ORIGINAL PAPER

People and Pixels 20 years later: the current data landscape and research trends blending population and environmental data

Tracy A. Kugler[1] · Kathryn Grace[2] · David J. Wrathall[3] · Alex de Sherbinin[4] ·
David Van Riper[1] · Christoph Aubrecht[5] · Douglas Comer[6] · Susana B. Adamo[4] ·
Guido Cervone[7] · Ryan Engstrom[8] · Carolynne Hultquist[7] ·
Andrea E. Gaughan[9] · Catherine Linard[10] · Emilio Moran[11] · Forrest Stevens[9] ·
Andrew J. Tatem[12] · Beth Tellman[13] · Jamon Van Den Hoek[3]

Abstract

In 1998, the National Research Council published *People and Pixels: Linking Remote Sensing and Social Science*. The volume focused on emerging research linking changes in human populations and land use/land cover to shed light on issues of sustainability, human livelihoods, and conservation, and led to practical innovations in agricultural planning, hazard impact analysis, and drought monitoring. Since then, new research opportunities have emerged thanks to the growing variety of remotely sensed data sources, an increasing array of georeferenced social science data, including data from mobile devices, and access to powerful computation cyberinfrastructure. In this article, we outline the key extensions of the *People and Pixels* foundation since 1998 and highlight several breakthroughs in research on human–environment interactions. We also identify pressing research problems—disaster, famine, drought, war, poverty, climate change—and explore how interdisciplinary approaches integrating people and pixels are being used to address them.

Keywords Remote sensing · Population data · Human dimensions of global change · Data integration · Mobile device data

Introduction

The seeds of a data revolution in the study of interactions between human society and the environment were sown in the 1960s and 1970s as Earth observation (EO)

✉ Tracy A. Kugler
    takugler@umn.edu

Extended author information available on the last page of the article

satellites orbiting the planet began to amass volumes of global data on land cover. Researchers began to use these new data to answer new sets of research questions about human-induced changes in the landscape, such as urbanization, deforestation, and crop production. The data were also used to gain insight into the environmental drivers of problems for human societies, such as drought, famine, and disasters. In the early 1990s, as a research community began to coalesce around human–environment interactions, a Human Dimensions of Global Change (HDGC) Committee was established under the National Research Council. At the same time NASA was working to expand the community of Landsat users to demonstrate the utility of EO data. NASA sponsored a workshop through the HDGC Committee to explore the current state of the connection between remote sensing and the social sciences and illustrate applications for Landsat data to advance empirical and theoretical understanding in social science domains. The primary legacy of that workshop was the publication of the classic volume, *People and Pixels: Linking Remote Sensing and Social Science* (Liverman et al. 1998), which compiled, synthesized, and reflected on the early stages of this data revolution.

Twenty years later, we are several steps into a second data revolution marked by the wide availability of environmental and social data at finer spatial and temporal resolutions, analysis-ready data products, and new analytical and computational tools. Researchers from a wide range of disciplines are leveraging these technical advances to pursue a new generation of research on coupled human–environment interactions, linking people and pixels to address new matters of global concern. In the early stages of *People and Pixels'* second generation, we see value in revisiting and updating the original hypotheses, exploring continuing limitations, and considering the implications of such progress for the future. The purpose of this article is to identify and document advances in data, software, coding, and processing that continue to facilitate the connection between people and pixels and to highlight empirical work at the intersection of people and pixels that advances our understanding of human–environment interactions. This article is primarily targeted to demographers, sociologists, economists, and other social scientists who are interested in incorporating environmental data into their work to explore human–environment interactions.

We organize this article into three main sections. In the first section, we highlight changes in the data landscape from 1998 to the contemporary period. We discuss advances in remote sensing, the emergence of large-scale in situ data gathered by mobile devices, and the growing class of derived data products for a range of topics of interest to social scientists. In the second section, we discuss tools and software that are making this profusion of data more readily accessible and applicable. In the third section, we present examples and case studies of social science research areas demonstrating applications of remotely sensed data. Within these examples, we first explore how remotely sensed data are used to measure the context of social phenomena, and then we address how remotely sensed data are being used to measure social phenomena and their effects.

Changes in the data landscape

Advances in remote sensing

Over the course of the past 20 years, space-based remote sensing has evolved dramatically, and innovation continues to proceed ever-more rapidly (Mathieu and Aubrecht 2018). Systematically collected, geographically comprehensive, and rigorously calibrated EO data serve as a crucial backbone for analytics of population–environment interactions.

The Landsat imagery that prompted *People and Pixels* now represents the beginning of a long-term archive of moderate resolution imagery. Landsat data now encompass more than 40 years of relatively consistent data at 30 to 60 m resolution, enabling investigations of change stretching over multiple decades. Examples of studies taking advantage of this long-term data archive include annual tracking of forest cover (Hansen et al. 2013) and urban development (Li et al. 2015; Sexton et al. 2013), assessing land use impacts on water quality (Olmanson et al. 2008; Zheng et al. 2015), tracking the types of land cover that are converted to agricultural use (Gibbs et al. 2010), and characterizing fire events and post-fire succession (Röder et al. 2008). Landsat data are complemented by data from the Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS) programs, which also span multiple decades. While Landsat has revisit periods of 8 to 16 days, AVHRR and MODIS generate data with coarser spatial resolution (250 m to 1 km), but with revisit periods of a day or less. The higher frequency data more readily capture rapidly changing phenomena such as wildfire, floods, and other hazards.

As satellite and sensor technology has advanced, much higher spatial resolution imagery has become commonplace. Commercial satellites such as SPOT, QuickBird, IKONOS, and the WorldView sensors have been capturing imagery with 3-m down to 30-cm resolution and 1- to 3-day revisit periods since the turn of the millennium. In the public realm, the European Copernicus program's Sentinel satellite constellation, launched between 2014 and 2017, provides both radar and optical data at 10-m resolution with approximately 5-day revisit periods. All Copernicus data are free and open under the joint European Union/European Space Agency data policy. The combination of high spatial resolution and frequent revisit times, as well as availability of both passive and active sensor technology, enables monitoring of local- to global-scale trends and dynamics in land cover, supporting the identification of anthropogenic impact on the environment and contextual features that influence humans. For example, such imagery is being used to examine changes in the landscape in conflict zones to assess the direct and indirect consequences of civil wars (see "Measuring social phenomena and their effects" section).

Since the mid-2010s, microsatellite constellations have gained considerable commercial attention. By deploying swarms of small satellites on a regular basis, data providers (e.g., Planet Labs) maintain networks of many dozens of high-resolution (0.5–5 m) sensors that together can provide daily, global coverage. Because microsatellites are less expensive to produce and launch than traditional satellites, many new operators continue to come online, including some with hyperspectral (e.g., Satellogic) and radar (e.g., Iceye) capabilities. The competition may serve to reduce the cost barrier for research applications of imagery from these commercial

sources. Belward and Skøien (2015) provide a comprehensive listing of civilian EO satellite launches and operating systems, and Denis et al. (2017) describe new innovations and disruptive technology in EO systems and associated markets.

Innovations on the horizon include space-based video and high-altitude pseudo satellites (HAPS). Organizations such as Skybox Imaging, UrtheCast, and Earth-i have begun providing full motion HD video from space (Earth-i provides 2-min color video clips). HAPS are large drones able to stay in the stratosphere (~ 20 km altitude) over a fixed point on Earth for several months, providing platforms for telecommunication, navigation, and remote sensing. HAPS would therefore be able to provide very high-frequency observations for an area of interest. While the concept of HAPS has been around for decades, operational systems appear to have come within reach only recently (d'Oliveira et al. 2016). Both space-based video and HAPS are generating interest in commercial and operational domains, for applications such as real-time traffic and pedestrian monitoring (Esch et al. 2017). The potential for these technologies to contribute to human–environment interactions research, however, remains largely unexplored.

With these advances in remote sensing, the ultimate goal of a "digital Earth," real-time continuous monitoring at very high resolutions, seems almost within reach, and may prove to be beneficial in a range of research domains involving populations and their activities across various spatial and temporal scales (Aubrecht et al. 2015). The very high spatial and temporal resolution of these data, potentially combined with data from in situ sensor networks, may eventually "even [allow] for the monitoring of living species" (Association of European Space Research Establishments 2017, p. 9). Social scientists have not yet fully incorporated very high-resolution data into their work, in part due to remaining barriers in terms of commercial data costs, complex data processing, and extracting information of use to social scientists from such data. However, the increasing availability of free or low-cost high-resolution imagery, accessibility of distributed and high-performance computing, new analysis toolkits, and derived data products may well enable very high-resolution data to drive a second phase in the people and pixels revolution.

Derived/modeled data products

Derived and modeled data products address the challenge of extracting information of use to social scientists from raw EO data products. While remotely sensed data capture physical properties of the Earth, they do not directly capture phenomenon of interest to social scientists, such as land cover change, vegetation condition, or the distribution and characteristics of human populations. Derived and modeled products use remotely sensed data, typically in combination with other data sources such as agricultural or population censuses, as inputs into a model, classification process, or other workflow. The outputs are analysis-ready gridded representations of social or biophysical properties of interest. Below we describe key advances in the development of data products notable for their potential to advance knowledge about human–environment interactions. We also discuss the uncertainty inherent in modeled data, which researchers using these data products must keep in mind when drawing conclusions.

Modeled environmental products

Over the past 20 years, a variety of modeled environmental products have emerged. These products use remotely sensed data as a model input, potentially integrated with other data, to provide pixel-based measures of environmental features such as forest cover, surface water, and air quality that are important links in the human–environment interface. General land cover classification products such as Global Land Cover 2000 (Bartholome and Belward 2005) and GLOBCOVER (Arino et al. 2007) form the foundation for much of this work, and these products continue to be refined and produced at higher resolutions (Jun et al. 2014). A more recent generation of products focuses on specific types of land cover or land use, such as forest cover (Hansen et al. 2013; Pengra et al. 2015), agricultural land use (Ramankutty et al. 2008), and urban land use (Esch et al. 2014; Pesaresi et al. 2013). Another class of products draws on additional human activity data to map human alterations or pressures on the landscape (Ellis et al. 2013; Sanderson et al. 2002; Venter et al. 2016). Finally, several recent products consider nonterrestrial conditions such as air quality (Geddes et al. 2017; van Donkelaar et al. 2016, 2018) and surface water dynamics (Pekel et al. 2016).

While a comprehensive review of these data products is beyond the scope of this paper, it is important to recognize that such products provide significant new data streams for linking people and pixels and studies at the human–environment interface. The fact that many of these products are GIS-ready also promotes their use by a wider community of researchers, and represents an important development since the publication of *People and Pixels*, when remotely sensed data processing was the province of a select group of experts. Modeling, however, brings risks that uncertainties in the data may not be fully understood by those seeking "plug and play" data. It behooves users to read dataset documentation fully to understand fitness for use.

Gridded population data

As noted in the original *People and Pixels* volume, one contribution of remote sensing to social science is to extend the reach of social science data. Social science data based on censuses and surveys are limited in their spatial and temporal resolution because the spatial representation typically relies on administrative units and data are gathered only periodically, such as the decadal time frame common for censuses. They may also exclude hard to reach populations or locations near potential conflict zones where safety of the enumerators might be at risk. Countries where population data are not collected via censuses or registries, or where data are not made available, are especially challenging to incorporate into quantitative social science research. The wall-to-wall coverage, high spatial resolution, and relatively frequent observations of remotely sensed data can be leveraged to fill some of these gaps.

Gridded population data are typically created by starting with national census data linked to administrative boundaries, then distributing the population in each unit across grid cells falling within the unit. Techniques used to spatially distribute the population within administrative units span a spectrum from simple areal weighting methods to more complex techniques that utilize ancillary data to allocate population in more "realistic" patterns (Leyk et al. 2019). The areal weighting approach equally distributes the total count tied to a census unit across all grid cells within the boundary of the unit

(Doxsey-Whitfield et al. 2015). Slightly more sophisticated is a pycnophylatic approach—an approach which weights the redistribution of census population data to avoid abrupt discontinuities at boundaries between census units (Tobler 1979). Dasymetric mapping approaches rely on ancillary data, such as land cover, to disaggregate census counts (Balk et al. 2006; Eicher and Brewer 2001; Leyk et al. 2019). Starting with data like satellite-derived urban/rural land cover distributions (Balk et al. 2006), many current techniques use a variety of ancillary data, mainly derived via remote sensing (Sorichetta et al. 2015; Stevens et al. 2015).

Dasymetric mapping approaches can generally be categorized as binary, weighted or modeled, and hybrid. Binary dasymetric techniques rely on ancillary data to divide pixels into those that are populated and those that are not (Mennis and Hultgren 2006). Typically, settlement data, derived from remotely sensed land cover and/or nighttime lights data, are used to identify grid cells that are likely to be populated. The census-based population counts are distributed evenly among those cells, with no population allocated to cells that are unlikely to be populated. A variety of spatially detailed (10–50-m resolution) settlement layers with global coverage are now available (Esch et al. 2013; Pesaresi et al. 2013). Weighted and modeled techniques define a range of population density classes, often based on multiple sources of ancillary data (e.g., elevation, land cover, nighttime lights). Statistical models are used to assign weights to each population density class, and the census-based population is distributed according to these weights. Hybrid approaches may rely on a statistical weighting layer combined with a binary constraint (Stevens et al. 2015). For example, an analyst could assign population only to built areas and create a weighting layer based on a statistical model to allocate population within those areas (Schroeder 2017).

Several recent developments have returned to binary dasymetric techniques, emphasizing improvements in the creation of settlement data, rather than complex statistical approaches. Previously, imagery-based settlement data had limitations, especially in sparsely settled areas in regions where buildings are made of the same materials as surrounding landscapes and are therefore difficult to detect, or where nighttime lights emissions are too sparse or of low intensity to detect. Increased image resolution combined with machine learning techniques has addressed these limitations, enabling creation of settlement data that identify individual buildings (Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN - Columbia University 2016). Other approaches leverage both optical and radar-based data to create more accurate settlement layers (see World Settlement Footprint (WSF) for example).

Some gridded population products go beyond population at place of residence to represent temporal dynamics of human mobility, including diurnal, weekly, and seasonal population fluctuations due to commuting patterns and travel and tourism flows. Daily movement patterns have been addressed in population models since about the early 2000s (McPherson et al. 2004). The first large-scale country-wide model was produced for the USA (LandScan USA), containing both a nighttime residential as well as a baseline daytime population distribution incorporating movement of workers and students (Bhaduri et al. 2007). Daytime–nighttime maps have proven to be particularly important in disaster risk applications to address exposure dynamics and evacuation needs for rapid-onset hazardous events such as earthquakes, tsunamis, and flash floods (Aubrecht et al. 2014; Freire et al. 2013) or terrorist incidents (Ahola et al. 2007).

Moving beyond case-study–specific application to a full conceptual modeling and processing framework for spatiotemporal population mapping, Martin et al. (2015) developed SurfaceBuilder247, which incorporates both daily and seasonal movements. Most recently, daytime/nighttime grids have been developed under the European Commission's ENACT project for all 28 countries of the European Union, considering the presence of residents, workers by sector, students, and tourists, along with the locations of residence and activity (Batista e Silva et al. 2018).

Given the variety of gridded population products available, it is becoming increasingly important for users to carefully consider which product is best suited for their requirements, necessitating extensive source and metadata information.[1] Metadata should help researchers assess the quality (e.g., recency and granularity) of the census data and attribute accuracy of other ancillary data. The metadata should also provide transparent information on the locational accuracy of ancillary data sources to help researchers determine if data products are commensurate with their scale of interest. Another significant challenge for producers of gridded population data is establishing an appropriate means of validating final model outputs. The nature of attempting to fill a gap between census and remotely sensed data means that reliable, accurate data are seldom available against which to assess accuracy and model fit.

Poverty mapping

Gridded population data provide spatial distributions of the human population as a whole, but generally do not provide information on population characteristics beyond age and sex distributions (e.g., Center for International Earth Science Information Network 2017). Recent efforts have begun exploring the use of remotely sensed data to map population characteristics or particular segments of the population (Bosco et al. 2017; Doxsey-Whitfield et al. 2015). Poverty mapping, in particular, is of interest because poverty data from national censuses and household surveys are often of poor quality, lack spatial detail, or are altogether absent. Between 2002 and 2011, as many as 57 developing countries failed to conduct more than a single survey capable of producing poverty statistics, and data are the most scarce in the poorest countries (Serajuddin et al. 2015).

The global coverage and high temporal frequency available from remotely sensed data makes them attractive as a potential supplement to fill the gaps in spatial poverty data. While remotely sensed data cannot measure poverty directly, they can provide indicators of variations in the level of development over space. Nighttime lights (NTL) imagery has been used to describe variations in poverty at national and continental scales (Ghosh et al. 2013; Henderson et al. 2012; Pinkovskiy and Sala-i-Martin 2016). However, the coarse spatial resolution of NTL data (750 m for VIIRS DNB; 1 km for DMSP-OLS data) limits their ability to assess variability at local spatial scales.

Researchers are now exploring the use of high-resolution imagery (spatial resolution less than 5 m) to map poverty both within urban areas by identifying slum versus

---

[1] The POPGRID Data Collaborative, a consortium of the major gridded population and settlement data producers, was established to aid users in making these assessments. The POPGRID website (www. popgrid.org) includes data documentation, tables comparing data products, and a web map service that facilitates visualization and comparison across data products.

nonslum areas and for entire countries. This research is based on the assumption that spatial patterns of buildings and roads discernable from satellite imagery can characterize slums and poverty levels. A number of approaches to mapping slums have been tested, including simple visual interpretation; object-oriented approaches; machine learning based on spectral and spatial features; and others (Engstrom et al. 2015; Kohli et al. 2016; Kuffer et al. 2016; Taubenböck and Kraff 2014). Jean et al. (2016) used high spatial resolution data to map poverty across multiple sub-Saharan African countries at the village level using a convolutional neural network approach. Methods currently under development incorporate spatial and spectral features from the field of computer vision to detect areas with potential indicators of poverty (Engstrom et al. 2016; Graesser et al. 2012; Sandborn and Engstrom 2016). Rather than mapping objects (e.g., buildings, roads, cars), the spatial features approach characterizes variability in the image related to spatial pattern, structure, orientation, texture, and irregularity (Herold et al. 2003) that can be related to poverty (Engstrom et al. 2017a, 2017b).

Caveats for modeled data

Modeled and derived data products are attractive because they provide ready access to variables of interest, such as land cover, precipitation, and human population distribution, with complete spatial coverage. However, because these data products are not based on direct measurements of these variables at every point in space, they are inherently uncertain. Uncertainties are introduced when combining multiple sources of data, interpolating between locations or times, modeling relationships, applying machine learning, and any other approaches that essentially make educated guesses about variables of interest based on observable variables.

Uncertainty is present in all types of modeled data products, and different products that nominally describe the same variables inevitably differ from each other. For example, in classified land cover maps, the classification of a pixel at a given location may not be consistent from one product to another, particularly in transitional zones or for mixed-type classes (Congalton et al. 2014). Among the many gridded population products, the population density estimated for a given location may vary as a result of the particular input data, ancillary variables considered, and the estimation methods applied (Leyk et al. 2019).

Due to these uncertainties, users of modeled data products are well-advised to treat them with an appropriate degree of caution. Users should be sure to review documentation and published discussion of uncertainty for any modeled data products and consider the implications of such uncertainty for their analysis. When multiple products representing the same variable are available, users should carefully consider which product is best fit for their use case (Leyk et al. 2019). Ideally, research incorporating modeled data products should include some form of uncertainty or sensitivity analysis to evaluate the validity of conclusions in light of data uncertainties. At a minimum, researchers should be careful not to consider modeled data products as absolute truths.

In situ data from mobile devices

While advances in satellite and airborne sensors are increasing the spatial and temporal resolution of remotely sensed data, another data revolution is underway on the ground.

The introduction and rise to near-ubiquity of mobile devices over the past 20 years has generated entirely new streams of in situ data. In 2000, there were fewer than 700 million mobile phone subscriptions, and now there are more cell phones in the world than people. The World Bank (2016) has estimated that by 2020 there will be over 50 billion Internet-connected devices with the most significant growth and impact in the developing world. According to a survey by the Pew Research Center in the USA, while 65% of respondents had cell phones in 2004, this figure increased to 92% by 2015 (Anderson 2015). While mobile devices are most prevalent in the developed world, the developing world is rapidly adopting mobile technologies, and there is global recognition of the opportunity afforded by data produced by these devices.

Mobile devices are able to record the time and their location via GPS sensors or triangulated from the cellular towers or Wi-Fi networks with which they communicate. Therefore, observations based on mobile device activity are easily tagged with timestamps and spatial coordinates. Data may be gathered from mobile devices either intentionally by their users, for example through organized citizen science projects, or incidental to other mobile device usage, such as social media posts and cell phone calls. Such data have proven useful for assessing post-disaster situations (Cervone et al. 2016; Horanont et al. 2013; Lu et al. 2016a) and are increasingly being used to understand population flows such as daily commuting, travel, and mobility patterns (Mocanu et al. 2013; Wood et al. 2011).

Citizen science projects for environmental monitoring make use of "citizens as sensors" to intentionally collect data through crowdsourced observations (Hultquist and Cervone 2018). The data are stored and quickly disseminated using social networks or dedicated services and made available to study specific phenomena. Such data can be particularly valuable during and after natural disasters, when official sensors are not available or are unable to provide detail commensurate with rapid changes over time and space (Tidball and Krasny 2012). Ongoing work is developing techniques to incorporate these data into spatial analysis frameworks and integrate them with other types of spatial data (de Albuquerque et al. 2015; Schade et al. 2013). For example, images of flooded areas collected by citizens can be combined with remotely sensed data to determine accurate and detailed inundation boundaries (Panteras and Cervone 2018; Sava et al. 2017).

Data derived from mobile device usage also have the potential to increase the spatial and temporal resolution of data on human activities. Established sources of population data such as censuses and surveys have limited spatial and temporal resolution, typically describing resident populations in terms of administrative units and, at best, annual time steps. In contrast, data gleaned from traces of mobile device activity, such as locations of cell phone calls, text messages, or social media posts, can record specific movements and engagements with precise geographic coordinates and timestamps. These data therefore have the potential to extend population dynamics mapping toward near real-time representation of human mobility (Aubrecht et al. 2018; Deville et al. 2014). Unlike citizen science observations and other explicitly volunteered geographic information, mobile device users have not intentionally created these data points for the primary purpose of generating observations on users' location through time. Such data are nonetheless captured by service providers and can be made available to researchers (Lu et al. 2016a). Some social media data sources, such as Twitter and Foursquare, include publicly available text or image

content, which are also tagged with time and location (Patel et al. 2017; van Zanten et al. 2016).

However, using these incidental data for research carries challenges. While very widespread, access to and use of mobile devices are not universal. Connectivity can be limited by environmental and situational conditions (Lu et al. 2016b), and social media usage may be concentrated among particular population subgroups (Cesare et al. 2018). In addition, because data derived from mobile devices are incidental to other activities and not originally designed for research use, the data are unlikely to follow consistent sampling patterns, which presents challenges for removing noise and appropriately analyzing the data (Aubrecht et al. 2017). Another challenge for researchers with limited budgets is that access to mobile device data often requires proprietary data licensing agreements. Finally, ethical and legal standards on the use of data from mobile devices are still evolving. Even if the data are collected with users' explicit consent, researchers should evaluate ethical implications and ensure that individuals are not identified (Mikal et al. 2016).

New tools to facilitate data integration

Another important evolution in the data landscape has been the development of tools to facilitate the integration of data across people and pixels. These new tools include both custom-built platforms designed specifically for integrating these types of data and general purpose analytical software and programming languages that are becoming both more powerful and easier to use. These tools reduce the effort required to assemble and process large swaths of remotely sensed data, and facilitate derivation of information from remotely sensed data in formats and data structures more familiar to social scientists. In this section, we briefly describe three data integration platforms that provide data and functionality that are particularly useful for cross-disciplinary work combining people and pixels: IPUMS Terra, which was developed out of the social science community to facilitate the integration of physical variables derived from remotely sensed data with census-based population data; Giovanni and AppEEARS, both of which were developed primarily for the Earth science research communities but which offer advantages for social scientists; and Google Earth Engine, which provides a range of datasets and a powerful processing platform.

The Appendix provides more details about each of these platforms (sections "IPUMS Terra," "Giovanni and AppEEARS," and "Google Earth Engine"), as well as related information that is otherwise beyond the scope of this paper. The "Software for handling spatial data" section of the Appendix briefly surveys the state of spatial analysis software and developments in support for spatial data in more general-purpose data analysis tools. The "Spatial analysis methods" section in the Appendix provides key references addressing statistical and mathematical spatial analysis methods.

IPUMS Terra is a product of the IPUMS Data Center, which focuses primarily on population data from censuses and surveys. Building on IPUMS' large body of population data, IPUMS Terra is designed to integrate population and environmental data. IPUMS Terra utilizes location-based integration to transform among three data structures: individual-level population microdata, area-level data describing places defined by geographic boundaries, and pixel-based raster data (Kugler et al. 2015;

Ruggles et al. 2015). IPUMS Terra currently includes population microdata and associated area-level data from IPUMS International (Minnesota Population Center 2018); basic area-level population data from countries not participating in IPUMS International; and environmental data on climate (Harris et al. 2014; Hijmans et al. 2005), land cover (European Commission, Joint Research Centre 2003; Friedl et al. 2010), and agricultural land use (Monfreda et al. 2008; Ramankutty et al. 2008).

Giovanni and AppEEARS are tools developed by NASA to facilitate remote sensing data discovery, extraction, processing, and visualization (Adamo and de Sherbinin 2018). Giovanni's focus is largely on remotely sensed atmospheric data of relatively low spatial resolution and very high temporal resolution (including hourly datasets), concentrated primarily in the areas of atmospheric composition, atmospheric dynamics, global precipitation, hydrology, and solar irradiance. Giovanni enables online mapping and visualization, as well as data downloads. AppEEARS' focus is largely on remotely sensed terrestrial data of moderate spatial and temporal resolution (30 m to 1 km; daily time steps or monthly composites). Datasets available in AppEEARS include Terra & Aqua MODIS, the Shuttle Radar Topography Mission (SRTM v3), Web Enabled Landsat Data (WELD), Gridded population of the World (GPW), and NASA data products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument. AppEEARS allows users to subset large geospatial datasets based on user-defined polygons or points.

In addition to tools developed by the research community, Google is leveraging their extensive hardware platform and computing resources to manage and process increasingly large volumes of remotely sensed data. Google Earth Engine (GEE) serves as a free platform (free for noncommercial use) where users can easily visualize and download a relatively wide range of spatial data (Gorelick et al. 2017). GEE hosts satellite and other geospatial data gathered from a range of sources including Landsat, Sentinel, some MODIS data, Climate Hazards Infrared Precipitation with Stations (CHIRPS), and many other datasets. GEE also supports simple geospatial analysis through its own Explorer product and more complex analyses through an application programming interface (API).

People and pixels in practice

People and Pixels outlined several ways in which remote sensing had the potential to support social science research (Rindfuss and Stern 1998). Among the modes they discussed, we have described advances in how remotely sensed data are being used to provide additional measures for social science by extending the spatial and temporal detail of population data and measuring features and phenomena such as poverty, forest cover change, and climate variables. We now describe examples of social science research that leverage remote sensing's capabilities to measure social phenomena, their contexts, and their effects, particularly incorporating time-series data. These examples demonstrate how researchers use remotely sensed data to refine understanding of how people interact with the environment and to broaden perspectives from individual locations and events to broader scale patterns.

Measuring the environmental context of social phenomena

Many social science disciplines have foundational theories that consider how individuals or groups respond to changes in context over space and time. The increasing availability and accessibility of remotely sensed data has enabled researchers to more readily incorporate measures of agricultural production, surface water conditions, urbanization, rates of deforestation, and other factors to test these theories. Remotely sensed data have proven particularly valuable in the poorest countries, where land use and agricultural censuses and climate monitoring systems are otherwise lacking. The ability to measure changes in context over time with remotely sensed data allows researchers to isolate and investigate the actual social phenomena of interest.

Researchers have taken advantage of detailed contextual data from remote sensing in combination with individual-level health data to investigate interactions between food security, climate extremes, and health outcomes. In investigations related to food security, remotely sensed vegetation metrics such as the normalized difference vegetation index (NDVI) or the water requirement satisfaction index (WRSI) can be used to estimate cultivated area, vegetation, or drought at fine spatial and temporal scales. In some settings (namely rainfed and subsistence farming communities), these indices may be used to estimate agricultural production, livestock fodder, or food availability in a given community (Bakhtsiyarava et al. 2018; Brown et al. 2014). As such, remotely sensed data complement sources of health data on individuals, such as the Demographic and Health Surveys or the World Bank's Living Standards Measurement Study, by providing estimates for agricultural extent and growing season conditions in the area surrounding the community where the individual lives (Johnson and Brown 2014; Shively et al. 2015). Researchers can then examine how individual health outcomes vary in relationship to variability in the growing season or agricultural production across time and space. Similar approaches can be used to link climate extremes such as heat waves and drought to human health outcomes (Davenport et al. 2017; Grace et al. 2012; Isen et al. 2017), or climatic conditions to vector-borne disease outbreaks (Chretien et al. 2007). These approaches can be extended to examine multilevel interactions, incorporating individual, community, and environmental characteristics. Disciplines within public health, including epidemiology and environmental health sciences, have also explored linkages between health and context using remotely sensed data. Remotely sensed data have been used in investigations of infectious disease outbreaks, cancer occurrence, and linkages between air pollution and health (Buczak et al. 2012; Dey et al. 2012; Maxwell et al. 2010; Stefani et al. 2013).

Remotely sensed data have also been applied to better understand the environmental contexts relevant to historical populations. For example, archeologists have used high-resolution synthetic aperture radar (SAR) data to map ancient irrigation systems that now lie underground (Wiig et al. 2018). Anthropologists have also employed viewshed analysis utilizing a SAR-based digital elevation model to gain insight about settlement and activity locations on San Clemente Island (Comer et al. 2013). These analyses revealed settlements situated within view of both each other and expanses of ocean frequented by marine mammals, facilitating hunting activities, and ritual sites within view of sites on other islands, reinforcing cultural and social relationships among islands. Remotely sensed data has also helped explain the paradox of Petra, a great ancient city built by the Nabateans. For centuries prior to the construction of Petra, the

Nabateans were nomadic, amassing great wealth by transporting precious goods through the vast deserts of the Arabian Peninsula (Comer 2013). Overlaying Nabatean temples and terraces on NDVI images derived from satellite data reveal that they appeared only in areas where agriculture is possible, suggesting a lifestyle transition from nomadic to agricultural, probably driven by Roman actions disrupting the Nabateans' control of trade in the region. By exploiting environmental information gleaned from multispectral data, Sever explored the relationship between Mayan settlements and wetlands (Sever and Irwin 2003) and argued convincingly for overexploitation of the environment as the cause of the collapse of the Mayan civilization (Sever 1998).

Measuring social phenomena and their effects

A major thrust of *People and Pixels* was understanding the human drivers of land cover change, especially in rapidly changing regions like the Amazon. At the time, the available spatial and temporal resolution of remotely sensed data enabled identification of land cover classes and terrain features, and researchers used this information to link changes in land cover to demographic changes. This kind of research still features prominently (de Sherbinin 2016) including studies relating population change to land cover change (Aide et al. 2013), biodiversity loss (Williams 2013), and fire activity (Hantson et al. 2015). Yet, in recent years, the range of application areas has expanded. With increased spatial and temporal resolution, researchers are moving beyond land cover and beginning to infer land use and other human activities through changes detectable in remotely sensed data. Researchers have used remotely sensed data to infer agricultural management practices (Gómez Giménez et al. 2017; Jain et al. 2016) and evaluate the effect of governance policies on forest protection (Blackman et al. 2017).

Researchers are also beginning to use remotely sensed data to investigate phenomena, such as narcotics trafficking and population responses to civil war, that are difficult to detect in traditional census and survey data. High-resolution, high-frequency remotely sensed data enable social scientists to detect the landscape signature of these activities, and thereby use human–environment interactions as a means of developing and testing theories about clandestine and otherwise difficult to measure phenomena. For example, over the past decade, social scientists working in Central American contexts began to notice new patterns of large, rapid, forest clearing in remote regions, including protected and indigenous areas (Grandia 2013; McSweeney et al. 2014, 2017). They hypothesized that these patterns may be linked to narcotics traffickers acquiring land over which to move cocaine and for money laundering activities such as cattle and oil palm production (McSweeney et al. 2017). This hypothesis was tested by using the analysis-ready Hansen forest change data (Hansen et al. 2013) to develop spatial and temporal pattern metrics for patches of forest loss and identifying statistically "unique" groups of deforestation (Sesnie et al. 2017). Anomalous deforestation in several departments of Nicaragua, Honduras, and Guatemala was found to be significantly correlated to cocaine flow data available at the departmental level, accounting for between 15 and 30% of total forest loss in the region between 2000 and 2014.

The time-series data available from repeat imaging by satellite also holds promise for understanding otherwise difficult to detect consequences of civil wars. Groups such as UNOSAT, Human Rights Watch, Amnesty International, and satellite data-savvy

journalists have used sequences of very high-resolution imagery to measure urban and infrastructural damage following armed conflict across a range of contexts and countries. While these satellite mapping efforts have been essential at raising awareness and supporting humanitarian relief, their focus on acute damage is only one component of the war economy taxonomy, which includes "combat," "shadow," and "coping" economies (Goodhand 2004). The combat economy directly supports military objectives, including territorial capture and control over natural resource exploitation; the shadow economy involves commodity smuggling and high-value resource extraction; and the coping economy encompasses adaptations by the noncombatant population, such as labor migration and changes in agricultural production (Auty 2001; Collier and Hoeffler 2002; Fearon and Laitin 2003; Hegre 2004; Ross 2004). Of the three aspects, the coping economy has been the least studied (Ballentine and Nitzschke 2005), yet is well suited for investigation by remote sensing. Moderate-resolution imagery with subweekly repeat coverage and high-resolution weekly imagery now being made available through the Copernicus program support examination across seasons, years, and war periods (Witmer 2015). Land cover changes including urban construction, deterioration, and damage and fluctuations in agricultural production can therefore be detected across broad spatial extents (Eklund et al. 2017; Klaasse et al. 2018). These changes may, in turn, signal population relocation or changes in agricultural practices that reflect the coping economy.

Conclusion

When *People and Pixels* was published 20 years ago, it showcased the work of a relatively small group of social scientists beginning to use newly available remote sensing data. These researchers glimpsed the possibilities that connecting people to pixels offered for understanding patterns and processes of human–environment interactions and invested significant effort in making connections across scientific domains and data structures. Key institutions also recognized the possibilities and supported these efforts with funding, including NASA's Land Use Land Cover Change Program and NIH/NIMH funding for social and population scientists to use GIS and remotely sensed data in their work (Entwisle and Stern 2005).

Over the past two decades, technological and cultural transformations have continued. Remote sensing data continues to be collected at ever-finer spatial and temporal resolutions. Processing methods to extract useful information from these high-volume data continue to develop, and new software and tools make the data more accessible. These technical advances drive, and are driven by, increasing spatial awareness throughout the social sciences and allied fields. Economists, public health researchers, anthropologists, demographers, and political scientists are all increasingly interested in not just the locations and times of specific events but the spatiotemporal patterns and processes that underlie series of related events and interactions.

The technical advances and the increased interest in spatial approaches throughout the social sciences are undoubtedly improving the connection between people and pixels. However, important challenges remain. First, we must not lose sight of ethical issues surrounding the collection and use of spatial data in social science applications. Interest in linking traditional social science data with remotely sensed data has driven

increased demand for more precise spatial locations in census and survey data. However, such spatially explicit data raise concerns and introduce tensions surrounding confidentiality, data preservation, and data sharing for reproducibility (VanWey et al. 2005). Current approaches, such as that employed by the Demographic Health Surveys, address these concerns by introducing uncertainty in spatial locations, which then necessitates accounting for the uncertainty in analyses.

Second, uncertainty deliberately introduced to protect confidentiality is by no means the only source of uncertainty when making connections between people and pixels. Modeled and derived data products, such as those discussed in the "Derived/modeled data products" section, have been a boon to facilitating work across disciplinary boundaries. However, because the final data products effectively mask the complexities of the processes used to create them, users of such data may be tempted to treat them as "truth." Researchers must maintain awareness of the multifaceted uncertainties in any data they use and their potential effect on analyses. Developing readily adopted means of accounting for and communicating uncertainty, therefore, continues to be a fruitful direction for research.

Finally, another major challenge lies in the balance of research effort between technical and substantive approaches. While technological advances will undoubtedly continue, the questions posed by social scientists will not be answered by higher resolution data or more processing power alone. We must ensure that the social science research efforts receive adequate attention and support and are not overwhelmed by the lure of new technology for its own sake. The trend toward greater interdisciplinarity to address issues and concerns that cut across social and environmental science is beneficial in this regard. Such complex questions are of great concern to policy makers and citizens, and addressing them requires both theoretical frameworks encompassing social and environmental domains as well as data on both people and pixels to provide the empirical underpinning. In addition to the social and environmental sciences, a key aspect of people and pixels is that remote sensing sciences must be closely involved in the conversation. Continued dialog will enable remote sensing scientists to understand what advances would be most beneficial to better understand landscape variability relevant to socio-environmental challenges and enable social and environmental scientists to better understand the limitations and advantages of the array of remotely sensed data products.

Compliance with ethical standards　No experiments were performed in the preparation of this work, other than those performed for previously published work.

Conflict of interest　The authors declare that they have no conflict of interest.

Appendix

IPUMS Terra

IPUMS Terra aims to foster more widespread use of linked population and environmental data by taking the heavy lifting of data integration and transformation off the hands of individual researchers. IPUMS Terra utilizes location-based integration to transform among three data structures: individual-level population microdata, area-level data describing places defined by geographic boundaries, and pixel-based raster data. The IPUMS Terra data access system provides a workflow through which users first select the data structure in which they would like to receive their output data. They then select data of interest from one or more of the source data structures and make basic decisions about how transformations are conducted. The system processes the requested data, handling any necessary transformations, and delivers a customized, integrated dataset in the user's chosen data structure.

The transformations performed by IPUMS Terra depend on the output data structure and the types of input data selected (Kugler et al. 2015; Ruggles et al. 2015). Users may choose to receive area-level output data (describing administrative units) and include both census-based aggregate population data and raster data on land cover or climate as inputs. IPUMS Terra then transforms the raster data by calculating summary measures (e.g., % area of forest cover, mean annual precipitation) for the administrative units described in the aggregate population data. Area-level population and environmental data can also be attached as contextual variables to individual microdata records, based on the administrative unit in which the person lives. Conversely, IPUMS Terra can distribute values from area-level population data across the pixels located within each administrative unit to produce raster output.

The data available through IPUMS Terra encompass a range of population and environmental characteristics. IPUMS Terra's population data are primarily drawn from IPUMS International (Minnesota Population Center 2018), which provides over one billion microdata records from 94 countries, with multiple census years available for many countries. IPUMS Terra has also pre-tabulated a series of area-level variables from the microdata. On the pixel side of the data collection, IPUMS Terra includes data on land cover (European Commission, Joint Research Centre 2003; Friedl et al. 2010), harvested area and yield of 175 individual crops (Monfreda et al. 2008; Ramankutty et al. 2008), long-term climate baselines (Hijmans et al. 2005), and time-series climate data (Harris et al. 2014). Transformations between population microdata, area-level data, and raster data are supported by IPUMS Terra's collection of geographic administrative unit boundary data. For most countries, IPUMS Terra provides first- and second-level administrative unit boundaries.

Giovanni and AppEEARS

Giovanni and AppEEARS are tools developed by NASA to facilitate remote sensing data discovery, extraction, processing, and visualization. Giovanni's focus is largely on remotely sensed atmospheric data of relatively low spatial resolution and very high temporal resolution, including hourly datasets. The tool uses server-side processing to process requests "on the fly," and produces map and graph visualizations. AppEEARS'

focus is largely on remotely sensed terrestrial data of moderate spatial and temporal resolution (30 m to 1 km; daily time steps or monthly composites). AppEEARS delivers results within minutes or hours (depending on the degree of processing) by email notification. Both tools enable analysis of time trends—Giovanni for one pixel or averaged over a larger area comprising many pixels, and AppEEARS for multiple points or polygons. Additional tools and use cases of Giovanni and AppEEARS are described in Adamo and de Sherbinin (2018).

Giovanni's remotely sensed data assets are concentrated primarily in the areas of atmospheric composition, atmospheric dynamics, global precipitation, hydrology, and solar irradiance. More than 1,850 variables are available, searchable via a keyword and faceted search. Example variables of potential interest to social and health scientists include precipitation, flooding, air temperature, PM2.5 particulate pollution, and UV exposure. Mapping options include time averaging, animation, precipitation accumulation, time-averaged overlay of two datasets, and user-defined climatology. For time-series plots, options include area averaged, differences, and seasonal data (e.g., July–August for daily max temperatures over multiple years). Visualization features include animation, interactive scatterplots, contouring, and scaling (linear or log).

AppEEARS provides access to more than 100 datasets from Terra & Aqua MODIS, the Shuttle Radar Topography Mission (SRTM v3), Web Enabled Landsat Data (WELD), Gridded population of the World (GPW), and NASA data products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument. Selected variables of interest to social scientists include land surface temperature, land cover type, population counts and densities, elevation, and vegetation indices. A key feature of AppEEARS is the ability to subset large geospatial datasets and to extract data for specific locations or regions around the world. Data can be extracted using an area sample via vector polygons (bounding box, Shapefile, or GeoJSON) or a point sample with geographic coordinates. In addition to data values, the user also receives quality values. The results preview includes interactive visualizations, decoded quality information, and summary statistics.

Google Earth Engine

In addition to tools developed by the research community, Google is leveraging their extensive hardware platform and computing resources to manage and process increasingly large volumes of remotely sensed data. Google Earth Engine (GEE) serves as a free platform (free for noncommercial use) where users can easily visualize and download a relatively wide range of spatial data (Gorelick et al. 2017). GEE hosts satellite and other geospatial data gathered from a range of sources including Landsat, Sentinel, some MODIS data, Climate Hazards Infrared Precipitation with Stations (CHIRPS), and many other datasets. Satellite images are processed daily and GEE provides data for some areas covering up to 40 years. Users can specify their locations and temporal scale of interest and download data for analyses in other spatial or statistical analyses packages.

Beyond hosting satellite and geospatial data, GEE also supports simple geospatial analysis through its own Explorer product and more complex analyses through an application programming interface (API). The API allows researchers and developers to leverage both the GEE data collection and Google's cloud-based computing

resources. The combination of data and processing power makes the development of modeled environmental products at high resolution and global extent feasible (Hansen et al. 2013; Pekel et al. 2016). For example, the Hansen et al. forest change data, based on 650,000 30-m resolution Landsat scenes, would have taken 15 years to process on a single computer, but was completed on GEE in a matter of days (University of Maryland 2013). The API also enables near real-time and interactive tools for detecting and monitoring land cover change (Bey et al. 2016; Hansen et al. 2016).

Software for handling spatial data

In the 20 years since the publication of *People and Pixels*, growth in the availability of high spatial and temporal resolution remotely sensed data has coincided with growth in both software designed to work with spatial data and the ability of general purpose software to handle spatial data. As of the late 1990s, remotely sensed and other spatial data were manipulated and analyzed in specialized software packages, such as Esri's ArcInfo, Erdas Imagine, ENVI, and GRASS, on stand-alone computers or servers. These packages, with the exception of GRASS, are commercial products with attendant cost considerations.

The original spatial software packages have continued to evolve, expand, and become more accessible to nonspecialists. Open-source geospatial software has also expanded greatly, leading to the development of organizations like the Open Source Geospatial Foundation (www.osgeo.org) that, in turn, foster further growth. As a result, awareness and use of geographic information systems (GIS) and other spatial software now extend across a broad range of disciplines and educational levels. For example, GIS is increasingly used in secondary schools in subjects ranging from natural sciences to history (Kerski et al. 2013).

In addition to specialized spatial software, scholars increasingly use general-purpose, open-source software, such as R, Python, and SQL, to process, analyze, and integrate remotely sensed data. The availability of support for spatial data in these general-purpose tools enables researchers to conduct both spatial and nonspatial analysis seamlessly in a single environment. R and Python are free, widely used, open-source languages that may be modified and enhanced by any programmer in the world. R is a programming language and environment focused mainly on statistical analysis and graphing. The R community has developed spatially oriented packages supporting geographic data structures, geospatial processing operations, geostatistics, and geovisualization. These packages allow researchers to integrate remotely sensed and other spatial data into analytical workflows, which may include statistical modeling or other statistical techniques, within a single software ecosystem (Lovelace et al. 2018). R packages for handling spatial data and analysis, like *sp* (supporting spatial data) and more recently *sf* (supporting "simple features") *raster*, *rgdal*, *GISTools*, *spdep*, *gstat*, and many others, have helped to make R an ideal platform to merge, manipulate, and display spatial data. Python is an open-source, general-purpose, object-oriented programming language used in a wide variety of domains. Software developers have written Python libraries dedicated to geographic data, including remotely sensed data. These libraries support spatial data input/output, geoprocessing, visualization, and spatial statistics and modeling (Rey 2017). Support for spatial data and spatial queries

has also been incorporated into relational databases, both commercial (e.g., SQL Server, Oracle) and open source (e.g., PostGIS).

Spatial analysis methods

This paper has primarily focused on the data landscape and the conceptual linkages and challenges related to contemporary "people and pixels" analyses. Beyond the scope of this project are the many discussions related to spatial statistics and other statistical or mathematical models and approaches that are used to analyze these data. As with any other spatial analysis work, researchers should be aware of the available methods and their capabilities and limitations for particular applications. Those interested in this area of research are advised to consider work by Schabenberger and Gotway (2017), Maxwell et al. (2018), Tong and Murray (2012), Long and Nelson (2013), and many others.

References

Adamo, S., & de Sherbinin, A. (2018). Remote sensing for social scientists: tools to facilitate the use of remote sensing data in social and health sciences research. Accessed at http://sedac.ciesin.columbia.edu/guides

Ahola, T., Virrantaus, K., Krisp, J. M., & Hunter, G. J. (2007). A spatio-temporal population model to support risk assessment and damage analysis for decision-making. *International Journal of Geographical Information Science, 21*(8), 935–953.

Aide, T. M., Clark, M. L., Grau, H. R., López-Carr, D., Levy, M. A., Redo, D., et al. (2013). Deforestation and reforestation of Latin America and the Caribbean (2001–2010). *Biotropica, 45*(2), 262–271.

Anderson, M. (2015). Technology device ownership: 2015. Pew Research Center, October 2015. Retrieved from http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/. Accessed Aug 2018

Arino, O., Gross, D., Ranera, F., Leroy, M., Bicheron, P., Brockman, C., ... & Bourg, L. (2007). GlobCover: ESA service for global land cover from MERIS. *2007 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 2412–2415.

Association of European Space Research Establishments (ESRE) (2017). Selected trends and space technologies expected to shape the next decade. *ESRE Whitepaper*, The Association of European Space Research Establishments (ESRE), November 2017.

Aubrecht, C., Özceylan Aubrecht, D., & Steinnocher, K. (2015). New perspectives on population data in spatial modeling for crisis management: where are we headed? In J. D. VanVactor (Ed.), *Crisis management: a leadership perspective* (pp. 87–92). Hauppauge, NY: Nova Science.

Aubrecht, C., Özceylan Aubrecht, D., Ungar, J., Freire, S., & Steinnocher, K. (2017). VGDI–advancing the concept: volunteered geo-dynamic information and its benefits for population dynamics modeling. *Transactions in GIS, 21*(2), 253–276. https://doi.org/10.1111/tgis.12203.

Aubrecht, C., Steinnocher, K., & Huber, H. (2014). DynaPop—population distribution dynamics as basis for social impact evaluation in crisis management. In S.R. Hiltz, M.S. Pfaff, L. Plotnick, & P.C. Shih (Eds.), *ISCRAM 2014 conference proceedings*, 11th international conference on information systems for crisis response and management, 319-323.

Aubrecht, C., Ungar, J., Özceylan Aubrecht, D., Freire, S., & Steinnocher, K. (2018). Mapping land use dynamics using the collective power of the crowd. In P. P. Mathieu & C. Aubrecht (Eds.), *Earth observation open science and innovation* (Springer. ISSI scientific report series) (Vol. 15, pp. 247–253).

Auty, R. M. (2001). The political economy of resource-driven growth. *European Economic Review, 45*(4–6), 839–846.

Bakhtsiyarava, M., Grace, K., & Nawrotzki, R. J. (2018). Climate, birth weight, and agricultural livelihoods in Kenya and Mali. *American Journal of Public Health, 108*(S2), S144–S150.

Balk, D. L., Deichmann, U., Yetman, G., Pozzi, F., Hay, S. I., & Nelson, A. (2006). Determining global population distribution: methods, applications and data. *Advances in Parasitology, 62*, 119–156.

Ballentine, K., & Nitzschke, H. (2005). The political economy of civil war and conflict transformation. In M. Fischer & B. Schmelzle (Eds.), *Transforming war economies: dilemmas and strategies* (pp. 11–33). Berlin: Berghof Research Center for Constructive Conflict Management.

Batista e Silva, F., Rosina, K., Schiavina, M., Marín-Herrera, M., Freire, S., Ziemba, L., Craglia, M., & Lavalle, C. (2018). *From place of residence to place of activity: towards spatiotemporal mapping of population density in Europe*. Lund: Proceedings of the AGILE conference 2018.

Bartholome, E., & Belward, A. S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing, 26*(9), 1959–1977.

Belward, A., & Skøien, J. O. (2015). Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing, 103*, 115–128.

Bey, A., Sánchez Paus Díaz, A., Maniatis, D., Marchi, G., Mollicone, D., Ricci, S., et al. (2016). Collect Earth: land use and land cover assessment through augmented visual interpretation. *Remote Sensing, 8*(10), 1–24.

Bhaduri, B., Bright, E., Coleman, P., & Urban, M. L. (2007). LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal, 69*(1–2), 103–117.

Blackman, A., Corral, L., Santos Lima, E., & Asner, G. P. (2017). Titling indigenous communities protects forests in the Peruvian Amazon. *Proceedings of the National Academy of Sciences, 114*(16), 4123–4128.

Bosco, C., Alegana, V., Bird, T., Pezzulo, C., Bengtsson, L., Sorichetta, A., et al. (2017). Exploring the high-resolution mapping of gender-disaggregated development indicators. *J R Soc Interface, 14*(129), 20160825.

Brown, M. E., Grace, K., Shively, G., Johnson, K. B., & Carroll, M. (2014). Using satellite remote sensing and household survey data to assess human health and nutrition response to environmental change. *Population and Environment, 36*(1), 48–72.

Buczak, A. L., Koshute, P. T., Babin, S. M., Feighner, B. H., & Lewis, S. H. (2012). A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data. *BMC Medical Informatics and Decision Making, 12*(1), 124.

Cesare, N., Lee, H., McCormick, T., Spiro, E., & Zagheni, E. (2018). Promises and pitfalls of using digital traces for demographic research. *Demography, 55*(5), 1979–1999.

Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J., & Waters, N. (2016). Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *International Journal of Remote Sensing, 37*(1), 100–124.

Chretien, J. P., Anyamba, A., Bedno, S. A., Breiman, R. F., Sang, R., Sergon, K., … & Linthicum, K.J. (2007). Drought-associated chikungunya emergence along coastal East Africa. Am J Trop Med Hyg 76 (3), 405–7.

Center for International Earth Science Information Network - CIESIN - Columbia University. (2017). Gridded population of the world, version 4 (GPWv4): basic demographic characteristics, revision 10. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H45H7D7F

Comer, D. C. (2013). Petra and the paradox of a great city built by nomads: an explanation suggested by satellite imagery. In D. C. Comer & M. J. Harrower (Eds.), *Mapping archaeological landscapes from space* (pp. 73–83). New York: Springer.

Comer, D. C., Blom, R. G., & Megarry, W. (2013). The influence of viewshed on prehistoric archaeological site patterning at San Clemente Island as suggested by analysis of synthetic aperture radar images. In D. C. Comer & M. J. Harrower (Eds.), *Mapping archaeological landscapes from space* (pp. 159–171). New York: Springer.

Collier, P., & Hoeffler, A. (2002). On the incidence of civil war in Africa. *Journal of Conflict Resolution, 46*(1), 13–28.

Congalton, R. G., Gu, J., Yadav, K., Thenkabail, P., & Ozdogan, M. (2014). Global land cover mapping: a review and uncertainty analysis. *Remote Sensing, 6*(12), 12070–12093. https://doi.org/10.3390/rs61212070.

Davenport, F., Grace, K., Funk, C., & Shukla, S. (2017). Child health outcomes in sub-Saharan Africa: a comparison of changes in climate and socio-economic factors. *Global Environmental Change, 46*, 72–87.

Denis, G., Claverie, A., Pasco, X., Darnis, J. P., de Maupeou, B., Lafaye, M., & Morel, E. (2017). Towards disruptions in Earth observation? New Earth observation systems and markets evolution: possible scenarios and impacts. *Acta Astronautica, 137*, 415–433.

de Albuquerque, J. P., Herfort, B., Brenning, A., & Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science, 29*(4), 667–689.

de Sherbinin, A. (2016). Remote sensing and socioeconomic data integration: lessons from the NASA Socioeconomic Data and Applications Center. In D. A. Quattrochi, E. A. Wentz, N. Lam, & C. Emerson (Eds.), *Integrating scale in remote sensing and GIS* (pp. 371–388). Boca Raton: CRC.

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., et al. (2014). Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences, 111*(45), 15888–15893.

Dey, S., Di Girolamo, L., van Donkelaar, A., Tripathi, S. N., Gupta, T., & Mohan, M. (2012). Variability of outdoor fine particulate (PM2.5) concentration in the Indian subcontinent: a remote sensing approach. *Remote Sensing of Environment, 127*, 153–161.

d'Oliveira, F. A., Melo, F. C. L. D., & Devezas, T. C. (2016). High-altitude platforms—present situation and technology trends. *Journal of Aerospace Technology and Management, 8*(3), 249–262.

Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolesi, L., Squires, J., Borkovska, O., & Baptista, S. R. (2015). Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4. *Papers in Applied Geography, 1*(3), 226–234.

Eicher, C. L., & Brewer, C. A. (2001). Dasymetric mapping and areal interpolation: implementation and evaluation. *Cartography and Geographic Information Science, 28*(2), 125–138.

Eklund, L., Degerald, M., Brandt, M., Prishchepov, A. V., & Pilesjö, P. (2017). How conflict affects land use: agricultural activity in areas seized by the Islamic State. *Environmental Research Letters, 12*(5), 054004.

Ellis, E. C., Goldewijk, K. K., Siebert, S., Lightman, D., & Ramankutty, N. (2013). *Anthropogenic biomes of the world, version 2: 2000*. Palisades: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4D798B9.

Engstrom, R., Copenhaver, A., Newhouse, D., Hersh, J., & Haldavanekar, V. (2017a). Evaluating the relationship between spatial and spectral features derived from high spatial resolution satellite data and urban poverty in Colombo, Sri Lanka. *Joint Urban Remote Sensing Event (JURSE 2017)*. https://doi.org/10.1109/JURSE.2017.7924590.

Engstrom, R., Copenhaver, A., & Qi, Y. (2016). Evaluating the use of multiple imagery-derived spatial features to predict census demographic variables in Accra, Ghana. In *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International* (pp. 7318-7321). IEEE.

Engstrom, R., Hersh, J., & Newhouse, D. L. (2017b). *Poverty from space: using high-resolution satellite imagery for estimating economic well-being. (Policy Research WPS 8284)*. Washington: World Bank Group.

Engstrom, R., Sandborn, A., Yu, Q., Burgdorfer, J., Stow, D., Weeks, J., & Graesser, J. (2015). Mapping slums using spatial features in Accra, Ghana. In *2015 Joint urban remote sensing event (JURSE)* (pp. 1–4). IEEE.

Entwisle, B., & Stern, P. C. (Eds.). (2005). *Population, land use, and environment: research directions*. Washington: National Academies Press.

Esch, T., Kurz, F., Pacini, F., Ottavianelli, G., & Vojtek, D. (2017). Capitalising high altitude pseudo satellites for urban applications—first results from the TEP Urban Project. *HAPS4ESA workshop*. Noordwijkerhout, The Netherlands.

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., et al. (2013). Urban footprint processor—fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. *IEEE Geoscience and Remote Sensing Letters, 10*(6), 1617–1621.

Esch, T., Marconcini, M., Marmanis, D., Zeidler, J., Elsayed, S., Metz, A., et al. (2014). Dimensioning urbanization—an advanced procedure for characterizing human settlement properties and patterns using spatial network analysis. *Applied Geography, 55*, 212–228.

European Commission, Joint Research Centre. (2003). *Global land cover 2000 database*. Retrieved from http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php. Accessed Jan 2013

Facebook Connectivity Lab & Center for International Earth Science Information Network - CIESIN - Columbia University. (2016). High resolution settlement layer (HRSL). Source imagery for HRSL © 2016 DigitalGlobe. Retrieved from https://www.ciesin.columbia.edu/data/hrsl/. Accessed Jul 2018

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review, 97*(1), 75–90.

Freire, S., Aubrecht, C., & Wegscheider, S. (2013). Advancing tsunami risk assessment by improving spatio-temporal population exposure and evacuation modeling. *Natural Hazards, 68*(3), 1311–1324.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sensing of Environment, 114*(1), 168–182.

Geddes, J. A., Martin, R. V., Boys, B. L., & van Donkelaar, A. (2017). *Global 3-year running mean ground-level nitrogen dioxide (NO2) grids from GOME, SCIMACHY and GOME-2*. Palisades: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4JW8BTT.

Ghosh, T., Anderson, S., Elvidge, C., & Sutton, P. (2013). Using nighttime satellite imagery as a proxy measure of human well-being. *Sustainability, 5*(12), 4988–5019.

Gibbs, H. K., Ruesch, A. S., Achard, F., Clayton, M. K., Holmgren, P., Ramankutty, N., & Foley, J. A. (2010). Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. *Proceedings of the National Academy of Sciences, 107*(38), 16732–16737.

Gómez Giménez, M., de Jong, R., Della Peruta, R., Keller, A., & Schaepman, M. E. (2017). Determination of grassland use intensity based on multi-temporal remote sensing data and ecological indicators. *Remote Sensing of Environment, 198*, 126–139.

Goodhand, J. (2004). From war economy to peace economy? Reconstruction and state building in Afghanistan. *Journal of International Affairs, 58*(1), 155–174.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18–27.

Grace, K., Davenport, F., Funk, C., & Lerner, A. M. (2012). Child malnutrition and climate in sub-Saharan Africa: an analysis of recent trends in Kenya. *Applied Geography, 35*(1–2), 405–413.

Grandia, L. (2013). Road mapping: megaprojects and land grabs in the northern Guatemalan lowlands. *Development and Change, 44*(2), 233–259.

Graesser, J., Cheriyadat, A., Vatsavai, R. R., Chandola, V., Long, J., & Bright, E. (2012). Image based characterization of formal and informal neighborhoods in an urban landscape. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*(4), 1164–1176.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science, 342*(6160), 850–853.

Hansen, M. C., Krylov, A., Tyukavina, A., Potapov, P. V., Turubanova, S., Zutta, B., et al. (2016). Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters, 11*(3), 034008.

Hantson, S., Pueyo, S., & Chuvieco, E. (2015). Global fire size distribution is driven by human impact and climate. *Global Ecology and Biogeography, 24*(1), 77–86.

Harris, I. P. D. J., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. *International Journal of Climatology, 34*(3), 623–642.

Hegre, H. (2004). The duration and termination of civil war. *Journal of Peace Research, 41*(3), 243–252.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *The American Economic Review, 102*(2), 994–1028.

Herold, M., Liu, X., & Clarke, K. C. (2003). Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering and Remote Sensing, 69*(9), 991–1001.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25*(15), 1965–1978.

Horanont, T., Witayangkurn, A., Sekimoto, Y., & Shibasaki, R. (2013). Large-scale auto-GPS analysis for discerning behavior change during crisis. *IEEE Intelligent Systems, 28*(4), 26–34.

Hultquist, C., & Cervone, G. (2018). Citizen monitoring during hazards: validation of Fukushima radiation measurements. *GeoJournal, 83*(2), 189–206.

Isen, A., Rossin-Slater, M., & Walker, R. (2017). Relationship between season of birth, temperature exposure, and later life wellbeing. *Proceedings of the National Academy of Sciences, 114*(51), 13447–13452.

Jain, M., Srivastava, A., Joon, R., McDonald, A., Royal, K., Lisaius, M., & Lobell, D. (2016). Mapping smallholder wheat yields and sowing dates using micro-satellite data. *Remote Sensing, 8*(10), 860.

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*(6301), 790–794.

Johnson, K., & Brown, M. E. (2014). Environmental risk factors and child nutritional status and survival in a context of climate variability and change. *Applied Geography, 54*, 209–221.

Jun, C., Ban, Y., & Li, S. (2014). China: open access to Earth land-cover map. *Nature, 514*(7523), 434. https://doi.org/10.1038/514434c.

Kerski, J. J., Demirci, A., & Milson, A. J. (2013). The global landscape of GIS in secondary education. *Journal of Geography, 112*(6), 232–247. https://doi.org/10.1080/00221341.2013.801506.

Klaasse, A., Haas, E., Dost, R., Riffler, M., Shamsiev, B., & Garrett, K. (2018). Satellites for Syria: new methods for assessing agricultural production in conflict areas to support productivity assessments, rehabilitation efforts, and (post conflict) assistance. *Land and Poverty Conference 2018: Land Governance in an Interconnected World,* World Bank. Retrieved from https://www.conftool.com/landandpoverty2018/index.php/10-11-Klaasse-339_paper.pdf?page=downloadPaper&filename=10-11-Klaasse-339_paper.pdf&form_id=339&form_version=final. Accessed Aug 2018

Kohli, D., Stein, A., & Sliuzas, R. (2016). Uncertainty analysis for image interpretations of urban slums. *Computers, Environment and Urban Systems, 60*, 37–49.

Kugler, T. A., Van Riper, D. C., Manson, S. M., Haynes, D. A., II, Donato, J., & Stinebaugh, K. (2015). Terra Populus: workflows for integrating and harmonizing geospatial population and environmental data. *Journal of Map & Geography Libraries, 11*(2), 180–206.

Kuffer, M., Pfeffer, K., & Sliuzas, R. (2016). Slums from space—15 years of slum mapping using remote sensing. *Remote Sensing, 8*(6), 455.

Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., et al. (2019). Allocating people to pixels: a review of large-scale gridded population data products and their fitness for use. *Earth System Science Data Discussions*. https://doi.org/10.5194/essd-2019-82.

Li, X., Gong, P., & Liang, L. (2015). A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data. *Remote Sensing of Environment, 166*(Sept.), 78–90. https://doi.org/10.1016/J.RSE.2015.06.007.

Liverman, D., Moran, E. F., Rindfuss, R. R., & Stern, P. C. (Eds.). (1998). *People and pixels: linking remote sensing and social science*. Washington, DC: National Academies Press.

Long, J. A., & Nelson, T. A. (2013). A review of quantitative methods for movement data. *International Journal of Geographical Information Science, 27*(2), 292–318.

Lovelace, R., Nowosad, J., & Muenchow, J. (2018). *Geocomputation with R*. Retrieved from https://geocompr.robinlovelace.net/. Accessed June 2018

Lu, X., Wrathall, D. J., Sundsøy, P. R., Nadiruzzaman, M., Wetter, E., Iqbal, A., et al. (2016a). Unveiling hidden migration and mobility patterns in climate stressed regions: a longitudinal study of six million anonymous mobile phone users in Bangladesh. *Global Environmental Change, 38*, 1–7.

Lu, X., Wrathall, D. J., Sundsøy, P. R., Nadiruzzaman, M., Wetter, E., Iqbal, A., et al. (2016b). Detecting climate adaptation with mobile network data in Bangladesh: anomalies in communication, mobility and consumption patterns during cyclone Mahasen. *Climatic Change, 138*(3–4), 505–519.

Mathieu, P.-P., & Aubrecht, C. (Eds.) (2018). *Earth observation open science and innovation*. ISSI scientific report series, vol. 15. SpringerOpen.

Martin, D., Cockings, S., & Leung, S. (2015). Developing a flexible framework for spatiotemporal population modeling. *Annals of the Association of American Geographers, 105*(4), 754–772.

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing, 39*(9), 2784–2817.

Maxwell, S. K., Meliker, J. R., & Goovaerts, P. (2010). Use of land surface remotely sensed satellite and airborne data for environmental exposure assessment in cancer research. *Journal of Exposure Science & Environmental Epidemiology, 20*(2), 176.

McPherson, T. N., Ivey, A., Brown, M. J., & Streit, G. E. (2004). Determination of the spatial and temporal distribution of population for air toxics exposure assessments. American Meteorological Society fifth conference on urban environment, American Meteorological Society. Retrieved from https://ams.confex.com/ams/pdfpapers/80347.pdf. Accessed Nov 2018

McSweeney, K., Nielsen, E. A., Taylor, M. J., Wrathall, D. J., Pearson, Z., Wang, O., & Plumb, S. T. (2014). Drug policy as conservation policy: narco-deforestation. *Science, 343*(6170), 489–490.

McSweeney, K., Richani, N., Pearson, Z., Devine, J., & Wrathall, D. J. (2017). Why do narcos invest in rural land? *Journal of Latin American Geography, 16*(2), 3–29.

Mennis, J., & Hultgren, T. (2006). Intelligent dasymetric mapping and its application to areal interpolation. *Cartography and Geographic Information Science, 33*(3), 179–194.

Minnesota Population Center. (2018). Integrated public use microdata series, international: Version 7.1 [dataset]. Minneapolis, MN: IPUMS. doi: https://doi.org/10.18128/D020.V7.0

Mikal, J., Hurst, S., & Conway, M. (2016). Ethical issues in using Twitter for population-level depression monitoring: a qualitative study. *BMC Medical Ethics, 17*(1), 22. https://doi.org/10.1186/s12910-016-0105-5.

Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., & Vespignani, A. (2013). The Twitter of Babel: mapping world languages through microblogging platforms. *PLoS One, 8*(4). https://doi.org/10.1371/journal.pone.0061981.

Monfreda, C., Ramankutty, N., & Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochemical Cycles, *22*(1). https://doi.org/10.1029/2007GB002947.

Olmanson, L. G., Bauer, M. E., & Brezonik, P. L. (2008). A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sensing of Environment, 112*(11), 4086–4097.

Panteras, G., & Cervone, G. (2018). Enhancing the temporal resolution of satellite-based flood extent generation using crowdsourced data for disaster monitoring. *International Journal of Remote Sensing, 39*(5), 1459–1474.

Patel, N. N., Stevens, F. R., Huang, Z., Gaughan, A. E., Elyazar, I., & Tatem, A. J. (2017). Improving large area population mapping using geotweet densities. *Transactions in GIS, 21*(2), 317–331.

Pekel, J. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature, 540*(7633), 418–422.

Pengra, B., Long, J., Dahal, D., Stehman, S. V., & Loveland, T. R. (2015). A global reference database from very high resolution commercial satellite data and methodology for application to Landsat derived 30m continuous field tree cover data. *Remote Sensing of Environment, 165*, 234–248.

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., ... & Marin-Herrera, M. A. (2013). A global human settlement layer from optical HR/VHR RS data: concept and first results. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6(5), 2102–2131.

Pinkovskiy, M., & Sala-i-Martin, X. (2016). Lights camera income. Illuminating the national accounts-household surveys debate. *Quarterly Journal of Economics, 131*(2), 579–631.

Ramankutty, N., Evan, A. T., Monfreda, C., & Foley, J. A. (2008). Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Global Biogeochemical Cycles, 22(1). https://doi.org/10.1029/2007GB002952.

Rey, S.J. (2017). Python for GIS. In J.P. Wilson. (Ed.), *The geographic information science & technology body of knowledge* (3rd Quarter 2017 Edition). University Consortium for Geographic Information Science. doi: https://doi.org/10.22224/gistbok/2017.3.4

Rindfuss, R. R., & Stern, P. C. (1998). Linking remote sensing and social science: the need and the challenges. In D. Liverman, E. F. Moran, R. R. Rindfuss, & P. C. Stern (Eds.), *People and pixels: linking remote sensing and social science* (pp. 1–27). Washington, DC: National Academies Press.

Röder, A., Hill, J., Duguy, B., Alloza, J. A., & Vallejo, R. (2008). Using long time series of Landsat data to monitor fire events and post-fire dynamics and identify driving factors. A case study in the Ayora region (eastern Spain). *Remote Sensing of Environment, 112*(1), 259–273.

Ross, M. L. (2004). What do we know about natural resources and civil war? *Journal of Peace Research, 41*(3), 337–356.

Ruggles, S., Kugler, T. A., Fitch, C. A., & Van Riper, D. C. (2015, November). Terra Populus: integrated data on population and environment. *2015 IEEE International Conference on Data Mining Workshop (ICDMW).* doi: https://doi.org/10.1109/ICDMW.2015.204.

Sanderson, E. W., Jaiteh, M., Levy, M. A., Redford, K. H., Wannebo, A. V., & Woolmer, G. (2002). The human footprint and the last of the wild: the human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. *BioScience, 52*(10), 891–904.

Sandborn, A., & Engstrom, R. N. (2016). Determining the relationship between census data and spatial features derived from high-resolution imagery in Accra, Ghana. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*(5), 1970–1977.

Sava, E., Clemente-Harding, L., & Cervone, G. (2017). Supervised classification of civil air patrol (CAP). *Natural Hazards, 86*(2), 535–556.

Schabenberger, O. & Gotway, C.A. (2017). Statistical methods for spatial data analysis. Chapman and Hall/CRC.

Schade, S., Diaz, L., Ostermann, F., Spinsanti, L., Luraschi, G., Cox, S., et al. (2013). Citizen-based sensing of crisis events: sensor web enablement for volunteered geographic information. *Applied Geomatics, 5*(1), 3–18.

Schroeder, J. P. (2017). Hybrid areal interpolation of census counts from 2000 blocks to 2010 geographies. *Computers, Environment and Urban Systems, 62*, 53–63. https://doi.org/10.1016/j.compenvurbsys.2016.10.001.

Sesnie, S. E., Tellman, B., Wrathall, D., McSweeney, K., Nielsen, E., Benessaiah, K., et al. (2017). A spatio-temporal analysis of forest loss related to cocaine trafficking in Central America. *Environmental Research Letters, 12*(5), 054015.

Serajuddin, U., Uematsu, H., Wieser, C., Yoshida, N., & Dabalen, A. (2015). Data deprivation: another deprivation to end. Policy Research Working Papers, The World Bank doi: https://doi.org/10.1596/1813-9450-7252.

Sever, T. L. (1998). Validating prehistoric and current social phenomena upon the landscape of the Petén, Guatemala. In D. Liverman, E. F. Moran, R. R. Rindfuss, & P. C. Stern (Eds.), *People and pixels: linking remote sensing and social science* (pp. 145–163). Washington: National Academies Press.

Sever, T. L., & Irwin, D. E. (2003). Landscape archaeology: remote-sensing investigation of the ancient Maya in the Petén rainforest of northern Guatemala. *Ancient Mesoamerica, 14*(1), 113–122.

Sexton, J. O., Song, X. P., Huang, C., Channan, S., Baker, M. E., & Townshend, J. R. (2013). Urban growth of the Washington, DC-Baltimore, MD metropolitan region from 1984 to 2010 by annual, Landsat-based estimates of impervious cover. *Remote Sensing of Environment, 129*, 42–53.

Shively, G., Sununtnasuk, C., & Brown, M. (2015). Environmental variability and child growth in Nepal. *Health & Place, 35*, 37–51. https://doi.org/10.1016/j.healthplace.2015.06.008.

Sorichetta, A., Hornby, G. M., Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. *Scientific Data, 2*, 150045.

Stefani, A., Dusfour, I., Corrêa, A.P.S.A., Cruz, M.C.B., Dessay, N., Galardo, A.K.R., … & Roux, E. (2013). Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data. Malaria Journal, 12(1), 192.

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS One, 10*(2), e0107042.

Taubenböck, H., & Kraff, N. J. (2014). The physical face of slums: a structural comparison of slums in Mumbai, India, based on remotely sensed data. *Journal of Housing and the Built Environment, 29*(1), 15–38.

Tidball, K. G., & Krasny, M. E. (2012). A role for citizen science in disaster and conflict recovery and resilience. In J. L. Dickinson & R. Bonney (Eds.), *Citizen science: public participation in environmental research* (pp. 226–234). Ithaca: Comstock Publishing Associates.

Tobler, W. R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association, 74*(367), 519–530.

Tong, D., & Murray, A. T. (2012). Spatial optimization in geography. *Annals of the Association of American Geographers, 102*(6), 1290–1309.

University of Maryland. (2013). First detailed map of global forest change. ScienceDaily. Retrieved December 4, 2018 from www.sciencedaily.com/releases/2013/11/131114142123.htm. Accessed 4 Dec 2018

van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., et al. (2016). Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites. *Environmental Science & Technology, 50*(7), 3762–3772.

van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., … & Winker, D.M. (2018). Global annual PM2.5 grids from MODIS, MISR and SeaWiFS Aerosol Optical Depth (AOD) with GWR, 1998-2016. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC) https://doi.org/10.7927/H4ZK5DQS.

VanWey, L. K., Rindfuss, R. R., Gutmann, M. P., Entwisle, B., & Balk, D. L. (2005). Confidentiality and spatially explicit data: concerns and challenges. *Proceedings of the National Academy of Sciences of the United States of America, 102*(43), 15337–15342. https://doi.org/10.1073/pnas.0507804102.

van Zanten, B. T., Van Berkel, D. B., Meentemeyer, R. K., Smith, J. W., Tieskens, K. F., & Verburg, P. H. (2016). Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences, 113*(46), 12974–12979.

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., et al. (2016). Global terrestrial human footprint maps for 1993 and 2009. *Scientific Data, 3*, 160067.

Wiig, F., Harrower, M., Braun, A., Nathan, S., Lehner, J., Simon, K., et al. (2018). Mapping a subsurface water channel with X-band and C-band synthetic aperture radar at the Iron Age archaeological site of 'Uqdat al-Bakrah (Safah), Oman. *Geosciences, 8*(9), 334.

Williams, J. N. (2013). Humans and biodiversity: population and demographic trends in the hotspots. *Population and Environment, 34*(4), 510–523.

Witmer, F. D. (2015). Remote sensing of violent conflict: eyes from above. *International Journal of Remote Sensing, 36*(9), 2326–2352.

Wood, J., Slingsby, A., & Dykes, J. (2011). Visualizing the dynamics of London's bicycle-hire scheme. *Cartographica: The International Journal for Geographic Information and Geovisualization, 46*(4), 239–251. https://doi.org/10.3138/carto.46.4.239.

World Bank. (2016). *World development report 2016: digital dividends*. Washington: World Bank. https://doi.org/10.1596/978-1-4648-0671-1.

Zheng, Z., Li, Y., Guo, Y., Xu, Y., Liu, G., & Du, C. (2015). Landsat-based long-term monitoring of total suspended matter concentration pattern change in the wet season for Dongting Lake, China. *Remote Sensing, 7*(10), 13975–13999.

Affiliations

Tracy A. Kugler[1] · Kathryn Grace[2] · David J. Wrathall[3] · Alex de Sherbinin[4] ·
David Van Riper[1] · Christoph Aubrecht[5] · Douglas Comer[6] · Susana B. Adamo[4] ·
Guido Cervone[7] · Ryan Engstrom[8] · Carolynne Hultquist[7] · Andrea E. Gaughan[9] ·
Catherine Linard[10] · Emilio Moran[11] · Forrest Stevens[9] · Andrew J. Tatem[12] ·
Beth Tellman[13] · Jamon Van Den Hoek[3]

1   Institute for Social Research and Data Innovation, University of Minnesota, 50 Willey Hall, 225 19th Ave S., Minneapolis, MN 55455, USA

2   Department of Geography, Environment and Society, University of Minnesota, Minneapolis, USA

3   Department of Geography, Environmental Sciences, and Marine Resource Management, Oregon State University, Corvallis, USA

4   Center for International Earth Science Information Network (CIESIN), Columbia University, New York, NY, USA

5   European Space Agency (ESA) and World Bank, Washington, DC, USA

6   Cultural Site Research and Management, Baltimore, MD, USA

7   Geoinformatics and Earth Observation Laboratory, Department of Geography and Institute for CyberScience The Pennsylvania State University, University Park, PA, USA

8   Department of Geography, George Washington University, Washington, DC, USA

9   Department of Geography and Geosciences, University of Louisville, Louisville, KY, USA

10   Department of Geography, University of Namur, Namur, Belgium

11   Center for Systems Integration and Sustainability, Michigan State University, East Lansing, USA

12   WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton, UK

13   The Earth Institute, Columbia University, New York, NY, USA