Automatic Generation of Census Enumeration Area and National Sampling Frame to Achieve Sustainable Development Goals

Authors : Sarchil H. Qader, Andrew Harfoot, Mathias Kuepie, Sabrina Juran, Attila Lazar, Andrew J. Tatem

Abstract : The need for high-quality, reliable, and timely population data, including demographic information, to support the achievement of the sustainable development goals (SDGs) in all countries was recognized by the United Nations' 2030 Agenda for sustainable development. However, many low and middle-income countries lack reliable and recent census data. To achieve reliable and accurate census and survey outputs, up-to-date census enumeration areas and digital national sampling frames are critical. Census enumeration areas (EAs) are the smallest geographic units for collection, disseminating, and analyzing census data and are often used as a national sampling frame to serve various socio-economic surveys. Even for countries that are wealthy and stable, creating and updating EAs is a difficult yet crucial step in preparing for a national census. Such a process is commonly done manually, either by digitizing small geographic units on high-resolution satellite imagery or walking the boundaries of units, both of which are extremely expensive. We have developed a user-friendly tool that could be employed to generate draft EA boundaries automatically. The tool is based on high-resolution gridded population and settlement datasets, GPS household locations, building footprints and uses publicly available natural, man-made and administrative boundaries. Initial outputs were produced in Burkina Faso, Paraguay, Somalia, Togo, Niger, Guinea, and Zimbabwe. The results indicate that the EAs are in line with international standards, including boundaries that are easily identifiable and follow ground features, have no overlaps, are compact and free of pockets and disjoints, and the boundaries are nested within administrative boundaries.