



Report No. WI-2002-07  
31<sup>st</sup> October 2002

## The Watershed Institute

Earth Systems Science and Policy  
California State University  
Monterey Bay  
<http://watershed.csumb.edu>

100 Campus Center, Seaside, CA,  
93955-8001  
831 582 4452 / 4431.

*Central  
Coast  
Watershed  
Studies*

**CCoWS**

## A remote sensing map of Yellowstone's geothermals

Dr Fred Watson<sup>1</sup>  
Wendi Newman<sup>1</sup>  
Thor Anderson<sup>1</sup>  
Ann Rodman<sup>2</sup>  
Doug Ouren<sup>3</sup>  
Dr Joseph Coughlan<sup>4</sup>  
Dr Bob Garrett<sup>5</sup>

<sup>1</sup> Watershed Institute, California State University Monterey Bay

<sup>2</sup> National Parks Service, Yellowstone

<sup>3</sup> USGS, Bozeman

<sup>4</sup> NASA Ames Research Laboratory

<sup>5</sup> Ecology Department, Montana State University Bozeman

Lead author contact details:  
[fred\\_watson@csumb.edu](mailto:fred_watson@csumb.edu)

## Acknowledgements

Primary funding for this work was provided by NASA Grant NCC2-1186. Supporting funding was provided by the National Parks Service, and the National Science Foundation (NSF Grant DEB 0074444). Large format maps were printed with assistance from The David and Lucille Packard Foundation.

We are especially grateful to PJ White (NPS) for recognizing the broader utility of this work from its early beginnings, and for continuing to facilitate the necessary connections between staff at CSUMB, MSU, and NPS that have brought it to maturity. We appreciate the support and advice of Park Geologists Dr Hank Heasler and other Park staff including Dr Cheryl Jaworowski and Steve Miller.

We thank the 2001-2 MSU elk tracking crew for sharing their intimate knowledge of backcountry thermals, and for helping with the interpretation of our maps - Adam Messer, Eric Bergman, Claire Gower, and Mark Johnson.

Thanks also go to the CSUMB field crew: Joel Casagrande, Julie Hager, Dr Susan Alexander, Don Kozlowski, and volunteers Dr Rob Burton and Kristy Uschyk.

Finally, thanks to Dr Don Watson and Randal Watson for statistical advice.

## Table of Contents

<b>Acknowledgements</b> .....	<b>1</b>
<b>Table of Contents</b> .....	<b>2</b>
<b>Introduction</b> .....	<b>4</b>
1.1 Background.....	4
1.2 Existing data .....	5
1.2.1 Hutchinson’s polygons.....	5
1.2.2 Shovic’s polygons .....	6
1.2.3 MSU elk-tracking points.....	6
1.2.4 YNP Thermal Inventory points .....	6
1.2.5 Heat flow bore holes .....	7
1.3 Limitations of existing data .....	7
1.4 Study area .....	8
<b>2 Methods</b> .....	<b>10</b>
2.1 Initial image processing .....	10
2.2 Raw thermal imagery.....	10
2.3 Contrast between geothermal and snowpack influence.....	11
2.4 Effects of solar heating.....	16
2.5 Elevation .....	17
2.6 Combining the data using unsupervised classification.....	20
<b>3 Results</b> .....	<b>23</b>
3.1 The final image .....	23
3.2 Interpretation of explanatory variables .....	28
<b>4 Evaluation</b> .....	<b>30</b>
4.1 General observations.....	30
4.1.1 Upper Geyser Basin .....	30
4.1.2 Midway and Lower Geyser Basins .....	30
4.1.3 Mary Mountain area .....	33
4.1.4 Norris Geyser Basin.....	33
4.1.5 Madison Junction area.....	35
4.2 Qualitative comparisons with other data.....	37
4.2.1 Sensitivity of present image .....	37

4.2.2	Pre-existing polygon data: Hutchinson, and Shovic.....	37
4.2.3	Pre-existing point data: MSU elk, and NPS Thermal Inventory.....	38
4.2.4	Visual comparison: all large data sets overlaid .....	39
4.2.5	Purposely acquired polygon data: CSUMB.....	47
4.3	Quantitative comparisons with other data .....	50
4.3.1	Methods.....	50
4.3.2	Results.....	53
4.3.3	Sensitivity to spatial error .....	56
<b>5</b>	<b>Inferring heat flux .....</b>	<b>58</b>
<b>6</b>	<b>Summary and further work .....</b>	<b>60</b>
6.1	Summary.....	60
6.2	Uses.....	61
6.2.1	Thermal monitoring .....	61
6.2.2	Exploration and inventory .....	62
6.2.3	Geologic processes .....	62
6.2.4	Landscape mapping .....	62
6.2.5	Snowpack modeling .....	62
6.2.6	Integrated visualization of the Park .....	62
6.3	Further work .....	63
<b>7</b>	<b>References .....</b>	<b>64</b>

## Introduction

### 1.1 Background

Yellowstone National Park contains the world's largest active geothermal area. We describe the first published attempt to map this area using large-scale satellite remote sensing techniques. The results provide a new level of information on the Park's geothermals, suitable for a wide range of purposes. Previous mapping efforts have either relied on subjective air-photo interpretation, or have focused on mapping of point features rather than the spatial extent and intensity of geothermal influence.

Simultaneous measurement of both the total area and intensity of geothermal influence benefits a surprising diversity of scientific endeavor. Perhaps most obviously, the center of the Park is a caldera that is highly dynamic in both space and time (Waite & Smith, 2002). Objective mapping of this activity contributes to the understanding of the underlying processes, and to predictions of future events. *Thermus aquaticus*, that most famous Archaic life form responsible for the multi-billion-dollar DNA industry, was discovered in a thermal in the Park (Brock, 1994). The rarest and 'newest' kingdom of life, the Korarchaeota, are known to exist only from a single pool in the Park (Madigan et al., 2000). Contemporary bio-prospecting is drawn by such profound discoveries, and relies on existing thermal maps to guide sampling expeditions. Three million tourists are drawn to the Park annually from around the world for its many attractions (YNP, 2000), not least of which is an opportunity for wildlife viewing unmatched in North America. That Yellowstone's geysers form a backdrop to the herds of the 'Serengeti of America' is no coincidence. Forage production proceeds year round in the snow free havens provided by the Earth's heat. The home ranges of the Park's only non-migratory elk population are defined by proximity to the Park's thermal refugia (Craighead et al., 1973; Garrott et al., 2002 in press.). Some 60 million bison once roamed the nearby states. The largest wild herd of their descendants in existence, comprising some 3200 animals, migrates annually along an alley of geothermally influenced meadows. The total sustainable population of these ungulates is hypothesized to be dependent on the forage area made available by the absence of snow (Garrott et al., 2002 in review), a relationship that promises to become more complex now that the first re-introduction of wolves to the conterminous United States has succeeded in spectacular fashion (Smith & Guernsey, 2002).

A complete understanding of all the above systems relies on knowing where the geothermals are, how large they are, how hot they are, and whether they are changing. Objective, quantitative mapping of the spatial extent and intensity of all geothermals satisfies many direct research needs. It serves as a geologic tool for spatial monitoring of the intensity, extent, and location of geothermally affected areas. If repeated, it may facilitate the discovery of newly developing thermal areas, especially in rarely visited backcountry areas. The total area or total heat flux mapped can serve as a critical dependent variable in models of population dynamics and wildlife–habitat relationships.

One of the most novel potential contributions of a geothermal map with continuous coverage across the landscape is as a component of a new expression of the integrated Park ecosystem that is being developed by the authors. We argue that the bulk of scientific knowledge and communication relating to the Park is currently limited to its individual components, rather than the system as a whole. And yet the complex interactions that comprise the whole are perhaps its most fascinating aspect. Scientists, managers, and the public alike are denied the ultimate appreciation for the Park until we can understand and express it holistically. We propose a dynamic visualization of the Park, underlain by geothermal influence, and highlighting dynamic links with terrain, vegetation, fire, snow, ungulates, and predators. This requires data and models, all dependent on geothermal mapping in a form that is compatible with large–scale landscape analysis.

## 1.2 Existing data

There are numerous existing data that include some form of information on the location of geothermal areas in the park. The most basic information is the inclusion of selected point features such as geysers and hot–springs on standard USGS topographic maps. These are available digitally, and include most of the obvious features mapped as blue circles on 1:24,000 topographic maps.

### 1.2.1 *Hutchinson's polygons*

The first serious geothermal mapping effort was conducted by Park Geologist Rick Hutchinson, who after years of experience working in the backcountry was

able to manually delineate the outline of thermal areas on a series of aerial photographs of the Park sometime prior to 1988. These were digitized some years later by other workers, producing the first polygon coverage that specifically delineates thermal areas. The coverage supplied to the present project includes 268 polygons, 263 of which are inside the study area.

### *1.2.2 Shovic's polygons*

More recent field mapping and air-photo interpretation lead to the creation of the 'landforms' digital polygon layer for the Park (Shovic, 1996). Although this study mainly addressed landforms from a geologic and geomorphic perspective, one of the possible polygon attributes in the coverage is 'hydrothermal'. We selected all such polygons (63 in total), thus yielding an additional source of spatial geothermal data.

### *1.2.3 MSU elk-tracking points*

The most recent spatial data are two point-feature (as opposed to polygon) data sets. The first of these is the Montana State University (MSU) Bozeman wildlife team's elk-tracking database. These data contain 11 years of elk locations obtained by daily tracking of randomly selected VHF-collared animals on foot during every winter since November 1991 (Garrott et al., 1997, 2002 in review). Each time a selected animal is located, the adjacent habitat is recorded, with one of the attributes being 'thermal'. The 10,238 points (through May 2002) thus include 1421 thermal locations and 8817 non-thermal locations. Because the points were selected by animals not humans, they are free from any biases humans may introduce in where they chose to survey for thermals. The locations were computed using human field navigation with 1:24,000 topographic maps. GPS's were used in latter years for occasional checks of navigation accuracy. Location errors of up to 50 m are expected in the data, drifting to 100 m in rare instances.

### *1.2.4 YNP Thermal Inventory points*

The second point-feature data set is currently being produced by the Park's 'Thermal Inventory' project. To collect these data, field survey crews are visiting all known thermal areas and using GPS to record accurate locations, temperatures, descriptions, photography, and other attributes. The data are

highly accurate from a spatial perspective, but emphasize the point-locations of specific thermal features, rather than the spatial extent of thermally influenced areas. The present work uses the 4,127 points from this database that had been collected through mid-summer 2002.

#### *1.2.5 Heat flow bore holes*

A larger, continental data set is also of interest. The USGS and other federal agencies have been mapping geothermal resources for development potential for several decades. The result is a large database of ground heat flow measurements based on temperature gradients observed in deep bore holes. The data can be queried using online tools ([http://map.ngdc.noaa.gov/website/NVDS/hot\\_springs/viewer.htm](http://map.ngdc.noaa.gov/website/NVDS/hot_springs/viewer.htm)). A informative set of interpolated maps based on these data are available from the Southern Methodist University web site at <http://www.smu.edu/geothermal/heatflow/heatflow.htm> (Blackwell et al., accessed 2002). The surface heat flow map from this publication shows the greater Yellowstone area as being subject to an average of 120–150 mW/m<sup>2</sup> of upward heat flow. Beneath the central Yellowstone Plateau (including the present study area), the mean heat flow increases to 2,000 mW/m<sup>2</sup> with localized flux as high as 58,600 mW/m<sup>2</sup> (Fournier, 1989; cited by Smith, 2002).

### **1.3 Limitations of existing data**

The key limitation of the existing data is the lack of information on the spatial extent of geothermally influenced areas, and on the total heat flux integrated over an area. This information is essential when considering efforts such as large-area geothermal monitoring and the influence of geothermals on large-scale landscape and wildlife processes.

Where spatial extent has been mapped, it has required the use of expensive and subjective techniques such as air-photo interpretation. The resulting data are of limited use mainly because of the difficulty of repeating the measurements either between laboratories, or between two points in time. Further, unless thermal infrared photography is used, the bare, sintered surface of remnant geothermal areas can easily be misinterpreted.

There is a need for more rapid and objective methods that quantify the intensity and extent, as well as the location of geothermally affected areas.

#### **1.4 Study area**

In order to optimize the time spent developing a suitable technique, we limited our mapping to the western central portion of the Park where the most geothermal activity takes place (Fig. 0.1).

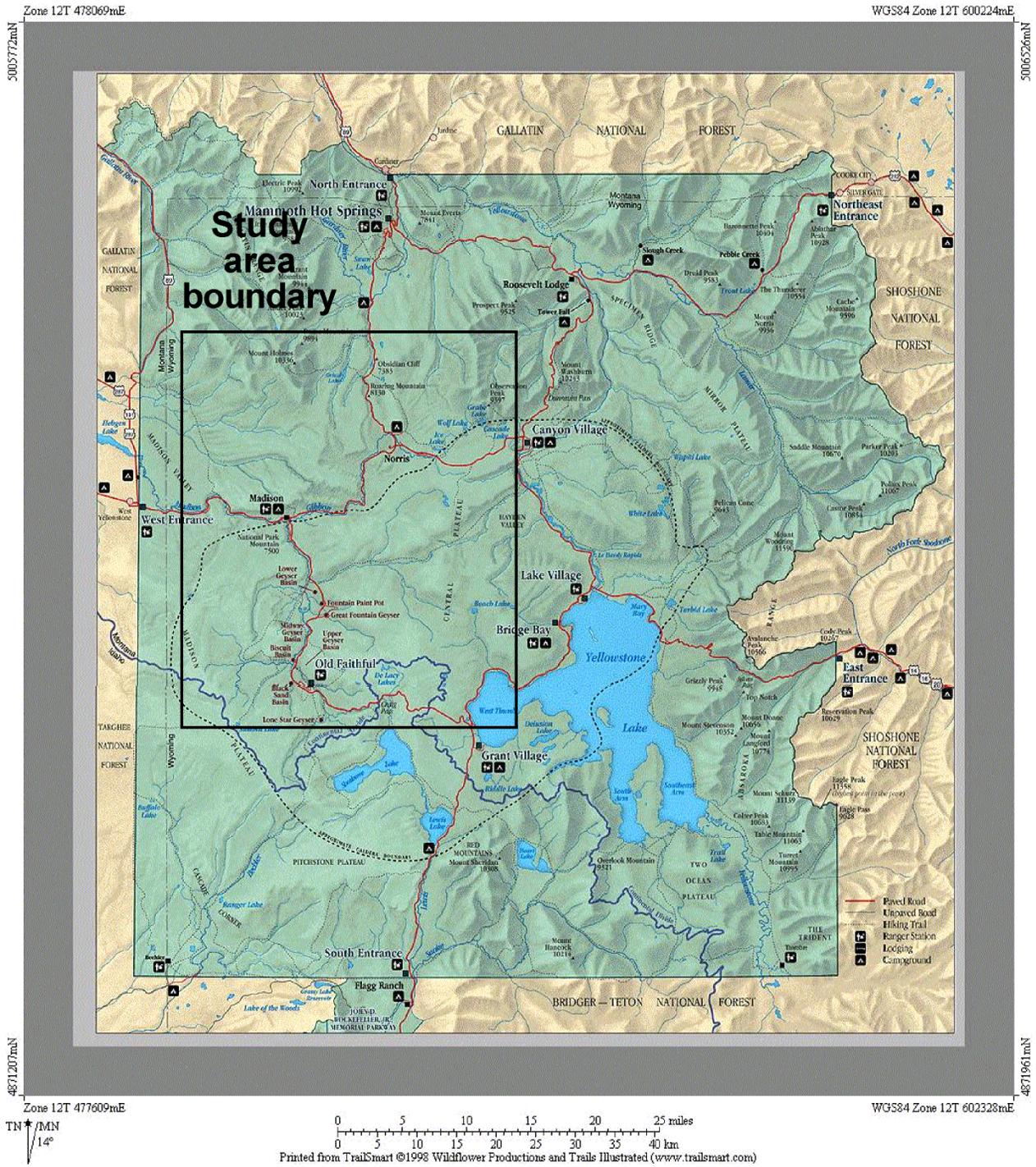


Figure 0.1. Study area: western central Yellowstone National Park.

## 2 Methods

The following topics describe the steps involved in developing a suitable technique for remote-sensing-based geothermal mapping in the Park. The initial processing of raw Landsat data are first described, and the raw thermal bands are examined. Then, a standard image of snow cover is computed, to see the extent to which geothermal influence may be mapped by estimating presence or absence of snow. The snow image corresponded well with known geothermal influence, but could not differentiate between areas that were snow-free due to thermals, or snow-free due to rocks and solar radiation – false positives. It was decided to use a multivariate classification of various images to estimate geothermal influence. This classification used Landsat thermal bands, as well as theoretical images of solar radiation influence, and terrain elevation. The details follow.

### 2.1 Initial image processing

The primary source data for the mapping was a Landsat-7 Thematic Mapper image acquired on March 25<sup>th</sup> 2000 and obtained from the USGS EROS Data Center with the assistance of the USGS lab at Montana State University Bozeman. The imagery was terrain corrected by USGS. This is a high level of correction made possible by recent developments in widespread digital terrain mapping. It accounts for the parallax error associated with vertically displaced terrain (e.g. mountains) being laterally displaced in the image due to the oblique observation angle of the satellite for all observations other than the nadir directly below the satellite. The data were obtained in UTM projection on the WGS84 datum, and were not re-sampled to any other projection by the authors. The thermal bands (6L & 6H) were, however, re-sampled from their original 57 m pixel size, down to half this size (28.5 m) in order to match the resolution of the visual bands. No smoothing was associated with this process.

### 2.2 Raw thermal imagery

The Landsat-7 Thematic Mapper scans two thermal bands 6H and 6L (10.4 – 12.5 micron). Given that an appropriate scene acquisition date is chosen, these bands are able to discriminate most of the major thermal features of the park without any sophisticated correction. Each band is discretized to a maximum of 255 discrete radiometric values, known simply as ‘digital numbers’ (DN). Figure

2.1 shows the spatial distribution of the highest radiometric values in Band 6H over the western central Park. All major thermal features are included, but so too are numerous ‘false positives’ – non-thermal areas that have an equal satellite radiometric response to true thermal areas. It is not possible to select radiometric values in the image that include all geothermal areas without also including a significant number of non-geothermal areas.

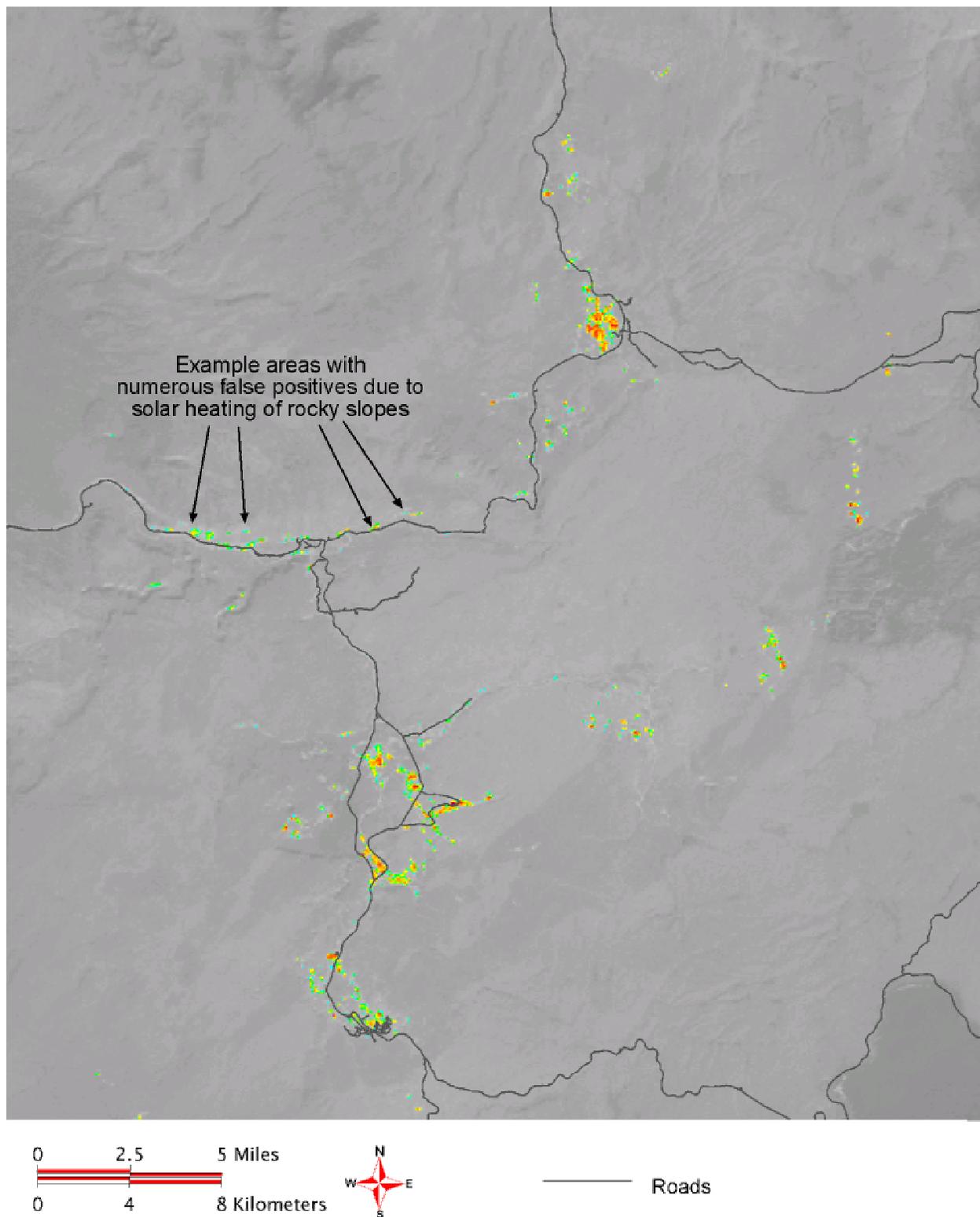
### 2.3 Contrast between geothermal and snowpack influence

The scene was acquired, near the time of peak snowpack accumulation in this part of the Park (Fig. 2.2). Fresh snow in mid-winter, because of its high albedo relative to soil and vegetation, provides a uniformly cold background that enhances the recognition of thermal features by longwave radiative heat flux. The judicious choice of acquisition date to make best use of snow explains the good thermal sensing results obtained with a single raw thermal band. Suitable clear-sky winter dates are rare. Frequently, winter images are either too cloudy, or spoiled by plumes of steam rising into the cold winter air from hydro-thermal vents.

The notion of using presence or absence of snow as a direct indicator of geothermal influence was examined explicitly by computing the normalized difference snow index (NSDI) (Dozier, 1989):

$$NSDI = \frac{Band2 - Band5}{Band2 + Band5}$$

This is mapped in Figure 2.3, and again with a common ‘snow indicator’ threshold value of 0.4 in Figure 2.4. The index is a useful discriminator of the presence of snow in the Park. As such, however, it includes non-geothermal areas such as south-facing cliffs and low-elevation meadows.



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 2.1. Landsat-7 thermal band 6H over western central Yellowstone National Park, March 25<sup>th</sup> 2000 – only the highest radiometric values are colored.**

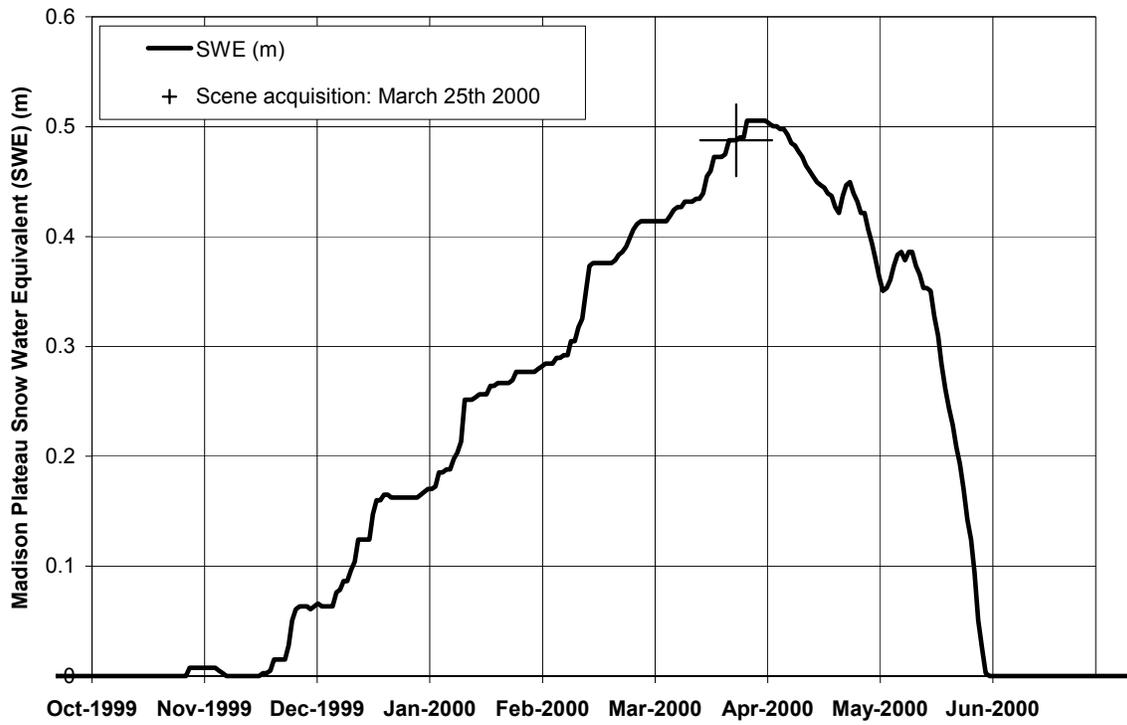
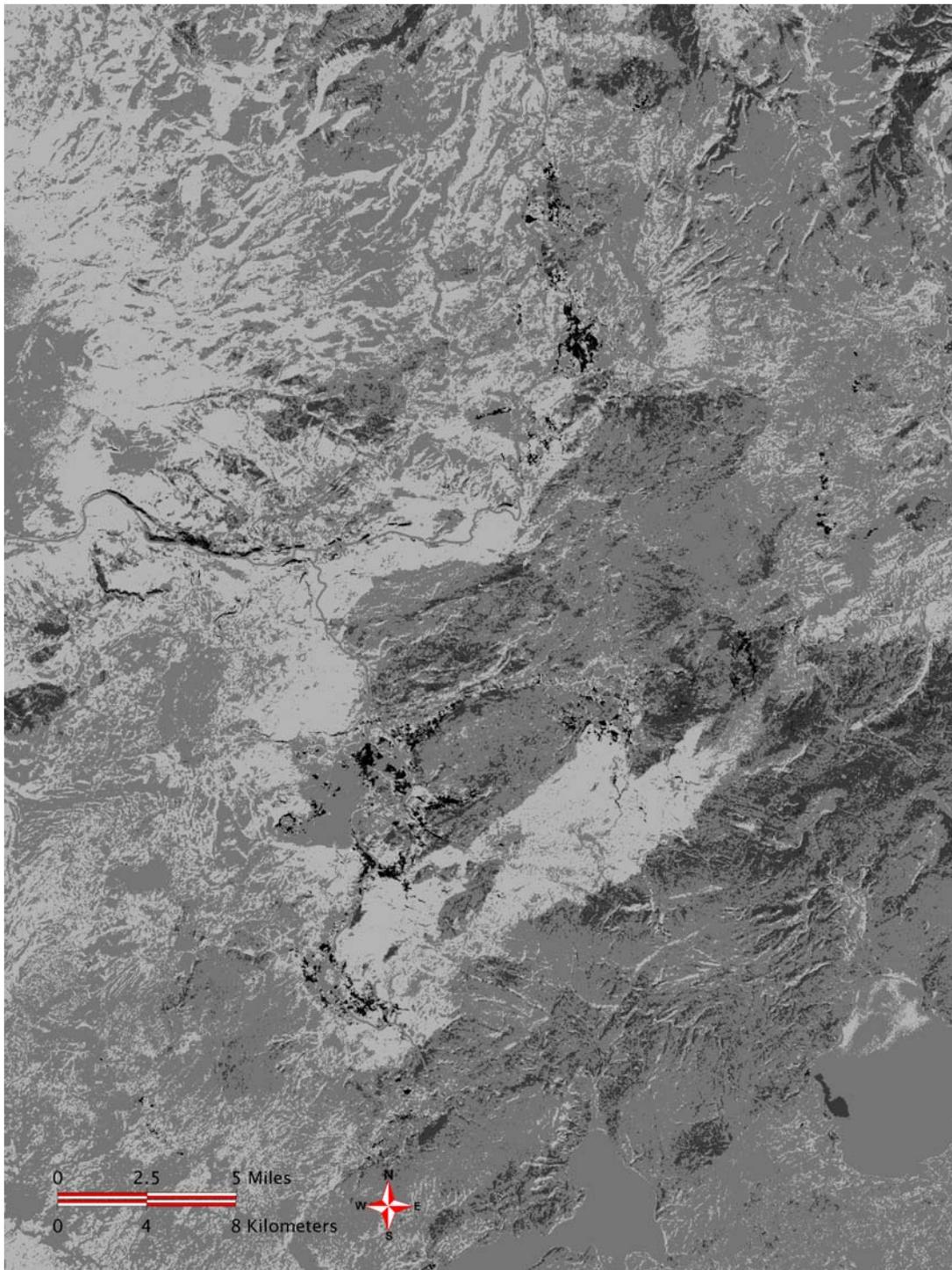


Figure 2.2. Snowpack accumulation around the time of scene acquisition.



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 2.3. Landsat-7 normalized difference snow index (NDSI) for western central Yellowstone National Park – March 25th 2000.**



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

Figure 2.4. Presence of snow estimated by thresholding the NSDI index above 0.4.

## 2.4 Effects of solar heating

False positives in the raw thermal image are most prevalent on steep, rocky, south-facing slopes. The implication is that they are primarily due to solar heating of these slopes, and latter re-radiation of longwave radiation, particularly from sun-bathed rocks. A technique known as Minnaert shading correction was used to compensate for these effects in a systematic manner. Minnaert correction was originally developed for inverse estimation of lunar terrain using visual images (Minnaert, 1941) and has been more recently applied in shading compensation of visual images of Earth (Smith et al., 1980; Watson et al., 1999).

If one assumes a diffusely reflective landscape, one can predict the component of a remotely sensed visual-wavelength image that is due solely to differences in solar illumination of different slopes in the terrain. One can then divide an real image by this predicted image to get an image containing just information about the visual landscape itself, and not about the interaction between the topography and the direction the sun happened to be shining when the image was acquired.

Real natural surfaces are not, however, entirely diffuse reflectors. The reflectance from a landscape surface depends not just on the angle of illumination, but also upon the angle of observation. This latter dependence is due to what is known as specular reflection, a mirror being the most extreme example. The Minnaert reflectance function is:

$$F = \cos^{k(\lambda)} i \cos^{k(\lambda)-1} e$$

where  $F$  is the Fraction of light that is reflected,  $i$  is the angle of incidence of sunlight,  $e$  is the angle of exitance (the observation angle), and  $k(\lambda)$  is the Minnaert constant, which can be determined for a given image and wavelength (band) by linear regression involving observed image values and values computed using the terrain and solar geometry (Smith et al., 1980).

By calibrating  $k$  against observed images, one can compute a map of the way and landscape would look if its “average” surface cover were distributed uniformly across the landscape, and then lit by the sun. The real image can then

be divided by the average image to compute an anomaly image, revealing much about spatial variation in surface properties, and removing most effects due to solar shading at the time of image acquisition.

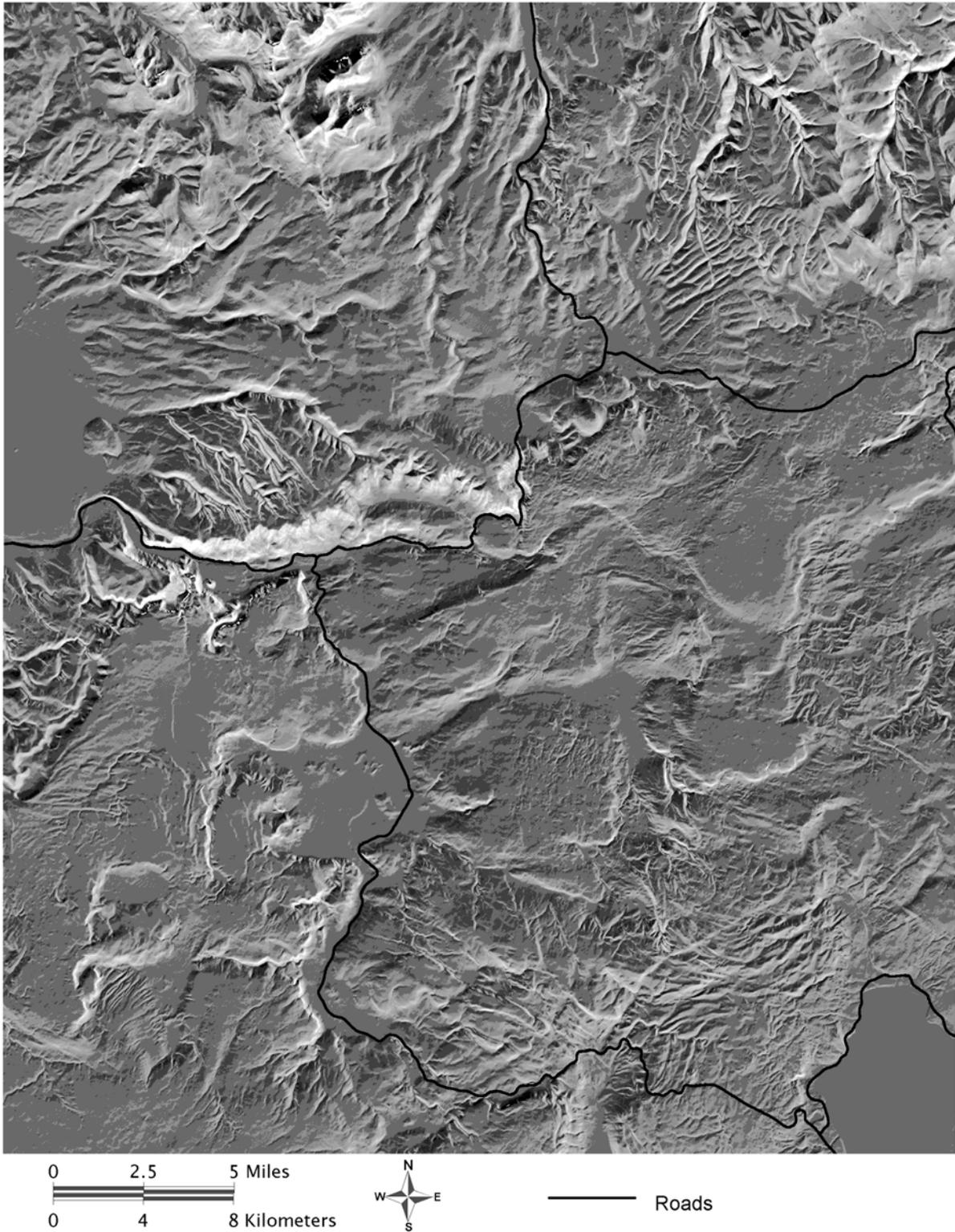
Thermal images of landscapes measure longwave radiation emitted from surfaces in the landscape (and the air above it) in proportion to the fourth power of their temperature (the Stefan–Boltzmann law). Solar radiation has a significant shortwave component, the absorbed fraction of which heats up surfaces, thus changing their longwave emission. Longwave radiation from the sun is largely absorbed by natural surfaces, also heating them up. Thus, in order to use the Minnaert correction technique for thermal images, we effectively make the assumption that the thermal emission of landscapes is governed in some way by its solar reflection properties. This is a weak assumption, but our experience has been that incorporating a theoretical Minnaert shading image into land cover classification improves the accuracy of results considerably (see results in Section 3.1).

The Minnaert shading image was computed for each band of the Landsat ETM image using the Tarsier environmental modeling framework (Watson et al., 2001). Shade–corrected versions of each band were also computed, although as described below, better mapping results were obtained using the shading image (Figure 2.5) directly, rather than the shade–corrected bands.

The use of these computed images is described in Section 2.6.

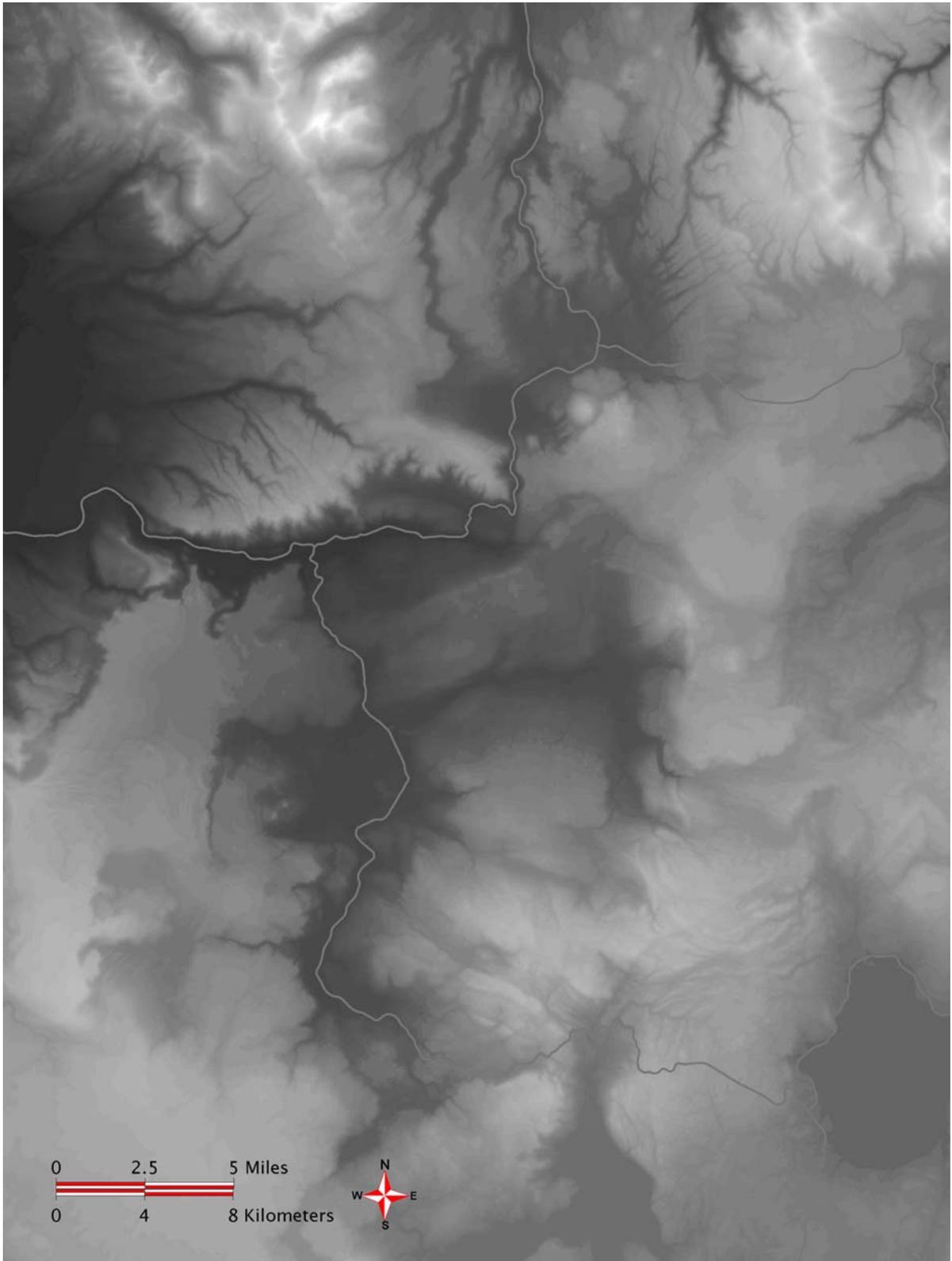
## 2.5 Elevation

From visual examination of the raw thermal imagery, an additional cause of false–positives appeared to be solar heating of slopes (flat or otherwise) at elevations low enough to be free of snow for non–geothermal reasons. Thus the digital elevation model supplied with the imagery (Fig. 2.6) was also included in the analysis.



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 2.5. Minnaert shading image of western central Yellowstone National Park, calibrated against 'reflectances' measured by Landsat-7 Thematic Mapper Band 6L.**



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 2.6. Digital elevation model (DEM) of western central Yellowstone National Park.**

## 2.6 Combining the data using unsupervised classification

A number of techniques can be envisioned for deriving a single geothermal image by using information from multiple sources such as thermal imagery, theoretical shading imagery, and elevation models. Perhaps the most direct is to construct a physical model of the radiation and heat balance of the terrain. Such a model would need to take into account the vegetation, snow, ground heat flux, radiative properties of surfaces and the atmosphere, and the heat capacity of different materials. This is rather involved, and unwarranted before simpler dimensional-reduction techniques.

Unsupervised classification assumes nothing about the physical processes that combine to bring about radiometric heat flux. Rather, this technique takes multivariate data sets and attempts to identify discrete groups of data points as they exist in the multivariate space defined by the variables of interest.

Our approach was to run standard unsupervised classifications on various sets of variables until a suitable set of variables was identified. In each test, an image would be created of all the classes identified by the automated procedure. We would then manually divide the classes in to two groups: those classes containing known thermal areas, and those containing non-thermal areas respectively. The choices available to us in each test included:

- The number of classes to be generated by the classification (50, 100, 200, or 400)
- Inclusion/exclusion of individual Thematic Mapper bands (1, 2, 3, 4, 5, 6L, 6H, 7) in the unsupervised classification
- Inclusion/exclusion of DEM
- Inclusion/exclusion of terrain slope image
- Inclusion/exclusion of terrain aspect image
- Inclusion/exclusion of vegetation index (NDVI)
- Inclusion/exclusion of various land cover images
- Inclusion/exclusion of theoretical shading images (calibrated to each band)
- Inclusion/exclusion of shade-corrected bands
- Inclusion/exclusion of snow index (NDSI)

- Weightings associated with any included image

In early tests, our first-hand knowledge of the Park was sufficient to judge certain tests as poor and discard them. Similarly, the attributes of the tests that tend to lead to good results could be identified. Over the course of about 25 classifications, some good candidate images were produced that could be submitted to more substantive verification. Initially this involved producing integrated maps containing the candidate imagery and showing them to Park experts such as the science staff at Park HQ, and the wildlife tracking crew at MSU. Earlier digital data sets containing geothermal information were also overlaid on these maps, including polygons of 'thermal' landscapes from the Park's land cover database, point features from early USGS maps, and 'thermal' points from the 11,000-point MSU elk-tracking database.

The merits of each data source for the mapping exercise are summarized. Inclusion of the DEM was useful in eliminating false positives at low elevations. Slope and aspect images help eliminate false positives on sunny slopes, but the theoretical shading image was superior for this purpose. Shade-corrected bands were not as useful in the classification as the separate inclusion of the raw band and the theoretical shading image. Any image involving vegetation or land cover contributed nothing of use. Similarly, images from the visual and near-infrared bands (1-5) contained too many vegetation effects. Weightings were important in order to bring the relative influence of each variable into suitable balance. Using small numbers of classes is helpful because it is quicker, but the best results are obtained when one takes the time to interpret very large classifications containing 400 classes.

For the penultimate image, a number of field excursions were conducted to specifically confirm and map the existence of suspected 'new' positives, to deny the existence of suspected false positives, to investigate suspected false negatives, and to clarify the boundaries of true positives. A statistical discussion of these ground-truthing and other verification procedures is given in below Section 4. The final image arose from the corrections made to the penultimate image after these field excursions.

The data flow used in the generation of the final image is illustrated in Figure 2.7.

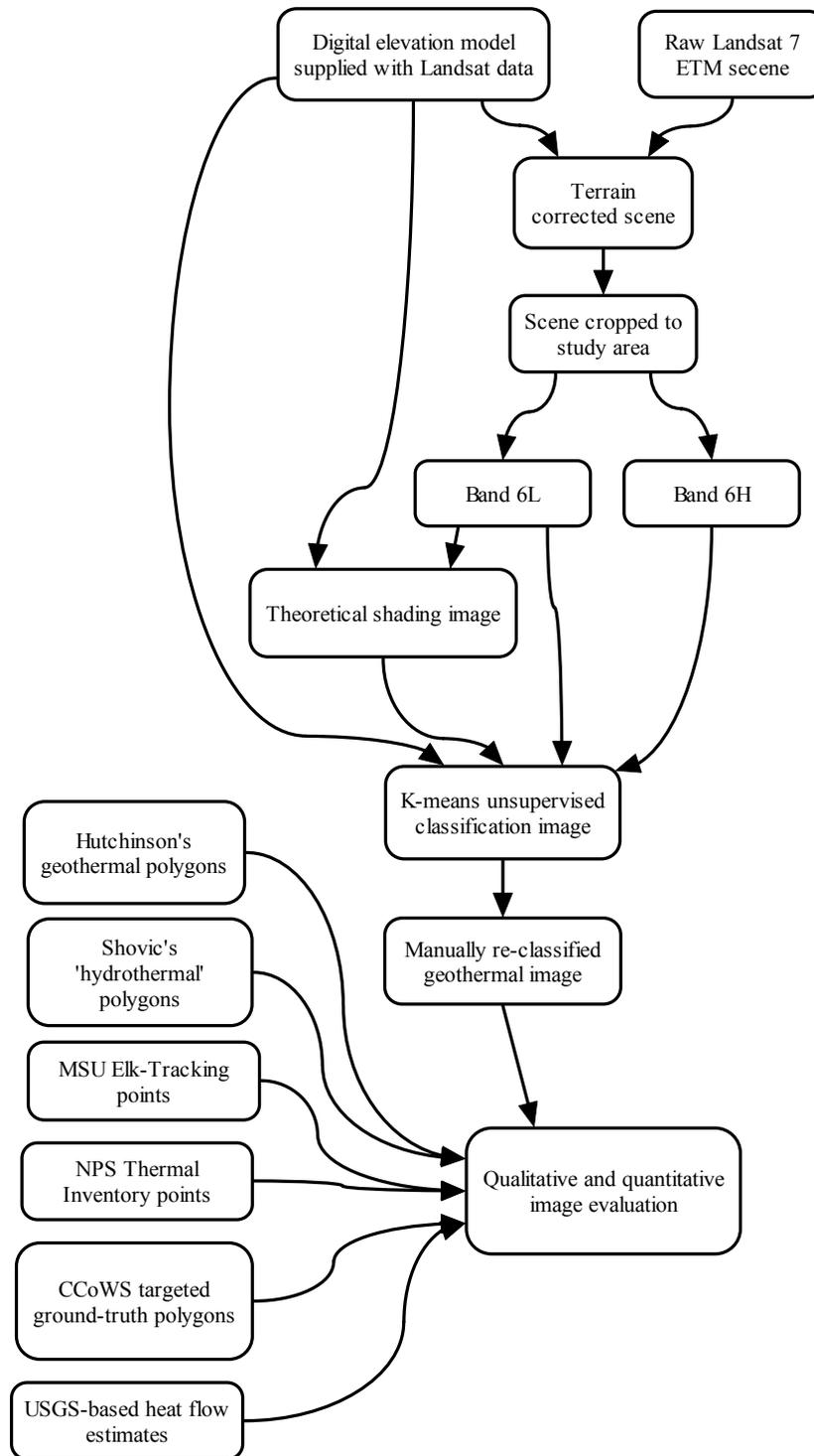


Figure 2.7. Flow of all data used to produce and evaluate the final geothermal image.

## 3 Results

### 3.1 The final image

The technical details of the final image are as follows.

K-means unsupervised classifications (Tou & Gonzales, 1974; see also Likas et al., 2002) were run using the the GIS software TNTMips V6.4 (Microimages, 2000). The most superior combination was the following set of variables input to a classification delineating 400 classes:

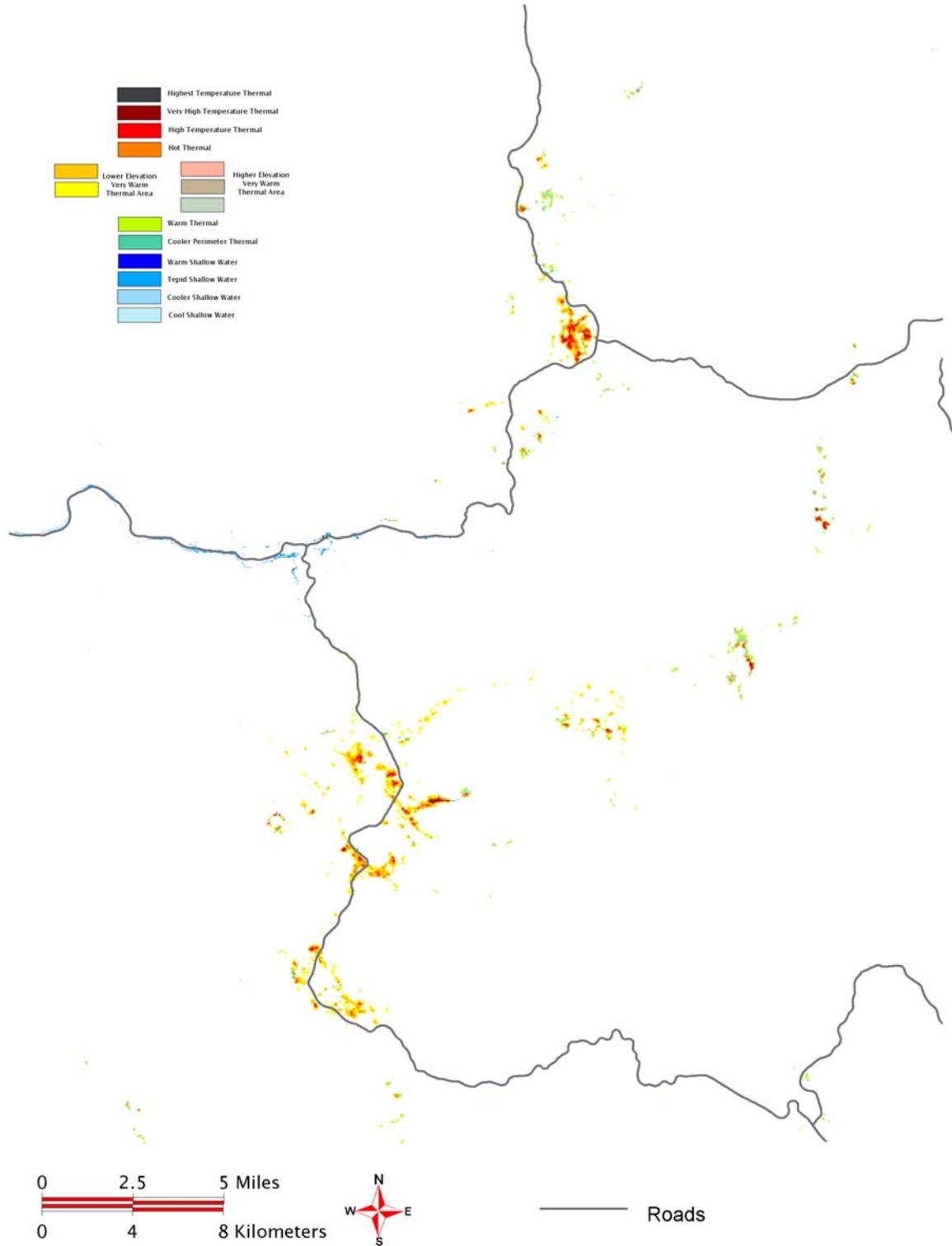
- March 25<sup>th</sup> 2000 ETM Band 6L re-sampled from 57 m down to 28.5 m, supplied as digital numbers (DN) in the range 0–255
- March 25<sup>th</sup> 2000 ETM Band 6H re-sampled from 57 m down to 28.5 m, supplied as digital numbers (DN) in the range 0–255
- Minnaert shading image based on 28.5 m DEM supplied with imagery, calibrated against Band 6L ‘reflectances’, values multiplied by 90 for weighting, then discretized to integers in the range 0–255
- Elevation image from 28.5 m DEM supplied with imagery, values in metres, multiplied by 0.05 for weighting, discretized to integers

The resulting image is shown in Figure 3.1. A total of 15 classes from the original 400 unsupervised classes were manually labeled as thermal. These are approximately aligned along a heat gradient from extremely hot, down to areas so weakly thermally influenced that they only thin the snowpack, rather than completely melt it. Note that the image has no way of discriminating direct sub-surface geothermal influence from heat advected away from direct geothermal sources in seeps and rivers. Thus, the Firehole, Madison, and Gibbon Rivers are correctly identified as geothermally *influenced*, downstream from respective geothermal heat sources.

The interpretation of several levels of heat is based mainly on the concentric patterns observable in large thermal basins. We infer that, where one class completely surrounds another, and so on, that the center of such a pattern is hotter than the periphery. We were able to qualitatively confirm the validity of this inference by examining the attributes of points from the recent Park thermal inventory based on exhaustive ground survey. Significant, very hot features identified in this inventory (such as boiling pools) were typically located

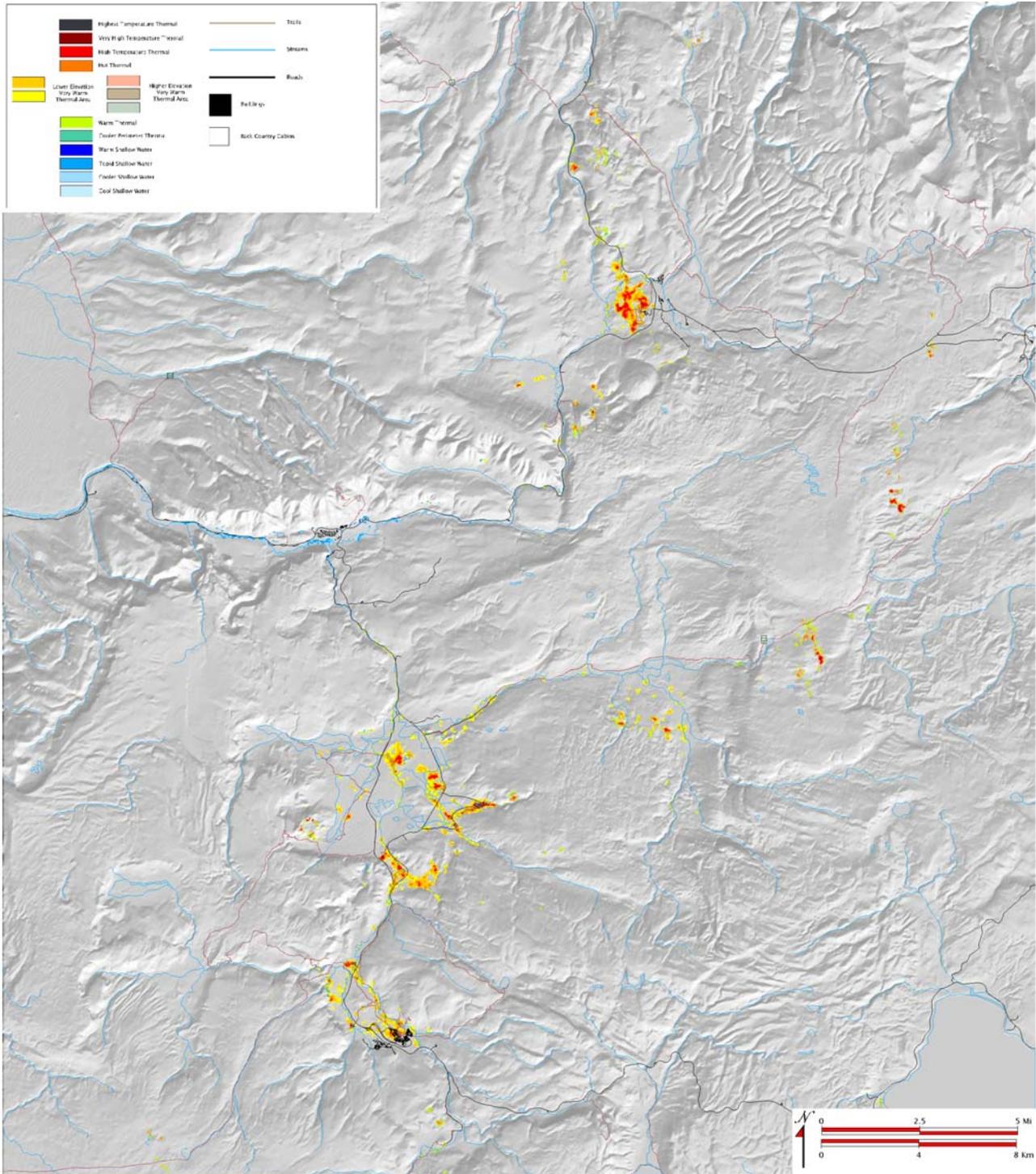
at the centers of concentric patterns indicated as 'Highest temperature' in the final map.

The mapped location of geothermal areas relative to terrain and vegetation type is shown in Figures 3.2 and 3.3.



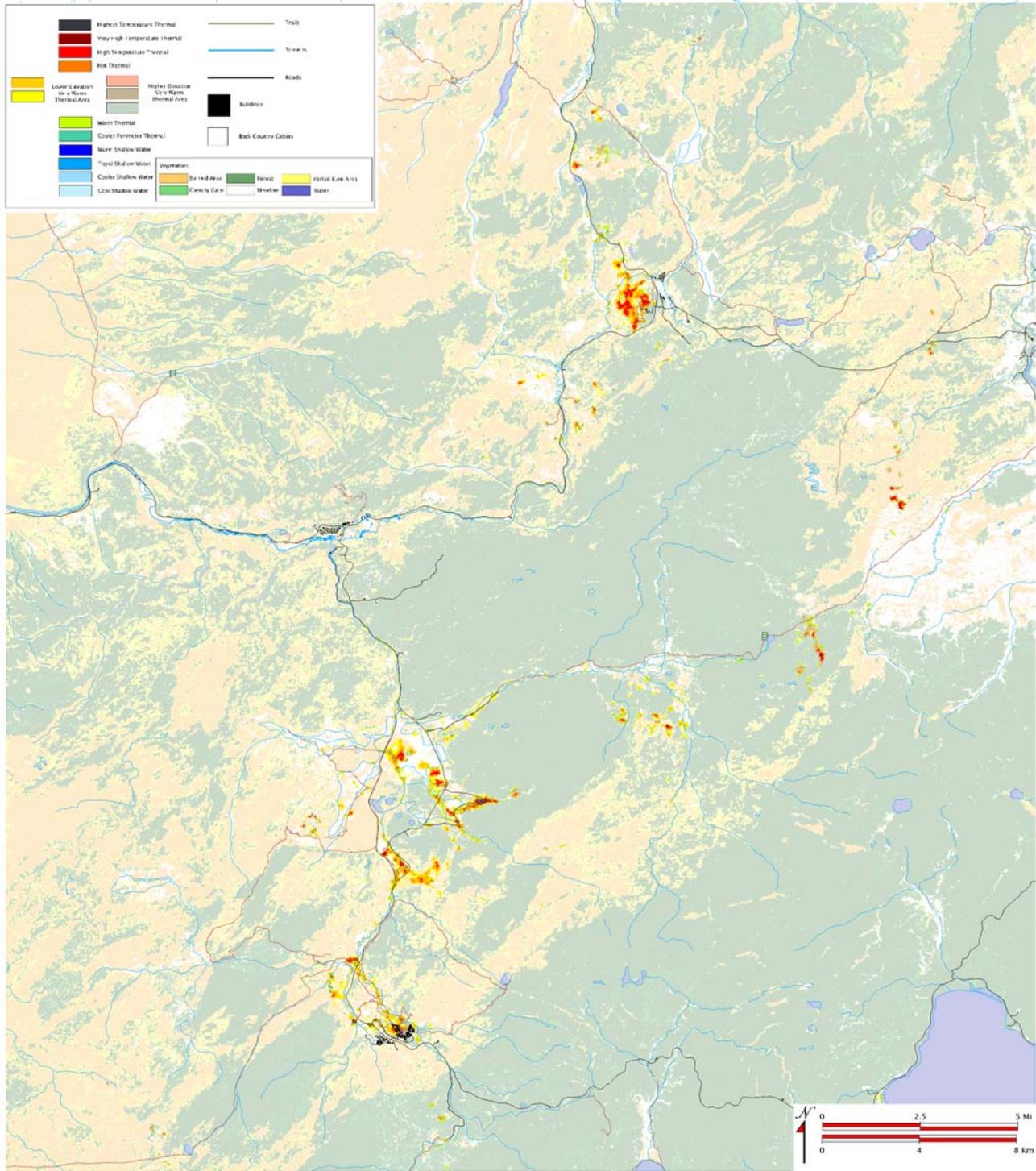
Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 3.1. Final map of geothermal areas of western central Yellowstone National Park in the year 2000, derived by classification of composite remote sensing and terrain data.**



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 3.2. Final map of geothermal areas – set against a terrain backdrop.**



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

Figure 3.3. Final map of geothermal areas – set against a backdrop of vegetation type.

### 3.2 Interpretation of explanatory variables

Additional inference about the meaning of the classes was gained by examining their 'hyper-location' in the 'hyper-space' defined by the four classification variables: Band 6L, Band 6H, Minnaert Image 6L, and the DEM. The region of hyper-space occupied by each class is illustrated in Figure 3.4, which shows three two-dimensional 'views' through the space.

The first view reveals a very high degree of correspondence between the two thermal bands, the subtleties of which are hard to discern, but are sufficient to yield a better result than if only one of them were included. The second view shows the dominant influence of the thermal bands (Band 6H on X-axis) in delineating the gradient from highest temperature (right) to warm (left). The DEM is shown as a secondary influence, essentially dividing cooler geothermal areas into three different elevation bands, and thus allowing better discrimination of cooler areas. This relates to the fact that temperature differences due to elevation in the study area (typically about 5 degrees Celsius) are similar to temperature differences amongst weakly geothermal areas. If it were not for the inclusion of the DEM, low-elevation-low-geothermal areas would not be able to be separated from high-elevation-high-geothermal areas. The third view shows the effect of the shading image. There are three included classes with a slightly higher Minnaert value, indicating sunny slopes that are geothermal, but whose level of geothermal influence is inferred to be somewhat lower than would be indicated by the thermal bands alone. Note that the 385 non-geothermal classes are represented on these diagrams as a single white circle representing their average. When these non-geothermal classes are plotted individually (not shown), the existence of non-geothermal classes with high values in both the shading image and the thermal bands indicates that the shading image provided the necessary information to correctly deem them as non-geothermal.

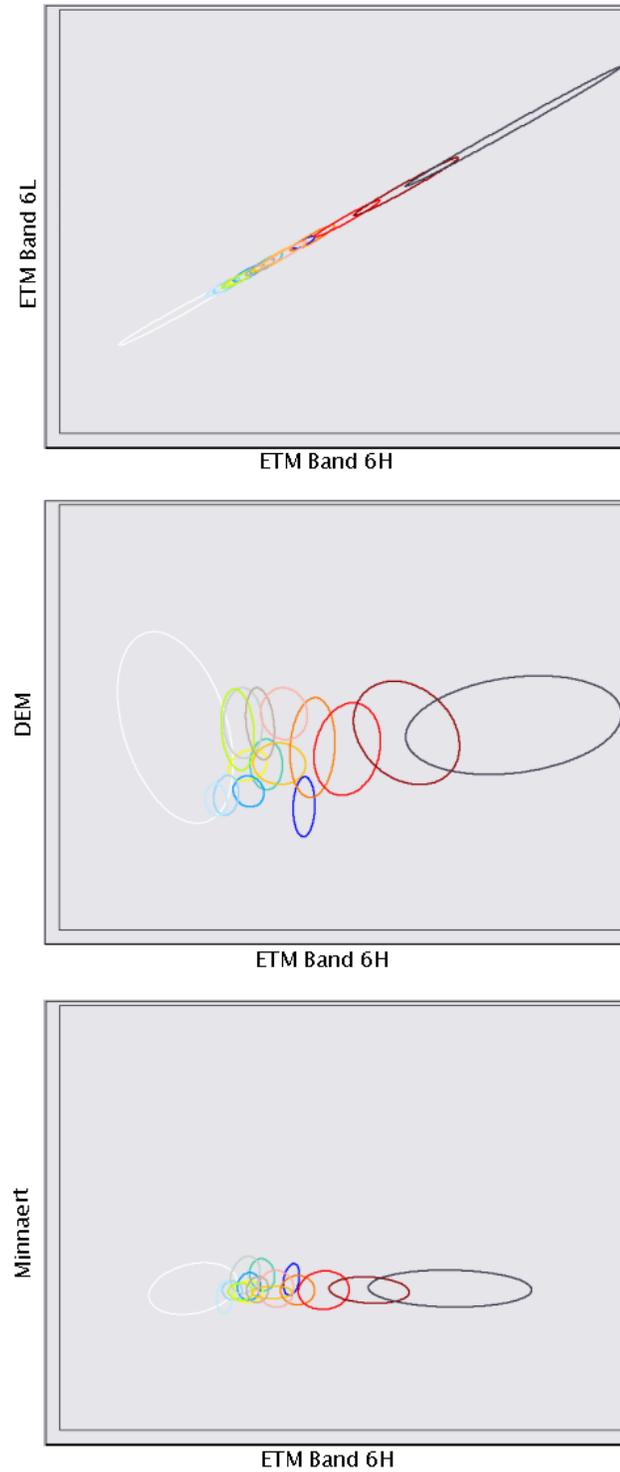


Figure 3.4. Hyper-locations within the four-dimensional classification space (6L, 6H, Minnaert, and Elevation) for each of the 15 remotely sensed classes of thermal landscape.

## 4 Evaluation

The accuracy of the final image was evaluated by comparison with a number of different types of data. A narrative ‘tour’ of the thermal image is given first, discussing its correspondence to features that are easily observed in the field. Then, comparative maps are given, showing the correspondence between the image and pre-existing digital data sets. Comparisons with a limited amount of targeted ground-truth data are also described. Finally, a set of quantitative, statistical measures of agreement with pre-existing data are designed, computed, and discussed.

### 4.1 General observations

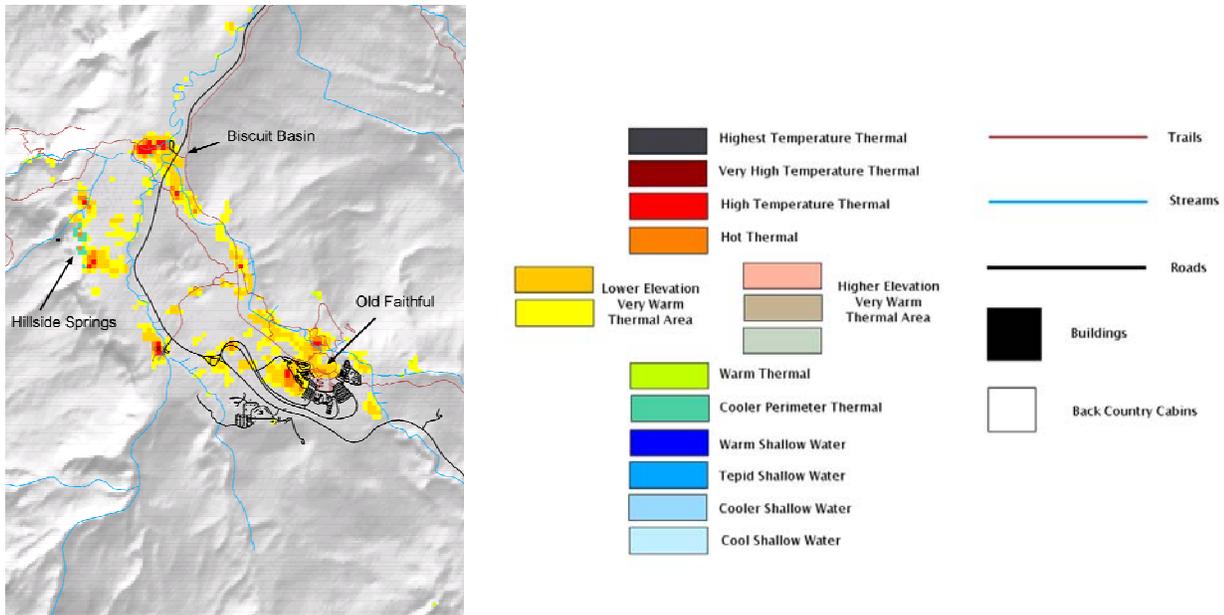
#### *4.1.1 Upper Geyser Basin*

The Upper Geyser Basin area, containing Old Faithful, is mapped in Figure 4.1. All the well-known thermal features of the basin appear in the areas mapped as thermal. This is a broad confirmation of the accuracy of the image. Old Faithful geyser itself is a relatively localized hydrothermal surrounded by a large bare apron. It appears in a small area of the image classified as ‘high temperature’ (red). The largest area of intense heat (red) in the image of the Upper Geyser Basin area occurs at Biscuit Basin, including Sapphire Pool. In the field, this area has a more extensive concentration of boiling surface water than the Old Faithful area – again a broad confirmation of accuracy. To the west, some sensitivity to more subtle thermal influence is indicated by the classification of the well-drained slopes of Hillside Springs as ‘cooler perimeter thermal’ (green).

#### *4.1.2 Midway and Lower Geyser Basins*

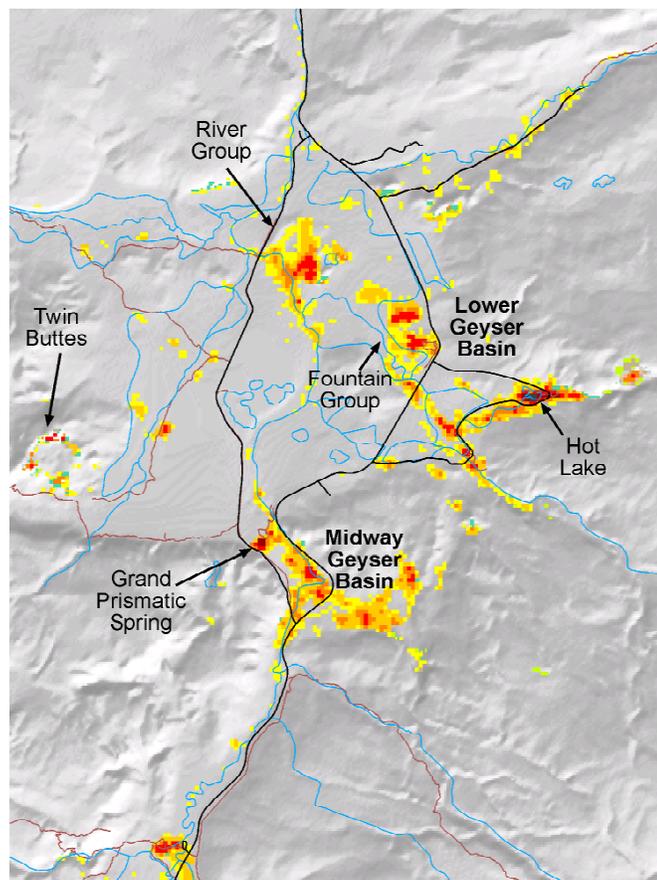
Moving north down the Firehole Valley, the Midway and Lower Geyser Basins are encountered (Fig. 4.2). Again there is broad agreement between well-known concentrations of geysers and other hot features, and the thermal image. Some features are identified as ‘highest temperature’ features (black), such as the Grand Prismatic Spring area in Midway Geyser Basin and Hot Lake in Lower Geyser Basin. The Hot Lake area is mapped in the image as having the largest expanse of highest temperature thermal influence in the study area. An interesting feature is the mapping of the Firehole River below Midway Geyser Basin in the ‘very warm’ categories. The mapping process does not attempt to

discriminate between geothermal source areas, and the areas where heat is advected downstream by hot flowing water.



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

Figure 4.1. Final thermal image of the Upper Geyser Basin area.



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

Figure 4.2. Final thermal image of the Midway and Lower Geyser Basins.

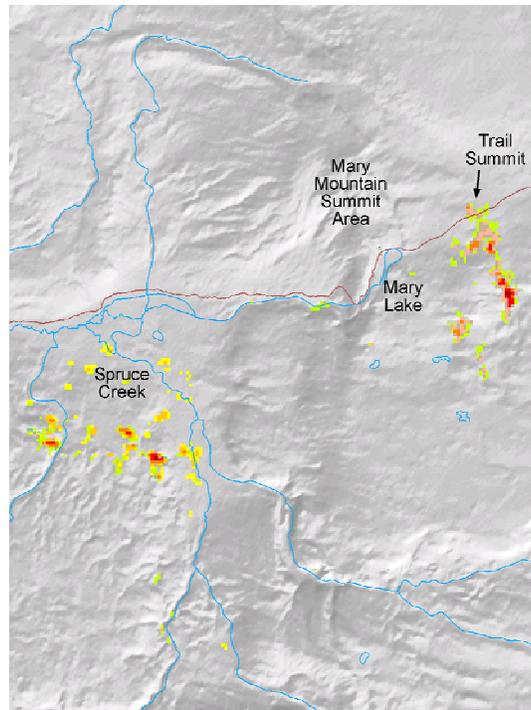
Two other large high-temperature areas are indicated at the Fountain Group and the River Group in the south and west of Fountain Flats respectively. An interesting narrow ring of heat about 700 m in diameter is indicated in the far west of this area, bordered on its northwest by the southeastern slopes of Twin Buttes ('hot'), and on its southeast by the lakes uphill of Spray Geyser ('cooler').

#### *4.1.3 Mary Mountain area*

Heading west, a number of 'very warm' areas are passed along lower Nez Perce Creek. In the field, these are mainly broad sintered areas with a few fumaroles and small pools. Half way to Mary Mountain lie the broad meadows at the confluence from Spruce Creek. The 'very warm' thermal areas here are easily missed amongst the swamps and stands of trees as one walks up to Mary Lake. Just to the south, much hotter areas are indicated that do not appear on 1980s USGS 1:24,000 topographic maps. At its summit, the trail crosses the northern top of a large thermal area (Fig. 4.3) that includes some quite localized 'very high' temperature sources, surrounded immediately by a sequence of 'high elevation, very warm', 'warm', and 'cooler' areas. At the trail crossing, these areas include sintered and bare soil, with some fumaroles. This is an area where the use of the elevation model in the classification procedure proved necessary in order to obtain a reasonable discrimination of thermal and non-thermal areas.

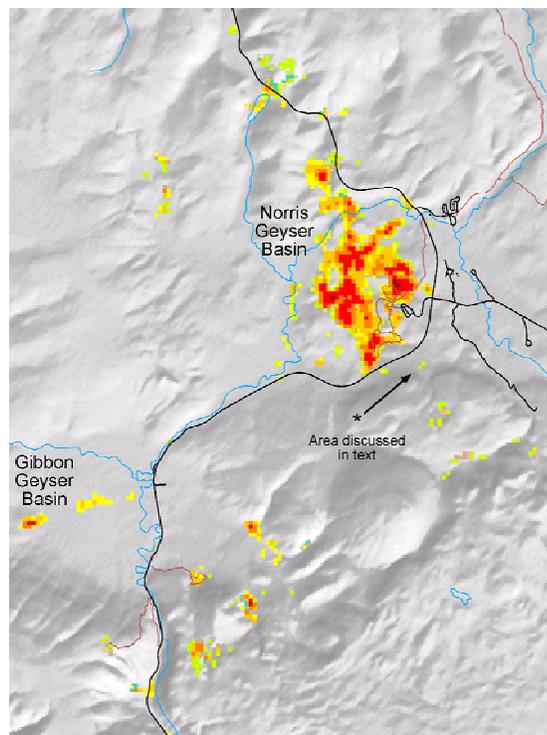
#### *4.1.4 Norris Geyser Basin*

The largest contiguous 'high temperature' area indicated by the image is the Norris Geyser Basin. From the overlook to the Basin's north, a myriad plumes of steam emanate from a mile long complex of geysers nestled amongst streams and wooded hills. Two areas of 'very high' heat are indicated. Nearby, some much more subtle thermal features are identified in the image. In contrast, the small area marked on Figure 4.4 with an asterisk (\*) is barely noticeable from a distance. Up close, it is a thinning of the unburned lodgepole pine forest with some clear patches where the snow in winter is shallow, and in places absent. The image classifies the area as containing 'warm' and 'cooler' influences. This, and other nearby 'warm' patches are known by Park workers, but do not appear on any retail maps. They are typical of the fringe features that are generally removed the road. They may be persistent, isolated features, or early indications of more prominent activity.



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.3. Final thermal image of the Mary Mountain area.**



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.4. Final thermal image of the Norris Geyser Basin area.**

#### 4.1.5 Madison Junction area

Figure 4.5 shows the thermal image of Madison Junction, where the Gibbon and Firehole Rivers combine to form the Madison River. The major thermal feature in the area is Terrace Spring, which consistently billows steam from a series of boiling pools just to the north of the road. There is also a large bare area to the south of the road. The thermal image appears to downplay the significance of the area, mapping it as either non-thermal or 'lower elevation, very warm' surrounded by extensive drainages of 'warm', 'tepid', and 'cooler' water. Under certain training classifications, a larger thermal area could be identified, but only at the expense of falsely classifying all the cliffs to the north as thermal. These cliffs, extending for about five miles on the high south-facing slopes north of the Rivers, were the most difficult false positives to remove from the image. They remain snow free for much of the winter, and their volcanic composition, heat storage characteristics, and color appear to combine to produce a significant source of longwave radiation. In the final classification, some of them remain classified as thermal, which may not be an error. Field experience in the area suggests an upslope thermal origin for some of the wet seeps observed on the lower slopes nearer the road.



**Figure 4.5. Final thermal image of the Madison Junction area.**

The Rivers themselves are mapped as being warmer than the surrounding terrain, as indicated by the large linear areas of blue on the image. Water has a high heat capacity and by its very nature, it is warmer than snow and many other winter-morning surfaces. The larger bodies of flowing water in the study area are to be found at Madison Junction, which may be sufficient to explain their prominence in the image there. However, field observations suggest that melting of riverbank and floodplain snow is more pronounced here than in other places. The possibility exists that the image quantifies the slight increase in water temperature that persists for many miles downstream of large thermal basins, and the way in which this advected heat is able to warm areas within about 100 metres of the central river channel in places where the river slows, flattens, braids and anastomoses.

## 4.2 Qualitative comparisons with other data

None of the pre-existing data measure exactly the same phenomena as the geothermal mapping presented here. Each differs in its detail, structure, underlying aims, properties mapped, or applicable dates. Therefore, the aim in comparison between the present image and pre-existing data is not to achieve a 100% match. Rather, the pre-existing data provide various types of information on the distribution of thermal influence in the Park that can be used, in an informed manner, to measure various aspects of the accuracy of the present image.

### *4.2.1 Sensitivity of present image*

We would expect at the outset that the spatial extent and intensity of heat flux would be best mapped by using the present image with one major caveat: small, localized geothermal areas are excluded. The resolution of the Landsat-7 Enhanced Thematic Mapper used here is about 60 metres, or about 3600 m<sup>2</sup> in areal terms. If we assume that the radiometric sensitivity of the instrument over a large area is about 1 degree Celsius, then for every factor that a localized geothermal is smaller than a 3600 m<sup>2</sup> pixel, it would need to be that many times hotter in order to be detectable against a uniformly cooler background. A lone pool of water boiling at 100 °C would need to be at least 6 m x 6 m, which is quite large. A small super-heated vent would be missed.

### *4.2.2 Pre-existing polygon data: Hutchinson, and Shovic*

Hutchinson's polygons were intended to map the spatial extent of thermal areas, using visual interpretation of aerial photography backed by field experience. In many cases, the data are expected to match the present heat-based image, with the following exceptions. Firstly, the shape of polygons is smoothed as an approximate boundary around fragmented visual features. Secondly, the visual appearance of thermal areas sometimes reflects previous thermal activity that may have subsided.

Shovic's 'hydrothermal' polygons form part of a larger analysis of 'landforms', and were interpreted for field and air-photo data as geomorphic features as much as geothermal features. They are expected to be less accurate indications

of thermally affected area than the more targeted work by Hutchinson. They were also outlined some years after the Hutchinson effort, and so may differ where new thermal activity is being revealed.

#### *4.2.3 Pre-existing point data: MSU elk, and NPS Thermal Inventory*

The MSU elk-tracking (MSUET) data are unique. They include both thermal and non-thermal locations – therefore offering a test of the negative thermal mapping accuracy of the image as well as the positive ability. They are also selected from the landscape by elk, with no regard to where humans might otherwise go, or look. At the same time, most individuals of the resident elk population spend their entire lives in an area only a few miles in breadth. Maternal groups can be considered to have discovered all thermal areas whose habitat advantage outweighs the locomotive effort in getting there. Uniquely, even though the locations of all 10,238 of these points were determined by elk, they were measured and described by humans in the field, and given attributes of ‘thermal’ or otherwise.

There may be some limitation in the accuracy of locations in the MSU database, which were all obtained by manual map reading in the field. The program started in 1991, well before reliable portable GPS, and at present only makes occasional use of GPS for checking purposes. In uniform, featureless terrain, location error is thought to be within 50 to 100 m at worst.

The NPS Thermal Inventory (NPSTI) database is much more recent, entirely using GPS locations established after the recent military de-encryption of the GPS signal. Locations are thus accurate to about 10 m. The database contains a larger number of thermal points, 4,127 as opposed to 1,421 in the MSU database, but there are no non-thermal points. Many of the points have useful attribute information, such as the temperature of a pool or lake. We have queried this information manually from time to time, but a comprehensive quantitative scheme for making use of these attributes in ground-truthing remote sensing imagery has not yet been devised.

Neither of the point-feature databases gives any quantitative indication of the spatial extent or perimeter of thermal features, other than where, for some reason, there may be numerous points within a single contiguous thermal area.

#### *4.2.4 Visual comparison: all large data sets overlaid*

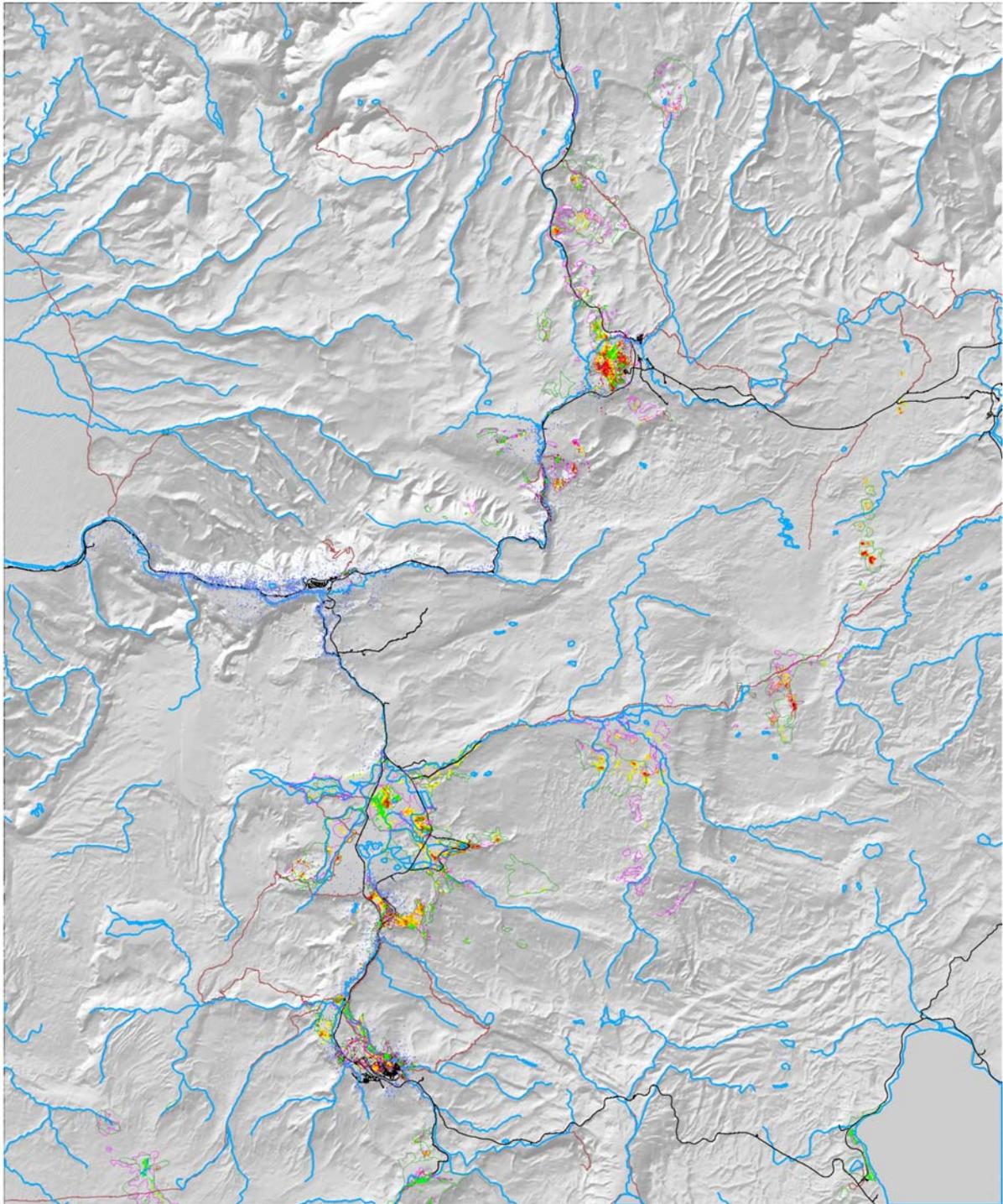
All pre-existing digital data are overlaid on the present image in Figure 4.6 and various areas of interest are highlighted in Figures 4.8 to 4.11. The Firehole Valley overlay (Figure 4.8) shows that, of the two polygon data sets, the Hutchinson data (pink) are more detailed, and match the thermal image more closely than the broad outlines of the Shovic data (green). Both sets of polygons are usually simpler in shape than the image data, and typically cover a slightly larger area. There are many areas where Hutchinson indicates thermal activity, but the present image does not. This could be indicative of a lack of sensitivity in the image, or of an over-statement of thermal influence by Hutchinson in the broad flats with continuous white soil. There are very few areas where the thermal image indicates thermal activity but the polygon data sets do not – false positives – suggesting that the image does not over-state thermal influence. Almost all the NPS Thermal Inventory points fall in areas mapped as ‘very warm’ and above. There are very few false negatives in the comparison with this data set. The MSU Elk Tracking points reveal many false negatives however, as would be expected given that this data set records numerous small, isolated thermal features away from the large Inventory areas. By and large however, there are far more true negatives indicated by the MSUET data set than false negatives. Indeed, in areas with high elk visitation, the boundary between thermal and non-thermal areas in the image closely corresponds to the boundary between MSU thermal points and MSU non-thermal points.

To the east, at Spruce Creek and Mary Mountain (Fig. 4.9), one of the largest areas of discrepancy occurs. Before 1988, Hutchinson mapped a field of geothermal patches (pink) several miles long extending south from Nez Perce Creek in the Spruce Creek area. Nearly a decade later, Shovic listed only one broad area, that disagreed almost exclusively with the Hutchinson data. Almost another decade later, the present image reveals a complex of very warm to hot thermal areas in the vicinity of Shovic’s polygon, but also extending beyond it toward Hutchinson’s polygons. The dominant cause of this disagreement may be real change in thermal activity over a large area. Further investigation is warranted. Further east past Mary Mountain, the thermal image matches Hutchinson’s polygons very closely, with Shovic’s polygons forming an all-encompassing outline. This is a typical pattern whereby the polygon data enclose a larger area that appears to be due to the mechanics of air-photo

interpretation and digitization at a particular scale. At the time of publication, no NPS or MSU point data has been collected in this area.

To the north, at Norris and Gibbon Geyser Basins (Fig. 4.10), similar comparisons are evident. The polygon data form broad outlines that generally agree with the thermal image. The NPS points fall in the most intense thermal areas, and the MSU points confirm a predominantly correct thermal/non-thermal boundary. Again, the MSU points reveal numerous false negatives associated with small, isolated thermal areas that cannot be revealed using Landsat data. Some evidence for change over time is also indicated. For example, the road just southwest of Norris Geyser Basin has recently been undermined due to new thermal activity. In this area, the Hutchinson data are absent, the more recent Shovic data show a large polygon, the MSUET and NPSTI data show a number of thermal points, and the thermal image shows a complex of 'warm' to 'very warm' thermal patches.

Down the valley at Madison Junction (Fig. 4.11), a much more subtly thermal environment is revealed. The major thermal area in the region is at Terrace Spring, to the right of the Figure. Here, Hutchinson and Shovic agree on a large apron of de-vegetated land south of the Springs. However, there are two non-thermal MSU points in the middle of this apron, and the thermal image detects very little heat. This may be due to change over time, or an over-estimation in the polygon data. All data sets agree about the location and extent of the intense thermal activity at the Springs themselves. To the south and west, the thermal image records most of the surface water of the three rivers (Madison, Firehole, and Gibbon) as being thermally influenced. Longitudinal thermometer data (not presented) agree with this assertion, and occasionally so too do the MSU Elk Tracking points, particularly in the areas where the thermal image indicates slightly warmer influence. The south-facing cliffs immediately north of the Madison River were problematic to map due to false positives created by solar heating. A particularly encouraging observation is that the very few pixels along these cliffs that remain classified as thermal are close to some isolated MSU thermal points (see area marked on map). This occurs in high gullies opposite Mt Haynes where some steep thermal seeps are 'discovered' by elk and their trackers from time to time. There are also some characteristic 'small thermal' false negatives indicated by the MSU data in isolated backcountry areas.



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.6. Comparison between the geothermal image and pre-existing digital data – whole study area – see later Figures for close-ups (legend given in Figure 4.7).**

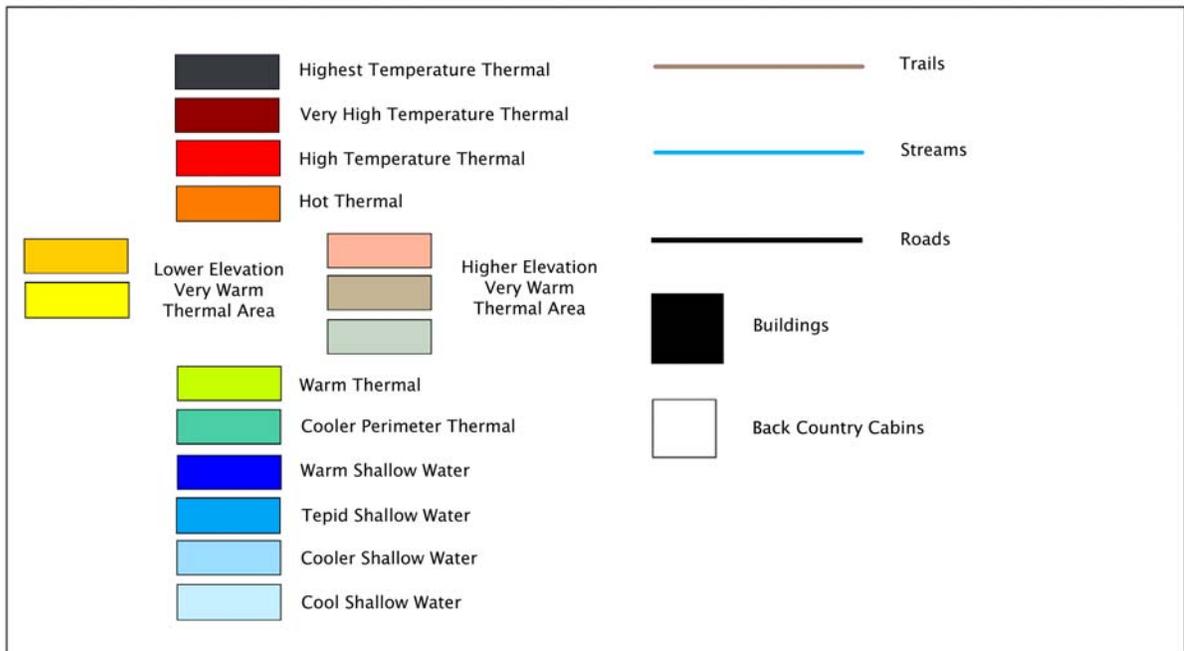
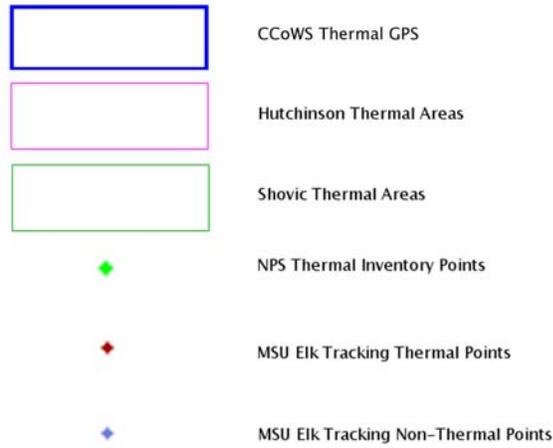
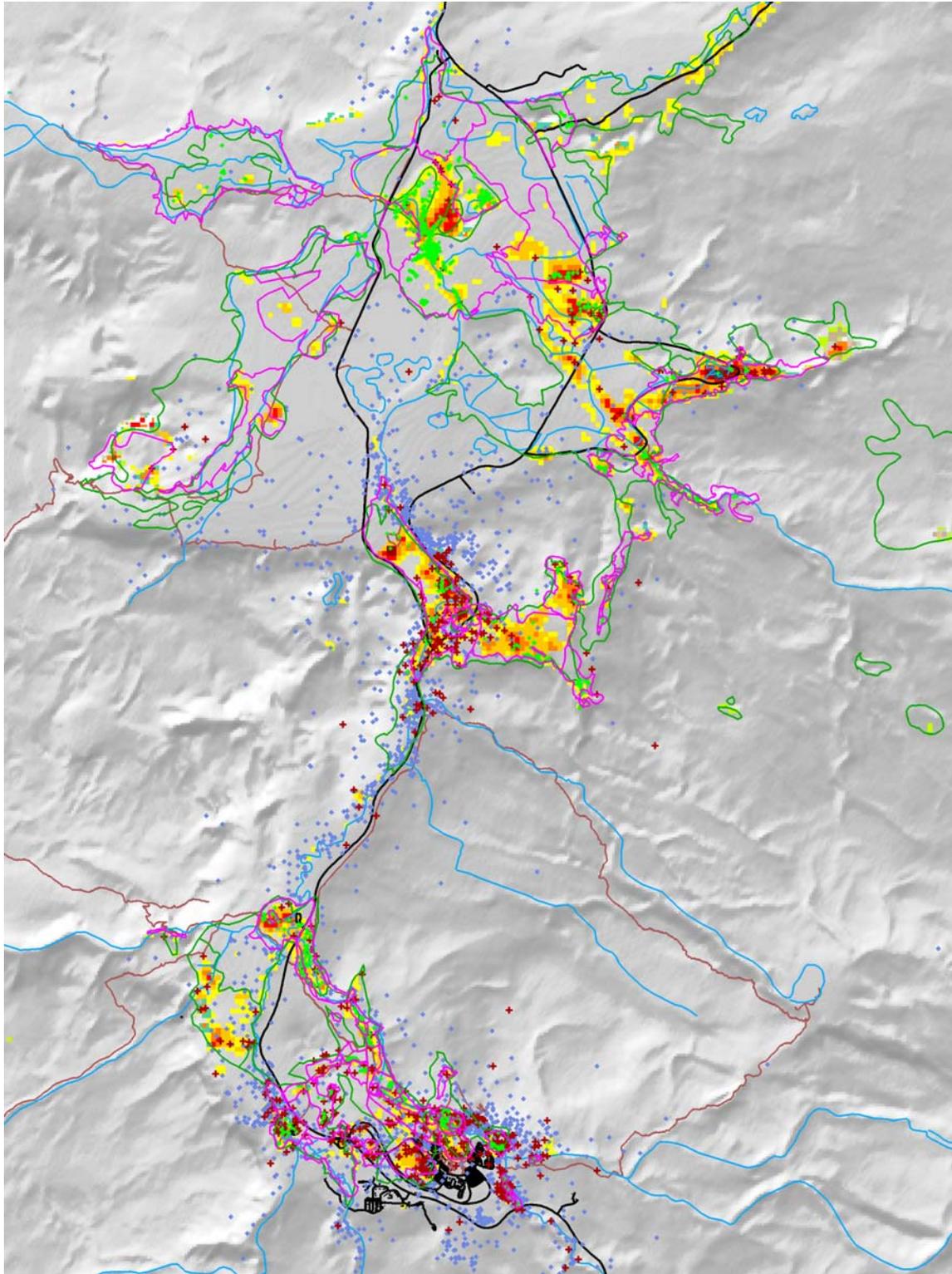
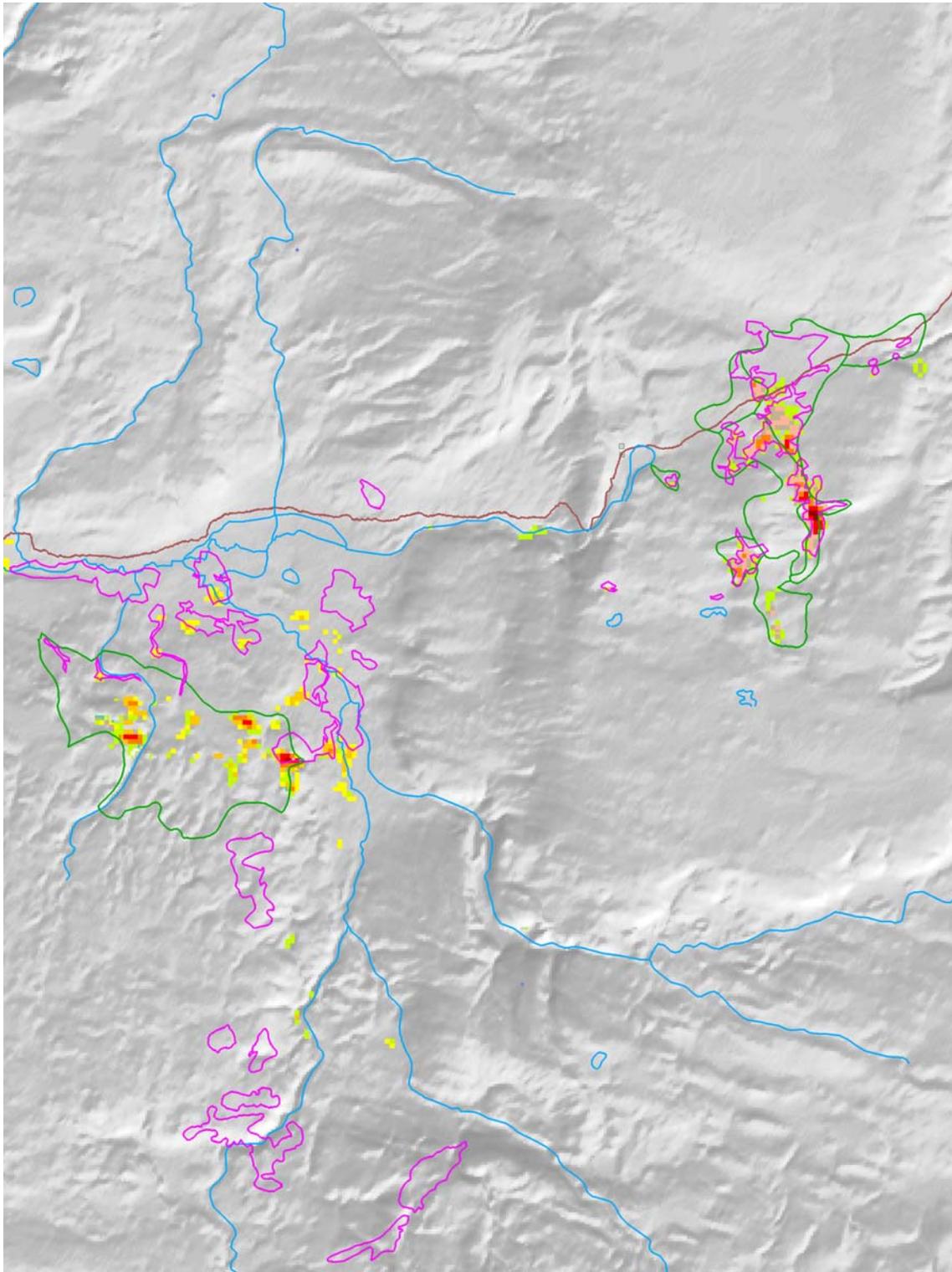


Figure 4.7. Legend for maps overlaying pre-existing data with thermal image.



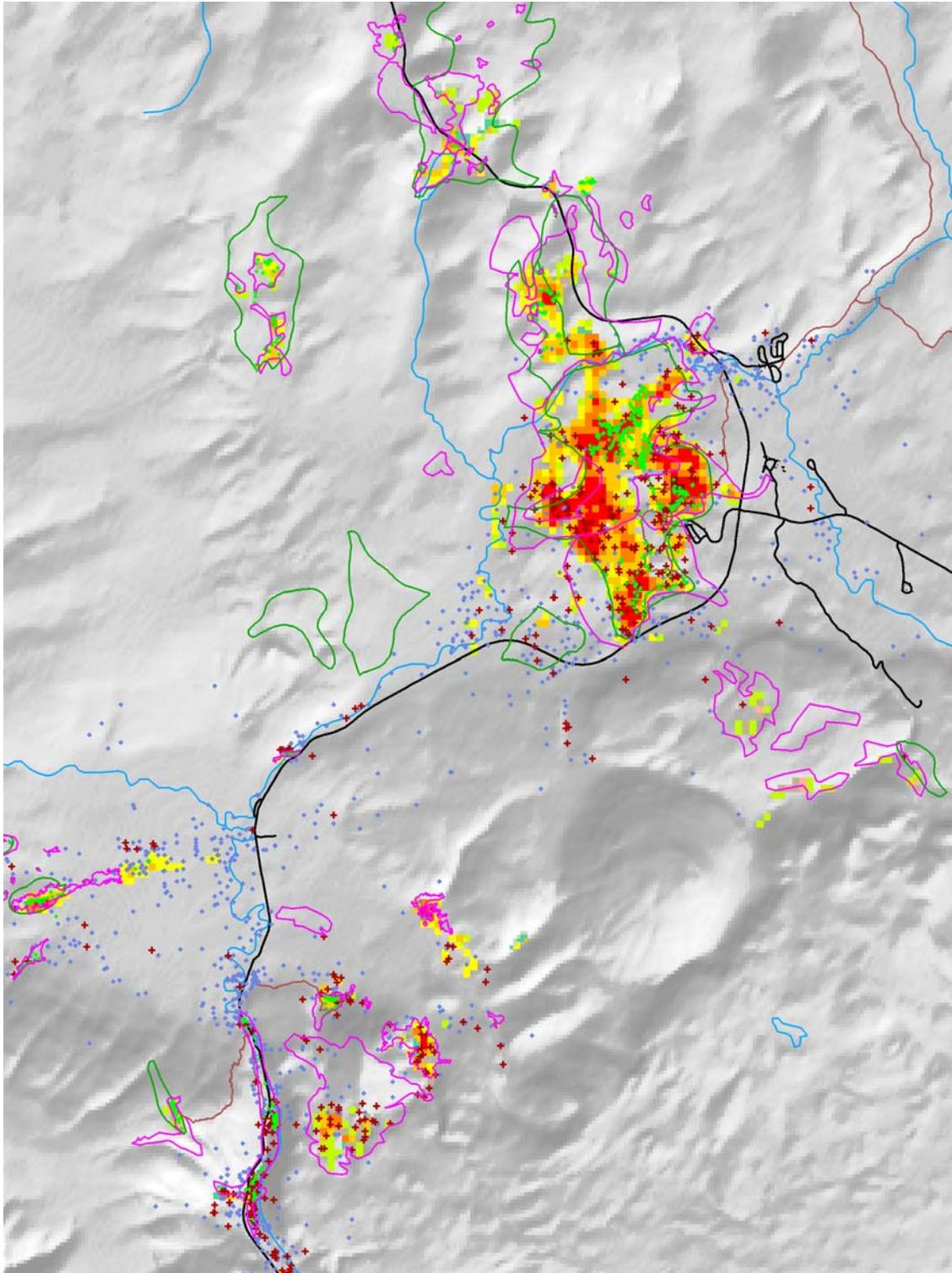
Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.8. Comparison between the geothermal image and pre-existing digital data – Firehole valley.**



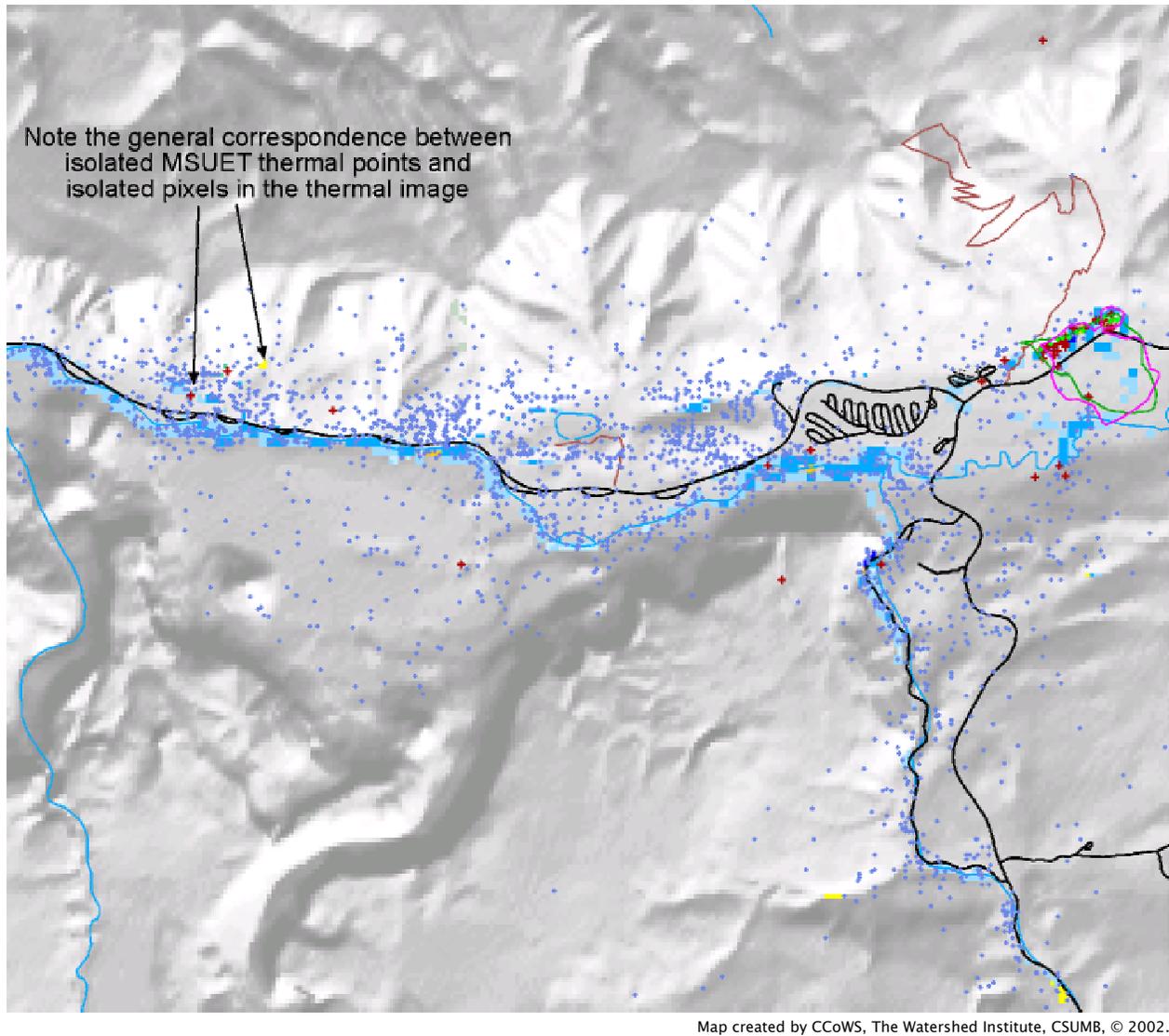
Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.9. Comparison between the geothermal image and pre-existing digital data – Spruce Creek / Mary Mountain area.**



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.10. Comparison between the geothermal image and pre-existing digital data – Norris/Gibbon area.**



**Figure 4.11. Comparison between the geothermal image and pre-existing digital data – Madison Junction area.**

#### *4.2.5 Purposely acquired polygon data: CSUMB*

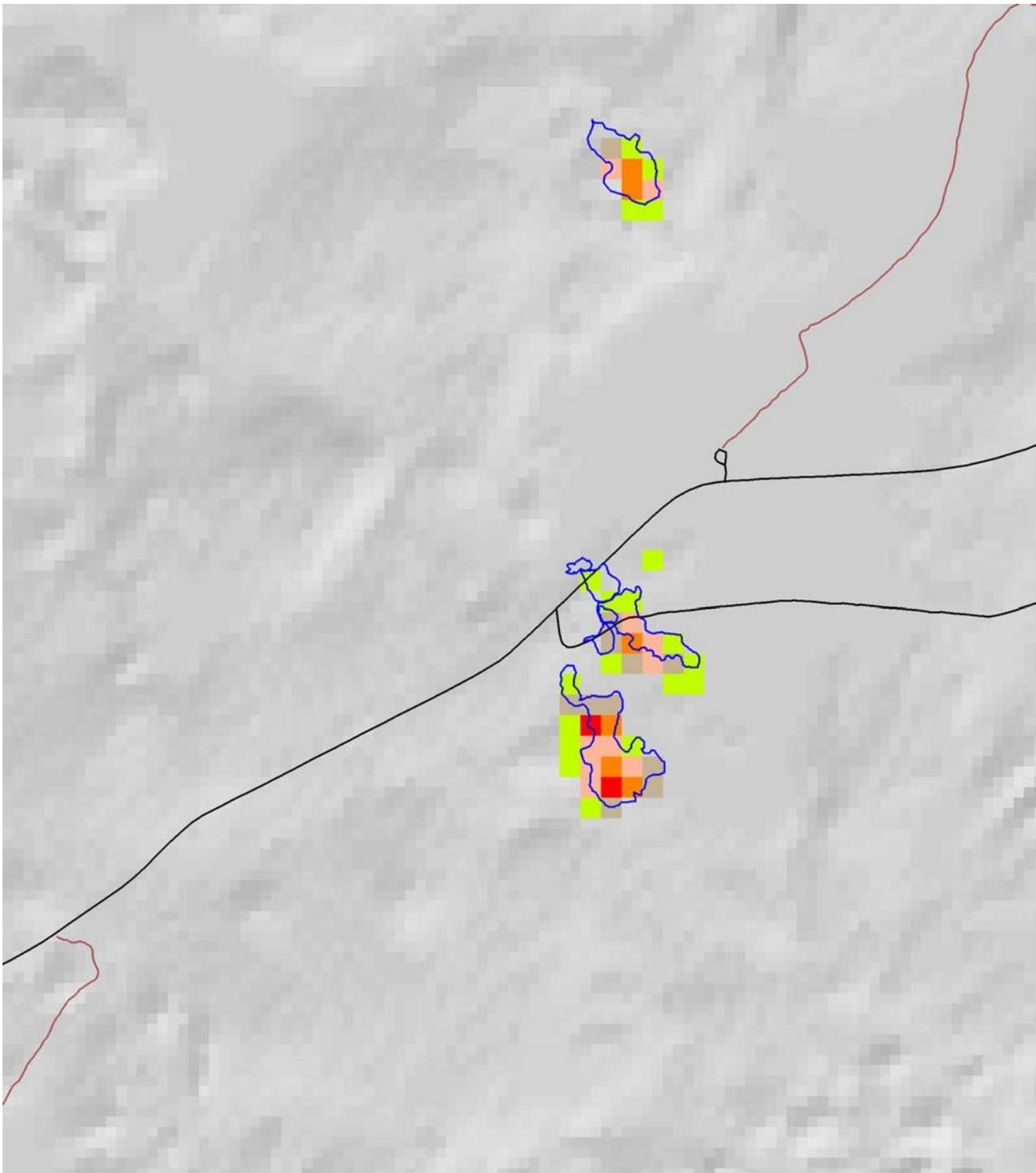
None of the pre-existing data sets can be considered a direct ground-based measurement of exactly what is remotely sensed in the thermal image. Therefore, some specific, targeted ground-truthing was performed. Two examples are discussed. In the first example, the penultimate thermal image indicated three thermal areas that were absent from all pre-existing data sources available to the authors<sup>1</sup> – and these areas were visited and measured in early Spring. In the second example, aerial photography indicated a thermal area that was absent from the penultimate thermal image (A. Messer, pers. comm.) – again, the area was visited and measured, this time in Summer.

In all targeted ground-truthing efforts, the main aim was to walk the perimeter of thermally influenced terrain, using a GPS to create a perimeter polygon. Attributes of any point features within these polygons can be obtained from the NPSTI program, which currently uses the present thermal image for exploration. Various protocols were devised for recognizing the boundary of ‘thermally influenced terrain’. In the Spring example, the snow/no-snow boundary was used in a high-elevation area that was otherwise uniformly snowbound in burned lodgepole pine forest. In the Summer example, two perimeters were walked, one on the bare soil boundary, and the other on the full-canopy-height boundary. The area was otherwise uniformly forested with old-growth lodgepole pine.

The results of both targeted ground-truthing exercises are shown in Figures 4.12 and 4.13. Allowing for slight shifts due to systematic error in the terrain-correction of Landsat imagery, the polygons overlap the corresponding areas in the thermal image very well. The area and shape are very closely matched within the limits of precision of the cell-based thermal imagery. This is an encouraging confirmation that the thermal image accurately measures not just the location and intensity of thermal influence, but also the total spatial extent of that influence as experienced at the ground surface.

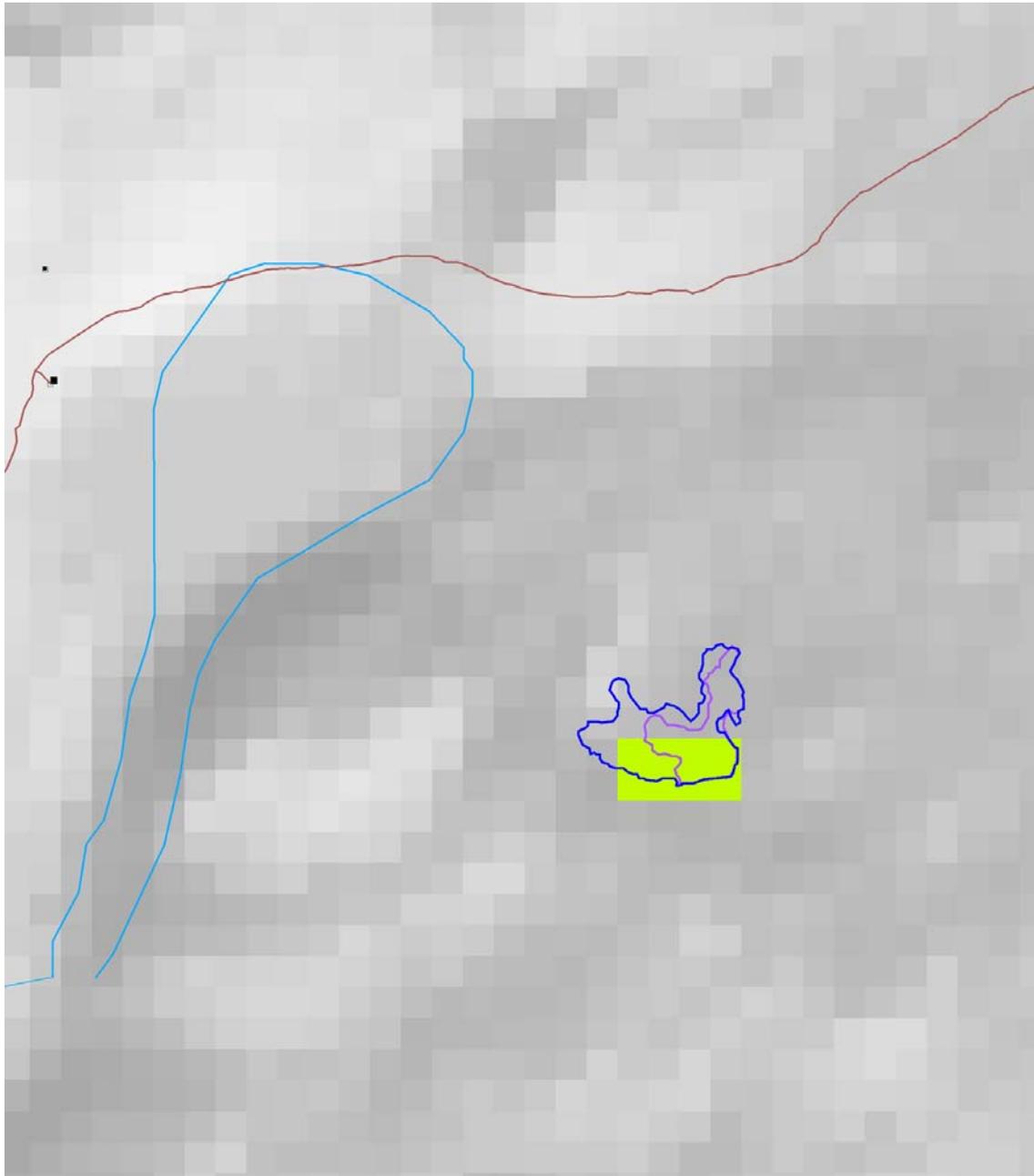
---

<sup>1</sup> Note that the full Hutchinson data set may include these features, but they lay beyond the extent of the Hutchinson data that was available to the authors at the time.



Maps created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.12. Targeted ground-truthing of the thermal image: field-surveyed snow/non-snow perimeters of three areas in the thermal image otherwise not indicated to the authors by pre-existing data. Location is Norris Canyon Road at the eastern edge of Solfatara Plateau.**



Map created by CCoWS, The Watershed Institute, CSUMB, © 2002.

**Figure 4.13. Targeted ground-truthing of the thermal image: field-surveyed perimeters of a thermal area absent from the penultimate thermal image, but indicated by aerial photography: dark blue line = perimeter determined by canopy boundary; purple line = perimeter determined by bare-soil boundary.**

### 4.3 Quantitative comparisons with other data

#### 4.3.1 Methods

The degree of coincidence of the thermal image with each of the pre-existing data sets was computed. In the case of polygon data, the polygons were first rasterized so as to have the same grid-based spatial structure as the thermal image. Then, the number of cells in agreement with the thermal image was tallied. For the point data, the number of points in agreement with the thermal image was tallied. For all but the NPSTI data, four outcomes are possible from any individual comparison:

- True positive – where both the pre-existing data and the thermal image indicate a thermal area
- True negative – where both sources indicate a non-thermal area
- False positive – where the thermal image indicates thermal, but the pre-existing data indicate non-thermal
- False negative – where the thermal image indicates non-thermal, but the pre-existing data indicate thermal

For the NPSTI data, only true positives and false negatives are possible, as there are no points in that data set that indicate non-thermal areas.

A simple measure of agreement is the ratio of true positives (TP) to the total number of pre-existing data positives (DP):

$$\text{Positive agreement } (PA) = \frac{TP}{DP}, \quad 0\% < PA < 100\%$$

A problem with this measure is that it can be made 100% simply by declaring the entire study area as being thermal, a false impression of accuracy.

It is thus useful to measure agreement on true negatives (TN) as well as positives, relative to the number of data negatives (DN):

$$\text{Negative agreement } (NA) = \frac{TN}{DN}, \quad 0\% < NA < 100\%$$

A simple integrated measure could be computed by averaging the two ratios:

$$\text{Mean agreement (MA)} = \frac{\frac{TP}{DP} + \frac{TN}{DN}}{2}, \quad 0\% < MA < 100\%$$

A property of this measure is that it gives equal weight to agreements based on positive or negative data, regardless of how much of each type of data are available. For example, if only 10% of the data points were negative, each one would count for nine times more information than a positive point.

A separate ratio may be computed that gives equal weight to the information contributed by each pre-existing data point:

$$\text{Combined agreement (CA)} = \frac{TP + TN}{DP + DN}, \quad 0\% < CA < 100\%$$

This measure is biased in a different manner. Because geothermals make up such a small fraction of the landscape, very high total agreements can be obtained simply by correctly estimating a non-thermal status for vast areas of the landscape that are easily determined as such. The measure de-emphasizes our interest in correctly estimating the spatial extent of thermals in particular, rather than correctly identifying the attributes of any cell in the landscape.

A third integrated measure focuses more on the agreement across all areas indicated as thermal in either of the two data sources (pre-existing data or thermal image), and ignores information about areas that neither data source indicates as thermal. We term it the ‘positive agreement disagreement index’ (PADI):

$$PADI = \frac{TP}{TP + FP + FN}, \quad 0\% < PADI < 100\%$$

where false positives, FP, are equal to:

$$FP = DN - TN, \quad 0\% < FP < 100\%$$

and false negatives, FN, are:

$$FN = DP - TP, \quad 0\% < FN < 100\%$$

A PADI value of 100% represents total agreement (no disagreement). A value of 0% represents no agreement about thermal areas (saying nothing with respect to possible agreement about non-thermal areas). A value of 50% means the area agreed as thermal equals the area disagreed as thermal (i.e. either the pre-existing data indicated thermal, or the thermal image, but not both).

The residuals of PADI values below 100% are comprised of either normalized false positives (NFP) or normalized false negatives (NFN):

$$NFP = \frac{FP}{TP + FP + FN}, \quad 0\% < NFP < 100\%$$

$$NFN = \frac{FN}{TP + FP + FN}, \quad 0\% < NFN < 100\%$$

$$PADI + NFP + NFN = 100\%$$

It should be kept in mind when interpreting language used here such as ‘true’ and ‘false’, that the pre-existing data are themselves fallible, and any disagreement about ‘truth’ may be due to error in either the thermal image or the pre-existing data. Further, the fallibility of pre-existing data may be due to errors of method and interpretation, or due to spatial location errors in the digitizing process, or due to changes in the landscape over time.

The PADI index is the most suitable integrated measure of agreement for comparison with polygon data, and other coverages that completely map the study area, but it cannot be used in comparison with point data sets. This is because point data sets only give values where points have been surveyed, whereas the thermal image asserts properties of all cells, regardless of their inclusion in the point data set. A misleadingly large number of false positives would be expected.

An analog of PADI for point data sets might thus remove the FP term from the PADI equation. This gives:

$$\frac{TP}{TP + FN} = \frac{TP}{DP} = PA$$

and we are back to the positive agreement, PA, defined earlier.

#### 4.3.2 Results

The above statistics are summarized for comparisons with all pre-existing data sets in Table 4.1. Positive statistics are given for all comparator data sets. Negative and compound statistics are only given where applicable. Of all the areas indicated as thermal by Hutchinson and Shovic, the thermal image only agrees with 28.0% and 19.4% of them respectively. However, agreement on non-thermal areas is much higher, at 99.7% for both polygon data sets. This is not unexpected, as the earlier qualitative analysis suggested that polygons are generally drawn by human interpreters fairly loosely around the outer perimeter of geothermal areas. This leads to numerous false negatives. The correspondence of thermal/non-thermal boundaries in the field-interpreted MSU point data sets gave a good indication that these false negatives were not generally due to wholesale under-estimation of thermal areas in the thermal image. The low positive agreement with polygon data sets propagates to low mean agreement, but is diluted in the combined agreement statistics (98.5% and 97.5%) by virtue of the fact that most of the study area is easily mapped as non-thermal. This high value should not be under-valued. It means that for any given location in the 2167 km<sup>2</sup> study area, Landsat remote sensing is only wrong about its 'thermality' 1.5% to 2.5% of the time.

The sources of error in predicting thermal areas themselves are clarified by the PADI statistics. Of all areas declared thermal by either data set in a given comparison, agreement is only reached 23.2% and 17.8% of the time respectively. The associated normalized false positive (NFP) percentages are fairly low, but the normalized false negatives (NFN) are very high – 59.7% and 74% respectively. Again, the disagreement is largely due to large polygons being drawn around patchy thermal areas, and thus most likely over-estimating the total extent of thermal influence.

Considerably better positive statistics are achieved with the MSU Elk Tracking points, which have a positive agreement of 55.7%. Given the field-interpreted nature of these data, the inference is drawn that the thermal image matches

them better than it matches the polygon data because the polygon data are not good estimators of the extent of geothermal influence. Negative statistics, however, are poor under comparison with the MSU points. This is also expected, being due to the numerous small, isolated thermals discovered by elk, trackers, but not Landsat.

Slightly higher positive statistics are achieved using the NPS Thermal Inventory points. Again, this is expected because the NPSTI points are biased toward intensely thermal areas that are rarely missed in the remote sensing analysis.

The best positive statistics, 92.4%, arise from comparisons with the only data set targeted specifically at measuring the extent of surface heat flux – the targeted GPS perimeters. The comparison with these data has a high PADI index, 55.0%, indicating good general agreement about specifically thermal areas, and a high NFP index, indicating that most of the disagreement about these areas is due to false positives – the thermal image predicting a larger thermal area than was measured by the field perimeters.

Pre-existing data											
Source	Content	Date	Data positives (DP)	Data negatives (DN)	PA: (TP/DP)	NA: (TN/DN)	MA: (PA+NA)/2	CA: (TP+TN)/(DP+DN)	PADI: TP/TP+FP+FN)	NFP: FP/TP+FP+FN)	NFN: FN/(TP+FP+FN)
Hutchinson thermals	263 polygons	pre-1988	43,266 cells	2,624,103 cells	28.0%	99.7%	63.8%	98.5%	23.2%	17.1%	59.7%
Shovic landforms	83 polygons	1996	74,216 cells	2,593,153 cells	19.4%	99.7%	59.6%	97.5%	17.8%	8.3%	74.0%
MSU elk tracking	10,238 points	1991-2002	1,421 points	8,817 points	55.7%	82.2%	69.0%	78.7%	-	-	-
NPS Thermal Inventory	4,127 points	2001-2002	4,127 points	0 points	63.6%	-	-	-	-	-	-
Targeted GPS perimeters	8 polygons	2002	131 cells	-	92.4%	-	-	-	55.0%	40.5%	4.5%

**Table 4.1. Agreement statistics between pre-existing data and the thermal image.**

### 4.3.3 Sensitivity to spatial error

Improvements to certain evaluation statistics are achieved by making some allowance for geo-referencing errors between the various data sets. Polygon data are generally hand digitized, and follow smooth boundaries around visually complex features. Point data located using both topographic maps and GPS contain a certain amount of location error. The thermal image data are subject to satellite parallax error. It is thus likely that numerous data-positives lie just outside the thermal areas in the present image because of geo-reference error, rather than error in the measurement of thermal extent and intensity *per se*.

To test this, we examined the change in evaluation statistics associated with successive enlargements of the geothermal areas. Each enlargement involved a uniform 'growth' of all thermal areas by one pixel (28.5 m), up to a maximum growth of 10 pixels (285 m). The fraction of total study area declared 'thermal' grew from 0.79% to 1.45% after a single enlargement, and to 7.37% after 10 enlargements.

Although only 28.0% of the data-positives in the Hutchinson data were matched by positives in the final thermal image, the percentage grows to 82.3% after ten enlargements. This confirms the suggestion that numerous false positives lie just outside the extent of the thermal image, and supports the notion that these may be 'false' not because of thermal mapping error, but because of geo-referencing and digitizing differences. Of course, there is a concomitant decrease in true negatives from 99.7% down to 93.9%. In balance, the mean agreement rises with thermal enlargement, but the combined agreement falls, due to its weighting by total numbers of data points, these being generally non-thermal. The PADI index also falls, conveying that for every false negative removed by expanding the mapped thermal area, there are a greater number of false positives included. Two conclusions may be reached. Firstly, enlargement does not lead to a better correspondence with pre-existing data. But secondly, the improvement in positive agreement with enlargement suggests that a significant portion of the lack of correspondence may be attributed to geo-referencing differences rather than mapping error.

Similar analysis using comparisons with the Shovic polygon data leads to the same conclusions.

Point data statistics can be similarly improved by enlargement of the thermal areas. The positive agreement with the NPSTI data set rises from 63.6% to 90.6% after four enlargements. In other words, the majority of false negatives can be removed (converted to true positives) by moving them 100 m away from their (possibly erroneously) mapped location.

The same degree of enlargement improves the positive agreement with the MSUET data set from 55.7% to 85.1%, although this reduces the negative agreement from 82.2% to 45.1%. Averaging the two gives a net decline in mean agreement from 69.0% to 65.1% over four enlargements. However, an optimal mean agreement of 70.1% is obtained after the first enlargement. Weighting by data points rather than thermals versus non-thermals, leads to a net decline in combined agreement from 78.7% to 50.5%, with no optimum along the way.

The conclusion from the point data comparisons is thus similar to that for the polygon data comparisons – that a numerical experiment involving enlargement of thermal areas highlights the potential importance of geo-referencing error over strict thermal mapping error. The fact that enlargement does not actually produce a better image (more true positives *and* true negatives) counters any suggestion that the statistical improvements after enlargement are because the original thermal image under-estimated total thermal area.

## 5 Inferring heat flux

The techniques presented are potential pre-cursors to a remote sensing estimate of geothermal heat flux from the Park. The following calculations demonstrate this potential, based on some simple assumptions about surface energy balance.

The energy balance of a point on the Earth's surface involves solar and longwave radiation exchanges with the atmosphere, ground heat flux conducted through the soil, heat flux advected with precipitation and groundwater movement, latent heat flux associated with evaporation, and sensible heat flux between the surface and the atmosphere. In a non-geothermal system, sub-surface heat flux terms are often ignored, and inputs from the sun are generally balanced in the long term by outputs of sensible, latent, and longwave radiant heat flux. In a geothermal system, there is a significant net input of heat from the ground, which is balance by increased losses to the atmosphere.

Assume, for a moment, that the Park's land surface and atmosphere is uniformly at 0 °C on a snowbound winter's morning, except for the surfaces of geothermal areas, which are at 1 °C. Assuming all surfaces have an emissivity ( $\epsilon$ ) of 1, the geothermal surfaces will emit more radiation, as calculated by the Stefan-Boltzmann equation:

$$\text{Difference in longwave emission} = \epsilon\sigma(274.15^4 - 273.14^4) = 4.6 \text{ Wm}^{-2}$$

Where  $\sigma=5.67 \times 10^{-8}$  is the Stefan-Boltzman constant, and 273.15 is the Kelvin equivalent of 0 °C.

The sensible heat flux between the geothermal surface and the atmosphere depends strongly on the aerodynamic resistance,  $r_a$ , of the atmospheric boundary layer:

$$\text{Sensible heat flux} = H = \rho_a c_{p,a} \frac{(T_{surf} - T_{air})}{r_a}$$

where  $\rho_a$  is the density of air ( $\sim 1 \text{ kg/m}^3$  in the Park),  $c_{p,a}$  is the specific heat of air (1003 J/kg/K), and  $T_{surf}$  and  $T_{air}$  are the respective temperatures (K). Aerodynamic resistance is determined by wind speed, and the roughness characteristics of the surface. Typical values of  $r_a$  are between 10 and 100 s/m,

which lead to sensible heat flux estimates of  $100 \text{ W/m}^2$  and  $10 \text{ W/m}^2$  respectively.

For each degree the geothermals are hotter than the atmosphere and the surrounding terrain, they therefore contribute approximately  $15$  to  $105 \text{ W/m}^2/\text{K}$  more heat to the atmosphere. The total mapped geothermal area in the thermal image is  $17 \text{ km}^2$ , which is  $0.79\%$  of the  $2167 \text{ km}^2$  study area. This scales the geothermal area surface heat flux per degree Kelvin from  $250$  to  $1784 \text{ MW/K}$ . Similarly, the mean contribution of the mapped geothermal areas to the average surface heat flux of the study area is  $116 \text{ mW/m}^2$ .

These numbers are comparable to estimates of the total thermal output of the Yellowstone region ( $\sim 6,100 \text{ MW}$ , Ingebritsen et al., 2001), and to the mean ground heat flow of the region ( $120\text{--}150 \text{ mW/m}^2$ , Blackwell et al., accessed 2002) and the plateau itself ( $2000 \text{ mW/m}^2$ , Fournier, 1989), as determined by measuring temperature gradients in a network of deep boreholes.

This is a crude analysis that serves merely to indicate a way forward toward more rigorous modeling of the surface energy balance. It introduces the idea that ground heat flow, as mapped by geothermal resources investigators, can be matched against as a surface expression of that heat flow sensed from space.

## 6 Summary and further work

### 6.1 Summary

A map of the geothermal areas of western central Yellowstone National Park was produced by novel remote sensing techniques based on Landsat data. The map delineates thermal areas in 15 classes along a heat scale from extreme down to barely detectable. The map is a raster, comprised of 28.5-meter pixels – with an effective resolution of 28.5 to 57 meters. Of the numerous raster data tested, the final mapping process utilized four rasters: 57 m ETM Bands 6L and 6H from March 25<sup>th</sup> 2000, a 28.5 m digital elevation model, and a 28.5 m theoretical terrain shading image calibrated to the Landsat imagery.

The techniques that were developed in producing the map are largely objective, being based on unsupervised classification of thermal bands and terrain information. A certain degree of subjective input is required when assigning meaning to individual classes, which involves fairly simple rule-based interpretation using comparison with pre-existing data. The technique is also relatively rapid. A map can be produced in approximately one day, once all data are obtained and converted into the necessary format. The latter caveat should not be underestimated, however, as it can take an unexpectedly long time to obtain digital spatial data from disparate sources, and to massage it into the required form.

Detailed comparison with pre-existing data and purposely collected field data suggests that the map effects the most accurate large-scale measure of the spatial distribution of *relative* surface heat flux for the Park to date. Future work could refine this into accurate, *absolute* measures of spatial heat flux (in W/m<sup>2</sup>) that could be compared with ground heat flow measurements made using the nation-wide well database. Seven different types of pre-existing data were compared with the map: Hutchinson's thermal polygons, Shovic's landform polygons, MSU elk-tracking point data, NPS Thermal Inventory data, some purposely acquired GPS thermal perimeter data, regional ground heat flow measurements based on bore hole data, and general observations.

There was general agreement with all data sets. The principal disagreement was where the air-photo interpreted polygon data sets drew a large boundary

around mixed thermal/non-thermal terrain, whereas the thermal image responds continuously at the 60 m scale to each local area's thermal/non-thermal status. We argue that this is a deficiency in the polygon data sets. Stronger agreement was reached with point data collected from field survey. Very few areas were declared thermal by the image without a surveyed thermal area co-located or very nearby. However, numerous small, isolated thermals present in the elk tracking data set were absent in the thermal image. The best agreement was reached with field data collected as part of the present study, and targeted at measuring the perimeters of thermal influenced areas. The accuracy of the thermal image estimates of total extent of thermal influence is confirmed by good boundary correspondence with the targeted data (after allowing for GPS and satellite translation error), and by good boundary correspondence with elk tracking data in areas of high elk visitation. All positive agreement statistics improved substantially after a numerical experiment was conducted that simulated the removal of the effects of location error in the image of the pre-existing data.

A single quantitative measure of the accuracy of the image cannot be computed using present testing data, because in every comparison with large pre-existing data sets, the accuracy of those data sets themselves can be brought into question. The inferences drawn from these comparisons, however, suggest that the map may be the most accurate large-scale estimator of the extent and intensity of surface heat flux for the major geothermal region of Yellowstone National Park.

In the final map, some false positives may remain on a limited number of south-facing rhyolite cliffs. Further ground-truthing should confirm this. False negatives are more prevalent. Although large, weakly thermal areas are detected in the map, weak to moderately thermal areas smaller than a Landsat thermal pixel (57 m) are largely absent from the map.

## 6.2 Uses

### *6.2.1 Thermal monitoring*

The map has a wide range of potential uses. As a thermal monitoring tool, it quantifies the total thermally influenced surface area of western central Yellowstone National Park (17.1 km<sup>2</sup> over a 2167 km<sup>2</sup> study area). If the same

remote sensing techniques were applied to Enhanced Thematic Mapper data from following years, the difference is anticipated to indicate any areas of significant localized change in thermal activity.

#### *6.2.2 Exploration and inventory*

The map has the potential to discover previously unknown thermal areas. At present, it is being used to guide NPS field surveyors compiling the Thermal Inventory of the Park.

#### *6.2.3 Geologic processes*

Yellowstone thermal features are the surface expression of the most active volcanic system on Earth. The Caldera has surged about a meter in the past century, as evidenced by shoreline and elevation changes (Waite & Smith, 2002). Over the past two million years, its eruptions have produced the largest volumes of volcanic material known to science (Fritz, 1985). By sequencing a number of maps obtained from many years of satellite data, a 'video' of surficial heat flux change over time could be produced. Geologists could use such information to gain new insight into the volcanic processes that underlie the Park.

#### *6.2.4 Landscape mapping*

The map can be used for input to landscape mapping efforts, by immediately delineating thermal landscapes, and thereby easing the complexity of mapping remaining landscapes (forest types, meadows, etc.).

#### *6.2.5 Snowpack modeling*

The map already forms part of the input to a daily spatial snowpack simulation model for the study area. Previously, there were no suitable data that could provide a useful indication of ground heat flux for this purpose.

#### *6.2.6 Integrated visualization of the Park*

The authors are working towards expressing an integrated understanding and visualization of the Park's ecosystems through a combination of data on terrain, vegetation, and wildlife, models of snowpack and herd behavior, and a new

integrated environmental software framework. The thermal map underlies many of these components.

### 6.3 Further work

Further work may include:

- mapping the whole Park
- mapping of historic & future images for change detection
- further, targeted ground truthing
- methodological improvements, such as the development of a mapping scheme based on inversion of a physical model of heat balance and radiative heat flux
- use of night-time Landsat thermal imagery
- use of higher resolution and/or hyper-spectral sensors, such as the airborne AVIRIS sensor

## 7 References

- Blackwell, D.D., Wisian, K.W., Steele, J.L. (accessed 2002) Geothermal resource/reservoir investigations based on heat flow and thermal gradient data for the United States. Department of Geological Sciences, Southern Methodist University. <http://id.inel.gov/geothermal/fy97/explore/exp-16.html>.
- Brock, T.D. (1994) Life at High Temperatures, Yellowstone Association for Natural Science, History & Education, Inc., 31 pp.
- Craighead, J.J., Craighead, F.C. Jr., Ruff, R.L., & O'Gara, B.W. (1973) Home ranges and activity patterns of nonmigratory elk of the Madison drainage as determined by biotelemetry. *Wildlife Monographs* 33.
- Dozier, J. (1989) Spectral signature of alpine snow cover from the Landsat Thematic Mapper, *Remote Sens. Environ.*, 28:9-22.
- Fournier, R. O. (1989) Geochemistry and dynamics of the Yellowstone National Park hydrothermal system. *Annual Reviews of Earth and Planetary Sciences*, v. 17, p. 13-53.
- Fritz, W.J. (1985) Roadside geology of the Yellowstone Country, Mountain Press Publ. Co., Missoula, Montana, 149 pp.
- Garrott, R.A., L.L. Eberhardt, P.J. White, and J. Rotella. (2002 in review) Climate-induced limitation of a large herbivore population. *Canadian Journal Zoology*.
- Garrott, R.A., L.L. Eberhardt, J.K. Otton, P.J. White, and M.A. Chaffee. (2002 in press) A geochemical trophic cascade in Yellowstone's geothermal environments. *Ecosystems*.
- Garrott, R.A., J.G. Cook, J.G. Berardinelli, P.J. White, S. Cherry, and D.B. Vagnoni. (1997) Evaluation of the urinary allantoin:creatinine ratio as a nutritional index for elk. *Can. J. Zool.* 75:1519-1525.
- Ingebritsen, S.E., Galloway, D.L., Colvard, L.M., Sorey, M.L., and Mariner, R.H. (2001) Time-variation of hydrothermal discharge at selected sites in the western United States: Implications for monitoring: *Journal of Volcanology and Geothermal Research*, v. 111, p. 1-23.
- Likas, A., Vlassis, N., & Verbeek, J.J. (2002 in press) The global *k*-means clustering algorithm, *Pattern Recognition*, in press.
- Madigan, M.T., Martinko, J.M., & Parker, J. (2000) Brock Biology of Microorganisms (9<sup>th</sup> ed.), Prentice Hall, New Jersey, USA, 991 pp.
- Microimages (2000), TNT Products V6.40, Reference Manual, Lincoln, NE, V 3482.2.2

- Minnaert, M. (1941) The reciprocity principle in lunar photometry. *The Astrophysical Journal*, 93:403–410.
- Shovic, H.F. (1996) Landforms and Associated Surficial Materials of Yellowstone National Park, Yellowstone National Park, Center for Resources, Box 168, Mammoth, WY, 82190.  
See also: [http://www.nps.gov/gis/metadata/yell/yell\\_landform.html](http://www.nps.gov/gis/metadata/yell/yell_landform.html).
- Smith, D.W. & Guernsey, D.S. (2002) Yellowstone Wolf Project: Annual Report, 2001. National Park Service, Yellowstone Center for Resources, Yellowstone National Park, Wyoming, YCR–NR–2002–04, 21 pp.
- Smith, J., Lin, T., & Ranson, K. (1980) The Lambertian assumption and Landsat data. *Photogrammetric Engineering and Remote Sensing*, 46:9:1183–1189.
- Smith, R.B. (accessed 2002) Yellowstone Hotspot.  
<http://www.mines.utah.edu/~rbsmith/RESEARCH/YellowstoneHotspot.html>.
- Tou, J.T. & Gonzales, R.C. (1974). *Pattern Recognition Principles*, Addison –Wesley, Reading, MA, USA, pp. 94–97.
- Waite, G. R. & Smith, R.B. (2002) Seismic evidence for fluid migration accompanying subsidence of the Yellowstone Caldera, *J. Geophysical Res.*, (in press).
- Watson F.G.R., Vertessy, R.A. & Grayson, R.B. (1999) Large-scale modelling of forest hydrological processes and their long-term effect on water yield. *Hydrological Processes*, 13:689–700.
- Watson, F.G.R., Rahman, M.V.J.R., & Seaton, S. (2001) Deploying environmental software using the Tarsier modelling framework. Proc. 3rd Australian Stream Management Conference, Brisbane 27–29 August, 2001, pp. 631–637.
- Yellowstone National Park (YNP) (2000) *Yellowstone National Park: Resources & Issues Handbook 2000*. Division of Interpretation, YNP, 165 pp.