General Examinations - Contextual Area - Question 5 Response

Jack Reid

August 26, 2021

5 - In many parts of the world data simply doesn't exist. Many countries have not taken basic census data for decades, which makes it extremely hard to make decisions, the World Bank calls it another deprivation to end. Explain the primary pathways for data collection in these environments. How can these forms of data collection and analysis result in inequitable outcomes or be used for outright oppression? What are ways in which data analysts can address the need to collect data to make decisions while also ensuring they protect the public from potential harms?

1 The Problem of Unequal Data

The lack of data, particularly geospatial data, as long plagued centralized decision-makers, for whom "legibility [is] a central problem," one that must be solved prior to the creation of policy [1]. Lack of data, and the resulting poor maps and understanding of regions, has resulted in misdrawn borders (e.g. the Northwest Angle of Minnesota [2]) and even brought nations to the edge of war [3]. Even in our current era, the availability of data is painfully uneven across the various regions of the world. As Scott argued at length, the collection of data is closely related to the power of the state [1]. Taylor and Johnston pointed out that "*statistics* and *state* come from the same root" [4]. In general, the best data, particularly the best longitudinal data, tends to come from states that are wealthy, stable, have centralized authority, and have little incentive (either internal or external) to fabricate. The converse of this is what Taylor and Overton call "the first law of geographical information: the poorer the country, the less and the worse the data" ([5] as paraphrased by [4]).

The 'first law' obviously applies to basic demographic and geographical data, such as that collected by standard decennial censuses or by national mapping services, but it also applies to data derived by analysis [6], including global and multi-regional datasets. This is only exacerbated by the fact that such derived datasets are typically created by individuals and organizations based in the wealthier states and thus subject to their particular interests and language limitations. An example of this can be seen in the primary composite dataset of ecosystem services valuations. The Ecosystem Services Valuation Database (ESVD) is maintained by research organizations based in continental Europe and is primarily funded by a UK government agency. The database organizes the studies that it references by the location of the specific ecosystem services being valued. Table 1 shows the breakdown of the target continents of these studies, in which a bias can be seen both towards wealthier regions and towards those regions of more interest to the researchers and funding sources.

Table 1: Regions studied by publications compiled by ESVD

Region	**Number of Studies**	**Percent of studies**
Africa	309	7.7
Asia	1140	28.4
Europe	1639	40.8
North America	594	14.8
Oceania	223	5.6
South America	109	2.7

This trend tends to be exacerbated in smaller sub-disciplines. Approximately 63% of studies of mangrove-related ecosystem services are focused on parts of Asia despite these regions constituting providing only 38% of the world's mangrove coverage [7].

This uneven availability of data has severly hampered international, expert-led development efforts, including such major initiatives as the United Nations Millenium Development Goals (MDGs) which were a global coordinated effort spanning 2000 to 2015. One of the more prominent of MDGs was Goal 6, "Combat HIV/AIDS, malaria, and other diseases," with its associated Target 8: "Have halted by 2015, and begun to reverse the incidence of malaria and other major diseases" [8]. It is difficult to measure progress towards this goal when, in 2005, some researchers indicated that that episodes of malaria globally may have been as much as 50% higher than those reported by the World Health Organization (WHO) [9]. The availability and quality of data (commonly referred to as 'monitoring') was thus a common element of critiques of the MDGs [10, 11, 12]. While many of these critiques used this as an argument for different fundamental models of international development that were not as data-reliant (something that I will return to later in this piece), proponents of the goals instead viewed this as a specific challenge to be overcome. As the MDGs were concluded and their successors, the Sustainable Development Goals (SDGs), were created, the World Bank labeled the lack of a data a "deprivation" on par with the lack of food, shelter, or healthcare [13]. The United Nations General Assembly, in its commissioning of the SDGs, "called upon the United Nations system, including the Statistics Division of the Department of Economic and Social Affairs of the Secretariat and the regional commissions and international agencies, to support national efforts in building and strengthening national statistical capacity, in particular that of developing countries, and called upon all international agencies to improve the coverage, transparency and reporting on all indicators." For better or worse, efforts to improve data collection are here to stay for the foreseeable future. The next few sections will examine how data is collected in traditionally data sparse areas, what harms it can cause, and how such harms might be avoided.

2 Data Collection in Data Sparse Areas

The first and most obvious method of data collection in areas where the state is not already taking charge of such activities are targeted international development efforts as discussed in the previous section. These bring to bear international expertise, funding, labor, and technology to areas that may otherwise lack such resources. They also apply political pressure to governments that may otherwise be reluctant to collect this data themselves. While such data collection efforts can be detailed, high quality, and consistent, they are typically targeted on a particular region and only last for the length of the particular program. For instance, detailed geospatial data on Liberia's demographics, road system and quality, bridges, schools, health facilities, and more were created by the United Nations via its ebola response program and its 2003-2018 peacekeeping mission [14]. Now that the ebola epidemic has largely subsided in the eyes of the world (outbreaks have continued, though with less international attention) and the peacekeeping mission has concluded, it is unlikely that this data will continue to be updated.

Another method of data collection is the targeted, international, academic study. These can take a variety of forms across innumerable fields of research. That said, some common types include:

1. A geniune interest, either personal or professional, in a particular region, ecosystem, or community. This is commonly found among field anthropology and environmental sciences researchers, but examples can be found in other fields as well. To use a colleague as an example, the economist Suhyun Jung, who has specialized in ecosystem services and land management decisions, has increasingly (but no exclusively) focused on the landowners of the Brazil Amazon [15, 16, 17].
2. Case studies used to test or demonstrate a new methodology, perhaps with the intent to inform future, global applications of the same technique. Example: Modeling mangrove biomass in a particular Mozambique river delta [18] prior to modeling mangrove biomass in river deltas globally [19].
3. A single, one-off study in the area, often based on expertise developed elsewhere. To use Jung as an example again, he recently conducted a study in Liberia that was informed by his similar work on forest concessions in Brazil [20].

Despite the wide variety of such studies, however, certain commonalities exist. For one, data generated tends to be limited in terms of area, topic, and duration to the specific interests and duration of the study. Researchers do not typically have the time, resources, or interest to collect census-style data from an entire region. Additionally, even in the series of studies seen in Category 1, the types of data collected are often not consistent across the multiple studies in the same area. As a result, later longitudinal comparison is difficult or impossible. Similarly, inconsistent methodologies across academic studies, coupled with poor documentation practices, can make comparisons across studies or regions difficult. A review of the aforementioned ESVD found that significant numbers of the included studies lacked documentation regarding the sources of their data, used a methodology that prohibited comparison with other studies, or failed to clarify if the authors were using constant year dollars or not (and if so, what year) [21]. A third concern is how sites are selected. While a personal motive based on a genuine need for the selected area can play a role, many studies are driven by a combination of convenience and funding. For example, airborne image processing techniques are usually demonstrated using some close-to-home target: a paper by researchers based in Guangdong researchers studied mangroves in Guangdong [22]; a paper by researchers in Brisbane studied mangroves in Brisbane [23]. Jung does most of his work in the Brazil Amazon in part because there is significant government and NGO funding sources for that area. His foray into a Liberia study was the result of a commission by a Liberian-based NGO [20]. This means that academic studies are themselves largely subject to the same organizations (national governments and major NGOs) running the international development efforts discussed previously. The implications of this are discussed in the next section.

A third method of data collection is remote observation, particularly space-based remote observation. Satellites do not have to meet with the difficulties of political boundaries or impassable terrain. Due to the nature of orbital mechanics, it is actually more difficult to create an earth observation (EO) system that *does not* image most of the world than one that does. This fact, coupled with interest of the earth science community in the global-scale dynamics of our planet, mean that many EO datasets include most of the surface of the Earth. Remote observation thus has the potential to help upend the first law of geography by providing at least some base level of data globally, with no distinctions for borders or wealth. This is particularly visible in connection to the SDGs, with remote sensing boosters both within academia [24, 25, 26] (including an entire upcoming issue of the journal *Remote Sensing* [27]) and within government agencies, as seen in the reports by the Group of Earth Observations (GEO) [28] and the UN Committee on the Peaceful Uses of Outer Space (COPUOS) [29].

The use of remote observation data for such applications has a long history. While many of the initial efforts at remote observation from air and space were done with military objectives in mind, scientific, commercial, and social applications soon became apparent. An enormous amount of EO satellite data is freely available to the public through 20+ National Aeronautics and Space Administration (NASA) earth science satellites [30], the European Space Agency (ESA) Copernicus Programme (which includes both the six Sentinel satellites and in-situ measurements), the various satellites managed by the Japan Aerospace Exploration Agency (JAXA) Earth Observation Center (EOC), the China-Brazil Earth Resources Satellite Program (CBERS), and the satellites of other space agencies. While this data is largely free currently, this has not consistently been true, nor is it guaranteed to continue in the future [31]. For most of the early history of satellite observation, imagery was kept highly classified and zealously guarded [32, 33]. Even when the data was available to the public, it was not always freely available, as various countries have made attempts to monitize remote observation data. In the 1970s and early 1980s, for instance, Landsat data was a government-managed operation that provided products at a low-cost, based primarily on the cost of reproduction. In the 1980s, however, the program was transfered to a private entity and prices were increased by more than an order of magnitude and significant copyright restrictions were put in place [34]. Currently the data is once again freely available after the monitization efforts met with limited success [35], but this may not remain the case moving forward [36]. In general we see the consistent pattern repeating itself: data is generated about places but is not necessarily available or accesssible to people in those places. Speaking anecdotally, it is not uncommon for international colleagues to request help accessing data from US government satellites that is nominally already freely available to anyone anywhere.

The use patterns of remote observation data has varied for reasons beyond cost and military secrecy, however. Social applications were being considered from quite early on. By the early 1970s five rationales for using satellite imagery in city planing had become widespread [37]:

1. It offers a synoptic, total view of the complex system in a given area.
2. Satellites provide repetitive, longitudinal coverage.
3. Satellite inventories were more efficient and up-to-date than ground surveys.
4. Remote sensing was objective.
5. Satellites produced digital imagery that could be easiliy combined with ground-based data in novel geographic information systems (GISs).

Despite these rationales, actual social applications of such data remained rare for several decades. The reasons for this are many, but probably include that many of these rationeles were overstated for their day. Insufficient resolution and inconsistent coverage limited use in local areas. While satellite imagery provides a wonderful decades-long longitudinal dataset now, it did not at the time. Satellite imagery was still heavily dependent on human photointerpretation, undermining the argument that the data was "objective" in any meaningful sense. Finally the cost and specialization required to effectively use the data limited its ability to be combined with other datasets. Internationally, these difficulties were compounded by the fact that that the satellite systems were designed to serve US government and scientific aims first and foremost. Satellites were thus not designed with the priorities of other nations in mind, including agriculture, forest management, and disease monitoring.

Finally, due partially to limitations in what satellites can observe and partially to the priorities of the designers, applications of EO data, particularly that which is not straightforward visual imagery, remain squarely focused on characterizing specific, usually environmental, phenomena, such as wildfires [38], aquatic bacterial growths [39], or deforestation [19], with only limited excursions into studying human wellbeing or the the connections between environmental phenomena and human wellbeing (such as using nightlight imagery as a proxy for economic development [40]).

One final method of collecting data in traditionally data sparse regions is the internet and telecommunications. The penetration of connectivity in such regions has advanced enormously over the past few decades [41] and increasingly this can be used to generate useful data. Telecommunications mobility data was used extensively during the 2013-2016 ebola epidemic in Western Africa [42] and is currently being used in connection with the COVID-19 pandemic [43, 44]. This data is predominantly in the hands of private corporations and thus tends to focus on different domains than the academic and government data collection methods discussed above.

3 The Harms of Data Collection

While the methods of data collection in the previous section vary in numerous important ways, it is possible to identify certain common ways in which these methods may result in inequitable outcomes or be used for outright oppression. The first is that all data collection is purposeful. As Scott wrote, "state simplications... have the character of maps. That is, they are designed to summarize precisely those aspects of a complex world that are of immediate interest to the mapmaker and to ignore the rest. To complain that a map lacks nuance and detail makes no sense unless it omits information necessary to its function" [1]. These purposes can be (and often are) ones of control or oppression by a state of its own populace. For a further discussion of this, see my response to Question 1 of this exam. Relevant to us here is also the concern about external purposes, namely those of other states, of foreign academics, of corporations. All of these groups have ends of their own and pursue data collection efforts that satisfy those ends, not necessarily the ends of the people being studied. The data collectors often ignore or conceal this fact. As Taylor and Johnston put it, "Data are usually treated unproblematically except for technical concerns about error. But... every data set represents a myriad of social relations... There is an implicit power relation; in general, the more powerful do the finding out about the less powerful " [4].

Some critiques of the MDGs pointed to the disagreements about the purpose of the quantiative indicators, specifically that they pulled attention and resources away from issues of higher priority to locals [11, 12], something that is a recurrent problem with simplifying metrics in development. "Many studies involve ranking places on one or more criteria, and allocating policy benefits accordingly. At its crudest this applied geography merely provides a list of winner and losers with no understanding of why the differences occur" [4]. This is arguably repeating in our current era of the SDGs. Campbell wrote that "The pessimistic thought is that sustainability has been stripped of its transformative power and reduced to its lowest common demoninator. After all, if both the World Bank and radical ecologists now believe in sustainability, the concept can have no teeth: it is so malleable as to mean many things to many people without requiring commitment to any specific policies" [45]. While this is arguably true, another interpretation is that the World Bank may have a *different* definition of sustainability than others and that definition may not fit the needs of different peoples. Furthermore, some developing nations believe that the contemporary interest of wealthy Western nations in global environmental protections comes at the cost of the economic development of those developing nations [46], pointing towards a difference in priorities between those doing the studying and those being studied.

This difference in priorities can, in addition to diverting resources elsewhere, cause active harms through collateral damage. After all "Haussmann's Paris was, *for those who are not expelled*, a far healthier city" (emphasis mine) [1]. Similarly, the use of mobile telecommunication device location data during the 2013-2016 ebola epidemic was done with little regard for privacy concerns, no real protections of individual data, and without even any significant relevance for addressing the epidemic [42].

These various data collection efforts are also often quite incomplete, as discussed in the previous section. In some situations, such incomplete data can be as dangerous (or more dangerous) than no data at all. Gaps in the data are not random or uniform. As just discussed, data collective is purposive and thus can lead to misleading biases. Data collection can also be purposely impeded or fabricated for either internal or external reasons. The Chinese government famously failed to track or make available data regarding their so-called Ghost Cities. This hampered the informed decision-making of local planners and community members, leading them to turn to alternative ways of generating this data developed by academic researchers [47]. Similarly, refusing entry to international journalists, academics, and inspectors is an infamous and all too common tactic of despotic governments seeking to cover up human rights abuses. Other gaps in data are more coincidental but still disproportionately impact certain areas. For example, the tropics experience more frequent and denser cloud cover than other parts of the globe, thereby resulting in less frequent remote sensing images of these regions (except for imagery that can penetrate clouds).

Finally, data collection and generation may be done with 'good intentions' but still suffer from cultural biases or ignorance. As mentioned in the previous section, data collection and analysis methods that are intended to be applied globally are often first trialed and calibrated in specific case studies which are conveniently available (i.e. nearby) to the researchers. This can result in methods that work well in one cultural context going on to generate misleading results in other contexts. This is particularly worrisome in machine learning methodologies, due to their black box nature and their common use for processing remote observation data. The latter is due to a combination of the sheer amount of data generated by EO systems and the difficulty for humans to interprete such data, particularly data that is not visual red-green-blue. The problems of machine learning and big data with regards to equity and bias are well documented and numerous, including such issues as bias in the training data, application of models outside of their domain of training, hiding existing discrimintary practices behind of veil of mathematics, a lack of transparency, privacy and surveillance capital concerns, and development and use by authoritarian actors. For a full treatment of these issues, see *Weapons of Math Destruction* [48], *Algorithms of Oppression* [49], and *Automating Inequality* [50]. Here I will confine myself to a single illustrative example.

In a paper published in the journal *Remote Sensing* in 2019, several researchers used the Faster Region Based Convolutional Neural Network (Faster R-CNN) machine learning algorithm on imagery from Google Earth to identify brick kilns with the goal of developing maps of these kilns to inform on-the-ground inspections. This is because, in a large region of Asia, such kilns are often the workplaces of illegal slave labor. In this way, they would be helping to address SDG 8.7 [51]. Several issues are noticeable in the paper. First, the authors frame their work as a method for identifying brick kilns across the "Brick Belt," an $\approx$ 1,550,000 km2 region including parts of India, Pakistan, Nepal, and Bangladesh [52]. Nonetheless, for their training and validation data, the authors drew entirely from a single 120 km2 region of northwestern India. This runs the risk of both false positives and false negatives if the algorithm is applied outside of this region, since both kiln construction and the surrounding terrain may vary in appearance across the Brick Belt. This is particularly concerning because the authors themselves acknowledge that the region they chose "is known to have a higher than average density of brick kilns," which indicates that this region is not, in fact, representative of the broader Brick Belt. This is compounded by the fact that, in order to avoid false negatives, the authors calibrated the system have a high false positive rate (it identified approximately twice the number of kilns as were actually present). While the authors say that this is an acceptable false positive rate for their application, when the algorithm is applied to other regions with lower densities of brick kilns, the false positive rate is likely to be even higher.

Second, for validation, the authors state that they used visual inspection of higher resolution imagery from the WorldView-2 satellite system. The authors do not provide any grounded explanation of how their expertise in visually identifying brick kilns from satellite imagery justifies not using in-situ verified data. In particular, they state that the "region used for accuracy assessment contained 178 brick kilns, all identified by visual classification." But this assumes that the visual inspector was actually able to identify all the brick kilns in the region and did not misidentify any.

None of this is to cast aspersions on the goals of the authors, to claim that the issues pointed to are irredeemable, or even to guarantee that the possible negative outcomes would actually manifest. Instead I wanted to point out some of the issues of bias that can creep into even a technically competent study focusing on a particular region. Not all such studies are technically competent, cognizant of potential negative impacts, and limited to a particular region for both development and application.

Now that some of the potential harms in data collection and use have been identified, the next question is, how can we avoid them?

4 How to Avoid Harms

First and most obviously, we can be careful about our collection and use of data so as to avoid technical mistakes. This includes such actions as characterizing gaps in the data rather than assuming them to be uniformly random, examining the generation process so as to identify potential errors, and not applying a data-based model out of its domain of calibration. In machine learning of remote observation data, for example, training data should typically be based on in-situ observations that are selected to be representative of the entire application area. These correctives are all important and should certainly be implemented, but they are insufficient on their own to avoid all the harms laid out in the previous section.

Second, we may refuse to do data collection and analysis in areas with authoritarian governments or other unsavory decision-makers, though this would certainly neglect many in dire need. This would thereby avoid one of Scott's conditions for technocratic social engineering disasters (the presence of an authoritarian government) [1]. We may also reject the high modernist ideology in our planning activities (another of Scott's conditions). When it comes to data collection and use, this can be done by being critical of the providence, applicability, and original purposes of datasets, and by being willing to take action to fill gaps in the data rather than just relying upon what is available. That said, this is not a trivial undertaking. In addition to the extra work required, in many ways such the high modernist ideology is the default of the technologist, and active self-reflection is required to avoid falling into that trap. Furthermore, while you may have avoided working with despots and are not ideologically blinded yourself, data, once collected, has some degree of permanence and it is not always clear who will use it in the future.

And the unfortunate matter is, even if we assume that Scott is correct in that his conditions are the necessary and sufficient conditions, what are they conditions for? "The *most tragic* episodes of state-initiated engineering" (emphasis mine) [1]. The egregiously racist influence that Robert Moses had the design of New York City [53] happened in an at least somewhat democratic society, not an authoritarian one. While it did not directly lead to mass famine and death, it is hardly something that we would want to replicate. I daresay that we want to do more than avoid the most tragic outcomes and instead want to do active good. We must therefore look beyond merely avoiding Scott's conditions.

Third, we may argue that data collection and analysis methods have developed over time, are now more objective, and are thus no longer vulnerable to historical biases and gaps. This is essentially what proponents of remote observation data advocated as far back as the 1970s, as discussed earlier. While improvements are real and remote observation certainly represents a way of checking claims made by deceptive actors, the previous section made clear that not only are contemporary methodologies still vulnerable to intentional exploitation and unintentional misuse, but these issues are inherent in the act of data collection itself, regardless of the methodology.

Fourth, we may change our framework of development. Since data is collected for a purpose and mediated by technology, if we change the purpose, we change both the use and the collection. This is the approach many critics of metrics such as the targets and indicators of the MDGs and SDGs have taken: "The solution cannot therefore be to seek fully to overcome the limitations in our knowledge (which are incapable of being fully overcome), but rather lies in adopting structures for decision-making which address these limitations" [12]. Such altered frameworks include Bayesian cost estimates [12], a focus on human rights [10], qualitative rather than quantiative objectives [11], a focus on freedom of choice [54], and Easterly's "Searchers" (those who seek for bottom-up solution to specific, addressable needs in local areas and thus do not need immense amounts of standardized data collection) [55]. One of the more popular thrusts in this vein are participative and collaborative development activities. The benefits and limitations of such frameworks, in the context of GIS data, are discussed further in my response to Question 1 of this exam.

Ultimately there does not seem to be a single, clean answer to how to generate and analyze data in a safe way. Rather it is through a combination of these (avoiding work with dictators, being careful with the technical details, remaining humble, and working with a non/less exploitative framework) that we can maximize the benefits of data while minimizing the harms.

References

[1] J. C. Scott, *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed.* Yale University Press, Mar. 2020.

[2] Boggs S.wihittemore, *International Boundaries A Study Of Boundary Functions And Problems.* 1940.

[3] E. C. Coleman, *The Pig War: The Most Perfect War in History.* Stroud: The History Press, May 2010.

[4] P. Taylor and R. Johnston, "Geographic Information Systems and Geography," in *Ground Truth: The Social Implications of Geographic Information Systems* (J. Pickles, ed.), pp. 51–67, New York: The Guilford Press, 1st edition ed., Dec. 1994.

[5] M. Overton and P. Taylor, "Further thoughts on geography and GIS: A preemptive strike?," *Environment and Planning A*, vol. 23, pp. 1087–1094, 1991.

[6] P. J. Taylor, "Full circle, or new meaning for the global," *The challenge for geography: a changing world: a changing discipline Blackwell, Oxford*, pp. 181–97, 1993.

[7] T. Vegh, M. Jungwiwattanaporn, L. Pendleton, and B. Murray, "Mangrove Ecosystem Services Valuation: State of the Literature," Tech. Rep. NI WP 14-06, Nicholas Institute for Environmental Policy Solutions, Duke University, Durham, NC, July 2014.

[8] Inter-Agency and Expert Group on MDG Indicators, "The Millennium Development Goals Report," tech. rep., United Nations, 2015.

[9] R. W. Snow, C. A. Guerra, A. M. Noor, H. Y. Myint, and S. I. Hay, "The global distribution of clinical episodes of Plasmodium falciparum malaria," *Nature*, vol. 434, pp. 214–217, Mar. 2005.

[10] P. Alston, "Ships Passing in the Night: The Current State of the Human Rights and Development Debate Seen through the Lens of the Millennium Development Goals," *Human Rights Quarterly*, vol. 27, no. 3, pp. 755–829, 2005.

[11] S. Fukuda-Parr, A. E. Yamin, and J. Greenstein, "The Power of Numbers: A Critical Review of Millennium Development Goal Targets for Human Development and Human Rights," *Journal of Human Development and Capabilities*, vol. 15, pp. 105–117, July 2014.

[12] S. Reddy and A. Heuty, "Global Development Goals: The Folly of Technocratic Pretensions," *Development Policy Review*, vol. 26, no. 1, pp. 5–28, 2008.

[13] "Data deprivation : Another deprivation to end." https://documents.worldbank.org/en/publication/documents-reports/documentdetail/700611468172787967/Data-deprivation-another-deprivation-to-end.

[14] UNMEER, "UNMEER (Liberia) Geo-Information Management Services (GIMS)." https://www.arcgis.com/apps/webappviewer/index.html?id=40514fa5569e4892a89e009b9f582c12.

[15] S. Jung, L. V. Rasmussen, C. Watkins, P. Newton, and A. Agrawal, "Brazil's National Environmental Registry of Rural Properties: Implications for Livelihoods," *Ecological Economics*, vol. 136, pp. 53–61, June 2017.

[16] S. Jung and S. Polasky, "Partnerships to prevent deforestation in the Amazon," *Journal of Environmental Economics and Management*, vol. 92, pp. 498–516, Nov. 2018.

[17] L. V. Rasmussen, S. Jung, A. D. Brites, C. Watkins, and A. Agrawal, "Understanding smallholders' intended deforestation behavior in the Brazilian Cerrado following environmental registry," *Environmental Research Letters*, vol. 12, p. 094001, Sept. 2016.

[18] T. Fatoyinbo, E. A. Feliciano, D. Lagomasino, S. K. Lee, and C. Trettin, "Estimating mangrove aboveground biomass from airborne LiDAR data: A case study from the Zambezi River delta," *Environmental Research Letters*, vol. 13, p. 025012, Feb. 2018.

[19] D. Lagomasino, T. Fatoyinbo, S. Lee, E. Feliciano, C. Trettin, A. Shapiro, and M. M. Mangora, "Measuring mangrove carbon loss and gain in deltas," *Environmental Research Letters*, vol. 14, p. 025002, Jan. 2019.

[20] S. Jung, C. Liao, A. Agrawal, and D. G. Brown, "Evidence on Wealth-Improving Effects of Forest Concessions in Liberia," *Journal of the Association of Environmental and Resource Economists*, vol. 6, pp. 961–1000, Sept. 2019.

[21] S. Jung, E. Joiner, J. Reid, and D. Wood, "Gaps in Mangrove Forest Data and Valuation Methods Limit Understanding of Socioeconomic Benefits," *Review of Environmental Economics and Policy*, Publication Pending.

[22] J. Cao, W. Leng, K. Liu, L. Liu, Z. He, and Y. Zhu, "Object-Based Mangrove Species Classification Using Unmanned Aerial Vehicle Hyperspectral Images and Digital Surface Models," *Remote Sensing*, vol. 10, p. 89, Jan. 2018.

[23] M. Kamal and S. Phinn, "Hyperspectral Data for Mangrove Species Mapping: A Comparison of Pixel-Based and Object-Based Approach," *Remote Sensing*, vol. 3, pp. 2222–2242, Oct. 2011.

[24] M. Burke, A. Driscoll, D. Lobell, and S. Ermon, "Using Satellite Imagery to Understand and Promote Sustainable Development," Working Paper 27879, National Bureau of Economic Research, Oct. 2020.

[25] B. Ferreira, M. Iten, and R. G. Silva, "Monitoring sustainable development by means of earth observation data and machine learning: A review," *Environmental Sciences Europe*, vol. 32, p. 120, Sept. 2020.

[26] M. Tassopoulou, N. Verde, G. Mallinis, H. Georgiadis, D. Kaimaris, and P. Patias, "Demonstrating the potential of remote sensing to support sustainable development goals implementation: Case studies over Greece," in *Seventh International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2019)*, vol. 11174, p. 111740D, International Society for Optics and Photonics, June 2019.

[27] *Special Issue "Earth Observations for Sustainable Development Goals"*. Remote Sensing, Basel, Switzerland: Multidisciplinary Digital Publishing Institute, Apr. 2022.

[28] Group on Earth Observations, "Earth Observations in support of the 2030 Agenda for Sustainable Development," Mar. 2017.

[29] COPUOS, "Contribution of the Committee on the Peaceful Uses of Outer Space to the United Nations Conference on Sustainable Development: Harnessing space-derived geospatial data for sustainable development," Tech. Rep. A/AC.105/993, United Nations, Rio de Janeiro, Brazil, June 2011.

[30] G. Shirah, "NASA's Earth Observing Fleet: March 2017." https://svs.gsfc.nasa.gov/4558, 2017.

[31] M. Borowitz, *Open Space : The Global Effort for Open Access to Environmental Satellite Data*. Cambridge, MA: MIT Press, 2017.

[32] M. Healy, "Rep. Brown Quits House Intelligence Panel, Claims Abuses of Secrecy," *Los Angeles Times*, Nov. 1987.

[33] J. Barry, "Mappings–a chronology of remote sensing," *Incorporations Zone*, vol. 6, pp. 570–1, 1992.

[34] P. McHaffie, "Manufacturing Metaphors," in *Ground Truth: The Social Implications of Geographic Information Systems* (J. Pickles, ed.), pp. 113–129, New York: The Guilford Press, 1st edition ed., Dec. 1994.

[35] M. M. Waldrop, "Landsat commercialization stumbles again.," *Science (New York, N.Y.)*, vol. 235, p. 155, Jan. 1987.

[36] G. Popkin, "US government considers charging for popular Earth-observing data," *Nature*, vol. 556, pp. 417–418, Apr. 2018.

[37] J. S. Light, *From Warfare to Welfare: Defense Intellectuals and Urban Problems in Cold War America*. Baltimore, Md.: JHUP, Aug. 2005.

[38] W. Schroeder, P. Oliva, L. Giglio, and I. A. Csiszar, "The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment," *Remote Sensing of Environment*, vol. 143, pp. 85–96, 2014.

[39] S. Stroming, M. Robertson, B. Mabee, Y. Kuwayama, and B. Schaeffer, "Quantifying the Human Health Benefits of Using Satellite Information to Detect Cyanobacterial Harmful Algal Blooms and Manage Recreational Advisories in U.S. Lakes," *GeoHealth*, vol. 4, no. 9, p. e2020GH000254, 2020.

[40] J. P. Hochard, S. Hamilton, and E. B. Barbier, "Mangroves shelter coastal economic activity from cyclones," *Proceedings of the National Academy of Sciences*, vol. 116, pp. 12232–12237, June 2019.

[41] H. K. Chavula, "Telecommunications development and economic growth in Africa," *Information Technology for Development*, vol. 19, pp. 5–23, Jan. 2013.

[42] S. McDonald, “Ebola: A Big Data Disaster,” tech. rep., The Centre for Internet & Society, Bengaluru, India, Jan. 2016.

[43] R. Doorley, A. Berke, A. Noyman, L. Alonso, J. Ribó, V. Arroyo, M. Pons, and K. Larson, “Mobility and COVID-19 in Andorra: Country-scale analysis of high-resolution mobility patterns and infection spread,” Feb. 2021.

[44] Google, “COVID-19 Community Mobility Report.” https://www.google.com/covid19/mobility?hl=en.

[45] S. Campbell, “Green Cities, Growing Cities, Just Cities? Urban Planning and the Contradictions of Sustainable Development,” in *Readings in Planning Theory* (S. Fainstein and J. DeFilippis, eds.), Hoboken, NJ: Wiley-Blackwell, fourth ed., Jan. 2016.

[46] N. Persadie and R. Ramlogan, “Pursuing Development and Protecting the Environment: Dilemma of the Developing World,” *Electronic Green Journal*, vol. 1, Dec. 2005.

[47] S. Williams, W. Xu, S. B. Tan, M. J. Foster, and C. Chen, “Ghost cities of China: Identifying urban vacancy through social media data,” *Cities*, vol. 94, pp. 275–285, Nov. 2019.

[48] C. O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* New York: Crown, 1st edition ed., Sept. 2016.

[49] S. U. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism.* New York: NYU Press, illustrated edition ed., Feb. 2018.

[50] V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* New York, NY: St. Martin’s Press, Jan. 2018.

[51] G. M. Foody, F. Ling, D. S. Boyd, X. Li, and J. Wardlaw, “Earth Observation and Machine Learning to Meet Sustainable Development Goal 8.7: Mapping Sites Associated with Slavery from Space,” *Remote Sensing*, vol. 11, p. 266, Jan. 2019.

[52] “How experts are using satellite images from Google Earth to tackle modern slavery,” Mar. 2018.

[53] L. Winner, “Do Artifacts Have Politics?,” *Daedalus*, vol. 109, no. 1, pp. 121–136, 1980.

[54] A. Sen, “Freedom of choice: Concept and content,” *European Economic Review*, vol. 32, pp. 269–294, Mar. 1988.

[55] W. Easterly, *The White Man’s Burden: Why the West’s Efforts to Aid the Rest Have Done So Much Ill and So Little Good.* New York City, NY: Penguin Books, 2007.