Using satellite data to track socio-economic outcomes: a case study of Namibia

THOMAS FERREIRA

Stellenbosch Economic Working Papers: WP12/2018

www.ekon.sun.ac.za/wpapers/2018/wp122018

June 2018

KEYWORDS: Satellites, Night Lights, Normalised Differenced Vegetation Index, Agriculture, Poverty, Namibia
JEL: I32, O13, Q56

ReSEP (Research on Socio-Economic Policy)
http://resep.sun.ac.za

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
SOUTH AFRICA

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

www.ekon.sun.ac.za/wpapers

Using Satellite Data to track Socio-Economic Outcomes: A Case Study of Namibia

Thomas Ferreira

Abstract

Efforts to improve the livelihoods of the poor in sub-Saharan Africa are hindered by data deficiencies. Surveys on socio-economic outcomes, for example, are generally conducted infrequently and are only statistically representative for relatively large geographic areas. To overcome these data limitations, researchers are increasingly turning to satellites which capture data for small areas at high frequencies. Night lights satellite data has particularly drawn interest and growth in lights have been shown to be a useful proxy for GDP growth (Henderson et al., 2012). However, in poor agricultural regions, night lights data might be less useful in explaining variation in socio-economic outcomes because such regions are generally under-electrified. Daytime satellite data measuring land use and vegetation quality, have been used to model socio-economic outcomes across regions, but no studies have explored whether daytime satellite data can be used to track welfare longitudinally. This paper argues that indicators of vegetation quality can be used to track welfare over time in agriculturally dominant areas. Such indicators are used extensively to predict agricultural yields and thus should correlate with welfare, as agriculture is an important source of income. This paper presents results from a small study in Namibia, that explores whether this is the case. Firstly, it is shown using classification of cropland, that daytime satellite data can identify areas of economic activity where night lights cannot. Secondly the relationship between vegetation quality and welfare is studied. Cross-sectionally, increases in vegetation quality correlate negatively with welfare. This is expected as the poor are more likely to live in rural areas. Within rural areas, however, vegetation quality correlates positively with welfare. This study thus supports the hypothesis that satellite based indicators of vegetative health can be used to track welfare over time in areas where night lights are not present.

Thomas Ferreira
Department of Economics
University of Stellenbosch
Private bag X1, 7602
Matieland, South Africa
E-mail: thomasf@sun.ac.za

The author is grateful to comments provided by members of the ReSEP team with special mention for Servaas van der Berg and Dieter von Fintel. Valuable comments were also received by Derek Yu, Janneke Pieters, Nick Vink and at the 3rd ISIbalo Centre for regional and urban innovation and statistical exploration (CRUISE) conference (2017).

1 Introduction

A lack of household survey data limits what can be said about economic development in much of sub-Saharan Africa (SSA). Two ways in which household survey data limits analysis, is the frequency of its release and the geographic scale at which it is representative. For example, surveys used to measure poverty and inequality, generally called Income and Expenditure Surveys, are only undertaken around every 5 years and are not longitudinal. These surveys are also only representative at the first sub-level of geographic administration in most countries.

These factors limit what researchers can say about poverty changes over shorter time periods, and also about the geographic location of poor households. Small-area estimation techniques, such as the method proposed by Elbers et al. (2003) do allow researchers to map poverty at smaller geographic areas, but the method can only map poverty to smaller areas for the same time periods as household surveys, with the added requirement that census data (which is more spatially disaggregated) be released around the same time.

Realising these limitations, researchers have started to investigate alternative data sources, such as mobile phone data, as additional sources to support analysis. Blumenstock et al. (2015) have shown, for example, how people's record of cellphone data usage can be used to predict poverty.

Another alternative is to use satellite data. Satellite data is recorded at high frequency and at spatial resolutions usually much smaller than the smallest statistically representative area in a household survey or census. Currently there are many satellite series available free of charge that go back for decades. These properties make them highly desirable. The process of capturing data with satellites falls under the broader science of remote sensing and this paper will refer to such data as remotely sensed (RS) data.

In terms of RS data, the night lights series is well-known in economics, and changes in the amount of light emitted in an area have been shown to correlate well with changes in GDP per capita (Henderson et al., 2012). It has been noted, however, that the night lights data adds more value as a proxy in countries with low levels of statistical capacity (Chen and Nordhaus, 2011; Henderson et al., 2012). It is not unreasonable to assume that countries with low levels of statistical capacity are also under-electrified. Thus, while at national level, night lights might correlate with GDP, the ability of night lights to proxy any economic activity at smaller geographic levels within these countries needs to be questioned.

The question that arises then, is whether there is a RS data source that can be used in under-developed areas where no light is observed? Research has shown that daytime RS data does hold promise. Weeks et al. (2012), Watmough et al. (2013) and Watmough et al. (2016) have all noted that measures of land cover and vegetation quality correlate with socio-economic outcomes between areas. Combinations of daytime RS data and night lights have also been used to map poverty in a number of African countries (Jean et al., 2016) and GDP in Guangdong, China (Cao et al., 2016).

There is, however, a fundamental shortcoming in the literature which this paper addresses. No research has focussed on the relationship between daytime RS data and socio-economic outcomes over time. This is an important omission as cross-sectional correlations can be misleading in making causal claims. Selection effects, for example, need to be considered. In SSA, the poor are more likely to be located in rural areas where vegetation quality is higher. However, it cannot be assumed that increases in vegetation quality lead to increases in poverty over time. Differences in agricultural practices, agro-ecological zones and general use of land also implies that relationships between the remotely sensed environment and socio-economic outcomes would be heterogeneous across regions. Henderson et al. (2012) note regional heterogeneity as the reason for only using growth formulations in examining the relationship between night lights and GDP. The establishment of plausible relationships between daytime RS data and welfare over time, would imply that welfare can be tracked within smaller regions and over higher frequencies than household surveys allow and can do this where night lights are not present.

This paper proposes that measures of vegetation quality, derived from RS data, can be used to track changes in socio-economic outcomes over time in rural regions in SSA. The reasoning is that most under-developed, inhabited areas are dominated by smallholder agriculture, and RS based vegetation indices are widely used to predict agricultural yields (Lobell, 2013). This implies that RS data can proxy for an important source of economic activity in such regions. Thus, welfare should increase in regions as vegetation quality improves.

This paper shows how existing data sources can be extended to inform on welfare. It presents results from a small-scale study that explores these ideas. Namibia was selected as the country of study because it has large regional inequalities. It has well-developed urban regions, as well as underdeveloped agricultural regions where poverty is prevalent and electrification rates are low.

The paper firstly shows, using classification of cropland, that in rural agricultural areas, daytime satellite data can indeed identify economic activity where night lights cannot. Secondly, using regression analysis it explores the correlation across time and space between vegetation quality and welfare. The results show that welfare correlates negatively with vegetation quality across geographic areas. Within rural areas, however, vegetation quality correlates positively with welfare over time. This supports the idea that RS based vegetation indices can track welfare over time and map welfare where night lights cannot.

The next section gives a review of the use of satellite data in economic studies. Section 3 discusses the data and methodology used. Section 4 presents the results. Section 5 concludes.

2 Economic Studies and Satellite data

The appeal of RS data was noted in the introduction - broad coverage at high spatial resolution and high frequency. Satellites orbit the earth and capture light reflected to it from the earth's surface. This process of capturing light from a distance is known as remote sensing. The light captured is stored and digitized to create an image with geo-referenced points that contain data of some sort.

Daytime RS data captures electro-magnetic radiation (EMR) from the sun, reflected back off the earth's surface. The EMR is captured by sensors on the satellites which store them in different spectral bands depending on their wavelengths. These sensors generally capture a broader spectrum of EMR than what the human eye can capture.[1]

Each object on the earth's surface absorbs and reflects different wavelengths and these properties can allow researchers to identify different objects based on their spectral signatures - how they reflect EMR. Vegetation for example, absorbs EMR in the red band and reflects EMR in the infra-red band. Ratios of the EMR reflected in these two bands can also give indications of the health of vegetation. One popular measure of vegetative health, the Normalised Differenced Vegetation Index (NDVI), is discussed in Section 3.

Satellites capturing images of the earth at night work in the same manner, with the exception that much of the light captured is artificially created by humans and not reflections from the sun. The night lights series made available by the Defence Meteorological Satellite Program (DMSP) is a free, user-friendly source and is widely used in economic studies.

[1] The human eye captures the visible spectrum which includes the blue, green and red bands of EMR.

2.1 The Use of Night Lights Data

Elvidge et al. (1997) showed that the area lit up in countries was positively correlated with gross domestic product (GDP) and electricity consumption, suggesting that the data could be used to model GDP. The correlation between lights and GDP at national level led researchers to conclusions that the data could be used to obtain estimates of GDP at sub-national levels. Sutton and Costanza (2002) modeled the relationship between total light emitted in a country and GDP at the national level, and created a 1km^2 grid of GDP for the world. Ebener et al. (2005) and Sutton et al. (2007) showed that night lights can be used to model GDP per capita at sub-national levels.

Chen and Nordhaus (2011) compared sub-national GDP estimates based on night lights with estimates based on the G-Econ data which is a data product with small-area GDP estimates.[2] They found that night lights data added valuable information in countries with low levels of statistical capacity, but not in countries with good capacity. This was due to the fact that measurement error in the night lights data is larger than standard economic data.

Henderson et al. (2012) showed how *growth* in luminosity was a good proxy of GDP *growth* in the long- and short-run. They then provided a revised measure of income growth as a combination of national accounts data and predicted income growth from night lights. They calculated the optimal weights to be placed on each measure. Similar to Chen and Nordhaus (2011), they found that in countries with good statistical capacity, night lights did not add value to the national accounts data, but in countries with low capacity, the optimal weighting was roughly an equal combination of national accounts and night lights.

One example of a study exploiting the correlations between GDP and night lights is Villa (2016). He used the night lights as a sub-national proxy of economic activity to measure the effect of a cash transfer programme on economic growth in Colombia. The program was phased in over time at the municipal level allowing for the implementation of a difference-in-difference identification strategy. Night lights data was used to provide estimates of GDP at municipal level.

Pinkovskiy and Sala-i-Martin (2016) used night lights as an independent source of data to study whether national accounts or survey data provide more reliable estimates of poverty. They estimated a proxy for income by calculating the optimal weights to be placed on national accounts and survey estimates, by exploiting the assumption that the measurement error in

[2] For more information on the G-Econ data visit http://gecon.yale.edu/.

night lights is unrelated to the errors of national accounts and survey data. They concluded that national accounts provide more reliable estimates of true income than survey data.

Night lights have also been used to create and map global poverty indices. Elvidge et al. (2009) estimated a global poverty map using night lights data. They created a wealth index based on gridded population data and night lights. Elvidge et al. (2012) developed a Night lights Development Index (NDLI) combining night lights and gridded population data. Anthony (2015) also calculated regional inequality measures by combining night lights data, with gridded population data and the G-Econ data.[3]

2.2 The Use of Daytime Remotely Sensed Data

Remotely sensed data has been used substantially in the studies on the trade-off between development and the environment, commonly referred to as the environmental Kuznets curve. These studies combine data on economic growth or changes in welfare with satellite data measuring environmental outcomes such as the area of forest cover. Pfaff (1999) has studied this for the Brazilian Amazon and Foster and Rosenzweig (2003) as well as Bhattacharya and Innes (2012) have done such studies in India. For this paper, the findings are not of concern, but the measures from satellite data, in combination with survey data, are.

Pfaff (1999) used Landsat images classified by Skole and Tucker (1993). The images were classified into different categories of forest using visual interpretation. This was combined with Brazilian county level data obtained from a number of sources to estimate relationships between deforestation and socio-economic variables.

Foster and Rosenzweig (2003) combined census, household survey and RS data spanning 29 years to study the relationship between income and forest cover around villages in India. They used the NDVI to measure forest cover in a 10km radius from villages. The proportion of pixels with NDVI values above 0.2 multiplied by the average NDVI of those pixels was used as their measure of forest cover. Bhattacharya and Innes (2012) studied the relationship using district level data. For environmental quality they also used NDVI. They used average NDVI over the districts as a measure of total biomass. They also constructed a proxy for forest cover based on the number of pixels that are above a certain NDVI threshold on average over a month.

[3] For gridded population data the LandScan series was used in all three studies. The data is created by combining census population numbers with spatial data to create population counts which reflect where people are located during the day. This is referred to as the ambient population.

Klemens et al. (2015) tested whether RS data can help explain poverty at small-area level. They used local-level census and survey data to model poverty rates at local administrative levels. They then added variables derived from satellite data to explore whether they help explain local poverty rates. Night lights were found to be significant and add explanatory power in explaining variation in rural but (surprisingly) not urban poverty rates. Measures of local vegetation were insignificant, however. For their measures of vegetation, they used the Leaf Area Index - a measure of leaf coverage - and the fraction of photosynthetically active radiation (FPAR) - which measures how much incoming radiation from the sun is absorbed by plants.

Imran et al. (2014) use RS data to map poverty in Burkina Faso. They started at the observation that poor households in rural areas produce too little to meet their consumption needs. Using household level data, representative at community level, they created a communal level asset index from variables related to food production. They then identified variables from available remote sensing data that represent agro-ecological stressors which affect food production. Variables included NDVI, rainfall data, topographical data, length of the growing season, livestock data, distance to markets and population density. At community level, these variables were regressed on a communal asset index to model welfare. This model was used to predict pixel level welfare in the country.

Weeks et al. (2012) explore the use of RS data to map health outcomes within cities. They argue that health outcomes are correlated within neighbourhoods, and that by combining data on health outcomes with neighbourhood characteristics from remotely sensed data, health outcomes can be mapped for cities. They used Ghana census data for Accra, with geo-references to neighbourhoods in the city. They created a housing index for each neighbourhood based on the census data. They then used RS data to map land cover in the city. They found that neighbourhoods with the lowest quality housing generally also had the least vegetation cover and the worst health outcomes while neighbourhoods with high quality housing had more vegetation cover and better health outcomes. This, they argue, implies that RS data can be used to identify areas of cities with bad health outcomes.

Watmough et al. (2013) study whether RS data could be used to predict female literacy levels in Assam, India. They classified RS images into nine different classes. Census data was used to estimate the level of female literacy in villages. They found that female literacy was positively correlated with the amount of woodland in an area and negatively correlated with the amount of winter cropland. Watmough et al. (2016) expanded on the previous analysis by estimating the relationship between remotely sensed data and a general welfare index derived from variables in census data. Land cover classes and NDVI were regressed on welfare. Woodland cover was again positively correlated with welfare. Surprisingly, NDVI was not an important factor but the authors note that this was possibly due to a high correlation with woodland cover.

Morikawa (2014) used daytime RS data to explore the effects that a poverty alleviation project had on communities and the environment. Based on the idea that poverty and land degradation reinforce each other, he used NDVI to evaluate the environmental effects of the project.

Researchers have also started to combine daytime and night lights imagery. Jean et al. (2016) used a machine learning technique to combine daytime RS data and night lights data to explain a large proportion of regional variation in poverty for a number of countries in SSA. Cao et al. (2016) produced a GDP map of Guangdong, China, using night lights to proxy the secondary and tertiary sectors, and land use data for the primary sector. Additionally they used NDVI data to limit the effect of light overflow in the night lights data. Without addressing light overflow, the extent of urban areas, based on night lights, is overestimated. (Small et al., 2011).

3 Data and Methodology

The reviewed literature has shown that daytime RS data can be used to predict socio-economic outcomes. However, the research fails to address relationships between socio-economic outcomes and daytime RS data over time. Thus, while the research can help to show where poverty is located, it cannot say how welfare will change as the RS variables change. This is an important omission as regional heterogeneity and selection effects could lead to misleading correlations and false causal interpretations. This argument was also made with regards to night lights by Henderson et al. (2012). To estimate the relationship between night lights and GDP, they argued that cross-country heterogeneity in, for example, the nature of production and lighting technology made them stick to growth formulations in their models.

This paper addresses this limitation by exploring whether RS data can be used to proxy welfare, with the added focus of variation over time. For this, it is argued that there is potential for the use of remotely sensed indicators of vegetation quality in agriculturally dominant regions. This argument is based on the fact that indicators of vegetation quality are used substantially to predict agricultural harvests (Lobell, 2013). Welfare should thus increase with increases in vegetation quality over time because livelihoods are dependant on it. In SSA, a large proportion of the population lives in rural areas, and thus RS based vegetation indicators can hold potential to track welfare for a large proportion of the population. It can also do so in under-electrified regions.

The added benefit of vegetation indices is that there are many user-friendly, easily obtainable sources available. If indeed daytime remotely sensed data correlates well, it holds potential for tracking outcomes at small-area level over time. This small scale-study explores whether RS based vegetation indices hold promise. If the results hold, a larger scale study should be undertaken.

3.1 Methodology

The analysis proceeds in two steps. Firstly, visual exploration is used, to identify whether daytime RS data does capture economic activity in an under-developed, agriculturally dominant area better than night lights. This is done by comparing the amount of cropland observed from daytime RS data with the amount of night light observed in both an industrialised developed region, and a under-developed agriculturally dominant region.

After establishing that daytime RS data does indeed capture economic activity in under-developed areas better than luminosity data, it is explored whether correlations exist between indicators of vegetation quality and welfare. Regional poverty rates and average expenditure levels are regressed on vegetation quality both cross-sectionally and over time using regional fixed effects. The use of fixed effects implies that the estimates will measure the relationship using deviations over time. It is expected that, cross-sectionally, higher levels of vegetation quality should be negatively correlated with welfare. This is because poverty is higher in rural areas. However, using regional fixed effects, which implies only using variation in vegetation quality over time within regions, the correlation between welfare and vegetative health should be positive. Thus, as vegetation quality improves, which implies higher levels of agricultural production, welfare should increase.

It would be ideal to also explore the vegetation health of land specifically identified as cropland. The process of land classification is tedious, however. Raw RS images need to be processed in a number of ways before they can be classified, and classification itself is also a timely proses. For these reasons vegetation quality for all land in an area is used.

The focus of the study is to establish whether indicators of vegetation quality from RS data correlate with welfare and not to determine the model with the highest explanatory power to predict welfare with vegetation quality. As noted, the relationship will also be heterogeneous over different ecological zones which implies that the optimal measure would differ across regions. The discussion proceeds by describing the study area, followed by the data sources and vegetation indicators used.

3.2 Study Area

Namibia was selected as a country of study as it has large regional differences in terms of socio-economic outcomes.[4] This presents the perfect opportunity to highlight that cross-sectional and time-series relationships can differ. Namibia has well-developed urban regions and under-developed rural regions. This should show that vegetation quality correlates negatively with welfare across regions. This should not be the case within regions over time.

Namibia is one of the least densely populated countries in the world and is the driest country in sub-Saharan Africa (Devereux and Naeraa, 1996; Fara, 2001). The north of Namibia is more densely populated than the rest of the country, with 6 regions covering an area of less than 15% (Devereux and Naeraa, 1996) that held more than 50% of the population in 2011 (Own calculations using Census 2011). The areas are Omusati, Oshana, Ohangwena, Oshikoto, Okavango and Caprivi (see Figure 1). The country has a diverse population of roughly 2.5 million people and is defined as an upper middle income country by the World Bank (2015). This hides the fact that it is one of the most unequal countries in the world, with a large share of the population impoverished due to a history of colonialism and racial discrimination.

[4] See Section 4.1 for descriptive statistics of regional differences.

Figure 1: Regions of Namibia

3.2.1 Colonial Rule

After nearly 90 years of colonial rule, Namibia gained independence in 1990. The period of colonial rule was violent. After the Herero and Nama wars between 1904 and 1908 Germany gained greater control of Namibia. German rule ended during World War 1 after occupation by South Africa in 1915. South Africa gained administrative control in 1920 (World Bank, 2008). In 1966 South Africa rejected a United Nations (UN) mandate to place the country under a trusteeship arrangement (thus continuing to rule the country) and as a result the UN cancelled South Africa's mandate.

In that same year the South West Africa People's Organization (SWAPO) declared war to free Namibia from colonial rule. SWAPO established bases in southern Angola after Angola gained independence in 1975. The organisation used guerilla tactics and northern Namibia became the centre of a war between South Africa and SWAPO. A UN-sponsored peace deal was finally reached in 1989 when Cuban troops who were supporting SWAPO left Angola and South African troops left Namibia. Elections were held in 1989, which SWAPO won easily. A constitution was adopted in February 1990 and in March independence was granted to the country (World Bank, 2008).

Under South African rule and just as in South Africa, apartheid was also enforced in Namibia.

Under German administration the country was divided into two areas - the "Police Zone" and the rest. The "Police Zone" was the area of land that was vacated by white farmers. Its northern border stretched across the north-centre of the country from east to west. Essentially only areas south of Namibia's northern border were excluded from the "Police Zone". Under South African rule these policies were continued. In the "Police Zone" the government created "native reserves" or "homelands" where many natives were sent after being dispossessed of their land in favour of white farmers. The government greatly supported white farmers by, for example, granting loans, drilling boreholes, and assisting with drought relief. On the other hand, the government did little to support farmers in the "homelands". In 1968 more segregation laws were instated with the government establishing 10 "homelands". These areas were also granted self-rule. These included areas north of the police zone (Odendaal, 2011). Werner (1993) points out that the reserves were essentially created to provide labour to the colonial economy.

3.2.2 Current Economic Climate

The policies led to a highly unequal society in Namibia. The north of the country, which contains a large share of the population, is dominated by smallholder agriculture on communal lands while urban centres such as Windhoek (in Khomas), Walvisbay and Swakopmund (both in Erongo) are well developed and industrialised. Poverty is lower in the urban centres than in the communal lands. In 2009/2010 the rural poverty rate was 37.4% while the urban poverty rate was 14.6%. Since independence, poverty has declined in the country. Between 2003/04 and 2009/10, poverty decreased by 9 percentage points to 28.7% (NSA, 2012).

3.3 Data

3.4 Socio-Economic Data

Two sources of socio-economic data are used in the analysis. The first is Namibia's Household Income and Expenditure Surveys (NHIES). NHIES data for 2003/04 and 2009/10 is used. The surveys serve as the official source of poverty estimates in the country. In Namibia, poverty is officially defined using per adult equivalent household expenditure and the upper poverty line is equal to N$4535.52 (±US$533.59) per year in 2009 prices (NSA, 2012). Both poverty rates and average regional per adult equivalent expenditure is used as welfare measures in the analysis.

One problem with the NHIES data is that it is only representative for large administrative regions. At this level of aggregation, the benefits of daytime RS data over night lights will not necessarily be highlighted as these regions all contain urban centres which will be electrified. However, daytime RS data will provide benefits at area level where night lights are not present. Regional level regression analysis would also only be based on 13 observations (see Figure 1) which also leaves the statistical analysis with little power.

Thus, the second source of data is constituency level poverty rates for 2001 and 2011 which were estimated using the small-area estimation (NPC, 2015). These rates were estimated using the method proposed by Elbers et al. (2003). Consumption was modelled in the NHIES surveys using variables also found in Namibia's census data and using these models, consumption was predicted in the census data which is geographically more refined. NHIES 2003/04 was used to impute consumption levels in the 2001 Census and NHIES 2009/10 was used to predict consumption levels in the 2011 Census. The 2011 Census is also used to provide descriptive statistics in this study. With regards to regions of Namibia, note that after the release of the 2011 Census, Kavango was divided into two regions - Kavango East and Kavango West. The name of Caprivi was also changed to Zambezi. The analysis presented here uses regions as defined in the 2011 Census and surveys prior to that.

3.5 Remotely Sensed Data

3.5.1 Indicators of Vegetation Quality

Various measures of vegetation quality exist, some of which were briefly referred to in Section 2 when discussing other studies. This study uses the Normalised Differenced Vegetation Index (NDVI) for the main results. NDVI correlates well with a number of vegetation properties which makes crop yield forecasting possible (Huang and Han, 2014). It is also the index used most extensively to forecast yields (Mkhabela et al., 2011). It is calculated as a ratio of two bands of EMR reflectance: the red (R) and near-infrared (NIR) bands. It is based on the fact that healthy vegetation absorbs EMR in the red band and reflects EMR in the infrared band. It is calculated (Myneni and Hall, 1995) as follows:

$$NDVI = \frac{NIR - R}{NIR + R}. \tag{1}$$

The index ranges between -1 and 1 and it is positively correlated with vegetative health. NDVI for water bodies is negative while rocks and bare soil have values close to zero (Mkhabela et al., 2011).

In the prediction of crop yields, NDVI has been incorporated in various different ways. Research varies on the period prior to harvest for which to include NDVI, and also, whether averages or sums (integrals) over an identified period should be used. Mkhabela et al. (2005) tested various specifications in Swaziland for the prediction of maize yields, and was guided by the explanatory power of different models to identify the best specification of NDVI. Explanatory power of models also differed across regions in the country. Huang and Han (2014) provide an overview of how studies have used NDVI for prediction of yields for a range of crops in different settings. The fact that this study only attempts to establish plausible relationships between vegetation and welfare implies that there is no attempt to find the measure with the highest explanatory power.

As a robustness check results are also shown for series based on the Enhanced Vegetation Index. The EVI is an alternative to NDVI. It was developed to improve vegetation monitoring by removing atmospheric influences and canopy background signals and it is more sensitive in areas with high biomass (Huete et al., 2002).[5]

For this study we use the 16-day composite product from the MODIS Terra Vegetation Indices (MOD13A1), version 5, dataset which includes NDVI and EVI. The dataset contains the highest daily value for NDVI/EVI within a 16 day period which also satisfies a set of data quality measures (NASA LP DAAC, 2017a).

The 16 day period selected was 6 - 22 March for each of the study years. This is based on the fact that the harvest period for millet, the most abundant crop grown in Namibia (NSA, 2015), begins, on average, in April or May each year, depending on the agro-ecological zone (FAO, 2010). As a robustness check the results for NDVI are also shown when the 16-day period 18 February - 6 March is used.

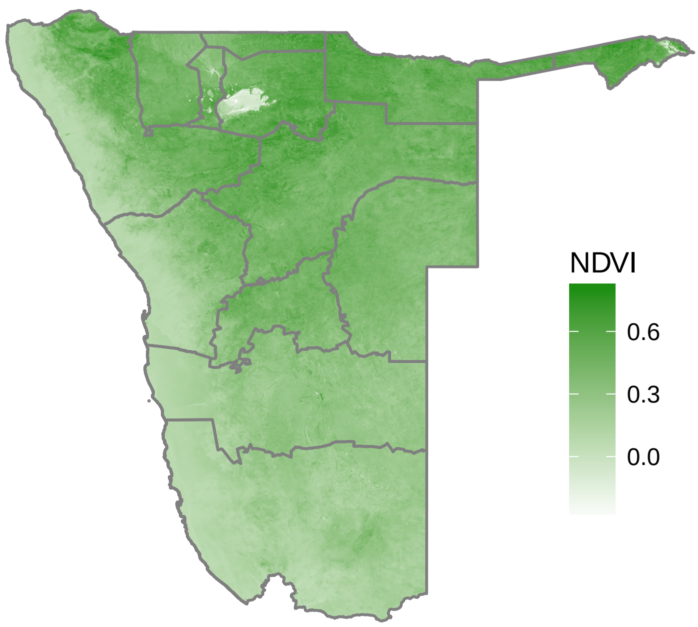
The last source of vegetation is a series based on EVI which calculates the integral of EVI (INEVI) for the primary growing season in a particular year. Zhang et al. (2006) shows that INEVI, based on the RS series NBAR MODIS, correlates well with another RS measure of vegetation productivity - Net Primary Productivity. The INEVI series used, comes from the V005 MODIS Land Cover Dynamics (MCD12Q2) product covering a resolution 500m^2 per pixel. The measure is based on algorithms developed in Zhang et al. (2003) to identify four key phenological phases of vegetation in a year. These phases can be used to identify the primary

[5] See Huete et al. (2002) for the formula for EVI

growing season in a year over which the integral of EVI is taken. Ganguly et al. (2010) provides an overview of the product. The product contains a large number of missing values across Namibia, however. This is due to the products algorithm excluding pixels that do not meet all the quality requirements (NASA LP DAAC, 2012b).[6] Thus, the results demonstrate the use of alternative specification of vegetation quality but should be interpreted with caution.

Data for 2001, 2003, 2009 and 2011 is used. The timing of the NHIES surveys (which stretch over two years) does complicate the choice of which year of data to use. Ideally the vegetation index should capture the season of production which provides for consumption and income that is captured in the surveys. In the NHIES data, the date of interviews of households does not vary systematically by region, and thus it is not possible to identify the reference agricultural season. As noted earlier, harvest of millet starts in April/May. For NHIES 2003/04 data collection started in September 2003 and ended in August 2004 (CBS, 2006). For NHIES 2009/10 data collection started in June 2009 and ended in July 2010 (NSA, 2013). It is thus clear that using 2003 and 2009 RS data should capture the periods of primary production that served consumption for most survey respondents. Figure 2 shows the NDVI for Namibia for 6-22 March 2009.

Figure 2: NDVI (6-22 March 2009) in Namibia



NDVI Source: MODIS Terra Vegetation Indices (MOD13A1), Version 5. Published by NASA LP DAAC.

[6] See NASA LP DAAC (2012a) for other know issues.

3.5.2 Crop Land Data

For the visual exploration of economic activity identified by RS data in under-electrified settings, Erongo and Kavango were selected for study. Erongo, a previously advantaged region, has high levels of wealth and electrification. Kavango was initially north of the Police Zone, which implies it was subject to apartheid era policies. It has high levels of crop-farming and is under-electrified. Farming in the region happens mostly on communal lands.

For the classification of cropland, Landsat data was used. The Landsat program is a joint program by the U.S Geological Survey (USGS) and National Aeronautics and Space Administration (NASA). There have been 8 Landsat satellites launched. This study uses data from Landsat 5 and Landsat 7. Landsat 7 captures images of the same point every 16 days, while Landsat 5 does it every 18 days, both at resolution of 30m^2 (USGS, 2015).

Images for 2003 were obtained from Landsat 7 imagery while Landsat 5 was used for 2009. The reversion back to Landsat 5 was due to a fault on the scanline corrector on Landsat 7, which caused a loss of data in images occurring after May 2003. Given that cloud-free images are required for the classification, the date of each image is not the same. Land use is slow to change, so this is not a problem in identifying cropland within a season. A satellite analysis team at Stellenbosch University - the Centre for Geographic Analysis (CGA) - was hired to process and classify the Landsat images into cropland or not.[7].

3.5.3 Night Lights Data

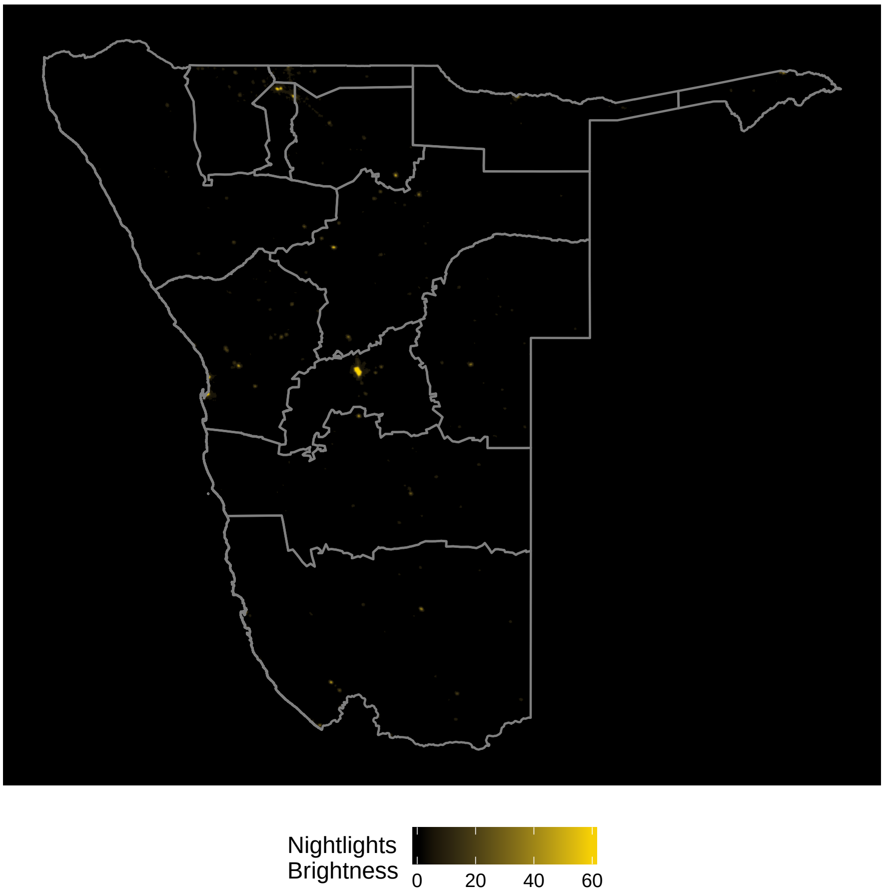
The night lights data is captured by the Defence Meteorological Satellite Program (DMSP) Operational Linescan System. The data is collected by the US Air Force Weather Agency and processed at the DMSP data centre. Raw data includes a range of radiance emitted at night, including moonlit objects, gas flares and man-made light. The DMSP releases a stable lights set for each year which excludes events such as fires and moonlit objects but includes gas flares. However, based on data available from DMSP, no gas flares are present in Namibia. The data is released in the form of 30 arc-second grids which is approximately 1km^2. Values range from 0 (which means that no stable light was detected there) in the year to 63 (which is a top censored value).

There are comparability problems over time as new satellites are launched and censors become

[7] For the CGA's report on how classification was conducted, please contact the author at thomasf@sun.ac.za

older. A number of methods have been proposed to make the series comparable over time. This report uses calibration parameters proposed by Elvidge et al. (2014) which was calculated by identifying areas where the brightness of lights has not changed over time. Using this as a base, the brightness of lights of the different satellite series are calibrated to each other. Night lights data was used for the years corresponding to the survey and poverty mapping data. For 2001 and 2003, the F14 series is used, for 2009 the F16 series is used and for 2011 the F18 series is used. Figure 3 presents the brightness of light at night for Namibia for 2009.

Figure 3: Brightness of Light at Night in Namibia - 2009



Night lights Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

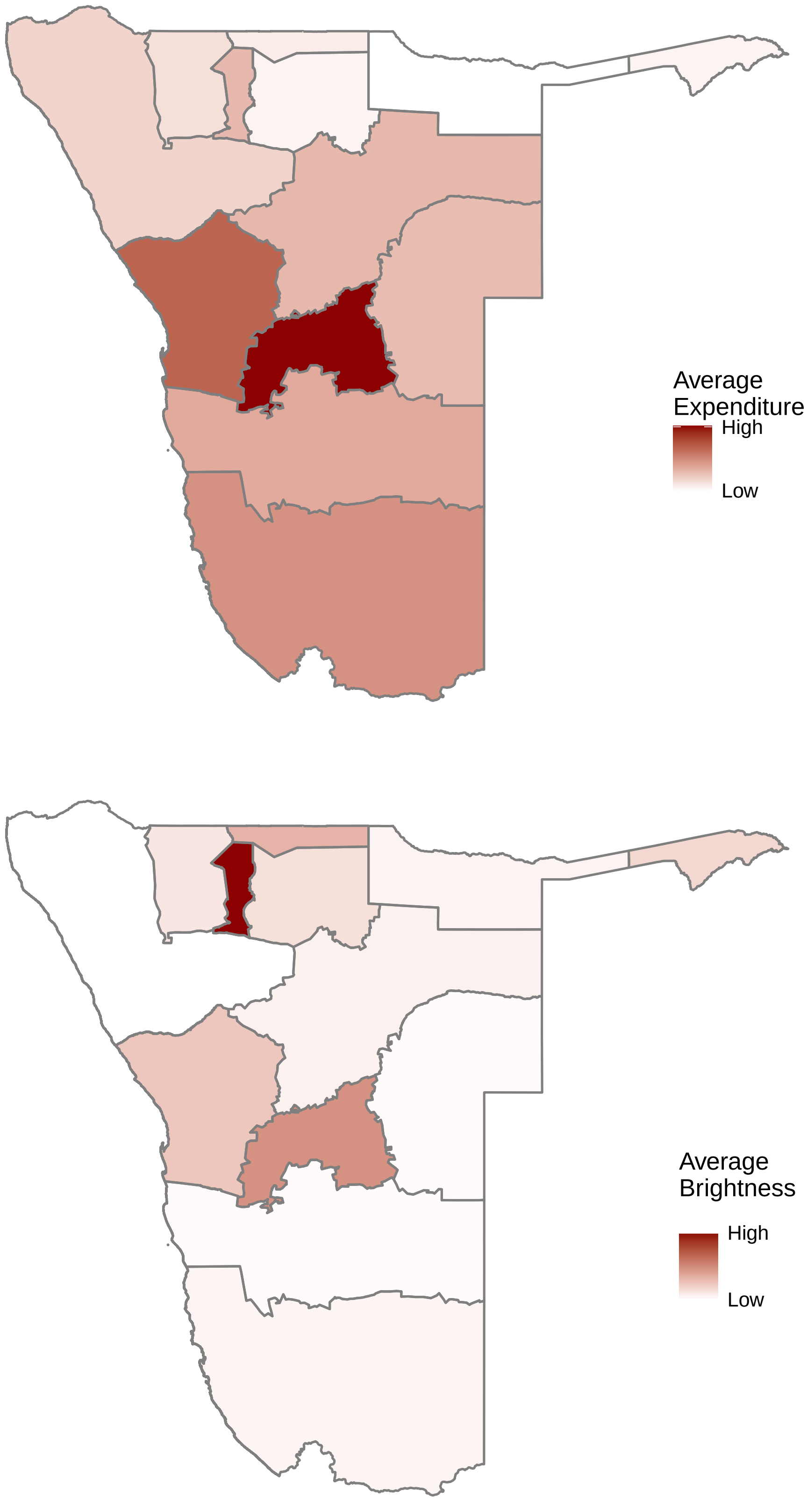
4 Results

As a start to this section, descriptives highlighting the regional inequalities in Namibia are shown. This is followed by a section which explores the extent of economic activity observed, and, finally the regression results. For all analyses that follow, the brightness of light and vegetation quality are expressed as regional averages.

4.1 Descriptives

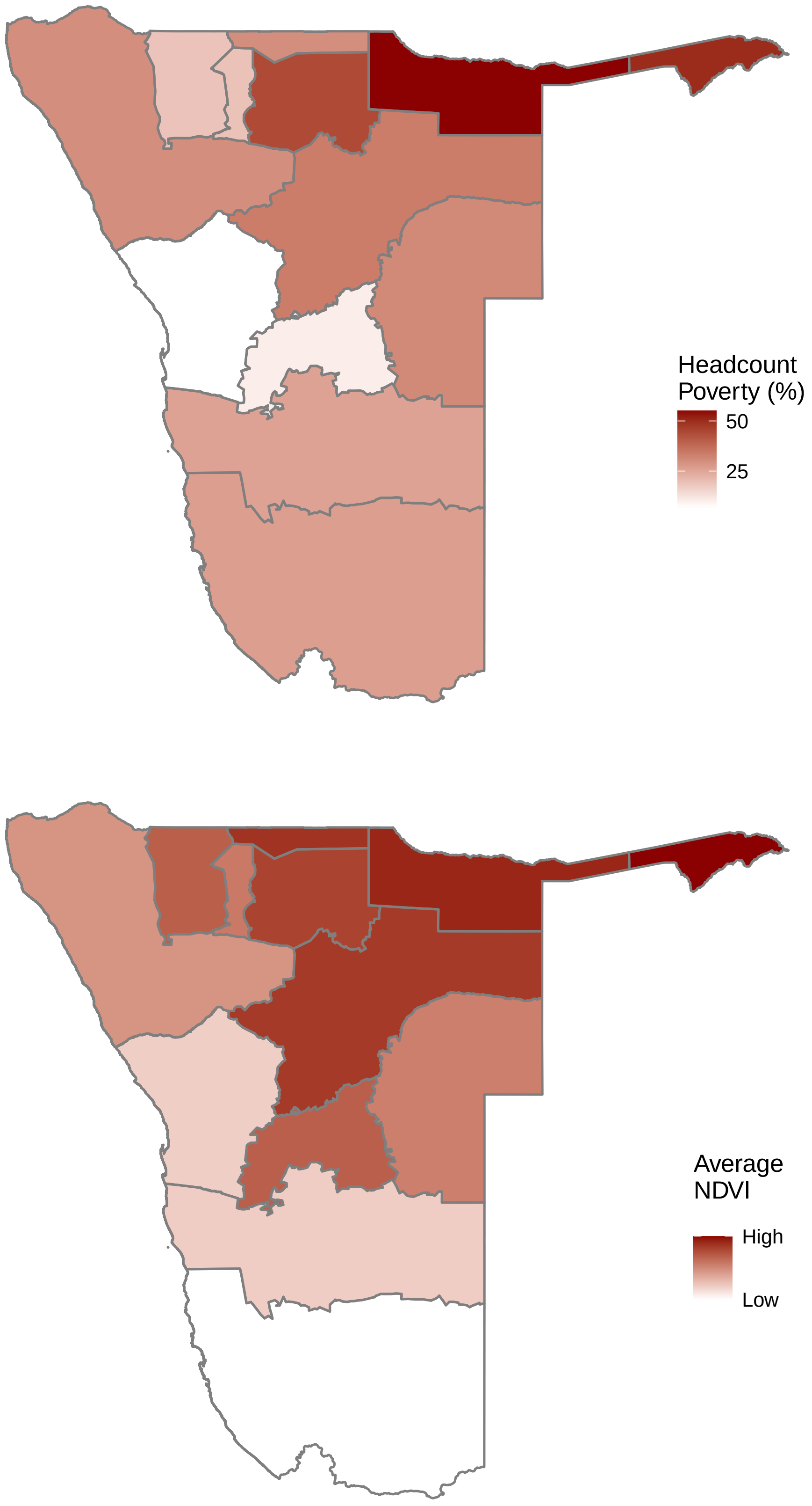
Figure 4 shows average expenditure and average brightness of lights, and Figure 5 shows average NDVI and headcount poverty by region for Namibia in 2009. It is evident that the highest poverty rates, and also the lowest average expenditure, is located in the north east of Namibia which are communal agricultural areas. Kavango, specifically, had headcount poverty higher than 50% in 2009. The regions with the main urban centres - Khomas and Erongo - had the highest levels of welfare.

Figure 4: Regional Differences in Average Expenditure and Night lights - 2009



Night lights Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency. Average Expenditure Source: Own Calculations using NHIES 2009/10.

Figure 5: Regional Differences in Headcount Poverty and NDVI - 2009



NDVI Source: MODIS Terra Vegetation Indices (MOD13A1), Version 5. Published by NASA LP DAAC. Poverty Rate Source: Own Calculations using NHIES 2009/10.

The regions with higher levels of poverty, also tended to have higher average NDVI levels, and these regions are once again located in the the north-east of Namibia. It needs to be pointed out, however, that Khomas has a relatively high level of average NDVI as well. Windhoek, the urban centre, covers a small proportion of the total area, which implies that the rural areas increase the average NDVI observed. Average brightness of light is low in the poor north-east regions and higher in Erongo and Khomas, which have the lowest levels of poverty. Oshana, in the north-centre of the country had the highest levels of average brightness. Oshana contains a number of large towns and the region is much smaller than both Erongo and Khomas. This highlights the importance of controlling for the size of a region.

To quantify regional differences in agricultural activities and electrification, table 1 presents results from Namibia's 2011 Census. The percentage of households in a region that use electricity for lighting varies from 12% in Omusati to 83% in Erongo. The percentage of households that farm also varies from 11% in Erongo and Khomas to 81% in Ohangwena. This indicates, that night lights would likely identify greater areas of economic activity in urbanised regions such as Erongo and Khomas. Daytime RS should identify greater areas in under-electrified regions with high levels of agriculture.

Table 1: Household Electrification and Farming for Across Regions in Namibia

	Use electricity for lighting	Farming	Farming is Main Source of Income	Farms Crops
Caprivi	0.35	0.48	0.21	0.44
Erongo	0.83	0.11	0.02	0.04
Hardap	0.69	0.22	0.07	0.04
Karas	0.71	0.14	0.05	0.03
Kavango	0.25	0.56	0.43	0.54
Khomas	0.70	0.11	0.01	0.05
Kunene	0.35	0.48	0.32	0.34
Ohangwena	0.13	0.81	0.26	0.81
Omaheke	0.38	0.30	0.21	0.10
Omusati	0.12	0.74	0.22	0.73
Oshana	0.34	0.49	0.13	0.47
Oshikoto	0.22	0.67	0.33	0.66
Otjozondjupa	0.58	0.24	0.10	0.11

Source: Own Calculations using Namibia Census 2011

4.2 Observed Economic Activity in Erongo and Kavango

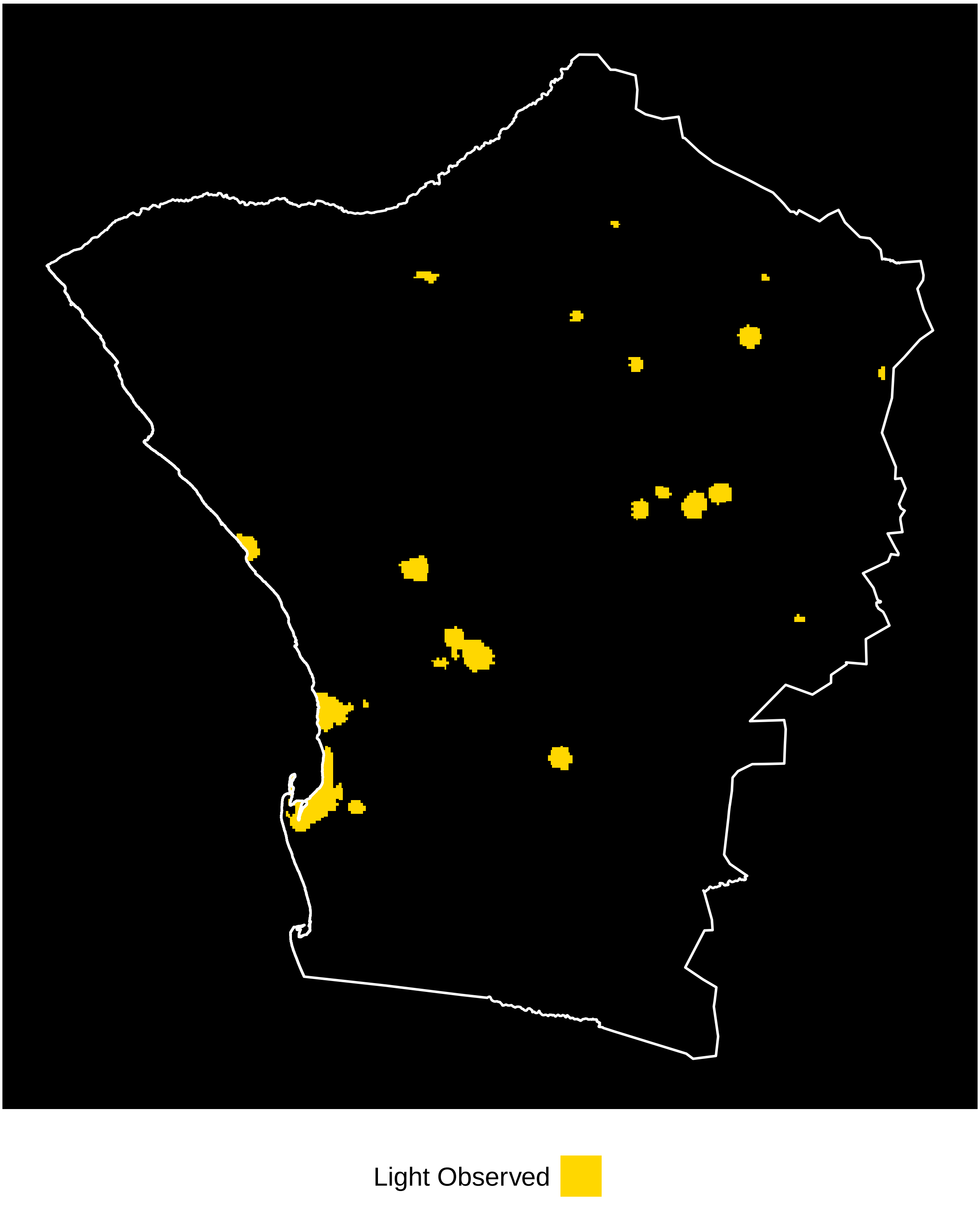
This section explores visually how observed lighting and crop farming differ across Erongo and Kavango. In Erongo, 83% of households use electricity for lighting while in Kavango only 25% do. On the other hand, only 4% of households farm crops in Erongo compared to 54% in Kavango.

For the figures depicting cropland, pixels expressed the percentage of the area covered in cropland.[8] Figures 6 and 7 show areas observed to have nighlights and cropland in Erongo respectively. A large, and spread out area of Erongo is observed under lighting while limited crop farming is observed in the north east of the region. Lights are observed for 1.6% of the total area while crop farming is observed for 0.046% of the total area. In Erongo, it is thus evident that night lights data has greater ability to locate areas of economic activity.

Figures 8 and 9 show observed lighting and cropland for Kavango. The figures confirm the expectation that daytime RS data observes a greater area of economic activity than night lights in under-electrified areas. Only urban areas, mostly the town of Rundu, are observed using night lights. Quantitatively, 4.6% of the total area was observed to be cropland. Night lights was only observed for 0.5% of the area.

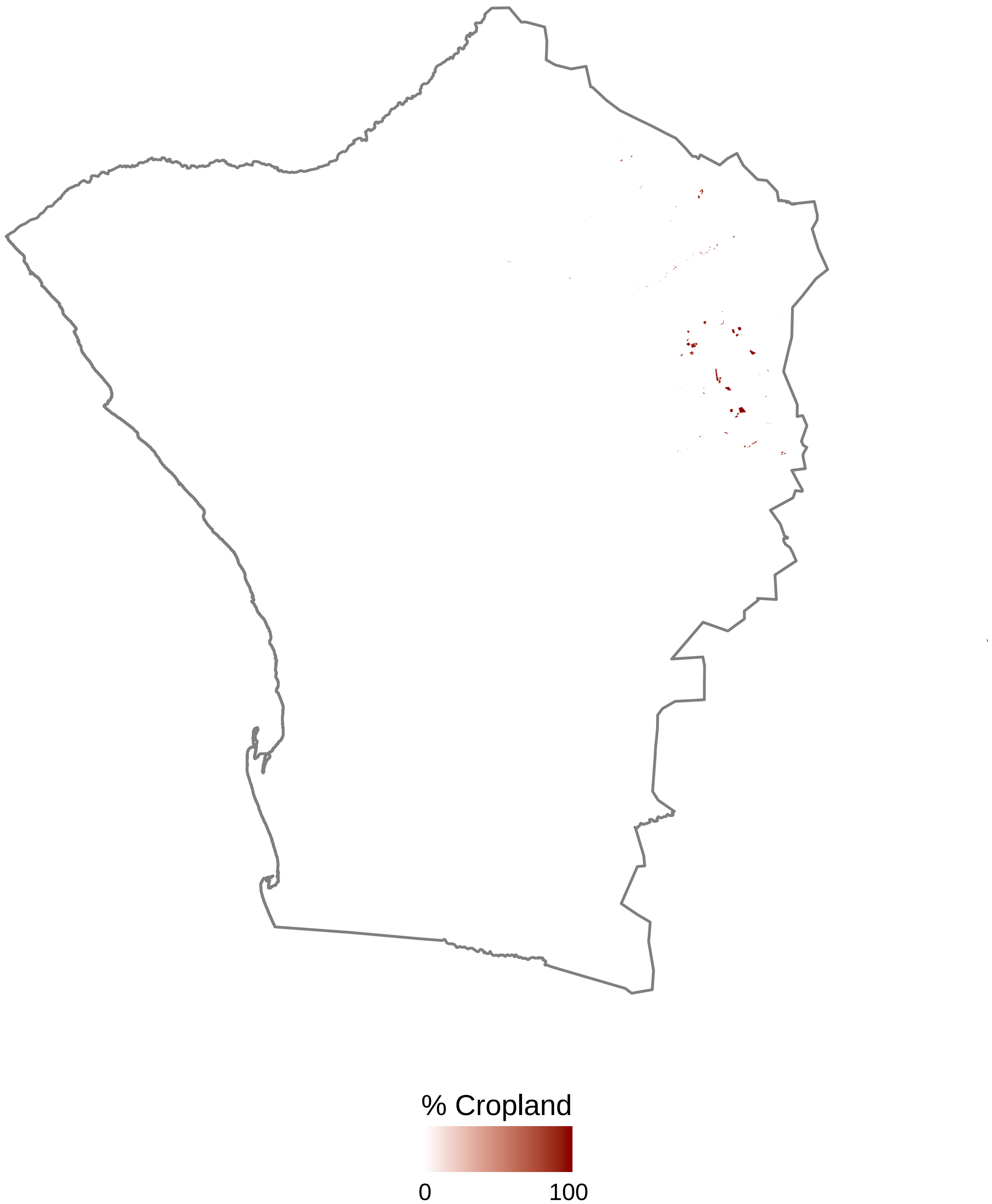
[8] This was done because there are too many data points in Landsat images to plot each individually. Thus, images were aggregated to a smaller number of pixels where each represents the percentage of cropland for the area it covers.

Figure 6: Area Observed by Night Lights - Erongo



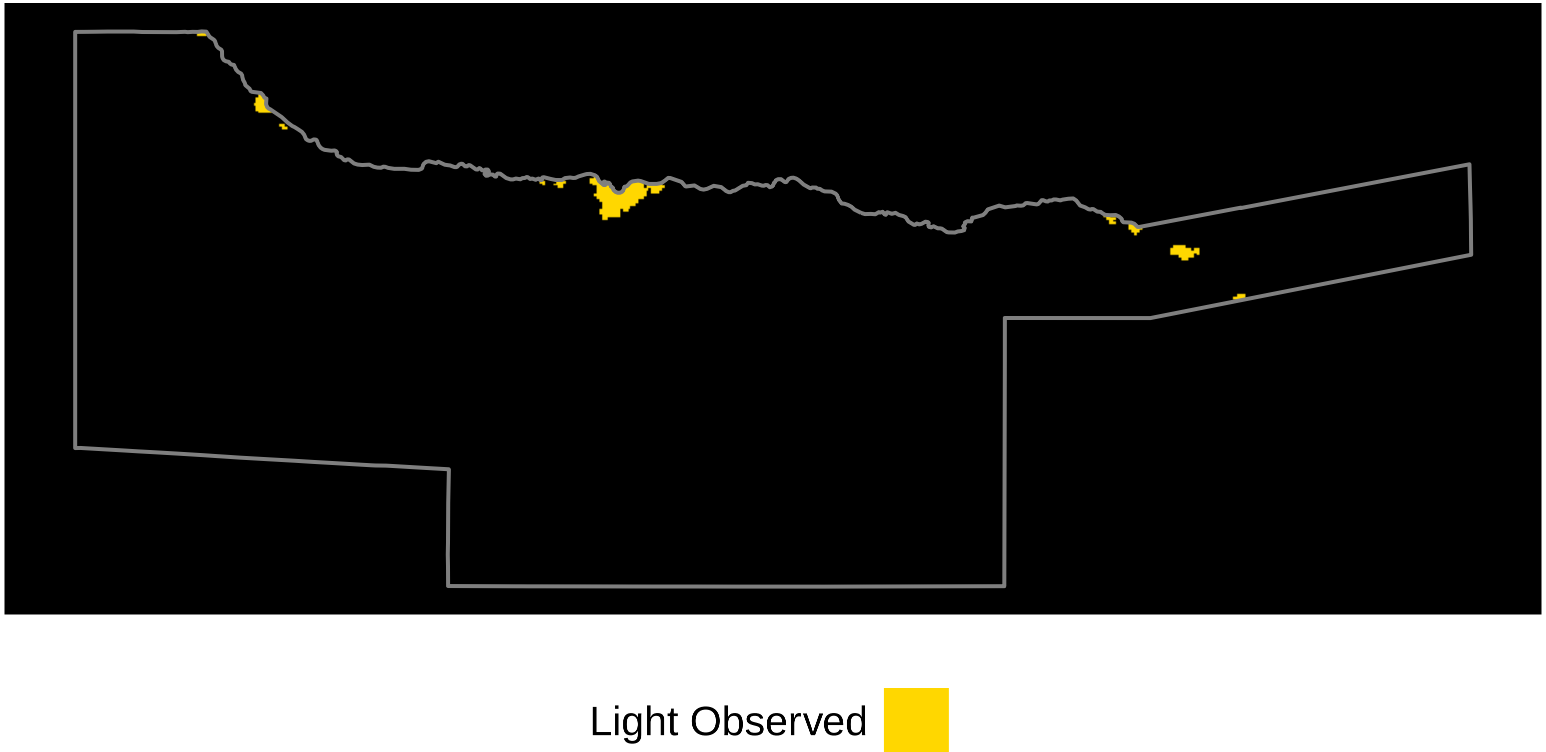
Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

Figure 7: Area Observed under Cropland - Erongo



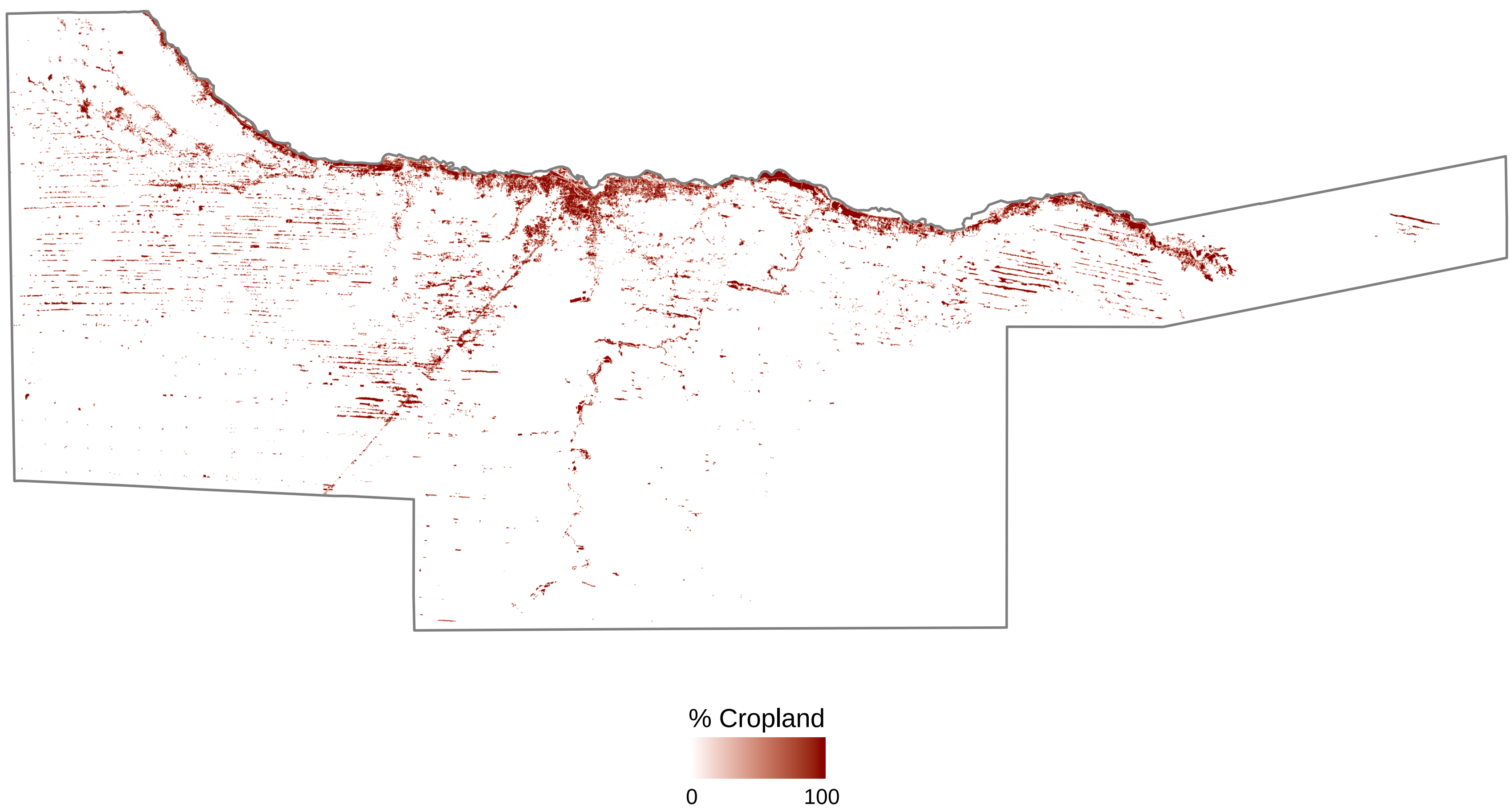
Classification using Landsat 5. Data available from the U.S. Geological Survey.

Figure 8: Area Observed by Night Lights - Kavango



Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

Figure 9: Area Observed under Cropland - Kavango



Classification using Landsat 5. Data available from the U.S. Geological Survey.

Figure 10: Location of Kahenge in Kavango

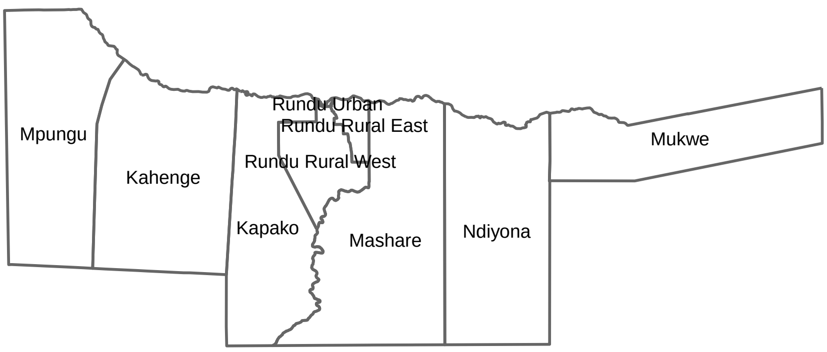


Table 2: Household Electrification and Farming for Across Constituencies in Kavango

	Use electricity for lighting	Farming	Farming is Main Source of Income	Farms Crops
Kahenge	0.08	0.68	0.61	0.67
Kapako	0.14	0.64	0.59	0.63
Mashare	0.08	0.68	0.53	0.67
Mpungu	0.10	0.82	0.61	0.82
Mukwe	0.16	0.74	0.60	0.73
Ndiyona	0.13	0.64	0.48	0.63
Rundu Rural West	0.38	0.37	0.22	0.36
Rundu Urban	0.81	0.17	0.04	0.15
Rundu Rural East	0.21	0.41	0.32	0.39

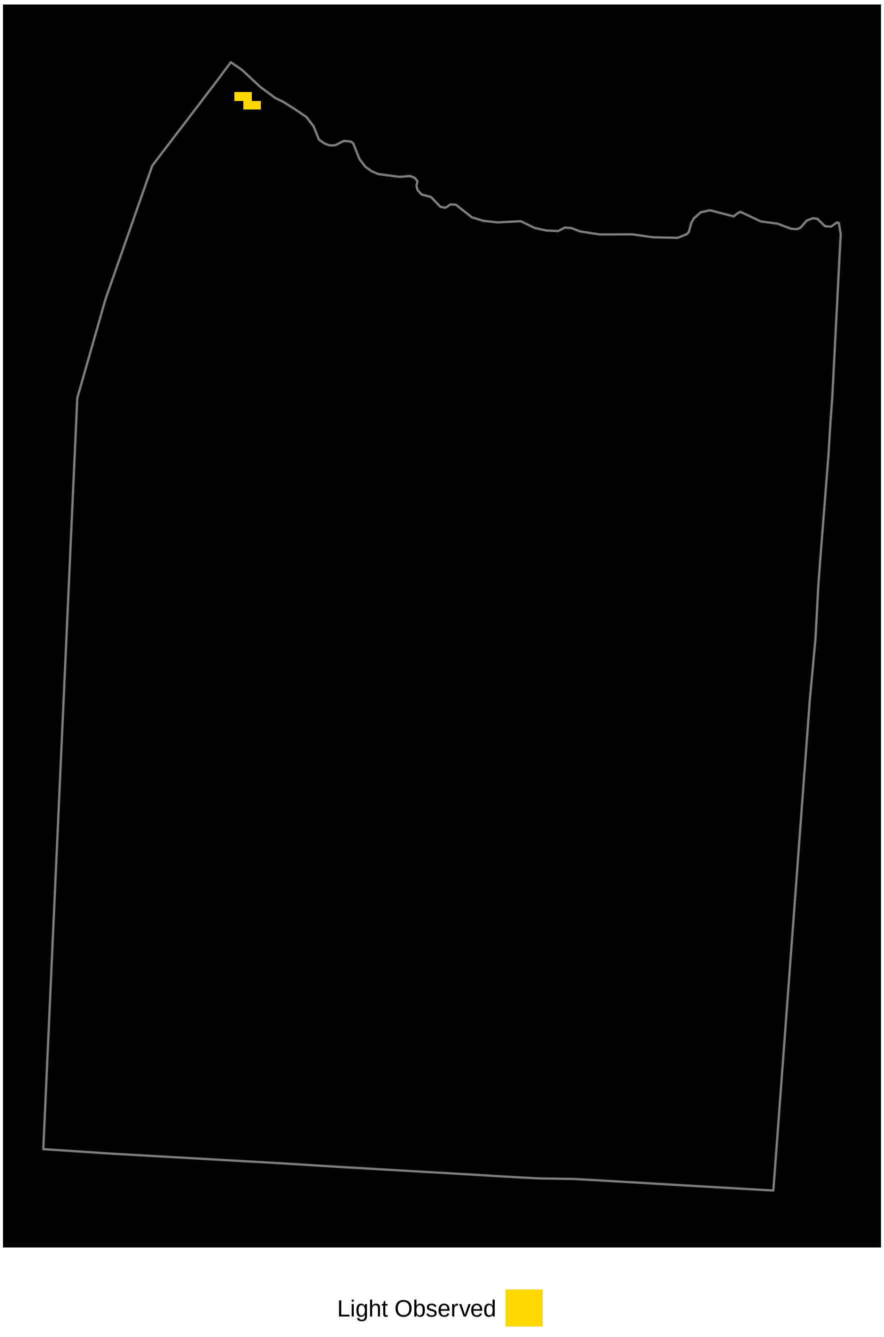
Source: Own Calculations using Namibia Census 2011

The figures for Kavango, at the scale which they can be presented here, fail to show the extent to which daytime RS data can pick up activity in rural areas. To illustrate this further, a sub-region in Kavango - Kahenge - is shown. Table 2 show the rates of farming and electrification for constituencies in Kavango and it shows that in Kahenge only 8% of households use electricity for lighting, 68% of households farm and 67% of households farm crops. See figure 10 for the location of constituencies in Kavango. Figures 11 and 12 show observed night lights and cropland for Kahenge. Night lights were observed in only 0.04% of the region while cropland was observed for 5.7% of the region.

Given that this paper is concerned with tracking socio-economic changes, the area that changed status between 2003 and 2009 was also compared for night lights and cropland in Kahenge. It is to be noted that cropland is not the measure used to track socio-economic outcomes in the next section, but the observed changes drive home the point that daytime RS data does indeed capture changes at much smaller and more disaggregated levels than the night lights data in rural under-developed areas. The areas where night lights changed are shown in figure 13 and areas where cropland changed are shown in figure 14.

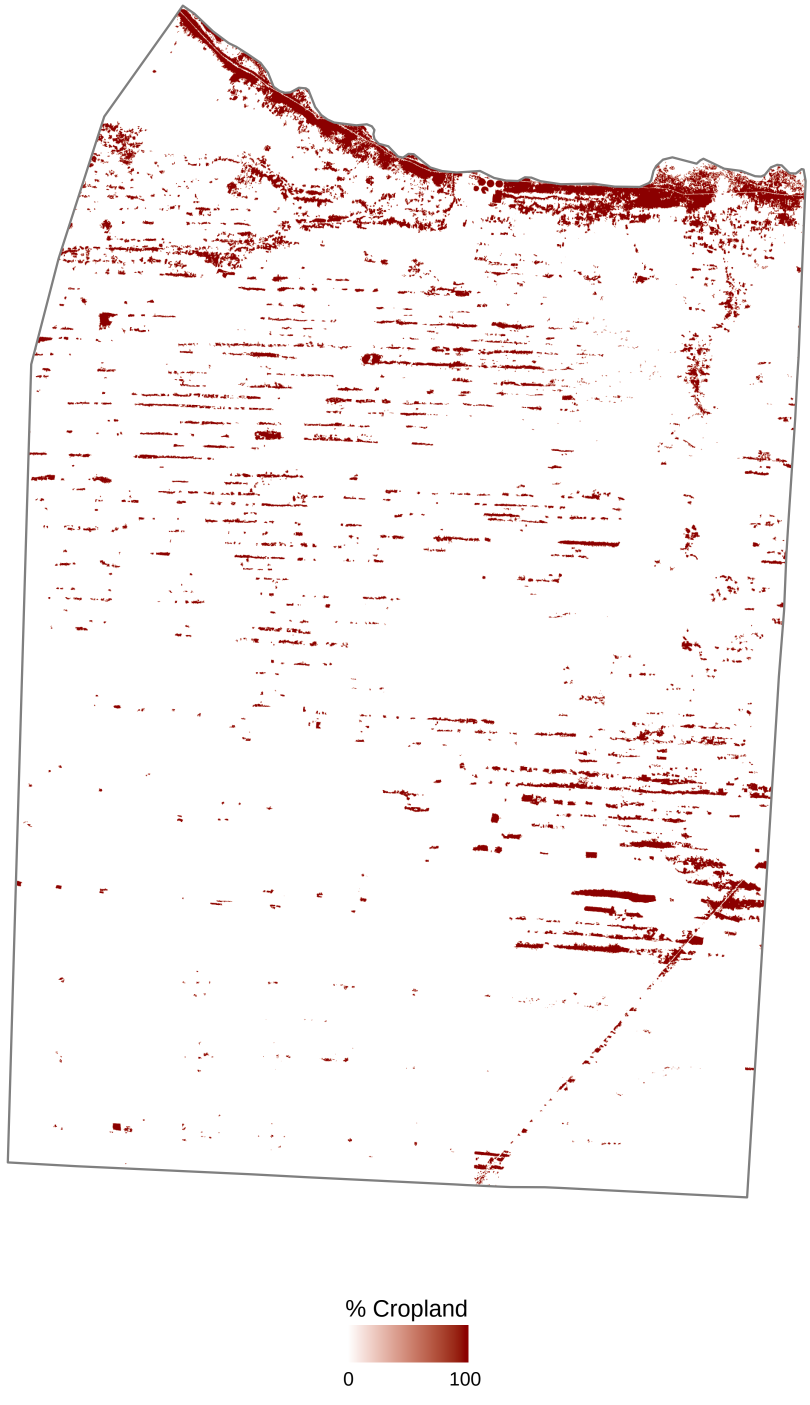
The fact that night lights are observed for a greater area than cropland in Erongo highlights that different satellite measures will be more effective in different settings. This analysis has indeed shown, however, that daytime RS data observes more economic activity and changes in underdeveloped agricultural areas. This does not imply that night lights data should not be used. At aggregate regional level it is still plausible that it correlates well with welfare. Daytime RS data has the potential, however, to track welfare changes at much smaller levels of aggregation within such regions. This has potential to, for example, accurately identify areas that suffered poor agricultural harvests and thus might be in need of economic relief in a given year.

Figure 11: Area observed by Night Lights in Kahenge, Kavango, 2009



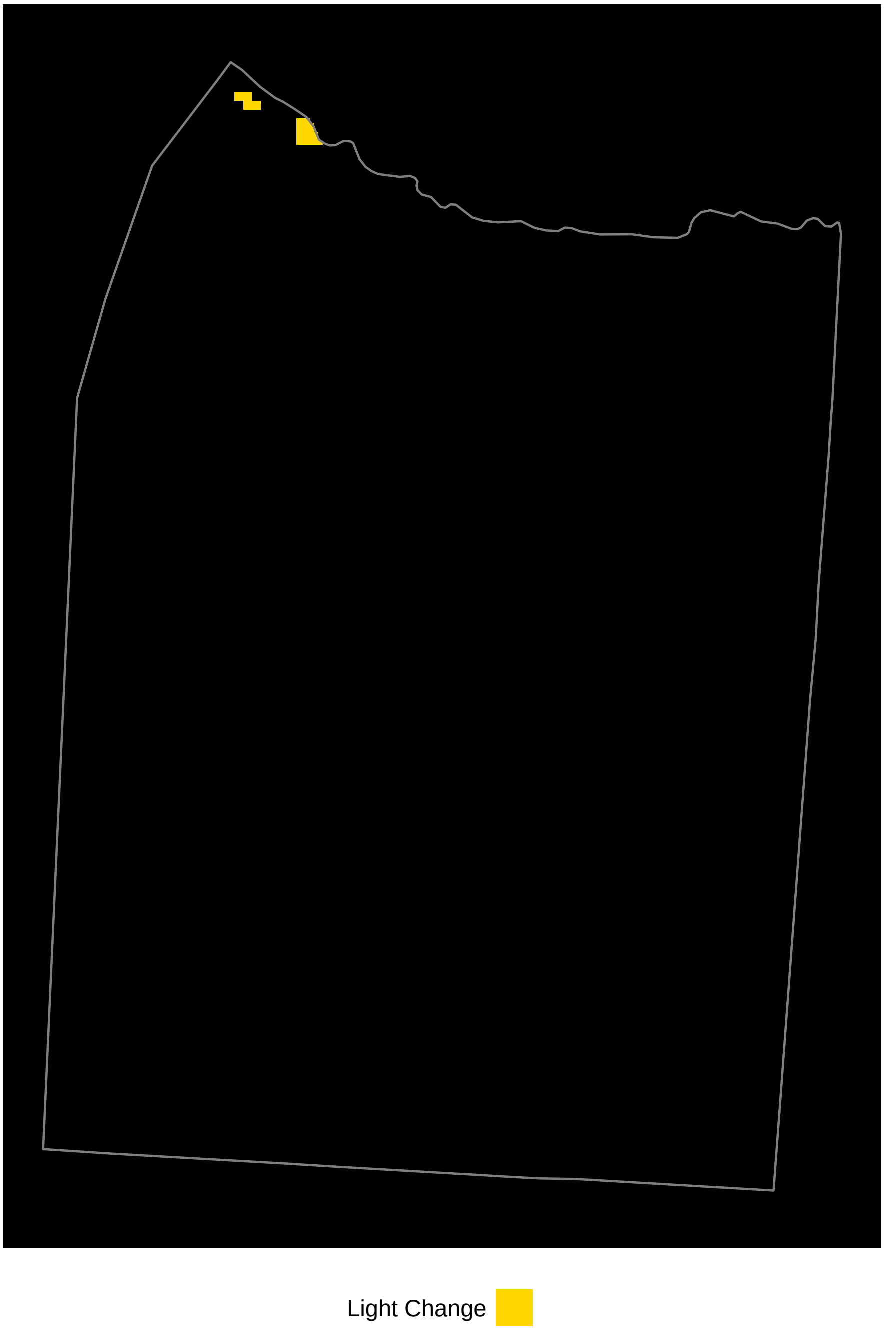
Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

Figure 12: Area observed by Night Lights in Kahenge, Kavango, 2009



Source: Classification using Landsat 5. Data available from the U.S. Geological Survey.

Figure 13: Area Night Lights Observed that Changed Between 2003 - 2009 in Kahenge, Kavango



Source: Classification using Landsat 5. Data available from the U.S. Geological Survey.

Figure 14: Area Cropland Observed that Changed Between 2003 - 2009 in Kahenge, Kavango

Cropland Source: Classification using Landsat 5. Data available from the U.S. Geological Survey.

4.3 Correlation between Remotely Sensed Data and Welfare

The regression results are presented next. The average brightness of light is logged and vegetation indices are standardized. All standard errors are clustered at the geographic level of observation - either regions or constituencies.

Table 3 shows results for regressions on logged average regional expenditure. Model 1 shows the results when no controls are introduced. As is consistent with our expectation, NDVI is negatively correlated with average expenditure cross-sectionally, but once regional fixed effects are introduced, NDVI has a positive effect on average expenditure. Coefficients are statistically significant even with only 26 observations. The fixed effects point estimate suggests that as average NDVI increases by a standard deviation, average expenditure increases, on average, by 15% in a region.

In model 2, population density is included. This is included to control for the number of people that are dependant on a specific area of land for production. After controlling for it, NDVI coefficients do not change by much and remain statistically significant.

Model 3 shows the relationship between average expenditure and the log of average lights when no controls are added and in model 4 population density is once again added. The results show that average expenditure is higher in regions with higher levels of brightness and that within regions, increases in average brightness are associated with increases in average expenditure. This is consistent with the literature. A 1% increase in average brightness is associated, on average, with a 42% increase in average expenditure. For the pooled OLS estimates, the point estimate is initially 0.12 and insignificant, but once population density is controlled for, it increases to 0.43 and becomes significant.

In model 5 both NDVI and night lights are included with population density. As the focus of the study is mainly on the use of indicators of vegetation quality, it is useful to consider night lights as a proxy for urbanization in a region. Thus, even after controlling for the extent of urbanisation, the negative relationship between NDVI and expenditure remains significant. Also, the positive relationship over time remains significant, but the coefficient decreases slightly suggesting that expenditure increases by 11% with a one standard deviation increase in NDVI.

Finally, year dummies are added to control for year specific effects. For the pooled OLS specification this increases the adjusted R^2 to 0.69. For the fixed effects estimates this decreases the coefficients on both the average brightness of lights and NDVI. This is consistent with the

Table 3: Regional Regressions on ln(Average Regional Expenditure)

	1		2		3		4		5		6	
	POLS	FE	POLS	FE	POLS	FE	POLS	FE	POLS	FE	POLS	FE
NDVI	-0.25***	0.15***	-0.23***	0.18***					-0.21***	0.11**	-0.33***	0.05
	(0.06)	(0.04)	(0.07)	(0.04)					(0.07)	(0.05)	(0.07)	(0.14)
ln(Light)					0.12	0.42***	0.43***	0.46***	0.41***	0.25	0.35**	0.19
					(0.11)	(0.12)	(0.13)	(0.12)	(0.11)	(0.18)	(0.13)	(0.19)
Pop Density			-0.00	-0.07			-0.07***	-0.04	-0.05***	-0.08	-0.04**	-0.07
			(0.01)	(0.05)			(0.01)	(0.04)	(0.01)	(0.05)	(0.02)	(0.05)
Year	No	No	No	No	No	No	No	No	No	No	Yes	Yes
$adj.R^2$	0.21	0.50	0.18	0.53	0.03	0.50	0.44	0.50	0.58	0.60	0.69	0.59
R^2: Within		0.52		0.57		0.52		0.54		0.65		0.66
R^2: Between		0.48		0.00		0.06		0.32		0.16		0.17
R^2: Overall		0.24		0.01		0.07		0.33		0.17		0.18
ICC		0.97		0.98		0.97		0.94		0.96		0.96
Observations	26	26	26	26	26	26	26	26	26	26	26	26
Groups		13		13		13		13		13		13

Socio Economic Data Source: Own Calculations using NHIES 2003/04 and 2009/10. NDVI Source: MODIS Terra Vegetation Indices (MOD13A1), Version 5. Published by NASA LP DAAC. Night lights Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.
Regression also controls for year fixed effects.
POLS refers to Pooled OLS and FE refers to Fixed Effects regression.
ICC refers to the Intra-class correlation.

finding that NDVI and expenditure are positively correlated. The intra-class correlation is 0.96 suggesting that almost all of the variation is contained in the fixed effects of the model. The fact that the linear time trend causes the coefficient on NDVI to become insignificant suggests that NDVI was capturing the long-term trend in expenditure over time.

The initial results suggest that indicators of vegetation quality do indeed correlate with expenditure as expected. The fact that the sign on NDVI changes between the pooled OLS and fixed effects regressions, highlight that there is selection amongst differing regions. The poor are more likely to be located in rural regions with higher levels of vegetation quality. Robustness checks are discussed next.

4.3.1 Robustness checks

Table 4 shows the estimates when regional average expenditure is regressed on alternative vegetation indices. The model specification used, is similar to specification 2 in Table 3. In the first specification the NDVI from the 16-day composites for 18 February - 6 March (instead of 6 March - 22 March) is used. The second model uses EVI from the 16 day composite for 6 March - 22 March. The final model uses INEVI. In all cases the results are consistent with previous results - vegetation is negatively correlated with welfare across regions and positively correlated with welfare over time within regions. Note that model coefficients do not change dramatically when different indicators are used. The explanatory power of the fixed effects estimator when INEVI is used, is lower though. This can be due to the large amount of missing data.

Table 4: Regional Regressions on ln(Average Expenditure) using Alternative Vegetation Indices

	POLS	FE	POLS	FE	POLS	FE
NDVI(earlier date)	-0.26**	0.22***				
	(0.10)	(0.06)				
EVI			-0.27**	0.19***		
			(0.10)	(0.05)		
INEVI					-0.26**	0.15*
					(0.10)	(0.07)
Pop. Density	-0.00	-0.09	0.00	-0.07	-0.00	-0.01
	(0.01)	(0.06)	(0.01)	(0.06)	(0.01)	(0.07)
$adj.R^2$	0.24	0.03	0.21	0.10	0.22	-0.52
R^2: Within		0.57		0.60		0.33
R^2: Between		0.00		0.00		0.31
R^2: Overall		0.00		0.00		0.18
ICC		0.98		0.98		0.95
Observations	26	26	26	26	26	26
Groups		13		13		13

Socio Economic Data Source: Own Calculations using NHIES 2003/04 and 2009/10. NDVI Source: MODIS Terra Vegetation Indices (MOD13A1), Version 5. Published by NASA LP DAAC. Night lights Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

Regression also controls for year fixed effects.

POLS refers to Pooled OLS and FE refers to Fixed Effects regression.

ICC refers to the Intra-class correlation.

For comparison with based results in Table 3 the model specification used is model 2.

Table 5 show the results when regional headcount poverty rates (instead of average expenditure) are regressed on average NDVI and average brightness. The estimates show, as expected, that NDVI is positively correlated with poverty across regions, but negatively correlated over time within regions. Average brightness is negatively correlated with poverty across regions and within regions over time. Based on the R^2 values, the RS data performs better in explaining variations in average expenditure levels than poverty rates.

Finally, Table 6 shows the results when constituency level poverty rates are regressed on NDVI and lights. As the rates are based on imputed income, the standard errors need to be bootstrapped for correct inference.[9] For nine constituencies light was not observed at all in either 2001 or 2011.[10] This also highlights the limitations of night lights for tracking economic activity in under-developed areas. Point estimates for NDVI are consistent with estimates based on regional level poverty rates. They are similar in magnitude and sign and are statistically significant. The relationship between night lights and poverty rates over time is not as clear at constituency as regional level. The point estimate, after controlling for population density, was

[9] Pooled OLS results are clustered at constituency level, but the fixed effects estimates are not.

[10] For these constituencies, a value of 0.0001 was assigned before the average level of brightness was logged.

Table 5: Regional Regressions on Headcount Poverty

	1		2		3		4		5		6	
	POLS	FE	POLS	FE	POLS	FE	POLS	FE	POLS	FE	POLS	FE
NDVI	5.30*	-6.54**	6.29**	-9.13***					5.54**	-5.08	10.16***	1.20
	(2.54)	(2.54)	(2.73)	(2.71)					(2.20)	(4.38)	(2.78)	(10.17)
ln(Light)					-5.90*	-20.08**	-12.14**	-24.69***	-11.61***	-15.41	-9.37**	-8.86
					(2.74)	(8.19)	(4.20)	(7.56)	(3.46)	(14.00)	(3.32)	(13.49)
Pop Density			-0.37	5.57			1.34***	4.17	0.98**	5.93*	0.52	5.49
			(0.36)	(3.72)			(0.43)	(2.44)	(0.34)	(3.26)	(0.37)	(3.65)
Year	No	No	No	No	No	No	No	No	No	No	Yes	Yes
$adj.R^2$	0.09	0.29	0.08	0.37	0.16	0.35	0.33	0.39	0.43	0.44	0.61	0.44
R^2: Within		0.32		0.42		0.38		0.44		0.50		0.53
R^2: Between		0.38		0.02		0.19		0.30		0.02		0.02
R^2: Overall		0.13		0.01		0.19		0.29		0.03		0.02
ICC		0.86		0.97		0.90		0.87		0.96		0.96
Observations	26	26	26	26	26	26	26	26	26	26	26	26
Groups		13		13		13		13		13		13

Socio Economic Data Source: Own Calculations using NHIES 2003/04 and 2009/10. NDVI Source: MODIS Terra Vegetation Indices (MOD13A1), Version 5. Published by NASA LP DAAC. Night lights Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

Regression also controls for year fixed effects.

POLS refers to Pooled OLS and FE refers to Fixed Effects regression.

ICC refers to the Intra-class correlation.

-24.69 in regional level regressions but -1.45 in constituency level regressions.

On concern is the model specification 6 in Table 6. The coefficient on NDVI when fixed effects and year dummies are introduced is positive, suggesting that NDVI and poverty are positively correlated over time within small regions. This result is driven by urban constituencies as is shown in Table 7. Table 7 introduces the percentage of the population that reside in rural areas within regions or constituencies in regressions on the headcount poverty rate. In the first specification, the percentage of the population residing in rural areas is interacted with NDVI at constituency level and regional level. This shows that over time, within more rural constituencies, poverty decreases with increases in NDVI. This interaction is also statistically significant when the sample of constituencies is used. In the second specification, the sample of constituencies is limited to those where more than 50% of the population reside in rural areas. In this specification, the NDVI is insignificant - a finding that mirrors specification 6 in Table 5.

The initial positive correlation between NDVI and poverty over time within constituencies was thus driven by predominantly urban constituencies. The results at regional level were clearly not as sensitive to urban areas because of the relative size of urban areas within regions. Applying the environmental Kuznets curve hypothesis to the results for urban areas suggests that a large number of urban constituencies are at initial stages of development where increases in development are associated with environmental deterioration.

The results support this paper's premise that RS based vegetation quality indices can be used to track socio-economic welfare over time in rural areas. The coefficients are large and have the correct sign based on the idea that livelihoods are dependant on nature and agriculture. The results have also highlighted why cross-sectional results can be misleading. The night lights results are also consistent with previous findings which is that night lights are positively correlated with welfare. However, as pointed out in the previous section, daytime RS data presents opportunities to track changes at much smaller geographic levels as night lights are completely absent in many rural areas.

Table 6: Constituency Level Regressions on Headcount Poverty

	1		2		3		4		5		6	
	POLS	FE	POLS	FE	POLS	FE	POLS	FE	POLS	FE	POLS	FE
NDVI	7.14***	-9.99***	6.52***	-10.28***					4.99***	-10.61***	6.20***	6.65**
	(1.43)	(2.28)	(1.55)	(2.24)					(1.25)	(2.07)	(1.36)	(3.01)
ln(Light)					-3.80***	-1.45	-3.57***	-1.45	-3.29***	-1.83	-3.13***	1.88
					(0.40)	(1.73)	(0.45)	(1.74)	(0.49)	(1.52)	(0.51)	(1.36)
Pop Density			-0.00	0.00			-0.00	0.00	-0.00	0.00	-0.00	0.01**
			(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Year	No	No	No	No	No	No	No	No	No	No	Yes	Yes
$adj.R^2$	0.12	-0.74	0.20	-0.75	0.34	-0.98	0.34	-1.00	0.40	-0.72	0.47	-0.06
R^2: Within		0.13		0.14		0.01		0.01		0.16		0.49
R^2: Between		0.19		0.28		0.40		0.39		0.07		0.12
R^2: Overall		0.12		0.20		0.34		0.33		0.04		0.06
ICC		0.86		0.88		0.70		0.70		0.86		0.95
Observations	214	214	214	214	214	214	214	214	214	214	214	214
Groups		107		107		107		107		107		107

Constituency Level Poverty Rates: NPC (2015). NDVI Source: MODIS Terra Vegetation Indices (MOD13A1), Version 5. Published by NASA LP DAAC. Night lights Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

Regression also controls for year fixed effects.

POLS refers to Pooled OLS and FE refers to Fixed Effects regression.

ICC refers to the Intra-class correlation.

Table 7: Regressions on Poverty Rates Accounting for the Rural Population

	Rural Interaction				>50% Rural	
	Regional Estimates		Constituency Estimates		Constituency Estimates	
	POLS	FE	POLS	FE	POLS	FE
NDVI	7.16**	4.38	3.33***	14.91***	7.21***	3.50
	(2.69)	(11.78)	(0.93)	(2.86)	(1.39)	(4.36)
NDVI*% Rural	1.03	-4.25	3.08*	-14.45***		
	(4.29)	(9.22)	(1.72)	(3.37)		
% Rural	22.42	28.97	33.09***	12.45		
	(13.62)	(36.88)	(2.23)	(12.43)		
ln(Light)	-4.36	-5.19	-0.64	1.49	-1.34*	2.02
	(3.51)	(14.10)	(0.40)	(1.30)	(0.78)	(1.36)
Pop Density	-0.24	3.76	0.00	0.01	0.06	-1.78***
	(0.44)	(4.37)	(0.00)	(0.01)	(0.07)	(0.65)
Year	Yes	Yes	Yes	Yes	Yes	Yes
adj.R^2	0.66	0.42	0.69	0.05	0.38	0.28
R^2: Within		0.56		0.55		0.67
R^2: Between		0.10		0.01		0.02
R^2: Overall		0.11		0.00		0.02
ICC		0.93		0.93		0.98
Observations	26	26	214	214	146	146
Groups		13		107		75

Socio Economic Data Source: Own Calculations using NHIES 2003/04 and 2009/10. NDVI Source: MODIS Terra Vegetation Indices (MOD13A1), Version 5. Published by NASA LP DAAC. Night lights Source: Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by the United States Air Force Weather Agency.

Regression also controls for year fixed effects.

POLS refers to Pooled OLS and FE refers to Fixed Effects regression.

ICC refers to the Intra-class correlation.

For comparison with based results in Table 3 the model specification used is model 2.

5 Conclusion

In most countries in SSA, household survey data is released at low frequency and is only representative at high levels of geographic aggregation. This is specifically the case with surveys designed to measure poverty. This limits what researchers can say about the location and short-term changes in poverty.

A new body of research attempts to overcome these limitations by using satellite data. Satellites capture data at high frequencies and small areas. The night lights data series, has particularly been shown to correlate well with GDP. However, a large share of the poor in sub-Saharan Africa are located in under-electrified, agriculturally dominant regions. Night lights cannot proxy economic activity where it is not present. Daytime RS data has also been used to model welfare cross-sectionally but the research and conclusions have not dealt with variation over time. Henderson et al. (2012) noted that night lights data is more suitable to estimate within region changes, and the same case can be made for daytime RS indicators. Regional heterogeneity can imply that cross-sectional correlations could be misleading if welfare changes over time are studied.

This paper has suggested the use of indicators of vegetation quality, derived from daytime RS as a proxy for welfare in rural regions. The reasoning is as follows: indicators of vegetation quality have been used substantially to predict crop harvests, and given that smallholder agriculture is an important source of income for households in rural areas, it should be that case that vegetation quality can proxy welfare. Indicators of vegetation quality would also overcome the ineffectiveness of night lights in underdeveloped areas.

Namibia was selected to explore theses ideas as it has large regional heterogeneity which could possibly highlight why cross-sectional results could be misleading. To illustrate that daytime RS data can identify economic activities where night lights cannot, the area of economic activity observed, proxied either night lights or cropland, was identified in Kavango - an impoverished, agricultural region. It was indeed observed that cropland was identified for a much larger area than night lights was observed for, indicating that daytime RS data can track economic change where night lights cannot.

Secondly, regression analysis was used to calculate correlations between indicators of vegetation quality and welfare, between regions, and within regions over time. Results showed that vegetation quality and welfare are negatively correlated between regions but are positively correlated over time within rural areas. This highlighted cross-sectional selection effects and also that RS based indicators of vegetation quality hold promise in tracking poverty over time and can do so in areas that night lights cannot. The study was based on a small number of observations, and a larger scale study is needed to verify the findings. The results in this study do justify that a large scale study should be undertaken.

The results also highlight the dependence of the poor in SSA on climate. It raises the question whether decreases in poverty rates were actually due to economic changes or coincidentally bad weather in one survey period and good weather in another. It also suggests that poverty alleviation and prevention policies should focus on poor households' vulnerability to climate change. Of course, effective policies of this kind would imply weaker correlations between vegetation quality and welfare, but RS based vegetation indices could then be used to measure the effectiveness of such policies.

References

Anthony, M., 2015. Night lights and regional income inequality in Africa. Working Paper Series UNU-WIDER Research Paper, World Institute for Development Economic Research (UNU-WIDER), available at `https://ideas.repec.org/p/unu/wpaper/wp2015-085.html`.

Bhattacharya, H., Innes, R., nov 2012. Income and the Environment in Rural India: Is There a Poverty Trap? American Journal of Agricultural Economics 95 (1), 42–69.

Blumenstock, J., Cadamuro, G., On, R., nov 2015. Predicting poverty and wealth from mobile phone metadata. Science 350 (6264), 1073–1076.

Cao, Z., Wu, Z., Kuang, Y., Huang, N., Wang, M., jan 2016. Coupling an Intercalibration of Radiance-Calibrated Nighttime Light Images and Land Use/Cover Data for Modeling and Analyzing the Distribution of GDP in Guangdong, China. Sustainability 8 (2), 108.

CBS, November 2006. 2003 / 2004 Namibia Household Income & Expenditure Survey: Main Report. Central Bureau of Statistics, Windhoek, Namibia.

Chen, X., Nordhaus, W. D., may 2011. Using luminosity data as a proxy for economic statistics. Proceedings of the National Academy of Sciences 108 (21), 8589–8594.

Devereux, S., Naeraa, T., 1996. Drought and Survival in Rural Namibia. Journal of Southern African Studies 22 (3), 421–440.

Ebener, S., Murray, C., Tandon, A., Elvidge, C. C., 2005. From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery. International Journal of Health Geographics 4, 5 – 17.

Elbers, C., Lanjouw, J. O., Lanjouw, P., 2003. Micro-level Estimation of Poverty and Inequality. Econometrica 71 (1), 355 – 364.

Elvidge, C., Hsu, F.-C., Baugh, K., Ghosh, T., may 2014. National Trends in Satellite-Observed Lighting: 1992–2012. In: Remote Sensing Applications Series. CRC Press, pp. 97–120.

Elvidge, C. D., Baugh, K. E., Anderson, S. J., Sutton, P. C., Ghosh, T., 2012. The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data. Social Geography 7 (1), 23–35.

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R., Davis, C. W., 1997. Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. International Journal of Remote Sensing 18 (6), 1373–1379.

Elvidge, C. D., Sutton, P. C., Ghosh, T., Tuttle, B. T., Baugh, K. E., Bhaduri, B., Bright, E., 2009. A global poverty map derived from satellite data. Computers & Geosciences 35 (8), 1652 – 1660.

FAO, 2010. Crop calendar - An information tool for seed security. Online, available at `http://www.fao.org/agriculture/seed/cropcalendar/cropcalendar.do[Accessedon20/07/2017]`.

Fara, K., 2001. How Natural Are 'Natural Disasters'? Vulnerability to Drought of Communal Farmers in Southern Namibia. Risk Management 3 (3), 47–63.

Foster, A. D., Rosenzweig, M. R., 2003. Economic Growth and the Rise of Forests. The Quarterly Journal of Economics 118 (2), 601–637.

Ganguly, S., Friedl, M. A., Tan, B., Zhang, X., Verma, M., 2010. Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. Remote Sensing of Environment 114 (8), 1805 – 1816.

Henderson, J. V., Storeygard, A., Weil, D. N., April 2012. Measuring Economic Growth from Outer Space. American Economic Review 102 (2), 994–1028.

Huang, J., Han, D., 2014. Meta-analysis of influential factors on crop yield estimation by remote sensing. International Journal of Remote Sensing 35 (6), 2267–2295.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., Ferreira, L., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment 83 (1), 195 – 213, the Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring.

Imran, M., Stein, A., Zurita-Milla, R., 2014. Investigating rural poverty and marginality in Burkina Faso using remote sensing-based products. International Journal of Applied Earth Observation and Geoinformation 26, 322 – 334.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., Ermon, S., 2016. Combining satellite imagery and machine learning to predict poverty. Science 353 (6301), 790–794.

Klemens, B., Coppola, A., Shron, M., Jun. 2015. Estimating local poverty measures using satellite images : a pilot application to Central America. Policy Research Working Paper Series 7329, The World Bank.

Lobell, D. B., 2013. The use of satellite data for crop yield gap analysis. Field Crops Research 143, 56 – 64.

Mkhabela, M., Bullock, P., Raj, S., Wang, S., Yang, Y., 2011. Crop yield forecasting on the Canadian Prairies using {MODIS} {NDVI} data. Agricultural and Forest Meteorology 151 (3), 385 – 393.

Mkhabela, M. S., Mkhabela, M. S., Mashinini, N. N., 2005. Early maize yield forecasting in the four agro-ecological regions of Swaziland using {NDVI} data derived from NOAA's-AVHRR. Agricultural and Forest Meteorology 129 (1–2), 1 – 9.

Morikawa, R., 2014. Remote Sensing Tools for Evaluating Poverty Alleviation Projects: A Case Study in Tanzania. Procedia Engineering 78, 178 – 187, humanitarian Technology: Science, Systems and Global Impact 2014, HumTech2014.

Myneni, R., Hall, F., Mar 1995. The interpretation of spectral vegetation indexes. Geoscience and Remote Sensing, IEEE Transactions on 33 (2), 481–486.

NASA LP DAAC, August 2012a. MOD12Q2 Land Cover Dynamics – Known Issues. Available at `http://www.bu.edu/lcsc/files/2012/08/MCD12Q2_Known_Issues.pdf`.

NASA LP DAAC, August 2012b. User Guide for the MODIS Land Cover Dynamics Product (MCD12Q2). Available at `http://www.bu.edu/lcsc/files/2012/08/MCD12Q2_UserGuide.pdf`.

NASA LP DAAC, 2017a. MODIS Terra Vegetation Indices 16-Day L3 Global 500m SIN Grid. Version 5. (MOD13A1). NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (https://lpdaac.usgs.gov), accessed on 23 May 2017 at `https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13a1`.

NASA LP DAAC, 2017b. V005 MODIS Land Cover Dynamics (MCD12Q2).

National Geophysical Data Center, 2016. Version 4 DMSP-OLS Nighttime Lights Time Series. Available at `https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html`.

NPC, April 2015. Namibia Poverty Mapping. Republic of Namibia, National Planning Commission, Windhoek, Namibia.

NSA, November 2012. Poverty Dynamics in Namibia: A comparative study using the 1993/94, 2003/04 and the 2009/10 NHIES surveys. Namibia Statistics Agency, Windhoek, Namibia.

NSA, July 2013. Namibia - National Household Income and Expenditure Survey 2009-2010: Metadata. Namibia Statistics Agency, Windhoek, Namibia.

NSA, November 2015. Namibia Census of Agriculture 2013/2014: Communal Sector Report. Namibia Statistics Agency, Windhoek, Namibia.

Odendaal, W., 2011. Land Grabbing in Namibia: A case study from the Omusati region, northern Namibia. Paper presented at the international conference on global land grabbing. 6-8 april 2011, Land Deals Politics Initiative, University of Sussex.

Pfaff, A. S., 1999. What Drives Deforestation in the Brazilian Amazon? Journal of Environmental Economics and Management 37 (1), 26 – 43.

Pinkovskiy, M., Sala-i-Martin, X., feb 2016. Lights, Camera ... Income! Illuminating the National Accounts-Household Surveys Debate. The Quarterly Journal of Economics 131 (2), 579–631.

Skole, D., Tucker, C., jun 1993. Tropical Deforestation and Habitat Fragmentation in the Amazon: Satellite Data from 1978 to 1988. Science 260 (5116), 1905–1910.

Small, C., Elvidge, C. D., Balk, D., Montgomery, M., 2011. Spatial scaling of stable night lights. Remote Sensing of Environment 115 (2), 269 – 280.

Sutton, P. C., Costanza, R., June 2002. Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. Ecological Economics 41 (3), 509–527.

Sutton, P. C., Elvidge, C. D., Ghosh, T., 2007. Estimation of Gross Domestic Product at Sub-national Scales Using Nighttime Satellite Imagery. International Journal of Ecological Economics and Statistics 8, 5 – 21.

USGS, November 2015. Landsat - Earth Observation Satellites. Fact Sheet 2015-3081, U.S Geological Survey.

Villa, J. M., oct 2016. Social Transfers and Growth: Evidence from Luminosity Data. Economic Development and Cultural Change 65 (1), 39–61.

Watmough, G. R., Atkinson, P. M., Hutton, C. W., 2013. Predicting socioeconomic conditions from satellite sensor data in rural developing countries: A case study using female literacy in Assam, India. Applied Geography 44, 192 – 200.

Watmough, G. R., Atkinson, P. M., Saikia, A., Hutton, C. W., 2016. Understanding the Evidence Base for Poverty Environment Relationships using Remotely Sensed Satellite Data: An Example from Assam, India. World Development 78, 188 – 203.

Weeks, J. R., Getis, A., Stow, D. A., Hill, A. G., Rain, D., Engstrom, R., Stoler, J., Lippitt, C., Jankowska, M., Lopez-Carr, A. C., Coulter, L., Ofiesh, C., sep 2012. Connecting the Dots Between Health, Poverty and Place in Accra, Ghana. Annals of the Association of American Geographers 102 (5), 932–941.

Werner, W., 1993. A brief history of land disposession in Namibia. Journal of Southern African Studies 19 (1), 135.

World Bank, Oct. 2008. Republic of Namibia - Addressing Binding Constraints to Stimulate Broad Based Growth : A Country Economic Report. World Bank Other Operational Studies 12601, The World Bank.

World Bank, 2015. World Development indicators. Available at http://databank.worldbank.org/data/views/variableSelection/selectvariables.aspx?source=world-development-indicators.

Zhang, X., Friedl, M. A., Schaaf, C. B., dec 2006. Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. Journal of Geophysical Research: Biogeosciences 111 (G4).

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C., Gao, F., Reed, B. C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sensing of Environment 84 (3), 471 – 475.