Night Lights and Noisy Data - Using Machine Learning to Better Detect Human-Generated Light

Johannes Coetsee[a]

[a]*Stellenbosch University*

Abstract

Abstract to be written here. The abstract should not be too long and should provide the reader with a good understanding what you are writing about. Academic papers are not like novels where you keep the reader in suspense. To be effective in getting others to read your paper, be as open and concise about your findings here as possible. Ideally, upon reading your abstract, the reader should feel he / she must read your paper in entirety.

Keywords: Remote Sensing, Night Lights, Random Forest

1. Introduction

The use of remote sensing data, and more specifically, satellite nighttime light data, presents potential for new and diverse applications in socioeconomic research. Nightlight data remains a largely objective measure, and is thereby suitable to use as a proxy in a broad array of studies that require the usage of potentially unreliable or otherwise lacking data. This advantage is especially pertinent in parts of the developing world, where socioeconomic research can prove most beneficial. There exists different night lights products which can be utilized towards this end, the most common of which is the 'Stable Lights' product, derived from the Defense Meteorological Satellite Program's (DMSPs) Operational Linescan System (OLS). This paper emphasises the usage of this product specifically focusing on its shortcomings. Most prominently, DMSP-OLS has difficulty in separating background noise from night lights generated from human-generated light, especially in areas that display lower levels of night light intensity. This presents an obvious problem: analyses that attempt to use Stable Lights as a proxy for economic activity, for instance, would exaggerate or understate economic activity in these low-luminous areas.

This paper attempts to address the challenge of inaccurate measurement of night lights by applying a filtering technique to identify and separate nightlights emitted by humans from those emitted by

Email address: `19491050@sun.ac.za` (Johannes Coetsee)

anything else. This filtering process is based on the methodology for deriving the 'Local Human Lights' product by Määttä & Lessmann (2019), and relies on a Random Forest (RF) Machine Learning algorithm for classification.

2. Explication of the Problem

2.1. Stable Lights and Economic Activity

The most prominent difficulty, however, relates to the amount of noise in the lower end of the light distribution due in part to the blooming effect mentioned above. Standard practice using the stable lights data set is to discard these values from analysis, thereby removing a large proportion of cell observations.

DMSP-OLS

2.2. Problems with Stable Lights

3. Method and Data

The filtering process used by Määttä & Lessmann (2019) necessitates the following steps:

-
-
-
-
-

3.1. Data

Although our methodology largely follows that presented by Määttä & Lessmann (2019), there are some distinct differences.

- Size of area
- Jitter size

- Probability of human-generated
-

3.2. Algorithm

4. Results

5. Discussion

This paper explicated on the usage of a filtering methodology proposed by Määttä & Lessmann (2019) with which to separate background noise from human-generated nightlight data.

Our results echo those by Määttä & Lessmann (2019): the RF algorithm introduces great improvements in classification accuracy and thus greater accuracy in filtering out background noise from the 'Stable Lights' product. This allows the researcher to do away with the need for quick-and-easy type fixes to noisy data at the lower end of the luminosity distribution.

However, it is important to note what this method does not achieve. For instance, the 'Human Lights' product does not address the issue of blooming or oversaturation at the high end of the luminosity spectrum. Likewise, its spatial resolution remains low in comparison to more modern products such as the Visible Infrared Imaging Radiometer Suite (VIIRS), and it is recommended to use these products rather than the DMSP-OLS 'Stable Lights' or 'Human Lights' products if possible.

References

10 Määttä, I. & Lessmann, C. 2019. Human lights. *Remote Sensing.* 11(19):2194.

Appendix

Appendix A