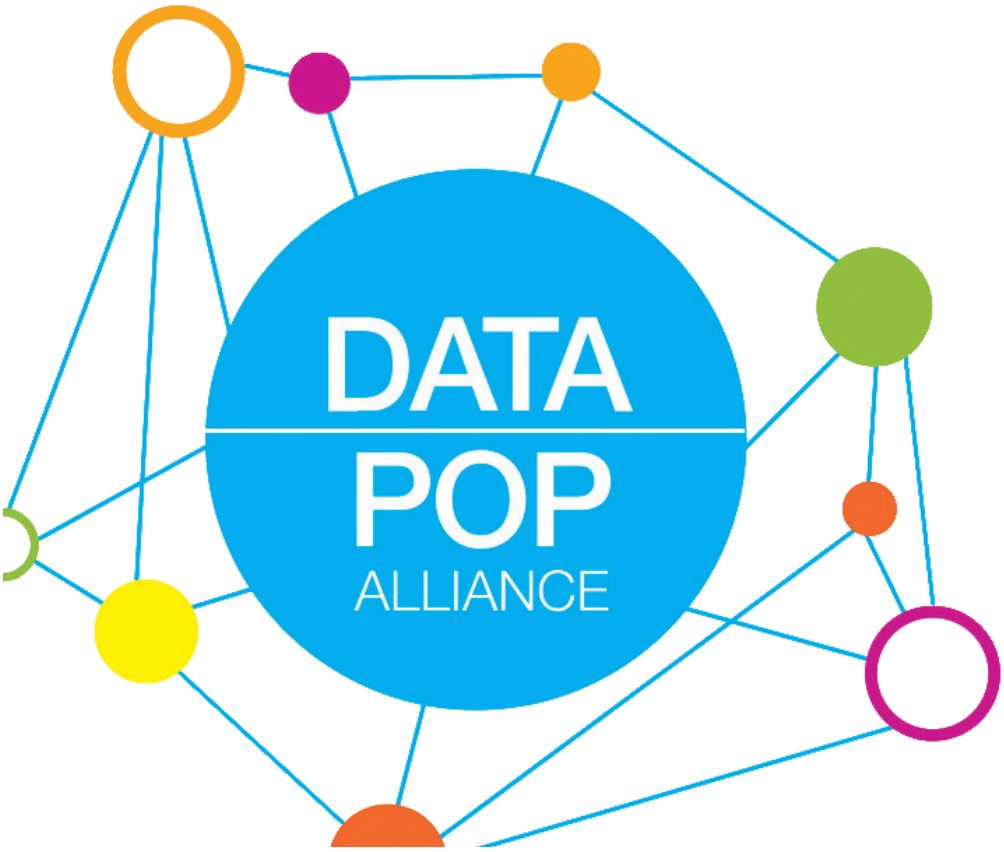
DATA-POP ALLIANCE

Inputs for World Development Report 2016 "Digital Dividends"

Big Data for Climate Resilience

October 2015

Input 2



DATA-POP ALLIANCE

Inputs for World Development Report 2016

Big Data for Climate Resilience

October 2015

Table of Contents

Figures

Boxes

Foreword

About this document

This document is one of three inputs for the World Development Report (WDR) 2016 "Digital Dividends" commissioned by the World Bank Group to Data-Pop Alliance.

Data-Pop Alliance is a coalition on Big Data and development jointly created by the Harvard Humanitarian Initiative (HHI), the MIT Media Lab, and the Overseas Development Institute (ODI) to promote a people-centered Big Data revolution.

About the co-authors

This paper was written by the following co-authors:

- **Emmanuel Letouzé** (co-lead and corresponding author), Director and co-Founder, Data-Pop Alliance; Visiting Scholar, MIT Media Lab. eletouze@datapopalliance.org
- **Patrick Vinck**, (co-lead author), Assistant Professor, Harvard University; Director, Harvard Humanitarian Initiative (HHI); Co-director, Co-founder, Data-Pop Alliance
- **Bessie Schwarz,** Research Affiliate, Data-Pop Alliance
- **Simone Sala,** Research Affiliate, Data-Pop Alliance
- **David Sangokoya**, (co-lead author), Research Specialist, Data-Pop Alliance
- **Elizabeth Tellman,** Research Affiliate, Data-Pop Alliance

Acknowledgements

This paper benefited from advice from Marc Levy (Columbia University) as well as inputs and editing support from Lauren Barrett (Communication Strategist, Data-Pop Alliance), Carson Martinez (Research Intern, Data-Pop Alliance) and Natalie Shoup (Program Manager, Data-Pop Alliance),

Funding

Funding for this paper was provided by the World Bank Group whose support is gratefully acknowledged, as well as the Rockefeller Foundation as part of their core support to Data-Pop Alliance's activities.

Disclaimer

The views presented in this paper are those of the authors and do not represent those of their institutions.

Suggested Citation

"Big Data for Climate Resilience." Inputs for the World Development Report 2016 "Digital Dividends"; Input 2. Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute). October 2015.

Introduction

As poor populations living in developing countries face frequent risks and shocks, such as droughts and floods, they are becoming increasingly vulnerable to the threat of global climate change. Humanitarian and development efforts have evolved in recognition that a long-term approach to climate and disaster risk reduction is needed.[1] Attention is given to what makes individuals, communities, and systems resilient to disasters–what makes them able to endure, adapt, and transform in response to uncertainty and adversity. However, building resilience in response to long-term climate changes is especially complex; strengthening these abilities will require innovative approaches. These approaches, in turn, must be rooted in accelerated learning and expanded understanding of critical resilience, vulnerability and hazard factors.[2]

Climate science has long relied on large amounts of fine-grained and comprehensive spatial data and computational power. But over the past decade, there has been significant growth of various kinds of high-frequency data and associated analytical capacities resulting from the sophistication and spread of digital technology lumped under the term 'Big Data.' Big Data constitutes a sea change for climate resilience: with the 'rise' of Big Data, the opportunities for building a better understanding of hazards, vulnerabilities, and resilience factors have vastly expanded.

Big Data as a new socio-technical phenomenon is not reducible to satellite imagery or other data sources; its many intertwined applications and implications are creating entirely new opportunities and challenges for our collective ability to analyze and potentially act on climate resilience features and factors, which has become a major concern of humanitarian and development actors and activities.[3]

Conceptually, following a previously developed framework,[4] we define Big Data as an ecosystem made up of digital "crumbs"[5]—exhaust data, web-based data and sensing data, "capacities"—including computational, methodological and human, and "communities"—as both emitters and users. (i.e. not just 'big data' sets and streams). Such an ecosystem (described as the '3 C's of Big Data') has four main functions or roles:

> **Descriptive**, involves narrative or early detection such as using data from satellite imagery to identify flooded areas or identifying areas in need from crisis maps;
>
> **Predictive**, includes what has been called 'now-casting' - to make real-time inferences on population distribution based on cell-phone activity before, during or after a shock, for example – as well as forecasting sudden and slow onset hazards;

Prescriptive (or diagnostic), goes beyond description and inferences to establish and make recommendations on the basis of causal relations, for instance, by identifying the effects of agricultural diversification on resilience;

Discursive (or engagement), concerns spurring and shaping dialogue within and between communities and with key stakeholders about the needs and resources of vulnerable populations such as crowdsourcing maps to assist disaster relief efforts;

In the context of climate resilience, sensor data such as satellite remote sensing (SRS) helps monitor indicators that are key to anticipating precipitation extremes such as terrestrial and sea-surface temperatures, atmospheric temperatures, sea-level rise, ice melt and glacier retreat, and changes in precipitation regimes.[6] SRS data is also fundamental in assessing numerous environmental variables including detecting trends in growing season parameters and their impacts on crops,[7] soil degradation[8] or key variables of the water cycle.[9]

Additionally, imagery is used to assess settlement patterns in countries or regions facing the most critical vulnerability challenges, particularly the presence of slums, informal settlements, and highly variable sizes and densities of built-up structures and sparse and scattered rural-urban areas.[10] Not all important resilience variables can be remotely sensed using satellites. Physical sensors on the ground provide a rapidly growing source of Big Data, enabling for example a better monitoring of carbon emissions or patterns of consumptions of electricity, which in turn can be used to make power supply systems more resilient and efficient.

Beyond data about the physical environment, Big Data also yields insights about social, demographic, behavioral, and psychological dimensions of resilience and climate change. Big Data about social ties and social networks at the neighborhood or community level were found to be strong factors determining immediate loss and long-term recovery from disaster.[11] Knowing where people are is arguably the simplest and most crucial statistic for any stage of climate analysis, adaptation, and mitigation—certainly for social vulnerability assessment and response. Call Detail Records (CDRs) and GPS locations have been compiled and analyzed for descriptive purposes such as real time information on patterns of movement and migrations.[12] The same data (CDRs) also provide insight on potential vulnerabilities and behaviors during and after disasters, including social cohesion, an important component of resilience.[13]

These examples show that Big Data has potential to improve and operationalize humanitarian efforts and develop an understanding of resilience and social vulnerability to climate change impact. New data sources along with new techniques to garner insights enable researchers and practitioners to move beyond census data and other official statistics that are difficult and time intensive to collect, to more dynamic, responsive and effective information for resilience at all scales.

However, Big Data's potential to help us better understand climate resilience does not mechanistically bridge the gap between information and action. Big Data runs the risk of becoming an extractive industry as individuals, communities and systems passively contribute information. Researchers and practitioners, perhaps even decision-makers, may extract value from Big Data analytics, but what about the communities themselves? The concept of resilience is firmly rooted in the capabilities of local communities, not 'externally engineered processes.'[14] Evidence shows that direct access to Big Data extends individual decision-making powers, such as when accessing individual and aggregated data about power consumption, climate trends, market prices, or transport conditions.[15]

How to best unlock the potential of Big Data for resilience remains an important and yet largely unaddressed question. Technical barriers include access to data 'crumbs,' global coordination for climate data sharing, data reliability and representativeness, and others. In addition, Big Data introduces new ethical and political risks that warrant careful consideration in its use. A key message is that Big Data must involve engagement (the fourth function of Big Data as an ecosystem) of several actors. Approaches should focus on not only ingesting data from people, but more importantly on getting data to the people. Building a resilient community involves fostering accountability, empowerment, and communication in all stakeholders and integrating social vulnerability at each level. In order to adjust to the changing needs of those most vulnerable in an ever-changing environment, we must put data into their hands. Making the data system a more participatory process empowers individuals to make the leap from data to action.

The rest of this paper, one the very first in-depth synthesis and analysis of Big Data for climate resilience of vulnerable populations, is organized as follows. Section 1 focuses on understanding, defining and conceptualizing resilience and Big Data. Section 2 highlights the contours of the Big Data, Climate Risk and Resilience Nexus, identifying how Big Data as an ecosystem contributes to understanding hazards, vulnerability, and exposure—the key components of climate risk—and operationalizing a climate risk management framework. Section 3 describes the barriers and gaps in using Big Data for climate resilience decision-making. Finally, the last section describes how Big Data can help vulnerable populations better understand climate data and harness their abilities to foster social resilience.

1 Conceptualizing and Understanding Resilience

> *"The idea of resilience… [is] really about how systems and settlements stand up to shock from the outside...and they don't just unravel and fall to pieces... Looking through the lens of resilience we really question how we let ourselves get into a situation that is so vulnerable. Resilience runs much deeper; it is about building modularity - building surge breakers into the basic things that support us."* -Rob Hopkins, Founder of the Transition Movement

There is overwhelming scientific evidence demonstrating the reality and impact of climate change. As some of the hottest years in history have occurred in the past two decades, it has become increasingly clear that human action has contributed significantly to this process. Climate change is increasing the frequency and intensity of weather-related hazards in varying, complex ways that can only be expected to worsen situations.[7] In 2009, the American Meteorological Society projected in the Journal of Climate that there is a 90% probability of global temperatures rising by 3.5 to 7.4 degrees Celsius in less than 100 years. These seemingly minor and predictive changes in temperature, although uncertain, could generate widespread disasters: sea levels could rise, threatening coastal communities; water supplies could dwindle as snow and ice evaporate; hurricanes and typhoons could become more powerful; droughts and floods could become more common; agricultural production could decrease; and incidences of disease could increase.

As increased climate variability linked to climate change produces more dangerous environmental disasters, developing countries—in particular, those living in extreme poverty in these countries—are most at risk. Longstanding patterns of economic development make poorer communities more vulnerable to natural hazards and leave them less resilient than developed countries when disasters occur. The Overseas Development Institute (ODI) predicts that up to "325 million extremely poor people will be living in the 49 most hazard-prone countries in 2030, the majority in South Asia and sub-Saharan Africa." The confluence of risk factors and the possibility for widespread disaster has led to increasing calls to make disaster risk reduction a core development concern and to increase efforts to address climate change at its core as it impacts sustainable development and impoverished communities.

While international efforts discuss and implement global climate change mitigation and adaptation strategies (i.e. emissions reduction and flood defenses), resilience has increasingly become a part of the global conversation to help vulnerable populations face imminent hazards.

Box 1: Key terms and concepts at a glance

Climate Change
"A change in the state of the climate that can be identified (e.g. using statistical tests) by changes in the mean or variability of its properties, and that persist for an extended period, typically decades or longer…whether due to natural variability or as a result of human activity."[16]

(Natural) hazard
A sudden or slow-onset natural event or process that may cause harm to humans or other organisms. Examples are floods, drought, earthquake, desertification, landslide, epidemic, and locust invasion.

(Natural) Disaster
As distinct from a hazard, a disaster consists of the combination of a natural hazard and its effect on the population and assets.

Vulnerability
The characteristics and circumstances of a community or asset that make it susceptible to the damaging effects of a hazard.

Exposure
The actual extent to which assets or populations are likely to experience a given set of hazards over time. Exposure plus vulnerability equals risk.

Risk
The combination of the probability of an event and its negative consequences.

Impact
The sum of the consequences if the risk does occur.

Defining resilience

Although resilience is used very differently across contexts (and sometimes conflated with sustainability), we define resilience as the capacity of a system (or community ecosystem) to absorb disturbance, reorganize and then retain essential functions, structures and feedbacks.[17] Conceptually, it is worth noting that what 'resilience' precisely means and entails differs between authors and disciplines—just as much as the meaning and features of vulnerability and its relation to resilience remain contested (we conceptualize vulnerability not just as the converse of resilience, but as "the susceptibility of [a] system or any of its constituents to harmful external pressures,"[3] while resilience characterizes a process or set of processes).[5]

In most of the policy and social science literature and discourse, resilience is understood as the (normatively desirable) capacity of a social system to 'bounce back' in the face of a shock; from its ecological science origins, the concept rather describes a system's ability and ways of leveraging its adaptive capacity to "maintain structure and function" without crossing critical thresholds into a state with new sets of system dynamics —while 'bouncing back' in the face of shock would be closer to a system's *robustness*.[6]

In the context of climate change, resilience involves a system withstanding variability in climate (and non-climate) stressors through its adaptive capacities (e.g. frequent, intense drought conditions as climate stressors for agriculture and the capacities of farming communities to adjust, adapt or cope). These stressors can be either *familiar* disturbances (i.e. a hurricane during its regular season), or *unfamiliar* disturbances (i.e. new, unknown climate trends). In this context, agriculture systems in resilient communities are able to "maintain structure and function" without crossing critical thresholds by which they would be unable to cope in the presence or aftermath of the stressor. A community would therefore be resilient to the extent to which it can cope and avoid critical thresholds, but not resilient if after the disturbance, community members have not regained any form of livelihood.

Similarly, like a plant that attempts to resist a disease and remain robust, a farming community who remains in a disturbed state, a resistant state, after a drought, has more difficult recovering, and eventually may not be able to recover at all (Walker, 2011). Being in this resistant (robust) state erodes the capacity to respond and organize in a beneficial and productive way. Many times, robustness can cause peripheral blindness, a preference to focus on only those hazards that favor quick responses and a tendency to declare success on a superficial basis. In time, these longer-term vulnerabilities that result from hazards are ignored. In this manner, resilience affords the capacity of tolerance and improvisation, and robustness does not.

Box 2: History of international efforts to address climate change

1979 - the first World Climate Conference

1992 - the UN "Earth Summit" produced the *United Nations Framework Convention on Climate Change* (UNFCCC), a first step in considering what could be done to limit climate change including how to cope with its inevitable impacts.

1997 - the Kyoto Protocol was adopted by the UNFCCC, which legally binds countries to reduce emission level to certain targets. Today, there are 195 partied to the UNFCCC and 192 parties to the Kyoto Protocol.

1998 - the Intergovernmental Panel on Climate Change (IPCC) was created by the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP) as an objective source on scientific information

2009 - The UN Climate Change Conference released the political agreement, the Copenhagen Accord, in which the world's major economies offered explicit international climate pledges.

2010 - The *Cancun Agreements* is adopted to reduce carbon emissions and build a system which made all countries accountable to each other for those reductions

2011 - *Durban Platform for Enhanced Action*, a comprehensive plan to draw up the blueprint for a fresh universal, legal agreement to deal with climate change beyond 2020

2012 - Rio+20: Re-thinking Sustainability and Sustainable Energy for All

2014 - COP 20 – Lima and defining Post-2015 SDGs

2015 - COP 21 – Paris New Climate Change agreement and Post-2015 SDGs

A distinctive, defining, feature of a resilient system is adaptability, which emphasizes the role of societal learning over mere technical fixes that may reduce system vulnerability in the short term but decrease the kind of flexibility key to resilient systems in the long term. Response strategies for hazards must also consider building resilience against possible future disturbances, specific and general.

Specific resilience describes resilience to a specific kind of change, such as relatively well-understood natural disasters—e.g. recurring droughts and floods due to ENSO. General resilience refers to a system's capacity to deal with unfamiliar and qualitatively new disturbances that lie outside of our past experience (another useful related conceptual dichotomy is Eakin et al.'s, which distinguishes generic vs. specific adaptation).

Resilience strategies have a tendency to focus on specific resilience to familiar disturbances, as opposed to addressing more challenging resilience building activities to possible, unknown disturbances. Carlson and Doyle (2000) suggest a potential tradeoff in resource allocation between specific and general resilience-building activities: the more resistant a system becomes in one direction, the less overall resilient it becomes in others.[18]

To illustrate the point further, one can think of a water balloon vs. a glass bottle filled with the same volume of water. The former is definitely less resistant, robust to specific stressors, such as being pocked by a needle, or compressed by hand; both quickly 'bounce back' if prodded with a larger stick (the latter would simply not change until it breaks at some point); overall the balloon can be said to be more adaptive therefore more resilient. If exposed to extreme temperatures, it will expand alongside its content, move up; likewise if the air pressure increases or drops, it will adapt. The glass bottle is sturdier but—therefore—less adaptive; it is less able to "absorb disturbance, reorganize and then retain essential functions, structures and feedbacks under changing circumstances". It is less resilient. The balloon can certainly 'bounce back' if conditions go back to what they were, but what characterizes its resilience—whether being resilient is desirable or not—is the fact that it retained function while conditions had changed.

In the context of climate change, these seemingly minor distinctions between resilience and robustness, and general and specific resilience, have important consequences, especially when combined with other considerations of geographical and temporal scales and the specificity of the hazard considered. In discussing these, we rely on resilience frameworks proposed by Turner et al (2010) and Carpenter et al (2012). Turner et al. stress the role of *science* in enhancing society's ability to act to avoid disasters along the cycle prevention-mitigation; monitoring-prediction; response-recovery while Carpenter et al. highlight the role of *information* in empowering people to act and hold each other accountable.

Big Data for resilience

The last decade has seen a rapid growth in the volume and variety of available data and geospatial data as a result of the spread of Internet and mobile devices and the increased availability of Big Data, including satellite imagery at high resolution. This historical growth has significantly expanded the potential of Big Data for climate change and resilience. The proposition is not new: climate science has always relied on large amounts of spatial data and computational power. With the 'rise' of Big Data, however, the opportunities and challenges to building a better understanding of hazards, vulnerabilities, and resilience factors have vastly expanded.

The term 'Big Data' encompasses both the new technologies that generate large-scale data on social outcomes and the opportunities created by data. Unfortunately, Big Data is too often reduced to 'big data' characterized by the volume, velocity, and variety of the data, which overlooks most of its novelty and complexity. Instead, Big Data must be conceptualized as a new socio-technological phenomenon resulting from the emergence and development of an ecosystem made up of the new kinds of data 'crumbs' about behaviors

and beliefs generated and collected by digital devices and services, ever more powerful computing

Together, these three parts, the crumbs, capacities, and communities, form a complex system in which feedback loops in data generation, use, and assessment produce new data and techniques. At the most basic level, organizations generate new kinds of data that lead to the development of new kinds of analytical tools, and then various actors interact with those tools.

In the following section, we will explore the Big Data, Climate Risk and Resilience nexus and evaluate the added value of Big Data as an ecosystem toward each component of climate risk—hazards, exposure and vulnerability—as well as its contribution to the climate risk management framework.

2 The Big Data, Climate Risk, and Resilience Nexus: Added Value of Big Data towards Climate Resilience

In this section we will explore the Big Data, Climate Risk and Resilience Nexus and identify how Big Data as an ecosystem contributes to understanding each of the components of climate risk—hazards, vulnerability and exposure—as well as the climate risk management framework.

2.1. The Big Data-Climate Risk-Resilience Nexus

The increasing scientific evidence of global climate change highlights the need to not only acknowledge and disseminate the relevance of incorporating adaptation in projects, programs and sectoral policies, but also ensure the effectiveness of investments aimed at reducing poverty and boosting development. In this sense, the approach of the World Bank is rooted in the expansion of a *climate risk management* culture, which is firmly based on the acknowledgment that the rising risks of climate change and their implications for development are seldom explicitly addressed. Embedded in the World Bank's approach, there is the acknowledgment that climate change is only one element in the spectrum of risks facing an investment. Within this framework, climate-related risk should be first associated to the current climate conditions, including variability and weather extremes, and so should its management. Big Data is key to determine climate baselines and identify the vulnerabilities to climate-induced hazards through 'crumbs' (the first 'C').

A step forward is required to assess whether existing or potential vulnerabilities could fluctuate because of climate change, as well as of other types of significant modifications of the surrounding conditions (i.e. population growth and/or shifting trade patterns) in order to incorporate appropriate measures into the design and implementation of actions, if necessary. Data-driven analytical tools are necessary to spot vulnerabilities that may be worsened by the effects of climate change, making adaptation to climate change an integral part of the climate risk management approach. In such a case, big data applied to climate risk management helps structuring an iterative process that starts with determining coping

strategies for current climate variability, tries to anticipate changes in climate change through the implementation of *ad hoc* adaptation management, and seeks to evolve new coping strategies as necessary. Since climate and climate change are not only an environmental issue, but also a major economic and social risk to national economies, data-powered tools for integrated analysis of vulnerabilities and cross-sector strategy planning are needed in both the assessment and implementation phases of development initiatives. In this sense, big data tools and capacities (the second 'C') can help increase the availability and relevance of climate-related information through integrating and interpreting historical series of data and short-term weather forecasts and taking into account seasonal climate predictions and projections of long-term climate change effects - especially when considering highly vulnerable areas.

The collection and systematization of these types of climate-related data contributes to better managing climate risks at a broad time scale and at the same time tailoring such information to bring together user communities and local institutions for integration into the planning processes. The application of the new generation of analytical tools can facilitate the interpretation of climate risk, especially enabling donor agencies, policy makers, and development stakeholders assessing initiatives that may be subject to climate risks. Improved visualization tools can help ensure the effective use of data as the basis for climate risk management. *Section 2.2* will focus on the way Big Data can characterize exposure and vulnerabilities to climate change in the context of climate-affected sectors in developing countries. Lastly, *Section 2.3* evaluates Big Data's contribution to both adaptation and mitigation programs across the main sectors affected by climate in the developing world, including health; water supply and sanitation; energy; transportation; industry, mining and construction; trade and tourism; agriculture, forestry and fisheries; environmental protection; and disaster management.

2.2. The Interaction of Big Data and Components of Climate Risk: Hazards, Vulnerability and Exposure

a. Climate change monitoring and hazard detection

Data can help monitoring and understanding the drivers of climate change, detecting variations that can lead to negative impacts. SRS data are an extremely important and growing resource to this aim, helping to monitor terrestrial and sea-surface temperatures, atmospheric temperatures, sea-level rise, ice melt and glacier retreat, and changes in precipitation regimes.[19] In doing so, they help validate other monitoring systems. For example, SRS temperature data confirmed the positive link between surface and atmospheric temperature, while showing some difference in the magnitude of the link, which orients new research.[20] Satellite altimetry data has shown a higher rate of sea-level rise than in-situ gauges. Both data sources have their limitations and sources of errors, but by combining them, a fuller sea-level budget is being pieced together.[21] With a longer period of observation, this will allow us to distinguish long-term trends from decadal variability.

SRS data also help validate climate models. A powerful example is the use of satellite observations to test critical feedbacks between water vapor and temperature at different atmospheric layers.[22] Although precipitation data has not resolved the question of whether there is a global mean trend of precipitation, it has given scientists a keener understanding of the climate drivers of precipitation variability. For example, it has confirmed that there are positive feedbacks ('wet-gets-wetter' and 'dry-gets-drier' trends) in tropical regions, knowledge that is key to anticipating precipitation extremes.[23]

SRS data is also fundamental in assessing the impact of climate change on environmental systems and society, as well as creating baselines for mitigation and adaptation. Over the last decades, SRS retrieval algorithms (algorithms that convert remote sensors' signals into measurements of useful variables) have been developed to make inferences on numerous environmental variables:

1) Land cover, at increasing levels of spatial resolution[24]
2) Biozones
3) Vegetation productivity[25] and various indicators of vegetation health, such as dryness, forest disturbances at fine spatial scales[26] and carbon fluxes driven by photosynthesis[27]
4) Phenology, detecting trends in growing season parameters and their impacts on crops[28]
5) Soil physical and chemical characteristics[29], including reliable indices of soil degradation[30] advances, which can help overcome the "soil data crisis"[31]
6) Key variables of the water cycle, including evapotranspiration[32] and soil moisture.[33]

These spatially and temporally continuous streams of data on the natural environment present an enormous potential, paving the way for real-time and integrated sensing at multiple spatial scale (local to global) and temporal scale (from short-term shocks to slow environmental change). As an example, SRS data has increased our ability to monitor the overall water cycle and understand heat and water fluxes at regional and global scales. Increasingly, this also allows us to anticipate changes in water availability at finer spatial scales. For instance, Forsythe et al. (2012) combined river gauge data from the lower reaches of the Indus Basin with remote sensing data of snow cover and surface temperatures to derive remote sensing data proxies of water availability at the valley scale, useful to farmers in the upper reaches of the Indus Basin.

The combined analysis of social sensors can also offer the capacity to monitor the impact of climate-driven hazards. Big Data offers the ability to generate insights to make sense of the dynamic interaction of the multiple catastrophes that emerge over time through social dynamics, the status of the infrastructure, and the triggering of other hazards. First, as the length and the spatial resolution of satellite records grow, the integration of remotely sensed climate data with terrestrial data will improve our ability to detect climate driven changes in ecosystems, species distribution, water availability, organic carbon sinks, among the others.

Second, much can be learned on the socio-economic impacts of short-term climate shocks by combining environmental and climate hazard data with Big Data sources on social dynamics and infrastructure that have proven their worth in the domain of disaster crisis response.[34] An interesting approach is that of earthquake detection, which offers lessons to improve the detection of climate-induced hazards as well. Leveraging complementary "social" sensors (such as Twitter, Instagram, and YouTube) and physical sensors (such as USGS seismometers and TRMM satellite) can be particularly relevant to improve hazard detection. Lastly, analyzing the social impacts and social responses to slower-changing environmental variables can help assess impacts and understand how populations adapt to a changing environment.

However, not all of the important environmental variables can be remotely sensed.. A significant frontier is expanding the use of volunteered geographic information in documenting changes in ecosystems and natural resources: citizen-science. Citizen-science involves research conducted or assisted by non-expert public volunteers. This growing field of participatory research is already common in ecology[35] having been used to detect changes in the diversity of organisms[36], phenology and range shifts.[37] It is now scaled up with the help of platforms such as Tomnod[38], smart-phone aps (e.g. WeSenselt) and mobile environmental sensors (e.g. Citclops collects data on ocean water quality from mobile phones) and to provide relevant measures of water quality, weather, and light pollution. Further scaling of mobile sensing and citizen-science will arise from standardizing field protocols[39], and from computational innovations in merging these complex datasets.[40]

b. Assessing exposure and vulnerability

Identifying the exposure to climate risks requires comprehensive and up-to-date information to classify the main attributes of areas where hazard events may occur. Exposure is a necessary (though not sufficient) determinant of risk, and understanding how societies are exposed to climate-driven hazards is the first step towards identifying vulnerabilities. Particularly, data help to describe the main determinants of exposure – i.e. the elements that can be affected by climate change (e.g. populations and assets), the change in climate parameters (e.g. precipitation, temperature, etc.), and climate-driven components of ecosystems (e.g. sea level rise). This section specifically focuses on the former elements and the key factors for reducing exposure.

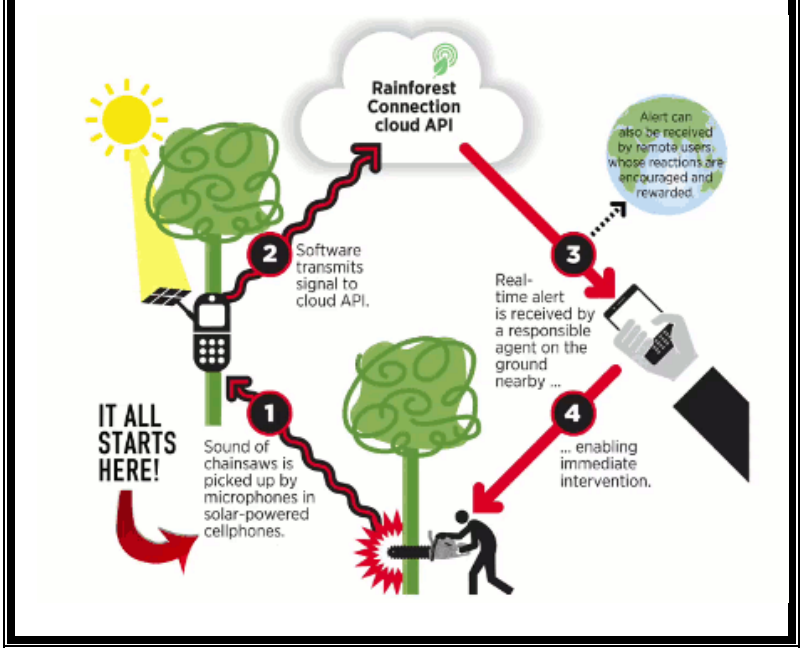
Geospatial data, including SRS, are essential to characterize land use, which in turn offers the opportunity to classify the level of exposure of communities to climate-driven hazards due to their spatial distribution. The analysis of lights at night through remote sensing, for example, enable institutions to track the distribution of electrical services in potentially-affected areas.

Big data leveraging data crumbs is helpful to understand population dynamics in potentially dangerous areas. Understanding the way population grows in a territory can provide insights about what actions can mitigate risk. By coupling data from mobile phones and statistical

inference, Wesolowski and Eagle (2010) modeled the dynamics of informal settlements in Kenya to provide policy makers with predictive insights related to urban planning.

Big data can also help detect changing patterns of basic environmental conditions that can worsen exposure to risks or make territories become exposed to risk.

Figure 1: Description of the system developed by Rainforest Connection



Combining crowdsourced data with SRS data can help detect changes in a territory. An example of this combination can be seen through the Rainforest Connection[41], a social enterprise based in the US that successfully tested the use of old smartphones to detect deforestation and alert authorities about illegal logging activity in Sumatra.

The application of Call Detail Records (CDRs) from Telecom operators to map population dynamics can help identify suitable approaches for assessing the exposure of populations to disasters. The combination of CDR, remote sensing, and census data have successfully been applied in Europe to map time-sensitive population dynamics[42]; the case was also notably successful in being able to preserve the anonymity of mobile phone users. It should be highlighted that at present CDRs can be particularly useful if coupled with census and/or survey data, especially for interpolation when data is lacking or scarce, as well as for those cases where populations are not typically covered by official surveys (e.g. informal settlements). CDRs alone are thus a magnifier of survey and census data, and not a substitute for them.

To assess vulnerability, Big Data offer innovative approaches to determine the (potential) negative impact of climate-induced phenomena (e.g. natural disasters) on communities, their livelihoods, and assets. Blending distributed data from the factors that can determine adaptive capacity through ad hoc analytics can offer institutions with a dynamic assessment of vulnerability, driving policy action in key hotspots that may compromise the ability of societies to cope with the adverse impacts of climate change. Big Data sets gathered via mobile phones can complement surveys in mapping vulnerability and exposure of populations to climate-driven disasters. Leveraging network science, for example, may provide inferences about the most stable market hubs or the best-kept transportation and service infrastructures based on population mobility around or on them.

Social vulnerability aims at determining who is more vulnerable to climate change because of their underlying social conditions. The social, demographic, behavioral, and psychological dimensions that make an individual and group more vulnerability to the impacts of climate change fall along a spectrum of universality or generalizability; some dimensions are fairly well documented and consistent across geographies[43], while others vary significantly across time, place, and context. Starting with the most generalizable, people who have more financial resources, who are not especially young or old, and have strong community support are less vulnerable. Many studies have proven that social ties and social networks, primarily observed at the neighborhood or community level, are strong factors determining immediate loss and long-term recovery from disaster.[44]

Over the last two decades, social vulnerability researchers have begun to distill the dimensions of social vulnerability into empirically based indicators. When combined in summary indices, typically using demographic information, these tools describe who is most vulnerable and where the most vulnerable are located before, during, and after a crisis.[45] If measured using benchmarks and monitored over time, these indicators may serve as diagnostic tools.

Big Data has immense potential to improve and operationalize our understanding of social vulnerability to climate change impact. Large and fast data along with new techniques to garner insights enables researchers and practitioners to move beyond census data and other official statistics that are difficult and time intensive to collect to more dynamic, responsive, more information for resilience at all scales. Remote sensing, user generated data and other big data can replace old data. Mobile phones have proved to be a useful source of data that can serve as a proxy to profile the socio-economic status of a population.[46] In East Africa, data from airtime credit purchases proved to be a good indicator for food spending[47], and could be used to spot vulnerabilities. Patterns of mobile phone usage were also found to be correlated with socioeconomic levels in both urban areas of Latin America[48] and Ivory Coast[49] using standard econometric techniques and machine learning approaches.

Knowing where people are is arguably the simplest and most crucial statistic for any stage of climate analysis, adaptation and mitigation—certainly for social vulnerability assessment and response. The current population data used to track populations for disaster work and for long term climate impact preparation are severely limited, slow, biased, coarse and overall insufficient for understanding and coordinating the needs of people on the ground. Scientists and relief coordinators currently rely on eyewitness accounts, manual headcounts, registration in refugee camps, satellite or aerial images of shelters, or changes in vegetation. Real-time population counts enable the location of vulnerable population during and right after climate-induced disasters. The Flowminder[50] and WorldPop[51] projects are leading examples of this work.

User-generated geospatial information from Geo-located Tweets, Cell Phone Records data and other sources indicate a new frontier understand more demographic attributes at the household, neighborhood, and community level to order to have a more robust and accurate picture of vulnerable.[52] More complex and context-dependent dimensions of social vulnerability, or the least generalizable variables, require richer data. Social media gives us a better view into the lives of those on the ground, or what some refer to as the "Big Data Shadow", as well as the psychological factors behind perception and behavior. In studying how relationships on Twitter evolved before, during, and after the Hurricane Sandy disaster, researchers at the University of Colorado found that geographically vulnerable Twitter users propagate more information during the disaster than before or after, and are more likely to re-tweet locally. These users also re-tweet differently, resulting in a relatively higher proportion of tweets that receive between 10 and 80 re-tweets.[53]

Overall, as climate change threats of all kinds increase, we need a better understanding not just of their physical impacts but the social drivers and impacts behind vulnerability. Current vulnerability science is overwhelming qualitative, which is immensely valuable but limited, though we have a robust body of case study literature and some other qualitative assessments. However, the lacuna of quantitative information of what makes people more vulnerable demonstrates a critical need for Big Data within climate science and response. On the other hand, Big Data is also expanding the ability of vulnerable communities to provide feedbacks to institutions and development stakeholders. This can help their society maintaining functions under the stress posed by climate change impacts. As crowdsourcing data through ad hoc platforms can improve relief during disasters, similar approaches may be employed to assess vulnerabilities of communities under climate change (such as MIT Climate CoLab[54]).

c. Big data for evaluating resilience

Projects have so far demonstrated the potential of Big Data for measuring factors that can contribute to assess resilience. The two potentially biggest advantages in Big Data for resilience measurement seem to relate to: (a) the opportunity to help quantifying many of the intangible elements that are important to model resilience at the household level e.g. social capital, institutional norms and behavior, cultural marginalization); and (b) the ability to fill data gaps in those regions or countries where available socio-economic data is modest and there are gaps in historical data. This is typically the case at the sub-national level, where accurate historic data is rarely available, and in areas where it is difficult to conduct surveys (e.g. due to conflicts). Ubiquitous sensing through mobile phones proved to be a valid technique to provide poverty estimates at higher spatial resolutions than those available in data-scarce regions.[55]

CDR data particularly present an unprecedented source of information to better understand this aspect of resilience and vulnerability. CDR data allow us to reconstruct social networks on a large scale and obtain information about both their local and global structure.[56] Mobile phone data also enabled effective characterization of cognitive constructs of social groups, other than their observable behavior[57], and the linkages between the socio-economic status

of a community and the relationships among its members.[58] Onnela et al. (2007) analyzed a society-wide network of communication and demonstrated that CDR data provides crucial information about the strength of ties and the ability of information to spread. Nevertheless, there are not yet attempts at using Big Data to measure the institutional components of resilience (i.e. state capacity, trust in government, responsiveness). Moreover, a general trend of current research shows a siloed approach, focusing on specific hazard tracking and mapping while dedicating much less attention to slow-evolving factors (e.g. land degradation, status of ecosystem services) that could co-operate to generate disasters and/or exacerbate their impacts by weakening resilience properties of an ecosystem or a society. Fostering interdisciplinary research through improved interoperability among databanks and systems seem pivotal to advance the field in such sense.

Big Data can also be used to build more resilient landscapes. Scientists from CGIAR develop a model incorporating different sources of data on ecological, economic and social variables, as well as the qualitative insights born from the long experience of practitioners to help improve agro-ecosystem interventions[59]. The key to this approach is that all factors are represented as probability distributions so that all sources of uncertainty are taken into account. As a result, the risk-return of different interventions can be compared. For the more, the model assesses how much our uncertainty would be reduced by additional data streams – providing guidance to data-driven policy making and action. This model could become a powerful tool in agricultural decision-making if connected to spatial probabilistic databases (rather than aggregate averages), which is something Big Data can clearly contribute to. This would be a significant contribution to building resilience since the sustainable intensification of agriculture can mitigate climate change, the water crisis, underemployment, extreme weather events, biodiversity loss, inequality and food crises.

2.3. Big Data and its added value toward climate change adaptation and mitigation strategies

In this section, we explore the possible contributions of big data to support climate change adaptation initiatives as well as the efforts to mitigate climate change. Big data can help strengthening both adaptation and mitigation programs across the main sectors affected by climate change in the developing world, which include: (a) human health; (b) water supply and sanitation; (c) energy; (d) transport; (e) industry, mining, and construction; (f) trade and tourism; (g) agriculture, forestry, and fisheries; (h) environmental protection; and (i) disaster management.

d. Mitigation

Data is the basis for accurate quantification of the emissions of individuals, economies, urban centers and landscapes. Big data can therefore drive higher efficiency of mitigation actions, particularly by increasing the accuracy in monitoring GHG emissions of industries and consumers, tracking carbon cuts and sequestration through integrated sensors networks, determining the transition to renewable energies and more energy- efficient practices. At the same time, it should be underlined that mitigation is not possible without political action and

collective action. As a result, we also highlight how more open and more fine-grained data may facilitate the social change that is needed to transition to less carbon intensive societies.

Big Data contributes to low carbon systems in the energy sector both on the supply side and on the demand side. On the supply side, companies that produce and sell energy can use remote sensors and big data-stream analytics to reduce carbon emissions by making real-time efficiency improvement on the short term, identifying ways to improve existing infrastructure such as power distribution grids on the medium term, and better planning future projects and resource management on the long term.[60] Using Big Data to make this kind of assessment eliminates the need to wait for field visits, which are often costly, but also leads to "more precise and reliable data, since direct data measurements through sensors do not suffer from recall or courtesy biases common to data collected in surveys. On the demand side, the main mechanism for reducing carbon emissions through Big Data involves extending consumer decision-making powers. Smart meters (which relay information bilaterally and in real-time between power companies and homes) are expected to help consumers make more informed decisions about their power consumption. For the time being, the evidence is mixed as to their effectiveness[61], since users must be sufficiently motivated to reduce their consumption independently of the information provided.[62] The possibilities abound however: allow users to set up real-time email or SMS notifications when their consumption reaches a certain level; keep consumers informed of shortages or peak usage times to help them reduce their consumption at key moments; create cost-saving opportunities by allowing users to respond to time-specific pricing, such as recharging electric vehicles during the night when excess power supply results in lower prices.[63] While the discussion of smart meters currently tends to focus on residential applications, installation of smart meters also has benefits in commercial and industrial buildings, which account for the bulk of electricity consumption usage: smart meters can help favor low carbon systems by providing incentives for commercial and industrial consumers who seek to reduce operating costs.

Big data can also support mitigation in the transportation sector, one that mostly contributes to climate change, producing roughly 23% of the global emissions of greenhouse gases. The sector is also the fastest growing consumer of fossil fuels and the fastest growing source of CO2 emissions.

In the area of transportation, data from a range of new sources, including Call Detail Records and GPS locations, can be compiled and analyzed for descriptive purposes - such as gathering real time information on traffic volumes and congestion levels -, predictive purposes - such as predicting traffic conditions and travel time[64] and prescriptive purposes - such as implementing dissuasive toll systems with automatic billing during rush hour. CDRs were also employed by AT&T Labs affiliates to "calculate home-to-work travel distances and estimate the median carbon emission per home-to-work commute of hundreds of thousands of people living in the Los Angeles, San Francisco, and New York metropolitan areas. Their estimates showed that in New York, living further from the center was correlated with an increased carbon footprint since most people commute into central Manhattan for work.

However, this is not the case for Los Angeles, since there is no single geographical concentration of jobs. The result for San Francisco is somewhere in the middle of the two cases."

As part of the Orange D4D Challenge 2013, IBM Ireland analyzed 2.5 billion CDRs from Côte d'Ivoire to develop the AllAboard Project, which seeks to maximize ridership and minimize waiting and travel time under budget and existing fleet constraints, yielding alternative transit routes that reduced journey times system-wide by 10%.[65] Other sources, such as traffic cameras, data from GPS devices in personal vehicles, and open data from sources such as OpenStreetMap can also be employed to understand travel patterns and suggest more energy efficient routes. Many projects combine data from several of these sources, such the collaborative effort between IBM and Singapore's Land Transport Authority, which uses data from video cameras, G.P.S. devices in taxis and sensors embedded in streets to generate hourly reports on traffic conditions geared towards city administrators and commuters. The model predicts traffic volume and speed for 500 locations up to 60 minutes in advance with accuracy ranging between 85-95%.[66]

Land use and land-use change (LULC) represents a major portion (16-20%)[67] of green house gas (GHG) emissions. Monitoring agricultural land and practices in precision farming, deforestation and illegal logging, and urban expansion and form[68], are being understood as the spatial and temporal resolution of remote sensing data increases. New cloud computing architecture such as Google Earth Engine and Amazon Cloud Services has expanded the ability for scientists and citizens to analyze LULC over new spatial and temporal scales with publicly available data. A landmark example is the new version of Global Forest Watch (www.globalforestwatch.org), a decades old project of The World Resources Institute focused on illegal deforestation. In partnership with the University of Maryland and 40 others, the project analyzed almost 700,000 Landsat images, using a total of 20 terra-pixels of Landsat data and one million CPU hours on 10,000 computers working in parallel on Google Earth Engine. A standard computer would have taken 15 years to complete this analysis. The analysis determined annual global forest change in forest cover since 2000.[69]

In another example, a human settlement mapping project of unprecedented size will determine the urban area extents of the entire globe. Using Landsat imagery in Earth Engine, the team, lead by Dr. Paolo Gamba at the University of Pavia, aims to eventually produce a high-resolution global map at 30m, showing change over the last 10-20 years.[70] Understanding these historical patterns helps better quantify the influence of LULC in carbon budgets, but more importantly, allows us to understand major drivers of LULC, and monitor changes at local levels that can improve management and inspire action. Software such as CLAS lite expands remote sensing analysis beyond scientists and enables governments and NGOs to tap the power of remote sensing of deforestation and LULC. The free software, along with a MOOC in Spanish and English, has trained over 300 organizations, many of which have used CLASlite to track oil palm expansion, gold mining, and map carbon[71]. Other groups such as Amazon have created their own forest monitoring tools in platforms like Earth Engine to track Amazon deforestation when it suspected the

government's data was inadequate. The group now produces monthly reports[72], as well as near real time alerts when forest lost is detected. This allows Amazon's team to independently monitor ecosystem changes in order to better respond to deforestation threats, and increase transparency about where and when forests are being lost.

In addition to new computing power and software expanding the power of remote sensing among scientists and decision makers, the potential of new satellite systems such as Planet Labs will may transform our ability to monitor LULC at ever high spatial and temporal resolution. Planet Labs, launched in 2014, uses a suite of tiny satellites that will provide 3-5 meter spatial resolution with near daily global coverage. While this data is not free, it has the potential to greatly improve detailed monitoring for changes in crops, forests, infrastructure, and environmental change. This stream of Big Data is being leveraged by private endeavors such as Redd Systems to support REDD (Reduction in Emissions from Degradation and Deforestation), a UN effort to create financial incentives for countries to reduce emissions from forested lands. REDD Systems leverages Planet Lab's imagery to track forest degradation daily, predicts future deforestation from this high resolution time series, and uses SMS to alert natural resource managers. This enables decision makers on the ground to intervene and add value to their carbon credits. REDD systems is just one example of how Big Data in satellite imagery can be translated to improving monitoring and mitigation of climate change by analyzing LULC.

Climate change mitigation is at least as much a political problem as it is a technical solution and concurrent problem. A modern society's dependence on fossil fuel is often entrenched within the political process, in the form of tight-knit policy and industrial networks.[73] The promotion of alternative technologies and policies require the displacement of these coalitions.[74] What then is the contribution of big data to uncovering the political economy of climate mitigation? So far, the contribution is slim, but there is untapped potential stemming from the very large amount of digitized text data, the resurgence of social network analysis in political science and social media data.

The largest study on the social networks underlying energy policy dates back from the 1980s when Laumann and Knoke (1987) collected extensive network data from interviewing industrial lobbies, trade associations, citizen groups, labor union, state government agencies and congressional staff. With modern network analysis methods, Carpenter et al. 2004 were able to derive important insights on the policy-making process yet such studies are rare because such data are very expensive to obtain by traditional methods (interviews and surveys). The increasing availability of digital records of political participation can be used to more efficiently reconstruct networks, as demonstrated by Onnela and Lazer's (2009) reconstruction of the network linking campaign donors and politicians, and this should be leveraged to understand the political changes surrounding transitions to a lower-carbon society. Leveraging digitized political archives and social media is also possible to understand communication and power. Adapting machine learning algorithms to political processes, Grimmer (2010) analyzed senators' press releases to map their political agendas, while King et al. (2013) automated the analysis of millions of social media posts from China to analyze

patterns of political censorship. Such applications open the possibility of analyzing the influence of different actors in framing policy.

e. Adaptation

This section explores how Big Data can feed into adaptation efforts, primarily by buttressing our ability to identify infrastructure and communities that are particularly vulnerable to the increasing climatic variability and increasing frequency of extreme events. It can also strengthen societies' capacity to anticipate and prepare for specific shocks.

Climate adaptation planning requires an accurate mapping of the built environment and infrastructures. Recent advance in remote sensing analytics and open access to global high-resolution satellite data have enabled novel method of automatically and accurately delineating built-up areas using VHR satellite imagery.[75] This approach allows for faster and less labor-intensive processing to identify the evolution of urban form over time. This approach also helps assessing settlement patterns in countries or regions facing the most critical vulnerability challenges, particularly the presence of slums, informal settlements, and highly variable size and densities of built-up structures and sparse.[76] The global coverage and the potential for rapid re-assessment could be used to establish baseline layers and monitor changes.

In addition to the growing capacity to access, process, and interpret SRS data, crowdsourcing and user-generated information is increasingly used to map out the built environment and infrastructures. Among many Web 2.0 mapping applications, the OpenStreetMap Project (OSM) is the most striking example of utilizing Voluntary Geographical Information (VGI). The speed at which OSM was populated with volunteer generated data on detailed roadmaps and locations of critical assets immediately after the Haiti earthquake in 2010 is astonishing.[77] The quality and validity of VGI datasets vary based on the number of volunteers per unit area and the socio-economic factors of the area being mapped.[78]

Agriculture is among the sectors that will be hit first and foremost by climate change. Adaptation strategies are thus already in place in the sector, and the most widely used big data stream for this purpose is SRS. Remote sensing instruments such as the Moderate Resolution Spectroradiometer (MODIS) and the Advanced Very High Resolution Radiometer (AVHRR) have been used to assess the impacts of drought on crop productivity. Early warning systems can compare the current vegetation index (VI, a measure of greenness) compared to the long-term average.[79] The Famine Early Warning System (FEWSNET) uses VI changes in combination with field data such as market prices to predict the likelihood of regional food deficits. More recent soil moisture sensors, the Soil Moisture and Ocean Salinity (SMOS) and NASA's forthcoming Soil Moisture Active Passive (SMAP), provide additional information that can help analysts to understand the length of drought.

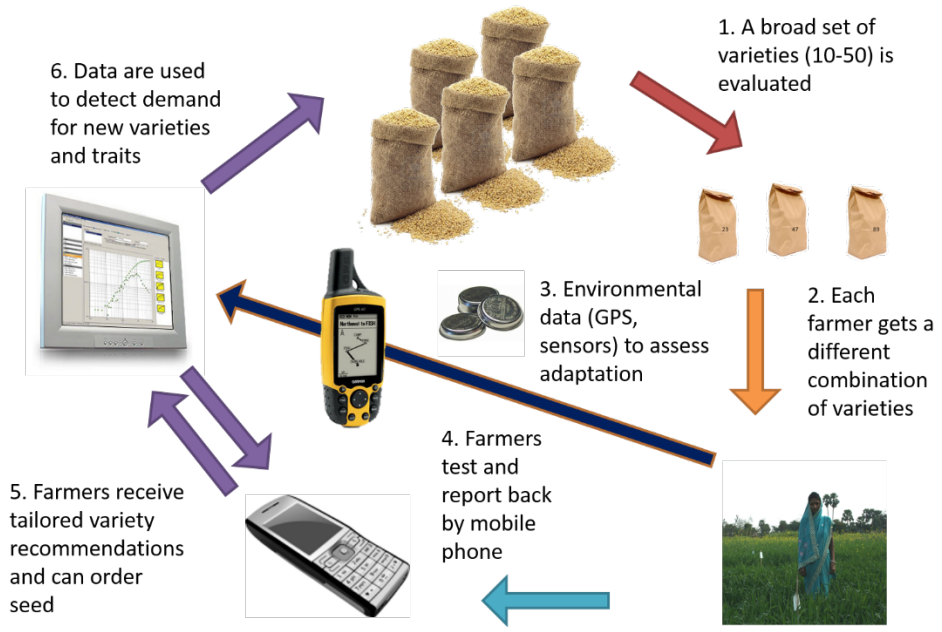
MODIS data have also been used to monitor floods[80] and as an input for broader flood modeling. This can be an important predictive data layer to assess the risk of floods.[81] While SRS imagery cannot necessarily be used in early warning systems, since the flood is actually

occurring, the models can help to identify areas that are likely to be at risk from floods in the future, and populations in these areas that may be vulnerable to climate impacts.[82] Crop losses from flooding are an important form of climate change loss and damage.[83]

Mobile phones and text messaging has become an important means for farmers in remote rural areas to collect information on market prices in market towns so as to plan trips to those towns for when prices are optimal. This can be an important form of resilience building for farmers who have traditionally suffered from information deficits with regard to market prices. Platforms such as Ushahidi could be used in the case of food security crises. Mobile platforms relying on citizens for data collection such as U-Report have been used to track the spread of pests threatening food supply.[84] The CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) proved that big data is also promising for generating new insights out of large-scale and distributed collection of observations of the behavior of plants under different factors.

Researchers from International Center for Tropical Agriculture (CIAT) demonstrated the viability to leverage non-experimental crop, soil and climate data to provide farmers with ad hoc advise about what, when and where to plant. Researchers from CIAT successfully tested this approach in five countries with eleven staple and cash crops.[85] Researchers from Bioversity International prototyped a software application to test the behavior of 300 crop varieties under different conditions with 15,000 Indian smallholder farmers in less than three years. Researchers provided the farmers with different seeds, while the farmers informed him about the performance of such varieties through a simple SMS-based system. This way it was possible to monitor the performance of crop varieties under different conditions, identifying the ones resulting more adapted to climate stressors.[86]

Figure 2: Workflow of the information system setup by Van Etten et al. (Bioversity, CCAFS)



Big data is being leveraged also in the private sector to offer tailored information to smallholder farmers that could strengthen their adaptive capacity. Climate Corporation[87] is currently developing an information system based on MapR's distribution of Hadoop to improve the accuracy of weather patterns prediction across the United States, and such an approach may be implemented to data-scarce region of the world for providing tailored meteorology advise. Ignitia, a Swedish social enterprise, undertook an innovative approach blending weather data from different sources (including SRS and ground stations) to develop a tropical weather forecast model in West Africa. Thanks to the model, farmers receive daily forecast via SMS. Ignitia[88] is also currently investigating how to predict trajectories, location and patterns of thunderstorms in West Africa through a Storm Tracking Model (STM).

3 Barriers and Gaps in Mainstreaming Big Data for Climate Risk Management

In the previous section, we explored the Big Data, Climate Risk, and Resilience nexus and identified Big Data's contributions towards a climate risk management framework: in monitoring climate and spotting climate-driven hazards; characterizing and assessing exposure and vulnerabilities to climate change in developing countries; and finally, improving both adaptation and mitigation strategies.

A number of barriers and gaps affect the feasibility of mainstreaming Big Data as a complex ecosystem with climate resilience strategies and frameworks. Though many of these challenges reflect issues common to Big Data applications for development, this section contextualizes them for climate resilience.

We identify the following challenges across the ecosystem of Big Data, specifically the barriers and gaps related to:

1.) **Big Data 'crumbs,'** including access to, sharing, and interoperability of various climate data types; data reliability, representativeness and replicability; and existing scientific challenges related to climate research;
2.) **Big Data capacities**, both technical and human (including existing technical infrastructures as well as human and technological capacity gaps);
3.) **Big Data communities**, including political and governance issues; and ethical concerns.

3.1. Barriers and gaps related to Big Data 'crumbs'

Access to data 'crumbs'

An obvious barrier to using Big Data for resilience is access. Much of the 'crumbs' of Big Data remain on the servers of private companies. The question of opening up CDRs, in particular, has received a lot of attention. Grameenphone, the leading telecommunications provider in Bangladesh, shares its mobile call data records to understand climate impacts by mapping population flows before an after extreme weather events. The Mobile Data, Environmental Extremes and Population (MDEEP) Project is a partnership between the United Nations University Institute for Environment and Human Security (UNU-EHS), the International Centre for Climate Change and Development (ICCCAD), Flowminder.org, and Telenor Group.[89]

Box 3: Data Collaboratives: NOAA's Big Data Project

In February 2014, NOAA issued a request for information (RFI) to businesses and researchers in order to solicit ideas from the private sector on opening access to its 30 petabytes of annual environmental data sourced from satellites, radars, ships, aircraft, river gauges and weather forecast products. The February RFI focused on developing physical infrastructure for storing and sharing NOAA's data within a publicly accessible cloud that would allow industry experts and scientists to extrapolate vast amounts of historical data, schedule satellite launches, and make better predictions. After receiving 70 responses from individuals, academia and industry organizations, NOAA launched its Big Data Project (BDP), providing an opportunity for NOAA to partner and collaborate with infrastructure-as-a-service providers to expand access to its data resources for both private industry and the general public.

The fundamental question of legal data ownership and control will become more urgent in the next few years.[90] Private corporations protect their commercial interests and their reputations by refusing to share CDRs and similar data, despite their potential social value. Though some telecom operators, such as Orange, Telefónica and Grameenphone, have opened up data in a controlled and time-bound manner (through research partnerships, challenges, etc.), there are no comprehensive guidelines to facilitate such data sharing and ensure stability and predictability of access.

Accessing social media data can be challenging as well. Though many social media platforms such as Twitter and Facebook offer limited access to portions of their data through APIs (Application Program Interfaces), many allow only querying of an archive of past messages. Only a few platforms offer public, "real-time" streaming. Recognized as the most widely used source of social media data, Twitter[91] provides a random sample of 1% of all postings that can be filtered by keyword or location. One way of increasing the quantity of data pulled from Twitter in a disaster is to increase the number of terms being searched by, for example, using a disaster lexicon.[92] Another approach is to aggregate feeds from various sites to increase the volume of data being analysed.

Using new-generation nano-satellites and drones could provide greater access to satellite data through cheaper aerial imagery for hazard monitoring. They can also help establish distributed wireless sensor networks that would lower the costs of monitoring hazards by radically reducing infrastructure, management, and physical connectivity requirements. However, these UAVs also raise ethical and privacy questions as discussed below.

Data sharing and interoperability across sources

International cooperation should be engaged towards the definition of a coordinated global data collection, sharing and analysis. Surely interoperability among datasets does constitute a challenge; however, more pressing issues to be tackled involve political interference preventing data sharing as well as the push towards commercialization of data (even among public institutions of the same country), particularly in a moment when accountability and transparency have permeated the political discourse. Despite these difficulties, international initiatives — such as the Global Observing Systems Information Center (GOSIC), the Global Geodetic Observing System (GGOS) and the Global Earth Observation System of Systems (GEOSS) — have nevertheless been successfully implemented to produce and disseminate climate data.[93]

Box 4: Climate Data Partnerships

In March 2014, the Obama Administration launched the Climate Data Initiative with the goal of leveraging the government's data resources to stimulate innovation and entrepreneurship in climate change. As a part of the President's Climate Action Plan, the Climate Data Initiative launches a number of significant commitments from federal agencies, private-sector collaborators and research institutions to combat climate change through data-driven innovation. Participating companies and organizations (such as Facebook, PepsiCo, and Microsoft Research) have made commitments to creating new data platforms and organizing cross-sectoral collaborations in data analysis and visualization.

Data reliability and representativeness

A related and well-established set of challenges pertains to the analysis of and via Big Data – both in terms of internal and external validities.[94] Although strictly speaking, both kinds of validities refer to the ability to make and generalize *causal* claims, i.e. to the realm of *prescriptive* uses of Big Data, we expand their meaning and scope to include predictive and descriptive uses as well, i.e. the extent to which useful insights can be gleaned from these data.

The basic question researchers must always answer is, 'What does the data tell us?' Big Data doesn't magically answer this question; in fact, the answer is usually harder to find than in the case of controlled data collection. Tweets and text messages collected during or after an emergency may be deliberately misleading or false. More often, the challenge is finding the 'signal' in the 'noise.' Automated processes for dealing with large quantities of unstructured data try to answer the key question through careful extraction, verification, and classification

of data.

Another main challenge to analysis is statistical bias, which comes in several forms. One big problem during disasters is selection bias resulting from attrition. For example, in the aftermath of an earthquake, more phone signals are likely to come from less affected areas than from areas that have been devastated. Assessing need based on the number of phone signals may send aid to the wrong places. Selection bias highlights the need to verify conclusions, most likely by correlating additional data sources, before acting.

Another issue is sampling bias. Despite huge sample sizes, most Big Data sets are not representative. People's decisions to use the technology in the first place are largely determined by characteristics that affect the behaviours under study; self-selection biases the sample. For example, factors such as age and income affect whether people use mobile phones. Using CDRs to study, say, mobility patterns after a disaster may give a misleading picture because the sample – mobile users – skews younger and richer than the general population[95]—although in some cases it may.[96] These biases will tend to be greater with technologies that have lower penetration rates because there is then more room for highly skewed usage between different social groups.[97] It is unclear, nonetheless, whether adoption patterns and trends of newer and future technologies across groups will mirror those observed in the past. Techniques to correct for sampling bias use standard statistical models and methods to control for mobile or internet penetration rates in, for example, a given area or age group.[98] Refining such approaches requires calibrating new data with reliable target data from official or vetted sources. Even then, the ability to generalize the models and their results to other times and places is limited.

Similarly, predictive modelling based on a large number of variables (known as features) may have a very high predictive power in a given setting, but their data requirements and changes in the relationships at play will make them hard to replicate over time and space.

Scientific challenges in climate research

We should start by saying that the complexity of the climate system does not allow to simulate the behavior of the climate system with a deterministic approach. Gaps in our capacity to model the climate system are reflected into the variables we do not take into consideration to tackle climate change - i.e. data we (do not) collect and parameters we do not include into our models.

Broadly speaking, the degree of uncertainty linked with climate scenarios (and even more to the related impacts) should be taken into account not only before designing mitigation initiatives and developing adaptation solutions, but also before even beginning.[99] Strengthening communities' resilience in natural resources management through consideration of uncertainty proves to foster their resilience to climate shock as well. Equipping institutions and communities with capacity to monitor natural resources effectively constitutes a first starting point to face such challenges.

Scientific challenges are particularly linked with the use of remote sensing data. The uncertainty arising from retrieval algorithms that convert the electromagnetic signals from sensors to variable of interest also needs to be reduced. This can be done starting with a deeper evaluation of the quality of datasets that are used as inputs to these algorithms, such as the variety of different global land-cover datasets.

An important challenge is that sensors with both high spatial and temporal resolution are needed to measure climate processes happening at fine scales, such as trends in the properties of clouds and small-scale turbulence affecting heat transport. Small satellites (e.g. Skybox[100], Planet Labs[101]) may prove to be particularly useful to enable the observation of the same location over a given time interval, rather than seeking exhaustive coverage. Another key issue is the length of remote sensing data records. More work will be needed to assess the uncertainty of estimates of inter-annual variability, given that these datasets rarely cover more than 30 years. Furthermore, the satellite missions need to be sustained to ensure future availability of long-term datasets.

Another issue in developing climate-related research is represented by the way research data is shared. As Borgman (2012) wrote, "research data take many forms, are handled in many ways, using many approaches, and often are difficult to interpret once removed from their initial context. [Interoperability in] data sharing is thus a conundrum."[102] This issue seems particularly relevant to develop impact assessment and design related responses.

3.2. Barriers and gaps related to Big Data 'capacities'

Existing technical infrastructures and availability of cost-effective tools

At the technical level, an obvious difficulty in studying climate change and related impacts is that of processing large mass of spatial data (e.g. 5 TB per day from NASA, 25 PB per day from Google). The need for new tools to organize and analyze such data is clear.

Computational power to conduct in-depth assessment is also very expensive and thus constitutes a bottleneck. Very-high resolution images hold the potential of identifying many important objects, such as critical infrastructure (nuclear power plants, cooling towers), or distinguishing formal and informal settlements. However, there still are technical challenges to overcome to realize this potential. Indeed, per-pixel based methods do not work well for very high-resolution, requiring the use of methods that are currently computationally very expensive.[103]

The availability of cost effective and distributable tools to be used by those countries has been an issue. Software applications like Google Earth Engine have proven to be effective tools to perform scientific studies as well as disseminate findings to the general public.[104] An additional problem is the availability of adequate tools to allow visualization of the huge data sets allowing the development of climate change scenarios and related impacts. Such tools can be pivotal to inform policy makers and raise awareness on climate impacts among the public. Recent advances in data visualization include the ability to develop web-based

interactive maps to show data at the micro-level in the United States[1]. To do so, a 17TB raw dataset from NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) were channeled in a pipeline capable of interactively compare the projected impacts on rainfall and temperature due to low-emission scenarios and high-emission scenarios. Further expansion of the concept recently led to the development of a global map[2] as well.

Human and technological capacity gaps

Additional barriers to mainstreaming Big Data include local human and technological capacities. Generally, exposure to hazards, income, human capital and technological capacities tend to be correlated, both cross-sectionally and longitudinally. These relationships are not linear and straightforward, but apart from what may appear as exceptions—e.g. the oil-rich desert Gulf countries—places and people that are most exposed to sudden and slow onset natural hazards are typically those least able to leverage the opportunities of Big Data.

Limited data transfer capabilities (i.e. low bandwidth), a problem often experienced in low-income countries, will hinder the implementation of Big Data applications that rely on cloud computing. More generally few institutions in developing countries can afford the kinds of equipment needed to perform Big Data analytics and therefore the field is primarily dominated by top global universities and corporations. Low-income countries have only 1% of the world's capacity to transmit data via Internet and phone capacities. Progress is being made, however, as in the case of Africa, which has enjoyed a 20-fold increase in bandwidth between 2010 and 2015 thanks to submarine cables).[105]

Another limitation is human capacities, perhaps best exemplified by the well-known dearth of skilled staff in statistical offices in developing countries—cause and effect of a brain drain—for which Big Data may seem like a distraction. Several proxies, including standard literacy rates around the world, point to the skills gap. A popular analytics software such as R may be entirely free and open source, but other barriers limit its adoption.

This has obvious short-term effects—Big Data techniques will simply not be part of the potential toolkit of at-risk populations and their institutions. In the face of emergencies where time is of the essence, this simple fact may lead to decisions with detrimental long term effects: the parachuting or even distant engagement of external Big Data experts bypassing local structures.

For Big Data to have a significant and lasting positive impact, investing in technical and above all human capacities will of course be key. In recent months, the notion of 'data literacy' has received increasing attention—a welcome development. But it must be clear that enhancing data literacy, a key requirement for building resilience through Big Data, is not reducible to training world-class computer science PhDs in developing countries.
Technical barriers also separate subject-matter experts from data scientists. This can undermine the development of models leveraging big-data to assess climate change and plan

[1] See the website: http://climatedata.us/
[2] See: http://www.climateinternational.org/

adaptation and mitigation initiatives accordingly. Indeed, models are directly or indirectly developed by software specialists who may have a relatively-limited knowledge of the problems being modeled. The issues faced in downscaling techniques, allowing to bridge the global scale of climate change scenarios and the site-specific information required for impact assessment, are quite paradigmatic of such an issue: "statistical downscaling has received considerable attention from statisticians. Their contributions have, however, largely been unrecognized by the climate community, although they attempt to address important end user needs."[106]

For poor countries, a digital divide exists in the capacity to gather and analyze data for informed decision-making. A broad lack of capacity still exists among governmental staff to leverage big data, despite the fact "that governments are swimming in more data than they have ever had" as stated by Jon Gosier, D8A Group CEO.[107] This is particularly true for poorer nations, as Claire Melamed, Overseas Development Institute, stated: "the explosion of big data has far-outpaced our ability to make sense of it in poorer nations that already lack human and technical capacity." A strategic partnership should be sought with the governments and people of traditionally poor and less technically advanced countries.[108]

Financial constraints can impede adequate monitoring of climate parameters and climate-induced natural resources. Such difficulties can jeopardize the adequate monitoring of pivotal elements to develop adaptation plans - such as the water cycle with in-situ and satellite data.[109]

3.3. Barriers and gaps related to Big Data 'communities'

Politics and governance issues

A balance between data-informed decisions should always take into account expert knowledge, as human evaluation can be as powerful as machine-based decision-making (Heffernan, 2011). Mitigation and adaptation initiatives are shaped through comprehensive policies, rather than being the results of technical procedures, and the design of information architectures - which nowadays have been defined as "*politics in code*" (Stark, 2012)- can hugely impact the accuracy of the decision-making they have been built for.

The background approach that institutions employ to tackle climate change differs, and it strongly determines the way institutions and societies reacts to climate impacts. This happens for example in the way disasters induced by extreme climate events are modeled. Disasters are indeed defined as "a sudden, calamitous event that seriously disrupts the functioning of a community or society and causes human, material, and economic or environmental losses that exceed the community's or society's ability to cope using its own resources."[110] Basing disaster modeling on historical data can be particularly difficult as it hugely depends on the characterization of disaster itself: many small events (e.g. death for malnutrition or disease) indirectly induced by climate change tend not to be reflected in such accounts.[111]

Often in developing and emerging regions the political will and technical capacity to share data with neighboring countries can jeopardize the development of regional models and the planning of adequate countermeasures. At the same time, being climate change a long-term problem sometimes posing security issues can represent a powerful driver for cooperation even in regions affected by protracted crises. The Executive Action Team (EXACT-ME), a roundtable of institutions whose mandate is managing water resources in Israel, Jordan and Palestinian territories, has proven that institutions can effectively cooperate and share data on sensitive resources (i.e. water) to develop data-informed solutions to climate change.[112]

Broadly speaking, countries often lack adequate institutional data governance mechanisms to be able to fully leverage the potential of big data. Data is indeed rarely effectively shared among the whole set of stakeholders of a national sector, and even less rarely among international stakeholders - e.g. water authorities of a transnational basin. Moreover, data (often referred as to oil for a reason) does indeed represents a powerful asset within organizations- and it can be used as such within a single institutions. This is one of the reasons why typically data is not shared among different departments of the same administration, and the development of comprehensive information systems leveraging organizational data is particularly complex.[113]

Privacy concerns also make problematic to share and exchange data across organizations, especially when it can allow profiling citizens and/or organizations.[114] Such data can though be pivotal to adequately track the impact of mitigation initiatives as well as to drive recovery and adaptation plans. Both organizational and technical issues thus need to be addressed to allow the collection and sharing of such data while respecting the privacy of citizens and organizations.

Cooperation among public and private sector represents a crucial issue to be addressed to fully capitalize the potential of big data. Private companies hold data that can help monitoring GHG emissions as well as determining the impact of climate change on supply chains. Telecommunication infrastructures can directly help collecting weather data as well as supply proxy indicators to monitor weather behavior. Data from mobile phone activity of users in a region could provide accurate and timely estimates about population density, which is crucial element for monitoring climate vulnerabilities (Letouzé et al. 2015). Adequate mechanisms to integrate data from the private sector should thus be identified. The Carbon Disclosure Project (CDP) represents a valuable hub for climate data collected from the private sector, integrating data from 4500 corporations up to 2014 (CDP, 2014), even if the quality of the data on which CDP's reports are based had been questioned.[115]

Ethical concerns

Ethical concerns constitute one of the biggest challenges in developing the next generation of Big Data crisis tools. In recent years, for instance, the notion of 'anonymized' data has been torn to pieces; with sufficiently large connected datasets, whether from cell-phones[116] or credit card transactions[117], re-identification is almost always possible. Critically, this

development was hardly foreseeable a few years ago; by extension, it is nearly impossible to foresee future technological advances.

Another risk is that some leaders could consider Big Data to be the perfect technical fix for all of the world's problems, looking for a '30,000 feet view' that ignores the critical need for local engagement and investment.[118] Vague discussions about the ethical use of data, often reduced to 'anonymization', overlook deeper ethical considerations that ought to shape the future of Big Data for resilience, especially in complex and volatile contexts. Arguments over core aspects of data use such as the "right to be forgotten" are still unsettled and ethical norms are evolving quickly, necessitating a cautious and ethically conscious approach.

4 Recommendations for Making Big Data Useful and Meaningful for Social Resilience in Vulnerable Populations

A resilient community can and must do more than face one discrete disaster or hazard. It learns from past exposure to hazards about its own vulnerabilities. It increases its adaptive capacity to detect and monitor emerging climate change hazards and vulnerabilities – some of which may be caused by the recovery process itself – that may build up over time.[119] It then acts to reduce its vulnerability by improving the management of natural systems, strengthening infrastructure, and strengthening social networks. These activities are fundamentally political, requiring that members of the public have three attributes: awareness, capacity for collective action, and ability to weigh in on decisions.

As discussed earlier, mainstreaming Big Data for climate risk management requires an integrated approach involving multiple actors, both vertically and horizontally. Big Data can help decision makers to identify barriers to climate risk reduction, define the optimal positioning of levees and shelters, or design robust evacuation routes. Furthermore, these new sources of data have great potentially to increase *understanding* of social resilience and vulnerability in relation to climate change. For example, they can help decision makers see whether exposure maps and vulnerability indicators predict the actual impacts on the ground during the crisis.

Case studies that use Big Data to demonstrate the importance of social ties in a disaster[120] show how analysis of social responses to disasters can improve future assessments. Beyond academic studies, information about exposure and vulnerability, followed by data on disaster response and impacts, should feed into the next cycle of risk assessment and preparedness to refine the field's understanding of vulnerability and resilience.

Resilience requires that all actors who make decisions – from households to ministries – be empowered to understand the risks they face and act on them.[121] One of the four pillars of the UN's International Strategy for Disaster Reduction approach to people-centred early-warning systems is understanding risk,[122] because people tend to act on early warnings only if they already understand the risk. Social science research also shows that, in order to act together effectively, communities need a shared understanding of goals, risks, and options,

all of which requires data and analysis.[123] At a minimum, then, building resilience requires all citizens to have access to data, even those whoa re most vulnerable (see Box 1). Several global efforts aim to make large amounts of environmental and social data available to the public. They also aim to facilitate the integration of these data for use in risk assessment. However, information alone is not sufficient. Communities must be able to make sense of the data in a language they can understand in order to apply their knowledge and abilities appropriately. For that, they need analytical, organizational, and educational support.

Box 5: Case Studies: Challenges to Identifying Socially Vulnerable Groups

In the Chicago heat wave of 1995, black communities that were equally exposed to violence, similarly impoverished, and located in the same area experienced significantly different death rates (33 vs. 3 deaths for every 100,000 residents in one example), depending on how frequently their community members interacted with each other. Neighbors knowing each other from church, talking on the street, or meeting in local stores proved to be life saving (Kleinberg, 2003). When Katrina hit land, New Orleans had a relatively low elderly population. Sixteen percent of the city's residents were over 60 according to the 2005 US Census. Yet 75% of deaths from the hurricane were people in this age bracket (Louisiana Department of Health, 2006). While social vulnerability is intuitive as a concept and highly apparent when manifested in examples on the ground, it is challenging to analyze, quantify, and operationalize as a practical tool for disaster risk management and resilience.

Social vulnerability

Although vulnerability and resilience sciences have advanced immensely in the last two decades, the science - particularly for developing countries - lags considerably behind the science of the geophysical dimensions of hazards.[124] Even in developed countries there are still difficulties in recognizing socially vulnerable groups in disaster risk managements (see Box 2). Despite its conceptual and scientific limitations, integrating the social dimensions of hazards into the disaster cycle is necessary for fully successful emergency planning and response.[125]

The demographic, behavioral, and psychological dimensions of vulnerability fall along a spectrum of universality or generalizability; some dimensions are fairly well documented and consistent across geographies[126], while others vary significantly across time, place, and context. Starting with the most generalizable, people who have more financial resources, who are not especially young or old, and have strong community support are less vulnerable to climate change hazards. As in the Chicago heat wave example (Box 2), many studies have proven that social ties and social networks, primarily observed at the neighborhood or community level, are strong factors determining immediate loss and long-term recovery from disaster.[127]

Each dimension of social vulnerability requires individual consideration in scientific assessment, emergency response, recovery, and long-term adaptation, though there are no perfect antidotes for any dimension. An emergency plan that does not consider elderly

residents who have more difficulty evacuating or a recovery plan that does not account for poor residents with fewer resources to rebuild is incomplete, whether or not the plan has an impeccable biophysical assessment of the hazard. Likewise, climate adaptation plans that do not consider how marginalized communities are more at risk will clearly not be fully prepared for the impact of future hazards.[128]

Over the last two decades, social vulnerability researchers have begun to distill the dimensions of social vulnerability into empirically based indicators through analysis of Big Data sets. When combined in summary indices, typically using demographic information, these tools describe who is most vulnerable and where the most vulnerable are located before, during, and after a crisis.[129] If measured using benchmarks and monitored over time, these indicators may serve as diagnostic tools.[130]

Because the dimensions of vulnerability are highly interrelated, even interdependent, and vulnerabilities are often path dependent, not all indicators are equal—they change between different contexts. Measuring vulnerability through in-depth community level case studies is still the best approach for fully understanding vulnerability and for detecting the less generalizable indicators. Strictly country level indices for instance are generally not contextualized. On the other hand, detailed, community level indices are contextual, but place-specific and non-generalizable. No matter the approach, there are a few basic principles that should guide the creation of social vulnerability indicators at each step - research, quantification, and operationalization:

1. Each indicator should use appropriate unit analyses (*e.g.* individual, neighborhood, community), and no unit should stand alone as an indicator of vulnerability.

2. All indicators within an index should be weighted according to their influence on the overall vulnerability.

3. Indicators and indices must be designed, communicated, and used within the contexts that they are attempting to analyze.

In order to achieve these principles, the climate scientists must engage local citizens and governments in the creation and use of indices. Being participatory at each stage of research, decision-making, and planning helps ensure that the indices ask the right questions, incorporate local knowledge and the plans have local buy-in. Complete social vulnerability should be layered with appropriate economic, physical, infrastructural, and other assessments. Integration will require the creation of new layers within the index to account for feedback loops between social and natural systems. Finally, indices must be validated against actual damage data and questions of uncertainty resolved.[131]

Interpreting Climate and Disaster Data and Making Decisions

Understanding and integrating climate science into the perceptions of those who are not climate scientists can be difficult because the data involved is complex and impersonal. This lack of understanding of the data generates uncertainty and distrust of not only the data and the probability of future climate change disasters, but also the decisions made in disaster

preparedness. Resolving uncertainty requires us to tackle two main problems; quantifying the uncertainty of future outcomes and communicating the quantified uncertainties.[132]

For instance, people have a difficult time accurately interpreting probability, which is a key element of large data analysis, making them often relying on inappropriate heuristics to formulate counterproductive choices.[133] This assumption is valid for all non-expert audiences–from political decision makers to the general lay public—and it is particularly true when it comes to climate forecasts and related decisions that build on risk and uncertainty. Analytic products can provide decision makers with rational solutions to expected changes, but personal and collective experience, as well as emotions and cultural values, should be taken into account when designing climate-related initiatives.[134]

Therefore a key challenge is convey the data and its meaning in simple, personalized and concert manner so that people can integrate the information accurately and effectively into their actions to cope with climate changes. In this way, effective communication can shift perception of the problem, increase risk perception and inspire prudent behavior around mitigation, adaptation, and political/ behavior.

Particularly difficult is to share climate knowledge to remote and rural communities - who are likely to be hit first and foremost by climate impacts. An interesting case is that of communicating climate information to farmers, who are on the frontline of both climate change mitigation and adaptation initiatives.

To overcome the sound dissemination of climate information, FAO piloted Climate Field Schools—adapting its Farmer Field School approach[135]—to increase farmers' knowledge of climate and promote the adoption of weather forecasts and climate data to plan their work.[136] Particularly, the workflow behind CFS starts with the translation of climate information from scientific language into local language and then into farmers' language.[137] It is pivotal to shed light on the way farmers understand and respond to climate information,[138] as well as integrating indigenous knowledge to strengthen climate observations[139] and make communities cope with climate change.[140]

As we have seen over the last few years and most prominently during the recent Nepal earthquake, digital humanitarians have made great strides in collecting and analyzing real-time information to improve situational awareness of a developing crisis. In many cases, simply having a live map with as much information about damage and needs as possible is enough to empower an agency or an affected individual to make a better decision. However, more often than not, real value comes from translating this situational information into actionable messages and targeting it to the right individuals and groups.

These messages should be going to those who are affected, giving clear information about where and how to get help, evacuate, etc. They should be going to "first-responders" –the crowd of citizens ready to help, but who need to know how to most effectively do that and need to coordinate. And they should also be going to relief organizations, to assist in coordination as opposed to what has in the past often been a fragmented response. Meesters and van de Walle[141] conducted a very preliminary impact evaluation of the DHN's

information work in the 2010 earthquake in Haiti and compared it to the information system deployed by an on-the-ground NGO via more traditional means. They found that while DHN had quicker deployment, it was less responsive to the informational needs of affected vulnerable communities and first responders.

Creating Resilience Tools for Vulnerable Populations

Many believe that digital platforms and Big Data hold unique potential to design effective people-centered early-warning systems. The idea of people-centered early warning systems emerged out of the Hyogo Framework, and was described as "systems whose warnings are timely and understandable to those at risk (...) including guidance on how to act upon warnings". It emerged from the recognition that even if hazard detection capacities are good, warnings themselves do not reach all people at risk, and if they do, they are not necessarily clear or do not necessarily address people's concerns.[142] Furthermore, the aid received often does not address the most important needs that vulnerable individuals face at the time.[143]

Insofar as they are open, decentralized and interactive, Big Data and digital technologies could be key in increasing the ability of affected individuals to mobilize their social networks and obtain help from the 'crowd' of first-responders, as well as to coordinate the efforts of relief organizations. Indeed, Sarcevic et al.[144] showed that health workers in the Haiti earthquake were spontaneously attempting to coordinate their activities on Twitter, thereby revealing a latent demand for information products and platforms to facilitate this action and merge it with the crisis mapping work described above. The field of digital humanitarianism is currently experimenting with technologies to meet these organizational goals, but methods have not yet matured nor been validated. In addition, only a handful of relief organizations currently have incorporated social media monitoring and communication during mass emergency situations.[145]

Some cases already demonstrate the potential for Big Data and digital platforms to coordinate response. In the Russian wildfires of 2010, a group of digital volunteers used social media to coordinate volunteer fire fighters on the ground and provide relief assistance.[146] However, the approach taken is not easily scalable, as it relied on a full-time team, the staffing of a hotline and human processing of each message. One approach to automating coordination is to use algorithms to mine tweets or other social media posts to match spontaneous requests for help and offers of help.[147] This approach is challenging because many tweets are unstructured –which lowers classification performance –and because new classifiers need to be built for different situations.

Nonetheless, some promising results have been obtained during typhoon Yolanda in the Philippines in 2013. Another approach is to develop a specific application that structures the dialogue between affected individuals, volunteer responders, as well as the overall 'supply-chain' of relief organizations.[148] For such applications to truly meet the needs of vulnerable communities, vulnerable individuals themselves should be able to shape them. To this end, a team of engineers from MIT, Rensselaer Polytechnic Institute and Qatar Computing Institute are creating an application development platform that would make it easy for non-developers to build disaster-response applications, specifically apps to match donations and

requests and apps to report on the ground information. This platform will allow the reuse, extension and integration of heterogeneous structured data from distributed sources, including public datasets, thus building on the innovations reviewed earlier.[149]

The next generation of Big Data crisis management tools should be built with and around different communities' needs and need to be tested and evaluated in light of operational goals. Hughes[150] describes participatory methods for the design of such tools. We argue that in developing these tools, it is essential to take into account what we know about the barriers to understanding in non-expert communities and appreciate immediate risk. How do people process the physical reality of a hazard and how is this modified by the availability of large amounts of complex data? Answering these questions is essential to developing effective communication strategies.

Hurricane Sandy in the fall of 2012 in the US clearly demonstrates the pivotal role of outreach. The storm was accurately forecasted as an unprecedented threat to the Northeastern seaboard days before it hit landfall, a threat that would eventually lead to billions of dollars in damage and 157 deaths.[151] Governmental and other responding institutions attempted to communicate the risk and ordered emergency actions.[152] Yet the public response during the disaster was confused and slow. Simply presenting conclusive data and prescribing a response based on ever more precise and complex data is insufficient for disaster response. A review of Sandy concluded that progress in forecasting may have reached a point of diminishing returns, and that the critical need now is progress in risk communication.[153]

Awareness is key to resilience building in any context. It is here where Big Data approaches offer unique opportunities for disseminating and communicating at scale. In particular, crisis management tools allow for people to be accessed directly at high speed. However, in order to capitalize on these opportunities a number of factors are required, such as: adequate technical capacity to design suitable dissemination platforms and tools as well as considered thought as to the framing delivery mechanism. Above all, clear standardized guidelines for the dissemination of disaster-related information are crucial in order to facilitate rapid, safe and ethical release of data, particularly in the context of emergencies for vulnerable communities.

Conclusion

In this paper, we explore the intersection among Big Data, Climate Risk, and Resilience and identify Big Data's contributions towards a climate risk management framework in monitoring climate and spotting climate-driven hazards, assessing exposure and vulnerabilities to climate change in developing countries, and finally improving both adaptation and mitigation strategies.

Big Data is key to determining climate baselines and identifying vulnerabilities to climate-induced hazards through 'crumbs' (the first 'C'). As SRS data, CDRs, geospatial data, etc., Big Data sources can contribute to monitoring and understanding the drivers of climate change; detect variations that can lead to negative impacts; validate climate models; help to describe the main determinants of exposure; assess the impact of climate change on environmental systems and societies; and assess vulnerabilities within communities. In building analytical, organizational and educational support through Big Data, vulnerable populations are able to access and make sense of climate data in a language they can understand in order to apply their knowledge and abilities appropriately, thus improving their adaptive capacities (the second 'C') to climate stressors.

An effective implementation of climate risk management will affect and involve several different sectors within the government (including infrastructure, health, natural resources, agriculture, and water management), and therefore, would provide an effective means to improve the impact of development efforts as the result of coordinated actions by governmental and nongovernmental actors. Through its discursive function, Big Data as an ecosystem has the potential to foster awareness-raising and institutional development by feeding risk assessment tools coupled with economic analysis, towards the implementation and incorporation of climate risk management into the priority agendas of countries and regional entities.

Another key element for climate risk to be effectively managed is the broad involvement of society and sectors (the third 'C') in climate change adaptation, and disaster risk management initiatives, which should be both considered as facets of a comprehensive climate management risk approach. Big data technologies can help establish a conducive environment for climate risk management and affords the ability to identify priorities and barriers to climate risk reduction by offering a unique, dynamic birds-eye view, by fostering an inclusive process of bottom-up engagement with tailored top-down planning.

Such an integrated approach would also support the mainstreaming of climate risk management, both vertically and horizontally. On one hand, data would help integrate climate risk management into development planning as part of sectoral, national, and regional development planning, and engage the full spectrum of government operations. On the other hand, big data would offer the opportunity for high-level coordination through the collaboration of institutions and entities at different levels from all the major sectors of economies.

In the context of climate change, Big Data as an ecosystem encourages communities to endure, adapt, and transform in response to current stressors and future shocks—to become resilient. Resilient communities not only prepare for the potential of specific climate disturbances (specific resilience), but also a whole array of hazards they had not previously encountered (general resilience). Additionally, resilient communities are able to bridge the gap between Big Data's potential to identify and monitor climate hazards and empower vulnerable populations to weigh in on decisions.

Endnotes

[1] UN System Task Team on the Post-2015 UN Development Agenda, "Disaster Risk and Resilience," UNISDR and WMO, 2012 http://www.un.org/millenniumgoals/pdf/Think%20Pieces/3_disaster_risk_resilience.pdf, p. 7

[2] Gunderson, Larry, "Ecological and Human Community Resilience in Response to Natural Disasters," Ecology and Society, 2010, http://www.ecologyandsociety.org/vol15/iss2/art18/

3 UN System Task Team on the Post-2015 UN Development Agenda, "Disaster Risk and Resilience," UNISDR and WMO, 2012, http://www.un.org/millenniumgoals/pdf/Think%20Pieces/3_disaster_risk_resilience.pdf, p. 7

[4] Letouzé, 2014 and 2015

5 See Pentland, 2012

[6] Yang et al. 2012

[7] Brown et al. 2012

[8] Chikhaoui et al. 2005 ; McBratney et al. 2006

[9] Mu et al. 2011 ; Rodriguez-Fernandez et al. 2014

[10] Pesaresi et al., 2011, JRC 2014

[11] Pelling and High, 2005; Tellman, 2010

[12] Source to be added

[13] Shelton et al., 2014; Kogan, Palen, and Anderson, 2015.

[14] Milliken, "Resilience: From Metaphor to an Action Plan", p. 6

[15] Sources?

[16] https://www.ipcc.ch/publications_and_data/ar4/syr/en/mains1.html

[17] Carpenter, Stephen R. "Resilience to Cope with Extreme Events." Sustainability, 2012, 4, 3248-3259; doi:10.3390/su4123248

[18] Walker, B., and F. Westley. 2011. Perspectives on resilience to disasters across sectors and cultures. *Ecology and Society* 16(2): 4. [online] URL: http://www.ecologyandsociety.org/vol16/iss2/art4/

[19] Yang, Jun, et al. "The role of satellite remote sensing in climate change studies." Nature climate change 3.10 (2013): 875-883.

[20] Christy, J. R., Spencer, R. W., & Norris, W. B. (2011). The role of remote sensing in monitoring global bulk tropospheric temperatures. International journal of remote sensing, 32(3), 671-685

[21] Leuliette, E. W. (2011). The Balancing of the Sea-Level Budget. Current Climate Change Reports, 1-7.

[24] 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.

[25] Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in ecology & evolution, 20(9), 503-510.

[26] Kennedy et al. 2010

[27] Garbulsky, M. F., Peñuelas, J., Gamon, J., Inoue, Y., & Filella, I. (2011). The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies: a review and meta-analysis. Remote Sensing of Environment, 115(2), 281-297.

[28] Brown, M. E., K. M. De Beurs, and M. Marshall. "Global phenological response to climate change in crop areas using satellite remote sensing of vegetation, humidity and temperature over 26years." Remote Sensing of Environment 126 (2012): 174-183.

[29] Palm, C., Sanchez, P., Ahamed, S., & Awiti, A. (2007). Soils: A contemporary perspective. Annu. Rev. Environ. Resour., 32, 99-129

[30] Chikhaoui, M., Bonn, F., Bokoye, A. I., & Merzouk, A. (2005). A spectral index for land degradation mapping using ASTER data: Application to a semi-arid Mediterranean catchment. International Journal of Applied Earth Observation and Geoinformation, 7(2), 140-153.

[31] McBratney, A. B., Minasny, B., & Rossel, R. V. (2006). Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. Geoderma, 136(1), 272-278.

[32] Mu, Qiaozhen, Maosheng Zhao, and Steven W. Running. "Improvements to a MODIS global terrestrial evapotranspiration algorithm." Remote Sensing of Environment 115, no. 8 (2011): 1781-1800.

[33] Rodriguez-Fernandez, N., Richaume, P., Aires, F., Prigent, C., Kerr, Y., Kolassa, J., ... & Mahmoodi, A. (2014, July). Soil moisture retrieval from SMOS observations using neural networks. In Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International (pp. 2431-2434). IEEE.
[34] Meier, P. (2014). Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response. CRC Press.
[35] Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., ... & Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. Frontiers in Ecology and the Environment,10(6), 291-297.
[36] Root et al. 2003
[37] Crimmins, T. M., Crimmins, M. A., & David Bertelsen, C. (2009). Flowering range changes across an elevation gradient in response to warming summer temperatures. Global Change Biology, 15(5), 1141-1152
[38] See http://www.tomnod.com
[39] Newman, G., Graham, J., Crall, A., & Laituri, M. (2011). The art and science of multi-scale citizen science support. Ecological Informatics, 6(3), 217-227.
[40] Kelling, S., Hochachka, W. M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., & Hooker, G. (2009). Data-intensive science: a new paradigm for biodiversity studies. BioScience, 59(7), 613-620
[41] Web site: https://rfcx.org/about
[42] Deville, P., C. Linard, S. Martin, M. Gilbert, F.R. Stevens, A.E. Gaughan, V.D. Blondel and A.J. Tatem, 2014. Dynamic population mapping using mobile phone data. PNAS, vol. 111, no. 45: 15888–15893.
[43] Cutter, S. L. "Vulnerability to Environmental Hazards." Progress in Human Geography 20 (1996): 529–539.
[44] Pelling, Mark and High, Chris 2005). Understanding adaption: What can social capital offer assessments of adaptive capacity? Global Environmental Change, 15(4). Tellman, 2010.
[45] Tate, E., 2012. Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. Natural Hazards. 63: 325-347 doi: 10.1007/s11069-012-0152-2
[46] Blumenstock, J.E. 2014. Calling for better measurement: Estimating an individual's wealth and wellbeing from mobile phone transaction records. Proceedings of Knowledge Discovery in Data.
[47] UN Global Pulse, 'Using Mobile Phone Data and Airtime CreditPurchases to Estimate Food Security', Global Pulse ProjectSeries no. 14, 2015.
[48] Frias-Martinez, Vanessa, Virsesa, J. (2012) On the relationship between socioeconomic factors and cell phone usage, International Conference on Information and Communication Technologies and Development.
[49] Smith-Clarke, Christopher, Afra Mashhadi, and Licia Capra. "Poverty on the cheap: Estimating poverty maps using aggregated mobile communication networks." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 511-520. ACM, 2014.
50 Flowminder project website: http://www.flowminder.org/
51 WorldPop project website: http://www.worldpop.org.uk/
52 Shelton, Taylor, Ate Poorthuis, Mark Graham, and Matthew Zook. "Mapping the data shadows of Hurricane Sandy: Uncovering the sociospatial dimensions of 'big data'." Geoforum 52 (2014): 167-179.
53 Kogan, Marina, Leysia Palen, and Kenneth M. Anderson. "Think Local, Retweet Global: Retweeting by the Geographically-Vulnerable during Hurricane Sandy." In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pp. 981-993. ACM, 2015.
54 http://climatecolab.org/
55 Smith, C., Mashhadi, A., and Capra, L. 2013. Ubiquitous Sensing for Mapping Poverty in Developing Countries. Orange Data for Development Challenge, NetMob 2013. 1-3 May 2013, Cambridge (MA), United States.
56 Onnela, J.-P., and D. Lazer. 2009. "The Evolving Co-contribution Network ofthe U.S. Electoral System." Paper presented at the Political NetworksConference II, Harvard University.
57 Eagle, N., Pentland, A.S., and Lazer, D. 2009. Inferring friendship network structure by using mobile phone data. PNAS, vol. 106, no. 36, pp. 15274–15278.
58 Eagle, N., Macy, M., Claxton, R. 2010. Network Diversity and Economic Development. Science DOI: 10.1126/science.1186605, 1029; 328.
59 http://www.hubbardresearch.com/cigar-global-intervention-decision-model/
60 Farhangi, H. (2010). The path of the smart grid. Power and Energy Magazine, IEEE, 8(1), 18-28.
61 Attari, S. Z., Gowrisankaran, G., Simpson, T., & Marx, S. M. (2014). Does information feedback from in-home devices reduce electricity use? Evidence from a field experiment (No. w20809). National Bureau of Economic Research.
62 Darby, S. (2010). Smart metering: what potential for householder engagement?. Building Research & Information, 38(5), 442-457.

63 Palensky, P., & Dietrich, D. (2011). Demand side management: Demand response, intelligent energy systems, and smart loads. Industrial Informatics, IEEE Transactions on, 7(3), 381-388.
64 Toole, J. L., Colak, S., Alhasoun, F., Evsukoff, A., & Gonzalez, M. C. (2014). The path most travelled: Mining road usage patterns from massive call data. arXiv preprint arXiv:1403.0636
65 Berlingerio, M., Calabrese, F. Di Lorenzo, G., Nair, R. , Pinelli, F. , Sbodio, M "AllAboard: a System for Exploring Urban Mobility and Optimizing Public Transport Using Cellphone Data".
66 IBM (Aug 01, 2007). IBM and Singapore's Land Transport Authority Pilot Innovative Traffic Prediction Tool. Retrieved from http://www-03.ibm.com/press/us/en/pressrelease/21971.wss
67 IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151.
68 Seto K. C., S. Dhakal, A. Bigio, H. Blanco, G. C. Delgado, D. Dewar, L. Huang, A. Inaba, A. Kansal, S. Lwasa, J. E. McMahon, D. B. Muller, J. Murakami, H. Nagendra, and A. Ramaswami, 2014: Human Settlements, Infrastructure and Spatial Planning. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
69 Hansen, Matthew C., Peter V. Potapov, Rebecca Moore, Matt Hancher, S. A. Turubanova, Alexandra Tyukavina, David Thau et al. "High-resolution global maps of 21st-century forest cover change." science 342, no. 6160 (2013): 850-853.
70 Patel, N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, G., Tatem, A., Trianni, G. 2015. Multitemporal settlement and population mapping from Landsat using Google Earth Engine. International Journal of Applied Earth Observation and Geoinformation . 35: 199–208 http://dx.doi.org/10.1016/j.jag.2014.09.005
71 See CLASlite success stories: http://claslite.carnegiescience.edu/en/success-stories/borneo-palm-oil-and-carbon-analysis.html
72 See imazon.org.br/mapas/?lang=en
73 Aklin, M., & Urpelainen, J. (2013). Political competition, path dependence, and the strategy of sustainable energy transitions. American Journal of Political Science, 57(3), 643-658.
74 Sabatier, P. A., & Weible, C. (Eds.). (2014). Theories of the policy process. Westview Press.
75 Pesaresi, M., Ehrlich, D., Caravaggi, I., Kauffmann, M., & Louvrier, C. (2011). Toward global automatic built-up area recognition using optical VHR imagery.Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 4(4), 923-934.
76 Pesaresi et al., 2011, JRC 2014.
77 Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: a case study of the Haitian earthquake. Available at SSRN 2216649n-turner.htm
78 Haklay, M., Basiouka, S., Antoniou, V., & Ather, A. (2010). How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. The Cartographic Journal, 47(4), 315-322.
79 Atzberger, Clement. "Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs." Remote Sensing 5, no. 2 (2013): 949-981.
80 Brakenridge and Anderson, 2006
81 Herold, Christian, and Roberto Rudari. "Improvement of the Global Flood Model for the GAR 2013 and 2015."
82 de Sherbinin, Alex. "Climate change hotspots mapping: what have we learned?." Climatic Change 123, no. 1 (2014): 23-37.
83 Warner, Koko, Kees van der Geest, Sönke Kreft, Saleemul Huq, Sven Harmeling, Koen Kusters, and Alex de Sherbinin. Evidence from the frontlines of climate change: loss and damage to communities despite coping and adapation. UNU-EHS, 2012.
84 Kumar, R. 2014. "How Youth Saved Bananas in Uganda". Blog post. World Bank http://blogs.worldbank.org/youthink/how-youth-saved-bananas-uganda
85 Jiménez, D., Cock, J., Satizábal, F., Barreto, M., Pérez-Uribe, A., Jarvis, A. and VanDamme, P.,(2009). Analysis of Andean blackberry (Rubus glaucus) production modelsobtained by means of artificial neural networks exploiting information collected by small-scale growers in Colombia and publicly available meteorological data Computers andElectronics in Agriculture. 69 (2): 198–208.

Cock, James, Thomas Oberthür, Camilo Isaacs, Peter Roman Läderach, Alberto Palma, Javier Carbonell, Jorge Victoria et al. "Crop management based on field observations: Case studies in sugarcane and coffee." Agricultural Systems 104, no. 9 (2011): 755-769.

Jiménez, D., Cock, J., Jarvis, A., Garcia, J., Satizábal, H.F., Van Damme, P., Pérez-Uribe, A.andBarreto-Sanz, M.,(2011).Interpretation of Commercial Production Information: Acasestudy of lulo (Solanum quitoense), an under-researched Andean fruit.Agricultural Systems.104(3):258-270

86 Gupta, Arnab, Jacob van Etten, Sarika Mittra, Prem Mathur. 2015. "Broadening the genetic base and knowledge of cultivated crops for climate adaptation: A citizen science approach." Our Commonfuture in Climate Change, UNESCO, Paris, France.

87 See on Datamation: http://www.datamation.com/applications/hadoop-makes-a-big-data-splash-1.html

88 See website of Ignitia: http://www.ignitia.se/

[89] http://www.ehs.unu.edu/article/read/press-note-new-project-uses-mobile-data-to-understand-climate

[90] For a review and discussion, see Letouzé, Vinck and Kammourieh, 2015

[91] Imran et al. 2014a

[92] Olteanu et al., 2014

[93] Yang, Jun, et al. "The role of satellite remote sensing in climate change studies." Nature climate change 3.10 (2013): 875-883.

[94] Internal validity refers to the degree to which a causal relationship can be established between two variables in a study, technically speaking, "internal validity exists if the observed effects of the independent variable on the dependent variable are real and not caused by extraneous factors"; external validity on the other hand refers to the degree to which the findings can be generalized to the general population outside of the sample; technically speaking it is "the ability to generalize the study results to other groups and settings beyond those in the current experiment". http://www.flashcardmachine.com/internal-vsexternal-validity.html

[95]"Sampling bias is systematic error due to a non-random sample of a population, causing some members of the population to be less likely to be included than others, resulting in a biased sample, defined as a statistical sample of a population (or non-human factors) in which all participants are not equally balanced or objectively represented. (S)ampling bias (…) undermines the external validity of a test (the ability of its results to be generalized to the rest of the population)." Source: Wikipedia

[96] For instance, as noted in Letouzé, Pestre and Zagheni, 2015, 2015, "asking 100 billionaires whether they would prefer to have dinner with friends or be sent to jail would yield results that would look very similar to those found in the general population, because income is not correlated with aversion to jail." Studying the impact of biases in mobile phone ownership on estimates of human mobility in Kenya, Wesolowski et al (2013) for example found that "that mobility estimates are surprisingly robust to the substantial biases in phone ownership across different geographical and socioeconomic groups".

[97] In other words, data from a technology that is used by 99% of the population are more likely to be representative of the general population than from one that is used by 1% of the population, although data from a representative sample of 0.1% the population will be more by design representative that the data from the 99% of the population described above.

[98] See Zagheni and Weber, 2012, Letouzé, Zagheni, Pestre et al, 2015

[99] (Tompkins and Ager, 2004)

[100] See website: http://www.skyboximaging.com/

[101] See website: https://www.planet.com/

[102] Borgman, Christine L. "The conundrum of sharing research data." *Journal of the American Society for Information Science and Technology* 63.6 (2012): 1059-1078.

[103] Vatsavai, Ranga Raju, et al. "Spatiotemporal data mining in the era of big spatial data: algorithms and applications." *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*. ACM, 2012.

[104] Schwarz et al., 2015

[105] http://qz.com/420219/in-just-5-years-submarine-cables-have-brought-a-20-fold-increase-in-bandwidth-in-africa/

[106] (Knutti et al. 2010)

[107] (Scidev, 2014)

[108] (Schwarz et al. 2015).

[109] Hegerl, 2014

[110] IFRC, 2015

[111] (Hilbert, 2013).

[112] (Jägerskog, 2007).

[113] Fedorowicz et al., 2010
[114] (Otjacques et al., 2007).
[115] Lippert, 2013).
[116] http://www.nature.com/srep/2013/130325/srep01376/full/srep01376.html
[117] http://www.sciencemag.org/content/347/6221/536.full
[118] See notably boyd and Crawford, 2011
[119] Kates et al. 2006
[120] Blumenstock et al. 2014; Bengtsson et al. 2011
[121] Carpenter et al. 2012; Tompkins and Adger, 2004
[122] ISDR, 2005
[123] Dietz, 2003; Ostrom, 2005
[124] Cutter, 2003).
[125] (U.S. National Research Council, 2012)
[126] Cutter, 2003),
[127] (Pelling and High, 2005; Tellman, 2010).
[128] Adger, 2010
[129] Tate, 2012
[130] The more contextual dimensions of vulnerability, the third category presented above, do not lend themselves to generalizable proxies like census data and are therefore cannot be full captured by a quantitative index.
[131] Tate, 2012
[132] Webster, 2003
[133] Patt et al., 2005.
[134] Marx et al., 2007.
[135] Berg, 2004
[136] FAO, 2010
[137] Motha et al, 2004).
[138] Roncoli, 2006),
[139] (Cruikshank, 2005)
[140] Kihupi et al., 2003).
[141] (Meesters and Van de Walle 2013)
[142] UNISDR, 2006).
[143] State of Humanitarian Aid, OCHA, 2015).
[144] (Sarcevic et al 2012)
[145] Imran et al. 2014a).
[146] Asmolov, 2014).
[147] Puhorit et al. 2014).
[148] World Disasters Report, 2013).
[149] (Shih et al. 2013)
[150] Hughes 2014
[151] NOAA/NWS 2013).
[152] (Freeman 2013).
[153] Marlon et al., 2015).

References

Adger, W.N., 2006. Vulnerability. Glob. Environ. Change 16, 268–281.

Adger, N., "Social Capital, Collective Action, and Adaptation to Climate Change**."** Der Klimawandel**:** Sozialwissenschaftliche Perspektiven. Ed. Martin Voss. Wiesbaden: VS Verlag für Sozialwissenschaften. 2010. *327-345*.

Adger, W.N., Barnett, J., Brown, K., Marshall, N., O'Brien, K., 2013. Cultural dimensions of climate change impacts and adaptation. Nature Climate Change 3: 112–117. doi:10.1038/nclimate1666

Bolin, B., 2007. Race, Class, Ethnicity, and Disaster Vulnerability, in: Handbook of Disaster Research, Handbooks of Sociology and Social Research. Springer New York, pp. 113–129.

Cutter, S.L., Boruff, B.J., Shirley, W.L., 2003. Social Vulnerability to Environmental Hazards. Soc. Sci. Q. 84, 242–261. doi:10.1111/1540-6237.8402002

Fothergill, A., 1996. Gender, risk, and disaster. Int. J. Mass Emergencies Disasters 14: 33–56.

Fothergill, A., Maestas, E.G.M., Darlington, J.D., 1999. Race, Ethnicity and Disasters in the United States: A Review of the Literature. Disasters 23: 156–173.

Pimm, S. L. 1991. The balance of nature? University of Chicago Press, Chicago. doi:10.1111/1467-7717.00111

Fothergill, A., Peek, L., 2004. Poverty and Disasters in the United States: A Review of Recent Sociological Findings. Nat. Hazards 32: 89–110. doi:10.1023/B:NHAZ.0000026792.76181.d9

Gallopín, G., 2006. Linkages between vulnerability, resilience, and adaptive capacity. Glob. Environ. Change 16: 293–303. doi:10.1016/j.gloenvcha.2006.02.004

Gray, C., Frankenberg, E., Gillespie, T., Sumantri, C., Thomas, D., 2014. Studying Displacement After a Disaster Using Large-Scale Survey Methods: Sumatra After the 2004 Tsunami. Ann. Assoc. Am. Geogr. 104: 594–612 **doi:**10.1080/00045608.2014.892351

doi:10.1080/00045608.2014.892351

Hahn, M., Riederer, A., Foster, S., 2009. The Livelihood Vulnerability Index: A pragmatic approach to assessing risks from climate variability and change—A case study in Mozambique. Glob. Environ. Change 19: 74–88 doi:10.1016/j.gloenvcha.2008.11.002

Holling, C. S. 1973. Resilience and stability of ecological systems. Annual Review of Ecological Systems 4: 1–23.

IPCC, 2014: Summary for Policymakers. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32.

Jonkman, S., Kelman, I., 2005. An analysis of the causes and circumstances of flood disaster deaths. Disasters, 29(1): 75−97. © Overseas Development Institute, 2005

Klineberg, E. 2003. Heat Wave: A Social Autopsy of Disaster in Chicago. Chicago Press

Louisiana Department of Health and Hospitals. 2006. Reunited the Families of Katrina and Rita: Louisiana Family Assistance Center. http://dhh.louisiana.gov/office/publicans/pubs-303/Full%20Report.pdf. Accessed 10/27/06

Miller, F., Osbahr, H., Boyd, E., Thomala, F., Bharwani, S., Ziervogel, G., Walker, B., Birkmann, J., van der Leeuw, S., Rockström, J., Hinkel, J., Downing, T., Folke, C., Nelson, D., 2010. Resilience and Vulnerability: Complementary or Conflicting Concepts?. Ecology and Society 15(3): 11.

Neumayer, E., Plümper, T., 2007. The Gendered Nature of Natural Disasters: The Impact of Catastrophic Events on the Gender Gap in Life Expectancy, 1981–2002. Ann. Assoc. Am. Geogr. 97: 551–566. doi:10.1111/j.1467-8306.2007.00563.x

U.S. National Research Council. 2012. Disaster Resilience: A National Imperative. Washington D.C.: The National Academies.

Ngo, E., 2001. When Disasters and Age Collide: Reviewing Vulnerability of the Elderly. Nat. Hazards Rev. 2: 80–89. doi:10.1061/(ASCE)

Pelling, M., High, C., 2005. Understanding adaptation: What can social capital offer assessments of adaptive capacity?. Glob. Environ. Change Part A 15: 308–319 doi:10.1016/j.gloenvcha.2005.02.001

Tellman, Elizabeth. 2011. "Community Resilience and Hurricane Ida: How Marginalized Salvadorans Lacking NGO and Governmental Support Cope with Climate Shock." Climate Change and Migration: Rethinking Policies for Adaptation and Disaster Risk Reduction. UNU-EHS (United Nations University Environment and Human Security). SOURCE issue 15. Pp. 32-46.

Tellman, B., Alaniz, R., Rivera, A. and Contreras, D., 2014. Violence as an obstacle to livelihood resilience in thecontext of climate change. UNU-EHS Working Paper Series, No.3. Bonn: United Nations University Institute of Environment and Human Security (UNU-EHS).

Turner, B., 2010. Vulnerability and resilience: Coalescing or paralleling approaches for sustainability science? Global Environmental Change. 20: 570–576 doi:10.1016/j.gloenvcha.2010.07.003

Werg, J., Grothmann, T., Schmidt, P., 2013. Assessing social capacity and vulnerability of private households to natural hazards – integrating psychological and governance factors. Nat Hazards Earth Syst Sci 13: 1613–1628. doi:10.5194/nhess-13-1613-2013