# SATELLITE IMAGERY FOR STUDYING DEVELOPMENT? THE ITALIAN CASE STUDY

Francesco Maria Chelli, Alessio Guandalini

# Introduction

The number of active satellites currently orbiting earth is about 1,100. Among them, about 60% are used for communications while the remaining aids in weather forecasting, national defense, science and agriculture. For the high volume, velocity and variety of information gathered by them, satellite data are commonly referred as big data.

Since when, in 1972, the Earth Resources Technology Satellite (later re-named LANDSAT) was launched, several projects for cartography, geology, foresting, hydrology, agriculture have been started. Also statisticians and National Statistical Institutes began to use satellite imagery for mapping territory. For instance, the Italian National Institute of Statistics (ISTAT) used them in occasion of the 4<sup>th</sup> Agricultural Census (Benedetti and Ciovatella, 2006) and the Australian Bureau of Statistics (ABS) recently evaluated some methodological approaches for estimating crop yield and land cover (Marley et al., 2014). Another application field, still close to mapping territory, firstly attempted in developed and recently in developing country, is the measure of the urban sprawl (Sudhira et al., 2004).

The advantage in using satellite data is strictly related to their characteristics. They overcome some problems of the traditional data collection process in these fields, such as difficulty in gathering data in hard to reach places, homogeneity, rough approximation of measures and also non-response. Moreover, satellite data are inexpensive<sup>1</sup>, periodically collectable<sup>2</sup>, spatially explicit<sup>3</sup> as well as globally available.

What has been just said is clear and intuitive for land cover problems, but it is more cumbersome to understand for socio-economic parameters. However, in a preliminary work Doll et al. (2000) and Doll et al. (2006) demonstrate that

<sup>&</sup>lt;sup>1</sup> Several organizations, owners of satellite data, release them for free. For Instance, they are free downloadable from NASA (National Aerounatics and Space Administration, http://www.nasa.gov), ESA (European Space Agency, http://www.nasa.gov), etc.

The release are carried out periodically.

<sup>&</sup>lt;sup>3</sup> The satellite data are georeferenced. They are aligned to a known coordinate system, that is each value is unequivocally linked to a point in the map with specific latitude and longitude.

nighttime satellite data can be used to estimate global urban population, gross domestic product (GDP), total carbon dioxide and economic activity level. In fact, these parameters are highly related to brightness of nocturnal lighting (see Elvidge *et al.*, 2012 and references therein).

Elvidge *et al.* (2012) derived an empirical measurement of human development solely from nighttime satellite imagery and population density. The Night Light Development Index (*NLDI*), they proposed, has strong correlation with the Human Development Index (*HDI*) that is a key statistic for United Nation, World Bank and other International Agencies to properly analyse the development of countries.

*NLDI* is based on global satellite maps of lighting that can be produced in a consistent, repeatable manner on an annual basis and, therefore, it overcomes some limitations of *HDI*. The *HDI* is a well-known index proposed in 90s to go beyond the simple idea of development measured in terms of national income (see e.g. UNDP, 2011). *HDI* can be affected by differences in data collection among countries. Moreover, for some countries old information only are available and for few others not even that (Wolff *et al.*, 2011). Finally, through *HDI*, the difference existing within a country can be more complicated to appreciate.

In the present work, we present a slight modification of *NLDI* for overcoming some drawbacks. Both indexes, *NLDI* and our proposal, have been computed not only at national level, but also for each geographical area (NUTS-1), region (NUTS-2). In particular for Marche region the indexes have been computed at provincial level (NUTS-3). Some questions about the interpretation of the indexes in a particular context like the Italian one, especially at subnational level, arose.

The paper is organized as follows. In Section 1, the methodology on the basis of *NLDI* is presented. In Section 2, devices for overtaking some limits in the actual *NLDI* definition are introduced, while the data used are extensively illustrated in Section 3. Finally, in Section 4, the values of *NLDI* are shown and an interpretation for the specific Italian case is given. In this section also a brief focus on the Marche region is presented. Section 5 contains concluding remarks.

#### 1. The Night Light Development Index (NLDI)

The Night Light Development Index (*NLDI*) has been introduced by Elvidge *et al.* (2012). The authors showed that it is strongly correlated with the Human Development Index (*HDI*) at country level. *NLDI*'s peculiarity is due to the data used. In fact, *NLDI* does not use monetary measures of wealth, but nighttime satellite and population density data.

*NLDI* is based on an idea built up over the years, since when in 1994 the National Geophysical Data Center (NOAA) started to derive and study satellite

maps of nighttime lights from data collected in the US Air Force Defence Meteorological Satellite Program (DMSP). Nocturnal lights are proxy of public goods, services, pavements, built infrastructures and economic activities. It can be assumed that people living in brightly lit areas have easier access to goods and services than people living in "dark" areas. Therefore, they have better living conditions. The more brighter and diffused – with respect to the number of "lit inhabitant" – the light, larger will be the development.

Ultimately, Elvidge *et al.* (2012) assume that the level of equality in the distribution of outdoor lighting among inhabitants, measured through the Gini index (Gini, 1914), can be considered a development index, indeed the *NLDI*. To take a deeper dive into the meaning of the index, let us show how the *NLDI* is derived.

The inputs of *NLDI* are two geospatial grids. The one of the nighttime lights contains the radiance level for each cell derived from the satellite imagery<sup>4</sup>, while the one of the related number of inhabitants, the population count in each cell available from Census. What is really important is that the two grids must be spatially co-registered, that is, cells must have same size and coordinate.

<b>Figure 1</b> – Radiance l	ight level	grid (a) and	population a	lensity g	rid (b).
------------------------------	------------	--------------	--------------	-----------	----------

	R	adi	anc	Gr æle	id a evel	<b>i</b> l pe	er co	ell		<b>Grid b</b> Population count pe				er cell					
o	o	o	o	o	o	o	o	o	0	0	0	0	0	0	0	0	0	0	0
o	0	0	o	0	0	o	0	o	o	0	0	1	1	2	1	2	1	4	1
0	0	0	0	0	0	0	0	o	0	1	1	3	2	3	4	4	4	3	1
0	0	0	0	0	0		0	o	0	1	1	4	5	5	5	10	5	5	2
0	0	0	0			10	1		o	1	2	5	5	10	10	10	10	10	3
0	0	0	0		50	150	50		0	1	3	5	5	10	30	50	30	10	4
0	0	0	0	10	150	255	150	10	2	2	2	5	10	10	50	100	50	10	10
5	0	0	0	2	50	150	50		3	15	5	5	5	10	30	50	30	10	3
0	0	0	0	0	5	10		2	0	4	3	5	5	10	10	10	10	10	2
0	0	0	0	0	0	0	0	0	0	0	2	5	5	5	5	10	5	5	1

In Figure 1, the example in Elvidge *et al.* (2012, p. 25) is reported. It is related to a very typical urban situation, in which, from the city center to suburbs, brightness and population density decrease. Grid a is related to the radiance level. The values in grid a are between 0 ("black" pixels, minimum radiance) and 255 ("white" pixels, maximum radiance). Grid b contains the number of inhabitants in each cell.

<sup>&</sup>lt;sup>4</sup> In some cases some processes to clean up the map and bring out outdoor lights are needed.

The data in the two grids are aggregate to have a tabular list with both radiance level and population count. Then, the data are sorted increasingly by brightness light level and aggregated in radiance level classes. A tabular form such as in Table 1 is obtained.

	Light level	Population	Cum %	Cum %
1	DN	in DN	of light	of population
1	0	205	0.0	25.0
2	1	10	0.1	26.2
3	2	90	1.5	37.2
4	5	55	3.7	43.9
5	10	40	7.2	48.8
6	50	120	24.6	63.4
7	150	200	77.6	87.8
8	255	100	100.0	100.0
Total	1,137	820		

**Table 1** – Population count per light level for the example in Figure 1.

Table 1 is a frequency distribution and the Gini index (Gini, 1914),

$$R = 1 - \frac{2\sum_{i=1}^{n-1} Q_i}{n-1}, \qquad 0 \le R \le 1$$
(1)

is used to measure the equality in distribution of the light. R = NLDI, where  $Q_i = \sum_{j=1}^{i} x_j / \sum_{j=1}^{n} x_j$  is the proportion of lights in the hands of the proportion  $P_i$  of inhabitants,  $P_i = \sum_{j=1}^{i} x_j / n$ . The closer is NLDI to 0, more developed the area can be considered. In the example NLDI=0.672 that, following the interpretation given by Elvidge *et al.* (2006), it denotes a middle and lower development in the area.

# 2. An improvement of NLDI

The index, as defined in the previous section, assumes its lower value when the lights are evenly distributed among inhabitants. Instead, it is equal to its maximum, when one person has lights and the rest lives "in the dark".

Table 2 gives a partial idea of the behavior of *NLD1* in extreme cases. It is noticeable that *NLD1*, as proposed in Elvidge *et al.* (2012), assumes the same value (0 or 1) in very different situations, independently of light brightness.

Therefore, to overcome this drawback, a penalization that takes into account the average of light brightness, is introduced.

146





*Note:*  $\mu(x_i)$  *is the weighted mean (with respect to the population count of the cell) of radiance level registered in the area.* 

Among the several possible solutions for the correction of NLDI, one of the simpler has been chosen.  $NLDI^*$  is defined as follow:

$$NLDI^* = \left(\frac{\mu(x_i)}{255}\right) NLDI + \left(1 - \frac{\mu(x_i)}{255}\right) \qquad 0 \le NLDI^* \le 1$$
 (2)

where  $\mu(x_i)$  is the weighted mean (with respect to the population count in the cell) of radiance level registered in the area. The index in (2) discounts *NLDI* in (1) by the ratio between  $\mu(x_i)$  and his maximum, 255 – all the population units live in an area with 255 of radiance level. Therefore, the weaker the light is, less important the inequality in the lights distribution is. Furthermore, a term that takes into account the brightness level in the area is added. In these way, it is possible to discriminate the scenarios in case 1. In fact, *NLDI*<sup>\*</sup> is equal to 0 for scenario 1a, 1 for 1c and 0.5 for 1b. The same happens for case 2, in which scenarios 2a (*NLDI*<sup>\*</sup>=0) is related to a better situation than 2b (*NLDI*<sup>\*</sup>>0). Scenario 2c has the same *NLDI*<sup>\*</sup> of 1c. In case 3a and 3b, *NLDI*<sup>\*</sup>=1.

As *NLD1*, *NLD1*<sup>\*</sup> is defined between 0 and 1. In general, in case of maximum dispersion of lights (*NLD1*=1) and when  $\mu(x_i)=0$  – the area is completely "in the dark" – *NLD1*<sup>\*</sup> =1. Moreover, same level of inequality (same *NLD1*) corresponds to different level of development depending on the level of light brightness in the area measured by  $\mu(x_i)$ .

# 3. Data

To derive *NLD1* and *NLD1*<sup>\*</sup>, data on nighttime light and on population density are needed. NASA makes the nighttime satellite imagery available for free from its website (http://earthobservatory.nasa.gov/Features/NightLights/page3.php). The imagery (see Figure 2, a) have been caught through the Suomi NPP satellite between April and October 2012. The satellite is equipped with a spectroradiometer called VIIRS (Visible Infrared Imaging Radiometer Suite). VIIRS detects photons of light in 22 different wavelength bands and filters them to distinguish even isolated highways lamp, fishing boat, faint and nocturnal atmospheric light. In fact, the use of VIIRS allowed to improve ten to fifteen times the resolution with respect old systems. For each pixel, corresponding to a square of 742 meters (0.46 miles), the outdoor light have been isolated and the radiance light level have been derived.

The data on population density are provided by EUROSTAT. Tanks to GEOSATA 1A project the results of the 2011 year European censuses have been plotted on km<sup>2</sup> grids. One of these grids gathers the data on population count, in

which each cell contains the numbers of inhabitants per km<sup>2</sup>. The grid is downloadable for free from EUROSTAT website (http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/populationdistribution-demography). In Figure 2(b) a representation of the population density for the Italian territory is given.

**Figure 2** – Nighttime satellite imagery (a) and population by square kilometer (b). Italy, year 2011..



Snippet from NASA Night LightsOur elaboration on GEOSTAT Population grid (Eurostat -2012, Flat MapOctober 2011).

The difference in time between the two datasets is slight and it can be assumed negligible. Therefore, the two grids have been properly overlapping through the open-source software QGIS and a double distribution of radiance level and population count in cells to derive the indexes has been produced.

### 4. The *NLDI* in Italy

The *NLDI* and the proposal modification have been applied to the Italian case. The indexes have been computed for the whole territory, the geographical areas (NUTS-1), the regions (NUTS-2) and few provinces (NUTS-3). The regional values of *NLDI* and *NLDI*<sup>\*</sup> have been compared with the *HDI* (UNDP, 2011) and the *QUARS*<sup>5</sup> (Sbilanciamoci!, 2011).

Looking at Table 3, *NLDI* and *NLDI*<sup>\*</sup> produce similar rankings for the regions even if their values are quite different. The correlation between the two indexes is

<sup>&</sup>lt;sup>5</sup> The QUARS is an index that takes into account 41 macro-indicators related to 7 dimensions: environment, economy and work, Rights and citizenship, health, education and culture, equal opportunities, participation. It is computed for all the Italian regions (NUTS-2).

obviously close to 1, because  $NLDI^*$  is a linear transformation of NLDI. since at nuts-1 level the original interpretation of nldi in Elvidge seems do not hold. The original interpretation of the NLDI state that lower NLDI is, more developed the area is (the same should hold for  $NLDI^*$  in light of above). The most developed regions would be Campania, whilst among the less developed would be Bolzano and Trento.

	NLDI		NLDI	*	ны		QUARS		
				N LD I			enviroment		
Piedmont	0.375	(17)	0.717	(17)	0.876	(11)	0.19	(16)	
Aosta Valley	0.421	(16)	0.905	(20)	0.885	(16)	1.15	(21)	
Lombardy	0.139	(7)	0.331	(5)	0.895	(19)	-0.30	(5)	
Bolzano	0.594	(21)	0.932	(21)	0.883	(14)	0.96	(19)	
Trento	0.531	(20)	0.888	(19)	0.877	(12)	0.96	(20)	
Veneto	0.185	(11)	0.410	(10)	0.884	(15)	-0.08	(10)	
Friuli Venezia Giulia	0.285	(15)	0.608	(14)	0.890	(17)	0.00	(11)	
Liguria	0.409	(18)	0.750	(18)	0.878	(13)	-0.33	(4)	
Emilia-Romagna	0.130	(4)	0.330	(4)	0.875	(10)	0.00	(12)	
Toscany	0.151	(8)	0.357	(8)	0.874	(9)	0.26	(17)	
Umbria	0.155	(9)	0.384	(9)	0.894	(18)	0.01	(13)	
Marche	0.110	(3)	0.296	(3)	0.870	(8)	0.03	(14)	
Lazio	0.082	(2)	0.228	(2)	0.852	(7)	-0.34	(3)	
Abruzzo	0.132	(6)	0.344	(7)	0.833	(1)	0.32	(18)	
Molise	0.264	(13)	0.576	(13)	0.840	(4)	-0.27	(6)	
Campania	0.059	(1)	0.183	(1)	0.849	(5)	-0.43	(2)	
Apulia	0.131	(5)	0.339	(6)	0.840	(3)	-0.69	(1)	
Basilicata	0.273	(14)	0.625	(15)	0.836	(2)	0.03	(15)	
Calabria	0.220	(12)	0.507	(12)	0.850	(6)	-0.10	(9)	
Sicily	0.179	(10)	0.410	(11)	0.897	(20)	-0.27	(7)	
Sardinia	0.324	(16)	0.634	(16)	0.897	(21)	-0.16	(8)	
Nord-West	0.257		0.519		0.885		-		
Nord-East	0.226		0.476		0.887		-		
Centre	0.116		0.293		0.881		-		
South	0.137		0.333		0.848		-		
Island	0.218		0.472		0.843		-		
Italy	0.193		0.423		0.874		-		

Table 3 – NLDI, NLDI\*, HDI and QUARS for Italian geographical areas (NUTS-1) and<br/>regions (NUTS-2), year 2011.

Note: in brackets the position of regions in the ranking ordered increasingly for each index.

It seems wrong. In fact, the two indexes are weak correlated with the regional HDI (correlation lower than 0.3) and moreover, in this case we have a positive correlation, while in Elvidge *et al.* (2012) is stressed be a negative correlation (-

150

0.71). *NLDI* and *NLDI*<sup>\*</sup> do not seem measure the development when applied to the Italian situation. The main reason could be that they barely see the development in areas with high electrification rate<sup>6</sup>.

Figure 3 – Italian regions by quartile classes of NLDI\*, HDI and QUARS – environment. Italy, year 2011.



Higher values of NLDI\* (and of course of NLDI) are reached by very hilly regions such as Bolzano, Trento, Aosta Valley, Piedmont, or at least by regions with large uninhabited areas, for instance Sardinia. On the contrary, lower NLDI\* are registered for regions easy to be occupied, that is flat, bordering on the sea and ultimately without environmental barriers. Among these regions there are Campania, Lazio, Emilia-Romagna. In conclusion, NLDI\* is lower for regions in which population is scattered on all the territory, while is higher for regions in which population is gathered around cities that, however, are medium or small in size. In this regard, the NLDI\* seems to be more related to ecological footprint, land consumption or urban sprawl. In fact, the correlation among NLDI\* and the synthetic index on environment of QUARS7 computed a regional level, is quite strong, around 0.7. Furthermore, the two indexes produce very similar rankings. Therefore, it can be assumed that for Italian case study, NLDI and NLDI\* measure the environmental impact of the human activity and more specifically the urban saturation of the territory. Focusing on NLDI\*, it possible to state that lower NLDI\* is, higher the environmental impact of the human activity and the urban saturation

<sup>&</sup>lt;sup>6</sup> The percentage of population with access to electricity. In Italy the electrification rate is equal to 100% (source: World Bank).

<sup>&</sup>lt;sup>7</sup> The synthetic index on environment of QUARS considers: population density, CO<sub>2</sub> Emissions, Use of fertilizer, protected areas (source: Istat), percentage of recycling (source: Ispra), renewable energy production (Terna Spa), Eco-management (source: Legambiente), organic farming, sustainable mobility (source: Sbilanciamoci!).

is. The differences among the indexes ( $NLDI^*$ , HDI and QUARS – environment) can be better appreciate looking in Figure 3 the representation of regions in quartile classes.

#### 4.1. Marche region

Marche region is in a high position in the ranking of *NLDI*<sup>\*</sup>. It means that the environmental impact of the human activity and the urban saturation is high.

 Table 3 – Comparison of Marche region provinces with other Italian provinces. Italy 2011

NLDI	NLDI*
0.083	0.241
0.077	0.218
0.073	0.228
0.131	0.349
0.152	0.380
0.111	0.296
0.245	0.600
0.01	0.096
0.019	0.105
0.037	0.141
0.277	0.625
0.193	0.423
	NLDI           0.083           0.077           0.073           0.131           0.152           0.111           0.245           0.01           0.019           0.037           0.277           0.193

This could be due to the features of its territory: hills settled a few kilometer from the sea. Therefore, population and cities grew up and scattered along the coast. The  $NLDI^*$  for Marche provinces is much closer to the values of big provinces like Rome, Milan and Naples than to the provinces with themselves population sizes.

### 5. Conclusions and remarks

Elvidge *et al.* (2012) proposed the Night Light Development Index (*NLDI*). This index is computed using nighttime satellite imagery and density population data. They demonstrated that their index, at country level, has a strong negative correlation with the Human Development Index (*HDI*) and, can help to solve the data collection problems that, for some countries, can affect the *HDI*.

In this paper is shown that *NLD1* gives the same value for several scenarios that reasonably refer to different development levels. In sake of simplicity, a linear transformation of *NLD1* that takes into account the brightness light level has been introduced. The new index, *NLD1*<sup>\*</sup>, is still defined between 0 and 1.

152

The two indexes are applied to the Italian case. The values of NLDI and  $NLDI^*$  for Italy, geographical areas (NUTS-1), regions (NUTS-2) and some provinces (NUTS-3) have been derived. However, the correlation between NLDI or  $NLDI^*$  and HDI at regional level is weak and, moreover, is positive. The two indexes. probably because they have been applied to an area with high electrification rate like Italy, do not seem to measure development. They are more correlated with the synthetic index on environment of QUARS therefore, it can be assumed that, for Italian regions, they measure the environmental impact of the human activity and more specifically the urban saturation of the territory. Focusing on  $NLDI^*$ , it possible to state that lower  $NLDI^*$  is, higher the environmental impact of the human activity and the urban saturation is. However, further studies to better understand the real meaning of the index are needed.

# References

- BENEDETTI R., CIAVATELLA M. 2006. L'uso dei dati telerilevati nelle stime di superficie e produzione. *Agriregionieuropa*, Anno 2, No. 7.
- DOLL C.H.N., MULLER J.-P., ELVIDGE C.D. 2000. Night-Time imagery as a tool for global mapping of socioeconomic parameters and greenhouse. *Ambio*, Vol. 29, No. 3, pp. 157-162.
- DOLL C.H.N., MULLER J.-P., MORLEY J.G. 2006. Night Mapping regional economic activity from night-time light satellite imagery. Ecological Economics, Vol. 57, pp. 72-92.
- 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, Vol. 7, pp. 23-35.
- GINI C. 1914. Sulla misura di concentrazione e della variabilità dei caratteri. Atti del R. Istituto Veneto di SS.LL.AA., Vol. 73, parte II, pp. 1203-1248 (English traslation in *Metron*, 2005, Vol. 63, No. 1, pp. 3-38).
- MARLEY J., ELAZAR D., TRAEGER K. 2014. Methodological approaches for utilising satellite imagery to estimate official crop statistics. *Research Paper*, Australian Bureau of Statistics. Camberra: June 2014.
- SUDHIRA H.S., RAMCHANDRA T.V., JAGADISH K.S. 2004. Urban sprawl: metrics, dynamics and modelling using GIS. *International Journal of Applied Earth Observation and Geoinformation*, Vol. 5, pp. 29-39.
- SBILANCIAMOCI! 2012. Come si vive in Italia? La qualità sociale e ambientale regione per regione. Rapporto QUARS 2011 Indice di qualità Regionale dello Sviluppo. Roma: Grafica Giorgetti (available on-line at http://www.sbilanciamoci .org/wp-content/uploads /2012/03/quars-2011.pdf).

- UNDP 2011. Sustainability and Equity: A Better Future for All. Human Development Report 2011. New York: Oxford University Press (available on-line at http://www.er.undp.org/content/dam/eritrea/docs/HDR/HDR\_2011\_EN\_Complete .pdf)
- WOLFF H., CHONG H., AUFFHAMMER M. 2011. Classification, detection and consequences of data error: evidence from the Human Development Index. Economic Journal, Vol. 121, pp. 843–870.

#### SUMMARY

# Satellite imagery for studying development? The Italian case study

The Human Development Index (*HDI*) is a well-known measure of development published since the 90s by the United Nation. Among the criticisms of *HDI*, there are undeniable problems in data collection that can frustrate comparisons between countries. In fact, for some countries old data only are available and few others have not even that. Elvidge *et al.* (2012) proposed the Night Light Development Index (*NLDI*), that can be computed solely from nighttime satellite imagery and population density, therefore without monetary data and with ease in data collection. The *NLDI*, that is a inequality measure of light distribution among inhabitants, has a strong correlation with the *HDI* at country level.

In this paper we show that NLDI can produce the same values for very different development levels. Therefore, a simple correction  $(NLDI^*)$  for overcoming this drawback is introduced. The original NLDI and our correction have been computed for the Italian case study, that is, the whole territory, the geographical areas (NUTS-1), regions (NUTS-2) and few provinces (NUTS-3) have been derived. The values obtained have been compared with those of others indexes to better understand the meaning of  $NLDI^*$  in a particular context like the Italian one.

Francesco Maria CHELLI, Università Politecnica delle Marche, f.chelli@univpm.it Alessio GUANDALINI, Italian National Institute of Statistics – ISTAT, alessio.guandalini@istat.it