Traffic Flow at Night: a custom algorithm for identifying basal nighttime radiance levels of roadways

Gabriel Da Rocha Bragion[1], Gabriel C. Gonçalves[1], Ana Paula Dal'Asta[1], Ana Carolina De Faria Santos[1], Lucas Maia De Oliveira[1], Antônio Miguel Vieira Monteiro[1], Silvana Amaral[1].

[1]Remote Sensing – National Institute for Space Research (INPE)
Av. Dos Astronautas – São José dos Campos – SP – Brazil

{gabriel.bragion,gabriel.goncalves,ana.dalasta,lucas.maia,
miguel.monteiro, silvana.amaral}@inpe.br, anacarolina.fs@outlook.com

Abstract. *The recent COVID-19 outbreak drove the attention to methods for monitoring the flow between settlements, including traffic flow. Although the remote sensing of nighttime lights is a viable option to estimate traffic flow derived indicators, changes on radiance levels at night are not all associated with traffic. This paper presents the theoretical approach proposed on the development of an algorithm able to identify spectrally unbiased control samples for regions of interest (ROI), namely roadway sections. Firstly, an overview of the algorithm is presented, followed by an empirical estimation of its time complexity. The results showed that the algorithm has an O(n) time complexity and that control samples and ROIs can have similar time series features, indicating that an analysis without the use of control samples can lead to biased results.*

1.Introduction

Typifying and monitoring regional road traffic spatiotemporal patterns might be crucial to better understand the possibilities of COVID-19 spread between human settlements. The monitoring of phenomena associated with the road traffic via remotely sensed data is mostly restricted to very high-resolution sensors or on-road measurements, which are often neither accessible nor systematically distributed [Tuerner et al. 2013]. Some studies exploited images and composites from the Visible Infrared Imaging Radiometer Suite (VIIRS) – Day/Night Band (DNB) to successfully detect and monitor light sources at night in a sub-pixel level, such as boats, gas-flares, and biomass burning, but only a few approached the traffic of land vehicles [Elvidge et al. 2015a, Polivka et al. 2016, Elvidge et al. 2015b, Chang et al. 2019]. Road traffic lights can be relatively dim and arguably hard to resolve from space, given the anisotropic factor and oblique emission angles of auto headlights [Kyba et al. 2014]. Therefore, the characterization of spatiotemporal patterns associated with road traffic radiance at night would be more robust if supported by methods for identifying patterns strictly associated with environmental changes, rather than the road traffic itself.

Albeit diverse, on-road sensors are most usually fixed and used for monitoring and fining purposes. Despite registering the total number of vehicles along a roadway section in a very fine spatial and temporal resolution, this type of sensor is not to be found in many smaller roadways. Smaller roadways are often the only vehicle route available to less prominent towns and play an important role in the spreading of contagious diseases

to areas that are closer to the base of the human settlement hierarchy [Fortaleza et al. 2020]. The use of satellite remote sensing methods comes in handy, for it can produce spectral information in a regular and extensive form, surpassing the drawbacks of the on-road driven methods.

Although daytime high-resolution imagery matches these criteria, studies of this sort are mostly focused on object-oriented techniques, limiting the recognition of features that have a similar spatial scale to a given sensor ground sample distance (GSD) [Batz and Schäpe 2000]. High-resolution imagery still has a high cost of acquisition, generally covering areas only on demand, and lacking the higher availability usually meat by moderate resolution sensors. However, sub-pixel target detection based on the reflectance factor analysis, through moderate resolution sensors, usually requires a higher spectral resolution [Change and Heinz 2000]. In this sense, the detection of targets at night is a suitable approach, for it does not depend on a higher spectral resolution, neither a finer GSD.

Target detection from nightly imagery is mostly based on the expected level and frequency of radiance associated with optical radiation sources [Elvidge et al. 2015]. The radiation amount from headlights measured by the DNB sensor can be lower than the amount expected from other typical artificial light sources, potentially lower than high albedo features and background areas near lit sites [Kyba et al. 2014]. Chang et al. (2019) presented a study analyzing the correlation between traffic flow and a DNB derived metrics from accumulated pixels overlapping freeways in China, and found out a broad correlation degree (R^2 ranging from 0.267 to 0.818), depended on the metric and vehicle type. Despite the outcome, factors like the higher density of roadhouses on higher flow freeways could lead to similar results, putting in check the assumption that these correlations are strictly due to the vehicle flow.

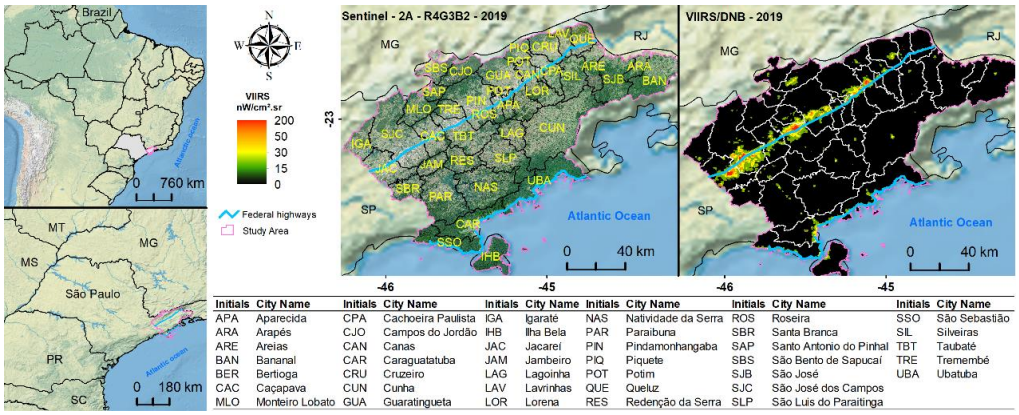
Previous studies showed that the monthly nighttime lights (NTL) average radiance is correlated to factors like the vegetation cover and changes in albedo and that some of these influences can be found even in the annual NTL composites [Levin 2017, Levin and Zhang 2017]. In order to identify different spatiotemporal patterns of road traffic from the monthly NTL composites, one must first investigate what is the contribution of other side parameters to the changes observed in the average radiance levels at night. Part of this problem could be analyzed by comparing the monthly average radiance associated with pixels lacking the presence of light sources to pixels overlapping roads of similar spectral response. This paper presents the theoretical approach applied to develop an algorithm able to identify spectrally unbiased unlit areas. These areas shall be used as control samples of the radiance levels from roadways sites, allowing a systematic estimation of the basal radiance level of a given roadway. An analysis of the empirically estimated time complexity is presented, given that early drafts of the code tended to present an exponential time complexity growth.

2. Material and methods

2.2. Study Area

The Metropolitan Region of the Paraíba Valley and North Coast (RMVPLN) is located in the State of São Paulo and, along the BR-116 highway, it connects the metropolitan regions of São Paulo (RMSP) and Rio de Janeiro (RMRJ) (Figure 1). The region comprises 39 municipalities divided into five subregions, holding a high variety of

economic activities and a heterogeneous demographic distribution, most concentrated in the urban areas. [Gomes, Reschilian, and Uehara 2018]. In a fresh reading of the RMVPLN centered regional planning, Gomes, Reschilina, and Uehera (2018) pointed out that the strategic location of the RMVPLN seemed to led the political vision for development towards the exploiting of the distinct local advantages in a competitive way. Although logic, without the population, society, and political engagement, this historical approach for the development resulted in sub-regional inequalities, without the carrying for the urban-regional needs as well. Smaller town workers often adopt a daily routine of traveling across different municipalities in carpool systems or public transportation, increasing the probability of transmission and spreading of the COVID-19 from hub cities to smaller towns.



Initials	City Name	Initials	City Name	Initials	City Name	Initials	City Name	Initials	City Name	Initials	City Name
APA	Aparecida	CPA	Cachoeira Paulista	IGA	Igaraté	NAS	Natividade da Serra	ROS	Roseira	SSO	São Sebastião
ARA	Arapés	CJO	Campos do Jordão	IHB	Ilha Bela	PAR	Paraibuna	SBR	Santa Branca	SIL	Silveiras
ARE	Areias	CAN	Canas	JAC	Jacareí	PIN	Pindamonhangaba	SAP	Santo Antonio do Pinhal	TBT	Taubaté
BAN	Bananal	CAR	Caraguatatuba	JAM	Jambeiro	PIQ	Piquete	SBS	São Bento de Sapucaí	TRE	Tremembé
BER	Bertioga	CRU	Cruzeiro	LAG	Lagoinha	POT	Potim	SJB	São José	UBA	Ubatuba
CAC	Caçapava	CUN	Cunha	LAV	Lavrinhas	QUE	Queluz	SJC	São José dos Campos		
MLO	Monteiro Lobato	GUA	Guaratinguetá	LOR	Lorena	RES	Redenção da Serra	SLP	São Luis do Paraitinga		

Figure 1. Municipalities, Average Nighttime Radiance of the Metropolitan Region of the Paraíba Valley and North Coast and samples.

2.1. Data and algorithm procedural approach

Monthly cloud-free nightly composites ("vcm" version), which were the main input data of this work, are processed and made available by the Earth Observation Group (EOG), at Payne Institute for Public Policy website (https://eogdata.mines.edu/download_dnb_composites.html). The composites represent the monthly average radiance at the surface from daily cloud-free pixels belonging to images retrieved by the Day-Night Band (DNB) sensor. The Visible Infrared Imaging Radiometer Suite's (VIIRS) DNB sensor, onboard the joint NASA/NOAA Suomi National Polar-orbiting Partnership (Suomi NPP) satellite, retrieves daily radiance values at night, approximately at 1h:30min, local time. The instrument collects data on a constant 742x742 m footprint, but its monthly composites are binned to a global 15 arc-second geographic grid (~463m at the equator) [Elvidge et al. 2017]. Nightly sensed radiance values were needed for two different algorithm processes. Firstly, they are required to assess the presence of light sources under a pixel footprint. For this purpose, it was determined that every pixel with an average radiance value higher than 2ηW/cm².sr is not to be considered as background by the algorithm, a value considered higher than the average background radiance for latitudes between 10 and 50º [Elvidge et al. 2017].

Road network data was retrieved from the National Cartographic Base, made available by the Brazilian Institute for Geography and Statistics (IBGE) [IBGE 2019]. Although unlit roadways might be the closest available targets to be selected as control samples, a preliminary analysis showed that the magnitude of a cluster of pixels'

radiances values overlapping a roadway can be as dim as completely unlit areas, in some cases. Moreover, there is no general optimal radiance value to specify if there are no light sources in a roadway, since traffic, noise and background radiance values are not stationary, both in space and time [Elvidge et al. 2017]. Therefore, the algorithm must automatically assume that a pixel overlapping a roadway is a non-background area.

Surface reflectance values were extracted from the MODIS MCD43A4 collection, band 1 (620-670nm), 2 (841-876nm), and 4 (545-565nm), a collection of images containing the best pixels of a 16-days-moving-window that have been modeled as if they were taken from a NADIR instantaneous field of view [Schaaf and Wang 2015]. The selected bands correspond to all the available bands in between the VIIRS/DNB spectral coverage (500 - 900nm). To increase the probability of high-quality pixels and proceed with the analysis with a more temporal compatible data, the quality assessment (QA) band of the MCD43A4 collection was consulted to produce a 30-day single composite for each month, ranging from January 2013 to January 2020. Finally, the processed MCD43A4 30-day images were reprojected to match the VIIRS/DNB grid.

VIIRS/DNB monthly NTL composites, MCD43A4 surface reflectance, and road network data are ingested into the algorithm (Figure 2a). Once the datasets are processed, the MCD43A4 data is associated to a region of interest (ROI) (Figure 2b). The goal of the algorithm is to find a cluster of 3x3 pixels, namely a control sample candidate, with the closest spectral response to a specific ROI sample, given a series of restrictions (Figure 2c and 2d). The proceedings illustrated by Figure 2 were implemented in a Python 3.5 environment.

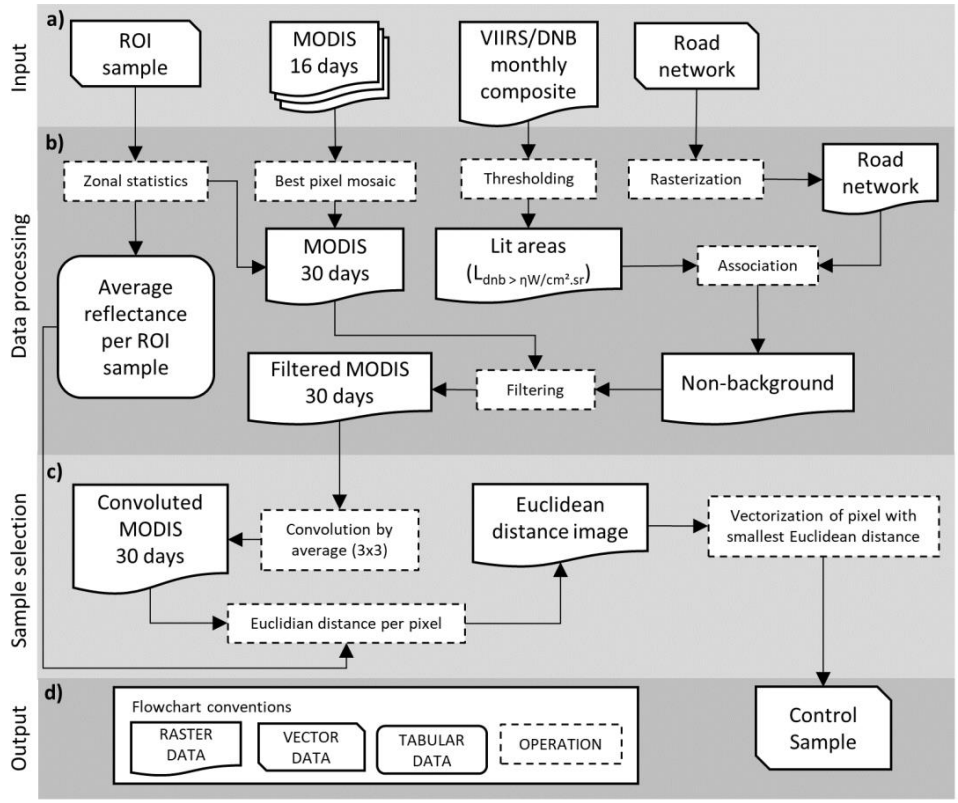


Figure 2. Flowchart of the proposed algorithm.

ROIs are clusters of pixels overlapping a roadway section. Currently, each ROI comprises nine pixels and can sum up to 5.9 km of roadway. A larger amount of pixels would result in a higher spectral mixture, difficulting the search and validity of a control sample. They were selected considering the functional role of the section, prioritizing roadways that give access to the municipalities but are not embedded in the lit urban area. The spectral response is represented by the average reflectance of a control sample candidate or ROI, based on the processed MCD43A4 bands, while its likehood is given by the Euclidean distance between those metrics (Equation 1).

$$\tilde{\lambda} = \sqrt{(\overline{B1}_{ROI} - \overline{B1}_{ctrl}) + (\overline{B2}_{ROI} - \overline{B2}_{ctrl}) + (\overline{B4}_{ROI} - \overline{B4}_{ctrl})} \quad \textbf{(Equation 1)}$$

Where $\tilde{\lambda}$ is the spectral likehood, $\overline{Bn}_{ROI}$ is the ROI's average reflectance from the n'th band of the MCD43A4 product, and $\overline{Bn}_{ctrl}$ is the control sample's average reflectance from the n'th band of the MCD43A4 product.

Apart from the restrictions aforementioned, a control sample must not contain an invalid pixel. An invalid pixel is a pixel whose value has no true physical meaning, either due to instrument problems or cloudy atmosphere conditions during the acquirement of data. Based on these restrictions, the algorithm must test every sample candidate and then calculates their $\tilde{\lambda}$. Finally, the control sample candidate with the smallest spectral likehood is elected as a control sample for that specific ROI, given a specific month.

3. Algorithm time complexity and overview

A profiling of the algorithm identified the functions related to the calculation of average radiances and reflectances as the most time-consuming ones. Both functions increase the number of operations as the number of images or samples is increased. Those specific operations are dependent on third-party functions, making it difficult to determine the complexity of the algorithm in a theoretical approach. Therefore, we empirically tested the time demanded by the algorithm while increasing the number of images and ROI in 90 different combinations (Figure 3).

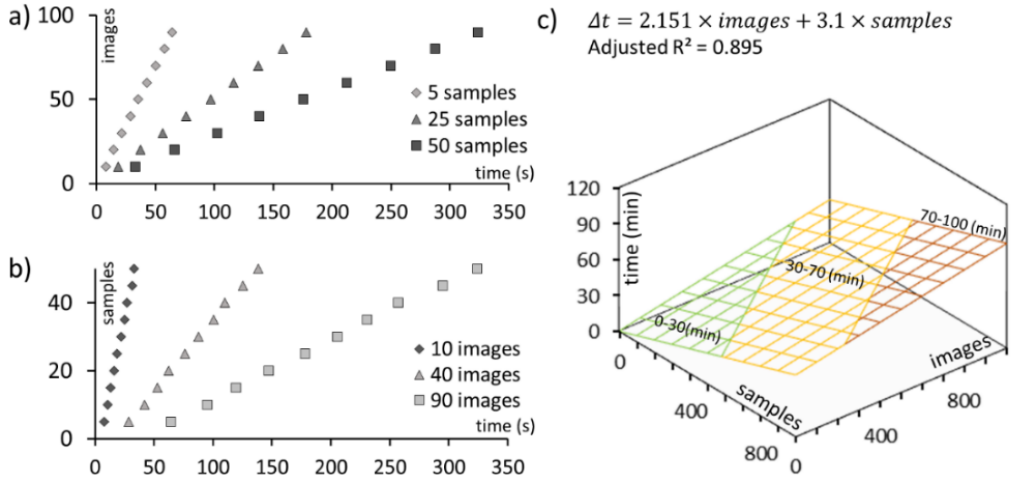


Figure 3. Algorithm time demand for different conditions regarding the number of images and samples, and fitted multivariate linear model.

In addition to the visual evidence, all the tested conditions presented a strong linear correlation ($0.996<R^2<0.999$; Figure 3.a and 3.b), suggesting that a linear multivariate regression model is appropriate to represent the time demand of the algorithm, regarding the number of images and samples (adjusted $R^2 = 0.895$; Figure 3.c). The ratio of the β_1 coefficient to the β_2 coefficient indicates that the number of images increases the time demand 44.12% slower than the number of samples. This is a positive outcome, since the number of images is the only dimension that grows indefinitely in a typical monitoring scenario. Another theoretical concern of the algorithm is the time complexity related to the sample selection method. Rather than testing the fitness of a given window as a potential control sample, and only then calculating the average reflectance, the method applied takes advantage of the algebraic implementation of the convolution filter [Tomilieri and Lu, 1997]. After reading all images as two-dimensional arrays, the algorithm defines all unelectable entries as non-numeric data. Therefore, all subsequentially operations result in a non-numeric entry, which is automatically excluded from the identification of the pixel with the smallest Euclidean distance, avoiding the need for multiple restriction tests.

Due to the association of restrictions criteria to non-numeric data, the algorithm is already able to filter off samples where there is a lack of good quality DNB or MODIS data, resulting in a gap in the time series. Moreover, the restrictions are all individually stored in arrays that can be retrieved based on the selected control sample coordinates. This allows the user to set quality flags to the output data, indicating what step has coerced the data to a non-numeric format, or even retain the values of the metrics needed for the processing methods and sample selection. This approach results in a series of metadata that can be used to further investigate the algorithm outcome and eventually investigate more precise alternatives to deal with problems concerning the input data quality. Currently, the output is a vector (or spatial table) comprising the average monthly radiance VIIRS/DNB values for both ROI and independent control samples, but the aforementioned metrics can be assigned to the table on demand.

Figure 4 displays examples of the algorithm's main outputs. Several relevant observations can be pointed out through a visual inspection of both averaged radiance time series (Figure a.2 and b.2). Regarding the averaged radiance level of the ROIs, it is clear that different roadways have distinct nominal radiances. While the seven-year time series of the BR-116 roadway shows values ranging from about 2.5 to 7.5 $\eta W/cm^2.sr$, BR-353's has averaged radiance values barely higher than 0.4 $\eta W/cm^2.sr$, the very same range observed in most of the control sample's time series. Likewise, Cao and Bai (2014) found averaged radiance values ranging from 2 to 4 $\eta W/cm^2.sr$ after analyzing daily DNB's radiances from a bridge section over the San Francisco Bay, California. These results indicate that it might be meaningless to define thresholds in order to separate lit from unlit areas, since there is a relatively wide range of mixture in radiance levels from background and dim or transient lit areas.

When it comes to heavy traffic roadways, such as the BR-116, the difference between the ROI's and the control samples' averaged radiance does not seem to change the time series' aspect (Figure 4a.2). Even though, after the subtraction, the resulting time series does present some relevant differences if compared with the original one, mainly expressed as shifts in the direction of the series in several pairs of months. The changes in the time series' aspect are clearer when observing roadways with a nominal dimmer averaged radiance. In Figure 4b.2, a major increase in the average monthly radiance is

observed from 2017 onwards. Without taking account of the control samples' time series, it could be wrongly interpreted there was a relevant change in the vehicles' regime; or even a restructuring of the outdoor lights of nearby settlements.

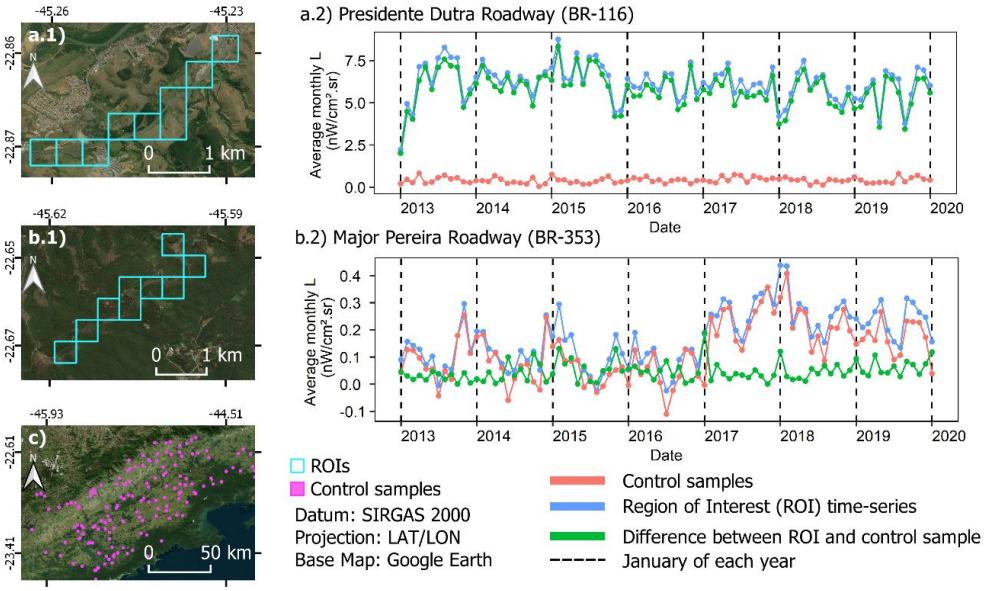


Figure 4. Location (a.1, b.1, c) and radiance level time series (a.2, b.2) of different ROIs and their respective control samples.

4. Conclusion

This paper presents the theoretical approach and the implementation strategy used on the elaboration of an algorithm able to access meaningful unlit control samples of monthly NTL composites based on its spectral response and restrictions criteria. An overview of the algorithm's empirically estimated time complexity showed that all the operations established by the code can be executed in a linear form. Currently, 85 monthly NTL composites are available to be ingested in a time series analysis, but this number will grow undefinedly. Moreover, daily processed nighttime images from the DNB sensor will be made available by NASA's Black Marble project soon (blackmarble.gsfc.nasa.gov/), stressing the need for algorithms that can be applied efficiently. In this instance, different datasets will certainly call for different likehood metrics and restriction criteria, nonetheless, they all can take advantage of the conceived structure of the presented algorithm.

By comparing the time series of roadway sections and spectrally-similar background areas identified by the proposed algorithm, it was confirmed that changes in radiance levels of roadways are not all associated with traffic flow at night. The results show that in the consideration of the VIIRS/DNB monthly composites as a dataset able to express quantitative information about the traffic of vehicles at night, the analysis of control samples is a necessary step. Whether variations in traffic flow can be detected by the VIIRS/DNB monthly composites, and what is the effect of the COVID-19 outbreak over the traffic flow at night are scientific questions that will be addressed in future works.

Acknowledgments

A Python 3.5 version of the algorithm discussed in this paper is available at the Laboratory for Investigations of Socio-Environmental Systems website (www.lissinpe.com.br/c%C3%B3digos). The authors are not aware of any conflict of interest. This study was financed in part by the Higher Education Improvement Coordination (CAPES) – Finance Code 001.

References

Baatz, M. and Schäpe. (2000). "Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation". In: *Angewandte Geographische Informations-Verarbeitung, XII*, Karlsruhe, Germany, p. 12-23.

Cao, C. and Bai, Y. (2014). "Quantitative Analysis of VIIRS DNB Nightlight Point Source for Light Power Estimation and Stability Monitoring". *Remote Sensing*, v. 6, n. 12, p. 1 - 16.

Chang, Y.; Wang, S.; Zhou, Yi; Wang, L.; Wang, F. (2019). "A Novel Method of Evaluating Highway Traffic Prosperity Based on Nighttime Light Remote Sensing". *Remote Sensing*, v. 12, n. 1, p. 1-22.

Change, C., Heinz, D. C. (2000). "Constrained Subpixel Target Detection for Remotely Sensed Imagery". *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, n. 3, p.1144-1159.

Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C., Ghosh, T. (2017). "VIIRS nighttimes lights". *International Journal of Remote Sensing*, v. 38, n. 21, p.5860-5879.

Elvidge, C. D., Zhinzhin, M., Baugh, K., Hsu, F. (2015a). "Automatic Boat Identification System for VIIRS Low Light Imaging Data". *Remote Sensing*, v. 7, p. 3020 – 3036.

Elvidge, C. D., Zhinzhin, M., Baugh, K., Hsu, F., Ghosh, T. (2015b). "Methods for Global Survey of Natural Gas Flaring". *Energies*. v. 9, n. 14, p. 1-15.

Fortaleza, C. M. C. B., Guimarães, R. B., Almeida, G. B., Pronunciate, M. and Ferreira, C. P. (2020). "Taking the inner route: spatial and demographic factors affecting vulnerability to COVID-19 among 604 cities from inner São Paulo State, Brazil". *Epidmiology and Infection*, v. 148, n. 118, p. 1-5.

Gomes, C., Reschilian, P. R., Uehara, A. Y. (2018). "Perspectives of the Regional Planning of Paraíba Valley and North Coast: Milestones and Institutionalization of the Metropolitan Region in the Action Plan of the Macro-metropolis of São Paulo". *Brazilian Journal of Urban Management*, v.10, n.1, p.154-171.

Instituto Brasileiro de Geografia e Estatística (IBGE). (2019). Base Cartográfica Contínua do Brasil, escala 1:250.000, version 2019. Accessed in: 28. Ago. 2020. Avaiable in: https://www.ibge.gov.br/geociencias/downloads-geociencias/cartas_e_mapas/bases_cartograficas_continuas/bc250.html.

Kyba, C., Garz, S., Kuechly, H., Miguel, A., Zamorano, J., Fischer, J., Hölker, F. (2014). "High-Resolution Imagery of Earth at Night: new sources, opportunities and challenges". *Remote Sensing*, v. 7, n. 1, p. 1-23.

Levin, N. (2017). “The impact of seasonal changes on observed nighttime brightness from 2014 to 2015 monthly VIIRS DNB composites”. *Remote Sensing of Environment*, v. 193, p. 150-164.

Levin, N. and Zhang, Q. “A global analysis of factors controlling VIIRS nighttime light levels from densely populated areas”. (2017). *Remote Sensing of Environment*, v.190, p.266-282.

Polivka, T. N., Wang, J., Ellison, L. T., Hyer, E. J., Ichoku, C. M. (2016). “Improving Nocturnal Fire Detection with the VIIRS Day–Night Band”. *IEEE Transactions on Geoscience and Remote Sensing*, v. 54, n. 9, p. 5503-5519.

Schaaf, C., Z. Wang. MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006. 2015, distributed by NASA EOSDIS Land Processes DAAC, https://doi.org/10.5067/MODIS/MCD43A4.006. Accessed 2020-09-24.

Tolimieri, R., An, M. ., Lu, Chao. (1997). “Algorithms for discrete Fourier transform and convolution”, New York, Springer.

Tuerner, S.; Kurz, F., Reinartz, P., Stilla, U. (2013). “Airborne Vehicle Detection in Dense Urban Areas Using HoG Features and Disparity Maps”. IEEE Journal of Selected in Applied Observations and Remote Sensing, v. 6, n. 6.