Poverty Mapping in Sub Saharan Africa Using Night Time Light Pollution Data

Kathleen Davies
Mechanical and Aerospace Engineering
University of Strathclyde
Glasgow, Scotland
kathleen.davies.2017@uni.strath.ac.uk

Alfred Alsop
Electronic and Electrical Engineering
University of Strathclyde
Glasgow, Scotland
alfred.alsop@strath.ac.uk

Jonathan Bowes
Electronic and Electrical Engineering
University of Strathclyde
Glasgow, Scotland
jonathan.bowes@strath.ac.uk

***Abstract*— Night Time Light data from satellite imagery can give an indication of access to electricity, a key developmental enabler as recognised by UN Sustainable Development Goal 7. Given the well documented relationship between energy access and economic development, this paper investigates the relationship between night time light pollution and poverty in Sub Saharan Africa. Night time lights could offer an alternative, quick and visually impactful method of identifying under-developed areas, and facilitate more optimal allocation of resources. This paper details the methods employed to define the relationship between poverty and night time lights and to understand any influential factors in the correlation, the results of which are also discussed. The results showed that under certain circumstances Night Time Light pollution can be used as a proxy for poverty in Sub Saharan Africa. Factors such as the resolution and zoning of data and the type of indicator used to represent development all had an impact on the accuracy of Night Time Lights as a poverty proxy.**

Keywords—GIS, remote sensing, SDG7, night time light pollution

I. Introduction

According to the World Bank, just 44.6% of Sub Saharan Africa had access to electricity in 2017 [1]. For the United Nations' Sustainable Development Goals to be achieved, this must be increased to 100% access by 2030 [2]. Mapping night time light (NTL) pollution data allows for electricity access to be mapped and therefore gives rise to a new and efficient way to identify areas which would benefit the most from investment. This would allow for better distribution of resources and enhance the lives of those living in these areas. With the deadline for achieving the United Nations'(UN) Sustainable Development Goals (SDGs) swiftly approaching [2], it is important that areas in need of development are identified and prioritised for assistance to enhance living conditions and quality of amenities. If found to be a suitable proxy for poverty, NTL mapping is an ideal candidate for the identification of underdeveloped areas, as it provides a clear visual representation of locations requiring intervention in order for human development to be increased and for the SDGs to be achieved. This paper focuses on determining if NTL data can be used as a proxy for poverty in Sub Saharan Africa (SSA) and the extent to which the two are correlated.

II. Literature review

A Global Poverty Map Derived from Satellite Data [3] uses NTL as a proxy for affluence and combines population data with NTL data to create a poverty index which is then compared to poverty data from World Development Indicators. *The Economics of Global Light Pollution* [4] establishes that population and GDP per capita both have significant impact on light pollution levels experienced by a region. *The Suitability of Different Nighttime Light Data for GDP Estimation at Different Spatial Scales and Regional Levels* [5] attempts to determine the extent to which NTL can be used as a proxy for GDP and finds that it is a more appropriate proxy at a provincial scale than on a city-wide scale in China and that resolution of data plays a key role in the correlation.

A. Light Pollution and Poverty in Provincial China

A previous study was conducted examining China on a provincial scale and documented in a paper by Wang et al [6]. This study created an Integrated Poverty Index (IPI) using Principal Component Analysis (PCA) and compared this to the average light index in each province. The average light index was calculated by first calculating the total luminance and then dividing this value by the sum of the number of pixels with a digital number of value between 1 and 63, given that each pixel in an image would have a digital number of between 0 and 63 where higher values correspond to higher intensity light. Utilising this type of representation of poverty is a good alternative to using a single development indicator as often this can be unreliable or give a narrow view of development by only considering one aspect (e.g. income). For example, as discussed in this paper, Gross Domestic Product (GDP) can give a good approximation of poverty from an economical aspect but would be less appropriate if it is desirable to consider the social aspects of poverty such as education and healthcare standards. However, these developmental indicators can be more subjective than GDP in that there is not a universal standardised way of collecting data for these indicators.

B. Light Pollution and Gross Domestic Product

A paper by Wu et al [7] explores the link between GDP and NTL data and confirms that a relation exists between the two entities but that this relation is complex in nature. This particular paper also splits light consumption into two categories: agricultural and non-agricultural. The study also highlights that there are factors which influence this correlation such as latitude, where it was discovered that higher latitudes generally consume more light and so this impacts the relationship. This indicates that poverty and NTL pollution are linked to some extent as GDP is an indicator of development. The paper found that while GDP per capita was less than 62.4 thousand international dollars (a unit which allows for comparison between countries), the

This is a peer-reviewed, accepted author manuscript of the following article: Davies , K., Alsop, A., & Bowes, J. (2021). Poverty mapping in sub Saharan Africa using night time light pollution. In 2020 IEEE Global Humanitarian Technology Conference (GHTC) IEEE. https://doi.org/10.1109/GHTC46280.2020.9342875

light consumption per capita increased commensurate with this. However, if the GDP per capita is above this value then the light consumption per capita decreases. This relationship was found in this study to follow a parabolic, inverted-U trend. However, GDP can only give an economic perspective of poverty, so it cannot be determined from this paper if NTL data can act as a suitable proxy for poverty when it is considered from a more holistic stance.

C. Night Time Lights and Poverty in Sub Saharan Africa

[8] by Noor et al also considered the use of NTL data as a proxy for poverty within SSA. PCA is used to create poverty indices as a representation of poverty to compare to NTL data. The outcomes of this study found that the NTL data could differentiate between the most and least impoverished populated areas. This work was released in 2008 and uses some data recorded as far back as 1995 and so, with the ever-changing nature of development, the results are perhaps less relevant to the current developmental state of some countries in SSA.

III. Methodology

A. Integrated Poverty Index

An IPI was created by applying PCA [9] to a set of development indicators [6]. PCA is a statistical analysis method which reduces the quantity of variables in a dataset while preserving as much of the data as possible [9]. A varied cross section of indicators was selected in order to encapsulate the three pillars of development – health, education and income. This is similar to the way in which Human Development Index (HDI) attempts to capture poverty by combining Life Expectancy, Education and Income Indices [10]. Data for as many indicators as possible was collated for completeness and input into a correlation matrix. This allowed for redundant indicators to be highlighted and removed, whilst retaining as much information as possible. If a pair of indicators shared a correlation coefficient of greater than 0.7, indicating the two are strongly correlated and so both convey a large amount of same information about the poverty level, one was eliminated from the dataset. In a strongly correlated pair, the indicator with the most information missing was eliminated. Any remaining gaps in data were filled using K-Nearest Neighbour [11]. The final IPI included data for the following development indicators: water stress [12], government efficiency [12], food security [12], homicide rate [12], unemployment [12], literacy rate [1], life expectancy [1], access to electricity [1], gross national income (GNI) per capita [1], measles immunisation rate [1], and individuals using the internet [1]. Originally income inequality [12], was also included in the IPI. However, it was found to be difficult to capture via NTL data, as it is possible that both the poor and affluent could be living in the same area, particularly in urban locations. PCA was then employed through the use of a modified version of a free to use, online code [13] to return a unique value for each country in SSA. The higher the IPI, the more developed the country. The IPI for each country was then compared to the mean value of NTL to establish if a correlation exists.

B. Night Time Light Data

The Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) NTL data [14] was collated and had already undergone an outlier removal process with background lights set to zero [14]. To make the data more robust, areas of zero population according to the EU Commission [15] were removed so that any data present in non-populated areas did not influence the mean value. The NTL data was also divided by the population [15] using Raster Calculator to give an approximation of NTL per capita. This data should only be treated as an approximation as, for example, a street with 5 houses containing 5 people could have the same amount of street lighting (and therefore light pollution) as a street with 5 flat blocks containing 500 people. As well as these processes being conducted for SSA, the methods were also repeated for Malawi on a regional level.

C. Assessment of Correlation

NTL data was put into a correlation matrix with the IPI values and the results graphed to assess the strength and nature of any correlation. NTL data was input into a correlation matrix and graphed against HDI [16] and a multidimensional poverty index (MDPI) [16] on a global level to understand the impact that indicator type has on the relation between development and NTL.

IV. Results and Discussion

A. Considering Sub Saharan Africa

The outcome of applying the PCA aligned with expectations. The highest scoring countries were Mauritius, South Africa and Botswana respectively while the least developed countries were Chad, South Sudan and the Central African Republic respectively. The linear correlation between the IPI and NTL data for SSA was positive but fairly weak, giving a value of 0.334 for the correlation coefficient. The mean NTL data for SSA was also graphed against the IPI data as shown in Figure 1.

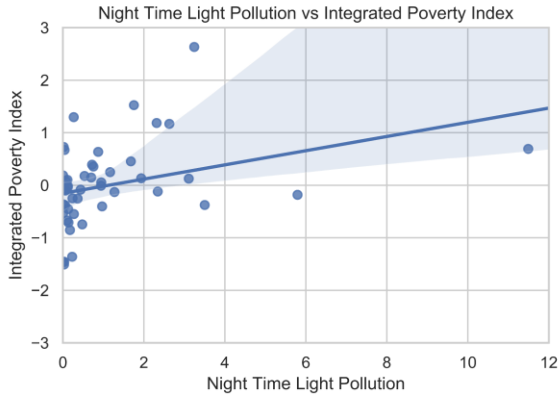

Figure 1: Graph with Regression Line of Night Time Lights [14] vs Integrated Poverty Index in Sub Saharan Africa

Figure 1 clearly shows that a relationship exists between NTL and poverty in SSA, however, the relation is clearly not linear and could instead be exponential. It would be difficult to accurately fit any higher order function to this data given the scarcity of data points at higher scores. In general, it can be seen from the plot in Figure 1 that a higher value for mean NTL goes hand in hand with a higher IPI value. This requires further research, such as an investigation into public lighting policies and how these vary from country to country, to fully understand.

B. Isolating Gross Domestic Product

Isolating a single development indicator, GDP [17], resulted in a strengthened correlation coefficient of 0.866 between the mean NTL data and development. This is likely a result of the fact that the IPI contained a broad cross section of indicators which all have different data collection methods. The way in which the data is collected for some of these development indicators may vary from country to country while GDP is a standardised indicator in that it has a standardised method of collecting information. This result suggests that using indicators where data collection methods can vary may result in a weaker correlation with NTL than using indicators that have a standardised method such as GDP [17] and HDI [16].

C. Isolating Gross Domestic Product in Malawi

When the area of interest was localised to consider only Malawi on a county basis, the correlation between NTL and GDP strengthened again to a value of 0.922. This implies that the method of zoning has an impact on the suitability of NTL as a proxy for poverty and that the localisation of data, and therefore finer resolution of data, has resulted in this increase in correlation coefficient. It appears that the more localised the area of interest, the more robust the correlation. This could mean that NTL becomes a more suitable proxy for poverty on smaller, regional scales if high accuracy is desirable. Additionally, it is likely that the strength of the corrrelation between NTL and poverty varies from country to country. Differences between countries like their public lighting policies and the proportion of land considered rural are most likely influential factors in the relationship shared by NTL and development. Through mapping both the GDP and NTL in

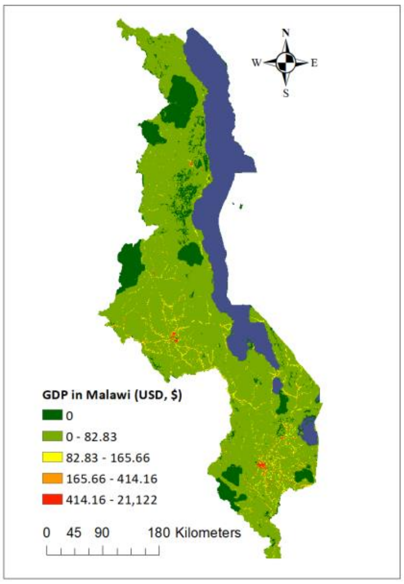

Figure 2: Gross Domestic Product [17] in Malawi, thematic layers include water bodies [20] and borders [19]

Malawi, a visual representation of the correlation was obtained. Figure 2 displays the GDP in Malawi and Figure 3 shows the NTL in Malawi. As is evident, the maps share significant similarities. This shows that when a standardised development indicator is used on a localised level, NTL can be a good proxy for poverty. The two maps marry well together and show knots of activity in the same locations, indicating a correlation exists between NTL and developement. Further work is needed to understand the extent to which both NTL data and development are related to population density, and how this may influence the use of NTL for developmental proxies.

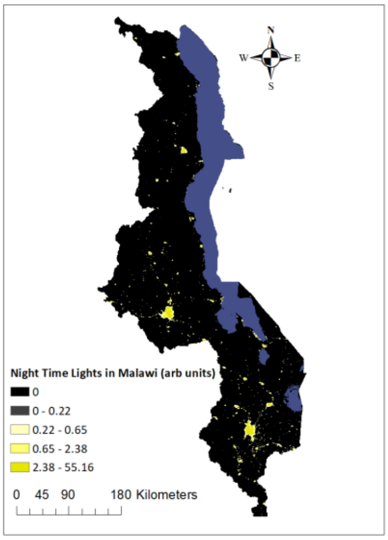

Figure 3: Night Time Lights [14] in Malawi, thematic layers include water bodies [20] and borders [19]

D. Considering the Globe

Broadening horizons to encompass the whole world, the mean NTL data was compared to the MDPI and HDI. Figure 4 plots the relation between NTL and the MDPI and clearly demonstrates that a relationship exists between NTL and poverty. However, this relation is complex so requires further work such as research on public lighting policies and how these differ from country to country and the impact of land use and quantity of rural land on the relationship between NTL and poverty.

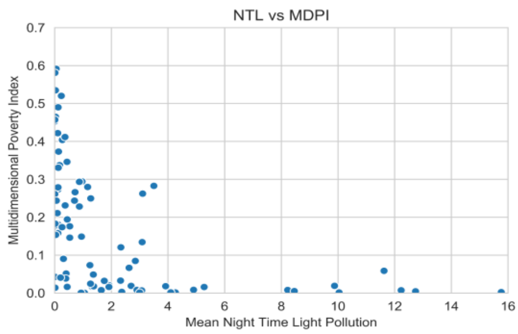

Figure 4: Graph of Global Night Time Lights [14] vs Multidimensional Poverty Index [16]

The correlation between HDI and NTL is displayed in Figure 5 and shows that the two entities are associated and share a non-linear relationship. This graphical representation of results in Figure 5 is of particular interest due to its striking resemblance to a well-known relationship between HDI and per capita energy consumption [18], shown in Figure 6. This demonstrates that the relationship HDI shares with NTL is very similar to that which it shares with energy consumption. Furthermore, there is the possibility that the amount of land considered to be rural has an impact on the suitability of NTL as a proxy for poverty. Further research could be conducted to assess the impact of rurality on the correlation between NTL and development.

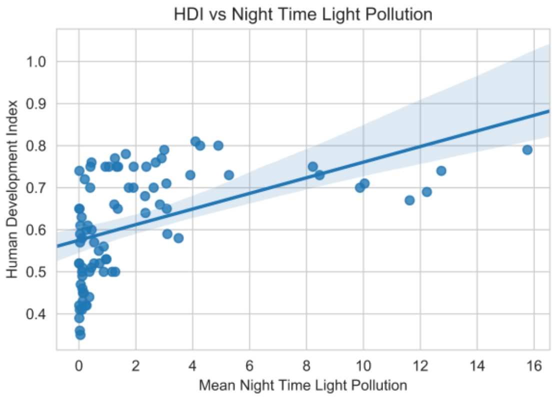

Figure 5: Graph of Global Night Time Light Pollution [14] vs Human Development Index [16] with Regression Line

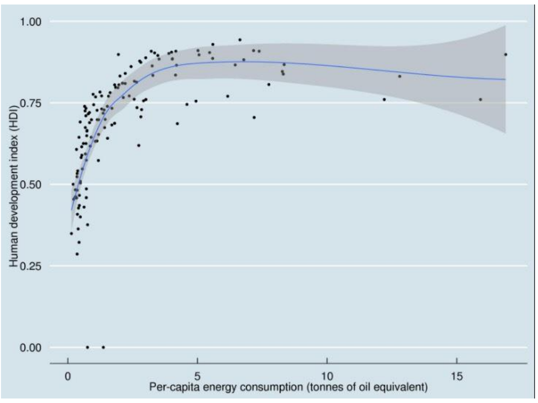

Figure 6: Graph of Per-Capita Energy Consumption vs Human Development Index [18]

Conclusion

This paper demonstrates that NTL may be a suitable proxy for poverty under certain conditions. The strongest correlation coefficient was found when data was examined within a single country, using a single development indicator which used a standardised data collection method. The zoning and resolution of data are very important considerations if NTL is to be used as a proxy for development. This is demonstrated by the increase in correlation coefficient when a single, standardised indicator was isolated and compared to NTL on a regional level rather than a national one. When the IPI was compared to NTL in SSA on a country wide basis, the correlation coefficient was found to be 0.33, indicating a relatively weak correlation or that the relationship may be non-linear. However, when GDP alone was compared to NTL in Malawi on a county basis, this coefficient strengthened to 0.92 suggesting that the two entities share a very close relationship. If high accuracy poverty estimation is desired then results suggest that NTL could be used to evaluate poverty if there is a high resolution of data, considered on a localised scale and using a standardised indicator. Further work is required to fully comprehend the complex relation NTL shares with development and to determine the influential factors in this correlation

Acknowledgment

This work was funded by the EPSRC (project EP/R007756/1) as part of the Strathclyde Game Changers program. With thanks also to Dr Edward Hart.

References

[1] The World Bank, "Data Bank," 2020. [Online]. Available: https://databank.worldbank.org/source/world-development-indicators.

[2] United Nations, "Sustainable Development Goal 7," 2019. [Online]. Available: https://sustainabledevelopment.un.org/sdg7.

[3] C. D. Elvidge, P. C. Sutton, T. Ghosh, B. T. Tuttle, K. E. Baugh, B. Bhaduri and E. Bright, "A global poverty map derived from satellite data," *Computers and Geosciences,* no. 35, pp. 1652-1660, 2009.

[4] T. Gallaway, R. N. Olsen and D. M. Mitchell, "The economics of global light pollution," *Ecological Economics,* no. 69, pp. 658-665, 2010.

[5] Z. Dai, Y. Hu and G. Zhao, "The Suitability of Different Nighttime Light Data for GDP Estimation at Different Spatial Scales and Regional Levels," *Sustainability,* vol. 9, no. 305, 2017.

[6] W. Wang, H. Cheng and L. Zhang, "Poverty assessment using DMSP/OLS night-time light satellite imagery at a provincial scale in China," *Advances in Space Research,* pp. 1253-1264, 2012.

[7] J. Wu, Z. Wang, W. Li and J. Peng, "Exploring factors affecting the relationship between light consumption and GDP," *Remote Sensing of Environment,* pp. 111-119, 2013.

[8] A. M. Noor, V. A. Alegana, P. W. Gething, A. J. Tatem and R. W. Snow, "Using remotely sensed night-time light as a proxy for poverty in Africa," *Population Health Metrics,* no. 6, 2008.

[9] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A ,* vol. 374, no. 2065, 2016.

[10] M. Roser, "Human Development Index (HDI)," Our World in Data, November 2019. [Online]. Available: https://ourworldindata.org/human-development-index. [Accessed 19 March 2020].

[11] L. E. Peterson, "K-nearest neighbor," Scholarpedia, 2009. [Online]. Available: http://scholarpedia.org/article/K-nearest_neighbor. [Accessed 19 March 2020].

[12] European Commission, "Global Conflict Risk Index," European Commission, 7 October 2019. [Online]. Available: https://ec.europa.eu/knowledge4policy/dataset/ds00151_en. [Accessed 19 March 2020].

[13] M. Heo, "python_tutorial/data_science/pca/PCA.ipynb," 26 November 2017. [Online]. Available: https://github.com/minsuk-heo/python_tutorial/blob/master/data_science/pca/PCA.ipynb. [Accessed July 2019].

[14] National Oceanic and Atmospheric Administration, "Version 1 VIIRS Day/Night Band Nighttime Lights," 20 June 2019. [Online]. Available: https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html.

[15] European Commission, "GHS POPULATION GRID (LDS)," 2019. [Online]. Available: https://ghsl.jrc.ec.europa.eu/ghs_pop.php.

[16] United Nations, "Human Development Reports," 2018. [Online]. Available: http://hdr.undp.org/en/data.

[17] Global Risk Data Platform, "Gross Domestic Product (GDP)," 2012. [Online]. Available: https://preview.grid.unep.ch/index.php?preview=data&events=socec&evcat=1&lang=eng.

[18] R. Wilson, "Energy versus well being: How much energy do we need to consume?," 8 February 2015. [Online]. Available: https://carboncounter.wordpress.com/2015/02/08/energy-versus-well-being-how-much-energy-do-we-need-to-consume/. [Accessed July 2019].

[19] GADM, "GADM data," 6 May 2018. [Online]. Available: https://gadm.org/data.html.

[20] Global Land Cover by National Mapping Organizations, "GitHub," 20 February 2017. [Online]. Available: https://github.com/globalmaps/gm_lc_v3. [Accessed 24 March 2020].