CHALLENGES AND LIMITATIONS OF THERMAL INFRARED REMOTE SENSING WITH UNOCCUPIED AERIAL SYSTEMS

by

Hayden Robinson Yates

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Earth Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

December 2023

ACKNOWLEDGEMENTS

To start, I would like to thank my advisor, Dr. Andrew K. Laskowski, for providing me with this opportunity and supporting me throughout this journey, especially during the challenging periods. I am also grateful to Dr. Mary Hubbard and Frank Dougher, both of whom served on my committee. To Frank - thank you for enthusiastically engaging with my research and providing a space for collaboration.

Several phases of this project relied on the generosity and contribution of fellow MSU researchers. In particular, I would like to thank Dr. Joseph A. Shaw and Erica Venkatesulu of the Electrical & Computer Engineering Department for welcoming me into their lab to run calibration experiments. Additional thanks to Dr. Eric Sproles for sharing equipment necessary for the field component of this project.

I would also like to acknowledge the Montana Association of Geographic Information Professionals (MAGIP), MontanaView, and the Department of Earth Sciences for supporting this work through scholarships, fellowships, and awards.

Lastly, I am endlessly thankful for the love and support of my parents, Nancy and Carter Yates. They cultivated my love for learning, which has been the richest gift of all.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

ABSTRACT

Unoccupied aerial systems (UAS) thermal infrared (TIR) remote sensing is emerging as a novel alternative to satellite remote sensing across many scientific applications, due to its unique advantages for spatial and temporal resolution. However, much is still unknown regarding its utility for geothermal resource assessment. In particular, the nascence of this technology means that technique development, especially for individual camera models, is still needed to generate reliable thermal data. Therefore, to evaluate whether UAS TIR remote sensing is a viable option for geothermal reconnaissance, the objectives of this study included: 1) assessing the accuracy of the FLIR Duo Pro R radiometric thermal camera, 2) testing the ability of various photogrammetric workflows to preserve the thermal signature of geothermal anomalies, 3) modeling the most influential explanatory variables of remotely-sensed temperatures using a machine learning (ML) model in ArcGIS Pro, and 4) determining whether our ML model can predict the thermal variation in our remotely sensed data, based on a curated set of explanatory variables. To accomplish these objectives, this study employed a three-phased approach: lab experiments, field surveying, and computational modeling. The results from our lab experiments indicate that the FLIR Duo Pro R overestimates stable temperatures at varying degrees (6-7°C) when it is operating at steady-state conditions. Our results also suggest that it exhibits measurement instability when exposed to simulated changes in environmental conditions, i.e. wind and heat. For instance, our wind experiment found that the camera underestimated the blackbody temperature by 8°C when exposed to wind and jumped nearly 16°C after the transition from wind to steady-state conditions. We also found that photogrammetric techniques strongly influence the signature (presence and magnitude) of thermal anomalies associated with geothermal system in UAS TIR imagery. Lastly, our ML model cannot confidently identify the most influential explanatory variables of remotely-sensed thermal anomalies or reliably predict remotely-sensed ground temperature based on explanatory variables. We conclude that the current limitations associated with UAS TIR remote sensing hinder its ease-of-use for geothermal applications, especially where thermal anomalies are subtle and require accurate temperature readings to differentiate them from their surroundings. Despite the challenges discussed in this paper, these techniques may still serve as a tool for assessing the spatial distribution of thermal properties of geothermal systems if future work continues to improve and develop these methodologies. Most notably, this technique will benefit from advancements regarding the temperature drift of the sensor during image acquisition and the photogrammetric processing of thermal imagery.

1

INTRODUCTION

Geothermal energy is an increasingly attractive, low-carbon alternative to fossil fuels, but current global production only represents a fraction of the total global potential (Jolie et al., 2021). As interest in geothermal development grows, there is a pronounced need to improve current reconnaissance methods, as they are labor intensive and financially prohibitive. For example, drilling is a common tool used for prospecting, but poses significant economic risk, as it can account for nearly 50% of the capital cost required for developing a geothermal site (Faulds et al., 2021; Robins et al., 2022). To meet this objective for low-cost reconnaissance techniques, nontraditional solutions such as the integration of thermal infrared (TIR) unoccupied aerial system (UAS) remote sensing and machine learning may be a cost-effective approach for future assessment of geothermal resources.

Capturing TIR imagery via UAS provides a quantifiable and repeatable method for assessing the spatial and temporal variability of ground surface temperatures at unprecedented resolution. This emerging technique is applicable across a broad range of disciplines, including precision agriculture, civil engineering, security & surveillance, archaeology, zoology, and healthcare (Dlesk et al., 2022). Within this wide range of applicability, its value for geothermal investigations is still growing. Previous studies that have leveraged UAS thermal mapping for geothermal systems have explored diverse topics: assessing the geologic hazards associated with geothermal features, inventorying remote and/or inaccessible geothermal systems, monitoring thermal changes over time, managing surface water resources, monitoring production and operations at existing power plants, and quantifying heat flux (Allahbakhshi et al., 2023; Bjornsson et al., 2019; Bunker et al., 2021; Harvey et al., 2016; Haselwimmer et al., 2013; Nishar et al., 2016; Sedano-Cibrián et al.,

2022; Silvestri et al., 2020). Additionally, a recent study at the Geysir geothermal field in Iceland demonstrated that detailed drone mapping can correlate near-surface thermal signatures with structural controls (Walter et al., 2020). However, the nascence of UAS TIR remote sensing means that method development is still needed to generate reliable thermal data, especially considering the myriad variables between differing applications (i.e. study objective, camera model, environmental conditions, spatial resolution, etc).

In conjunction with drone-based thermal mapping, geothermal resource assessment may benefit from advanced computational methods such as machine learning, which optimizes algorithms to associate certain features with multiple sets of input data (Faulds et al., 2021; Jolie et al., 2021). In this study, we explored whether machine learning can effectively determine which surficial factors, if any, are diagnostic of UAS-imaged thermal anomalies at the earth's surface. To do so, we combined drone surveying, in situ ground sampling, and machine learning to explore a new framework for constraining thermal anomalies related to geothermal systems.

To investigate the utility of UAS TIR remote sensing and machine learning for geothermal resource characterization, this research employed a three-phased approach: lab experiments, field surveying, and computational modeling. The first part of this project investigated the accuracy and environmental sensitivity of a radiometric thermal camera, the FLIR Duo Pro R (FDPR), in a lab setting. During the second phase, we captured optical and thermal imagery with a tricopter drone and collected in situ field data immediately post-flight. For the final chapter, we created orthorectified optical and thermal mosaics using photogrammetry, followed by the analysis and modeling of spatial data in a Geographic Information Systems (GIS) software to determine whether any surficial, geologic, or climactic factors lead to hot spots in drone thermal imagery.

A spring-dominated geothermal system in southwest Montana, Norris Hot Springs, served as the case study for the development of this technique. Known geothermal systems

are scattered throughout southwest Montana as a result of the Basin and Range extensional tectonics of this region, which created an ideal setting for the high heat flow and permeability required of geothermal systems. Despite this, geothermal resources remain underexplored and underdeveloped in the state of Montana. Thus, this work contributes towards the development of novel and interdisciplinary methods for characterizing geothermal systems, as well as the understanding of geothermal resources in Montana.

Thermal Infrared Remote Sensing

Remote sensing is the process of deriving information about an object or surface by measuring its reflected or emitted electromagnetic (EM) energy, from a distance, via a sensor (Campbell and Wynne, 2011). As EM radiation moves through the atmosphere, it reaches the earth's surface, where it either a) bounces off the surface back towards the atmosphere or b) is absorbed by the surface and then later re-emitted by the same surface (Figure 1.1). Whether incoming energy is reflected or absorbed primarily depends on its wavelength as well as the properties of the given surface.

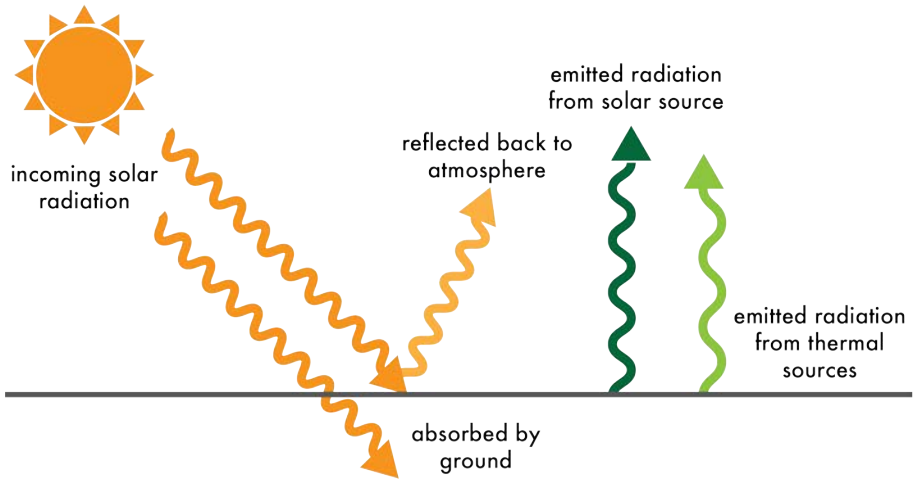

Figure 1.1: Schematic diagram showing the behavior of incoming solar radiation as it reaches the earth's surface, as well as the emission of radiation from thermal sources.

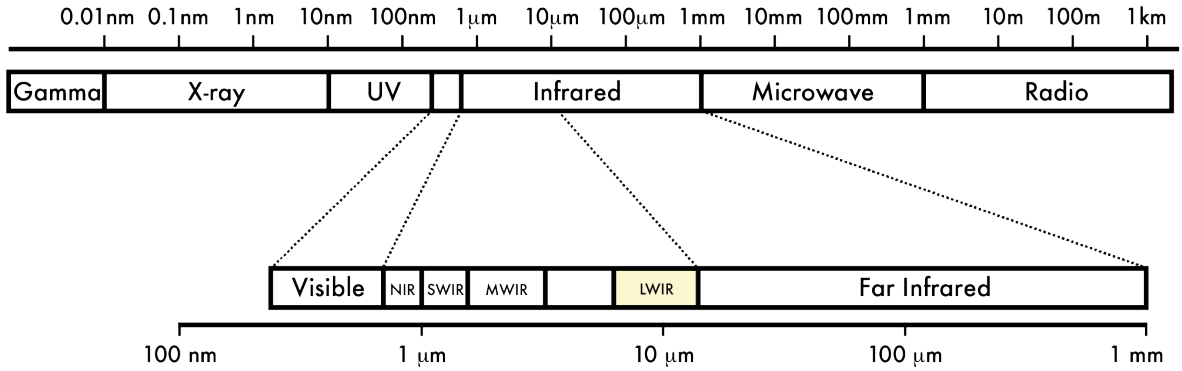

Figure 1.2: The electromagnetic spectrum; longwave infrared (LWIR) region highlighted.

A subset of remote sensing, TIR remote sensing, specifically measures the emitted longwave infrared (LWIR) radiation, as opposed to the reflected radiation measured in the visible and near infrared (NIR) portions of the EM spectrum (Figure 1.2). When thermal energy is emitted by an object, it either originates from a) the absorption and subsequent re-emission of incident energy or b) by the object itself (Figure 1.1). The majority of the emitted LWIR radiation detected at the earth's surface is attributed to the former, i.e. the energy transfer process where the earth's surface absorbs incoming solar radiation and then re-emits a corresponding amount of that radiation back into the atmosphere at longer wavelengths. While a significant amount of LWIR radiation detected at the earth's surface is re-emitted solar radiation, TIR remote sensing can also measure other sources of emitted TIR radiation. These sources include regions of anomalous heat flow within the earth's crust, which are of particular interest in this study.

Platform Types

Historically, satellites have served as the traditional platform for the acquisition of remotely sensed thermal imagery. However, unoccupied aerial vehicles (UAVs) are now considered a desirable platform alternative for many applications, across many disciplines. Acquiring LWIR data by UAS has many advantages to that of satellite, for both spatial

and temporal resolution. It is now possible to obtain thermal imagery at a much finer grain size (i.e. centimeter scale) than previously possible with satellite platforms, which typically capture thermal imagery at very coarse grain sizes (i.e. a ground sampling distance (GSD) of 30x30m). At such coarse resolutions, the emitted radiation recorded in a satellite pixel may represent an average value of thermal features and non-thermal features, effectively reducing the thermal signature of an anomalous feature as well as dulling the contrast between the anomaly and the surrounding area (Haselwimmer and Prakash, 2013). Conversely, the enhanced resolution achievable with UAS TIR remote sensing may yield more accurate thermal data on a pixel-by-pixel basis and provide the level of detail needed for thermal mapping in localized studies. In addition to superior spatial resolution, the repeatability of UAS missions enables greater temporal frequency (i.e., multiple flights per day) than currently available by satellite, providing the user with increased autonomy and flexibility.

Thermal Cameras

There are two types of thermal cameras available for UAS TIR remote sensing: non-radiometric and radiometric, which both operate within 8-14 μm of the EM spectrum (Figure 1.2) (Kelly et al., 2019). While each camera captures surface radiance, non-radiometric cameras record data that represent relative differences (i.e. the range of brightness values across an individual scene). Radiometric cameras, once calibrated, record data that represent absolute measurements of radiance or temperature (Kelly et al., 2019; Nugent et al., 2013).

To produce a digital image that represents the thermal properties of a given object or landscape, radiometric cameras absorb emitted infrared radiation detected within their field of view (FOV) (Aragon et al., 2020). The incoming radiation within the FOV moves through the camera's optical system (lens) towards the Focal Plane Array (FPA), which is a matrix of individual thermometers (Figure 1.3) (Aragon et al., 2020). As the radiation reaches the FPA, it uniquely alters the electrical resistance of each individual element (Aragon et al.,

2020). Based on the electrical signal recorded by each element in the array, a digital number (DN) is assigned to each pixel and the matrix of DNs forms the final image. The DNs from the raw images are then used to calculate temperature during post-processing (Aragon et al., 2020).

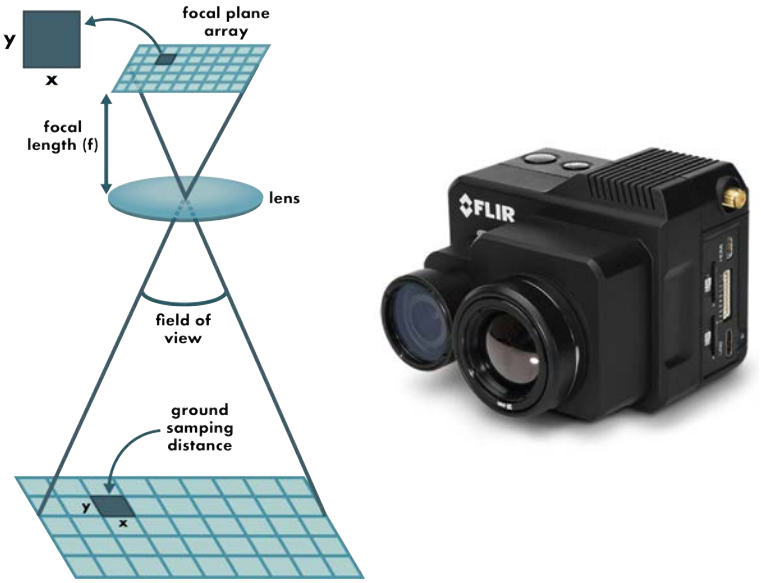

Figure 1.3: *Left:* Schematic diagram of the focal plane array and its relationship to the ground surface.
Right: The camera used in this study, the FLIR Duo Pro R (FDPR) radiometric thermal camera. Note the dual lenses, which concurrently capture optical and thermal imagery.

In addition to the two models of thermal cameras, there are two distinct types of FPAs: cooled and uncooled microbolometers. The functional difference between the two is important - cooled microbolometers have a thermo-electric cooler (TEC) acting on the detector array during camera operation, allowing it to maintain a stable temperature, whereas the uncooled version lacks this component and cannot maintain temperature stability (Nugent and Shaw, 2014). The advanced capability of cooled microbolometers is offset by the fact that they cost more and weigh more than their counterpart. Thus, the uncooled model currently dominates the UAS market, due to its lighter payload and greater affordability.

Despite these specific benefits for UAS applications, uncooled microbolometers may yield inaccurate thermal data since they are not temperature-stabilized (Kelly et al., 2019). Since the internal sensor temperature is not stabilized during operation, the FPA may produce an inconsistent real-time response as ambient temperature fluctuates (Nugent et al., 2013). In theory, the FPA should independently record an object's thermal signature despite changes in ambient temperature. However, this has not been observed for the following camera models: Apogee SI-11, FLIR A655sc, TeAx 640, and FLIR Vue Pro (Aragon et al., 2020; Kelly et al., 2019).

Ideally, the temperature drift of uncooled microbolometers should be rendered negligible through internal firmware corrections implemented by the camera manufacturer. As a compensatory measure, most cameras employ a Non-Uniformity Correction (NUC) or Flat Field Correction (FFC), which aim to standardize the response of each individual FPA element by capturing an image with the shutter closed, allowing the camera to register any radiation associated with the camera's internal components (Kelly et al., 2019). The camera then calibrates subsequent images by accounting for internal contributions of radiation (Budzier and Gerlach, 2015). The NUC and FCC use the shutter temperature for these corrections, assuming that it is a temperature proxy for all parts of the camera (Kelly et al., 2019). However, when a thermal camera is used for UAS missions, the lens experiences greater fluctuations in temperature than does the rest of the camera interior (Kelly et al., 2019). This likely renders the NUC and FCC insufficient in these specific scenarios. Given that these types of thermal cameras (including the FDPR of this study) are intended for use with drones, their potential sensitivity to varying environmental conditions is of concern.

Although UAS remote sensing is an increasingly prevalent tool for geospatial science, there are still notable limitations to procuring accurate thermal data with this method, owing to the temperature instability of the uncooled microbolometer. Due to this inherent problem, the calibration of thermal cameras with uncooled microbolometers is an active area

of research in the field of remote sensing. Previous studies have established camera-specific, empirical calibrations by using ground temperature reference points, employing ambient temperature as a proxy for FPA temperature, and measuring the temperature response of the FPA in a controlled thermal chamber (Aragon et al., 2020; Budzier and Gerlach, 2015; Kelly et al., 2019; Nugent et al., 2013). Unfortunately, these solutions are mostly prohibitive as they are labor-intensive and require specialized equipment. Another restrictive complication is that end-users still lack a uniform method for calibrating thermal cameras with uncooled microbolometers, as the calibration is unique to each individual model. This means that the correction relationships previously established in the literature cannot be applied to the camera used in this study, the FDPR (Figure 1.3). As such, the first part of this project tested this specific camera model in a lab setting to determine its baseline accuracy and its response, if any, to simulated environmental conditions.

Machine Learning

Recent work has explored the value of machine learning for certain geothermal investigations, including determining regional-scale suitability for very shallow geothermal systems and detecting hidden signatures of geothermal features (Assouline et al., 2019; Vesselinov et al., 2020). None, however, have integrated TIR remote sensing, ground-sampled training data, and machine learning as a potential approach for future geothermal research. Thus, one of the overarching objectives of our study focused on evaluating the effectiveness of machine learning in conjunction with TIR remote sensing and in situ field data for identifying which surficial and structural parameters influence UAS-imaged surface thermal anomalies of geothermal systems. To do so, we tested whether a machine learning model can algorithmically establish the most causal variables of thermally anomalous signatures recorded by UAS at the ground surface as well as predict remotely sensed temperature based on explanatory variables.

Geothermal Systems

Energy Development

As global energy dependence moves towards renewable resources, geothermal systems are a strong candidate for the development of low-carbon energy. To date, geothermal only represents a small percentage of the global green energy market, whereas other renewables such as wind and solar occupy a larger share (Brooks, 2022). Besides the primary benefit of being low-carbon, another important advantage of geothermal energy is its ability to generate power regardless of changing weather patterns, unlike wind and solar (Jolie et al., 2021). While all three renewables are poised for growth, geothermal power has a particularly significant opportunity since it is currently underutilized as a resource. Accordingly, its global potential is projected to reach 150 GWe by 2050 (Jolie et al., 2021). Although the United States already produces 22% of the annual global geothermal capacity (3.5 of 15.5 GWe), making it one of the largest global players, its potential far exceeds current production (Faulds et al., 2021; Jolie et al., 2021). This vast untapped potential is geographically concentrated in the western half of the contiguous U.S., due to the tectonic history of this region and associated heat flow regime (Figure 1.4).

Geologic Context

There are three major factors required for the formation of geothermal systems: high heat flow, sufficient permeability, and fluid convection (Jolie et al., 2021). As cold meteoric water percolates through the the earth's crust, it heats up at depth, and then buoyantly rises back towards the surface, typically along conduits such as faults (Siler and Pepin, 2021). Depending on the geologic setting, the heated groundwater may remain trapped at depth within a reservoir or it may eventually reach the surface of the earth and manifest as a geyser, hot spring, fumarole, or other diagnostic hydrothermal feature.

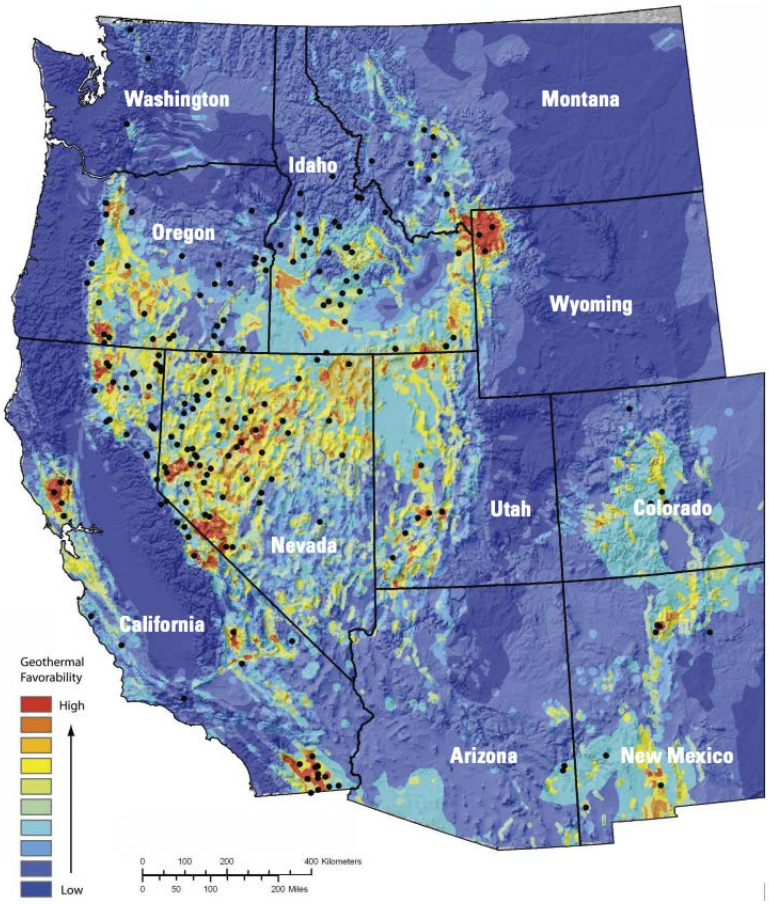

Figure 1.4: Geothermal potential map of the western United States. Black dots denote existing geothermal plants (USGS, 2008).

At the most fundamental level, the heat for geothermal systems originates from the Earth's core, which actively radiates heat from its center out towards the surface. While most of this heat is stored at depths too deep for current extraction techniques, some of this heat is transferred to the near-surface through the process of hydrothermal convection, bringing it to depths shallow enough for contemporary extraction methods (Jolie et al., 2021). Regions of anomalously high heat flow within the earth's crust are categorized as magmatic or amagmatic, where the former refers to upper to mid-crustal magma chambers and the latter refers to areas of thinned crust resulting from extensional tectonic processes (Faulds

et al., 2019). Although amagmatic sources have a shallower temperature-depth profile than magmatic sources, they still provide sufficient heat for geothermal activity under the right conditions (Jolie et al., 2021). For example, thermal outflows related to crustal thinning located near our study area in southwestern Montana range from 23 to 84°C (Sondregger et al., 1981). This phenomenon of high heat flow resulting from tectonic extension is of particular importance for the western United States, where the region's geothermal potential is largely related to episodes of crustal thinning throughout the Cenozoic (Figure 1.4).

Geologic Background

Regional Tectonics

The late Cenozoic Basin and Range province of the western United States is characterized by regional-scale crustal extension. Crustal thinning developed in two distinct phases: the Late Paleogene transition from convergence to extension and subsequent Basin and Range extension during the Miocene (Constenius, 1996). The first episode began during the Late Cretaceous, after the cessation of crustal thickening during the Sevier and Laramide orogenies. Due to the high gravitational potential associated with the overthickened crust, the unstable mass started to collapse, which was facilitated by large-scale normal faulting (Constenius, 1996). This period of extension also exhumed mid-crustal rocks from depth via the formation of metamorphic core complexes (Figure 1.7) (Constenius, 1996; Snoke and Chapman, 2021).

The second phase of extension initiated circa 29 Ma as the Farallon plate finished subducting beneath the North American plate, changing the margin from a subduction zone to a transform boundary. As the system evolved to form the San Andreas fault, the continental block of the North American plate underwent transtensional rotation (Dickinson, 2006). This transition in relative plate motion initiated crustal thinning throughout much of the western United States, with the resultant distribution of faulting indicating clockwise

rotation around an axis to the north (Janecke, 2007). The deformation style of this period of extension is categorized as Basin and Range, which refers to a distinct faulting pattern and associated topographic sequence of alternating ridges and basins. In this type of regime, large-scale normal faults create features known as horst & grabens or half-grabens (Figure 1.5). In these structures, the footwall drops down relative to the hanging wall, creating a topographic low where river drainages develop and valley fill accumulates.

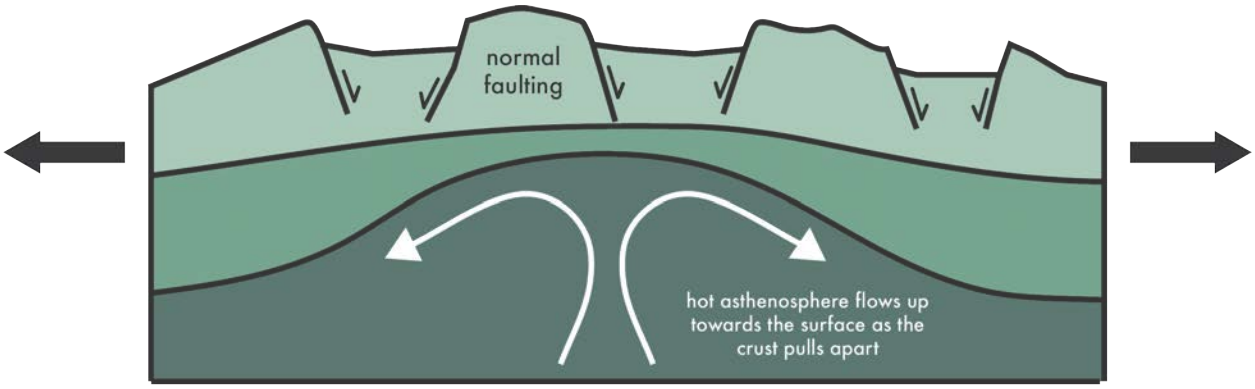

Figure 1.5: Schematic diagram showing crustal thinning accommodated by horst and graben structures in the crust. The rise of the asthenosphere at depth increases heat flow towards the surface.

The predominant physiographic feature associated with Cenozoic thinning is the Great Basin, which is a tectonically active region of extension and transtension within the broader Basin and Range structural province (Faulds et al., 2019). Although southwest Montana is geographically segmented from the Great Basin by the Snake River Plain, it is characterized by the same basin and range-style faulting and is closely related in tectonic style and history (Reynolds, 1979). In southwest Montana and eastern Idaho, NE-SW oriented normal faults related to Late Cenozoic Basin and Range extension are scattered above the eastern part of the Snake River Plain (Janecke, 2007). In this region, present-day Basin and Range orientations and extent developed during the Miocene and still remain seismically active (Reynolds, 1979).

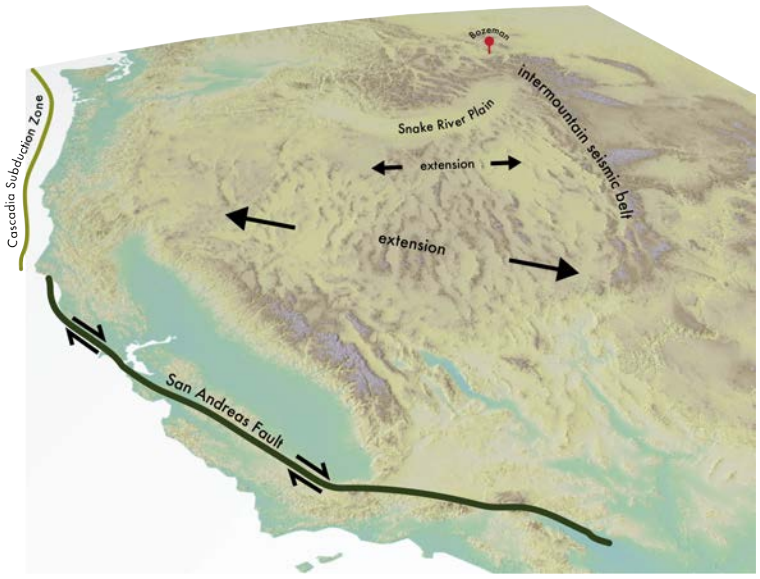

Figure 1.6: Relief map of the western United States, marked with features of interest. Since the Oligocene, the sense of motion along the San Andreas Fault has led to oblique extension throughout the Basin and Range province.

The tectonic boundary between the official Basin and Range structural province and the North American craton & northern Rocky Mountains is delineated by regional-scale structural discontinuities in Southwest Montana. Here, faulting related to Cenozoic extension terminates at the intermountain seismic belt to the east and the Lewis and Clark Line to the north & northeast (Reynolds, 1979; Sears and Fritz, 1998). The eastern margin of Basin and Range faulting is defined by the northern section of the intermountain seismic belt, a prominent zone of intraplate seismicity that extends from western Montana down into southern Utah (Figure 1.6) (Snoke and Chapman, 2021). The northern margin is defined by the Lewis and Clark fault zone, a middle to late Cenozoic right-lateral transfer fault, which accommodated extension as the subplate moved to the west (Sears and Fritz, 1998).

As a result of the twofold crustal extension within the Basin and Range province, there is sufficient heat flow and permeability to form geothermal features throughout this region. As evidence for this, the Great Basin is already a large producer of geothermal energy and the

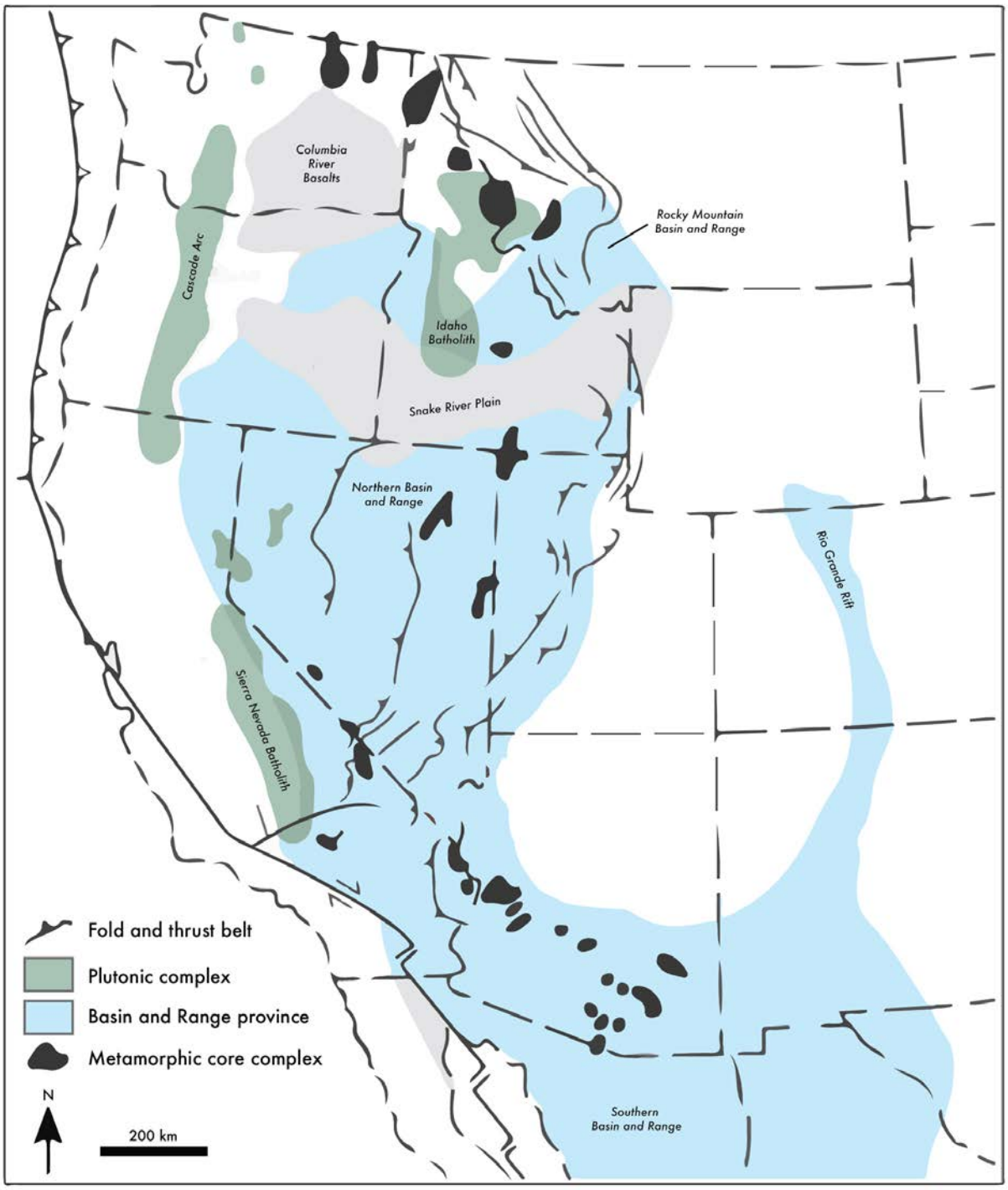

Figure 1.7: Map of the western United States showing significant tectonic features and the subregions of the Basin and Range Province. Southwest Montana is part of the Rocky Mountain section (modified from Newmann, 2019; Whitney et al., 2013).

broader province as a whole has strong potential for continued development (Faulds et al., 2021). Given its similar structural style to the Great Basin, it is hypothesized that the geothermal systems of southwest Montana are also related to recent Basin and Range style faulting. However, despite its similar structural setting, southwest Montana has received comparatively little geothermal resource attention. For instance, previous assessments of geothermal potential by the Department of Energy (DOE) have only focused on the states of Nevada, Utah, Idaho, Washington, and Hawaii (Faulds et al., 2021).

Structural Settings

As discussed, extensional tectonics lead to regions of anomalous heat flow, but this only fulfills one of the three key components required for geothermal activity. In addition to a high heat gradient, another critical factor for geothermal production is permeability, which is controlled by structural complexity in Basin and Range settings (Faulds and Hinz, 2015). Structurally complex systems, i.e. dense networks of intersecting faults and fractures, lead to enhanced permeability and thereby facilitate upwelling (Siler et al., 2019). This phenomenon is particularly prominent in areas where fault segments are not continuous, i.e. structural discontinuities such as fault terminations, step-overs, intersections, and other fault interaction areas (Faulds et al., 2020; Faulds & Hinz, 2015; Siler et al., 2018). Since structural complexity increases the potential for hydrothermal circulation, it therefore increases the potential for production (Faulds et al., 2019; Siler et al., 2018). Although faults and fault interaction density are key drivers for structural controls on geothermal systems, not all areas that meet these criteria result in significant hydrothermal fluid flow (Siler and Pepin, 2021). This is particularly important for traditional geothermal operations, as fluid must flow at a sufficient rate in order for it to be sustainably extracted (Jolie et al., 2021).

Another important variable governing fluid flow in geothermal systems is fault age. Younger aged (i.e. Quaternary) faults are more likely to circulate fluids, as they aren't yet

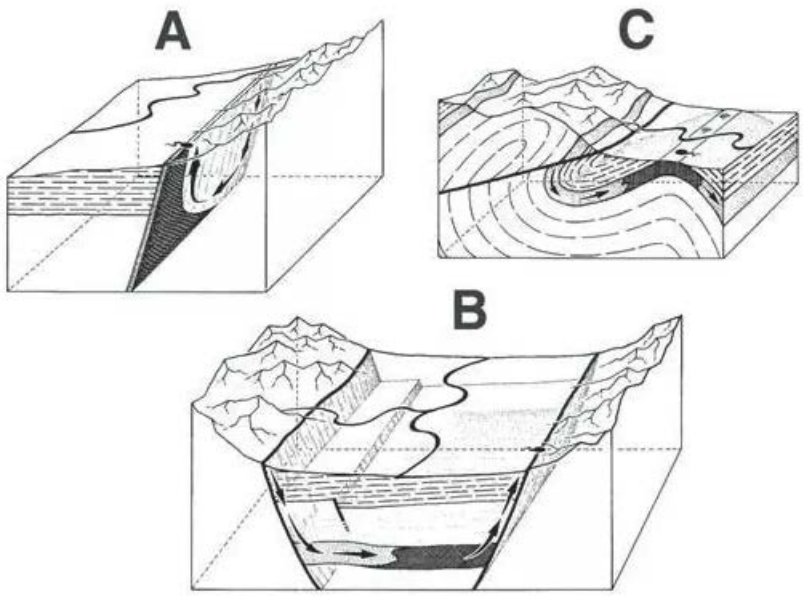

Figure 1.8: Conceptual examples of structurally controlled geothermal systems in amagmatic regimes (Sorey et al.,1983).

sealed by secondary mineralization, which restricts permeable pathways and inhibits fluid flow (Siler et al., 2018). Considering the influence of fault timing and fault connectivity on hydrothermal upwelling, it is unsurprising that overlapping and/or intersecting Quaternary faults are the most frequent structural setting of geothermal systems catalogued within the Great Basin (Faulds and Hinz, 2015).

Given the correlation between structural complexity and geothermal systems, it is important to constrain structural settings to gain insight into fluid flow (Siler and Faulds, 2013). Previous work has applied 3D modeling, drones, and traditional field mapping to constrain structural settings at various geothermal systems (Siler et al., 2019; Walter et al., 2020). Others have identified fault intersection control through traditional structural analysis of faults and shallow meter temperature surveying of thermal anomalies (Molisee and Bell, 2015). A question, however, is whether increased thermal coverage from drone surveying can more reliably identify structural controls. Therefore, a key part of this study will use

a machine learning model to investigate if UAS-imaged thermal properties are correlative with faulting in the area.

Study Location: Norris Hot Springs

Norris Hot Springs is a geothermal site developed for recreational use that is located 45 minutes southwest of Bozeman near Norris, MT in Madison County (Figure 1.9). We selected this geothermal system as our case study area for two distinct reasons. The primary motivation being it shows a clear hydrothermal surface expression that is rooted within a deeper geothermal system. Secondly, existing geologic mapping of the Norris Quadrangle indicated the potential of structural controls on the system via normal faulting (Figure 1.10).

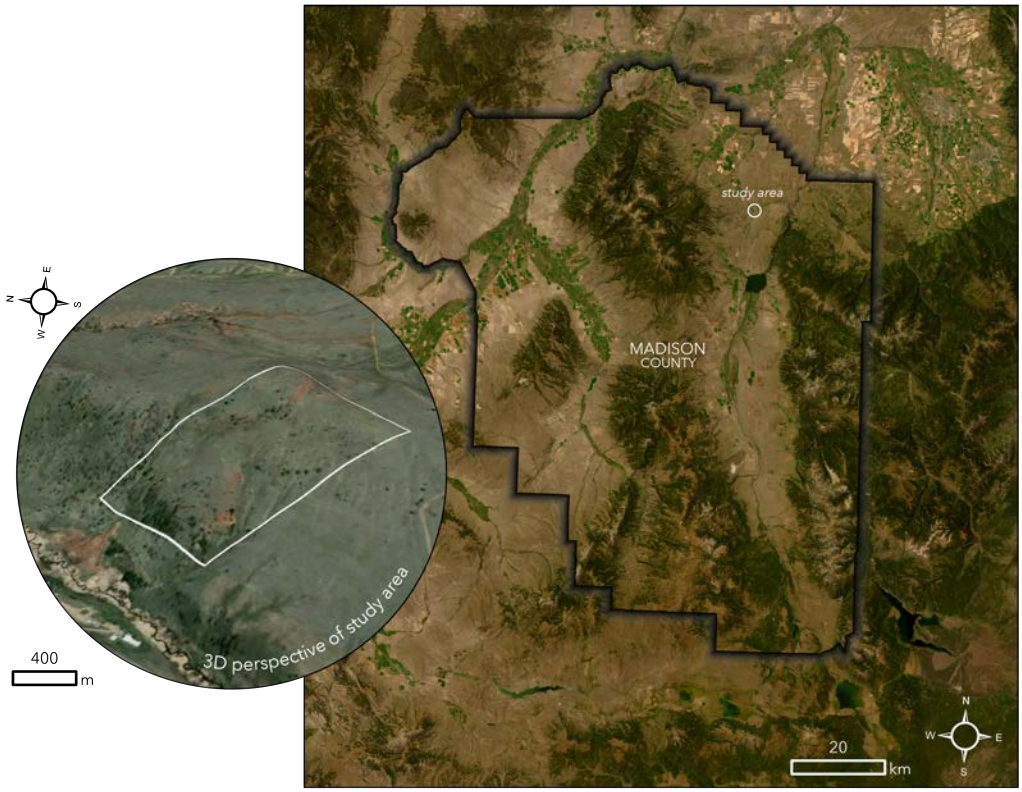

Figure 1.9: Satellite imagery of Madison County featuring the study locality of Norris, MT. *Inset:* 3D perspective of study area.

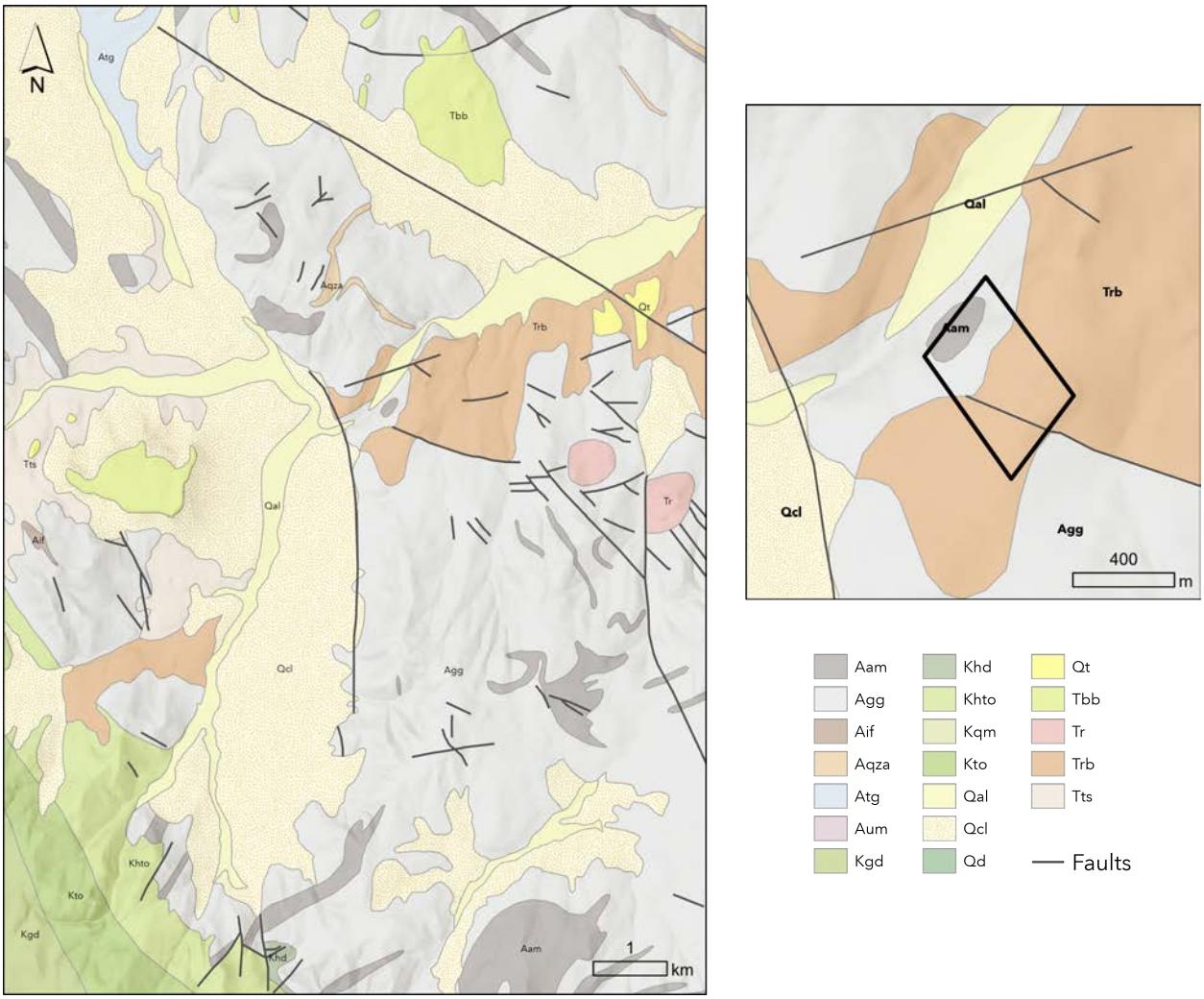

Figure 1.10: *Left:* Digitized geologic map of the USGS Norris Quadrangle. *Right:* Close up of study area and immediate surroundings.

METHODS

Lab Experiments

Acquisition

For the thermal mapping component of this research, we used a radiometric thermal camera designed for drone compatibility - the FLIR Duo Pro R (FDPR). This camera has two individual lenses that capture optical and thermal images at the same time, although at different resolutions (Table 2.1). The data are stored as a composite natural color image and a composite false color thermal image. The natural color image contains three bands - red, green, blue - while the false color thermal image contains four bands - red, green, blue, and a thermal radiance (LWIR) band as the fourth (Figure 2.1).

Specification	**FDPR value**
Thermal imager	Uncooled vox microbolometer
Spectral band	5.5 - 13.5 μ m
Thermal sensitivity	$<$ 50 mK
Thermal sensor resolution	640 x 512
Thermal lens	13 mm: 45°x 37°
Thermal frame rate	30 Hz
Visible sensor resolution	4000 x 3000
Visible camera FOV	56°x 45°
Measurement accuracy	$\pm$ 5°C or 5% of readings in the -25°C to 135°C range

Table 2.1: The manufacturer specifications of the FLIR Duo Pro R dual optical and radiometric thermal camera.

The camera's internal detector is an uncooled microbolometer which, as previously discussed, may yield an inconsistent real-time response when the camera is subject to changes in ambient temperature (Nugent and Shaw, 2014). Since ambient temperatures are likely to fluctuate during UAS missions, it is important to understand the capabilities of these types of thermal sensors, specifically in a field setting. While previous studies have evaluated

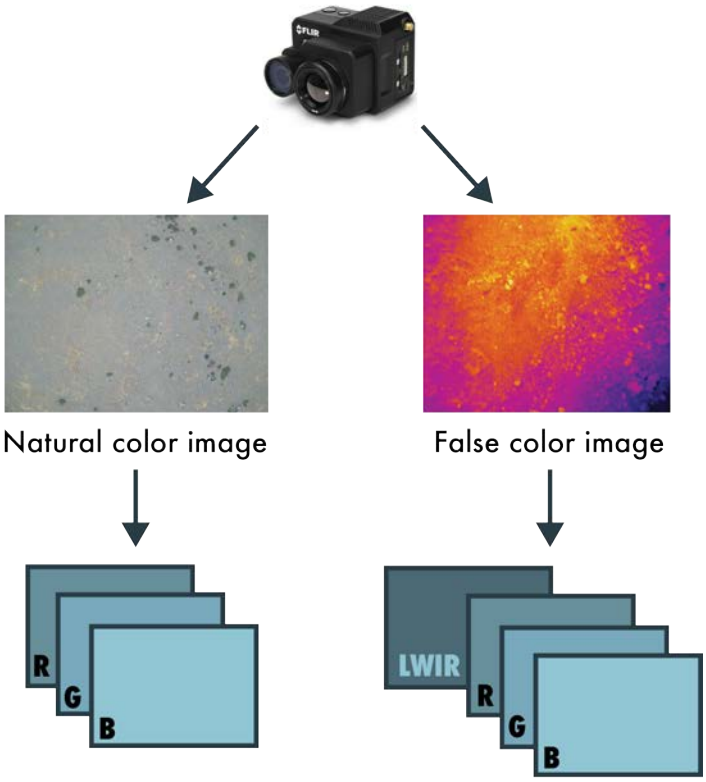

Figure 2.1: Schematic representation of the data captured and stored in one RJPEG by the FDPR. The simultaneous acquisition of a natural color image and a thermal image yields a total of seven bands available for post-processing.

other thermal cameras with uncooled microbolometers, the sensitivity is specific to each camera model. Therefore, to assess the calibration accuracy of our camera, the FDPR, and its response to varying environmental conditions, we conducted four experiments in a controlled laboratory setting in the Electrical & Engineering Department at Montana State University.

Each of the four tests involved a blackbody unit, which is an excellent and necessary tool for the radiometric calibration of thermal sensors. Conceptually, a perfect blackbody absorbs all incoming radiation and then re-emits 100% of that radiation (Campbell and Wynne, 2011). While a perfect blackbody does not exist in the real world, manufactured

blackbody units have an emissivity as close to 1 as possible, where emissivity describes the ratio between the emitted energy of an object and the emitted energy of a perfect blackbody under the same parameters (Campbell and Wynne, 2011). For our experiments, we treated the blackbody unit in the lab as a perfect emitter with an emissivity value of one.

The premise of all four experiments involved setting the blackbody unit to a known temperature and then evaluating the camera's ability to measure the established temperature. For each test we positioned the camera approximately 0.5 meters from the blackbody unit and surrounded the camera with insulating foam to shield the camera and blackbody from one another (Figure 2.2). We oriented the camera 90° from normal for stable positioning and streamed the camera feed to a display monitor to visualize the thermal imagery in real-time. We also enabled the spot meter, which is a FLIR feature that calculates the mean temperature of a 2x2 pixel array at the center of the camera's FOV.

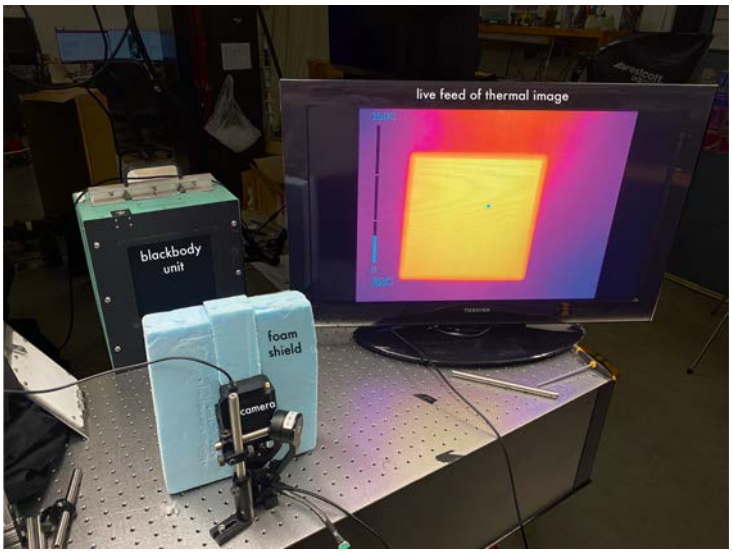

Figure 2.2: Experimental setup that served as the basis for each lab test.

We captured images through the FLIR UAS app, which allows the user to set values for the following radiometric parameters: sky condition, humidity, air temperature, emissivity,

and subject distance. These values are later used for the conversion of digital numbers to temperature. We maintained the same app parameters for all four experiments: clear skies, low humidity, 20°C, 1, and 1m, respectively. The app also permits the user to select between one of two file types for storing images – either an RJPEG (radiometric JPEG) or a combined TIFF & JPEG package. The RJPEG contains both the optical and thermal images within one single file as a packaged 8-bit visible JPEG and 14-bit thermal TIFF. The TIFF contains raw thermal data that represent absolute radiance, which can be used to calculate temperature during post-processing. The other file option, the partitioned JPEG & TIFF combination, contains a TIFF whose thermal data is pre-processed according to the radiometric parameters set in the app at the time of acquisition.

There are advantages and disadvantages to choosing either file type. The RJPEG format is proprietary to FLIR and thus not easily read by most external software, but the benefit is that the user is able to adjust the radiometric parameters in post-processing, which is not possible for the TIFF & JPEG combination. For all of our lab experiments, we collected images as the combined TIFF & JPEG file format to ensure easy readability of the data. However, for one experiment, we collected a set of RJPEG images in addition to the aforementioned file type. Collecting both file formats during the same experiment generated two directly comparable datasets for further analysis of the difference, if any, between the two configurations.

The four experiments, outlined below, tested the accuracy of the camera in stable ambient conditions as well as changing ambient conditions. The stabilization and increment tests explored the performance of the camera in a controlled environment whereas the wind and heat experiments assessed whether the camera's measurements are influenced by external factors. We ran each test independently of the others - that is, the results of one experiment did not influence the experimental design of another.

Stabilization Previous studies indicate that thermal cameras with uncooled microbolometers require a stabilization period upon initial powering before the camera is able to record valid data (Aragon et al., 2020; Kelly et al., 2019). Accordingly, we tested the performance of the FDPR upon immediate activation, without any prior use. We set the blackbody to 25°C and registered a room temperature of 22.4°C using a Kestrel weather meter at the beginning of the experiment. Immediately after activating the camera, we recorded images for two full hours, captured at an automatic 30-second interval.

Increment This experimental series tested the camera's accuracy when recording thermal properties across a range of object temperatures. Considering the findings of Kelly et al., 2019, we allowed the camera (and blackbody unit) to stabilize for 30 minutes prior to taking any measurements. The Kestrel meter registered a room temperature of 22.6°C at the beginning of the experiment. We started the test with a blackbody temperature of 20°C, followed by 25°C, 30°C, 35°C, and lastly, 40°C. After each 5°C increase in the blackbody temperature, we waited 15 minutes before recording new measurements, to allow the blackbody unit to fully reach its new temperature setting. As a secondary objective of this experiment, we captured each image in both of the file formats previously discussed, with the end goal of determining the optimal file type for future data acquisition during field surveys. At each temperature stage we collected a set of ten TIFF & JPEG images immediately followed by a second set of ten RJPEG images. Each set of photos automatically triggered at a 10-second interval.

Wind For the wind test, we simulated wind acting on the camera body to evaluate the camera's sensitivity to potential wind exposure during drone flights. We maintained a blackbody temperature of 25°C for the duration of the test and captured images at a 10-second interval, for a total period of 90 minutes. The Kestrel weather meter registered a room temperature of 23.1°C at the beginning of the experiment. At the start of the test

we aimed a source of cool air at the camera exterior for 30 minutes (Figure 2.3). When placed directly between the wind and the camera, the Kestrel meter registered a wind speed of 12 m/s. When placed downwind of the camera, it registered a wind speed of 2.5 m/s. To measure the camera body temperature throughout the experiment, we attached a thermocouple at two separate contact points on the camera body and manually recorded the camera's temperature at 5-minute intervals. After a period of 30 minutes, we removed the wind simulation and continued capturing images and recording the camera body temperature to assess the FDPR's response, post-wind exposure.

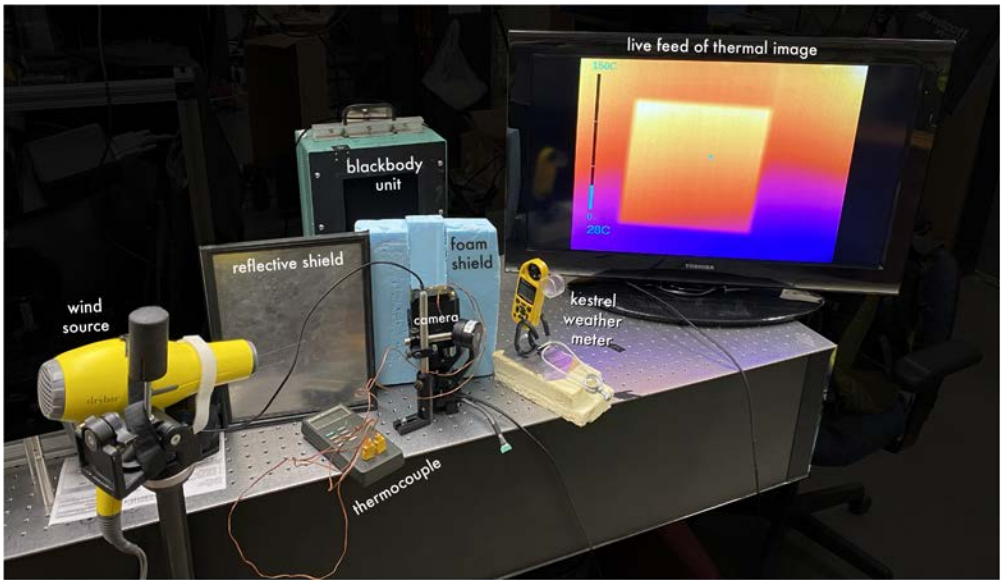

Figure 2.3: The experimental set-up for the wind test.

Heat For the final experiment, we set the blackbody to 25°C, consistent with the stabilization test and the wind test. To simulate an increase in ambient temperature, we aimed a halogen lamp at the camera, which increased the temperature of the camera body through heat exposure. We placed reflective shields between the heat source and the blackbody unit to target the heat solely on the camera and to prevent any incidental heating

to the blackbody surface. This test ran for a duration of 8 minutes and the photos triggered at a 10-second interval.

Processing

After acquiring data from all four lab experiments, we processed the raw thermal images. While it is relatively straightforward to obtain the raw thermal values from the TIFF & JPEG combination, it is difficult to obtain them from the RJPEGs since this file type is proprietary to FLIR. Although FLIR offers a software (FLIR Tools) for reading the images, it essentially operates as a blackbox, meaning that the user uploads images and the software outputs temperature values without providing any information regarding the methods used to convert the data. In this case, the program only displays the converted temperature values and does not reveal the raw DNs, which are necessary for a direct comparison of the two file types. To circumvent this limitation, we extracted the raw thermal data from the RJPEG images by running ExifTool (version 12.44) at the command line (Harvey, n.d.). After extracting the raw thermal data from both file formats, we used MATLAB R2022a in conjunction with ExifTool to retrieve the exchangeable image file format (EXIF) metadata from the RJPEGs. These tags are included in the captured images and contain pertinent information such as the specifications of the camera model, time stamps, GPS positions, etc.

To exclude the margins of the blackbody unit from our analysis, we cropped each image to a standardized region of interest (ROI) that focused on the actual blackbody material (Figure 2.4). Since the camera and blackbody positions remained constant throughout all four experiments, the ROI is standardized to the same pixel array for every single image. To convert the raw DNs to temperature values, we used two distinct equations: Equation 2.1 for the TIFF files and Equation 2.2 for the RJPEG files (Irujo, 2022). We obtained the parameter values for Equation 2.2 from the aforementioned EXIF metadata extracted from the RJPEGs. These included the sensor's gain, offset, and calibration parameters, which

are necessary to convert raw sensor counts to temperature values (Irujo, 2022). We also modified Equation 2.2 to yield an answer in Celsius, rather than Kelvin.

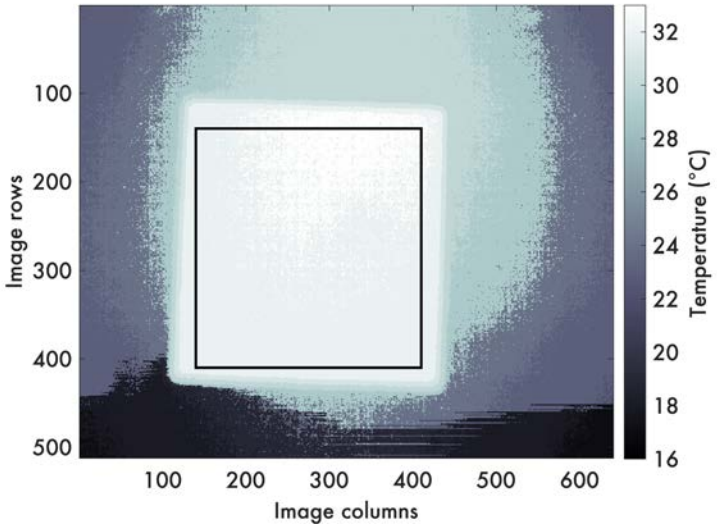

Figure 2.4: Example image of the blackbody unit featuring the ROI (outlined by black square) used to calculate the lab results presented in this paper.

$$T_C = 0.04 * S - 273.15, \tag{2.1}$$

$$T_C = \frac{B}{\ln(\frac{R_1}{R_2(S-O)} + F)} - 273.15, \tag{2.2}$$

Symbol	**Parameter**
T_C	Temperature (Celsius)
B	Planck B constant
R_1	Planck R1 constant
R_2	Planck R2 constant
S	Digital number
O	Planck O constant
F	Planck F constant

Field Work

The Study Area

We selected an area adjacent to and immediately upslope of Norris Hot Springs for our case study. Here, we established a study boundary totaling 0.1 km^2 in size (Figure 2.5). This extent served as the testing area for our data collection framework: an early morning thermal flight immediately followed by ground temperature and thermal conductivity measurements. We established the sampling locations for the ground measurements within the study area by randomly designating 30 point features with the *Create Random Points* geoprocessing tool in ArcGIS Pro, version 2.8.2. These points eventually served as the training sites for our machine learning model.

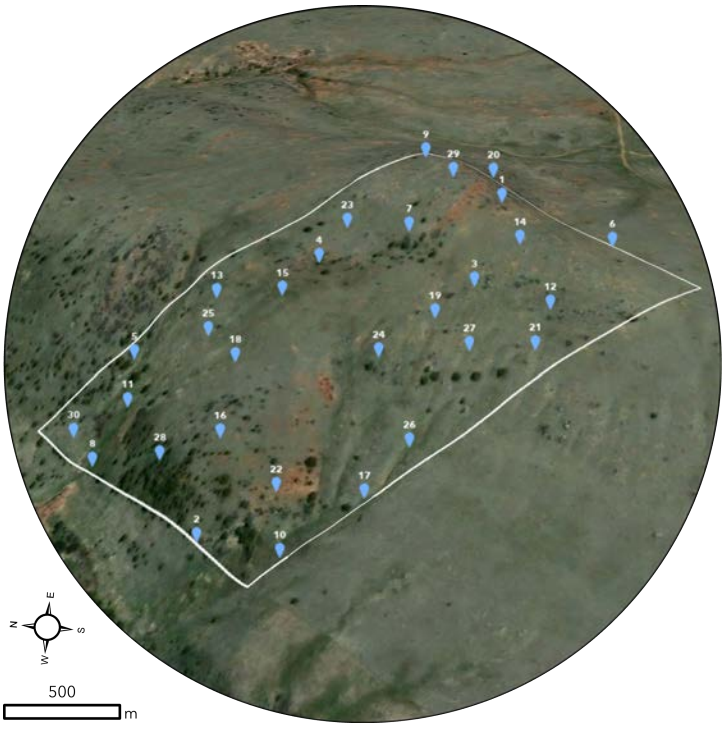

Figure 2.5: 3D perspective of the study area, featuring the 30 randomly distributed sampling locations.

To find the real-world ground locations of each point feature, we surveyed the site at the beginning of the field season with an Emlid Reach GNSS (global navigation satellite system) receiver, which offers centimeter-scale precision. After surveying a given sampling site, we marked the ground with a pinflag, survey stake, and survey whisker to enhance the visibility of each location. Staking the sampling locations expedited the data collection workflow for future field work. We needed to make the workflow as efficient as possible due to the time-sensitive nature of thermal mapping - the best time of day to conduct a UAS flight targeting thermal anomalies is in the early morning, before the sun rises and starts to warm the ground. However, at the time of this research, Part 107a of the Code of Federal Regulations (CFR) did not permit takeoff until 30 minutes prior to sunrise, for safety concerns regarding visibility (Title 14, CFR, Part 107, 2021). Since these limitations constrained the window for thermal mapping on both ends, staking the sampling locations expedited the post-flight in situ data collection.

UAS Surveying

For our thermal surveys, our equipment included the aforementioned FDPR radiometric thermal camera as well as the VisionAerial SwitchBlade-Elite tricopter UAV (Figure 2.6). We created flight plans with the QGroundControl flight planning software, which syncs with the drone's handheld controller. Before beginning a mission, we distributed four ground control points (GCPs) within the study area at sampling locations previously surveyed with the Emlid receiver (Figure 2.5). We utilized bucket lids covered with aluminum foil as our GCPs. By covering each GCP with foil we decreased its emissivity, thereby increasing its reflectance, and decreasing its thermal signature in the flight imagery. Manipulating the surface material of each GCP enabled us to clearly differentiate them from their surroundings during the georeferencing process.

Although we attempted numerous field days over a wide time period, the majority of

Figure 2.6: *Left:* GCP covered in aluminum foil, overlying one of the sampling locations. *Right:* The VisionAerial SwitchBlade-Elite tricopter UAV with the FDPR mounted on the gimbal.

our attempts were unsuccessful due to uncontrollable obstacles such as changing weather patterns and equipment issues. One key issue that quickly emerged is that Norris, MT is located in a wind corridor, which tended to create prohibitive flying conditions even if the daily forecast predicted calm weather. We eventually succeeded in acquiring a full dataset (both flight mission and ground sampling) on August 2nd, 2022. The weather conditions on August 2nd were overcast and cloudy, with a haze of smoke from regional wildfires. Temperatures hovered around 20°C. Although we observed a light breeze interspersed by stronger gusts, it was still possible to fly safely.

For the August 2nd mission, we placed the GCPs at the following locations: 01, 18, 23, and 24. Although this isn't the ideal spatial distribution for typical GCP placement, our choices were limited due to the large expanse and steep terrain of the field area as well as the previously discussed time constraints. However, the GCP distribution is a non-issue because we also employed a real-time kinematic (RTK) receiver during all missions. RTK

is a newer technology that improves GNSS accuracy during UAS missions as the receiver communicates with the drone to perform GPS corrections during flight. Some argue that it even renders GCPs negligible (Štroner et al., 2020). We decided to use the RTK receiver in conjunction with GCPs in this study for maximum positioning accuracy. Therefore, our GCPs still augmented the positional accuracy of our flight data, despite their placement locations.

Since thermal images have a lower spatial resolution compared to optical images, it is imperative to operate a UAS mission at a high forward- and side-overlap to ensure that there will be enough tie points for photogrammetric processing. Accordingly, we flew the August 2nd mission with 80% overlap (Figure 2.7. We kept most of the default radiometric parameters in the FLIR app, but changed the sky conditions to cloudy and the subject distance to 140 meters to match the above ground level (AGL) flight height of the mission. Due to the limited battery capabilities of the UAV, we changed the battery three separate times during flight, each time returning to the landing pad before resuming the flight pattern where it previously left off. The total mission time lasted approximately one hour. After collecting the thermal imagery, we immediately collected the in situ field data.

Post-Flight Ground Sampling

Within the detailed study area next to Norris Hot Springs, 30 random sampling points served as training sites for our machine learning model. We collected multiple sets of ground-based measurements at each sampling location. These measurements included thermal conductivity and temperature, which we collected immediately post-flight in order to inventory the ground's thermal properties at the same time as the thermal flight. We recorded measurements at every site except 05, 12, and 30, whose locations coincide with boulders.

To measure the temperature and thermal conductivity of the ground, we used a

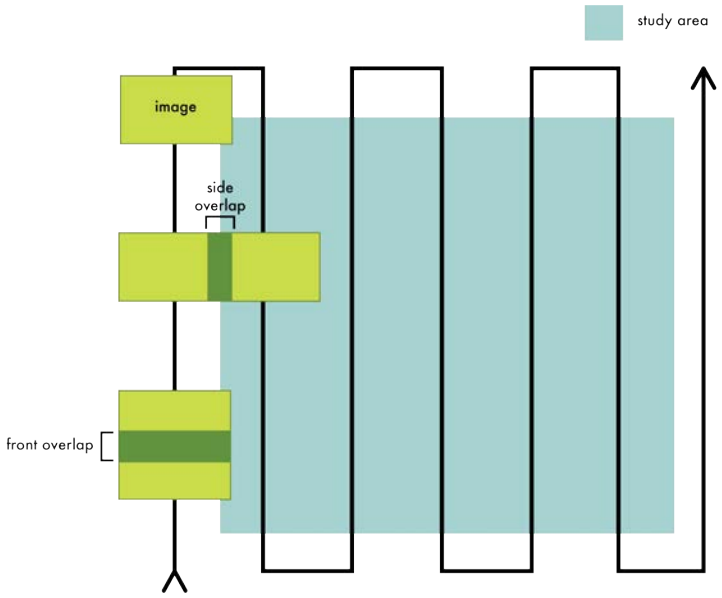

Figure 2.7: A schematic of the grid pattern flown during a UAS flight, depicting the front- and side-overlap parameters.

Thermtest Transient Line Source TLS-100 portable meter. We insterted the sensor needle within a 10 cm radius of the survey stake, without making contact between the two pieces of equipment. We allowed the probe to isothermally stabilize with the ground before recording any measurements. Once the temperature stopped fluctuating, we recorded it in a field notebook and proceeded to run the conductivity test. The conductivity test works by delivering a current to the ground surrounding the probe and then measuring the rise in temperature over a period of time ("TLS-100 Transient Line Source Thermal Conductivity Meter", n.d.). At the conclusion of the conductivity test, we recorded the data and repeated the process at the next sampling location, until we finished acquiring data at all 27 sites.

Soil Collection and Analysis As part of our field data collection, we also conducted a one-time soil survey for a sieve-based grain size analysis. These data served as an additional set of training data for our machine learning model. We collected these samples on November 11th, 2021. We used an auger to extract a core of soil at each sampling location, with the

average core depth equaling 12.9 cm. After allowing the soil samples to dry for a few weeks, we conducted the grain size analysis using a W.S. Tyler RO-TAP RX-29 sieve shaker (Figure 2.8). Our sieve pan stack consisted of four sieve classes: granule and larger, very coarse and coarse sand, medium, fine and very fine sand, and coarse silt and smaller (Figure 2.9). We lumped multiple phi sizes into a given class due to limited sieve availability in the lab. (Figure 2.9).

Figure 2.8: The stack of sieves secured within the sieve shaker.

We started each analysis by taring the Mettler Toledo PB3001 balance with a paper weighing boat. We then transferred the respective soil sample to the boat and weighed the entire sample. At the beginning of the sieving process and between every sample we cleaned each sieve. This included visually inspecting the sieves for trapped particles and clearing any trapped debris with a brush (either soft or hard bristle) or a metal pick. Next, we poured the entire sample into the coarsest sieve at the top of the sieve stack. We secured the stack in the machine, and then shook the sample for 4-5 minutes to sort the grains into their respective classes (Figures 2.8 and 2.9). After removing the entire stack, we individually weighed the contents of each pan. Starting with the coarsest sieve, we transferred any remaining material to the paper boat and weighed the amount of soil corresponding to the

respective phi class(es). After weighing, we visually inspected the sieve for trapped particles and cleared any trapped debris. We repeated this process for each successive sieve pan. After weighing the material from each sieve in the stack, we reassembled the entire sample and recorded its weight, post-sieving. We then repeated this entire process for each of the soil samples.

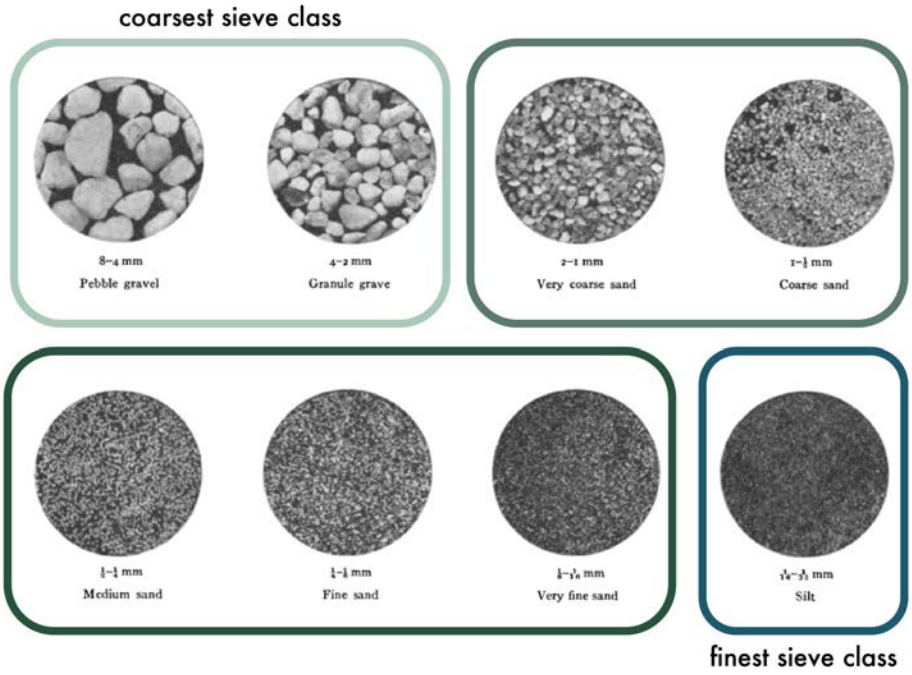

Figure 2.9: The four sieve classes used in this study, from coarsest to finest: a) granule and larger, b) very coarse and coarse sand, medium, c) fine and very fine sand, and d) coarse silt and smaller.

Computational Modeling

Photogrammetry

After capturing images with a UAS, the typical objective is to stitch the photos together to create an orthorectified mosaic, commonly shortened to orthomosaic. This is achieved through the process of photogrammetry, a technique that calculates 3D geometric

information from individual photos, corrects image distortions, and places images into real-world locations. Once the photos are geometrically corrected, calibrated, and georeferenced, they are stitched together to form one comprehensive orthomosaic.

As previously mentioned, the FDPR simultaneously captures optical and thermal imagery. While a plethora of information exists regarding the photogrammetric processing of optical imagery, a similarly robust body of knowledge does not yet exist for the processing of thermal imagery, due to the novelty of UAS TIR remote sensing. As such, mosaicking our thermal data proved to be challenging.

We acquired the August 2nd flight data in the RJPEG format. We used the natural color images to construct the natural color orthomosaic and used the false color thermal images to construct false color mosaics representing relative thermal patterns (Figure 2.1). These thermal mosaics also included the LWIR band as a fourth channel. Given the multidimensionality of the images captured by the FDPR, we approached the mosaicking process from a few angles, which are explained in detail later in this section.

We began our initial attempts at photogrammetry with Pix4Dmapper, version 4.8.1. Despite significant effort, numerous trials, and its ability to read RJPEG images, Pix4D could not stitch all of the thermal data from our August 2nd flight. This was surprising, considering the high forward- and side-overlap of our flight pattern. For our second attempt, we used DroneDeploy's cloud-based solution for processing thermal flights. In this scenario, the user uploads thermal images to the DroneDeploy website and the company produces a mosaic available for download. While DroneDeploy successfully constructed a false color mosaic with the thermal RGB bands, it dropped the fourth thermal band (LWIR) during processing. Since the LWIR band is essential for this research, this solution did not suffice for our needs. Additionally, the blackbox nature of this method left a lot of unknowns regarding the processing steps involved.

For our final attempts, we switched to Agisoft Metashape, version 2.0.0. We successfully

aligned our entire dataset upon first trial, unlike Pix4D. To better understand the Metashape software, as well as the thermal mosaicking process more broadly, we decided to thoroughly explore multiple processing methods and their associated results. Accordingly, we tested three different mosaicking techniques within Metashape. For each of the three methods, we created thermal mosaics of the study area from the imagery collected on August 2nd, 2023. The three methods are discussed in the following order: RJPEG blending, TIFF blending, and a swath-based approach loosely inspired by Malbéteau et al., 2021.

For all three methods, we trimmed the entire dataset prior to any alignment efforts by removing any images captured while the drone flew to or from the landing location. This ensured we only used photos explicitly captured along the individual flight paths, at the correct flight height, for any reconstructions. After trimming the dataset, we followed the baseline thermal imagery processing guide from Metashape (Agisoft LLC, 2023). Our general workflow for each method followed these steps: imagery upload, alignment, GCP georeferencing, DEM construction, orthomosaic generation, and lastly, thermal visualization.

We uploaded the imagery to the software with the 'multi-camera system' option to accommodate the dual optical and thermal imagery captured by our camera. We calibrated the focal length and pixel size parameters in Metashape to match those of our specific camera model: 13 mm and 0.017 x 0.017mm, respectively. We also locked the focal length parameter to prevent the software from changing the value during processing (Agisoft LLC, 2023).

After calibrating the camera parameters, we aligned the photos. During the alignment stage, we tested various limits for the key point and tie point thresholds. Key points are identifiable features within the images that the program is able to detect algorithmically. Setting the threshold for this task limits the number of features detected in each photo - without a limit, unreliable features may be included. We settled on a 40,000 key point limit. Key points that are identified and matched across multiple images are called tie points. For this one, we chose a limit of 10,000. We also selected the highest accuracy option for this

step. After the alignment step, we incorporated the GCPs for georeferencing.

Once we aligned and georeferenced the dataset, we built the DEM. We set the depth filtering to mild, which ensured that important features were not scrapped as outliers. After building the DEM, we created the orthomosaic. Since multiple UAS photos coincide with a given pixel location, the user must choose a blending mode that dictates what pixel value will represent that specific location in the constructed mosaic. Metashape offers three types of blending modes: *mosaic*, *average*, and *disabled* (Figure 2.10). The default mode, *mosaic*, compartmentalizes the data into low and high frequency domains (Agisoft LLC, 2023). The *average* mode calculates the weighted average of all overlapping pixels. Lastly, the *disabled* mode uses the pixel value from the image with the best resolution and camera view closest to normal. We retained the default blending mode of *mosaic* and projected the mosaic surface to the DEM built in the previous step.

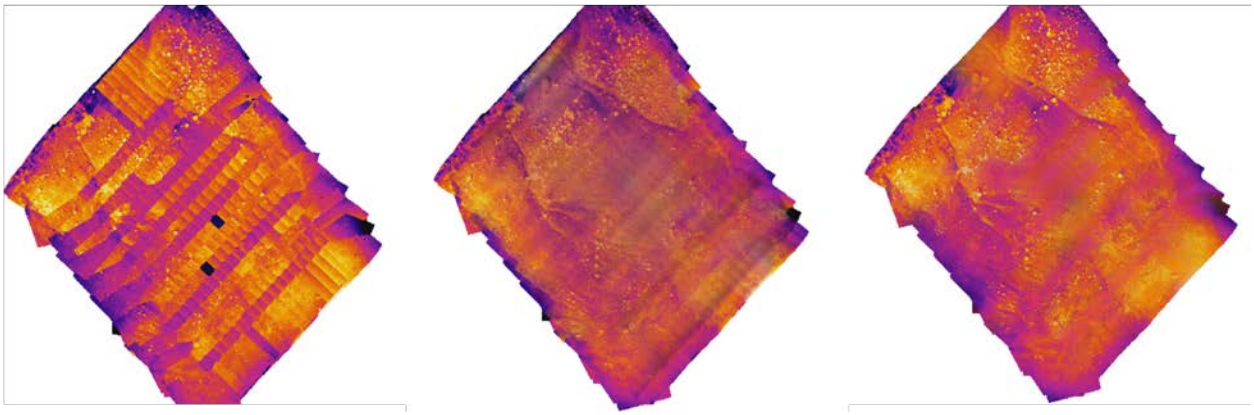

Figure 2.10: False color mosaics showing each of the three blending modes available in Agisoft Metashape. From left to right: disabled, average, mosaic. We generated the three mosaics from the RJPEG images (n=396) from all flight lines (n=19).

Next, we improved the positional accuracy of the mosaic. Although our original mission included RTK correction, we further refined the model by re-adjusting the mosaic to each of the four GCPs (Figure 2.11). We reduced positional error by manually adjusting marker locations within individual photos, beginning with the photos associated with the highest

error. Lastly, we used the raster calculator to change the mosaic display to Band 4, to visualize the LWIR band instead of the composite false color infrared image. After finishing the Metashape processing workflow, we exported the four-band mosaic as an uncompressed TIFF and moved the data into the ArcGIS Pro and MATLAB environments for analysis and interpretation.

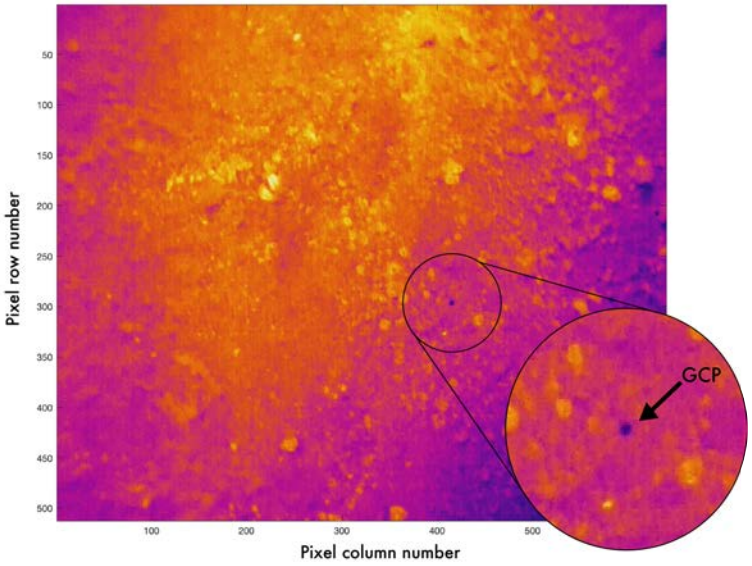

Figure 2.11: Image of a raw RJPEG from the August 2nd flight, displayed as the false color composite. *Inset*: A GCP is clearly visible in the thermal image due to the contrast between the low emissivity of the aluminum foil covering the GCP surface vs. its surroundings.

Thermal Mosaic Generation: RJPEG Blending This method included all flight lines (n=19) from the August 2nd mission. Although we successfully generated a thermal mosaic of four bands (RGB + LWIR) with this method, we noticed that Metashape altered the digital numbers associated with the LWIR band, through some unknown transformation. Despite our efforts to figure out the applied conversion, we did not find any answers in the Metashape manual, published documentation, or online help forums. Therefore, we decided to test a second method where we extracted the 14-bit TIFFs from the RJPEGs prior to

uploading them into a new Metashape project, following the troubleshooting suggestions in an alternate workflow (JK REF).

Thermal Mosaic Generation: TIFF Blending For this method, we started the mosaicking process with TIFFs extracted from the original RJPEGs for all 19 flight lines. We copied the original RJPEGs to a new folder and then extracted the 14-bit thermal TIFF from each copied RJPEG by running Exiftool at the command line. After this, we uploaded the newly extracted TIFFs to Metashape and followed the same steps outlined in Method #1, with the exception of the GCP step. When we extracted the TIFFs, the GPS information disappeared from each photo, which meant that Metashape could not place the images into their real-world locations. However, we still generated the final mosaic and then manually georeferenced it in ArcGIS Pro by using identifiable features as control points.

Thermal Mosaic Generation: Alternating Swaths For both of the previous methods, we used all available flight lines (n=19) for each of the mosaicking processes. However, we observed a distinct striping pattern in the rendered mosaics, which we hypothesize to be an artifact from wind acting on the camera lens during certain flight lines, as observed by Malbéteau et al., 2021. To explore whether UAS flight direction impacted the thermal imagery of a given flight line, thereby impacting the final mosaic, we tested our third Metashape method. For this technique, we intentionally excluded certain flight lines during the mosaicking process. We used the selection tool in Metashape to select all the cameras from the first flight line and subsequently disabled them. We then systematically disabled all images from all odd-numbered flight lines. After selectively removing the odd-numbered flight lines, we constructed a mosaic with only even-numbered flight lines, following the same workflow outlined in Method #1. We repeated this process for the even-numbered flight lines to produce a mosaic with only odd-numbered flight lines. Even though we excluded roughly half of our original flight data from each process, we were still able to stitch the

dataset because we prioritized an 80% overlap during the original mission. Removing the alternating flight lines reduced the spatial overlap between flight lines to 60%, which was still sufficient.

Thermal Mosaic Generation: Pix4D For this method, we followed the RJPEG processing workflow provided by Pix4D ("Processing thermal images", 2018). Like the Metashape methods, we used the trimmed dataset (n=396). Per the suggested workflow, we used the 'thermal camera' template for the initial processing step. We then updated the camera calibration values to reflect the focal length and pixel size of our camera. During our first attempts we kept most of the suggested parameters, but continued to encounter calibration issues that led to processing failures. After many iterations trying different settings, we reached small success by modifying the internal parameter setting to 'all prior' and the calibration method to 'standard,' rather than the suggested method of 'alternative.' Next, we ran the program's second and third processing steps to build an index map, which is the Pix4D term for the mosaicked LWIR band. The software uses the FLIR SDK to convert the raw digital numbers to temperature values, so the index map only contains the converted temperature values, not the raw sensor counts.

Training Data Rasters

Along with the data we collected at each of the 30 randomly sampling points, we expanded our explanatory variables dataset for our machine learning model by creating surface rasters in ArcGIS Pro. These rasters included classified land cover, solar insolation, aspect, and slope.

Land Cover Classification In addition to recording the land cover type at each sampling location at the beginning of the field season, we ran a supervised classification in ArcGIS Pro to systematically classify land cover across the entire field area. Classification is a data

categorization technique where the user provides a model with predetermined classes and training samples and then the algorithm assigns labels to unclassified observations based on statistical similarities to the training samples. For our analysis, we chose a supervised rather than unsupervised classification due to its unique advantage that feature labels are created by the user prior to running the model.

We used our natural color orthomosaic for the input data for the classification. We designated a schema of various land cover classes (water, vegetation, rocks, and barren) and created training samples for each member. We fed these data to the classification algorithm and the tool assigned each of the pixels to one of the designated classes, based on the spectral response pattern of each pixel relative to that of the land cover classes. We ran multiple iterations of the tool, tweaking the training samples until reaching a satisfactory final result.

Solar Radiation We also created a raster of the solar insolation acting on the study area on August 2nd, 2023. Using the *Area Solar Radiation* tool in ArcGIS Pro, we generated 12 rasters - one for every 30 minute interval from 6am to 12pm, corresponding to the time period of our data acquisition in the field. Each raster represents the amount of solar radiation (in WH/m^2) reaching the study area, which spatially varies across the site due to the relationship between the sun angle and the topography.

Modeling

For the modeling phase of this project, we used two approaches: linear regression and machine learning. To test the relationship between ground-sampled temperature and remotely sensed temperature, we ran simple linear regression models for each of the four thermal mosaics created in Metashape, as well as the partial thermal mosaic created in Pix4D. In each model, we tested the relationship between the August 2nd in situ temperature and the thermal data recorded in each of the five rasters. We ran all five regression models in MATLAB.

To investigate whether machine learning can identify the surficial factors diagnostic of thermal patterns observed in UAS TIR imagery, we iteratively ran the Forest-based Classification and Regression tool in ArcGIS Pro. Broadly speaking, there are two phases involved with a Random Forest machine learning model: training and prediction. The training phase explores the relationship between the explanatory variables and the variable to predict. In this study, the explanatory variables consisted of our ground-sampled data such as temperature, thermal conductivity, and soil grain sizes, in addition to raster data such as aspect, slope, solar radiation, dominant soil order, geologic unit, and distance to faults. Our variables to predict included the five thermal mosaics as well as the August 2nd in situ temperature data. Thus, we tested five different prediction variables and treated each one as an independent model.

Prior to modeling, we calculated the mean mosaic value at each sampling location, for each of the five mosaics. We used the *Buffer* tool to create a two meter buffer around each sampling point. We chose a two meter buffer to overcome any inherent spatial error by ensuring overlap between the ground-surveyed sites and the mosaic. We then proceeded to calculate the mean value (either in DN or temperature) within each buffer zone using the *Zonal statistics* tool. This gave us the corresponding values for our predictor variable at each sampling location. We repeated this process for each mosaic and then joined the results to the sampling points feature layer.

Before training any of the models, we set a random seed (57 ACM599) in the tool's environment settings to standardize the inherent randomness between runs, allowing for a direct comparison of performance metrics between model iterations and across the five independent models. For each model, we adhered to the following procedure: to assess the initial stability of the model, we began the training phase with multiple validation runs per attempt. Multiple validations enabled us to examine the distribution of variable importance across multiple iterations of a given model. A larger range in boxplot distributions

indicates that variable importance fluctuates from run to run, suggesting a less stable model. Conversely, smaller boxplot ranges indicate a more stable model that performs consistently across all runs. Therefore, assessing the variable importance boxplot provides key information about the model stability and the influence of hyperparameter values on model performance, before fully training the model.

During the initial stages of training the model, we did not exclude any of the training data. This decision allowed us to iteratively fine tune the model's hyperparameters, including the number of trees and maximum tree depth. While finessing the model, we continued to assess its stability through the variable importance boxplot. After determining the best hyperparameter values, we then excluded 10% of the training data since it is standard practice to withhold a portion of the training data to use for validation purposes. We also reduced the number of validation runs to one. Once we modified the amount of withheld training data and the number of validation runs, we proceeded to run the model again to test its accuracy. The resulting metrics described the overall performance of the model, i.e. its ability to explain the variation in thermal patterns based on our explanatory variables.

RESULTS

Lab Experiments

For all four lab experiments, we calculated the following metrics to analyze our data: mean temperature, standard deviation, and mean absolute error (MAE). We chose the latter two metrics for a few reasons. The standard deviation describes the dispersion of the individual measurements around the mean temperature, providing information about the variability of the data recorded by the camera. The mean absolute error describes the difference between the measured values and the expected value (the blackbody temperature), which reflects the accuracy of the radiometric calibration of the camera. Thus, both statistical measures provide valuable insights regarding the baseline performance of the camera.

Increment

For the increment test, we had two objectives: a) to assess the FDPR's ability to measure a range of object temperatures and b) to determine the optimal file type for data acquisition. We set the blackbody to five temperature settings: 20, 25, 30, 35, and 40°C. At each temperature stage we acquired 20 images: 10 TIFFs and 10 RJPEGs. We calculated the mean temperature, standard deviation, and mean absolute error for each of the ten images per file type, at each temperature stage (Table 3.1). The results varied across the blackbody temperature increments (Figures 3.1 and 3.2). The camera overestimated blackbody temperatures at 25, 30, 35, and 40°C, but underestimated temperatures at 20°C (Table 3.1). All of the standard deviations were below 1°C and the highest value (0.9208) was associated with the 20°C RJPEG measurements.

The results also varied depending on the file format. At blackbody temperatures of 30 and 40°C, the TIFF and RJPEG files recorded similar values (Table 3.1). Yet at 20, 25, and

35°C, we observed a discrepancy between the temperatures recorded by the two file types. For four of the five increments, the RJPEG files had a lower MAE than did the TIFF files. This supports the finding that the mean RJPEG temperatures more closely align with the actual blackbody temperatures across the temperature stages (Table 3.1 and Figure 3.2).

	TIFF			RJPEG		
Blackbody temperature (°C)	Mean temperature (°C)	MAE	STD	Mean temperature (°C)	MAE	STD
20	14.930	-5.062	0.1276	19.2449	-0.7551	0.9208
25	30.330	5.330	0.1222	27.7229	2.7229	0.1555
30	31.914	1.914	0.1165	32.0465	2.0465	0.0833
35	41.146	6.146	0.0572	36.5220	1.522	0.0672
40	41.206	1.206	0.0853	40.9621	0.9621	0.0427

Table 3.1: The mean temperature, mean absolute error, and standard deviation calculated at each of the five blackbody settings, for both of the file formats (n=100).

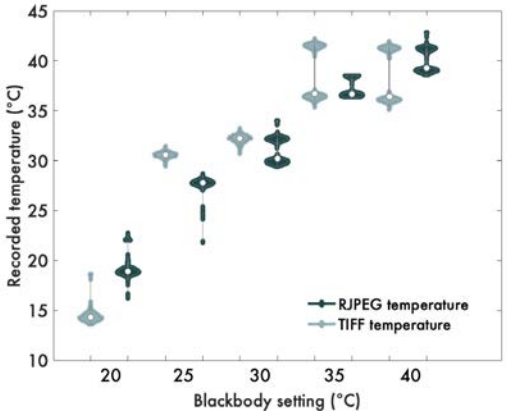

Figure 3.1: Violin plot showing the distribution of temperature measurements from each of the 10 images per file type, at each blackbody increment. (n=100).

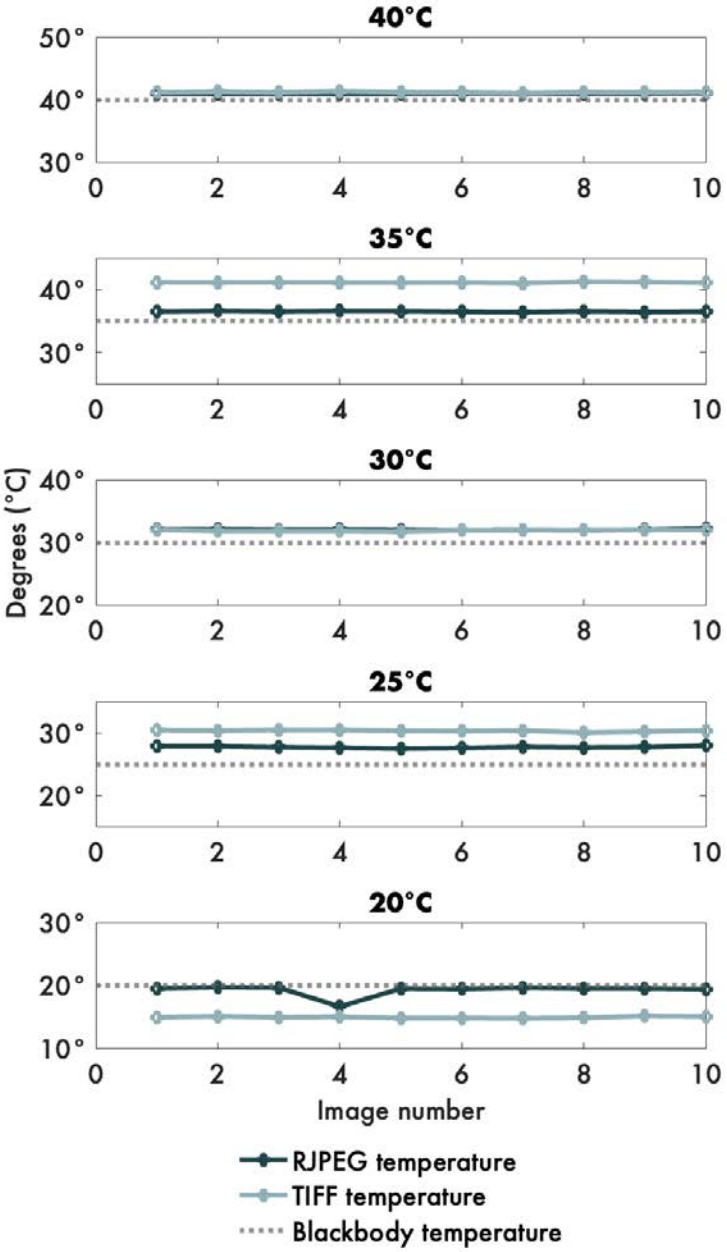

Figure 3.2: Mean temperature per image, per file type, at each blackbody increment (n=100).

Stabilization

For the stabilization experiment, we tested the FDPR's ability to record consistent measurements immediately after powering on. We maintained a blackbody temperature of 25°C for the entire experiment, capturing a total of 241 images over a period of 120 minutes. We calculated a mean temperature of 31.9011 ± 0.2051°C. Our results do not show large temperature fluctuations after turning on the camera, which is evidenced by the small standard deviation (Figure 3.3). Although the standard deviation was small, the MAE was 6.9011°C, which indicates an overestimation of the blackbody temperature by almost 7°C.

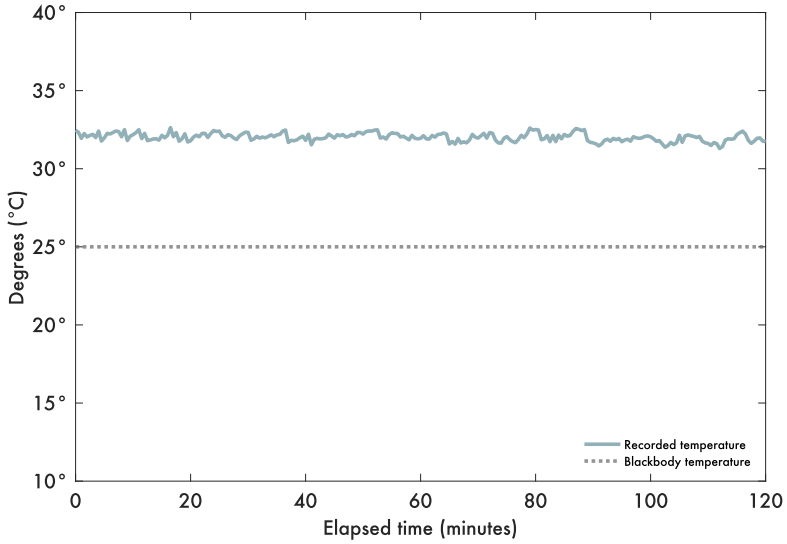

Figure 3.3: Results from the stabilization experiment. Temperatures recorded by the FDPR compared to the established blackbody temperature of 25°C (n=241).

Wind

In the wind experiment, we tested the FDPR's measurement sensitivity when exposed to wind, i.e. cooler ambient temperatures. We subjected the camera to a simulated source of wind for the first 30 minutes, followed by a period of 60 minutes without wind exposure. We maintained a blackbody temperature of 25°C for the duration of the experiment and

captured 499 images over a period of 90 minutes.

For the first 30 minutes, we calculated a mean temperature of 16.5892 ± 2.62°C and an MAE of 8.4108°C. This error suggests that when the camera is exposed to wind, it underestimates object temperatures of 25°C by roughly 8°C. Additionally, the recorded temperatures were highly variable for the first 15 minutes of the test before stabilizing (Figure 3.4). Immediately following the removal of the wind source at the 30 minute mark, the temperature measurements steeply increased. At this transition, the camera recorded a temperature increase of nearly 16°C, despite the blackbody remaining constant at 25°C. During the final 60 minutes of the experiment, corresponding to the recovery period, we recorded a mean temperature of 31.0060 ± 0.2890°C and an MAE of 6.006°C. This error suggests that under steady state conditions the camera overestimates object temperatures of 25°C by 6°C, which is relatively consistent with the error calculated in the stabilization experiment.

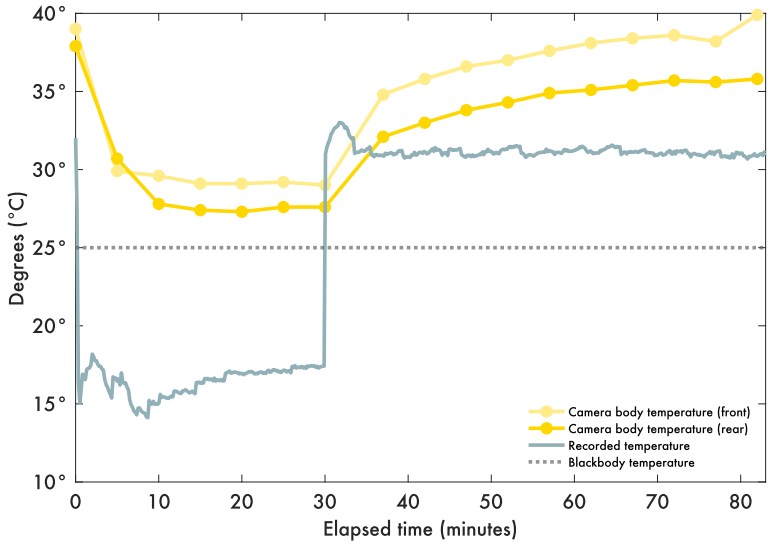

Figure 3.4: Results from the wind experiment, featuring temperatures recorded by the FDPR and camera body temperatures measured with the thermocouple. We removed the wind source at the 30 minute mark; n=182 for the wind exposure phase and n=317 for the post-wind phase.

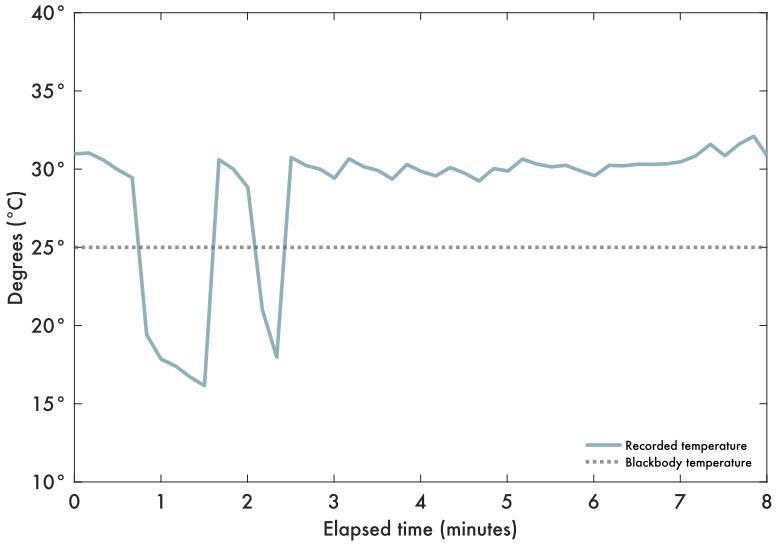

Figure 3.5: Results from the heat experiment, featuring temperatures recorded by the FDPR compared to the known blackbody temperature of 25°C (n=49).

Heat

For the heat experiment, we exposed the camera body to a halogen lamp to simulate an increase in ambient temperature. We maintained a blackbody temperature of 25°C and collected 49 images over a period of 8 minutes. Upon exposure to the heat source, the camera recorded large fluctuations for the first three minutes, followed by relative stabilization for the remainder of the test (Figure 3.5). The mean temperature calculated for the entire suite of images was 28.4318 ± 4.22 and the MAE was 3.4318°C. The larger standard deviation provides evidence for the strong fluctuations in recorded temperatures.

Sensor Non-Uniformity

In addition to testing the FDPR's radiometric accuracy and sensitivity to environmental temperatures, we also explored sensor noise at the pixel-level. This included a vignetting analysis in MATLAB using the images from the stabilization experiment. Vignetting is a common phenomenon in optical and thermal cameras where there is a gradual reduction in

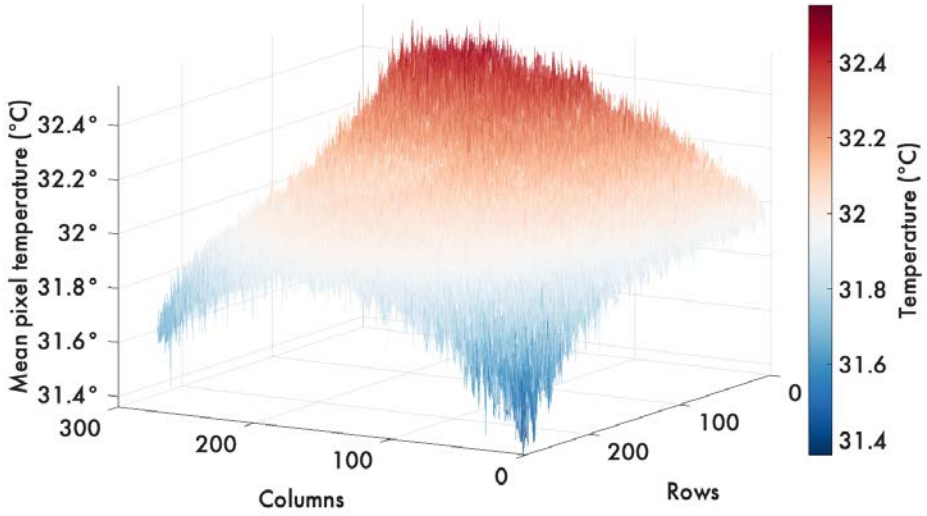

Figure 3.6: 3D surface rendering of the mean temperature value per pixel within the ROI (n=241).

pixel brightness or intensity from the center of the image out towards the periphery. This effect typically manifests as a radial pattern from the center of the image out towards the corners. To visualize the relative distribution of pixel accuracy across the detector array in the FDPR, we calculated the mean temperature at each pixel location for all 241 images (Figure 3.6).

Our results confirm that thermal images captured by the FDPR experience a subtle vignetting effect. The temperature values varied spatially across the ROI, with an approximate difference of 1°C from the warmest temperature to the coolest temperature. However, our results show a lateral, rather than radial, reduction in pixel temperatures across the image. This non-radial vignetting is similar to the findings of Smigaj et al., 2017, who suggest that this is the result of nonuniform degradation of the sensor over time.

Field Work

Ground-sampled Temperature

Immediately after the August 2nd thermal flight, we collected in situ temperature measurements at 27 of the 30 randomly designated sampling locations. The results are presented in Table 3.2. The values ranged from 19.3 to 24.5°C with a mean temperature of 22.5°C.

Soil Collection and Analysis

As part of our field data collection, we acquired soil samples at each of the 30 randomly designated sampling locations within our study area. We then analyzed these samples for their distribution of grain sizes and used these data as explanatory variables in our machine learning model. We could not collect soil samples at the exact coordinates for sites 05, 12, and 30 since they coincide with boulders. However, we collected samples immediately adjacent to each boulder to obtain a sample within the two meter buffer zones that we established in ArcGIS Pro. Additionally, we collected two samples at a few locations with limited soil cover to increase the chance that one of them contained a sufficient amount of material for sieving. These sites are denoted with either 'a' or 'b' appended to the end of their ID, indicating which of the two samples we used for the grain size analysis at the respective sampling site (Table 3.3).

The results from our analysis are are broken out by four phi groupings and are listed in Table 3.3. Thirteen of the samples displayed a symmetrical distribution, with the two intermediate phi classes representing the largest percentage of the total sample. However, there were 17 samples with a positively skewed distribution, where the largest phi class (granule and larger) comprised the majority of the bulk material.

Sample ID	**Ground-sampled temperature (°C)**
1	24.3
2	21.4
3	24.5
4	21.7
6	24.0
7	22.7
8	19.6
9	22.6
10	21.3
11	19.3
13	21.2
14	22.7
15	21.8
16	24.1
17	23.0
18	21.7
19	23.0
20	23.8
21	24.4
22	22.9
23	22.2
24	23.4
25	20.8
26	21.2
27	23.7
28	22.4
29	24.1

Table 3.2: The in situ temperatures collected immediately after the August 2nd thermal flight. The sampling locations that coincided with boulders are not included.

	x > -1 ϕ	-1 ϕ > x > 1 ϕ	1 ϕ > x > 4 ϕ	4 ϕ > x
	granule & larger	very coarse sand & coarse sand	medium sand, fine sand, very fine sand	coarse silt & smaller
sample ID	**individual weight (%)**			
111121_01	27	40	29	2
111121_02	54	30	15	0
111121_03	24	36	38	2
111121_04	50	26	21	3
111121_05	31	33	31	5
111121_06	31	41	26	2
111121_08	57	26	16	1
111121_09a	2	27	46	16
111121_10	41	36	22	1
111121_11	46	37	16	1
111121_12	34	36	27	2
111121_13	23	36	36	4
111121_14a	39	34	24	3
111121_14b	48	27	21	5
111121_15	28	38	30	3
111121_16	48	33	18	2
111121_17	48	39	11	1
111121_18	52	32	14	1
111121_19	44	38	17	2
111121_20	35	34	27	5
111121_21	27	35	29	4
111121_22	37	38	24	2
111121_23	23	35	40	1
111121_24	41	30	26	3
111121_25	36	40	22	2
111121_26	34	44	21	1
111121_27	21	47	30	2
111121_28	81	13	6	1
111121_28b	58	32	8	1
111121_29	32	27	34	7
111121_30a	75	21	4	1
111121_30b	38	37	21	4

Table 3.3: Results of the soil grain size analysis, broken out by individual weight percent for each group of phi sizes.

Computational Modeling

Photogrammetry

One advantage of the FDPR is its ability to simultaneously capture optical and thermal imagery. Therefore, we were able to construct a high-resolution natural color orthomosaic and a digital surface model (DSM) of the study area from the dataset acquired on August 2nd (Figure 3.7). The orthomosaic served as the basis for our land cover classification and the DSM served as the input data for the slope, aspect, and solar radiation rasters that we created in ArcGIS Pro.

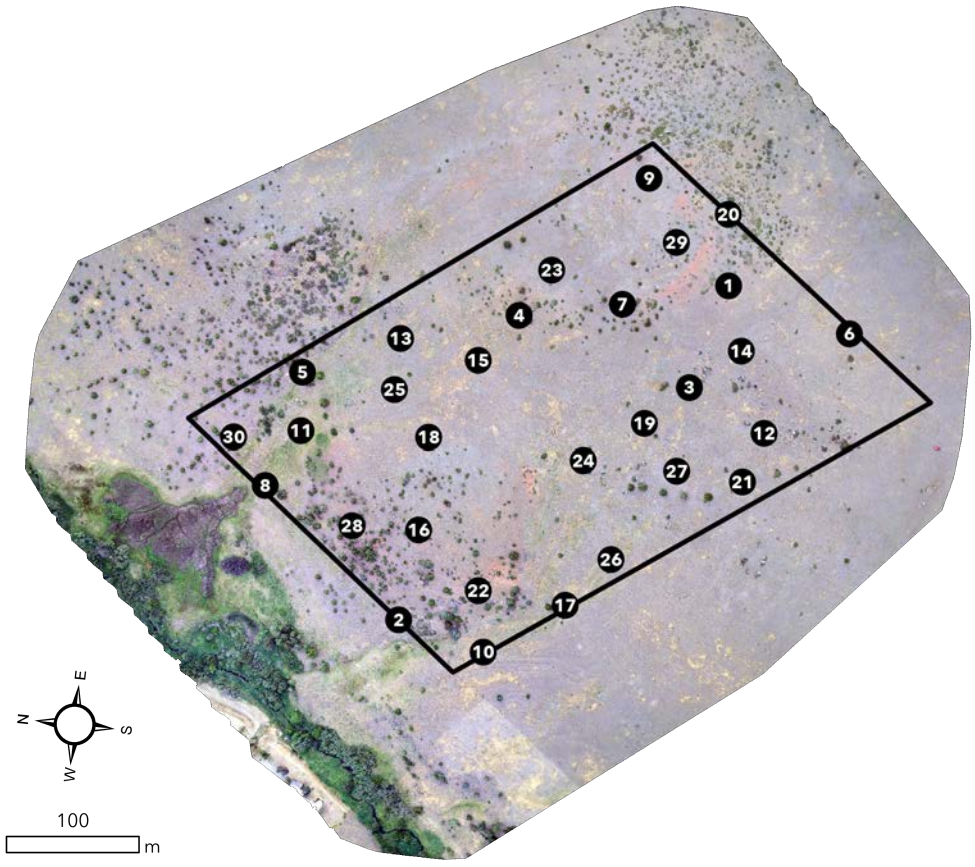

Figure 3.7: Natural color orthomosaic constructed from the August 2nd flight data, featuring the 30 sampling locations and official study area boundary.

In addition to the natural color orthomosaic, we generated five thermal mosaics using different processing methods - one in Pix4D and four in Metashape. We tested multiple techniques to determine the most accurate and spatially comprehensive method for the photogrammetric processing of our thermal data. We refer to the four methods as follows: Pix4D, RJPEG blending, TIFF blending, and lastly, the alternating swaths method. Although we constructed each of the five thermal mosaics using the same source data, the different processing methods yielded distinctly different thermal patterns in each mosaic. To assess the distribution of pixel values within each mosaic, we calculated descriptive statistics and plotted individual histograms for each dataset and (Table 3.4 and Figure 3.8). These metrics helped us evaluate and compare the results of the various processing methods against one another.

Processing method	Min. value	Max. value	Mean	STD	Kurtosis	Skewness
Raw images	1,034	5,119	2,955.23	485.9	3.87	0.179
Pix4D (°C)	-24.35	42.63	17.18	5.74	5.19	-0.42
RJPEG blending	25,906	31,182	29,111.14	484.69	7.06	-0.036
TIFF blending	2,078	3,920	2,983.03	230.49	4.10	-0.253
Odds mosaic	26,878	31,266	29,468.03	582.21	4.42	-0.064
Evens mosaic	25,136	29,907	28,414.46	641.94	3.00	-0.721

Table 3.4: Descriptive statistics for the August 2nd raw images and each of the five mosaics produced by the four different processing methods. All are reported in DNs except for the Pix4D method, which is reported in degrees (Celsius).

Raw Images The raw thermal data for all flight lines (prior to any photogrammetric processing) ranged from 1,034 to 5,119 (Table 3.4). When converted to absolute temperature via Equation 2.2, this range equals -2.77 to 90.96°C.

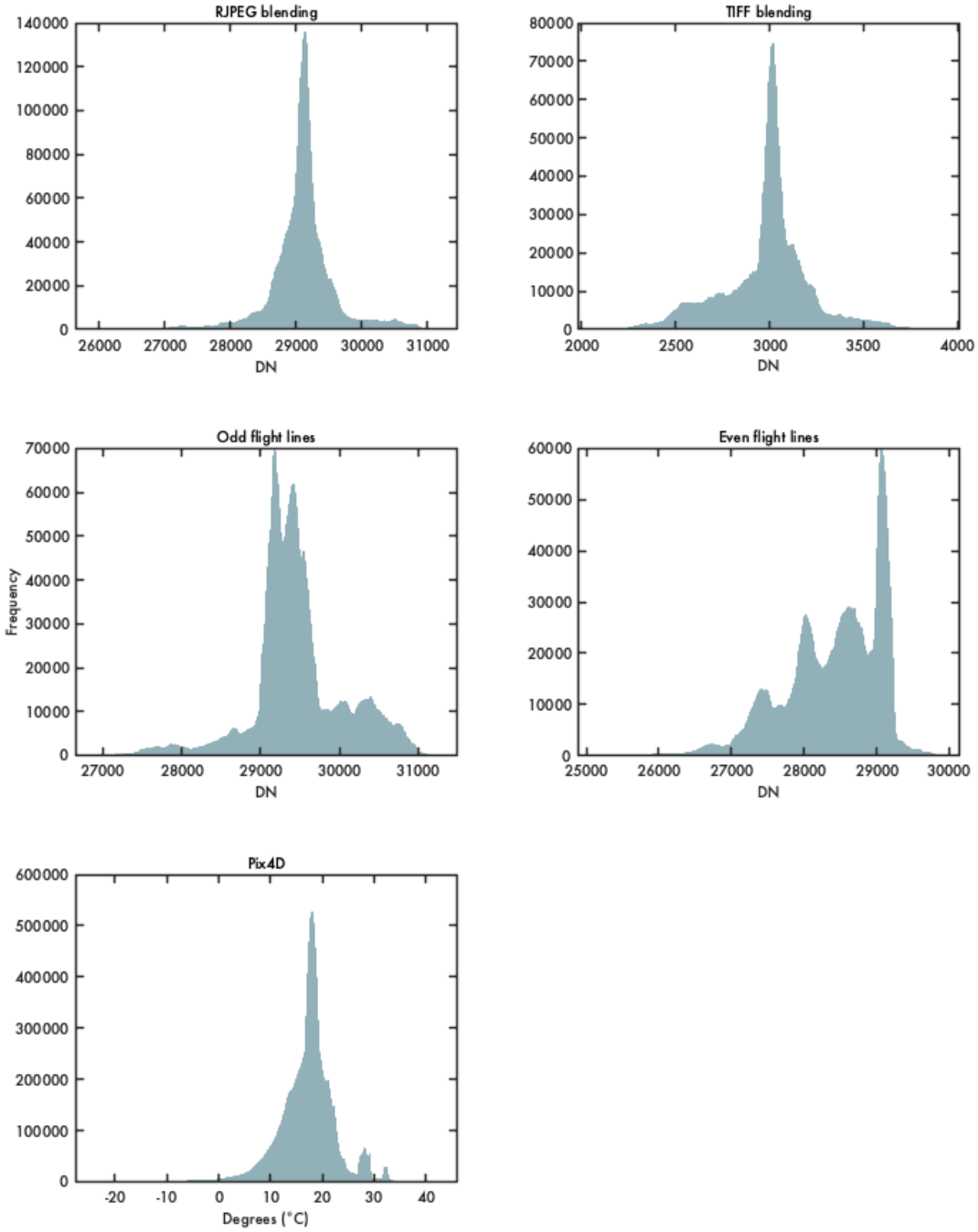

Figure 3.8: Histograms showing the shape and distribution of the pixel values in each of the five mosaics created from the four different processing methods.

Thermal Mosaic Generation: Pix4D The Pix4D method produced a thermal mosaic of absolute temperatures that ranged from -24.35 to 42.63°C (Table 3.4). This mosaic range is much cooler than the temperature range recorded in the raw images (-2.77 to 90.96°C). A minimium temperature of -24°C is also unrealistic, since we captured our data during the peak of summer.

Additionally, the spatial coverage of this mosaic is limited - it only covers two thirds of the original flight area. One of the weaknesses that we encountered with Pix4D is that it struggles to calibrate thermal images when the ground sampling distance (GSD) varies across the dataset. The steep topography of our field area and inability to implement terrain-following during flight made this variation unavoidable. It appears that the software was more successful at stitching the regions of larger GSD (the northwestern portion) as opposed to the regions of smaller GSD (the southeast corner). This makes sense, as spatial overlap increases with flight height and the northwestern region of our mission had a higher AGL flight height than did the southeastern part. Although Pix4D could not generate a full mosaic, we still included the partial mosaic for comparison since it retained the thermal values of the LWIR band (Figure 3.9).

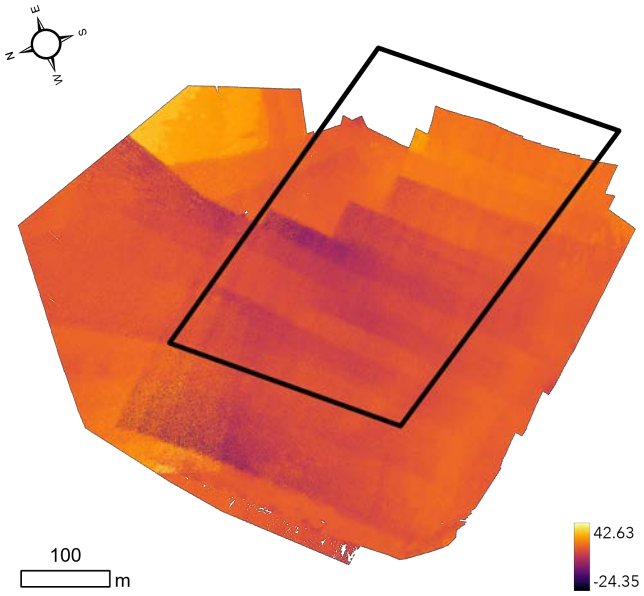

Figure 3.9: The partial thermal mosaic generated by Pix4D.

Thermal Mosaic Generation: RJPEG Blending Since we could not create a full coverage thermal mosaic with Pix4D, we employed the Agisoft Metashape software for our next attempt, the RJPEG processing method. In this method, we successfully stitched all of the raw data into a mosaic that covered the entire flight area (Figure 3.10). The resulting mosaic ranged in value from 25,906 to 31,182 (Table 3.4). However, as previously discussed, the Metashape software altered the LWIR DNs through an unspecified transformation, so it is unknown how these DNs relate to absolute temperature. It is suspected that Metashape applied a bit depth conversion and a non-linear histogram stretch, complicating the relationship between the original values and the output values (Figure 3.11). Despite this limitation, it is still possible to compare these results against those of the other Metashape mosaics, in terms of relative temperatures.

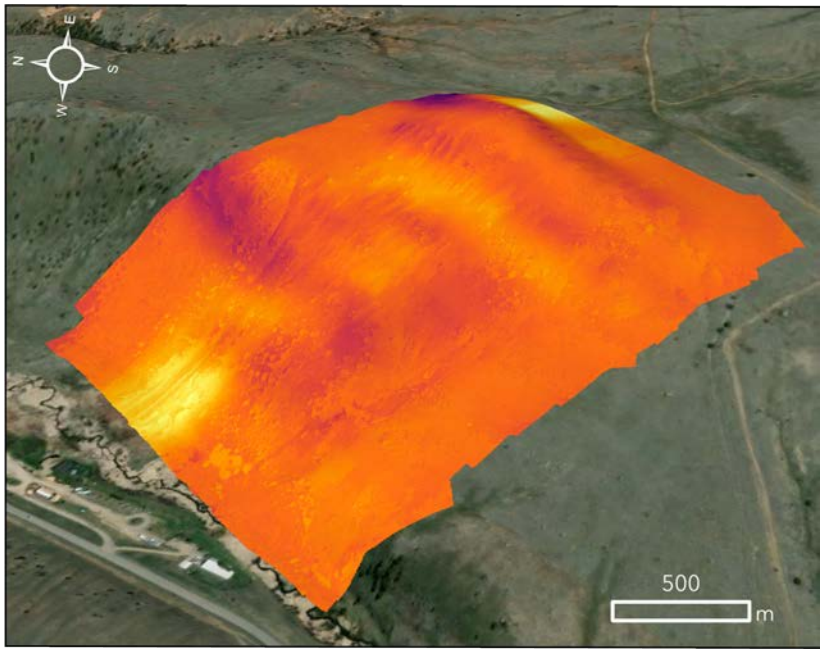

Figure 3.10: 3D perspective of the thermal mosaic created with the RJPEG blending method, using all flight lines (n=19).

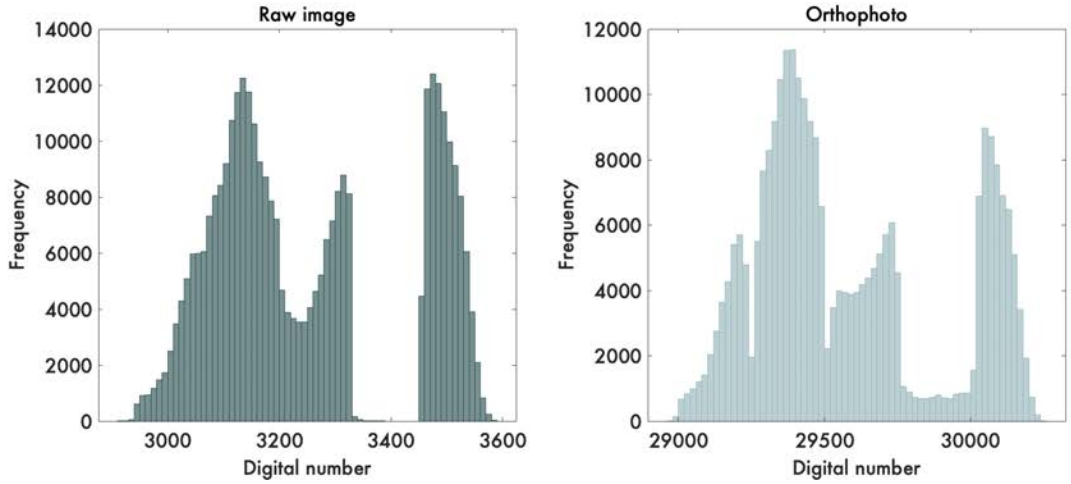

Figure 3.11: Comparison of the DN distribution between a raw thermal image from the August 2nd dataset and the corresponding orthophoto from Metashape.

Thermal Mosaic Generation: TIFF Blending Using the methods described in Section 2, we constructed a thermal mosaic from the TIFFs extracted from the original RJPEGs. There are a few important distinctions between the results of this method and the RJPEG blending method. Unlike the RJPEG method, this mosaic retained the original LWIR DNs, which can be related to absolute temperature via Equation 2.2. The raw DNs ranged from 2,078 to 3,920 and once converted to temperature, ranged from 32.32 to 71.48°C. An additional distinction between the results of the RJPEG and TIFF methods is the spatial coverage of each mosaic. The Metashape software was less successful at stitching together all of the TIFFs, which means that the areal extent of this mosaic is slightly smaller than that of the RJPEG blending method. This minor issue is most noticeable in the southeast corner of the mosaic (Figure 3.12).

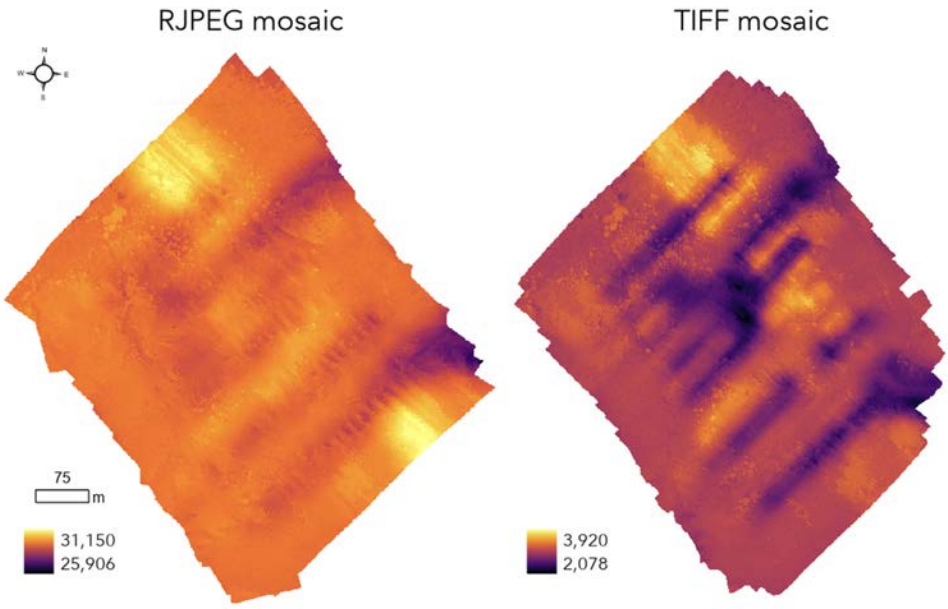

Figure 3.12: Comparison of the thermal mosaics constructed with the RJPEG blending method vs. the TIFF blending method. Both mosaics contain all flight lines (n=19). Note: the mosaics are not displayed with the same range of values.

Regarding thermal patterns, each mosaic has a northeast-southwest trending linear

pattern of alternating warmer and cooler temperatures (Figure 3.12). We suspect that this is an artifact related to the temperature sensitivity of the camera. As suggested by Malbéteau et al., 2021, it is not just the presence of wind that affects a thermal sensor, but rather the angular relationship between the wind direction and the camera. Therefore, we hypothesized that the linear artifact observed in our results is due to the interaction between the camera and wind, but only along certain flight lines. That is, we see a different response whether the UAS flew with a tailwind vs. a headwind along a given flight path. We tested whether we could reduce this noise by manipulating the processing technique, leading us to the alternating swaths method.

Thermal Mosaic Generation: Alternating Swaths For this method, we used the same processing approach to create two thermal mosaics: one comprised of odd-numbered flight lines (the 'Odd mosaic') and one comprised of even-numbered flight lines (the 'Even mosaic') (Figure 3.13). For this technique, we standardized each mosaic to one flight direction. By limiting the source data to a common flight direction, we investigated whether the resulting mosaic would have a more accurate thermal pattern and reflect the true thermal properties of the ground.

In theory, excluding certain flight lines (if the overlap is still sufficient) should not affect the thermal pattern in an orthomosaic. However, due to the wind sensitivity of the thermal camera, as evidenced in our lab experiments, we found that the mosaic results are distinctly different when comparing the one of odd flight lines vs. the one of even flight lines (Figure 3.13). In the Odd mosaic, the DNs ranged from 26,878 to 31,266 and in the Even mosaic, the DNs ranged from 25,136 to 29,907 (Table 3.4). The Even mosaic had a lower (i.e. cooler) range of values than did the Odd mosaic. This was expected because the Even mosaic contained the headwind flight lines whereas the Odd mosaic contained the tailwind flight lines. Morever, the distribution of pixel values in the Even mosaic had a negative

skewness of 0.72, indicating that there were more pixels containing cooler values than we would expect in a symmetric distribution (Table 3.4 and Figure 3.8).

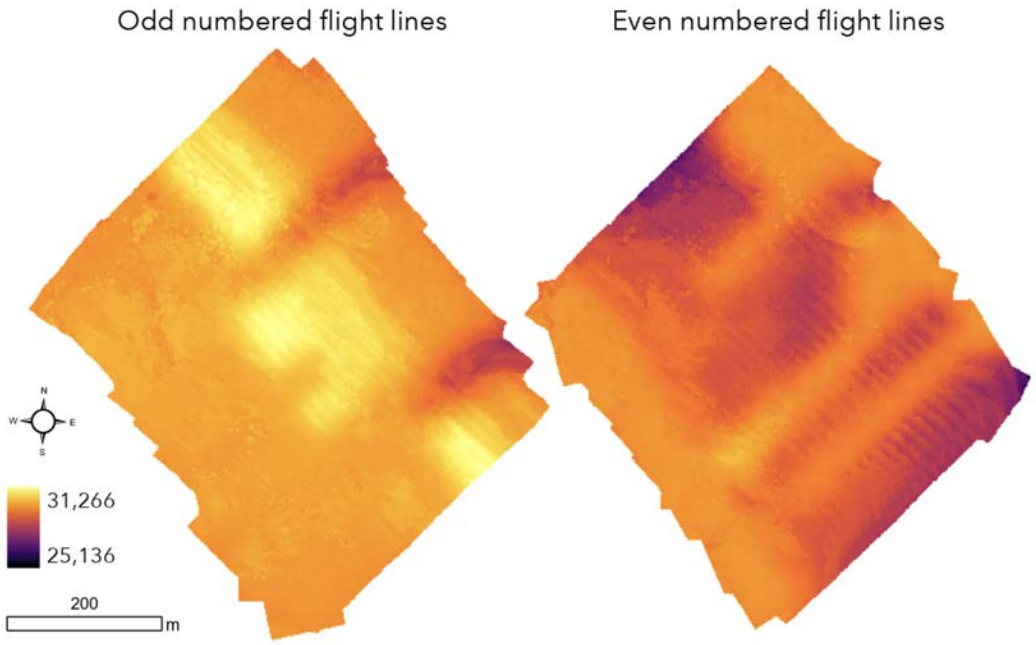

Figure 3.13: The two mosaics generated via the alternating swaths method. Both mosaics are displayed with the same DN range to facilitate direct comparison between the two.

To further explore the relationship between flight line direction and the camera measurements, we created a time series of the range of DNs captured in all raw images (n=396) from the August 2nd flight (Figure 3.14). With the progression of image number (and thereby, time), we observed trends of decreasing DNs for a sequence of images, followed by stable or increasing DNs for a subsequent set of images. Isolating the time series to just the headwind flight paths (the even numbered swaths) provides even more evidence for this pattern (Figure 3.15). Almost every single flight path headed into the wind corresponds to cooler temperatures recorded by the camera. Conversely, the time series that includes only the odd numbered flight lines (the tailwind swaths) shows stable or increasing temperatures for almost all of the flight lines (Figure 3.16). That being said, there are two instances of

low DNs near image numbers 20 and 145.

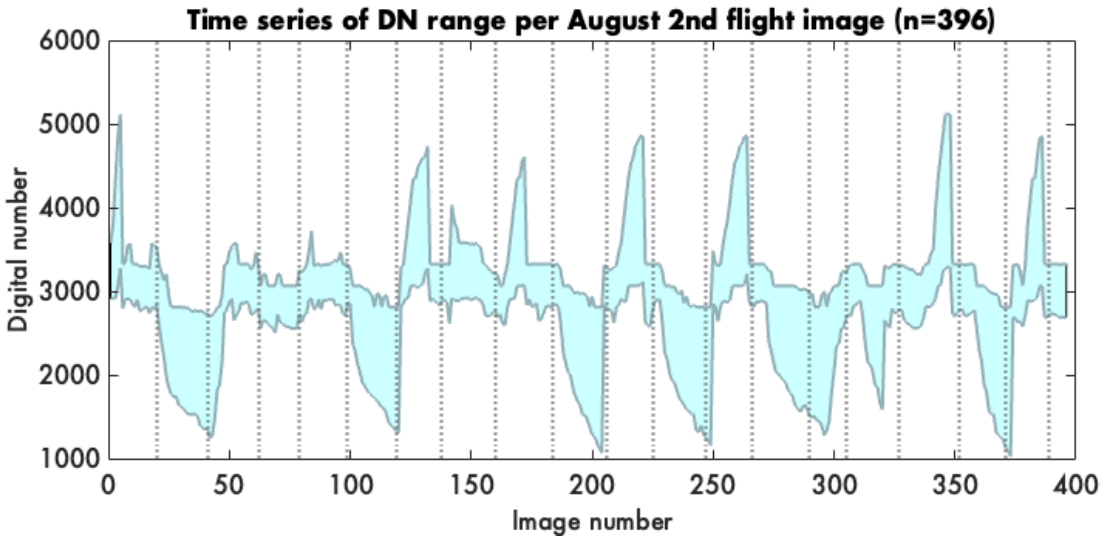

Figure 3.14: Time series of the DN range of every image (n=396) for the entire August 2nd flight, including all flight lines. The vertical dashed lines represent the turning point of the drone from one flight line to the next.

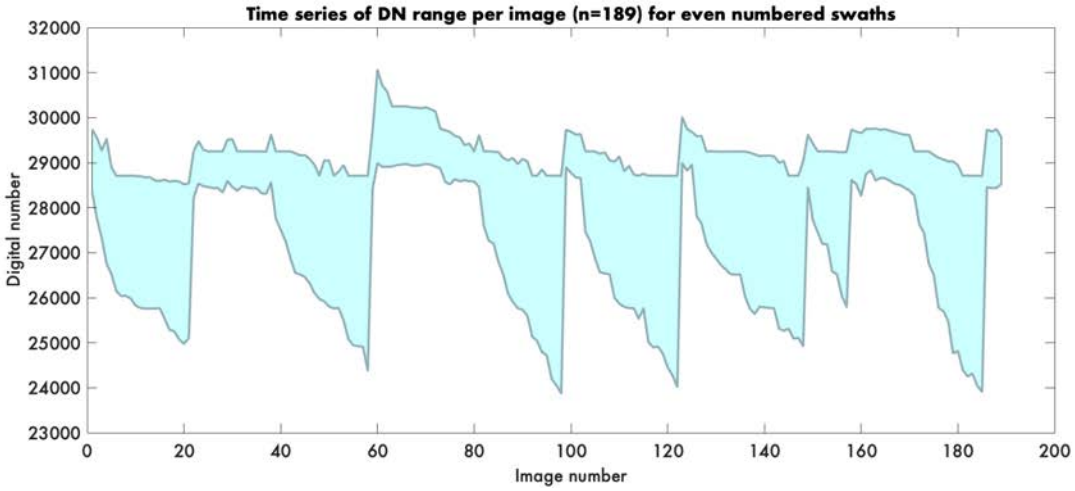

Figure 3.15: Sequence of DN ranges for the images (n=189) captured in the even numbered flight lines.

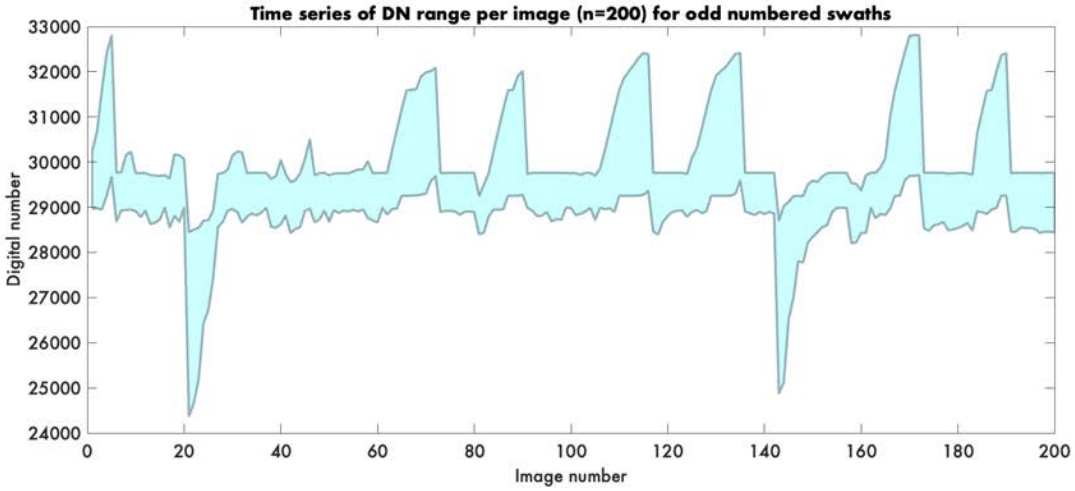

Figure 3.16: Sequence of DN ranges for the images (n=200) captured in the odd numbered flight lines.

Thermal Outflow

One of the primary objectives of this study concerns the utility of UAS TIR remote sensing for geothermal reconnaissance. Since the location of the Norris Hot Springs surface expression is known, we focused on this specific feature when comparing the results of each of the mosaicking methods. When qualitatively assessing the results, it appears that the TIFF mosaic shows the highest contrast between the thermal outflow and the surrounding ground (Figure 3.17). This contrast is dulled in the RJPEG, Odd, and Even mosaics. Surprisingly, the thermal outflow is not preserved at all in the Pix4D mosaic.

To quantitatively compare the results, we measured the DN (or temperature) value of the pixels corresponding to the start of the surface expression. For reference, the thermal source in the raw images is represented by a DN of 3,999, which equals 72.9°C. In the RJPEG mosaic, we measured a value of 30,919. In the TIFF mosaic, we measured a value of 3,824, which represents a temperature of 69.67°C. However, this is an overestimation of the thermal outflow temperature, which the Montana Bureau of Mines and Geology measured at 53°C (Sondregger et al., 1981). For the Odd and Even mosaics, we measured values of 30,432 and 27,071, respectively. Lastly, we measured a temperature of 17.76°C in the Pix4D mosaic (we evaluated the same pixel location as the other mosaics, even though the thermal outflow is not visible as an anomaly in this one).

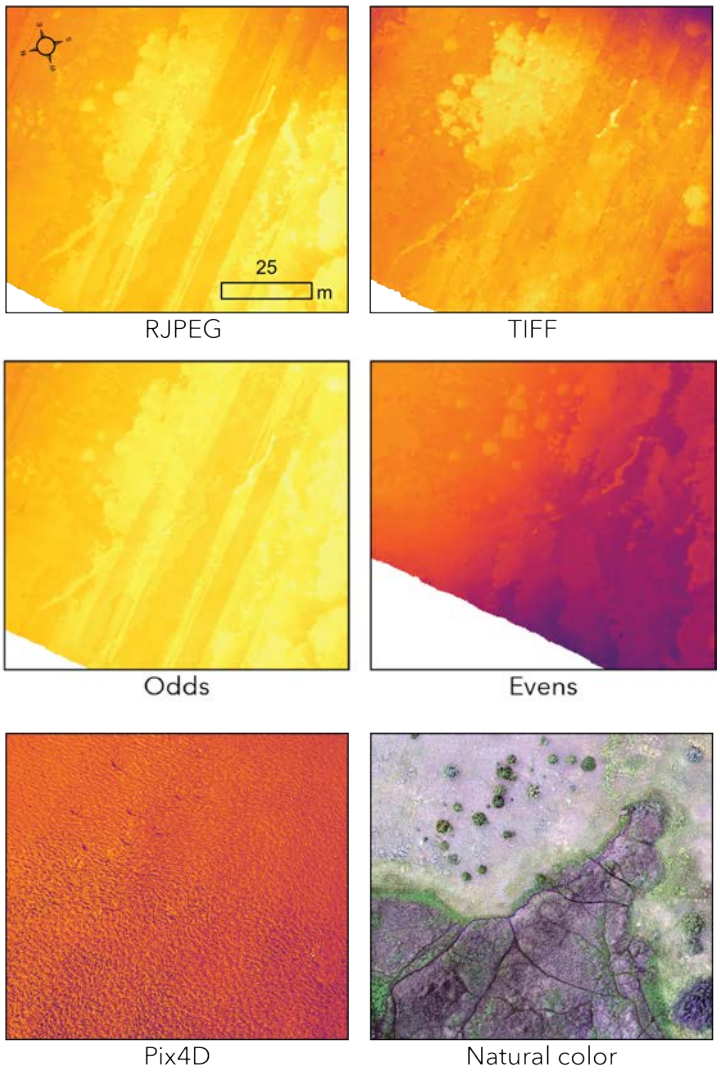

Figure 3.17: The start of the thermal outflow that feeds Norris Hot Springs, as seen in each of the thermal mosaics and the natural color orthomosaic. All thermal mosaics are symbolized for their respective minimum and maximum values (i.e. no histogram stretch).

Training Data Rasters

Using ArcGIS Pro, we created a series of rasters that served as explanatory variables in our machine learning model. These included land cover classification, solar radiation, aspect, slope, and elevation (Figures 3.18, 3.19, and 3.20). For the land cover classification raster, we iteratively adjusted the results before settling on a final version. The final raster contains four schema classes: vegetation, rocks, barren/grass, and water (Figure 3.18).

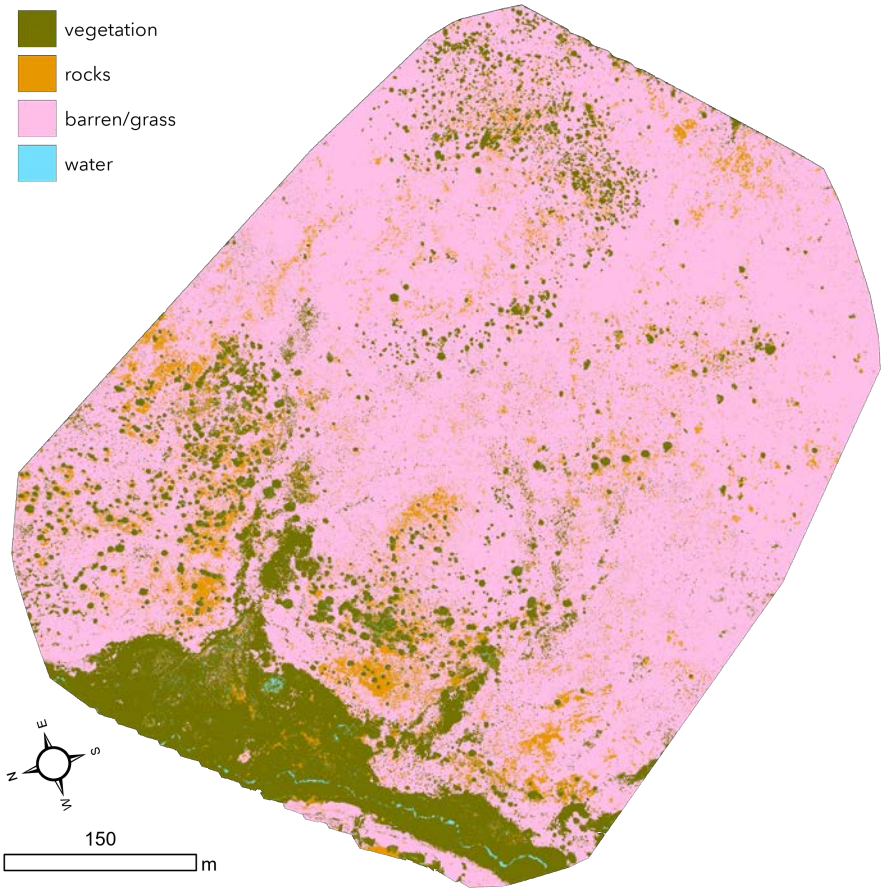

Figure 3.18: The final result of our land cover classification.

To account for solar radiation in our model we created 12 individual rasters - one for every 30 minute interval from 6am to 12pm, corresponding to the time period of our data acquisition in the field (Figure 3.19). Although we flew the mission at dawn to limit the

amount of solar radiation acting on the field area during the time of flight, solar insolation changes over time from 6am to 12pm and it also varies spatially across the field area, due to changes in topography (Figure 3.19).

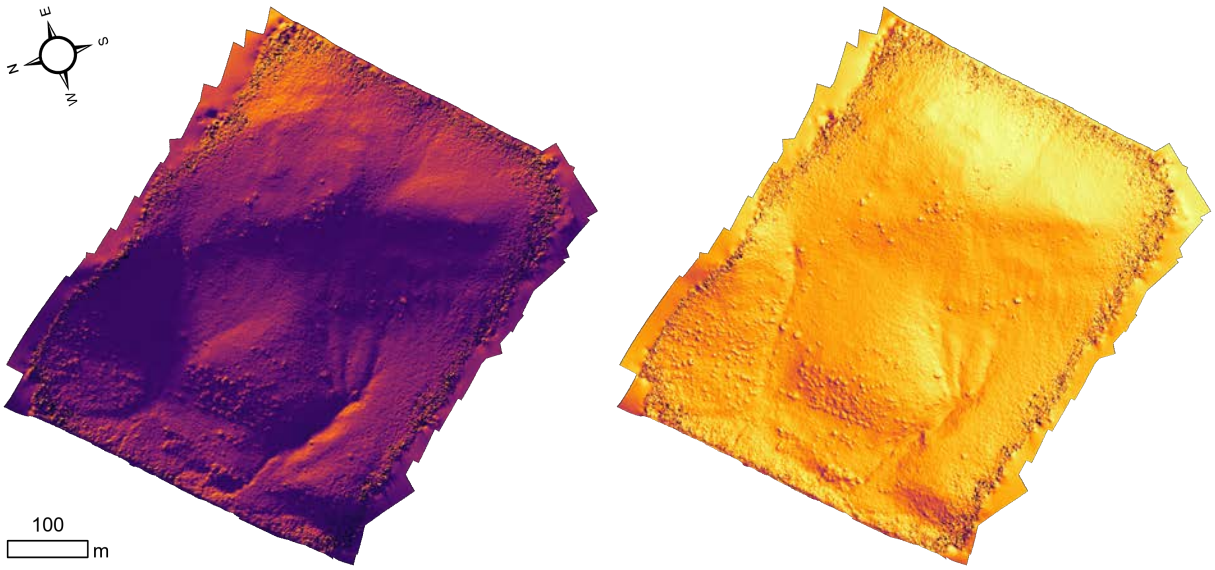

Figure 3.19: The two end members of the 12 solar radiation rasters generated in ArcGIS Pro. *Left*: Solar radiation at 6am on August 2nd, 2022. *Right*: Solar radiation at 12pm on August 2nd, 2022.

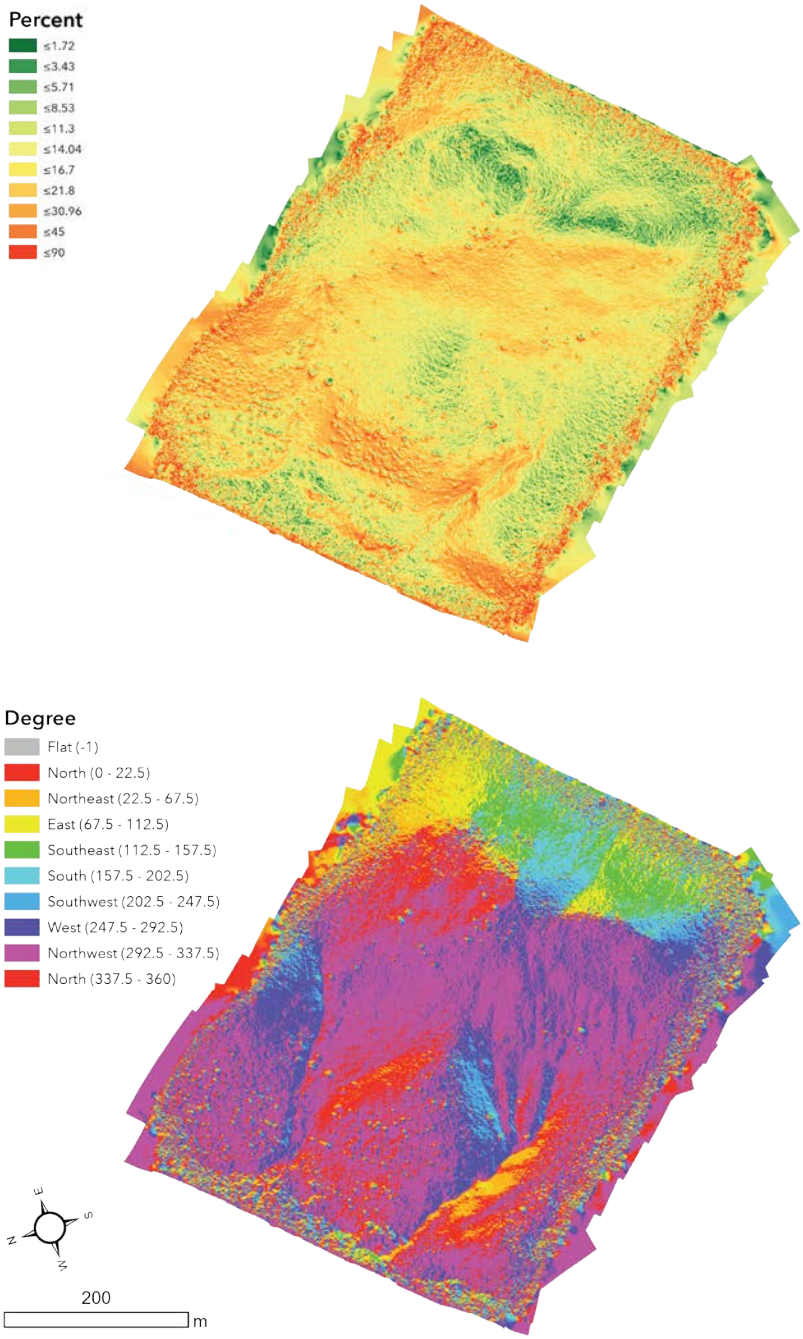

Figure 3.20: *Top*: Slope map of the study area. *Bottom*: Aspect map of the study area.

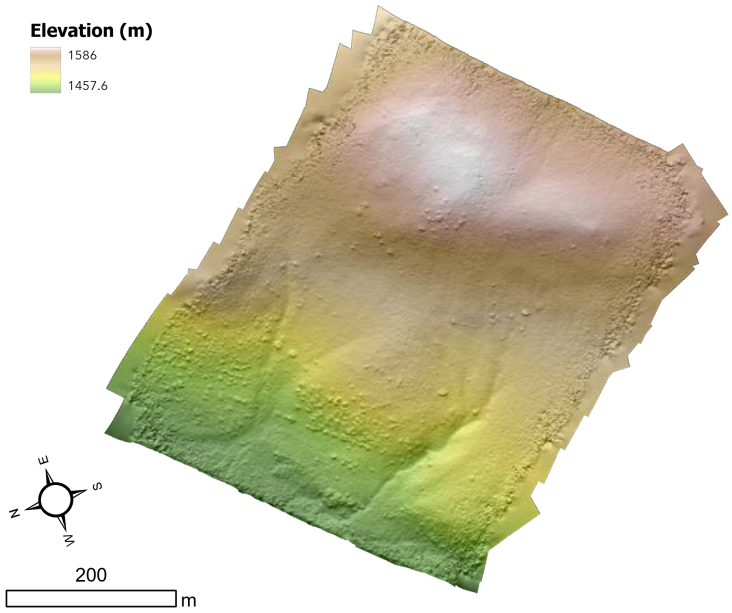

Figure 3.21: DEM of the study area.

Modeling

Linear Regression To test the relationship between the August 2nd in situ temperature and the sets of remotely sensed temperature, we ran a simple linear regression for each of the five mosaics. We evaluated these results through their coefficients of determination (R^2), which describe the amount of variation in the dependent variable that is explained by the independent variable. Based on this metric, our results do not show a strong relationship between any of the mosaics and ground-sampled temperature (Table 3.5). This means that our models fail to explain the variation in remotely sensed temperature. Additionally, only two of the models (Pix4D and the Even mosaic) were statistically significant (P-value <0.05). The other three regression models did not yield statistically significant results, meaning that they do not provide enough confidence to correlate a relationship between the in situ temperature and the remotely sensed temperatures.

The strongest relationship exists for the mosaic of even flight lines, which had the largest R^2 value (Table 3.5). However, it is still a relatively weak result - it suggests that the in situ temperatures only explain approximately 35% of the variation in temperature obtained from the Even mosaic. The second strongest model was the Pix4D one (Table 3.5). Again, this is a relatively weak connection and indicates that ground-sampled temperature only explains roughly 28% of the variation in the Pix4D temperatures.

Predictor variable	R^2	P-value
Pix4D	0.285	0.009
RJPEG mosaic	0.127	0.068
TIFF mosaic	0.039	0.321
Odd flight lines	0.005	0.735
Even flight lines	0.322	0.001

Table 3.5: Results of the linear regression models, including coefficients of determination (R^2) and measures of statistical significance (P-value).

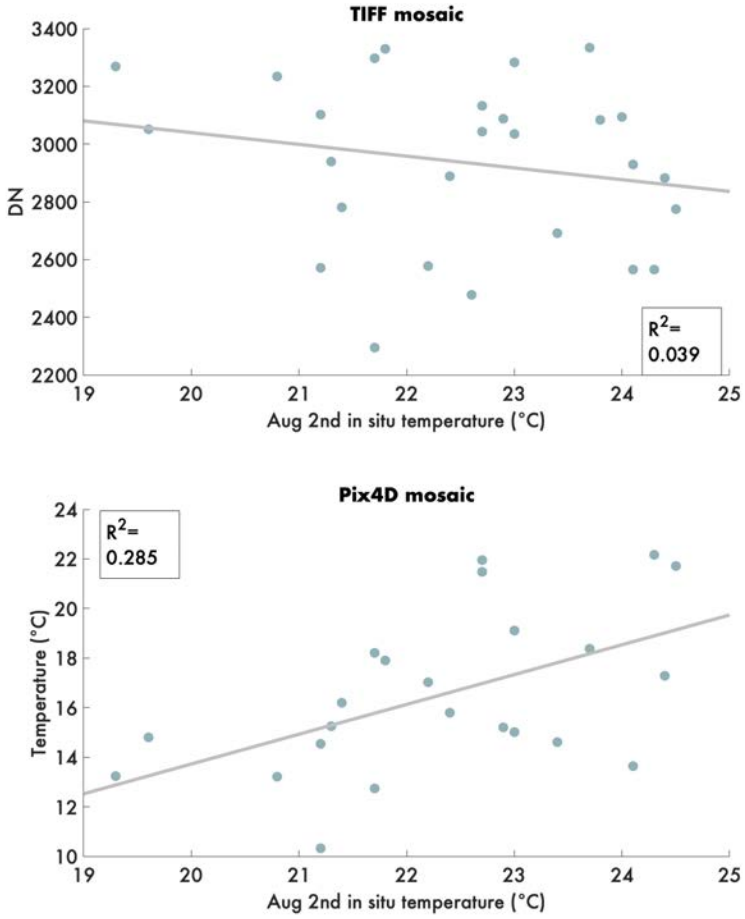

Figure 3.22: Linear regression results for the TIFF and Pix4D mosaics.

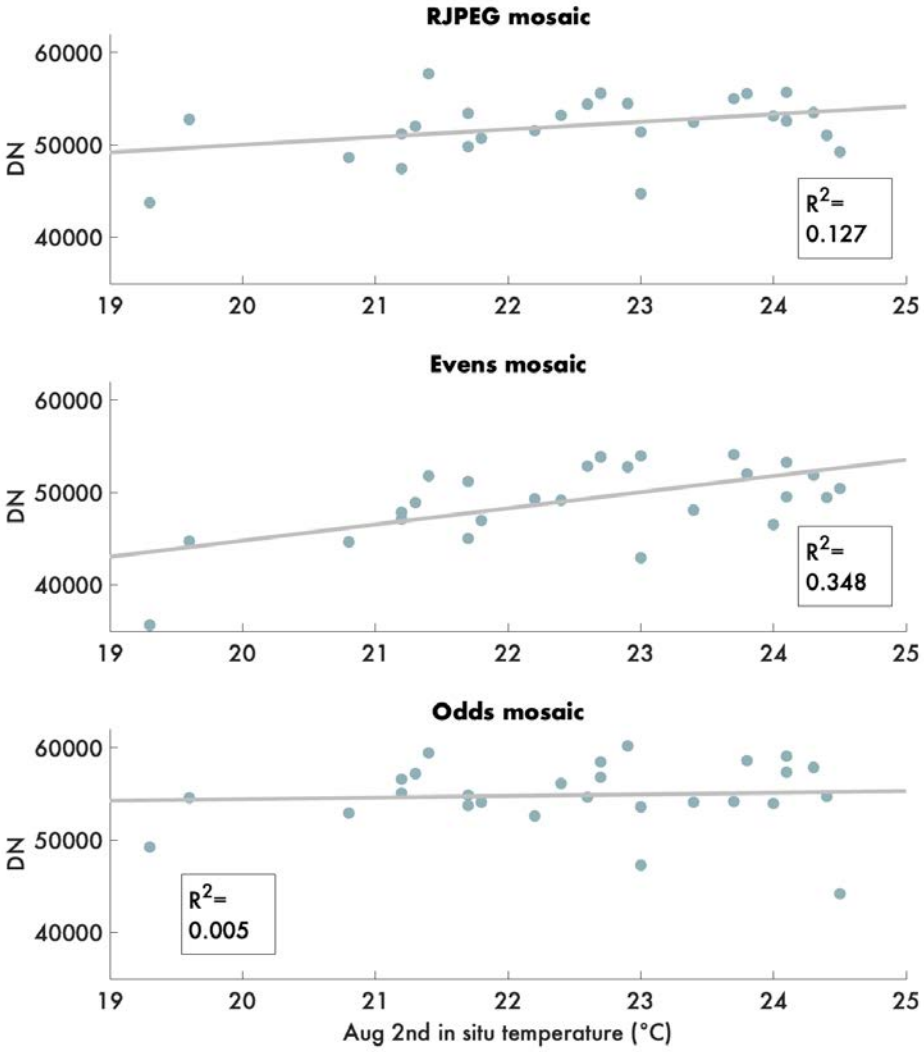

Figure 3.23: Linear regression results for the RJPEG, Even, and Odd mosaics.

Forest-based Regression For the machine learning portion of our modeling work, we investigated the correlation between a curated set of explanatory variables (i.e. thermal conductivity, soil grain size distribution, solar radiation, aspect, slope, land cover classification, distance to faults, etc.) and multiple sources of thermal data, including remotely-sensed temperature and ground-sampled temperature. We tested multiple predictor variables to explore whether a certain processing method would yield a higher performing model than other processing methods. We also tested the August 2nd ground temperature as a predictor variable to remove remote sensing as a model factor, altogether. By omitting the remotely sensed data, we tested the relationship between the explanatory variables and the temperatures we measured directly.

As part of our modeling process, which is explained in detail in Section 2, we created a variable importance box plot for each model. These boxplots allowed us to assess the stability of each model while tuning the hyperparameters, such as number of trees, number of randomly sampled variables, tree depth, etc. The boxplot ranges provide information about the consistency of the model: smaller ranges indicate less variability across five validation runs, thereby indicating a more stable model. If the model results change with each validation run, then it means that the importance of a given explanatory variable changes with each model iteration. In a stable model, the same explanatory variable should maintain its relative importance from run to run.

The final boxplots for each of the five models are provided below. Note - some of the explanatory variable names are abbreviated to conserve space. LCC (field) represents the land cover classification data we groundtruthed at each sampling location. LCC (supervised) represents the land cover classification we ran in ArcGIS Pro, which we trained to produce a land cover map of the entire study area, based on imagery analysis of our natural color orthomosaic. The geologic unit and soil type data originate from the USGS Norris Quadrangle, and the Soil Survey Geographic Database (SSURGO), respectively. The grain

size variables represent the results of our soil grain size analysis, where 'granule' represents the largest sieve class, 'very coarse' represents the second largest, 'medium' represents the third to last, and 'coarse silt' represents the smallest sieve class (Figure 2.9). The thermal conductivity and temperature variables represent the post-flight in situ data we collected on August 2nd at the 30 sampling locations.

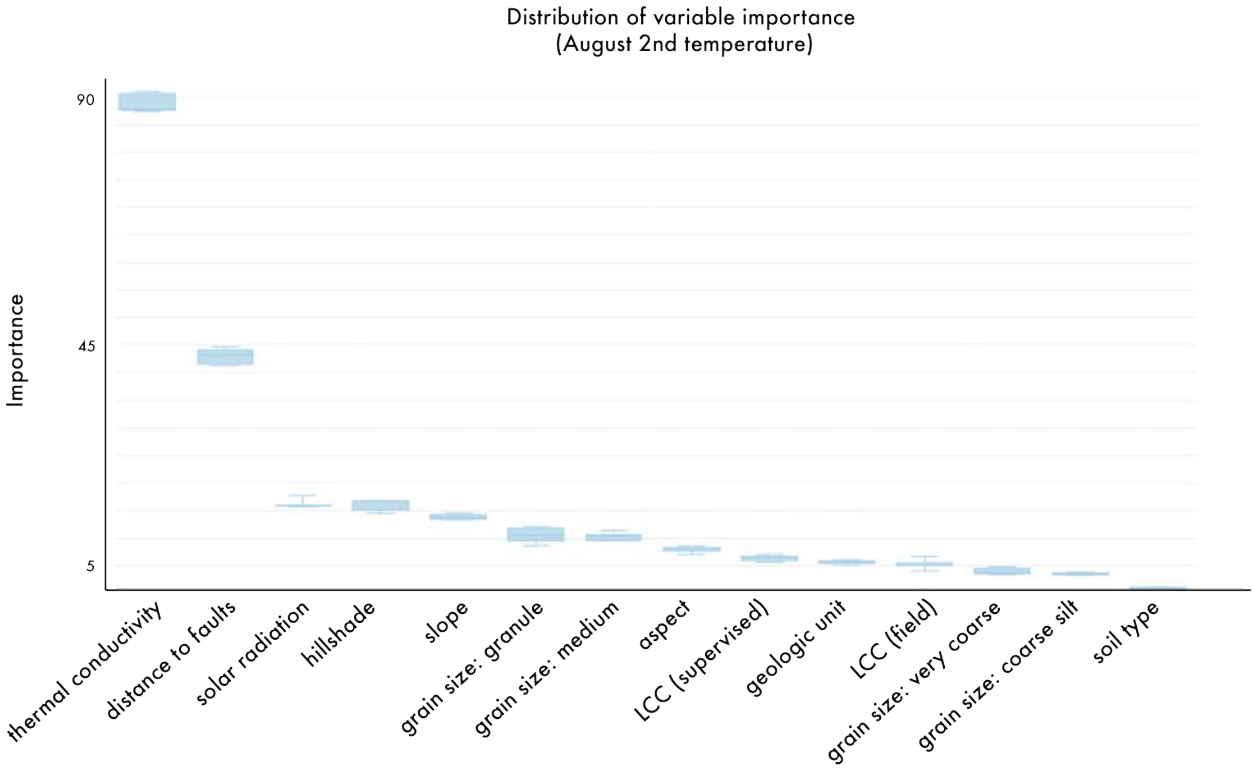

Figure 3.24: Variable importance boxplot for five runs of model validation when using the August 2nd in situ temperature as the predictor variable. No data excluded during validation.

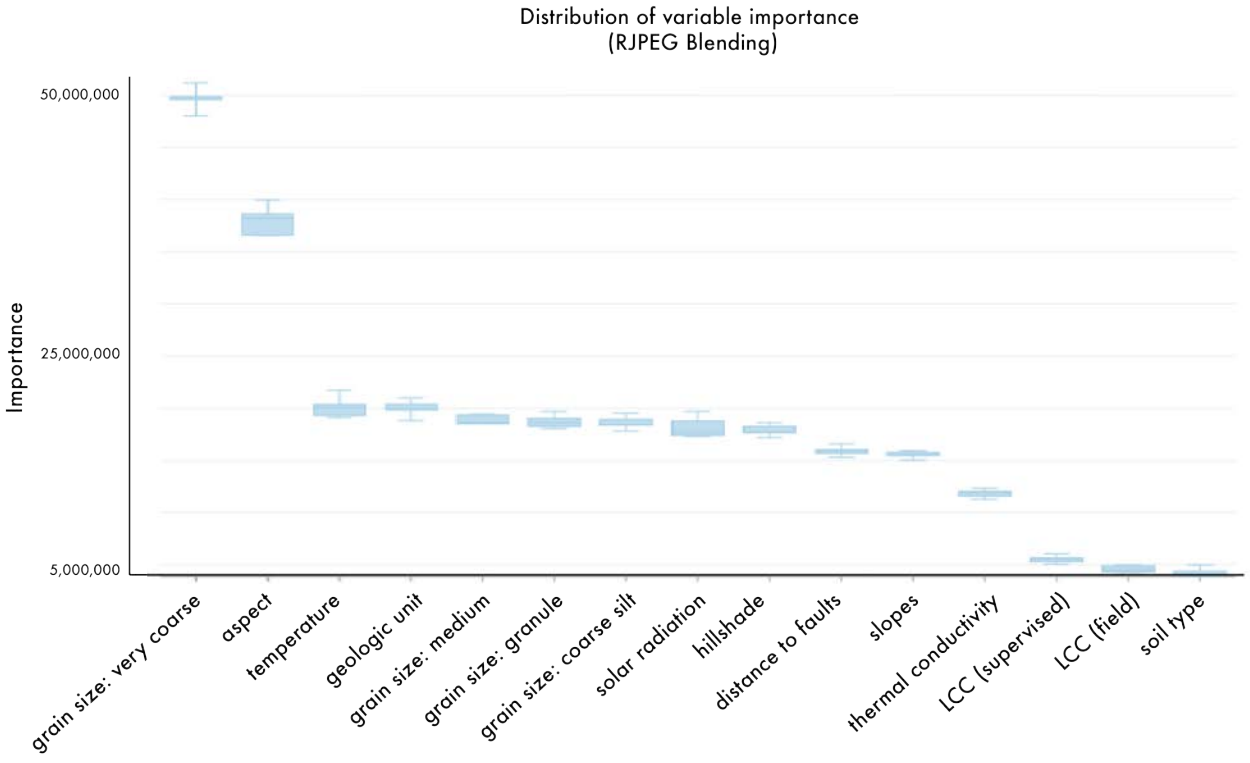

Figure 3.25: Variable importance boxplot for five runs of model validation when using the RJPEG mosaic as the predictor variable. No data excluded during validation.

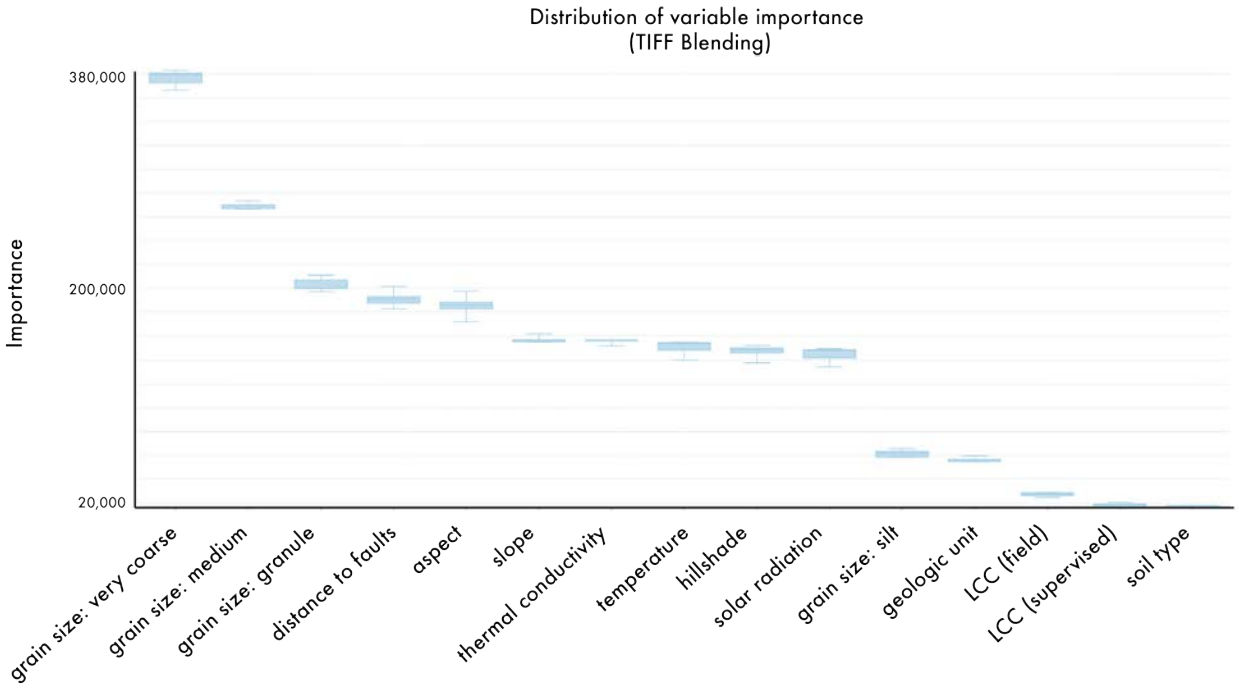

Figure 3.26: Variable importance boxplot for five runs of model validation when using the TIFF mosaic as the predictor variable. No data excluded during validation.

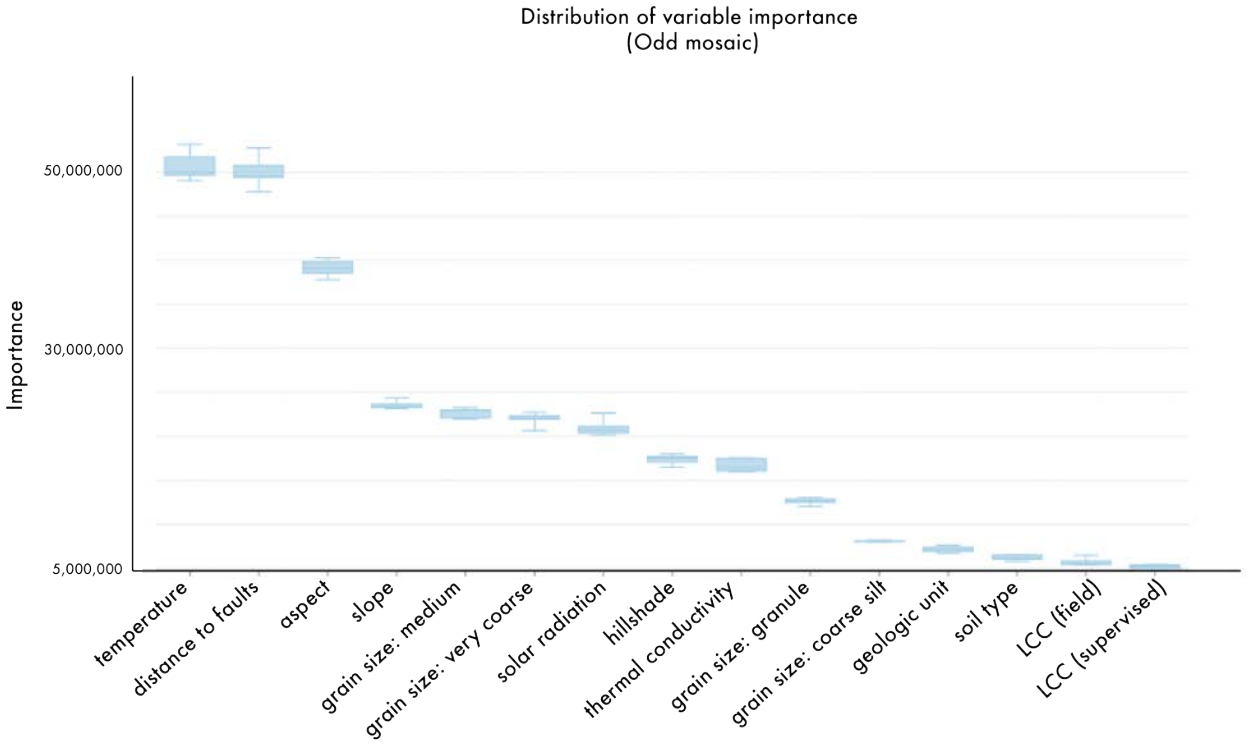

Figure 3.27: Variable importance boxplot for five runs of model validation when using the Odd mosaic as the predictor variable. No data excluded during validation.

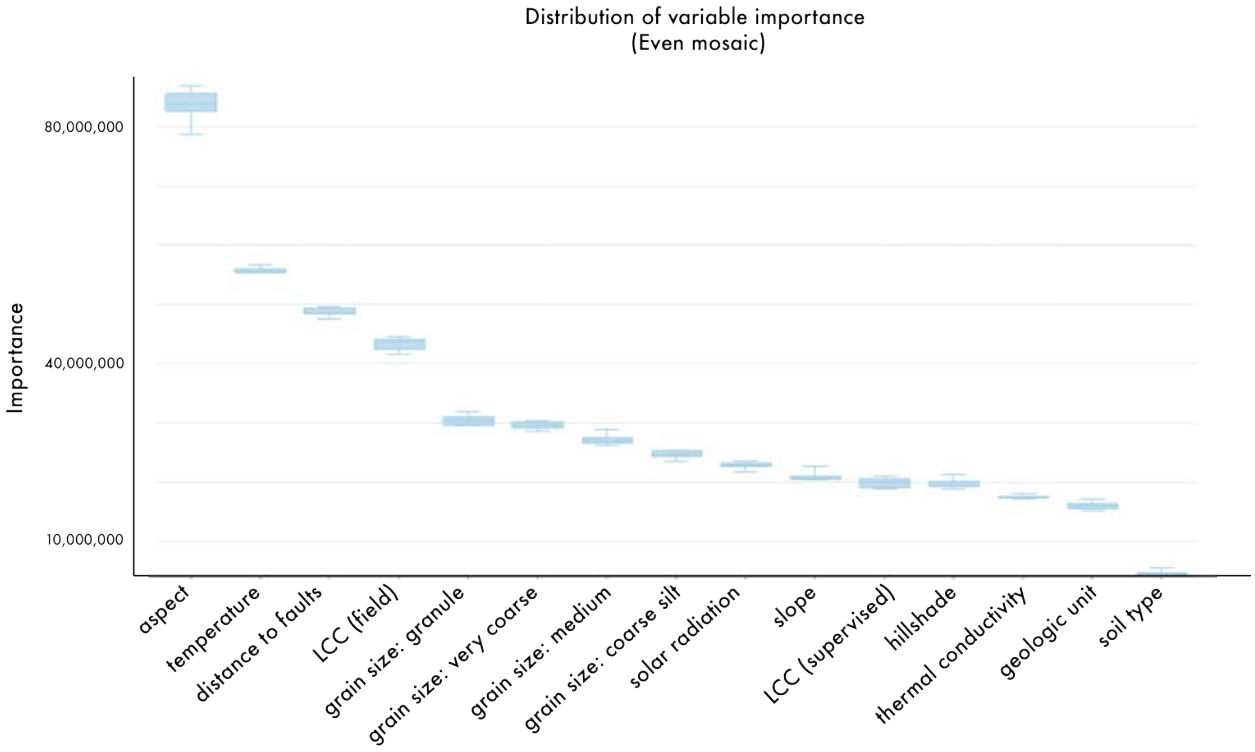

Figure 3.28: Variable importance boxplot for five runs of model validation when using the Even mosaic as the predictor variable. No data excluded during validation.

After finetuning the hyperparameters and achieving the most stable version possible, we ran each model. The *Forest-based classification and regression* tool in ArcGIS Pro yields several performance metrics after running a given model, the first being out-of-bag errors. These include mean square error (MSE) and percent of variation explained (R^2 converted to percentage) (Table 3.6). Additional performance metrics include R^2, P-values, and standard errors for both the training and validation phases of the model (Tables 3.7 and 3.8).

Table 3.6 contains the out-of-bag errors associated with each model, which represent the model's predictive capabilities. The MSE describes the model's performance by calculating the average of the squared differences between the model's predicted values and the observed values provided in the training data. Since a lower MSE indicates a better model, it is clear from our results that most of the models are not good predictors. The percent of variation explained is either low or negative for most of our models. R-squared is typically a positive value, so a negative number indicates very poor performance. It essentially means that the model is a worse fit for the data than the fit of a simple horizontal line. This suggests that all three of the models with a negative R^2 are inherently worse at predicting the remotely sensed temperature than if we predicted results at random.

However, the model using the August 2nd in situ temperature as the predictor variable had much better results than any of the other models that used remotely sensed temperature. In this instance, the MSE was 5.2°C and the percent of variation explained was 44.3%. The latter indicates that the explanatory variables accounted for 44.3% of the variation in the ground-sampled temperature. Despite these results, the measure of statistical significance (P-value of 0.442) indicates that these results are not actually significant.

	Out-of-bag errors	
Predictor variable	MSE	% of variation explained
RJPEG mosaic	10,875,088.996	1.177
TIFF mosaic	89,786.980	-13.941
Odd flight lines	13,038,098.521	-23.171
Even flight lines	18,860,237.166	-7.303
Aug 2nd in situ temperature	5.155	44.297

Table 3.6: Performance metrics

	Regression diagnostics: Training		
Predictor variable	R^2	P-value	Standard Error
RJPEG mosaic	0.948	0.000	0.028
TIFF mosaic	0.929	0.000	0.031
Odd flight lines	0.953	0.000	0.024
Even flight lines	0.927	0.000	0.033
Aug 2nd in situ temperature	0.960	0.000	0.029

Table 3.7: Training diagnostics

	Regression diagnostics: Validation		
Predictor variable	R^2	P-value	Standard Error
RJPEG mosaic	0.170	0.730	0.548
TIFF mosaic	0.026	0.898	0.718
Odd flight lines	0.425	0.548	0.069
Even flight lines	0.363	0.588	0.190
Aug. 2nd in situ temperature	0.591	0.442	0.512

Table 3.8: Validation diagnostics

DISCUSSION

Lab Experiments

Increment

For the first objective of this experiment, we tested the radiometric accuracy of the FDPR across a range of blackbody temperatures. The results varied across the five different temperature stages. At the lowest blackbody temperature of 20°C, the camera underestimated the blackbody temperature (Table 3.1). However, at the higher temperature regimes, it overestimated blackbody temperatures by varying magnitudes. We did not identify an overall trend relating the magnitude of error to the degree set for the blackbody. For example, the TIFF results indicate that the camera overestimated the blackbody at 35°C by nearly 6°C, but only overestimated the 40°C blackbody by approximately 1°C (Table 3.1). Based on these results, there isn't a uniform post-processing correction that we can apply to account for variation in the camera's calibration accuracy.

The manufacturer specifications state that the FDPR is accurate within ± 5°C, but we found that the camera did not maintain this level of accuracy in three instances (Table 3.1). However, despite the different magnitudes of error that we observed, the standard deviation at each temperature increment, for either file type, was less than 1°C. The low standard deviations across all of the image sets imply that the camera's measurements are precise when ambient conditions are stable. This supports the idea that under stable environmental conditions, the camera records data that is precise, but not always accurate.

For the second objective of the experiment, we compared the two file types available for storing images. Surprisingly, the respective file type and associated processing method produced differing results. When comparing the results between a corresponding set of RJPEGs and TIFFs, we observed a discrepancy in mean temperature and MAE. For example, at the 25°C increment, we calculated an MAE of 5.330°C for the TIFFs files but an MAE

of 2.7229°C for the RJPEGs. Even though we captured both sets of images immediately after one another, the TIFF data overestimate the 25°C blackbody temperature more than the RJPEG data. This was the case for four of the five increments. The one exception was the 30°C increment, where the TIFF results had a slightly better MAE than did the RJPEG results. However, we conclude from this experiment that the RJPEG file type (and associated temperature conversion) yields more accurate data than its counterpart. As a result of this finding we decided to capture all images in the RJPEG format for future UAS missions.

Stabilization

In our stabilization experiment, we calculated a standard deviation of 0.2051°C, which indicates measurement consistency for the duration of the test. We did not observe large fluctuations in recorded temperature upon powering the FDPR, unlike the observations of Kelly et al., 2019 with the FLIR Vue Pro or Smigaj et al., 2017 with the Optris PI-450 (Figure 3.3). One explanation for the dissimilar findings is that our camera likely reached equilibrium with the room's ambient temperature prior to activation. It is plausible that the camera may perform differently in a situation where it is transferred from a hotter or colder environment immediately prior to this type of test.

Although we set the blackbody unit to a constant temperature of 25°C, the FDPR calculated a mean temperature of 31.9011°C. The large error of 6.9011°C suggests that the camera overestimates a 25°C object by nearly 7°C. This is slightly higher than the error we calculated from the 25°C stage of the increment test (Table 3.1). However, both findings are similar in magnitude to the results of a previous study that demonstrated a MAE of –5.965°C associated with a FLIR A655sc thermal camera (Aragon et al., 2020).

Wind

In the wind experiment, we recorded a mean temperature of 16.5892 ± 2.62°C with an MAE of 8.4108°C during the simulation and a mean temperature of 31.0060 ± 0.2890°C with an MAE of 6.006°C after removing the simulation. At the transition where we removed the wind source (30-minute mark), the camera recorded a sharp increase of nearly 15°C, despite the blackbody unit remaining at a stable temperature of 25°C (Figure 3.4). All else being equal, the camera recorded temperatures much cooler than the true blackbody temperature when subjected to wind vs. no wind.

Immediately following the removal of the wind source, we observed an over-correction by the camera, where it recorded the warmest temperatures of the experiment, prior to leveling out. This observation has implications for our August 2nd flight data - if the camera over-corrects after being exposed to wind, then the initial images in a tailwind flight line may display higher temperatures than expected. Lastly, the MAE of the recovery period suggests that the camera overestimates object temperatures of 25°C by 6.006°C, which is similar to the mean bias of 6.9011°C observed in the stabilization experiment.

The results of this experiment suggest that the camera's response is strongly impacted by wind exposure. This proves problematic, given that the FDPR is marketed for outdoor use. Wind variability during UAS missions will therefore have a significant effect on thermal measurements made in the field. Although our results provide valuable information regarding the temperature-dependency of our camera, a solution remains elusive. It would be quite complicated to correct for wind, either in post-processing or in the field. In this controlled experiment, we maintained a steady stream of air at 12 m/s for the duration of the 30-minute simulation. However, this is not representative of true field conditions, which do not typically remain at steady state. Due to the dynamic nature of wind speed and direction in real-world conditions, it would be challenging to compensate for the effects of wind during post-processing. Even if we empirically established the relationship between wind speed and

the camera response, it would be difficult to identify which flight images correspond to which wind conditions and to then apply these corrections to the data. A more plausible approach to account for variable wind conditions in the field is to physically buffer the thermal camera with insulating foam or a 3D-printed shield. While both modifications have potential, they are physically limited by the presence of the gimbal, which allows the camera to rotate during flight (Figure 2.6). Future studies exploring modified hardware and gimbal components are needed to test the feasibility and effectiveness of this idea.

Heat

In the heat experiment, we investigated the relationship between an increase in ambient temperature and the temperature measured by the FDPR. While the blackbody temperature stayed constant at 25°C, the camera overestimated this temperature and recorded a mean of 28.4318°C. The results also show that the camera registered large oscillations in temperature upon exposure to the external heat source (Figure 3.5). This is evidenced by the standard deviation of 4.22°C. After the initial period of instability, it achieved a more stable response for the remainder of the test. The period of instability only lasted for the first three minutes, which suggests that when the camera body is exposed to heat, it is eventually able to record stable measurements, albeit at an overestimation. This overestimation of blackbody temperature is similar to the results of our steady-state experiments. Interestingly, the MAE of 3.4318°C was lower than the MAE calculated in other experiments when conditions were stable. In those tests, the MAE ranged from 6-7°C. In this experiment, the camera fluctuated between underestimating and overestimating the blackbody temperature, whereas any deviations in the stabilization or increment experiments remained above the blackbody temperature at all times. Like our findings from the wind experiment, the heat experiment indicates that the camera measurements are influenced by changing ambient conditions.

Sensor Non-Uniformity

It is evident from our vignetting visualization that there is pixel-level noise unique to our camera's detector array (Figure 3.6). The mean pixel temperature across the array is spatially heterogeneous. Similar to the findings of Smigaj et al., 2017, our results show a temperature gradient in the shape of a semicircle, as opposed to a gradient that radiates equally in all directions from the center of the detector (Figure 3.6). While it is known that most cameras have an inherent vignetting artifact related to their optical components, we did not expect a vignetting pattern that varies laterally, rather than radially. Smigaj et al., 2017 propose that this atypical pattern is due to imperfections in the optics of the camera and/or the degradation of the sensor's calibration over time. Considering that our FDPR is only a few years old, we suspect that the former is to blame.

The preliminary phase of this research tested the calibration accuracy and environmental sensitivity of the FLIR Duo Pro R camera through lab experiments with a blackbody calibration unit. Overall, our experiment results indicate that the thermal camera overestimates stable blackbody temperatures by varying degrees when operating in stable conditions and exhibits large error when exposed to changes in ambient conditions, such as wind and heat. These results suggest that temperature measurements recorded by the FDPR are strongly influenced by multiple internal and external factors. Without the mitigation or consideration of these factors, the camera cannot record accurate surface temperatures, which reduces its effectiveness for scientific applications.

Computational Modeling

Photogrammetry

Compounded by the aforementioned issues associated with uncooled microbolometers, there are many challenges associated with orthomosaicking thermal imagery in commercial photogrammetry software. Since the environmental sensitivity of these types of thermal sensors leads to a drift in measurements across a flight plan, the thermal data may vary with time, distance, and even flight direction. Malbéteau et al., 2021 found that individual flight lines are not only affected by the presence of wind, but also influenced by the relationship between the flight direction and wind direction, i.e. whether the UAS is flying into the wind or away from the wind. With this in mind, we developed our third mosaicking method (the alternating swaths approach) in an attempt to standardize the dataset to the same flight direction for all flight lines. In this technique, we generated two mosaics - one containing the even numbered flight lines and one containing the odd numbered flight lines. The even swaths represent the flight lines from our August 2nd flight that coincided with a headwind and the odd swaths coincided with a tailwind.

An additional issue that we encountered during the mosaicking process was the preservation of the raw thermal data. When stitching the images in Pix4D, the software retained the absolute temperature values but could not mosaic the entire dataset. Although we were able to successfully mosaic all of the thermal images when using Metashape, the software altered the LWIR DNs via an unknown transformation. Consequently, three of our mosaics have altered values. We can discuss these in the context of relative comparison, but we cannot convert them to absolute temperature. The complications associated with mosaicking thermal images while retaining radiometric information is a significant limitation of the Metashape software. Conversely, the Pix4D limitations include an inability to compile the entire dataset and discarding thermal anomalies in the final mosaic. Therefore, none

of the processing entities that we attempted (DroneDeploy, Pix4D, Metashape) worked perfectly - each one was limited by unique issues. Although photogrammetry is an excellent tool for optical imagery, more work is needed to improve the photogrammetric processing methods for thermal infrared data.

Thermal Mosaic Generation: Pix4D Pix4D could not stitch the entire August 2nd dataset, only producing a map that covered approximately 60% of the flight footprint. Additionally, the resulting mosaic had a distinct texture that we interpret as an interference pattern from the blending method employed by the software, as this pattern is not observed in the images prior to mosaicking. However, the primary benefit of the Pix4D method is that it retained the original radiometric data, meaning we can interpret the results in terms of absolute temperature, rather than DNs. Despite this, the mosaic temperatures do not fall within a realistic range that we would expect for ground surface temperatures during the summer. For instance, the results show a minimum temperature of -24.35°C, which is well below freezing, which is certainly not the case for this field area at that time of year. Additionally, the maximum temperature of 42.63°C is much lower than the temperature of the thermal outflow as captured by the raw images (72.9°C). Moreover, this maximum temperature does not spatially coincide with the thermal outflow, indicating that the highest temperatures in the Pix4D mosaic are artifacts.

Thermal Mosaic Generation: RJPEG Blending The thermal mosaic for this method, which included all flight lines (n=19), had the widest range of temperature values when compared to the two other mosaics with the altered DNs (the Odd and Even mosaics) (Table 3.4). The mean value of this mosaic fell in between the respective means of the Odd and Even mosaics. Upon visual inspection of the RJPEG mosaic, it is apparent that there are SW-NE trending stripes of alternating warmer and cooler thermal trends across the entire flight mission (Figure 3.12). We interpret these trends as artifacts, rather than a

true thermal pattern, for a few reasons. First, it is unlikely that a striped thermal pattern would be found across a landscape in nature. Second, the stripes spatially correspond to the individual flight lines. Lastly, a pair of side-lapping images from two different flight lines show relatively different thermal properties, despite covering almost identical ground areas.

Thermal Mosaic Generation: TIFF Blending This method produced a thermal mosaic with a slightly smaller extent of than the RJPEG, Even, or Odd mosaic. We suspect that Metashape was unable to match the edges and corners of the mosaic due to the lower spatial resolution of the input files - in the RJPEG blending and alternating swaths methods, Metashape used the false color infrared images to stitch the mosaics together, but in this method we only provided it with the single LWIR band for stitching.

This method retained the DNs of the raw images, which ranged from 2,078 to 3,920. However, this range of values is much lower than the range from the raw images. The standard deviation is also lower than that of the raw images. This may imply that thermal outliers are removed during the blending process. If so, this is a problem for detecting subtle thermal anomalies. If the mosaicking process dulls the magnitude of thermal features, it draws into question the effectiveness of this method for scientific applications, including geothermal reconnaissance.

Thermal Mosaic Generation: Alternating Swaths After noticing the SW-NE trending stripes present in the mosaics from the previous two methods, which we interpreted as artifacts related to the respective flight direction along individual flight lines, we created the Odd and Even mosaics. Since we had a high side overlap between flight lines, we were able to eliminate every other flight line and still have enough data to construct a mosaic. We can directly compare the Odd mosaic and Even mosaics to observe the differences between the thermal patterns when including or excluding flight lines with a headwind or a tailwind. The Even mosaic had a lower range of DNs than did the Odd mosaic (Table3.4). The

Even mosaic also had a negatively skewed distribution, meaning that the dataset had a larger concentration of lower values, representing cooler temperatures. Since we standardized this mosaic for all headwind flight paths, and we know that the camera records cooler temperatures when it is subjected to wind, it makes sense that this mosaic retained more pixels representing cooler temperatures. This is also supported by the time series of this mosaic, which shows a repeating pattern of decreasing DNs per flight line (Figure 3.15).

We also observed a single northwest-southeast trending channel of warmer temperatures that is quite prevalent across the Even mosaic and also detectable in the RJPEG and TIFF mosaics (Figures 3.10 and 3.13). We have two hypotheses to potentially explain this linear region of high temperatures. One idea is that this is an artifact from the Non-Uniformity Correction (NUC), which re-calibrates the camera mid-flight and triggers at an automatic interval. If this is the case, we suspect the NUC triggered at the same distance along each flight line for two reasons: a) the constant speed of the UAS throughout the mission and b) the standardized timing of the NUC. The second and more convincing hypothesis is that this relates to the camera's tendency to over-correct temperatures immediately following a period of wind exposure, as evidenced by the results of our lab experiment targeting the effects of wind on the camera. Since the spots of brightness all correspond to the same location along each flight line, it is plausible that the camera over-corrected temperatures immediately after turning from a headwind swath to a tailwind swath. This is also supported by the flight pattern, where the UAS flew towards the NE side of the study area along a headwind swath, turned, and then proceeded towards the SW side of the study area along a tailwind swath. As a result of this pattern, the camera captured the first images along each tailwind swath on the NE side of the study area, which is where we observe the channel of brightness. This is further evidenced by the pattern observed in Figure 3.14, where an increasing trend of warm temperatures immediately follows a decreasing trend of cooler temperatures.

Thermal Outflow

The Norris Hot Springs outflow is an ideal target for comparing the results of the multiple processing methods we tested. We found that the mosaic results varied depending on the respective processing technique. Since some of the mosaics retained the raw radiometric data, we were able to directly compare the outflow temperatures recorded by the raw images, the TIFF mosaic, and the RJPEG mosaic. The raw images captured a temperature of 72.9°C. The temperature recorded in the TIFF mosaic was slightly lower, at 69.67°C. The thermal outflow in the Pix4D mosaic is not visible and the temperature at that location registers 17.76°C, which is much lower than the temperatures recorded in the raw image and TIFF mosaic. Since it has a similar temperature to its surroundings, it appears that the Pix4D method did not retain the thermal anomaly present in the raw image during the course of processing. From this, we conclude that photogrammetric choices have a significant impact on the results produced via thermal mapping. This has strong implications for the utility of UAS TIR for geothermal applications, particularly in regions where thermal anomalies are subtle. Depending on the photogrammetric software and processing choices, the mosaic results may or may not preserve thermal anomalies present at the ground surface, which could lead to inadvertently missing features in a thermal survey.

Modeling

Linear Regression The results of our linear regression analysis were statistically insignificant, with the exception of the Pix4D mosaic and the Even mosaic. Despite their statistical significance, neither model exhibited a strong connection between ground-sampled temperature and remotely sensed temperature. We suspect that there are a few reasons for the weak nature of this relationship. First and foremost, the temperature drift issues with uncooled microbolometers may render inaccurate absolute and relative temperatures. This may result in a mismatch between the remotely sensed thermal signature at a given location

vs. the ground-sampled temperature. Secondly, as an inherent part of the mosaicking processes, the algorithm chooses a value from multiple overlapping pixels to represent the thermal signature at that coordinate. Depending on the blending method, the resulting thermal value may vary at a given location. Lastly, while we made a concerted effort to collect the in situ temperature data immediately following the UAS flight, multiple hours passed between the beginning of the UAS flight and the end of the ground sampling due to the time-consuming nature of both tasks. This resulted in an offset in the timing of our data, meaning that the temperatures in the thermal imagery and the temperatures recorded via field sampling may not correlate as highly as if they were collected at the same time.

Forest-based Regression The results of our forest-based regression were poor. None of the models were statistically significant, meaning that we cannot confidently draw conclusions regarding the relationships between our explanatory variables and our predictor variables. The lack of statistical significance and the poor performance of our models may result from limitations within our study design. One of the primary issues is that we did not have enough training features. Random forest models are best suited for hundreds of features, but we could only accommodate 30 sampling sites in our study design due to time constraints for ground-sampling, post-flight. Additionally, there are issues with our remotely sensed thermal data, due to the aforementioned complications of uncooled sensors as well as the way that photogrammetric software process the data. Since the validity of the input data controls the validity of the model, our results might have been stronger if our remotely sensed data were more accurate.

Future Work

As evidenced by this research, there are currently many limitations to capturing accurate thermal data with a UAS. Many of these issues stem from the camera equipment and the photogrammetric workflows available to end-users. Accordingly, there are several

opportunities for continued investigation into improving these techniques. Based on the results of this study, we have multiple suggestions for the development of these methodologies moving forward.

To improve the radiometric accuracy of the thermal images, we propose the following ideas. The first suggestion concerns the cooling effect of wind on the thermal sensor. One option is to establish an empirical relationship between the camera's response and different wind speeds and to then develop a post-processing correction to account for this empirical relationship. This correction should be applied to the flight lines affected by wind prior to uploading to a photogrammetry program. However, as previously discussed, this potential solution may be labor-intensive and complicated. It also depends on access to certain equipment, i.e. a blackbody unit. Our second solution for this fundamental issue is the addition of a physical buffer on the camera during flight. However, further evaluation is needed to assess the compatibility of these components during flight and whether the buffer provides enough insulation in order to offset the temperature drift of the sensor.

An additional issue that we encountered during our research is an artificial increase in temperatures recorded by the camera immediately following exposure to wind. Like the previous suggestion, we propose establishing an empirical relationship and then applying a post-processing correction to the raw thermal images. Again, this may be time-consuming and dependent on specialized equipment. Our third suggestion expands upon the alternating swaths method that we used in this study, as well as the two previous ideas. As in the alternating swaths method, we propose removing the flight lines impacted by wind and then applying an empirical correction to account for the increase in warmer temperatures at the start of the tailwind flight lines.

To improve the timing between in situ temperature collection and the UAS flight, we propose the use of multiple ground temperature probes that remain in place and log the temperature data during the flight. This would yield a more synchronous dataset between the

ground-sampled temperatures and remotely-sensed temperature. By improving the collection timing of these datasets, it would potentially improve the calibration of the remotely sensed imagery and thereby improve the outcomes of the linear regression and machine learning models.

Lastly, we would like to explore an improved algorithm for stitching the thermal images via photogrammetry. The photogrammetric workflows currently available for TIR imagery are one of the significant limitations in this project. Future work should prioritize a more sophisticated blending method. After correcting the raw images to account for the temperature drift effects, a new blending method should prioritize outliers, rather than calculating mean temperatures. By placing a higher weight on minimum and maximum values that are detected in the raw imagery, the mosaicking results are more likely to preserve thermal anomalies, if present.

CONCLUSIONS

Although UAS remote sensing is an increasingly prevalent technique in spatial sciences, there are still many limitations to achieving accurate thermal data by UAS. In this study, we assessed the utility of UAS thermal infrared remote sensing for scientific applications, including geothermal reconnaissance. To evaluate the validity of this technique from multiple angles, our project focused on a comprehensive suite of lab experiments, field surveying, and modeling.

Our lab experiments indicate that the thermal data captured by our camera has a notable amount of error, for both stable conditions and changing environmental conditions. Our findings suggest that temperatures measured by the FDPR are particularly influenced by wind. The impact of wind greatly reduces the camera's accuracy - the FDPR underestimates temperatures when exposed to wind, followed by a period of temperature overestimation after the cessation of wind.

Our results from the field work phase of this research suggest that the temperature errors we identified during the lab experiments are also present in our field data. The temperature drift of the camera's thermal sensor influenced the data collected during UAS flight and these issues are evidenced by multiple artifacts in our thermal mosaics. Not only does the environmental sensitivity of the thermal sensor affect the absolute temperatures recorded in the raw images, but the associated errors also propagate through the photogrammetric processing stages. Since current photogrammetric workflows are designed for optical imagery, the options for stitching thermal imagery introduce even more uncertainty into the final results. Without mitigating or correcting for these sources of error, the artifacts introduced by this issue render UAS TIR remote sensing ineffective for certain research purposes requiring high accuracy of absolute temperatures.

The results from our linear regression and machine learning models did not yield strong

correlations between ground-sampled temperature and remotely sensed temperature, nor a strong relationship between our explanatory variables and temperature. These results indicate that the aforementioned limitations must be addressed in order to produce more accurate input data to generate models with better performance and more statistical significance. Additionally, future work should consider modifying the study design used in this research to include a higher number of training features.

Although UAS TIR remote sensing is marketed as easy-to-use for the general public, we found that there are extensive considerations for this technique and that it is not a simple or reliable tool at this point in time. Despite the challenges discussed in this paper, UAS TIR remote sensing may still serve as a useful tool for assessing the spatial distribution of thermal properties in the future. However, to fully realize the potential of this technique, more work is needed to explore solutions for these challenges - both for image acquisition and photogrammetric processing.

REFERENCES

Allahbakhshi, M., Shevchenko, A. V., Belousov, A. B., Belousova, M. G., Kämpf, H., & Walter, T. R. (2023). Geothermal Explosion at the 2014 Landslide-Covered Area of the Geyser Valley, Kamchatka, Russian Far East. *GeoHazards*, *4*(1), 60–76. https://doi.org/10.3390/geohazards4010005

Aragon, B., Johansen, K., Parkes, S., Malbeteau, Y., Al-Mashharawi, S., Al-Amoudi, T., Andrade, C. F., Turner, D., Lucieer, A., & McCabe, M. F. (2020). A Calibration Procedure for Field and UAV-Based Uncooled Thermal Infrared Instruments [Number: 11 Publisher: Multidisciplinary Digital Publishing Institute]. *Sensors*, *20*(11), 3316. https://doi.org/10.3390/s20113316

Assouline, D., Mohajeri, N., Gudmundsson, A., & Scartezzini, J.-L. (2019). A machine learning approach for mapping the very shallow theoretical geothermal potential. *Geothermal Energy*, *7*(1), 19. https://doi.org/10.1186/s40517-019-0135-6

Bjornsson, G., Grímsson, G., Sigurdsson, A., & Laenen, V. (2019). Thermal Mapping of Icelandic Geothermal Surface Manifestations with a Drone.

Brooks, A. (2022). *Renewable Energy Resource Assessment Information for the United States* (tech. rep. EXEC-2020-003533, 1855910, 8837). https://doi.org/10.2172/1855910

Budzier, H., & Gerlach, G. (2015). Calibration of uncooled thermal infrared cameras. *Journal of Sensors and Sensor Systems*, *4*, 187–197. https://doi.org/10.5194/jsss-4-187-2015

Bunker, J., Nagisetty, R. M., & Crowley, J. (2021). sUAS Remote Sensing to Evaluate Geothermal Seep Interactions with the Yellowstone River, Montana, USA [Number: 2 Publisher: Multidisciplinary Digital Publishing Institute]. *Remote Sensing*, *13*(2), 163. https://doi.org/10.3390/rs13020163

Campbell, J. B., & Wynne, R. H. (2011). *Introduction to Remote Sensing, Fifth Edition* [Google-Books-ID: NkLmDjSS8TsC]. Guilford Press.

Constenius, K. N. (1996). Late Paleogene extensional collapse of the Cordilleran foreland fold and thrust belt. *GSA Bulletin*, *108*(1), 20–39. https://doi.org/10.1130/0016-7606(1996)108⟨0020:LPECOT⟩2.3.CO;2

Dickinson, W. R. (2006). Geotectonic evolution of the Great Basin. *Geosphere*, *2*(7), 353–368. https://doi.org/10.1130/GES00054.1

Dlesk, A., Vach, K., & Pavelka, K. (2022). Photogrammetric Co-Processing of Thermal Infrared Images and RGB Images [Number: 4 Publisher: Multidisciplinary Digital Publishing Institute]. *Sensors*, *22*(4), 1655. https://doi.org/10.3390/s22041655

Faulds, J., & Hinz, N. (2015). Favorable Tectonic and Structural Settings of Geothermal Systems in the Great Basin Region, Western USA: Proxies for Discovering Blind Geothermal Systems, 6.

Faulds, J., Hinz, N., Coolbaugh, M., Ramelli, A., Glen, J., Ayling, B., Wannamaker, P., Deoreo, S., Siler, D., & Craig, J. (2019). Vectoring into Potential Blind Geothermal Systems in the Granite Springs Valley Area, Western Nevada: Application of the Play Fairway Analysis at Multiple Scales.

Faulds, J., Shervais, J., Wannamaker, P., Forson, C., & Lautze, N. (2021). Challenges and opportunities for geothermal exploration and hydrothermal research: Recent advances utilizing geothermal play fairway analysis in the Western USA. In *Symposium on the Application of Geophysics to Engineering and Environmental Problems 2021* (pp. 167–168). Society of Exploration Geophysicists; Environment; Engineering Geophysical Society. https://doi.org/10.4133/sageep.33-083

Harvey, M. C., Rowland, J. V., & Luketina, K. M. (2016). Drone with thermal infrared camera provides high resolution georeferenced imagery of the Waikite geothermal area, New Zealand. *Journal of Volcanology and Geothermal Research*, *325*, 61–69. https://doi.org/10.1016/j.jvolgeores.2016.06.014

Haselwimmer, C., & Prakash, A. (2013). *Thermal Infrared Remote Sensing, Sensors, Methods, Applications*. https://doi.org/10.1007/978-94-007-6639-6_22
ch 22 haselwimmer

Haselwimmer, C., Prakash, A., & Holdmann, G. (2013). Quantifying the heat flux and outflow rate of hot springs using airborne thermal imagery: Case study from Pilgrim Hot Springs, Alaska. *Remote Sensing of Environment*, *136*, 37–46. https://doi.org/10.1016/j.rse.2013.04.008

Irujo, G. (2022). IRimage: Open source software for processing images from infrared thermal cameras. *PeerJ Computer Science*, *8*, e977. https://doi.org/10.7717/peerj-cs.977

Janecke, S. (2007). Cenozoic extensional processes and tectonics in the northern Rocky Mountains. Retrieved September 3, 2023, from https://www.academia.edu/18384325/Cenozoic_extensional_processes_and_tectonics_in_the_northern_Rocky_Mountains

Jolie, E., Scott, S., Faulds, J., Chambefort, I., Axelsson, G., Gutiérrez-Negrín, L. C., Regenspurg, S., Ziegler, M., Ayling, B., Richter, A., & Zemedkun, M. T. (2021). Geological controls on geothermal resources for power generation [Bandiera_abtest: a Cg_type: Nature Research Journals Number: 5 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Energy access;Geochemistry;Hydrology;Structural geology;Volcanology Subject_term_id: energy-access;geochemistry;hydrology;structural-geology;volcanology]. *Nature Reviews Earth & Environment*, *2*(5), 324–339. https://doi.org/10.1038/s43017-021-00154-y

Kelly, J., Kljun, N., Olsson, P.-O., Mihai, L., Liljeblad, B., Weslien, P., Klemedtsson, L., & Eklundh, L. (2019). Challenges and Best Practices for Deriving Temperature Data from an Uncalibrated UAV Thermal Infrared Camera. *Remote Sensing*, *11*. https://doi.org/10.3390/rs11050567

Malbéteau, Y., Johansen, K., Aragon, B., Al-Mashhawari, S. K., & McCabe, M. F. (2021). Overcoming the Challenges of Thermal Infrared Orthomosaics Using a Swath-Based Approach to Correct for Dynamic Temperature and Wind Effects [Number: 16 Publisher: Multidisciplinary Digital Publishing Institute]. *Remote Sensing*, *13*(16), 3255. https://doi.org/10.3390/rs13163255

Molisee, D. D., & Bell, J. W. (2015). Structural Constraints of Buffalo Valley Hot Springs, North-Central Nevada, 8.

Newmann, J. R. (2019). *Structural Evolution of an Extensional Terrane Margin: Case Studies from the Colorado River Extensional Corridor, Southeastern California, USA* (Doctoral dissertation). UC Santa Barbara. Retrieved September 3, 2023, from https://escholarship.org/uc/item/1g9002nb

Nishar, A., Richards, S., Breen, D., Robertson, J., & Breen, B. (2016). Thermal infrared imaging of geothermal environments and by an unmanned aerial vehicle (UAV): A case study of the Wairakei – Tauhara geothermal field, Taupo, New Zealand. *Renewable Energy*, *86*, 1256–1264. https://doi.org/10.1016/j.renene.2015.09.042

Nugent, P. W., & Shaw, J. A. (2014). *Calibration of uncooled LWIR microbolometer imagers to enable long-term field deployment* (Vol. 9071) [Journal Abbreviation: Proceedings of SPIE - The International Society for Optical Engineering Publication Title: Proceedings of SPIE - The International Society for Optical Engineering]. https://doi.org/10.1117/12.2053082

Nugent, P. W., Shaw, J. A., & Pust, N. J. (2013). Correcting for focal-plane-array temperature dependence in microbolometer infrared cameras lacking thermal stabilization [Publisher: SPIE]. *Optical Engineering*, *52*(6), 061304. https://doi.org/10.1117/1.OE.52.6.061304

Processing thermal images. (2018). Retrieved October 17, 2023, from https://support.pix4d.com/hc/en-us/articles/360000173463-Processing-thermal-images

Reynolds, M. W. (1979). Character and Extent of Basin-Range Faultin, Western Montana and East-Central Idaho [Publisher: Rocky Mountain Association of Geologists]. Retrieved October 10, 2021, from https://archives.datapages.com/data/rmag/BasRanSym79/reynolds.htm

Robins, J. C., Kesseli, D., Witter, E., & Rhodes, G. (2022). 2022 GETEM Geothermal Drilling Cost Curve Update: Preprint. *Renewable Energy*.

Sears, J. W., & Fritz, W. J. (1998). Cenozoic tilt domains in southwestern Montana: Interference among three generations of extensional fault systems. In J. E. Faulds & J. H. Stewart (Eds.), *Accommodation zones and transfer zones; the regional segmentation of the Basin and Range Province* (p. 0). Geological Society of America. https://doi.org/10.1130/0-8137-2323-X.241

Sedano-Cibrián, J., Pérez-Álvarez, R., de Luis-Ruiz, J. M., Pereda-García, R., & Salas-Menocal, B. R. (2022). Thermal Water Prospection with UAV, Low-Cost Sensors and GIS. Application to the Case of La Hermida [Number: 18 Publisher: Multidisciplinary Digital Publishing Institute]. *Sensors*, *22*(18), 6756. https://doi.org/10.3390/s22186756

Siler, D., & Faulds, J. (2013). Three-dimensional geothermal fairway mapping: Examples from the western Great Basin, USA. *Transactions - Geothermal Resources Council*, *37*, 327–332.

Siler, D., Faulds, J., Hinz, N., Dering, G., Edwards, J., & Mayhew, B. (2019). Three-dimensional geologic mapping to assess geothermal potential: Examples from Nevada and Oregon. *Geothermal Energy*, *7*(1), 2. https://doi.org/10.1186/s40517-018-0117-0

Siler, D., Hinz, N., & Faulds, J. (2018). Stress concentrations at structural discontinuities in active fault zones in the western United States: Implications for permeability and fluid flow in geothermal fields [Institution: Univ. of Nevada, Reno, NV (United States); Ormat Technologies, Inc., Reno, NV (United States); Univ. of California, Oakland, CA (United States) Publisher: Geological Society of America]. *Geological Society of America, Bulletin*, *130*(7-8). https://doi.org/10.1130/b31729.1

Siler, D., & Pepin, J. (2021). 3-D Geologic Controls of Hydrothermal Fluid Flow at Brady geothermal field, Nevada, USA. *Geothermics*, *94*, 102112. https://doi.org/10.1016/j.geothermics.2021.102112

Silvestri, M., Marotta, E., Buongiorno, M. F., Avvisati, G., Belviso, P., Bellucci Sessa, E., Caputo, T., Longo, V., De Leo, V., & Teggi, S. (2020). Monitoring of Surface Temperature on Parco delle Biancane (Italian Geothermal Area) Using Optical Satellite Data, UAV and Field Campaigns [Number: 12 Publisher: Multidisciplinary Digital Publishing Institute]. *Remote Sensing*, *12*(12), 2018. https://doi.org/10.3390/rs12122018

Smigaj, M., Gaulton, R., Suarez, J. C., & Barr, S. L. (2017). Use of Miniature Thermal Cameras for Detection of Physiological Stress in Conifers [Number: 9 Publisher: Multidisciplinary Digital Publishing Institute]. *Remote Sensing*, *9*(9), 957. https://doi.org/10.3390/rs9090957

Snoke, A. W., & Chapman, J. B. (2021). North America: Central Cordillera. In D. Alderton & S. A. Elias (Eds.), *Encyclopedia of Geology (Second Edition)* (pp. 157–172). Academic Press. https://doi.org/10.1016/B978-0-12-409548-9.12124-4

Sondregger, J., Kovacich, S., & Bergantino, R. (1981). Geothermal resources map of Montana: Montana Bureau of Mines and Geology Hydrogeologic Map. Retrieved December 4, 2023, from https://www.mbmg.mtech.edu/mbmgcat/public/ListCitation.asp?pub_id=10352#gsc.tab=0

Štroner, M., Urban, R., Reindl, T., Seidl, J., & Brouček, J. (2020). Evaluation of the Georeferencing Accuracy of a Photogrammetric Model Using a Quadrocopter with Onboard GNSS RTK [Number: 8 Publisher: Multidisciplinary Digital Publishing Institute]. *Sensors*, *20*(8), 2318. https://doi.org/10.3390/s20082318

TLS-100 Transient Line Source Thermal Conductivity Meter. (n.d.). Retrieved October 18, 2023, from https://thermtest.com/tls-100

Vesselinov, V., Mudunuru, M., Ahmmed, B., Karra, S., & Middleton, R. (2020). Discovering Signatures of Hidden Geothermal Resources based on Unsupervised Learning.

Walter, T. R., Jousset, P., Allahbakhshi, M., Witt, T., Gudmundsson, M. T., & Hersir, G. P. (2020). Underwater and drone based photogrammetry reveals structural control at Geysir geothermal field in Iceland. *Journal of Volcanology and Geothermal Research*, *391*, 106282. https://doi.org/10.1016/j.jvolgeores.2018.01.010

Whitney, D. L., Teyssier, C., Rey, P., & Buck, W. R. (2013). Continental and oceanic core complexes. *GSA Bulletin*, *125*(3-4), 273–298. https://doi.org/10.1130/B30754.1