# WEARABLE SENSORS FOR PERSONAL TEMPERATURE EXPOSURE ASSESSMENTS: A COMPARATIVE STUDY

A Thesis
by
ELIZABETH F. BAILEY

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTER OF ARTS

May 2019
Department of Geography and Planning

WEARABLE SENSORS FOR PERSONAL TEMPERATURE EXPOSURE ASSESSMENTS: A COMPARATIVE STUDY

A Thesis
by
ELIZABETH F. BAILEY
May 2018

APPROVED BY:

Margaret M. Sugg, Ph.D.
Chairperson, Thesis Committee

Christopher Fuhrmann, Ph.D.
Member, Thesis Committee

Jennifer Junkle, Ph.D.
Member, Thesis Committee

Max C. Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

**Abstract**

WEARABLE SENSORS FOR PERSONAL TEMPERATURE EXPOSURE ASSESSMENTS: A COMPARATIVE STUDY

Elizabeth F. Bailey
B.A., Queens University of Charlotte
M.A., Appalachian State University

Chairperson: Margaret M. Sugg, Ph.D.

Heat exposure is the leading weather-related cause of death in the United States. The impacts of heat on human health has sparked research on different approaches to measure, map, and predict heat exposure at more accurate and precise spatiotemporal scales. Personal heat sensor studies rely on small sensors that can continuously measure ambient temperatures as individuals move through time and space. The comparison between different types of sensors and sensor placements have yet to be fully researched. The objective of this study is to assess the validity of personal ambient temperature sensors. To accomplish this objective, we evaluate the performance of multiple low-cost wearable sensors (HOBOs, iButton Thermochrons, iButton Hygrochrons, and Kestrel DROP D3FW Fire) for measuring ambient temperature in a (1) field exposure study by varying the placement on human subjects and in a (2) field calibration study by co-locating sensors with fixed site weather station monitors. A secondary aim involved investigating consensus between validation metrics that can be used in future sensor comparison studies. Bland-Altman analysis, correlation coefficients, and index of agreement statistics were used to quantify the difference between sensor and weather station ambient temperature measurements. Results demonstrated significant differences in measured temperatures for sensors based on sensor type and placement on participants. Future research should account for the differences in personal ambient temperature readings based on sensor type and placement.

## Acknowledgments

I owe thanks to numerous individuals and organizations for their assistance with this thesis. There are not enough words to properly convey my gratitude to my advisor, Dr. Maggie Sugg. She has been the most influential part of my graduate career and my thesis would not be possible without her constant support and guidance. Dr. Sugg's genuine passion for her work is evident in her incredible work ethic and dedication to her students. Dr. Sugg's guidance and support throughout my graduate education has had a profound impact on my life both academically and personally and will shape my life's path for years to come. I also want to thank my committee members, Dr. Christopher Fuhrmann, and Dr. Jennifer Runkle, who offered their guidance and expertise throughout the thesis process. A special thanks to Dr. Christopher Fuhrmann and his team at Mississippi State University for running the Mississippi State University data collection.

I would like to recognize all the participants who volunteered to wear sensors in the Mississippi State University and Appalachian State University Studies. This research would not have been possible without the data collected from their participation. I would like to thank Scott Stevens for his work merging and cleaning data from the participant study. I would also like to thank Dr. Baker Perry for allowing us to utilize his weather station at Appalachian State University.

I also owe thanks to Appalachian State's GSAS and Office of Student Research funding for enabling me to present the culmination of my research at the Association of American Geographers Annual Meeting and the SouthEastern Division of the American Geographers Annual Meeting. I am thankful for all the opportunities provided to me during my past two years at Appalachian State. I am especially fortunate to have been a part of the Department of Geography & Planning, where I had the opportunity to get to know and work with many of the wonderful faculty members. I would like to thank Dr. Saskia van de Gevel, Dr. Derek Martin, and Dr. Baker Perry for supervising my graduate research and teaching assistantship and for providing me the opportunity to take place in their research and courses.

I am grateful for the support of my fellow graduate students in the Department of Geography & Planning for their continued support during my graduate education. I would also like to give a special thanks to Lauren Andersen for all of her advice, support, and help with editing throughout my thesis process. Finally, I owe my greatest appreciation to my parents, Douglas and Paula Bailey for always being my biggest advocates. I would also like to thank my family, Amy and Alan Elam, Geb Bailey, Lee and Marvin Cook, and Valarie Oliver for their continued support throughout my life. My family has made uncountable personal sacrifices to support my educational endeavors and my educational pursuits would not be possible without their unwavering love and support.

## Table of Contents

Abstract ........................................ iv
Acknowledgments ........................................ vi
List of Tables ........................................ ix
List of Figures ........................................ x
Foreword ........................................ xi
Introduction ........................................ 1
Journal Article: *Wearable Sensors for Personal Temperature Exposure Assessments: A Comparative Study* ........................................ 4
  Abstract ........................................ 5
  Key Words ........................................ 6
  Highlights ........................................ 6
  Introduction ........................................ 7
  Data and Methods ........................................ 11
  Results ........................................ 17
  Discussion ........................................ 22
  Conclusion ........................................ 27
  Acknowledgments ........................................ 28
References ........................................ 29
Vita ........................................ 54

## List of Tables

Table 1. Participant demographics for the MSU and ASU study sites. .................................. 34

Table 2. Temperature Exposure Assessment Manufacturer Guidelines. ................................ 35

Table 3. Summary statistics for sensor temperature readings during the ASU and MSU participant study. ........................................................................................................ 36

Table 4. Summary statistics for three types of microenvironments (outdoor, indoor, in-transit) experienced during the participant study at ASU. ................................................................. 37

Table 5. Summary statistics for different activity intensity levels (low, moderate, high) during the participant study at ASU.............................................................................. 38

Table 6. Quantitative analysis of participant activities during the maximum temperatures reached by each device during the ASU participant study. .................................................... 39

Table 7. Summary statistics for the weather station and wearable sensor devices in Boone, NC and Starkville, MS. ......................................................................................... 40

Table 8. Measures of agreement between the stationary sensors and weather station. .......... 41

Table 9. A Linear mixed model with a random intercept for each participant comparing the fixed weather station data to participant sensor measurements at ASU. ................................ 42

Table 10. A Linear mixed model with a random intercept for each participant comparing the fixed weather station data to participant sensor measurements at MSU. ............................... 43

## List of Figures

Figure 1. Study area map showing the locations of Boone, NC and Starkville, MS. ............ 44
Figure 2. Placement of sensors for ASU and MSU participant studies. ................................ 45
Figure 3. Sensors attached to the weather station located in Starkville, Mississippi and Boone, North Carolina. ........................................................................................................ 46
Figure 4. Bland Altman plots for the sensors with the highest and lowest mean differences at ASU and MSU. ............................................................................................................... 47
Figure 5. Box plots showing mean temperatures during low, moderate, and high activity levels throughout the ASU participant study. ...................................................................... 49
Figure 6. A comparison of the Appalachian State University weather station and personal sensor temperature readings from July 26th, 2018 through July 29th, 2018. ......................... 51
Figure 7. Correlation values between weather station in-situ temperature readings and sensor temperature readings from sensors attached to the weather station in Boone, NC................ 52
Figure 8. Correlation values between weather station in-situ temperature readings and sensor temperature readings from sensors attached to the weather station in Starkville, MS. .......... 53

## Foreword

The main body of this thesis is formatted to the guidelines for manuscript submission to *Environmental Research*, a multidisciplinary journal of environmental sciences, ecology, and public health.

**Introduction**

Environmental heat exposure is a prevalent human health hazard. Heat-related illness, also called hyperthermia, is a condition that occurs when core body temperatures exceed ranges needed for physiological functioning (Kuras et al. 2017). Studies show a common trend of enhanced morbidity and mortality in populations when heat deviates above average. This effect is especially prevalent when heat ratings exceed the 95th percentile of experienced temperatures for a specific geographic location (Gosling et al. 2009). On average, heat waves cause more weather-related fatalities than any other type of natural disaster in the United States (Bernard and Mcgeehin 2004).

Due to the adverse effects of heat on public health, a substantial reservoir of heat and health-related articles have been published in the last decades (e.g., Fuhrmann et al. 2016, Gosling et al. 2009, Hondula et al. 2012, Hondula et al. 2015, McGeehin and Mirabelli 2001, Noe et al. 2012). The impacts of heat on human health has sparked research on the best approach to map and predict heat exposure on more accurate and precise scales. This has led to a variety of personal heat sensor studies that have identified many factors, both physical and social, that contribute to an increased risk of increased heat exposure (Basu and Samet 2002, Bernard et al. 2015, Kuras et al. 2017, Sugg et al. 2018). An analysis of previous research illustrates the gaps in the literature that currently exist, limiting the accuracy of current heat alert systems.

One of the most common difficulties in the creation of adequate alert systems exists in the identification of spatial, physiological, and social differences in heat vulnerability and resilience (Kuras et al. 2017). People experience heat differently based on age, gender, body mass and surface area, hydration status, metabolic rate, preexisting health conditions, psychological state, acclimatization, indoor heat exposure, time patterns, and outdoor microclimates (Bernhard et al. 2015, Chan et al. 2001; Chen and Ng 2012; Kuras et al. 2017). Personal exposure, time-activity patterns, and differential factors are not considered in population-based, large-scale studies. The understanding of personal heat exposure is essential to the identification of groups that are the most susceptible to heat-related illnesses and to prevent easily mitigated health impacts. Moreover, the identification of vulnerable populations and spatial vulnerability trends can provide useful tools for future policy and mitigation measures involving public health and heat (Hondula et al. 2012).

Personal heat exposure studies are pilot in nature and focus on individual populations or sub-groups. These small scales are beneficial for identifying geographic, socioeconomic, and physiological vulnerabilities within populations. However, the small scales and inconsistent methods between studies make comparisons across space and time difficult. This lack of comparison eliminates the ability to establish larger spatial or socioeconomic trends between ambient temperatures and heat-related illness. A comprehensive set of standards within the field are needed to create a model that would allow for comparisons between studies, allowing broader personal trend questions to be answered (Kuras et al. 2017).

Another limitation to the comparability between studies exists in the literature gap of personal sensor validation. A substantial literature gap exists in combining meteorological data and personal heat exposure methodology and technology (Kuras et al. 2017). To date, few studies compare personal temperature sensors (HOBO, Kestrel, iButton) with meteorological or remote sensed data. This literature gap creates difficulties in comparing past studies to each other that utilize different sensor types or placements on participants. Quantifying the differences in sensors will allow for error to be reduced in future studies.

There is also a literature gap in the variance between the different personal monitors and a gap in knowledge regarding what device is best suited for accurately mapping personal heat exposure.

While no known studies have focused solely on personal heat sensor validation, multiple studies have been published in the field of personal air quality sensor validation. Air sensor validation studies have developed a comprehensive set of methodology for the identification of differences between studies (Dons et al. 2017, Lewis and Edwards 2016, Ueberham and Schlink 2018). Past validation studies have centered on Bland Altman analysis to quantify the differences between sensors and determine when sensors deviate from each other (Dons et al. 2017, Stahl et al. 2016, Ueberham and Schlink, 2018). Studies have also utilized Pearson's correlations (Stahl et al. 2016), mean absolute percentage error (Stahl et al. 2016, Ueberham and Schlink, 2018), Lin's concordance correlation coefficients (Dons et al. 2017) and Taylor diagrams (Ueberham and Schlink, 2018). To date, no studies utilize this set of methodology on the validation of personal temperature sensors.

The objective of this study is to assess the literature gap in the validity of personal ambient temperature sensors. To accomplish this objective, we evaluate the performance of multiple low-cost wearable sensors (HOBOs, iButton Thermochrons, iButton Hygrochrons, and Kestrel DROP D3FW Fire) for measuring ambient temperature in a (1) field exposure study by varying the placement on human subjects and in a (2) field calibration study by co-locating sensors with fixed site weather station monitors. A secondary aim involved investigating consensus between validation metrics that can be used in future sensor comparison studies.

# Wearable Sensors for Personal Temperature Exposure Assessments: A Comparative Study

Elizabeth F. Bailey[1], Dr. Margaret Sugg[1], Dr. Christopher Fuhrmann[2], and Dr. Jennifer Runkle[1]

[1]Department of Geography and Planning, Appalachian State University, Boone, NC, USA

[2]Department of Geosciences, Mississippi State University, Mississippi State, MS, USA

## Abstract

Heat exposure is the leading weather-related cause of death in the United States. The impacts of heat on human health has sparked research on different approaches to measure, map, and predict heat exposure at more accurate and precise spatiotemporal scales. Personal heat sensor studies rely on small sensors that can continuously measure ambient temperatures as individuals move through time and space. The comparison between different types of sensors and sensor placements have yet to be fully researched. The objective of this study is to assess the validity of personal ambient temperature sensors. To accomplish this objective, we evaluate the performance of multiple low-cost wearable sensors (HOBOs, iButton Thermochrons, iButton Hygrochrons, and Kestrel DROP D3FW Fire) for measuring ambient temperature in a (1) field exposure study by varying the placement on human subjects and in a (2) field calibration study by co-locating sensors with fixed site weather station monitors. A secondary aim involved investigating consensus between validation metrics that can be used in future sensor comparison studies. Bland-Altman analysis, correlation coefficients, and index of agreement statistics were used to quantify the difference between sensor and weather station ambient temperature measurements. Results demonstrated significant differences in measured temperatures for sensors based on sensor type and placement on participants. Future research should account for the differences in personal ambient temperature readings based on sensor type and placement.

**Key Words**

Personal Ambient Temperature, Wearable Sensors, Heat Exposure, Sensor Validation

**Highlights**

- Differences in temperatures occur based on sensor placement on the participant's body.
- Differences in temperature occur between different types of sensors.
- The Kestrel is the most correlated to the weather station out of all sensors.
- Devices attached to the weather station facing towards solar radiation have higher rates of error.

## 1. Introduction

Due to the adverse effects of heat on public health, a wealth of heat and health-related articles have been published in the last decades (e.g., Fuhrmann et al. 2016, Gosling et al. 2009, Hondula et al. 2012, Hondula et al. 2015, McGeehin and Mirabelli 2001, Noe et al. 2012). The impacts of heat on human health has sparked research on the best approach to measure, map, and predict heat exposure at the individual scale. Human exposure assessments are currently dominated by the use of data from expensive and fixed measurement stations that may be analyzed with modeling techniques like interpolation, land-use regression or dispersion models (e.g., Dong et al. 2014, Rhea et al. 2012, Sheridan and Kalkstein 2010, Williams et al. 2012, Zhou et al. 2014). These data are helpful for general conclusions related to population-level health inferences but have limitations when assessing individual exposure due to the diverse microclimates, difference in individual physiological conditions, and different activity levels that people experience in their daily lives that cannot be fully captured by coarse weather station data. Recent advances in wearable sensor and GPS technology has led to a variety of personal heat sensor studies that have identified many factors, both physical and social, that contribute to an increased risk of heat-related health outcomes (Basu and Samet 2002, Bernhard et al. 2015, Kuras et al. 2017, Sugg et al. 2018). Personal heat sensor studies often utilize small wearable technology that tracks an individual's microclimate and temperature throughout the day. These devices allow for the collection of data at more precise temporal and spatial scales than weather station data can provide, and results have shown significant heterogeneity among participants depending on job, demographics, and/or geographic location.

One of the most common difficulties in the mitigation of heat health impacts and the creation of adequate heat alert systems is the identification of spatial, physiological, and social differences in heat vulnerability and resilience at the individual level (Kuras et al. 2017). Some of the differences that impact how people thermoregulate include, age, gender, body mass and surface area, hydration status, metabolic rate, preexisting health conditions, psychological state, acclimatization, and patterns of exposure (Bernhard et al. 2015, Chan et al. 2001; Chen and Ng 2012; Kuras et al. 2017). Personal exposure and time-activity patterns are not considered at the individual level for large population-based studies. The understanding of personal heat exposure is essential to the identification of individuals that are the most susceptible to heat-related illnesses in order to prevent easily mitigated health impacts. Moreover, the identification of vulnerable populations can provide useful information for future policy and mitigation measures involving heat and public health (Hondula et al. 2012).

To-date, the majority of personal heat exposure studies have been pilot in nature and focused on specific populations or sub-groups. These small scales are beneficial for identifying geographic, socioeconomic, and physiological vulnerabilities within populations. However, the small scales and inconsistent methods between studies make comparisons across space, time, and different populations difficult. This lack of comparison eliminates the ability to establish broader spatial or socioeconomic trends between ambient temperatures and heat-related illnesses (Kuras et al. 2017). A comprehensive validation study comparing different types of sensors and sensor placements is needed to compare studies that utilize different types of sensors. Broader personal temperature exposure trend questions cannot be answered until studies can be compared (Kuras et al. 2017).

A substantial literature gap not only exists in the comparison of different types of sensors but also in the comparison between standard meteorological data and personal heat sensor technology (Kuras et al. 2017). To date, few studies compare observations from personal temperature sensors (HOBO, Kestrel, Thermochron, Hygrochron) to meteorological or remote sensing data (Bernhard et al. 2015, Sugg et al. 2018, Kuras et al. 2015). While there are some studies that compare personal temperature sensors to meteorological data, the studies have not yielded consistent results. Bernhard et al. (2015) found that outdoor personal ambient temperature readings (PAT) predicted a 0.5 ℃ increase in temperature for each 1℃ increase at a neighboring weather station. Kuras et al. (2015) and Sugg et al. (2018) reported similar finding with less significant results than Bernhard et al. (2015). The influencing factors lending to significant differences between different types of sensors and weather station data has not yet been quantified. This literature gap creates difficulties in comparing past studies that utilize different sensor types or placements on participants. Quantifying the temperature differences will allow comparisons of different studies and the generation of recommendations as to which device is best suited for accurately mapping personal heat exposure during participant based studies.

While no known studies have focused solely on personal heat sensor validation, multiple studies have been published in the field of personal air quality sensor validation. Air quality sensor validation studies have developed a comprehensive set of methodology for the identification of differences between sensors (Dons et al. 2017, Lewis and Edwards 2016, Ueberham and Schlink 2018). Past validation studies have centered on Bland Altman analysis to quantify the differences between sensors and determine when sensors deviate from each other (Dons et al. 2017 and Ueberham and Schlink, 2018). Studies have also utilized mean absolute percentage error (Ueberham and Schlink, 2018), mean bias error (Ueberham and Schlink, 2018), Lin's concordance correlation coefficients (Dons et al. 2017) and Taylor diagrams (Ueberham and Schlink, 2018). Ueberham and Schlink (2018) utilize mean absolute error (MAE) to gauge the precision of the sensors, and mean bias error (MBE) to determine the accuracy of the sensors (Ueberham and Schlink 2018). Accuracy is defined as a measurement that is very close to the correct value, while precision is a measurement exactness. Ueberham and Schlink (2018) tested two personal temperature sensors in their study (Testo testostor 171 and TSI Q-Trak 7565) illustrating that their methodology is well suited for validating personal ambient temperature sensors. The study found moderate to high agreement between devices outdoors (IA and r = 0.5-0.97). Bland Altman plots illustrate that the lowest levels of precision occur during higher temperatures, demonstrating a limitation of these sensors when they are potentially needed most. To date, no studies utilize this set of methodology on the validation of HOBOs, Kestrel DROP D3 Fires or iButtons.

This study aims to fill these literature gaps by identifying which personal monitors and placements are the most valid and replicable for more extensive research that examines the best ways to map and predict personal temperature exposure. To accomplish this objective, we evaluate the performance of multiple low-cost wearable sensors for measuring ambient temperature in a (1) field exposure study by varying the placement on human subjects and in a (2) field calibration study by co-locating sensors with fixed site weather station monitors. A secondary aim involved investigating consensus between validation metrics that can be used in future sensor comparison studies. Results will provide recommendations to inform future personal heat studies.

## 2. Data and Methods

### *2.1. Participant Recruitment*

Twenty-one participants from Appalachian State University (ASU) and seventeen participants from Mississippi State University (MSU) consented and were enrolled to participate in this study. Participants were recruited via recruitment emails, fliers, and an online ASU bulletins. The ASU participants signed up for one of two 5-day data collection periods from July 23rd through August 3rd (occurring Monday-Friday), while MSU participants signed up for a study period of September 21st through September 28th, or September 28th through October 5th. Devices were activated the night before and delivered to participants on the first day of their assigned study period. At the time of device drop off, all equipment was explained and baseline surveys, activity logs, and written equipment instructions were distributed and consent forms were signed. Researchers made sure to stress the importance of keeping devices in their designated placements and to record any changes in the daily activity logs. Participants did not receive any monetary compensation for their participation. This project received human subject's approval from the institutional review board (IRB) at ASU (IRB #18-0325) and MSU (IRB #18-383).

### *2.2. Data Collection*

#### *2.2.1. Phase I: Participant Data Collection*

The data for this study were collected in two phases. In Phase I, personal environmental exposure data were collected for a total of thirty-eight participants from Appalachian State University (ASU) in Boone, North Carolina ($n = 21$) and Mississippi State University in Starkville, Mississippi ($n = 17$). Table 1 highlights the demographics of the participant portion of the study. Overall participants at ASU were older and less diverse than MSU participants, who were predominantly undergraduate students.

The two locations are climatologically diverse rural towns with populations less than 26,000 (US Census 2010). Boone, NC is located in the Appalachian Mountains with an elevation of 1016 meters, while Starkville, MS is located at 102 meters above sea level (Figure 1). The average daily maximum temperature during the study period (July-September), using 30-year climate normals from 1981 through 2010, is 89.7°F/32.1°C in Starkville, MS and 76.3°F/24.6°C in Boone, NC. The average temperature during was 20.01°C during the Boone, NC study period and 23.98°C during the Starkville, MS study. The first week (July 23rd-27th) of the Boone, NC study period was predominantly sunny, while the second week of the study period (July 30th-August 3rd) was cloudy with heavy rain for four out of five days. There was no major precipitation events during the Starkville, MS study period (September 21st-October 5th).

Participants wore up to six devices including 1.) A Garmin Vivoactive HR watch 2.) Two iButton Thermochrons 3.) An iButton Hygrochron 4.) A Kestrel DROP D3 Fire and 5.) A HOBO pendant. Participants at ASU were equipped with the highest number of devices. HOBO devices (Pendant Temperature/Light Data Logger UA-002-64) were worn on the top of the shoe, and the Kestrel and Thermochron devices were attached to a backpack or purse with a carabineer, which served as a control for sensor readings on the body. An additional Thermochron was worn on the collar with a safety pin, a common practice in other personal monitoring studies (e.g., Sugg et al. 2018, Runkle et al. 2019, Sugg et al. 2019).

MSU participants wore the following devices: a HOBO (on the shoe), a Kestrel (on a carabiner), a Hygrochron (on shirt collar), and a Thermochron (on shirt collar), however, no participants were equipped with multiple Thermochrons due to a resource shortage (Figure 2). All devices at both universities were set to record data at five-minute intervals. Different sensor models are equipped with varying capabilities of monitoring (Table 2). All sensors utilized in this study record temperature data. Participants at MSU and ASU completed activity logs at thirty-minute intervals and baseline surveys that outline essential socioeconomic data. These activity logs provide individual context to the data and temperature anomalies.

*2.2.2. Activity Logs*

To quantify differences in microenvironments (e.g., indoor, outdoor, in-transit) and activity type (e.g., exercise, stationary), participants labeled their activities as low intensity, moderate intensity, or high intensity in their daily activity logs. Participants also indicated if they were outside, inside, or in-transit throughout their activity logs. Activity log information was collected from 7:00 AM to 7:00 PM at a thirty-minute time scale. Responses to activity logs were numerically coded in excel and microenvironments/activity levels were stratified for statistical analysis.

*2.2.3. Phase II: Weather Station Data Collection*

In Phase II, at least two of each type of personal sensors (HOBOs, iButton Thermochrons, iButton Hydrochrons, and Kestrel Drop D3FWs) were attached to a weather station in Starkville, MS, and Boone, NC during the same time periods as the corresponding participant studies (Figure 3). One device was attached to the weather station facing up in direct sun exposure and one was attached hanging down at indirect sun exposure. The different placements test the differences in measurements based on direct or indirect sun exposure. Data collection from this test is used to examine the accuracy of personal monitoring devices in relation to meteorological data from the two weather stations that are located in climatologically and geographically diverse locations (Starkville, MS, and Boone, NC), and the impact of direct solar radiation on sensor accuracy. The Boone, NC weather station records measurements at hour and one-minute intervals, while the Starkville, MS weather station only records data at an hourly rate. The weather stations in Starkville, MS and Boone, NC utilize thermistors that are housed in aspirated radiation shields.

### *2.3. Statistical Analysis*

#### *2.3.1. Comparison between Weather Station and Sensor Measurements*

Sensors in Boone, NC, and Starkville, MS, were set to record temperatures every five-minutes. Weather station data from Boone, NC was recorded at a one-minute scale and averaged to a five-minute scale. Sensor data from Starkville, MS was averaged to the nearest hour to match the weather station data that was recorded at an hourly scale.

To initially evaluate the accuracy of personal wearable sensors in approximating weather station temperature measurements, Pearson's product moment correlation coefficients were calculated ($\alpha = 0.05$) (Stahl et al. 2016). The mean absolute error (MAE) was also calculated to gauge the precision of the sensors, while mean bias error (MBE) and root mean square error (RMSE) were used to determine the accuracy of the sensors (Ueberham and Schlink 2018). Both accuracy and precision are important components in the assessment of which sensor is performing the best, and ideally, the best sensor should be both highly accurate and highly precise (Ueberham and Schlink 2018). RMSE measures the average magnitude of error, weighting larger errors at a higher rate than MBE or MAE.

Following the methodology of Stahl et al. (2016) and Ueberham and Schlink (2018), a Bland Altman plot was used to quantify the differences between sensors and the weather station. Bland Altman analysis allowed for both the appraisal of disagreements and the determination of when the disagreements occurred. For the sensors to be considered accurate, its line of equality within the Bland Altman plot must fall within the 95% confidence interval of the mean difference (Myles and Cui 2007). Bland Altman analysis and notched box plots were also utilized to identify the most significant differences within the data, both between sensors and by determining what times of day the sensors differ from each other and the weather station data the most.

Lin's concordance correlation coefficient (*Pc*) is calculated to determine the degree of variation between the various types of sensors and the weather station data (Dons et al. 2017). Concordance correlation coefficients test how well a new measurement (personal temperature sensors) reproduce a gold standard (weather station data) (Lin 1989).

### *2.3.2. Participant Personal Temperature Exposure Data*

Participant data was first examined using basic summary statistics, t-tests, and boxplots. These tests are used to determine which sensors observed temperatures warmer than others and to identify data outliers. Bland Altman analysis is then utilized to determine the average temperatures where the largest deviations between sensor temperature readings occur. Significant temperature differences occur between the various types of sensors and sensor locations when being worn by participants. Finally, a linear mixed effects model for repeated measures with a random intercept for each participant is utilized to compare the participant sensor observations to the weather station readings at the same time. Linear mixed effect models are used to determine the degree of difference between the weather station and the participant sensor readings.

## 3. Results

### *3.1. Participant Study*

Summary statistics illustrate that the Kestrel device has the lowest mean temperature for both the ASU and MSU participant studies. At ASU, the highest maximum temperature (57.2 °C) was recorded by the iButton 2, which was attached to a participant's collar while the participant was indoors preparing a research lab. The highest temperature at MSU (48.9 °C) was recorded by the HOBO device (Table 3). All T-tests between different types of sensors and placements were significant at both locations, even after adjusting for a bonferroni correction (0.05/sample size).

Bland Altman plots show that iButton 1 (attached to bag) and iButton 2 (on collar) have the largest absolute mean difference at ASU (3.09), while the Hygrochron and HOBO have the largest absolute mean difference at MSU (1.76). The devices that have the lowest absolute mean difference are the iButton 1 (attached to bag) and Kestrel at ASU (0.18) and the iButton 2 (on collar) and HOBO at MSU (0.50). The largest measured temperature differences between sensor types occur between $32 - 49$ °C at ASU and $24 - 29$ °C at MSU (Figure 4).

#### *3.1.1. ASU Participant Study: Activity Log Analysis*

Summary statistics show that there is no type of environment (i.e. indoor, outdoor, in-transit) that has the highest mean or max temperature readings. In-transit has the highest mean temperature for Hygrochrons, HOBO, and Kestrel devices. iButton 1 devices (attached to a bag) have the highest mean temperatures outdoors, and iButton 2 (attached to the collar) has the highest mean and max temperatures indoors (Table 4). The highest overall correlation coefficients between different sensors types occurred in the category of "in-transit"

($R = 0.513 - 0.796, p - value < 0.001$). Indoor and outdoor microclimates displayed average correlation values below $R = 0.5$. There are no clear patterns as to what microenvironments experience the largest degrees of error between different types of sensors.

Summary statistics illustrate that high intensity activity levels have the lowest mean and maximum temperatures for all devices except the HOBO maximum temperature. iButton 1 (attached to bag) has the highest mean and max temperatures during moderate activity ($mean = 23.25\ °C, max = 53.00\ °C$), while iButton 2 (attached to collar) has the highest maximum and mean temperatures during low activity levels. The Kestrel, HOBO, and Hygrochron have highest max temperatures during low activity levels, and the highest mean temperatures during moderate activity levels (Table 5).

To identify if activity type influenced sensor type and location temperature readings, we examined activity log data classifications of activity intensity. While the least extreme maximum and mean temperatures are recorded during high activity levels, the sensors have the lowest correlations during this category of activity with an average correlation of $R = 0.366$ ($R = 0.056 - 0.510$). The lowest correlation ($R = 0.056, p = 0.202$) in this category is between the Kestrel and iButton 1 (attached to Kestrel on a bag). Correlation coefficients are highest for low and moderate intensity activities ranging between $R = 0.26$ to $R = 0.829$, with an average correlation coefficient of $R = 0.520$ for low activity and $R = 0.449$ for moderate activity. Box Plots illustrate that high activity intensity has the most variability between days throughout the study, while mean temperature during low activity levels are consistent across the study period (Figure 5).

No patterns were found in the deviations between devices based on location (i.e. outdoor, indoor, in-transit). Results show that high intensity activities have the lowest mean and maximum temperatures band the lowest correlations between sensors. This could be due to the majority of high intensity level activities occurring in air conditioned gyms. Although past studies have found differences in sensor accuracy based on outdoor and indoor microclimates, our study found inconclusive results for these sensor types (Ueberham and Schlink 2018).

Table 6 illustrates the locations of participants during the maximum temperatures reached by each device. The majority of the temperature spikes recorded by iButton 1, iButton 2, and Kestrel devices match to the participant A7 during a variety of activities and locations. These spikes in activities were predominantly during moderate activity walking and low activity teaching and lab preparation. Participant A7 was not equipped with a Hygrochron, explaining why the participant was not listed for any spikes in this category. The participant was equipped with a HOBO and it is unclear why the device did not react with the same spikes in temperatures as the other devices worn by participant A7. These spikes throughout various locations and activity types show that there is not a specific environment that generates the differences in sensor readings.

### *3.1. Participant Sensors and Weather Station Data Analysis*

#### *3.1.1. Comparisons of Stationary Sensor Measures with Weather Station*

Wearable sensors have similar mean temperatures to weather stations at both locations, however, sensors have much higher maximum temperatures compared to the weather stations for both sites (Table 7). The HOBO sensor reported the highest peaks in both Boone, NC (46.72 °C) and Starkville, MS (46.08 °C). All of the sensors in Boone, NC observed lower minimum temperatures

($12.5\,°C - 13.61\,°C$) than the weather station (15.16 °C), illustrating that the devices deviate from the weather station temperatures at multiple times of the daytime and nighttime hours (Figure 6).

Testing the level of similarity between the sensors attached to the weather station and the in-situ weather station data in Boone, NC shows that all the devices, except the iButton Thermochron facing upwards ($R = 0.06, p - value < 0.01$) were highly correlated. Although most correlations were highly significant, the correlation coefficients were low, ranging from 0.21 to 0.39. Stationary Kestrel measures were highly correlated with the weather stations, with values of 0.39 ($p - value < 0.001$) for the device hanging down (i.e., away from direct solar radiation) from the weather station and 0.25 ($p - value < 0.001$) for the device facing upwards (i.e., in direct solar radiation) attached to the weather station. While the correlations between the weather station and the sensors were low, the correlations between different types of sensors were much higher. The correlations between different types of sensors ranged from 0.81 to 0.98 (Figure 7).

The Starkville, MS weather station and sensor correlations were significantly higher than those in Boone, NC ($p - value < 0.01$). The lowest correlation coefficient between a sensor and the Starkville, MS weather station is the HOBO facing up (0.88), while the highest are the Hygrochron (0.98) and Kestrel (0.98) that were protected by solar radiation shields (Figure 8). The results are congruent with results in Boone, NC, illustrating that the devices that are attached to the top of the weather station in direct sunlight have lower correlations in relation to the weather station. Both study areas illustrate the HOBO devices have the lowest correlation values while the Kestrel and Hygrochron have the highest correlation values.

The HOBO devices (facing upwards and downwards) have the highest MAE, MBE, and RMSE values at both study sites. HOBO devices also have the lowest $Pc$ values at both Boone, NC, and Starkville, MS, with all values rating as "poor" by both Altman (1991) and McBride (2005) standards. The Kestrel facing down and all Starkville, MS, devices protected by radiation shields have the highest $Pc$ values. The highest $Pc$ value in the Boone, NC, study is 0.36 (poor) while the highest $Pc$ values in Starkville, MS are $0.93 - 0.98$ (excellent) (Altman 1991). While the Starkville, MS, and Boone, NC, values differ drastically, the same devices display the highest and lowest accuracies at both sites. During both studies Kestrel devices perform the best and HOBO devices perform the worst in comparison to the weather station data (Table 8).

*3.1.2. Comparison between Sensors Worn by Participants and Weather Station Data*

Mixed effects models show that weather station temperature measurements are significantly associated with personal heat exposure measured by all sensor types and placements. Across all participants at ASU, model results predict that personal sensor temperature readings increased by $0.20 - 0.27$ for every 1 °C increase in temperature recorded by the weather station. The model results predicted that personal sensor temperature readings increases by $0.28 - 0.50$ for every 1 °C increase in temperature recorded by the weather station at MSU. The marginal $R^2$ values ($0.040 - 0.059$) for all devices at ASU are low, showing that the weather station data does not explain much of the deviations in PAT values (Table 9). Marginal $R^2$ values are much higher at MSU ranging from $0.044 - 0.210$ (Table 10).

## 4. Discussion

The first objective of this study was to quantify the differences in temperature measurements between different types of personal wearable sensors and sensor placements when worn by participants. The second objective of this study was to examine how closely personal ambient temperatures match weather station data in order to make recommendations on the best type of sensor for participant-based monitoring studies. To the author's knowledge, this is the first study that compares HOBO, iButton Thermochron, iButton Hygrochron, and Kestrel sensor data during participant based studies and relative to fixed weather station data. This study adds to the body of knowledge that works to validate personal temperature exposure sensors and draw comparisons between sensors and weather station data (Bernhard et al. 2015, Kuras et al. 2015, Sugg et al. 2018, and Ueberham and Schlink, 2018). We found significant differences between different types of sensors and sensor placements that underpin the need to quantify differences between devices before results from studies utilizing different devices or device placements can be compared.

### *4.1. Sensor and In-Situ Weather Station Data*

Comparisons between sensor temperature readings and weather station readings show that sensors run predominantly warmer than the weather station during daytime hours, and colder than the weather station through the nighttime hours. While deviations in temperature between the sensors and weather station occur during the day and night, deviations are much greater during the daytime spikes than the nighttime hours. The HOBO has the highest overall recorded temperatures and deviations from the weather station at both locations.

These differences between temperature readings based on sensor placements show the sensitivity of the devices to direct solar radiation. To the author's knowledge, this is the first study that attached these sensor types directly to a weather station to monitor when deviations occur.

The Kestrel devices performed the best in relation to the weather station, with the lowest MAE, MBE and RMSE values and highest *Pc* value. The HOBO demonstrated the most variation from the weather station, with the highest degrees of error across all tests and a *Pc* values that rank below excellent at MSU and poor at ASU according to Altman (1991) standards and poor at both universities by McBride (2005) standards. Although both studies display the same trends, correlation coefficients and Pc values were significantly higher during the MSU study than the ASU study. This could be caused by a variety of factors including time of year, altitude, the direction that the sensors were oriented (i.e. north, south, east, west), weather conditions, and land use variability around the different weather stations.

### *4.2. Participant Study*

The participant portion of the study shows larger differences between temperature sensor readings when worn by participants. This variability is caused by the variety of diverse microclimates and activity types that participants experience throughout their daily lives. Bland Altman plots show that the greatest level of difference between sensors occurs from 32-49 °C at ASU and 24-29 °C MSU. This disagreement could be due in part by the weather differences between MSU and ASU. The study period at ASU included multiple precipitation events, while MSU did not experience any precipitation during the study period.

This disagreement illustrates a need for more validation studies that examine when differences between different sensor types are most pronounced.

Results from the participant study at ASU demonstrate the importance of both device type and device placements. iButton 1 (located on a bag) and iButton 2 (on the participant's collar) only have a correlation coefficient of 0.515. This placement of the device was the only differential factor, illustrating the importance of device location on temperature readings. The iButton Thermochrons and Hygrochrons located on participants collars ran warmer and had a larger standard deviation than any other placement. The devices with the lowest variability and mean temperatures were the iButton Thermochron and Kestrel devices attached to participants bags. This contradicts the findings of Dumas et al. (2016). Dumas et al. (2016) found consistent temperature readings between temperature placements on the shoe, collar, and waist. This difference may be due to differences in sample size between Dumas et al. (2016) ($n = 2$) and our study ($n = 21$), along with the longer length of our sampling periods.

Participants were also more willing to carry devices on bags or keys then attached to shoes or clothing. Similar to the findings of Sugg et al. (2019), the HOBO devices were more burdensome to participants than any other device, which likely caused more participants to fail to wear the HOBO devices correctly. This sensor was especially problematic due to the number of participants that did not frequently wear shoes that had shoelaces that the HOBO could attach to. Multiple participants expressed frustration that they were limited to shoes with shoelaces for the week, especially participants that worked in a professional work environment with a business dress code. Devices that attach to keys or bags that are carried with the participant were most favored by participants in professional work settings due to their more discrete nature and lack of impact on clothing and footwear options.

### *4.2. Participant Sensors and Weather Station Data*

Results are comparable to other PAT and weather station model predictions. Bernhard et al. (2015) found that PAT sensors predict an average increase in temperature of 0.37°C (95% CI 0.35, 0.39) for each 1°C increase in temperature was recorded at the weather station (Bernhard et al. 2015). Our results found that sensors readings increased by $0.20 - 0.27$ at ASU and $0.28 - 0.50$ at MSU for every 1 °C increase in temperature recorded by the weather station. This trend is comparable for all studies that completed similar analysis (e.g. Kuras et al. 2015 and Sugg et al. 2019). This suggests that the weather station temperatures are overall warmer than the temperatures experienced by participants in their daily lives. This is likely partly due to the time that participants spend indoors and the temperature differences across microclimates that scarce weather station data cannot account for.

The marginal $R^2$ values $(0.040 - 0.059)$ for all devices at ASU are lower than the values for MSU $(0.044 - 0.210)$. This is likely due to the diverse topography around the Boone, NC, weather station. Participants from Boone, NC, also traveled further distances around the study area than MSU participants and traveled through a greater variety of geographically diverse areas throughout their daily routines.

### *4.3. Limitations*

The inherent limitations to this study exist in the realm of human error. Personal sensor studies rely on voluntary compliance and adherence to the study methods. Adherence to methodology is especially important to this study when considering the placement of sensors and how that placement impacts temperature readings. It is not possible to ensure that participants wore their sensors correctly at all times. To account for this, we explained the importance of adherence to study methodology to participants when devices were distributed and asked participants to mark any changes in device locations on their activity logs. The quantification of human error in this study is nearly impossible and can only be addressed in the form of qualitative analysis through time-activity logs. One way we attempted to limit this potential bias is through the use of a large sample sizes from a repeated measures research design.

## 5. Conclusion

In agreement with Sugg et al. (2019), our study found significant differences between sensor types and placements. Our study found that the Kestrel is the most accurate in relation to meteorological data. The major limitation to the Kestrel is the size of the device. The Kestrel device is the largest out of all devices in this study and therefore is best attached to a bag or keys that the participant carries with them. While the placement of devices on a bag was appreciated by participants at work, it is important for researchers to stress the importance of keeping the devices with them at all times and to not put the devices inside of a bag at any time. The iButton and Kestrel attached to each other were highly correlated, suggesting that placement of devices is a driving factor in the differences in temperature readings by personal sensors. This study illustrates the need for unity in the placement of devices by different studies and highlights a need for continued research on what device is not only the most accurate, but most convenient for participants to wear. Future work should also use sensor data to examine other heat stress metrics like heat index, wet bulb globe, and physiological equivalent temperature.

**Acknowledgments**

The authors thank all of the participants that contributed to the data collection at Appalachian State University and Mississippi State University. The authors would also like to thank Scott Stevens for his generous contributions to data merging.

## References


Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman and Hall. Retrieved from https://books.google.com/books?hl=en&lr=&id=v-walRnRxWQC&oi=fnd&pg=PR11&dq=Altman+DG+(1991)+Practical+statistics+for+medical+research.+London:+Chapman+and+Hall.&ots=SxZUAcyl_j&sig=X7tIsjPj2fXCOQ4mvh7zowXVJFE#v=onepage&q=Altman DG (1991) Practical

Basu, R., & Samet, J. M. (2002). An exposure assessment study of ambient heat exposure in an elderly population in Baltimore, Maryland. *Environmental Health Perspectives*, *110*(12), 1219–1224. https://doi.org/10.1289/ehp.021101219

Bernard, S. M., & McGeehin, M. A. (2004). Municipal Heat Wave Response Plans. *American Journal of Public Health*, *94*(9), 1520–1522. https://doi.org/10.2105/AJPH.94.9.1520

Bernhard, M. C., Kent, S. T., Sloan, M. E., Evans, M. B., McClure, L. A., & Gohlke, J. M. (2015). Measuring personal heat exposure in an urban and rural environment. *Environmental Research*, *137*, 410–418. https://doi.org/10.1016/J.ENVRES.2014.11.002

Chan, N., Stacey, M., Smith, A., Ebi, K., & Wilson, T. (2001). An empirical mechanistic framework for heat-related illness. *Climate Research*, *16*, 133–143. https://doi.org/10.3354/cr016133

Chen, L., & Ng, E. (2012). Outdoor thermal comfort and outdoor activities: A review of research in the past decade. *Cities*, *29*(2), 118–125. https://doi.org/10.1016/J.CITIES.2011.08.006

Dong, W., Liu, Z., Zhang, L., Tang, Q., Liao, H., Li, X. (2014). Assessing Heat Health Risk for Sustainability in Beijing's Urban Heat Island. *Sustainability*, *6*(10), 7334–7357. https://doi.org/10.3390/su6107334

Dons, E., M. Laeremans, J. P. Orjuela, I. Avila-Palencia, G. Carrasco-Turigas, T. Cole-Hunter, E. Anaya-Boig, A. Standaert, P. D. Boever, T. Nawrot, T. Götschi, A. D. Nazelle, M. Nieuwenhuijsen, and L. I. Panis. (2017). Wearable Sensors for Personal Monitoring and Estimation of Inhaled Traffic-Related Air Pollution: Evaluation of Methods. *Environmental Science & Technology*, *51*(3), 1859–1867. https://doi.org/10.1021/acs.est.6b05782

Dumas JS, Jagger MA, Kintziger KW. (2016). Where to wear ibuttons: Individual level temperature and humidity observations for public health surveillance [Abstract]. In: American Meteorological Society Annual Meeting. https://ams.confex.com/ams/96Annual/webprogram/Paper280434.html [accessed 15 December 2018].

Fuhrmann, C. M., Sugg, M. M., Konrad, C. E., & Waller, A. (2016). Impact of Extreme Heat Events on Emergency Department Visits in North Carolina (2007–2011). *Journal of Community Health*, *41*(1), 146–156. https://doi.org/10.1007/s10900-015-0080-7

Gosling, S. N., Lowe, J. A., McGregor, G. R., Pelling, M., & Malamud, B. D. (2009). Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change*, *92*(3–4), 299–341. https://doi.org/10.1007/s10584-008-9441-x

Hondula, D. M., Davis, R. E., Leisten, M. J., Saha, M. V, Veazey, L. M., & Wegner, C. R. (2012). Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: a case-series analysis. *Environmental Health*, *11*(1), 16. https://doi.org/10.1186/1476-069X-11-16

Hondula, D. M., Davis, R. E., Saha, M. V., Wegner, C. R., & Veazey, L. M. (2015). Geographic dimensions of heat-related mortality in seven U.S. cities. *Environmental Research*, *138*, 439–452. https://doi.org/10.1016/j.envres.2015.02.033

Kuras, E. R., Hondula, D. M., & Brown-Saracino, J. (2015). Heterogeneity in individually experienced temperatures (IETs) within an urban neighborhood: insights from a new approach to measuring heat exposure. *International journal of biometeorology*, *59*(10), 1363-1372.

Kuras, E. R., Richardson, M. B., Calkins, M. M., Ebi, K. L., Hess, J. J., Kintziger, K. W. Jagger, M. A., Middel, A., Scott, A. A., Spector, J.T., Uejio C.K., Vanos, J.K., Zaitchik, B. F., Gohike, J. M. and Hondula, D. M. (2017). Opportunities and Challenges for Personal Heat Exposure Research. *Environmental Health Perspectives*, *125*(8), 085001. https://doi.org/10.1289/EHP556

Lewis, A., & Edwards, P. (2016). Validate personal air-pollution sensors. *Nature*, *535*(7610), 29–31. https://doi.org/10.1038/535029a

Lin, L. I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, *45*(1), 255–268. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2720055

McBride, R. B. (2005). A proposal for strength-of-agreement criteria for Lins Concordance Correlation Coefficient. Retrieved from https://www.scienceopen.com/document?vid=e3cfffed-a777-439b-b244-b4acbaf7f0c7

McGeehin, M. A., & Mirabelli, M. (2001). The potential impacts of climate variability and change on temperature-related morbidity and mortality in the United States. *Environmental Health Perspectives*, *109 Suppl*(Suppl 2), 185–189. https://doi.org/10.1289/ehp.109-1240665

Myles, P. S., & Cui, J. (2007). I. Using the Bland–Altman method to measure agreement with repeated measures. *British Journal of Anaesthesia*, *99*(3), 309–311. https://doi.org/10.1093/bja/aem214

Noe, R. S., Jin, J. O., & Wolkin, A. F. (2012). Exposure to Natural Cold and Heat: Hypothermia and Hyperthermia Medicare Claims, United States, 2004–2005. *American Journal of Public Health*, *102*(4), e11–e18. https://doi.org/10.2105/AJPH.2011.300557

Rhea, S., Ising, A., Fleischauer, A. T., Deyneka, L., Vaughan-Batten, H., & Waller, A. (2012). Using Near Real-Time Morbidity Data to Identify Heat-Related Illness Prevention Strategies in North Carolina. *Journal of Community Health*, *37*(2), 495–500. https://doi.org/10.1007/s10900-011-9469-0

Runkle, J., Sugg, M., Boase, D., Galvin, S. L., & C. Coulson, C. (2019). Use of wearable sensors for pregnancy health and environmental monitoring: Descriptive findings from the perspective of patients and providers. *DIGITAL HEALTH*, *5*, 205520761982822. https://doi.org/10.1177/2055207619828220

Sheridan, S. C., & Kalkstein, A. J. (2010). Seasonal variability in heat-related mortality across the United States. *Natural Hazards*, *55*(2), 291–305. https://doi.org/10.1007/s11069-010-9526-5

Stahl, S. E., An, H.-S., Dinkel, D. M., Noble, J. M., & Lee, J.-M. (2016). How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport & Exercise Medicine*, *2*(1), e000106. https://doi.org/10.1136/bmjsem-2015-000106

Sugg, M. M., Fuhrmann, C. M., & Runkle, J. D. (2018). Temporal and spatial variation in personal ambient temperatures for outdoor working populations in the southeastern USA. *International Journal of Biometeorology*, *62*(8), 1521–1534. https://doi.org/10.1007/s00484-018-1553-z

Sugg, M. M., Stevens, S. S., & Runkle, J. D. (2019). Estimating Personal Ambient Temperature in Moderately Cold Environments for Occupationally Exposed Populations. *Environmental Research* (In Review).

Ueberham, M., & Schlink, U. (2018). Wearable sensors for multifactorial personal exposure measurements – A ranking study. *Environment International*, *121*, 130–138. https://doi.org/10.1016/J.ENVINT.2018.08.057

Williams, S., Nitschke, M., Sullivan, T., Tucker, G. R., Weinstein, P., Pisaniello, D. L., Bi, P. (2012). Heat and health in Adelaide, South Australia: Assessment of heat thresholds and temperature relationships. *Science of The Total Environment*, *414*, 126–133. https://doi.org/10.1016/J.SCITOTENV.2011.11.038

Zhou, W., Ji, S., Chen, T.-H., Hou, Y., & Zhang, K. (2014). The 2011 heat wave in Greater Houston: Effects of land use on temperature. *Environmental Research*, *135*, 81–87. https://doi.org/10.1016/J.ENVRES.2014.08.025

**Table 1.** Participant demographics for the MSU and ASU study sites.

| | ASU | MSU |
|---|---|---|
| n | 21 | 17 |
| Age (mean / sd) | 32.6 (13) | 21.5 (3.0) |
| Sex (n / %) | | |
| Female | 11 (52%) | 8 (47%) |
| Male | 9 (42%) | 9 (53%) |
| Height (mean / sd) | 68.1 (5.4) | 66.1 (3.9) |
| Weight (mean / sd) | 159.6 (41.7) | 164.7 (32.4) |
| Race (n / %) | | |
| Asian | 1 (4.8%) | 0 |
| Caucasian | 20 (95.2%) | 15 (88.2%) |
| Biracial | 0 | 1 (5.9%) |
| Other | 0 | 1 (5.9%) |
| Education (n / %) | | |
| Some high school | 0 | 1 (5.9%) |
| High school diploma | 1 (4.8%) | 13 (76.5%) |
| Associate's degree | 0 | 1 (5.9%) |
| Bachelor's degree | 10 (47.6%) | 1 (5.9%) |
| Graduate or professional degree | 10 (47.6%) | 1 (5.9%) |

**Table 2.** Temperature Exposure Assessment Manufacturer Guidelines

| Sensor Type | Data Collected | Temperature Accuracy | Manufacturer Notes |
|---|---|---|---|
| iButton Thermochron | Temperature | $\pm 1°C$ / $\pm 1.80°F$ | N/A |
| iButton Hygrochron | Temperature, Relative Humidity | $\pm 0.5°C$ / $\pm 0.9°F$ | N/A |
| HOBO | Temperature, Light Intensity | $\pm 0.53°C$ / $\pm 0.95°F$ | N/A |
| Kestrel DROP D3 Fire | Temperature, Relative Humidity, Heat Stress Index, Dew Point, Wet Bulb Temperature, Station Pressure | $\pm 0.50°C$ / $\pm 0.9°F$ | "For greatest accuracy, avoid direct sunlight on the temperature sensor and prolonged sunlight exposure to the unit in low airflow conditions." |

**Table 3.** Summary statistics for sensor temperature readings during the ASU and MSU participant study.

| | ASU | | | | | MSU | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sensor Type | Mean (℃) | Max (℃) | Min (℃) | Standard Deviation | *n* | Mean (℃) | Max (℃) | Min (℃) | Standard Deviation | *n* |
| **Kestrel** | 22.51 | 39.11 | 14.78 | 2.59 | 20,196 | 22.76 | 40.39 | 18.11 | 2.26 | 21,356 |
| **iButton 1 (On bag)** | 22.52 | 53.00 | 15.00 | 2.42 | 24,986 | NA | NA | NA | NA | NA |
| **iButton 2 (On collar)** | 25.15 | 57.50 | 15.50 | 3.29 | 21,345 | 23.80 | 37.00 | 16.00 | 3.21 | 10,240 |
| **HOBO** | 23.31 | 43.83 | 15.94 | 2.70 | 23,191 | 23.18 | 42.89 | 16.44 | 2.78 | 33,188 |
| **Hygrochron** | 24.68 | 38.06 | 17.56 | 3.53 | 7,348 | 23.95 | 41.06 | 17.11 | 3.40 | 23,928 |

**Table 4.** Summary statistics for three types of microenvironments (outdoor, indoor, in-transit) experienced during the participant study at ASU.

| | iButton1 | | | iButton2 | | | Kestrel | | | HOBO | | | Hygro | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean (℃ ) | Max (℃ ) | Std. Dev. | Mean (℃ ) | Max (℃ ) | Std. Dev. | Mean (℃ ) | Max (℃ ) | Std. Dev. | Mean (℃ ) | Max (℃ ) | Std. Dev. | Mean (℃ ) | Max (℃ ) | Std. Dev. |
| **Outdoor** | 23.64 | 41.0 | 3.35 | 25.26 | 43.50 | 3.37 | 23.43 | 40.67 | 3.15 | 24.69 | 43.83 | 3.38 | 26.58 | 33.17 | 2.97 |
| **Indoor** | 22.83 | 48.0 | 3.35 | 25.61 | 57.50 | 3.37 | 22.86 | 49.11 | 3.15 | 24.03 | 38.83 | 2.71 | 26.41 | 35.06 | 2.97 |
| **In Transit** | 23.55 | 53.0 | 3.11 | 25.57 | 40.00 | 3.27 | 23.52 | 38.28 | 2.92 | 24.70 | 43.72 | 3.71 | 27.08 | 38.06 | 3.54 |

**Table 5.** Summary statistics for different activity intensity levels (low, moderate, and high) during the participant study at ASU.

| | iButton1 | | | iButton2 | | | Kestrel | | | HOBO | | | Hygro | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Intensity Level | Mean (℃ ) | Max (℃ ) | Std. Dev. | Mean /℃ ) | Max (℃ ) | Std. Dev. | Mean (℃ ) | Max (℃ ) | Std. Dev. | Mean (℃ ) | Max (℃ ) | Std. Dev. | Mean (℃ ) | Max (℃ ) | Std. Dev. |
| **Low** | 23.00 | 48.00 | 2.18 | 25.72 | 57.5 | 3.26 | 23.02 | 49.11 | 3.12 | 22.51 | 43.83 | 2.97 | 26.57 | 38.06 | 3.29 |
| **Moderate** | 23.25 | 53.00 | 3.08 | 25.21 | 39.61 | 2.79 | 23.16 | 37.78 | 2.78 | 25.46 | 38.39 | 3.03 | 26.83 | 33.17 | 2.83 |
| **High** | 22.57 | 37.00 | 2.41 | 24.12 | 34.5 | 2.95 | 22.31 | 34.5 | 1.71 | 23.76 | 39.39 | 2.60 | 24.15 | 30.11 | 2.29 |

**Table 6.** Quantitative analysis of participant activities during the maximum temperatures reached by each device during the ASU participant study

| Sensor Type | Participant ID | Date | Time | Max Temperature (℃) | Location | Activity | Description |
|---|---|---|---|---|---|---|---|
| iButton 1 | A7 | 7/25/2018 | 3:30 PM | 53.0 | In-transit | Moderate | Walking/Driving |
| | A7 | 7/25/2018 | 3:50 PM | 48.0 | Indoor | Low | Lab preparation |
| | A7 | 7/27/2018 | 1:05 PM | 44.5 | Indoor | Low | Teaching |
| | A7 | 7/27/2018 | 12:10 PM | 41.0 | Outdoor | Moderate | Walking |
| | A3 | 7/26/2018 | 5:50 PM | 38.0 | Outdoor | Low | Sat outside |
| iButton 2 | A7 | 7/25/2018 | 4:25 PM | 57.5 | Indoor | Low | Lab preparation |
| | A7 | 7/27/2018 | 1:00 PM | 44.0 | Indoor | Low | Teaching |
| | A7 | 7/27/2018 | 12:10 PM | 41.0 | Outdoor | Moderate | Walking |
| | A7 | 7/24/2018 | 4:00 PM | 40.5 | Indoor | Low | Lab preparation |
| | A2 | 7/26/2018 | 5:05 PM | 40.0 | In-transit | Moderate | Walking/Driving |
| Kestrel | A7 | 7/25/2018 | 4:40 PM | 49.1 | Indoor | Low | Lab preparation |
| | A7 | 7/27/2018 | 2:00 PM | 47.7 | Indoor | Low | Teaching |
| | A7 | 7/25/2018 | 4:50 PM | 47.5 | Indoor | Low | Working in lab |
| | A7 | 7/27/2018 | 12:15 PM | 40.7 | Outdoor | Moderate | Walking |
| | A7 | 7/27/2018 | 3:20 PM | 38.3 | Outdoor | Moderate | Walking/Driving |
| HOBO | A8 | 7/27/2018 | 2:45 PM | 43.8 | Outdoor | Low | Got tire replaced |
| | A8 | 7/27/2018 | 3:05 PM | 43.7 | In-transit | Low | Driving |
| | A4 | 7/26/2018 | 4:10 PM | 39.39 | Outdoor | Intense | Running |
| | A3 | 7/26/2018 | 2:45 PM | 38.8 | Indoor | Low | At dentist’s office |
| | A6 | 7/26/2018 | 6:45 PM | 38.4 | Outdoor | Moderate | Walking |
| Hygrochron | A1 | 7/25/2018 | 11:40 AM | 38.1 | In-transit | Low | Ate lunch in car |
| | A1 | 7/26/2018 | 2:50 PM | 35.6 | In-transit | Low | Driving |
| | A1 | 7/26/2018 | 4:05 PM | 35.1 | Indoor | Low | Indoor |
| | A1 | 7/26/2018 | 12:55 PM | 34.1 | In-transit | Low | Cleaning |
| | A1 | 7/25/2018 | 3:25 PM | 34.1 | Indoor | Low | Working at desk |

**Table 7.** Summary statistics for the weather station and wearable sensor devices in Boone, NC and Starkville, MS.

| Sensor Type | ASU | | | | MSU | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean (℃ ) | Max (℃ ) | Min (℃ ) | Standard Deviation | Mean (℃ ) | Max (℃ ) | Min (℃ ) | Standard Deviation |
| Weather Station | 20.3 | 26.9 | 15.2 | 2.5 | 24.5 | 33.5 | 16.1 | 3.6 |
| Hygrochron (Down) | 20.0 | 31.1 | 13.1 | 3.8 | 25.0 | 39.6 | 15.6 | 4.9 |
| Hygrochron (Up) | NA | NA | NA | NA | 25.3 | 42.6 | 14.6 | 5.6 |
| Hygrochron (Shield) | NA | NA | NA | NA | 24.9 | 36.0 | 21.0 | 4.1 |
| Thermochron (Down) | NA | NA | NA | NA | 25.5 | 38.0 | 19.0 | 4.8 |
| Thermochron (Up) | 20.9 | 40.0 | 13.0 | 5.1 | NA | NA | NA | NA |
| Kestrel (Down) | 19.5 | 30.7 | 13.2 | 3.2 | 24.5 | 36.9 | 15.4 | 4.2 |
| Kestrel (Up) | 20.1 | 37.6 | 12.5 | 4.6 | 25.3 | 45.4 | 14.4 | 6.1 |
| Kestrel (Shield) | NA | NA | NA | NA | 24.3 | 35.0 | 15.9 | 3.7 |
| HOBO (Down) | 21.6 | 35.2 | 13.6 | 5.2 | 26.5 | 46.1 | 15.8 | 6.4 |
| HOBO (Up) | 19.0 | 46.7 | 12.9 | 7.0 | 26.4 | 45.2 | 15.1 | 6.4 |

**Table 8.** Measures of agreement between the stationary sensors and weather station. MAE = mean absolute error, MBE = mean bias error (negative values signify under prediction and positive values signify over prediction in comparison to the weather station), RMSE = root mean square error, Pc = Lin's concordance correlation coefficient.

| Sensor Type | ASU | | | | MSU | | | |
|---|---|---|---|---|---|---|---|---|
| | MAE | MBE | RMSE | *Pc* | MAE | MBE | RMSE | *Pc* |
| Hygrochron (Down) | 5.92 | -0.59 | 7.13 | 0.26 | 2.11 | 0.86 | 3.30 | 0.91 |
| Hygrochron (Up) | NA | NA | NA | NA | 2.98 | 1.46 | 4.65 | 0.85 |
| Hygrochron (Shield) | NA | NA | NA | NA | 1.18 | 0.75 | 1.86 | 0.96 |
| Thermochron (Down) | NA | NA | NA | NA | 2.18 | 0.44 | 3.34 | 0.90 |
| Thermochron (Up) | 7.72 | 1.05 | 10.03 | 0.05 | NA | NA | NA | NA |
| Kestrel (Down) | 5.18 | -1.56 | 6.04 | 0.36 | 1.93 | 0.07 | 2.61 | 0.93 |
| Kestrel (Up) | 6.89 | -0.34 | 8.39 | 0.21 | 3.69 | 1.53 | 5.69 | 0.81 |
| Kestrel (Shield) | NA | NA | NA | NA | 0.87 | 0.21 | 1.20 | 0.98 |
| HOBO (Down) | 7.74 | 1.79 | 9.72 | 0.16 | 4.62 | 3.55 | 6.98 | 0.74 |
| HOBO (Up) | 9.74 | 3.30 | 12.69 | 0.16 | 4.78 | 3.43 | 7.08 | 0.73 |

**Table 9.** A Linear mixed model with a random intercept for each participant comparing the fixed weather station data to participant sensor measurements at ASU. $\sigma^2$ = variance of population values, $\tau_{00}$ = intercept variance, ICC= Intraclass Correlation Coefficient, Marginal $R^2$ = the proportion of variance explained by the fixed factor (weather station), Conditional $R^2$ = the proportion of variance explained by the fixed and random factors

| | **HOBO Temp C** | | | **iButton 1 Temp C** | | | **iButton 2 Temp C** | | | **Kestrel Temp C** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *p* | *Estimates* | *CI* | *p* | *Estimates* | *CI* | *p* | *Estimates* | *CI* | *p* |
| (Intercept) | 17.95 | 16.96 – 18.95 | **<0.001** | 18.65 | 17.79 – 19.51 | **<0.001** | 18.82 | 17.42 – 19.98 | **<0.001** | 17.56 | 16.45 – 18.66 | **<0.001** |
| Weather Station Temp C | 0.27 | 0.23 – 0.31 | **<0.001** | 0.20 | 0.16 – 0.23 | **<0.001** | 0.27 | 0.22 – 0.33 | **<0.001** | 0.25 | 0.21 – 0.29 | **<0.001** |
| **Random Effects** | | | | | | | | | | | | |
| $\sigma^2$ | 4.78 | | | 3.65 | | | 7.99 | | | 3.92 | | |
| $\tau_{00}$ | 1.42 $_{ID}$ | | | 1.24 $_{ID}$ | | | 1.30 $_{ID}$ | | | 2.20 $_{ID}$ | | |
| ICC | 0.23 $_{ID}$ | | | 0.25 $_{ID}$ | | | 0.14 $_{ID}$ | | | 0.36 $_{ID}$ | | |
| Observations | 1949 | | | 2099 | | | 1793 | | | 1697 | | |
| Marginal $R^2$ / Conditional $R^2$ | 0.059 / 0.275 | | | 0.040 / 0.284 | | | 0.040 / 0.174 | | | 0.050 / 0.392 | | |

**Table 10.** A Linear mixed model with a random intercept for each participant comparing the fixed weather station data to participant sensor measurements at MSU. $\sigma^2$ = variance of population values, $\tau_{00}$ = intercept variance, ICC= Intraclass Correlation Coefficient, Marginal $R^2$ = the proportion of variance explained by the fixed factor (weather station), Conditional $R^2$ = the proportion of variance explained by the fixed and random factors.

| | **HOBO Temp C** | | | **Hygro Temp C** | | | **iButton 2 Temp C** | | | **Kestrel Temp C** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *p* | *Estimates* | *CI* | *p* | *Estimates* | *CI* | *p* | *Estimates* | *CI* | *p* |
| (Intercept) | 13.04 | 12.18 – 13.90 | **<0.001** | 12.10 | 11.07 – 13.12 | **<0.001** | 18.63 | 17.36 – 19.89 | **<0.001** | 15.08 | 14.48 – 15.68 | **<0.001** |
| Weather Station Temp C | 0.42 | 0.41 – 0.44 | **<0.001** | 0.50 | 0.48 – 0.52 | **<0.001** | 0.28 | 0.22 – 0.33 | **<0.001** | 0.32 | 0.31 – 0.34 | **<0.001** |
| **Random Effects** | | | | | | | | | | | | |
| $\sigma^2$ | 5.59 | | | 8.11 | | | 6.93 | | | 4.30 | | |
| $\tau_{00}$ | 2.21 $_{ID}$ | | | 2.98 $_{ID}$ | | | 1.39 $_{ID}$ | | | 0.90 $_{ID}$ | | |
| ICC | 0.28 $_{ID}$ | | | 0.27 $_{ID}$ | | | 0. 17 $_{ID}$ | | | 0.17 $_{ID}$ | | |
| Observations | 24025 | | | 19492 | | | 1689 | | | 17179 | | |
| Marginal $R^2$ / Conditional $R^2$ | 0.210 / 0.434 | | | 0.208/ 0.421 | | | 0.044 / 0.204 | | | 0.194/ 0.334 | | |

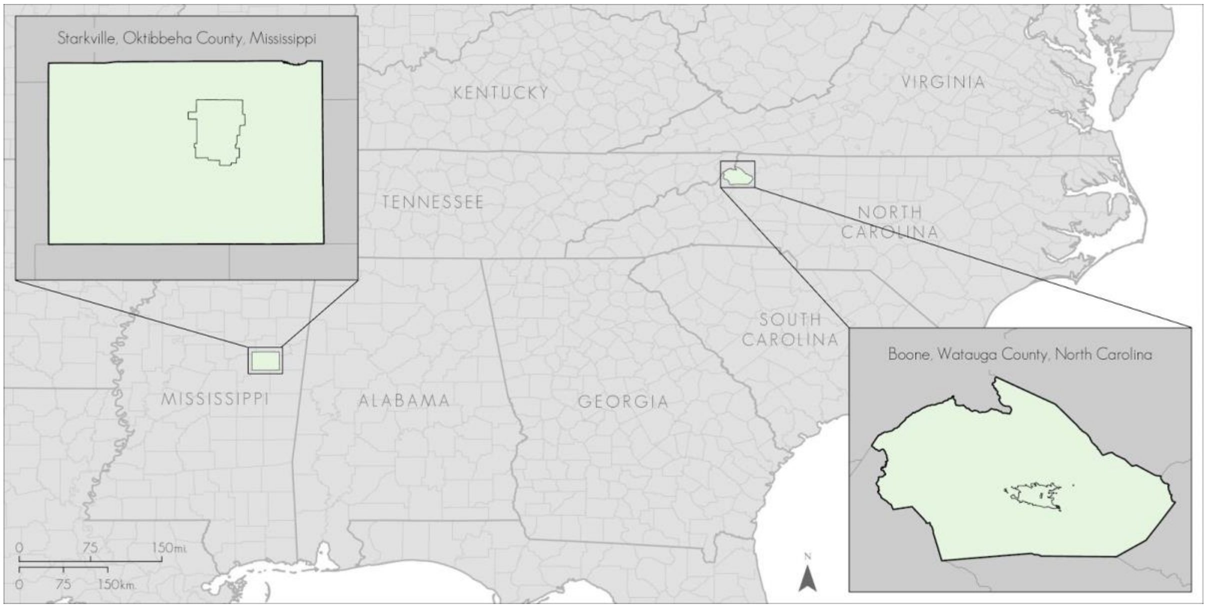


**Figure 1.** Study area map showing the locations of Boone, NC and Starkville, MS.

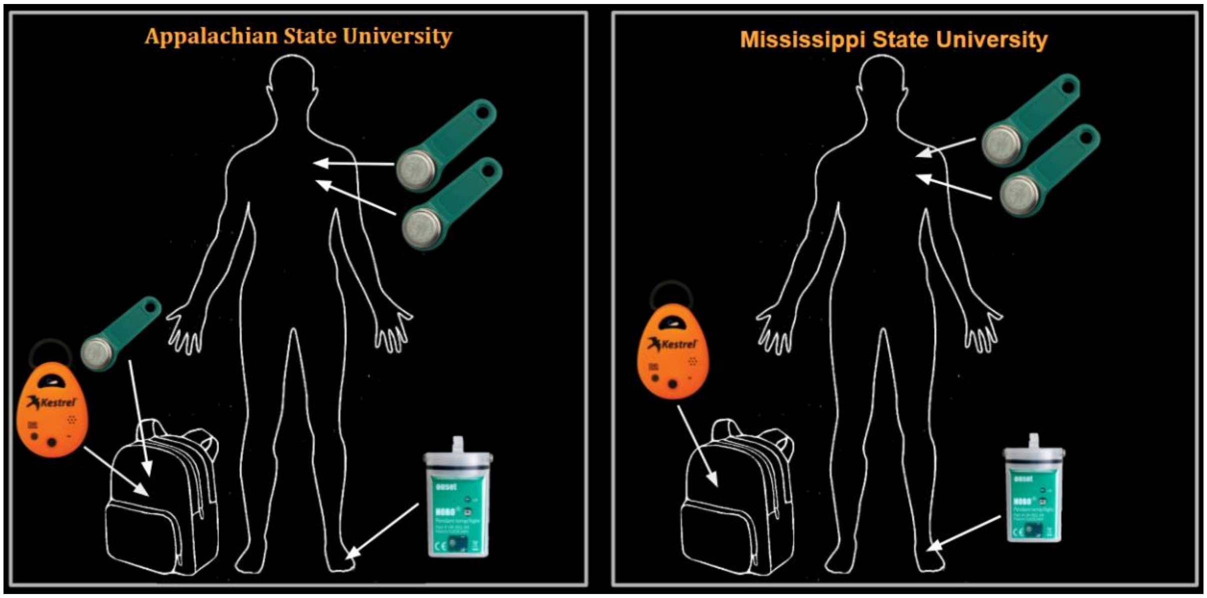


**Figure 2.** Placement of sensors for ASU and MSU participant studies. At ASU, a HOBO was placed on the shoe, a Kestrel and an iButton Thermochron were placed on a backpack, and an iButton Thermochron and Hygrochron were placed on the shirt collar. At MSU, a HOBO was placed on the shoe, a Kestrel on a backpack, and an iButton Thermochron and Hygrochron were placed on the shirt collar.

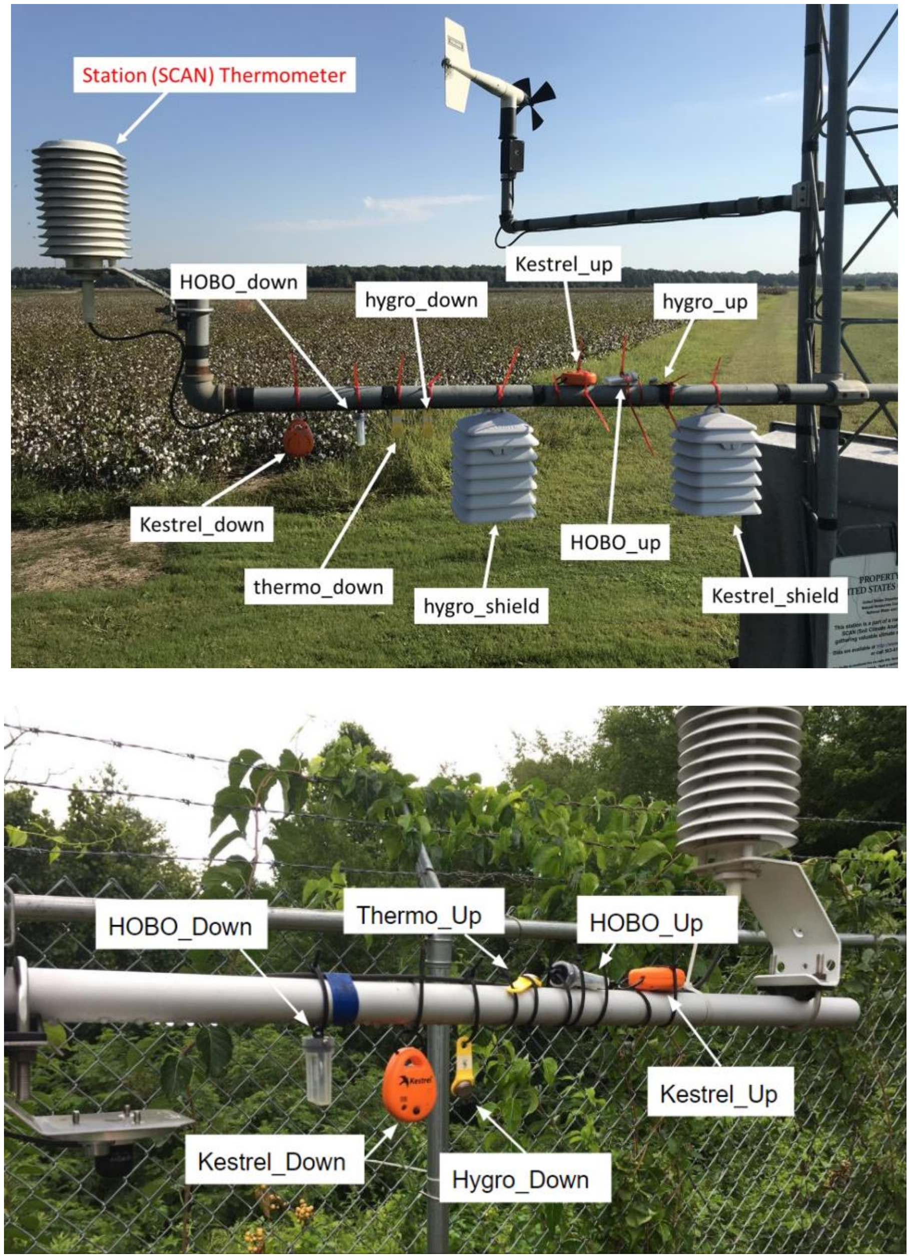


**Figure 3.** Sensors attached to the weather station located in Starkville, Mississippi (above) and Boone, North Carolina (below).

Figure 4.1: iButton 1 and Kestrel at ASU

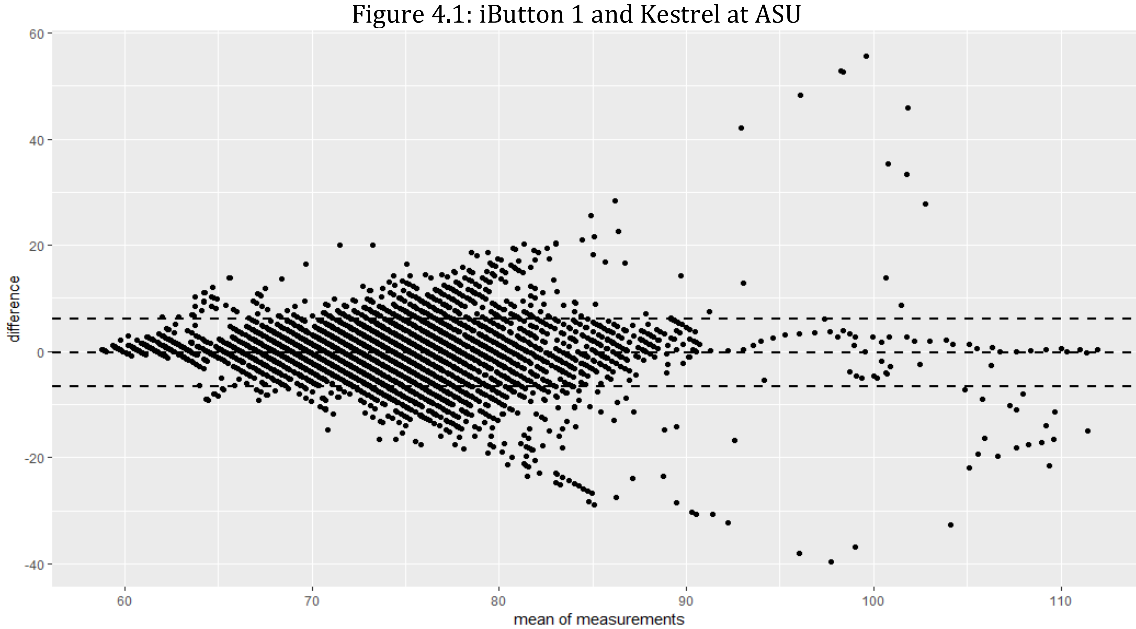


Figure 4.2: iButton 1 and iButton 2 at ASU

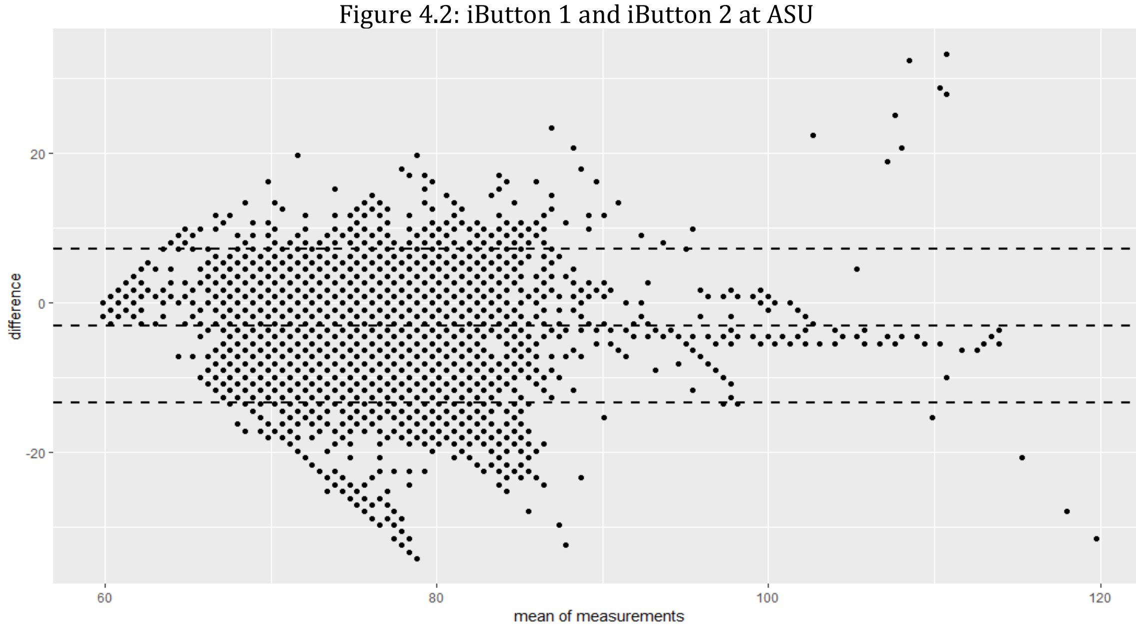

Figure 4.3: iButton 2 and HOBO at MSU

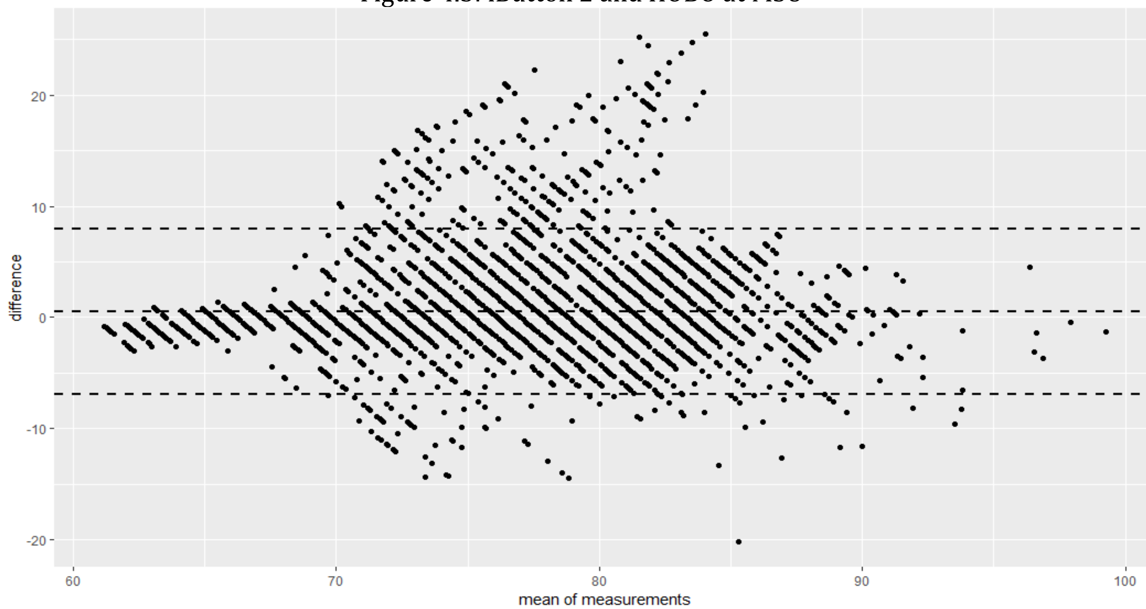


Figure 4.4: Hygrochron and HOBO at MSU

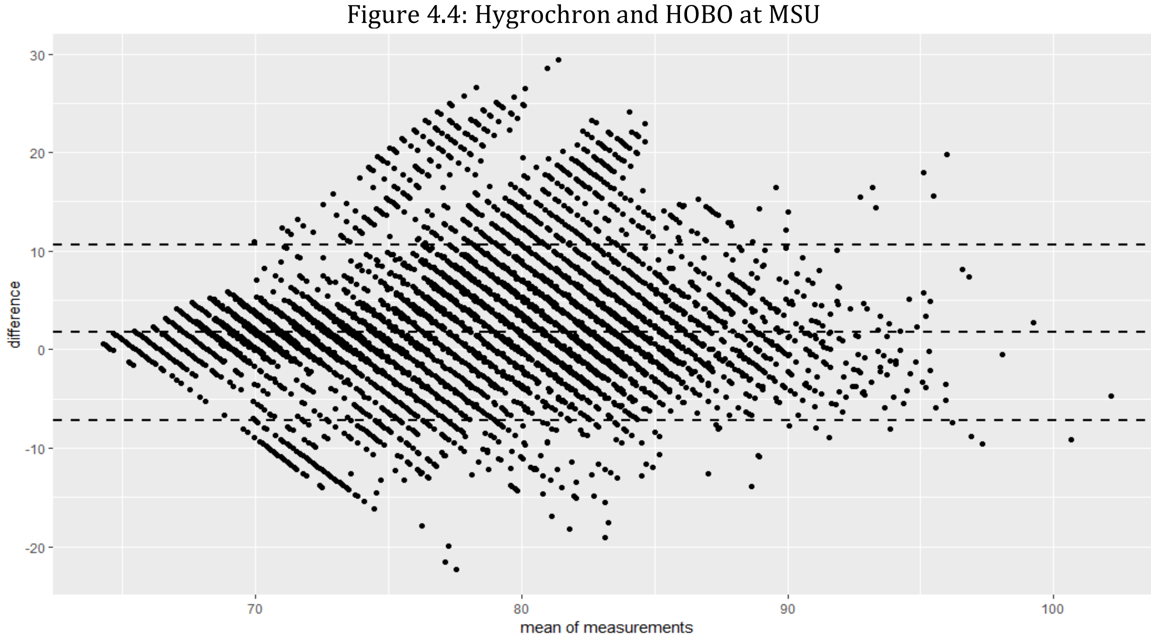


**Figure 4.** Bland Altman plots for the sensors with the highest (Figures 4.2 and 4.4) and lowest (Figures 4.1 and 4.3) mean differences at ASU and MSU.

Figure 5.1: iButton 1 during all activity levels

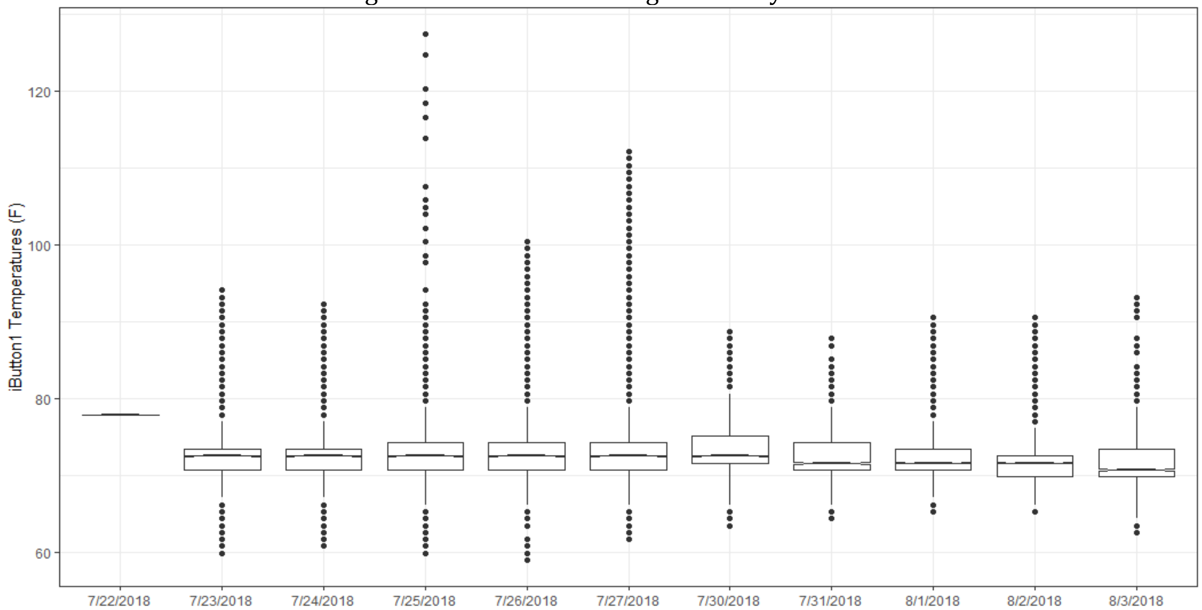


Figure 5.2: iButton 1 during low activity levels

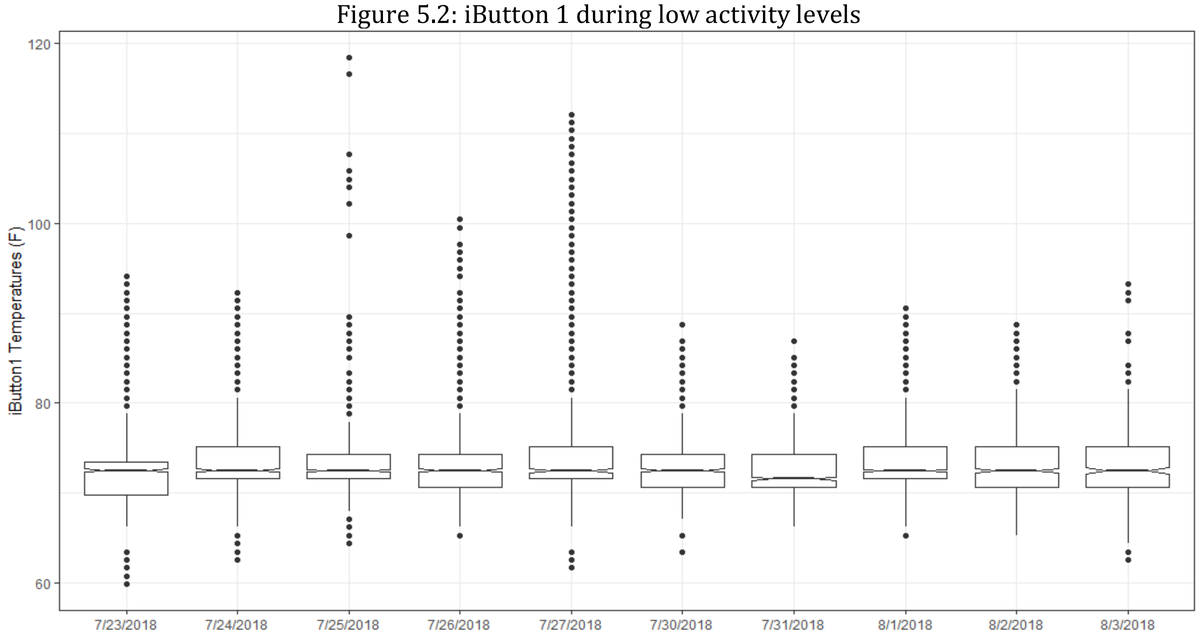

Figure 5.3: iButton 1 during moderate activity levels

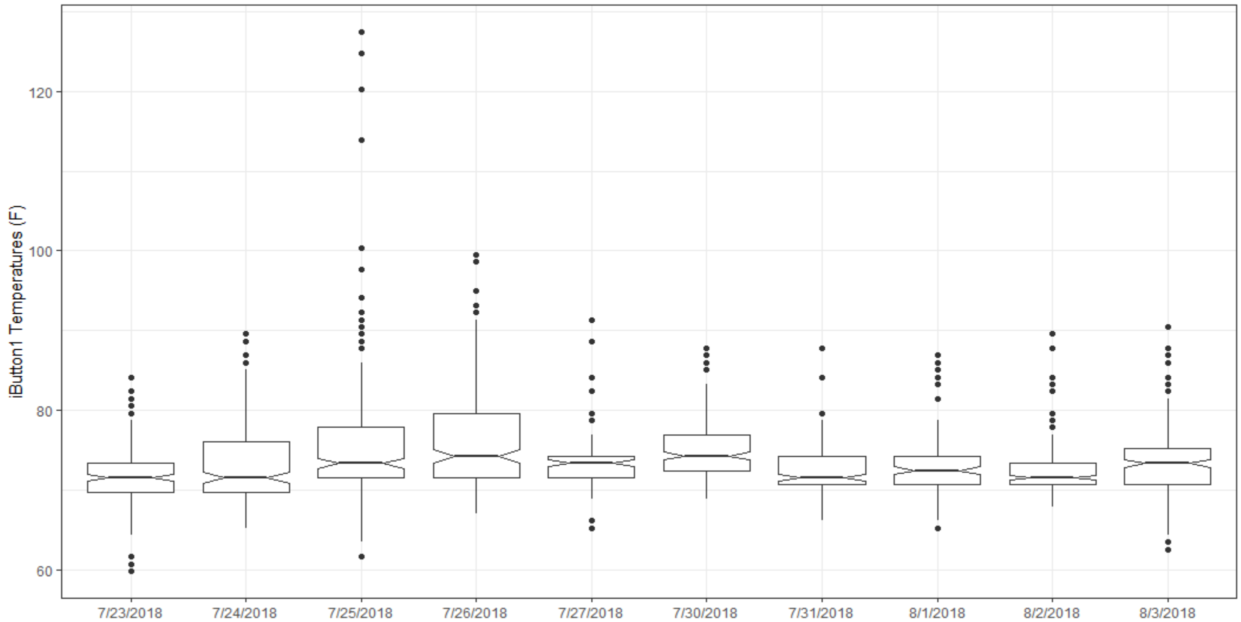


Figure 5.4: iButton 1 during high activity levels

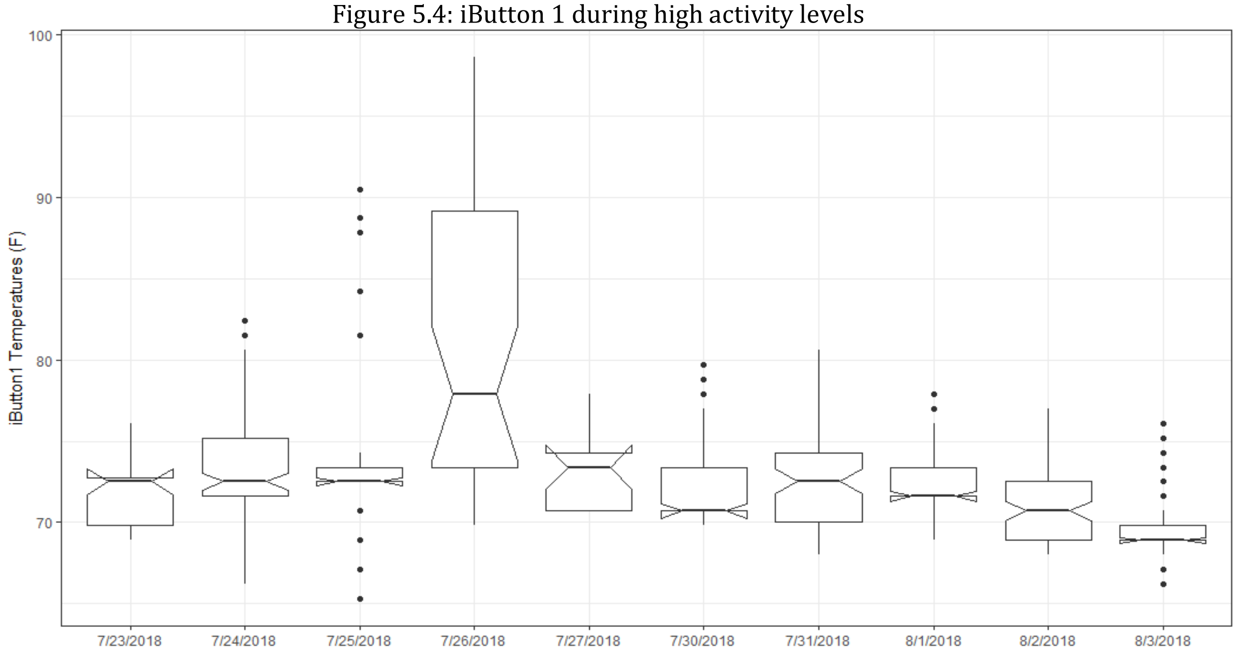


**Figure 5.** Box plots showing mean temperatures during low, moderate, and high activity levels throughout the ASU participant study.

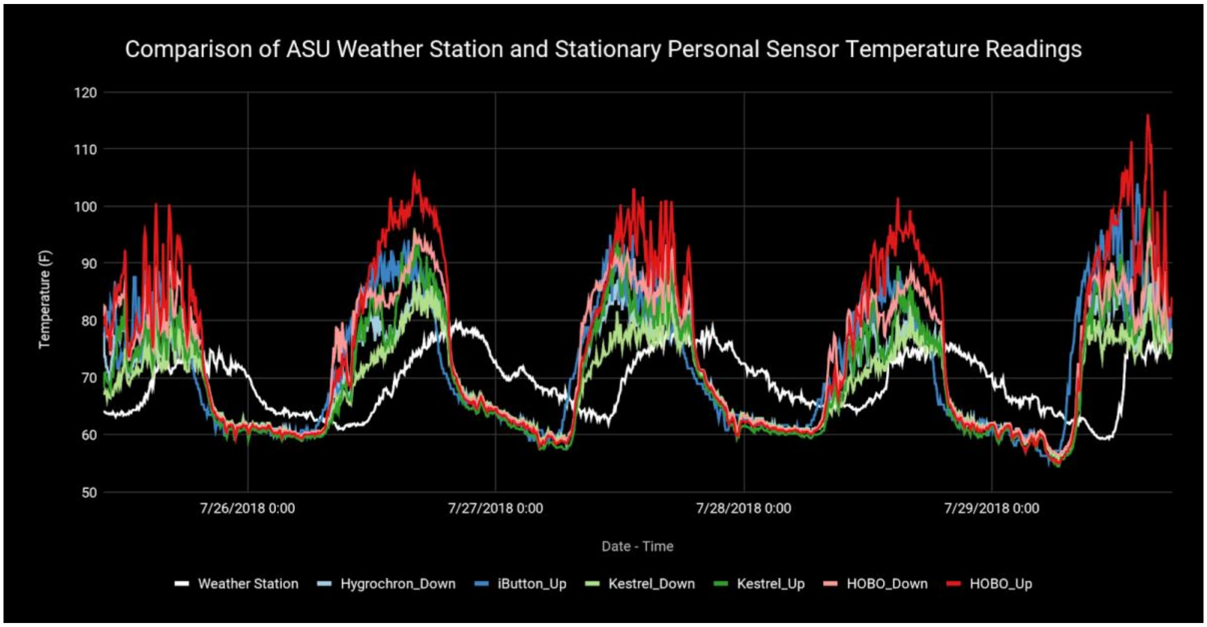


**Figure 6.** A comparison of the Appalachian State University weather station and personal sensor temperature readings from July 26th, 2018 through July 29th, 2018.

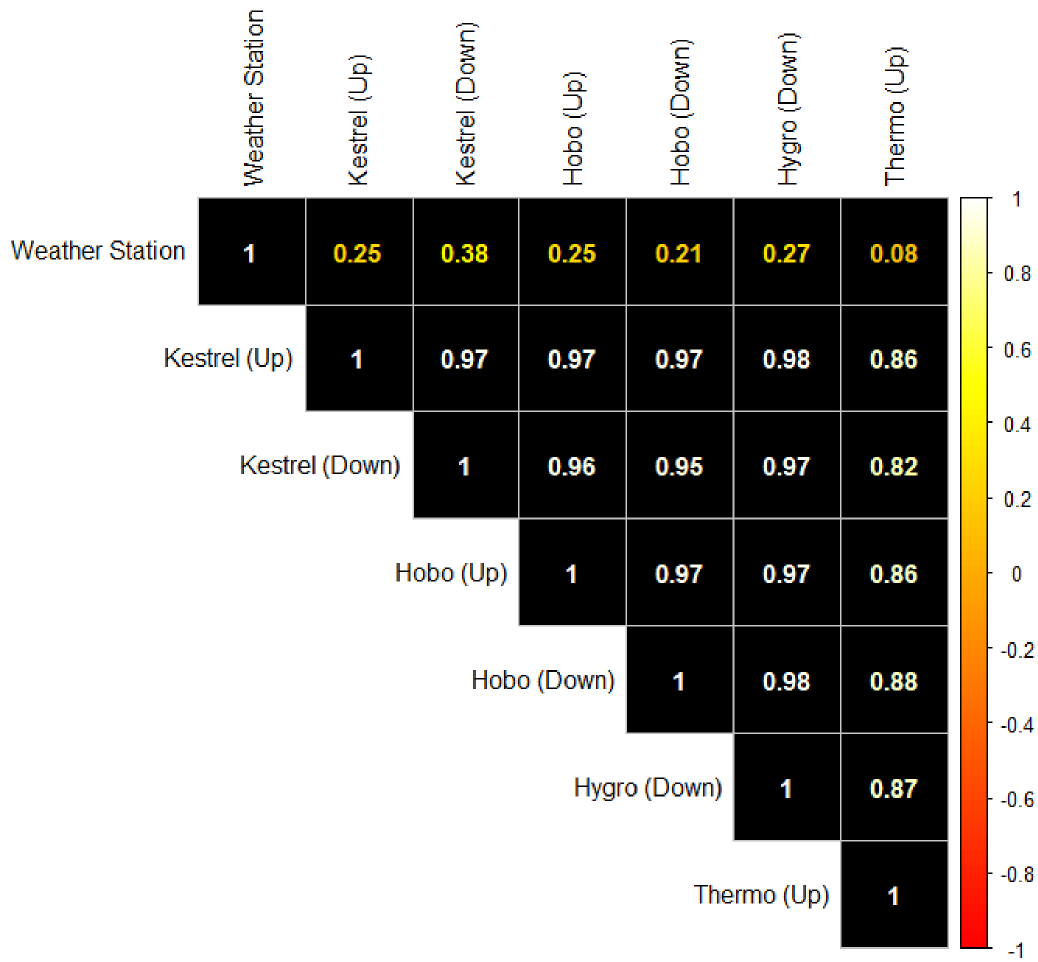


**Figure 7.** Correlation values between weather station in-situ temperature readings and sensor temperature readings from sensors attached to the weather station in Boone, NC.

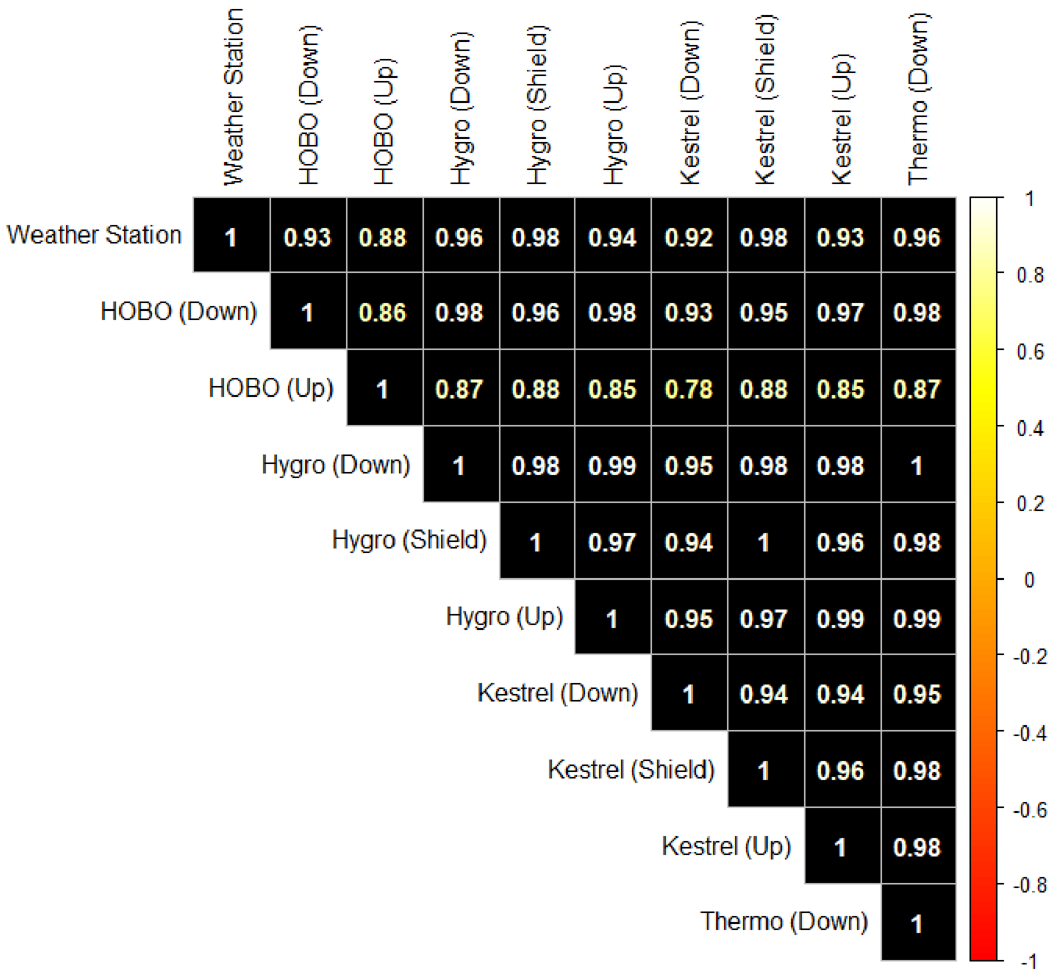


**Figure 8.** Correlation values between weather station in-situ temperature readings and sensor temperature readings from sensors attached to the weather station in Starkville, MS.

## Vita

Elizabeth Frances Bailey was born and raised in Tallahassee, Florida. Her parents, Douglas and Paula Bailey, instilled an appreciation for the outdoors in her from a young age. After graduating from Leon High School in Tallahassee, Florida, in May 2013, Elizabeth decided to move to her mother's hometown of Charlotte, North Carolina, to attend Queens University of Charlotte.

As an undergraduate, Elizabeth studied political science and international relations. As a junior, Elizabeth took a Geographic Information Systems class under the instruction of Dr. Reed Perkins. This opportunity inspired Elizabeth to pursue a master’s degree in Geography at Appalachian State University. Elizabeth graduated Summa Cum Laude with a Bachelor of Arts in Political Science in May 2017.

Elizabeth began her graduate education at Appalachian in August 2017. As a graduate student, Elizabeth became involved in heat sensor research with Dr. Maggie Sugg. Throughout her graduate career, Elizabeth was able to present her research at 2019 Annual Meeting of the American Association of Geographers, the 2018 Annual Meeting of the Southeastern Division of the American Association of Geographers, and the 2018 and 2019 Celebration of Student Research and Creative Activities Symposium.

Upon graduating with a Master of Arts in Geography in May 2018, Elizabeth plans to pursue a career in GIS with the consideration of pursuing a Ph.D. in Geography in the future.