<Agung Andiojaya>

Utilizing Nighttime-light (NTL) Data to Measure Indonesia's Gross Domestic Regional Product (GDRP)

Agung Andiojaya[1]; R.R. Nefriana[1]; Riana Rizka[1]; Titi Kanti Lestari[1]

[1]Badan Pusat Statistik

Abstract:
Gross Domestic Regional Product (GDRP) is one of the widely used macroeconomic indicators to measure the achievement of socio-economic development. To obtain a comprehensive GDRP calculation requires complete data availability on the national level to the smallest administrative area level. The challenge that is often faced by institutions that carry out GDRP calculations is that the smaller the administrative level, the more complex and difficult the effort to obtain the data. Responding to this challenge, this paper tries to explore the use of Nighttime-light (NTL) data to be utilized as an alternative data in estimating the amount of GDRP and its change. The analysis results based on panel data model with random effect show that on average, NTL data is statistically significant in describing the GDRP in various Indonesian provinces. However, during the pandemic (2020), NTL was not significant in describing the GDRP and showed movements that were not in line with the GDRP. The different results of the analysis in 2020 are likely due to Indonesian economics, which tends to be dominated by the informal sector rather than the formal sector.

Keywords:
Indonesia Nighttime-light, Night Light Data, Day/Night Band VIIRS, panel data model, random effect.

1. Introduction:

Gross Domestic Regional Product (GDRP) is one of the widely used macroeconomic indicators to measure socio-economic development . This figure is the sum of value-added produced by all production units in a region or a total value of final goods and services (net) made by all economic units. To obtain a comprehensive GDRP calculation, Statistics Indonesia requires complete data on the national level to a minor administrative area level.

The difficulty often faced by institutions that carry out GDRP calculations is that the smaller the administrative level, the more complex and challenging the effort to obtain the data. For instance, the low response rates persist over time for both household and business types of respondents [1]–[4]. As for local governments as data providers, it is difficult to elicit the data for goods purchased by the government because the information is not documented in detail and makes allocating national accounts burdensome. Meanwhile, the taxation data commonly obtained by other National Statistical Offices (NSOs) from each country's taxation office is not yet sufficiently available to the appropriate extent in Indonesia [5], [6]. Ipso facto, Statistics Indonesia reached a cooperation agreement with the Directorate General of Taxation in late 2019 via Statistics Indonesia's Statistical Business Register secretariat's approach after a four-year negotiation process. Nonetheless, due to a taxation law that protects individual data, the

detailed data needed, such as annual revenue or output, cannot be drawn just yet, even for a governmental agency like Statistics Indonesia. In lieu of this administrative data and other regular surveys, National Accounts conduct ad-hoc surveys to fill the data gap. Unfortunately, due to resource constraints (mainly supported by the non-availability of Subject Matters Areas responsible for covering some economic industries), the sample sizes are often insufficient and drawn purposively. The non-availability of these Subject Matter Areas also insists that Statistics Indonesia estimates some required numbers, such as the numbers for financial, real estate, business, and other services economic industries. These data gap challenges are getting more salient during the pandemic in which traditional data collection approaches, especially surveys, are not viable and arduous.

At the same time as the challenges as mentioned earlier, an opportunity emerges. Some success stories show the potency of using satellite imagery data as a proxy for some economic indicators. In 2012 Henderson, Storeygard, and Weil reported that they fruitfully utilized observed night-light as an indicator of income across countries and sub- and supranational regions [7]. Prakash, Shukla, Bhowmick, and Beyer in 2019 similarly found that night-light data correlate robustly with GDP and other critical macroeconomic indicators such as industrial production and credit growth at a national level in India [8]. The night lights in India were also strongly correlated with their gross state domestic product (GSDP). Also, only recently, in 2020 Adhikari and Dhital, by exploiting night-time lights captured by the U.S. Air Force satellites combined with fiscal, political, and administrative decentralization databases, revealed that decentralization hindered regional convergence between first and second subnational regions within a country [9].

With these empirical success shreds of evidence in mind, we want to investigate if there is a positive cross-section correlation between the levels of nightlights luminous intensities, measured by satellites from outer space, and levels of GDRP in Indonesia. This paper explores the use of Nighttime-light (NTL) data as alternative data in estimating the amount of GDRP and its change. By approximating GDRP by night lights data, we hope to alleviate the data collection challenges mentioned earlier.

2. Methodology:

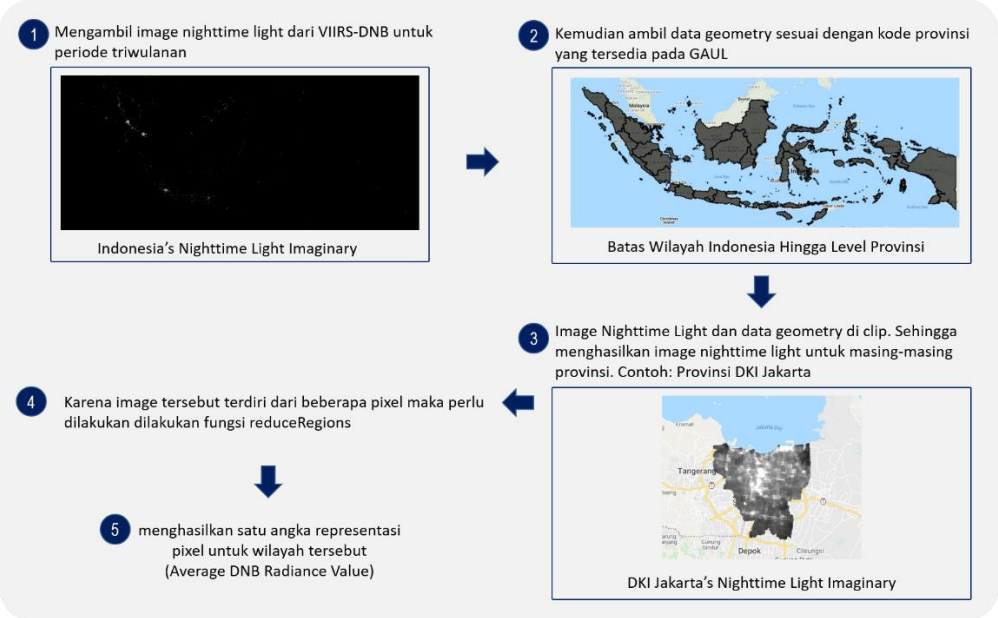
Study Area

This study analyzes data of 33 provinces in Indonesia quarterly starting from Quarter-1 2018 to Quarter-4 2020.

Data

The data used in this study include Indonesia's Gross Domestic Regional Product (GDRP) at Constant Price by Province sourced from BPS-Statistics Indonesia. Then, for nighttime-light (NTL) satellite data employ average day-night-band (DNB) values of Monthly Cloud-free DNB Composite NPP-VIIRS data sourced from https://eogdata.mines.edu/products/vnl/. To obtain spatial data for each province in Indonesia, this study uses Indonesian geometric data in the form of Global Administrative Unit Layers (GAUL), First-Level Administrative Units sourced from FAO UN.

In generating NTL data per province, this study utilizes Google Earth Engine (GEE). GEE is a cloud-based platform for planetary-scale geospatial analysis, which allows the process of a variety of geographical data at scale and handles large geographical datasets. GEE also provides access to numerous remotely sensed datasets and derived products, including VIIRS DNB and Global Administrative Unit Layers.



Method

By considering that each province's GDRP data has varying values and trends, this study will investigate the relationship between nightlight data and GDRP using a panel regression model with random effect. In general, the random effect model can be expressed in equation (1) [10].

$$Log(GDRP_{it}) = \alpha + \beta\, Log(NTL_{it}) + w_{it} \quad (1)$$

where

$$w_{it} = \varepsilon_{it} + u_i \quad (2)$$

Note:

$GDRP_{it}$ = GDRP at Constant Price of i^{th} province, t^{th} quarter
α = Join intercept
β = Coefficient of regression or slope
NTL_{it} = Nighttime-light Data of i^{th} province, t^{th} quarter
u_i = error term of i^{th} province
ε_{it} = error term of i^{th} province, t^{th} quarter
i = 1 to N
t = 1 to T

Apart from the investigation of NTL and GDRP model, this paper is also exploring the relationship between NTL data and GRDP before and during a pandemic, this study generates a binary dummy variable in the model to mark the quarter before and after the pandemic. This dummy variable marks the 2018 Q1-2019 Q4 period before the covid-19 pandemic and makes this period the baseline for the dummy variable. Thus, the model built in this paper becomes:

$$Log(GDRP_{it}) = \alpha + \beta\, Log(NTL_{it}) + DummyPandemic_t + DummyPandemic_t * Log(NTL_{it}) + w_{it} \quad (1)$$

3. Result:

Table 1 shows three essential findings. First, we get an optimistic estimator of Log NTL and statistical significance in the model. These results indicate that NTL can statistically be used to estimate changes in GDRP. A positive estimator value means that when there is an increase in light intensity, it will increase in GDRP and vice versa.

Second, although the p-value of the DummyPandemic variable is not statistically significant, the negative value of the estimator can be a sign that the GDRP value during the pandemic tends to be lower than before the pandemic. This situation confirms that there has been a decline in economic growth caused by the Covid-19 pandemic that hit Indonesia throughout 2020.

The third finding, which is shown from the value of the estimator of the interaction results of the LogNTL and DummyPandemic variables, confirms the second finding. The estimator of the LogNTL and DummyPandemic variables is -0.0124, which means that the relationship between nightlight growth and GDP growth during the pandemic was 0.0124 points lower than before the pandemic.

Table 1. A panel regression model with the random effect result

	Estimate	Std.Error	Pr(>|z|)
Intercept	17.5338	0.1548	< 2.2e-16 ***
Log_NTL	0.0201	0.0032	3.379e-10 ***
DummyPandemic	-0.0030	0.0054	0.5830503
Log_NTL * DummyPandemic	-0.0124	0.0036	0.00057 ***
R^2			0.1083
Adjusted R^2			0.1015
Chi Square on 3 DoF			47.631
p-value			2.5514e-10
Note on Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1			

Figure 1 presents the value of real GRDP, while Figure 2 shows the predicted value of GDRP calculated based on the model. The two figures show that the GDRP values in some areas have similar movement with the observed GDRP values (such as West Sumatra, Riau, Jambi, Bengkulu, and North Sulawesi). As for other provinces, the predicted value of GDRP generated by the model does not reveal a promising result.

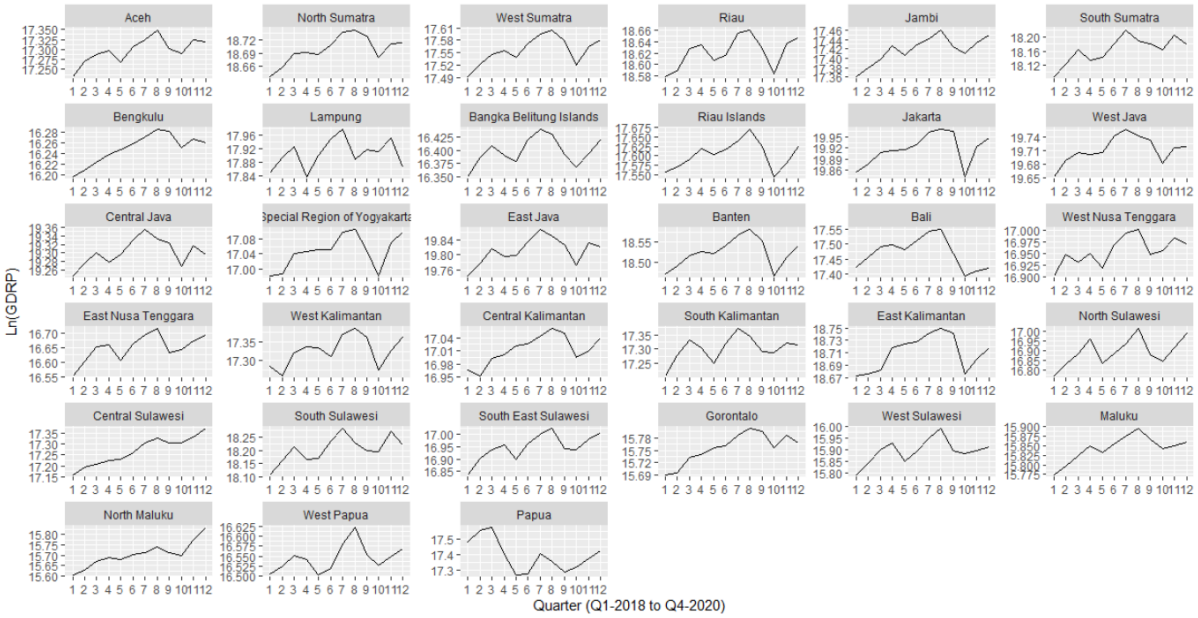


Figure 1. Trend line of observed GDRP by Province from Q1-2018 to Q4-2020

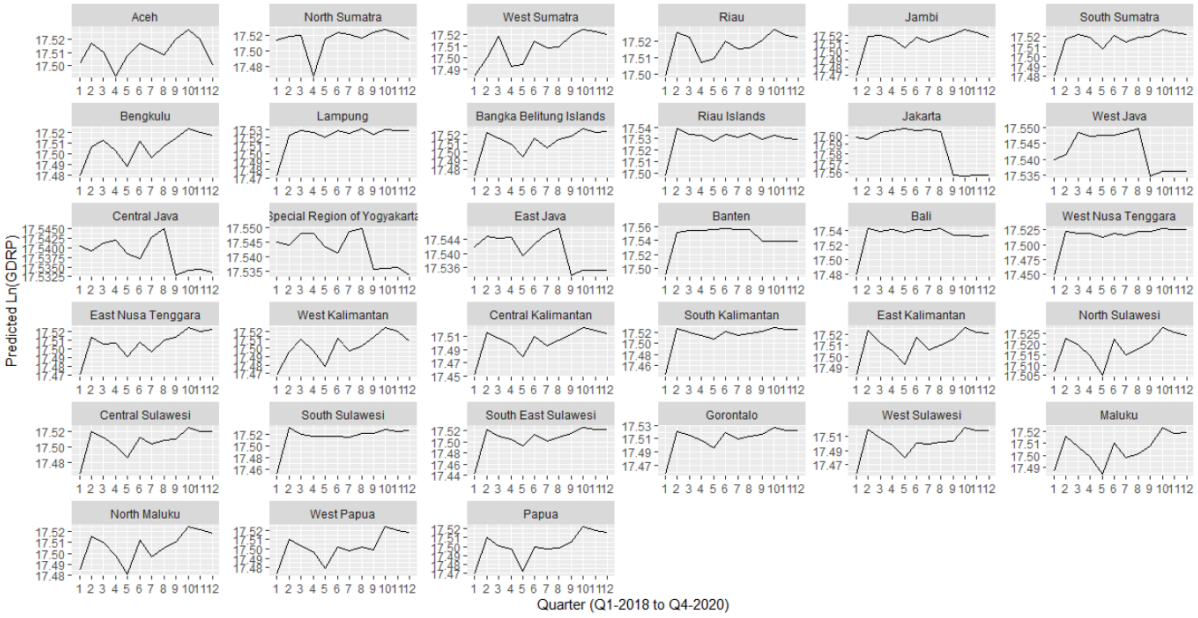


Figure 2. Trend line of predicted GDRP by Province from Q1-2018 to Q4-2020

4. Discussion and Conclusion:

In the model generated in this study, the NTL variable is statistically significant to explain variations in the value of GRDP. In several provinces, the model can provide a good predicted value of GDRP whose trend has a similar movement as the observed value of GRDP. However, in the other provinces, the model's predicted value is prone to move in a different direction from the observed value of GRDP.

Henderson et al. (2012) stated that there is a positive and robust relationship between NTL and GRDP. Therefore, in future research, it is necessary to consider other data as the control variables in the model. Despite these shortcomings, NTL can still be a promising alternative data to measure a region's economy.

References:

[1] BPS, “Laporan Kinerja BPS Tahun 2016,” Jakarta, 2016.
[2] BPS, “Laporan Kinerja BPS Tahun 2017,” 2018.
[3] BPS, “Laporan Kinerja BPS Tahun 2018,” Jakarta, 2019.
[4] BPS, “Laporan Kinerja BPS Tahun 2020,” Jakarta, 2021.
[5] R. Luisa, C. Thompson, and J. Jones, “A Statistical Business Register Spine as a New Approach to Support Data Integration and Firm-level Data Linking: An ABS Perspective,” *Stat. J. IAOS*, vol. 1, pp. 767–774, 2020.
[6] F. Statistics, “Use of Register and Administrative Data Sources for Statistical Purposes: Best Practices of Statistics Finland,” Helsinki, 2004.
[7] J. V. Henderson, A. Storeygard, and D. N. Weil, “Measuring economic growth from outer space,” *Am. Econ. Rev.*, vol. 102, no. 2, pp. 994–1028, 2012, doi: 10.1257/aer.102.2.994.
[8] A. Prakash, A. K. Shukla, C. Bhowmick, R. Carl, and M. Beyer, “Night-time Luminosity : Does it Brighten Understanding of Economic Activity in India ?,” *Reserv. Bank India Occas. Pap.*, vol. 40, no. 1, pp. 1–24, 2019.
[9] B. Adhikari and S. Dhital, “Decentralization and Regional Convergence: Evidence from Night-Time Lights Data,” Illinois, 2020.
[10] R. C. Hill, W. E. Griffiths, and G. C. Lim, *Principles of Econometrics*, 5th Ed. United States: Wiley, 2018.