Title
The Final Frontier: Embedding Networked Sensors in the Soil

Authors
N. Ramanathan
T. Schoellhammer
D. Estrin
et al.

Publication Date
2006

The Final Frontier: Embedding Networked Sensors in the Soil

Nithya Ramanathan, Tom Schoellhammer, Deborah Estrin, Mark Hansen,
Tom Harmon, Eddie Kohler, and Mani Srivastava

Center for Embedded Networked Sensing, UCLA

Abstract

This paper presents the first systematic design of a robust sensing system suited for the challenges presented by soil environments. We describe three soil deployments we have undertaken: in Bangladesh, and in California at the James Reserve and in the San Joaquin River basin. We discuss our experiences and lessons learned in deploying soil sensors. We present data from each deployment and evaluate our techniques for improving the information yield from these systems. Our most notable results include the following: in-situ calibration techniques to postpone labor-intensive and soil disruptive calibration events developed at the James Reserve; achieving a 91% network yield from a Mica2 wireless sensing system without end-to-end reliability in Bangladesh; and the *javelin*, a new platform that facilitates the deployment, replacement and in-situ calibration of soil sensors, deployed in the San Joaquin River basin. Our techniques to increase information yield have already led to scientifically promising results, including previously unexpected diurnal cycles in various soil chemistry parameters across several deployments.

1 Introduction

Soil ecosystems are complex, elusive, and still largely misunderstood: *Science* has called them the "final frontier" [10]. This paper presents the first systematic design of a robust sensing system for soils. We discuss three deployments we have undertaken over the past year, from a rice paddy in Bangladesh to the San Joaquin River basin in California. In these deployments, we experienced first hand the unique challenges that arise in working with soil systems. Soil processes are difficult to observe because i) long-lived, dependable sensors do not exist for many important modalities, and ii) below-ground soil characteristics (unobserved) introduce significant latent spatial variability in sensor data that can be difficult or impossible to adequately resolve. Thus, collecting every data point and ensuring it is usable is important because there is less data to be collected even in the best case.

In this paper we focus on techniques we have used to improve this quantity of scientifically usable data, which we call the *information yield*. We present experiences, lessons learned and data from each deployment, and evaluate techniques we designed to improve information yield in the context of each particular deployment. These techniques address:

- **Maximizing sensor yield.** We evaluate techniques to improve the *sensor yield*, the percent of data received from the network that is usable for scientific purposes.

  For example, *in-situ calibration* techniques can be used in the field to identify only those sensors that need to be removed and calibrated, avoiding premature calibration to minimize soil disturbance and identifying sensors that require immediate calibration to increase the amount of usable data delivered from the system.

- **Maximizing network yield.** We evaluate techniques to improve the *network yield*, the percent of expected data that is collected at the basestation. Our wireless deployment in Bangladesh incorporated a disruption-tolerant networking layer to maximize network yield. This deployment achieved a 91% delivery ratio without relying on latency- and power-consuming end-to-end reliability.

- **Maximizing interactivity.** We evaluate the impact of interactivity, or taking actions as soon as possible, on information yield. For example, collecting physical samples from the field for lab testing when sensors recorded questionable data enabled us to validate potentially faulty nitrate and chloride data and increase our sensor yield by 51% from these sensors.

  Based on our experience, we have designed the *javelin*, a platform for very wet soils or shallow groundwater that enables sensor interactivity by minimizing soil disturbance during deployment, and facilitating the deployment, replacement and in-situ calibration of sensors.

Due to the success of our techniques to increase information yield, the sensing systems we have deployed have collected measurements interesting for scientists. Previously unexpected diurnal cycles in various soil chemistry parameters across several deployments have led to further studies and investigation. Much work remains, but this paper presents the beginning in what we hope is a new direction for sensor system deployments.

The paper is organized as follows. Sections 2 and 3 discuss soil environments and our three case study deployments. Section 4 describes how we calculate the information yield and presents numbers for each deployment. The remainder of the paper focuses on the evaluation of novel techniques we have employed in our soil deployments to improve information yield. We discuss maximizing sensor yield in Section 5, network yield in Section 6, and real-time interactivity in Section 7.

Related Work Several recent papers have described other experiences with sensor network deployments. Tolle et al. describe a deployment measuring microclimate in a coastal redwood in Sonoma County, California [4], focusing primarily on multidimensional data analysis. The authors observe that their sensors—temperature, humidity, and photosynthetically active radiation (PAR)—required little calibration beyond that performed by the manufacturers, and that real-time monitoring of network quality could have helped to improve their network yield, which was 49%.

Werner-Allen et al. describe a seismic deployment on a volcano in Ecuador [5]. Their network yield, 51%, was similar to the redwood deployment's, although in the absence of certain systematic failures it might have been higher. Some of their sensor network data is compared to data collected in more conventional ways, allowing them to evaluate data fidelity. The authors evaluate event detection accuracy, finding that their sensing system detected 5% of the volcanic eruptions that took place. Low detection

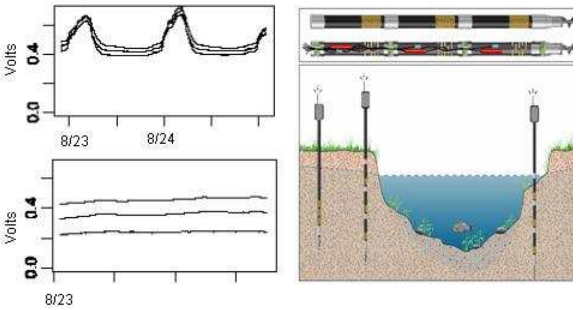

Figure 1: San Joaquin River Deployment: The three curves in the top and bottom panel on the left correspond to observations at one, two and three feet below the San Joaquin riverbed (right). The two sensor "stacks" were positioned within feet of each other.

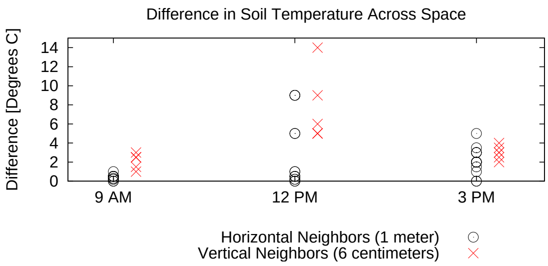

Figure 2: Temperature at James Reserve, CA: Difference between temperature readings from sensors that are i) 1 meter apart at the same depth, and ii) 6 cm apart, at 2 cm and 8 cm depths, at the same surface location.

accuracy is attributed to parameter settings. They also evaluate data fidelity, finding that selected acoustic and seismic signals are consistent with ground truth and expected readings.

Both of these deployments utilize sensors that are fundamentally more reliable and require less maintenance than the sensors used in soil deployments. The unique challenges that arise in working with soil systems have driven the design of our system and our deployment experiences in different directions.

Musaloiu-E et al. report on an end-to-end soil deployment undertaken in Baltimore involving soil temperature and moisture sensors. The authors describe a sensor calibration process undertaken before deployment, but only briefly mention data faults and do not focus on issues relating to information yield. Instead, they focus their analysis on energy consumption, database design and post-deployment analysis infrastructure.

2 Difficulty with Soil Deployments

Soils are fundamentally different from most environments, impacting the design of a sensing system deployed in this environment.

Soils exhibit spatial heterogeneity of physical and chemical properties even at small scales. Thus, measurements taken from proximal locations in soil are often not redundant, and can even differ significantly in structure and amplitude. In Figure 1, for example, we present data from a series of nitrate sensors; the top panel contains two days of data, and the bottom panel contains only data for the first day. The three curves in each panel correspond to observations at one, two and three feet below the San Joaquin riverbed. While the two sensor "stacks" were positioned within feet of each other, we notice very different diurnal patterns in the data. The full deployment consisted of 6 such triples, with roughly half exhibiting each pattern.

In part, this behavior is to be expected. While mathematical models can be formulated to describe, for example, the diffusion

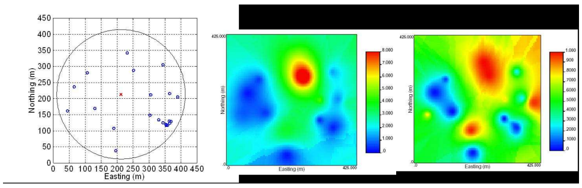

Figure 3: Output of a Kriging model based on data collected from one of our soil deployments in Palmdale, CA. The left panel shows the monitored locations in the field, and the right two panels are two possible views of the soil output from the model based on two different settings of a parameter controlling uncertainty thresholds.

of chemicals through "ideal" soils, measurements of real soils to determine the parameters necessary for these models reveal unpredictable variation even across short distances. Figure 3 demonstrates how the resulting uncertainty in parameter choice can change a scientist's view of the soil. This argument applies to a variety of other sensing modalities. Figure 2 shows heterogeneity of temperature measurements, this time from stacks spaced 1 meter apart in a 3×2 grid with temperature sensors at 2 cm and 8 cm below the surface. The measurements are taken from this patch on a typical day in October. The figure is a scatter plot of the differences between temperatures in neighboring locations: circles correspond to differences between temperature sensors at the same depth but 1 meter apart in the grid, and crosses correspond to differences between temperature sensors at the same location but separated by 6 cm in depth. Temperature differences up to 14 °C can be seen between sensors separated by 6 cm.

In many environments, this kind of spatial variability can be addressed with some form of dense sampling. We might literally deploy a large collection of fixed sensors, or instead, move a small number around with a robotic system. A class of so-called Networked Infomechanical Systems [2] has proved successful in lakes and streams, for example. In soils, however, dense sampling is too invasive, disrupting the very systems we would like to observe. The uncertainty associated with modeling heterogeneous physical processes using point measurements poses a distributed signal processing challenge that is as interesting but quite different from the more common acoustic and seismic applications that are perhaps more familiar to the sensor network community.

To further complicate matters, many of the sensors used in contaminant tracking are short-lived, prone to faults and require frequent calibration. For example, the sensors mentioned in connection with the data in Figure 1 employ ion-selective electrodes (ISEs), a class of so-called active sensing elements. In this case, an electrical potential is generated when there is a difference in ion concentrations between an internal reference sample and the soil being tested. A chemically treated membrane acts as a filter for the specific ion being measured. These membranes, however, are not field robust and are the major source of reliability problems for ISEs, resulting in short, error-rich deployments.

One standard approach to measurement reliability problems involves deploying redundant (dense) systems. As mentioned above, however, this strategy is not feasible for soils. This is not to say that techniques do not exist to increase a user's confidence in the data; for example, physical soil samples can be extracted and analyzed in a lab to perform point checks of sensor readings; but such techniques are also not fool proof. In addition, sensor maintenance involves removing sensors from the ground. This process is disruptive to data collection as soils require anywhere from a day to several months to settle once disturbed, during which time the data from the sensors are not always usable.

To summarize, soils present a challenging new frontier for embedded sensing. The phenomena under study exhibit complex spatial variability, and the sensors can be noisy and unreliable, forcing relatively short-term deployments.

3 Soil Deployment Case Studies

In this section we describe three soil deployments studied throughout this paper: the initial James Reserve deployment, and follow-on deployments in Bangladesh and the San Joaquin River basin. Each deployment improved upon the previous either in robust design or functionality.

James Reserve The purpose of the transect at the James Reserve is to explore the spatial and temporal scales at which sub-surface measurements should be taken, and to study the relationship between soil CO_2 fluxes and moisture and temperature conditions in the soil. The network has collected over 7 million points since October 2005.

This transect spans approximately 80 meters. At each of 10 sites, 13 above and below ground measurements are taken. Above ground, air temperature, relative humidity, barometric pressure, and photosynthetic active radiation (PAR) are measured. Below ground, temperature, moisture, and CO_2 concentration measurements are taken at depths of 2 cm, 8 cm, and 16 cm in the soil. Sensors are connected to Campbell scientific dataloggers, and powered using deep cycle marine batteries that can last several months before requiring replacement. More recently, we augmented this deployment with temperature and moisture sensors that communicate over a wireless Mica2 network in order to test sensor placement and experimental methodology in soil environments. One of the primary reasons that this deployment has been able to collect data for over a year is that it does not include ISEs.

The dataloggers in this transect are not equipped with wireless radios, and store data locally. Thus, problems can persist for long periods of time before they are fixed.

Bangladesh In January 2006, we deployed a wireless sensing system in a rice paddy in Bangladesh (Figure 4) to help scientists evaluate the relationship between irrigation and arsenic contamination in the groundwater [11]. Tens of millions of people in the Ganges Delta drink well water impacted by arsenic, a massive environmental poisoning projected to cause approximately 3,000 deaths per year [13]. The experiment was designed and deployed with scientists and civil engineers from the Bangladesh University of Engineering and Technology and MIT. We deployed 42 ISEs to monitor ammonium, calcium, carbonate, chloride, pH, oxidation-reduction potential, and nitrate, and 8 soil temperature, moisture and pressure sensors at 3 different depths in 3 locations. The network collected 26,000 measurements over a period of 12 days. This deployment was short-lived because it primarily relied on ISEs.

Influenced by the problems at the James Reserve, in Bangladesh we employed a wireless network of sensors that provided real-time access to data and network parameters. In order to improve the amount of usable data collected by this wireless network we incorporated a delay-tolerant networking (DTN) layer into our network stack for reliable data delivery. This DTN layer was quickly put to use when after the first day in the field, the landowner informed us that our basestation ran the risk of being stolen if we left it in the field over night. Without a networking layer tolerant to basestation absences our network would not have captured much as most of the diurnal activity took place in the early hours of the morning. In Section 6 we discuss how our DTN layer enabled a 91% network yield even though the basestation was absent more than half the time.

The most surprising discovery from this deployment was the diurnal variations observed in ammonium, chloride, and carbonate (graphs in Figure 4). While data flattened around day 7 as a result of a scheduled irrigation event, the diurnal trends (also seen in hydraulic parameters) indicate that diurnal, possibly plant-induced, processes may be important in the mobilization of arsenic. The scientists are returning to the field in December 2006 to further study this phenomenon. We will join them in 2007 to deploy a more extensive and robust wireless sensing system.

While the wireless connectivity enabled real-time interaction with the network, allowing us to find and fix problems when they occurred, other problems remained. A pylon in this deployment had up to 24 sensors, so deploying a single pylon took all day, and replacing a faulty sensor or moving a pylon was nearly impossible once deployed.

San Joaquin The purpose of the deployment in the San Joaquin River was to characterize the transport and mixing phenomena at the confluence of two distinctly different rivers: the Merced River (relatively low salinity) and the agricultural drainage-impacted San Joaquin River (relatively high salinity). Soil measurements were supplemented with measurements in the river taken by an autonomous robotic node [2]. Six sets of 3 nitrate ISEs connected to Hobo dataloggers (Onset computers) were deployed at one foot increments below ground, alongside soil temperature and salinity sensors, in the first week of August, 2006. 48,000 measurements were collected from the nitrate sensors over this 5 day deployment. The key to this short-term deployment's success was the *javelin* (Figure 1), a sampling platform designed to ease the type of deployment effort we experienced in Bangladesh. Javelins enabled each set of sensors to be easily deployed at multiple depths, and redeployed in multiple locations over the short deployment. This spatial coverage enabled by the javelins was especially useful in that region because several kilometers of homogeneous soil was separated with random patches of heterogeneity. Section 7 discusses the javelin further.

Interestingly, the most surprising discovery in this deployment was again diurnal trends, though this time in nitrate data (graphs in Figure 1). The scientists are unsure about what could be causing these trends when a second array of sensors just a few meters away showed no such fluctuations. Others have noticed similar patterns in river nitrate and suggested that this may have been caused by photosynthetic activity [3]. However, the diurnal behavior here is in the sediments beneath the river and the peaks are synchronized, suggesting that a sudden fluctuation in river water concentrations is not the cause.

4 Calculating Information Yield

Given these experiences, we will spend the remainder of the paper discussing techniques we employed to improve the information yield in each of these deployments. In this section we describe how we calculate *information yield*, which is the percent of received data which are usable for scientific purposes. Information yield is made up of the *network yield*, the percent of expected readings received, and the *sensor yield*, the percent of received readings that are usable.

Data are classified as *usable* if they fall within the operational range of a sensor, or the concentration range where the sensor is most capable of distinguishing between concentrations. It is worth mentioning that while data in this range are usable for analysis, they are not equivalent to verified, quality measurements; bias or other faults could still impact the readings. The operational range is defined through *calibration*, the process of mapping a sensor's measured output to an estimate of the property being sensed. The calibration for many sensors does not change over time, and in these cases we can use the manufacturing supplied calibration equation to obtain this operational range.

For example, CO_2 concentration is reported in parts-per-million (ppm) and obtained using $(625S - 2500)/CPT$, where S is the

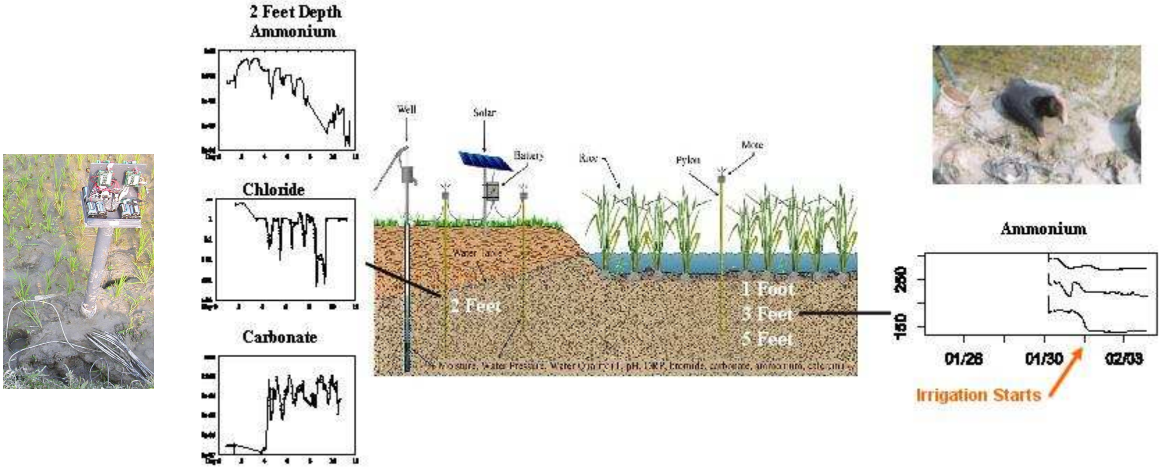

Figure 4: Bangladesh Deployment Image of soil pylons deployed in February of 2006 in a rice field in Bangladesh. Diurnal variations in ammonium, calcium, and carbonate at a depth of 2 feet (left graphs), and in ammonium 1, 3, and 5 feet (right graphs) are apparent Figure by Jason Fischer, UC Merced. Left photograph is of a soil pylon, and right photograph is of the sensor deployment process.

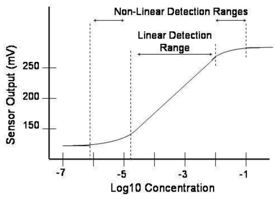

Figure 5: Idealized calibration curve. Linear and Non-linear detection ranges are labelled.

sensor output, C is a constant, P is pressure and T is the temperature. It is impossible for CO_2 concentration to be negative, so in order to find the lower bound for the sensor's operational range we set the equation to 0 and solve the equation for S. Scientists discard all data that occur below this threshold. A similar process was used for most of the sensors in our deployments.

The process for identifying the operational range for ISEs was different. This process is based on the physical limitations, or sensitivity, of the sensor, instead of the physical limitations of the phenomenon being measured. In addition, we could not use the factory calibration for ISEs because their calibration changes over time and must be updated.

As with most sensors, ISEs are calibrated by exposing the sensor to a range of standard concentrations to identify the function that relates the electrical potential output of the sensor to the observed phenomena. For ISEs, the sensor output is plotted against the logarithm of the concentrations used in the calibration. The resulting calibration curve takes the form of a stretched out S (Figure 5). This curve contains a *linear detection range*, which covers the range of concentrations where the sensor responds linearly. The calibration equation used to translate sensor output voltages to concentrations is defined by the slope and intercept of the line in the linear-detection range. The linear detection range is bounded above and below by a *non-linear detection ranges* (NLDR), characterized by the range of concentrations where the sensor responds non-linearly to changing concentrations so the slope decreases. Error associated with readings increase as the slope decreases [12], thus readings in the NLDR have lower associated confidence. The sensor is not sensitive to concentrations above and below the NLDR, and so the slope for the calibration curve approaches zero in these regions.

Scientists define usable data from an ISE as data that occur inside the linear-detection range, and data in the NLDR that can be validated [12].

ISEs are currently the primary method of obtaining time-series concentrations of contaminants in water and soil environments. Nevertheless, they are fragile and require extensive care. Given the tedium associated with these sensors, the question arises: Why use such unreliable sensors? Some of the issues inherent to these sensors will be overcome over time, but the sensors are not likely to significantly improve in the next five to ten years. Moreover, there will always be new and less reliable sensors that require careful monitoring. Many of the techniques suggested in this paper can be applied to a general class of sensors. Finally, as the number of sensors in standard deployments increase, the manual labor required to make highly reliable deployments could become the limiting factor in deployment size. Systems solutions to these problems that can scale with the size of the deployment may allow deployments to scale up without sacrificing data quality.

Yields Using the operational ranges defined for each sensor, we calculated the overall information, network and sensor yields for each deployment (shown in Figure 6). The deployment at James Reserve had the highest information yield at 87%, compared to 52% at San Joaquin, and 59% at Bangladesh (first set of bars in the figure). The network yields (second set of bars) were all close to 100%, so the information yield is primarily dictated by a deployment's sensor yield. The sensor yield (3rd set of bars) was primarily dictated by the ISE yields. This is easy to see if we look at the sensor yield separately for different sets of sensors: moisture and temperature sensors (4th set of bars in the figure) had a yield of almost 100%, while the ISEs (last set of bars in the figure) had yields closer to 55%. Thus, it is not surprising that the deployment at James Reserve, which did not use ISEs, has a much higher sensor yield, and as a result a much higher information yield, than the other deployments.

Best Practices We highlight several best practices. First, sensor bias varies with hardware, and each data acquisition board has its own bias factor. Thus, sensors should be calibrated with the entire data acquisition system (e.g. the mote and sensor board) that will be deployed in the field, not just the sensor. Second, ISEs must be calibrated before, after, and even during a deployment, depending on the duration, because their calibration parameters change over time. Third, we designed an end-to-end check during the deployment to ensure that nothing had significantly changed during the rough deployment process. After digging the hole, we

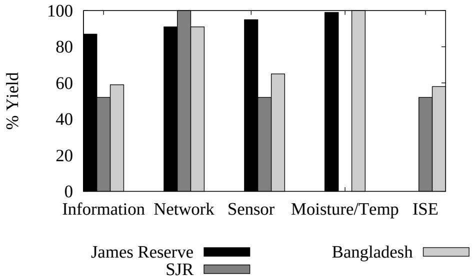

Figure 6: Average Yields Average information, network, and sensor yield for all three deployments. Sensor yields for moisture and temperature sensors are also presented separately from ISE yields. Moisture and temperature sensors (shown separately) have much higher sensor yields than ISEs. San Joaquin River had no moisture/temperature sensors, and James Reserve had no ISEs.

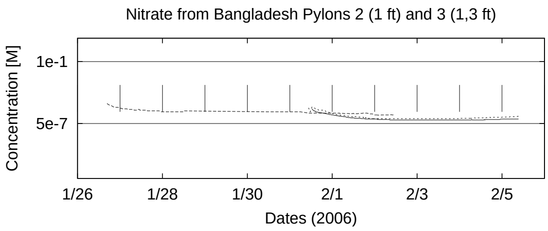

Figure 7: Nitrate data taken from three different locations in Bangladesh. Vertical lines cover the linear detection range, and horizontal lines delineate the extent of the NLDR. Most of the data occur within the NLDR. Nitrate data from remaining three sensors not shown because a majority of them were not usable.

stick the sensor in the water that pools at the bottom and take a measurement, and then take another measurement immediately after the sensor is buried. We expect all sensors of the same type to report the same concentration because they are all measuring the surface water that has pooled in the hole at that point. We also expect the measurements taken before and immediately after the sensor is buried to remain relatively constant.

5 Improving Sensor Yield

The first dimension to improving information yield is to maximize sensor yield. Broadly speaking, there are three classes of problems we observed to reduce sensor yield. In the first instance, data occurs in a sensor's non-linear detection range (NLDR) and may or may not be usable, requiring further validation. In the second instance, sensors produce readings that do not reflect reality (e.g. a wire breaks). In this case, the faulty hardware should be fixed. In the third instance, a sensor slowly transitions from producing usable data to producing readings that are difficult to interpret (i.e. calibration drift). In this case, the sensor should be calibrated. In the following three subsections we discuss these three classes of problems that occurred in the field, and the actions we took in the field to either fix or validate them.

5.1 Validating Questionable Data

Data that occur in a sensor's NLDR are typically discarded, as they can indicate a problem with the sensor. However, it is also possible that the sensor is not faulty and that the ion concentration is truly outside of the sensor's linear detection range. Much of the data we collected from chloride and nitrate sensors in Bangladesh occurred in the NLDR of the sensor and fell into this second category. Figure 7 is a graph of data from three nitrate sensors; the vertical lines indicate the linear detection range, and the horizontal lines delineate the NLDR. In order to determine if the data were usable, the scientists we were working with in Bangladesh extracted several physical soil samples for lab analysis. We used the results from this analysis in conjunction with a computer model of the soil chemistry for that region to confirm that the levels for nitrate and chloride were expected to be in the NLDR of the sensor. As a result of this lab analysis, the scientists were able to use this data. This conclusion is validated by the data we collected. Of the 3400 data points recorded in the NLDR of a sensor, 2850 of these points are from either nitrate or chloride sensors; i.e. most of the data in the NLDR are from sensors measuring concentrations that we expect to fall in this range. In addition, most of the points recorded from a nitrate or chloride sensor in the NLDR (2000 of the 2850 points) were corroborated by at least one other sensor of the same type also reporting data in the NLDR. This is further validation that the sensors were not faulty and in fact representative of the environment.

5.2 Sensor and Hardware Faults

Large Time Gradients Large time *gradients* in the data, or large changes in sensor output with respect to time, are usually not representative of the environment because: i) physical phenomena are limited by natural laws, so we can place an expectation on how rapidly they are able to change, and ii) the sensor is not capable of measuring large magnitudes of change over a short time period. Thus, time gradients usually indicate a problem in the sensor or hardware.

We have observed large gradients in data from all of our deployments. In most cases, a gradient only spans several points and the data is simply discarded. In data collected in Bangladesh, large gradients persisted for up to several days in some instances, and often indicated a problem. In one instance, we found an exposed sensor connector from one of the sensors reporting high time gradient data sitting in a muddy pool in the field. Moving the cable and connector to a dry enclosure addressed the problem. However, a wet connector was not the cause of all of our high time gradient data. The top panel in Figure 8 is a graph of data from an ammonium sensor with high gradient data for 2 days starting on February 2.

In order to better understand this phenomenon, after returning from Bangladesh we redeployed part of the system we had deployed there, including the sensor in Figure 8. In one of these deployments, a little after hour 2 (bottom panel in Figure 8) we noticed large gradients in the data from the ammonium sensor. We disconnected the sensor from the sensor board (captured in the data stream as the sharp peak immediately after the dip in readings around hour 2.5) and connected it to a pH meter, an independent meter used in the field to measure the sensor output. The meter verified the large gradients reported by the mote, and led us to discover that the output was caused by an electrical short in the internal sensor wiring. By contorting the sensor cable we were able to temporarily fix the short. The problem recurred several times during this day long deployment (a little after hour 3, just before hour 6, and just before hour 7 in the graph), and each time was temporarily fixed by adjusting the sensor cable. Several of the sensors that reported data with high gradients have been sent back to the manufacturing company for investigation.

Stuck-At Value The *stuck-at* fault represents a sensor getting stuck at a particular value. Often this is a value at the high or low end of the sensor's operational range. These faults are dangerous because the measurement can tell you nothing about the underlying phenomenon. Yet, when they are in-range, simple out-of-range detection does not help [1]. In our James Reserve transect, at least some of the stuck-at faults were easy to identify, as the values occurred outside of the sensor's operational range. For example, a soil temperature sensor connected to node 7 reported 27,000

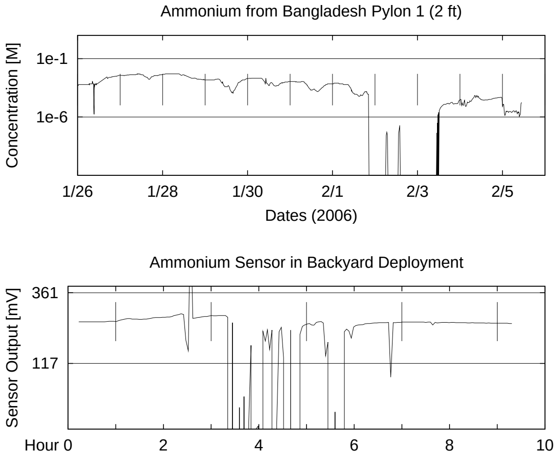

Figure 8: Top panel is a plot of data collected from an ammonium sensor in Bangladesh. Faults in the form of excessive time gradients beginning around Feb. 2 for two days are easily identified visually. Bottom panel is a plot of ammonium data that captures a similar time gradient fault observed during a deployment in our backyard.

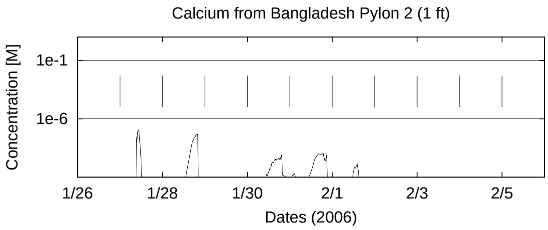

Figure 9: All data from this calcium sensor are outside of the NLDR, indicated by the horizontal line, and thus not usable.

continuous measurements of either -91.9 or $-89.4°$C from April to August, 2006. This sensor should have been replaced.

Broken Sensor Deployments are a chaotic and rough process, and sensors are not always as field-robust as we would like. One example is a calcium ISE deployed in Bangladesh which had reported several consecutive days of data well outside of its operational range (Figure 9). In order to identify the cause of the problem, we connected the sensor to a pH meter. The readings from the pH meter corroborated the data returned by the mote, isolating the problem to the sensor. Ideally we could have replaced this calcium sensor with another one. However, given our pylon design, it was too difficult to replace just one sensor (or the whole pylon for that matter); moreover, deploying and redeploying sensors was so labor intensive and destructive that instead we decided to leave the sensor in place and hope that it improved (which it never did). Further evidence that the sensor should have been replaced came during the post-deployment calibration of this calcium sensor, where the sensor displayed little to no reaction to changing concentrations; i.e. the slope of the calibration curve was essentially 0, evidence of a faulty sensor. While some sensors failed the check we did immediately after placing them in the ground (e.g. two chloride sensors that were wired in reverse), this calcium sensor was not in this group. We believe that the membrane for this sensor was damaged during or immediately after the somewhat rough deployment process. This sensor is currently being investigated by the manufacturing company to isolate the problem.

We should have replaced this sensor after receiving a 0% sensor yield for several days. This conclusion is supported by our data as well. 11 of the 42 ISEs reported a sensor yield of less than 25%, and all of these sensors reported data outside of their operational range for at least 2.5 consecutive days. Thus, replacing a sensor that was not working for 2.5 days would have targeted sensors that ultimately did not collect very much usable data.

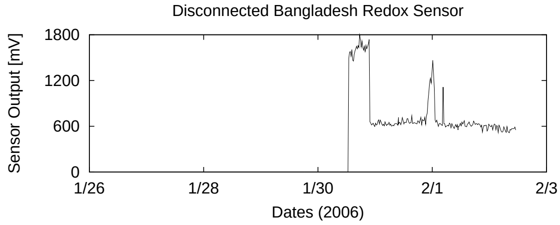

Figure 10: Data from an oxidation-reduction potential (redox) sensor is 0 millivolts until the morning of Jan 30, when the connector breaks, and the board reports noise sampled from the disconnected port.

Disconnected Sensor Often sensor connections are not secure, and sensors can become intermittently or persistently disconnected. This disconnection manifests differently on different data acquisition hardware. In Bangladesh we used MDA300 sensor boards, which report all data from a port, regardless of the status of sensor connection. In one instance, we noticed that data from an oxidation-reduction potential sensor was suddenly extremely noisy (Figure 10). Upon checking the sensor with our pH meter we discovered that the sensor's BNC connector was broken. Unfortunately, both fixing the connector and replacing the sensor were too difficult given our platform design, so we left the sensor in place.

In the James Reserve deployment, sensors are connected to Campbell dataloggers. In most instances, the datalogger will report a null value if the sensor is disconnected, making it easy to detect such problems. 211,545 data points, or most of the 362,000 faulty data points collected at James Reserve, were caused by such a disconnected sensor. In the case of a CO_2 sensor, if the sensor is not properly powered, the datalogger reports a default value of .0025, which occurred 93,000 times in this deployment, accounting for 25% of the faulty data collected in this deployment.

In most cases, if a sensor appears disconnected for over a day, the data, power, and ground wire connections should all be checked.

Low Battery The hardware component that translates analog sensor signals to digital values (the ADC) requires a minimum battery voltage of 2.7 V for correct operation. Battery voltages below this level have been observed to impact data quality [4, 7]. On Mica2 based systems, this requires monitoring the battery level directly, since the CC1000 radio can continue to transmit data even when the battery voltage is at 2.2 V. Thus, the Mica2 may continue to transmit data, but it may not be usable.

We encountered this problem in Bangladesh, where 40 minutes after being deployed, the battery for mote 11 plummeted from 3.2 V to 2.4 V over a period of 20 minutes. Unfortunately, replacing the battery did not improve the rate of faulty data recorded from this mote. With an overall sensor yield of only 35%, this mote reported 352 faulty data points a day (as compared to the median value of 205 faulty data points per day from all motes), the highest rate of all motes in that deployment. In addition, almost half of all data with high time gradients was collected from this mote. We believe that an electrical problem on the board caused the sudden drain in battery voltage in addition to the low sensor yield.

Lightning Dataloggers are vulnerable to lightning strikes when they are connected to sensors buried in the soil, which serve as a direct path to ground, enabling current flow. Electrical problems such as this one usually manifest as faulty readings from all sensors connected to that datalogger. For example, all sensors connected to node 7 in our James Reserve deployment reported readings outside of their operational range for an hour starting on August 11, 2006 at 1:30AM. If the datalogger does not recover, it should be replaced.

Elusive Problems There were several problems that we could not track down. For example, all PAR sensors connected to every node reported a faulty reading of -888.88 from July 22–26, and again on August 10, 2006. We do not know what caused this behavior.

Not all unusable data collected from our deployments can be explained by the criteria above. We suspect that some of this unusable data can be explained by drift in the sensor calibration. In the next section, we describe how calibration issues impacted the sensor yield in our deployments, and describe two techniques being developed to address this problem.

5.3 Calibration Drift

In this subsection we discuss the third class of problems that impact sensor yield, instances where sensors slowly transition from producing usable data to producing readings that do not seem to reflect reality as a result of calibration drift.

Calibration parameters (i.e. the slope and intercept) for most sensors *drift*, or change over time. Many soil sensors, such as the ISEs, are especially prone to drift [12] and are thus calibrated before, after, and often during a deployment. In a particularly bad example, sensors connected to mote 11 in our Bangladesh deployment averaged a change in the calibration offset of 100 mV when comparing calibration equations obtained before and after the deployment. Given the average operational range for a sensor of 300 mV, an offset change of 100 mV is a significant change.

Drift impacts a sensor's perceived operational range, and thus the amount of data usable for scientific purposes. The top and bottom panels in Figure 11 are each plots of three CO_2 sensors buried at 2, 8, and 16 cm. The CO_2 sensor's operational range is defined as concentrations above 0 ppm (indicated by the horizontal line on the graph). The top panel is representative of readings taken from 9 of the 10 locations where CO_2 sensors were deployed, which remain above this threshold. The bottom panel contains data collected from one node in the deployment where the sensors at 2 and 8 cm were good, but the readings from the sensor at 16 cm slowly dip below this line between December 2005 and March of the following year. This steady trend is not characteristic of a faulty sensor, and likely indicates that the calibration for the sensor was gradually drifting. Recall from the previous section that the millivolt output from the CO_2 sensor is converted to concentration (ppm) using a calibration equation. If the parameters of this equation change over time, the translation from millivolts to concentration will be incorrect. In this instance, resulting in negative concentrations. Without further measurements taken during the time of drift, it is nearly impossible to identify the change in calibration parameters. Instead, the scientists discarded these 9,000 points collected during this three month period from the CO_2 sensor.

Accurately capturing drift is necessary to correctly identify usable data. But this is not a simple problem because there are competing interests influencing the decision of how best to capture drift. One approach to capture sensor drift is to re-calibrate a sensor occasionally, and model the drift between calibration events. In a one week lab experiment we calibrated a set of sensors daily and found that calibration does not change linearly with time, nor does it change in a constant direction. This experiment argues for calibrating sensors as *frequently* as possible in order to capture calibration parameters. However, calibration itself is labor-intensive. This problem is exacerbated for soil deployments where sensors are buried underground and inaccessible. Every time a sensor needs to be taken out of the ground, calibrated, and put back in, the soil needs time to settle back into a compacted state. This period can extend from a day to several months, depending on

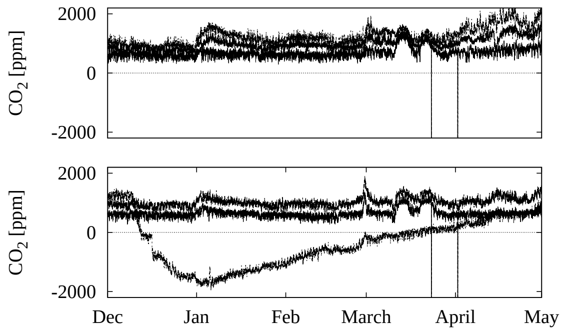

Figure 11: Drift of CO_2 Sensor in James Reserve Plots of three CO_2 sensors buried at 2, 8, and 16 cm. Horizontal line at 0 indicates the lower threshold for usable data from these sensors. Top panel is representative of the 8 other monitoring locations. Bottom panel contains data from one outlier sensor that gradually dips below this line beginning in December.

the soil structure and moisture content. This argues for performing calibration as *infrequently* as possible, especially for short-lived deployments (similar to that in the San Joaquin River) which can be severely impacted by several days of sensor down-time. We discuss two techniques to either calibrate a sensor or determine when a sensor must be calibrated, while the sensor is in-situ in order to minimize unnecessary soil impact.

The first technique is to use a known relationship between a sensor in the soil and a phenomenon that is easier to measure. The challenge in using this technique is to define a model between these two. For example, at James Reserve, because the amount of CO_2 in the air (which is easy to measure) is tightly correlated with the amount of CO_2 that is in the soil, these two measurements can be compared. Using a recently factory-calibrated sensor to measure CO_2 concentration in the air, scientists have found that soil sensors that differ by more than 10% from this reading are candidates for calibration. This threshold is based on over a year of field experience with the James Reserve transect.

In-Situ Calibration We could not apply this above technique to the ISEs in Bangladesh and the San Joaquin River valley because there is no known relationship between ions in muddy water and some more easily measured phenomenon. To address such cases we are experimenting with *in-situ calibration* in order to capture changing calibration parameters while the sensor is buried in the soil, and avoid premature calibration. A Teflon tube is attached to a sensor, with one opening of the tube positioned just above the sensor membrane and the other end exposed above ground. Periodically, the sensor is spiked through this tube with several milliliters of a standard solution. The solution concentration is chosen to be higher than that of the environment so that a pulse can be seen in the sensor data as the solution is delivered and then absorbed into the environment. Significant changes in the amplitude or slope of this resulting pulse across spikes could be used as an indication that the sensor is drifting and should be re-calibrated.

Preliminary results are encouraging. Figure 12 contains results from one experiment we performed on newly installed nitrate sensors at James Reserve. The solid arrow corresponds to a 5 mL nitrate injection, and the dashed arrow corresponds to a 5 mL water injection used to flush the nitrate solution. The sharp dip in voltage indicates the nitrate spikes (voltage is inversely related to concentration for nitrate ISEs), and the curve decays as the nitrate slowly absorbs into the soil. The graph shows that the slope and response time for spikes administered to the sensor over the course of a day are relatively consistent. This idea is still relatively new. Next steps to validate this approach include administering spikes

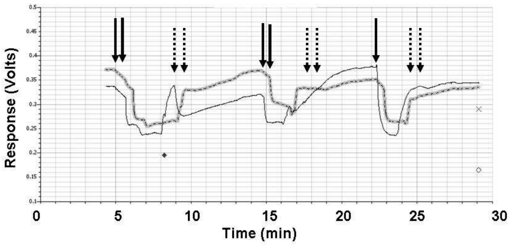

Figure 12: In-Situ Calibration of Nitrate Sensor at James Reserve Data from nitrate (solid line) and moisture (dashed line) sensors when spiked with 5 mL of nitrate solution (solid arrow) and 5 mL of water (dashed arrow). Nitrate concentration varies inversely with voltage, so nitrate data pulses down upon injection of nitrate solution.

regularly over the course of a week and then pulling the sensor out to re-calibrate when the response changes, and quantifying the impact that the small injections of solution do not significantly alter the environment.

Lessons Learned We learned two main lessons.

First, through the actions we took and their results we found that in-field actions enabled us to improve the amount of usable data we collected from our sensors. We were able to isolate problems by validating data that would have been otherwise discarded by extracting and analyzing physical soil samples as in the case of the nitrate and chloride sensors, measuring sensors in the field as in the case of the broken calcium sensor, and calibrating sensors in-situ. Even when we were not able to fix problems in the field, such as replacing a sensor we knew to be broken, in-field actions enabled us to isolate problems and take validating measurements that we could only have suspected at after the deployment. In Section 7 we quantify the impacts of immediate actions on sensor yield.

Second, through the actions we were *not* able to take, we discovered the importance of designing a robust platform. The platform should facilitate sensor replacement when sensors are faulty, and in-situ calibration. It should also be easy to deploy and re-deploy, especially important for short-term deployments. In Section 7 we describe the *javelin*, a platform designed to address these issues and successfully deployed in the San Joaquin River deployment.

6 Improving Network Yield

The second dimension to improving the information yield is to maximize the *network yield*. In this section, we discuss the techniques employed in each soil deployment in order to obtain high network yields ranging from 91%–100%. Our wireless network also achieved a network yield of 91%. This is higher than previous wireless deployments which averaged around 50% [4, 5].

Nodes in the transect at James Reserve achieved an average network yield of 91% from each of the 10 wired dataloggers. 77% of the 960,000 missing measurements occurred as a result of the battery running out (determined because measurements are missing from every sensor). The remaining 23% occur as a result of a disconnected sensor (dataloggers do not report measurements from disconnected sensors). Both of these problems are simple to detect and fix, but they were not immediately addressed because the nodes were not equipped with wireless connectivity. A tradeoff exists: Wireless communication enables realtime interaction but traditionally introduces lower network yield and increased complexity.

Nodes in the San Joaquin deployment achieved a 100% network yield. While significantly higher than the James Reserve yield, a 100% yield is not completely unexpected as the dataloggers were continually attended, data was downloaded every day over a wired channel, there were fewer nodes and sensors, and the deployment was much shorter than the James Reserve deployment.

We expected our wireless sensing system in Bangladesh to achieve the lowest network yield of all of the deployments because it relied on an unreliable wireless channel instead of a wired medium for data communication. However, this deployment achieved an average network yield of 91% per node. Most surprising is that the network achieved this high yield without using an end-to-end reliability layer. The reliability was attained in two ways. First, we used Sympathy [6], a system designed to systematize data management by monitoring data flow from each node, and identifying actions a user can take in the field when the network yield falls below a user specified threshold. Second, we incorporated a delay-tolerant networking [8] (DTN) layer on the mote and basestation [9]. This layer sits directly on top of the distance-vector routing layer. If a node does not have a valid route to the basestation the node stores all packets to the local EEPROM until a route becomes available.

We analyzed our data in order to understand the 9% packet loss. Since packet storage and transmission is persistent across node reboots, we believe the loss was not caused by nodes rebooting. Instead, we suspect a bug in the DTN layer. It took approximately 20 minutes for a route to time out on nodes once the basestation disappeared. Until then, nodes operate under the assumption that a route still exists, and continue to attempt packet transmission to the basestation. After exceeding MAC-layer retransmissions, packets not acknowledged by the basestation are dropped at the final hop instead of being stored in local storage. This bug has since been fixed.

Time stamps While the DTN layer was extremely effective at improving network yield, an unfortunate result of nodes storing packets locally was that packets were received out-of-order at the basestation. Since we could not rely on time of reception at the basestation to order packets, nodes needed to be able to accurately time stamp data packets. No node in our network, including our basestation, had access to the Internet or GPS. Network nodes obtained the time from the basestation which flooded the time every 20 minutes. The clock on the basestation needed to be manually set every morning when it was booted up. Problems with this centralized protocol occurred when a node re/booted and the basestation was not available to initialize the clock. Thus, to order the data we used mote time stamps by default, but when the time stamp was incorrect, we used a combination of sequence numbers to order the data with linear regression based on manually chosen time stamps.

Lessons Learned For wireless networks, a DTN layer that enables nodes to store packets when a valid route to the basestation is not available can be extremely effective at protecting against unexpected basestation outages in addition to unreliable wireless links.

In order to address the issues introduced by out-of-order data and networks subject to disruptions and loss of connectivity with a basestation, we need a new time-synchronization protocol. This protocol should be fast converging, to initialize nodes' clocks soon after start-up, and because delay-tolerant networking enables nodes to store data in the absence of a basestation, this protocol should be distributed so that nodes can obtain the correct time, even in the absence of a basestation.

7 Interactivity

In Sections 5 and 6 we discussed the sensor and network failures we have observed. Left unchecked, such problems can cause severe data loss. Identifying the presence of a problem and fixing it as quickly as possible can significantly improve information yield. The wireless connectivity in our Bangladesh deployment provided us the opportunity to interact with the hardware while the data was gathered. We begin this section by discussing and quantifying the impact of real-time interaction on the information yield from sensors deployed in Bangladesh, and present several representative examples from other deployments where interactivity could have greatly improved information yield. Based on these experiences, we discuss the *javelin*, a platform for soil monitoring that enables real-time interactivity.

In some instances, in-field interaction improved the sensor yield. For example, data that falls in a sensor's NLDR should be validated. By extracting physical soil samples in Bangladesh and analyzing chloride and nitrate levels, we were able to use 2850 of 3400 data points that occurred in the non-linear detection range of sensors. This action improved our sensor yield by 10% overall, and by 51% for those sensors. However, there were 550 data points recorded from sensors other than chloride and nitrate of the 3400 total points in the NLDR that we had to discard. Systematic interactivity could have enabled us to extract physical samples for these sensors as well.

In several instances, while in-field interactivity did not aid us in improving the sensor yield, we were at least able to definitively isolate the problem. Broken sensors accounted for 58% of the faulty data points collected in Bangladesh; while we could not replace them, similar to the calcium sensor described in Section 5, we were able to validate that they were broken during the deployment by checking the sensor with a pH meter during the deployment. Disconnected sensors accounted for 2% of faulty data in Bangladesh, and 58% of faulty data in our transect at James Reserve; in Bangladesh by checking the connector in the field we were able to validate the problem. High gradient data accounted for 60% of faulty data collected in Bangladesh; while we did not isolate all of the problems, we were able to fix one of the instances by moving a connector to a dry area, and isolate another instance by checking the sensor with a pH meter.

In several instances, all in the James Reserve transect, we did not interact with the sensors because the nodes did not have wireless connectivity. Thus, even faults that were easy to identify and fix were not addressed. Stuck-at faults at one temperature sensor during a 5 month period accounted for 7% of faulty data; and a failed battery resulted in 3 months of lost data from one node, resulting in a network yield of 69% from that datalogger.

7.1 Javelin

As we have learned through our deployments and discussed, a platform for soil monitoring must enable interactivity with the deployment in several ways. First, because many soil monitoring sensors require frequent calibration and are unreliable, testing and replacing individual sensors should be easy. In addition, the platform should support in-situ calibration. Second, because dense sampling is often next to impossible in heterogeneous soil environments, the platform should be quick to (re)deploy, and minimize the impact on the soil to keep soil settling times at a minimum. Third, the platform should facilitate the extraction of physical samples near sensors as soil sensors are often faulty and data from these sensors require validation.

The *javelin* pylon depicted in the top of Figure 1 and deployed at the San Joaquin River was designed to address these issues. Sensors are housed inside of a 1.25 inch PVC tube. Slits are cut around the circumference of the tube to allow moisture in, but keep out soils and other particles that may damage the sensor membranes. Communication hardware resides in a PVC enclosure attached to the top of the tube. A javelin is not designed to handle more than 5 sensors, addressing a mistake we made in the design of the pylon we used in Bangladesh which became extremely difficult to maneuver when holding its maximum of 24 sensors. The end of the tube ends in a point. In contrast to the up to 5 holes required to deploy the pylon in Bangladesh, the javelin can be driven into the ground in a single hole, minimizing environmental impact and making it easier to replace bad sensors.

The javelin is also designed to support Teflon tubes attached to each sensor for in-situ calibration. This tube can also be used to extract water near the tip of the sensor, or physical samples, useful in validating questionable data.

Not all soil systems can utilize the same platform. The javelin does not perform well in environments that are not moisture saturated as the sensors are shielded by the column and do not come into contact with sufficient moisture. However, in wet soils like in Bangladesh or the San Joaquin River, the javelin performs well. Figure 1 is a graph of diurnal nitrate trends detected by 3 nitrate sensors deployed at 1 foot intervals inside of a javelin.

Lessons Learned In-field interactivity is key to improving both network and sensor yield, and is enabled in two ways. First, nodes should be equipped with wireless communication to enable real-time data analysis. Real-time communication is required to notify users immediately when problems arise; wireless communication could have notified users in the James Reserve deployment immediately when a node went down, instead of having to wait for several months to discover it during regular maintenance. Second, it is impossible to manually monitor data from all sensors in a deployment. Like Sympathy does for network quality, we need a tool to systematize the monitoring and management of the data quality. The basestation should be equipped with software to systematically monitor data and notify users of actions they can take in the field to fix faults, validate questionable data, or address mis-calibrated sensors. The node should be equipped with software to enable immediate feedback, so that once an action is taken, a user can request subsequent samples to ensure that the problem has been fixed, instead of having to wait for the next sampling period. We are working to develop such a system based on the data we collected in Bangladesh.

8 Conclusions and Future Work

Soils are challenging environments for sensing systems due to their short duration, the measurement uncertainty, and the sensing uncertainty. We discuss the techniques we employed to improve information yield in three deployments undertaken in Bangladesh and in California. Through these deployments we have learned three lessons.

First, in-field interactivity significantly improves sensor and network yield. Actions such as in-situ calibration, validating potentially faulty data, and fixing broken hardware can improve the quantity and quality of usable data collected from a network. Such interactivity is enabled in two ways: 1) Nodes should have wireless connectivity to enable real-time communication; and 2) software should be installed on the motes and the basestation to enable systematic and timely monitoring of the data quality. We are working to design such a system to monitor data quality, and plan to deploy it with our soil monitoring system in various locations in Palmdale, and in Bangladesh in December, 2008.

Second, network yield for wireless networks is significantly improved using delay-tolerant networking techniques, but requires a robust time-synchronization protocol to handle the resulting out-of-order packet delivery.

Third, based on our experience, we have designed the *javelin*, a platform for very wet soils or shallow groundwater that enables sensor interactivity by minimizing soil disturbance during deployment, and facilitating the deployment, replacement and in-situ calibration of sensors.

Finally we wish to acknowledge our collaborators. Deployments such as these are not possible without collaboration with partnering institutions in the host country. Our deployment in Bangladesh was made possible through collaborations with engineering students in Bangladesh and scientists in the US who had been travelling yearly to Bangladesh over the past 5 years.

References

[1] L. Balzano and M. Srivastava. Fault in sensor networks. Technical Report 06, NESL, 2006.

[2] Amarjeet Singh et al. Autonomous robotic sensing experiments at san joaquin river. Technical Report 66, CENS, October 2006.

[3] D. Scholefield et al. Concerted diurnal patterns in riverine nutrient concentrations and physical conditions. *Science of the total environment*, may 2005.

[4] G. Tolle et al. A macroscope in the redwoods. In *Proc. SenSys*, 2005.

[5] G. Werner-Allen et al. Monitoring volcanic eruptions with a wireless sensor network. In *Procs. of EWSN*, 2005.

[6] N. Ramanathan et al. Sympathy for the sensor network debugger. In *Procs. of SenSys*, 2005.

[7] R. Szewczyk et al. Lessons from a sensor network expedition. In *Proc. EWSN*, January 2004.

[8] K. Fall. A delay-tolerant network architecture for challenged internets. In *Procs. of ACM SIGCOMM*, August 2003.

[9] R. Kapur, D. Estrin, and R. Guy. Reliability and storage in sensor networks. Master's thesis, UCLA, 2005.

[10] J.R. McNeill and V. Winiwarter. Breaking the sod: Humankind, history, and soil : Soils: The final frontier. *Science*, June 2004.

[11] N. Ramanathan and L. Balzano et al. Rapid deployment with confidence: Calibration and fault detection in environmental sensor networks. In *CENS Tech Report #62*, 2006.

[12] C. C. Rundle. A beginners guide to ion-selective electrode measurements, 2005.

[13] W. Yu. Socio-hydrologic approaches for managing groundwater contamination problems: strategies for the arsenic problem in bangladesh. doctoral thesis, division of engineering and applied sciences. 2003.