Development and evaluation of an open-source, low-cost distributed sensor network for environmental monitoring applications

To cite this article: N Gunawardena et al 2018 Meas. Sci. Technol. 29 024008

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2018 IOP Publishing Ltd.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.

As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.
Development and Evaluation of an Open-Source Low-Cost Distributed Sensor Network for Environmental Monitoring Applications

N Gunawardena, E R Pardyjak, R Stoll, and A Khadka
University of Utah, Department of Mechanical Engineering, 1495 East 100 South (1550 MEK), SLC UT, 84112
E-mail: pardyjak@utah.eng.edu

May 2017

Abstract. Over the last decade there has been a proliferation of low-cost sensor networks that enable highly distributed sensor deployments in environmental applications. The technology is easily accessible and rapidly advancing due to the use of open-source microcontrollers. While this trend is extremely exciting, and the technology provides unprecedented spatial coverage, these sensors and associated microcontroller systems have not been well evaluated in the literature. Given the large number of new deployments and proposed research efforts using these technologies, it is necessary to quantify the overall instrument and microcontroller performance for specific applications. In this paper, an Arduino-based weather station system is presented in detail. These Low-cost Energy-budget Measurement Stations, or LEMS, have now been deployed for continuous measurements as part of several different field campaigns, which are described herein. The LEMS are low-cost, flexible, and simple to maintain. In addition to presenting the technical details of the LEMS, its errors are quantified in laboratory and field settings. A simple artificial neural network based radiation-error correction scheme is also presented. Finally, challenges and possible improvements to microcontroller-based atmospheric sensing systems are discussed.


Keywords: Micrometeorology, Automated Weather Station, Arduino
1. Introduction

The atmosphere is comprised of many layers such as the stratosphere or mesosphere, but the layer that most directly affects daily life is the lowest layer, or the troposphere [1]. The troposphere itself can be broken into smaller layers including the atmospheric surface layer (ASL), or typically the lowest 100 m of atmosphere directly above the earth’s surface [2].

Since most humans live and work within the ASL, it is important for us to understand it thoroughly. This is not a trivial task. Since the ASL is the closest to the earth, it is not only affected by the surrounding atmosphere, but also by the energy and mass transfer associated with near-surface processes. This includes heat transfer to and from soils and buildings, evaporation, plant transpiration, emissions from human activity, and more [2]. While the ASL is relatively well understood for homogeneous terrain [3], a number of outstanding knowledge gaps related to complex terrain continue to confound scientists. Complex terrain may be defined as surfaces with substantial heterogeneity over a range of scales. Examples include cities and mountains, but also surfaces where heterogeneity may exist in surface land-use or land-cover (e.g., agricultural fields). These knowledge gaps limit the predictability of current operational weather forecasting models and atmospheric simulation codes. A fundamental understanding of the ASL in complex terrain will allow for better climate and weather prediction, which can lead to improved flood prediction, pollutant dispersion prediction, wind-energy yields, and numerous other benefits [4, 5, 6].

High-quality field-experiment data are critical for model development and evaluation. Atmospheric measurement techniques is a rich and mature field that continues to develop (see [7] for an excellent overview of standard techniques). There are many ways to make ASL measurements. Some measurements depend on airborne platforms such as tethered balloons, radiosondes, full-size aircraft, and more recently drones. Other measurements can be made using remote sensing instruments such as LIDAR (LIght Detection And Ranging), RASS (Radio Acoustic Sounding System), or SODAR (SOnic Detection And Ranging) equipment (for more details, see the review of Pardyjak and Stoll [8] in the present MS&T). However, for the surface layer, the foundation of atmospheric measurements are meteorological towers (met towers). Typically constructed of aluminium or steel tubing and standing anywhere from two to three hundred meters, these towers are fixed to the ground or platforms.
using a variety of methods and are usually guyed in place.

To make measurements, many different instruments are affixed to towers. While most towers contain temperature, pressure, humidity, and wind measurement devices, they are not limited to these. Ultimately, the tower’s only limitation is what can be physically affixed to it. As a result, they may contain less common instruments such as soil moisture sensors, gas analysers, or infrared cameras. Most met towers include at least one temperature and relative humidity instrument, and a wind measurement device [9]. The instruments are often placed at regular heights along the tower so atmospheric profiles (wind profiles, for example) can be gathered.

Since the instruments used are usually scientific grade, they are quite expensive. An uninstrumented 3-meter tower costs from $500 - $600 and a 10-meter tower can cost up to $4000.‡ A commonly used temperature and relative humidity sensor, the Campbell Scientific, Inc. HMP155, costs $800. A standard 3D sonic anemometer costs $5,000 to $7,000, while a sonic anemometer combined with a gas analyzer costs $19,000. Meanwhile, a Campbell Scientific CR6 datalogger used to store the data is $1900. After accounting for batteries, solar panels, cables, and other support hardware, a basic 3-meter met tower ends up costing at least $1,000, while a fully instrumented 10-meter flux tower can easily cost upwards of $50,000.

While capable, the paradigm of large, heavily-instrumented towers has drawbacks, primarily cost. Their high cost makes it difficult for researchers to purchase and deploy more than a few. As a result, the spatial coverage of a group of towers is usually limited. This is a problem as complex terrain has high spatial heterogeneity, requiring many measurements. There are some practical drawbacks as well. Due to their size, transporting and constructing the towers consumes significant hours of labor and support equipment, especially if the deployment location is in a remote area. Their size also precludes them from being deployed in urban or suburban areas, where space is at a premium and aesthetics are valued. Finally, small-number tower deployments have no redundancy, hence equipment failures have devastating consequences to an experiment.

One alternative to traditional tower-based methods is the use of large numbers of small, distributed sensor stations. These stations are usually self contained and small enough to be carried by a single person. The instruments, housing, datalogger,‡ All costs in this document come from manufacturer websites or direct quotes, but may have changed since the time of writing. US dollars are used for all prices throughout the document.
Evaluation of Low-Cost Distributed Sensor Systems

and power systems (usually batteries and solar) are all included as a part of the package, and they cost anywhere from tens to thousands of dollars [10]. While they can be networked, it is not necessary. One example of a common distributed sensor station is the Onset HOBO U23 Pro v2. [11]

While they are typically not as capable as more sophisticated and expensive systems, in terms of variety of measurements and sampling rates, distributed sensors have many advantages. Their size makes them much easier to deploy in a wider variety of environments. For example, a traditional tower is very difficult to deploy in an urban environment [12], but trivial for a small distributed-type station. Their lower costs mean researchers can buy dozens of stations for the cost of one tower system, greatly increasing spatial resolution. While most of these types of stations are not explicitly designed to measure wind profiles like taller tower-based systems do, their small size means they can simply be attached vertically on a structure to emulate a wind profile measurement.

Distributed sensor systems are used in many different fields and are a popular subject of ongoing research — usually as wireless sensor networks (WSNs) [13]. Ultimately, any large system that needs to be observed can benefit from wireless sensor networks. Mainwaring et al. [14] used wireless sensor networks for wildlife habitat monitoring, while Yu et al. [15] used them for forest fire detection. Kim et al. [16] used a WSN on the Golden Gate Bridge to monitor structural health. Structural health monitoring is a big field for WSNs and often uses state-of-the-art techniques and technology such as compressed sensing and sensing skins [17]. In addition, Abbasi et al. [18] have written an overview of the many applications of WSNs within agricultural applications while Bhattacharya et al. [19] and Liu et al. [20] have presented systems for indoor and outdoor air quality, respectively.

In fact, an older review [10] shows that there are several different applications, including military, health, and home related applications. It is also an active area of research within the computer science domain, where researchers try to develop efficient, robust data routing algorithms [21]. Finally, it is studied within the electrical engineering domain, where researchers are developing low-cost, durable, and low-power sensor nodes [22].

WSNs bring even more advantages over traditional tower-based measurements. Their connectivity allows for network calibration, reducing required human resources [23]. The data have multiple redundancies since they can be stored on-device, and transmitted to a central location [24]. Both WSNs and distributed sensor systems
are more immune to data loss when compared to traditional measurement stations. Losing one sensor out of twenty or one hundred is not as bad as losing one tower out of three. Even if one sensor is lost, their low-cost means replacement is trivial.

One of the largest technical problem facing distributed sensors and WSNs is power consumption [13, 25, 26]. This has been described in the literature but was also experienced during the development and testing of the systems described herein. Due to the stations smaller size and cost, the batteries used for power are smaller and cheaper, leading to reduced available power. While many stations utilise energy harvesting technology such as solar panels, not all do [13]. As a result, it is imperative that stations consume as little power as possible.

From a researcher’s point of view, another large problem with low-cost instrumentation is calibration. A cheaper instrument may not be tested to the same extent as a research-grade instrument, and the accuracy, precision, and reliability may not be well known. This leads to a lack of confidence in the sensor by the end user.

An additional limitation with wireless sensor networks and distributed sensor stations is measurement flexibility. Typical high-end meteorological dataloggers can handle many different kinds of equipment, whereas distributed sensors are often singular in their tasks and are not easily modified for different tasks. Some stations, such as the Berkeley Telos, are fully customised hardware that requires significant specialised knowledge to build and use [27]. However, open-source hardware can change that.

Arduino is a line of open-source microcontroller development boards created by the eponymous Italian company Arduino. Arguably one of the most successful examples of open-source hardware, each Arduino board typically contains a microcontroller (usually made by Atmel), voltage regulator, LEDs, pin-breakout headers, and a USB interface to upload code [28]. Along with the board, the Arduino company provides open-source software to interact with the board. Since this software is fully compatible with C/C++, the user does not have to learn any proprietary languages.

Previous work with Arduino-based environmental monitoring stations includes Jiao et al.’s [29] deployment of a 4-node air quality sensor network in Decatur, Georgia. While they were able to test several types of air quality sensors, there were drawbacks. The stations were large (0.4 m$^3$) and powered by mains electricity (AC Power), limiting their portability. In addition, the manufacturing of the stations
is very manual, which limits the scalability and therefore the proliferation of the stations. More work needs to be completed in building devices that retain high accuracy but are easily manufactured in a scalable manner. Ultimately, four nodes is not very distributed.

Young et al. [30] used a commercial Wi-Fi connected sensor station by Aginova called the Sentinel Microi. After building a custom radiation shield for it, they were able to produce accurate results at a low-cost. Nevertheless, there are drawbacks. Despite the cost of $150, they were only able to measure one parameter: temperature, and only in the range of 0°C to 70°C. Stations have been built at the University of Utah where temperature, humidity, and surface temperature can be measured for approximately $90, with a range of -40°C to 120°C [31]. In addition, Young et al.’s station is only able to take temperature data at \( \approx 0.02 \) Hz, which is far too slow for many environmental applications [8]. Young et al.’s paper demonstrates that it is possible to build low-cost, low-power sensor stations, but additional development is required to make them more capable.

Fisher et al. [32] described multiple open-source Arduino-based sensor systems. These systems were used to measure water use in agricultural applications and environmental parameters in a forested setting. However, Fisher et al.’s stations have some shortcomings. Fisher et al.’s stations do not have many sensors, and there is not much flexibility in the sensors that can be used. They also do not have solar panels, which necessitates regular battery changes. Further, the accuracy of the sensors were not quantified and explored in depth.

Similar to Fisher et al. [32], Vellidis et al. [33] created a low-cost WSN for agricultural purposes. However, this system was not open-source or expandable, limiting its flexibility, yet it is low-cost and communicates via active RFID, allowing for easy data retrieval and collection.

While not a comparison of WSNs specifically, Mittelbach et al. [34] have compared expensive soil volumetric water content (VWC) sensors to inexpensive VWC sensors. This study was very well conducted in that it was very thorough and in-depth, but was limited to VWC measurements. Repeating these types of experiments for other sensors packages is necessary to better understand their limitations. Ultimately, open-source hardware and software can offer everything that commercial distributed sensor systems can, but with increased flexibility and possibly lower cost.
2. Local energy-budget measurement stations

The Local Energy-budget Measurement Stations, or LEMS, are small, open-source weather stations built to measure atmospheric properties in the ASL and ultimately provide low-cost estimates of the surface energy balance [35]. They were designed to be low-cost, durable, and easy to deploy in various types of terrain. While there are custom components to them, they were designed to use as much off-the-shelf components as possible. The material for a fully instrumented LEMS costs approximately $1000 USD at the time of manufacturing. Component costs are presented in the following subsections and in Tables 1 and 2.

While the LEMS can and have been used for many different applications, they were originally developed to inexpensively obtain parameters needed to estimate the various components of the surface energy balance (SEB) [35]. To illustrate this, consider a simplified version of the SEB, which can be written as [36]:

\[ R = H + L + G. \]

In the SEB equation, \( R \) is the net radiation, \( H \) is the sensible heat flux, \( L \) is the latent heat flux, and \( G \) is the ground heat flux. The equation follows the standard sign convention where radiative fluxes going toward the surface are taken as positive, while the responsive fluxes (terms on the right hand side of the equation) are taken as positive if they are leaving the surface. The LEMS can be used to estimate each of these four terms. There are multiple ways this can be done, but one method can be described with equation (1) [37], which is the expanded SEB equation for a thin saturated horizontally homogeneous surface. The relationship between each term of equation (1) and the LEMS sensors is described below.

\[
\begin{align*}
R_{S\downarrow}(1 - \alpha) + R_{L\downarrow} + \sigma \epsilon T_s^4 &= \rho c_p g_H (T_s - T_a) - k_{soil} \frac{\partial T}{\partial z} + \lambda g_M e_s(T_s) - e_a \\
&= I + II + III + IV
\end{align*}
\]

I Net radiation

(a) \( R_{S\downarrow} \) is the incoming shortwave radiation, as measured by the LI200 pyranometer.

(b) \( \alpha, \epsilon, \) and \( \sigma \) are the surface albedo, emissivity, and the Stephan-Boltzmann constant, respectively, which are used to compute the outgoing surface longwave radiation.

(c) \( R_{L\downarrow} \) is the incoming longwave radiation, which is determined from a model that is dependent on the air temperature and relative humidity, which is
measured by the SHT15 air temperature and relative humidity sensor.
(d) $T_s$ is the surface temperature, as measured by the TN9 surface temperature sensor.

II Sensible heat flux
(a) $\rho$, $c_p$, and $g_H$ are the air density, specific heat, and the boundary layer conductance to heat. $g_H$ is generally a function of wind speed, which is measured with the Davis anemometer.
(b) $T_s$ is the surface temperature, as measured by the TN9.
(c) $T_a$ is the air temperature, as measured by the SHT15.

III Ground heat flux
(a) $k_{soil}$ is the thermal conductivity of the soil, which is dependent on soil moisture content. The soil moisture is measured by the 5TM soil moisture sensors.
(b) $\frac{\partial T}{\partial z}$ is the soil temperature gradient, as measured by the 5TM soil temperature sensors.

IV Latent heat flux
(a) $\lambda$ is the latent heat of vaporisation of water, while $g_M$ is the boundary layer conductance to moisture. $g_M$ is generally a function of wind speed, which is measured with the Davis anemometer.
(b) $e_s(T_s)$ is the saturated vapour pressure, at the surface temperature as measured by the TN9.
(c) $e_a$ is the air vapour pressure, which is calculated from the relative humidity and temperature as measured by the SHT15, and the barometric pressure as measured by the BMP085 pressure sensor.
(d) $P_{atm}$ is the barometric pressure, as measured by the BMP085.

There are some instances of LEMS being used to calculate terms of the SEB in the literature. For example, Bailey et al. [37] developed a method to estimate ground heat fluxes ($G$) using the two Decagon 5TM sensors (instead of the typical heat-flux plates), which was then used as part of a method for computing evapotranspiration rates in an urban vegetated region. In addition, Jensen et al. [38] implemented a version of the Penman-Monteith method [39] using LEMS data acquired on the slope of a desert mountain to estimate the SEB. Sensible heat fluxes ($H$) were calculated using the Decagon 5TM soil sensors, the Licor LI200 incoming solar radiation sensor,
Evaluation of Low-Cost Distributed Sensor Systems

the Sensirion SHT15 air temperature and relative humidity sensor, and the Davis anemometer. Jensen et al. [38] also showed that latent heat fluxes (L) can be estimated using the Bosch BMP085 pressure sensor, and the Sensirion SHT15 air temperature and relative humidity sensor [39, 38].

As illustrated in equation (1) and the examples above, the surface temperature sensor is crucial to making many of these flux measurements. While this sensor is low-cost, many other automated weather stations do not include this measurement, whereas the LEMS do. SEB validation and other application-oriented science involving the LEMS will be the subject of future publications.

The following subsections describe the various subsystems of the LEMS. A photograph and a descriptive schematic can be seen in Figure 1.

2.1. Data logger and controller

The brain of the LEMS is the Arduino Mega 2560. As described on the Arduino website [28], this Arduino is based upon Atmel’s ATmega2560, which is an 8-bit microcontroller that contains 256 kilobytes of flash memory for program storage, 4 kilobytes of EEPROM for low-volume data storage, and 8 kilobytes of static RAM. Running at 16 MHz, it also contains 54 digital IO pins, 16 analog inputs, four UARTs, six hardware interrupts, one SPI port, and one I^2^C port allowing for many different sensors and connections. The Arduino Mega 2560 contains two voltage regulators to power it and any external peripherals: one at 5 V and one at 3.3 V. However, for the LEMS, the Arduino’s 5-V regulator was bypassed with a separate 5-V regulator that was more compatible with the rest of the hardware.

2.2. Sensors

Tables 1 and 2 display the sensors used in the LEMS, their reported accuracy, and a typical corresponding research-grade version of a sensor designed to make the same measurement. Sensors were mainly selected by maximising accuracy while minimising cost, while still maintaining Arduino compatibility. At the time of design, there were not many low-cost soil moisture and temperature sensors available on the market, so the Decagon 5TM was used. The 5-V operating voltage and serial communication method of the 5TM also made it very easy to integrate with the Arduino. While there were inexpensive light sensors available (any photodiode with the correct frequency response would have worked), they had not been characterised
Figure 1: LEMS schematic (left) and image (right). In the schematic, wires are not shown and components are not to scale. A description and typical height of each instrument is as follows: (a) Two Decagon 5TM soil volumetric water content sensors, installed at a depth of 5 cm and 25 cm. (b) Zytemp TN9 surface temperature sensor, installed at a height of 50 cm. (c) Waterproof case containing a Bosch BMP085 pressure sensor, installed at a height of 1.75 m. (d) 5.2-Watt solar panel installed at a height of 2 m. (e) Sensirion SHT15 temperature and relative humidity sensor, installed at a height of 2 m. (f) Licor LI-200 pyranometer, installed at a height of 2.1 m. (f) Davis anemometer, installed at a height of 2.2 m. The image shows a LEMS deployed on the University of Utah campus grounds. The white box on the grass contains a larger battery for extended run time.

at the time, so the LiCor LI-200 was chosen for the task [40]. Even though these two sensors are low-cost compared to some other options, they are expensive compared to the rest of the components of the station. Their combined cost accounts for over half of the station cost. Using the LI-200 also required an ADC and conditioning circuit, further increasing costs (See subsection 2.3.3 for more information).

The Davis anemometer was also expensive compared to the other components - though it was the cheapest off-the-shelf wind sensor found at the time. The Davis anemometer is compared to the much more capable sonic anemometer because a sonic anemometer was the most precise instrument we had available at the time.
for comparison, and also because sonic anemometers are considered the “standard” instrument for high-quality velocity measurements in the atmospheric surface layer [41]. However, there are lower-cost, lower-accuracy alternatives, though they are still more expensive than the Davis anemometer. The R. M. Young 03002-L Wind Sentry Set is a cup-and-vane anemometer that costs $700. The R. M. Young 05103-L is a propeller type anemometer with a torpedo-shaped vane that costs $1,050. (For comparison, the Davis anemometer is $107.)
Table 1: Table of costs and accuracies for research-grade instruments. This table does not include prices for other components such as housings, loggers, and miscellaneous hardware. All sensors except for the Campbell Scientific CSAT3 have the worst case accuracy presented. The accuracy presented for the CSAT3 is the typical accuracy, and assumes a 5 ms\(^{-1}\) wind speed purely in the \(U_x\) direction. More information about the CSAT3 error can be found in its manual [42] and in [43].

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Name</th>
<th>Cost ($)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Vaisala HMP155[44]</td>
<td>800</td>
<td>± 0.397°C</td>
</tr>
<tr>
<td>Humidity</td>
<td>&quot;</td>
<td>&quot;</td>
<td>± 1.8%RH</td>
</tr>
<tr>
<td>Solar Radiation</td>
<td>Eppley SPP[45]</td>
<td>2700</td>
<td>±10.0 W/m(^2)</td>
</tr>
<tr>
<td>Wind Direction</td>
<td>Campbell CSAT3[42]</td>
<td>8544</td>
<td>± 0.70°</td>
</tr>
<tr>
<td>Wind Speed</td>
<td>&quot;</td>
<td>&quot;</td>
<td>± 0.18 ms(^{-1})</td>
</tr>
<tr>
<td>Barometric Pressure</td>
<td>Vaisala PTB110[46]</td>
<td>634</td>
<td>± 1.5 hPa</td>
</tr>
<tr>
<td>Surface Temperature</td>
<td>Apogee SI-111[47]</td>
<td>720</td>
<td>± 0.5°C</td>
</tr>
<tr>
<td>Soil Temperature</td>
<td>Type-K Thermocouple[48]</td>
<td>20</td>
<td>± 2.2°C</td>
</tr>
<tr>
<td>Soil Moisture</td>
<td>Campbell CS650[49]</td>
<td>238</td>
<td>± 3.0% VWC</td>
</tr>
<tr>
<td>Heat Flux</td>
<td>Hukseflux HFP01[50]</td>
<td>595</td>
<td>± 3.0%</td>
</tr>
</tbody>
</table>

Table 2: Table of costs and accuracies for low-cost instruments. The Decagon 5TM is the analogue of the heat flux plates presented in Table 1

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Name</th>
<th>Cost ($)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Sensirion SHT15[51]</td>
<td>42</td>
<td>± 1.6°C</td>
</tr>
<tr>
<td>Humidity</td>
<td>&quot;</td>
<td>&quot;</td>
<td>± 4.0%RH</td>
</tr>
<tr>
<td>Solar Radiation</td>
<td>LiCor LI-200[52]</td>
<td>270</td>
<td>± 3.0% of Eppley</td>
</tr>
<tr>
<td>Wind Direction</td>
<td>Davis Anemometer[53]</td>
<td>107</td>
<td>± 7.0°</td>
</tr>
<tr>
<td>Wind Speed</td>
<td>&quot;</td>
<td>&quot;</td>
<td>± 1.0 ms(^{-1})</td>
</tr>
<tr>
<td>Barometric Pressure</td>
<td>Bosch BMP085[54]</td>
<td>2</td>
<td>± 3.5 hPa</td>
</tr>
<tr>
<td>Surface Temperature</td>
<td>Zytemp TN9[55]</td>
<td>20</td>
<td>± 2.0°C</td>
</tr>
<tr>
<td>Soil Temperature</td>
<td>Decagon 5TM[56]</td>
<td>186</td>
<td>± 1.0°C</td>
</tr>
<tr>
<td>Soil Moisture</td>
<td>&quot;</td>
<td>&quot;</td>
<td>± 3.0% VWC</td>
</tr>
</tbody>
</table>
2.3. Arduino shield

An Arduino Shield is a printed circuit board (PCB) that is designed to plug into the Arduino headers to prevent the clutter of jumper wires and breadboards. For the LEMS, a custom shield was designed to accommodate the sensors and sensor hardware, as outlined below. Due to time, monetary, and manpower constraints at the time of manufacturing, the shield was designed to avoid using IC packages that were hard to solder by hand, such as ball grid array or quad flat no-leads packages. While the shield PCB was not manufactured by hand, the PCB assembly was all completed by hand. For twenty LEMS, this took a significant amount of time. With quantity pricing, each LEMS shield cost about $83 each. A block diagram of the LEMS shield can be seen in Figure 2. The LEMS hardware (and software) is fully open source, hence, electrical schematics and design files can be found on Github at https://github.com/madvoid/LEMS.

2.3.1. Data storage  All data are stored on a 2 GB Secure Digital (SD) card that is inserted into an SD card socket located on the shield. SD cards have several advantages over other methods of storage. They are cheap, durable, light, and ubiquitous. In addition, since they can communicate via SPI, they are supported by most microcontrollers, including the Arduino. On the LEMS shield, there are two key components that allow the Arduino to interface with the SD card. The first, of course, is an SD card socket. This is a standard spring-loaded SD card socket found in consumer electronics. The second is a level-shifting IC. Since the Arduino runs at 5 V and the SD card runs at 3.3 V, level shifting is required to prevent the Arduino from applying an overvoltage to the SD card. At the time of design, only cards of up to 2 GB were supported by the LEMS, due to FAT limitations of the SD card library used. In addition, software SPI was used instead of hardware SPI on the Arduino to aid with circuit layout and other design constraints.

2.3.2. Charging and power  To charge the 6-V sealed lead acid (SLA) batteries that are used with the LEMS, a solar panel (or other source of energy) and a charging circuit are required. For the solar panel, a 5.2-W 8-V open voltage solar cell is used, which cost ≈ $45. While there are battery charging ICs available, many of them are in surface mount packages, so a simple charging circuit was devised. The charging circuit consists of two diodes. The first diode is a Schottky diode placed
Evaluation of Low-Cost Distributed Sensor Systems

Figure 2: Schematic of the LEMS electronics. Most power connections are not explicitly shown, nor are small passive components such as capacitors, resistors, and LEDs. All other major components and connections are shown. The extra terminals allow for future expansion or are there as a result of retired sensors at the battery terminal that prevents current from flowing into the solar panel. The second diode is a Zener diode that shorts the solar panel terminals together when the battery voltage reaches 7.5 V. The charging circuit is in parallel with the power circuit, which simply consists of a 5-V linear low-dropout (LDO) voltage regulator. There is also a 3.3-V LDO regulator to provide the 3.3-V components power.
2.3.3. Analog sensors  To accurately make measurements with the LiCor LI-200 radiation sensor, several components were added to the LEMS shield. First, a single-stage transimpedance amplifier was built with an operational amplifier to convert the current output of the LI-200 to a usable voltage. This step is necessary regardless of the ADC used. The transimpedance amplifier performs several functions. It simultaneously converts the current output from the sensor into a voltage, and amplifies the signal so it can be accurately read by the ADC. It also filters noise from the signal. In this instance, it converts an approximately 90 $\mu$A maximum signal to 4 V, and low-pass filters anything below $\approx 22$ Hz. This filtering frequency was chosen because it involved easy to source components and produced satisfactory results for our application. Each LI-200 has a different maximum signal, so each shield has a specific high-precision resistor used to set the gain. As a result, each LEMS has a specific LI-200 paired with it for the lifetime of the LEMS. Then, a 4-channel 12-bit ADC was added to the shield, which allowed for higher resolution and accuracy than the Arduino’s built-in ADC could provide. This ADC communicates with the Arduino via SPI. While only the first channel was needed to measure the output of the transimpedance amplifier, the other 3 were left open for future expansion. Finally, a precision 4.5-V reference was added for the ADC. Together, these three components comprise the analogue section of the LEMS shield.

2.3.4. Connections and other electrical components  To connect all the sensors and external equipment to the Arduino shield, two types of connectors are used. Standard female 2.54 mm headers that allowed access to all of the Arduino pins if necessary and screw terminals that allow access to the commonly used microcontroller features such as power and communications pins. The waterproof connectors used on the case are wired to the screw terminals (See subsection 2.4 for more details). Between these two types of connectors, many different connections can be made with the LEMS.

There are a few other major components that are located on the LEMS shield. The first is the pressure sensor. Since the housing (further described in 2.4) had ventilation holes, the pressure sensor could be included on the shield and not in the external radiation shield. A real-time clock (RTC) was also included on the shield to provide a timestamp for the data. A power switch was soldered to the shield so the battery would not have to be connected and disconnected repeatedly. Finally, two LEDs, one green and one red, were located on the shield to show the status of the board. The red LED was used solely for problems, and the green LED would blink
Evaluation of Low-Cost Distributed Sensor Systems

every time a measurement was made. Though simple, this two LED system proved to be very effective at showing the status of the logger.

2.4. Other hardware and packaging

As illustrated in Figure 1, a significant amount of other hardware is necessary to actually deploy the LEMS. To house the Arduino and shield, a waterproof Seahorse SE120 case was used. Durable, waterproof, and easy to machine and transport, this case works very well for typical long-term micrometeorology deployments. The case has a strip of aluminium attached to it, to which large hose clamps are attached. The hose clamps are a inexpensive and widely available way to attach the LEMS to any post or pipe. The LEMS are attached to 2-meter tall studded T-posts (commonly used to erect low-cost fencing), which are also inexpensive, durable, and easy to deploy. The solar panel is attached to the case using a threaded rod and custom aluminium brackets. A radiation shield (purchased from Ambient Weather), that houses the temperature and humidity sensors, is attached to the case via a strip of aluminium. Attaching all of these items to the case results in a bulkier case, yet it was still easier to transport everything at once instead of worrying about several different components to assemble in the field. All sensor wires that are outside the case use plastic waterproof connectors. These connect to the matching panel waterproof connectors that are mounted on the case. All external sensors are mounted onto the T-post using plastic zip ties. Finally, a small housing was built out of PVC pipe joints for the TN9 infrared temperature sensor. This was necessary because the TN9 did not come with its own housing. The total cost of the support hardware is $\approx 75\), with $50 of it being waterproof connectors.

2.5. Expandability

One of the primary design requirements of the LEMS was easy expandability. Every single pin of the Arduino is passed through the shield so that sensors can be added in the future if necessary. Thanks to the flexibility of the Arduino programming environment and hardware, the code is easily modifiable, even in the field. Finally, the waterproof electronics housing case is easily machined. A battery-powered hand drill was often used to modify the case as necessary in the field. The ability to easily make in situ modifications has been very useful. The LEMS have been modified and customised for nearly every field experiment. Sensors and different batteries have
been added or removed. The packaging has been modified for better installation in certain locations. For example, for one experiment, the solar panel was extended and mounted away from the rest of the equipment so the LEMS could be mounted in shady areas. Finally, the code was often modified and updated to better suit the needs of the experiment on hand.

2.6. Code

The Arduino programming was done using the Arduino IDE. It is much simpler than traditional IDEs such as Eclipse, Visual Studio, or Xcode. The Arduino can be programmed using C/C++ and a set of specially defined, hardware specific functions that come included within the Arduino IDE. This allows the programmer to avoid writing directly to memory/registers, and to avoid writing in assembly. Thanks to the open-source community, it was not necessary to write code from scratch for most of the sensors used by the LEMS. Many open-source sensor libraries are available online, some even provided by the sensor suppliers. The Arduino IDE also comes with many built-in libraries which were used extensively. The built-in SD card library was modified to work properly with the LEMS. The code was designed to be modular alongside the LEMS. That is, if the LEMS are deployed using different sensor configurations, changing the code to accommodate this is trivial.

3. Field deployments

Due to the flexibility and ease of deployment of the LEMS, they can be used in many different types of experiments. The following paragraphs describe experiments where the LEMS have been deployed. Table 3 summarizes the key features of the different experiments.

3.1. Arid mountain experiment

In Fall 2012 and Summer 2013, the LEMS were deployed at the U.S. Army Dugway Proving Ground in and around the Great Salt Lake Desert in western Utah as part of The Mountain Terrain Atmospheric Modeling and Observations (MATERHORN) experiment [9]. The LEMS were deployed alongside five instrumented towers, 15 automated portable weather instrumentation data systems (called PWIDS), LIDARs, a tethered meteorological balloon, a thermal camera, and a ceilometer.
Table 3: Table summarizing the different LEMS field-experiment deployments. More detail and references can be found in subsections 3.1 - 3.4

<table>
<thead>
<tr>
<th>Location</th>
<th>Terrain</th>
<th>Number Deployed</th>
<th>Season</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dugway, UT</td>
<td>Desert/mountainous</td>
<td>19</td>
<td>Fall/spring</td>
</tr>
<tr>
<td>Monmouth, OR</td>
<td>Agricultural</td>
<td>3</td>
<td>Summer</td>
</tr>
<tr>
<td>Heber, UT</td>
<td>Mountainous</td>
<td>11</td>
<td>Winter</td>
</tr>
<tr>
<td>SLC, UT</td>
<td>Urban/suburban</td>
<td>11</td>
<td>Year-long</td>
</tr>
</tbody>
</table>

The LEMS measured all the variables listed in Table 2 except for wind. For this experiment, the main purpose of the LEMS was to increase the spatial resolution of the measurements. The data were used to further understand slope flows, particularly during the day-night transition time.

While a few of the LEMS were deployed in flat, desert conditions, most were deployed in dry, desert steppe, mountainous terrain. In the desert terrain, the LEMS were easy to deploy and performed well due to the abundant sunlight. The mountainous deployment was more difficult to execute due to a lack of roads. Nevertheless, the mountainous deployment was successful and the LEMS were much easier to deploy than traditional equipment. For example, the other type of automated weather stations used (PWIDS) required off-road vehicles to deploy, and didn’t include any surface or sub-surface measurements. In this type of terrain, the biggest obstacle to deployment was the hardness and rockiness of the dirt. This made mounting the T-posts and installing the soil measurement sensors very difficult. Furthermore, in high-salinity soils, such as those present during this experiment, additional precautions need to be taken when correlating the soil dielectric constant with volumetric water content [57, 56].

3.2. Vineyard experiment

During the summer of 2013, the LEMS were deployed in an Oregon vineyard as part of an experiment designed to better understand spore transport and canopy microclimate. The LEMS measured all the variables listed in Table 2 except for wind. While two of the LEMS were used to increase spatial resolution, the third LEMS was used to specifically monitor the microclimate conditions in the near-leaf environment.
During this experiment, the modularity and expandability characteristics of the LEMS were useful since additional sensors had to be added to the LEMS in the field. Moreover, the LEMS portability was helpful since mobility within the vineyard was limited by the plants and trellis structures. More information on the experiment can be found in Bailey et al. [37].

A separate experiment, conducted in the same vineyard during the summer of 2014, used the LEMS in a unique way. Multiple LEMS were affixed to meteorological towers at different heights so profiles could be obtained. These profiles were particularly useful because they contained data from above and below the canopy. Previously, the LEMS had only been deployed in a distributed sense, one at each location. More information on this experiment can be found in Miller et al. [58].

3.3. Mountain/Valley fog experiment

In Winter 2014, the LEMS were deployed in the small town of Heber, Utah for an additional MATERHORN experiment called MATERHORN Fog [59]. For this experiment, the LEMS measured all the variables listed in Table 2. Since there was only one meteorological tower for this experiment, the sole purpose of the LEMS was to increase spatial resolution and gather data from the surrounding area. The data were used to help understand the physics behind fog formation in complex terrain. This deployment was very straightforward and there were few requirements that were out of the ordinary. However, the experiments were conducted in the winter under cold (low temperature of $\approx -11^{\circ}C$) and snow covered conditions. Larger batteries were used in case of several consecutive days without sunlight. These larger batteries have continued to be used for convenience, and can be seen in Figure 1. Since some of the LEMS were deployed on public land, the possibility of public interference with the equipment existed. The only precaution taken was the attachment of tags that described the equipment and contact information. An overview of the complete experiment is provided in Gultepe et al. [59].

3.4. Urban field experiment

During the summer of 2013 and from the summer of 2015 until present, LEMS were deployed on the University of Utah campus in and around building and trees to investigate various aspects of urban micrometeorology. In both cases, the deployment was straightforward and not very difficult. Challenges included lack of sufficient solar
availability due to shading and lawn-care workers who accidentally cut cables on several occasions. In one application, the data from the LEMS were used to develop an energy balance model for complex plant canopies, while they were used in another application to provide data to validate a high-resolution urban wind model. More information can be found in the papers of Bailey et al. [37] and Eshagh et al. [60].

4. Sensor performance tests

Several experiments and analyses were conducted to quantify the accuracy and performance of the LEMS instrumentation in an idealised laboratory situation and in the field. Characterising the low-cost sensors increases confidence in them for future research applications. These experiments and their results are described below.

4.1. Laboratory comparison

To establish a baseline error, laboratory tests were conducted for the SHT15 temperature and relative humidity sensor. Independent of the radiation shield or other environmental factors, these tests quantify the measurement error due to the sensors themselves. While the manufacturer provides estimates of the expected accuracy, these experiments provide an independent validation. The LiCor LI-200 radiation sensor and the Decagon 5TM soil temperature/VWC sensors were not tested because they are considered research grade, and are well characterised. The TN9 infrared surface temperature sensor was not characterised due to the lack of a sufficient experimental test setup. The experimental methodology and results are presented below.

4.1.1. Relative humidity test

A controlled laboratory humidity test was set up using binary saturated aqueous solutions. When a binary salt solution mixed with water attains its equilibrium vapour pressure in a closed environment, a precise relative humidity is achieved [61]. For this experiment, four salts were used to span a large range of humidity values: LiCl, K₂CO₃, NaCl, and KCl, producing humidity values of 10.9%, 43.2%, 75.4%, and 84.8% respectively. For each experiment, the given salt solution, four SHT15 sensors (labelled 1-4), and one HMP155 sensor were placed within a insulated Styrofoam cooler, which was then placed in a basement room which maintains a relatively constant temperature throughout the year. The
Figure 3: Laboratory relative humidity test. The corresponding statistics can be seen in Table 4. The circles represent four different SHT15s and the stars represent the HMP155. Each data point is the average of 1.5 hours (n = 816) of raw equilibrated measurements for a given salt. Vertical bars represent the standard deviation during the test period. The red diagonal line is the 1-to-1 line, not a fit line for any of the sensors.

Table 4: Table of error statistics for the four SHT15s used for the laboratory relative humidity test. The corresponding data can be viewed in figure 3. Fit slope, fit intercept, and $R^2$ are linear regression parameters for each individual sensor; they are not parameters related to the red 1-to-1 line displayed in Figure 3. The maximum error is the largest difference between the sensor reading and the theoretical value. The RMSE is the root-mean-square error.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Fit Slope</th>
<th>Fit Intercept</th>
<th>$R^2$</th>
<th>Max. Error (%RH)</th>
<th>RMSE (%RH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHT15 1</td>
<td>0.895</td>
<td>6.321</td>
<td>0.983</td>
<td>6.3</td>
<td>4.6</td>
</tr>
<tr>
<td>SHT15 2</td>
<td>0.954</td>
<td>3.748</td>
<td>0.998</td>
<td>3.8</td>
<td>2.2</td>
</tr>
<tr>
<td>SHT15 3</td>
<td>0.927</td>
<td>6.396</td>
<td>0.998</td>
<td>5.5</td>
<td>3.5</td>
</tr>
<tr>
<td>SHT15 4</td>
<td>0.893</td>
<td>11.296</td>
<td>0.997</td>
<td>9.1</td>
<td>6.5</td>
</tr>
</tbody>
</table>
Figure 4: Laboratory temperature test. The corresponding statistics can be seen in Table 5. The measurements from four separate SHT15s are plotted against the measurements from a single finewire temperature sensor. Each data point is a 5-minute average of the raw measurements. No error bars or standard deviations are displayed. The red diagonal line is the 1-to-1 line, not a fit line for any of the sensors.

solution was given two days to equilibrate. Five-minute averages were computed from 1.5 hours of equilibrated 0.1 Hz measurements. These data were used to compute statistics comparing the low-cost sensors to precise humidities determined from the aqueous solutions. As seen in Figure 3, the HMP155 typically outperforms the SHT15s, though the advantage decreases at the higher humidity values. The results of this experiment help support the hypothesis that the large relative humidity errors discussed in section 4.2.1 are not the result of the sensor, but the radiation shield. The statistics for the data shown in Figure 3 can be seen in Table 4. This experiment was conducted during the fall season of 2016, and the HMP155 sensor was calibrated during the summer of 2015.

4.1.2. Air temperature test A controlled laboratory temperature test was also setup for the SHT15 sensors. For this experiment, four SHT15 sensors and one E-type finewire thermocouple probe were placed in a mechanically aspirated insulated box (to prevent radiation errors) and were subjected to a wide range of temperatures.
Table 5: Table of error statistics for the four SHT15s used for the laboratory air temperature test. The corresponding data is shown in figure 4. The fit slope, fit intercept, and $R^2$ value are the parameters of a regression line created for each individual sensor; they are not parameters related to the red 1-to-1 line displayed in Figure 4. The maximum error is the largest difference between the SHT15 measurement and the finewire measurement. The RMSE is the root-mean-square error. All data are 5-minute averages.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Fit Slope</th>
<th>Fit Intercept</th>
<th>$R^2$</th>
<th>Max. Error ($^\circ$C)</th>
<th>RMSE ($^\circ$C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHT15 1</td>
<td>0.994</td>
<td>0.403</td>
<td>1.000</td>
<td>0.70</td>
<td>0.34</td>
</tr>
<tr>
<td>SHT15 2</td>
<td>0.992</td>
<td>−0.115</td>
<td>1.000</td>
<td>0.90</td>
<td>0.21</td>
</tr>
<tr>
<td>SHT15 3</td>
<td>0.988</td>
<td>0.434</td>
<td>1.000</td>
<td>0.59</td>
<td>0.31</td>
</tr>
<tr>
<td>SHT15 4</td>
<td>0.994</td>
<td>0.188</td>
<td>1.000</td>
<td>0.57</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Temperature data from the finewire thermocouple were sampled at 1 Hz using a Campbell Scientific CR5000 Datalogger. Five-minute averages were then calculated for equilibrated temperatures and are shown in Figure 4. Despite the lack of data in the 15-23 $^\circ$C range, the SHT15 sensors performed very well. The statistics for the data shown in 4 can be seen in Table 5.

4.2. Field deployment evaluation

During the MATERHORN Fog experiment (see section 3.3), a LEMS was placed adjacent to a meteorological tower for instrument validation purposes. The LEMS contained an SHT15 within a radiation shield purchased from Ambient Weather to measure air temperature and relative humidity, whereas the meteorological tower had an HMP45 within a Campbell Scientific, Inc. 10-plate naturally aspirated radiation shield. In addition, the LEMS had a Davis cup and vane anemometer, while the tower had an R. M. Young 81000 sonic anemometer mounted at the same 2-m height. Data from these sensors are compared against the more expensive tower-based instruments. All data presented are 5-minute averages. All data from the LEMS were originally sampled at 0.1 Hz. On the meteorological tower, the wind speed and direction data were gathered at 20 Hz, while the air temperature and relative humidity data were gathered at 1 Hz.
Figure 5: Field experiment humidity results. All data are 5-minute averages. (a) Scatter plot from an HMP45 and SHT15. The solid diagonal line is the unity slope line. \( n = 8020 \). (b) The distribution of the absolute difference between the HMP45 and the SHT15. \( n = 8020 \). (c) Mean absolute difference between the HMP45 and SHT15 with respect to the mean sun elevation angle and mean sun azimuth angle. There are 288 points on the graph, each being the average for a given time of 31 days of data. (d) Absolute difference between the HMP45 and SHT15 with respect to the sun azimuth angle. \( n = 8020 \).

4.2.1 Relative humidity measurement comparison The relative humidity measurements from the Sensirion SHT15 on the LEMS were compared to the 2-m relative humidity measurements from the Campbell HMP45. As seen in figure 5(a), the SHT15 measurements correlate well (\( r = 0.96 \)) with the HMP45 measurements, though there is significant deviation from the 1-to-1 line (RMSE = 5.1195 %RH). As seen in the error histogram in figure 5(b), the majority of measurement errors are quite small - the mean absolute error is 3.7 %RH, within the error specifications of the SHT15 (\( \pm 4.0\%\text{RH} \)). However, the large measurement errors are very large (maximum absolute error = 40.8%RH), and appear to be dependent on the sun’s position (figures 5(c) and 5(d)). Thirty-three percent of all measurements are larger than the SHT15 error specifications, and eight percent of all measurements are larger than twice the SHT15 error specifications. This error is indicative of the
ineffectiveness of the radiation shield used for the LEMS. The LEMS radiation shield is open-bottomed, which has been previously shown to produce large errors [62].

In addition, there were two factors that likely made the radiation shield performance worse. The first was tilting of the instrument, which was observed during the field experiment. This can allow for more radiation to enter the shield, and may explain the asymmetry observed in figures 5(c) and 5(d). This tilt can be contributed to user deployment error and design shortcomings. While the LEMS are easy to deploy, the positioning of the radiation shield is less flexible than more expensive shields. Therefore, the angle of the radiation shield is dependent on the zenith angle of the T-post that the LEMS is mounted on. Due to installation difficulties originating from poor soil and inadequate equipment, the zenith angles of the T-posts are often non-zero. The second factor was snow. The instrument was deployed in the winter when there was a substantial amount of snow on the ground. Huwald et al. [63] showed that snow cover can significantly contribute to radiation errors, and this may have occurred during this deployment.

4.2.2. Air temperature measurement comparison  Since the HMP45 on the tower and the SHT15 on the LEMS also measure temperature, the experiment described in the previous section can also be used to validate the SHT15’s temperature measurements. As shown in figure 6(a), the SHT15 almost always measured a temperature that was equal to or greater than that that recorded by the HMP45. The SHT15 has a maximum temperature error specification of ±1.6°C, whereas the mean absolute error is 0.89°C. Out of all the measurements, 18.8% of them are larger than the SHT15 error specification, and 8.3% are larger than twice the error specification. Figures 6(c) and 6(d) show how the temperature absolute error correlates with the elevation and azimuth angles. Figure 6(d) also does not show the imbalance in error that 5(d) does, though it is unknown why. The same conclusions reached about the humidity error in relation to the open-bottom radiation shield, the tilting, and the snow are applicable to the temperature error as well.

4.2.3. Wind measurement comparison  The Davis cup-and-vane anemometer on the LEMS (at a height of 2.2 m) was compared to the R. M. Young 81000 sonic anemometer placed at a height of 2 m on the meteorological tower. As seen in Figure 7a, the Davis anemometer tends to over estimate the wind compared to the R. M. Young anemometer, consistent with the overspeed behaviour documented in
Figure 6: Field experiment temperature results. All data are 5-minute averages. (a) One to one plot of HMP45 and SHT15. The solid diagonal line is the unity slope line. \( n = 8020 \). (b) The distribution of the absolute difference between the HMP45 and the SHT15. \( n = 8020 \). (c) Mean absolute difference between the HMP45 and SHT15 with respect to the mean sun elevation angle and mean sun azimuth angle. There are 288 points on the graph, each being the average for a given time of 31 days of data. (d) Absolute difference between the HMP45 and SHT15 with respect to the sun azimuth angle. \( n = 8020 \).

The over estimate could also be a result of the small sensor height differences and wind shear. Using the log-law wind profile, and a surface roughness of 2-5 cm, we estimate wind speed differences, due to shear, of 2.1% to 2.6%. In addition, the Davis anemometer does not perform well at low speeds. To further demonstrate this, the comparison data set was split into two subsets: one containing wind speeds below 0.5 ms\(^{-1}\) (low wind speeds) as measured by the R. M. Young, and the other containing wind speeds above 0.5 ms\(^{-1}\) (high wind speeds). The value of 0.5 ms\(^{-1}\) was chosen because it is the documented start-up speed of the Davis anemometer [53]. Figure 8 shows histograms of wind speed measurements with magnitudes above 0.5 ms\(^{-1}\). It is evident that the Davis anemometer and the R. M. Young agree relatively well. Using the Kolmogorov-Smirnov (K-S) test statistic as an inter-histogram comparison metric, the restricted wind speed comparison (data
Figure 7: Field experiment wind speed sensor comparison. (a) Plot of the Davis cup and vane anemometer vs. the R. M. Young sonic anemometer along with a solid one-to-one line. The set of points forming a horizontal line on the abscissa from 0 ms\(^{-1}\) to 1 ms\(^{-1}\) shows the R. M. Young measuring non-zero wind speeds while the Davis records zero wind speeds. The Davis Anemometer literature indicates a lower limit to the measurement range of 0.5 ms\(^{-1}\) \[53\]. (b) Histogram of the absolute difference between the R. M. Young 8100 and the Davis Anemometer. \(n = 8021\). All data are 5-minute averages.

in Figure 8) has a K-S value of 0.052 and the directions have a K-S value of 0.0625. However, figure 9 shows the histograms of the wind measurements below 0.5 ms\(^{-1}\) do not match up very well. The low-speed speed comparison has a K-S value of 0.228 and the low-speed direction comparison has a K-S value of 0.112. Hence, there is relatively high confidence that two distributions come from the same sample for the higher winds speeds, but not for the lower winds speeds. While this was expected due to the Davis specifications, it limits the abilities of the LEMS to measure very low wind speeds, which is often desirable.

4.2.4. Pressure measurement comparison The Bosch BMP085 barometric pressure sensor located at 1.6 m on the LEMS was compared to the Vaisala PTB110 pressure sensor placed at a height of 7.8 m on the meteorological tower. For the analysis, the
Figure 8: Wind speed and direction distributions from the field experiment for wind speeds >0.5 ms\(^{-1}\). (a) and (b) show the wind speed distributions for the R. M. Young and Davis anemometers respectively. (b) and (c) show the wind direction distributions for the R. M. Young and Davis anemometers respectively. All data are 5-minute averages.

Pressure measurements from the PTB110 were adjusted down to the height of the BMP085 using the barometric formula [67]. The BMP085 has a positive offset error of 188 Pa when compared to the PTB110, but otherwise performs very well. It was also found that some error present is due to the increased temperatures experienced by the BMP085. Since the BMP085 is located in a weatherproof box, the temperature measured by the BPM085 was higher than the ambient temperature during the day when solar radiation exposure heats the box. This results in a pressure measurement bias of \(\approx 4 \text{ Pa/}^\circ\text{C}\).

5. Correcting radiation errors with artificial neural networks

Since low-cost hardware can be less accurate than expensive hardware, it is worth investigating error correction methods. Here we show that an artificial neural network (ANN) can be used to correct large air temperature radiation errors using other measurements taken from the station and a training period with an independent
Evaluation of Low-Cost Distributed Sensor Systems

Figure 9: Field wind speed comparison of wind speeds lower than 0.5 ms⁻¹. (A) and (B) show the wind speed histograms for the R. M. Young and Davis anemometers respectively. (B) and (C) show the wind direction histograms for the R. M. Young and Davis anemometers respectively. All data are 5-minute averages.

sensor. Previously, Yang et al. [68] showed that small radiation errors of less than 2°C can be corrected using a combined computational fluid dynamics and genetic algorithm method. In addition, Nakamura et al. [69] showed that small radiation errors can be corrected using an empirical model.

5.1. ANN details

A standard feed forward ANN was implemented using MATLAB’s neural network toolbox; all implementation details described in this paragraph are expanded upon in the Matlab Neural Network Toolbox User’s Guide [70]. The ANN has one hidden layer with fifteen hidden nodes. A sigmoidal activation function was used for the hidden layer, while an identity activation function was used for the output layer. The performance function was the mean squared error function and the backpropagation optimization method used was Levenberg-Marquardt. A standard LEMS was set up on the University of Utah campus adjacent to a finewire thermocouple, and the data were used to train and test the ANN. The following features were used as the inputs to the ANN, all gathered from the LEMS: time of day, SHT15 air temperature, incoming
Evaluation of Low-Cost Distributed Sensor Systems

Figure 10: ANN corrected air temperature results. All data are 5-min averages and the solid diagonal lines are the 1-to-1 lines (n = 1319 for both plots). (a) Original SHT15 temperature data vs. finewire thermocouple data. $R^2 = 0.916$, RMSE = 0.85 °C. (b) Corrected SHT15 temperature data vs. the finewire thermocouple data. $R^2 = 0.977$, RMSE = 0.45 °C.

solar radiation, sun elevation angle, sun azimuth angle, and wind speed. The ANN target was the air temperature as measured by the finewire thermocouple mounted adjacent to the LEMS. One month of 5-minute averaged measurements, from 24 February 2017 to 24 March 2017 was passed into Matlab for training. Matlab was set to internally split the data into 90% training data, 5% validation data (used as a stopping condition), and 5% test data (which was unused). Then, a completely separate five days of measurements, from 24 March 2017 to 29 March 2017 was used to evaluate the ANN performance. Training typically took less than 5 seconds on a 2013 Macbook Pro, and training was stopped based on the network’s performance on the validation set.

5.2. Correction results

The ANN performed the radiation correction on the test data well. As seen in figure 10(a), the uncorrected SHT15 measured higher temperatures than the finewire thermocouple due to solar radiation influences. The data presented in figure 10(a)
has an $R^2$ value of 0.934, a maximum absolute error of 4.20 °C, and a root-mean-square error of 0.849 °C. The data presented in figure 10(b) has an $R^2$ value of 0.978, a maximum absolute error of 2.09 °C, and a root-mean-square error of 0.448 °C.

When using neural networks to perform radiation correction, transferability is an important question. In other words, can an ANN trained on one instrument or location perform corrections on a different instrument or location? The experiment described above was repeated with some changes. An ANN was trained on a LEMS and an HMP45 located in Heber, Utah, and was used to correct the temperature data gathered with the same LEMS on the University of Utah Campus. This was a test of location transferability to understand if an ANN trained on a given instrument in one location could correct errors on the same instrument in a different location? The ANN correction did not perform well and transferability remains an open question. A thorough study of the computer science sub-field of transfer learning would be beneficial for this application [71] but is beyond the scope of the present paper.

6. Challenges and future improvements

Building and deploying the LEMS revealed much about embedded system design. The following list summarizes many of these lessons learned. Some challenges that all environmental equipment face were not included (e.g. adverse weather, human, and animal interference, general bad luck, etc.).

- Battery life is the most important factor to consider when developing remote stations. Many problems and most data loss could be attributed to low battery life. The original goals for the LEMS was accurate, low-cost, and durable. The new goals for the LEMS are accurate, low-cost, durable, and energy efficient.
- Sourcing parts can be difficult. Environmental sensing is not a large field, so there is not a surplus of choices for instrumentation and equipment, especially when other requirements such as price are factored in. To minimise costs, one often needs to depend on custom hardware, which brings up its own set of problems.
- Scalability is difficult. It is simple to make one custom part for one station. But when several custom parts need to be made, manufacturing time rapidly adds up. It is best to depend on commodity parts, if possible.
• Choose sensors with simple communication methods. The infrared temperature sensor used for the LEMS had a non-standard communication protocol that made development more difficult. Even if a sensor with a non-standard protocol is found that has a library, the library could interfere with the normal operation of the rest of the device. Conversely, purchasing a sensor that communicates via I2C all but guarantees it will work with the system.

• The world is moving to surface mount ICs. Despite the difficulty with soldering surface mount ICs, there are many more components available if they are considered. One solution to the soldering problem is to build a surface mount soldering station. In fact, outsourcing is an important option, a task that is necessary anyways if the stations are to be scaled to quantities.

• Analog sensing is difficult. It is much easier to purchase digital sensors and depend on the manufacturer’s tolerance than design an analog system of your own. Analog sensors added cost, complexity, and development time to the LEMS.

• Sensor measurements are not always straightforward. An example of this is the Decagon 5TM. This soil temperature and moisture sensor does not actually measure volumetric water content. It measures the dielectric permittivity of the soil, which is then converted to a volumetric water content value using the Topp equation [72]. While this is fine for many soils, it may not be fine for soils with a high salinity, such as some of the soils found at Dugway Proving Grounds, Utah.

These lessons and resulting recommendations were used extensively in the design and fabrication of LEMS Version 2. LEMSv2 was built in the Fall/Winter of 2016, and was deployed January 2017 in the Cadarache Valley in southeastern France [73]. LEMSv2 has a completely redesigned electronics system that lead to an order of magnitude less power consumption, better timing, and many other improvements over LEMSv1. There are still some things that can be improved for Version 3. First, wireless networking needs to be added in some form to make them easier to monitor. In addition, development of several lower-cost sensors is needed. For example, less expensive soil temperature/VWC and sunlight sensors since they are currently the most expensive components on the LEMS. Finally, the manufacturing process needs to become more scalable. While LEMSv2 is much easier to build than LEMSv1, it still is not fully outsourced, leading to substantial in-house labour requirements.
Evaluation of Low-Cost Distributed Sensor Systems

7. Summary

A group of open-source Arduino-based weather stations were built and deployed in several different field experiments. These Local Energy-budget Measurement Stations, or LEMS, measured soil temperature and moisture, surface temperature, air temperature, relative humidity, barometric pressure, incoming solar radiation, wind speed, and wind direction. The LEMS cost an order of magnitude less than a similarly instrumented tower, used mostly commodity parts, and is fully open source. Laboratory tests show that the temperature and relative humidity measurements made by the LEMS sensors were nearly as accurate as the research grade sensors. However, field tests show that radiation errors were common in the temperature and relative humidity measurements. In addition, the anemometer used at the time did not perform well at low wind speeds. In regards to the radiation error experienced, it was shown that an artificial neural network (ANN) could be used to correct the radiation errors. The largest practical problem with the LEMS power consumption. Future stations will need to reduce power consumption, and for operator ease, improve connectivity/data gathering.

Acknowledgments

This research was funded by the Office of Naval Research Award N00014-11-1-0709, Mountain Terrain Atmospheric Modeling and Observations (MATERHORN) Program. It is also funded by U.S. National Science Foundation grants IDR CBET-1512740, CBET-PDM 1134580 and 1133590, and AGS 1255662. Special thanks go to Christian Holbert, Justin Johnson, Nate Miller, Brian Bailey, Chaoxun Hang, Alexei Perelet, Vignesh Sivaramakrishnan, Heather King, Sawyer Hill, and Hanieh Eshagh at the University of Utah for help with the LEMS. Thanks to Thierry Hedde at CEA, and Pierre Durand at Laboratoire d’Aerologie - University of Paul Sabatier III for help with LEMSv2.

References

Evaluation of Low-Cost Distributed Sensor Systems


Evaluation of Low-Cost Distributed Sensor Systems


26 quality monitoring system based on wireless sensor networks. In Sensing technology (icst),


[22] R Min, M Bhardwaj, S Cho, E Shih, A Sinha, A Wang, and A Chandrakasan. Low-power


32 in wireless sensor networks: a survey. WSEAS Transactions on Computers, 8(4):705–714,
33 2009.


[29] W Jiao, G Hagler, R Williams, R Sharpe, R Brown, D Garver, R Judge, M Caudill, J Rickard,
40 Network (CAIRSENSE) project: Evaluation of low-cost sensor performance in a suburban
41 environment in the southeastern United States. Atmospheric Measurement Techniques


45 Microclimate and Energy Use with a Low Cost Wall Mounted Measurement System. In


[34] H Mittelbach, F Casini, I Lehner, A J Teuling, and S I Seneviratne. Soil moisture
51 monitoring for climate research: Evaluation of a low-cost sensor in the framework of the
52 Swiss soil moisture experiment (SwissSMEX) campaign. Journal of Geophysical Research

Evaluation of Low-Cost Distributed Sensor Systems


[52] Li-Cor, Lincoln, Nebraska. LI-200R Pyranometer: Instruction manual, 2017.


Evaluation of Low-Cost Distributed Sensor Systems


