

Chapter 15

Sensing Change: Measuring Cookstove Adoption with Internet-of-Things Sensors



Daniel L. Wilson

1 The Challenge of Measuring Impact

Practitioners of development engineering aspire to create new technologies, and the goal of these new technologies is to have a positive impact for people in need. Perhaps the technology affords cleaner water, better governance, faster medical diagnosis, or easier access to public services. Regardless of the technology, we became development engineers because we want our work to have positive impacts.

In this chapter, I'll discuss cleaner cooking appliances that were intended to reduce exposure to air pollution and the drudgery associated with traditional wood-fired cooking methods. I'll discuss the trials and tribulations of creating new kinds of sensor technologies to make it easier to measure cookstove adoption. This process spans the course of about 8 years, three major technology variants, and thousands of households in dozens of countries. In the beginning, I'll discuss the challenges of collecting sensor data using clunky industrial data loggers in Sudan. Then, we will move onto a first round of custom-built data loggers deployed on advanced cookstoves in India. Then, in a four-country trial with thousands of households participating, we will explore how an entirely custom Internet of Things (IoT) sensor, survey, and analytics platform were used to address many of the issues that arose in the first two experiments. Finally, I'll share how this cookstove sensor technology was spun off into a startup sensor company. But first, why all this fuss about measuring cookstoves?

Cooking is a critical global development challenge because roughly three billion people rely on traditional wood- or other solid-fueled fires to cook their daily meals (The World Bank, 2011). The smoke from this practice is one of today's greatest

D. L. Wilson (✉)
Geocene Inc., Berkeley, CA, USA
e-mail: danny@geocene.com

© The Author(s) 2023
T. Madon et al. (eds.), *Introduction to Development Engineering*,
https://doi.org/10.1007/978-3-030-86065-3_15

environmental health risk factors and is responsible for some 1–4 million premature deaths per year (Forouzanfar et al., 2015; Lim et al., 2013). Cleaner cookstove technologies could positively affect this situation, but we will discuss that more in later sections.

What is impact, and how do we measure it? This is a major topic that development engineers are grappling with today (Thomas & Brown, 2020). Suppose you have developed a new novel cookstove technology that reduces emissions by 95%. If that cookstove is only used by a few dozen farmers as part of your research project, did it have an impact? Alternatively, suppose you make a tiny improvement to a traditional earthen cookstove; your innovation drives down emissions by 5%, and you estimate that 10 million cookstoves are built per year using your technology. Did your small improvement on a traditional technology have an impact?

I would like to propose a framework to think about impacts throughout this case study. That is, impact is proportional to the product of performance, adoption, and scale:

$$\text{Impact} \propto \text{Performance} \times \text{Adoption} \times \text{Scale}$$

$$I \propto PAS$$

This specific framework, which is pronounced “IPAS” (people are better at remembering acronyms that have to do with beer), is useful for remembering how multiple, often conflicting, considerations must be taken into account when measuring impact. For the purposes of this chapter, we can think of the definition of each of these variables as follows:

- **Impact** is the realized positive global utility that exists as a result of the existence of the technology.
- **Performance** is the marginal difference in outcome between the intervention state, on an individual product or person level, and the status quo. For example, for cooking, a good metric of performance might be the difference in personal relative risk of disease between breathing smoke emissions cooked on a traditional and cleaner cookstove.
- **Adoption** is the average degree to which the user of an individual unit of technological intervention (e.g., a single cookstove) utilizes that intervention per unit time. For example, this could be the average number of cleaner cookstove uses per week. The study of adoption is part of a broader category of development practice activities often referred to as “monitoring and evaluation.”
- **Scale** is the number of users of your intervention.

Of course this model is a gross oversimplification. It is important to remember that, despite the linear simplicity of the IPAS framework, many real-world systems have nonlinear relationships. For example, if a cookstove that reduces emissions by 50% will not reduce negative health impacts by 50% in the same way that cutting a smoking habit from two packs per day to a single pack per day will not reduce lung cancer risk by 50% (Burnett et al., 2014). The IPAS model also ignores important considerations such as disadoption of harmful baseline practices, impacts of training

and time on performance, temporal changes in scale, and second-order effects such as how adoption of one technology may lead to virtuous second-order adoption of more beneficial technologies (adopting solar panels leads to adopting lighting that leads to better educational outcomes) or negative second-order effects (adopting microwave cooking leads to consumption of less healthy foods) (Pillariseti et al., 2014).

Despite the drawbacks of this simplified model, it does emphasize two aspects of technological interventions that engineers are notorious for downplaying: adoption and scale. As engineers, many of us are only trained to improve and measure performance, not adoption or scale. Whether designing a cookstove, water filter, generator, or other intervention technology, we tend to be very diligent about quantifying the performance of our interventions. The desire to measure and report metrics such as efficiency, power, and throughput comes naturally to many engineers. However, it is important to remember that a great-performing intervention that people do not like or that just a few people use will not have a significant impact. In fact, sometimes, higher performance reduces impact because higher performance is often correlated with factors that drive down adoption and scale such as cost, durability, multipurpose use, and difficulty of distribution (Jetter et al., 2012).

So, measuring impact is critically important. We know that measuring impact means more than just measuring performance. We also need to measure adoption (to what extent an average user utilizes a particular innovation) and scale (the quantity of users). For the purposes of this chapter, we will focus exclusively on a case study of how to measure the adoption of cookstoves using Internet of Things (IoT) sensors. This case study will span my PhD research and then the company that grew out of that research. So, let's start by thinking about how to measure adoption.

1.1 Surveys Collect Bad Data

Although most practitioners of development engineering are motivated by good intentions, our positive intentions do not necessarily lead to positive impacts. When we create a new technology, we cannot just hope (or expect) that it will improve lives. Before we can scale up or make any claims about the impacts of technological interventions, we must carefully measure if and how customers use technological interventions. We call these patterns of individual customer interaction with a new technology “adoption.”

Historically, surveys have been the most common “instrument” for monitoring adoption (Burwen, 2011; Lewis & Pattanayak, 2012). At first glance, survey instruments seem like an attractive option. After all, most intervention programs are interested in some basic understanding of questions like the following: Do customers use our intervention? Do they like it? Would they buy it? Would they recommend it to a friend? Is the intervention improving their lives in some meaningful way?

Most of these questions feel like they could be reasonably answered in a thoughtful discussion with a customer. Surveys can be implemented with as little material as a clipboard and pen, training staff to enumerate a survey is relatively straightforward, and respondents can give rich contextual color, anecdotes, and insights in their answers which can be difficult to gather otherwise. These factors can make surveys seem like a low-cost and high-reward instrument for collecting data about technological impacts.

However, surveys suffer from two critically important problems: recall errors and social-desirability bias (Das et al., 2012; Edwards, 1957; Methodology, n.d.; Thomas et al., 2013; Zwane et al., n.d.). Recall errors result from the difficulty surveyees have in recalling facts or events, even if the surveyee is trying in earnest to respond truthfully. In research, my colleagues and I have been interested in understanding if, why, and to what extent people used their traditional and intervention cooking appliances.

However, getting quantified metrics about cookstove use through surveys can be challenging. For example, let me ask you, the reader, how many times and for how many total minutes did you operate your microwave last month? Well? This is a difficult question for anyone to answer. Most of us don't think much about our personal microwave utilization stats, so you might respond to me by just guessing a number that "feels right," or if you're good with mental math, you might do some quick thinking to try and estimate an accurate response. Or, more likely, maybe you won't say much at all. When I ask this question to students in university lectures, most students just freeze and say "I . . . I don't know." In terms of measuring quantified impact, these kinds of answers are not very useful.

While recall errors are problematic, they do not necessarily introduce systematic bias to survey results. In answering the microwave question, we might expect roughly the same number and degree of overestimated and underestimated answers. By contrast, the other main drawback of surveys, social desirability bias, does indeed create problems with systematic bias. Social desirability bias is the tendency for research subjects to offer the normative or desirable response. This kind of normative response, where the respondent tells the survey enumerator the "right" answer that the enumerator "wants to hear," tends to overstate "positive" behaviors and understate "negative" behaviors (Nunnally & Bernstein, 1994).

In our microwave example, imagine I had sent that microwave to you as a gift a few months back. It's a weird gift, for sure, but then, I'm a weird guy who is obsessed with cooking appliances. Today, I sit you down with a pad of paper and a pencil, and I put you on the spot. "How often do you use that microwave?" I ask. "Well," you might be inclined to say, "I use that microwave all the time. Every morning in fact. It's perfect for oatmeal, which I just love by the way, so I use it every morning. Yep."

Nope. You didn't even take it out of the box because you already have a perfectly good stove. Also, you hate oatmeal.

Unfortunately, this is usually the social dynamic in development engineering field studies. Intervention products like pumps, lights, cookstoves, water filters, and mosquito nets are donated or heavily subsidized by the research project. Then,

after some time, study participants are asked to self-evaluate their adoption of the technology (e.g., “how many nights last week did you sleep under the mosquito net?”) and other qualitative aspects of the technology’s appropriateness (e.g., “do you feel like this mosquito net does a good job keeping your family safe from malaria?”). Multiple research studies have demonstrated that responses to these kinds of questions are weakly correlated (or not correlated at all) with actual user behavior (Wilson et al., 2015, 2016a; Wilson et al., 2018). So, how do we measure actual user behavior?

1.2 Sensors Collect Good Data, But It’s Hard to Do It Right

Sensors are a great choice for collecting cold unbiased data about the physical world. Sensors can be used to objectively measure facts about the environment such as the heat from a cookstove, the flow of water through a pump, the flip of a switch, or the opening and closing of a door (Clasen et al., 2012; Thomas et al., 2020; Turman-Bryant et al., 2020). When sensors are designed well and implemented in a thoughtful manner, they can collect data streams that offer insights about users’ behaviors that could not be realized through surveys alone (Iribagiza et al., 2020; Thomas et al., 2020; Wilson et al., 2017).

As a quick side note: sensors are not immune to bias issues either. Research has shown that, when users know their behavior is being observed, even by sensors, the way they behave changes (e.g., they might use an intervention cookstove more). This effect, where users change behavior when they are aware of being observed, is called the “Hawthorne Effect” (Landsberger, 1958; McCarney et al., 2007; Methodology, n.d.; Thomas et al., 2016).

However, sensors alone do not deliver insights or even intelligible data. There is a wide gap between the raw data collected by data-logging sensors and the insights most of us expect from sensor-based data collection systems (Kipf et al., 2015; Wilson et al., 2020). When it comes to interpreting the data from sensor systems, we have been spoiled by exceptionally vertically integrated and user-friendly sensor systems, especially in the Internet of Things (IoT) age. Take a modern IoT pedometer like Fitbit. At its core, a pedometer is a multi-axis accelerometer measuring acceleration and thousands of samples per second. In the case of Fitbit, millions of person-hours have gone into turning that raw data stream of thousands of acceleration samples per second into a beautiful, intuitive, and comprehensible hardware product and mobile application that allows anyone to collect and visualize metrics about the number of steps a user has taken of the course of time.

So, where’s the Fitbit of cookstove adoption tracking? Therein lies the problem. Development engineering practitioners want to ask questions about adoption of novel technologies, and usually, there are no off-the-shelf sensor systems or IoT products available to simplify the data collection and interpretation process. As development engineering practitioners, we find that we’re usually not in a situation where we can buy an elegant hardware and software solution to solve our adoption

Table 15.1 Example data logger data

Datalogger_id	Timestamp	Temperature
Logger-314	2020-05-01 1:00:00	20.0 °C
Logger-314	2020-05-01 1:01:00	20.2 °C
Logger-314	2020-05-01 1:02:00	21.5 °C
Logger-314	2020-05-01 1:03:00	22.1 °C
...

monitoring and evaluation needs. Using the pedometer analogy, we're back in the situation where we need to collect the thousands of accelerations reading per person per second and then make sense of the data later on. This means collecting raw sensor data about flow, temperature, humidity, acceleration, voltage, power, pH, or whatever other environmental variable correlates well with adoption of the technology you care about.

The tool of the trade for collecting raw sensor data streams is called a “data logger.” A sensor is a device that turns some property of the physical world, such as temperature, into a machine- or human-readable format such as voltage or the position of a dial. By contrast, a data logger is a device that records the values of a sensor to memory for later transmission to a computer. Data loggers typically collect and store sensor data into large flat archives such as comma-separated value files (Table 15.1).

In the cookstove example, temperature data loggers could be used to collect large quantities of temperature time series data. Essentially, the output of a temperature data logger would look like this.

Just raw data, data point after data point and file after file, typically for hundreds of data loggers and millions of observations. For example, imagine a 100-household cookstove adoption monitoring and evaluation program. This program monitors two cookstoves per household for 3 months using data loggers that record temperature once per minute. This monitoring campaign would collect roughly 26 million individual data points.

Managing the collection and analysis of this data looks like this can be a major headache. First, as you can probably surmise from the small collection of top Google-ranked data loggers in Fig. 15.1, the data logger industry is industrial and clunky. These devices usually do not come with slick mobile apps and intuitive fleet management tools. Instead, data from data loggers is typically downloaded by hand using proprietary cables and dongles using a piece of industrial software which often only runs on a legacy version of Microsoft Windows as shown in Fig. 15.2. Training field staff to reliably operate this kind of software has been a significant challenge in my research.

Successful projects using traditional data loggers must ensure quality across many critical steps in the data chain. Figure 15.3 shows 12 important steps in the sensor data chain spanning the time period before data collection begins, while data collection is taking place and after collection of data has been completed. At any step along this path, significant data quality issues can be introduced which can dramatically impact the success of a project. Here are just a few real-world examples

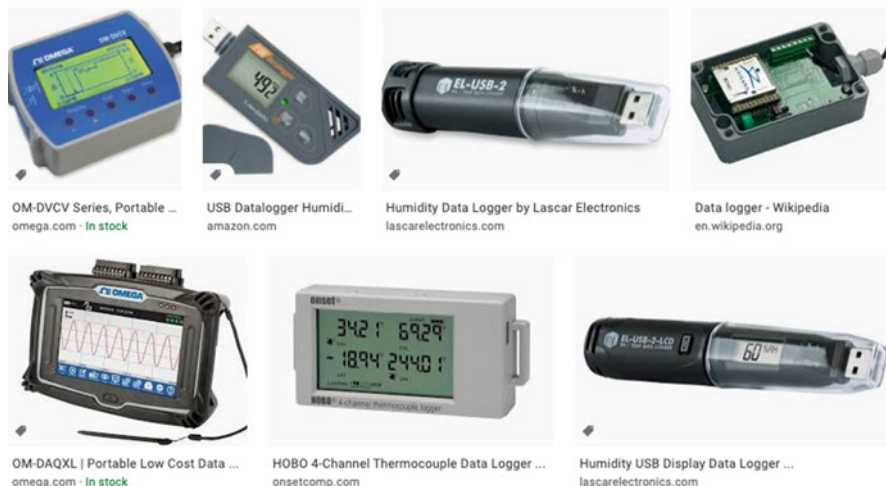


Fig. 15.1 The first images for the Google Image Search term “data logger.” Accessed on May 30, 2020

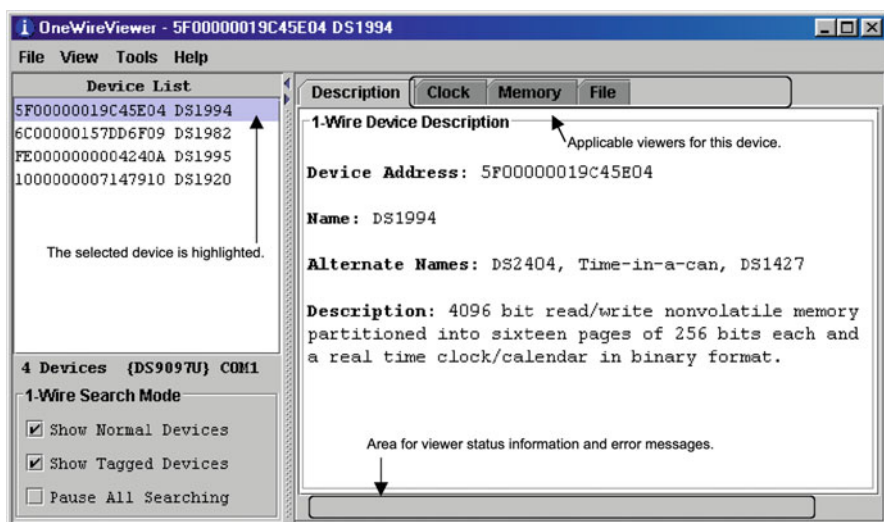


Fig. 15.2 Example of a typical data logger software interface from Maxim OneWireViewer User’s Guide, version 1.3

of small mistakes that create significant problems in sensor-based data collection campaigns:

1. Incorrect dates on staff laptops leading to confusion about when data were actually collected. The data logger gets its clock time from the computer that

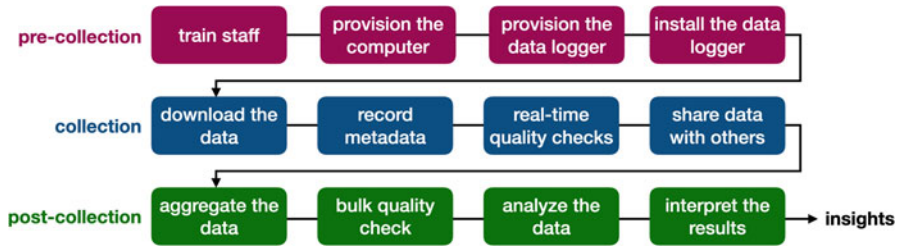


Fig. 15.3 The data chain for monitoring and evaluation of technology adoption with data-logging sensors

provisioned it, and the staff's computers had the wrong time (or even the wrong date) for some unknown amount of time.

2. Staff incorrectly record metadata when downloading data from data loggers. For example, data is downloaded, and the file from this data is mistakenly named household-124.csv instead of household-123.csv; there is no easy way to even know there was a mistake, let alone correctly know where this data was downloaded from.
3. Staff do not have adequate training to interpret data quality in real-time, and research leadership does not have systems in place to oversee data quality. Therefore, data quality problems persist unseen throughout a study. We have observed several examples where research staff diligently visit households every few weeks for months on end to repeatedly download garbage data from a data logger with a broken sensor. This problem is not noticed until months after the data collection part of the study is complete.
4. Graduate students with little to no data science experience are tasked with analyzing huge datasets from these kinds of experiments. It takes months or years for the students to complete the analysis, and often, the analysis suffers from statistical or analytical errors.

If there is any broken link in the data chain, an expensive multi-year study can be significantly damaged or even ruined entirely.

2 Monitoring Cookstove Adoption with Sensors

2.1 Darfur

2.1.1 Implementation Context

In 2012, the war in Darfur, Sudan, had been ongoing for 9 years. As a result of the conflict, millions of people from rural Darfur had concentrated into camps of internally displaced persons (IDPs). People in these camps traditionally cooked on



Fig. 15.4 Darfuri women carrying fuelwood back to IDP camp. Photo credited to Ashok Gadgil

firewood that was freely collected in the rural areas. However, these new large-population centers had put enormous pressure on the woody biomass resource in and around the camps. Over the years, the radius of complete biomass denudation increased to the point where, in 2005, it was estimated that women had to walk a 7-hour round trip from the camp to collect enough firewood to cook for just 2 or 3 days (Galitsky et al., 2006) (Fig. 15.4).

These trips were difficult and dangerous with many reports of sexual violence occurring while women were outside the relative safety of the camps. In many camps, particularly in the arid North Darfur, fuelwood had become out of reach—and collection trips on foot had become impossible. The situation had become desperate enough that many women reported that they had begun to trade their relief food rations to shady businessmen who would truck in firewood from rural Darfur and Chad. Women would use the firewood they bartered for to cook what food they had remaining for their families.

Since 2005, Professor Ashok Gadgil and team at UC Berkeley and Lawrence Berkeley National Laboratory had been developing and distributing an improved cookstove called the Berkeley-Darfur Stove (Fig. 15.5). This cookstove had been demonstrated to reduce fuel use by roughly 50% (Jetter et al., 2012; Rapp et al., 2016). The hope was this reduction in fuel use would lead to a commensurate reduction in time, cost, and risk spent collecting or purchasing fuel wood. By 2012, the work of distributing this cookstove had been transferred to a Berkeley-based nonprofit organization called Potential Energy. Potential Energy was interested in quantifying the impact of the 50,000 or so cookstoves that had been distributed to

Fig. 15.5 Berkeley-Darfur Stove



date. At that time, Potential Energy had just changed its name away from the Darfur Stoves Project. Potential Energy's executive and assistant directors, Andree Sosler and Debra Stein, knew me through my relationship with my PhD advisor, Potential Energy board member, and Berkeley-Darfur Stove inventor Ashok Gadgil.

Our goal was to assess the adoption of about 150 Berkeley-Darfur Stoves just after they had been distributed for free in the IDP camp near the town of Al Fashir in North Darfur. The plan was to compare adoption of the cookstove measured by sensors to self-reported adoption data gathered through surveys. With these two data sources, our hope was that we could build some sort of regression in order to make sense of some previously collected survey data about adoption. We hoped to run this study in two contexts: internally displaced peoples' camps near Al Fashir as well as unorganized rural settlements further away from Al Fashir.

Part of Potential Energy's motivation to perform this adoption study was to build a case for carbon financing (Wilson et al., 2016b). At the time this study was conceived, Potential Energy was already working with an organization, Impact Carbon, that validated carbon credits. Still, additional validation about cookstove adoption could have supported Potential Energy's case that the fuel savings of the Berkeley-Darfur Stove helped to offset anthropogenic CO₂ emissions. To measure the adoption of the cookstoves, Potential Energy was interested in using a sensor-based stove usage monitoring system (SUMS). Seminal work on SUMS had been



Fig. 15.6 El Haj Adam of SAG in a sea of Berkeley-Darfur Stoves outside Al Fashir assembly workshop

performed by Ilse Ruiz-Mercado during her doctoral research at UC Berkeley just a few years earlier, and Potential Energy decided that they would like to pursue this approach (Ruiz-Mercado et al., 2008, 2011, 2012, 2013).

Our plan was to work with a local Sudanese nongovernmental organization (NGO) partner, Sustainable Action Group (SAG), to run the study. SAG would offer up one of their staff as the study leader and coordinate all of the day-to-day activities and management of the field staff related to the study. SAG was already intimately involved with the Berkeley-Darfur Stove Project because they were the NGO responsible for the local assembly and distribution of cookstoves after the cookstoves arrived as flat kits to Darfur (Fig. 15.6).

Following in the footsteps of Ilse Ruiz-Mercado, our team had decided to use Maxim iButton temperature loggers as our stove usage monitoring system (SUMS). I'll refer to this term as "SUMS" throughout this chapter whenever discussing a sensing device that is specifically employed to track cookstove adoption. The Maxim iButton is a self-powered data logger with all its electronics, temperature sensor, and battery contained in a metal button of the size and shape of a watch battery. Some models could withstand temperatures as high as 140 °C. The idea was that we could install the iButton SUMS on the outside of the cookstoves, and as the temperature of the cookstove rose, these temperatures would be recorded by the iButtons, and we could later correlate spikes in temperature with cooking events.



Fig. 15.7 Custom iButton case. The iButton data logger itself is the object that looks like a coin cell battery, second from the left

However, these iButton data loggers were never designed to be used on cookstoves. They had some major weaknesses including a short battery life, and when the battery died, all of the data was lost. Maxim iButtons are most often used in the food industry, mounted to the side of huge milk containers inside refrigerated trucks or inside buffet service stations to ensure food remains cold or hot enough to be safe for consumption. To adapt the iButtons for use on a cookstove, we designed an aluminum case that would hold the iButton and keep it firmly pressed against the surface of the cookstove. This aluminum case also allowed us to clearly stamp a large alphanumeric code onto the case of the SUMS, for example, “A-12,” which made it possible to track which data came from which data logger (e.g., “2013-08-12 A-12” was data from A-12 downloaded on August 12, 2013) (Figs. 15.7 and 15.8).

Throughout late 2012, our team attempted to obtain a visa to travel to Darfur to kick off this project with SAG, but due to political turmoil in Sudan and restrictive policies about visitors from the United States, our team was not able to secure visas by early 2013. Instead, our team decided that in January 2013, Angeli Kirk and I would fly to Addis Ababa, Ethiopia, where we would meet up with a representative of SAG, Abdel Rahman, to do a multiday training session. Angeli, a research colleague, PhD student, and friend from Berkeley’s Agriculture and Resource Economics program, would meet up for the latter half of the training session (and to assess Potential Energy’s expansion opportunities in Ethiopia).

During this training session, we would familiarize Abdel with the design of the experiment as well as how to use the sensors we would need to implement the study. I had brought all of the equipment with me to transfer to Abdel who would then take them into Sudan (Fig. 15.9). We planned to wait until the summer to start the study, and we were still hopeful that we would be able to secure a future visa to perform an in-person training with field staff in Darfur sometime in the late spring of 2013. Therefore, this training served as a kind of “kickoff” where we would get familiar with the sensors and identify the personnel, facility, and equipment resources that SAG would need to complete the study.

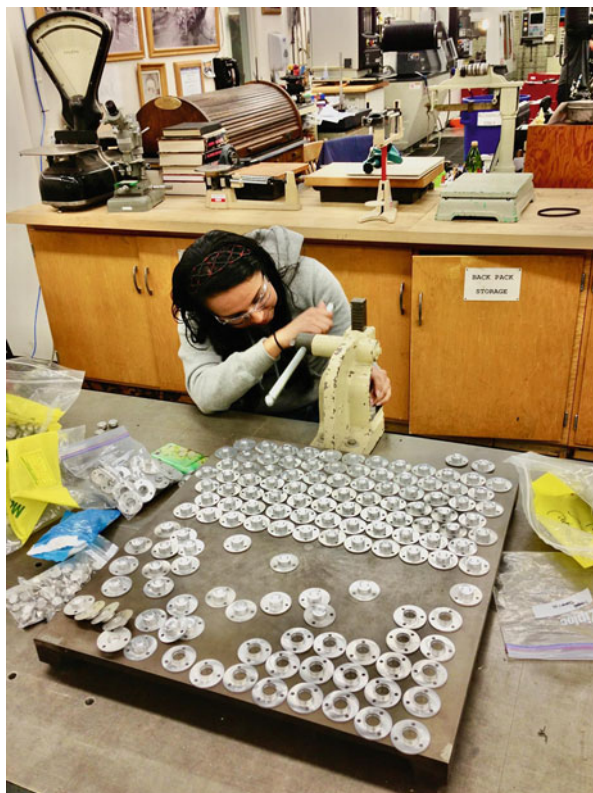


Fig. 15.8 Areidy Beltran, undergraduate research assistant extraordinaire, assembles iButton SUMS in the mechanical engineering machine shop during finals week, December 2012

2.2 Innovate, Implement, Evaluate, and Adapt

2.2.1 The Lead-Up

As the early months of 2013 wore on, it became clear that our Berkeley-based team was never going to be able to visit Darfur. In late February, we got an email from Jan Maes, a consultant who worked for Potential Energy. He let us know that a member of the Impact Carbon team, Ellie Gomez, had been abducted from her hotel in Darfur by armed assailants. This was the same hotel I was planning to stay at. Miraculously, Ellie escaped her kidnappers just minutes after the abduction. This close encounter happened in the context of frequent stories about aid workers and NGO employees being kidnapped, held for ransom, and sometimes murdered in Darfur.

The abduction incident threw cold water on our whole study. We had originally planned to do some studies in the rural settlements surrounding the IDP camps, but considering the security situation, these plans were scrubbed. Also, we decided we



Fig. 15.9 Supplies brought from Berkeley to Ethiopia in January 2013 to be transferred to Darfur

could not even visit individual domiciles in the IDP camps; instead, we would have to ask women to congregate at an IDP Women’s Center.

These meetings at the Women’s Center were partly necessary to administer qualitative surveys, but they were also a result of the way iButton sensor data needed to be downloaded: to access the data, our field staff needed to carry a laptop computer, a set of cables and dongles, and screwdrivers and wrenches to detach the data logger from the cookstove. This whole process took about 10 minutes per cookstove and certainly created quite a spectacle in the IDP camps. If we could have discretely and wirelessly transmitted the data, it is likely that our study could have been designed differently.

In addition to creating a custom case for securely holding and attaching the SUMS to the cookstoves, it was extremely important to identify a good placement location for the SUMS. Because these iButtons were a fully integrated miniature data logger in a small case, that meant that the sensitive microcontroller, memory, and battery would get just as hot as the temperature sensor. Want to measure a 100 °C temperature? Well, the whole data logger has to get that hot.

This posed a significant optimization challenge when measuring cookstove use in the hot Sudanese desert. The bounds of this design challenge were as follows:

1. In order to easily identify temperature spikes in temperature time series data, the temperature when cooking should be as high as possible.
2. In order to maximize battery life and minimize the risk of damaging the data logger and losing all the data, the temperature should be as low as possible.
3. The hot summer sun in Darfur could easily heat a cookstove's surface to 50 °C or more even when a user was not cooking.

This “get the data logger as hot as possible but not so hot that it destroys itself or drains the battery before the study is complete but also hot enough to unambiguously identify cooking apart from solar heating” design challenge required some significant engineering effort. This challenge was unique to the design of the iButton SUMS because of the quirk that the sensor was co-located with the battery, memory, and microcontroller. We used parameterized testing, thermal cameras, and some optimization calculations around battery lifetime vs. temperature to select a best-possible mounting location (Fig. 15.10).

Not only was this position-optimization process cumbersome and overwrought, it had very little potential for scalability or impact. How could any team of other than well-resourced engineering PhD students replicate this technique for future studies?

Double-Click vs. Single Tap

It's a muggy January 2013 day in Addis Ababa, Ethiopia. I am sitting on a loveseat in my hotel room with Abdel Rahman at my side. We're both looking at the screen of a laptop that Abdel has brought with him from his NGO office in Sudan. Together, we're working through how to use a particularly byzantine piece of data logger software, but we keep getting hung up on one particular step. Abdel is having a

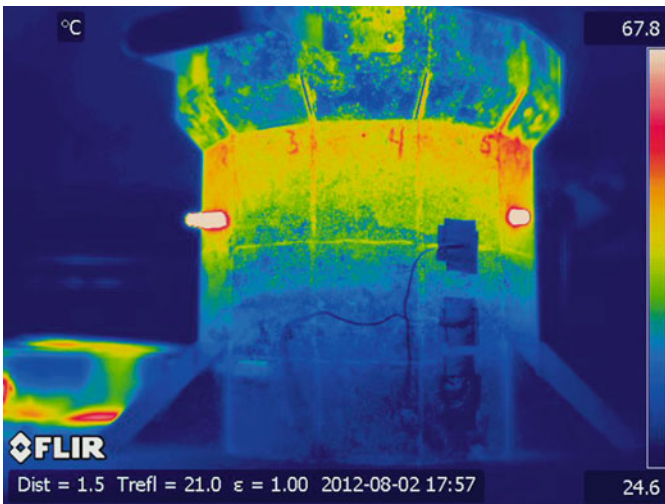


Fig. 15.10 A Berkeley-Darfur Stove imaged by an infrared camera to determine optimal iButton SUMS placement, August 2012

very difficult time double-clicking certain icons in the software's user interface. As in, he just cannot properly execute the double-clicking action. It's not clear what is leading to this confusion, because Abdel is clearly computer literate. But he can't quite seem to get the timing or positioning of the cursor quite right. Perhaps this is a new laptop, or maybe he is used to using a wired mouse and not fussing with the flimsy trackpad and buttons of this low-quality laptop.

We're spending long stretches of time, sometimes minutes on end, simply trying to open files, click menus, and generally get the cursor to behave the way Abdel wants. He's getting frustrated, and I can feel it. I'm starting to wonder to myself, "if Abdel, as the senior lead of this project, cannot succeed at this task from the quiet comfort of a hotel room, how are his dozens of less-educated field staff going to be successful in the middle of a Darfuri IDP camp in the blazing hot sun surrounded by countless challenges and distractions?" Abdel is fidgeting next to me in the loveseat and getting increasingly agitated. He asks for a break, pulls out his phone, and starts effortlessly navigating the apps and screens, checking email, messaging friends, and reading articles. I think to myself, "well, that's interesting."

This was a formative experience in my journey to build a better system for monitoring cookstove adoption. Whatever we did long term, we needed it to work on mobile phones and mobile phones alone.

My experience struggling to use the laptop in the hotel room in Ethiopia with Abdel Rahman caused me to become increasingly concerned that a laptop-based sensor data acquisition system would not be appropriate for the field staff in Darfur. In the spring of 2013, Javier Rosa, an electrical engineering and computer science PhD student at Berkeley, and I spent considerable effort trying to hack the iButton dongles to work with Android phones.

However, after much consternation, we were never able to read an iButton with a phone. But this desire for a system that would allow for a phone-only ecosystem to provision, deploy, and collect data from SUMS stayed with us (Kipf et al., 2015). For the time being, we resigned ourselves to the idea that one or two exceptionally well-trained field staff would need to run the laptop computer to download data from the iButtons, while the survey enumeration staff would administer the survey in another area of the Women's Center.

Given the ubiquity of literacy with mobile phones, we decided to design the survey-based data collection for our study around an open-source survey tool called Open Data Kit (ODK). We also designed a sidecar paper survey with identical questions just in case the phone survey failed. Also, in true mechanical engineering grad student overengineered fashion, we even designed a little jig to allow field staff to repeatably take high-quality photos of the survey since the SAG offices in Darfur lacked any kind of scanner (Fig. 15.11).

Part of the reason we were anxious about the ODK system not working correctly (and thus implemented the paper survey as a backup plan) was our discovery that global sanctions on the Sudanese government were blocking Internet traffic originating from Sudan to ODK's servers, which were hosted by Google at that time. We discovered this in late spring 2013, and the experiment was scheduled to



Fig. 15.11 The Open Data Kit (ODK) mobile phone and backup jig to take photos of paper surveys

start that summer. This incident led to a scramble to redeploy the open-source ODK backend system on Javier's own server. Pointing traffic originating in Sudan to a server on US soil was probably still a violation of some sort of international law, but our attempts to deploy a server in Sudan were not successful. This entire incident, which could have significantly delayed our experiment if it would have been discovered a few months later, helped me to understand the value of controlling the entirety of the data platform when performing field research in heavily sanctioned countries or countries with extremely restrictive Internet access. We also owe our success to the open-source feature of the ODK backend system, so it could be ported to Javier's server at UC Berkeley.

Meanwhile, around April of 2013, it was becoming clear to our team that Abdel Rahman was not going to be able to execute this study as we had planned. It had become increasingly difficult to maintain consistent lines of communication and continue momentum, progress, and meet schedules with Abdel and SAG. It seemed like the SUMS study was not a top-of-mind priority for SAG and was losing momentum. Additionally, it was unclear if Abdel was prepared to lead the complex administration and oversight of the study from Darfur. Abdel remained a critical member of our SUMS study team, especially as it related to cookstove assembly and distribution, but we needed to make an important decision about bringing in additional help.

On April 10, 2013, Jan Maes, Potential Energy's consultant, introduced me to Dr. Mohammed Idris Adam. Dr. Moh (as he liked to be called) was a professor at Al Fashir University in Darfur. He had been assisting Impact Carbon and Potential Energy with enumerators of surveys related to carbon credits, and he had become a trusted collaborator. Over the course of the coming months, Dr. Moh, Angeli, and I collaborated via email and planned to meet in Addis Ababa in July to coordinate final plans for the experiment.

Around this time, another important hire was made. Potential Energy hired a Sudanese woman then living in the Bay Area named Omnia Abbas. Omnia was an incredible resource at Potential Energy. She was fluent in Arabic and English, intimately familiar with Sudanese culture, and could travel with relative ease back and forth to Darfur.

I met with Dr. Moh, Omnia, and Potential Energy’s associate director, Debra Stein, in Addis Ababa in July 2013. The content of this meeting was largely a repeat of the meeting just 6 months earlier. Unfortunately, the timing of this meeting was over Ramadan, and the fasting and prayer schedule made it difficult for Dr. Moh to participate past the midafternoon (this is a common rookie scheduling oversight made by researchers). But still, the effect of this second meeting was transformational for the project. In retrospect, the difficult decision to pivot project leadership just months before the study was slated to begin was one of the most important decisions we made for the overall success of this project. Before hiring Dr. Moh and Omnia, our team did not realize the incredible importance of having an experienced, motivated, and trusted champion for your research study stationed in the field. Today, I strongly believe that it is impossible for field research to be administered solely from outside the study site. To this day, one of the first questions I ask students who are planning field work is “Who is the champion who is located in the field?” (Fig. 15.12).



Fig. 15.12 Left: Dr. Mohammed “Moh” Idris Adam (top row, third from left) and Abdel Rahman (top row, fifth from left) with the SUMS field team they assembled in Darfur. Right: In the foreground, field staff administer surveys to Darfuri IDP women using paper and mobile phone-based techniques. In the background, Dr. Moh and an assistant interact with SUMS using a laptop computer

2.2.2 The Study

In August of 2013, the study in Darfur began in earnest. One of the first things we noticed was that far more of our sensors were failing due to overheating than expected. We had invested such careful attention in the placement of the sensors (e.g., the thermal imaging), but now, we noticed that about 20% of our data loggers were coming back with symptoms of overheating.

In our survey data, we found that participants who self-reported charcoal as a primary or secondary fuel source were much more likely to have burned-out sensors. After some sleuthing by SAG field staff, we discovered that some women like to burn charcoal in the Berkeley-Darfur Cookstoves even though the cookstove was only designed to burn wood. In order to adapt the cookstove to the charcoal fuel, women were flipping the cookstove upside down and packing the bottom of the stove (close to where the iButton was mounted) with charcoal. This behavior caused the iButtons to overheat and become unresponsive. It was very valuable to discover how common this charcoal-cooking behavior was, but it was at the heavy cost of losing almost 20% of our data. Additionally, this data was lost in a way that biased study results (participants who don't use their cookstove can't burn out their sensors and therefore were overrepresented in our surviving sensor data).

Throughout the fall, our team performed the hard work of administering a sensor-based data collection program from overseas. Dr. Moh and our team had regular weekly meetings late at night California time. At first, these meetings covered mundane issues such as how to buy new data plans for the SIM cards in the ODK cell phones, if we would be reimbursing the field staff for gasoline, and how to download and install Dropbox. However, as time went on, an increasing number of our conversations covered mounting issues related to data quality. "Where's the data from household 10?" "I looked through the data, and there are 3 baseline surveys that are all marked as being from the same respondent." "The sensor data from household 153 says that it was collected in the year 1970—what's going on there?" Our team was so busy administering the day-to-day activities of the study that we did not have the time or resources to preprocess any of the data coming back from Darfur.

While I was occupied with administration work, Angeli Kirk was looking through some of the early results from sensor data that had been downloaded and sent back to California via Dropbox. She noticed that a small subset of households had barely used their stoves yet. She wondered what impact, if any, returning to the Women's Center for the follow-up survey (and data download) would have on future adoption for this group. However, the plan had always been that the data collection would stop after the follow-up survey visit. Personally, I was struggling to hold the administration of the project together as it was, and I was not interested in mission creep to answer a new and unplanned research question. However, after some prodding from Angeli and an amendment to our institutional review board (IRB) protocol, I acquiesced. We asked Dr. Mohammed to send future tranches of follow-up survey takers home with SUMS-equipped stoves instead of taking the sensors off. A second and final follow-up was planned on the fly just to remove the

SUMS. This decision, which was enabled by an early peek at results, turned out to be vital for our research.

As the plan changes and late-night phone calls carried on, administering the Darfur SUMS program became a difficult contextual picture to maintain. The number of caveats and special cases and “oh yeah, that sensor is a different special case for ___ reason” issues began to mount. These gotchas were cataloged in ad hoc in emails, paper notebooks, spreadsheets, and mental notes. By the time the last of the sensor data was being collected in November 2013, we knew our team only had just a month or two of short-term memory acuity before it would become nearly impossible to stitch together all of the data into a coherent story.

2.2.3 After the Study

In the winter of 2013, we began assembling the data from the SUMS study. I personally had never analyzed data of this volume before. Up until that point, I had just “faked it until I made it” with data analysis. I was comfortable using Excel and had some really weak MATLAB skills. But the four million data point SUMS dataset was a huge step up from anything I, or anyone else on our team, had analyzed before. There were hundreds of files, each with about ten thousand rows of raw temperature and timestamp data, and our job was to create a coherent story about how women in Darfur adopted the Berkeley-Darfur Cookstove. We didn’t know where to start.

A few months earlier, I had carpooled to Burning Man with another graduate student named Jeremy Coyle. Jeremy was a PhD student in biostatistics at Berkeley, and he lived across town from me in the student coops. Our team needed some help, so I rode over to his place on my bike to see if he could help me learn to use R, the statistical programming language, to analyze the Darfur SUMS dataset.

I had thought that Jeremy and I would just spend an afternoon getting me up speed with R, and I would be on my way. But eventually, to both of our surprises, Jeremy and I collaborated on the analysis of the Darfur dataset for the next 3 months. The amount of time that went into this analysis was far beyond my expectations. As an aside, Jeremy and I still collaborate on research, software development, and data analysis to this day.

The majority of this effort was spent creating a cooking-event-detection algorithm that could reliably find the start and end of individual cooking events from long stretches of temperature time series (Fig. 15.13). Also, there was significant data cleaning, organizing, and merging of survey and sensor data. Because the sensor and survey data acquisition systems were completely separate, all of the merging and comparison efforts needed to be hand-rolled in code. Our team estimated that we spent 400 person-hours on this analytics effort. In 2020, the going rate for quality data science consulting is about \$300/hour. Although a professional senior data scientist would probably be able to finish this analysis faster than we did, an analysis like this could easily cost on the order of \$100,000. However, more likely, a smaller NGO or development agency that wanted to deploy sensors to

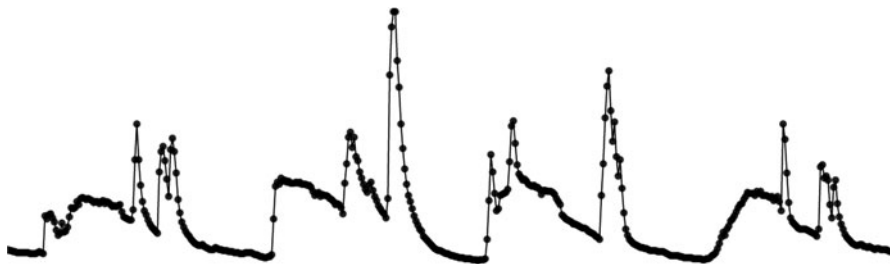


Fig. 15.13 An example of a temperature time series to illustrate the challenge of creating a deterministic algorithm to count and quantify “cooking.” Where do the cooking events start and stop? Axis scales intentionally left blank, but vertical (Y) is temperature, and horizontal (X) is time

monitor cookstove adoption would find themselves stuck at the data analysis step and might resort to a much simpler analysis that didn’t realize the full value of the data. This experience helped me to realize the massive investment in time and/or money to analyze sensor data for development research. This investment was likely a major barrier to ubiquitous deployment of sensors for research.

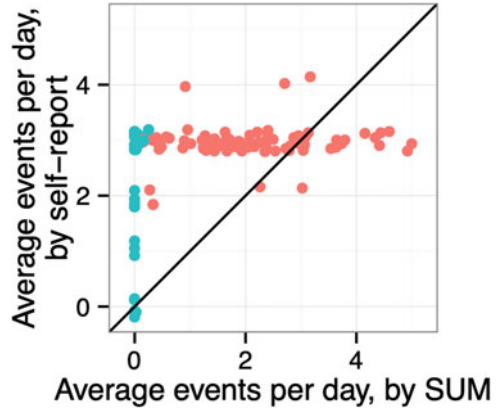
2.2.4 What We Learned

The Darfur SUMS study ended up teaching us many valuable lessons. We found that about 75% of women adopted the Berkeley-Darfur Stove (Wilson et al., 2016a). For the women who did not adopt it, we found out that about 80% of them could be converted into adopters simply through the act of conducting the first follow-up. Angeli was right. Without her early analysis and encouragement to change the research study, we would have never known this important insight that non-adopters just needed a small nudge to become adopters. In addition to this, we confirmed what we believed about the quality of survey data to assess adoption of technology; no matter how we asked the question, we were not able to assess adoption of the Berkeley-Darfur Stove through surveys. As shown in the figure below, when asked a question about how many times per day she uses her Berkeley-Darfur Stove, cooks almost always respond with the socially desirable response: three (every meal). The number of times someone self-reports cooking in surveys has essentially no correlation with their behavior measured by sensors (Fig. 15.14).

In addition to what we learned about how women in Darfur use cookstoves, we learned even more about what it takes to run an effect sensor-based impact evaluation. Some of the most important takeaways from Darfur were as follows:

1. Analytics is a major barrier to effective deployments of sensors for monitoring and evaluation.
2. Field staff are far more comfortable with mobile phones than with laptop computers, and mobile phones do not attract the attention that computers do.

Fig. 15.14 Cooking events measured by self-report vs. SUMS



3. Training field staff to deploy industrial data loggers using industrial tools is extremely difficult and error-prone.
4. Cookstove sensors that cannot survive cookstove temperatures are bound to break.
5. The higher a temperature a cookstove sensor can measure, the clearer its cooking signal will become, and the more easily analytics can be run on the data.
6. It's common for data loggers' batteries to die during a study. So, a dead battery should not erase all of the data from the logger (as happens with iButtons).
7. The ability to continuously and easily audit data in real-time as it is collected in the field is critical to maintaining data integrity and communicating feedback about data quality issues to your field team and for potentially asking interesting new questions during your study.
8. Every time a person needs to take an action in the sensor data chain, anything from naming a file to adding an attachment to an email, you are opening your project up to significant risks in terms of coordination, privacy, and data quality.
9. If it takes the undivided attention of a PhD student and 2 years of significant technical and administrative effort to execute a relatively small sensor-based impact evaluation, then sensor-based impact evaluations are too burdensome for all but the most well-funded academic research studies.

2.2.5 Where We Went Next

SUMSarizer One of the main takeaways from the seminal work in Darfur was that analyzing sensor data for impact evaluation was extremely difficult. Jeremy Coye, Ajay Pillarisetti from Public Health, and I were very interested in democratizing event detection from time series data. We wanted to make it easier for coding-naive users to summarize SUMS data, so we began work on a machine learning tool called SUMSarizer. SUMSarizer was a web-based “label and learn” tool that allowed users to import raw SUMS files from common SUMS data loggers. Once imported, users

by Advait Kumar, Abhinav Saksena, Meenakshi Monga, and me at the Indian Institute of Technology (IIT) Delhi during a 2014 Fulbright Fellowship to India. Unlike the iButton, the ASUM was compatible with multiple sensor input channels, had nonvolatile memory that could survive a dead battery, had room for billions rather than thousands of samples, and used microSD card storage instead of a proprietary interface that required custom dongles and Windows-only software. The ASUM was powered by an Arduino-compatible Atmel microcontroller, and for our research program in India, we integrated its multichannel analog frontend with an advanced cookstove's internal thermoelectric generator, USB charging port, battery, fan, and a proximity switch (Fig. 15.16).

The flexibility of the ASUM allowed us to perform novel research about how advanced cookstoves with inbuilt thermoelectric generators are adopted, namely, how the ability of a cookstove to power a USB port for charging a cellphone and other small appliances influenced adoption of the cookstove (Wilson et al., 2018). Also, the ASUM allowed our team, for the first time, to be able to administrate an entire cookstove impact evaluation using only mobile phones. Most Android phones at the time had a slot for microSD card expansion storage, and we used this port to read the cards and then transmitted the data to the cloud with the phones. The data from ASUMs were analyzed using SUMSarizer.

However, ASUM was not without its problems. It was a boutique purpose-built device for our study. The device could not maintain charge on a reasonable-size battery for more than a couple weeks, it was not very rugged and had no integrated tool for metadata collection (e.g., which household the data logger was deployed in), and the flexibility of the microSD card was as much a liability as a feature; on a few of the ASUM microSD cards we retrieved from the field, we found all of our data was missing and had been replaced by MP3 files of Bollywood music.

Geocene The most recent iteration of a stove usage monitoring system is Geocene. Today, in summer 2020, I am the CEO of Geocene. Geocene makes a fully integrated SUMS product. The three pillars of this system are as follows:

1. The Geocene Temperature Logger (also known as the “Dot”). This data logger runs on replaceable AAA batteries that can be found in almost any small town in the world, uses rugged thermocouple probes that can withstand temperatures up to 500 °C, has over a year of storage and battery life, and communicates over Bluetooth Low Energy (BLE).
2. The Geocene Mobile App runs on Android or iOS and is the only tool field staff need to provision and collect data from Dots. The mobile application also has an inbuilt survey feature. This allows survey data to be collected alongside sensor data, making easy the process of merging, filtering, and analyzing data with survey questions as covariates. The mobile application is designed to run even without access to the Internet, but when it does have access, it syncs its data to the cloud.
3. Geocene Studies is a cloud-based web application that organizes users' sensor and survey data and leverages the SUMSarizer engine to analyze data from the



Fig. 15.16 Top: Evolution of the ASUM from breadboard to final manufactured device. Bottom: ASUMs installed on BioLite cookstoves undergoing functional testing at IIT Delhi

Dots. Today, Studies is analyzing data from about 7000 cookstoves every day, has hundreds of active users, and contains about 1 billion individual temperature samples (Fig. 15.17).

Geocene has supported several very large field trials including the Household Air Pollution Investigation Network (HAPIN) with about 3200 households and LEADERS Nepal with about 2000 households. In addition to cookstove projects, Geocene's platform also supports electricity-monitoring and GPS asset-tracking sensors. Using these tools, Geocene has also supported international development



Fig. 15.17 Clockwise from top-left: a Geocene Dot-brand thermocouple data logger, the Geocene mobile application, and a screenshot of temperature time series data with cooking events detected and highlighted from Geocene’s web application, Geocene Studies

projects monitoring electric grids and tracking the leakage of aid through supply chains and onto the black market.

However, despite these large projects, Geocene does not generate enough revenue from supporting cookstove programs to employ the talented professional engineering team it takes to support such a modern and highly integrated IoT product. Today, most of Geocene’s revenue comes from consulting work for Silicon Valley companies building IoT products. Still, the founding and heart of Geocene are still sensor-based impact evaluation for the developing world.

3 Summary

As development engineers, we strive to develop technologies that will improve lives. However, measuring impact is a complex exercise that must consider product performance, adoption, and scale (IPAS). Historically, measuring adoption relied on self-reported surveys, but we have discussed and demonstrated in this chapter that sensor-based objective evaluation of product adoption is a more reliable and informative measure of adoption behaviors.

Still, deploying sensors came with significant challenges related to sensor provisioning, data collection, warehousing, and analytics. To solve these problems, we built three major iterations of cookstove sensor platforms from cookstoves and

ODK, to ASUM, to the Geocene Dot and Studies platform. In each subsequent iteration, we endeavored to move field work to intuitive mobile interfaces and analytics to cloud-based coding-naive point-and-click systems. In doing this, our goal was to make sensor-based monitoring and evaluation more accessible to broader groups of users.

This journey has been an incredibly enriching part of my life. I hope this chapter has helped the reader imagine some of the possibilities that a career in development engineering could afford. In your own work, my colleagues and I hope you take the lessons we learned into consideration. Aim for impact, measure reality, and keep on iterating toward a future with less poverty and more justice for all human beings.

Discussion Questions

1. When Internet-connected sensors are the dominant instruments used for data collection rather than surveys, how do we still engage local communities to participate in data collection? What happens to capacity building and job opportunities that used to be afforded to large teams of survey enumerators.
2. If you had \$25 K to spend on an impact evaluation of 100 households in rural India but the sensors you needed to perform this evaluation cost \$500 each, what would you do? Propose a rough budget that includes travel expenses, staff salaries, incidentals, and sensors.
3. Given what you have learned in this case study, if you were designing a sensor-based field study of technology adoption, what would be your top-three priorities in terms of designing the study?
4. Sensors can offer objective observations about the physical environment, but what are three ways you can imagine sensors fail to teach us what we need to know about technology adoption?

References

- Burnett, R. T., Pope, C. A. I., Ezzati, M., Olives, C., Lim, S. S., Mehta, S., et al. (2014). An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environmental Health Perspectives*, 122, 397–403.
- Burwen, J. (2011). *From technology to impact: Understanding and measuring behavior change with improved biomass cookstoves* (pp. 1–60). Energy and Resources Group, University of California.
- Clasen, T., Fabiani, D., Boisson, S., Taneja, J., Song, J., Aichinger, E., et al. (2012). Making sanitation count: Developing and testing a device for assessing latrine use in low-income settings – Environmental Science & Technology (ACS publications). *Environmental Science & Technology*, 46, 3295–3303. <https://doi.org/10.1021/es2036702>
- Das, J., Hammer, J., & Sánchez-Paramo, C. (2012). The impact of recall periods on reported morbidity and health seeking behavior. *Journal of Development Economics*, 98(1), 76–88. <https://doi.org/10.1016/j.jdeveco.2011.07.001>
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. Dryden Press.
- Forouzanfar, M. H., Alexander, L., Anderson, H. R., Bachman, V. F., Biryukov, S., Brauer, M., et al. (2015). Global, regional, and national comparative risk assessment of 79 behavioural,

- environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 386, 2287–2323. [https://doi.org/10.1016/S0140-6736\(15\)00128-2](https://doi.org/10.1016/S0140-6736(15)00128-2)
- Galitsky, C., Gadgil, A. J., Jacobs, M., & Lee, Y.-M. (2006). *Fuel efficient stoves for Darfur camps internally displaced persons report of field trip to north and South Darfur, Nov. 16 – Dec.17, 2005 – LBNL-59540* (pp. 1–43). Lawrence Berkeley National Laboratory.
- Iribagiza, C., Sharpe, T., Wilson, D. L., & Thomas, E. A. (2020). User-centered design of an air quality feedback technology to promote adoption of clean cookstoves. *Journal of Exposure Science & Environmental Epidemiology*, 1–12. <https://doi.org/10.1038/s41370-020-0250-2>
- Letter, J. J., Zhao, Y., Smith, K. R., Khan, B., Yelverton, T., DeCarlo, P., & Hays, M. D. (2012). Pollutant emissions and energy efficiency under controlled conditions for household biomass Cookstoves and implications for metrics useful in setting international test standards. *Environmental Science & Technology*, 46, 10827–10834.
- Kipf, A., Brunette, W., Kellerstrass, J., Podolsky, M., Rosa, J., Sundt, M., et al. (2015). A proposed integrated data collection, analysis and sharing platform for impact evaluation. *Development Engineering*, 1–9. <https://doi.org/10.1016/j.deveng.2015.12.002>
- Landsberger, H. A. (1958). *Hawthorne revisited*. Ithaca.
- Lewis, J. J., & Pattanayak, S. K. (2012). Who adopts improved fuels and cookstoves? A systematic review. *Environmental Health Perspectives*, 120(5), 637. <https://doi.org/10.1289/ehp.1104194>
- Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., & Shibuya, K. (2013). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380, 2224–2260.
- McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne effect: A randomised, controlled trial. *BMC Medical Research Methodology*, 7(1), 30. <https://doi.org/10.1186/1471-2288-7-30>
- Methodology, B. B. I. J. O. S. R., 2010. (n.d.). *The Hawthorne effect in community trials in developing countries*. Taylor & Francis. <https://doi.org/10.1080/13645570903269096>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw Hill.
- Pillariseti, A., Vaswani, M., Jack, D., Balakrishnan, K., Bates, M. N., Arora, N. K., & Smith, K. R. (2014). Patterns of stove usage after introduction of an advanced cookstove: The long-term application of household sensors. *Environmental Science & Technology*, 48(24), 14525–14533. <https://doi.org/10.1021/es504624c>
- Rapp, V. H., Caubel, J. J., Wilson, D. L., & Gadgil, A. J. (2016). Reducing ultrafine particle emissions using air injection in wood-burning cookstoves. *Environmental Science & Technology*, 50(15), 8368–8374.
- Ruiz-Mercado, I., Lam, N. L., & Canuz, E. (2008). Low-cost temperature loggers as stove use monitors (SUMs). *Boiling Point*.
- Ruiz-Mercado, I., Maser, O., Zamora, H., & Smith, K. R. (2011). Adoption and sustained use of improved cookstoves. *Energy Policy*, 39(12), 7557–7566. <https://doi.org/10.1016/j.enpol.2011.03.028>
- Ruiz-Mercado, I., Canuz, E., & Smith, K. R. (2012). Temperature dataloggers as stove use monitors (SUMs): Field methods and signal analysis. *Biomass and Bioenergy*, 47, 459–468.
- Ruiz-Mercado, I., Canuz, E., Walker, J. L., & Smith, K. R. (2013). Quantitative metrics of stove adoption using Stove Use Monitors (SUMs). *Biomass and Bioenergy*, 57, 136–148.
- The World Bank. (2011). *Household cookstoves, environment, health, and climate change* (pp. 1–94). The World Bank.
- Thomas, E. A., & Brown, J. (2020). Using feedback to improve accountability in global environmental health and engineering. *Environmental Science & Technology*, acs.est.0c04115–10. doi:<https://doi.org/10.1021/acs.est.0c04115>.
- Thomas, E. A., Barstow, C. K., & Rosa, G. A. (2013). Use of remotely reporting electronic sensors for assessing use of water filters and cookstoves in Rwanda. *Environmental Science & Technology*.

- Thomas, E. A., Tellez-Sanchez, S., Wick, C., Kirby, M., Zambrano, L., Abadie Rosa, G., et al. (2016). Behavioral reactivity associated with electronic monitoring of environmental health interventions—A cluster randomized trial with water filters and cookstoves. *Environmental Science & Technology*, 50(7), 3773–3780. <https://doi.org/10.1021/acs.est.6b00161>
- Thomas, E. A., Kathuni, S., Wilson, D. L., Muragijimana, C., Sharpe, T., Kaberia, D., et al. (2020). The drought resilience impact platform (DRIP): Improving water security through actionable water management insights. *Frontiers in Climate*, 2, 258–211. <https://doi.org/10.3389/fclim.2020.00006>
- Turman-Bryant, N., Sharpe, T., Nagel, C., Stover, L., & Thomas, E. A. (2020). Toilet alarms: A novel application of latrine sensors and machine learning for optimizing sanitation services in informal settlements. *Development Engineering*, 5, 100052. <https://doi.org/10.1016/j.deveng.2020.100052>
- Wilson, D. L., Adam, M. I., Abbas, O., Coyle, J., Kirk, A., Rosa, J., & Gadgil, A. J. (2015). Comparing cookstove usage measured with sensors versus cell phone-based surveys in Darfur, Sudan. In S. Hostettler, E. Hazbourn, & J.-C. Bolay (Eds.), *Technologies for development* (pp. 211–221). Springer International Publishing. https://doi.org/10.1007/978-3-319-16247-8_20
- Wilson, D. L., Coyle, J., Kirk, A., Rosa, J., Abbas, O., Adam, M. I., & Gadgil, A. J. (2016a). Measuring and increasing adoption rates of cookstoves in a humanitarian crisis. *Environmental Science & Technology*, 50(15), 8393–8399. <https://doi.org/10.1021/acs.est.6b02899>
- Wilson, D. L., Talancon, D. R., Winslow, R. L., Linares, X., & Gadgil, A. J. (2016b, January). *Avoided emissions of a fuel-efficient biomass cookstove dwarf embodied emissions*. Retrieved August 9, 2017, from.
- Wilson, D. L., Coyle, J. R., & Thomas, E. A. (2017). Ensemble machine learning and forecasting can achieve 99% uptime for rural handpumps. *PLoS One*, 12(11), e0188808. <https://doi.org/10.1371/journal.pone.0188808>
- Wilson, D. L., Monga, M., Saksena, A., Kumar, A., & Gadgil, A. J. (2018). Effects of USB port access on advanced cookstove adoption. *Development Engineering*, 3, 209–217. <https://doi.org/10.1016/j.deveng.2018.08.001>
- Wilson, D. L., Williams, K. N., & Pillarisetti, A. (2020). An integrated sensor data logging, survey, and analytics platform for field research and its application in HAPIN, a multi-center household energy intervention trial. *Sustainability*, 12(5), 1805–1815. <https://doi.org/10.3390/su12051805>
- Zwane, A. P., Zinman, J., & Van Dusen Proceedings of the, E., 2011. (n.d.). Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.1000776108/-/DCSupplemental>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

