**FINAL REPORT**

# EVALUATING THE USE OF LANDSAT IMAGERY FOR MONITORING GEOTHERMAL HEAT FLOW IN YELLOWSTONE NATIONAL PARK

by

Shannon Savage
Land Resources and Environmental Sciences
Montana State University

CESU TASK AGREEMENT NUMBER: J1580050584

November 2009

## EXECUTIVE SUMMARY

Yellowstone National Park (YNP) contains the world's largest concentration of geothermal features, and is legally mandated to protect and monitor these natural features. Remote sensing is a component of the current geothermal monitoring plan. Landsat satellite data have a substantial historical archive and will be collected into the future, making it the only available thermal infrared imagery for historical analysis and long-term monitoring of geothermal areas in the entirety of YNP. Landsat imagery from Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors was explored as a tool for mapping geothermal heat flux and geothermally active areas within YNP and to develop a change analysis technique for scientists to utilize with additional Landsat data available from 1978 through the foreseeable future.

Terrestrial emittance and estimates of geothermal heat flux were calculated for the entirety of YNP with fourteen Landsat summer images from 1986 to 2007. Difference images of terrestrial emittance for four of the summer images (2007 minus 2006, and 2002 minus 2001) were created. Finally, TM data from 2007, 2006, 2005, 1989, and 1986 were used to classify geothermally active areas inside the defined geothermal areas as well as throughout YNP and a 30-km buffer around YNP. Four change maps were produced from the five classified maps.

Estimations of geothermal heat flux were inaccurate due to inherent limitations of Landsat data combined with complexities arising from the effects of solar radiation and spatial and temporal variation of vegetation, microbes, steam outflows, and other features at each geothermal area. Terrestrial emittance, however, was estimated with acceptable results. The change analysis showed a relationship between absolute difference in terrestrial emittance and earthquake swarms, with 34% of the variation explained. Accuracies for the classifications of geothermally active areas were poor, but the method used for classification, RandomForest, could be a suitable method given higher resolution thermal imagery and better reference data.

Disclaimer: The views expressed are the authors and do not necessarily represent the views of Yellowstone National Park, the Department of Interior or the United States Government.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF TABLES - CONTINUED

## LIST OF FIGURES

## INTRODUCTION

Yellowstone National Park (YNP), located in Wyoming, Montana, and Idaho, became the world's first national park primarily because of its geothermal features. The land was set aside for the "benefit and enjoyment of the people" and to "provide for the preservation from injury or spoliation of all timber, mineral deposits, natural curiosities, or wonders within said park, and their retention in their *natural condition*" (Yellowstone Park Act, 1872) (*emphasis mine*). Currently there are recognized external threats on the geothermal features of YNP, including possible geothermal development in Idaho and Montana, and oil, gas, and groundwater development in Wyoming, Montana, and Idaho (Sorey, 1991; Custer *et al.*, 1993; Heasler *et al.*, 2004). Geothermal features are also found in two park units neighboring YNP, John D. Rockefeller, Jr. Memorial Parkway and Grand Teton National Park in Wyoming and might be affected by these external threats as well. Other potential drivers that likely influence changes in geothermal features include drought, mean annual temperature increases or decreases, changes in barometric pressure, earthquakes in the vicinity or on the other side of the planet, and caldera resurgence and deflation (Rojstaczer *et al*., 2003).

The National Park Service (NPS) is legally mandated to monitor and protect geothermal features within its units, and YNP is listed as a significant feature on its own (Geothermal Steam Act, 1970 as amended in 1988). More importantly, details of change within geothermal systems are poorly known. Knowledge about change might provide scientific insight into patterns that would help advance the understanding of processes in important geothermal systems. A better understanding of these systems would help with placement of visitor information and would be an important planning tool for placing infrastructure in YNP. Finally, there is a growing demand for alternative energy in the United States and the development of geothermal energy is imminent. The impact of this sort of development outside YNP on the geothermal features inside YNP will become an important issue. A geothermal monitoring plan that includes groundwater inventory, monitoring, and assessment, chloride flux inventory, monitoring, and assessment, and remote sensing of geothermal features, has been proposed for YNP to address these issues (Heasler et al., 2004). Remote sensing is an important element of the plan since it is an excellent way to assess historic change and has great potential for providing methods for future monitoring of geothermal areas, even where the spatial resolution might be too coarse for monitoring individual features.

## GOALS

The general purpose of this project was to evaluate the use of Landsat data for mapping and monitoring geothermal heat flow in YNP. There were three specific goals for this project. The first goal was to assess the utility of Landsat TM and ETM+ thermal infrared imagery for monitoring GHF within the boundaries of YNP by calculating estimates of terrestrial emittance and GHF with 14 Landsat images. The second goal was to conduct a change analysis of the spatial distribution of terrestrial emittance within

YNP for two pairs of years. Both Landsat TM (2007 and 2006) and Landsat ETM+ (2002 and 2001) images were to be inspected. The final goal was to assess the ability of Landsat TM imagery combined with RandomForest and target detection classification methods to classify active areas accurately within YNP's defined geothermal areas as well as throughout YNP and 30-km beyond its boundary. Five years of Landsat imagery were to be classified (2007, 2006, 2005, 1989, and 1985).

## AVAILABLE LANDSAT IMAGES

The first Landsat satellite, carrying the Multi Spectral Scanner (MSS), was launched in 1972. Subsequent satellites including those launched in 1984 (Landsat 5, carrying Thematic Mapper (TM)) and 1999 (Landsat 7, carrying Enhanced Thematic Mapper Plus (ETM+)) are still in orbit. Thermal data became available in 1978 with the MSS sensor on Landsat 3. This provides over 30 years of data that can be utilized for change detection and landscape studies of geothermal heat at YNP. Landsat scenes cover a swath of 185 km (YNP is in the center of one of these scenes: Path 38, Row 29) and currently collect data every 8 days (TM and ETM+ are in relatively opposite orbits and each have a repeat coverage of 16 days, thus the same location is imaged every 8 days).

Landsat sensors collect data from the visible portion of the electromagnetic spectrum along with several wavelengths from the infrared portion of the spectrum (near infrared, middle infrared, and thermal infrared). A panchromatic band is also collected by the ETM+ sensor. Landsat TM and ETM+ reflective spectral data have a spatial resolution of 28.5-m (typically resampled to 30-m on a side or 900 $m^2$), while all MSS data have a spatial resolution of 57-m x 79-m (resampled by the EROS Data Center to 60-m on a side or 3,600 $m^2$). thermal infrared data have a spatial resolution of 120-m on a side from the TM sensor and 60-m on a side from the ETM+ sensor.

Landsat data are now available for free download from the USGS EROS Data Center. Landsat TM data are available from July 1982 to present, while ETM+ data are available in complete form from April 1999 until May 2003 (prior to the failure of the scan line corrector) and with scan line gaps from the end of May 2003 to present. The TM images, despite 25 years of sensor degradation, are commonly used as replacements for ETM+ data after May 2003.

Geometric registration of all images is vital when comparing different images in a change analysis. A "master" TM image (7 September 2005) was chosen that aligned well with the National Agriculture Imagery Program (NAIP) imagery (root mean square error (RMSE) = 0.4128 pixels, or less than 15 m) and roads and trails data recorded with sub-meter GPS units by YNP staff. All images acquired for this project were geometrically registered to the master image with an RMSE for each registration of less than 0.5 pixels (15 m) (Table 1).

Table 1: Original data collected for CESU project included in the final data archive. Data marked with ** were purchased with project funding.

| Scene | Sensor | Acquisition Date | Format | Correction | Resample | Source |
|---|---|---|---|---|---|---|
| p38r29_mss_19810630 | MSS2 | 19810630 | NLAPS | Systematic | NN | Eros Data Center** |
| p38r29_tm_19860717 | TM5 | 19860717 | NLAPS | Precision | CC | YNP Spatial Analysis Center |
| p38r29_tm_19890802 | TM4 | 19890802 | NLAPS | Precision | CC | YNP Spatial Analysis Center |
| p38r29_tm_19910715 | TM5 | 19910715 | GeoTiff | Geometric | NN | YNP Spatial Analysis Center |
| p38r29_tm_19950912 | TM5 | 19950912 | NLAPS | Systematic | NN | Eros Data Center** |
| p38r29_tm_19960712 | TM5 | 19960712 | GeoTiff | Terrain | CC | GloVis |
| p38r29_tm_19970715 | TM5 | 19970715 | NLAPS | Precision | CC | WyomingView |
| p38r29_tm_19980718 | TM5 | 19980718 | NLAPS | Precision | CC | WyomingView |
| p38r29_etm_19990713 | ETM+ | 19990713 | NLAPS | Precision | NN | YNP Spatial Analysis Center |
| p38r29_etm_20000715 | ETM+ | 20000715 | GeoTiff | Systematic | CC | MontanaView |
| p38r29_etm_20010702 | ETM+ | 20010702 | GeoTiff | Systematic | CC | MontanaView |
| p38r29_etm_20020705 | ETM+ | 20020705 | GeoTiff | Precision | CC | MontanaView |
| p38r29_tm_20030801 | TM5 | 20030801 | NLAPS | Precision | CC | WyomingView |
| p38r29_tm_20050721 | TM5 | 20050721 | GeoTiff | Precision | CC | MontanaView |
| p38r29_tm_20050907 | TM5 | 20050907 | GeoTiff | Precision | CC | MontanaView |
| p38r29_etm_20050915 | ETM+ SLC-off | 20050915 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_tm_20060505 | TM5 | 20060505 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_etm_20060513 | ETM+ SLC-off | 20060513 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_tm_20060606 | TM5 | 20060606 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_tm_20060708 | TM5 | 20060708 | GeoTiff | Precision | NN | MontanaView |
| p38r29_etm_20060716 | ETM+ SLC-off | 20060716 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_etm_20060902 | ETM+ SLC-off | 20060902 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_tm_20061012 | TM5 | 20061012 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_tm_20061028 | TM5 | 20061028 | NLAPS | Precision | NN | Eros Data Center** |
| p38r29_tm_20070625 | TM5 | 20070625 | NLAPS | Precision | NN | Eros Data Center** |

## STUDY AREA

Three different study area boundaries were utilized for this study (Figure 1). The first study area was simply the boundary of YNP, covering approximately 890,000 ha in Wyoming, Montana, and Idaho, USA. The second, and smallest, study area was the boundaries of the currently defined geothermal areas, covering 6,343 ha, and comprising less than 1% of the entire area of YNP. Rick Hutchinson, geologist for YNP from 1976 to 1996, spent many years studying the geothermal areas in YNP and during that time produced maps of the geothermal boundaries. Those maps have subsequently been updated and checked for accuracy by staff at YNP's Spatial Analysis Center to produce the most up-to-date digital map of the defined geothermal areas (Spatial Analysis Center, 2005). The third study area included all of YNP plus a 30-km buffer around the YNP

boundary, covering approximately 2,400,000 ha. A 30-km buffer was delineated around YNP so that the Corwin Springs, Montana, and Island Park, Idaho, Known Geothermal Resource Areas (KGRAs) (Long *et al.*, 1976; Sorey, 1991) would be included in the classification process.

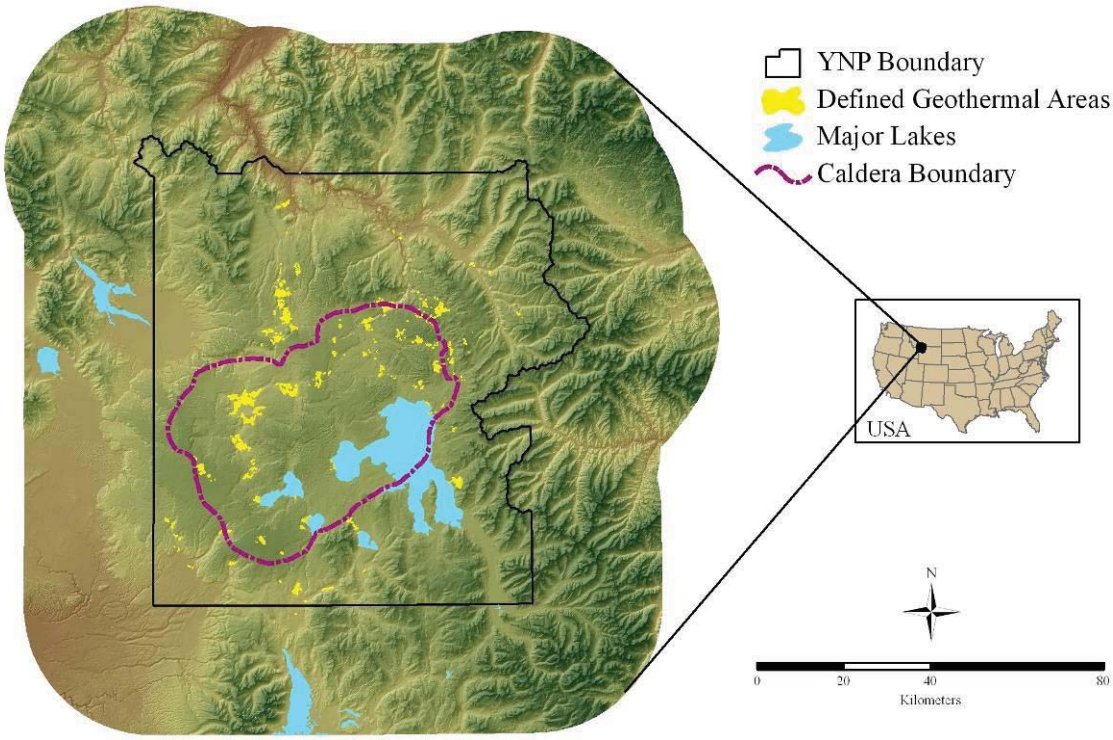


Figure 1: Location of study area boundaries displayed on a shaded relief map. Three different study area boundaries were utilized in this project: (1) the currently defined geothermal areas, (2) the boundary of Yellowstone National Park (YNP), and (3) a 30-km buffer around YNP.

## ESTIMATING TERRESTRIAL EMITTANCE AND HEAT FLUX

Geothermal heat flux (GHF) is the variation of heat in geothermal systems and is radiated, or emitted, from the surface of the Earth. It represents only heat coming from below the surface and it does not include any accumulated indirect or direct solar heating effects such as convection from air currents, and conduction of solar effects on soil (indirect), or solar heating due to variations in topography such as south-facing slopes (direct). GHF can be measured from bore holes (Sorey, 1991), by estimation from other indirect measurements such as chloride flux (Fournier *et al.*, 1975; Norton and Friedman, 1985; Friedman and Norton, 2007), or by utilizing thermal sensors (Boomer *et al*., 2002). Terrestrial emittance represents the heat emitted from the ground and is composed of GHF and includes direct and indirect solar radiation effects.

Multispectral Landsat satellite imagery has been used to map geothermal heat and activity in a variety of situations. Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) imagery have been used successfully to map and analyze volcanic features (Andres and Rose, 1995; Kaneko and Wooster, 1999; Flynn *et al.*, 2001; Urai, 2002; Patrick *et al.*, 2004). Many studies have used TM and ETM+ data to map lineaments (*e.g.*, fault lines) as part of the process of finding geothermal areas (Bourgeois *et al.*, 2000; Song *et al.*, 2005) and to map minerals such as iron oxide and hydrothermally altered soil (Carranza and Hale, 2002; Daneshfar *et al.*, 2006; Dogan, 2008).

Landsat thermal infrared imagery, however, has rarely been used to assess the spatial distribution of GHF in YNP, and in one instance, only one image was used for a snapshot of GHF (Watson *et al.*, 2008). The method developed by Watson *et al*., 2008 to quantify the intensity of surficial geothermal activity at YNP, was developed with 2000 Landsat ETM+ winter imagery, and the results suggested good potential for geothermal monitoring. Thermal radiance data from ETM+ imagery were utilized to estimate terrestrial emittance. Estimates of non-geothermal-related heat were incorporated with terrestrial emittance to subsequently measure and create a map of continuous variations in residual terrestrial emittance (*i.e.*, no solar effects) that was hypothesized to estimate a lower bound for GHF.

The Watson *et al*., 2008 method utilized a spectral library of "light yellowish brown loamy sand" from the NASA Jet Propulsion Laboratory (JPL) to estimate a single emissivity value for the entire image. This method might be improved upon by assigning emissivity on a pixel-by-pixel basis rather than using a single value. Emissivity can be estimated from a Normalized Difference Vegetation Index (NDVI) that uses the red and NIR Landsat bands to represent amounts of healthy green vegetation (Brunsell and Gillies, 2002). The estimated emissivity can be applied to the calculation of terrestrial emittance, and thus to estimations of GHF.

Landsat data can be valuable for calculation of GHF in YNP. The method suggested in this paper is not highly parameterized – it requires only three Landsat bands and some atmospheric correction coefficients. Emissivity is incorporated per pixel rather than as one value across the entire image, potentially increasing the precision of the GHF calculations. Finally, one Landsat image covers the entire area of YNP. Landsat data provide the means to calculate GHF for all of YNP and has the potential to enable scientists to identify locations that might need to be studied in more depth.

Using Landsat data to estimate GHF presents many challenges. Solar radiation and related topographic effects have substantial impacts on total emittance calculations since, for example, south-facing slopes that have no GHF will often have high terrestrial emittance values (Watson, 1975; Kohl, 1999; Gruber *et al.*, 2004). The effect of surface albedo is also an important component and problematic in the calculation of GHF, because dark areas such as large parking lots (*e.g.*, in the Old Faithful area) or recently burned areas absorb and re-emit large amounts of solar radiation than bright surfaces, resulting in high terrestrial emittance readings that might not include a GHF component (Watson, 1975; Coolbaugh *et al.*, 2007).

The Landsat ETM+ sensor is superior to the Landsat TM sensor because its thermal infrared sensor is kept calibrated by a more stable radiative cooler and it has finer spatial resolution (NASA, 2009). There are only four years of complete data available

from Landsat ETM+, while Landsat TM is 25 years old and its thermal infrared sensor has deteriorated over the years. This deterioration also might make changes in GHF more difficult to detect. The pixel resolution for both ETM+ (60 m) and TM (120 m) thermal infrared data is much coarser than for the reflective data from both sensors (30 m). When one pixel is 60 m on a side (3,600 $m^2$) or 120 m on a side (14,400 $m^2$), effects from small geothermal features or areas are averaged over the pixel.

The main purpose of this portion of the project was to evaluate the utility of Landsat TM and ETM+ thermal infrared data for monitoring GHF. An effective method would enable the calculation of terrestrial emittance and GHF covering the entirety of YNP that could be applied to additional Landsat images for use in monitoring and change analyses. Previous studies in YNP have been for a single date and/or over limited geographic areas.

## Methods

### Image Preprocessing

YNP is centered within one Landsat scene at Path 38 Row 29. Fourteen Landsat TM and ETM+ summer images from 1986 to 2007 were acquired from various sources (Table 2). These images were chosen based on snow-free summer anniversary dates and lack of clouds. Summer dates were selected because although winter images were preferred for comparison to other work, the majority of the available winter Landsat images were cloud-covered and therefore not practical for use in this project. Image dates range from 25 June to 2 August resulting in anniversary dates within 5-½ weeks of one another. Two of the images are cloud free, while the remaining 12 have less than 5% cloud cover.

Table 2: Landsat images used in this study. Images marked with * are cloud free.

| Acquisition Date (1980s) | Sensor | Acquisition Date (1990s) | Sensor | Acquisition Date (2000s) | Sensor |
|---|---|---|---|---|---|
| 17 July 1986 | TM5 | 15 July 1991 | TM5 | 15 July 2000 | ETM+ |
| 2 August 1989 | TM4 | 12 July 1996 | TM5 | 2 July 2001 | ETM+ |
| | | 15 July 1997 | TM5 | 5 July 2002 | ETM+ |
| | | 18 July 1998 | TM5 | 1 August 2003 | TM5 |
| | | 13 July 1999* | ETM+ | 21 July 2005* | TM5 |
| | | | | 8 July 2006 | TM5 |
| | | | | 25 June 2007 | TM5 |

Each image was clipped to the YNP boundary. Clouds and cloud shadows were masked by on-screen digitizing. Elevations greater than 2,700 m were masked to remove snow from the input data. The thermal infrared band for each of the 14 images was degraded to 120-m pixel size as requested by the NPS. Resampling was necessary because of differing resolutions among the source images; 120-m was the coarsest

resolution of the source images and resampling to a finer resolution would create false precision, therefore the coarsest resolution was selected. The COSine Transformation (COST) (Chavez, 1996) method of dark object subtraction atmospheric and radiometric correction was applied to the original raw data values of the six reflective bands of each image. The original Landsat raw data values are represented by digital numbers (or DNs) with values from 0 to 255 (8-bit radiometric resolution). The dark object DN values were chosen by examining the image histogram for each of the six reflective bands. The DN value where the histogram increased to more than 100 pixels was assigned the dark object value. These values along with information from the Landsat header files were used to convert the images to surface reflectance values for Landsat bands 1, 2, 3, 4, 5, and 7 at a 30-m pixel size (Utah State University, 2008).

The Normalized Difference Vegetation Index (NDVI) was used to estimate fractional vegetation (Fr, unitless) based on the method by Brunsell and Gillies (2002). Fractional vegetation represents the percentage of vegetation within a pixel and is derived from NDVI as follows:

$$Fr = [(NDVI - NDVI_0)/(NDVI_{max} - NDVI_0)]^2 \quad \text{(Equation 1)}$$

where $NDVI_0$ represents bare soil and $NDVI_{max}$ represents scene specific maximum vegetation. Assuming average broad-band emissivity for bare soil of 0.97 (from the "light yellowish brown loamy sand" and "white gypsum dune sand" JPL spectral libraries (NASA, 2008)) and emissivity for vegetation of 0.98 (from the "coniferous vegetation" JPL spectral library (NASA, 2008)), emissivity ($\varepsilon$, unitless) per pixel (excepting water pixels) was estimated from the Fr:

$$\varepsilon = Fr*\varepsilon_v + (1 - Fr)*\varepsilon_s \quad \text{(Equation 2)}$$

where $\varepsilon_v$ represents vegetation emissivity and $\varepsilon_s$ represents soil emissivity. Water pixels were assigned an average broad-band emissivity value of 0.99 (Shaw and Marston, 2000). To match the lower-resolution TIR imagery, the resulting emissivity image was subsequently degraded to 120-m pixels by averaging the 30-m pixel values.

Terrestrial Emittance and Heat Flux Calculation Procedures

The raw TIR data (band 6) for each image was converted to at-satellite radiance ($L_\lambda$, $Wm^{-2}sr^{-1}\mu m^{-1}$) using published calibration factors (Chander *et al.*, 2009). Radiance was converted to top-of-atmosphere emittance ($M_{toa}$, $Wm^{-2}$) by integrating over the bandwidth (from 10.4 μm to 12.5 μm = 2.1 μm) and the hemisphere ($\pi$ sr):

$$M_{toa,\,6H} = 2.1\pi L_\lambda \quad \text{(Equation 3)}$$

MODerate resolution atmospheric TRANsmission (ModTran) was utilized to estimate atmospheric transmittance ($\tau$) and upwelling atmospheric emittance ($M_{up}$, $Wm^{-2}$) for a "Mid Latitude Summer" model atmosphere. Following the Watson method

(Watson *et al.*, 2008), surface emittance integrated over band 6 ($M_{surf, 6H}$, $Wm^{-2}$) was estimated:

$$M_{surf,6H} = (M_{toa,6H} - M_{up})/\tau \quad \text{(Equation 4)}$$

where $M_{up} = 4.64$ $Wm^{-2}$ and $\tau = 89.39\%$. The fitted coefficients from Watson's regression model were utilized to estimate broad-band surface emittance ($M_{surf}$, $Wm^{-2}$):

$$M_{surf} = (0.004812M_{surf,6H})^2 + 2.653M_{surf,6H} + 181.8$$

Terrestrial emittance ($M_{terr}$, $Wm^{-2}$) was estimated using the NDVI-derived emissivity values and downwelling atmospheric emittance ($M_{down}$, $Wm^{-2}$) calculated with ModTran for a "Mid Latitude Summer" model atmosphere:

$$M_{terr} = M_{surf} - (1 - \varepsilon)M_{down} \quad \text{(Equation 5)}$$

where $\varepsilon$ ranges from 0.97 to 0.99, and $M_{down} = 240$ $Wm^{-2}$.

Potential annual direct incident solar radiation (SR) was calculated from a 30-m digital elevation model (DEM) of the study area (McCune and Keon, 2002) to take solar effects into account. This equation incorporated the slope, aspect, and latitude of the terrain and returns SR in units of MJ $cm^{-2}$ $yr^{-1}$:

$$SR = 0.339 + 0.808(\cos(L)*\cos(S)) - 0.196(\sin(L)*\sin(S)) - 0.482(\cos(A)*\sin(S)) \quad \text{(Equation 6)}$$

where L = latitude in radians, S = slope in radians, and A = folded aspect in radians east of north (this rescales 0-360° to 0-180°, so NE = NW, E = W, and so on, to emphasize north/south contrast). The output values were multiplied by 316.89 $Js^{-1}m^{-2}$ to arrive at SR in $Wm^{-2}$. This image was degraded to 60-m and 120-m pixel images.

Albedo was calculated from five of the six reflective Landsat bands. The green band (band 2) was excluded because it does not improve the $R^2$ of the regression. The DNs were first converted to surface reflectance then applied to the following shortwave albedo calculation (unitless):

$$\alpha_{short} = 0.356\alpha_1 + 0.130\ \alpha_3 + 0.373\ \alpha_4 + 0.085\ \alpha_5 + 0.072\ \alpha_7 - 0.0018 \quad \text{(Equation 7)}$$

where $\alpha_{\#}$ refers to the Landsat band (Liang, 2000).

The estimate of GHF was calculated by incorporating albedo and SR so that locations with low albedo and high absorption of solar radiation, for instance a recent fire scar, would not result in falsely high GHF:

$$GHF_{\alpha} = M_{terr} - (SR * (1 - \alpha_{short})) \quad \text{(Equation 8)}$$

where $1 - \alpha_{short}$ is absorption based on Kirchoff's law (Elachi, 1987).

Temperature in °C was calculated from $M_{terr}$ based on the Stefan-Boltzman Law ($M = \varepsilon\sigma T^4$, where M is emittance, $\varepsilon$ is emissivity, $\sigma = 5.67 \text{ X } 10^{-8}$ $Wm^{-2}K^{-4}$, and T is temperature in units Kelvin):

$$TempC = (M_{terr}/0.0000000567)^{0.25} - 273.15 \quad \text{(Equation 9)}$$

Comparison to Airborne Data

The $M_{terr}$ values for the July 2002 image in the Norris Geyser Basin area were compared to the summary statistics and heat flow values from a nighttime airborne thermal infrared image of the same area from October 2002 (Hardy, 2005; Seielstad and Queen, 2009). The Hardy (2005) data originally had a pixel resolution of 0.76 m on a side. These pixels were degraded to 60 m on a side to match the Landsat data. Two extents were examined: (1) the entire extent of the Hardy data, and (2) the boundary of Norris Geyser Basin according to the defined geothermal areas (Spatial Analysis Center, 2005). Summary statistics and total heat flow were calculated for the four images and compared. The moderate-resolution, daytime, summer, 10.4 to 12.5 μm Landsat data were compared to the high-resolution, nighttime, winter, 3 to 5 μm airborne data, despite their differences, in order to assess the quality of the Landsat-derived $M_{terr}$. The Hardy (2005) data were the only area-wide data available for such a comparison. Since both methods seek to estimate total flux based on the portions of the spectrum sampled, the comparison is relevant to an evaluation of whether they are of the same order of magnitude.

## Results and Discussion

Terrestrial Emittance and GHF in Yellowstone National Park

Estimated $M_{terr}$ values ranged from 216.98 $Wm^{-2}$ to 495.98 $Wm^{-2}$ for the 120-m resolution images (highest and lowest values from the 1991 image) (Table 3). The range of estimated $M_{terr}$ values for the 60-m resolution images was from 308.59 $Wm^{-2}$ to 448.16 $Wm^{-2}$. Mean values of $M_{terr}$ ranged from 319.93 $Wm^{-2}$ to 378.66 $Wm^{-2}$ for the 14 120-m resolution images and from 368.90 $Wm^{-2}$ to 382.02 $Wm^{-2}$ for the 4 60-m resolution images.

Table 3: Summary statistics for terrestrial emittance ($M_{terr}$) for 4 60-m resolution and 14 120-m resolution summer Landsat images of Yellowstone National Park ($Wm^{-2}$)

**60-m ETM+ Images**

| Year | Min | Max | Mean | Median | Mode | Std. Dev. |
|---|---|---|---|---|---|---|
| **2002** | 308.59 | 446.74 | 368.90 | 367.95 | 364.71 | 18.88 |
| **2001** | 310.29 | 445.97 | 375.14 | 374.42 | 376.54 | 20.01 |
| **2000** | 323.98 | 448.16 | 382.02 | 381.71 | 374.91 | 20.10 |
| **1999** | 314.56 | 447.01 | 378.51 | 377.68 | 373.02 | 20.17 |

**120-m TM and ETM+ Images**

| Year | Min | Max | Mean | Median | Mode | Std. Dev. |
|---|---|---|---|---|---|---|
| **2007** | 303.84 | 413.64 | 353.90 | 354.45 | 353.16 | 15.19 |
| **2006** | 302.21 | 410.62 | 351.47 | 350.49 | 351.76 | 11.93 |
| **2005** | 313.71 | 449.77 | 370.10 | 368.98 | 367.39 | 18.66 |
| **2003** | 287.95 | 427.6 | 362.51 | 361.05 | 363.78 | 17.68 |
| **2002** | 311.05 | 441.27 | 368.90 | 368.53 | 363.95 | 18.55 |
| **2001** | 311.14 | 445.85 | 375.12 | 374.81 | 370.6 | 19.68 |
| **2000** | 232.29 | 446.41 | 378.66 | 381.17 | 373.64 | 26.65 |
| **1999** | 316.02 | 445.99 | 378.60 | 377.96 | 373.89 | 19.81 |
| **1998** | 314.23 | 420.71 | 362.53 | 361.65 | 359.98 | 16.10 |
| **1997** | 265.95 | 448.95 | 349.23 | 347.44 | 345.3 | 14.40 |
| **1996** | 289.56 | 408.4 | 350.45 | 348.98 | 342.94 | 16.25 |
| **1991** | 216.98 | 495.98 | 355.34 | 353.21 | 342.31 | 16.99 |
| **1989** | 279.2 | 475.56 | 335.97 | 332.13 | 325.99 | 17.11 |
| **1986** | 298.25 | 379.36 | 319.93 | 317.53 | 315.99 | 7.86 |

Estimated $GHF_{\alpha}$ values ranged from -86.94 $Wm^{-2}$ to 377.36 $Wm^{-2}$ for the 14 120-m resolution images (highest and lowest values from the 2000 image) (Table 4). The estimated values of $GHF_{\alpha}$ for the 4 60-m resolution images ranged from 16.68 $Wm^{-2}$ to 382.59 $Wm^{-2}$. Mean $GHF_{\alpha}$ values ranged from 63.35 $Wm^{-2}$ to 128.06 $Wm^{-2}$ for the 14 120-m resolution images and from 118.02 $Wm^{-2}$ to 132.11 $Wm^{-2}$ for the 4 60-m resolution images.

Table 4: Summary statistics for albedo and potential annual direct incident solar radation corrected geothermal heat flux ($GHF_{\alpha}$) for 4 60-m resolution and 14 120-m resolution summer Landsat images of Yellowstone National Park ($Wm^{-2}$)

| **60-m ETM+ Images** | | | | | | |
|---|---|---|---|---|---|---|
| **Date** | **Min** | **Max** | **Mean** | **Median** | **Mode** | **Std. Dev.** |
| **2002** | 16.68 | 377.96 | 118.02 | 116.88 | 111.23 | 32.32 |
| **2001** | 26.44 | 360.22 | 125.60 | 125.53 | 122.93 | 32.09 |
| **2000** | 35.82 | 372.63 | 132.11 | 131.86 | 129.23 | 32.73 |
| **1999** | 26.03 | 382.59 | 126.58 | 126.31 | 122.13 | 33.42 |

| **120-m TM and ETM+ Images** | | | | | | |
|---|---|---|---|---|---|---|
| **Date** | **Min** | **Max** | **Mean** | **Median** | **Mode** | **Std. Dev.** |
| **2007** | 18.88 | 336.29 | 101.97 | 100.71 | 101.95 | 29.68 |
| **2006** | 24.06 | 344.83 | 100.00 | 97.99 | 96.73 | 26.99 |
| **2005** | 26.42 | 352.91 | 119.87 | 119.52 | 115.69 | 31.51 |
| **2003** | 0.79 | 333.60 | 110.30 | 109.99 | 109.99 | 30.46 |
| **2002** | 24.21 | 336.10 | 118.06 | 118.02 | 116.80 | 31.19 |
| **2001** | 27.31 | 326.88 | 125.63 | 125.61 | 122.10 | 31.12 |
| **2000** | -86.94 | 377.36 | 128.30 | 130.70 | 125.26 | 38.13 |
| **1999** | 27.48 | 365.07 | 126.53 | 126.39 | 119.79 | 33.31 |
| **1998** | 28.06 | 350.75 | 111.39 | 111.25 | 108.73 | 29.45 |
| **1997** | -3.32 | 341.20 | 95.87 | 93.57 | 92.23 | 28.97 |
| **1996** | 9.07 | 348.41 | 91.16 | 87.28 | 79.32 | 28.61 |
| **1991** | -36.82 | 337.91 | 105.71 | 103.71 | 97.85 | 30.45 |
| **1989** | 0.01 | 340.24 | 78.34 | 73.11 | 61.14 | 29.30 |
| **1986** | 4.08 | 306.08 | 63.35 | 58.34 | 50.09 | 24.72 |

Estimated temperature values ranged from -24.43 °C to 32.67 °C for the 14 120-m resolution images (highest and lowest values from the 1991 image) (Table 5). The estimated values of temperature for the 4 60-m resolution images ranged from -1.54 °C to 25.02 °C. Mean temperature values ranged from 0.91 °C to 12.63 °C for the 14 120-m resolution images and from 10.79 °C to 13.28 °C for the 4 60-m resolution images.

Table 5: Summary statistics for temperature calculated from terrestrial emittance ($M_{terr}$) for 4 60-m resolution and 14 120-m resolution summer Landsat images of Yellowstone National Park (°C)

| **60-m ETM+ Images** | | | | | | |
|---|---|---|---|---|---|---|
| **Date** | **Min** | **Max** | **Mean** | **Median** | **Mode** | **Std. Dev.** |
| **2002** | -1.54 | 24.78 | 10.79 | 10.70 | 10.90 | 3.65 |
| **2001** | -1.16 | 24.65 | 11.97 | 11.95 | 11.64 | 3.83 |
| **2000** | 1.79 | 25.02 | 13.28 | 13.31 | 12.68 | 3.80 |
| **1999** | -0.23 | 24.83 | 12.61 | 12.49 | 12.00 | 3.83 |

| **120-m TM and ETM+ Images** | | | | | | |
|---|---|---|---|---|---|---|
| **Date** | **Min** | **Max** | **Mean** | **Median** | **Mode** | **Std. Dev.** |
| **2007** | -2.59 | 19.10 | 7.88 | 8.00 | 7.83 | 3.04 |
| **2006** | -2.95 | 18.57 | 7.41 | 7.22 | 6.72 | 2.37 |
| **2005** | -0.42 | 25.29 | 11.02 | 10.83 | 9.82 | 3.60 |
| **2003** | -6.20 | 21.54 | 9.56 | 9.40 | 9.08 | 3.45 |
| **2002** | -1.00 | 23.87 | 10.79 | 10.76 | 10.27 | 3.59 |
| **2001** | -0.98 | 24.64 | 11.98 | 11.93 | 11.63 | 3.83 |
| **2000** | -20.15 | 24.73 | 12.57 | 13.16 | 12.10 | 5.39 |
| **1999** | 0.08 | 24.66 | 12.63 | 12.56 | 11.31 | 3.76 |
| **1998** | -0.03 | 20.34 | 9.57 | 9.37 | 9.13 | 3.15 |
| **1997** | -11.45 | 25.15 | 6.95 | 6.71 | 5.42 | 2.88 |
| **1996** | -5.83 | 18.17 | 7.18 | 7.02 | 4.96 | 3.24 |
| **1991** | -24.43 | 32.67 | 8.15 | 7.69 | 6.35 | 3.35 |
| **1989** | -8.25 | 29.48 | 4.23 | 3.39 | 2.21 | 3.46 |
| **1986** | -3.84 | 12.85 | 0.91 | 0.59 | 0.07 | 1.67 |

Estimating terrestrial emittance requires only three Landsat bands and atmospheric and radiometric corrections. Terrestrial emittance includes all the types of heat emitted from the ground: GHF, direct incident solar radiation, and indirect solar effects such as convection from air currents, and soil conduction of solar energy.

Mean annual air temperature is a good representation of ground water temperature. Mean temperature and $M_{terr}$ values were similar to, but slightly higher than, the average annual air temperature of YNP (4.64 °C, or 337.6 $Wm^{-2}$) (Western Regional Climate Center, 2005), demonstrating that the model includes geothermal as well as non-geothermal heat. Many of the hottest $M_{terr}$ pixels were located on low-elevation flat and south-facing slopes (in the Northern Range), and within 1988 fire scars that have been revegetated with thick stands of young lodgepole pine (*Pinus contorta*) intermixed with down and standing grey and white snags (Figure 2).

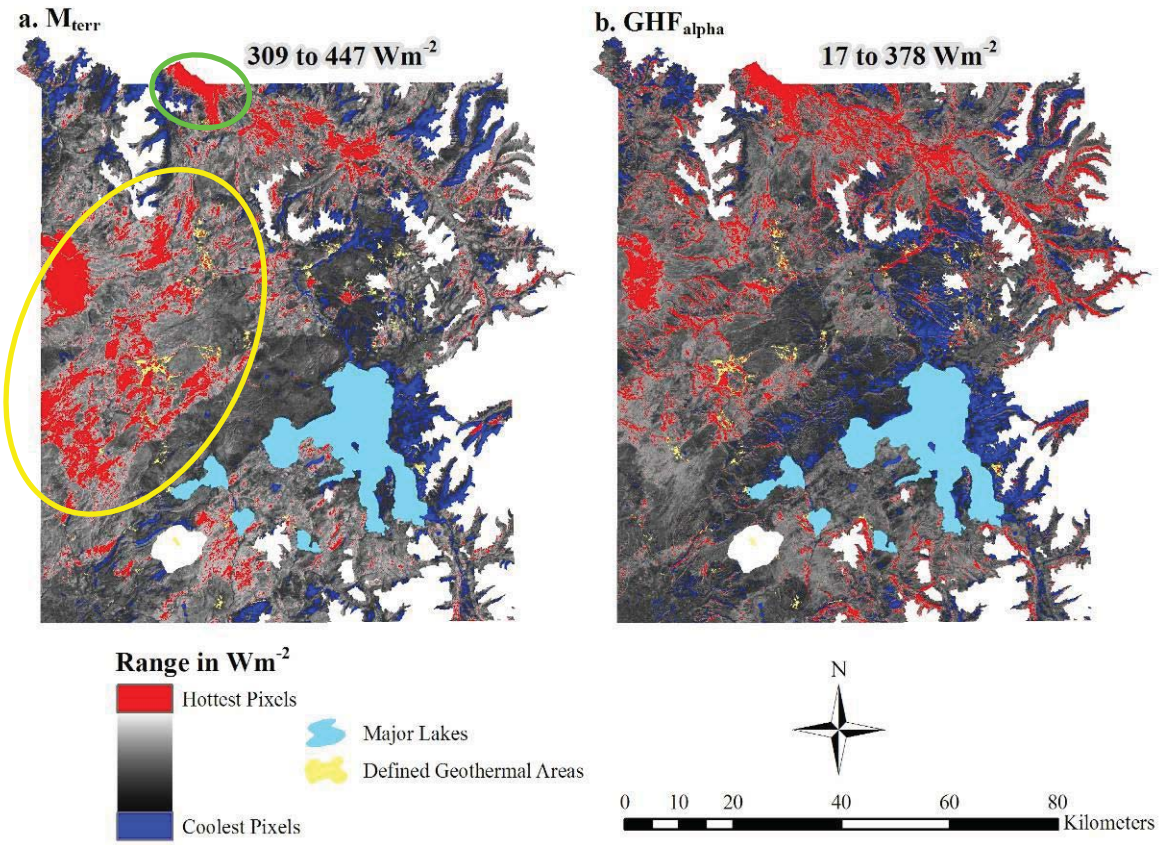


Figure 2: Range of values for (a) terrestrial emmitance (with many of the 1988 fire scars circled in yellow and a portion of the Northern Range circled in green ($M_{terr}$) and (b) albedo and potential annual direct incident solar radiation corrected geothermal heat flux ($GHF_{\alpha}$) in Yellowstone National Park on 5 July 2002 ($Wm^{-2}$). White areas were snow or cloud-covered.

Comparison to Airborne Data in the Norris Geyser Basin Area

The summary statistics of $M_{terr}$ values were similar to the Hardy heat flow summary statistics (Table 6). The Hardy (2005) data had higher values overall than the $M_{terr}$ data, with the maximum values much higher and the minimum values only slightly higher. The range of the Hardy data was more than double the range of the $M_{terr}$ data. The total heat flow values for $M_{terr}$ were within an order of magnitude of the Hardy heat flow data for both the full Hardy data extent and a subset that covers just Norris Geyser Basin (Table 7).

Table 6: Comparison of October 2002 Hardy (2005) heat data summary statistics to July 2002 estimated terrestrial emittance ($M_{terr}$) summary statistics (values in $Wm^{-2}$). Information from the full 2002 Hardy data and Norris Geyser Basin extents are displayed.

| Hardy full data extent | Min | Max | Mean | Median | Mode | Std. Dev. |
|---|---|---|---|---|---|---|
| **Hardy data 60 m** | 342.2 | 664.3 | 407.7 | 401.3 | 387.5 | 35.7 |
| **$M_{terr}$ 60 m** | 338.2 | 418.5 | 378.2 | 377.1 | 376.4 | 13.7 |

| Norris Geyser Basin | Min | Max | Mean | Median | Mode | Std. Dev. |
|---|---|---|---|---|---|---|
| **Hardy data 60 m** | 361.8 | 664.3 | 432.0 | 426.7 | 418.5 | 37.0 |
| **$M_{terr}$ 60 m** | 353.4 | 418.5 | 387.5 | 387.7 | 393.3 | 11.9 |

Table 7: Comparison of October 2002 Hardy (2005) total heat flow and power values to July 2002 estimated terrestrial emittance ($M_{terr}$) heat flow and power values. All $M_{terr}$ values are within an order of magnitude of the Hardy data.

| | Hardy Heat Flow | $M_{terr}$ Heat Flow |
|---|---|---|
| **Hardy data extent** | 407.7 $Wm^{-2}$ | 378.2 $Wm^{-2}$ |
| **Norris Geyser Basin** | 432.0 $Wm^{-2}$ | 387.5 $Wm^{-2}$ |

| | Hardy Power | $M_{terr}$ Power | Area of Analysis |
|---|---|---|---|
| **Hardy data extent** | 7.0 GW | 6.5 GW | 17,125,200 $m^2$ |
| **Norris Geyser Basin** | 1.5 GW | 1.4 GW | 3,502,800 $m^2$ |

The $M_{terr}$ values calculated for the July 2002 Landsat image were within the same order of magnitude of the heat flow values calculated by Hardy (2005) for the nighttime airborne October 2002 image of Norris Geyser Basin and surrounding area, providing some confirmation that the calculations used in this project were consistent with previous analysis. Solar radiation was not taken into account for either image. While there are fewer solar radiation effects during a nighttime image, there are still accumulated effects from the sun heating the ground the previous day, week, month, and year. The similarities between the daytime and nighttime readings are, therefore, expected, and show that Landsat can be used to calculate terrestrial emittance with comparable results to higher spatial resolution sensors.

The $M_{terr}$ and Hardy heat flow values were larger for Norris Geyser Basin than for the larger extent, as expected. Reducing the study area to the smaller Norris Geyser Basin included less non-geothermal-ground, thus more heat would be emitted per area than in a larger, mostly non-geothermal-ground study area.

The $M_{terr}$ and Hardy values were similar, but the differences are also noteworthy. Since the Landsat image was from July and solar radiation was not taken into account, $M_{terr}$ values were expected to be greater than the Hardy heat flow values, but they were not. This is most likely due to the different data collection and processing methods. The Hardy data were derived directly from raw DNs and temperature calibration data. No atmospheric corrections were needed since the data were collected with a low-elevation

airborne flight. The $M_{terr}$ data, on the other hand, were not calibrated to ground temperature, and needed atmospheric corrections since they were collected from space.

## CHANGE ANALYSIS OF TERRESTRIAL EMITTANCE AND HEAT FLUX

### Methods

GHF is an important aspect in the dynamics of geothermal features. While GHF represents only heat coming from below the Earth's surface, $M_{terr}$ represents all heat emitted from the ground and is composed of GHF as well as direct and indirect solar radiation effects. $GHF_{\alpha}$ attempts to account for solar effects and appears to do so, with caveats, relative to $M_{terr}$ at a YNP-wide scale. $M_{terr}$ has several advantages over $GHF_{\alpha}$ for analyzing change in YNP's geothermal areas. First, a version of $M_{terr}$ has been field verified (Watson *et al*., 2008), while $GHF_{\alpha}$ has had no field verification. Second, the spatial patterns of $GHF_{\alpha}$ are substantially different from those of the less variable $M_{terr}$, including data striping artifacts, and overly high values on north-facing slopes. Lastly, all things being equal, the level of uncertainty in the data increases with each additional processing step, and $M_{terr}$ requires less processing than $GHF_{\alpha}$.

YNP scientists can study changes in $M_{terr}$ values to examine changes in behavior of geothermal features or to monitor for changes in heat flux that might be occurring in response to land management practices within and outside of YNP. New features regularly emerge and active features become inactive. The geothermal areas of YNP must be monitored on a regular basis to be able to assess changes that might occur over days or decades. Two sets of YNP-wide $M_{terr}$ images were processed and are reported here to give one example of how change can be analyzed with data derived from Landsat images. In addition, two sets of YNP-wide $GHF_{\alpha}$ images were processed and provided to Yellowstone National Park's Geology Program. Lastly, further change analyses were performed during Shannon Savage's Ph.D. project. Please see her dissertation for details on those change analyses (Savage, 2009).

#### Image Differencing

The 25 June 2007 and 8 July 2006 TM $M_{terr}$ images (120-m pixel resolution) and the 5 July 2002 and 2 July 2001 ETM+ $M_{terr}$ images (60-m pixel resolution) were differenced by subtracting the earlier date from the later date (*e.g.*, 2007 minus 2006) for every pixel. Large positive output values indicated an increase in $M_{terr}$ from the earlier year to the later year. Large negative output values indicated a decrease in $M_{terr}$ from the earlier year to the later year. Output values near zero indicated little to no change in $M_{terr}$ from the earlier year to the later year.

## Results and Discussion

Image Difference

Approximately 1% of the estimated $M_{terr}$ values increased 10 $Wm^{-2}$ or more from 2001 to 2002, and approximately 1% of the estimated $M_{terr}$ values increased 10 $Wm^{-2}$ or more from 2006 to 2007. The largest increase from 2001 to 2002 was 72.98 $Wm^{-2}$. The largest increase from 2006 to 2007 was 52.98 $Wm^{-2}$. Nearly 82% of the estimated $M_{terr}$ values were within +/- 10 $Wm^{-2}$ from 2001 to 2002, and nearly 95% were within +/- 10 $Wm^{-2}$ from 2006 to 2007. Approximately 16% of the estimated $M_{terr}$ values decreased 10 $Wm^{-2}$ or more from 2001 to 2002, and approximately 4% of the estimated $M_{terr}$ values decreased 10 $Wm^{-2}$ or more from 2006 to 2007. The largest decrease from 2001 to 2002 was -61.79 $Wm^{-2}$. The largest decrease from 2006 to 2007 was -45.65 $Wm^{-2}$.

The majority of the study area did not change dramatically from year to year in the two test cases comparing 2007 to 2006 and 2002 to 2001. This was expected since changes in pixels geothermal activity in most cases are too small to detect with 60-m or 120-m pixel resolution. The areas that did change, however, are explained by fire and snow. The small portion of dramatic increase in $M_{terr}$ from 2001 to 2002 is shown within fire scars from the 2001 fire season (Figure 3). Much of the portion of dramatic decrease in $M_{terr}$ from 2001 to 2002 is near the Pitchstone Plateau area (circled in Figure 3) where there was little to no snow in 2001, but was mostly covered by snow in 2002. Although snow emits heat as it melts, it is likely that the snow in the 2002 image is relatively fresh.

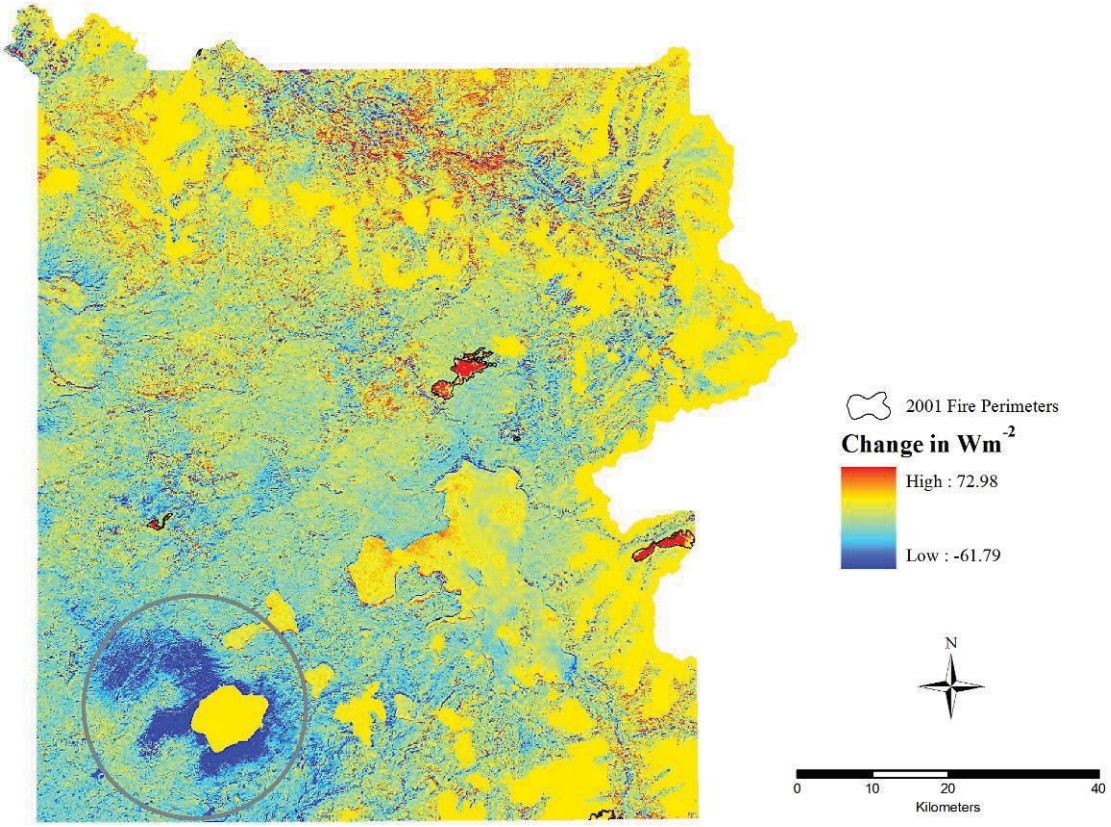


Figure 3: Changes in terrestrial emittance ($M_{terr}$) from 2 July 2001 to 5 July 2002 in Yellowstone National Park (values displayed are in $Wm^{-2}$). The Pitchstone Plateau is circled in grey, demonstrating where there was no snow in 2001 but fresh snow in 2002.

## CLASSIFYING GEOTHERMALLY ACTIVE AREAS

YNP is home to thousands of geothermal features and contains the highest concentration of geysers, hot springs, fumaroles, and mud pots in the world (Waring *et al.*, 1983). Greater than 12,000 individual geothermal features have been identified within the defined geothermal areas in YNP via the Thermal Inventory Project, a multi-year National Park Service-sponsored project with the goal of collecting precise GPS measurements of every geothermal feature in YNP (Spatial Analysis Center, 2008). The defined geothermal areas were delineated from historical data, field observations, data from the Thermal Inventory Project, and heads-up digitizing using one-meter resolution digital orthophoto quarter quadrangles (DOQQs) (Spatial Analysis Center, 2005). The defined geothermal areas include locations that are geothermally inactive, such as Brimstone Basin, near the southeastern arm of Yellowstone Lake. Geothermal activity has never been observed at Brimstone Basin, but the area appears geothermally influenced (Langford, 1972; Nordstrom *et al.*, 2009).

Monitoring geothermal features requires an accepted base map of the area, and while the defined geothermal areas are reasonably accurate, they need to be refined so that inactive geothermal barrens (*e.g.*, Brimstone Basin; ground that is not geothermally

active but appears geothermally influenced) are not included as active geothermal barrens (*e.g.*, white ground with little to no live vegetation and emitted geothermal heat), and active areas that do not appear geothermally influenced are included as geothermally active areas (GAA). For the purposes of this project, a GAA is defined as an area that has hot springs, geysers, fumaroles, and/or mudpots, and/or is emitting geothermal heat.

YNP covers a large area (approximately 890,000 ha), and the current defined geothermal areas cover less than 1% of that area. The size of YNP prevents personnel from being able to visit and monitor all of the geothermal areas each year as part of a monitoring program. Many geothermal areas are in remote backcountry areas not easily accessible on a day-hike, thus requiring multi-day excursions in order to monitor changes at these areas. Not only is this time-consuming, it is expensive, and not practical.

Remote sensing offers a possible alternative to endless field work for monitoring YNP's geothermal areas. Landsat Thematic Mapper (TM) multispectral satellite imagery covers the entirety of YNP and collects information from the visible (0.452–0.518 μm, 0.528–0.609 μm, and 0.626–0.693 μm), near infrared (NIR; 0.776–0.904 μm), middle infrared (MIR; 1.57–1.78 μm, and 2.1–2.35 μm), and thermal infrared (TIR; 10.45–12.42 μm) portions of the electromagnetic spectrum (EMS) (Chander *et al.*, 2009). All but the TIR band have 30-m spatial resolution (30 m on a pixel side, or 900 $m^2$), while the TIR band has 120-m spatial resolution (14,400 $m^2$). The current TM sensor on Landsat 5 has been in orbit since 1984 and a new TM sensor with a similar spectral range, including TIR, is expected to be sent into orbit on Landsat 8 in 2012 (NASA, 2009). Landsat TM images are collected over YNP every 16 days, allowing annual or seasonal classification of the GAAs.

One of the primary applications of remote sensing, and Landsat data in particular, is classifications of features on the landscape, often land cover/vegetation, but also minerals, water, human impacts, and geothermally influenced ground. Classifications of landscape features produce maps that can be compared over time to assess change. Thermal and terrestrial emittance anomalies, for example, have been identified successfully with Landsat data. Three Landsat TM scenes recorded high-temperature thermal anomalies, such as vertical ash eruptions and an active basaltic lava flow, from 1986 to 1988 over Santiaguito Dome and Pacaya Volcano in Guatemala (Andres and Rose, 1995). A method of quantifying the intensity of surficial geothermal activity in YNP was developed with 2000 Landsat ETM+ imagery (Watson *et al.*, 2008). This method utilized thermal radiance data to create a map of terrestrial emittance anomalies, proxies for geothermal heat flux (GHF).

Decision tree classification methods are recent additions to the image classification arsenal that allow analysts to utilize original imagery along with ancillary data without requiring expert knowledge to conduct highly accurate image classifications (Lawrence and Wright, 2001; Lawrence *et al.*, 2004). Random Forest (RF) is a decision tree classification method that grows hundreds of decision trees, where each tree is grown using a different bootstrapped (resampled with replacement) random subset of training data, and each split within each tree is based on a different random subset of predictor variables (Breiman, 2001; Lawrence *et al.*, 2006). The "forest" of trees then votes to assign a class to each input data point (Breiman, 2001; Prasad *et al.*, 2006). RF classifications have been shown to have accuracy rates as good as or better than any other classification method used for remote sensing, while being less sensitive to noise and

uneven classes in training sets (Pal, 2005; Gislason *et al.*, 2006; Lawrence *et al.*, 2006). Internal accuracies of RF datasets are calculated with out-of-bag (OOB) samples, that is, those training data excluded from the bootstrapped random subsets, potentially reducing the need for independent accuracy assessments (Liaw and Wiener, 2002; Gislason *et al.*, 2006; Lawrence *et al.*, 2006; Prasad *et al.*, 2006). When the reference data are biased, however, the OOB estimation might not be reliable (Lawrence *et al.*, 2006). RF has been applied to Landsat ETM+ data to classify agricultural land cover in Littleport, Cambridgeshire, UK (Pal, 2005) and to Landsat Multispectral Scanner (MSS) data to classify forest types in a mountainous area in Colorado (Gislason *et al.*, 2006) with 88% and 83% accuracy respectively. RF can handle high dimensional data, easily accommodates ancillary data, avoids overfitting, does not make assumptions about the distribution of data, and is particularly well suited to predictive mapping. It is, however, somewhat of a "black box" method since the resulting statistical model consisting of a forest of decision trees is not subject to easy interpretation of the relationship of predictor and response variables other than the relative importance of predictors (Friedl and Brodley, 1997; Liaw and Wiener, 2002; Gislason *et al.*, 2006; Prasad *et al.*, 2006). RF should be a reasonable classification technique for determining whether Landsat imagery is able to map the distribution of GAA in the defined geothermal areas of YNP since GAA have not yet been classified in YNP and there is no reported study of any other statistical method substantially outperforming RF for classification purposes.

The constrained energy minimization (CEM) method is a target detection classification technique, developed primarily for hyperspectral data and often used to identify minerals or very rare targets. The method only requires prior knowledge of the distinct target of interest and eliminates unidentified spectral signature sources and suppresses noise in the data (Du *et al.*, 2003). The distribution of mine tailings in Coeur d'Alene River Valley, Idaho in 1993 was mapped using CEM (Farrand and Harsanyi, 1997). Twelve false-alarm pixels were found out of 484 pixels determined to be rich in ferruginous sediments. Although designed for hyperspectral imagery, an empirical study showed that CEM can be used with SPOT imagery (multispectral imagery with spectral similarities to Landsat imagery) and correlated derivatives of these data for classification of individual targets (Chang *et al.*, 2000). Since GAA are rare targets of interest outside defined geothermal areas at YNP, and non-GAA information is highly variable and difficult to collect due to the size of and diversity of landcover types within YNP, the CEM target detection algorithm is an appropriate tool for classifications of GAA outside the defined geothermal areas.

The main purpose of this portion of the project was to assess the ability of Landsat TM data combined with RandomForest and target detection classifiers to classify GAA accurately. An effective method would enable the classification of GAA in and outside YNP that could be applied to additional Landsat images for use in monitoring and change analysis. A successful classification method would provide scientists with information on where to check for new geothermal areas in YNP and where to focus on ground-work or aerial flights to best assess change. This strategy would reduce the amount of time-consuming and expensive field monitoring or aerial image acquisition, especially in the backcountry.

## Methods

Image preprocessing

Five summer TM scenes were utilized for classification: 2007, 2006, 2005, 1989, and 1986. The first three scenes were chosen because they were the most recent, complete, and mostly cloud-free (less than 5%) consecutive years of imagery. The scene for 1989 was chosen because it should indicate major changes after the 1988 fire season. The 1985 scene was chosen because it was the only summer scene available prior to the 1988 fire season that had thermal infrared information.

Several ancillary data sets were required for analysis. A 30-m digital elevation model (DEM), digital spatial polygon data of the defined geothermal areas, and digital spatial point data of the Thermal Inventory Project data were provided by YNP. The defined geothermal area polygons include nearly all known geothermally active areas in YNP as well as some inactive areas. These data were provided as the starting point for classification with the aim of refining the boundaries so inactive areas would not be included. The Thermal Inventory Project includes greater than 12,000 precise (sub-meter accuracy) GPS locations of individual geothermal features. If a GPS point was collected at the edge of a feature, the distance and azimuth to the center was estimated and the point was moved to that center location. The Thermal Inventory Project points represent the most complete collection of all geothermal features within YNP, but do not include areas of hot ground with no geothermal features. These ancillary data were utilized as reference data for the classification process.

Slope and aspect were derived from the DEM, with slope in degrees and aspect as categorical data with 9 categories (N, S, E, W, NE, SE, SW, NW, and flat). These topographic data were utilized in subsequent calculations (see below) as well as in the classification process as additional predictor variables. Slope and aspect by themselves would not be able to classify GAA, but their interactions with other predictor variables might help the classification process.

The TM images and all ancillary data were clipped to the full study area. Clouds and cloud shadows were masked by on-screen digitizing, and elevations greater than 2,700 m were masked to remove snow from the study area as no records of geothermal features have been found for these areas and deep snow conceals thermal infrared signatures (if they exist). The methods described above (equations 1 through 5) to calculate $M_{terr}$ from the red, NIR, and TIR Landsat bands were followed. SR and albedo were calculated as described above (equations 6 and 7) and $GHF_{\alpha}$ was calculated from $M_{terr}$, SR, and albedo following the methods above (equation 8). NDVI, albedo, SR, and $GHF_{\alpha}$ were included in the classification process along with the original reflective spectral bands and $M_{terr}$ in order to provide the classification algorithm with many possible predictors and predictor interactions that might improve the final outcome.

Transformed datasets were derived from the original Landsat bands to potentially improve the classification process by adding additional predictors that can detect diverse landscape features. Principal components analysis (PCA) reduces the amount of data to be analyzed and accounts for the most variance in the original image (Singh, 1989). Correlations between the components and the input bands can be calculated and each PC

can be interpreted as representing certain combinations of Landsat bands and/or features on the ground (Jensen, 2005). A standardized PCA was performed on the original reflective bands and $M_{terr}$ resulting in 7 new components where the majority of the original variance can be found in the first 3 components.

A tasseled cap (TC) transformation was performed on the six reflective bands, resulting in 3 additional transformed components for the classification process. This is a physically based identification process similar to PCA in that it reduces the amount of information to be analyzed into the first three components. These components represent brightness (TCB) (soil brightness or total reflectance), greenness (TCG) (relative amounts of leafy green vegetation), and wetness (TCW) (soil moisture) (Crist and Cicone, 1984).

The six original reflective Landsat bands and 17 derived and ancillary data components were stacked to create one 23-component image for the study area (Table 8). This 23-component image was clipped to the defined geothermal area boundaries for use in an initial classification. The final images used in the classification processes had 30-m spatial resolution.

Table 8: Components used in the RandomForest and constrained energy minimization classification processes. Components 1 through 5 and 7 were original Landsat bands. Components 6, 8 through 18, 20, and 23 were derived from the original Landsat bands. Components 19, 21, and 22 were derived from topographic information.

| # | Component Name | # | Component Name |
|---|---|---|---|
| 1 | Band 1 – Blue | 13 | Principal Component 6 (PCA6) |
| 2 | Band 2 – Green | 14 | Principal Component 7 (PCA7) |
| 3 | Band 3 – Red | 15 | Tasseled Cap Brightness (TCB) |
| 4 | Band 4 – Near Infrared (NIR) | 16 | Tasseled Cap Greenness (TCG) |
| 5 | Band 5 – Middle Infrared (MIR1) | 17 | Tasseled Cap Wetness (TCW) |
| 6 | Terrestrial Emmitance ($M_{terr}$) | 18 | Normalized Difference Vegetation Index (NDVI) |
| 7 | Band 7 – Middle Infrared (MIR2) | 19 | Potential Annual Direct Incident Solar Radiation (SR) |
| 8 | Principal Component 1 (PCA1) | 20 | Albedo |
| 9 | Principal Component 2 (PCA2) | 21 | Aspect |
| 10 | Principal Component 3 (PCA3) | 22 | Slope in degrees |
| 11 | Principal Component 4 (PCA4) | 23 | Estimated Geothermal Heat Flux ($GHF_{\alpha}$) |
| 12 | Principal Component 5 (PCA5) | | |

RandomForest Classification Procedures

The classifications developed for this study had two classes: (1) GAA – anywhere that was geothermally active, and (2) non-GAA – anywhere that was not geothermally active. The RF classification was designed to refine the currently defined geothermal areas by distinguishing the non-GAA contained within the defined geothermal areas.

The GAA reference data for the RF classification processes were collected as a random selection of Thermal Inventory Project data points (Table 9) (Spatial Analysis Center, 2008). The Thermal Inventory Project began in 1998 and was completed in 2008. Over 12,000 GPS-located points were collected along with pH, conductivity, temperature, and a description of each geothermal feature. The data from the Thermal

Inventory Project were the most accurate available data for representing GAA, since the points are locations of active geothermal features and were collected with a sub-meter precision GPS unit. An inherent bias to these data is that the Thermal Inventory Project focused on *features*, not *areas*, so active geothermal barrens were not identified in this dataset, and therefore are absent from the reference data.

Non-GAA reference data were impossible to collect in the field due to the lack of time, money, and permission to place 10,000 temperature loggers throughout the defined geothermal areas that would be necessary to identify a large number of 900 $m^2$ locations known to not be emitting geothermal energy. Thus, the non-GAA training and validation data were collected with digital spatial data. The Thermal Inventory Project points were buffered with a 60-m radius (in order to exclude the 30-m pixel in which the points reside and the 8 surrounding 30-m pixels) and random points were generated in the areas outside those buffer zones but within the defined geothermal areas. Over 3,500 reference points were collected for this study (Table 9).

Table 9: Geothermally active area (GAA) and non-geothermally active area (non-GAA) reference data used in the RandomForest classification of the defined geothermal areas.

| GAA Training | Non-GAA Training | GAA Validation | Non-GAA Validation |
|---|---|---|---|
| 1,366 | 1,914 | 300 | 300 |

The five classifications were performed with the ModelMap package within R statistical software (Freeman and Frescino, 2009). ModelMap contains the randomForest function and in addition to producing an out-of-bag (OOB) error estimate and a graph of predictor importance, creates a text file that can be converted to a raster image. The training data set was used to create five different RF classifications. All 23 image components (listed in Table 8) were utilized as predictor variables for all five classifications. Each output text file was converted to a classified raster image based on the probability threshold of 0.5. In other words, any value greater than 0.5 was classified as GAA, while the rest were classified as non-GAA. The classified raster image was subsequently converted to polygon data for use as training data for classifications of study area locations outside the defined geothermal areas.

Five error matrices based on data withheld from the reference data were constructed to calculate the overall and class accuracies of each classification. Class accuracies are shown with user's accuracy (errors of commission) and producer's accuracy (errors of omission) (Congalton, 2001). Kappa statistics were calculated for each classification method. The Kappa statistic measures how much better (or worse) the classification is from a randomly generated classification and is more conservative than overall accuracy (Congalton and Green, 1999; Congalton, 2001). Kappa values range from -1 to 1 and the closer the value to 1, the more accurate the classification.

## Target Detection Classification Procedures

The target detection classification was designed to classify GAA outside of the defined geothermal areas, perhaps identifying previously unknown areas of geothermal activity.

The training data for this classification were taken from the final 2007 RF classified shapefile of the defined geothermal areas because, despite its limitations, it was the only dataset available for this purpose. All GAA polygons greater than 900 $m^2$ were selected for use as training for the target detection classification of the full study area. The CEM algorithm was utilized for target detection. CEM outputs an image containing continuous values on an arbitrary scale with higher values indicating pixels more similar to the training data. The resulting continuous data were converted to a binary classified image based on a threshold set by the CEM process (a value of 2), then merged with the defined geothermal area classifications to create a final classified image of the entire study area.

A complete error matrix could not be constructed since there were no non-GAA validation data available outside of the defined geothermal areas. 174 Thermal Inventory points were identified outside of the defined geothermal areas, however, and were used to evaluate the ability of CEM to identify GAA for the 2007 classification.

### Creation of Change Maps

Four change images were created by comparing one classified image to the previous date classified image (2007 vs. 2006, 2006 vs.2005, 2005 vs.1989, 1989 vs. 1986). The resulting change images had four classes: (1) GAA in both years, (2) non-GAA in both years, (3) non-GAA in year 1 (*e.g.*, 2006) changed to GAA in year 2 (*e.g.*, 2007), and (4) GAA in year 1 changed to non-GAA in year 2. The results of these processes are not discussed further but have been made available to the Yellowstone National Park Geology Program.

## Results and Discussion

### RandomForest Classification of the Defined Geothermal Areas

Overall accuracies of the five classifications are listed in Table 10. The semi-independent overall accuracies were consistently higher than the RandomForest OOB accuracies, however, none of the accuracies reached 70%. This substantial difference was likely because of the variability within the reference data as the randomly generated points covered a diverse and highly variable landscape. All of the Kappa statistics were greater than 0.3 which indicates that although the overall accuracies were not high, all of the classifications were better than random. Only the 2007 classification will be discussed further as all of the classifications were poor, none of the classifications were significantly better than the others, and the 2007 classification represents the most recent data in these analyses.

Table 10: Random forest out-of-bag (OOB) accuracies, semi-independent overall accuracies, and Kappa statistics for the five random forest classifications of the defined geothermal areas.

| Method | Random Forest OOB Accuracy | Semi-independent Overall Accuracy | Semi-independent Kappa Statistic |
|---|---|---|---|
| 25 June 2007 | 57.56% | 66.17% | 0.32 |
| 8 July 2006 | 56.12% | 67.83% | 0.36 |
| 21 July 2005 | 56.22% | 68.17% | 0.36 |
| 2 August 1989 | 55.46% | 65.50% | 0.31 |
| 17 July 1986 | 57.16% | 68.50% | 0.37 |

Only 16.67% of the defined geothermal areas were classified as GAA in 2007. Class accuracies ranged from 60.00% to 72.33% (Table 11). Variable importance plots illustrate each predictor variable's contribution to the mean decrease in the OOB error rate (Sesnie *et al.*, 2008), where those variables closer to the top of the plot are the most influential to the accuracy of the classification and those closer to the bottom are the least important to the accuracy. The RF variable importance plot for 2007 indicated that 11 of the top 12 variables were dominated by reflective spectral information (Figure 4). All of the original Landsat reflective bands were in the top 12. PCA3 was interpreted from the eigenvector weightings as representing primarily the NIR band with additional influence from the MIR2 band (Table 12), PCA1 was interpreted as representing all of the reflective spectral bands almost equally (in other words, the pixel brightness), and TCW, TCB, and albedo were derived from the reflective bands. These 5 components were in the top 12. The only non-reflective spectral information in the top 12 was topographic information (slope). Three of the 4 geothermal components, PCA4 (interpreted as representing primarily $M_{terr}$), $M_{terr}$, and $GHF_{\alpha}$, on the other hand, were among the bottom 6 variables and were of lesser importance than the majority of the included predictor variables (Figure 4). Aspect, a topographic variable, was one of the least influential predictor variables along with PCA5 and PCA6 (interpreted as representing primarily the visible bands).

Table 11: Semi-independent error matrix for the 25 June 2007 classification of the defined geothermal areas. Class accuracies are represented by user's accuracy (errors of commission) and producer's accuracy (errors of omission). The Kappa statistic is a measure of classification accuracy that is more conservative than overall accuracy.

| | | Reference Data | | |
|---|---|---|---|---|
| | Class | GAA | Non-GAA | User's Accuracy |
| **Classified Data** | GAA | 180 | 83 | 68.44% |
| | Non-GAA | 120 | 217 | 64.39% |
| | Producer's Accuracy | 60.00% | 72.33% | |
| | | Overall Accuracy = 66.17% | | |
| | | Kappa = 0.3233 | | |

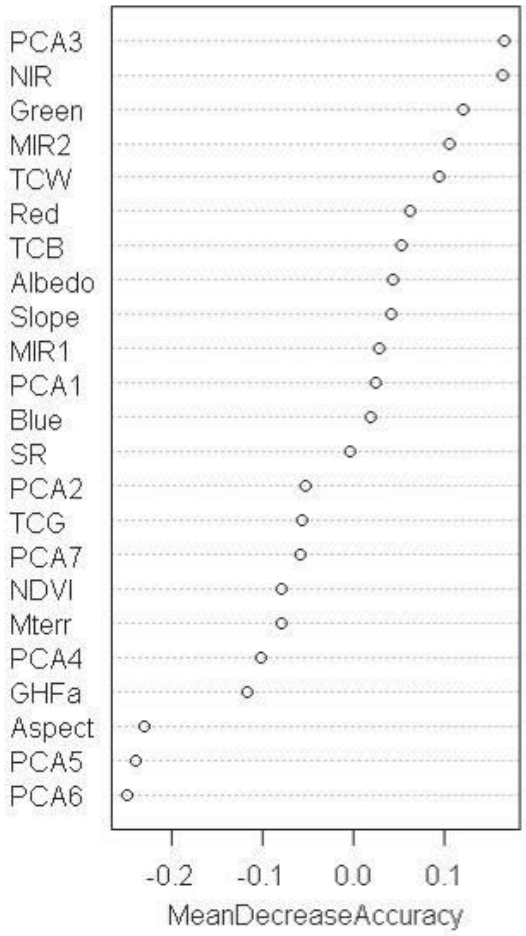


Figure 4: Predictor variable importance plot for the 25 June 2007 classification of the defined geothermal areas. Variables at the top of the plot were more influential to the accuracy of the classification than variables at the bottom.

Table 12: Principal component eigenvectors that show the weightings of each input band on each principal component (PCA) for 25 June 2007. PCA1 is highly weighted in the visible bands and the NIR and MIR bands. PCA2 and PCA4 are weighted high in $M_{terr}$ and MIR. PCA3 is weighted mostly in the NIR. PCA5 is highly weighted in blue and green, and PCA6 is highly weighted in green and red. PCA7 is mostly weighted in MIR, with some influence from red.

| | PCA1 | PCA2 | PCA3 | PCA4 | PCA5 | PCA6 | PCA7 |
|---|---|---|---|---|---|---|---|
| **Blue** | 0.443 | -0.307 | 0.191 | -0.093 | -0.791 | 0.190 | 0.052 |
| **Green** | 0.450 | -0.302 | 0.116 | -0.125 | 0.560 | 0.542 | 0.263 |
| **Red** | 0.463 | -0.260 | 0.205 | -0.108 | 0.242 | -0.621 | -0.470 |
| **NIR** | 0.372 | -0.011 | -0.900 | 0.037 | -0.036 | -0.161 | 0.170 |
| **MIR** | 0.326 | 0.496 | -0.035 | 0.391 | -0.007 | 0.394 | -0.582 |
| **$M_{terr}$** | 0.148 | 0.550 | 0.021 | -0.821 | -0.026 | 0.013 | 0.008 |
| **MIR** | 0.345 | 0.445 | 0.320 | 0.369 | 0.016 | -0.321 | 0.583 |

The 2007 classification identified only a small portion of the largest geothermal feature in YNP, Grand Prismatic in Midway Geyser Basin, as GAA, but classified all of Excelsior Geyser, a large feature adjacent to Grand Prismatic, as GAA (Figure 5). Much of the geothermal barrens were classified as GAA. The Firehole River was generally classified as non-GAA, with a few exceptions that might be related to pixel geometric registration. For example, a feature with high $M_{terr}$ might be located at the edge of a pixel, but with geometric registration errors might fall in a different pixel and be classified incorrectly.

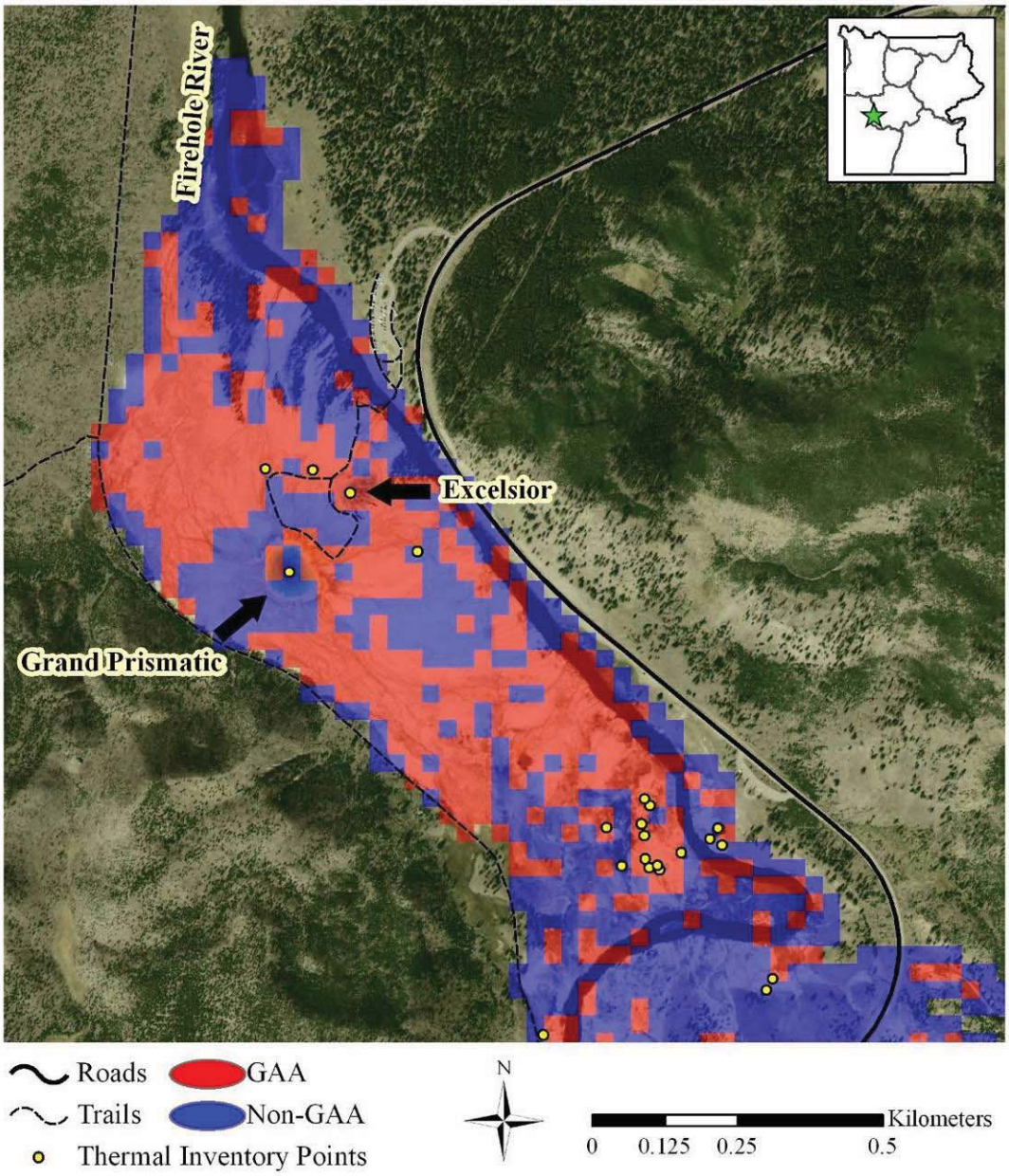


Figure 5: 25 June 2007 classified map of Grand Prismatic Spring and Excelsior Geyser in Midway Geyser Basin. Thermal Inventory Project points are displayed over the classification, with National Agriculture Imagery Program (NAIP) imagery in the background. Excelsior Geyser was successfully classified as a geothermally active area (GAA), but much of Grand Prismatic Spring was misclassified as a non-geothermally active area (non-GAA). Geothermal barrens were classified as both GAA and non-GAA throughout the area. The majority of the Firehole River was classified as non-GAA.

Target Detection Classification of the 30-km-Buffered Study Area

Continuous output values for the target detection classification ranged from -1 to 2. Pixels with a value of 2 were classified as GAA. All other pixels were classified as

non-GAA. The percentage of known GAA locations outside the defined geothermal areas (located during the Thermal Inventory Project, but not found within the boundaries of the defined geothermal areas) that the 2007 classification was able to detect was 6.3% (11 of 174 reference points were classified correctly).

Only 0.07% of the study area was classified as GAA in 2007. The map of the 2007 classification of the area near La Duke Hot Springs in Montana (just outside of YNP) showed very little area classified as GAA (Figure 6), however, two of the three geothermal features at La Duke Hot Springs fall on one pixel that was successfully classified as GAA.


Figure 6: 2007 classification map of a portion of Corwin Springs, Montana Known Geothermal Resource Area. Thermal Inventory Project points are displayed over the classification, with National Agriculture Imagery Program (NAIP) imagery in the background.

There are recognized potential threats to geothermal features in YNP, including possible geothermal development in the Corwin Springs, Montana, and Island Park, Idaho KGRAs (Sorey, 1991; Heasler *et al*., 2004), that make accurate maps of GAA outside of YNP increasingly important. The classifications created for this project were better than random, but accuracies were low. The area where the reference data were randomly generated was a highly variable landscape, and it is likely that many of the non-GAA training and validation points were in areas of high Mterr that had no geothermal activity, such as inactive geothermal barrens, and were thus confused with areas of high Mterr that *were* geothermally active, such as active geothermal barrens. This classification method was not robust enough, or more likely the geothermal landscape was too variable, to produce an acceptably accurate GAA classification.

Inherent geothermal characteristics such as biological and steam spectral signatures might affect classification of GAA in different ways. Grand Prismatic Spring, as seen in Figure 7, is brilliantly colorful and very different from the surrounding landscape. The colors are produced by different microbes within the hot water that survive (and thrive) in various chemical and temperature gradients (Brock, 1967; Brock, 1978), each of which are expressed as distinctive bands of color: Archaea in the blue/green center of the pool at the hottest temperature (75° C), the cyanobacterium, *Synechococcus*, at the pool's yellow edge (59° C to 70° C), a rusty-brown microbial mat produced by the cyanobacterium, *Phormidium*, outside the pool between 30° C and 59° C, and furthest away from the heat source in the center of the pool, the cyanobacterium, *Calthrix*, produces a dull-brown microbial mat (Kaplan and Bartley, 2000). Various photopigments are found in the different microbes in the geothermal features of YNP, including carotenoids (absoption peaks at 0.45 – 0.55 μm), chlorophyll *a* (absorption peaks at 0.43 μm and 0.67 μm), phycocyanin (absorption peak at 0.62 μm), and bacterioclorophylls *a* and *c* (absorption peaks at 0.80 – 0.90 μm and 0.73 μm, respectively) (Jorgensen and DeMarais, 1988; Ward *et al*., 1989). The absorption peaks for these photopigments might be related to the highly important predictor variables that are dominated by reflective spectral information and thus detectable with the reflective Landsat bands, however, the Landsat bandwidths are most likely too wide to distinguish different microbial communities from one another and from vegetation. Grand Prismatic is also often covered by a thick layer of steam as the average temperature in the middle of the pool is approximately 75° C (Kaplan and Bartley, 2000). Excelsior Geyser, whose average temperature is approximately 93° C (Thompson and Yadav, 1979; Brody and Tomkiewicz, 2002) is rarely seen without steam (behind Grand Prismatic in Figure 7). Steam has a distinctive spectral signature and is lower in temperature than the water that is producing it, likely obscuring geothermal signatures and causing confusion in the GAA classification with Landsat data.

Figure 7: Grand Prismatic Spring in Midway Geyser Basin, with Excelsior Geyser steaming in the background, demonstrating the extraordinary variability of geothermal areas in Yellowstone National Park. Photograph by Shannon Savage, taken on 22 June 2006.

Good training and validation data are fundamental to the success of remotely sensed classifications. The reference data used for this study were not ideal due to both the lack of resources and the lack of ability to collect proper information at a 30-m spatial resolution. Given much additional time and money (and permission from YNP), appropriate sample data could possibly be collected by placing thousands of temperature probes (based on common sampling approaches) across the defined geothermal areas at the same time the Landsat satellite collects data over YNP and interpolating the temperatures to fit 30-m pixels. By using ground probes, the influences of solar radiation and albedo can be better modeled and minimized, allowing the Landsat TIR band to be properly calibrated to ground temperature, and GAA could then be distinguished from non-GAA. Unfortunately, placing that many probes to get adequate training data is simply not practical, especially for an ongoing monitoring program.

## LESSONS LEARNED

Landsat data include information from the thermal infrared portion of the electromagnetic spectrum and were assessed for their ability to successfully map geothermally active areas and geothermal heat flux in geothermally influenced Yellowstone National Park (YNP). Landsat thermal infrared data from 1978 to present and into the foreseeable future are freely available from the United States Geological Survey, effectively providing an opportunity to study geothermal ground over the entirety of YNP for 30 years and beyond. An accurate, inexpensive, and reproducible method for mapping geothermal ground at YNP might be possible with Landsat data, but the results of this study indicate there are inherent limitations to Landsat data and issues with the nature of geothermal features that, combined, make accurately mapping geothermally active areas and geothermal heat flux difficult with the methods tested herein.

Terrestrial emittance represents all heat being emitted from the ground, including effects from direct and indirect solar radiation, as opposed to geothermal heat flux, which represents only geothermal heat emitted from below ground. The Landsat thermal infrared band, with assistance from the red and near infrared bands, was able to detect terrestrial emittance, but without rigorous solar radiation field data, was unable to accurately estimate geothermal heat flux. Incorporating estimated emissivity on a pixel-by-pixel basis rather than as an average over the entire image (as in the Watson *et al*., 2008 method) produced locally precise terrestrial emittance estimates by accounting for differences in emissivity due to varying amount of vegetation in each pixel. By deriving emissivity values from the finer spatial resolution reflective Landsat bands, more detail was incorporated into the $M_{terr}$ values than if the thermal band had been used alone. Additionally, incorporating estimates of potential direct incident solar radiation and surface albedo into the terrestrial emittance calculation produced a moderately acceptable estimation of geothermal heat flux covering all of YNP that, although unable to account for all inherent variability such as conductance of non-geothermal heat through the soil or diurnal, seasonal, or annual temperature oscillations, might help YNP scientists identify areas of interest for further study.

The random forest classification method was unable to produce an acceptably accurate classification of geothermally active areas in the currently defined geothermal areas at YNP, however, the classifications developed were better than random. The inherent limitations of Landsat data, the inherent nature of geothermal features and areas, and acknowledged biases within the reference data used were probable reasons for the unacceptable classifications. Appropriate reference data, however, are impractical to collect, and even with excellent reference data, the technological limitations of Landsat imagery might preclude producing highly accurate geothermally active area classifications. The target detection training data were based on the poor random forest classifications and did not perform well. Given a highly accurate classification of the defined geothermal areas, however, the constrained energy minimization target detection classification method might prove to be the appropriate tool to classify geothermally active areas in and around YNP.

## OBSERVATIONS AND RECOMMENDATIONS

While Landsat data are excellent resources for mapping many landscape features, the inherent technological limitations of Landsat data, in particular the thermal infrared band, impede accurate mapping of geothermal heat flux and geothermally active areas. The Landsat thermal infrared band has relatively low spatial resolution and is not sensitive enough to variations in terrestrial emittance that might be caused by individual geothermal features or areas of geothermally active ground, thus geothermal heat flux and geothermally active areas are difficult to detect accurately even when Landsat thermal infrared data are combined with the reflective data. Geothermal barrens, for example, include both geothermally active and inactive areas that have very similar reflective properties, resulting in confusion when attempting to classify geothermally active areas or improve estimations of geothermal heat flux with a combination of thermal infrared and reflective bands. Until a high spatial resolution thermal infrared scanner is developed that images the entirety of YNP on a regular basis, Landsat data will remain the only available thermal infrared data for historical and continuous monitoring of geothermal areas in YNP. The results of this study suggest that future studies of geothermal areas at YNP with Landsat data will be unsuccessful until Landsat data can be more accurately calibrated to geothermal heat flux and solar effects.

## REFERENCES CITED

Andres, R.J. and W.I. Rose. 1995. Description of thermal anomalies on two active Guatemalan volcanoes using Landsat Thematic Mapper imagery. *Photogrammetric Engineering & Remote Sensing*, 61(6): 775-782.

Boomer, S.M., D.P. Lodge, B.E. Dutton, and B. Pierson. 2002. Molecular characterization of novel red green nonsulfur bacteria from five distinct hot springs communities in Yellowstone National Park. *Applied and Environmental Microbiology*, 68(1): 346-355.

Bourgeois, O., O. Dauteuil, and B. Van Vliet-Lanoe. 2000. Geothermal control on flow patterns in the last glacial maximum ice sheet of Iceland. *Earth Surface Processes and Landforms*, 25(1): 59-76.

Breiman, L. 2001. Random forests. *Machine Learning*, 45(1): 5-32.

Brock, T.D. 1967. Life at high temperatures. *Science*, 158(3804): 1012-1019.

Brock, T.D. 1978. Thermophilic Microorganisms and Life at High Temperatures. Springer-Verlag, New York, New York, 465 pp.

Brody, M. and W. Tomkiewicz. 2002. Park visitors' understandings, values and beliefs related to their experience at Midway Geyser Basin, Yellowstone National Park, USA. *International Journal of Science Education*, 24(11): 1119-1141.

Brunsell, N.A. and R.R. Gillies. 2002. Incorporating surface emissivity into a thermal atmospheric correction. *Photogrammetric Engineering & Remote Sensing*, 68(12): 1263-1269.

Carranza, E.J.M. and M. Hale. 2002. Mineral imaging with Landsat Thematic Mapper data for hydrothermal alteration mapping in heavily vegetated terrane. *International Journal of Remote Sensing*, 23(22): 4827-4852.

Chander, G., B.L. Markham, and D.L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5): 893-903.

Chang, C.-I., J.-M. Liu, B.-C. Chieu, H. Ren, C.-M. Wang, C.-S. Lo, P.-C. Chung, C.-W. Yang, and D.-J. Ma. 2000. Generalized constrained energy minimization approach to subpixel target detection for multispectral imagery. *Optical Engineering*, 39(5): 1275-1281.

Chavez, P.S.J. 1996. Image-based atmospheric corrections - revisited and revised. *Photogrammetric Engineering & Remote Sensing*, 62(9): 1025-1036.

Congalton, R.G. 2001. Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire*, 10(4): 321-328.

Congalton, R.G. and K. Green. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers, Boca Raton, Florida, 137 pp.

Coolbaugh, M.F., C. Kratt, A. Fallacaro, W.M. Calvin, and J.V. Taranik. 2007. Detection of geothermal anomalies using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) thermal infrared images at Bradys Hot Springs, Nevada, USA. *Remote Sensing of Environment*, 106(3): 350-359.

Crist, E.P. and R.C. Cicone. 1984. A physically-based transformation of Thematic Mapper data - the TM tasseled cap. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22(23): 256-263.

Custer, S.G., D.E. Michels, W. Sill, J.L. Sonderegger, W. Weight, and W. Woessner. 1993. Recommended Boundary for Controlled Groundwater Area in Montana Near Yellowstone Park, U.S. Department of the Interior National Park Service, Fort Collins, CO.

Daneshfar, B., A. Desrochers, and P. Budkewitsch. 2006. Mineral-potential mapping for MVT deposits with limited data sets using Landsat data and geological evidence in the Borden Basin, Northern Baffin Island, Nanuvut, Canada. *Natural Resources Research*, 15(3): 129-149.

Dogan, H.M. 2008. Applications of remote sensing and geographic information systems to assess ferrous minerals and iron oxide of Tokat province in Turkey. *International Journal of Remote Sensing*, 29(1): 221-233.

Du, Q., H. Ren, and C.-I. Chang. 2003. A comparative study for orthogonal subspace projection and constrained energy minimization. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6): 1525-1529.

Elachi, C. 1987. Introduction to the Physics and Techniques of Remote Sensing. John Wiley & Sons, New York, NY, 413 pp.

Farrand, W.H. and J.C. Harsanyi. 1997. Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique. *Remote Sensing of Environment*, 59(1): 64-76.Freeman and Frescino, 2009

Flynn, L.P., A.J.L. Harris, and R. Wright. 2001. Improved identification of volcanic features using Landsat 7 ETM+. *Remote Sensing of Environment*, 78(1-2): 180-193.

Fournier, R.O., D.E. White, and A.H. Truesdell. 1975. Convective heat flow in Yellowstone National Park, Proceedings of the Second United Nations Symposium on the Development and Use of Geothermal Resources, pp. 731-739.

Friedl, M.A. and C.E. Brodley. 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61(3): 399-409.

Friedman, I. and D.R. Norton. 2007. Is Yellowstone losing its steam? - Chloride flux out of Yellowstone National Park. Chapter I in Professional Paper 1717, U.S. Geological Survey, pp. 272-297.

Gislason, P.O., J.A. Benediktsson, and J.R. Sveinsson. 2006. Random forests for land cover classification. *Pattern Recognition Letters*, 27(4): 294-300.

Gruber, S., M. Hoelzle, and W. Haeberli. 2004. Rock-wall temperatures in the Alps: Modeling their topographic distribution and regional differences. *Permafrost and Periglacial Processes*, 15(3): 299-307.

Hardy, C.C. 2005. "Characterizing thermal features from multi-spectral remote sensing data using dynamic calibration procedures", University of Montana, Missoula, Montana, 153 pp.

Heasler, H., C. Jaworowski, and D. Susong. 2004. A geothermal monitoring plan for Yellowstone National Park, Yellowstone Center for Resources, Yellowstone National Park, Wyoming, 24pp.

Jensen, J.R. 2005. Introductory Digital Image Processing - A Remote Sensing Perspective. Prentice Hall, Upper Saddle River, NJ, 526 pp.

Jorgensen, B.B. and D.J. Des Marais. 1988. Optical properties of benthic photosynthetic communities: Fiber-optic studies of cyanobacterial mats. *Limnology and Oceonography*, 33(1): 99-113.

Kaneko, T. and M.J. Wooster. 1999. Landsat infrared analysis of fumarole activity at Unzen Volcano: time-series comparison with gas and magma fluxes. *Journal of Volcanology and Geothermal Research*, 89(1-4): 57-64.

Kaplan, I.R. and J.K. Bartley. 2000. Global Biogeochemical Cycles: Carbon, Sulfur, and Nitrogen. In: W.G. Ernst (Editor), Earth Systems: Processes and Issues. Cambridge University Press, London, UK, 566 pp.

Kohl, T. 1999. Transient thermal effects below complex topographies. *Tectonophysics*, 306(3-4): 311-324.

Langford, N.P. 1972. The Discovery of Yellowstone Park: Journal of the Washburn Expedition to the Yellowstone and Firehole Rivers in the Year 1870. University of Nebraska Press, Lincoln, Nebraska, 125 pp.

Lawrence, R., A. Bunn, S. Powell, and M. Zambon. 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment*, 90(3): 331-336.

Lawrence, R.L., S.D. Wood, and R.L. Sheley. 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment*, 100(3): 356-362.

Lawrence, R.L. and A. Wright. 2001. Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric Engineering & Remote Sensing*, 67(10): 1137-1142.

Liang, S. 2000. Narrowband to broadband conversions of land surface albedo: I Algorithms. *Remote Sensing of Environment*, 76(2): 213-238.

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. *R News*, 2(3): 18-22.

Long, C.L., D.B. Hoover, and C.L. Tippens. 1976. Audio-magnetotelluric station location map Island Park Known Geothermal Resource Area, Idaho. Open-File Report No. 76-700E, U.S. Geological Survey, 19 pp.

McCune, B. and D. Keon. 2002. Equations for potential annual direct incident radiation and heat load. *Journal of Vegetation Science*, 13(4): 603-606.

NASA. 2008. ASTER Spectral Library. Last accessed on 3 March 2009. http://speclib.jpl.nasa.gov/.

NASA. 2009. Landsat 7 Science Data Users Handbook. Last accessed on 6 April 2009. http://landsathandbook.gsfc.nasa.gov/handbook.html.

Nordstrom, D.K., R.B. McCleskey, and J.W. Ball. 2009. Sulfur geochemistry of hydrothermal waters in Yellowstone National Park: IV Acid-sulfate waters. *Applied Geochemistry*, 24(2): 191-207.

Norton, D.R. and I. Friedman. 1985. Chloride flux out of Yellowstone National Park. *Journal of Volcanology and Geothermal Research*, 26(3-4): 231-250.

Pal, M. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1): 217-222.

Patrick, M., K. Dean, and J. Dehn. 2004. Active mud volcanism observed with Landsat 7 ETM+. *Journal of Volcanology and Geothermal Research*, 131(3-4): 307-320.

Prasad, A.M., L.R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, 9(2): 181-199.

Rojstaczer, S., D.L. Galloway, S.E. Ingebritsen, and D.M. Rubin. 2003. Variability in geyser eruptive timing and its causes: Yellowstone National Park. *Geophysical Research Letters*, 30(18): 2-1 to 2-4.

Savage, S.L. 2009. “Mapping changes in Yellowstone’s geothermal areas”, Montana State University, Bozeman, Montana, 210 pp.

Seielstad, C. and L. Queen. 2009. Thermal remote monitoring of the Norris Geyser Basin, Yellowstone National Park. Final Report for the National Park Service Cooperative Ecosystem Studies Unit, Agreement No. H1200040001, 38 pp.

Sesnie, S.E., P.E. Gessler, B. Finegan, and S. Thessler. 2008. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Remote Sensing of Environment*, 112(5): 2145-2159.

Shaw, J.A. and C. Marston. 2000. Polarized infrared emissivity for a rough water surface. *Optics Express*, 7(11): 375-380.

Singh, A. 1989. Review article: digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10(6): 989-1003.

Song, Y., H.-C. Kim, B.-W. Yum, and E. Ahn. 2005. Direct-use geothermal development in Korea: Country update 2000-2004, Proceedings of the World Geothermal Congress, pp. 1-7.

Sorey, M.L. Effects of potential geothermal development in the Corwin Springs Known geothermal resources area, Montana, on the thermal features of Yellowstone National Park. Water-Resources Investigations Report 91-4052, U.S. Geological Survey, Menlo Park, California, 204 pp.

Spatial Analysis Center. 2005. Hydrogeothermal Areas of Yellowstone National Park, Wyoming, Montana, Idaho. Digital Spatial Data, Yellowstone Center for Resources, Yellowstone National Park, Wyoming.

Spatial Analysis Center. 2008. Thermal Inventory point data for Yellowstone National Park, 1998 - 2008. Digital Spatial Data, Yellowstone Center for Resources, Yellowstone National Park, Wyoming.

Thompson, J.M. and S. Yadav. 1979. Chemical analyses of waters from geysers, hot springs, and pools in Yellowstone National Park, Wyoming from 1974 to 1978. Open-File Report 79-704, United States Geological Survey, Menlo Park, California, 51 pp.

Urai, M. 2002. Heat discharge estimation using satellite remote sensing data on the Iwodake volcano in Satsuma-Iwojima, Japan. *Earth Planets Space*, 54(3): 211-216.

Utah State University. 2008. Image Standardization: At-sensor Reflectance and COST Correction. Last accessed on 8 April 2009. http://earth.gis.usu.edu/imagestd/.

Ward, D.M., R. Weller, J. Shiea, R.W. Castenholz, and Y. Cohen. 1989. Hot spring microbial mats: anoxygenic and oxygenic mats of possible evolutionary

significance. In: Y. Cohen and E. Rosenburg (Editors), Microbial Mats: Physiological Ecology of Benthic Microbial Communities. American Society for Microbiology, Washington, D.C., 494 pp.

Waring, G.A., R.R. Blankenship, and R. Bentall. 1983. Thermal springs of the United States and other countries of the world - a summary. U.S.G.S. Professional Paper 492, U.S. Geological Survey, Alexandria, Virginia, 401 pp.

Watson, F.G.R., R.E. Lockwood, W.B. Newman, T.N. Anderson, and R.A. Garrott. 2008. Development and comparison of Landsat radiometric and snowpack model inversion techniques for estimating geothermal heat flux. *Remote Sensing of Environment*, 112(2): 471-481.

Watson, K. 1975. Geologic applications of thermal infrared images. *Proceedings of the IEEE*, 63(1): 128-137.

Western Regional Climate Center. 2005. Period of record monthly climate summary - Yellowstone Park, Wyoming (489905). Last accessed on 6 April 2009. http://www.wrcc.dri.edu/cgi-bin/cliRECtM.pl?wyyell.