OPEN ACCESS

1st International Electronic Conference on Remote Sensing
22 June - 5 July 2015

Conference Proceedings Paper – Remote Sensing

Top-down identification of mixed vs. residential use in urban areas: Evaluation of remotely sensed nighttime lights for a case study in Cuenca City, Ecuador

Christoph Aubrecht [1,] *, José Antonio León Torres [2]

[1] The World Bank - Social, Urban, Rural & Resilience (GSURR) - Washington DC, USA; E-Mail: caubrecht@worldbank.org

[2] National Autonomous University of Mexico (UNAM) - Institute of Engineering - Mexico City, Mexico; E-Mail: jleont@iingen.unam.mx

* Author to whom correspondence should be addressed; E-Mail: caubrecht@worldbank.org

Received: / Published:

Abstract: This paper introduces a novel top-down approach to geospatially identify and distinguish areas of mixed use from predominantly residential areas within urban agglomerations. Within the framework of the World Bank's Central American Country Disaster Risk Profiles (CDRP) initiative a disaggregated property stock exposure model has been developed as one of the key elements for subsequent disaster risk and loss estimation. Global spatial datasets are thereby used consistently to ensure wide-scale applicability and transferability. Residential and mixed use areas need to be identified in order to spatially link accordingly compiled property stock information. Impervious Surface Area (ISA) data based on remotely sensed nighttime lights are used as proxy to identify areas of peak human activity. Intense lighting in that context is associated with a high likelihood of commercial and/or industrial presence. Areas of low light intensity, in turn, can be considered more likely residential. Several ISA intensity thresholds are tested for Cuenca City, Ecuador, in order to best match a given reference situation on the ground based on cadastral land use data. Results are considered highly relevant not only for the CDRP initiative but more general underline the relevance of remote sensing data for top-down modeling approaches at wide spatial scale.

Keywords: top-down modeling; urban areas; nighttime lights; human activity; residential use; mixed use; global spatial data; CDRP

---

1. Introduction

Issues of urban development are increasingly being addressed at the global scale, with international NGOs and development institutions often setting the path and moving the public agenda forward. Regularly published reports such as the United Nations' World Urbanization Prospects (UN 2014) or the World Bank's World Development Reports (World Bank 2014) address fundamental issues and define key research questions to be tackled by the scientific and international development community. In that context it has become more and more evident that spatial data is playing a crucial role for consistent cross-regional analyses and unbiased evaluation of locally implemented actions. Remote sensing data in particular provide a rich and globally consistent source for analyses at multiple levels. At global scale different aspects have to be considered than for local-level spatial analyses, including consistency, scalability, retraceability etc. Several global project initiatives address these issues in various thematic domains. The World Bank's Global Urban Spatial Data Initiative, for example, has the objective of addressing pending issues of regional definition and data incompatibilities and supports the international collaborative setup and development of a consistent data set of global urban extents and associated population distribution patterns. In the same context, the Global Human Settlement Working Group, established under the umbrella of the Group on Earth Observations (GEO), aims at establishing a new generation of global settlement measurements and products based on consistent high-resolution satellite imagery analysis.

The presented study has been carried out within the framework of the World Bank's Country Disaster Risk Profiles (CDRP) project initiative which is currently being implemented at the continental scale for Central America (Gunasekera et al. 2014). With the clear aim at extending to other regions, global applicability and easy transferability are considered crucial for the model setup. Global spatial datasets are therefore used throughout the CDRP project, with the presented approach specifically developed for implementation in the setup of a disaggregated property stock exposure model, one of the key elements for subsequent disaster risk and loss estimation. While focusing primarily on natural hazards and risks, urban-rural identification and intra-urban classification aspects are highly relevant for setting the basic spatial framework for analysis (Aubrecht et al. 2014).

The paper introduces a novel approach to geospatially identify and distinguish areas of mixed use from predominantly residential areas within urban agglomerations. After initial urban-rural classification at a 1km grid level, that urban mask needs to be classified in residential and mixed use areas in order to spatially link accordingly compiled property stock information (e.g. from global tabular databases such as PAGER-STR). Impervious Surface Area (ISA) data based on remotely sensed nighttime lights from the DMSP-OLS sensor are used as proxy to identify areas of peak human activity, often associated with a high likelihood of commercial and/or industrial presence. Several light intensity thresholds are tested for a case study in Cuenca City, Ecuador, in order to best match a given reference situation on the ground, where local-level cadastral land use data is used to identify the actual distribution ratio of residential vs. mixed use areas.

Results are presented first-hand in this paper and future work is addressed highlighting the relevance of global remote sensing data for top-down modeling approaches at wide spatial scale. Outcome is considered relevant for global urban spatial modeling in a variety of topical domains.

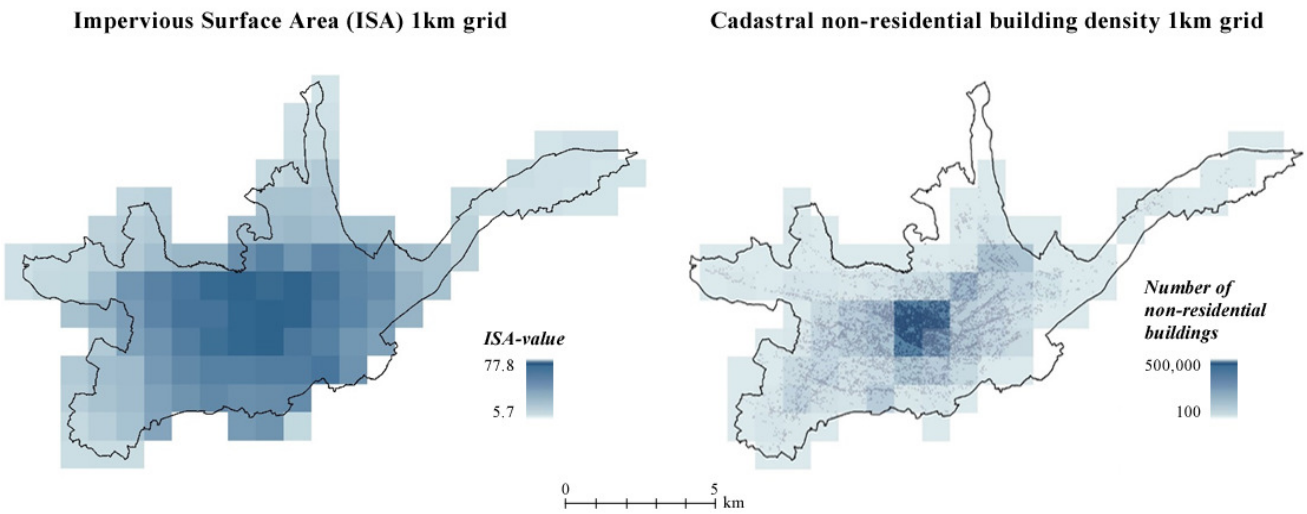
2. Methods

As briefly outlined above, the presented study has been carried out under the framework of the World Bank's Central American Country Disaster Risk Profiles (CDRP) Initiative. With that kind of continental and global models, the implemented scale level plays an important role in defining the basic spatial units of analysis. Working on a 1 km grid level - frequently used for global models - the spatial identification and distinction of unique inventory regions is often not unambiguously possible at the grid cell level due to the well-studied mixed pixel issue (e.g. Foody & Cox 1993, Fisher 1997). While large urban residential areas as well as certain dedicated industrial zones are still often built in rather compact manner and can thus indeed cover entire grid cells, particularly commercial areas are commonly intertwined with residences forming wider areas of mixed use. In order to appropriately identify urban non-residential areas in a spatial top-down model it is therefore considered reasonable to assume a certain share of residential occupancy throughout and consider grid cells that also include a non-residential share as areas of mixed use.

For identification of those built-up urban areas that also feature a share of non-residential use, we refer to satellite-derived nighttime lights data as proxy measure. Specifically, we use the global Impervious Surface Area (ISA) product, based on data from the DMSP-OLS (Defense Meteorological Satellite Program - Operational Linescan System) sensor, due to its strong inherent spatial correlation with built-up area (Elvidge et al. 2007). The assumption hereby is that intense lighting in that context is associated with a high likelihood of commercial and/or industrial presence, commonly clustered in certain parts of a city (such as central business districts and or peripheral commercial zones). Areas of low light intensity, in turn, can be considered more likely residential.

The main objective of this study is thus to identify the light intensity threshold that matches best the separated distribution of residential vs. mixed use areas on the ground. It should be noted that the presented approach is proposed only for pre-identified urban areas (Aubrecht et al. 2015), as for rural regions coarse-scale lighting intensity has reduced spatial correlation with built-up and other additional aspects come into play.

Figure 1. ISA (left) and cadastral non-residential building density (right) for Cuenca City.



In order to identify an exemplary ground reference status we analyze cadastral data for a test case area featuring Cuenca City, Ecuador (Municipalidad de Cuenca 2015). Building footprint data including information on occupancy and building use are aggregated to a 1km grid for comparative purposes. Figure 1 illustrates the non-residential building density at 1km grid level in comparison to the Impervious Surface Area (ISA) grid. Following the same approach as outlined above, i.e. distinguishing purely residential areas from areas of mixed use, the aggregated cadastral data show a 75% ratio of residential areas, complemented by 25% of mixed use areas. We can consequently use this bottom-up-determined distribution ratio to identify the appropriate lighting intensity threshold in the top-down model.

In order to define the relevant ISA data value histogram for the threshold identification, we select all the ISA cells within the pre-defined urban test case area of Cuenca City. The min-max ISA value range is thereby identified as 5.7-77.8. We then iteratively apply several threshold points in that value range and compare the resulting areas of relatively low and relatively high ISA values to the aggregated cadastral grid. The eventually selected final cut-off point is that threshold value that produces the best-matching output with regard to the 75-25 cadaster-based residential vs. mixed use area distribution ratio.

3. Results and Discussion

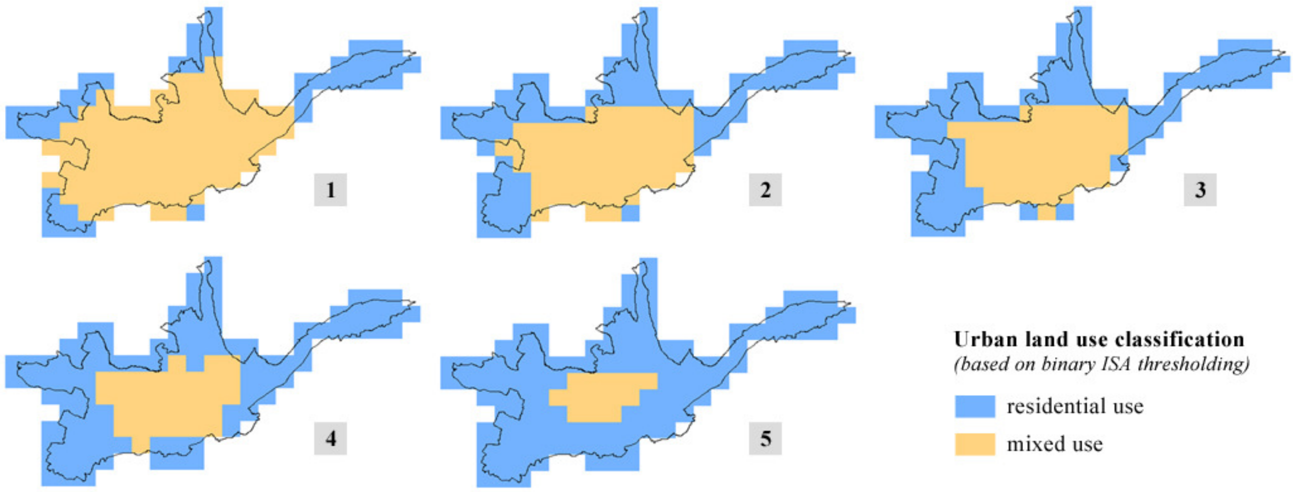
Table 1 illustrates the various tested ISA threshold values and the corresponding building use distribution ratios as derived from comparative spatial overlay with the aggregated cadastral data. ISA min-max range and respective threshold values are shown in the left part of the table, with the percentile column indicating the relative value distribution. Specifically, that means that for the highlighted ID 4 the ISA value of 42 indicates the median value (50^{th} percentile) in the distribution histogram. Half of the values in the study area under consideration thus feature an ISA value lower than 42 and the other half feature a higher value.

Spatially overlaid on the aggregated cadastral building use density grid (at the 1km level directly comparable to the ISA grid), that results in a 74% residential ratio and a 26% mixed use share, thus best-matching the bottom-up-derived 75% threshold value. Figure 2 maps the binary land use classification (residential use vs. mixed use) for the 5 tested ISA thresholds respectively. The share of mixed use area decreases thereby corresponding to the higher ISA cut-off points.

Table 1. ISA distribution thresholds and corresponding building use distribution ratios.

ID	ISA Distribution				Building Distribution	
	Min	*Threshold*	*Max*	*Percentile*	*Residential Use*	*Mixed Use*
1	5.7	15	77.8	13%	32.00%	68.00%
2	5.7	25	77.8	27%	52.00%	48.00%
3	5.7	35	77.8	41%	70.00%	30.00%
4	**5.7**	**42**	**77.8**	**50%**	**74.00%**	**26.00%**
5	5.7	51	77.8	63%	96.00%	4.00%

Figure 2. Binary land use classification of Cuenca City based on ISA thresholds from Table 1. Table record IDs are indicated in the figure as 1-5.



This result is very interesting as it in fact backs up the prior non-evaluated assumption implemented in the Central American CDRP model to use the ISA median value as threshold. Without ground reference data as available for this presented test case study, the use of the median value seemed most appropriate as it introduces the least possible subjectivity and merely separates a certain data set in high and low according to its histogram without additionally induced statistical skew.

Having the building-level cadastral data at hand for this test case study enables us not only to determine the real binary land use distribution ratio, but furthermore allows us to consequently evaluate the degree of spatial overlap as a measure of model output accuracy. In the presented case 82.8% of the total non-residential building stock of Cuenca City (3.6 of 4.3 million km^2) is indeed captured within the selected top-down-derived binary mixed use mask.

To further evaluate the model output the implemented approach was re-run with a shrunk urban mask. While in the above-outlined approach all the ISA grid cells were considered that include a certain share of the Cuenca City administrative area, for this re-run only a more central part of the urban agglomeration was selected. The resulting best-matching ISA threshold is now slightly higher (around the 55th percentile). This was expected as predominantly residential areas in the periphery of the city are now not included in the mask and those lower ISA values are thus missing in the histogram, thus pushing up the median value.

This exercise basically highlights the importance of correct spatial pre-identification of the urban area for subsequent inner-urban analysis. If the urban mask is spatially over- or under-defining, the appropriate ISA threshold value would de- or increase respectively.

4. Conclusions and outlook

With the initially assumed median value threshold for the binary ISA classification confirmed through comparative in situ data analysis for an accurately defined urban agglomeration the presented case study is considered very beneficial for the overall implementation process of the CDRP initiative. Also the second re-run of the model with a shrunk more central urban mask that showcased the correspondingly expected ISA threshold upward shifts provides another back-up for the model validity as well as underlines the importance of accurate urban delineation in the first place.

With the CDRP exposure and subsequent risk and loss models already implemented for all of Central America further test studies can be carried out to increase the sample size of the model evaluation and also test the approach in different regional settings. While in Central America basically no big deviations are expected with regard to the model applicability, it will be interesting to see testing results when extending to the Caribbean and across.

With the DMSP-OLS-based ISA data set seemingly working well as input data source for the residential-mixed identification model, still significant further accuracy improvements are expected when referring to more recent data of the newer VIIRS (Visible Infrared Imaging Radiometer Suite) sensor (Elvidge et al. 2013). Due to its higher spatial resolution as compared to OLS, VIIRS is superior in terms of identifying certain rough inner-urban spatial structures which could in return improve the identification of commercial areas. Furthermore, also featuring a better radiometric resolution, saturation issues in urban core areas that arise in OLS data are eliminated which, again, should add significant improvement to the presented classification approach. VIIRS data will therefore be considered as comparative input data in the next stage of the study.

Acknowledgments

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of its Executive Directors or the governments they represent.

Author Contributions

Christoph Aubrecht developed the conceptual framework for the presented top-down method and tested geospatial implementation in various Central American countries in the context of the CDRP project initiative. José Antonio León Torres then implemented the developed approach for the Cuenca City (Ecuador) case study and performed the comparative analysis with the bottom-up cadastral data. This paper was initiated by Christoph Aubrecht, with contributions from both authors.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Aubrecht, C.; Gunasekera, R.; Ishizawa, O. Urban characterization aspects and associated spatial scale considerations in the context of global-level exposure modeling: Comparative data evaluation. In *Proceedings of the Second European Conference on Earthquake Engineering and Seismology (2ECEES)*. Istanbul, Turkey, **2014**.
2. Aubrecht, C.; Gunasekera, R.; Ishizawa, O.; Pita, G. The flipside of 'urban' - A novel model for built-up-adjusted rural-urban pattern identification and population reallocation. *Internal working*

paper **2015**, The World Bank - Social, Urban, Rural & Resilience (GSURR), Disaster Risk Management (DRM): Washington DC, USA.

3. Elvidge, C.D.; Tuttle, B.T.; Sutton, P.C.; Baugh, K.E.; Howard, A.T.; Milesi, C.; Bhaduri, B.; Nemani, R. Global Distribution and Density of Constructed Impervious Surfaces. *Sensors* **2007**, *7(9)*, 1962-1979.
4. Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.-C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proceedings of the Asia-Pacific Advanced Network* **2013**, *35*, 62-69.
5. Fisher, P. The pixel: A snare and a delusion. *International Journal of Remote Sensing* **1997**, *18(3)*, 679-685.
6. Foody, G.M.; Cox, D.P. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing* **1994**, *15(3)*, 619-631.
7. Gunasekera, R.; Ishizawa, O.A.; Aubrecht, C.; Pita, P.; Pomonis, A.; Fane, K.; Murray, S.; Blankespoor, B. Developing an Adaptive Exposure Model to Support the Generation of Country Disaster Risk Profiles. In *Geophysical Research Abstracts, Vol. 16*, Proceedings of the EGU (European Geosciences Union) General Assembly, Vienna, Austria, **2014**, EGU2014-16168.
8. Municipalidad de Cuenca (GAD). *Catastro urbana de la ciudad de Cuenca*. Departamento de Avalúos y Catastros, **2015**. Available online: http://www.cuenca.gov.ec (accessed on 05 May 2015).
9. UN. *World Urbanization Prospects, the 2014 Revision*. United Nations, Population Division of the Department of Economic and Social Affairs: New York, NY, USA, **2014**.
10. World Bank. *World Development Report 2014: Risk and Opportunity - Managing Risk for Development*. World Bank, Development Economics Research Group: Washington DC, USA, **2014**.