IIAI Open Conference Publication Series
IIAI Letters on Business and Decision Science
Vol. 001, LBDS030
DOI: https://doi.org/10.52731/lbds.001.030

Estimates of CO_2 Concentrations Based on Satellite Data and Their Usefulness in Developing a More Accurate Picture of Economic Activity

Yoshiyuki Suimon [*] [†], Hiroto Tanabe [†], Kiyoshi Izumi [*]

Abstract

In this research, we propose a method of using data on CO_2 concentrations based on observational data from the Greenhouse Gases Observing Satellite (GOSAT, also known as Ibuki) to arrive at an accurate picture of macroeconomic activity with minimal lag. In doing so, we build on our conventional approach of using fast-breaking consumption-related indicators, putting together an estimation model that includes—as a newly added feature—the data on CO_2 concentrations derived from satellite observations, which are less susceptible to the sort of sampling bias that that can arise when drawing data from individual companies or other such narrow sources. What we find is that this updated model yields estimates of private consumption that are more accurate than what we have been able to produce with our conventional model based on consumption-related indicators, and that the updated model yields estimates of consumption that are reliably accurate even under the unusual circumstances created by the COVID-19 pandemic since 2020.

Keywords: Macroeconomic activity, CO_2 emissions, Satellites observation data, GOSAT

1 Introduction

Faced with the growing threats posed by climate change, countries around the world have come forth with targets for reductions in their emissions of carbon dioxide (CO_2) and other greenhouse gases [1]. This includes Japan, where a variety of efforts to address greenhouse gas emissions are being undertaken [2]. In the research we present here—initiated in response to the heightened attention being given to CO_2 emissions—we analyze the relationship between macroeconomic activity and the volume of CO_2 emissions, and then propose a method of estimating macroeconomic activity using timely information on CO_2 concentrations measured by means of observational data from the Greenhouse Gases Observing Satellite (GOSAT, also known as Ibuki) [3] [4]. Our aim is to improve the accuracy of estimates by augmenting conventional economic data with granular satellite data used as an additional feature in an estimation model.

In the field of economic analysis, a great deal of work has been done recently on bringing data of all sorts to bear in developing more accurate models of economic activity. These newly studied sources of data fall under the collective moniker of alternative data [5], and the data that has been analyzed is quite far-ranging.

[*] School of Engineering, The University of Tokyo, Japan
[†] Nomura Securities Co., Ltd., Japan

For example, remote sensing technology has featured in a number of pieces of research. One group of researchers has presented the idea of using data on nighttime light emissions collected by satellites to develop a better understanding of the state of China's economy [6]. In the same vein, another colleague of mine and I proposed measuring production activity in Japan's manufacturing sector based on an understanding of nighttime production activity gained from an analysis of what satellite photographs show about nighttime light emissions from factories and other facilities [7]. In multiple other studies as well, data on nighttime light emissions collected using remote sensing technologies has been put to work in a number of ways to measure various aspects of economic activity [8] [9] [10] [11].

2 The Use of Satellite Observational Data

In the research we present here, we proceed with an analysis of observational data recorded by the Greenhouse Gases Observing Satellite (GOSAT, also known as Ibuki), which was put into service chiefly to measure concentrations of carbon dioxide (CO_2) and methane (CH4), two major greenhouse gases. The GOSAT project is a joint effort of the Japan Aerospace Exploration Agency (JAXA), Japan's Ministry of the Environment (MOE), and the National Institute for Environmental Studies (NIES) [3] [4] [12]. GOSAT circles the earth about once every 100 minutes, orbiting at a high altitude of 666km and returning to the same point in space every three days. It takes some 56,000 measurements over those three days, covering the Earth's entirety.

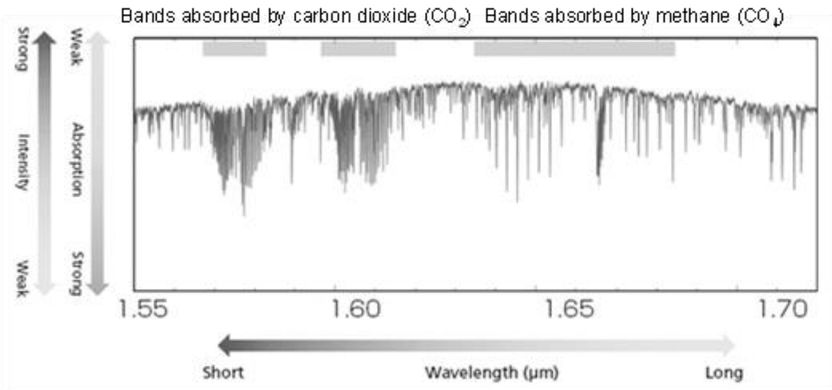

Figure 1: The observed solar spectrum and absorbed wavelengths [3]

GOSAT measures sunlight reflected by the Earth's surface, and in doing so can produce estimates of atmospheric concentrations of CO_2, making use of the fact that CO_2 absorbs light in particular wavelength bands. The CO_2 concentration readings thus measured for each unit of mesh can then be selected for analysis and compiled so as to produce rough estimates of CO_2 concentrations for a particular country or region of the world. Figure 2 compares estimated CO_2 concentrations over Japan based on the GOSAT data with the CO_2 emissions data for Japan reported by the World Bank. They track one another fairly closely.

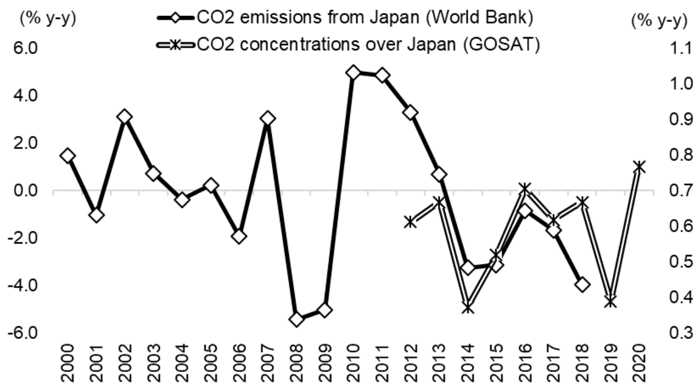

Figure 2: CO_2 emissions vs. estimated CO_2 concentrations

As explained above, GOSAT captures data over the entire globe once every three days, so the data it makes available for analysis is quite fresh. Given the apparent usefulness of the CO_2 concentration estimates yielded by the GOSAT data, below we put these estimates to work in an analysis of Japan's macroeconomic situation. Figure 6 shows the coefficients of correlation between estimated CO_2 concentrations based on the GOSAT measurements and various economic indicators related to GDP (using YoY changes in each indicator for each quarter over the period 2010 through 2019). The CO_2 concentration estimates generally correlate more highly with the real GDP-related indicators than they do with the nominal indicators. The upshot is that CO_2 concentrations are linked with real economic activity (that is, measurements of economic activity from which the effect of inflation has been stripped out), which is precisely what one would expect to see.

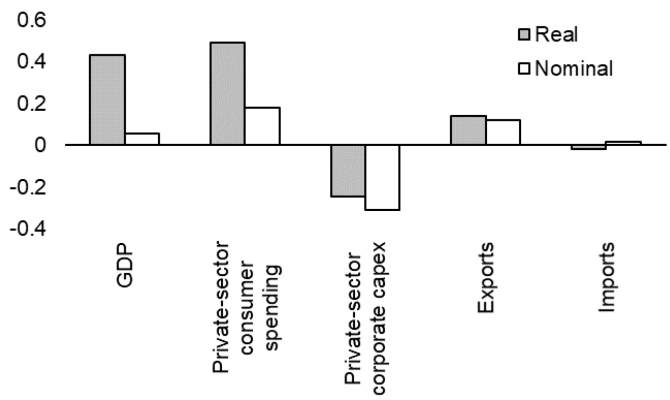

Figure 3: Coefficients of correlation between estimated CO_2 concentrations and GDP statistics

As Figure 3 makes apparent, real private-sector consumer spending is the component of GDP that shows the closest correlation with the CO_2 concentration estimates yielded by the GOSAT measurements. Here as well, it makes intuitive sense that the widely varied economic activity related to private-sector consumption on both the demand side and the supply side would correlate fairly closely with CO_2 concentrations, given the massive scale of such activity.

3 Estimation Model of Macrocosmic Activity

In this section, we build on what we have demonstrated thus far and assemble an estimation model of economic activity in Japan based on machine learning, making use of the estimated CO_2 concentrations yielded by the GOSAT data. The estimated CO_2 concentrations for Japan discussed in the simple model in the preceding section were calculated as the average of the readings for all of the twelve blocks shown in Figure 4 below. (In our research, we use a Level-3 GOSAT product that shows CO_2 column-averaged mixing ratio data projected on a global map, from which we take the block-specific data relevant to our study.)

Here we enhance the feature-richness of our estimation model of economic activity by using the more granular information on CO_2 concentrations taken from each of the 12 blocks separately. By way of illustration, Figure 5 plots YoY changes in estimated CO_2 concentrations at six different locations in Japan. The numbers given in the graph legend indicate the latitude and longitude of each unit of mesh.

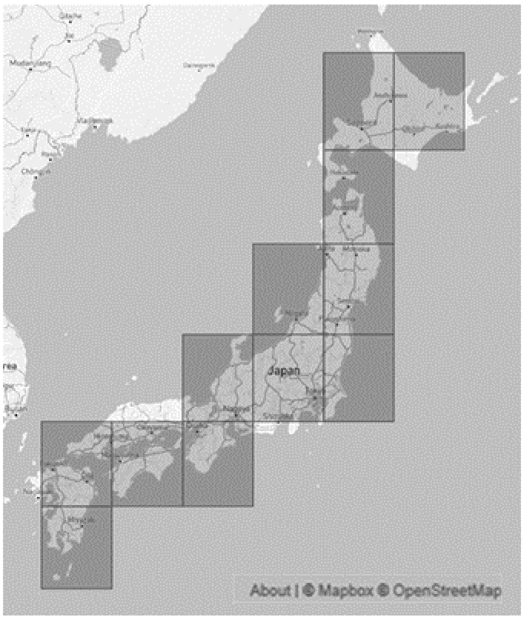

Figure 4: Units of mesh covering Japan selected for analysis

While it is apparent that patterns in changes in CO_2 concentrations vary from one region to another, the movements are broadly similar and broadly correlated. We therefore reduce the dimensionality of these variables by running a principal components analysis (PCA) on the data on YoY changes in CO_2 concentrations in the 12 blocks selected for study. Figure 6 shows the contributions made by each of the principal component factors. What this reveals is that the first three principal component factors together account for more than 90% of the variance in need of explaining.

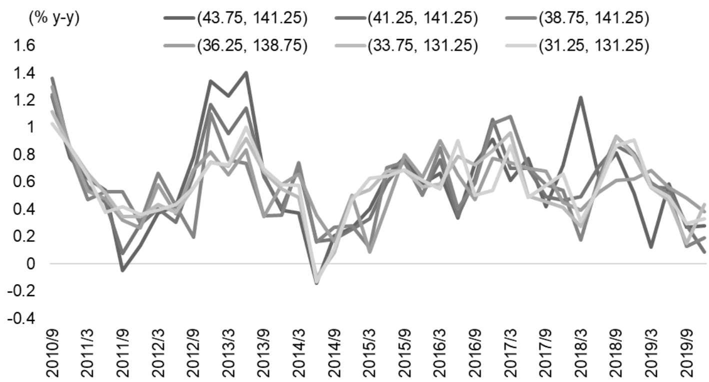

Figure 5: Estimated CO_2 concentrations in selected units of mesh covering parts of Japan

We then use this information to assemble an estimation model of real private-sector consumption. In our analysis, we use quarter-by-quarter data for the third quarter of 2010 through the second quarter of 2021, where the explained variable is the YoY change in real private-sector consumption. As the explanatory variable, we use information on the first three principal component factors for the quarter in question and the preceding quarter, as extracted from the data on YoY changes in estimated CO_2 concentrations for each block. We call this explanatory variable 1.

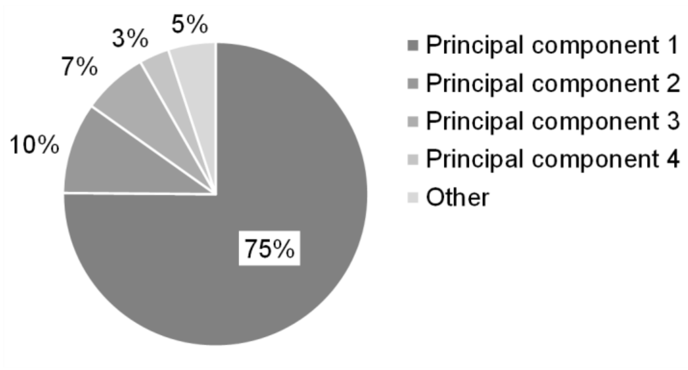

Figure 6: Contributions of principal component factors

In parallel, we look at credit card transaction data (the JCB Consumption NOW index), department store sales (Japan Department Stores Association statistics), and new automobile sales volumes (Japan Automobile Dealers Association statistics), all of which are indicators of consumption that are reported at a minimal lag. Each of these sets of numbers is released at the beginning of each month, offering early readings of consumption patterns in the month just ended. Thus, these data releases give readers a picture of consumer spending behavior about a month before the release of the main government statistics on consumption, such as the Ministry of Internal Affairs and Communications' Family Income and Expenditure Survey and the Ministry of Economy, Trade and Industry's Current Survey of Commerce.

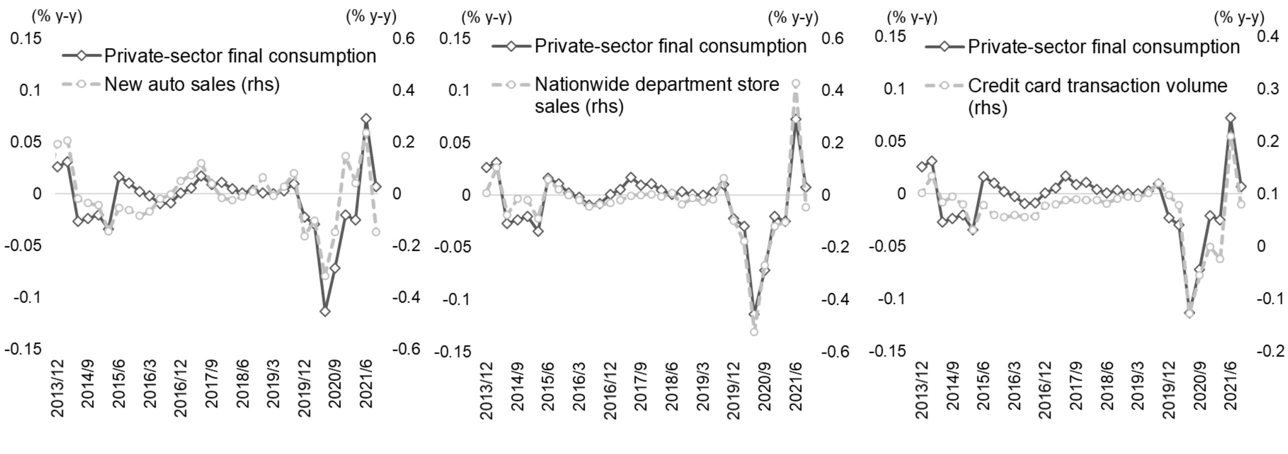

Figure 7: Consumption-related indicators vs. GDP-based private-sector final consumption

In an existing methodology of ours [13], we have shown that estimates of consumption made using these data sets are indeed useful in nowcasting consumption. We therefore deploy these data sources in our current analysis as well, giving these data sets the collective label of explanatory variable 2. Note that of these three data sources, the JCB Consumption NOW index has only been publicly available since April 2015, so for the period prior to that we instead use data on credit card transaction volumes from the Ministry of Economy, Trade and Industry's Survey of Selected Service Industries.

We then conduct an analysis of this consumption-related data (explanatory variable 2) together with the data based on satellite observations (explanatory variable 1) to produce a composite variable that we call explanatory variable 3. We can then test the explanatory power of these as features of a CO_2 concentration model for each of the 12 blocks.

Explanatory variable 1:

Satellite data series (the first three principal components of the CO_2 concentration estimates, for both the quarter in question and the preceding quarter)

Explanatory variable 2:

Consumption-related data series (credit card transaction data, department store sales, and new automobile sales, all for the quarter in question)

Explanatory variable 3:

Composite data series (a combination of explanatory variable 1 and explanatory variable 2)

We use a total of four mathematical models of real private-sector consumption. In addition to a standard linear regression model, we also employ a ridge regression, a lasso regression, and an elastic net regression.

In this research, we train our models using out-of-sample data for either the preceding three years or the preceding four years (excluding the data for the point in time for which we are attempting to produce an estimate). We then plug each of the explanatory variables at each point in time into the trained model to produce an estimate of real private-sector consumption for that point in time.

Table 1: Compared accuracy of estimates of real private-sector consumption output by three models (RMSE)

Pre-pandemic (through 2019 Q4)		Linear regression	Ridge regression	Lasso regression	ElasticNet
3-year training	Explanatory variable 1	3.64	3.50	2.33	2.44
	Explanatory variable 2	2.31	2.25	1.76	1.86
	Explanatory variable 3	2.82	1.46	1.57	1.40
4-year training	Explanatory variable 1	2.53	2.48	1.74	1.87
	Explanatory variable 2	1.13	1.12	1.00	1.01
	Explanatory variable 3	1.15	1.02	0.97	0.95

Post-pandemic (2020 Q1 onwards)		Linear regression	Ridge regression	Lasso regression	ElasticNet
3-year training	Explanatory variable 1	5.75	5.78	6.50	6.17
	Explanatory variable 2	4.25	3.07	1.93	2.12
	Explanatory variable 3	4.50	2.12	2.31	1.96
4-year training	Explanatory variable 1	6.60	6.61	6.94	6.81
	Explanatory variable 2	1.46	1.40	1.50	1.18
	Explanatory variable 3	1.28	0.96	1.43	1.02

Table 1 lays out the root-mean-square error (RMSE) between our estimates of real private consumption (YoY change, %) and the officially reported measurements. Given the tremendous impact that the COVID-19 pandemic has had on economic activity since early 2020, we break these measurements down into a pre-pandemic set (up through 2019 Q4) and a post-pandemic set (2020 Q1 onwards). What we find is that while estimates of consumption trends based on the satellite data series alone (explanatory variable 1) are less accurate than estimates derived from our conventional approach of using consumption-related data (explanatory variable 2), it is nevertheless the case that the satellite data does allow one to estimate real private consumption with some accuracy despite the fact that these estimates make no use whatsoever of any data directly related to consumption. In addition, we find that using our newly proposed method of combining the satellite data series with the consumption related data series to make a composite data series (explanatory variable 3) yields estimates of real private consumption that are more accurate than what we arrive at with our existing approach (explanatory variable 2). This remains the case regardless of which regularized regression model we employ (ridge regression, lasso regression, or elastic net regression). Figure 8 depicts the degrees of improvement in accuracy (RMSE) achieved through these means. Adding the satellite observational data on CO_2 concentrations to the mix of information used in estimating real private consumption turns out to yield an especially impressive improvement in accuracy when applied to the post-pandemic data set. Under pandemic conditions, department store sales and other conventional consumption-related data series are heavily impacted by such unusual states of affairs as voluntary suspensions of business during declared states of emergency. From a feature engineering standpoint, the approach we propose here (explanatory variable 3) improves our estimation model through the inclusion of data on CO_2 concentrations derived from satellite observations that are not subject to the sampling bias that shows up in data collected from individual corporations. The regularized regression models we have employed in this research yield improvements in the estimation modeling of consumption trends over our conventional approach [13], and serve as reliable ways of estimating consumption trends even under the unusual conditions that have arisen during the pandemic.

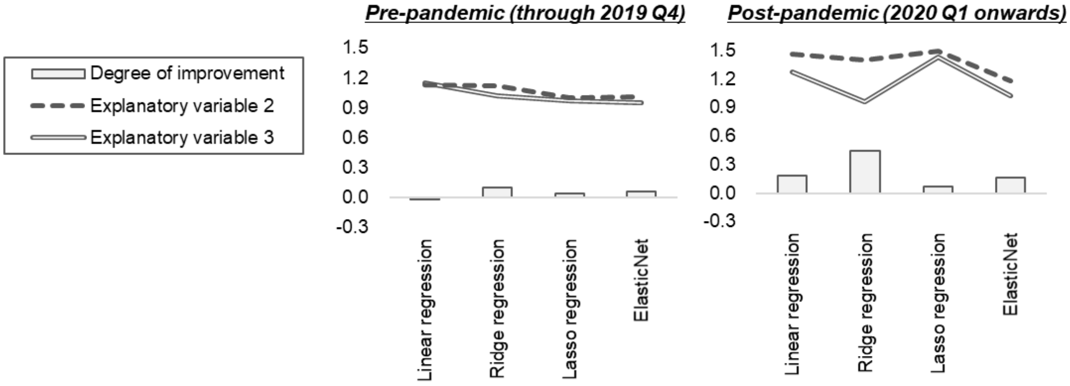

Figure 8: Degree of improvement in the accuracy of estimates (RMSE)
for each type of regression model

Figure 9 plots the estimates output by our model alongside the officially reported data. What this makes clear is that these estimates reliably predict YoY changes in real private consumption even in the midst of the pandemic. The more conventional measures we have historically used to estimate consumption, such as the new automobile sales volumes and department store sales shown in Figure 7, have been subject to large up-and-down swings during the pandemic, and they were also shaken up by rush demand ahead of the October 2019 consumption tax hike and the subsequent pullback in demand. We believe these large up-and-down swings arise because of a data sample that is biased towards data on purchases of durable goods (as exemplified by automobile purchases) and discretionary spending (which shows up strongly in department store sales). In contrast, the results presented in Figure 9 show that a more feature-rich model that incorporates satellite observational data on CO_2 concentrations yields estimates of changes in private consumption that are more reliably accurate, despite the fact that the training was performed using out-of-sample data, as discussed above. We therefore conclude that data on CO_2 concentrations derived from satellite observations are a useful tool that can help observers develop an accurate picture of economic activity.

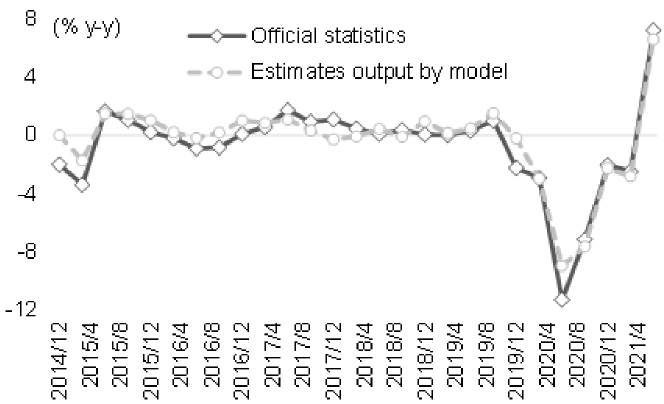

Figure 9: YoY change in real private consumption: official statistics vs. estimates output
by model (Explanatory variable 3, four-year training)

4 Conclusion

In the research we have presented here, we have investigated the relationship between macroeconomic activity and the volume of CO_2 emissions (partly as a response to the heightened attention that CO_2 emissions have been receiving), and we have proposed a method of using data on CO_2 concentrations based on satellite observations to arrive at an accurate picture of macroeconomic activity with minimal lag. In doing so, we have built on our conventional approach of using fast-breaking consumption-related indicators, putting together an estimation model that includes—as a newly added feature—data on CO_2 concentrations derived from satellite observations, which are not easily swayed by company-specific factors or other such narrow influences. What we have found is that this updated model yields estimates of private consumption that are more accurate than what we have been able to produce with our conventional model based on consumption-related indicators, and that the updated model yields estimates of consumption that are reliably accurate even under the unusual circumstances created by the COVID-19 pandemic.

Our research here has been focused on macroeconomic activity in Japan, but the satellite observational data that is available is not limited to any particular country or region of the world. We expect that the same method can be used to gain an understanding of economic activity in countries beyond Japan as well. We also note that making use of information from smaller units of mesh in the satellite observational data than we have studied here may allow for analyses that are based on a more detailed picture of the spatial distribution of CO_2 concentrations. In particular, we think that information on the spatial distribution of CO_2 concentrations may be especially helpful in analyses of countries within which the structure of industry varies by region. We see these as potentially fruitful avenues for further research.

References

[1] Ministry of Foreign Affairs of Japan, Intended Nationally Determined Contribu-tions (INDC): Greenhouse Gas Emission Reduction Target in FY2030, https://www.mofa.go.jp/ic/ch/page1we_000104.html

[2] Ministry of the Environment, Global Environment; http://www.env.go.jp/en/earth/index.html

[3] JAXA, GOSAT: Greenhouse Gases Observing Satellite "IBUKI", 2021 https://www.jaxa.jp/projects/pr/brochure/pdf/04/sat02.pdf

[4] GOSAT Project; https://www.gosat.nies.go.jp/eng/GOSAT_pamphlet_en.pdf

[5] Y. Suimon, How to Monitor Economic Activity by Using Alternative Data, Journal of the Japanese Society for Artificial Intelligence, 2021, vol. 36, no. 3, pp. 279-285.

[6] H. Clarck, M. Pinkovskiy and X. Sala-i-Martin, China's GDP Growth May be Understated, NBER Working Paper, 2017, no. 23323.

[7] Y. Suimon and T. Fukuma, Manufacturing Industrial Production Activity Analysis using Night-time light Image of Satellite, The Japanese Society for Artificial intelligence, 2019, 117th SIG-KBS, pp. 25-27.

[8] X. Chen and W. D. Nordhaus, Using luminosity data as a proxy for economic statistics, 2011, Proc Natl Acad Sci USA, vol. 108, no. 21, pp. 8589-8594.

[9] J. V. Henderson, A. Storeygard, and D. N. Weil, Measuring economic growth from outer space, American Economic Review, 2012, vol. 102, no. 2, pp. 994-1028.

[10] S. Michalopoulos and E. Papaioannou, Pre-colonial ethnic institutions and contemporary African development, Econometrica, 2013, vol. 81, no. 1, pp. 113-152.

[11] D. Donaldson and A. Storeygard, The View from Above: Applications of Satellite Data in Economics, Journal of Economic Perspectives, 2016, vol. 30, no. 4, pp. 171-198.

[12] T. Matsunaga and T. Yokota, Scientific and Other Achievements of Greenhouse Gases Observing SATellite (GOSAT), Journal of The Remote Sensing Society of Japan, 2019, vol. 39, no. 1, pp. 2-8.

[13] Y. Suimon and M. Yanai, Using business data to nowcast consumption, JCB Con-sumption NOW Report, 2019; https://www.jcbconsumptionnow.com/info/news-54