Measuring rural electrification with satellite data

Takwanisa Machemedze, Taryn Dinkelman, Mark Collinson, Wayne Twine and Martin Wittenberg

Overview

- How do we measure development in the absence of high quality social survey information?
- Argued that night light data captured by satellites can provide a reliable measure of development in contexts where:
  - other measures do not exist or
  - where the quality of the data is such that the statistics cannot be trusted.

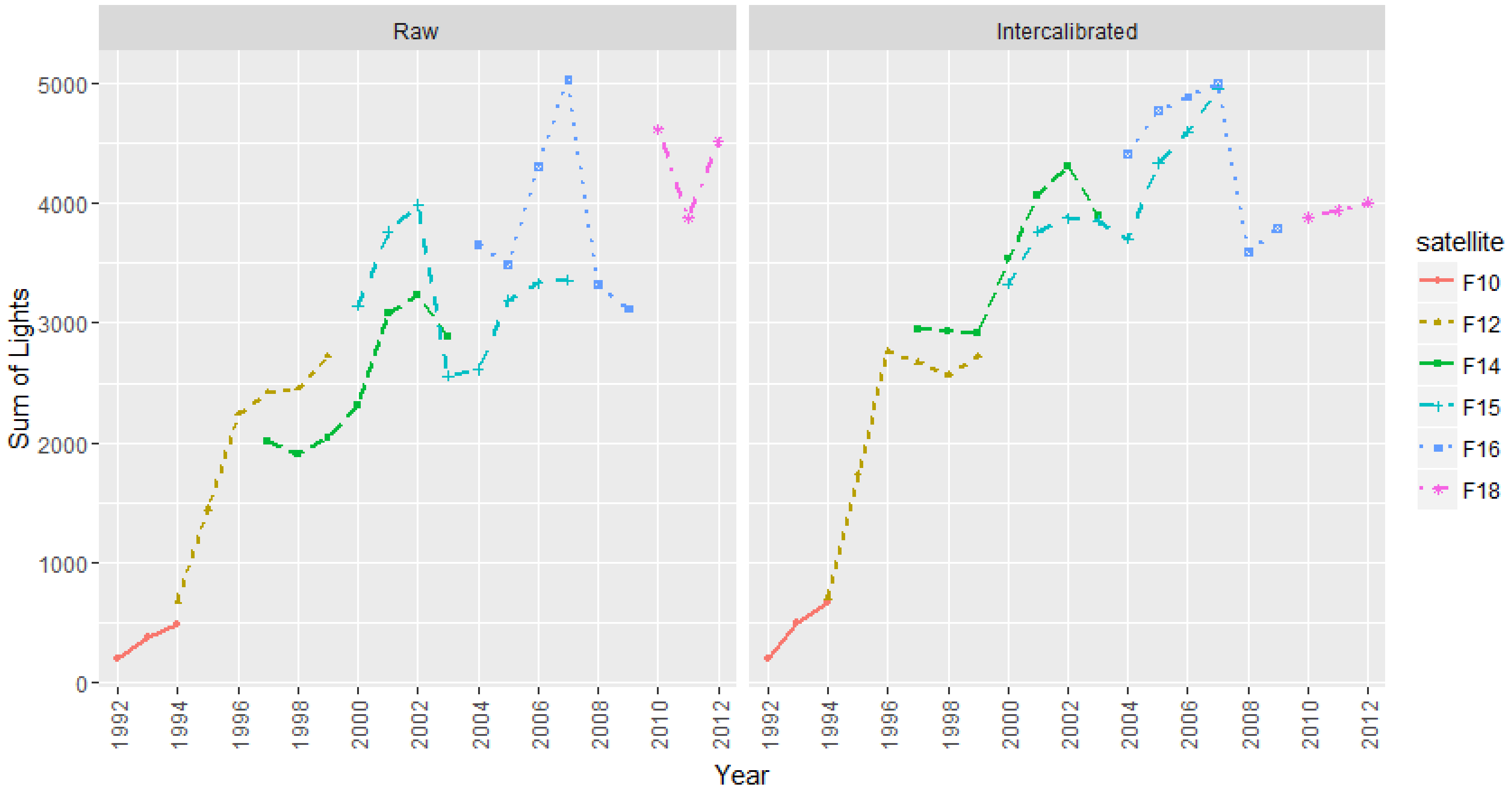
What do we want to test

a) Does the satellite data pick up the temporal patterns of rural electrification?
b) Does the satellite data pick up the spatial patterns of electrification? Can it pick up the difference between developed and undeveloped areas? And at what spatial resolution?
c) What is the correlation between the satellite data and the household electrification data?

Data

- Data
  - DMSP-OLS night light data
  - 6 satellites over 21 years (1992-2012)
  - Pixel values ranging from 0 (absents of light) to 63 (light saturation)
  - Inter-calibration – to a common range defined by a reference year
    - no on-board calibration,
    - there are differences in the performance of instruments,
    - different sensors had different detection limits and saturation radiances.
  - sum of the digital values for a given area as the measure of light for that area.
  - Also referred to as Sum of Light (SOL) or total night light (TNL)

Raw data vs Intercalibration



Data

- SOL - sensitive to the size of the area over which it is calculated
- For comparison between areas, we use standardised SOL, i.e. SOL/pixels
- weighted by the proportion of the pixel within the polygon
- Agincourt HDSS study site census data
  - household assets module conducted every second year since 2001
- Village typology (Hargreaves 2000) – classifying villages as:
  - a) “Central communities” b) “Established communities” c) “Undeveloped villages” and d) “Refugee settlements”
  - First two categories had electricity access in 2000

Methods - Temporal variation in the night light data

- First cut is to look at temporal trends in nightlight data
- However, increasing trends do not establish electrification
- Use a ready made counterfactual – the Kruger National Park
  - KNP not electrified except for isolated camp sites
  - difference-in-difference estimation strategy
- $\frac{sol}{pixel} = \beta_1 + \beta_2 t + \gamma Agincourt + \delta t * Agincourt + \varepsilon \qquad (1)$
  - Measure of interest is delta
  - Also used Nelspruit (an established urban area) as a counterfactual and pooled data (Nelspruit and KNP)

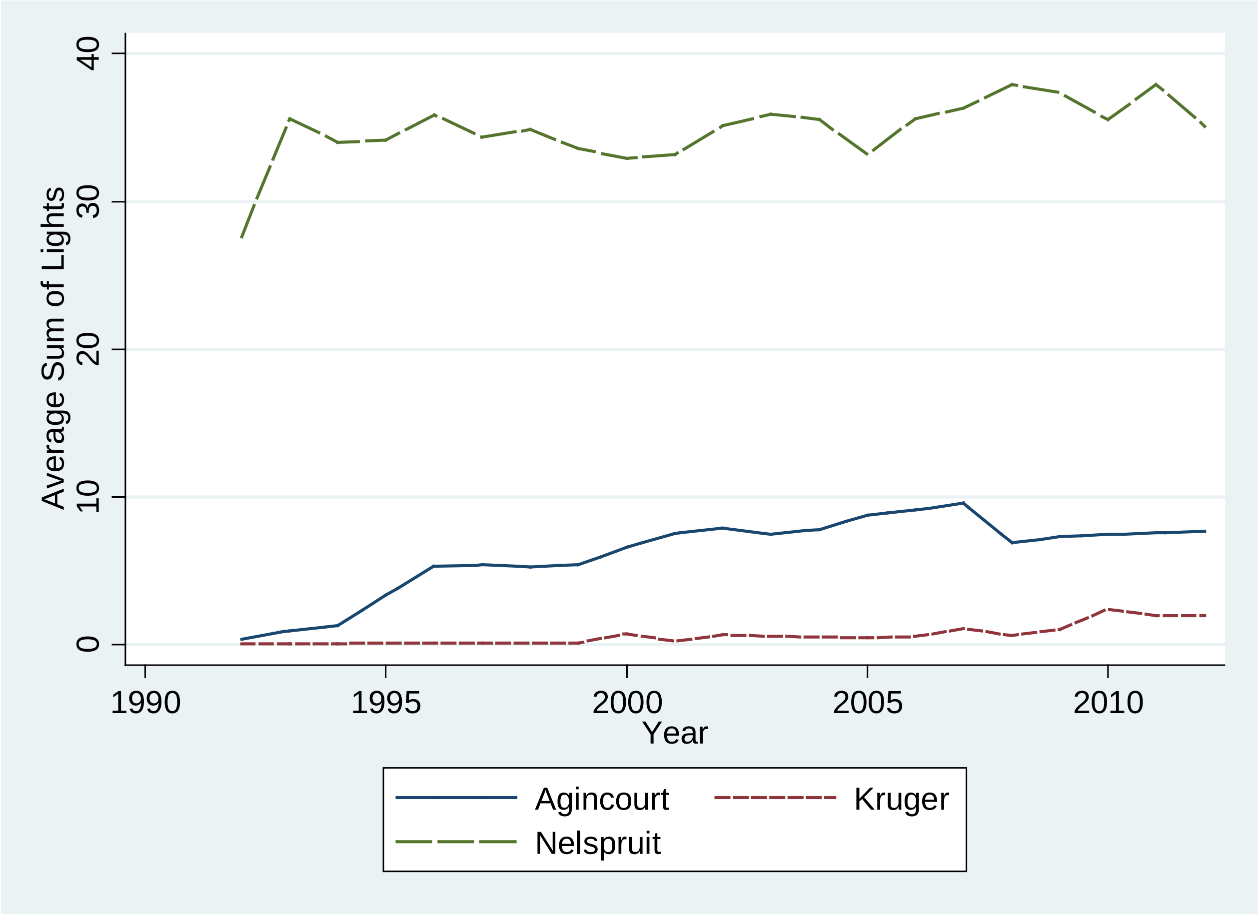
Methods -Spatial variation in the night light data

- circumstantial evidence of electrification
- Check differences between village typologies, i.e. electrified versus not.
- Hargreaves(2000) typology
  - type1-central, type2- established, type3 – underdeveloped, type4 -refugee
- $\frac{sol}{pixel} = \beta_1 + \beta_2 type2 + \beta_3 type3 + \beta_4 type4 + \varepsilon \qquad (2)$
- Also test nightlight data with direct measure of HH electricity connections
- $sol/pixel = \beta_1 + \beta_2 connections/pixel + \varepsilon \qquad (3)$

Methods - Variation in space and time

- A model that use both cross-sectional and temporal variation
- $sol/pixel_{it} = \beta_1 + \beta_2 connections/pixel_{it} + \theta_i + \eta_t + \varepsilon_{it}$ (4)
  - where i subscripts village and t year. The θi terms are village fixed effects, ηt year fixed effects and εit is an idiosyncratic error.

Results

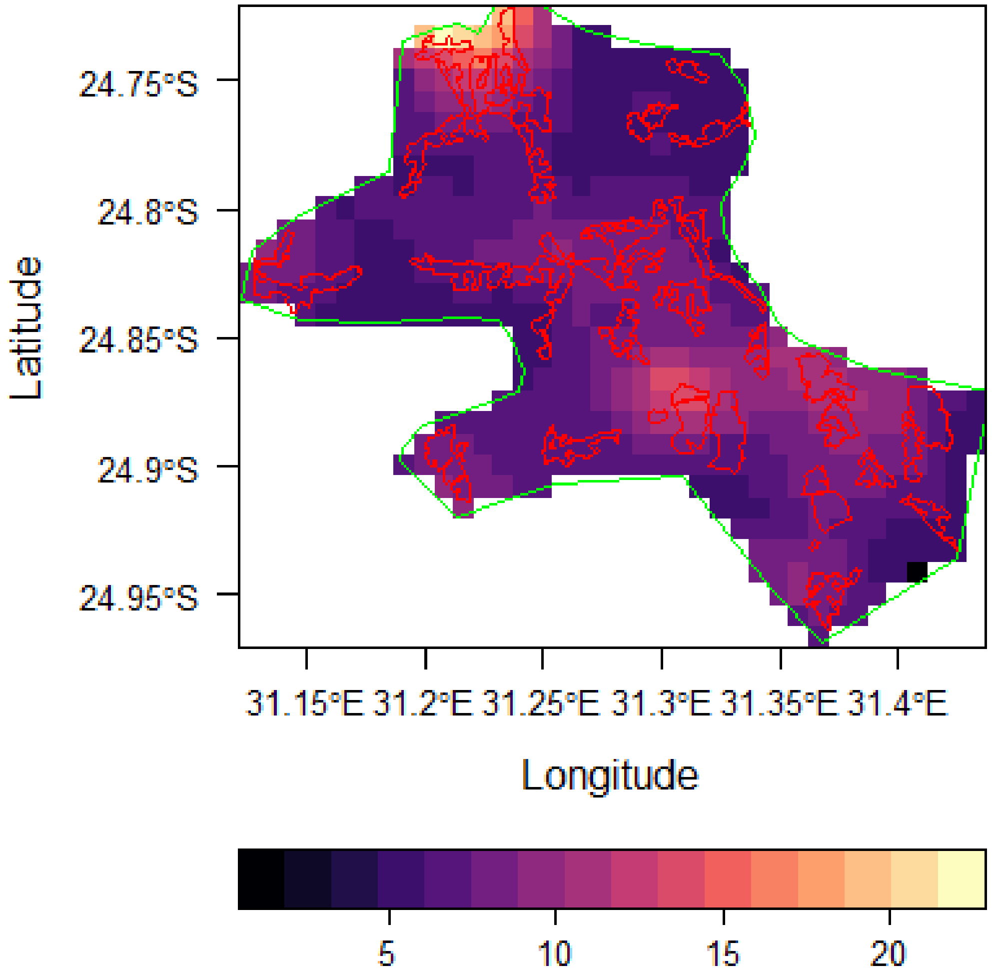


Average Sum of Lights
0
10
20
30
40
1990
1995
2000
2005
2010
Year
Agincourt
Kruger
Nelspruit

Results

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Agincourt	Kruger	Nelspruit	Pooled	Pooled
t	0.348**	0.0932**	0.211*	0.0932**	0.152**
	(0.0649)	(0.0164)	(0.0889)	(0.0167)	(0.0429)
Agincourt				2.690**	3.336**
				(0.923)	(0.899)
Nelspruit				32.88**	34.18**
				(1.312)	(0.453)
t*Agincourt				0.255**	0.196*
				(0.0776)	(0.0722)
t*Nelspruit				0.118	
				(0.0970)	
Constant	2.322**	-0.368*	32.52**	-0.368*	-1.014*
	(0.799)	(0.140)	(1.240)	(0.142)	(0.481)
Observations	21	21	21	63	63
R-squared	0.670	0.696	0.355	0.992	0.992
Robust standard errors in parentheses					
** p<0.01, * p<0.05					

Results – variation in space (2012 image)



Results – variation in space

	(1)	(2)	(3)	(4)
			without Village X	
VARIABLES	2000	Pooled	2000	Pooled
established	-0.0646	-0.0949	-0.0646	-0.0949
	(0.991)	(0.454)	(0.998)	(0.454)
undeveloped	-3.356*	-1.774**	-3.356*	-1.774**
	(1.156)	(0.455)	(1.164)	(0.455)
refugee	-1.829	-0.296	-3.086**	-1.219*
	(1.545)	(0.506)	(0.957)	(0.511)
Constant	8.842**	7.544**	8.842**	7.544**
	(0.797)	(0.321)	(0.802)	(0.321)
Observations	20	420	19	399
R-squared	0.360	0.041	0.568	0.049

Village types taken from Hargreaves (2000). Base category is “central community”
Robust standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

Results – variation in space

Panel A:	(1)	(2)	(3)	(4)	(5)	(6)
All villages	2001	2003	2005	2007	2009	2011
Connections/pixel	0.00875	0.00961*	0.00113	0.00475	0.00150	0.00231+
	(0.00564)	(0.00424)	(0.00260)	(0.00389)	(0.00123)	(0.00122)
Constant	7.249**	6.557**	9.766**	9.439**	7.581**	7.601**
	(0.928)	(0.985)	(0.638)	(0.952)	(0.375)	(0.451)
Observations	21	21	21	25	27	27
R-squared	0.124	0.233	0.007	0.068	0.016	0.045
Panel B:	(1)	(2)	(3)	(4)	(5)	(6)
Without Village X	2001	2003	2005	2007	2009	2011
Connections/pixel	0.0126*	0.0124**	0.00326	0.00750*	0.00144	0.00243+
	(0.00443)	(0.00334)	(0.00209)	(0.00303)	(0.00136)	(0.00126)
Constant	6.476**	5.865**	9.220**	8.732**	7.495**	7.440**
	(0.585)	(0.761)	(0.546)	(0.759)	(0.390)	(0.440)
Observations	20	20	20	24	26	26
R-squared	0.322	0.422	0.061	0.169	0.017	0.060
Robust standard errors in parentheses						
** p<0.01, * p<0.05, + p<0.1						

Results - both the temporal and spatial variation

	(1) Pooled	(2) Pooled	(3) Pooled
connect/pixel	0.00413**	0.00419**	0.00720**
	(0.00128)	(0.00128)	(0.00189)
2001b.year			
2003.year	-0.429	-0.431	-0.521
	(0.703)	(0.673)	(0.493)
2005.year	1.422*	1.417*	1.193**
	(0.607)	(0.598)	(0.426)
2007.year	1.777**	1.797**	1.890**
	(0.657)	(0.640)	(0.407)
2009.year	-0.764	-0.898	-0.837
	(0.577)	(0.566)	(0.426)
2011.year	-0.642	-0.778	-0.841
	(0.600)	(0.588)	(0.429)
Village Effects	N	N	Y
Quartic in pixels	N	Y	N.A.
Constant	7.779**	11.00**	7.151**
	(0.542)	(1.281)	(0.397)
Observations	142	142	142
R-squared	0.272	0.319	0.761
Robust standard errors in parentheses			
** p<0.01, * p<0.05			

Conclusion

- nightlight data seems to have captured the electricity roll-out in the Agincourt study site.
- shows marked increases in brightness over time and captures the broad differences between “developed” and “undeveloped” parts of the site
- Also shown some measurement issues which contaminate the relationship