

Goldblatt, Ran; Jones, Madeline; Bottoms, Brad

Article

Geospatial data for research on economic development

Development Engineering

Provided in Cooperation with:

Elsevier

Suggested Citation: Goldblatt, Ran; Jones, Madeline; Bottoms, Brad (2019) : Geospatial data for research on economic development, Development Engineering, ISSN 2352-7285, Elsevier, Amsterdam, Vol. 4, pp. 1-3,
<https://doi.org/10.1016/j.deveng.2019.100041>

This Version is available at:

<https://hdl.handle.net/10419/242299>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Geospatial data for research on economic development



It is estimated that in 2016 nearly 38 per cent of workers in the world's least developed countries were living with their families below the poverty line (approximately \$1.90 per person per day) (UN, 2017), often without access to schools, healthcare, electricity, safe water and other critical services. Yet, accurate data on the number of poor people and their living conditions are scarcely existent, especially in developing countries. Census counts comprise the primary source of this type of data, but are published infrequently, varying in resolution and precision, due in part to limited available resources for data collection.

The gap in the availability of data between the developing and the developed world is, at least partially, the result of technological advances and innovative solutions for data collection. Novel sources of data facilitate the development of new approaches for demographic modeling and measurement in wealthy nations, while in developing countries, fewer sources of big data exist (Blumenstock et al., 2015). For example, while nearly half of today's world population has internet access, more than 4 billion people - mostly in developing countries - still don't have access to the internet.¹ The result is an increasing gap in the availability of data for developing and developed countries. Consider, for example, the utilization of data gathered from online social media for demographic research. Social media platforms such as Facebook, Twitter and other collaborative virtual spaces are increasingly harnessed to track the *digital traces* (Bruno, 2007) of society, whether for monitoring stocks of migrants (Zagheni Emilio et al., 2017) or for generating taxonomies of happiness levels in US states (Mitchell et al., 2013). This type of data only provides insights about those who participate in such virtual spaces and does not always fully capture the social complexity of the ever-changing developing world.

The lack of accurate and timely data on the distribution of poverty also often forces policy-makers to make decisions based on limited (and sometimes inaccurate) data (Blumenstock, 2016). For example, development programs and policies are typically designed to change outcomes (e.g. to raise incomes, to improve learning, or to reduce illness). An impact evaluation is intended to evaluate the outcome of a program by answering one specific question: What is the impact of a specific program on an outcome of interest? (Gertler et al., 2016). Answering this question requires to track and monitor the impacts of the program - and this can only be done by a continuous and frequent collection of data from multiple sources. Such data provides the basis for any evidence-based policy making.

But poverty and other societal characteristics also hold a spatial dimension. Poverty can often be identified visually and reflected in the density of residential structures, patterns of roads, the material of roofs or the distribution and extent of green spaces. Many of these features are observable from space. From the perspective of social science, capturing this data may provide meaningful information about the context that shapes social phenomena (Council and Stern, 1988). There are more than 1400 satellites continuously collecting an unprecedented amount of data from Earth at ever-improving spatial, spectral and temporal resolutions, over increasingly short time intervals. Data collected by these satellites and other sensors are progressively shifting the way we measure and understand our world. These data allow researchers to study and map the physical environment and infer about social processes in near real time, in almost every corner of the world.

The heightened availability of satellite data has revolutionized our ability to map human settlements and understand the distribution and characteristics of the population on Earth. New methods for machine learning, predictive analytics and remote sensing are reforming the way we measure and address societal challenges, detect poverty and steer efforts to reduce it. Researchers are continually discovering new ways to convert remotely sensed data into meaningful information about the nature and pace of change of human and physical landscapes.

Over the last few decades, researchers have been working to detect and map poverty and wealth from space, primarily with nighttime light data. Based on the assumption that poorly lit areas with high population counts signify higher percentages of poor people and lower wealth, researchers have been able to estimate the number of individuals living in poverty, globally (Elvidge et al., 2009). Nighttime lights are also being extensively used to map urbanization. According to this approach, a pixel's nighttime light value exceeding a specified threshold (which may vary across regions or countries) signifies urban development. Moreover, because the intensity of light emitted at night is closely related to economic development, this proxy is also leveraged to predict national economic growth (Henderson et al., 2012) as well as regional economic development (Doll et al., 2006). However, inference of human activity with Nighttime lights is often inaccurate, particularly in low-density urban areas (Zhang and Seto, 2013). DMSP-OLS sensors, which until recently (with the availability of data collected by the Visible Infrared Imaging Radiometer Suite (VIIRS)) have been the primary source for nighttime light data can exaggerate the extent of urban areas (Henderson et al., 2003) while overlooking small or developing

¹ <https://www.weforum.org/agenda/2016/05/4-billion-people-still-don-t-have-internet-access-here-s-how-to-connect-them/>.

settlements. In addition, the extent and intensity of lit areas cannot directly delimit urban regions due to the “blooming” effect (the identification of lit areas as consistently larger than the settlements with which they are associated) (Imhoff et al., 1997) and “over saturation” of pixels (i.e., pixels in bright areas, such as in city centers, reach the highest possible intensity and no further details can be recognized) (Hsu et al., 2015). As a result, there has been a significant move towards mapping poverty with daytime satellite data at higher spatial resolutions. For example, Engstrom et al. (2017) showed that object and texture features extracted from high-resolution satellite imagery over Sri Lanka (e.g. the number and density of buildings, number of cars, density and length of roads, roof material, etc.) could explain nearly 60 percent of poverty headcount rates and average log consumption, compared to models built using nighttime lights that could explain only 15 percent of the variation in poverty or income. Jean et al. (2016) combined nighttime maps with high-resolution daytime satellite images to estimate consumption expenditure and asset wealth. The authors showed, that a convolutional neural network can be trained to identify image features that could explain up to 75% of the variation in local-level economic outcomes in Africa. Other studies develop novel methods designed to map slums and informal settlements in developing countries by capturing some of their unique spatial features (e.g. building density, building heights, and sizes), differentiating between informal from formal settlements (Taubenböck and Kraff, 2014). With around 1 billion people living today in slums (UN-Habitat, 2015) lacking access to drinking water, sanitation facilities and durable housing conditions, accurate and up-to-date information on the extent of slums and the number of slums dwellers is fundamental for the preservation of a sustainable human society.

These new sources of geospatial data are transforming our understanding of the modern world. Novel machine-learning algorithms are combined with satellite imagery from multiple sources to estimate road quality in developing countries, predict consumption expenditure, measure wealth and inequality, monitor small-holder farmers' yields in regions affected by drought, and monitor electricity provision in rural areas. Parallel computing platforms with large storage capacities are becoming increasingly available, allowing for the extraction of meaningful information about the developing world and encouraging new design of environmentally responsible approaches to measuring and monitoring economies.

This special issue of *Development Engineering* highlights 5 case studies that illustrate novel use of technology and new sources of big geospatial data to derive insights pertaining to societal processes, primarily in developing countries. Augusto Zagatti et al. utilize nationwide de-identified Call Detail Records (CDR) to investigate nighttime and daytime population densities and commuting patterns in Haiti. The authors construct origin-destination matrices of weighted connections between home and work locations to identify fragmented labor markets in the country. Eitzel et al. evaluate the potential of ‘medium-tech’ digital mapping technologies as solutions for empowering disenfranchised communities in rural Zimbabwe. The authors demonstrate that digital collaborative mapping with relatively simple technology allows spreading information among rural farmers in Zimbabwe. This case study demonstrates the potential of geospatial data as a tool to address poverty and sustainability and improve the well-being of the most vulnerable populations on our planet. Marty et al. apply a Geographic Simulation and Extrapolation (GeoSIMEX) approach to mitigate the spatial imprecision inherent in

the geoparsed data when evaluating the impact and success of international aid and interventions projects. The authors present analysis of the impact of Chinese aid on vegetative land cover in Rwanda and Burundi and find that Chinese investments have slowed the loss of vegetation in Rwanda. Rivera Ballesteros et al. utilize satellite imagery collected by Sentinel-2 satellite for supervised image classification to estimate the extent of tree cover in twelve gazetted forests in semi-arid Burkina Faso. The authors show this source of imagery allows to detect patches of forest cover in arid and semi-arid regions that could not be identified with lower resolution imagery and evaluate the extent and characteristics of deforestation in these regions. Finally, Goldblatt et al. discuss the challenges of measuring urbanization in developing countries that lack the resources and infrastructure needed to produce reliable data. The authors highlight the potential use of publicly available remotely sensed data for mapping changes in the built-up LC/LU in Ho Chi Minh City, Vietnam, in the period between 2000 and 2015.

The invention of computerized Geographic Information Systems (GIS) in the 1960s revolutionized how we collect, map and analyze spatial data and give meaning to the world around us. With advances in computer processing power and data storage, increased availability of personal computers, the invention of the internet (the World Wide Web) and the growing popularity of mobile and “smart” phones, the volume of data we generate and consume is growing exponentially. Devices on satellites, airplanes and UAVs collect massive amounts of imagery that capture the characteristics of Earth. All these studies featured in this special section illustrate novel research methods to analyze these unique sources of data to achieve a better understanding of the nature of developing countries.

References

- Blumenstock, J.E., 2016. Fighting poverty with data. *Science* 353, 753–754. <https://doi.org/10.1126/science.aah5217>.
- Blumenstock, J., Cadamuro, G., On, R., 2015. Predicting poverty and wealth from mobile phone metadata. *Science* 350, 1073–1076. <https://doi.org/10.1126/science.aac4420>.
- Bruno, L., 2007. Beware, your imagination leaves digital traces. *Times High. Lit.* (6).
- Council, R.R., Stern, P.C., 1988. Linking remote sensing and social science: the need and the challenges. In: *People and Pixels: Linking Remote Sensing and Social Science*. National Academies Press, pp. 1–27.
- Doll, C.N.H., Muller, J.-P., Morley, J.G., 2006. Mapping regional economic activity from night-time light satellite imagery. *Ecol. Econ.* 57, 75–92. <https://doi.org/10.1016/j.ecolecon.2005.03.007>.
- 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. *Comput. Geosci.* 35, 1652–1660. <https://doi.org/10.1016/j.cageo.2009.01.009>.
- Emilio, Zagheni, Weber, Ingmar, Krishna, Gummadi, 2017. Leveraging Facebook's advertising platform to monitor stocks of migrants. *Popul. Dev. Rev.* 43, 721–734. <https://doi.org/10.1111/padr.12102>.
- Engstrom, R., Hersh, J., Newhouse, D., 2017. Poverty from Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being (Policy Research Working Paper No. 8284). Poverty and Equity Global Practice Group, World Bank, Washington DC.
- Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., Vermeersch, C.M., 2016. *Impact Evaluation in Practice*. The World Bank.
- Henderson, M., Yeh, E.T., Gong, P., Elvidge, C., Baugh, K., 2003. Validation of urban boundaries derived from global night-time satellite imagery. *Int. J. Remote Sens.* 24, 595–609. <https://doi.org/10.1080/01431160304982>.
- Henderson, J.V., Storeygard, A., Weil, D.N., 2012. Measuring economic growth from outer space. *Am. Econ. Rev.* 102, 994–1028. <https://doi.org/10.1257/aer.102.2.994>.
- Hsu, F.-C., Baugh, K.E., Ghosh, T., Zhizhin, M., Elvidge, C.D., 2015. DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. *Rem. Sens.* 7, 1855–1876. <https://doi.org/10.3390/rs70201855>.
- Imhoff, M., Lawrence, W.T., Stutzer, D.C., Elvidge, C.D., 1997. A technique for using

- composite DMSP/OLS “City Lights” satellite data to map urban area. *Remote Sens. Environ.* 61, 361–370. [https://doi.org/10.1016/S0034-4257\(97\)00046-1](https://doi.org/10.1016/S0034-4257(97)00046-1).
- 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, 790–794. <https://doi.org/10.1126/science.aaf7894>.
- Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S., Danforth, C.M., 2013. The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS One* 8, e64417. <https://doi.org/10.1371/journal.pone.0064417>.
- Taubenböck, H., Kraff, N.J., 2014. The physical face of slums: a structural comparison of slums in Mumbai, India, based on remotely sensed data. *J. Hous. Built Environ.* 29, 15–38. <https://doi.org/10.1007/s10901-013-9333-x>.
- UN, 2017. Progress towards the Sustainable Development Goals: Report of the Secretary-General (Economic and Environmental Questions: Sustainable Development). (United Nations).
- UN-Habitat, 2015. *Slum Almanac 2015/2016: Tracking Improvement in the Lives of Slum Dwellers. Participatory Slum Upgrading Programme*, UN-Habitat.
- Zhang, Q., Seto, K.C., 2013. Can night-time light data identify typologies of urbanization? A global assessment of successes and failures. *Rem. Sens.* 5, 3476–3494. <https://doi.org/10.3390/rs5073476>.

Ran Goldblatt*, Madeline Jones, Brad Bottoms
New Light Technologies Inc., 1440 G Street Northwest, Washington, DC
20005, USA
E-mail address: ran.goldblatt@nltgis.com (R. Goldblatt).

* Corresponding author.