In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Sita Karki	Date

Geo-statistical and Remote Sensing Application for Thermal Research: Estimating Geothermal Heat Emittance Anomalies in Yellowstone National Park

by
Sita Karki

A thesis
submitted in partial fulfillment
of the requirements for a degree of
Master of Science in the Geographic Information Science
Idaho State University
Spring 2015

Committee Approval

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of **SITA KARKI** find it satisfactory and recommend that it be accepted.

Dr. Shannon Kobs Nawotniak, Major Advisor

Dr. Michael McCurry ,Committee Member

Dr. John Welhan, Committee Member

Dr. Chikashi Sato, Graduate Faculty Representative

TABLE OF CONTENTS

List of Abbreviations

AVIRIS	Airborne Visible/Infrared Image Spectrometer
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
DEM	Digital Elevation Dataset
ETM+	Enhanced Thematic Mapper Plus
IR	Infrared
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index
NDBSI	Normalized Difference Bare Soil Index
NDWI	Normalized Difference Water Index
NED	National Elevation Dataset
NIR	Near Infrared
SWIR	Short Wave Infrared
TM	Thematic Mapper
VNIR	Visible and Near Infrared

List of Figures

List of Tables

Abstract

I used Landsat 5 TM and ETM+ sensor data, including thermal infrared (TIR), to detect geothermal anomalies using multivariate regression. I calculated the contributions of background variables such as elevation, slope, aspect, shaded relief, vegetation and bare earth to the emitted energy in order to isolate true geothermal signals from the raw TIR. I trained the model around Yellowstone National Park (YNP) and tested it inside YNP. I applied my model in Coso and Tendaho geothermal regions in California and Ethiopia, respectively, to understand how it responds in areas different from YNP.

I determined the model coefficients by running the multivariate regression in training areas. The multivariate program was written using Fortran 90 code and was parallelized with openMP for faster solution. I calculated the background contribution using multivariate regression and Monte Carlo approaches to train the model to cool surrounding pixels. I subtracted the background from the raw emission calculated using thermal satellite bands and used standard deviation filtration to highlight the anomalies.

The multivariate model in YNP detected geothermal anomalies confirmed by earlier studies and highlighted features not mapped by earlier studies. Landsat ETM+ provided results with less noise, higher temperature and emission than Landsat TM 5. The developed model is robust and accommodates with the varying sizes of test area without changing the results significantly. The model results were broadly consistent with established and potential geothermal zones in Coso and Tendaho geothermal areas. This model is economical technique to detect geothermal anomalies over large areas. The future investigation includes rock and mineral indices and complex relationship among variables that could account for false positive anomalies.

Chapter 1

Introduction

Geothermal energy is potentially widely available in areas of young volcanism and tectonically active regions and is considered good for electricity production if temperatures exceed 302^{o}F (423 K) (Blackwell et al., 2007). It is a form of renewable energy that results from heat flow from within the interior of earth. Geothermal energy may also be economically viable for both direct use (e.g., to heat buildings) and for electrical power (DiPippo, 1991). It also reduces the demand for imported oil and decreases emission of carbon, sulfur and nitrogen oxides, and particulate matter into the atmosphere (Mock et al., 1997).

Field based studies and remote sensing applications have been used to locate and study the geothermally active areas. Satellite thermal infrared (TIR) from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors have been commonly used for geothermal resource exploration. Coolbaugh et al. (2006) and Eneva et al. (2006) used ASTER sensor coupled with field based studies to study geothermal areas. Hellman and Ramsey (2004) used ASTER and Airborne Visible/Infrared (IR) Image Spectrometer (AVIRIS) imagery to distinguish active and extinct springs in Yellowstone National Park. Savage et al. (2012) demonstrated the use of Landsat TM and ETM+ sensors. Vaughan et al. (2012) used ASTER and MODIS thermal infrared (TIR) images to monitor geothermal areas in the Yellowstone area and Watson et al. (2007) studied geothermal heat emittance using snowpack model inversion techniques and Landsat ETM+ sensor.

Satellite based remote sensing is less expensive than field based studies, as wide swath width and challenging topography can be covered in a single image. The quantification of geothermal energy emittance using remote sensing technology and statistical analysis has the potential to reduce cost by identifying the most promising potential geothermal targets for further field investigation.

A. Problem Statement

Background conditions have been a continuing problem for geothermal remote sensing applications (e.g., Hellman and Ramsey, 2004; Coolbaugh et al., 2006; Eneva et al., 2006; Watson et al., 2008; Savage et al., 2012; Vaughan et al., 2012;). Since all objects emit electromagnetic radiation (Campbell, 2008), the interference from background features cannot be ignored while studying geothermal distribution using remote sensing technologies. Previous studies (Savage et al., 2012; Vaughan et al., 2012; Watson et al., 2008) measuring geothermal energy in Yellowstone National Park (YNP) did not investigate in detail the effects of individual background variables. Coolbaugh et al. (2006) and Eneva et al. (2006) corrected thermal images by subtracting the heat due to topography, thermal inertia and albedo. They selected the variables to reduce the background temperature noise contributed by other environmental variables to reveal thermal anomalies in the final processed image. Vaughan et al. (2012) calculated net heat emittance by subtracting heat from non-geothermal surrounding with similar topographic and land cover characteristics. The determination of terrestrial emittance from geothermally active areas has been limited to the study of a few variables: slope, aspect, albedo and inertia. Also, the relative impacts were not calculated towards contribution to background, so old models are fairly inflexible. To solve this, I studied the contribution of each background variable including slope, aspect, elevation, shaded relief, and indices of vegetation and bare soil in selected training areas. By knowing the relative contributions of background variables in these training areas, I can develop a more adaptive model in other areas by more effectively eliminating their effects.

Study Area

I focus my study on YNP in western Wyoming, eastern Idaho and southwestern Montana (Figure 1) and apply the model developed in YNP to the Coso geothermal area in California and Tendaho geothermal area in Ethiopia. I selected YNP as my model development and testing area because of the relative abundance of data from previous studies (e.g., Watson et al., 2008; Savage et al., 2012; Vaughan et al., 2012), which makes it easier to compare and validate my results. YNP is one of the most active geothermal areas of the world. I selected the Coso and Tendaho geothermal areas because they are environmentally distinct from YNP, allowing me to evaluate the robustness of my model in different settings.

In this study, I developed a model for geothermal estimation of YNP by training the model via multivariate regression of data from geothermally cool areas around YNP. The model was subsequently used to investigate geothermal anomalies in YNP. I compared my modeled temperature with field measured temperature (Bergfeld et. al., 2011 and USGS, 2015) to understand the influence of pixel mixing and saturation. I applied the same model developed for YNP to examine Coso and Tendaho geothermal areas to investigate the 1st order robustness of my model in areas fundamentally different from YNP. To study the relative contributions of background variables to the total heat, training areas were chosen outside YNP, as shown in Figure 2. Data from IDWR and

USGS were used to identify non-geothermal training areas. No training areas were selected south of YNP in Wyoming because most of the area is classified as having geothermal potential, as identified by the Bureau of Land Management (BLM, 2012). The geothermal training areas were used to develop an algorithm for estimating emittance as a function of contributing variables.

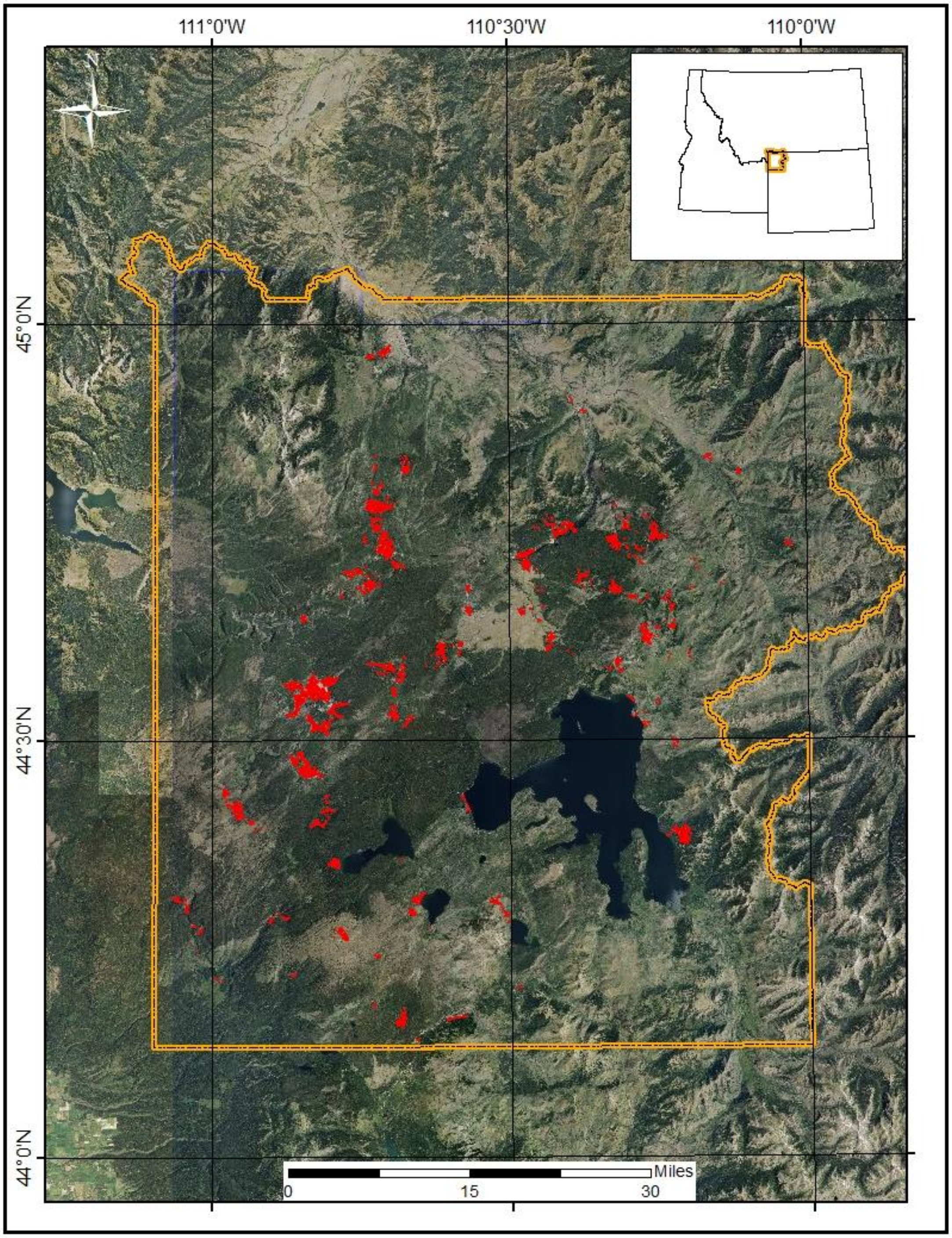


Figure 1: Known geothermal areas inside YNP (Data Source: Wyoming State Geological Survey; Background image: NAIP 2012 Image Service). The orange colored line is the YNP boundary and the red polygons are the geothermal areas.

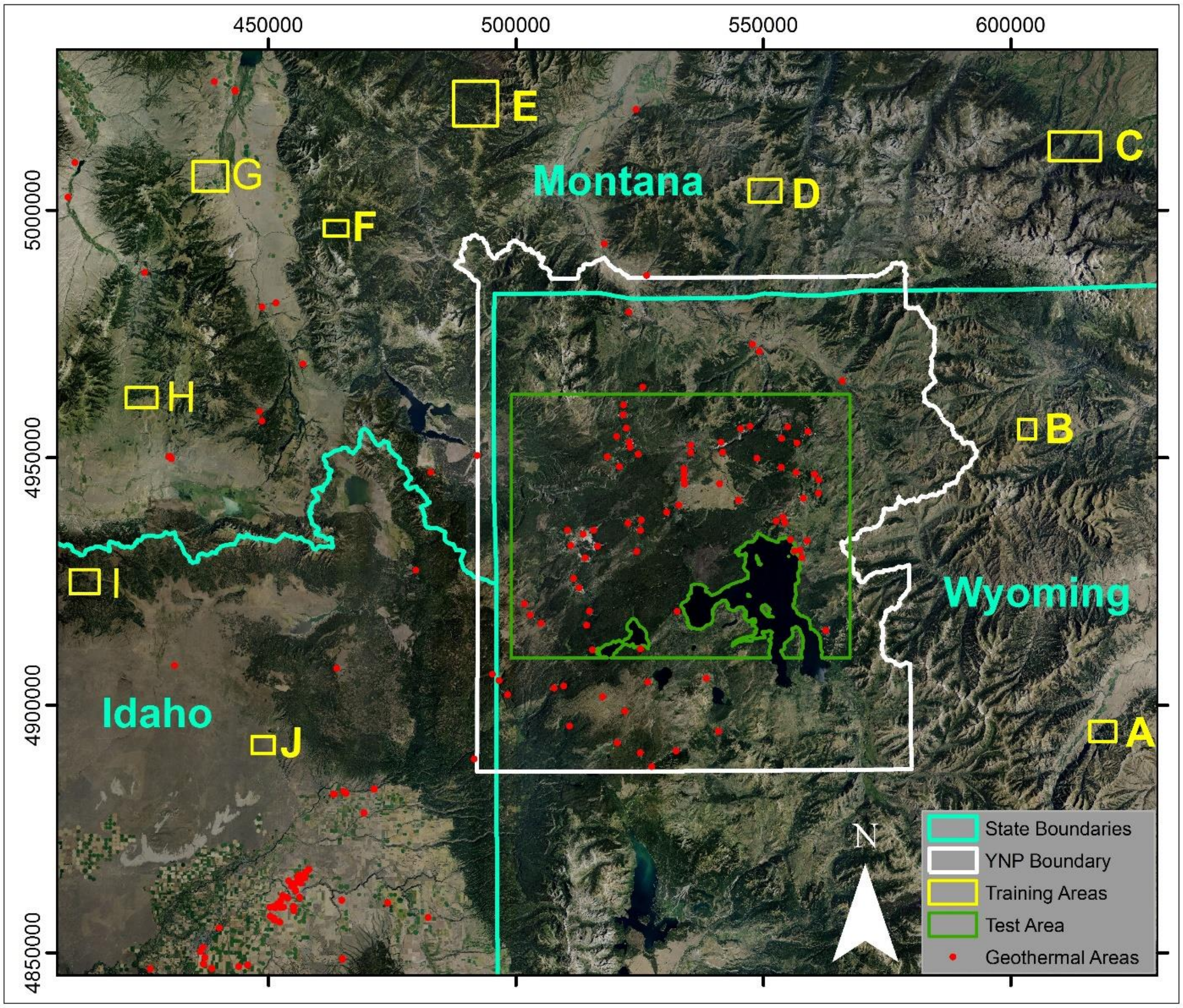


Figure 2: Map showing areas around YNP used as model training areas. The red points represent the geothermal features from IDWR and US Geological Survey (USGS); background image is from NAIP**,** 2012).

B. Objective of the Study

The objective of the study was to quantify, via multivariate regression, the relative effects of elevation, slope, aspect, insolation (incoming solar radiation), vegetation, water, soil moisture, and bare soil as independent variables that determine the total background emission. After determining the relationship between dependent and independent variables in the training areas, the same model was used to calculate emittance in YNP, Coso and Tendaho geothermal areas. The residual obtained by subtracting modeled from observed TIR values was used to identify areas of elevated emittance. I tested all of YNP to evaluate occurrence of false positives and false negatives in my approach and applied the model to Coso and Tendaho to evaluate its applicability in different geographic environments.

The hypothesis driving this study is:

Total TIR emittance is a result of all contributing variables on the earth's surface. As such, it is possible to identify geothermal anomalies by modeling background emittance from elevation, slope, aspect, vegetation, etc., and subtract the calculated background from the measured total emittance. The resulting residual image would comprise a background signal of randomly distributed low magnitude noise, while geothermal sources would be emphasized as areas of higher magnitude emittance.

This study attempted to isolate and identify TIR anomalies due to geothermal sources by analyzing the residual of the regression between emittance and background variables. The study differs from earlier studies in that it tries to quantify the relationship between thermal emission and background variables such as slope, vegetation, aspect and water. By taking into account the individual contribution, I developed a regression model that explained thermal anomalies and is more flexible in its application than previous studies that required direct background matching for successful subtraction.

This research takes advantage of preexisting satellite remote sensing technologies and helps to harness optimum benefit for resource exploration. I investigated the role of background variables and developed a model to detect geothermal anomalies over large area. I employed a geo-statistical approach to establish well-defined relationships with the background variables. This technique will increase the use of available remote sensor data without investing more financial resources in developing other platforms.

C. Organization of the study

As per the Master of GIS program thesis guidelines, my thesis consists of conventional thesis chapters such as introduction, literature review and conclusions in chapters 1, 2 and 5 respectively and two standalone papers in chapters 3 and 4. I used the first person pronoun “I” in these thesis chapters whereas the standalone papers follow the standard publication practice to use “we” to refer to the authors.

I studied YNP in two phase, first using Landsat TM and the second using Landsat ETM + sensors and analyzed the results in both phases. I present the first phase of the study in Chapter 3 and the second phase of the study and its application in other test areas in Chapter 4. These chapters represent the standalone publication document and contain some repeated background information. A glossary of terms and a copy of the final code are provided in the appendix. The code is also available for download from the ISU Geosciences website; it can be compiled and run from command line on Windows or Mac computers.

Chapter 2

Literature Review

Geothermal heat flux can be measured by installing shallow (1 – 100s of meters) thermal gradient holes, a high-quality but often expensive and time-intensive process. The investigation of large areas can be expedited by applying remote sensing technologies. Identifying geothermal features using satellite TIR imagery has the disadvantage that the desired signal is small relative to a large and variable background. This study focuses on quantifying emittance from geothermal features by accounting for background TIR emissions via multivariate regression. This approach can greatly reduce the contribution of background effects, making the model more robust.

A. Yellowstone National Park

YNP is a geothermally active area in northwest Wyoming (Fig.1). The geothermal area resides within a nested group of volcanic caldera, the most recent of which erupted 640,000 years ago (Christiansen, 2001), with small effusive eruptions dating to 70,000 Ka (Christiansen et al., 2007). The magma under the Yellowstone caldera is the source of crustal deformation and passive degassing (Aly and Cochran, 2011), with circulation of hot water through faults and fractures leading to the creation of over 10,000 thermal features at the ground surface, such as hot springs, geysers, mud pots, hot grounds of sizes ranging from centimeters to 10's of meters (Fournier, 1989). This makes this site ideal for developing and testing a TIS emittance regression model. The water temperature in these features ranges from near freezing to boiling, influenced by seasonal change, size of the feature, nature of the water pathways feeding the hydrothermal feature, etc.

B. Previous Work Using Geothermal Studies in Geographic Information System (GIS) and Remote Sensing to Detect Geothermal Features

Watson et al. (2008) studied geothermal emittance through snowpack model inversion techniques. In the study, residual terrestrial emittance was calculated after accounting for elevation, soil/bedrock, latitude/longitude, solar effects and soil/bedrock, assuming that only those variables contributed the background emission in geothermally passive sites. Watson et al. (2008) measured heat energy sufficient to melt peripheral snow to estimate the lower bound of heat emittance, and used the snow boundary to identify geothermal zones. Their calculated emittance ranged from 0 to 94 W/m^2, with good correlation between remotely sensed thermal anomalies and snowpack-inversion measurements ($R^2=0.82$). The consideration of more variables could have increased the model performance by eliminating more of the TIR background emission.

Vaughan et al. (2010) studied the saturation and pixel mixing thresholds for ASTER imagery where geothermal features are smaller. They identified many sources of uncertainty in sub-pixel thermal calculation such as pixel resampling, atmospheric correction, and background temperature and emissivity assumptions. The saturation and mixing detection thresholds were dependent on the percentage of hot target area and the radiometric temperature difference. The contribution of atmospheric correction to temperature uncertainty was highest followed by the contributions from emissivity and background assumptions. Validating satellite based thermal results with field-based measurements reduces these uncertainties and improves the sub-pixel modeling of thermal features.

Vaughan et al. (2012) used ASTER and MODIS TIR images to estimate heat emittance from geothermal areas in YNP using background subtraction. They selected the snowing season for their study to identify geothermal areas from non-geothermal areas. This approach has the risk of classifying snow free, windy and steep areas as geothermal areas that are not actually geothermal in nature. In that study, radiative heat emittance from proximal, non-geothermal areas with similar topographic and land cover characteristics of equal area was subtracted from total emittance to derive net emittance. Vaughan et al. (2012) assumed that thermal radiance emitted from non-geothermal areas is similar to the background heat from geothermal areas. There are several problems associated with this background emissivity subtraction technique: in a practical sense, it is difficult to find background areas with similar characteristics as test areas, and thermal areas will have little to no snow accumulation compared to background areas that are exposed to snow. Thus, the background emittance subtraction approach applied by Vaughan et al. (2012) is not appropriate in all conditions. Vaughan et al. (2012) estimated total heat emittance of around 2 GW (2.0 for ASTER and 2.3 for MODIS), which represents only about 30-45% of the geothermal heat emittance estimated from geothermometry and chloride flux balances (Fournier, 1989). One of the drawbacks of their study, however, is the poor spatial resolution of ASTER (90 m) and MODIS (1000 m) thermal pixels, which are unable to detect many small and/or low-temperature thermal features on such pixel scales.

Mia et al. (2012) used the Landsat 5 TM and ETM+ visible and near infrared (VNIR) bands to estimate vegetation cover as a proxy for ground temperature. They did not try to understand the contribution from background variables; instead, a topographically similar area on the same latitude (for solar radiation) was chosen 20 km west of the study area to estimate the background thermal contribution from soil, topography and aspect. They calculated NDVI-based spectral emissivity to detect bare soil, mixed land and stressed vegetation.

Warner and Chen (2001) attempted to suppress the effects of solar heating and topography in Landsat TM daytime thermal imagery of the Humboldt Range, Pershing, Nevada. They used a hyperspherical direction cosine (HSDC) transformation, with model correction, and statistical-empirical correction, but they could not find a single correction approach that could account for all the variables. In this sense, the use of multivariate regression has a clear advantage over individual variable correction techniques, and it also allows for cumulative effects to be modeled in a single equation.

Eneva et al. (2006) studied the geothermal features in the region between the Mammoth Geothermal Power and Coso Geothermal Power projects in California using ASTER thermal infrared (TIR-90m) imagery. They calculated net heat flux at the surface by considering variables such as albedo, topography and thermal inertia (the rate at which an object can gain or lose heat). They integrated heat flux over time to account for heat dissipation, creating a pseudo temperature image. This pseudo temperature image was subtracted from an ASTER temperature image (AST-08). The residual, or thermally corrected, image enhanced the thermal anomalies, reduced noise, and suppressed the false anomalies that appeared in the uncorrected image. The residual calculated by Eneva et al. (2006), therefore, accounts for the effect of albedo, topography and thermal inertia but not for other environmental variables like vegetation, bare soil and water.

Savage et al. (2012) examined the changing terrestrial emittance among spatial groups of thermal features from 1986 to 2007 in YNP. Relationships between thermal areas and factors such as such as distance to geologic faults, distance to large water bodies and distance to earthquake epicenters showed the strongest relationship with earthquake swarms, with 34% of the variation explained. Savage et al. (2012) indicated that Landsat imagery might be useful for monitoring geothermal responses in YNP, but it cannot be used as the sole monitoring tool because many geothermal areas are smaller than a Landsat TM pixel. Since ASTER (90 m) and MODIS (1000 m) thermal pixels are too coarse for the study areas, I used Landsat TM and ETM+ sensors.

Hellman and Ramsey (2004) used ASTER and Airborne Visible/IR Image Spectrometer (AVIRIS) imagery to distinguish active and extinct springs in Yellowstone National Park. Spring deposits have a distinct geometry and chemical signature, thus the study was done in VNIR, SWIR (Short Wave Infrared) and TIR bands. ASTER (spaceborne with 15, 30, 90 m pixel in VNIR, SWIR and TIR regions, respectively) and AVIRIS (airborne, hyperspectral, 20m resolution) images were analyzed to study reflectance, emissivity and temperature from thermal features, and band ratios 4/6 (for ASTER) and 139/195 (for AVIRIS) were used to identify minerals. Since, the AVIRIS data were acquired from airborne sources, they are not as consistently available as Landsat products, making it difficult to compare results over long periods of time and therefore were not selected for this study.

Sobrino et al. (2008) retrieved temperature and land surface emissivity from satellite and airborne sensors using a fractional vegetation cover (FVC) estimation technique and the temperature and emissivity separation. Retrieving surface emissivity is important to study land surface temperature and map mineral resources as it provides the native thermal signature of the objects. Traditionally, multispectral TIR has been used to calculate temperature and emissivity using the TES algorithm, but Sobrino et al. (2008) estimated emissivity using the normalized difference vegetation index (NDVI). The emissivity separation requires considerable information, such as atmospheric corrections, at least 4 TIR bands, and only works well in situations with high spectral contrast. The NDVI method can compensate these needs; it does not require accurate atmospheric correction and needs no or only one TIR band. The downside, however, is that *a priori* knowledge of soil emissivity is required, and the model cannot be applied to areas of water, snow, ice and bare rock. Because of these constraints, the NDVI approach is not applicable to the present study.

C. **Multivariate Regression in GIS and Remote Sensing**

Multivariate regression in geographic information systems (GIS) data processing has the ability to combine information from multiple layers to solve complex problems. Coolbaugh and Shevenell (2004) used regression tools to define favorable and unfavorable geothermal areas in Nevada using young volcanic rocks, earthquakes, global positioning systems (GPS) measurement of strain, northeast-trending young faults and regional gravity anomalies. Areas favorable for geothermal exploration were identified using digital maps of geothermal wells, temperature gradient holes, oil wells, water wells and depth to the water table. They did not consider background geothermal heat flux in their study because very few geothermal systems were known in the study area and because the inclusion of discovered geothermal systems based on subjective judgments contributed significant errors into their GIS-based predictions.

To avoid subjectivity in this study, the background TIR signal contribution calculated using multivariate regression was subtracted from the raw thermal image in order to detect thermal anomalies systematically. Multivariate regression can be used to develop a prediction model as well as reduce errors. Kunkel et al. (2011) developed a multivariate regression model for estimating soil and nitrogen stocks in semi-arid terrain in southwestern Idaho. They considered variables such as NDVI (vegetation index calculated using near infrared and red bands), insolation, precipitation, elevation, aspect and slope. They used univariate regression to investigate the relationship of individual variables to the dependent variable and to avoid multicolinearity (redundancy among variables). Though regression and cokriging produced similar results and were similar to field observations, the study found regression to be a simple and easy method of prediction at larger geographic scales. In contrast, kriging is more complex and cannot examine cause and effect relationships. Ranhao et al. (2008) demonstrated the use of interpolation to correct residual errors in multivariate regression. They performed multivariate regression to predict precipitation in the Daqing Mountains in northern China with latitude, longitude, and slope aspect as the determining variables. In my study, I investigated the dependency of temperature on several environmental variables to which kriging is not suitable because geothermal features do not represent continuous variables like snow depth, rainfall or temperature.

Chang and Li (2000) compared multivariate regression results with traditional interpolation methods such as Thiessen polygons, linear interpolation, inverse distance weighting (IDW) and kriging to construct snow water equivalent surfaces from snow course data by month in Idaho. Snow water equivalent was estimated using multivariate regression where topographic characteristics such as slope, aspect and curvature were the independent/determining variables. Multivariate regression outperformed the interpolations methods in the majority of the study area. In addition, interpolation techniques such as kriging or IDW are sensitive to the number and location of data points and therefore requires collecting numerous ground-temperature data, which is both impractical and expensive.

D. **Application of Multivariate Regression to Quantify Background TIR Emission**

The reason I considered multivariate regression to account for background TIR emission is because other methods, such as the normalization technique used by Warner and Chen (2001) and background area subtraction used by Mia et al. (2012), cannot quantify the effects of background influences in different locations. Warner and Chen (2001) showed that none of the normalization techniques are sufficient to suppress the effect of albedo and topography completely, so the multivariate regression approach taken in this study attempts to account for slope, aspect, elevation and shaded relief, vegetation and bare soil indices independent variables to avoid problems of variable topography and solar insolation.

This study took into account more background variables than earlier studies considered (e.g., Eneva et al., 2006; Watson et al., 2007) and determined geothermal heat emittance by subtracting background emission from total emission. Coolbaugh et al. (2006) and Eneva et al. (2006) took into account topography, albedo and thermal inertia but did not include variables like vegetation and water. Similarly, Weng (2009) accounted for vegetation and soil properties but did not consider other variables. Vaughan et al. (2012) did not study individual variables, but subtracted the emission from a geographically similar area nearby. None of these studies considered the individual determining variables or performed regression analysis to understand the effects on the dependent variable; instead, they calculated the contribution of a few limited variables and subtracted the effect of each individually. Since more variables were considered in this study, the resulting multivariate regression model is considered more robust.

Eneva et al. (2006) used albedo, atmospheric transmission and shaded relief to build a background heat energy model. Albedo was calculated using the satellite-derived reflectance after field correction. Since I am not collecting any field reflectance samples to correct my satellite measurements, I will not use albedo in my model. If I use satellite-derived albedo without field correction, it causes multicolinearity because albedo is a component of my dependent variable, temperature. I used slope and aspect to account for topographic variation and calculated shaded relief to account for brightness contributed by topography and insolation.

Savage et al. (2012) demonstrated that the use of Landsat TM and ETM+ sensor data, with 120 and 60 m spatial resolution in the thermal bands, respectively, are coarser to study small geothermal areas like those in YNP. Despite this, I used Landsat TM (120 m) and ETM+ (60 m) data to develop temperature image for my study area because these are the best available free thermal dataset to study YNP dominated with geothermal features ranging from centimeters to several meters. Landsat TM and ETM+ thermal pixels are finer than MODIS (1000 m) and ASTER (90 m) thermal pixels, which are more appropriate for studying large-scale geothermal features. The acquisition of finer resolution airborne thermal imagery is costly and comparative study over multiple years would be very difficult. The historical collection of Landsat TM (from 1982 to 2012) and ETM+ (1999 to 2003), makes it possible to compare my study to earlier studies done using the same sensors, as well as detect temporal changes and evaluate the results over long time spans.

I use satellite imagery from snow-free days in late summer to avoid the problem Vaughan et al. (2012) had in their study. They selected the snowing season for their study to identify geothermal areas from non-geothermal areas. This approach led to the inclusion of steep areas and snow-free wind swept areas that were not necessarily geothermal in nature. Since I do not use snow as a filter, I did not run into the problems due to snowpack accumulation.

Initially, univariate regression was used to identify and select the most statistically significant variables for multivariate regression. Kunkel et al. (2011) used univariate regression to investigate the relationship of individual variables to the dependent variable and to avoid multicolinearity. Though regression and co-kriging results were similar to field observations, Kunkel et al. (2011) found regression to be a simple and easy method of prediction at broader scales. In contrast, kriging is more complex and cannot examine cause and effect relationships. In my study, I want to understand the dependency of one variable over another, which is not possible through kriging. Further, I cannot use kriging because geothermal features are not continuous variables like snow depth or temperature.

Chapter 3

Determination of Geothermal Anomalies through Multivariate Regression of Background Variables at Yellowstone National Park using Landsat 5 TM Thermal Band Data

Sita Karki[1], Shannon Kobs Nawotniak[1]*, H. Carrie Bottenberg[1], Michael McCurry[1], John Welhan[1,2]
[1] Department of Geosciences, Idaho State University, Pocatello, ID 83209
[2] Idaho Geological Survey, Box 8072, Idaho State University, Pocatello, ID 83209
* Corresponding author: kobsshan@isu.edu

Keywords: *Landsat 5 TM, Yellowstone National Park, remote sensing, background subtraction, geothermal anomalies*

Abstract

Geothermal anomalies of Yellowstone National Park (YNP) are identified and quantified using Landsat 5 TM thermal band data. Multivariate regression of independent background variables that effect thermal emissivity, including elevation, slope, aspect, insolation, vegetation, water, soil moisture, and exposed land, were utilized in this study to create a comprehensive background filter for the raw imagery. Subtracting the multivariate background model from raw Landsat 5 TM data accentuates large geothermal anomalies such as Grand Prismatic and less thermally evident features such as the Old Faithful Geyser while removing significant false anomalies from the imagery. Geothermal anomaly emittances within YNP were calculated with a range of 40-120 W/m^2. False positives for geothermal activity were reduced in the scene, with remaining ones focused on bare earth slope, consistent with other studies. A differencing between known geothermal pool temperatures and model residual temperatures at 25 sites indicates an average difference of 347 K (stdev 12 K), suggesting scalability from residual output to corrected temperature detection. The methodology employed for

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

detecting known geothermal anomalies in YNP could be utilized to detect unknown geothermal potential in underexplored geothermal regions.

1. Introduction

All objects on the earth's surface emit electromagnetic radiation, which can be detected and measured using low-cost remote sensing techniques. Problematically, the raw emittance measured from satellite imagery includes the response from the target of interest as well as background or intervening features, which results in noise that can mask true anomalies. It is critical to differentiate signal sources when quantifying a specific thermal feature. In this work, we develop a multivariate background emittance model that we use as a filter for Landsat TM data across Yellowstone National Park (YNP).

YNP is one of the most geothermally active and well-studied locations in the world, allowing us to validate our model against established records of geothermal anomalies of varying scale. Yellowstone is a large caldera in Wyoming, Montana, and Idaho, located at the western extent of the Snake River Plain. The volcano that last erupted cataclysmically 640Ka, depositing the >1000 km^3 Lava Creek Tuff, an event that was followed by smaller effusive eruptions as recently as 70Ka (Christiansen and Blank, 1972). The magmatic system underlying Yellowstone caldera remains active, as evidenced by continued passive degassing and ground deformation (e.g., Aly and Cochran, 2011). Faults and fractures in the crust in YNP provide pathways for water circulation from the surface to deep, relatively hot crust (Morgan et al., 1977; Bargar, 1978). This hydrothermal circulation manifests itself at the surface as geysers, mud pots,

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

hot springs, and fumaroles; there are over 10,000 surficial geothermal anomalies within the park boundaries, with scales spanning centimeters to 10's of meters and temperatures up to the boiling point of water.

Despite the presence of such significant geothermal anomalies, thermal remote sensing analyses of the features have been hampered by noise from background emittance. Previous work by Coolbaugh et al. (2006) and Eneva et al. (2006) at YNP attempt to mitigate the noise problem in satellite-based remote sensing imagery by subtracting heat due to topography, thermal inertia and albedo. Similarly, Watson et al. (2008) calculated the geothermal emittance anomalies of YNP by correcting for solar and elevation effects using a snow covered Landsat 7 ETM+ scene to mask out other variables that emit thermal energy. While these approaches improve isolation of true positive thermal anomalies in satellite data, their success was limited by their *a priori* selection of a limited suite of contributing background variables. Following a different approach to noise filtering, Vaughan et al. (2012a and 2012b) calculated corrected net heat emittance at known geothermal sites in YNP by subtracting heat from nearby non-geothermal areas. While this approach yielded excellent results, the method is inherently limited to use in areas of pre-defined, known anomalies from which the user can identify targets of interest and appropriate neighboring non-geothermal pixels for subtraction. These limitations are problematic given the impact of geothermal anomalies on their surrounding conditions; for instance, elevated geothermal emittance can be very damaging to local vegetative health, which, in turn, influences overall thermal emittance as measured via remote sensing (Mia et al., 2012).

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

In this work, we evaluate the relative contributions of slope, aspect, elevation, vegetation, soil, and water to thermal emittance in geothermally inactive areas surrounding YNP to establish coefficient ranges appropriate for multivariate analysis. Minimum and maximum coefficients for each of the significant background variables define the solution bounds used in a Monte Carlo-based background filter for thermal anomalies within YNP boundaries. By defining coefficient ranges using non-geothermal zones, we reduce the risk of overfitting the algorithm for the thermal areas, thereby yielding false negatives in the final image.

2. Materials and Methods

We focused our study in and around YNP in northwestern Wyoming, eastern Idaho and southern Montana (Fig. 1). Training zones A through J represent areas similar to those within YNP but geothermally cold, as determined using data from the Idaho Department of Water Resources and Derkey and Johnson (1995). Diverse land conditions, such as steep slopes, vegetated areas and barren lands, were included to ensure model accommodation to wide range of environmental variability typical of the region. The independent variables were evaluated in the training areas in order to minimize overfitting of the model to the geothermal anomalies present inside YNP.

We used Landsat 5 TM satellite imagery (30 m spatial resolution, thermal resampled from 120 m), with imagery from September 24, 2011. The selected image did not contain snow or cloud cover, and was chosen to be outside of peak vegetation conditions; results from this image are consistent with output generated in August and October dates in other years. Although Landsat 7 ETM+ has 60 m resolution in the

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

thermal infrared bands, we did not use this sensor due to technical problems associated with it since 2003 (Sobrino et al., 2008).

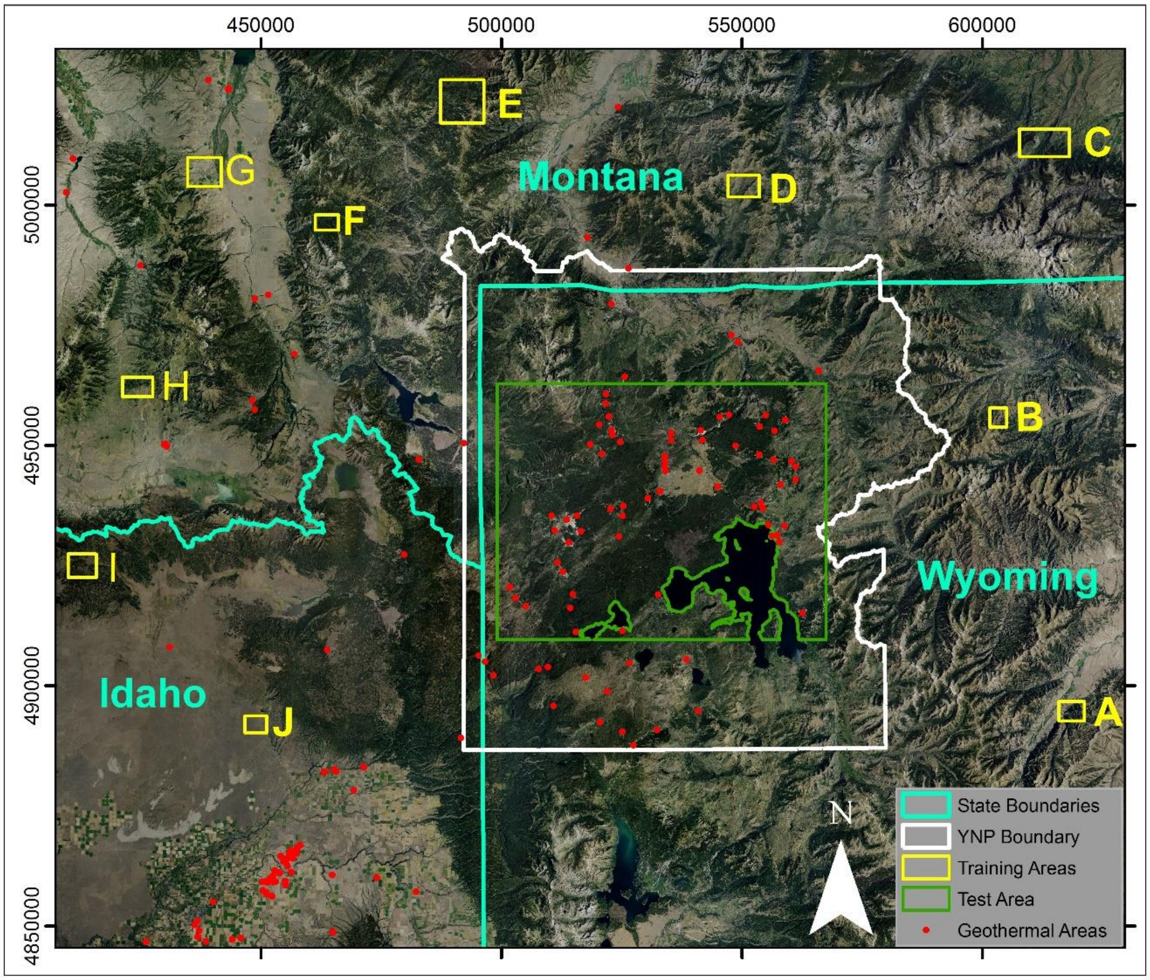


Figure 1: Map showing training areas (A to J; yellow boxes) around YNP (white outline) and test area (green outline) inside YNP. The red dots represent known geothermal anomalies. Points within the park boundary are from the polygons of R. Hutchinson (*unpublished*), points in Idaho are from Idaho Department of Water Resources (2001), and points in Montana are from Derkey and Johnson (1995). Background image from NAIP 2012/13 Image Services.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

We examined elevation, slope, aspect, insolation, vegetation, water, soil moisture, and exposed land as independent variables to calculate the background emission via multivariate regression. The primary datasets used for calculating these derived variables were Landsat 5 TM imagery and National Elevation Datasets (NED). The Landsat 5 TM data were converted to radiance and temperature (for band 6) in ENVI software using standard calibration parameters (NASA, 2007; Chander et al., 2009).

Normalized Difference Vegetation Index (NDVI) was used to measure vegetation greenness, a proxy for vegetation health and plant type from Landsat 5 TM imagery. NDVI values range from 1 to -1, with higher values representing more greenness. Near infrared (NIR) and red bands were used to calculate NDVI using the equation (Jensen, 1986):

$$NDVI = \frac{NIR - red}{NIR + red} \quad (1)$$

The Normalized Difference Bare Soil Index (NDBSI) uses Shortwave Infrared (SWIR) and NIR bands to measure bare soil area and it is expressed as (Roy et al., 1997):

$$NDBSI = \frac{SWIR - NIR}{SWIR + NIR} \quad (2)$$

NDWI, or Normalized Difference Water Index (NDWI) is used to delineate water features and enhance its presence in remotely sensed imagery (McFeeters, 1996). The equation for NDWI is given as:

$$NDWI = \frac{green - NIR}{green + NIR} \quad (3)$$

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

We also calculated the *modified*-NDWI, which suppresses the noise from built-up land, soil and vegetation because of the use of SWIR instead of NIR (Xu, 2006). We included both indices at the outset of the study rather than making a priori decision regarding which would be more useful in the multivariate thermal algorithm; variables were evaluated for redundancy and significance before being included in the final model. The equation for *modified*-NDWI is given as:

$$modifiedNDWI = \frac{green - SWIR}{green + SWIR} \tag{4}$$

National Elevation Datasets (NED; 10 m spatial resolution) were used to calculate slope, aspect, hillshade and insolation in ArcGIS. Slope is the rate of change of elevation with distance, ranging from 0 to 90 degrees above horizontal. Aspect, or the direction of that the local slope is facing, is recorded as azimuthal compass direction. Hillshade is a function of solar azimuth and elevation. This study used the metadata associated with the corresponding Landsat image to calculate the hillshade for the day of the year and time of day the scene was collected. Slope, aspect and hillshade were calculated using the algorithms by Burrough et al. (1998) native to ArcMap 10. Insolation, or solar radiation, was calculated using the hemispherical viewshed algorithm introduced by Rich et al. (1994) and developed by Fu and Rich (2000, 2002); the Area Solar Radiation tool in ArcMap 10 was used to calculate the insolation in Watt-hour/meter^2 at the time of day corresponding to the relevant Landsat image capture. Because date- and time-appropriate sun orientation was used in both the hillshade and insolation calculations, these variables are analogous to one another through they exist on different scales and use different units.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

We used univariate regression to establish the significance of the potential background variables on the total emittance, retaining significant variables for use in the multivariate calculation (Fig. 2). A *p*-test with 95% confidence identified background variables that do not significantly contribute to the total emittance; variables that failed the *p*-test for more than half of the training areas were removed from further evaluation in the study.

Independent variables were evaluated for multicollinearity, or variable redundancy. NDWI and *modified*-NDWI, for instance, are similar approaches to measuring vegetation greenness; while those terms will clearly exhibit multicollinearity, other relationships between independent variables may be less clear. To test for multicollinearity, variables were combined into groups for multivariate analysis in ArcMap 10. Variable groups were built sequentially, adding one variable at a time, with Ordinary Least Squares (OLS) used to identify improved coefficient fits. Variance Inflation Factor (VIF) values determined from best-fit solutions indicated which variables demonstrated multicollinearity (O'Brien, 2007). In cases of redundancy, the variable with the highest coefficient of determination was preserved while the others were excluded from further analysis. NDVI, NDWI, and *modified*-NDWI tested positive for redundancy relative to one another, with NDVI retained. Similarly, insolation and hillshade were redundant to one another; inclusion of hillshade resulted in better model fit, so insolation was dropped from further analysis. In both cases of multicollinearity, removal of the redundant variable(s) did not have a significant effect on the overall coefficient of determination.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

Figure 2: Figure showing the R-squared values corresponding to each background variables. Slope, elevation and NDVI have lower R-squared values than rest of the variables. NDBSI and shaded relief has the biggest influence on the total emittance as demonstrated by the color.

The univariate best-fit coefficients in the training zones for the remaining variables were used to create upper and lower bounds in the multivariate solution for YNP. By establishing coefficient bounds in geothermally cold but otherwise consistent zones, we restricted the degree to which the multivariate solution can overfit the geothermally active park. We did not use any predetermined weighting in selecting the coefficient ranges, as that would prejudice the model toward one set of land conditions over another. Large, cold water bodies were excluded from the final analysis, however, as they would otherwise force the model to preferentially fit to them rather than the targeted terrestrial sites. The multivariate model used a Monte Carlo approach to coefficient

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

selection within the established bounds. Given the size of YNP and the number of iterations necessary to converge on a stable solution, the multivariate solution was calculated using Fortran90 code with openMP for thread-scale parallelism. The resulting best-fit solution, which describes the background temperature, is subtracted from the original raw image to leave a residual that highlights the geothermal anomalies in YNP (Fig. 3). The residual image is converted to emittance from degrees Kelvin using the Stefan-Boltzmann equation.

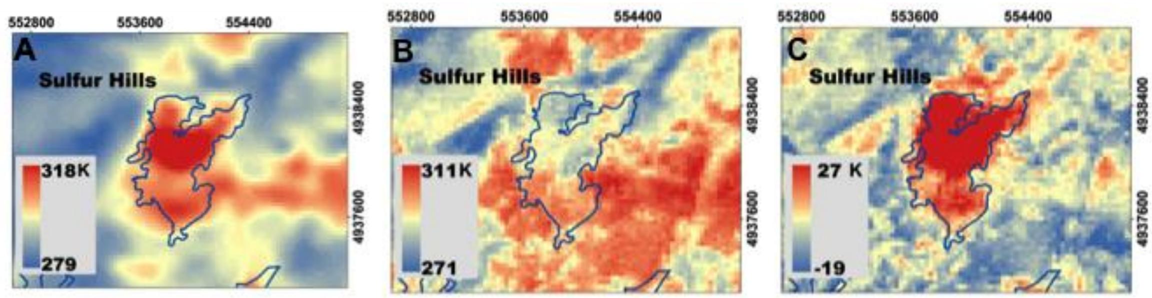


Figure 3. Close-up of the Sulfur Hills Thermal Area in YNP showing a) the raw Landsat 5 TM, b) multivariate background calculation, and c) residual (raw-background) images expressed in temperature (K). The blue polygons in the images are mapped geothermal zones by Hutchinson (unpublished). Subtracting the background from the raw image significantly highlights the geothermal anomaly relative to the false positive visible on the right side of the raw image. Location within YNP denoted on Figure 4; final output for Sulfur Hills in Figure 5e.

The residual image produced by this method contains significant low-level background noise in addition to the emphasized thermal anomaly. This noise is minimized by removing all pixels with values less than 3 standard deviations above the average emittance, leaving only large anomalies. This filter approach can only be applied

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

in circumstances where there are a large number of regular pixels relative to geothermally anomalous pixels, such as the park-wide analysis.

3. **Results and Discussion**

Application of the multivariate background model and 3 standard deviation filters results in very good agreement between modeled anomalies and field-evaluated geothermal anomalies across the park (Fig. 4). False anomalies occur in the output, mostly concentrated along northeast-trending ridges, as well as true anomalies that are not represented in the YNP polygon data but have been confirmed by Watson et al. (2008). The false anomalies that appear in the model output suggest that one of the variables may be underfitting the solution in certain circumstances, perhaps as a result of complex interplay between two or more variables, or that there may be a significant variable yet excluded from the analysis. Future work in this direction should include investigation of rock unit exposures, as they may be responsible for locally increased emittance.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

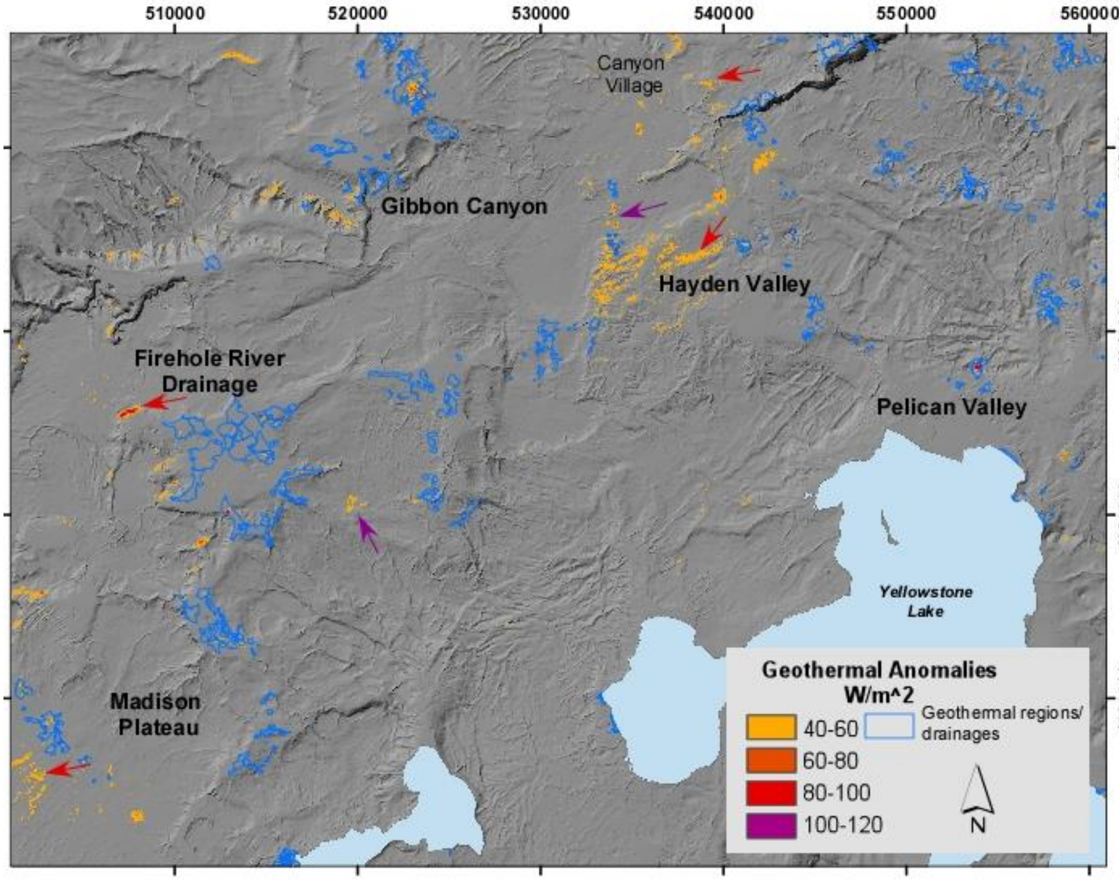


Figure 4: Residual emittance in YNP showing pixels greater than 3 standard deviations above the average for the zone. Blue polygons indicate geothermal zones as mapped by R. Hutchinson (*unpublished*). Red arrows indicate false anomalies in the image, while purple arrows denote positive identification of true anomalies consistent with ground truthing reports by Watson et al. (2008). False positives are preferentially located along northeast-trending ridges. Due to the scale of the test zone, presented here in overview, many of the positively identified thermal anomalies are not clearly visible in this image; see Figures 5 and 6 for closer views of anomalies.

The approach yielded particularly good fits with individual geothermal features within the anomaly polygons of R. Hutchinson (*unpublished*), available from the

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

Yellowstone Center for Resources GIS geodatabase (Fig. 5). For example, the model distinguished between Grand Prismatic Spring and Excelsior Geyser Crater in the Midway Geyser Basin while minimizing the surrounding runoff zones in the polygon. Similarly, the model identified individual anomalies at the Violet Hot Springs, including both spring and mud pot features, and various geysers and pools in the Norris Geyser Basin. While large hot springs are most readily visible in the residual imagery, terrestrial anomalies with relatively small footprints are also identifiable; Old Faithful and several other individual geysers are distinguishable from the background in the Upper Geyser Basin though it appears cold in the park-scale view (Fig. 6). The close-up perspective also shows some park infrastructure, such as buildings and parking lots, as positive anomalies just above the display threshold (Fig. 6).

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

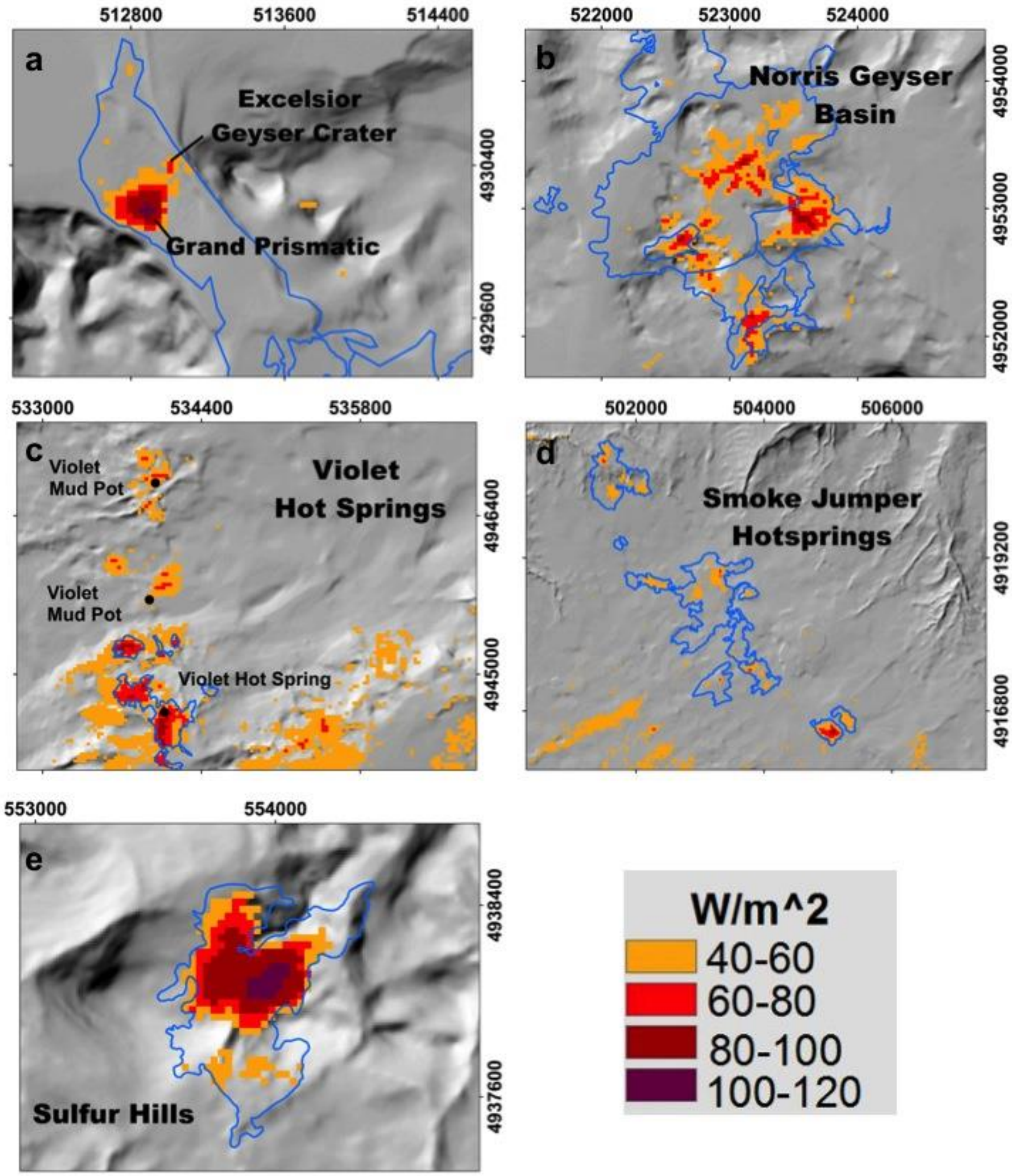


Figure 5: Close-up images of residual emittance for several geothermal anomalies in YNP. Features highlighted by the background filter model include hot spring pools, mud pots, and geysers. The method highlights the features within the broader mapped geothermal zones

This manuscript is published in *Geothermal Resources Council Transactions,* volume *38, pp.* 503-510.

denoted by the blue polygons (R. Hutchinson, *unpublished*). Locations within YNP are denoted on Figure 4.

By subtracting background values in order to highlight true anomalies this model unavoidably reduces the output temperature to degrees in excess of background rather than real temperature. Twenty-five field temperature measurements of YNP pools by Bergfeld et al. (2011) were compared with spatially coincident model output to determine if there was a baseline offset that could be applied to the residual model pixels to convert them to true temperatures (Table 1). For the imagery presented here, the average difference between measured and modeled temperatures was 347 K, with a standard deviation of 12 K. Given the reported variability of YNP geothermal features over time (e.g., Friedman and Norton, 1981; Vaughan et al., 2012a; Savage et al., 2012) and signal mixing in coarse Landsat 5 TM 120 m pixels, this is a narrow distribution of differences. Evaluation of imagery from other dates will be necessary to establish whether this coarse scaling change is broadly appropriate or is strongly influenced by intermediate diurnal and seasonal effects.

The background subtraction and 3 standard deviation pixel filter approaches used in this research pose a challenge for identifying relatively low-temperature thermal anomalies or anomalies with spatial footprints well under Landsat 5 TM pixel resolution. Comparing Figure 3c and Figure 5e, both of Sulfur Hills, illustrates the loss of low-grade thermal anomalies during the 3 standard deviation pixel filter used to minimize the visual impact of residual noise and low-confidence anomaly pixels. Further work will clarify the lower temperature and spatial limits of use for this model.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

The ability of the method to identify relatively small spatial features despite the coarse pixel size available for thermal data through Landsat 5 TM suggests that the approach is worth investigating at higher resolution scales, such as the 1m resolution Forward Looking Infrared (FLIR) surveys used by Jaworowski et al. (2010). In their work, Jaworowski et al. (2010) identified significant relationships between park infrastructure and geothermal anomalies, with road construction resulting in diverted hydrothermal runoff and elevated temperatures on pavement. This interaction, located outside of mapped geothermal polygons, is also visible in our results as an anomaly located directly northeast of the Overpass Group (Fig. 6). Their high-resolution imagery was able to capture features below the visible threshold in this study, including the Circle Pool group approximately 500 m southeast of Grand Prismatic Spring.

In contrast to the approach of Vaughan et al. (2012a and 2012b), which was designed for monitoring changes in YNP heat emittance, this method does not require *a priori* knowledge of geothermal anomalies and immediately proximal quiescent areas. As such, it can be more rapidly deployed as an exploratory tool over large areas. As presented, application of our proposed model requires identification of similar zones that are geothermally quiet for model coefficient training. This may be sidestepped, however, by processing large areas in which geothermally anomalous pixels make up a very small fraction of the total image, relaxing coefficient bounds, and allowing more iterations to achieve a convergent solution. In such an untrained case, the overwhelming number of cold pixels should prevent overfitting of the background model to the actual geothermal

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

anomalies; while conceptually sound, the untrained approach should be evaluated prior to extensive use.

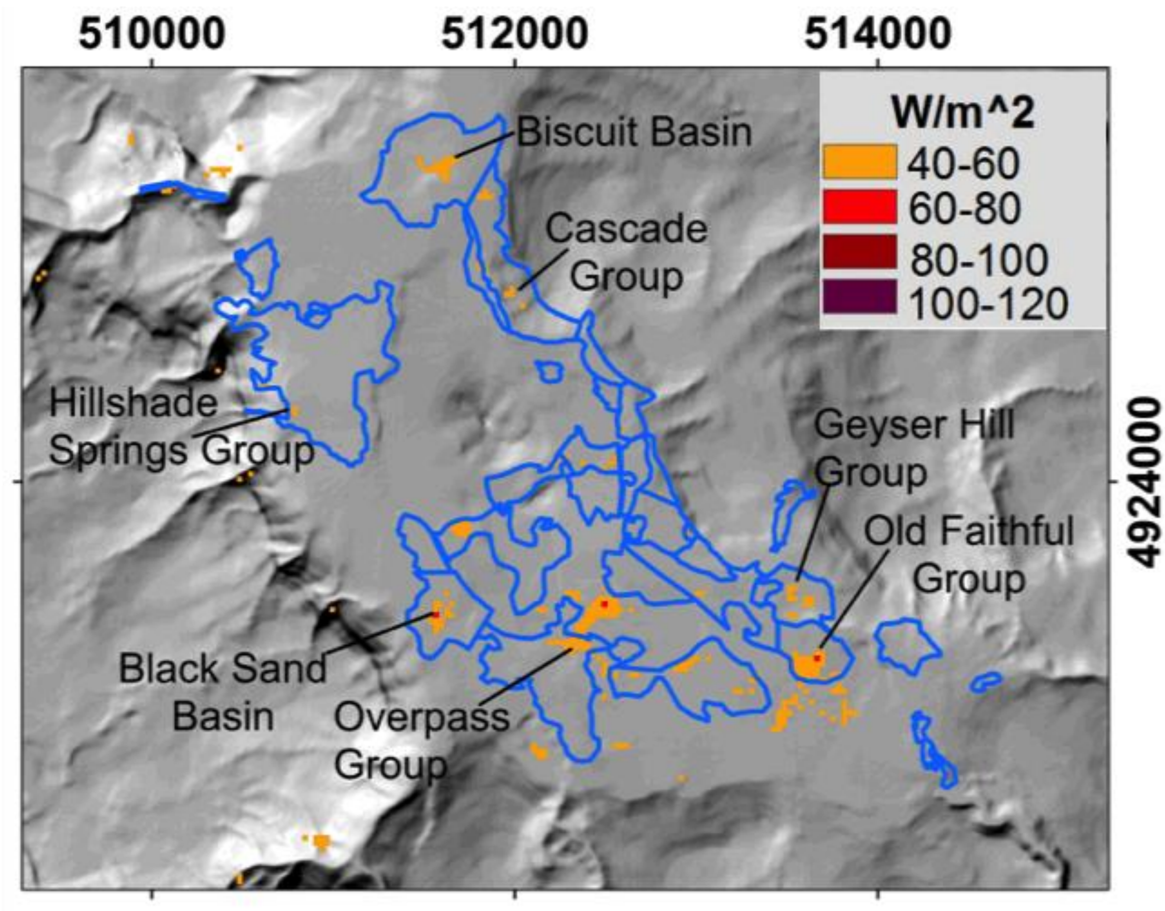
Table 1. Comparison points between measured field temperatures (Bergfeld et al., 2011) and residual temperature from this model. This table indicates an average baseline temperature of 347 K offsetting residual from measured temperatures in the processed image. Note: field measurements occurred during Augusts and Septembers during the years 2003-2009 and do not indicate fluctuations that may have occurred during that period.

Location	**Easting**	**Northing**	**Field Temperature (K)**	**Residual Temperature (K)**	**Difference (K)**
Back Basin 2	522963	4952193	340.4	5.1	335.3
Back Basin 4	523011	4952171	360.4	7.5	352.9
Bear Creek	558816	4932859	365.3	10.0	355.3
Behind Congress	523655	4952727	365.7	7.5	358.2
Black Pit	523588	4952139	355.3	1.7	353.6
Black Sands 1	511542	4923259	363	10.2	352.8
Black Sands 2	511628	4923190	349.1	3.1	346.0
Chocolate Pots	520496	4950780	325.5	2.6	322.9
Dishwater	523384	4952086	362.4	5.4	357.0
Green Dragon	523196	4951898	361.6	7.5	354.1
Hot Springs Basin 2	558553	4953761	364.9	2.0	362.9
Hot Springs Basin 5	558925	4955398	349.9	5.0	344.9
Hot Springs Basin 8	559347	4954788	341.1	8.2	332.9
Hundred Springs Plain	523113	4953330	362.1	3.5	358.6
NR Gibbon R1	523658	4954007	357	0.5	356.5
NR Gibbon R2	523680	4954101	345.8	4.5	341.3
Obsidian Pool	544530	4939794	362.2	5.2	357.0
Potts Basin 1	533421	4919761	341.8	1.0	340.8
Potts Basin 2	533505	4919689	360.5	-0.3	360.8
Potts Basin 3	533504	4919547	318.3	-5.0	323.3
Smokejumper 1	503793	4917530	358	12.2	345.8
Steam Valve	523494	4952561	341.8	2.4	339.4
Sulphur Caldron 1	544992	4941758	341.9	3.7	338.2
Terrace Springs	512184	4944102	336.3	10.0	326.3

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

W Nymph Lake Thermal Area 1	520335	4954609	355	4.9	350.1

Previous work by Vaughan et al. (2012a and 2012b) and Watson et al. (2008) use winter scenes in order to minimize the effects of intervening background emittance. While both clearly show the merits of this approach, their models are thereby limited in the regions in which they can be applied. While the multivariate background subtraction method presented here requires that more variables be constrained, it is accordingly more appropriate for use in areas without reliable winter snow accumulation.



This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

Figure 6: Upper Geyser Basin of YNP. Though the anomalies are indistinct when viewed at the park-wide scale in Figure 4, closer inspection of the Upper Geyser Basin reveals that the method discussed here identifies individual geysers within the group, including Old Faithful. Orange dots south of the Old Faithful Group correspond to park infrastructure, including buildings and parking lots. Blue polygons indicate geothermal zones as mapped by R. Hutchinson (*unpublished*).

4. Conclusions

The multivariate background subtraction method used in this study identified geothermal anomalies in YNP at multiple scales, from individual geysers to large hot springs and extensive geothermal anomaly clusters. This study is unique in that it successfully used multivariate regression analysis of Landsat TM 5 thermal infrared data to identify geothermal anomalies by developing a filter based on thorough explanation of background variables during snow-free conditions. By moving away from more traditional snow-filtering approaches, this model can be trained for use in potential geothermal areas in areas without regular snow accumulation. Future work will involve testing the inclusion of a geologic variable in the multivariate regression and investigation of complex relationships between independent variables that may contribute to false anomaly detection along northeast-trending ridges. The model will also be tested in geothermally active regions outside of YNP to evaluate robustness in different climate zones, with a focus on minimizing time required for coefficient training. Results from the

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

current model application in YNP suggest that it is a low-cost, solution for geothermal anomaly detection over large areas for both large- and small-scale features.

References

Aly, M.H. and Cochran, E.S., 2011. Spatio-Temporal Evolution of Yellowstone Deformation between 1992 and 2009 from InSAR and GPS Observations. *Bulletin of Volcanology* 73(9), 1407-1419.

Bargar, K.E., 1978. Geology and Thermal History of Mammoth Hot Springs, Yellowstone National Park, Wyoming. *Geological Survey Bulletin 1444*. 55p.

Bergfeld, D., Lowenstern, J.B., Hunt, A.G., Shanks, W.C.P, III, and Evans, W.C., 2011, Gas and isotope chemistry of thermal features in Yellowstone National Park, Wyoming: U.S. Geological Survey Scientific Investigations Report 2011-5012, 26 p. and data files.

Burrough, P. A., and McDonell, R. A., 1998. *Principles of Geographical Information Systems*, Oxford University Press, New York, 190p.

Chander, G., Markham, B., and Helder, D. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of the Environment*, 113, 893-903.

Christiansen, R.L., and Blank Jr., H.R., 1972. Volcanic stratigraphy of the Quaternary rhyolite plateau in Yellowstone National Park. *United States Geological Survey Professional Paper, 729-B*, p. 18.

Coolbaugh, M.F., Kratt, C., Fallacaro, A., Calvin, W.M. and Tranik, J.V., 2006. Detection of Geothermal Anomalies using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Thermal Infrared Images at Brady Hot Springs, Nevada, USA. *Remote Sensing of Environment*, 106, 350-359.

Derkey, P.D., and Johnson, B.R., 1995. Digital Maps of Low- to Moderate-Temperature Geothermal Spring and Wells in the Pacific Northwest: A Contribution to the Interior Columbia Basin Ecosystem Management Project. *US Geological Survey Open-File Report 95-689*, 11p. [online] Available at: < http://pubs.usgs.gov/of/1995/of95-689/ > [Accessed 20 October 2012].

Eneva, M., Coolbaugh, M. and Combs, J., 2006. Application of Satellite Thermal Infrared Imagery to Geothermal Exploration in East Central California. *GRC Transactions*, 30, 407-412.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

Friedman, I. and Norton, D.R., 1981. Ground Temperature Measurements: Part III, Ground Temperatures in and near Yellowstone National Park. *United States Geological Survey Professional Paper 120*, 23-39.

Fu, P., and Rich, P. M., 2000. *The Solar Analyst 1.0 Manual*. Helios Environmental Modeling Institute (HEMI), USA, 46p.

Fu, P., and Rich, P. M., 2002. A Geometric Solar Radiation Model with Applications in Agriculture and Forestry. *Computers and Electronics in Agriculture* 37, 25–35.

Idaho Department of Water Resources (IDWR), 2001. *Geothermal Resources*. [online] Available at: < http://www.idwr.idaho.gov/GeographicInfo/GISdata/geothermal.htm > [Accessed 20 October 2012].

Jaworowski, C., Heasler, H.P., Neale, C.M.U., and Sivarajan, S., 2010. Using Thermal Infrared Imagery and LiDAR in Yellowstone Geyser Basins. *Yellowstone Science* 18(1), 8-19.

Jensen, J. R., 1986. *Introductory Digital Image Processing*, Prentice-Hall, New Jersey, 379p.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. International Journal of Remote Sensing, 17(7), 1425-1432.

Mia, M.B., Bromley, C.J., and Fujimitsu, Y., 2012. Monitoring heat flux using Landsat TM/ETM+ thermal infrared data- A case study at Karapiti (‘Craters of the Moon’) Thermal Area, New Zealand. Journal of Volcanology and Geothermal Research. 235-236, 1-10.

Morgan, P., Blackwell, D.D., Spafford, R.E., Smith, R.B., 1977. Heat flow measurements in Yellowstone Lake and the thermal structure of Yellowstone Caldera. *Jounral of Geophysical Research* 82(26), 3719-3732.

National Aeronautics and Space Administration (NASA). *Landsat 7 Science Data Users Handbook*. Updated October 2007, [online] Available at: < http://landsathandbook.gsfc.nasa.gov/pdfs/Landsat7_Handbook.pdf> [Accessed 20 November, 2012].

O’Brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity,* 41, 673-690.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

Rich, P., Dubayah, M.R., Hetrick, W.A., and Saving, S.C., 1994. Using Viewshed Models to Calculate Intercepted Solar Radiation: Applications in Ecology. *American Society for Photogrammetry and Remote Sensing Technical Papers*, 524–529.

Roy, P.S., Miyatake, S. and Rikimaru, A., 1997. Biophysical Spectral Response Modeling Approach for Forest Density Stratification. *Proceedings of the 18th Asian Conference on Remote Sensing*, October 20-24, 1997, Malaysia. [Online]. Available at:<http://www.a-a-r-s.org/aars/proceeding/ACRS1997/Papers/FR97-7.htm>.

Savage, S. L., Lawrence, R.L., Custer, S.G., Jewett J.T., Powell, S.L. and Shaw, J.A., 2012. Analyzing Change in Yellowstone's Terrestrial Emittance with Landsat Imagery. *GIScience & Remote Sensing*, 49(2), 317-345.

Sobrino, J.A., Jiménez-Muñoz, J.C., Sòria, G., Romaguera, M. and Guanter, L., 2008. Land Surface Emissivity Retrieval from Different VNIR and TIR Sensors. *IEEE Transactions on Geoscience and Remotes Sensing*, 46(2), 316-327.

United States Department of Agriculture (USDA), 2012. *National Agriculture Imagery Program (NAIP).* [ArcGIS Image Service] Available at: < http://gis.apfo.usda.gov/arcgis/services > [Accessed 10 Oct 2012].

Vaughan, R.G., Keszthelyi, L.P., Lowenstern, J.B., Jaworowski, C. and Heasler, H., 2012a. Use of ASTER and MODIS Thermal Infrared Data to Quantify Heat Flow and Hydrothermal Change at Yellowstone National Park. *Journal of Volcanology and Geothermal Research*, 233–234, 72-89.

Vaughan, R.G., Lowenstern, J.B., Keszthelyi, L.P., Jaworowski, C., and Heasler, H., 2012b. Mapping Temperature and Radiant Geothermal Heat Flux Anomalies in the Yellowstone Geothermal System Using ASTER Thermal Infrared Data. *GRC Transactions* 36, 1403-1410.

Watson, R.G.R., Lockwood, R.E., Newman, W.B., Anderson, T.N. and Garrott, R.A., 2008. Development and Comparison of Landsat Radiometric and Snowpack Model Inversion Techniques for Estimating Geothermal Heat Flux. *Remote Sensing of Environment* 112, 471-481.

Xu, H., 2006. Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, 27(14), 3025-3033.

This manuscript is published in *Geothermal Resources Council Transactions*, volume *38, pp.* 503-510.

Chapter 4

Detecting Geothermal Anomalies in Yellowstone National Park using Landsat ETM+ Thermal Band Data and Application of the Model in Coso Geothermal Area in California and Tendaho Geothermal Area in Ethiopia

1. Introduction

Detection of geothermal anomalies in remote sensing data depends on segregation of background thermal emission from the actual anomaly signals. The contributions of background environmental variables such as vegetation or insolation to thermal emission vary over space and time, thereby requiring careful analysis to filter their effects from thermal remote sensing imagery. We use a multivariate regression model to quantify the effect of background variables and create a filter to remove their impacts from the raw thermal image. The resultant residual image, devoid of background-contributed false anomalies, displays the true anomalies present in the scene.

The nature of the residual and detection of geothermal anomalies is dependent on the background variables that occur in the study area. The effect of background variables is not uniform and the specific variables considered in the study determine the per-pixel value in the background image. The gross emission in the raw image and the deduction of variable background emission determine the magnitude and distribution of the residual (Fig. 1). The multivariate regression used to calculate the background either overestimates or underestimates the effect of background on a per-pixel basis; when the multivariate model accurately captures all of the contributing background emissions, this over- and underestimation should appear as random background noise.

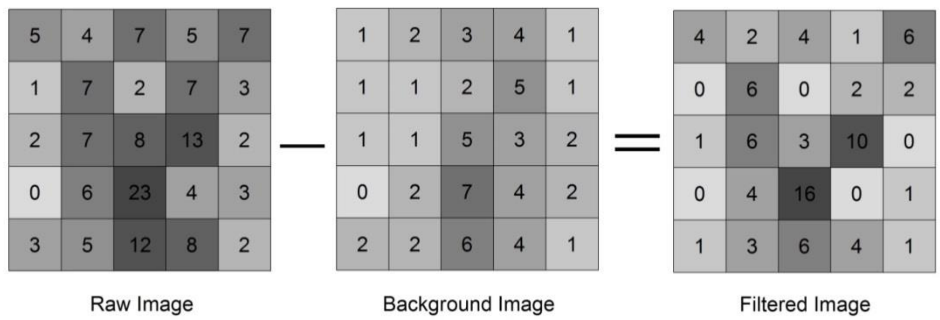


Figure 1: Illustration showing the background subtraction technique. Individual squares represent pixels and are numbered and shaded according to the signal strength associated with that position. The first image is the total heat emittance from thermal sensor, second is background image calculated from multivariate regression, and the third is residual image obtained after subtraction of background from the raw image. Different pixels are highlighted in the filtered image than in the raw image, due to the masking effects of the background noise that have been filtered out.

Background subtraction methods have been applied to geothermal remote sensing before with mixed success (e.g., Hellman and Ramsey, 2004; Coolbaugh et al., 2006; Eneva et al., 2006; Watson et al., 2008; Savage et al., 2012 and Vaughan et al., 2012). Significant limitations of past work have included using very few independent variables (Watson et al., 2008; Savage et al., 2012 and Vaughan et al., 2012) and requiring ground-covering snow in the imagery as a natural filter for background emissions (Watson et al., 2008).

Previous studies did not conduct model training in geothermally quiescent areas to quantify the relative importance of background variables; rather, the thermal image was corrected for the effect of thermal inertia, topography and albedo (e.g., Coolbaugh et al., 2006; Eneva et al., 2006). The albedo was calculated using visible and near-infrared ASTER bands, and the effect of topography was calculated using local slope. The field measured temperature and weighted average of day and night time ASTER thermal image were used to account for inertia.

Instead of quantifying the effect of individual background variables, Vaughan et al. (2012) calculated the total heat emittance at the geothermal sites and subtracted the emission from the surrounding non-geothermal area with similar topography and land cover characteristics. They assumed that the emission from surrounding non-geothermal area is similar to the background emission at the geothermal sites; this assumption is hindered by the known effects of geothermal anomalies on contributing background emittors, such as vegetation (Mia et al., 2012).

Earlier geothermal studies (e.g., Hellman and Ramsey, 2004; Coolbaugh et al., 2006; Eneva et al., 2006; Watson et al., 2008; Savage et al., 2012; Vaughan et al., 2012) conducted using satellite thermal bands either did not investigate the background variables (Hellman and Ramsey, 2004; Savage et al., 2012) or studied few background variables (Coolbaugh et al., 2006; Eneva et al., 2006; Watson et al., 2008) or subtracted the emission from surrounding environment without quantifying the effect of each variable (Vaughan et al., 2012). Our methodology differs from these approaches as we use training areas similar to the target zones to develop quantitative relationships between background variables and their contribution to raw thermal emissions, and then use those equations to adaptively filter background from the targets on a per-pixel basis.

We focus our study in and around Yellowstone National Park (YNP) in western Wyoming, eastern Idaho and southeastern Montana (Fig. 2). We selected YNP because of the number and range of geothermal features and the availability of data from previous studies for model comparison (e.g., Watson et al., 2008; Bergfeld et al., 2011; USGS, 2015). YNP is located at the eastern end of the Snake River Plain and covers an area of 8,983 km^2. The major volcanic events responsible for present physiographic features of the plateau occurred within the last million years, with voluminous ash-flow eruptions and caldera collapse ~640,000 years ago (Christiansen, 2001). The active magmatic system under Yellowstone causes degassing and ground deformation, while faults and fractures provide pathways for water circulation. There are more than 10,000 (Fournier, 1989) surficial geothermal features, including hot springs, mud pots and geysers, of varying size and temperature in YNP.

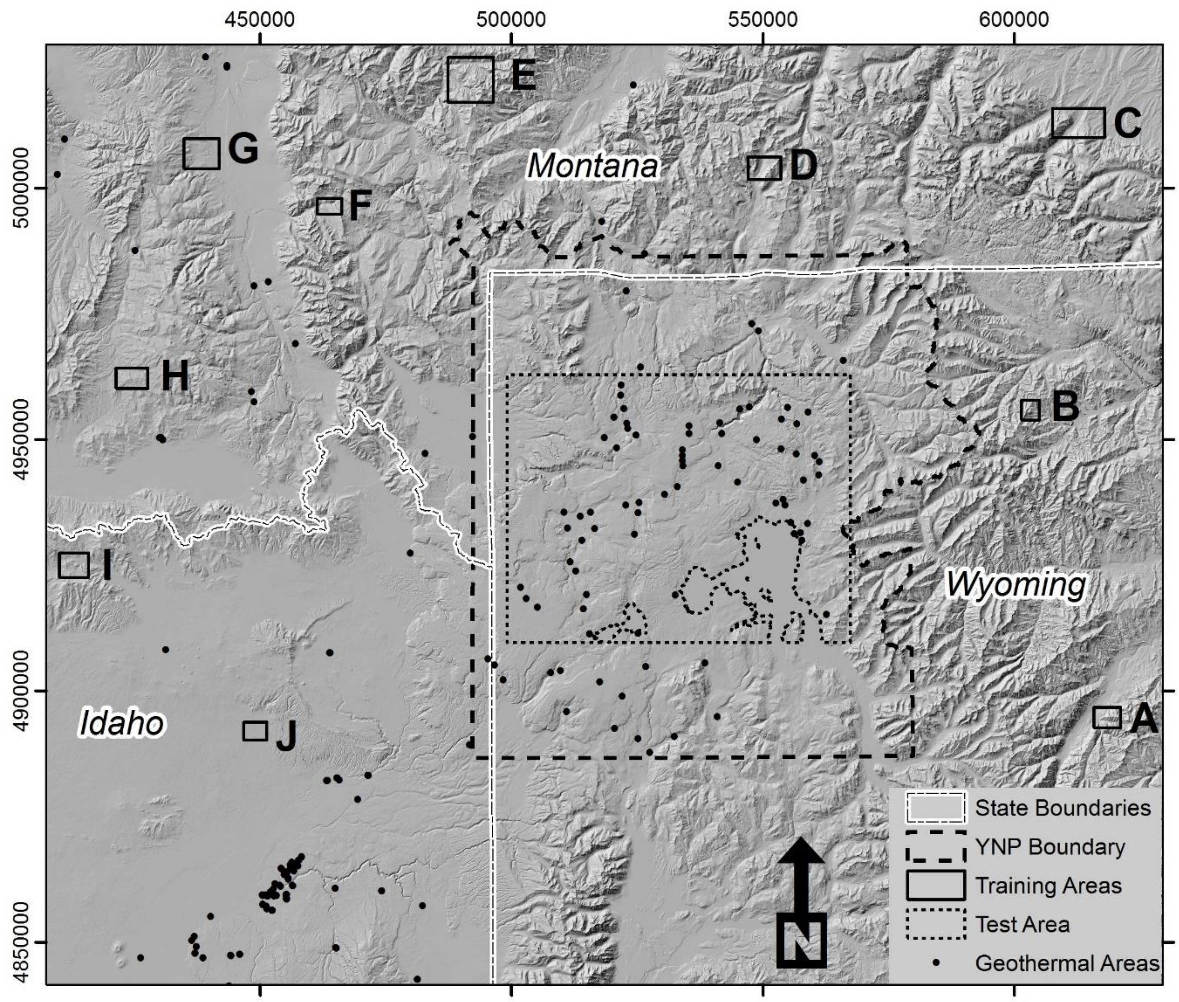


Figure 2: Map showing training areas (A to J) around YNP and test area inside YNP. The dots outside the park boundary represent geothermal feature locations from Idaho Department of Water Resources (2001) and US Geological Survey (USGS, 1995).

2. Materials and Methods

We selected training areas around YNP to determine quantitative relationships between independent background variables and measured heat (Fig. 2). The selected training sites represent geothermally inactive areas as indicated by the absence of geothermal features as mapped by Idaho Department of Water Resources (IDWR, 2001) and USGS (1995). We assume that the same background variables are significant in both training and test sites because of their proximity and similar environmental conditions. The training areas include barren lands, vegetated land, exposed slope, wetlands, and a variety of flat through rugged topography to make the model representative of a broad spectrum of land types.

Landsat ETM+ with 60 m thermal band resolution was used to study geothermal features in YNP as it provides greater detail than Landsat TM (120 m resolution) and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER, 90 m). The finer resolution of ETM+ resulted in less noise in the final image because of the decreased signal mixing in each pixel. ETM+ was available for a shorter time span than TM because of technical problems that began in 2003 (Sobrino et al., 2008). We chose late summer season to avoid peak vegetation saturation and used snow and cloud free images from 15 September 1999, 16 August 2000, and 23 September 2002. We used ETM+ thermal bands to calculate the raw temperature of the test area by using standard calibration parameters (NASA, 2007; Chander et al., 2009) in ENVI software. The temperature was used to calculate emittance using the Stefan-Boltzmann equation (Campbell, 2008).

Independent variables investigated

The environmental variables considered in the study were slope, aspect, shaded relief, curvature, bare soil, soil moisture, water and vegetation. From these independent variables, we selected a set of significant variables for model development (Fig 3), calculating the background emission using multivariate regression. Water, bare soil and vegetation indices were calculated using Landsat data, while digital elevation (10 m) data from USGS National Elevation Datasets (NED) was used to calculate slope, aspect, elevation, curvature and insolation values.

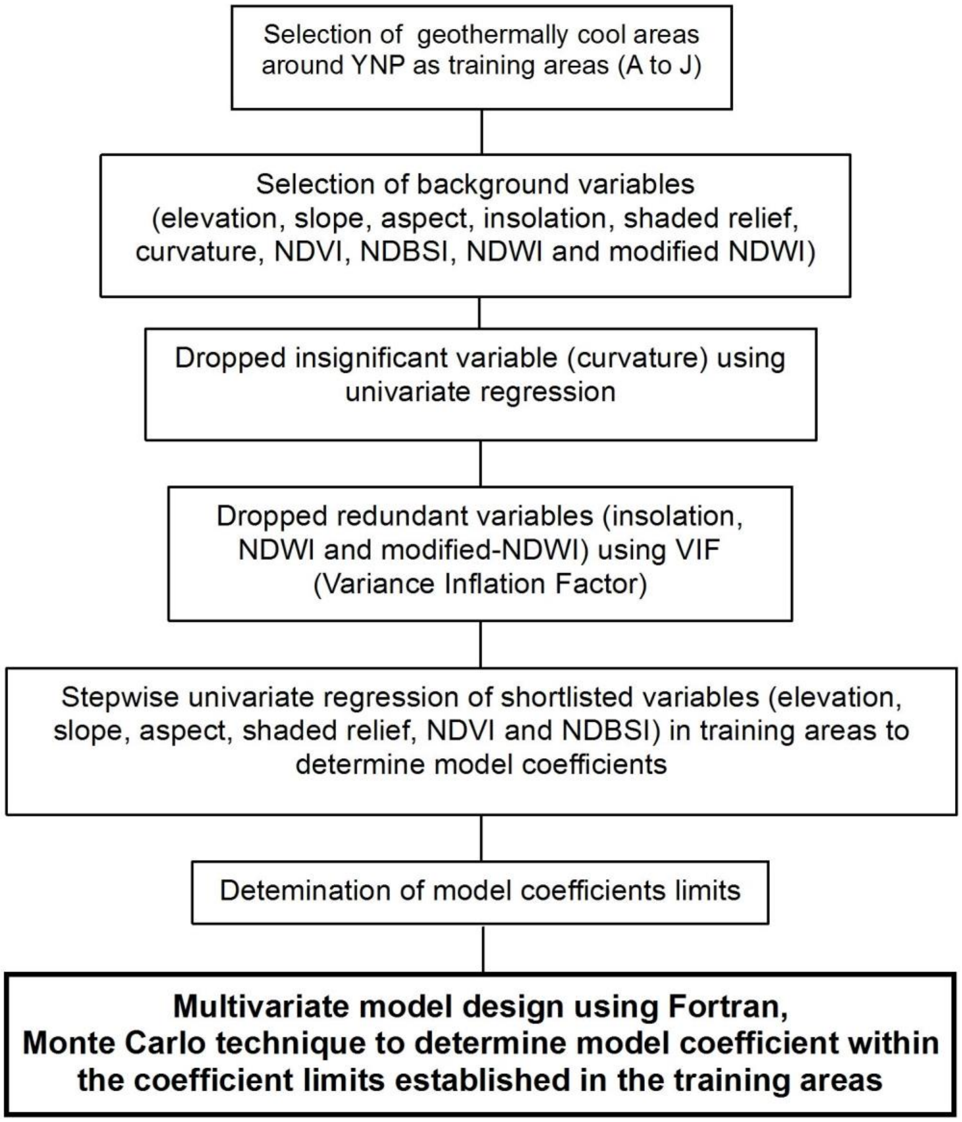


Figure 3: Methodology used in training areas to determine the significant variables. We started with 10 background variables and only 6 six variables were significant for final multivariate regression equation.

The Normalized Difference Vegetation Index (NDVI) is a proxy measure of vegetation greenness, health and plant type calculated using near infrared (NIR) and red bands in the equation by Jensen (1986):

$$NDVI = \frac{NIR - Red}{NIR + Red} \qquad (1)$$

Similarly, we used green and NIR bands to calculate the Normalized Difference Water Index (NDWI), which is used to delineate water bodies and enhance water presence through imagery (McFeeters, 1996):

$$NDWI = \frac{Green - NIR}{Green + NIR} \qquad (2)$$

We also computed *modified*-NDWI, which suppresses the noise from soil and vegetation with the use of Shortwave Infrared (SWIR) band (Xu, 2006):

$$modified\ NDWI = \frac{Green - SWIR}{Green + SWIR} \qquad (3)$$

NDWI and *modified*-NDWI were both included in the initial analyses, despite their similarity, in order to identify which had the stronger relationship with the dependent variable; the lesser performing one was later removed from further calculations.

Shortwave Infrared (SWIR) and NIR bands are used to calculate the Normalized Difference Bare Soil Index (NDBSI), which measures bare soil area and exposed land using the equation of Roy et al. (1997):

$$NDBSI = \frac{SWIR - NIR}{SWIR + NIR} \qquad (4)$$

The remaining independent variables were calculated using the 10 m DEM. Slope, aspect, shaded relief, insolation and curvature were calculated using ArcGIS software, which uses the algorithms by Burrough et al. (1998). Slope is the rate of change of elevation with the change of horizontal distance and its value ranges from 0 (horizontal) to 90 (vertical). Aspect indicates the direction that the sloping side is facing and varies from 0 to 360 degrees as compass direction. Shaded relief, or hillshade, represents the incoming solar radiation and shadow condition due to topography and is measured as a brightness value from 0 to 255; this was calculated using the same solar elevation and azimuth as represented in the corresponding Landsat scene. Insolation was calculated using the hemispherical viewshed algorithm by Rich et al. (1994) and Fu and Rich (2000, 2002) native to ArcGIS. For insolation, the solar radiation was calculated using the same sun orientation and elevation as the Landsat image acquisition and was expressed in Watt-hour/meter^2. Curvature is the second derivative of the surface and measures the slope of the slope. There are two types of curvature, profile and plan curvature; plan curvature measures the curvature in the direction of slope while plan curvature measures the curvature perpendicular to the direction of slope. A positive curvature indicates the surface is upwardly convex and a negative curvature indicates the surface is upwardly concave, a value of 0 indicates the surface is flat.

Univariate evaluation

We applied univariate regression using the ordinary least squares (OLS) tool in ArcGIS software to determine the significance of each variable's correlation to thermal emission. Temperature was the dependent variable**;** independent variables were evaluated one at a time in the OLS tool. Only those variables found to have a significant correlation with temperature (95% confidence p-test) were retained for multivariate regression. All of independent variables passed the significance test except for curvature (both profile and plan), which was removed from further consideration. Each variable found to be significant through univariate regression was then plotted against the dependent variable temperature to understand the actual nature of the relationship. Since aspect is expressed as an angle (0 to 360 degrees) from North, it has a quadratic relationship while all the other variables have linear relationships with temperature.

Redundancy

The Variance Inflation Factor (VIF; O'Brien, 2007) was used to identify redundant variables. The multivariate regression was run using the OLS tool in ArcGIS software with temperature as the dependent variable and all other variables as independent variables, returning VIF values. Variables with a VIF >7.5 are considered redundant while VIF less than 7.5 are considered significant while using OLS in ArcGIS software. The variable with highest VIF was dropped one at a time and OLS tool was run again to identify other redundant variables. In the situation where more than one variable had VIF values higher than 7.5, the one with highest VIF was removed first as it was redundant over rest of the variables. The test was continued until only variables with VIF less than 7.5 were left. NDWI and *modified*-NDWI were found to be redundant over NDVI so they were removed from further analyses. Similarly, insolation was redundant over shaded relief as it measured the same solar radiation in a different form as shaded relief. Six variables (Fig. 4) passed the tests of significance and redundancy and were considered in the final multivariate model.

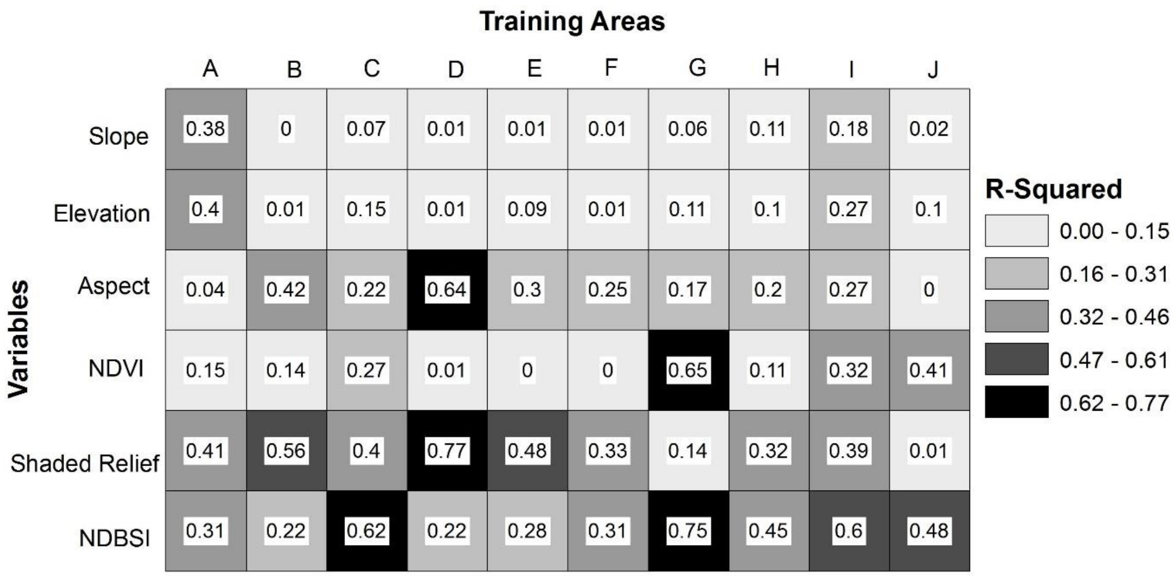


Figure 4: Figure showing the R-squared values corresponding to the background variables. Slope, elevation and NDVI have lower influence while NDBSI and shaded relief have the more influence on the dependent variable as indicated by the darker color. Curvature, not included here, was not significant. NDWI, modified-NDVI, and insolation were removed from further investigation due to redundancy and are thus also not included here.

Determining coefficients

The independent variable's coefficient represents its influence on the dependent variable. We ran the univariate regression in all the training areas to determine the coefficient for each variable. These coefficients specific to the training areas were used to identify the range of coefficients that define the bounds of the multivariate model. Since slope aspect has a quadratic relation with temperature, it has two coefficients while all the other variables have only one. The range of univariate regression coefficient values is shown in Table 1 for each independent variable. By establishing the values of these coefficients in the geothermally inactive areas, we tried to mitigate potential model overfit for application of the model in geothermally active areas. Determining the ranges of coefficient values in the training areas minimizes the potential false anomalies that could be caused by anomalous behavior of some background variables.

Table 1: Coefficient for each variable for the training areas. The minimum and maximum for each coefficient was used for as coefficient interval for multivariate model generation. There are two sets of coefficients for aspect as it has quadratic relationship with the temperature.

Training Areas	Coefficients for Independent Variables							
	Slope	Elevation	Aspect1	Aspect2	NDVI	Hillshade	NDBSI	Intercept
A:Hilly, sparse vegetation	-0.02037	-0.00285	0.01977	-0.00005	-3.777	0.00976	50.750	338.410
B: Mountain Valley, sparse vegetation	0.12190	0.00180	-0.01926	0.00004	-10.330	0.06808	61.430	327.275
C: Hilly, soil and exposed surface	0.00956	-0.00492	0.00201	0.00001	8.746	0.05397	61.660	346.223
D: Rocky surface, no vegetation	0.08805	0.00146	-0.04258	0.00009	6.034	0.07942	57.050	319.895
E: Hilly forest	0.06343	-0.00253	0.00021	0.00000	9.084	0.04175	37.190	319.229
F: Hilly forest	0.04575	0.00147	0.00013	0.00000	3.926	0.02439	25.800	302.889
G: Agricultural lands	0.01205	0.00228	0.00671	-0.00003	0.555	0.06236	51.300	320.935
H: Low lands, sparse vegetation	-0.07244	0.00277	0.01246	-0.00003	7.421	0.08450	49.530	311.485
I: River basin, medium vegetation	0.04394	-0.01439	0.01388	-0.00004	7.009	0.04021	33.480	338.165
J: Flat barren lands	-0.01431	-0.03677	0.00195	-0.00001	0.110	-0.01963	36.900	394.502
Maximum	0.12190	0.00277	0.01977	0.00009	9.084	0.08450	61.660	394.502
Minimum	-0.07244	-0.03677	-0.04258	-0.00005	-10.330	-0.01963	25.800	302.889

Multivariate model

Using the maximum and minimum value of each coefficients we developed for the training areas, we employed a Monte Carlo technique to generate the multivariate model (appendix) for the test area:

Background temperature=[*Slope Coefficient**Slope]+[*Aspect Coefficient 1**Aspect]+[*Aspect Coefficient 2**Aspect^2]+[*Hillshade Coefficient**Hillshade]+[*Elevation Coefficient**Elevation]+[*NDVI Coefficient**NDVI]+[*NDBSI Coefficient**NDBSI]+[*Intercept*] (5)

The Monte Carlo technique randomly selects values of each coefficient between the minimum and maximum values for each independent variable in Table 1. The model pairs the coefficients with each background variables to calculate the background temperature. The user has to specify the input text files containing the raw temperature and background variables and the program outputs the text file containing the multivariate equation. We determined the optimum number of iterations by increasing the number of iterations until the point of no further improvement, or stable solution. The number of iteration that provided stable solution was ~100,000 for our location. Among numerous multivariate equations with all possible combinations of coefficients, the model picks the one that results in the least residual (calculated temperature – raw temperature) for the test area. To minimize the time expense of so many iterations, the code was parallelized via openMP for faster solutions on desktop or laptop-scale computers.

Running the multivariate model in the test area

The inputs for the multivariate model are the background variables and the raw temperature of the test area. The model generates the multivariate equation appropriate for the test area that explains the background temperature (Fig. 5). The coefficients in the multivariate output equation were not pushing against the boundary limits for any of the variables, indicating that the coefficient ranges established in the training zones were not overly restrictive for the test area. We plugged in the value of background variables in the resultant multivariate equation and calculated the background temperature for the test area. This background image explains the effect of all independent variables considered in our study and is used to calculate residual temperature. We subtracted background image from the raw temperature image to calculate the residual image which is devoid of background contribution. The magnitude and distribution of the emittance in the residual pixels provided the basis for detecting thermal anomalies. We converted the temperature scale measurements to emittance using Stefan-Boltzmann equation. We ran our model in the test area about half the size of YNP (Fig. 2). The number of pixels required for appropriate fit is relative to the number of pixels anticipated to represent anomalies in the image; geothermal anomalies should make up a small fraction of the total space in the image to be able to see the contrast between cold and hot areas.

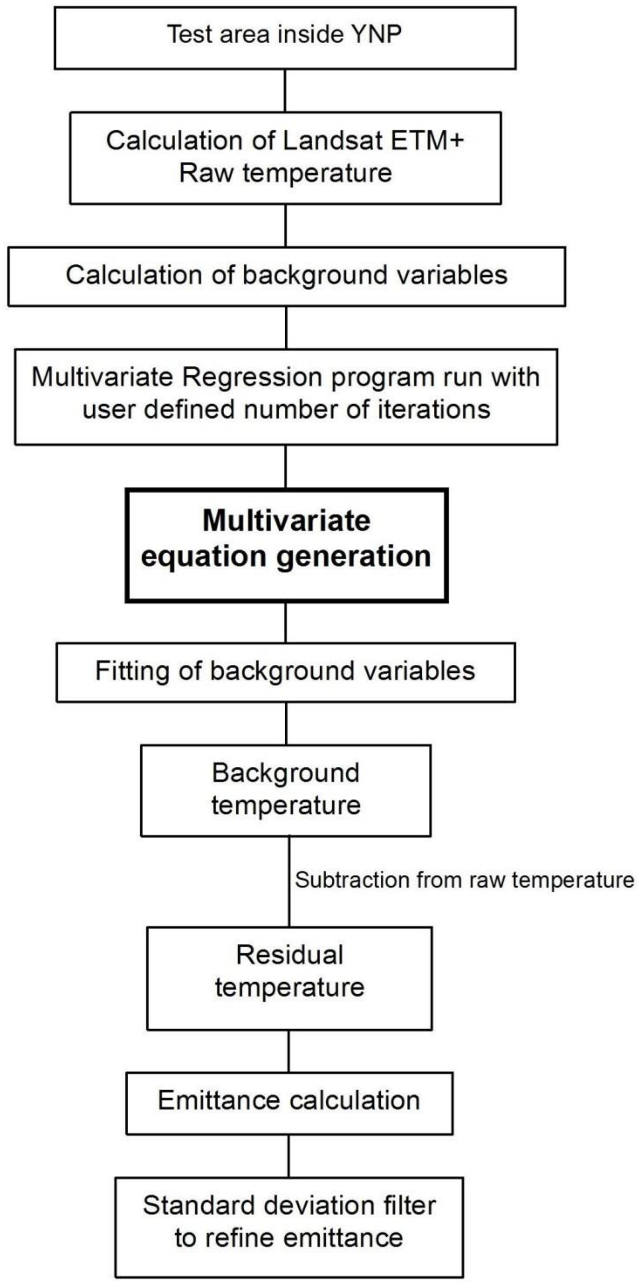


Figure 5: Methodology applied in test area to determine geothermal anomalies. The multivariate model used the background variables to generate background temperature for the test area.

Emittance was calculated using residual temperature which is the difference between raw and background temperature. The standard deviation filter was used to help refine geothermal anomalies.

3. Results and Discussions

The multivariate equation (Eqn. 5) generated for the test area gives the measure of background temperature explained by the environmental variables. The background temperature is subtracted from the Landsat raw thermal temperature (Fig. 6) to calculate the residual temperature. The residual temperature image shows the temperature above the background temperature caused by factors not considered in background calculation and could be the geothermal anomalies.

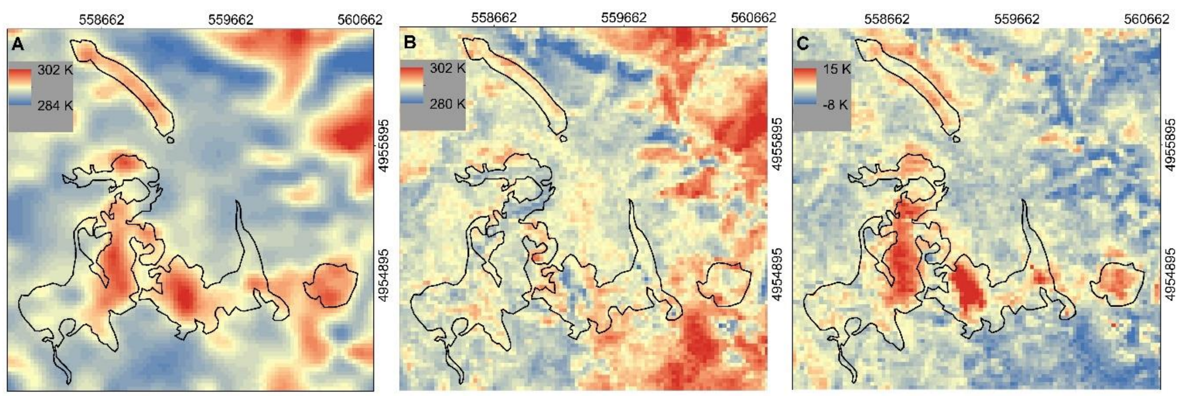


Figure 6: Close-up of the Hot Spring Basin in YNP showing a) raw Landsat 5 ETM+, b) background, and c) residual (raw-background) images expressed in temperature (K). The black polygons in the images are mapped geothermal zones by R. Hutchinson (*unpublished*). Subtraction of background from the raw image significantly removes the false positive areas present in the raw image and highlights the true anomalies.

The temperature was used to calculate emittance using the Stefan-Boltzmann equation (Campbell, 2008) for three late summer Landsat images considered in our study.

We averaged the emittance of the three images on a per-pixel basis to minimize small temporal fluctuation and detect geothermal anomalies that were present in all the scenes. This averaging technique increased the confidence of output by highlighting temporally persistent anomalies and suppressing transient anomalies.

The averaged residual image shows the emittance from geothermal as well as low-level non-geothermal noise (Fig. 7). For visualization purposes, we used a standard deviation-based filter to remove the low-level background noise, starting with one standard deviation and increasing until we reached the point where only the hot pixels were displayed. The filtered images presented in this work use a 3 standard deviation filter, similar to that used by Watson (2008); lower thresholds increased the inclusion of background noise in the image while higher thresholds removed anomalous pixels. The standard deviation filter was used for visualization purposes only, and was not used for any numerical alteration of the images.

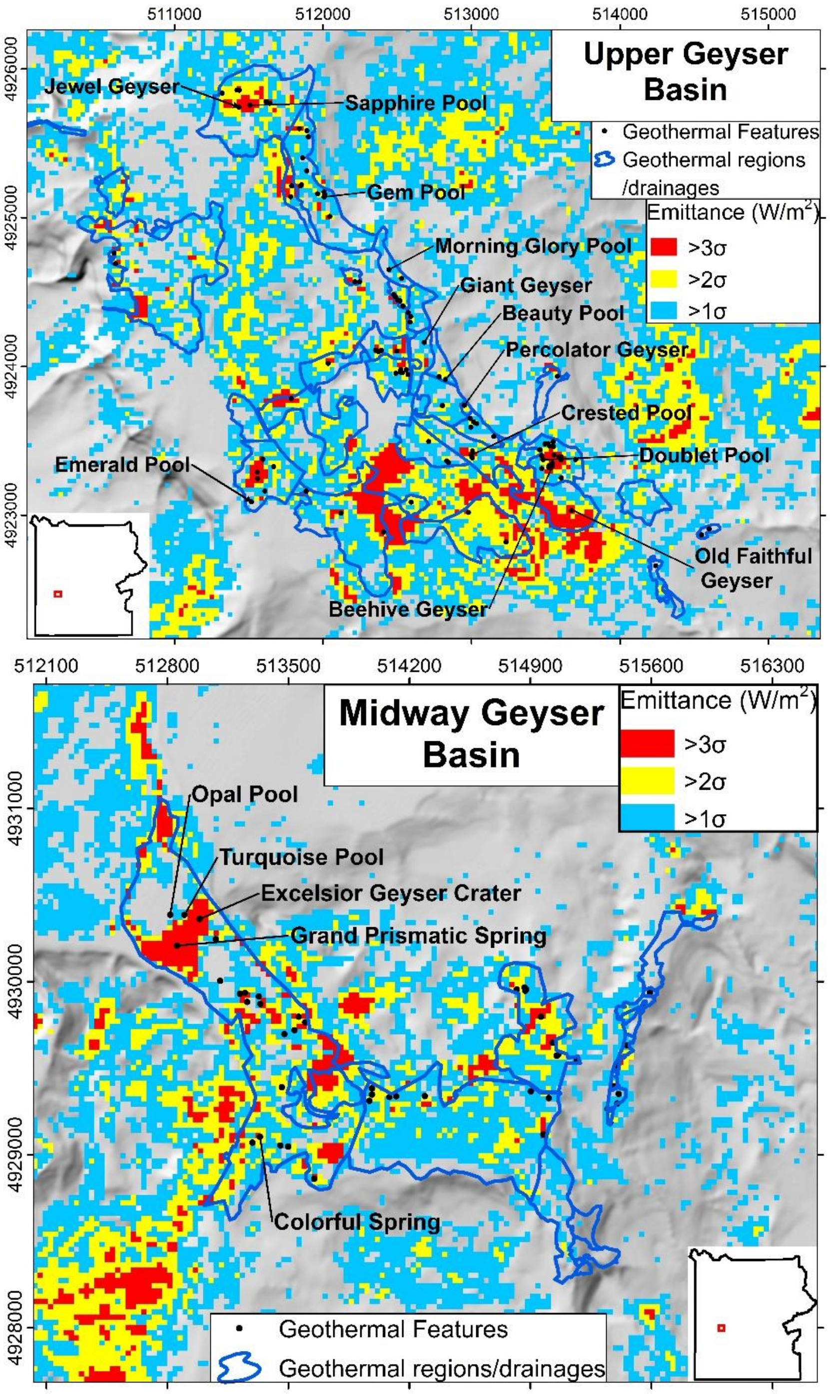


Figure 7: The standard deviation filter was used to display the anomalous pixel in relation to the surrounding pixels in Upper Geyser Basin and Midway Geyser Basin. More pixels were displayed when we used low standard deviation filter and fewer pixels were displayed when

standard deviation was increased as shown by blue, yellow and red color for 1, 2 and 3 standard deviations.

The multivariate regression model resulted in very good agreement between our model output and geothermal areas identified across the park mapped by R. Hutchinson (*unpublished*) available from the Yellowstone Center for Resources GIS geodatabase. The Hutchinson polygons include areas of geothermal influence, including travertine deposits in drainage zones, as well as individual geothermal features; our model identifies separate features within these broader zones. Figures 8 and 9 illustrate the locations of hot spring pools, mud pots, and geysers within the broader polygons. The model also identifies true anomalies, such as Violet Mud Pot 1 and 2 in Figure 8 E that are not represented in the Hutchinson (*unpublished*) polygons but have been field-confirmed by Watson et al. (2008). The model distinguished between Grand Prismatic Spring and Excelsior Geyser Crater in the Midway Geyser Basin while minimizing the effect of the surrounding runoff zones in the polygon (Fig. 8 C). Similarly, the model also identified various geysers and pools in Norris Geyser Basin (Fig. 8 F). False anomalies occur in the output, mostly concentrated along northeast-trending ridges. The existence of false anomalies in the model output suggest that one or more of the variables may be underfitting the solution in certain circumstances, perhaps as a result of complex interplay between two or more variables, or that there may be a significant variable yet excluded from our analysis, such as rock type.

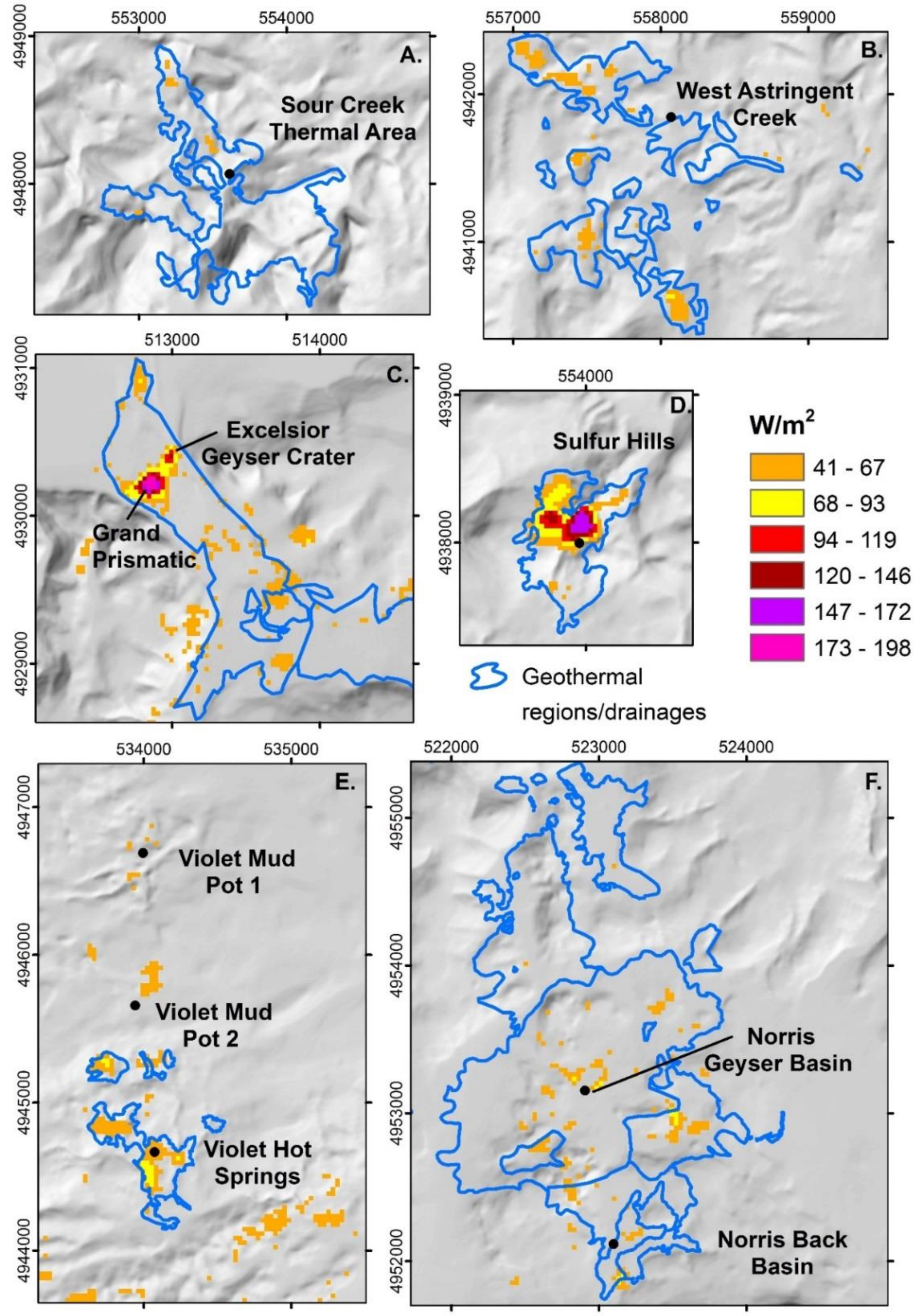


Figure 8: Residual emittance in (A) Sour Creek Thermal Area, (B) West Astringent Creek, (C) Excelsior Geyser Crater, (D) Sulfur Hills, (E) Violet Hot Springs and (F) Norris Geyser Basin of YNP showing pixels greater than 3 standard deviations above the average for the zone. Blue polygons indicate geothermal zones as mapped by R. Hutchinson (*unpublished*).

There are scattered elevated emittance outside the boundary of geothermal zones on the east and west side of Lower Geyser Basin, south of Smoke Jumper Basin and on the east of Upper Geyser Basin (Fig. 9). The patch of elevated emittance on the Fairy Creek group boundary in the Lower Geyser Basin corresponds to a pool visible in aerial imagery, but we were unable to locate the field data available to verify its hydrothermal nature. While large hot springs are most readily visible in the residual imagery, terrestrial anomalies with relatively small, time-sensitive footprints such as Old Faithful and other individual geysers are also identifiable in Upper and Lower Geyser Basins (Fig. 9 A, C).

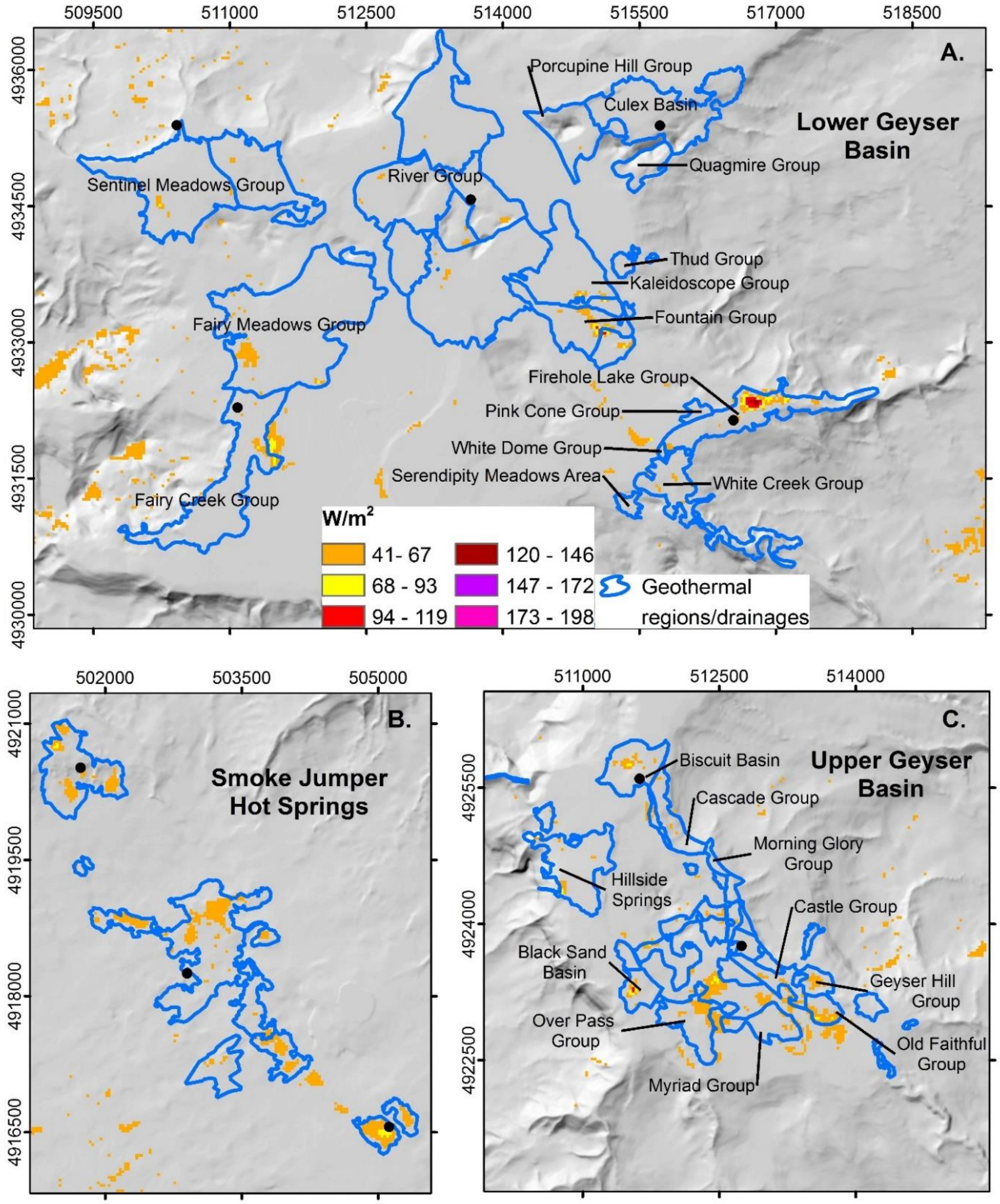


Figure 9: Residual emittance in A. Lower Geyser Basin, B. Smoke Jumper Hot Springs and C. Upper Geyser Basin of YNP showing pixels greater than 3 standard deviations above the average for the zone. Blue polygons indicate geothermal zones as mapped by R. Hutchinson. (*unpublished*).

Comparison of Multivariate model with field measurements

In addition to evaluating model output for spatial consistency with known anomalies, we compared our raw and residual temperature results with field temperatures measured by Bergfeld et al. (2011) to evaluate model output temperatures at individual sites. They classified their geothermal areas as fumaroles ($\bar{x}$ =368.53 °K), frying pans ($\bar{x}$ =365.65 °K) and pools ($\bar{x}$ =342.93 °K). We removed some geothermal locations of Bergfeld et al. (2011) that did not correspond to both the elevated thermal features in our model output and the identifiable features on the ground visible in the aerial images. This may be due to very small feature size or possible error in published latitude longitude information.

Field temperature measurements of pools, frying pans and fumaroles by Bergfeld et al. (2011) were compared with our model output to determine if there is a consistent relationship between field data and residual output (Fig. 10). The weak relation between residual and field temperature suggests that effect of background is not linear. Further, the weak relationship between raw thermal data and field data suggests that signal mixing in the pixels may have significant effects; at the very least, the model output does not greatly degrade temperature correlation to the field measurements relative to the raw, unprocessed IR data. The field temperature (range, 108.1) is more scattered than residual temperature (range, 20.06) and raw temperature (range, 17.92). The field temperature measurement (2003 to 2009) and satellite image acquisition (1999 to 2002) were not made contemporaneously, and high temperature hydrothermal regions may have been preferentially selected for the temperature survey. The field measured pool temperature is positively correlated with area of the pool (Fig. 11), which may have been caused by higher hot water discharge into the pool, longer residence time, or higher thermal inertia due to pool volume. Despite the positive correlation of field temperature with pool size, the percentage of the pixel occupied by the pool (Fig. 12) did not correlate well with model residual temperature. This suggests that although pixel mixing affects our study, we cannot account its effect because of the lack of enough data to define a clear trend. The majority of the pools occupy a small percentage of the hot pixel and have a random distribution of high, medium and low temperature.

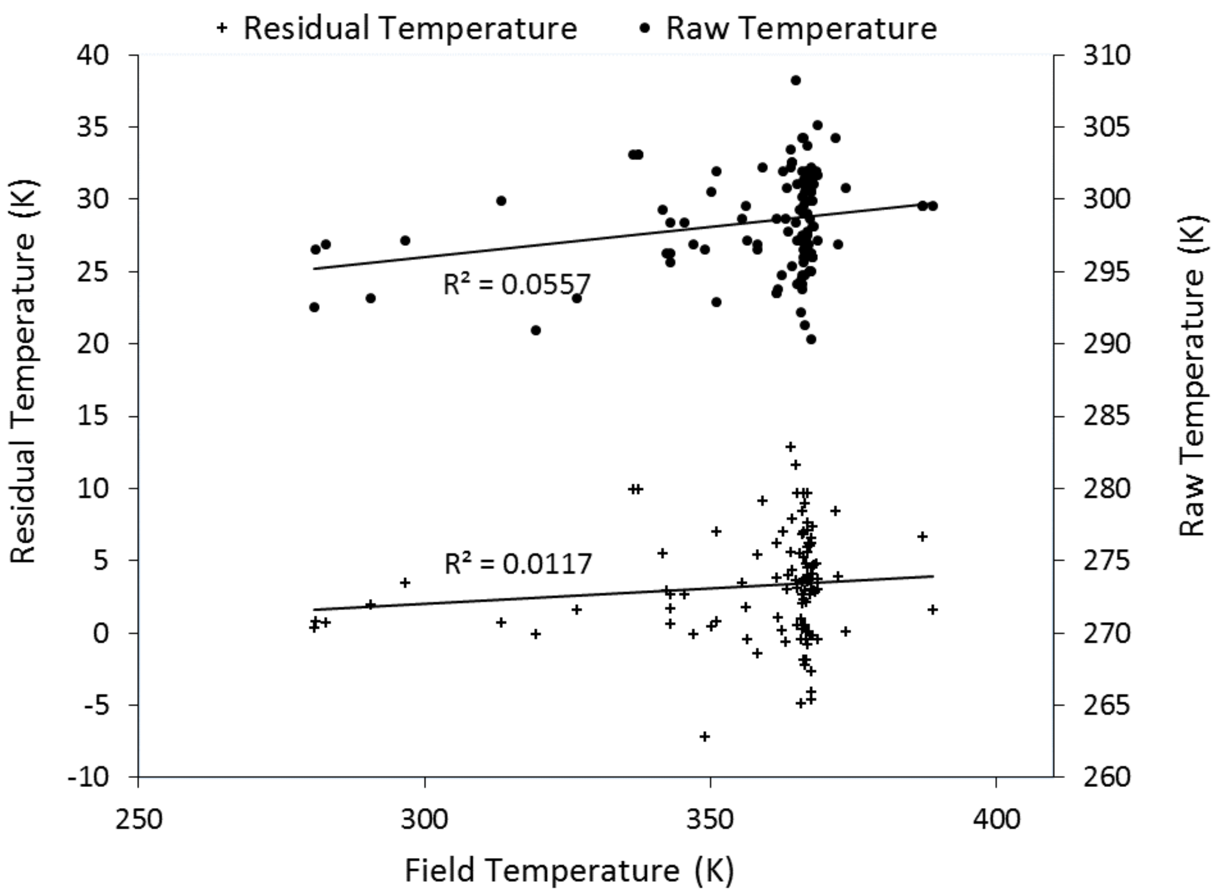


Figure 10: Correlation of raw and residual temperature with field temperature measured by Bergfeld et al. (2011). The correlation between residual temperature and field temperature is weaker than that of residual temperature.

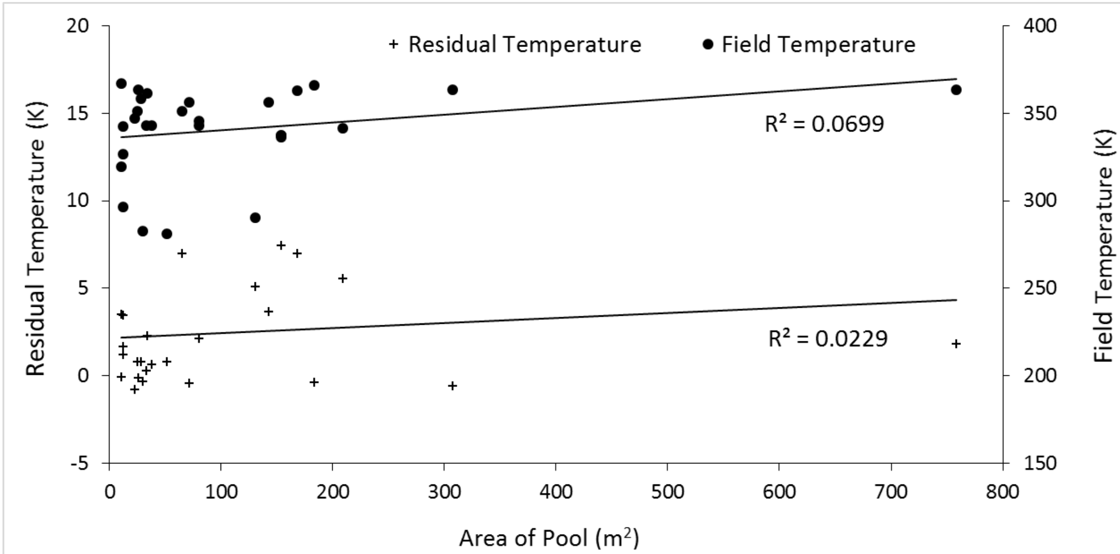


Figure 11: The area of pool has positive correlation with field temperature measured by Bergfeld et al. (2011). The bigger pool can retain more hot water and can stay warm for longer compared to smaller pools which can cool down faster.

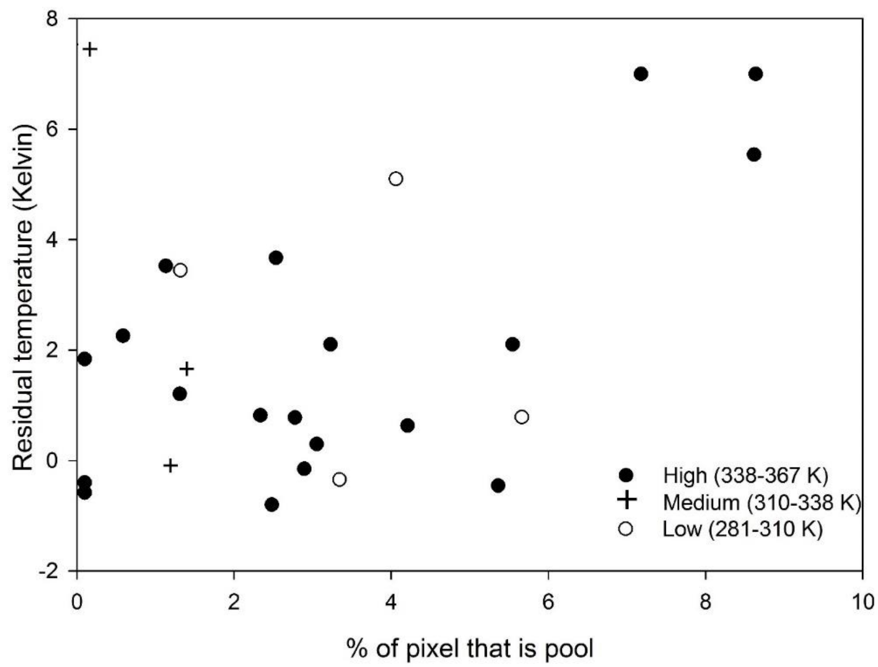


Figure 12: The correlation between the residual temperature and percentage of pixel that is pool. The random distribution of high, medium and low temperature points in relation to the percentage of pixel that is pool suggests that the size of the pool and its portion inside the hot pixel does not determine the residual temperature.

We compared our model calculated residual temperature with the temperature measured by USGS from 2010 to 2014 for ten different locations across YNP (Fig. 13). Since we used the Landsat imagery from August and September to calculate raw and residual temperature, we selected the same months for USGS field measurements and removed any measurements below 273 K because water freezes below that temperature. Out of ten sites, we removed varying percentages of bad temperature data from five sites: Nuphar Lake (8%), Gray Lake (0.02%), Opalescent Spring (0.004 %), Porcelain Outflow (2.9%) and Steamboat Geyser (0.7%). These eliminated records comprised about 1% of all the measurements from 10 sites (1,799,438 records). Most of the measurements eliminated were in the range of 270 to 271 K and were from a period of only a couple of days. The average difference between raw temperature and field temperature was 14 K with standard deviation of 19 K. Similarly, the average difference between residual and field measured temperature was 307 K with standard deviation of 19 K.

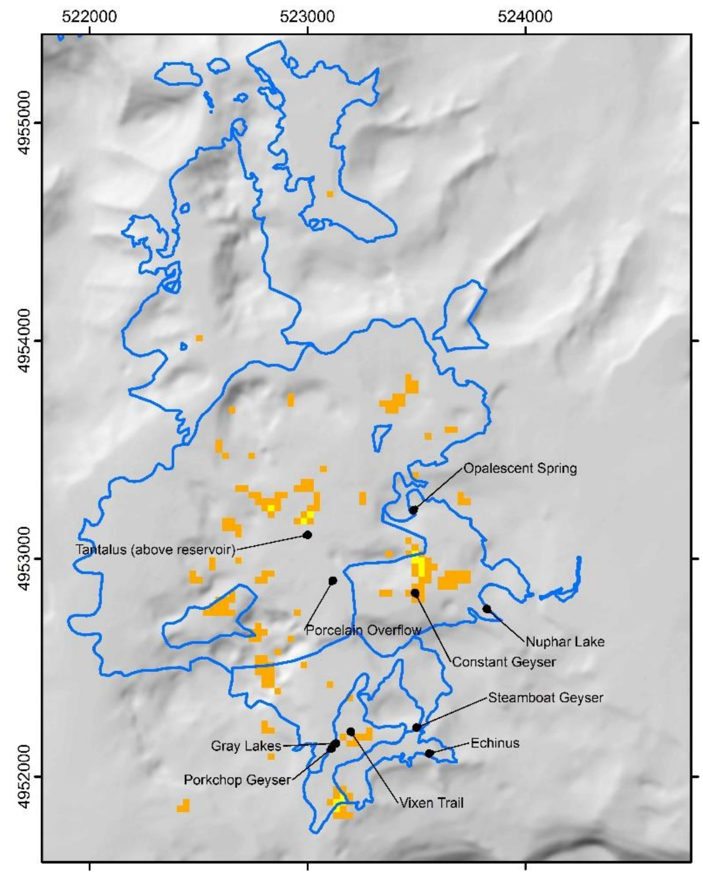


Figure 13: Location of field temperature measurement done by USGS in Norris Geyser Basin of YNP. Residual emittance pixels greater than 3 standard deviations above the average for the zone are shown with geothermal boundaries as mapped by R. Hutchinson (*unpublished*) shown in blue.

We subtracted background values from the raw in order to highlight true anomalies, the anomalies therefore represent degrees in excess of background rather than raw temperature. We found that adding a constant base temperature of 307 K to the residual temperature resulted in a better fit with field temperature measured by USGS (Fig. 14). The addition of constant offset to the residual temperature provided the same temperature fluctuation to the elevated temperature and made it easy to compare with field temperature.

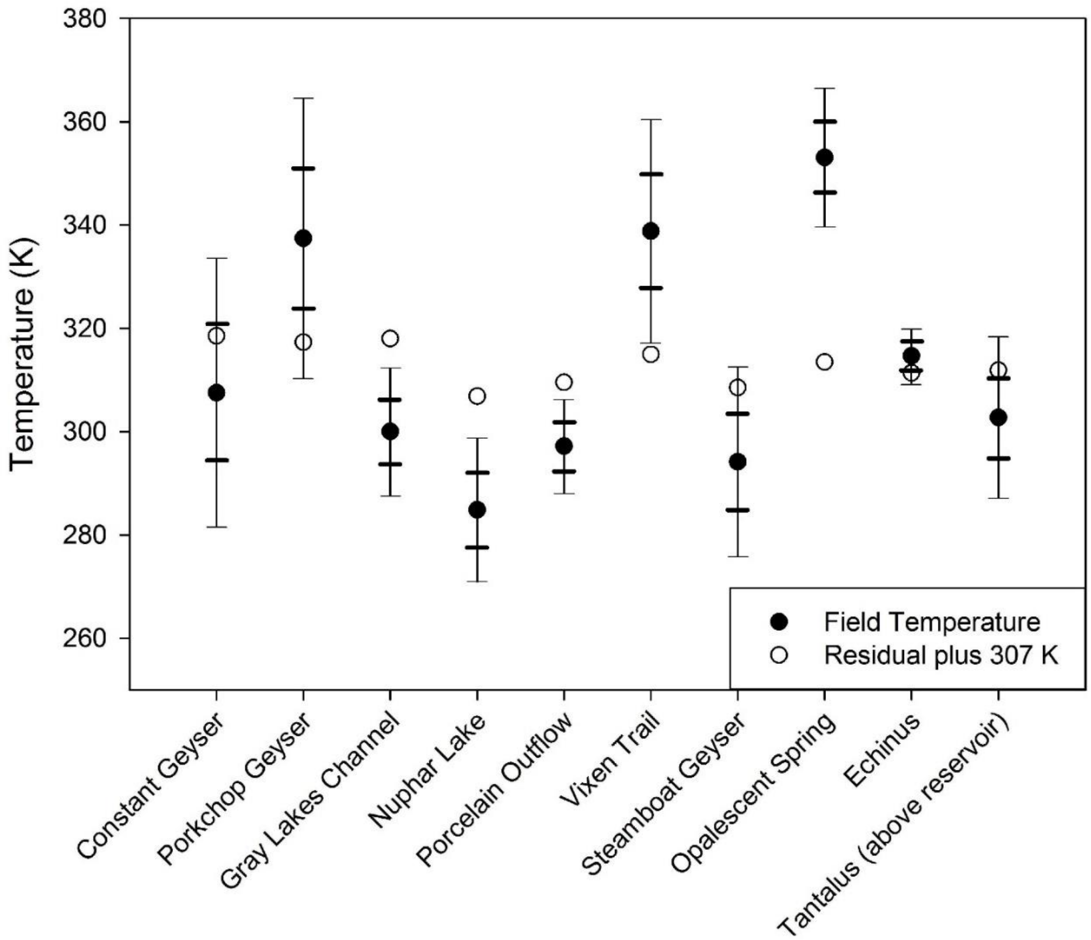


Figure 14: Graph showing the magnitude of field, residual and elevated residual temperature for ten locations of USGS temperature measurements. The elevated residual temperature was obtained by adding 307 K to all the residual temperature. The thick and thin lines represent the first and second standard deviation of field temperature, respectively.

Nature of emittance and the effects of elevation derived variables on emittance

The effects of background variables are not constant across the region and this gets reflected in the residual image when we subtract background from raw emittance. The anomalous pixels are highlighted in the residual image because it is devoid of background contributions. When we subtract the effect of background, the range of emittance decreases but it is compensated by increase in the frequency (Fig. 15) making it easy to detect hot anomalous pixels from surrounding cold pixels. Elevated residual emittance occurs not only in geothermal areas but in areas with higher elevation (Fig. 16), which suggests that we are either missing some variable in our model or our model is currently insufficient for very large elevation gains across a scene. The distribution of raw and residual emittance helped to illustrate how the subtraction of background emittance can help to identify the density and distribution of anomalous pixels in relation to elevation (Fig. 16). The scatter plot shows that anomalous pixels represented by residual emittance spikes get narrower with increasing elevation compared to raw emittance.

In order to correct for the effect of higher elevation and to determine the cut-off elevation above which geothermal anomalies do not exist in our study area, we plotted frequency of pixels for each elevation for known geothermal areas (Fig. 17). The frequency histogram showed that the highest elevation of known anomalies is 2700 m. We used this elevation to mask all the pixels above this range and reapply our multivariate model to calculate a new residual emittance. Despite masking these high elevation pixels, the model output was indistinguishable from the unmasked approach (Figs. 18), indicating that the overall model is fairly insensitive to the high elevation false anomalies.

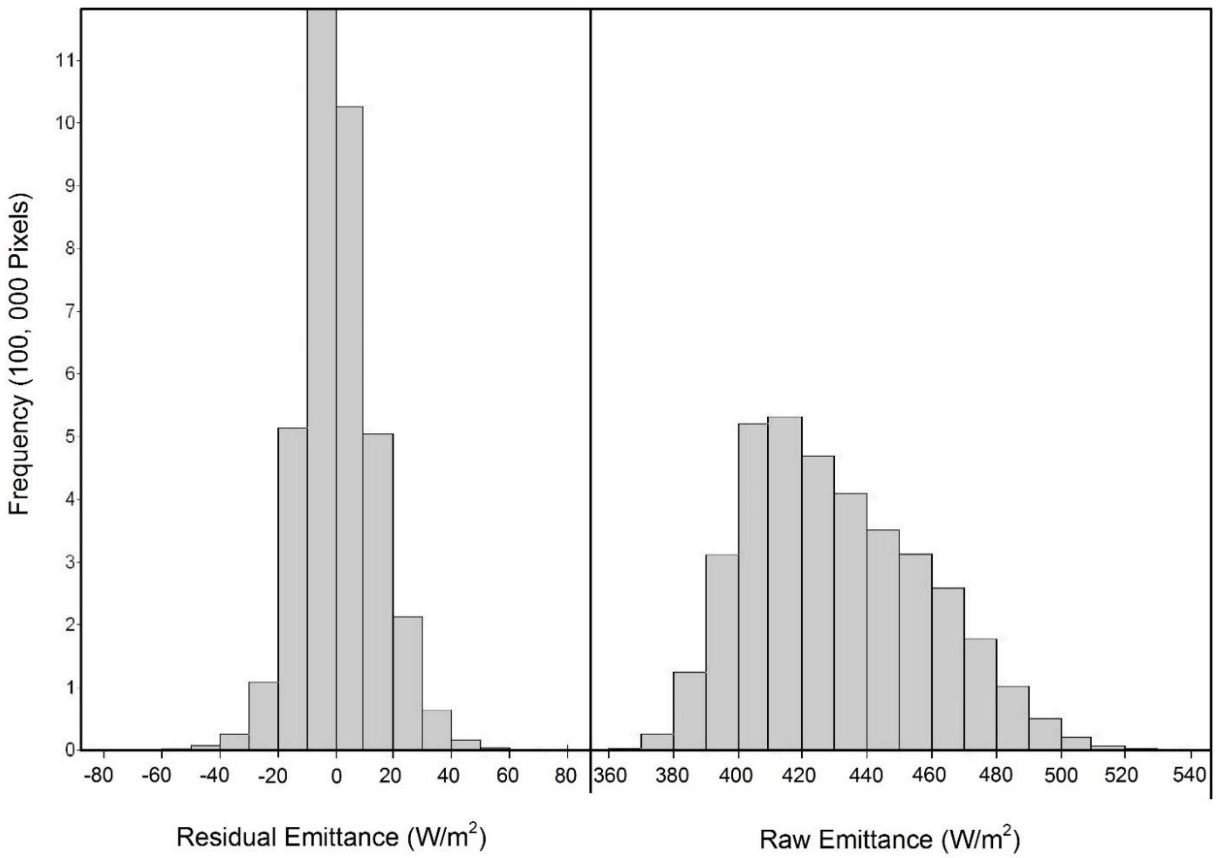


Figure 15: The frequency histogram of the raw and residual emittance of all the pixels in the test area at YNP. The lateral spread of histogram gets significantly minimized while the frequency of hotter pixels peaks when background emittance is subtracted from raw to get residual emittance.

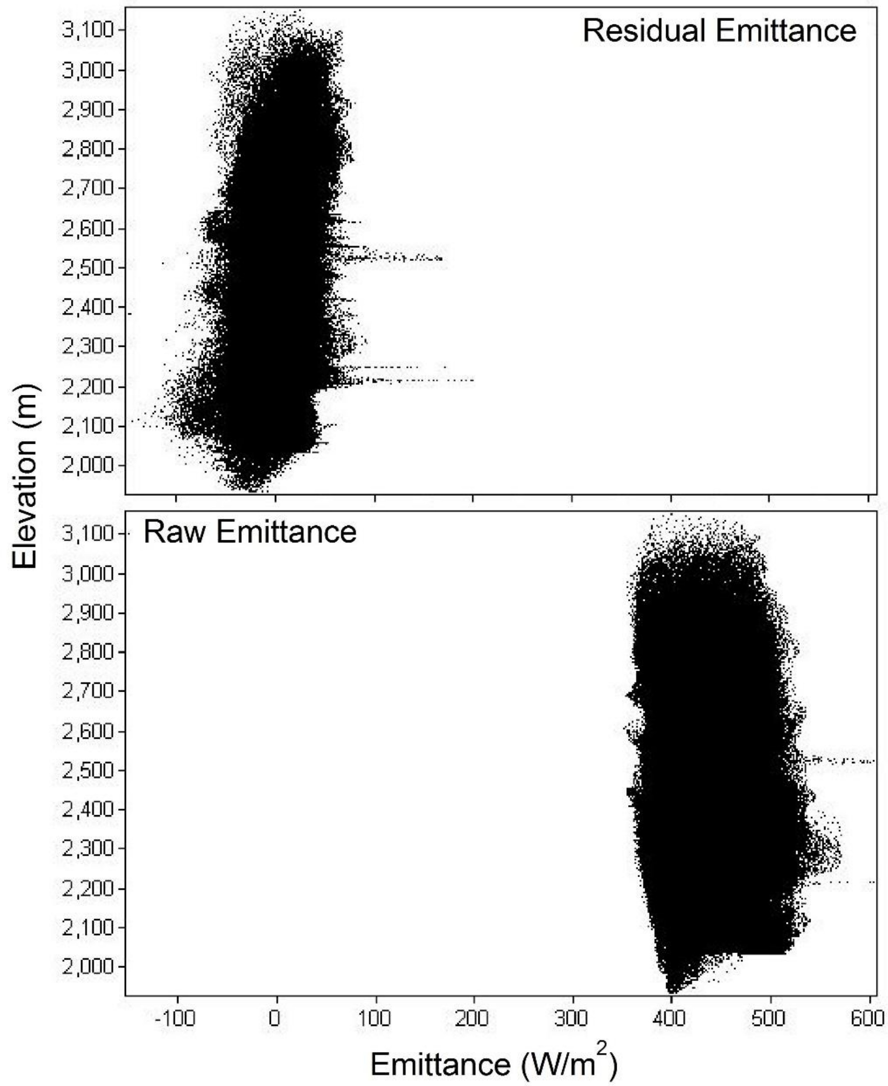


Figure 16: Raw and residual emittance in relation to elevation. The scatter plot of emittance in relation to elevation makes it easy to compare the spread of hot pixels in the same location.

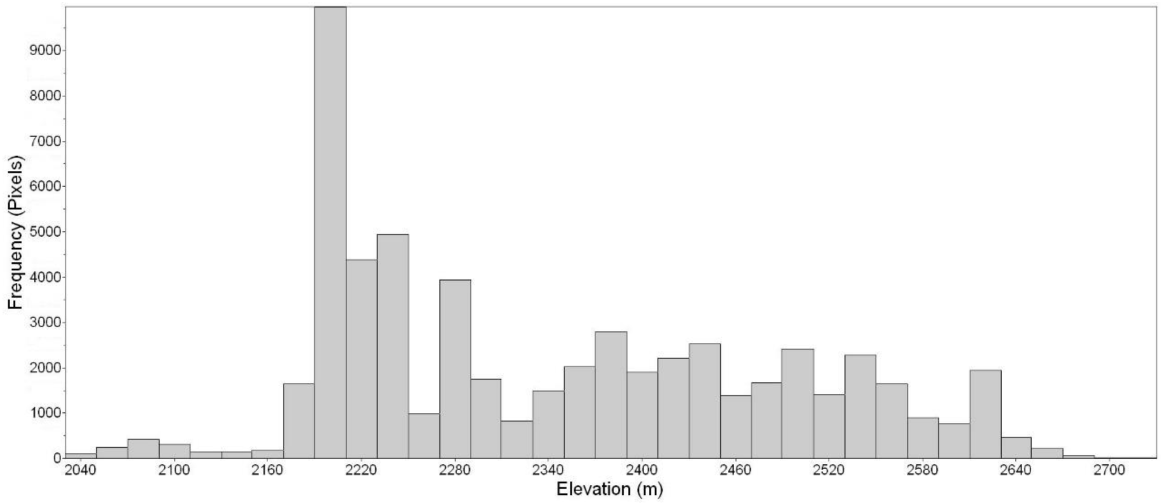


Figure 17: Histogram of the known anomalies showing pixel frequency. Most of the hot pixels exists between 2160 and 2640 m whereas no ho hot pixels exist above 2700 m.

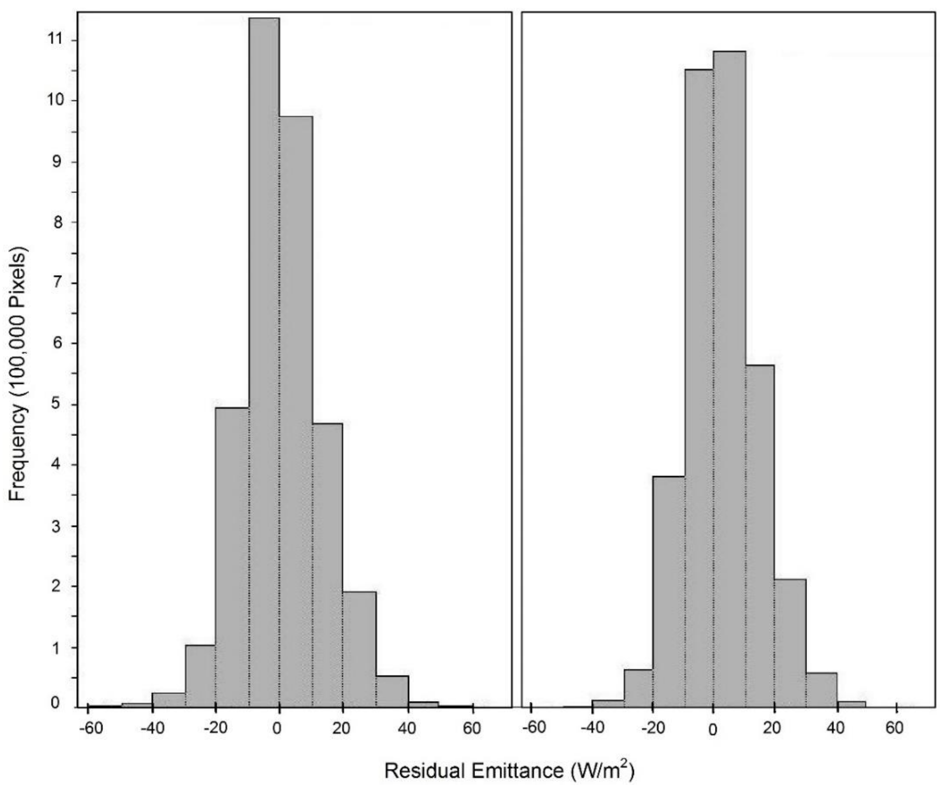


Figure 18: The frequency histogram of the residual emittance of the pixels in the test area at YNP. The first histogram shows the emittance when no pixels were masked and the second one shows the result when pixels with elevation greater than 2700 m were masked. We cannot see much

difference between these two histograms indicating that multivariate model results are not affected much by higher elevation false anomalies.

Effect of size of test area on emittance

The percentage of hot pixels and the size of the test area considered by the model determine the calculated emittance values. When we consider a test area with higher percentage of hot pixels, the multivariate regression model tries to fit to the majority of pixels representing the anomalies. The selection of fewer pixels in a smaller area forces the model to consider fewer warmer pixels which ultimately suppresses detection of the hotter pixels. To understand model behavior in response to size of test area represented by number of pixels, we ran the model for different test areas (Fig 19). When we fed large number of pixels, we got the higher anomalies (Fig.20, condition A), when we decreased number of pixels, the magnitude of anomalies also decreased (Fig. 20 condition B and C). We further decreased the number of pixel (condition E), the hotter pixel started reappearing resembling condition A but the cluster of colder pixels started appearing unlike previous conditions. From these several scenarios, condition A is the optimal condition where hotter pixels are highlighted and colder pixels are suppressed and randomized. When very small area was chosen as in condition E, the model fitted to the majority of pixels and highlighted few pixels as hot and cold cluster. The size of test area and the percentage of hot pixels inside the test area determined the temperature and distribution of these hot and cold pixels.

Figure 19: Size of the test area indicated by number of input pixels: (A) 3,673,402 (B) 1, 792, 879, (C) 744, 683 (D) 37, 601 and (E) 1, 559 and geothermal areas represented by black polygons. The percentage of hot pixels inside geothermal boundary in relation to test area increased as we decreased the test area size: (A) 1.84% (B) 1.71%, (C) 2.85% (D) 7.06 % and (E) 38.29%. The geothermal area inside study area E was common to all test areas.

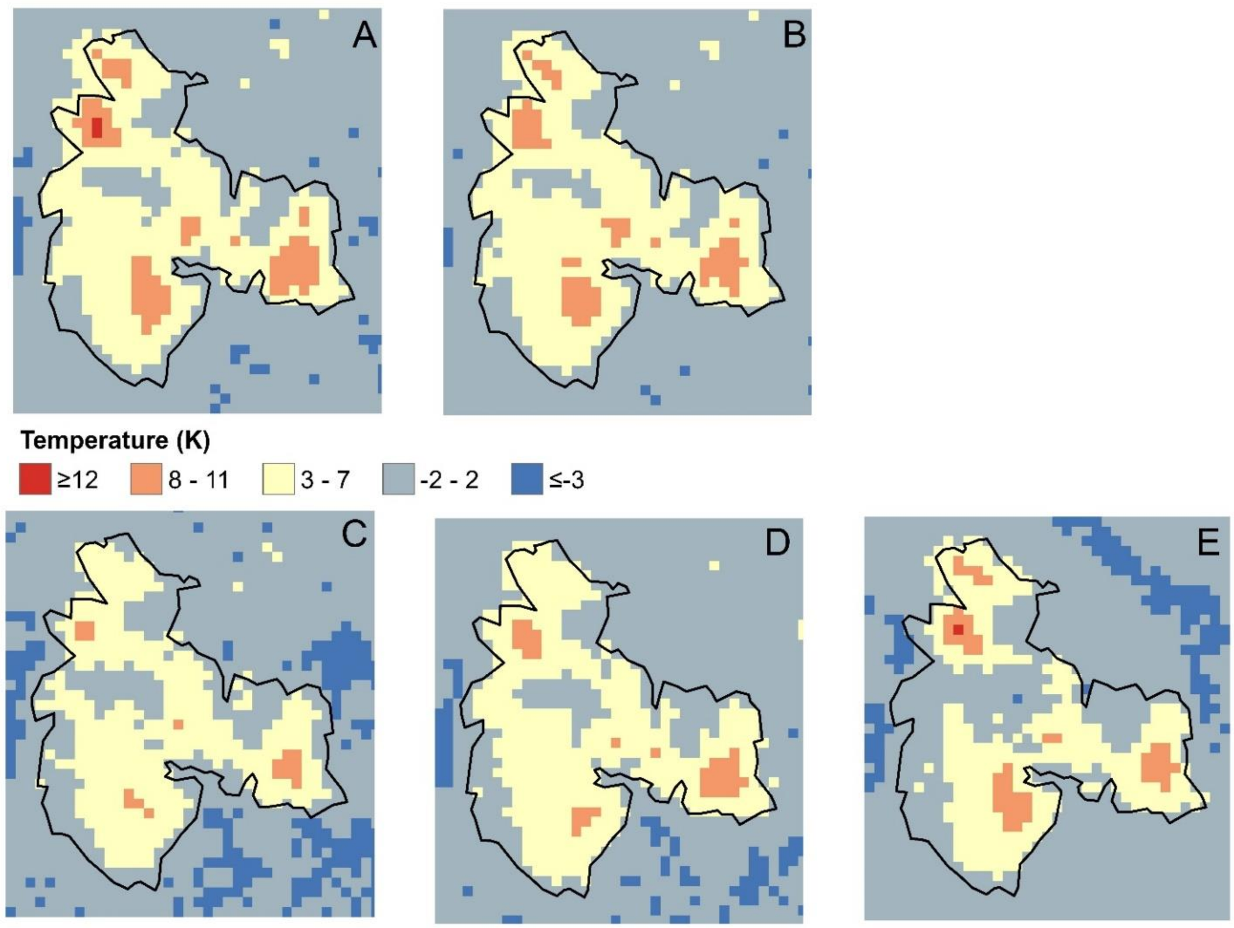


Figure 20: Residual temperature images of geothermal area that was common to all test areas. The percentage of hot pixels increased as we decreased the size of test area from A to E. The hot pixels decreased and started reappearing while we moved from A to E. The colder pixels were random for bigger test area and it started to cluster when size of the test area dropped.

Application of multivariate model in other areas

We applied our multivariate model in areas other than YNP, namely the Coso geothermal area in California and the Tendaho geothermal area in Ethiopia, in order to study the model's performance in diverse environmental conditions. Both Coso and Tendaho have much lower elevations than YNP and little to no vegetation and these conditions may help us understand more about the source of false anomalies. The response of our model in these very different types of areas has the potential to provide information on creating a more robust model suitable for all environmental conditions. The different magnitude of background variables will result in different multivariate regression equations for these test areas; we will not apply the same standard deviation filtration used in YNP but will determine it for each test area.

Coso geothermal area

The Coso geothermal area is located on the eastern side of the Sierra Nevada range in eastern California and covers about 400 km^2. It is mostly covered by lava flows and some rhyolite domes of late Cenozoic age (Wohletz and Heiken, 1992). The most recent eruption occurred about 40, 000 years ago and formed Volcano Peak basaltic cinder cone and lava flow (USGS 2012). Rhyolite domes are cut by numerous normal faults forming fumaroles and hot springs of high geothermal gradient (Wohletz and Heiken, 1992). The geothermal system at Coso reflects renewed magmatic activity (Adams et al., 2000) with temperature in excess of 325°C.

We applied the multivariate model trained around YNP in the Coso geothermal area to study the model's performance in a very different geographic setting. Final model coefficients were well within the ranges established at YNP, indicating that the continuation of the coefficient ranges was not a limiting factor in model fit despite the very different environmental conditions. After obtaining the residual emittance, we used the standard deviation filtration technique (Fig 21) to filter cooler emittance and display only the hot pixels. In this case, the two standard deviation filter was sufficient to visualize the elevated emittance. The model output indicates positive surface anomalies on the western side of Coso geothermal field which is mostly covered by pyroclastic materials as mapped by Duffield et al. (1980) and it conforms to the location of few geothermal wells. Since these wells do not indicate the surficial expression of geothermal anomalies, we did not expect our results to match with these well locations.

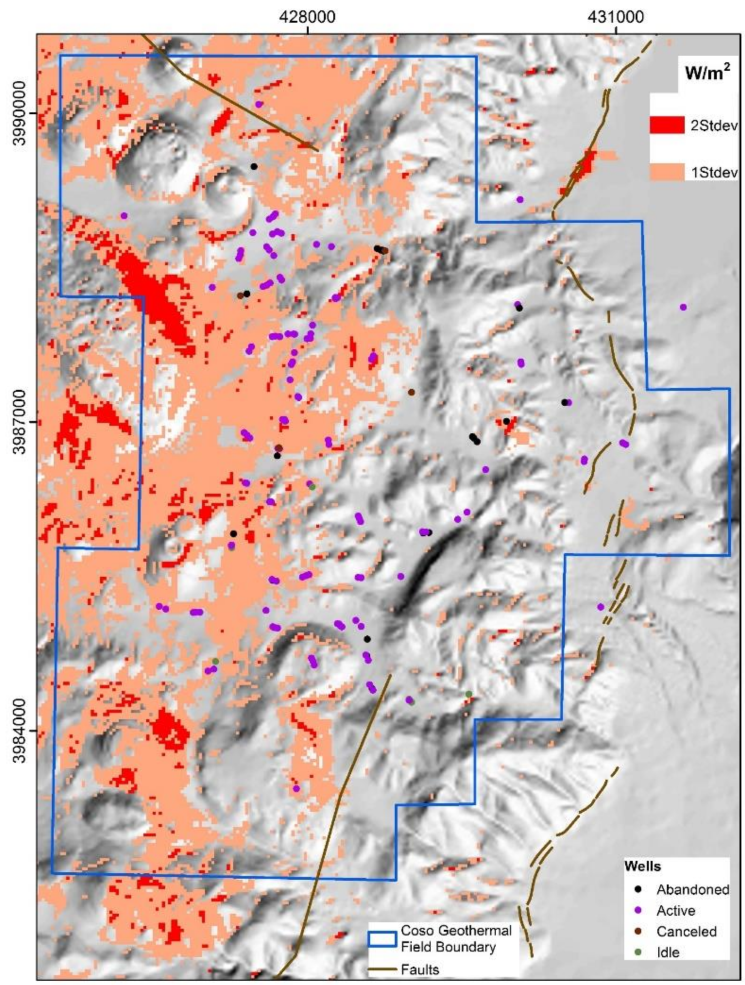


Figure 21: First and second standard deviation filtration of emittance in Coso geothermal areas. The model seems to be picking emittance in the fault zones in the northern part of the broader geothermal area while the faults in the southern zones does not seem to be conforming with the faults.

The model results in this Known Geothermal Resource Area (KGRA) were consistent with the studies done by Duffield et al. (1999) (Fig. 22). The elevated emittance on the southwestern part of the image coincides with the volcanic vents and the cluster of hot pixels on the northwestern side of the image marks the approximate boundary of Pliocene and Pleistocene vents. The presence of higher emittance pixels on the areas of basalt deposit on the northern side suggests that the model is not merely responding to areas of higher albedo.

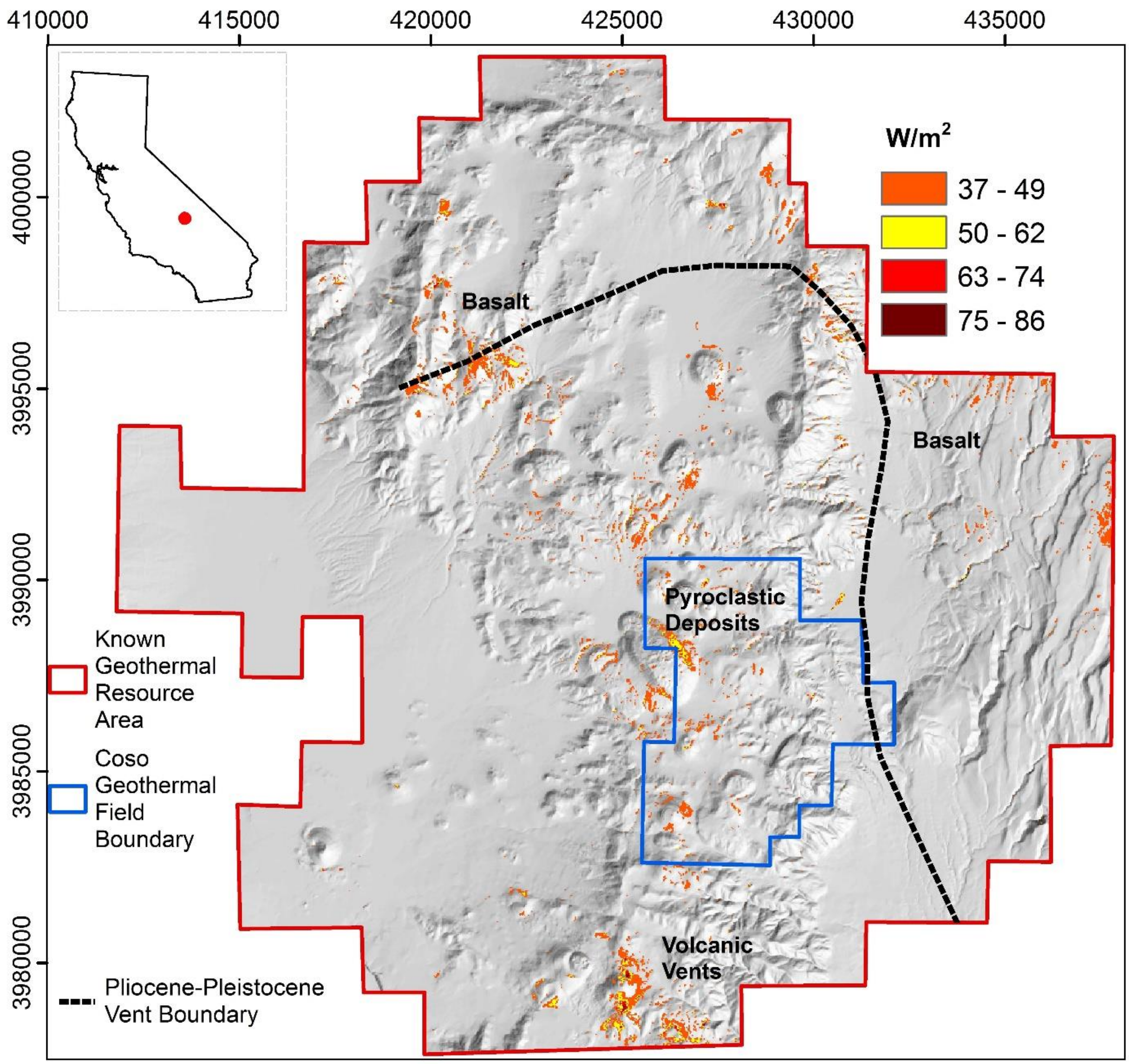


Figure 22: Figure showing residual emittance after filtering the cooler pixels using 2 standard deviation filter in Coso geothermal field and Known Geothermal Resource Area (KGRA) in California (California Department of Conservation, 2014). Higher emittance values are

concentrated on the northern side of Pliocene-Pleistocene vent boundary and in the areas of volcanic vents and pyroclastic material mapped by Duffield et al. (1980).

a. Application of multivariate model in Tendaho geothermal area, Ethiopia

The multivariate model for calculating residual emmitance was also applied in the Tendaho geothermal area in the Afar Depression of Ethiopia (Fig. 23). The Afar is a diffuse triple junction where the onland expression of the northwest-southeast trending Red Sea rift, the onland expression of the northwest-southeast trending Gulf of Aden Rift, and northeast-southwest trending Main Ethiopian Rift interact over a 200 square kilometer region (Abbate et al., 1995; Manighetti et al., 2001; Manighetti et al., 1998; Kidane et al., 2003; Beyene and Abdelsalam, 2005).

The Afar Depression is a region marked with steep horsts and graben, as well as geothermal areas such as in the Tendaho graben (Abbate et al., 1995). The Tendaho geothermal area is the southern expression of the Red Sea Propagator, and bimodal volcanism is prevalent in the area (Kidane et al., 2003). The Tendaho graben is ~45 km wide and filled with Quaternary sediments (Abbate et al., 1995).

We used our model to look for geothermal anomalies in the Tendaho area, again using the same coefficient ranges that came from the training areas outside of YNP. Like at Coso, the best-fit coefficients generated by the multivariate model were well within the limits established at YNP. After obtaining the residual emittance in Tendaho, we used a one standard deviation filter to remove cooler emittance and display only the hot pixels (Fig. 24). The higher emittance is directly focused on the central graben of the Main Ethiopian Rift (MER), the NE trending structure in the southwest part of the image, and the Red Sea Propagator, the SE trending structure in the NW part of the image (Fig. 25

and 26). The two grabens intersect in this region and this area has previously been noted as a potential region for geothermal expression (Mamo and Bekele, undated). It is worth noting that the basaltic rocks in northeast side of Tendaho are cool in the residual image while the basalts in the southwest are hot, indicating that the model is not simply registering lithology. The concentration of elevated emittance on western side corresponding to Red Sea Propagator also suggests that it is not caused by bare ground of higher albedo.

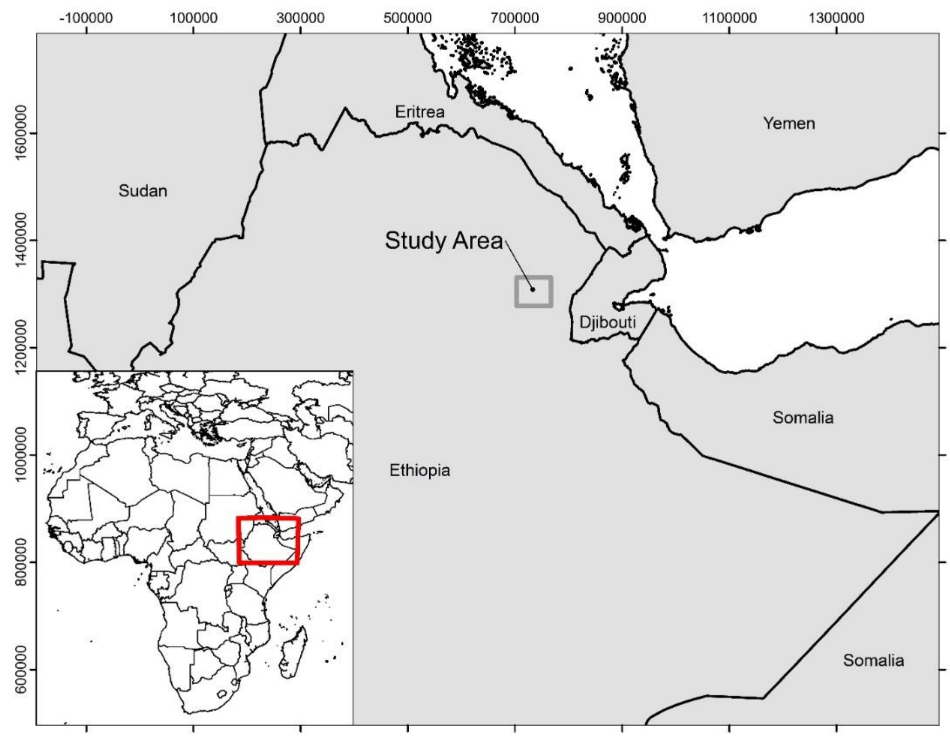


Figure 23: Location of Tendaho geothermal site in Afar depression in Ethiopia.

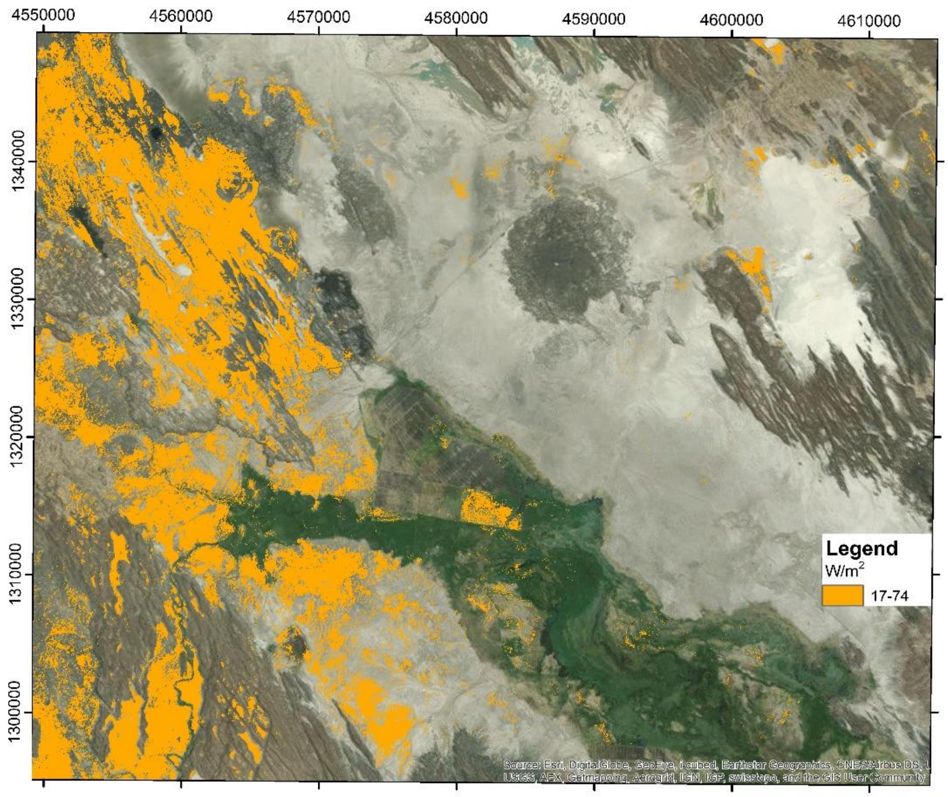


Figure 24: Aerial imagery (ESRI, 2014) of Tendaho geothermal test site in Afar depression in Ethiopia. This displayed test area corresponds with the area mapped by Mamo and Bekele (*undated*). Areas of thermal emittance greater than one standard deviation are displayed.

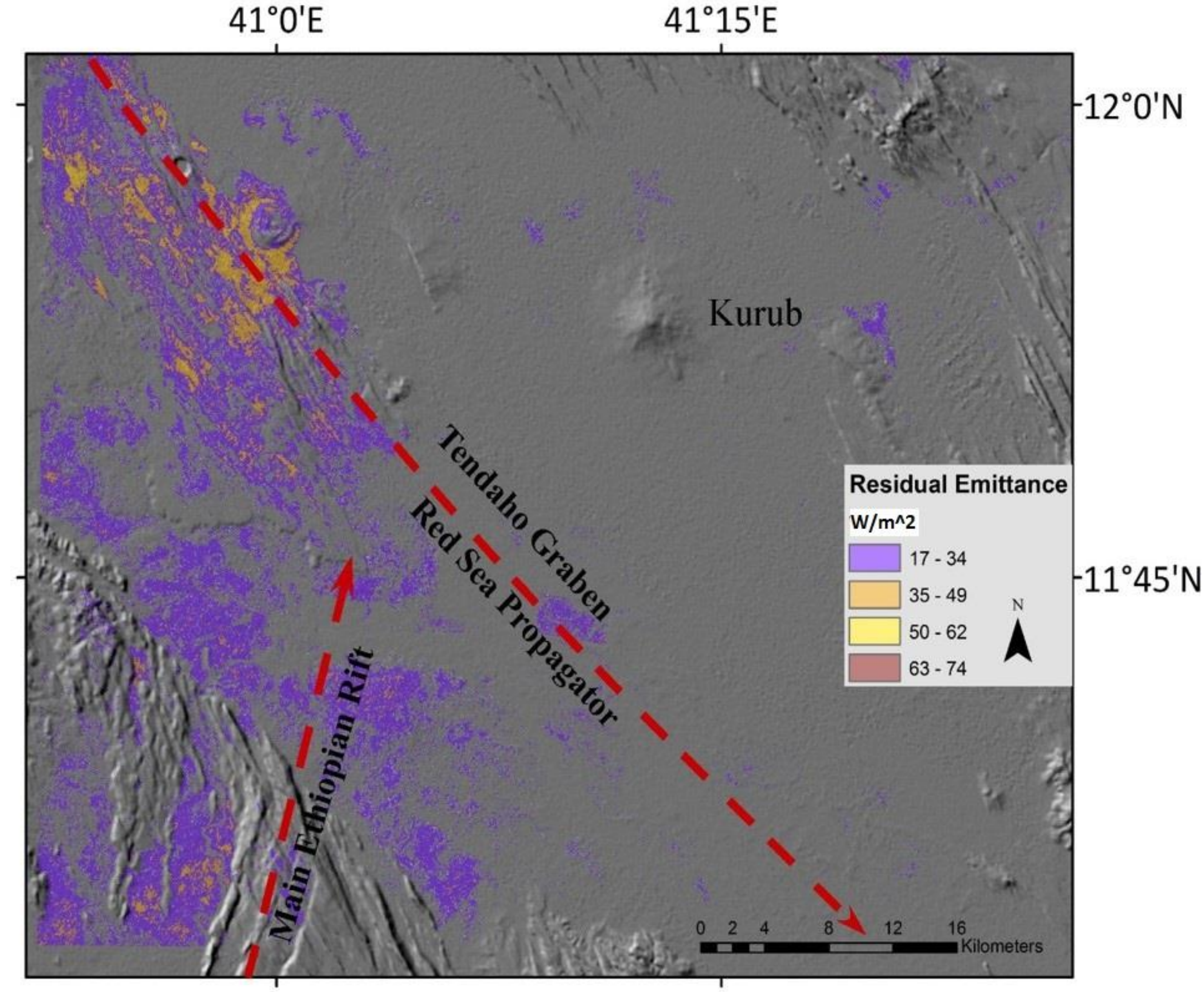


Figure 25: Tendaho geothermal site in Afar depression in Ethiopia. The test area corresponds with the area mapped by Mamo and Bekele (*undated*). The map marks the Main Ethiopian Rift, Red Sea Propagator, and Kurub (a basaltic volcano). Dama Ale, a rhyolitic volcano that terminates the Red Sea Propagator, is just outside of the frame to the southeast. (Basemaps after Abbate et al., 1995; Manighetti et al., 2001).

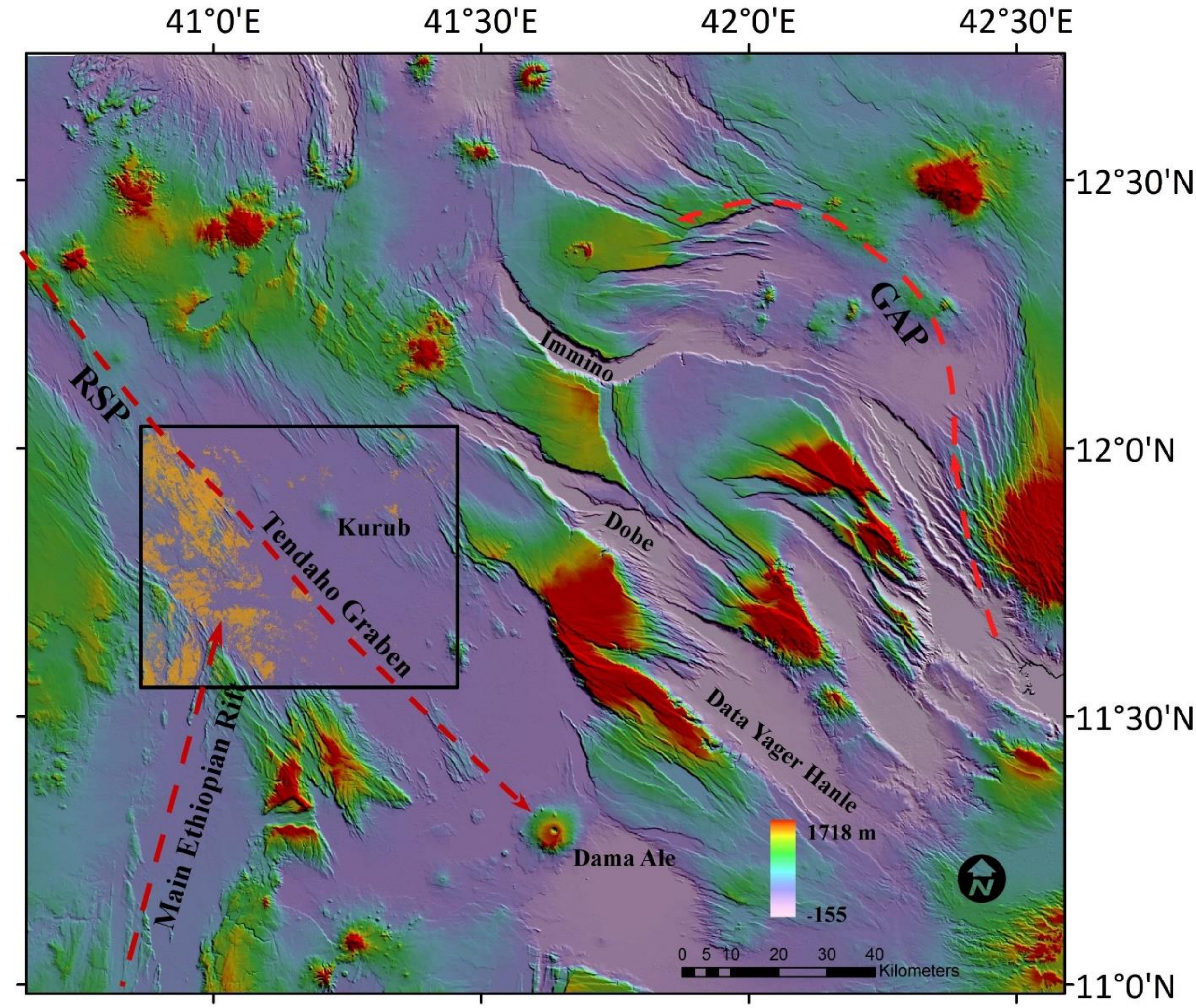


Figure 26: Tendaho geothermal site in Afar depression in Ethiopia. The map shows Main Ethiopian Rift Red Sea Propagator (RSP), Gulf of Aden Propagator (GAP), prominent graben in the region are the Immino, Dobe, and Data Yager Hanle. Volcanoes include Kurub (a basaltic volcano in the Tendaho graben) and Dama Ale,,which is a rhyolitic volcano that marks the terminating point of Red Sea Propagator (Basemaps after Abbate et al., 1995; Manighetti et al., 2001).

4. Conclusions

The multivariate regression model used in this study was used to identify geothermal anomalies in YNP and two other locations. The model detected several locations of elevated emittance such as pools, geyser, hot grounds, etc. Despite the fact that we used Landsat ETM+, which has scenes available only until 2003, the results were significant while comparing with the field based data from more recent years. The model was successfully used in different environments without re-training for preliminary detection of thermal anomalies, and doesn't rely on snow-filtering techniques or *a priori* knowledge of anomalies. The application of the model developed in YNP provided the evidence that we can use our model in other areas without retraining for preliminary investigation of thermal anomalies but detailed investigation may require that the model to be retrained in the specific test area. This technique is a low cost method to identify heat anomalies and can be used over large study sites as a first pass technique.

Future research work for this project includes expansion of the current research areas to include a broader view of the Afar region and may require re-training the model. The model can be improved by incorporating more independent variables such as rock types and mineral indices and coupling with more field data at the exact location during the same time of image acquisition such as using unmanned aerial vehicle (UAV) or satellite imaging. The field samples taken at the same time helps in investigating the spectral response of different rocks type, seasonal behavior of vegetation and location of short-lived pools and geothermal areas.

References

Abbate, E. Passerini, P. and Zan, L., 1995. Strike-Slip Faults in a Rift Area: A Transect in the Afar Triangle, East Africa. *Tectonophysics*, 241, p. 67-97.

Adams, M.C., Moore, J.N., Bjornstad, S. and Norman, D.L., 2000. *Proceeding World Geothermal Congress*, May 28-June 10, 2000, Kyushu-Tohoku, Japan.

Bergfeld, D., Lowenstern, J.B., Hunt, A.G., Shanks, W.C.P, III, and Evans, W.C., 2011, Gas and isotope chemistry of thermal features in Yellowstone National Park, Wyoming: U.S. Geological Survey Scientific Investigations Report 2011-5012, 26 p. and data files.

Beyene, A. and Abdelsalam, M.G., 2005. Tectonics of the Afar Depression: A Review and Synthesis. *Journal of African Earth Sciences*, 41, p. 41-59.

Blackwell, D.D., Negraru, P.T. and Richards, M.C., 2007. Assessment of the Enhanced Geothermal System Resource Base of the United States. *Natural Resources Research*, 15(4), p.283-308.

Burrough, P. A., and McDonell, R. A., 1998. *Principles of Geographical Information Systems*, Oxford University Press, New York, p.190.

California Department of Conservation, 2014. Geothermal Resources and Maps. [online] Available at:< http://www.conservation.ca.gov/dog/geothermal/maps/Pages/Index.aspx#g2> [Accessed 15 November 2014].

Campbell, J.B. 2008. Introduction to Remote Sensing, Fourth Edition. The Guilford Press, 626p.

Chander, G., Markham, B., and Helder, D. 2009. Summary of Current Radiometric Calibration cCoefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of the Environment*, 113, pp. 893-903.

Coolbaugh, M.F., Kratt, C., Fallacaro, A., Calvin, W.M. and Tranik, J.V., 2006. Detection of Geothermal Anomalies using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Thermal Infrared Images at Brady Hot Springs, Nevada, USA. *Remote Sensing of Environment*, 106, p.350-359.

Christiansen, R.L., 2000. The Quaternary and Pliocene Yellowstone Plateau Volcanic Field of Wyoming, Idaho and Montana. *USGS Professional Paper,* 729-G, p.G6.

DiPippo, R., 1991, Geothermal energy: electricity production and environmental impact, A worldwide perspective, energy and environment in the 21st Century: MIT Press, Cambridge, p. 741–754

Duffield, W. A., 1983. Geologic framework for geothermal energy in Cascade Range. *Geotherm. Res. Council Transactions*. 7, p. 234–246.

Eneva, M., Coolbaugh, M. and Combs, J., 2006. Application of Satellite Thermal Infrared Imagery to Geothermal Exploration in East Central California. GRC Transactions, 30.

Environmental Systems Research Institute (ESRI), 2014. *World Imagery*. Available at: < http://goto.arcgisonline.com/maps/World_Imagery > [Accessed 12 Nov 2014].

Fournier, R.O., 1989, Geochemistry and Dynamics of the Yellowstone National Park Hydrothermal System. *Annual Review of Earth and Planetary Sciences*, 17, p. 13-53.

Idaho Department of Water Resources (IDWR), 2001. *Geothermal Resources*. [online] Available at: < http://www.idwr.idaho.gov/GeographicInfo/GISdata/geothermal.htm > [Accessed 20 October 2012].

Jensen, J. R., 1986. *Introductory Digital Image Processing*, Prentice-Hall, New Jersey, p. 379.

Mamo, T. and B. Bekele., undated. Surficial Geological Mapping at Tendaho Geothermal Field, Ethiopia, Geological Survey of Ethiopia, 12 p.

Kidane, T., Courtillot, V., Manighetti, I., Audin, L., Lahitte, P., Quidelleur, X., Gillot, P.Y., Gallet, Y., Carlut, J., and Haile, T., 2003. New Paleomagnetic and Geochronological Results from Ethiopian Afar: Block Rotations Linked to Rift Overlap and Propagation, and Determination of a 2 Ma Reference Pole for Stable Africa. *Journal of Geophysical Research*, 108.

Manighetti, I., Tapponnier, P., Courtillot, V., and Gallet, Y., 2001. Strain Transfer Between Disconnected, Propagating Rifts in Afar. *Journal of Geophysical Research*, 106, p. 13613-13665.

Manighetti, I., Tapponnier, P., Gillot, P.-Y., Jacques, E., Courtillot, V., Armijo, R., Ruegg, J.C., and King, G., 1998. Propagation of rifting along the Arabia-Somalia plate boundary: Into Afar. *Journal of Geophysical Research*, 103, p. 4947-4974.

McFeeters, S.K.1996. The use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *International Journal of Remote Sensing*, 17(7), p.1425-1432.

Mia, Md. B., Bromley, C.J. and Fujimitsu, Y. 2012. Monitoring heat flux using Landsat TM/ETM+ thermal infrared data- A case study at Karapiti ('Craters of the Moon') thermal area, New Zealand. *Journal of Volcanology and Geothermal Research*, p. 235-236.

National Aeronautics and Space Administration (NASA). *Landsat 7 Science Data Users Handbook*. Updated October 2007. [online] Available at: < http://landsathandbook.gsfc.nasa.gov/pdfs/Landsat7_Handbook.pdf> [Accessed 20 November, 2012].

O'Brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity,* 41, p.673-690.

Rich, P., Dubayah, M.R., Hetrick, W.A., and Saving, S.C. 1994. Using Viewshed Models to Calculate Intercepted Solar Radiation: Applications in Ecology. *American Society for Photogrammetry and Remote Sensing Technical Papers*, 524–529.

Roy, P.S., Miyatake, S. and Rikimaru, A., 1997. Biophysical Spectral Response Modeling Approach for Forest Density Stratification. *Proceedings of the 18th Asian Conference on Remote Sensing*, October 20-24, 1997, Malaysia. [Online]. Available at:<http://www.a-a-r-s.org/aars/proceeding/ACRS1997/Papers/FR97-7.htm>.

Savage, S. L., Lawrence, R.L., Custer, S.G., Jewett J.T., Powell, S.L. and Shaw, J.A., 2012. Analyzing Change in Yellowstone's Terrestrial Emittance with Landsat Imagery. *GIScience & Remote Sensing,* 49(2), p.317-345.

Sobrino, J.A., Jiménez-Muñoz, J.C., Sòria, G., Romaguera, M. and Guanter, L., 2008. Land Surface Emissivity Retrieval from Different VNIR and TIR Sensors. *IEEE Transactions on Geoscience and Remotes Sensing*, 46(2).

United States Department of Agriculture (USDA), 2012. *National Agriculture Imagery Program (NAIP).* [ArcGIS Image Service] Available at: < http://gis.apfo.usda.gov/arcgis/services > [Accessed 10 Oct 2012].

United States Geological Survey (USGS), 1995. *Digital Maps of Low- to Moderate-Temperature Geothermal Spring and Wells in the Pacific Northwest: A Contribution to the Interior Columbia Basin Ecosystem Management Project*. [online] Available at: < http://pubs.usgs.gov/of/1995/of95-689/ > [Accessed 20 October 2012].

United States Geological Survey (USGS), 2012. *Volcano Hazard Program: Coso Volcanic Field.* [online] Available at: < http://volcanoes.usgs.gov/volcanoes/coso_volcanic_field/> [Accessed 20 November 2014].

United State Geological Survey (USGS), 2015. Field temperature data from 2010 to 2014 for ten stations in Norris Geyser Basin in Yellowstone National Park. [Personal Communication].

Vaughan, R.G., Keszthelyi, L.P., Lowenstern, J.B., Jaworowski, C. and Heasler, H., 2012. Use of ASTER and MODIS Thermal Infrared Data to Quantify Heat Flow and Hydrothermal Change at Yellowstone National Park. *Journal of Volcanology and Geothermal Research*, 233–234, p.72-89.

Watson, R.G.R., Lockwood, R.E., Newman, W.B., Anderson, T.N. and Garrott, R.A., 2008. Development and Comparison of Landsat Radiometric and Snowpack Model Inversion Techniques for Estimating Geothermal Heat Flux. *Remote Sensing of Environment*, 112, p.471-481.

Wohletz, K. and Heiken, G., 1992 *Volcanology and Geothermal Energy*. Berkeley: University of California Press. [online] Available at: http://ark.cdlib.org/ark:/13030/ft6v19p151/ [Accessed 21 November 2014].

Xu, H., 2006. Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, 27(14), p.3025-3033.

Chapter 5

Conclusion

Harvesting geothermal energy at industrial or homestead scales requires detection of significantly hot areas to facilitate exploration and identification of new resources. Previous attempts to develop remote sensing applications for geothermal anomaly detection have had limited success due to the impact of background environmental variables. These background effects result in false positive anomalies as well as the masking of true anomalies. I investigated the individual contributions of these background variables to the emitted energy in the area of YNP on a per pixel basis in order to isolate the signals of geothermal origin.

I hypothesized that the residual after removing the effects of background variables would be comprised of low magnitude noise surrounding areas of geothermal anomalies. For this study, I trained my model in the areas around YNP and tested the model in the geothermal locations inside YNP. My hypothesis was successfully validated as the cluster of higher magnitude pixels indicated the presence of geothermal anomalies that corresponded with the established boundaries and identified additional features not included in the YNP database. The model was effective to identify the actual geothermal features in contrast to the runoff and travertine deposits boundary as a basis to identify geothermal features. Further, I applied my model in Coso and Tendaho geothermal regions in California and Ethiopia, respectively. I studied YNP in two phases, first using Landsat TM and the second using Landsat ETM + sensors. I analyzed the outcomes of both phases and presented the results as two standalone papers in Chapter 3 and 4.

Earlier studies in detection of geothermal anomalies either depended on the accumulation of snowpack around the geothermal areas in order to detect geothermal boundaries (e.g., Watson et al., 2008; Vaughan et al., 2012) or assumed that background effects can be eliminated by subtracting the values of nearby, non-geothermal pixels (Vaughan et al., 2012). Some previous studies attempted to use a small number of background variables, such as topography, thermal inertia and albedo (e.g., Coolbaugh et al., 2006; Eneva et al., 2006) but did not focus on comprehensive understanding of the effect of all background variables through multivariate regression. My study is distinct from earlier work because I specifically calculated the background temperature contributed entirely by background variables, using multivariate regression and Monte Carlo approaches to train the model to fit the cool, ambient pixels surrounding the anomalies. I subtracted the background contribution from the raw temperature from thermal satellite bands to investigate the residual and compare with established geothermal boundaries and locations in YNP.

I started my investigation by selecting background variables like slope, aspect, insolation, curvature, elevation, water, and bare soil and vegetation indices. I determined the significance of each variable using univariate regression and then eliminated the insignificant ones such as curvature. Insolation and water index were eliminated because they were redundant with hillshade and vegetation index, respectively. Elevation, slope, aspect, shaded relief, vegetation and bare soil indices remained after downselection, becoming the final variables for inclusion into the multivariate regression. I used Landsat TM 5 and ETM+ sensors to calculate temperature in YNP. The model coefficients required by each background variables were determined by running the multivariate model in representative training areas around YNP. The distribution of model coefficients in all areas provided the upper and lower bounds for coefficients required by my multivariate model.

The multivariate program was written using Fortran 90 code (included in appendix) and was parallelized with openMP for faster solution. I calculated the raw temperature and all background variables for the test area in YNP and ran the multivariate model to obtain an equation for background temperature. The background temperature was subtracted from raw satellite temperature to obtain the residual. I used the Stefan-Boltzman equation to convert temperature (K) to emittance (W/m^2). A standard deviation filtration technique was used to filter out the surrounding pixels from non-geothermal areas and highlight the hot pixels present in hot areas for visualization purposes. I compared the residual temperature and emission with the established geothermal locations and field temperature measured at various locations. The results obtained using Landsat ETM+ sensor provided less noise with high temperature and emission magnitude than Landsat TM 5.

The multivariate model used in the YNP detected geothermal anomalies confirmed by earlier studies and even highlighted features not mapped by earlier studies. The modeled temperature was also in conformity with the field temperature, suggesting that my model is robust and can be used to detect geothermal anomalies in other areas. The model resulted in some false positive anomalies in areas of high altitude. Thus, I ran my model masking the high elevation pixels (more than 2700 m), resulting in slightly decreased emittance without eliminating the issue of false anomalies completely.

I applied my model in Coso and Tendaho geothermal areas to understand how the model performs in different conditions. The results obtained from both the Coso and Tendaho were consistent with established geothermal locations and showed more potential locations of geothermal origin. This suggests that my model, though trained around YNP, can be used in other areas for preliminary detection of thermal anomalies. However, detailed investigation requires training of the model in the areas around the specific test areas.

I tested if the size of test area, represented by number of input pixels, impacts the results of the model and found that it is not very sensitive to the number of pixels, demonstrating the flexibility of my model to test areas of various sizes. Overall, the model proved to be a very applicable and economical technique to detect thermal anomalies in large areas.

Future work

Though the multivariate model detected thermal anomalies and determine the magnitude of thermal anomalies in YNP and other test areas significantly different from YNP, I was not able to get rid of the false anomalies in the northeast facing slopes. These false anomalies could be the result of some missing variables or the interplay among the independent variables. I did not investigate the variables related to rock and mineral indices in my study because of its complex nature and its interaction with other variables.

Immediate future work for this research includes the expansion of the current test area to include a broader swath of the Afar, looking to link existing InSAR datasets to elevated heat. The detailed investigation will incorporate the training of the model in the Afar region.

Further studies related to this work would address the issues beyond the scope of the present study, such as improving model robustness through incorporating variables such as rock types, mineral indices, and thermal inertia. While I anticipate that lithology terms would improve the model, it should be noted that lithology appeared to have limited contributions in the Tendaho, where I examined heat from a rift system and associated bimodal volcanic rocks. The application of finer resolution thermal satellite imagery and collection of field data during the same time as image acquisition would provide more insights to understanding the nature of the individual variables and model constants.

Recommendations

The multivariate model I developed for this study can be used to detect thermal anomalies. The model works much better with surficial expression of the anomalies and cannot detect the subsurface and hidden features. Although I limited my study with the use of Landsat sensors, any thermal satellite sensors can be used in conjunction with higher resolution digital elevation models. The acquisition of nighttime thermal imagery could help to understand the effect of thermal inertia. I encountered a problem with false positive anomalies in northeast facing ridges; the investigation of more topography-derived descriptors could help to address and account for such false signals. The multivariate model I developed calculates the sum total of univariate contribution of independent variables and does not account for interaction between the variable. The inclusion of complex relationships among variables could strengthen the model by reducing the magnitude of background noise in the residual image. The application of the model in other areas such as Coso and Tendaho geothermal areas suggested that this multivariate model can be applied in other areas for preliminary detection of thermal anomalies and can be followed by training of the model in the test area if detailed investigation is warranted.

References:

Abbate, E. Passerini, P. and Zan, L., 1995. Strike-Slip Faults in a Rift Area: A Transect in the Afar Triangle, East Africa. *Tectonophysics*, 241, p. 67-97.

Adams, M.C., Moore, J.N., Bjornstad, S. and Norman, D.L., 2000. *Proceeding World Geothermal Congress*, May 28-June 10, 2000, Kyushu-Tohoku, Japan.

Aly, M.H. and Cochran, E.S., 2011. Spatio-Temporal Evolution of Yellowstone Deformation between 1992 and 2009 from InSAR and GPS Observations. *Bulletin of Volcanology*, 73(9), 1407-1419.

Bargar, K.E., 1978. Geology and Thermal History of Mammoth Hot Springs, Yellowstone National Park, Wyoming. *Geological Survey Bulletin 1444.* 55p.

Bergfeld, D., Lowenstern, J.B., Hunt, A.G., Shanks, W.C.P, III, and Evans, W.C., 2011, Gas and Isotope Chemistry of Thermal Features in Yellowstone National Park, Wyoming: U.S. Geological Survey Scientific Investigations Report 2011-5012, 26 p. and data files.

Beyene, A. and Abdelsalam, M.G., 2005. Tectonics of the Afar Depression: A Review and Synthesis. *Journal of African Earth Sciences*, 41, p. 41-59.

Bureau of Land Management (BLM). GIS Data: *Geothermal Potential Area*. [online] Available at: < http://www.blm.gov/wo/st/en/prog/energy/geothermal/geothermal_nationwide/Documents/GIS_Data.html> [Accessed 10 Oct 2012].

Burrough, P. A., and McDonell, R. A., 1998. *Principles of Geographical Information Systems*, Oxford University Press, New York, 190p.

California Department of Conservation, 2014. Geothermal Resources and Maps. [online] Available at:< http://www.conservation.ca.gov/dog/geothermal/maps/Pages/Index.aspx#g2> [Accessed 15 November 2014].

Campbell, J.B., 2008. *Introduction to Remote Sensing*. 4th ed. New York:The Guilford Press.

Chander, G., Markham, B., and Helder, D. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of the Environment*, 113, 893-903.

Chang, K.T. and Li, Z., 2000. Modelling Snow Accumulation with a Geographic Information

System. *International Journal of Geographical Information Science,* 14(7), p.693-707.

Christiansen, R.L., 2000. The Quaternary and Pliocene Yellowstone Plateau Volcanic Field of Wyoming, Idaho and Montana. *USGS Professional Paper,* 729-G, p.G6.

Christiansen, R.L., and Blank Jr., H.R., 1972. Volcanic stratigraphy of the Quaternary rhyolite plateau in Yellowstone National Park. *United States Geological Survey Professional Paper,* 729-B, p. 18.

Christiansen, R.L., Lowenstern, J.B, Smith, R.B, Heasler, H., Morgan, L.A, Nathenson, M., Mastin, L.G, Muffler, J.P. and Robinson, J.E., 2007. Preliminary assessment of volcanic and hydrothermal hazards in Yellowstone National Park and vicinity. *United States Geological Survey Open-file Report,* p.94.

Coolbaugh, M.F., Kratt, C., Fallacaro, A., Calvin, W.M. and Tranik, J.V., 2006. Detection of Geothermal Anomalies using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) thermal infrared images at Brady Hot Springs, Nevada, USA. *Remote Sensing of Environment*, 106, p.350-359.

Coolbaugh, M.F. and Shevenell, L.A., 2004. A Method for Estimating Undiscovered Geothermal Resources in Nevada and the Great Basin. *Geothermal Resources Council Transactions*, 28.

Derkey, P.D., and Johnson, B.R., 1995. Digital Maps of Low- to Moderate-Temperature Geothermal Spring and Wells in the Pacific Northwest: A Contribution to the Interior Columbia Basin Ecosystem Management Project. *US Geological Survey Open-File Report 95-689*, 11p. [online] Available at: < http://pubs.usgs.gov/of/1995/of95-689/ > [Accessed 20 October 2012].

Duffield, W. A., 1983. Geologic framework for geothermal energy in Cascade Range. *Geotherm. Res. Council Transactions*. 7, p. 234–246.

Eneva, M., Coolbaugh, M. and Combs, J., 2006. Application of Satellite Thermal Infrared Imagery to Geothermal Exploration in East Central California. *GRC Transactions*, 30.

Environmental Systems Research Institute (ESRI), 2014. *World Imagery*. Available at: < http://goto.arcgisonline.com/maps/World_Imagery > [Accessed 12 Nov 2014].

Fournier, R.O., 1989, Geochemistry and dynamics of the Yellowstone National Park Hydrothermal System. *Annual Review of Earth and Planetary Sciences*, 17, p. 13-53.

Friedman, I. and Norton, D.R., 1981. Ground Temperature Measurements: Part III, Ground Temperatures in and near Yellowstone National Park. *United States Geological Survey Professional Paper 120*, 23-39.

Fu, P., and Rich, P. M., 2000. *The Solar Analyst 1.0 Manual*. Helios Environmental Modeling Institute (HEMI), USA, 46p.

Fu, P., and Rich, P. M., 2002. A Geometric Solar Radiation Model with Applications in Agriculture and Forestry. *Computers and Electronics in Agriculture* 37, 25–35.

Hellman, M.J. and Ramsey, M.S., 2004. Analysis of Hot Springs and Associated Deposits in Yellowstone National Park using ASTER and AVIRIS Remote Sensing. *Journal of Volcanology and Geothermal Research*, 135, p.195-219.

Idaho Department of Water Resources (IDWR), 2001. *Geothermal Resources*. [online] Available at: < http://www.idwr.idaho.gov/GeographicInfo/GISdata/geothermal.htm > [Accessed 20 October 2012].

Jaworowski, C., Heasler, H.P., Neale, C.M.U., and Sivarajan, S., 2010. Using Thermal Infrared Imagery and LiDAR in Yellowstone Geyser Basins. *Yellowstone Science* 18(1), 8-19.

Jensen, J. R., 1986. *Introductory Digital Image Processing*, Prentice-Hall, New Jersey, 379p.

Kidane, T., Courtillot, V., Manighetti, I., Audin, L., Lahitte, P., Quidelleur, X., Gillot, P.Y., Gallet, Y., Carlut, J., and Haile, T., 2003. New Paleomagnetic and Geochronological Results from Ethiopian Afar: Block Rotations Linked to Rift Overlap and Propagation, and Determination of a 2 Ma Reference Pole for Stable Africa. *Journal of Geophysical Research*, 108.

Kunkel, M.L., Flores, A.N., Smith, T.J., McNamara, J.P. and Benner, S.G., 2011. A Simplified Approach for Estimating Soil Carbon and Nitrogen Stocks in Semi-arid Complex Terrain. *Geoderma*, 165, p.1-11.

Mamo, T. and B. Bekele., undated. Surficial Geological Mapping at Tendaho Geothermal Field, Ethiopia, Geological Survey of Ethiopia, 12 p.

Manighetti, I., Tapponnier, P., Courtillot, V., and Gallet, Y., 2001. Strain Transfer Between Disconnected, Propagating Rifts in Afar. *Journal of Geophysical Research*, 106, p. 13613-13665.

Manighetti, I., Tapponnier, P., Gillot, P.-Y., Jacques, E., Courtillot, V., Armijo, R., Ruegg, J.C., and King, G., 1998. Propagation of rifting along the Arabia-

Somalia plate boundary: Into Afar. *Journal of Geophysical Research*, 103, p. 4947-4974.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. International Journal of Remote Sensing, 17(7), 1425-1432.

Mia, M.B., Bromley, C.J. and Fujimitsu, Y., 2012. Heat Flux Monitoring Using Satellite Based Imagery at Karapiti ('Craters of the Moon') Fumarole Area, Taupo, New Zealand. In: Kyushu University, University of Dhaka, *Thirty-Seventh Workshop on Geothermal Reservoir Engineering*. Stanford, California, 30 Jan-1 Feb 2012.

Mock, J.E., Tester, J.W. and Wright, P.M., 1997. Geothermal Energy from the Earth: Its Potential Impact as an Environmentally Sustainable Resource. *Annual Review of Energy and Environment*, 22, p.305-356.

Morgan, P., Blackwell, D.D., Spafford, R.E., Smith, R.B., 1977. Heat flow measurements in Yellowstone Lake and the thermal structure of Yellowstone Caldera. *Jounral of Geophysical Research* 82(26), 3719-3732.

National Aeronautics and Space Administration (NASA). *Landsat 7 Science Data Users Handbook*. Updated October 2007, [online] Available at: < http://landsathandbook.gsfc.nasa.gov/pdfs/Landsat7_Handbook.pdf> [Accessed 20 November, 2012].

O'Brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity*, 41, 673-690.

Ranhao, S., Baiping, Z. and Jing, T., 2008. A Multivariate Regression Model for Predicting Precipitation in the Daqing Mountains. *Mountain Research and Development*, 28(3), p.318-325.

Rich, P., Dubayah, M.R., Hetrick, W.A., and Saving, S.C., 1994. Using Viewshed Models to Calculate Intercepted Solar Radiation: Applications in Ecology. *American Society for Photogrammetry and Remote Sensing Technical Papers*, 524–529.

Roy, P.S., Miyatake, S. and Rikimaru, A., 1997. Biophysical Spectral Response Modeling Approach for Forest Density Stratification. *Proceedings of the 18th Asian Conference on Remote Sensing*, October 20-24, 1997, Malaysia. [Online]. Available at:<http://www.a-a-r-s.org/aars/proceeding/ACRS1997/Papers/FR97-7.htm>.

Savage, S. L., Lawrence, R.L., Custer, S.G., Jewett J.T., Powell, S.L. and Shaw, J.A., 2012. Analyzing Change in Yellowstone's Terrestrial Emittance with Landsat Imagery. *GIScience & Remote Sensing*, 49(2), p.317-345.

Sobrino, J.A., Jiménez-Muñoz, J.C., Sòria, G., Romaguera, M. and Guanter, L., 2008. Land Surface Emissivity Retrieval from Different VNIR and TIR Sensors. *IEEE Transactions on Geoscience and Remotes Sensing*, 46(2).

Sokolowski, J.A. and Banks, C.M., 2010. *Modeling and Simulation Fundamentals*. New Jersey: John Wiley & Sons, Inc.

United States Department of Agriculture (USDA), 2012. *National Agriculture Imagery Program (NAIP)*. [ArcGIS Image Service] Available at: < http://gis.apfo.usda.gov/arcgis/services > [Accessed 10 Oct 2012].

United States Geological Survey (USGS), 1995. *Digital Maps of Low- to Moderate-Temperature Geothermal Spring and Wells in the Pacific Northwest: A Contribution to the Interior Columbia Basin Ecosystem Management Project*. [online] Available at: < http://pubs.usgs.gov/of/1995/of95-689/ > [Accessed 20 October 2012].

United States Geological Survey (USGS), 2012. *Volcano Hazard Program: Coso Volcanic Field.* [online] Available at: < http://volcanoes.usgs.gov/volcanoes/coso_volcanic_field/> [Accessed 20 November 2014].

United State Geological Survey (USGS), 2015. Field Temperature Data from 2010 to 2014 for Ten Stations in Norris Geyser Basin in Yellowstone National Park. [Provided through personal communication].

Vaughan, R.G., Keszthelyi, L.P., Davies, A.G. and Schneider, D.J., 2010. Exploring the Limits of Identifying Sub-pixel Thermal Features Using ASTER TIR Data. *Journal of Volcanology and Geothermal Research*, 189, p.225-237.

Vaughan, R.G., Keszthelyi, L.P., Lowenstern, J.B., Jaworowski, C. and Heasler, H., 2012. Use of ASTER and MODIS Thermal Infrared Data to Quantify Heat Flow and Hydrothermal Change at Yellowstone National Park. *Journal of Volcanology and Geothermal Research*, 233–234, p.72-89.

Vaughan, R.G., Lowenstern, J.B., Keszthelyi, L.P., Jaworowski, C., and Heasler, H., 2012b. Mapping Temperature and Radiant Geothermal Heat Flux Anomalies in the Yellowstone Geothermal System Using ASTER Thermal Infrared Data. *GRC Transactions* 36, 1403-1410.

Warner, T.A. and Chen, X., 2001. Normalization of Landsat Thermal Imagery for the Effects of Solar Heating and Topography. *Remote Sensing*, 22, p.773-788.

Watson, R.G.R., Lockwood, R.E., Newman, W.B., Anderson, T.N. and Garrott, R.A., 2008. Development and Comparison of Landsat Radiometric and Snowpack Model Inversion

Techniques for Estimating Geothermal Heat Flux. *Remote Sensing of Environment*, 112, p.471-481.

Weng, Q., 2009. Thermal Infrared Remote Sensing for Urban Climate and Environmental Studies: Methods, Applications, and Trends. *Journal of Photogrammetry and Remote Sensing*, 64, p.335-344.

Wohletz, K. and Heiken, G., 1992 *Volcanology and Geothermal Energy*. Berkeley: University of California Press. [online] Available at: http://ark.cdlib.org/ark:/13030/ft6v19p151/ [Accessed 21 November 2014].

Xu, H., 2006. Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, 27(14), 3025-3033.

Appendix:

This is the multivariate program code that was written using Fortran 90. The input for the program is the text document with array of X-coordinate, Y-coordinate, raw temperature, NDVI, NDBSI, elevation, slope, aspect and hillshade. The user has to specify the name of the input file and number of iterations. The optimum number of iterations can be determined by increasing the number of iterations until the point of no further improvement or stable solution. The number of iteration that provided stable solution for my study was 100,000. The program outputs “testOut.txt” file with resultant multivariate equation. The code has to be run with openMP parallelization.

```
Program Multivariate
!-------------------------------------------------------------------------------------------------------
---!
! Multivariate - A program to calculate dependent variable temperature based on the
random coefficient generated.  !
! The combination of coefficient and independent variable that generates the least average
residual will be selected.   !
!
!........................
! Explanation of variables:
! Dependent variable: Temperature
! Independent variables: Slope, Aspect, Hillshade, Elevation, NDVI, NDBSI
! Slope : slope of ground determined in ArcMap 10.1 using 10 m NED, range:0 to 90
degree                                                                            !
! Aspect: Direction of maximum slope determined in ArcMap 10.1 using 10 m NED,
range:0 to 360 degree                                                    !
! Hillshade: Illumination of ground, range: 0 to 255
! Elevation: Altitude of 10 m NED
! NDVI: Normalized Difference Vegetation Index, range: -1 to 1 determined in ENVI 4.8
using Landsat TM 5, band 3 & 4        !
! NDBSI: Normalized Difference Bare Soil Index, range: -1 to 1 determined in ENVI 4.8
using Landsat TM 5, band 4 & 5        !
! Intercept: Multivariate regression Model intercept
!-------------------------------------------------------------------------------------------------------
--!
implicit none
integer::k,istat
```

```
integer::i,n,seedsize,m
integer, dimension(8):: dateVals
integer :: omp_get_num_threads, omp_get_thread_num, nthreads, tid
integer :: t1, t2, rate

integer, dimension(:), allocatable :: seed
real(kind=8),dimension(:),allocatable::
slope,aspect,hillshade,elevation,NDVI,NDBSI,intercept,ET,Res,temp,xx,yy
real(kind=8)::maxSlope,minSlope,maxAspect1,minAspect1,maxAspect2, minAspect2,
maxHillshade,minHillshade,maxElevation,&
minElevation,maxNDVI,minNDVI,maxNDBSI,minNDBSI,maxIntercept,minIntercept
real(kind=8),dimension(:),allocatable::
coffSlope,coffAspect1,coffAspect2,coffHillshade,coffElevation,&
coffNDVI,coffNDBSI, coffIntercept
real(kind=8)::ranslope,ranAspect1,ranAspect2,ranHillshade,ranElevation,ranNDVI,ranN
DBSI,ranIntercept, minRes
real(kind=8),dimension(:),allocatable:: sumRes, meanRes
character(len=10)::filename
real::Start, finish,PT

call system_clock(t1, rate)
print*,'Enter the name of Input text file, Eg; filename.txt'
Read*,filename
open(unit=99,file=filename,status='old',action='read')
open(unit=23, file="testOut.txt")
write (23,*) 'Input file=',filename
k=0
istat=0
do while(istat.eq.0)
        read(99,*,iostat=istat)
        k=k+1
end do
write(*,*) 'Number of Rows=',k
write (23,*) 'Number of Rows=',k
rewind(99)

allocate(xx(k),yy(k),slope(k),aspect(k),hillshade(k),elevation(k),NDVI(k),NDBSI(k),inter
cept(k),ET(k),temp(k),Res(k))
do i=1,k
        read(99,*,end=50)
xx(i),yy(i),temp(i),NDVI(i),NDBSI(i),elevation(i),slope(i),aspect(i),hillshade(i)
end do
50 continue
close(99)
call DATE_AND_TIME(VALUES=dateVals)
call RANDOM_SEED(SIZE=seedSize)
```

```
allocate(seed(seedSize))
call RANDOM_SEED(GET=seed)
call RANDOM_SEED(PUT=dateVals((9-seedSize):8))

maxSlope=0.1219
minSlope=-0.07244

maxAspect1=0.01977
minAspect1=-0.04258

maxAspect2=0.000088
minAspect2=-0.000047

maxHillshade=0.0845
minHillshade=-0.01963

maxElevation=0.002773
minElevation=-0.03677

maxNDVI=9.084
minNDVI=-10.33

maxNDBSI=61.66
minNDBSI=25.8

maxIntercept=394.50159
minIntercept=302.889

print*,'How many combinations should I run for a solution?'
read*,m
write (23,*) 'Number of Run=',m
allocate(sumRes(m),MeanRes(m),coffSlope(m),coffAspect1(m),coffAspect2(m),&
coffHillshade(m),coffElevation(m),coffNDVI(m),coffNDBSI(m),coffIntercept(m))

sumRes(:)=0.
MeanRes(:)=0.


!$OMP parallel private(tid)
tid = omp_get_thread_num()

!$OMP do private(i, n, ET, ranSlope, ranAspect1, ranAspect2,
ranHillshade,ranElevation,ranNDVI,ranNDBSI,ranIntercept) !, coffSlope, coffAspect, &
!!$OMP coffHillshade, coffElevation, coffNDVI, coffNDBSI, coffIntercept)

do i=1,m
```

```
            call random_number(ranSlope)
            coffSlope(i)=ranSlope*(maxSlope-minSlope)+minSlope

            call random_number(ranAspect1)
            coffAspect1(i)=ranAspect1*(maxAspect1-minAspect1)+minAspect1

            call random_number(ranAspect2)
            coffAspect2(i)=ranAspect2*(maxAspect2-minAspect2)+minAspect2

            call random_number(ranHillshade)
            coffHillshade(i)=ranHillshade*(maxHillshade-minHillshade)+minHillshade

            call random_number(ranElevation)
            coffElevation(i)=ranElevation*(maxElevation-minElevation)+minElevation

            call random_number(ranNDVI)
            coffNDVI(i)=ranNDVI*(maxNDVI-minNDVI)+minNDVI

            call random_number(ranNDBSI)
            coffNDBSI(i)=ranNDBSI*(maxNDBSI-minNDBSI)+minNDBSI

            call random_number(ranIntercept)
            coffIntercept(i)=ranIntercept*(maxIntercept-minIntercept)+minIntercept

            do n=1,k
                  ET(n)=(coffslope(i)*(slope(n)))+(coffAspect1(i)*(aspect(n)))&

            +(coffAspect2(i)*(aspect(n))*(aspect(n)))+(coffHillshade(i)*(hillshade(n)))&
                  +(coffElevation(i)*(elevation(n)))+(coffNDBSI(i)*(NDBSI(n)))&
                  +(coffNDVI(i)*(NDVI(n)))+coffIntercept(i)
                  Res(n)=abs(ET(n)-Temp(n))
                  if(Res(n).le.0.)print*, "Not right."

                  sumRes(i)= sumRes(i) + Res(n)

            end do
            MeanRes(i) = sumRes(i)/k
            if(sumRes(i).le.0.)print *, "Not right.", sumRes(i)
            end do
!$OMP end do
!$OMP end parallel
call system_clock(t2, rate)
PT = (t2-t1)/real(rate)

print*,"The lowest average residual is:", minval(MeanRes), "@" ,minloc(MeanRes)
```

```
write (23,*) "The lowest average residual is:",minval(MeanRes), "@",minloc(MeanRes)
print*,"The Multivariate Equation
is:","[",coffSlope(minloc(MeanRes)),"*Slope]+[",coffAspect1(minloc(MeanRes)),"*Asp
ect]+[",&
coffAspect2(minloc(MeanRes)),"*Aspect*Aspect]+[",coffHillshade(minloc(MeanRes)),"
*Hillshade]+[",&
coffElevation(minloc(MeanRes)),"*Elevation]+[",&
coffNDVI(minloc(MeanRes)),"*NDVI]+[",coffNDBSI(minloc(MeanRes)),"*NDBSI]+["
,coffIntercept(minloc(MeanRes)),"]"
print*,'Parallel Computation Time :',PT,'seconds'

write (23,*)"The Multivariate Equation
:","(",coffSlope(minloc(MeanRes)),"*[Slope])+(",coffAspect1(minloc(MeanRes)),"*[Asp
ect])+(",&
coffAspect2(minloc(MeanRes)),"*[Aspect]*[Aspect])+(",coffHillshade(minloc(MeanRes
)),"*[Hillshade])+(",&
coffElevation(minloc(MeanRes)),"*[Elevation])+(",&
coffNDVI(minloc(MeanRes)),"*[NDVI])+(",coffNDBSI(minloc(MeanRes)),"*[NDBSI])
+(",coffIntercept(minloc(MeanRes)),")"
print*,'Parallel Computation Time :',PT,'seconds'
write (23,*)'Parallel Computation Time :',PT,'seconds'
close(23)
end program Multivariate
```

Glossary

Background Emittance	Emittance contributed by background variables, calculated using background temperature
Background Temperature	Temperature contributed by background variables
Background Variables	Independent variables, determining variables, environmental variables such as slope, aspect, elevation
Dependent Variables	Variables dependent on background variables , for example, temperature
Emittance	Total radiative energy emitted from the body per unit area, measured in Watt/m^2
Emissivity	Efficiency of the surface at which an object emits energy, expressed as the fraction of energy being emitted compared to black body (a perfect emitter, with emissivity of 1), value ranges from 0 to 1, unit less.
Raw Emittance	Emittance calculated using raw temperature
Raw Temperature	Temperature calculated using thermal bands
Residual Emittance	Emittance devoid of background emittance obtained by subtracting background from raw emittance
Residual Temperature	Temperature devoid of background temperature obtained by subtracting background from raw temperature