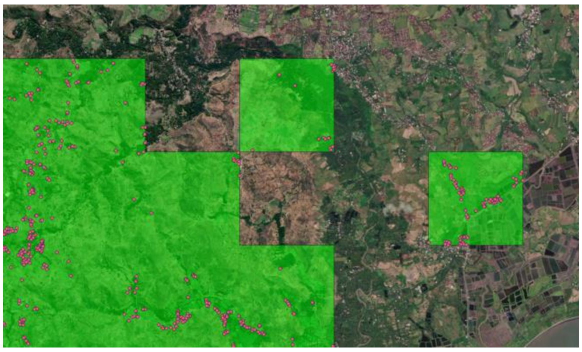

Figure 1. Rural area map (green part) with geotagging.

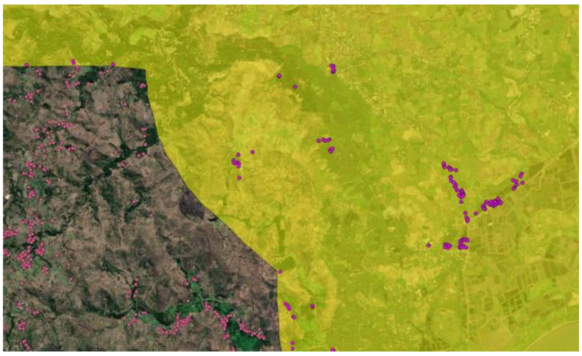

Figure 2. Geotagging in rural areas (yellow part) within 2 km of an all-season road.

While the method for the estimation was quite simple, the problem mainly lies in the street data. First, the official street data provided by RBI was outdated compared to OSM as seen in Figure 3,4, and 5. Although the OSM maps themselves were not perfect, in most cases, the data was more complete compared to RBI maps (locations 1 and 2) since there are some areas where both maps are lacking (location 3), some areas even show that RBI maps have some faulty lines or imaginary roads (location 2). Another weakness of OSM data was that it was collected by volunteers and as a result the data quality or even concept and definition would vary in different locations.

Figure 3. Google Maps image of three different locations

Figure 4. RBI Maps image of three different locations

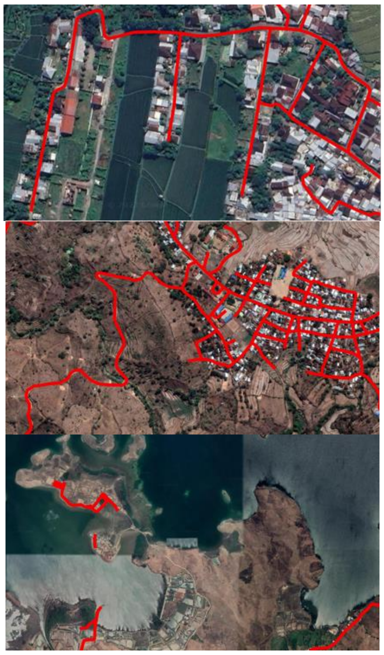

Figure 5. OSM image of three different locations

3.3. The Proportion of the Population Living in Slums

The housing sector has been a common problem for every country since it correlates with the country's economy. The biggest challenge in monitoring slums was the broad definition of slums itself since the term was highly politicized. To measure the Proportion of the urban population living in slums, informal settlements, or inadequate housing (SDGs indicator 11.1.1), UN-Habitat defines slums using criteria shown in Table 1 [44], [45].

While BPS already provided indicator 11.1.1 [1], it doesn’t follow the defining criteria from the metadata and only uses variables collected through SUSENAS such as access to water and sanitation, sufficient living area/overcrowding, and structural durability [46]. Meanwhile, variables such as accessibility (whether the disadvantaged and marginalized groups have their needs provided) and location (access to public services, employment opportunities, and not in risky areas) were obtainable by analyzing remote sensing, spatial data from Wilkerstat, and machine learning. Based on the metadata, risky areas encapsulate geological hazardous zones (earthquakes, landslides, floods, eruptions, etc.), polluted areas (high industrial pollution areas or under garbage mountains), and other unprotected high-risk zones (in the proximity of railroads, airports, energy transmission lines).

Table 1. Criteria defining slums, informal settlements, and inadequate housing.

	Slums	Informal Settlements	Inadequate Housing
Access to water	X	X	X
Access to sanitation	X	X	X
Sufficient living area, overcrowding	X		X
Structural quality, durability, and location	X	X	X
Security of tenure	X	X	X
Affordability			X
Accessibility			X
Cultural adequacy			X

Data such as LULC, soil type, slope, and elevation could be used to create a multi-hazard probability map [9] and combined with Wilkerstat data to map out housing in areas with a high probability of natural disasters. The multi-hazard risk map can also be made by feeding records of past natural disasters into a machine-learning algorithm [47]. Wilkerstat can also be combined with railroad/transmission lines data to map out housing in unprotected high-risk zones or with rivers connected to factories to map housing in polluted areas.

Albeit the past disasters information was available via Indonesia Disaster Information Data (DIBI) which is managed by the National Agency for Disaster Countermeasures, there are no available data about high pollution factories, garbage dumb, or other polluted areas. Meanwhile, pollution in Indonesia has become a quite serious matter. Several studies have shown that the Citarum River which runs through several cities and regencies and is even the main water source for the population in the riverbank and even for the majority of Jakarta was very heavily polluted [48] and way past the ability of the water body to receive pollution load without causing pollution [49]. The possibility of remotely monitoring pollution has been explored such as monitoring sewage outfall through the usage of Unmanned Aerial Vehicles (UAV) [50], soil heavy metal pollution monitoring using hyperspectral remote sensing and improved gradient boosting [51], and atmospheric air pollution modeling using Sentinel-5 and Google Earth Engine (GEE) [52].

3.4. The Proportion of Population with Access to Public Transportation

The goal of SDGs 11.2.1 (Proportion of the population that has convenient access to public transport, by sex, age, and persons with disabilities) was to monitor public transportation system usage and accessibility while reducing the population's reliance on private transportation. The population was considered to have convenient access when a recognized stop was within 500 meters of walkable distance from the reference point (home, school, market, hospital) for low-capacity public transportation systems such as bus and 1 kilometer for high-capacity public transportation systems such as train [53]. The calculation of the indicator itself relies on public transport stop location, road network, population data, number of residences per dwelling unit, and demographic data. This indicator was calculated using DEGURBA data, public transport stop location, and road network data which is used to calculate the area within walking distance [53], [54]. The public transport stop location could be acquired by the transportation provider or on OSM feature using the Overpass API can also be used as a comparison.

Even then, this method was still very limited in terms of the convenience of the public transport system. This is caused by the method not taking into consideration performance like service frequency, capacity, or comfort. Public transportation should also consider the population's affordability where according to the recommendation, no more than 5% of the poorest quintile of the population's net household income should not be spent on transport which is collected in SUSENAS. The biggest challenge comes from the security perspective which comes down to the safety or security aspects. While reports of thief or sexual harassment in public transportation are in no way uncommon, there are no official statistics of such cases.

3.5. Land Consumption and Population Growth Rate

In recent years, the population growth rate in Indonesia has followed a declining trend numbering 1.22 in 2021 to 1.17 in 2022, and 1.13 in 2023 [55]. While the population growth itself has been decreasing, the population still increases each year which means that the development needs to keep up. SDGs indicator 11.3.1 (Ratio of Land Consumption Rate to Population Growth Rate) was proposed to monitor the change in land consumption and population growth to better understand why a city grows, what is the effect, how long the transition takes, and others that may help in future investment and planning [56]. Based on the metadata for indicator 11.3.1 [57], [58], the indicator can be calculated using a built-up area map (from LULC), population data, administrative map, and DEGURBA.

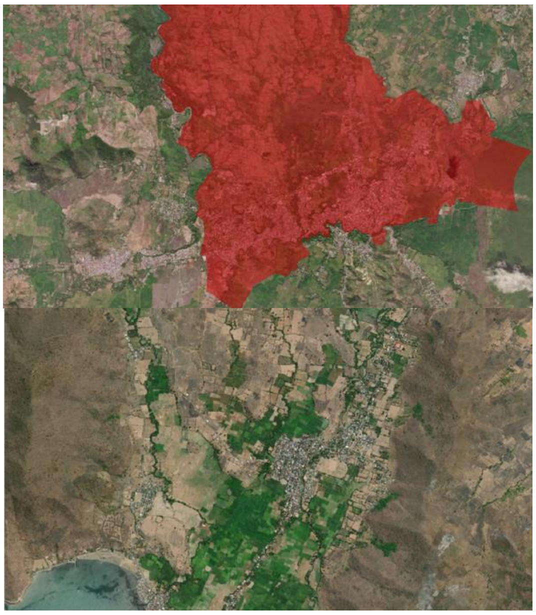

Figure 6. Urban area based on BPS definition

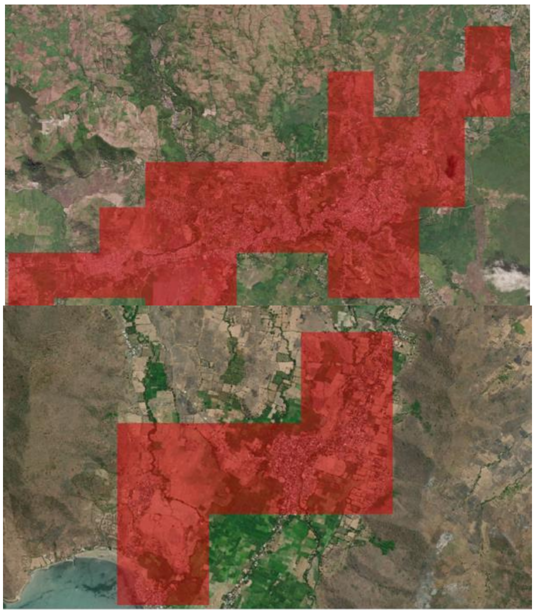

Figure 7. Urban area based on UN definition

Since in Indonesia, there is no official LULC map, this data would need to be processed from other data like satellite imageries using machine learning. Unlike normal image data which only consists of Red, Green, and Blue, satellite imagery data were much more complex. Sentinel-2 bands for example consist of 12 bands including Red, Green, and Blue ranging from 443 nm to 2190 nm [59]. The environment in Indonesia as an archipelago was also very heterogeneous which makes creating a robust model for classifying LULC quite challenging due to the different environment characteristics of each region.

If the metadata were to be followed, this indicator also relies heavily on DEGURBA classification since it is focused on urban areas where most of the population lives. If there are no urban areas detected, then that region is left out. Figure 6 and 7 shows an experimental result of a comparison of areas classified as urban using the current definition by BPS and the definition by the UN in Dompu Regency. Currently, BPS still classifies urban and rural areas by village based on the available infrastructure (image A) compared to the DEGURBA definition by the UN and used as metadata for SDGs indicator which is based on the number of population per kilometer (image B). Image A in the first location shows that a lot of uninhabited area was considered to be urban since it was part of a village that was classified as urban while leaving a big portion of the built-up area because it belongs to the neighboring village. The second location shows a region that doesn't have any urban area although it was densely populated because it doesn't have the necessary infrastructure. Therefore, to better understand how the population and cities grow, the DEGURBA in Indonesia would also need to be improved.

3.6. Public Open Spaces

SDGs indicator 11.7.1 (Average share of the built-up area of cities that is open space for public use for all, by sex, age, and persons with disabilities) refers to the share of a city built-up areas that is open public spaces and the share of the population with access to open public spaces [60], [61]. The share of a city's built-up areas that are open public spaces was estimated by comparing the total area of open public spaces and the total area of land allocated to streets with the total area of the city. The share of the population with access to open public spaces was calculated by the total population living within 400 meters of walking distance from open public spaces to the total population within a city. Since not every city or regency in Indonesia doesn’t have open public space inventories, satellite imageries combined with machine learning can be used to identify potential public open spaces. Open data sources such as OSM also have polygon data of open public spaces in some cities which could contribute to the effort.

Object detection could be applied to satellite imagery to help identify open public spaces. This information could then be combined with the road network to either calculate the share of a city's built-up areas that is open public space or used to estimate the number of people living within 400 meters of walking distance from the open public spaces. To estimate the number of people, the ratio of built-up areas within 400 meters of walking distance to the total built-up (from LULC) area could be used. An alternative method was using the ratio of geotagged houses within 400 meters of walking distance to the total geotagged houses. The advantage of this method was that other uninhibited built-up areas such as factories were excluded.

Based on the UN Statistical Commission in its 51st Session (March 2020), DEGURBA was used as a workaround method to delineate cities which means that it will face the same problems as indicator 11.3.1. Another consideration to take was the usage of road network data where the quality of the available data was quite limited as explained in indicator 9.1.1. On top of that, both RBI and OSM road data were in the form of lines instead of polygons and it doesn’t have road width information which makes calculating the total area of land allocated to streets quite challenging.

4. Conclusion

Opportunities and challenges of remote sensing, geospatial data, and machine learning have been explored in this paper. This shows that data from remote sensing and other geospatial data have the potential as a data source for monitoring sustainable development. To optimize this potential, an effective method for classifying LULC, especially built-up areas needs to be developed and better road network data is required. The usage of machine learning on satellite imagery could also be used to help improve the road network data [62], [63] which is crucial for estimating population with road access or public facility accessibility.

Better built-up area information was crucial for a better DEGURBA map since it can be used as a variable in dasymetric mapping to delineate the population density. Combining surveys with spatial data such as geotagging is also beneficial since some information like accessibilities and location was hard to get just by relying on tables.

Some recommendations for improving monitoring sustainable development in Indonesia are:

a. Combining optic and radar satellite imagery to better classify LULC.
b. Implementing DEGURBA per the UN 51st Statistical Commission for more comparable data on an International scale.
c. Exploring the usage of object detection for extracting built-up areas in Indonesia.
d. Exploring the usage of machine learning for extracting road network data in Indonesia.
e. Monitoring public facilities' ease of access instead of whether the population ever used certain facilities that year.

References

[1] BPS-Statistics Indonesia, “Indikator Tujuan pembangunan berkelanjutan indonesia 2022,” 2022, [Online]. Available: https://www.bps.go.id/publication/2022/12/22/500ea80678477e59232910a8/indikator-tujuan-pembangunan-berkelanjutan-indonesia-2022.html

[2] D. W. Triscowati, B. Sartono, A. Kurnia, D. Dirgahayu, and A. W. Wijayanto, "Classification of Rice-Plant Growth Phase Using Supervised Random Forest Method Based on Landsat-8 Multitemporal Data," *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, vol. 16, no. 2, p. 187, 2020, doi: 10.30536/j.ijreses.2019.v16.a3217.
[3] D. W. Triscowati, B. Sartono, A. Kurnia, D. D. Domiri, and A. W. Wijayanto, "Multitemporal remote sensing data for classification of food crops plant phase using supervised random forest," vol. 1131102, no. November 2019, p. 10, 2019, doi: 10.1117/12.2547216.
[4] H. L. Fitriana, S. Suwarsono, E. Kusratmoko, and S. Supriatna, "Mapping Burnt Areas Using the Semi-Automatic Object-Based Image Analysis Method," *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, vol. 17, no. 1, p. 57, 2020, doi: 10.30536/j.ijreses.2020.v17.a3281.
[5] V. Klemas, "Remote sensing of floods and flood-prone areas: An overview," *J Coast Res*, vol. 31, no. 4, pp. 1005–1013, 2015, doi: 10.2112/JCOASTRES-D-14-00160.1.
[6] O. M. Bello and Y. A. Aina, "Satellite Remote Sensing as a Tool in Disaster Management and Sustainable Development: Towards a Synergistic Approach," *Procedia Soc Behav Sci*, vol. 120, pp. 365–373, 2014, doi: 10.1016/j.sbspro.2014.02.114.
[7] R. Nicolau, J. David, M. Caetano, and J. M. C. Pereira, "Ratio of Land Consumption Rate to Population Growth Rate-Analysis of Different Formulations Applied to Mainland Portugal", doi: 10.3390/ijgi8010010.
[8] B. Calka, A. Orych, E. Bielecka, and S. Mozuriunaite, "The Ratio of the Land Consumption Rate to the Population Growth Rate: A Framework for the Achievement of the Spatiotemporal Pattern in Poland and Lithuania," *Remote Sensing 2022, Vol. 14, Page 1074*, vol. 14, no. 5, p. 1074, Feb. 2022, doi: 10.3390/RS14051074.
[9] A. D. Sakti *et al.*, "School location analysis by integrating the accessibility, natural and biological hazards to support equal access to education," *ISPRS Int J Geoinf*, vol. 11, no. 1, 2022, doi: 10.3390/ijgi11010012.
[10] L. Mao and D. Nekorchuk, "Measuring spatial accessibility to healthcare for populations with multiple transportation modes," *Health Place*, vol. 24, pp. 115–122, 2013, doi: 10.1016/j.healthplace.2013.08.008.
[11] BPS-Statistics Indonesia, *Pedoman Pengolahan Peta di BPS Kabupaten/Kota Pemetaan dan Pemutakhiran Muatan Wilayah Kerja Statistik Sensus Penduduk 2020*. BPS-Statistics Indonesia, 2019. [Online]. Available: https://sirusa.bps.go.id/webadmin/pedoman/2019 3569 ped Pedoman Pengolahan Peta di BPS Kabupaten-Kota.pdf
[12] Di. Ienco, R. Gaetano, C. Dupaquier, and P. Maurel, "Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1685–1689, 2017, doi: 10.1109 /LGRS.2017.2728698.
[13] N. M. Farda, "Multi-temporal Land Use Mapping of Coastal Wetlands Area using Machine Learning in Google Earth Engine," *IOP Conf Ser Earth Environ Sci*, vol. 98, no. 1, 2017, doi: 10.1088/1755-1315/98/1/012042.
[14] N. Kussul, G. Lemoine, F. J. Gallego, S. V. Skakun, M. Lavreniuk, and A. Y. Shelestov, "Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 9, no. 6, pp. 2500–2508, 2016, doi: 10.1109/JSTARS.2016.2560141.
[15] Y. Nurmasari and A. W. Wijayanto, "Oil Palm Plantation Detection in Indonesia Using Sentinel-2 and Landsat-8 Optical Satellite Imagery (Case Study: Rokan Hulu Regency, Riau Province)," *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, vol. 18, no. 1, p. 1, 2021, doi: 10.30536/j.ijreses.2021.v18.a3537.
[16] T. Devara and A. W. Wijayanto, "Machine Learning Applied To Sentinel-2 and Landsat-8 Multispectral and Medium-Resolution Satellite Imagery for the Detection of Rice Production

Areas in Nganjuk, East Java, Indonesia," *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, vol. 18, no. 1, p. 19, 2021, doi: 10.30536/j.ijreses.2021.v18.a3538.
[17] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 80, pp. 91–106, 2013, doi: 10.1016/j.isprsjprs.2013.03.006.
[18] X. Li and W. Zhou, "Dasymetric mapping of urban population in China based on radiance corrected DMSP-OLS nighttime light and land cover data," *Science of the Total Environment*, vol. 643, pp. 1248–1256, 2018, doi: 10.1016/j.scitotenv.2018.06.244.
[19] M. D. Su, M. C. Lin, H. I. Hsieh, B. W. Tsai, and C. H. Lin, "Multi-layer multi-class dasymetric mapping to estimate population distribution," *Science of the Total Environment*, vol. 408, no. 20, pp. 4807–4816, 2010, doi: 10.1016/j.scitotenv.2010.06.032.
[20] J. Mennis, "Dasymetric Apping for Estimating Population in Small Areas Dasymetric Mapping for Estimating Population in Small Areas," vol. 2, no. October, pp. 727–745, 2015, doi: https://doi.org/10.1111/j.1749-8198.2009.00220.x.
[21] L. Li and D. Lu, "Mapping population density distribution at multiple scales in Zhejiang Province using Landsat Thematic Mapper and census data," *Int J Remote Sens*, vol. 37, no. 18, pp. 4243–4260, 2016, doi: 10.1080/01431161.2016.1212422.
[22] Eurostat, "Methodological manual on territorial typologies. 2018 Edition.," *General and regional statistics*, no. Manuals and guidelines, p. 132, 2018, [Online]. Available: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-18-008
[23] Eurostat, "Applying the Degree of Urbanisation A NEW INTERNATIONAL MANUAL FOR DEFINING CITIES, TOWNS AND RURAL AREAS 2021 EDITION," 2021, [Online]. Available: http://ec.europa.eu/regional_policy/sources/docgener/work/2014_01_new_urban.pdf.
[24] Eurostat, *Applying the Degree of Urbanisation: A methodological manual to define cities, towns and rural areas for international comparisons 2020 edition*. 2021.
[25] K. Zhao, J. Kang, J. Jung, and G. Sohn, "Building extraction from satellite images using mask R-CNN with building boundary regularization," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2018-June, pp. 242–246, 2018, doi: 10.1109/CVPRW.2018.00045.
[26] W. Nurkarim and A. W. Wijayanto, "Building footprint extraction and counting on very high-resolution satellite imagery using object detection deep learning framework," *Earth Sci Inform*, no. 0123456789, 2022, doi: 10.1007/s12145-022-00895-4.
[27] S. Jung, K. Lee, and W. H. Lee, "Object-Based High-Rise Building Detection Using Morphological Building Index and Digital Map," *Remote Sens (Basel)*, vol. 14, no. 2, 2022, doi: 10.3390/rs14020330.
[28] Y. C. Putra and A. W. Wijayanto, "Automatic detection and counting of oil palm trees using remote sensing and object-based deep learning," *Remote Sens Appl*, vol. 29, no. December 2022, p. 100914, 2023, doi: 10.1016/j.rsase.2022.100914.
[29] K. Ayush, B. Uzkent, M. Burke, D. Lobell, and S. Ermon, "Efficient Poverty Mapping using Deep Reinforcement Learning," pp. 1–13, 2020, [Online]. Available: http://arxiv.org/abs/2006.04224
[30] Khairunnisah, A. W. Wijayanto, and S. Pramana, "Mapping Poverty Distribution of Urban Area using VIIRS Nighttime Light Satellite Imageries in D . I Yogyakarta , Indonesia," *Asian Journal of Business Environment*, vol. 13, no. 2, pp. 1–14, 2023, doi: 10.13106/ajbe.2023.vol13.no2.xx.
[31] S. R. Putri, A. W. Wijayanto, and S. Pramana, "Multi-source satellite imagery and point of interest data for poverty mapping in East Java, Indonesia: Machine learning and deep learning approaches," *Remote Sens Appl*, vol. 29, no. July 2022, p. 100889, 2023, doi: 10.1016/j.rsase.2022.100889.
[32] A. Muhajir Nasir, "Statistik Pendidikan," *Media Akademi*, 2016, doi: 10.31227/osf.io/judwx.

[33] BPS-Statistics Indonesia, "Proporsi rumah tangga dengan akses terhadap pelayanan dasar menurut provinsi (Persen), 2020-2022." Accessed: Aug. 18, 2023. [Online]. Available: https://www.bps.go.id/indicator/12/2016/1/proporsi-rumah-tangga-dengan-akses-terhadap-pelayanan-dasar-menurut-provinsi.html
[34] K. Kemendikbud, "Peraturan Menteri Pendidikan dan Kebudayaan (Permendikbud) Nomor 1 Tahun 2021 tentang Penerimaan Peserta Didik Baru Jenjang TK, SD, SMP, SMA dan SMK," *Permendikbud*, pp. 1–25, 2021, [Online]. Available: https://lpmpkaltara.kemdikbud.go.id/wp-content/uploads/2021/01/Permendikbud-Nomor-1-Tahun-2021.pdf
[35] BPS-Statistics Indonesia, "Jumlah Penduduk Menurut Provinsi di Indonesia (Ribu Jiwa), 2020-2022," BPS-Statistics Indonesia. Accessed: Aug. 18, 2023. [Online]. Available: https://sulut.bps.go.id/indicator/12/958/1/jumlah-penduduk-menurut-provinsi-di-indonesia.html
[36] F. P. Boscoe, K. A. Henry, and M. S. Zdeb, "A Nationwide Comparison of Driving Distance Versus Straight- Line Distance to Hospitals," *Prof Geogr*, vol. 2012 April, no. 64(2), 2012, doi: 10.1080/00330124.2011.583586.A.
[37] W. Luo, "Using a GIS-based floating catchment method to assess areas with shortage of physicians," *Health Place*, vol. 10, no. 1, pp. 1–11, 2004, doi: 10.1016/S1353-8292(02)00067-9.
[38] W. Luo and F. Wang, "Measures of spatial accessibility to health care in a GIS environment: Synthesis and a case study in the Chicago region," *Environ Plann B Plann Des*, vol. 30, no. 6, pp. 865–884, 2003, doi: 10.1068/b29120.
[39] M. R. McGrail and J. S. Humphreys, "Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method," *Applied Geography*, vol. 29, no. 4, pp. 533–541, 2009, doi: 10.1016/j.apgeog.2008.12.003.
[40] World Bank, "SDGs 9.1.1 Metadata," 2020. [Online]. Available: https://unstats.un.org /sdgs/metadata/files/Metadata-09-01-01.pdf
[41] M. Mikou, J. Rozenberg, E. Koks, C. Fox, and T. Peralta Quiros, "Assessing Rural Accessibility and Rural Roads Investment Needs Using Open Source Data," *Transactions in GIS*, vol. 25, no. 2, 2019, doi: 10.1596/1813-9450-8746.
[42] The World Bank Group, *Measuring Rural Access Using new technologies*. 2016. [Online]. Available: http://documents.worldbank.org/curated/en/367391472117815229/pdf/107996-REVISED-PUBLIC-MeasuringRuralAccessweb.pdf
[43] C. M. Ilie, M. A. Brovelli, and S. Coetzee, "Monitoring sdg 9 with global open data and open software – A case study from rural Tanzania," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 42, no. 2/W13, pp. 1551–1558, 2019, doi: 10.5194/isprs-archives-XLII-2-W13-1551-2019.
[44] UN-Habitat, "Metadata on SDGs Indicator 11.1.1 Indicator category: Tier I," no. March, pp. 1–12, 2018, [Online]. Available: https://unhabitat.org/sites/default/files/2020/06/ metadata_on_sdg_indicator_11.1.1.pdf
[45] UN-Habitat, "SDGs 11.1.1 Metadata," 2021. [Online]. Available: https://unstats.un.org/sdgs/metadata/files/Metadata-11-01-01.pdf
[46] BPS-Statistics Indonesia, "Persentase Rumah Tangga yang Memiliki Akses Terhadap Hunian Yang Layak Dan Terjangkau Menurut Daerah Tempat Tinggal (Persen)." Accessed: Aug. 18, 2023. [Online]. Available: https://www.bps.go.id/indikator/indikator/view_data /0000/data/1242/sdgs_11/3
[47] S. Yousefi, H. R. Pourghasemi, S. N. Emami, S. Pouyan, S. Eskandari, and J. P. Tiefenbacher, "A machine learning framework for multi-hazards modeling and mapping in a mountainous area," *Sci Rep*, vol. 10, no. 1, pp. 1–14, 2020, doi: 10.1038/s41598-020-69233-2.
[48] M. Sholeh, P. Pranoto, S. Budiastuti, and Sutarno Sutarno, "Analysis of Citarum River pollution indicator using chemical, physical, and bacteriological methods," vol. 020068, no. December 2019, 2023.

[49] M. R. Djuwita, D. M. Hartono, S. S. Mursidik, and T. E. B. Soesilo, "Pollution load allocation on water pollution control in the citarum river," *Journal of Engineering and Technological Sciences*, vol. 53, no. 1, pp. 1–15, 2021, doi: 10.5614/j.eng.technol.sci.2021.53.1.12.
[50] J. Zhang, T. Zou, and Y. Lai, "Novel method for industrial sewage outfall detection: Water pollution monitoring based on web crawler and remote sensing interpretation techniques," *J Clean Prod*, vol. 312, no. May, p. 127640, 2021, doi: 10.1016/j.jclepro.2021.127640.
[51] L. Wei, Z. Yuan, Y. Zhong, L. Yang, X. Hu, and Y. Zhang, "An Improved Gradient Boosting Regression Tree Estimation Model for Soil Heavy Metal (Arsenic) Pollution Monitoring Using Hyperspectral Remote Sensing," *Applied Sciences (Switzerland)*, vol. 9, no. 9, 2019, doi: 10.3390/app9091943.
[52] V. Tabunschik, R. Gorbunov, and T. Gorbunova, "Unveiling Air Pollution in Crimean Mountain Rivers: Analysis of Sentinel-5 Satellite Images Using Google Earth Engine (GEE)," *Remote Sens (Basel)*, vol. 15, no. 13, p. 3364, 2023, doi: 10.3390/rs15133364.
[53] UN-Habitat, "SDGs 11.2.1 Metadata," 2021. [Online]. Available: https://unstats.un.org/sdgs/metadata/files/Metadata-01-04-01.pdf
[54] L. Reviewed, *Metadata on SDGs Indicator 11.2.1 Indicator category: Tier II*, no. March. 2018.
[55] BPS-Statistics Indonesia, "Laju Pertumbuhan Penduduk (Persen), 2021-2023." [Online]. Available: https://www.bps.go.id/indicator/12/1976/1/laju-pertumbuhan-penduduk.html
[56] D. Mwaniki, "Regional Training Workshop on Human Settlement Indicators," *United Nation*, no. March, 2018, [Online]. Available: https://www.unescap.org/sites/default/files/6.Working_definition_of_a_city_for_SDG11_UN-Habitat_Wshop_26-29Mar2018.pdf
[57] UN-Habitat, "MODULE 3 LAND USE EFFICIENCY Monitoring and Reporting the SDGs | LAND USE EFFICIENCY," 2018, [Online]. Available: https://unhabitat.org/sites/default/files/2021/08/indicator_11.3.1_training_module_land_use_efficiency.pdf
[58] UN-Habitat, "SDGs 11.3.1 Metadata," 2021, [Online]. Available: https://unstats.un.org/sdgs/metadata/files/Metadata-11-03-01.pdf
[59] ESA, *Sentinel-2 User Handbook*, no. 1. 2015. doi: 10.1021/ie51400a018.
[60] UN-Habitat, "Metadata on SDGs Indicator 11.7.1 Indicator category: Tier II," pp. 1–12, [Online]. Available: https://unhabitat.org/sites/default/files/2020/11/metadata_on_sdg_indicator_11.7.1_02-2020_1.pdf
[61] UN-Habitat, "SDGs 11.7.1 Metadata," 2021, [Online]. Available: https://unstats.un.org/sdgs/metadata/files/Metadata-11-07-01.pdf
[62] Z. Zhang, Q. Liu, and Y. Wang, "Road Extraction by Deep Residual U-Net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749–753, 2018, doi: 10.1109/LGRS.2018.2802944.
[63] Y. Wei, Z. Wang, and M. Xu, "Road Structure Refined CNN for Road Extraction in Aerial Image," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 709–713, 2017, doi: 10.1109/LGRS.2017.2672734.

Acknowledgments
I am grateful to acknowledge the support and help from my colleagues in reviewing and providing valuable comments. I also thank my professor and parents for their words of encouragement in writing this paper. All persons and institutes who kindly made their data available to make this review happen are acknowledged.