Grand Challenges Explorations
Phase I Financial and Scientific Report

Phase I Final Scientific Report

I. Activities

Our proposed hypotheses:

1. Real-time verification of rainfall cross referenced with farmer input will provide a more consistent verification of satellite inferred rainfall than only one method alone.
2. Combining crowd-sourced and field sensor data together with satellite imagery at the start of each season will help calibrate satellite rainfall measurements for the upcoming season.

To test these hypotheses, we developed and refined our existing crowdsourcing and sensor systems, prepared for system deployment by visiting potential sites in Nagpur, India, deployed our observation system, and collected data.

<u>Development of crowdsourcing system</u>

We had previously developed an interactive voice response (IVR) system to make automated phone calls to farmers. This system consists of custom Python code running on Amazon Web Services (AWS) cloud servers interfacing with the Twilio (telecom aggregation service) API to make phone calls. These calls ask pre-recorded questions, and capture dial pad responses from the responder. In order to prepare this system for deployment in this study, we tested our system with test calls to both the US and India. During testing, we discovered that Twilio does not offer local numbers for outgoing calls in India, making any calls we make to India an international call from a neighboring country. This would be prohibitively expensive for the number of calls planned, so instead, we selected Awaaz as our IVR provider for this study. We tested our scripts automating the IVR calls using Awaaz's service successfully and purchased airtime in bulk. Awaaz was also able to administer replenishment of airtime minutes to villagers as an incentive to participating in our surveys.

<u>Development of sensors</u>

We refined an existing device design to include rainfall, temperature, and humidity sensors. We also added a solar panel to increase sensor lifetime and selected a 2G/3G cellular communications module for connectivity. Along with our partner at the University of Delaware, we tested the performance and ruggedness of our sensor compared to commercial weather stations in the Delaware Environmental Observation System (DEOS) network. After several design revisions, we developed an installation procedure for our deployment in India according to appropriate World Meteorological Organization (WMO) standards. We also used the same standards to inform our sensor enclosure design, minimizing heat soak (causing reporting of higher than ambient temperatures) and maximizing rain gauge catchment accuracy.

Figure 1: Sensor testing at DEOS site. Our weather station is highlighed in the foreground while the DEOS station is in the background

As we finalized our design, we manufactured 100 of our weather stations, using in-house engineers, fabricators, and fabrication facilities. Simultaneously, we selected several vendors to provide a cloud-based pipeline for data storage and post-hoc analysis.

Finally, we partnered with Orange S.A., Sierra Wireless, and Nokia to test EC-GSM, a next generation global cellular protocol specifically designed for low-power, low-bandwidth, rural IoT applications such as ours. We demonstrated our prototype device at Mobile World Congress 2017 and received positive feedback from manufacturers and telecom operators.

Preparations for deployment

We partnered with Dr. Pawan Labhasetwar, head of the Water Technology and Management Division at the National Environmental Engineering Research Institute (NEERI), a government research center located in Nagpur. In addition, we partnered with Sharada Consulting Services to help provide in-country logistical support and project assistance for the deployment.

To prepare for our deployment covering the 2017 kharif growing season, our team visited India in March 2017 to test and refine our deployment procedures. With input from our partners, we identified Vidarbha as a target region due to the large concentration of poor farmers in need of better risk management practices. We worked with Sharada to identify 100 target villages--covering a variety of farm sizes, crops, and climate patterns.

In the second half of our pre-deployment trip, we visited several farming villages to develop procedures for farmer enrollment and further refined our deployment strategy. Finally, we tested sensor connectivity by making sure our sensors could communicate with our cloud servers from the field.

Deployment

We worked with Sharada and the Nagpur district agricultural officer to contact the sarpanch of each target village and prepare them for the team's visit. Simultaneously, we trained a team of 12 enumerators on how to enroll farmers and install sensors. Once our sensors arrived and were assembled, our team of enumerators began traveling to villages enrolling farmers and installing sensors.

Figure 2 map showing village locations with sensors and enrolled villagers

In each village, our team would do the following:

- Meet with the sarpanch and at least 10 farmers at a common meeting space
- Introduce who we were and explain the project's goals
- Enumeration team members broke off to enroll farmers and administer surveys. Survey responses were captured using ODK running on android cell phones provided to each enumerator
- With the sarpanch's help, identify a farmer whose field would house the sensor, travel to that farmer's field, install the sensor, and administer the survey
- Sensors were installed in or very near to the field and sited according to WMO rain gauge siting standards.

Data collection and analysis

From June 2017 to October 2017, we called all of our enrolled farmers with weekly calls and collected their responses. Separately, we captured weather observations on our cloud server from each sensor every 4 hours. Monthly airtime top ups were sent through Awaaz to villagers who participated in our surveys. For satellite rainfall, we downloaded precipitation data from the Climate Hazards Group InfraRed Precipitation with

Station data (CHIRPS). We collected the rain gauge readings from the only weather station in the region, located at the Dr. Babasaheb Ambedkar International Airport, 8 km from the Nagpur city center. These readings were extracted from Weather Underground.

II. Outcomes and Outputs

As part of our data collection:

- Enrolled approximately 1,000 farmers from 71 villages in the Nagpur district to participate in our IVR system. We placed 50,000+ phone calls and have approximately 11,000 reports from villagers who answered at least one of the pre-recorded questions about rainfall.
- Recorded farmers' survey results indicating farm size, crops planted, household size, age, education level, irrigation type, best/worst years, farming practices, insurance usage, and weather forecast usage patterns.
- Installed sensors in ~70 villages (see Figure 2) and have approximately 10,000 sensor observations from the region over the growing season. Observations include rainfall, temperature, and humidity readings.

<u>Results and discussion</u>

First, we evaluate the performance of our rainfall data by computing the daily rainfall from our sensors averaged over all the villages and compare it against the daily rainfall measurements from the Nagpur airport weather station. This station is by the Indian Meteorological Department.

In Figure 3, we observe that our average readings across the region generally follow trends measured by the airport weather station–indicating that the sensors are well calibrated to measure ground truth. Some divergence is to be expected as the airport data are a gathered from one location while the sensor data shown are averaged across the entire Nagpur district.

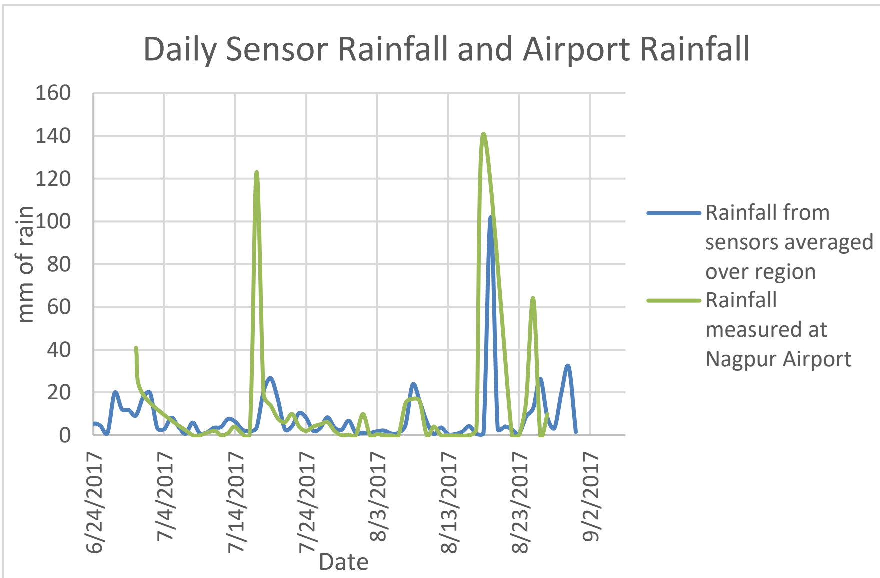

Figure 3: Sensor measurements compared to weather station data at the Nagpur airport

In addition to verifying that our sensors match local rain gauge trends, we can also see that trends in crowdsourced data (measured as reported number of hours rained per day) closely match our sensor data in Figure 4. Overall, there is reasonable convergence among our collected data and local rain gauge data. Since the crowdsourced data show trends in common with the sensor

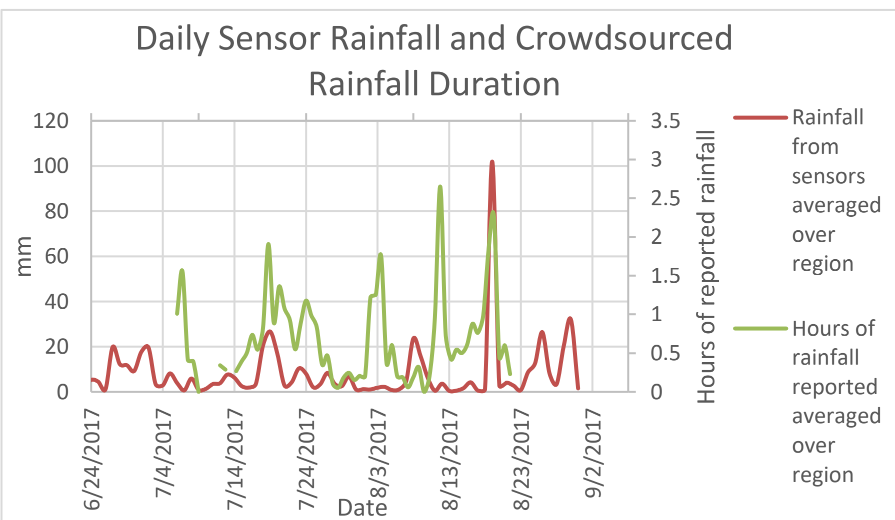

Figure 4: Sensor rainfall compared to reported hours of rainfall from crowdsourcing

data, we may be able to use the reported rainfall to reduce the number of sensors required.

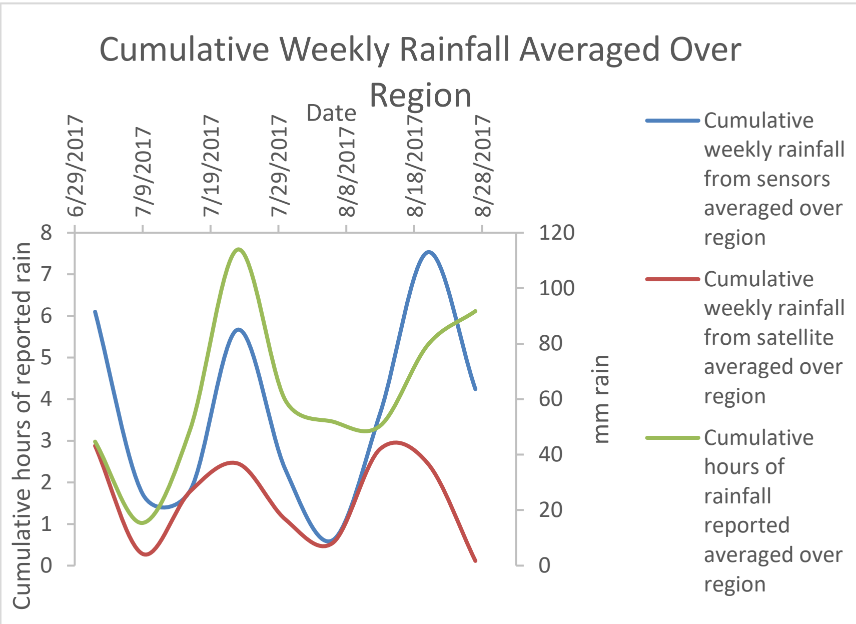

Figure 5: Weekly sensor, crowdsourced, and satellite results

Figure 5 plots total weekly rainfall for sensor, crowdsourced, and satellite data averaged across the region, where sensor and satellite data are in units of total mm, and crowdsourced data are in units of total reported hours– demonstrating, again, a reasonable fit to observed trends.

Given that both sensor and crowdsourced data are in line with the weather station and satellite data, our next aim is to show that our data better reflects the variation in microclimates that satellite data alone can not.

To evaluate the ability of our system to observe microclimates, we compute the standard error of daily rainfall across all village locations where we collected sensor data. Then, we compute the corresponding standard error of the satellite data over the same region. In this case, we use the standard error measure, not in the traditional use of measuring sample error from mean. Rather, we use standard error to indicate the *amount of variation* in rainfall over the region at a given time. The higher the standard error, the more varied the rainfall over the region. The lower the standard error, the less varied the rainfall over the region. By comparing the standard error of our measurements to the standard error of satellite-derived measurements, we can quantify how much more variation (from microclimates) our data contains compared to satellite data alone.

As shown in Figure 6, the variation in the sensor readings are higher than the satellite readings for most days, likely indicating that our data contain more information with regard to actual weather variability than satellite data do alone. <u>This validates our first hypothesis and demonstrates the effectiveness of our proof-of-concept.</u>

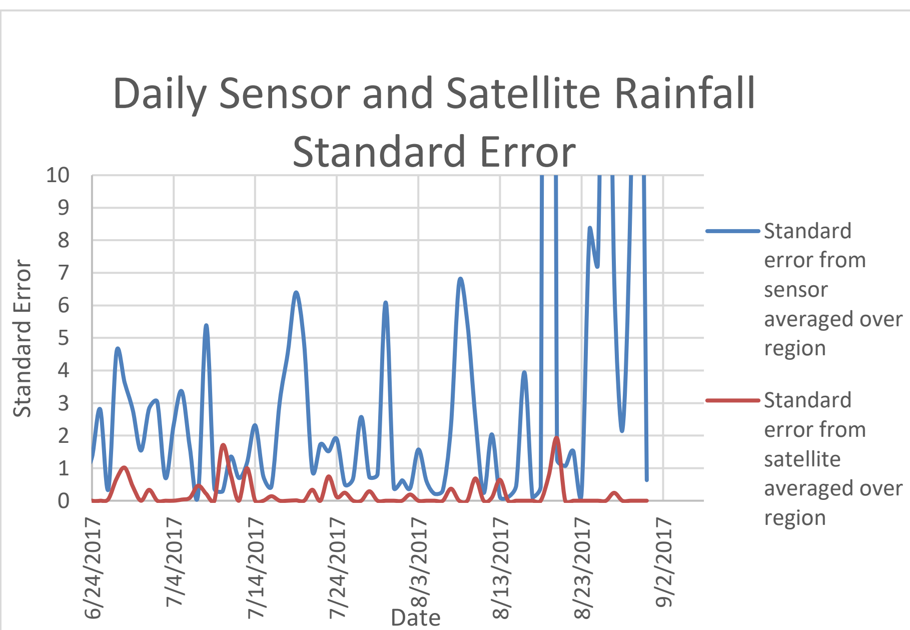

Figure 6: Sensor and satellite standard error

Note that, in this case, higher variation in rainfall measurements is a positive outcome. Increased variation captures the variability of microclimates that satellite data obfuscate because of their coarse resolution. We observe that the standard error for satellite-derived rainfall estimates is often near or at zero (generally corresponding to days where the satellite detected no rainfall in the region) and yet, our data exhibit rainfall variation across the region during the same time period.

This is most likely caused by precipitation falling in only a small area of the region that is not captured by the satellite. Furthermore, low variation seen in the satellite data does not match known monsoon rainfall patterns. Nor does it match anecdotal reports we gathered stating that the villages in the west routinely experience different rainfall patterns than villages in the east. Indeed, smallholders in this region have adapted their cropping patterns changes to match these variations, which are contrary to satellite data.

In many ways, the observations noted above match our understanding of satellite rainfall measures—they are generally optimized to provide sub-national monthly rainfall measurements. Accurate weather forecasts depend on granular weather measurements often lacking in the developing world[1]. As a result, the geographical and temporal resolution of satellite indices are too coarse to help smallholder farmers. They do not capture microclimate effects and monsoon onset, which are critical to smallholder farmers.

III. Challenges

- We enlisted 1,000 farmers (10 farmers per village) for IVR crowdsourcing calls. In each village, 5 farmers were called on Monday and Friday regarding the previous 2 days of rainfall, and 5 farmers were called on Wednesday and Sunday for the previous two days of rainfall. Via touch-tone, they responded to the following questions: Did it rain the last two days? How many hours yesterday? And how many hours the day before? This led to approximately 2,000 calls being made each week, for a total of 8,000 calls per month. Overall, the response rate to all questions remained at around 2,000 per month–a 25% response rate. This response rate increased slightly after two months, perhaps because the high frequency responders were continuously rewarded with air time. While we do not verify truthful reporting, this is something we would propose to address and test in Phase II.
- Shipping our sensors from the US to India proved to be challenging as our hardware was held by Indian customs for approximately 2 weeks. As a result, we had to pay a large duty penalty and had significantly less time than originally planned to complete deployment before the rains began. Although we were ultimately able to install sensors in 70 villages, we had to retain our enumerators and in-country logistics for much longer than anticipated, costing us additional funds.
- We originally proposed to place our sensors across 10 villages. However, during our pre-deployment meeting, we decided with our partners to increase the number of villages to 100. This meant that deployment would take significantly longer than originally planned. Combined with a reduced timeframe described above, this increase in scope required sensor installation by trained enumerators without staff supervision resulting in potentially faulty installation and erroneous survey data. Although we required photographs to be taken at each village, it was not possible to verify that all deployment procedures were followed. In the future, we may not need to install as many sensors to cover the same area.
- Once deployed, we had difficulty maintaining sensors that were not connecting to our cloud server. Once again, due to customs difficulties, we had significantly reduced training time for our in-country support team--leading to subsequent difficulties with sensor maintenance. In the future, we anticipate that this issue of connectivity will be solved by dedicating more resources towards monitoring and sensor service.

IV. Other Sources of Support

Source	**Amount** *(including in-kind)*	**Type**
None	$0.00	N/A
TOTAL	$0.00	

[1] S. Mourtzinis, et al., (2017). From grid to field: Assessing quality of gridded weather data for agricultural applications. *European Journal of Agronomy*, Vol 82(A), pp. 163-172.

Phase I Financial Report (Maximum 1 page)

Organization Name	Groundtruth, LLC
Project Title	Multimodal satellite rainfall measurement for index insurance
Report Type	Final
Reporting Period	May 1, 2016 — April 30, 2018

Project Expenses by Type	***Actual Expenses (USD)***	***Actual Expenses (USD)***	***Actual Expenses (USD)***	**TOTAL EXPENSES (USD)**
	2016	**2017**	**2018**	**2016-2018**
Personnel	7,500.00	42,510.00	0.00	**50,010.00**
Equipment*	0.00	0.00	0.00	**0.00**
Travel	1,416.41	8,583.59	0.00	**10,000.00**
Consultants	0.00	4,990.00	0.00	**4,990.00**
Supplies	423.01	24,576.99	0.00	**25,000.00**
Subcontracts	0.00	10,000.00	0.00	**10,000.00**
TOTAL EXPENSES	**9,339.42**	**90,660.58**	**0.00**	**$100,000****

** If equipment expenditures total over $5,000, please list each item separately*

***** We encourage spending all grant funds for the charitable purpose of your grant. If total project expenses are less than $100,000 please email GCE@gatesfoundation.org for further guidance.***