

A Sensor Array Network for Continuous Soil Microclimatic Monitoring & Snow Depth Modeling in Forest Environments

Technical Documentation & User Guide



Colin M Beier
Associate Professor
Department of Sustainable Resources Management
SUNY College of Environmental Science & Forestry
1 Forestry Drive, Syracuse NY 13210

James S Mills
Research Support Specialist
Adirondack Ecological Center
SUNY College of Environmental Science & Forestry
6312 NYS Route 28N, Newcomb NY 12852

With contributions from

Michael Mahoney

Patrick McHale

Will Saunders

Technical Documentation: A Sensor Array Network for Continuous Soil Microclimatic Monitoring & Snow Depth Modeling in Forest Environments

1	Summary	3
2	Objectives	4
3	Design	5
3.1	Vertically distributed thermal profiling	5
3.2	iButton thermochron sensors	5
3.3	Assembly design prototypes.....	6
3.4	Design testing: open versus forested sites.....	7
3.5	Final array design.....	10
4	Array Installation & Retrieval	12
4.1	Installation	12
4.2	Trail cameras	13
4.3	Data retrieval	14
5	Data Processing & Analysis	15
5.1	Pre-processing and quality control.....	15
5.2	Soil and forest floor microclimate	15
5.3	Snow depth modeling and monitoring.....	16
5.4	Data analysis & visualization using Program R	18

1 Summary

This manual describes the design and deployment of temperature profiling arrays for applications in terrestrial ecosystem monitoring, with a focus on continuous measurement and modeling of snow depth and ground-level microclimates. Steps outlined here include array design, fabrication, prototype testing, field installation, data retrieval and analysis/modeling. All work was completed as part of long-term monitoring programs at ESF's Huntington Forest from 2018-2022 with funding support from NYSEDA.

A profiling array is a series of measurement devices or sensors arranged along a fixed axis to quantify continuous variation of the measured variable(s) along that axis. For the profiling arrays described here, the devices were autonomous data-logging temperature sensors (iButton thermochrons) arranged with regular spacing along a vertical axis from the subsoil to one meter aboveground. Sensors would then be programmed to record temperatures all at the same time at a consistent time interval.

Vertical profiling arrays for snow measurement and monitoring operate on the principle that snow is a good to excellent insulator from air temperatures, which tend to fluctuate between daytime highs and nighttime lows. Therefore, a continuously recording thermometer located below the snow line in winter will measure overall warmer and more stable temperatures over time. As snowpack grows and shrinks during the winter, the corresponding shift in temperature regime from snow to air is measured along the vertical profile, allowing the estimation of snow height relative to the sensor position in the array.

We designed two prototypes for the array assembly and tested them simultaneously in open field and forested settings. Both designs enclosed the vertical sensor array in a PVC housing capped and weather-sealed at both ends and then installed into the ground. Sensors were arranged along a vertical axis from the subsoil in to one meter aboveground, with 20cm spacing. The primary difference between designs was how the sensors were assembled into the array housing. After field testing and comparing data collected by the prototype arrays, we selected the 'rod' design for its much greater efficiency and operability in the field. Over 50 of these arrays were ultimately fabricated for field deployments.

Profiling arrays were deployed across Archer Creek watershed for applications including measurement and monitoring of snow dynamics and forest floor microclimate. For snow monitoring, we installed automated trail cameras at a subset of array locations to produce a reference dataset of daily snow depth for modeling purposes. Machine learning models that were trained and tested using these data yielded promising results for snow depth estimation. Soil microclimatic monitoring, in which the temperature profiling arrays were coupled with soil moisture probes, was done as part of a soil gas flux study during the warm season. Sampling designs and results from these applications are described.

Overall, we demonstrated the viability of temperature profiling arrays for ecosystem monitoring in a remote forest setting. Although fabrication and field installation of the arrays required a substantial up-front investment of time and materials, their longer-term operation requires relatively little time or expense, except for sensor replacement, which will be needed every 3-5 years. We are exploring thermochrons with replaceable batteries in lieu of iButtons, which must be replaced their built-in power supply fails. Collaborators are pilot-testing alternative array designs, including higher-resolution arrays (more sensors) where sensors exposed to the elements instead of being housed in a sealed airspace.

2 Objectives

The goal of this project was to develop and implement a methodology to efficiently estimate snow cover and describe ground-surface microclimates over time in a remote forest setting, as part of an ongoing long-term ecosystem monitoring program. We drew upon research in vertically resolved temperature profiling and its applications for efficient monitoring of aquatic and terrestrial environments. We used autonomous data-logging temperature sensors, or thermochrons, to continuously (every four hours) record temperatures at 20 cm intervals along a vertical axis from 20 cm belowground to one meter aboveground. Arrays were deployed for different purposes, from comparing performance of design prototypes to extensive sampling for watershed-scale snow and microclimatic monitoring. This technical report describes the design, prototype testing, deployment and initial results of these efforts, based on work carried out at Huntington Wildlife Forest, a SUNY ESF field station in Newcomb, NY.

Supporting objectives documented in this report include:

- Designing vertical profiling arrays using programmable iButton thermochron sensors
- Building and pilot-testing array prototypes in open field and forest floor settings
- Selection of the final array design with specifications needed for fabrication
- Deployment of arrays including field installation, sensor retrieval, and data management
- Data processing and analysis for snow depth estimation and microclimate measurement



3 Design

3.1 Vertically distributed thermal profiling

In thermal profiling, a group or 'array' of thermometers is deployed at equal intervals, typically along a vertical axis, in order to record measurements at different heights or distances from a given surface or feature. By measuring temperature at regularly spaced points along the axis, the unmeasured space between thermometers can be statistically interpolated to generate a continuous estimate of how temperature varies along the entire profile. Due to the insulative properties of water, temperatures are generally more stable over time in deep waters, while surface temperature fluctuate more readily with diurnal (daily) and seasonal (monthly) variation. Vertical profiling is often used in lake monitoring to measure thermal stratification of the water column and other properties important for seasonal turnover and aquatic ecosystem processes.

The basic principles of thermal profiling are essentially the same, regardless of the specific setting. When deployed on land for the purposes of estimating snow depth, thermometers are arrayed vertically, either starting at or below the ground surface and ending at some distance above the ground. As with deep water, thermometers buried below ground are insulated by the soil and will experience much less variation in temperatures during a typical 24 hr period and over longer time horizons. Like liquid water, snow is an excellent insulator. Sensors buried in snow that is deep enough to insulate them from exposure to the fluctuating temperatures of the air will exhibit a similar thermal stability to those buried below ground. Sensor above the snow line will continue to fluctuate with daily high and nightly low temperatures, as well as overall variability in weather patterns. By comparing temperature regimes at regular points (heights) along the vertical profile, the level of snow at a given point in time can be estimated. This approach also provides continuous measurements of forest floor microclimates, including subsoil habitats, that act as controls on terrestrial ecosystem functions such as respiration.

3.2 iButton thermochron sensors

Autonomous data-logging sensors are the ideal devices for thermal profiling and related applications for microclimatic monitoring. Among many options, the iButton thermochron device has gained widespread adoption in recent years for a wide range of applications. An [iButton](#) is a computer chip with a globally unique identification number and integrated power supply enclosed in a 16mm weather-resistant stainless-steel case. iButtons can include read/write memory, real-time clocks and temperature and/or humidity data logging functions, among others. Because iButtons are small, portable, rugged, water resistant, self-powered, programmable and relatively inexpensive, they are ideal for variety of environmental monitoring applications including thermal profiling.

We used two iButton thermochron models in our temperature profiling arrays: the DS1921G-F5# and the DS1922L-F5#, manufactured by Analog Devices (formerly Maxim Integrated LLC). Both models record temperatures from -40 C to +85 C at user-defined intervals but differ in terms of measurement precision: DS1921G has 0.5 C resolution and DS1922L has 0.0125 C resolution. The higher resolution DS1922L model has drawbacks: it is roughly twice the unit cost of the lower resolution DS1921G, while it can store less than half of the observations (data) before it begins overwriting existing data. In pilot testing, we assigned the high-resolution model to the lowest array position to measure subsurface temperatures with greater precision, because we expected the diurnal (24 hr) variability belowground to be much less than aboveground and that slight changes in subsoil temperatures would only be detected

with high-resolution data. We also used lower-precision iButtons at the lowest position in several arrays for comparison purposes. Overall from this testing, we determined that the value of the high-precision measurements was limited for our applications, and the drawbacks of the DS1922L model outweighed the benefits. However, we note that several of these high-precision iButtons remained in use, but were programmed to record temperature at the same 0.5 C resolution as the rest of the DS1921G sensors for data consistency purposes.

3.3 Assembly design prototypes

We created two prototype array designs for vertical profiling to estimate snow depth by continuously recording soil and air temperatures (Figure 1). Both designs shared the same sampling approach: six iButton thermochron sensors at height intervals as follows: 0.2 m below the soil/air interface (-0.2m), at the soil/air interface (0.0 m), and then 0.2 m, 0.4 m, 0.6 m and 1.0 m above the soil surface. Both designs had sensors housed vertically in sections of 1.5" diameter PVC pipe. The two array designs differed in how sensors were set and fixed at their designated positions (heights relative to ground) in the PVC housing (Figure 3.3.1).

The first design used 'plungers' to situate iButtons in the desired position within the housing. Plungers were wooden dowls cut to precise lengths corresponding to sensor heights. In this design, iButtons were placed between two 1.5 in diameter, 0.5 in thick closed-cell polyethylene foam discs fabricated to fit snugly in the PVC housing. Plungers were used to push the foam discs and sensors into position, starting from the bottom position (-0.2 m) and working in sequence up the housing to the top position (+ 1.0 m). To collect the sensors for data retrieval and reprogramming, this design required

The second design used a PVC rod attached to a PVC cap so that when assembled, the rod vertically bisected the housing without touching the interior sides. The iButtons were inserted in standard plastic fobs, which were secured to the PVC rod at each position (height) with a small screw. The orientation of the fobs and sensors on the PVC rod were aligned along the same radial axis. To collect the sensor data, the PVC cap could then be unscrewed and the rod assembly removed from the housing, with the sensors in their fobs still attached to the rod.

Array housings were fabricated with 1.5 in PVC caps and silicone caulk on both ends to ensure that interior spaces were sealed from the ambient exterior. The bottom end of the PVC housing was capped and sealed with silicon caulk and a #8 rubber stopper. The upper end of the housing was closed with a threaded PVC cap and sealed with silicon caulk. For the rod design, silicone caulk was applied to both the interior and exterior of the cap to seal the location where the screw passed through the cap.

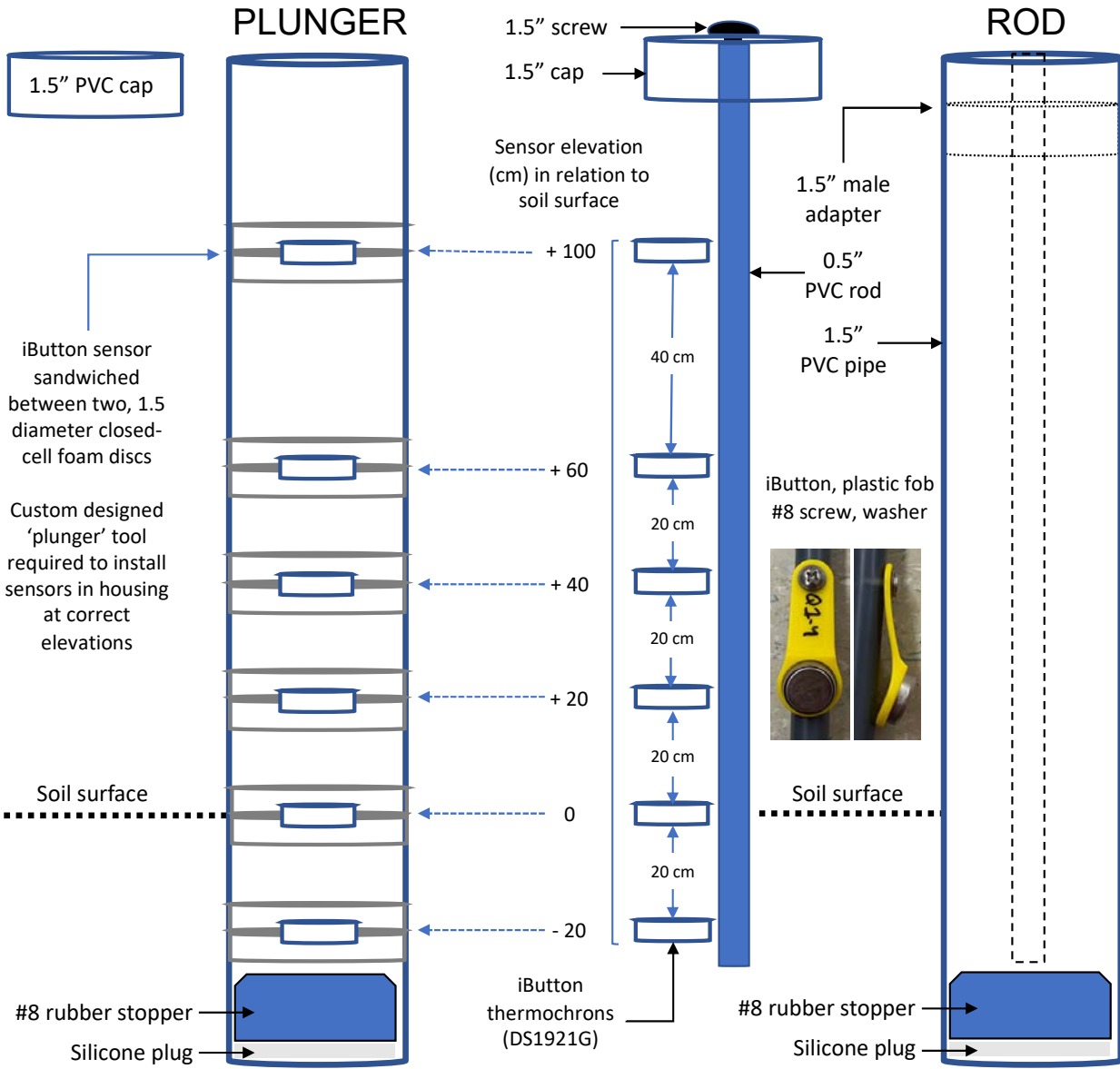


Figure 3.3.1. Schematics of the plunger and rod designs for temperature profiling arrays and housings.

3.4 Design testing: open versus forested sites

Both designs were deployed for pilot testing during the 2017-18 winter. For testing, one prototype of each design was deployed in a closed-canopy forest setting, approximately 3 meters apart. A replicate pair was deployed in an adjacent open meadow that was maintained for weather observation and atmospheric deposition sampling (NADP-NTN site NY20). The two pairs of arrays were located 15 m apart on a gentle south-facing slope at approximately 1610 ft (491m) elevation. Arrays were installed and began recording temperatures at four-hour intervals from 15 Aug 2017 and until 05 May 2018. Following retrieval of the arrays, we compared the data collected between locations and array designs.

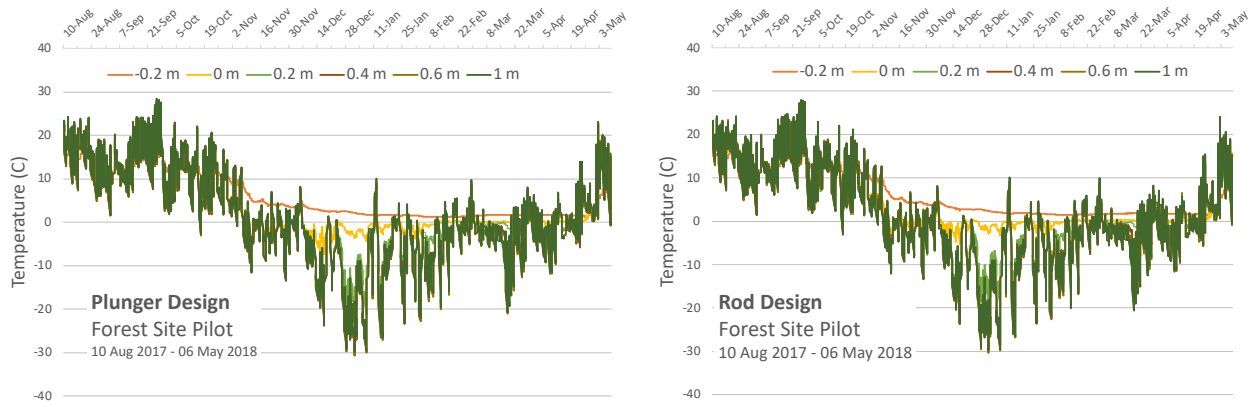


Fig 3.4.1. Comparison between prototypes of temperature profiling arrays using the plunger and rod designs, based on initial testing in forested sites from Aug 2018 to May 2019.

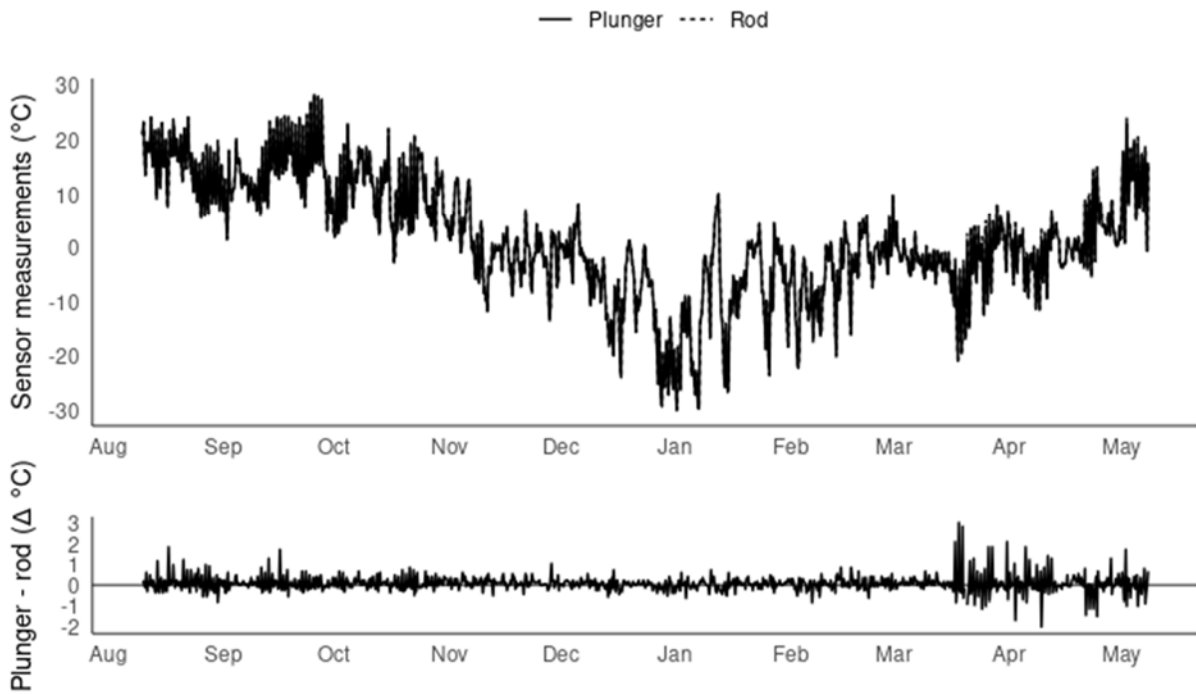


Fig 3.4.2 Average daily temperatures (upper panel) and corresponding average difference (ΔC) between plunger and rod designs, from testing in a forested sites from Aug 2018 to May 2019. Differences were estimated between arrays at the same profile positions (heights relative to soil surface) and aggregated to an average daily difference before plotting.

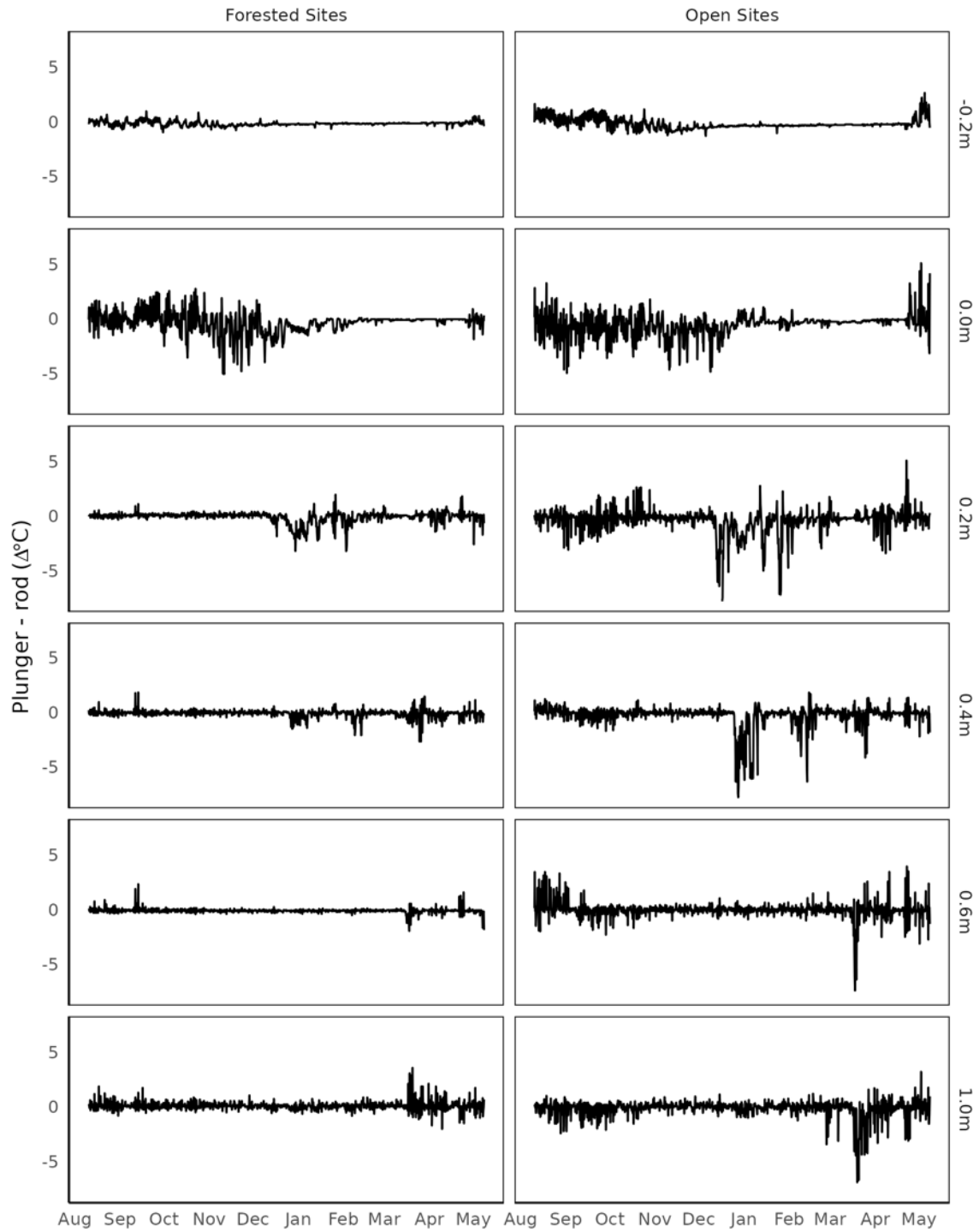


Figure 3.4.3. Detailed comparison of measurement bias (ΔC) between plunger and rod array designs at forested and open sites (and each sensor position) during the prototype testing (Aug 2018-May 2019).

3.5 Final array design

Following the pilot study, we selected the rod design (Fig 3.5.1) for its practicality and efficiency in terms of retrieving sensors, collecting data, reprogramming sensors and redeploying the array to resume data collection. All of these tasks could be completed rapidly in the field without significant disturbance to the installed housing, whereas the plunger design required uninstalling the arrays, returning them to the lab for sensor extraction and reinsertion, and then reinstalling the arrays each time data was collected.

Data from pilot testing also supported this design choice. Although some data discrepancies between design prototypes were evident, these were much less pronounced in the forested setting where our arrays would be deployed. On average the difference between designs was less than 1C during the winter season (Nov – Mar). We noted greater measurement bias (disagreement) between prototypes in the warmer months, especially during the transition from winter to spring (late Mar into early Apr), where air temperatures tend to fluctuate between freezing and thawing on a daily basis.

As a rule, the rod design yielded temperature time series that were more stable and less ‘noisy’ than the plunger design. Upon retrieval of the arrays, we noted that some slippage of the foam discs and sensors might have occurred in the plunger design, meaning that temperatures were not consistently recorded at the intended positions (heights). Such slippage was negated by the rod design itself and any dislocation of the sensors in the rod design, if it had occurred, would have been easy to observe upon retrieval. We also found no support for concerns about the rod design sensors measuring the same temperatures along the vertical profile, due to a lack of thermal stratification within the housing. Pilot data indicated that the air space in the rod housing was thermally stratified and that sensors captured diurnal (24 hr) variation as expected, with lower variability at the below ground position (-0.2 m) versus above ground positions (+0.2 to +1.0 m). Overall, lacking any compelling evidence that the plunger design yielded more reliable data, we selected the rod design for its much greater practicality.

The final array design (Fig 3.5.1) consisted of a 54 inch section of 1.5 inch diameter PVC pipe. The below-ground end of the PVC housing was sealed internally using a #8 rubber stopper with a 0.5 inch plug of cured silicone sealant and externally with a 1.5 inch PVC cap. The aboveground end of the PVC housing had a 1.5 inch (mm) threaded male adapter cemented to the pipe. A 1.5 inch (mm) female cap was threaded onto this male adapter end. A 48 inch length of 0.5 inch diameter PVC rod was attached to the cap with a 1.5 inch stainless steel screw, sealed with silicone caulk on the interior of the cap. The plastic fobs holding the iButton sensors were secured to the PVC rod using #8 stainless steel screws and washers. The PVC rod vertically bisected the housing and only touched the cap (at its attachment point) and the rubber stopper at the bottom where it rests freely (no fastener or adhesive was used).

High-contrast metric scales (Fig 3.5.1) were affixed to the subset of profiling arrays where trail cameras would also be deployed, to allow for visual estimation of snow depth from the daily images recorded. Scales were secured on the housing exterior using plastic zip-ties. For installation in uneven and rocky ground, steel garden stakes were driven into the ground adjacent to the array to secure the vertical orientation and overall stability of the installed housing.

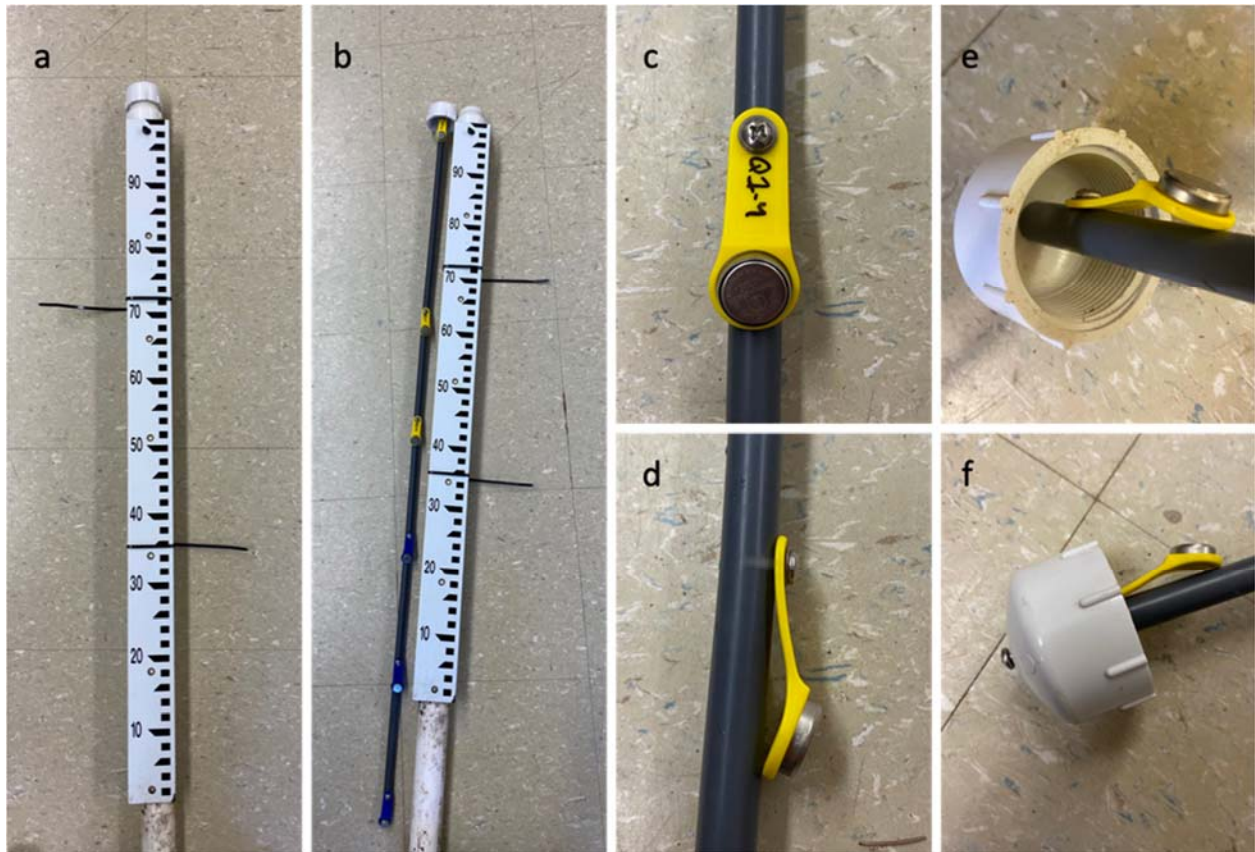


Figure 3.5.1. Final array and housing assembly; a) exterior view, fully assembled with meter scale used to interpret daily camera images of snow depth; b) exterior view, with rod sensor assembly removed; c) and d) detail on iButton and plastic fob attachment point to PVC rod; e) and f) detail on PVC rod attachment to cap and sensor located at the +1.0m position.

4 Array Installation & Retrieval

4.1 Installation

Installation of two or more arrays to sample a single location is recommended. This replication provides measurement redundancy in the event of sensor failure and/or physical disturbance to the installed array during the sampling period. As a rule, replication also improves overall sampling accuracy and supports the estimation of measurement uncertainty for analysis and modeling purposes.

Prior to array installation or any method of sensor deployment, we strongly recommend users create a basis for tracking sensor inventory and compiling related metadata by assigning each iButton a unique identifier that corresponds to each device's unique serial number. Although serial numbers can be used as unique IDs, we created our own unique ID numbers and named each sensor based on its array and profile position to make our metadata more intuitive. We assigned letters to individual arrays (A, B, C, D...) and numbers to represent positions (1-6 from lowest to highest), and these unique names (e.g., A-4, D-1, Q-6, etc) were related to sensor serial numbers in a reference table that included device model number, deployment history, etc. We then created several metadata tables and field logs based on this relational data structure. This approach supports overall management of the array network, including tracking sensor inventory through each deployment cycle, reassigning sensors to different or new arrays and/or positions, estimating sensor malfunction rates and replacement needs, and identifying faulty sensors and/or array housings (i.e., where failure/malfunction rates are much higher than average).

Table 4.1.1. Excerpt of metadata table for iButton inventory, including unique identified (iBID), serial number, assigned array, profile position (Pos), vertical location (Vloc), measurement interval (Rate) and measurement resolution (Res). Note that 'Rate' field represents the same four-hour interval, but contains units of either seconds (14400) or minutes (240) depending on the iButton model.

iBID	Serial Registration No	Model	Array	Pos	Vloc (m)	Rate	Res (C)
001	1400000049BB9E41	DS1922L	E	1	-0.2	14400	0.063
002	BA00000049C33641	DS1922L	E	2	0.0	14400	0.063
003	3100000042E56A21	DS1921G-F5	E	3	0.2	240	0.5
004	4D00000043316C21	DS1921G-F5	E	4	0.4	240	0.5
005	4700000043496F21	DS1921G-F5	E	5	0.6	240	0.5
006	3100000043392D21	DS1921G-F5	E	6	1.0	240	0.5
007	CC00000049CB2841	DS1922L	F	1	-0.2	14400	0.063
008	E800000049CC6841	DS1922L	F	2	0.0	14400	0.063
009	7600000043495B21	DS1921G-F5	F	3	0.2	240	0.5
010	DD00000043560221	DS1921G-F5	F	4	0.4	240	0.5
011	0A0000004356DC21	DS1921G-F5	F	5	0.6	240	0.5
012	5500000043466320	DS1921G-F5	F	6	1.0	240	0.5

For installation, a plot center point was established at each sampling location and two arrays were deployed 3 m from this point, along a north-south azimuth. Holes of 14 in (35 cm) depth were dug with soil augers or a section of steel pipe of similar diameter as the array housing, to minimize disturbance to the adjacent soil strata. Arrays were inserted into the holes and adjusted so that the 'zero line' on the

housing, corresponding to the 0 cm elevation sensor position inside, was aligned as closely as possible with the soil surface. We defined the soil surface as the boundary between leaf litter (O horizon) and the upper A horizon, which was commonly indicated by the location of a dense mat of fine roots. This boundary was located by brushing away litter from the forest floor, by hand, as part of the hole digging. Due to soil heterogeneity, to find sufficient soil depth for secure anchorage, arrays may need to be moved slightly from the measured location (based on plot center point). As needed for stability, a steel garden stake was driven into the ground next to the array, which was secured to the stake with zip ties.

The iButton sensors can be programmed to collect data at user-specified intervals defined by the specific application and requirements of the sampling design. For initial deployment, iButtons can be programmed in the office or lab using an interface device that connects the sensor to a computer or tablet (with necessary software and hardware drivers installed). Many options exist for sensor interface devices. We have primarily used the one-wire USB interface ([model DS1402D-DR8+](#)) manufactured by Analog Devices (formerly Maxim Integrated) for programming as well as field data retrieval.

Table 4.1.2. Excerpt from metadata table describing plot (site) locations, arrays and trail cameras.

Site	Arrays	Camera	Latitude	Longitude	Aspect	Date_Installed
1	Q1, Q3		43.994365	-74.243446	NW	2019-10-21T11:52:54-04:00
2	O2, O4		43.995234	-74.246177	SE	2019-10-21T13:25:41-04:00
3	Q2, Q4		43.996265	-74.243412	SW	2019-10-21T12:30:16-04:00
4	V1, V2	Yes	43.997189	-74.240727	W	2019-10-16T13:14:38-04:00
5	U1, U2		43.998425	-74.237785	W	2019-10-16T13:35:56-04:00
6	Z1, Z2		43.9987	-74.235518	N	2019-10-16T13:53:04-04:00

4.2 Trail cameras

For snow monitoring applications, trail cameras were installed on nearby trees and then positioned and programmed to record a daily image of the full length of the array. A high-contrast metric scale was secured to the PVC array housing using zip ties. Cameras were programmed to record an image once per day at noon. These images allow us to visually estimate snow depth daily at the array location. In turn, these visual estimates provide independent training data for model-based estimation of snow depth using the temperature profile measured by the sensors. Daily images are also useful for interpreting anomalous data caused by physical disturbance to the array during the sampling period.

The choice of camera make and model is left to the investigators. We tested several models over several years until a sufficiently reliable camera was identified. Our current model comes from the [GardePro A3 series](#) manufactured by Shenzhen Zhuopu Digital Technology Co, Ltd. While most trail cameras feature motion-activation, not all feature the ability to disable this function and instead program the device to collect images at regular (daily) intervals. If possible, we recommend either working with camera models that are known to be reliable in adverse conditions during the winter or testing two or more models for a full winter season before selecting a model for full-scale deployment. Power failure is a common issue in cold temperatures, so we recommend always using fresh alkaline (non-rechargeable) batteries each deployment. Each camera will also require a microSD card for data storage and transfer. Camera images are downloaded from the SD card when arrays are visited for data retrieval and sensor reprogramming.

4.3 Data retrieval

Collection of the logged temperature data from the profiling arrays is a straightforward process. First, the sensor array is accessed by unscrewing and removing the upper cap from the housing, which should remain installed in the ground. Care should be taken to minimize disturbance to the installed housing, unless adjustments are needed. Once the cap is removed with its attached PVC rod and sensor array assembly, the technician uses an i-Button compatible reader device to connect with each sensor. We used a [DS1402d-DB8](#) iButton reader (Maxim Integrated, LLC) for both indoor and outdoor purposes.

Once the sensor connection is established, the technician runs a diagnostic test, downloads/transfers logged data, clears the sensor memory, and reprograms the sensor for redeployment. User interface software provided with the iButtons supports these functions, as well as data management and reuse of mission parameters when reprogramming each sensor. In addition, most reader devices do not require removing sensors from plastic fobs, allowing the technician to efficiently proceed through each sensor on the array. Sensors need only be removed from fobs when being replaced. Once all sensors are processed and any replacements are programmed, the array is reassembled by inserting the rod into the housing. Silicone caulk may be reapplied at this time to weather seal the array housing.

Although reliable and durable, iButtons have a finite lifespan; some will experience malfunctions that cannot be repaired, and all will eventually deplete their built-in power supply. Replacement rates in this project were approximately 5-10% annually within the first 3 years of deployment. We anticipated a much higher rate of sensor failure after 3 years due to exhaustion of power supply and general 'wear-and-tear' over multiple deployment cycles. Overall, we recommend that users plan to replace their inventory of iButtons every 5 years, while maintaining a cache of new iButtons available for prompt replacement. For applications in remote or difficult to access locations, we recommend carrying these replacement iButtons on data retrieval trips (when malfunctioning devices are typically identified).

5 Data Processing & Analysis

5.1 Pre-processing and quality control

Using software provided with iButton 1-wire reader devices, data retrieved from sensors can be viewed and exported to a variety of common file formats for further processing and analysis, including Excel spreadsheets, statistical software platforms such as Project R, and relational databases such as Access and PostgreSQL. Screening the raw downloaded data for errors is an important step in quality control prior to data processing and analysis. In our experience these errors mostly correspond with a terminal malfunction of the sensor, resulting in continuous periods of invalid or missing data that are easy to identify during data retrieval and export. There can also be anomalous data points included in an otherwise valid time-series of measurements. Multiple options exist for locating and handling these errors as part of a data quality control (QC) procedure, including some basic functions of the sensor interface / data retrieval software provided with the iButtons. We recommend exporting the raw data to a spreadsheet or text (csv) file for archival purposes, then implementing QC filters or other automated procedures to generate a 'clean' set of individual sensor datasets that can be compiled into aggregate datasets as needed. For example, if all sensors were programmed to record at the same times each day (as recommended), data from one or more arrays can be compiled efficiently into a single data table with the same time/date field (i.e., individual sensors as columns, measurement times as rows).

5.2 Soil and forest floor microclimate

A basic application of the profiling arrays is the measurement and monitoring the soil and forest floor microclimate at one or more locations. Comparisons of subsoil, soil surface and near-surface air temperatures among multiple sites across a study landscape, such as a long-term monitoring watershed (Figure 5.2.1), can provide essential data for a variety of basic and applied research questions. In the example below, we deployed paired arrays at each plot, using a paired plot design to sample three areas of the Archer Creek Watershed: south-facing upland, northwest-facing upland and wetland (beaver meadow). For the south-facing upland we located plots in a well-drained upland and adjacent headwater wetland to estimate impacts of soil hydrologic regime on microclimate. We found the more saturated soil (S14HeadwaterWetland) was significantly warmer than better drained soils (at -20cm).

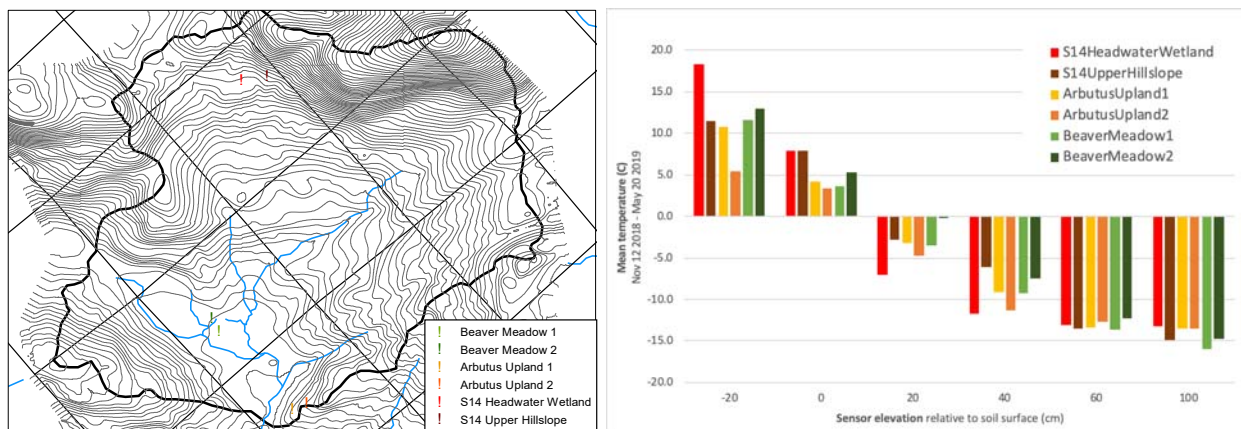


Figure 5.2.1. Microclimate monitoring during the 2018-19 winter at three locations (6 plots, 12 arrays) in the Archer Creek watershed at Huntington Forest, Newcomb, NY. Mean temperature chart compares plot-level data and is arranged L to R by sensor elevation(+, from subsoil (-20 cm) to near-surface (+1 m)).

5.3 Snow depth modeling and monitoring

Vertical profiling methods for modeling snow depth use observed differences in temperature along the profile to estimate the height at which snow cover has detectable insulative effects. Air temperatures typically oscillate between highs and lows on nominal 24 hr (daily) cycles, so thermometers exposed to the air will measure this diurnal pattern, while those insulated from the air will be more locally stable. Air and soil microclimates have different temperature regimes in terms of both average temperatures and their variability over short time periods. In our profile design, the top and bottom end members of the array represent temperature regimes for the ambient air (except if snow depth > 1m) and the subsoil (except if exposed to air) microclimates, respectively. For the remaining sensors, the relative similarity of their measured temperature regimes to these air and soil reference points indicates whether there is any insulative effect present at the height measured by the sensor. In other words, a sensor recording a temperature regime similar to the highest (+1m) position is exposed to the air, while one similar to the subsoil (-0.2m) is insulated from air temperatures to some degree.

Difference between temperature regimes along the profile may be quantified in various ways, but how well the arrays capture these differences can be tested by visualizing the temperature profile as a time series (Fig 5.3.1). Before snow begins to accumulate on the forest floor, the ground-level sensor (0m) will track with the other sensors positioned aboveground, although with less diurnal variability or 'noise', than the overall warmer and more temporally stable subsoil (-0.2m) temperature regime. Sensors that are all measuring air temperature will align very closely, but as snow accumulates above their profile height and insulates them, they will shift towards the warmer and more stable subsoil regime. These shifts can be observed in the separation among time-series plots, which indicate persistent snowpack, and the realignment of time-series plots when snow melts in spring.

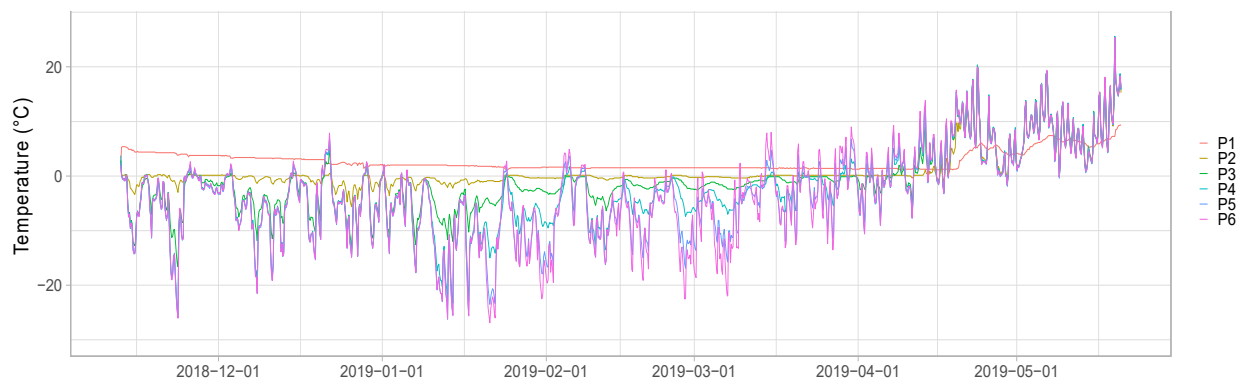


Fig 5.3.1. Temperature time-series measurements of a full profiling array during the 2018-19 winter. Array positions (in m) are P1 = -0.2 (subsoil), P2 = 0, P3 = +0.2, P4 = +0.4, P5 = +0.6, P6 = +1.0. Note that measurement began in mid-November 2018 when a thin ground layer of snow was already present.

Although snow depth can be approximated in this qualitative way, a more objective method that quantitatively relates temperature variation along the profile to snow depth is needed for multiple reasons. Among them, snow characteristics can vary between snowfall events and change over time with weather patterns, especially with thaw-freeze cycles. These characteristics directly affect the insulative capacity of the existing snowpack, meaning that the amount of snow needed to stabilize a sensor's temperature regime can and likely does vary during the winter. Also, neither the heat conducting properties of the PVC housings once installed, nor the consistency of thermal stratification of

the air space inside the housing, are well understood. These parameters relate to both the accuracy and precision of snow depth estimates derived from analysis of variation along the temperature profile.

A better approach for estimating snow depth from profiling data requires having continuous measures of snow depth at the array location. We deployed trail cameras at 6 arrays to take daily images of the array and interpreted these images to generate a reference snow depth dataset for comparison and analysis with the profile time-series (Fig 5.3.2). These reference data can be treated as ‘true’ measures for modeling how differences among temperature regimes along the vertical profile relate to observed snow depth. Models trained on these data (Fig 5.3.3) can then be used to generate predictions of snow depth for any array with full profiling data, regardless of whether snow depth imagery is available.

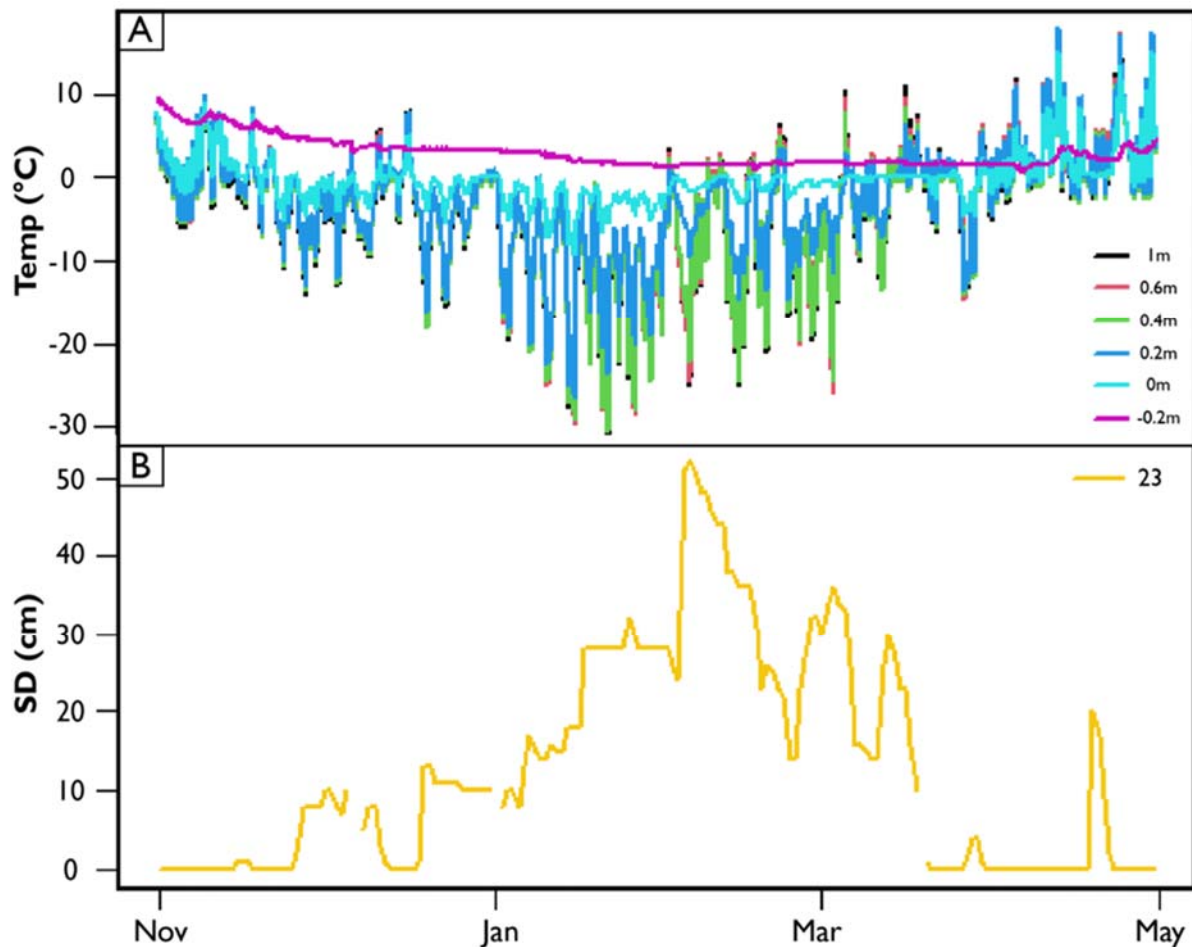


Figure 5.3.2. Comparison of array temperature profile measurements (A) and daily snow depth (B) based on trail camera images at the selected array, during the 2021-22 winter (01 Nov 2021 – 01 May 2022) in the Archer Creek watershed. Snow depth (SD) was estimated visually based on daily images.

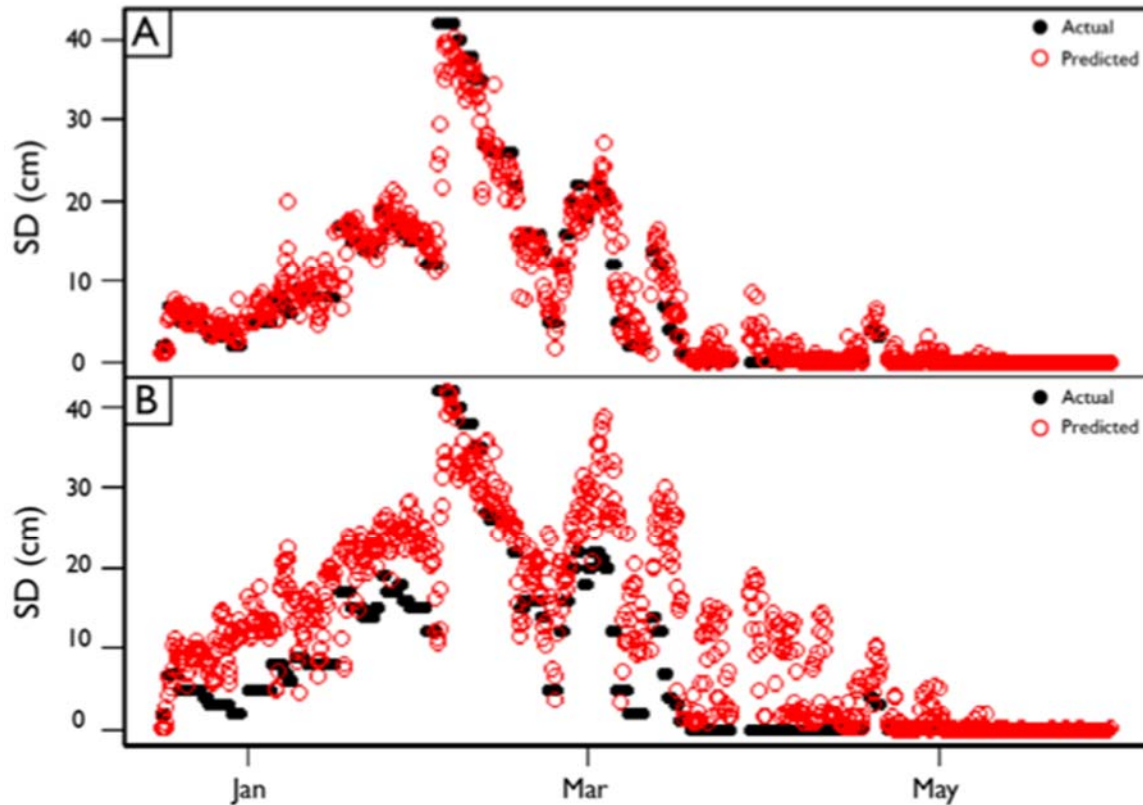


Figure 5.3.3. Snow depth modeling results, comparing daily estimates from camera images (black dots) with model predictions (red circles) for one array location in the Archer Creek watershed, from 15 Dec 2021 – 15 June 2022. Panels A and B show outcomes of different model training techniques using the same observation set with Random Forest regression using the *ranger* R package. Average model errors ranged between 2-4 cm snow depth for Model A and 5-8 cm snow depth for Model B.

5.4 Data analysis & visualization using Program R

All data compilation, analysis/modeling and visualization in this report was conducted using the open-source Program R. Custom R programming code (scripts) developed specifically for this purpose are provided below. A zipped folder that includes these code resources, R markdown files and example datasets is available upon request by emailing cbeier@esf.edu.

delta_graphic.R – compares two or more arrays via subtraction (delta) of sensor measurements at each vertical position and visualizes results

```
library(lubridate)
library(ggplot2)
library(tidyr)
forest_rod <- read.csv("arbutus-ibutton-data/pilot/forested_site_rod_design.csv")
forest_plunger <- read.csv("arbutus-ibutton-data/pilot/forested_site_plunger_design.csv")
open_rod <- read.csv("arbutus-ibutton-data/pilot/open_site_rod_design.csv")
open_plunger <- read.csv("arbutus-ibutton-data/pilot/open_site_plunger_design.csv")

forest_deltas <- cbind(forest_rod[1], forest_plunger[2:7] - forest_rod[2:7])
open_deltas <- cbind(open_rod[1], open_plunger[2:7] - open_rod[2:7])

rbind(
  cbind(
    setNames(forest_deltas, c("time", "-0.2m", "0.0m", "0.2m", "0.4m", "0.6m", "1.0m")),
    site = "Forested Sites"
  ),
  cbind(
    setNames(open_deltas, c("time", "-0.2m", "0.0m", "0.2m", "0.4m", "0.6m", "1.0m")),
    site = "Open Sites"
  )
) |>
pivot_longer(-c(time, site)) |>
mutate(time = as_datetime(time)) |>
ggplot(aes(time, value)) +
geom_line() +
scale_x_datetime(date_breaks = "1 month", date_labels = "%b") +
scale_y_continuous(limits = c(-8, 7.5)) +
facet_grid(name ~ site) +
theme_minimal() +
theme(panel.grid = element_blank(),
      axis.line = element_line(),
      panel.border = element_rect(fill = NA),
      plot.background = element_rect(fill = "white", color = NA)) +
labs(x = "",
      y = expression("Plunger - rod"~"("Delta ~ degree"C"))

ggsave("temp_deltas.png", width = 7, height = 9, units = "in")
```

granger_test.R -- uses the non-parametric Granger causality test to describe relationships among array measurements (means, standard deviation) across vertical positions and over time

```
library(lubridate)
library(ggplot2)
library(tidyr)
forest_rod <- read.csv("arbutus-ibutton-data/pilot/forested_site_rod_design.csv")
forest_plunger <- read.csv("arbutus-ibutton-data/pilot/forested_site_plunger_design.csv")
open_rod <- read.csv("arbutus-ibutton-data/pilot/open_site_rod_design.csv")
open_plunger <- read.csv("arbutus-ibutton-data/pilot/open_site_plunger_design.csv")

forest_deltas <- cbind(forest_rod[1], forest_plunger[2:7] - forest_rod[2:7])
open_deltas <- cbind(open_rod[1], open_plunger[2:7] - open_rod[2:7])

rbind(
  cbind(
    setNames(forest_deltas, c("time", "-0.2m", "0.0m", "0.2m", "0.4m", "0.6m", "1.0m")),
    site = "Forested Sites"
  ),
  cbind(
    setNames(open_deltas, c("time", "-0.2m", "0.0m", "0.2m", "0.4m", "0.6m", "1.0m")),
    site = "Open Sites"
  )
) |>
pivot_longer(-c(time, site)) |>
mutate(time = as_datetime(time)) |>
ggplot(aes(time, value)) +
geom_line() +
scale_x_datetime(date_breaks = "1 month", date_labels = "%b") +
scale_y_continuous(limits = c(-8, 7.5)) +
facet_grid(name ~ site) +
theme_minimal() +
theme(panel.grid = element_blank(),
      axis.line = element_line(),
      panel.border = element_rect(fill = NA),
      plot.background = element_rect(fill = "white", color = NA)) +
labs(x = "",
      y = expression("Plunger - rod"~"("Delta ~ degree"C"))

ggsave("temp_deltas.png", width = 7, height = 9, units = "in")
```

ml_snow_depth.R – snow depth modeling (training, testing, validation) using temperature array data via the ‘ranger’ implementation of Random Forests algorithm

```
library(pROC)
library(dplyr)
library(tidyr)
library(caret)
library(ranger)
library(ggplot2)
library(yardstick)
library(patchwork)
set.seed(1107)

### Set up functions
# Function to calculate lagged temps and rolling sd's:
shape_data <- function(input) {
  input |>
  mutate(
    temp_lag4 = temp - lag(temp, 1),
    temp_lag8 = temp - lag(temp, 2),
    temp_lag12 = temp - lag(temp, 3),
    temp_lag16 = temp - lag(temp, 4),
    temp_lag20 = temp - lag(temp, 5),
    sd_lag4 = zoo::rollapplyr(temp, FUN = sd, width = 2, fill = NA),
    sd_lag8 = zoo::rollapplyr(temp, FUN = sd, width = 3, fill = NA),
    sd_lag12 = zoo::rollapplyr(temp, FUN = sd, width = 4, fill = NA),
    sd_lag16 = zoo::rollapplyr(temp, FUN = sd, width = 5, fill = NA),
    sd_lag20 = zoo::rollapplyr(temp, FUN = sd, width = 6, fill = NA)
  ) |>
  drop_na()
}

# Function to apply shape_data to each height:
reshape_data <- function(input) {
  input <- input |>
  select(Date_Time, snow = Cam_SD, starts_with("T"))

  input <- input |>
  pivot_longer(-c(Date_Time, snow), values_to = "temp") |>
  select(-Date_Time, -snow) |>
  group_by(name) |>
  nest(data = -c(name)) |>
  mutate(data = purrr::map(data, shape_data)) |>
  pivot_wider(values_from = data) |>
  unnest(everything(), names_sep = "_") |>
  bind_cols(tail(input, -5))
}
```

```

### Reshape data
# Reshape K and fit random forest using default parameters:
# (note: to confirm no data leakage, you can move the P2 block above this line)
load("K_training_set.RData")
reshaped_K <- reshape_data(training_set_K)

trained_on_K <- ranger(
  snow ~ .,
  data = reshaped_K[setdiff(names(reshaped_K), "Date_Time")]
)

# Reshape P2 and fit random forest using default parameters:
load("P2_training_set.RData")
reshaped_P <- reshape_data(training_set_P2)

trained_on_P <- ranger(
  snow ~ .,
  data = reshaped_P[setdiff(names(reshaped_P), "Date_Time")]
)

###
# Predict data and check performance metrics:
reshaped_P$predictions <- predict(trained_on_K, reshaped_P) |>
  predictions()

reshaped_K$predictions <- predict(trained_on_P, reshaped_K) |>
  predictions()

# Predicting P is rough!
metrics(reshaped_P, snow, predictions)
# Predicting K is much easier:
metrics(reshaped_K, snow, predictions)
# Interestingly enough, RMSE/MAE aren't wildly different between the two.
# R2 is though.

# We get the rough trend adequately enough, but really mess up in the spring:
(p <- ggplot(reshaped_P) +
  geom_line(aes(Date_Time, snow), color = "blue") +
  geom_line(aes(Date_Time, predictions), linetype = 2) +
  ggtitle("Trained on K, predicted on P") +
  scale_y_continuous(limits = c(0, 55)))

# We miss peak heights in winter but otherwise are fine enough:
(k <- ggplot(reshaped_K) +
  geom_line(aes(Date_Time, snow), color = "blue") +
  geom_line(aes(Date_Time, predictions), linetype = 2) +
  ggtitle("Trained on P, predicted on K") +
  scale_y_continuous(limits = c(0, 55)))

```

```
p + k & theme_minimal()
```

```
### Measure accuracy on presence/absence predictions:
```

```
# Get a threshold to use for each model:
```

```
threshold_trained_on_P <- pROC::roc(  
  reshaped_P$snow > 0,  
  predict(trained_on_P, reshaped_P) |> predictions()  
) |>
```

```
  pROC::coords("best")
```

```
threshold_trained_on_P <- threshold_trained_on_P$threshold
```

```
threshold_trained_on_K <- pROC::roc(  
  reshaped_K$snow > 0,  
  predict(trained_on_K, reshaped_K) |> predictions()  
) |>
```

```
  pROC::coords("best")
```

```
threshold_trained_on_K <- threshold_trained_on_K$threshold
```

```
# On the question "is there any snow at all" we do alright, though not perfect.
```

```
# I'm thinking we'd probably benefit from more training data, both for non-snow
```

```
# conditions and for other sensors.
```

```
caret::confusionMatrix(  
  factor(reshaped_P$predictions > threshold_trained_on_K),  
  factor(reshaped_P$snow > 0)  
)
```

```
caret::confusionMatrix(  
  factor(reshaped_K$predictions > threshold_trained_on_P),  
  factor(reshaped_K$snow > 0)  
)
```