ACRS 2022
MONGOLIA

CLOUD DETECTION OF VIIRS DNB USING GLCM FEATURES IN EAST ASIA

Jonggeol Park[1] , Ichio Asanuma[2] and Kanichiro Mochizuki[3] ,
[1]Tokyo university of Information Sciences, 4-1, Onaridai, Yakabaku, Chiba-city, Chiba 265-8501, Japan,
Email: amon@rsch.tuis.ac.jp
[2]Tokyo university of Information Sciences, 4-1, Onaridai, Yakabaku, Chiba-city, Chiba 265-8501, Japan,
Email: i3asanuma@gmail.com
[3]PASCO Corporation, 1-7-1 Meguro, Meguroku, Tokyo, 153-0063, Japan,
Email: kiaknu9933@pasco.co.jp

KEY WORDS: VIIRS NDB, Cloud, GLCM, Random forest classification, Support Vector Classification

ABSTRACT: In this study, clouds were detected in urban areas using night time VIIRS DNB. When light passes through thin clouds, the image appears blurred because of Mie scattering. GLCM was used to determine if the edges of the image were blurred. We proposed a GLCM with background (Non-economic area) 0 processing to be unaffected by ROI size. SVC and RFC were used to compare the accuracy of cloud detection. Finally, we compared the VIIRS cloud mask with the RFC results.

1. INTRODUCTION

Optical sensors cannot accurately observe the ground surface due to the influence of clouds. In addition, since the state of the ground cannot be known if the analysis is performed while the influence of clouds remains, it is necessary to create a cloud mask and remove the areas affected by clouds. MODIS, which is often used for time-series data, provides highly accurate global cloud masks (MOD35). However, thin cloud detection is still a very difficult problem using visible, near-infrared, and thermal infrared (Ackerman et al., 2008; Holz et al., 2008). Because the thermodynamic properties of clouds depend on the scattering and absorption properties of cloud droplets and particles, several thermal infrared band data were used for cloud discrimination (Pavolonis et al., 2005; Platnick et al. 2017). Elvidge et al. (2017) created a VIIRS nighttime light composite data that considers moonlight, solar zenith angle, gas flares, etc., using a decision tree. Roman et al. (2018) created a VIIRS global dataset by inputting snow, clouds, vegetation, ozone, etc., using the moon, ground, and ground-to-sensor atmospheric transfer model. In addition, the data set may have cloud misclassification due to effects such as attenuation and scattering by thin clouds with an optical thickness of less than 1.0τ, and scattering by side lighting. Wang et al. (2020) used the Random Forest (RF) method, which is one of the machine learning methods with spectral features as input, for cloud mask and cloud thermodynamic phase detection using VIIRS on Suomi NPP. board. Cloud detection using satellite data includes naive Bayesian classification (Uddstrom et al., 1999; Heidinger et al., 2012), RF classification (Wang et al., 2020), and Convolutional Neural Network (CNN) classification (Marais et al., 2012) have been studied using various models. However, the above study is a cloud detection method for daytime satellite data, so cloud detection in nighttime light is not applicable.

This research uses GLCM features that focus on inter-pixel gradation changes in light intensity due to Mie scattering of light in order to investigate the effects of clouds and other atmospheric factors on the nighttime light of VIIRS DNB. It does not use physical features such as temperature or reflectance from spectral data. In this study, we investigated the following four processes to detect the influence of clouds on the nighttime light of VIIRS DNB.

1. Accuracy comparison between support vector machine classification (SVC) and random forest classification (RFC) methods in cloud classification,
2. Comparison of classification accuracy of GCLM feature quantity with GLCM and background influence as 0,
3. Determining the optimal Region Of Interest (ROI) size for GLCM,
4. Comparison between the cloud mask and the results of the proposed method in this study

2. DATA

In this study, we use the Black Marble nighttime lights product suite (VNP46) data, which is a high-order processed VIIRS DNB onboard NASA's Suomi National Polar-orbiting platform (SNPP) satellite. The resolution is 500m and it is available from January 2012 to the present (as of September 1, 2022). VNP46 is global data that has undergone advanced processing such as cloud removal, atmospheric correction, terrain correction, vegetation correction, snow correction, and moon correction by NASA's Land Science Investigator-led Processing System (SIPS).

3. GLCM FEATURES WITH THE EFFECT OF THE BACKGROUND AREA SET TO 0

In this study, GLCM features are used to capture the blur phenomenon at the boundary between the light area and the background area. Fig. 1 shows changes in GLCM features when the ROI size is changed for a clear night light and a blurred night light source due to clouds (Fig. 1a, b, c, d). Figure 1a) is an ROI image (5x7) with a value of 0 (dark pixels) in the periphery and a value of 3 (bright pixels) in the center, and 1b) is an ROI image in which the center pixel is blurred to 3x3. Figures 1c) and 1d) are ROI images obtained by extracting only 3 × 3 of the center of 1a) and 1b). ROI images were 4-level grayscale and GLCM used a sum of 4 angles for a distance of 1 (d=1). Since the target pixel in Fig. 1a) is 3, the difference from the surroundings (the DN value is 0) is 3. In Fig. 1b), the DN value changes in steps like 2, 1, 0 from the center, and the maximum DN difference is 2. Considering only the part with a high DN difference, 1a) should have a higher contrast, but 1b) has a higher contrast value because the background area of 1a) is wider (1a is 0.679, 1b is 0.755). Conversely, in the case of homogeneity, 1a) (0.932) is higher than 1b) (0.758). If the ROI image is small, 1c) has a high contrast (3.6 for 1c and 0.6 for 1d) and a small homogeneity (0.64 for 1c and 0.70 for 1d). 1a) and 1c) are images with different ROI sizes and a central pixel value of 3 and a background of 0. 1a) with a larger ROI has lower contrast and higher homogeneity. Since the feature calculation by GLCM is performed for the entire ROI pixels, the feature changes depending on the number of background elements with a value of 0. Therefore, in this study, we propose a GLCM to reduce the effect of the background region. By removing the change in the background area (change from DN value 0 to 0) from the ROI image, it is possible to emphasize the change in the light area such as edges (for example, DN value 3 to 0). Use the GLCM of Fig. 1a´ - d´ with the upper left (change in DN value from 0 to 0) value forced to 0 in the GLCM of FIG. 1a - d. As a result, the GLCM features of 1a') and 1c') have the same value even if the width of the background differs depending on the ROI size. The GLCM features in Figs. 1b') and 1d') are different, but if the size of 1d') is 5 × 5, they have the same value. By using GLCM that reduces the influence of the background area, the contrast is higher in 1a' than in 1b' (1a' is 9.00, 1b' is 1.54) and the homogeneity is higher in 1b' (1a' is 0.10, 1b' is 0.51) without being affected by the ROI size.

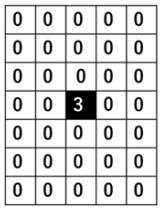
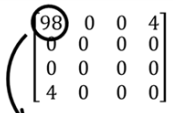
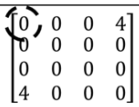
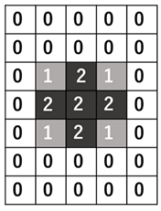
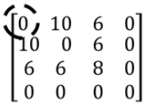
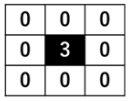
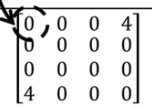
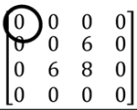
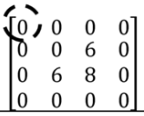
	ROI	GLCM	GLCM feature					
			cont	diss	homo	ASM	corr	ND(co,ho)
a)			0.68	0.28	0.93	0.88	-0.04	-0.97
a')	Backgroud 0 process		9.00	3.00	0.10	0.50	-1.00	-0.08
b)			0.76	0.53	0.77	0.3	0.4	-0.98
b')	Backgroud 0 process		1.54	1.08	0.51	0.15	-0.12	-0.96
c)			3.60	1.20	0.64	0.44	-0.25	-0.9
c')	Backgroud 0 process		9.00	3.00	0.10	0.50	-1.00	-0.08
d)			0.60	0.60	0.70	0.34	-0.43	-0.98
d')	Backgroud 0 process		0.60	0.60	0.70	0.34	-0.43	-0.98

Figure 1. Example of 4-level GLCM with background effect set to 0

For 1a' GLCM features were calculated with the background area of 1a set to 0 (same as 1b', 1c', 1d'). Here cont means Contrast, diss means Dissimilarity, homo means Homogeneity, ASM means Angular Second Moment, corr

means Correlation, ND(co,ho) means (Contrast/88.8 – Homogeneity/0.834) /(Contrast/88.8 + Homogeneity/0.834).

4. RESULT

4.1 Classification results by RFC

From the VNP46A1 images from January to December 2020, a total of 40577 locations with central pixels of 22736 cloud-free locations and 17841 cloudy locations were selected visually while comparing the VCM, DNB and M12 images. SVC and RFC were used as machine learning methods, and 70% of the training data (28403 locations) and the remaining 30% of the validation data (12174 locations) were selected by a random function. Figure 2 shows the accuracy of RFC for GLCM and GLCM with 0 processing for the background area, to investigate the influence of clouds on the ROI image. In the cloud area classification by RFC by ROI size change, the accuracy of the training data was 99.996%, and the accuracy of the validation data increased from 89.3% to 97.0%. In the case of GLCM with 0 processing of the background region, the accuracy of the training data was 99.996%, and the accuracy of the validation data increased from 91.0% to 97.5%. It was found that the accuracy was 0.5% to 2% higher when using the GLCM feature value with the background region processed to 0. In addition, when the ROI size exceeds 33, the accuracy of the validation data changes little from 97.5%, so the appropriate ROI size was set to 33. A small ROI size is better because a large ROI size may leave some small clouds and misclassify them. The RFC learning model for cloud detection uses an ROI size of 33.

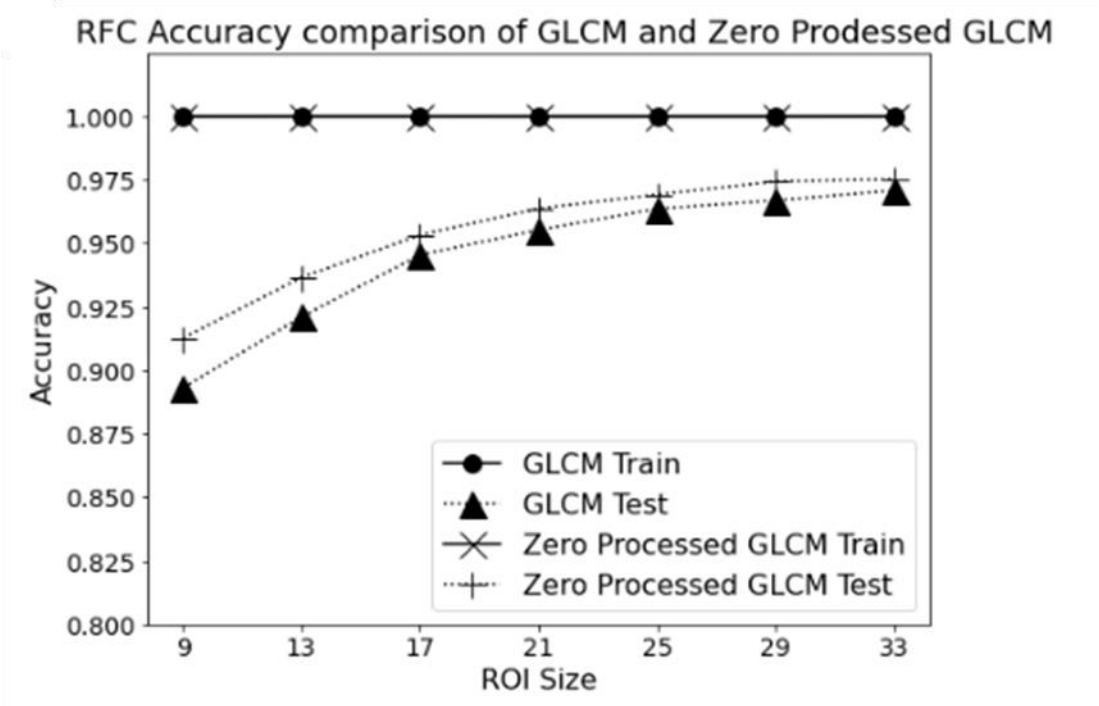
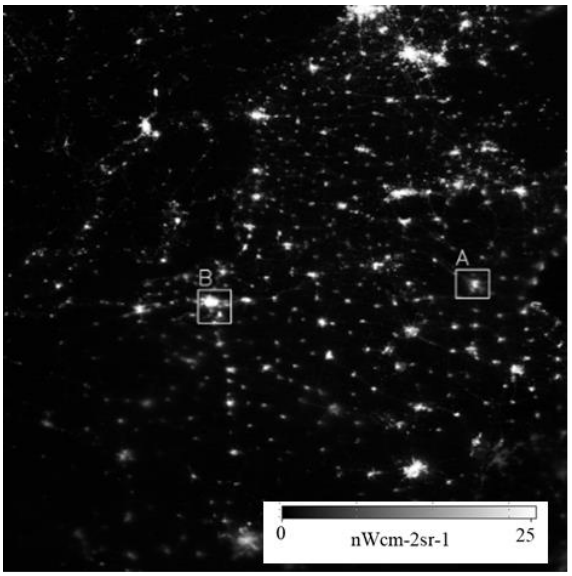


Figure 2. RFC Accuracy comparison of GLCM and Background 0 processed GLCM

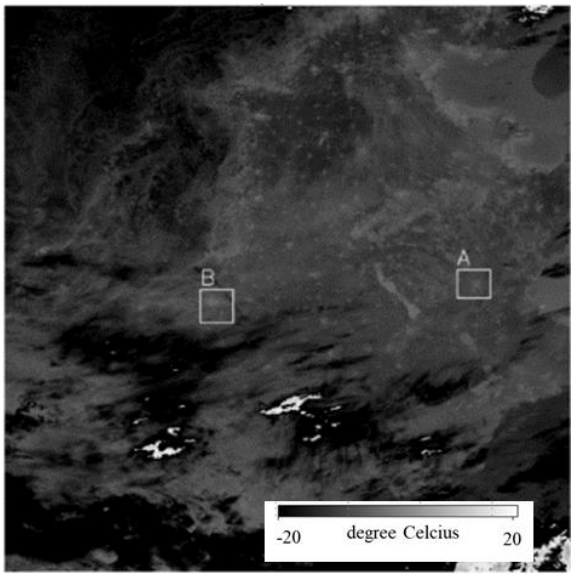
4.2 Comparison of cloud mask and RFC results

To investigate the effectiveness of cloud area detection by GLCM, which is proposed in this study with the influence of the background area set to 0, it was compared with a cloud mask (VIIRS Nighttime Cloud Mask; VCM). The ROI was cut out by raster scanning the DNB image, and the cloud area was classified by the RFC learning model. Figure

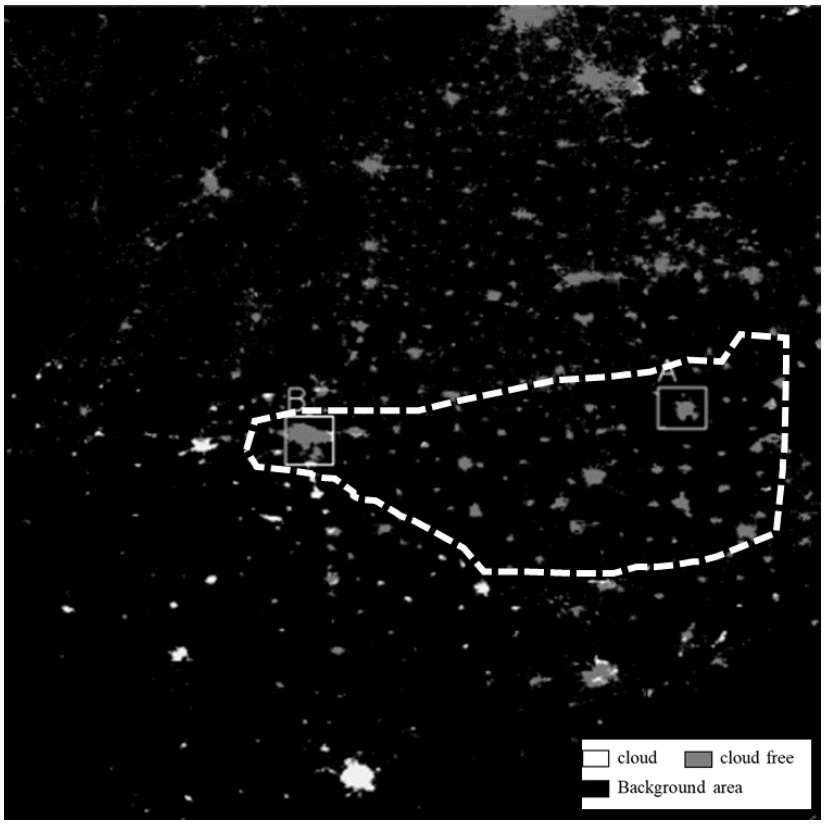
3 is the VNP46A1 image of tile number h29v05 on March 5, 2019. Fig. 3a) is a DNB image, and b) is a VCM image with cloud areas in white and sunny areas in black. Fig. 3c) is the radiance image (M12) with a center wavelength of 3700 nm, and d) is an image of clouds determined by the RFC learning model. White is the cloud area, gray is the sunny area, and black is the background area. In the cloud classification of RFC, the cloud area is broader than b) in southern and western China (dotted line area). Therefore, 3e) to h) show an enlarged view of Linyi City in Western China. Because the VCM and RFC results are different, we compared it with the DNB image on April 3, 2020, when Linyi was having a fine day (Fig. 3i-l). In the DNB of Fig. 3i), clouds cannot be visually confirmed because urban areas and square circulation roads can be seen. Both VCM and RFC results were classified as a fine day. Comparing nighttime light images 3i) and 3e) of Linyi City on a fine day, it can be seen that the DNB in 3e) is more blurred due to atmospheric influences such as clouds.



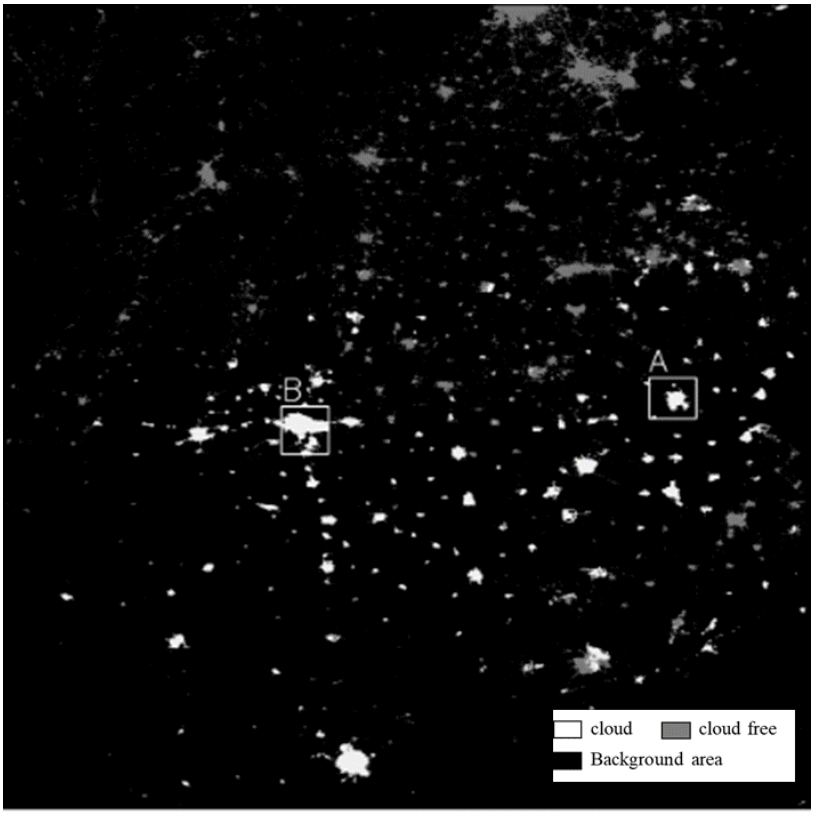
a) DNB



c) M12



b) DNB image masked by VCM



d) RFC result with background 0 processed GLCM

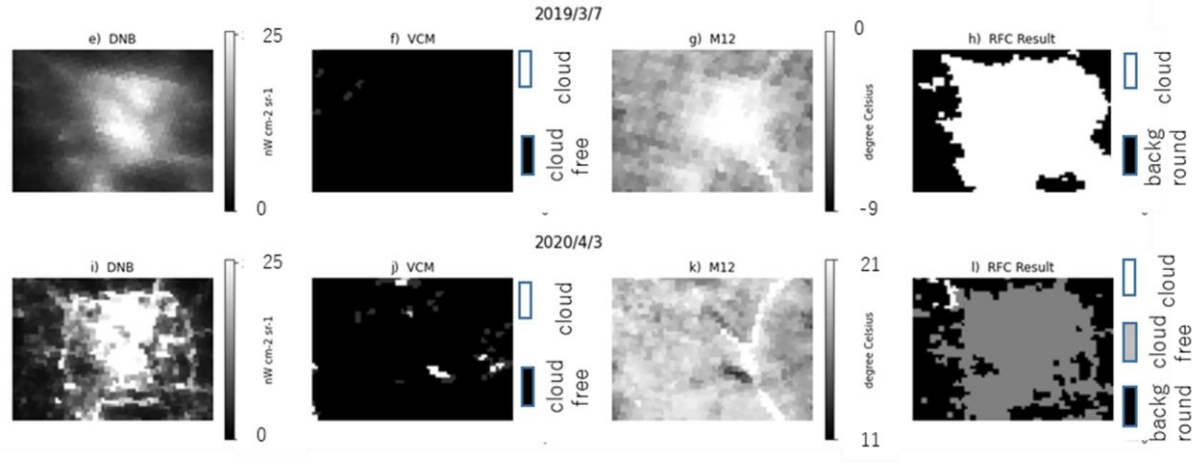


Figure 3. Comparison VIIRS Cloud Mask (VCM) and RFC result

This is the result of cloud discrimination using the DNB image on March 5, 2019. a) is the DNB image, b) is the night cloud mask (VCM) of VIIRS, c) is M12 with a center wavelength of 3.7 μm, and d) is the result of cloud detection result by RFC with the background effect set to 0. e) is an enlarged DNB image of the rectangular part (Linyi City) of a), f) is the VCM of the same area, g) is M12, and f) is the cloud judgment result by RFC. Figure 3i-l) are DNB, VCM, M12 and RFC results on April 3, 2020, when Linyi was sunny.

5. DISCUSSION

In this study, using VIIRS DNB images, we investigated the effects of clouds on economic activity areas at night time. We focused on the phenomenon that light from the ground is scattered and blurred when Mie scattering is affected by thin clouds. GLCM of texture analysis is used to investigate the features of the grayscale difference near the edge between economically active and non-economically active areas at night. Considering that the GLCM has a wide area of non-economic activity (background area) at night, we proposed a GLCM in which the background area

is treated as 0. The use of GLCM features with 0 processing in the background region improved the accuracy by 3-5% for SVC and by 0.5-2% for RFC. In addition, the effective feature quantities in the RFC classification were contrast, ND(co,ho), dissimilarity, and homogeneity in that order. According to the RFC results, even if the ROI size is increased from 33, the accuracy of the validation data does not rise above 97.5%, so we decided that 33 is the appropriate size. We compared cloud classification results by RFC and VCM image. The RFC classification was found to be effective only in the human economic activity area at night.

References

Ackerman, S. A., Holz, R. E., Frey, R., Eloranta, E. W., Maddux, B. C., McGill, M., 2008. Cloud Detection with MODIS. Part II: Validation, J. Atmos. Ocean. Technol., 25, pp. 1073–1086, https://doi.org/10.1175/2007JTECHA1053.1, http://journals.ametsoc.org/jtech/article-pdf/25/7/1073/3333780/2007jtecha1053_1.pdf

Elvidge, C. D., Zhizhin, M., Baugh, K., Hsu, F. C., 2015. Automatic Boat Identification System for VIIRS Low Light Imaging Data, Remote Sens. Vol. 7, pp.3020-3036; https://doi.org/10.3390/rs70303020

Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C., Ghosh, T., 2017, VIIRS night-time lights, Int. J. Remote Sensing, Vol. 38, 5860-5879; https://doi.org/10.1080/01431161.2017.1342050

Henderson, J. V., Storeygard, A., Weil, D. N., 2012. Measuring economic growth from outer space. The American Economic Review, Vol.102, No.2, pp. 994-1028.

Holz, R. E., Ackerman, S. A., Nagle, F. W., Frey, R., Dutcher, S., Kuehn, R. E., Vaughan, M. A., Baum, B., 2008. Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP, J. Geophys. Res. Atmos., 114, https://doi.org/10.1029/2008JD009837, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008JD009837

Kurota, Masamitu, 2017, Correlations between nighttime light and socio-economic indicators in low-income countries, Sophia Economic Review, Vol.62, pp19-26 (Japaness)

Mann M.L., E.K. Melaas, A. Malik, 2016, Using VIIRS day/night band to measure electricity supply reliability: preliminary results from Maharashtra, India, Remote Sens., p. 711 https://doi.org/10.3390/rs8090711

Pavolonis, M. J., Heidinger, A. K., Uttal, T., 2005. Daytime global cloud typing from AVHRR and VIIRS: Algorithm description, validation, and comparisons, J. Appl. Meteor. Climatology, 44, pp.804–826

Uddstrom, M. J., Gray, W. R., Murphy, R., Oien, N. A., Murray, T.: A Bayesian Cloud Mask for Sea Surface Temperature Retrieval, J. Atmos. Ocean. Technol., 1999, 16, 117–132, https://doi.org/10.1175/1520-0426(1999)016<0117:ABCMFS>2.0.CO;2,

Welch, R. and Zupko, S., 1980. Urbanized Area Energy Utilization Patterns from DMSP Data, Photogrammetric Engineering and Remote Sensing, Vol. 46, pp. 201 − 207

Xi Chen and William Nordhaus, 2015. A Test of the New VIIRS Lights Data Set: Population and Economic Output in Africa, Remote Sens. 2015, 7(4), pp. 4937-4947, https://doi.org/10.3390/rs70404937

Zhao, N., Hsu, F. C., Cao, G., Samson, E. L., 2017. Improving accuracy of economic estimations with VIIRS DNB image products, Int. J. Remote Sensing, Vol. 38, 38, pp. 5899-5918, https://doi.org/10.1080/01431161.2017.1331060