■受領№.1445

夜間光データを用いた物流発生集中量推定モデルの構築とその応用

代表研究者
川崎　智也
東京大学　講師

Estimating freight production/attraction using nighttime satellite imagery

Principal Researcher
Tomoya Kawasaki,
The University of Tokyo Lecturer

本研究は、夜間光データがゾーン別貨物発生集中量を推定するための重要な指標となり得るか検証するものである。分析の結果、夜間光データは GRP や人口などの一般的な社会経済指標よりも高精度に貨物発生集中量を推定できることが分かった。可住地面積を用いて夜間光を補正すると、さらに精度が向上することが示された。これらの知見は、特に統計データが限られている国や地域において、貨物の発生集中量の推定に有益なものとなることが考えられる。

This study aims to examine whether nighttime light intensity (NLI) data can be a significant indicator for estimating zonal freight production (FP) and attraction (FA) through a case study in Japan. As a result of the analysis, NLI exhibits better estimates than the popular indicators such as GRP and population. The use of habitable area for the estimation of total NLI provides the highest accuracy among the four types of proposed NLIs. Through the analysis, NLI data is sufficiently applicable for estimating zonal FP and FA, relative to other important socio-economic indicators. These findings could help estimating FP/FA particularly in countries and regions with limited statistical data available.

1. 研究目的

Nighttime light intensity (NLI) data is human-induced light emission data and the open alternative data used for estimating a variety of socio-economic activities and geographic conditions. Freight volume could be determined by the economics of production and consumption; based on this, NLI is possibly useful for FP and FA estimations. However, the applicability of NLI data for estimating freight volume has not been clarified in the existing studies. Therefore, this study aims to examine whether NLI can be a significant indicator for estimating the zonal FP and FA compared with other socio-economic indicators and to examine the types of NLIs that can accurately estimate the FP and FA. The case study is conducted in Japan, where both international and domestic FP and FA are available at the prefectural level. In the case study, we first retrieved monthly unit NLI data for 47 prefectures in Japan using the Google Earth Engine and proposed three total NLIs for each prefecture. As explanatory variables for estimating FP and FA, 196 socio-economic indicators were obtained from the central and municipal governments of Japan. Thereafter, these indicators, including NLI, were used as explanatory variables to develop FP and FA models by using an elastic net, enabling the sorting of significant and insignificant variables. In addition, we developed single regression models using each type of NLI and popular explanatory variables for the purpose of accuracy comparison for the FP/FA estimations.

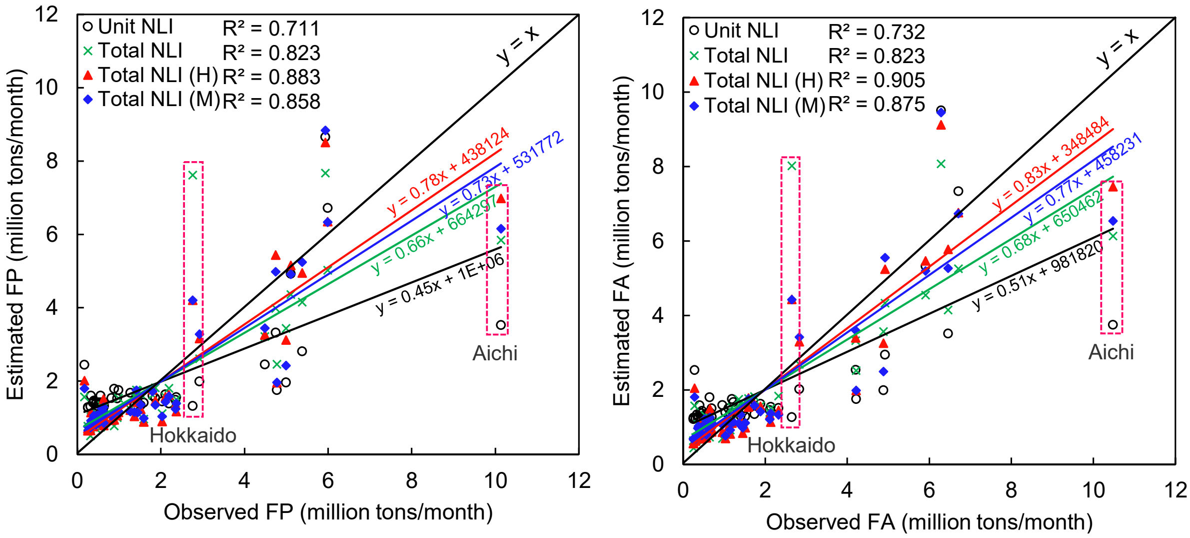
2. 研究内容

The parameter estimation of the elastic net for each FP and FA is shown in Table 1. We obtain 21 and 27 explanatory variables for the FP and FA model, when λ of elastic net is 29,364 and 59,687, respectively, as shown in Table 1. As a result, total NLIs were selected as significant variables, while the unit NLI was not selected in either model, as shown in Table 2. Since the explanatory variables are standardized, the relative impact between the variables can be discussed based on the magnitude of the coefficients. The total NLI (H) has a higher impact than the other total NLIs; in particular, it is only significant for the FP model. Although straightforward comparisons are difficult because the model structures for FP and FA are different, the relative importance of NLIs to the other variables is higher in the FA model. This is because areas with brighter lights are the consumption areas in many cases, whereas factories and other logistics-related facilities where high volume of cargo is produced are located in the suburb areas, not necessarily located in bright areas. Total NLI (M) was significant in the FA model for the same reason. In total NLI (M) calculation, the darker area was drastically eliminated compared with total NLI (H). Thus, total NLI (M) was not significant in the FP model. As a result, in comparison with other socio-economic variables, total NLIs are identified as relatively useful data for estimating freight volume at the prefectural level in Japan.

Table 2 Result of parameter estimation of elastic net (selected variables only)

Explanatory variable (x_i)	Coefficients	
	Production	Attraction
Total NLI		6,826
Total NLI (H)	13,300	328,113
Total NLI (M)		161,291
Total population		9,576
Day and night population ratio		-71,415
GRP (agriculture)	-21,434	-45,259
GRP (forestry)	-30,282	
GRP (mining)	117,211	
GRP (electricity, gas, water, waste processing industry)	103,276	
Prefectural income per capita	26,449	
Number of establishments (fishery)		-11,593
Number of establishments (secondary industry)		16,991
Number of employees (manufacturing industry)	11,181	182,736
Number of employees (secondary industry)		76
Product shipment amount	326,846	242,957
Manufacturing industry added value		187,844
Total amount of industrial water	178,052	78,021
Amount of industrial water (fresh water)	2,303	5,668
Amount of industrial water (seawater)	47,931	511
Air freight volume		-86,050
Passenger transported by ship	-84,322	
Sales (private sector) (manufacturing industry)	468,534	241,292
Sales (private sector) (electricity, gas, heat supply)	24,954	3,072
Residential area		66,781
Number of pollution complaints (typical 7 pollution)	178,932	104,573
Number of traffic accidents	71,676	215,864
Number of casualties by traffic accidents	76,356	44,106
Number of fatalities by traffic accidents	153,073	
Number of injured by traffic accidents	76,460	54,726
Number of driver's license holders (large vehicle)	183,205	101,271
Number of driver's license holders (middle vehicle)		184,867
Number of driver's license holders (regular vehicle)		30,289
Traffic accident matters	389,528	170,795
Sample size (N)	47	47
Adjustment parameter (α)	0.669	0.512
Regularization parameter (λ)	29,363.82	59,686.55

In many countries, especially developing countries, it is often not possible to implement analysis with multiple variables, as described in the previous section, since statistical data are not always rich. Therefore, in this section, we develop single regression models using the key variables that are important and frequently used variables for the estimation of FP and FA as the explanatory variables. The estimated values from the single regression models were compared with the observed values to examine the relative accuracy for the estimation of the FP and FA based on each explanatory variable. We compare the models using four types of NLIs (i.e., unit NLI, total NLI, total NLI (H), and total NLI (M)) and discuss which NLI accurately estimates the FP and FA. Using each indicator as an explanatory variable of the single regression model, the theoretical values of FP and FA are estimated and subsequently compared with the observed values. The results are presented in Figure 1. Due to the space limitation, we did not provide all statistics of single regression model but it is noted that all single regression models provided a coefficient of determination of more than 0.79, and all the estimated parameters satisfied more than the 5% significance level. R2 in Figure 1 and 2 indicate multiple correlation coefficients in terms of y = x (observed and estimated values). From Figure 1, the estimated values based on the total NLI (H) and the observed values have the highest R2 that is consistent with the result of the analysis based on elastic net. The accuracy of the estimation result of the FA model was higher than that of the FP model in all models. The slope of regression line between observed and estimated values for unit NLI is the smallest by 0.45 (FP) and 0.51 (FA) that tend to underestimate the observed values in prefectures that have relatively higher freights. For example, the actual FP volume of Aichi Prefecture is approximately 10 million tons per month, while the estimated value is 3 million tons, a discrepancy of 7 million tons. The estimated values for Hokkaido were approximately half of the actual values. This implies that these prefectures have a low NLI compared with their cargo production. The areas of these prefectures are large, with many dark areas, thus conversion to total NLI can mitigate this effect. Therefore, the results based on the total NLI improve the accuracy of estimation results, particularly in these prefectures. In addition, habitable areas are more effective adjustment factor than the other total areas for the estimation of FP and FA. In the elastic net analysis shown in Table 1, unit NLI was not selected as a significant variable for estimating FP and FA; however, R2 with respect to the unit NLI model exceeds 0.7 which indicates calculating FP and FA by unit NLI at the prefecture level is also feasible.

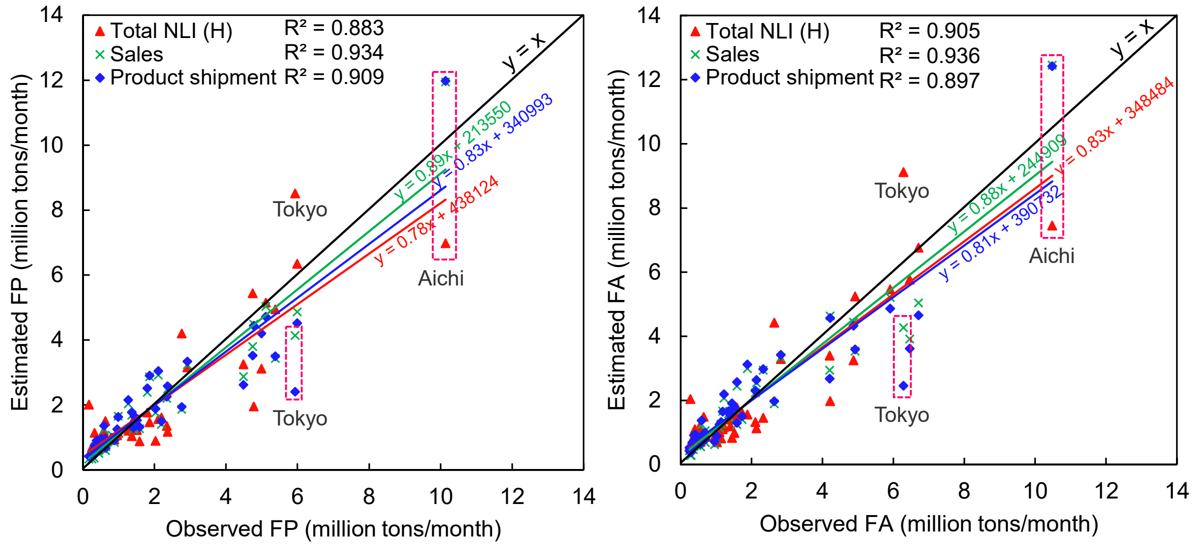


(a) FP (b) FA

Figure 1 Estimated and observed freight volumes based on four types of NLIs
*Note: The multiple correlation coefficient (R2) shown in the figure is R2 with respect to y

We additionally developed single regression models with the significant variables identified by elastic net (i.e., “sales” and “product shipment”) and compared them with the model developed by NLI (H) that exhibited the highest accuracy among the four NLIs in Figure 1. Similar to the models of NLIs, all estimated parameters exceed the 5% significance level and 0.75 of the coefficient of determination in single regression models with each “sales” and “product shipment.” In terms of R2 with respect to y = x for FP models, the models with sales and product shipment afford better estimation results than the total NLI (H), as shown in Figure 2.

In Aichi prefecture, the NLI (H) model underestimates the cargo volume in both the FP and FA models, whereas the models with “sales” and “product shipment” overestimate those values. This brings the regression line of “sales” and “product shipment” closer to y = x for FP model. Therefore, R2 increases in the FP models with “sales” and “product shipment”. However, FA model indicates slightly better estimation result on NLI (H) than product shipment. This also shows better utilization of nightlight imaginary data on estimating cargo volume particularly for freight attraction. In the comparison of “product shipment” and “sales” models, the “sales” model received a higher accuracy in terms of R2; in particular, Tokyo is well-estimated based on the “sales” model, while estimations by the models of the other two variables are less accurate. Consequently, it is shown that the “sales” model has the highest accuracy in terms of R2. For the model of NLI (H), it can be concluded that it is useful enough for FP/FA estimation because its accuracy is nearly equal to the models with the most significant variables (“product shipment” and “sales”) identified by elastic net.



(a) FP　　(b) FA

Figure 2 Estimated and observed freight volumes by NLI (H) and “sales”/ “product shipment”
*Note: The multiple correlation coefficient (R2) shown in the figure is R2 with respect to y = x.

2. 発表（研究成果の発表）

国内外の学会誌、学会講演会等における発表があれば 5 件程度記載。

記載内容：氏名、題目、誌名、巻、号、頁（年次）、学会名（場所、年次）

Kawasaki, T., Nakanishi, W., Hyodo, S., Namba, Y., Mori, H., Kishi, H.: Estimating freight production/ attraction using nighttime satellite imagery, Communications in Transportation Research, Vol.2, 100067, 2022.

Namba, Y., Kawasaki, T., Oka, H.: Estimation of freight trip distributions using spatial-temporal aggregated freight vehicle data, T-LOG 2022, Incheon, South Korea, 2022.9.

生本力丸・川崎智也・兵頭知：夜間光データを用いた地域生活圏レベルの発生集中貨物量の推計，第66回土木計画学研究発表会，琉球大学，2022.11.