ASSESSING THE ACCURACY OF LANDSAT-DERIVED STREAM TEMPERATURE FOR USE IN JUVENILE SALMONID HABITAT ASSESSMENTS ON THE ANCHOR RIVER, ALASKA

A Thesis

Presented to the Faculty of

Alaska Pacific University

In Partial Fulfillment of the Requirements

For the Degree of

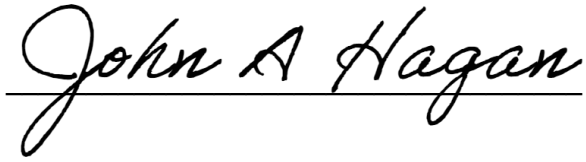
Master of Science in Environmental Science

By

John A. Hagan

April 2017

I grant Alaska Pacific University the non-exclusive right to use this work for the purpose of making single copies available to the public on a not-for-profit basis if the University's circulating copy is lost or destroyed.



John A Hagan

Date 4/25/2017

ACKNOWLEDGEMENTS

This work was completed in the partial fulfillment of a Master's degree of Environmental Science at Alaska Pacific University (APU). The guidance of my committee members was invaluable during the process, I'd like to thank Dr. Bradley Harris, Scott Smeltz, and Jason Geck from APU and Dr. Suresh Sethi from Cornell University. Graduate studies at APU were funded by the At-Sea-Processors Association and the Alaska Space Grant Program. Additional technical and scientific support came from Sue Mauger (Cook InletKeeper), Mike Booz (Alaska Department of Fish and Game), Dr. Mike Loso (APU), Dr. Nathan Wolf (APU), Aileen Nimick (APU), Sarah Webster (APU), Dr. Chris McGonigle (Ulster University), Jay Calvert (Ulster Univeristy), Tom Harris (Anchor Point resident), and Dr. Rebecca Handcock (Murdoch University). I would also like to thank my fellow graduate students at the Fisheries, Aquatic Science, and Technology Laboratory at APU. And lastly I would like to thank my wife, Karissa McLane, for her endless support and patience, without her, none of this would have been possible.

ABSTRACT

This thesis investigates the environmental variables that influence the accuracy of Landsat 8 thermal infrared (TIR) imagery-derived stream temperatures in a ground-truthing study on the Anchor River basin in Southcentral Alaska. From May 1st to September 30th, 2015, in-situ temperature data were collected concurrently with the remotely sensed Landsat 8 TIR images at ten field sites throughout the Anchor River watershed. At each field site, stream delineations were preformed to assess the influence of riparian vegetation, stream morphology, discharge, and air temperature on TIR-derived stream temperatures. The accuracy of TIR data was assessed by calculating the Landsat Thermal Offset (LTO), the difference between TIR-derived temperature and the in-situ stream temperature (TIR – In-situ =LTO). A Generalized Additive Model (GAM) was employed to examine the relationship between environmental variables and LTO. Model selection with Akaike information criterion (AIC) determined that Rosgen Stream Classification (*Rosclass*), 7-day trends in stream discharge (*Gagediff*), and daily average air temperature (*Airdaily*) were strongly associated with Landsat 8 thermal offset (LTO). Stability in stream discharge (*Gagediff* = ±0.2 ft.) in the week preceding a Landsat TIR data collection event, coupled with stream morphology characteristics consistent with *Rosclass* B and C reaches, were associated with the highest accuracy of raw Landsat TIR-derived stream temperatures. Across all sites and dates, GAM regressions accurately translated raw Landsat TIR data into temperatures reflecting in-situ water temperatures in small order streams. While regression analyses supported *Rosclass*, *Gagediff*, and *Airdaily* environmental covariates as being influential in explaining LTO, the underlying mechanisms that these variables reflect remains an open question. These variables may indicate conditions that lead to stream stratification, or alternatively, may reflect other environmental processes leading to stronger temperature differentials between temperature

in the stream and that on the surrounding riparian land. Continued validation efforts in other high latitude watersheds over longer time periods will be important for testing the robustness of methods to utilize Landsat TIR-based stream temperature predictions, progressing towards a goal of providing a broad-scale tool for fisheries and watershed managers to assess the degree of, and potential impacts from, climate change on Alaskan freshwater aquatic and fishery resources.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Alaska is experiencing the greatest regional warming of any state in the country. Since 1960, state-wide average annual air temperatures have increased by 2 °C and winter temperatures have risen by 4 °C (Chapin et al. 2014). This warming has led to earlier spring snowmelt, reduced sea ice, widespread glacier retreat, drier landscapes, and inland surface water temperature increase (Chapin et al. 2014). In parallel with warming, there have been decreases in adult Chinook salmon (*Onchorhynchus tsawytscha)* returns, shifts in predominant age at maturity, and decrease in size-at-age of juveniles (Evenson and Carlile 2014; Mantua et al. 2015). These trends are concerning, and it is plausible that warming climate may be impacting Chinook salmon in Alaska (McPhee 2015; Schindler et al. 2015). However, the direct and indirect effects of regional warming on Pacific salmon population dynamics are not well understood due to major gaps in stream temperature data throughout high-latitude watershed basins.

Water temperature structures fish species' distributions, abundance, life histories via direct (e.g. metabolic rates, dissolved oxygen) and indirect (prey species distribution and abundance) mechanisms (Bunn and Arthington 2002). For salmonids, stream temperature influences spawning, incubation success, egg maturation, growth, migration, and competitive ability (Armour 1991), and, as such, natal stream temperatures are correlated with fish survival and may thereby influence abundance (Griffiths and Schindler 2012).

Throughout Alaska, state and federal agencies monitor juvenile salmonid size, age, and species; however, a similar effort to monitor the temperature environments of salmon rearing streams at landscape scales, which may be a leading indicator of population trajectories, has not been a priority. This may be in part because monitoring stream temperature in Alaska is

particularly difficult due to the size and remoteness of salmonid-producing watersheds. For example, the Yukon River has 48 named tributaries and drains 13 basins totaling 832,700 km² (Brabets et al. 2000). Efforts to collect in-situ water temperature data (sampled directly from the stream) have proven to be costly, time consuming, and only represent a small portion of a watershed's thermal profile (Togerson et al. 2001).

Although several of these in-situ water temperature collection efforts have been made regionally, the amount of available stream temperature data for Alaska remains limited (Geist et al. 2014). An alternative solution may be found in collecting stream temperature data on a watershed scale using remote sensing techniques. Appropriately calibrated remotely-sensed TIR may be suitable for examining the relationships between stream temperatures and salmon growth and population fluctuations year-round with reduced cost and logistical demands relative to field sampling.

Background

Hypotheses regarding how landscape-scale temperature change affects fish physiology and habitat productivity have existed for several decades (Schlosser 1991). However, due to the challenge in collecting landscape-scale temperature data, understanding the relationships between stream temperatures and fish populations at large temporal and spatial scales remains nascent. Schlosser (1991) proposed that the interface between terrestrial and aquatic ecosystems are variable based on hydrologic conditions, thus fish refugia are influenced and created by a wide range of dynamic processes. Schindler et al. (2015) found that observing ecological heterogeneity and the processes that generate it across landscapes are a crucial component of management and conservation of juvenile salmonid watersheds. Thus, landscape-scale remotely-

sensed stream temperature data provide opportunity to analyze relationships between large-scale spatial and temporal trends and critical juvenile salmonid life stages.

Thermal infrared imagery (TIR) is widely available and can accurately characterize waterbody thermal heterogeneity (Torgersen et al. 2001; Cristea and Burges 2009). Satellite-derived TIR has been used to successfully map surface thermal heterogeneities of large water bodies such as lakes (Simon et al. 2014) and estuaries (Wawrzyniak et al. 2012), and has been used for mapping spatial variability of stream temperature (Torgersen et al. 2001; Handcock et al. 2006).

TIR has a broad range of remote sensing applications including glacial retreat, sea surface temperature monitoring, land use monitoring, vegetation distributions, deforestation, and forest fire extents, and, as such, thermal sensors have been deployed on several satellites since the 1970's (e.g. Nimbus 7 and NASA's Heat Capacity Mapping Mission). The Landsat Earth Observation Satellite (Landsat) program distributes free publicly available TIR data collected since 1982. The Landsat program has had seven satellites in orbit; currently, Landsat 7 (launched in 1999) and Landsat 8 (launched in 2013) are collecting data from a near-polar, sun-synchronous orbit, while Landsat 1-6 have been decommissioned (NASA 2011). Landsat 7 experienced a scan line corrector failure in May 2003, producing data gaps across all images, thus making Landsat 8 a more reliable source for thermal imagery (USGS 2013).

Landsat 8 TIR provides synoptic digital numbers that represent reflected solar radiation. These digital numbers can then be converted into degrees of temperature at a resolution suitable for predicting water temperatures for large to mid-size rivers (e.g. Handcock et al. 2006). To date, the accuracy of Landsat 8 TIR has not been tested for use in smaller streams suited for juvenile salmonid rearing. Since juvenile salmonids use a wide range of stream systems in

Alaska, further assessment is necessary to quantify the degree of spatial accuracy and the influence of surrounding habitat and temporal environmental processes on Landsat-derived stream temperatures.

Landsat Thermal Infrared Imagery

Landsat 8 TIR imagery provides radiometric surface temperature information at a landscape scale by capturing emitted thermal-infrared radiation. Over water, TIR provides values of reflected solar energy from the top 100 micrometers of the water column (Togerson et al. 2001). Radiant water temperature (emitted radiation) measurements are representative of kinetic water temperature (in-situ) under well-mixed water conditions typical of lotic systems (Togersen et al. 2001). Kinetic temperature is biologically important for salmon physiology and a driver of habitat productivity (Handcock et al. 2006).

Landsat 8 TIR imagery is publicly accessible through the U.S. Geological Survey Explorer. Images can be downloaded from the database, which is automatically updated within 24-hours of image collection. Once downloaded, images must be calibrated using quantized-scaled Digital Numbers (DN) representing the multispectral image data (Appendix 1; USGS 2013). Images must also be corrected for atmospheric transmission radiance using an atmospheric correction algorithm (Appendix 2; USGS 2013; Handcock et al. 2012; Simon et al. 2014).

Landsat 8 is equipped with 11 spectral bands. Each band represents a different portion of the electromagnetic spectrum. Bands 1-9, known as the Operational Land Imager, are spectral bands commonly used in the earth sciences. The remaining bands, 10 and 11, collect TIR data using the Thermal InfraRed Sensor (USGS 2013).

The thermal bands 10 and 11 acquire TIR imagery at 100-meter resolution and resample to 30-meter pixels (USGS 2013). Pixels are resampled using cubic convolution interpolation, a method used to determine values in an image through a weighted average of the 16 closest pixels to the input coordinates (Keys 1981). Spatial resolution of images acquired by the Landsat 8 TIR bands is expressed in meters. 30-meter resolution refers to one temperature value covering an area of 30 meters by 30 meters and displayed as a unique pixel in raster image (USGS 2013). Band 10 was used in this study to derive stream surface temperature as Band 11 is significantly more contaminated by stray light than Band 10 (USGS 2013).

Handcock et al. (2006) tested the accuracy and uncertainty of several TIR products, including Landsat 7 TIR, on inland streams in the Pacific Northwest with a range of stream wetted widths (10–500 m) and TIR pixel sizes (5–1000 m). Stream sizes were grouped into three classes and TIR was compared to in-situ temperatures. The three stream width size classes were tested with 30m pixel resolution from the Landsat TIR image. The size classes were: three or more pixels across a stream, one to three pixels across a stream, and less than one full pixel across a stream. The authors found that if more than three pixels spanned the wetted width of a given stream (width $\geq$ 90m), the TIR temperature overestimated the in-situ temperature on average by 1.2 °C. If one to three pixels spanned the wetted width, influence from riparian and terrestrial ground cover increased the overestimation of TIR temperatures by 2.2 °C on average. In the last size class, a fraction of a pixel across the stream, TIR temperatures overestimated by 7.6 °C on average. Handcock et al. (2006) found that overestimation of the TIR temperatures is due to ground cover influence, thermal stratification between the stream surface and the location of the in-situ temperature measurements deeper in the water column, and that reliable satellite TIR measurement of stream temperatures is limited to large rivers with greater discharge.

Wawrzyniak et al. (2012) looked at the longitudinal and temporal variations in thermal patterns of the lower Rhône River in France using Landsat 8 TIR. Citing Handcock et al. (2006), these authors only used stream reaches with one or more pixels spanning the wetted width. A water extraction technique was used to remove pixels containing terrestrial ground cover. The average accuracy of Landsat 8 TIR was found to be ±1.1 °C for reaches with more than 3 pixels across and ±1.4 °C for reaches with one to 3 pixels across. Wawrzyniak et al. (2012) concluded that TIR accuracy is more influenced by terrestrial surfaces than number of pixels spanning a stream reach.

Simon et. al (2014) characterized the thermodynamics of waterbodies, specifically of irregularly-shaped inland reservoirs using Landsat 8 TIR. The authors proposed that images must be corrected, mainly for atmospheric effects, to be useable, and the study's objective was to validate the standard correction algorithm for Landsat 8 TIR. Two freshwater reservoirs in central France were selected as study sites. Results showed a significant difference between corrected and uncorrected image data and recommended atmospheric corrections be made to Landsat 8 TIR images.

Scope of Study

Previous studies have assessed the overall accuracy of Landsat TIR and found that larger streams yield more accurate stream temperatures (Handcock et al. 2006; Wawrzyniak et al. 2012). However, most juvenile salmonids occupy relatively shallow, low-velocity areas in small streams averaging 1 m deep and with average velocity of 40 cm/s during rearing (Bjornn and Reiser 1991). These streams, often 1st to 5th order systems (stream order is a measure of the relative size of streams, smallest tributaries are referred to as first-order streams**)**, have wetted widths that occupy just a fraction of one Landsat 8 TIR pixel (Vannote et al. 1980). This study

quantifies the accuracy of Landsat 8 TIR on stream systems smaller than the 30 m Landsat pixel by assessing the ability of Landsat to derive stream temperature in a small salmonid rearing stream in Southcentral Alaska. In contrast to Handcock et al. (2006) and Wawrzuniak et al. (2012), this study has also analyzed the influence of environmental variables on TIR-derived stream temperature.

METHODS

This study was framed as a ground-truth exercise - collecting in-situ temperatures from an Alaskan watershed and comparing the data to TIR-derived temperatures sampled at the same time and location. The watershed was chosen based on its accessibility, diversity of river morphology, presence of common Alaskan vegetation, and use by juvenile salmonids, including Chinook salmon. Due to the stream only occupying a small portion of a Landsat 30 m pixel, percent of water within the pixel, the remaining ground cover (everything not-water), and stream morphology were surveyed in ten 30-meter by 30-meter field sites. In addition, stream discharge and air temperature data were collected from local sensors maintained by the United States Geologic Survey (USGS) and the Natural Resource Conservation Service (NRCS). Following the field exercise, the accuracy of TIR-derived temperatures were evaluated and quantified based on environmental variables collected in the field during the same time scale.

Study Area

The Anchor River watershed is located on the southern portion of the Kenai Peninsula; its headwaters are in the Caribou hills to the east of its mouth in the lower Cook Inlet (Figure 1). The total area of the watershed is 583 km^2 including 185 km of streams that are classed as

anadromous fish spawning habitat by Alaska Department of Fish and Game and used by four species of Pacific salmon (ADF&G 2015). The watershed is classified as a 5th-order peat wetland-supported system typical of a non-glacial watershed. Maximum stream discharge occurs from August to November with low flows occurring during June and July (Rinella et al. 2008).

In-situ stream temperatures were collected at ten field sites within the Anchor River watershed, with 3 on the South Fork Anchor River, 5 on the North Fork Anchor River, and 2 on the Chakok River, a North Fork tributary (Figure 1). To avoid thermal input from potential heat sinks, the field sites were chosen at least 30 meters from bridges, roads, or buildings. Sites were selected to represent habitat types, gradients, and wetted widths typical of salmon-rearing systems. No sites were chosen on the main stem Anchor River to avoid tidal influence. Since much of the Anchor River watershed is privately owned, site selections were also constrained to locations with public access. One site on private land was made available by a local contact.

Temperature Collection

In-situ Data

In-situ steam temperature data were collected from 5 May 2015, through 1 November 2015, to coincide with Landsat TIR data collection and historical Chinook salmon incubation, emergence, rearing, and spawning periods (ADF&G 2013). Data were collected using HOBO TidbiT v2 Water Temperature Data Loggers set to collect stream temperature in degrees Celsius at bi-hourly intervals (Hoboware 2014.). Loggers were mounted 5 cm from the streambed on embedded rebar and roughly mid-way between the thalwag and wetted edge in each of the ten field sites. Data loggers were deployed, calibrated, and checked monthly for accuracy using methods outlined by Mauger et al. (2014).

Landsat Thermal Imagery

Landsat 8 overflew the study area 11 times during the sampling period. TIR (band 10) images were selected for 6 cloud-free dates to avoid atmospheric influence (USGS 2015). Once downloaded, images were calibrated using quantized-scaled Digital Numbers (DN) representing the multispectral image data (Appendix 1; USGS 2013). Images were also corrected for atmospheric transmission radiance using the appropriate atmospheric correction algorithm (for details see Appendix 2; USGS 2013; Handcock et al. 2012; Simon et al. 2014). Temperature was converted from Kelvin to Celsius before analysis for consistency with in-situ sensor temperature output. At least one Landsat TIR image was used from each month during the field study (May – September) anticipating seasonal trends in both stream discharge and air temperature.

Field Site Delineations

All 10 field sites were characterized based upon a suite of stream and riparian attributes. Each field site was treated as a "stream reach" typical of fisheries habitat surveys. Environmental data were collected on-site except for stream discharge and air temperature data, which were collected using USGS stream gage and NRCS Snotel databases respectively (Table 1, Table 2).

Rosgen Stream Classification System

Rosgen (1996) described a stream classification system for riverine ecosystems, commonly referred to as the "Rosgen Stream Classification System." This system is used throughout the primary literature and has been adopted by the Environmental Protection Agency (EPA) and many other federal and state agencies for watershed management. It is considered the standard for river morphology surveys and assessment. The System is broken down into 3 levels (I, II, and III). The Level I System is the most general, and is based on stream characteristics

resulting from width-to-depth ratio, sinuosity, gradient, deposition, entrenchment and, channel pattern; stream types are classified from A to G based on these characteristics.

The Rosgen Stream Classification - Level I System was used to define the stream type at each in-situ sampling site. Based on its size and watershed attributes, the Anchor River sites were expected to be Types B, C, and E. Type B sites have a moderately steep channel slope, moderately entrenched, have a cross-section width/depth ratio, low sinuosity, and pool-riffle dominated bed morphology. Type C sites feature a well-developed floodplain, moderate channel slope, slightly entrenched with sequencing riffles and pools, moderately sinuous, and characteristic "point bars" within the active channel. Type E sites have very low elevational relief, are slightly entrenched, very low channel width/depth ratios, very high sinuosity, and high number of pools (Rosgen 1996; Appendix 3).

Ground Cover

Ground cover (everything not water in a 30m x 30m pixel), including terrestrial and riparian vegetation, exposed substrate, and debris, was characterized at all 10 of the field sites. Ground cover was classified into 3 categories - Light, Medium, and Dark - based on thermal output following Handcock et al. (2012). Within each category, vegetation was classified to sub-categories using Boggs Indices (Boggs et al. 2014). Boggs Indices are ground cover descriptions specific to Alaska and aerial imagery analysis (Viereck et al. 1992) using variations of the Alaska Vegetation Classification levels III and IV (Boggs et al. 2014).

Percent Water Class

The percent water by area within the 30m pixel was calculated for each field site, using a categorical variable called Percent Water Class (PWC). Percent Water Class took on one of four

values corresponding to the percent of pixel that was water (**1**: 0-25%; **2**: 26-50%; **3**: 51-75%; **4**: 76-100%).

Gage Height

Stream stage data were collected from the USGS stream gage site located on the South Fork Anchor River (Gage Number: 15239900; Latitude 59.4450, Longitude -151.4511; Figure 1; USGS 2016). Stream stage data are collected every 15 minutes and are publicly available (USGS 2016). Stream stage, measured in feet above the river bed, is used to describe discharge in this study (data were not converted into cubic feet per second using a rating curve). Data were obtained for the entire field season at 15 minute intervals to account for trends, fluctuations and anomalies. Stream stage data were assessed for each Landsat 8 TIR image date as follows:

Gage Daily Average (*Gagedaily*)- The average stream stage from the 24-hour period preceding a given Landsat TIR pass-over date.

Gage Weekly Average (*Gageweekly*)- The average stream stage from the 7 days preceding the Landsat TIR image date.

Gage Difference (*Gagediff*)- The difference between *Gagedaily* and *Gageweekly*. This provides context for fluctuations or stability within the water table in the week preceding the Landsat TIR image. A negative *Gagediff* suggests a declining water level, a positive *Gagediff* suggests a rising water level, and a *Gagediff* close to 0 implies stability.

Air Temperature

Air temperature data were obtained from the NRCS Snotel Site at the Anchor River Divide (Site Number**:** 1062; Latitude 59. 8597, Longitude -151. 3111; Figure 1; NRCS 2015).

Air temperature, in degrees Celsius, is collected every hour and is publicly available (NRCS 2016). For the purposes of this study, air temperature data were obtained for the entire field season. Air temperature data were assessed for each Landsat 8 TIR image date as follows:

Air Daily Average (*Airdaily*)- The average air temperature from the 24-hour period preceding a given Landsat TIR pass-over date.

Air Weekly Average (*Airweekly*)- The average air temperature from the 7 days preceding the Landsat TIR image date.

Air Difference (*Airdiff*)- The difference between *Airdaily* and *Airweekly*. This provides context for fluctuations or stability in air temperature in the week preceding the Landsat TIR image date. A negative *Airdiff* suggests a declining air temperature, a positive *Airdiff* suggests a rising air temperature, and an *Airdiff* close to 0 implies stability.

Temperature and Stream Delineation Database Construction

A Geographical Information System (GIS) database was constructed in ArcGIS software (ESRI 2014) to allow the in-situ temperature, site characteristic, and Landsat information to be joined based on spatiotemporal concurrence. The WGS84 horizontal datum and the Alaska State plane projection were used to georeference the location information for each dataset. Landsat 8 TIR scenes for each date were imported to the GIS and each pixel converted to a polygon corresponding to the Landsat image. This allowed for stream delineations to be made within the same boundaries as the 10 field sites.

The position of each in-situ stream temperature sensor was collected with a survey-grade (accuracy: 1–100 cm) GPS unit (Trimble Geo7X). Location data were post-processed using the GPS Pathfinder Office Software (GPS Pathfinder Office 2014). All in-situ temperature data were

downloaded and checked for quality using Hoboware software (Hoboware 2014) before being added to the GIS as a single *x, y* point with temperature data spanning the entire study period. Each in-situ point fell into its corresponding Landsat polygon and was thereby joined to the Landsat 8 TIR temperature value, thus creating a database with both temperature values and stream delineation information. Each field site was assigned a Site Name (based on its location; Figure 1), Rosgen Stream Classification, Ground Cover Class, Percent Ground Cover Class, Percent Water Class, Air Temperature, and Gage Height values (Table 1; Table 2).

Data Analysis

Landsat Thermal Offset

The accuracy of Landsat 8 TIR is measured using the Landsat Thermal Offset (LTO), the difference between the TIR image-derived temperature (T_i) and the in-situ temperature (T_l):

$$LTO = T_i - T_l$$

An LTO of 0℃ means Landsat TIR and in-situ temperatures are equal. A positive LTO results when Landsat TIR has a higher temperature than the in-situ measurement. A negative LTO results when the Landsat TIR was lower than the in-situ temperature.

Generalized Additive Model

Initial data explorations revealed non-linear relationships between some of the environmental explanatory variables and LTO. As such, a Generalized Additive Model (GAM) was employed to examine the relationship between environmental variables and LTO associated with each in-situ site (6 dates, 10 sites, n = 60). GAMs are an extension of Generalized Linear Models (GLMs), which use a smoothing function to describe the relationship between predictor

variables and response data, like those found in weighted regressions (Bio et al. 1998). Their flexibility in modeling potentially nonlinear relationships between explanatory variables and response variables makes GAMs attractive for exploratory analyses for which a priori understanding about the associations between covariates and modeled data is limited. GAMs were fit to explore relationships between site-specific environmental covariates and LTO (Wood 2004). All models were run using R, a language and environment for statistical computing and graphics (R Core Team 2015). In model formulation notation, the global model tested was:

LTO ~ Rosclass + Water + Light + f(Gagedaily, k=4) + Medium + Dark + f(Gagediff, k=4) + f(Airdaily, k=4) + f(Airdiff, k=4) +f(Airweekly, k=4) + f(Gageweekly, k=4)

The *f(variable, k=4)* terms indicate a given explanatory variable was modeled as a cubic spline with four knots. GAM-produced LTO predications were generated using the {mgcv} package (Wood 2001). Since temporal data can have non-linear complexity, smoothing functions were used on temporal data sets (stream stage and air temperature) and were run as cubic splines with four knots. All other covariates entered GAM models in level form (Wood 2004; Appendix 4).

Model Fitting and Evaluation

To account for model uncertainty in assessing the relationship between stream and riparian explanatory variables and LTO, Akaike information criteria corrected for small sample sizes (AICc) based multi-model inference was used for every combination of explanatory environmental variables contained in the global model (Burnham and Anderson 2002). Model selection was based upon the complete set of possible models, treating all covariates on level

form or as a cubic spline function (with four knots), and without any interaction terms (Symonds and Moussalli 2010).

RESULTS

Model Selection

The global model containing all environmental covariates had high goodness of fit (R^2_{adj} = 0.873). AICc- based model selection identified the two top models had comparable support (Table 3). The models with the lowest AICc scores and root mean squared errors (RMSE) included both Rosgen Stream Classification (*Rosclass)* and Gage Difference (*Gagediff) (AICc=253.3, RMSE* = 1.76℃*;* Table 2, Table 3; Figure 2*).* The next best model had only a marginally higher AICc score and a slightly lower RMSE and included *Rosclass*, *Airdaily*, and *Gagediff* (AICc = 253.6, RMSE = 1.72℃; Table 2, Table 3; Figure 3).

Explanatory Variable Support

The relative importance (RI) of the independent variables were assessed using AICc weights. *Rosclass* (RI = 93%) and *Gagediff* (RI = 91%) were the two most important variables. *Airdaily* (RI = 52%), Air Difference (*Airdiff*; RI = 36%), Gage Daily (*Gagedaily*; RI = 39%), dark ground cover (*Dark*; RI = 30%), medium ground cover (*Medium*; RI = 27%), light ground cover (*Light; RI = 20%),* and percent water (*Water*; RI = 24%) had somewhat lower relative variable importance support when comparing across all models.

Corroborating relative importance (RI), the top model (LTO ~ *Rosclass* + *f(Gagediff)*) indicated *Rosclass* was statistically significant (p= 0.015) as was *Gagediff* (p < 0.001) based

upon traditional frequentist hypothesis testing. The second top model (LTO ~ *Rosclass* + *f(Gagediff)* + *f(Airdaily))* found *Rosclass* ($p=0.014$) and *Gagediff* ($p < 0.001$) to be statistically significant. *Airdaily* ($p=0.25$) was not statistically significant.

Predictive Performance

The AICc top model (LTO ~ *Rosclass* + *f(Gagediff)*) predications had a mean absolute error (MAE) of 1.38 with 88.6% deviance explained (Figure 4). The second best AICc supported model (LTO ~ *Rosclass* + *f(Gagediff)* + *f(Airdaily))* predications had a MAE of 1.36 with 89.1% deviance explained (Figure 5). The top model containing *Rosclass* and *Gagediff*, when assessed by date (Figure 6), displayed a tendency to over-predict stream temperatures (Landsat measured warmer than in-situ), with exception of September 12 2015 (9/12/2015), when temperatures were under-predicted (in-situ measured warmer than Landsat). The widest margin of errors was measured on May 30 2015 (5/30/15), producing both the min, -5.69℃, and the max, 3.46℃, residuals. When assessed by field site (Figure 7), LTO residuals showed much more variability, with some sites always over-predicating and some always under-predicting. Field sites that were a *Rosclass* E, CR-1 and CR-2, experienced the highest and lowest mean residuals, indicating they have the least accurate LTO. The minimum residual occurred at field site NF-3. The site with the most accurate LTO was SF-3, a *Rosclass* B.

The second-best model had similar residual patterns in fit performance. Residuals, when assessed by date (Figure 8), showed over-predicting LTO. The widest margins of error occurred on May 30 2015 (5/30/2015) and July 1 2015 (7/1/2015; with 5/30/2015 producing the largest negative residual, -5.99℃). The largest positive residual occurred on August 11 2015 (8/11/2015; 3.5℃). When assessed by site (Figure 9), residuals revealed that the model tended to over-predict LTO at most sites, apart from NF-4 and NF-5.

Effects of Covariates on LTO

To provide further insight into the influence of specific covariates on LTO performance, GAM models with single covariates were fit based upon the top AICc supported explanatory variables. A model run with LTO ~ *Rosclass* only (R^2_{adj} = 0.0391, RMSE= 5.01℃) performed poorly, with MAE for model predictions of 3.9℃ with 7.17% deviance explained (Figure 10). Conversely, a model run with LTO ~ *f*(*Gagediff)* only (R^2_{adj} = 0.805, RMSE= 2.24℃) performed relatively strongly, with MAE for model predictions of 1.70℃ and 81.4% deviance explained (Figure 11).

Examining the top two AICc supported covariates (*Rosclass*, *Gagediff*), *Rosclass* B and C type field sites have a lower mean LTO than E type sites. Calculating means for each class, B type sites had a mean LTO of 4.44℃, C type streams had a mean LTO of 3.06℃, and E type streams had a mean LTO of 7.07℃. This would imply C type sites had the most accurate raw LTO, followed by B type streams. E type streams had the highest mean LTO, indicating lower accuracy (Figure 12).

Gagediff had a wide range of variability when examined with LTO through the duration of the field season. As expected, LTO was lowest when *Gagediff* was closest to 0 ft., meaning no increase or decrease in stream discharge in the week preceding the Landsat image collection. LTO decreased in accuracy most significantly during the June 15 2015 (6/15/2015) sampling event, when stream discharge had drastically dropped in the preceding week (*Gagediff* = 0.28 ft.). This outlier yielded the most inaccurate Landsat TIR temperature (largest LTO). Mean LTO was highest on June 15 2015 (6/15/2015; 13.92℃) when *Gagediff* was negative (-0.28 ft.), and

the lowest mean LTO was July 1 2015 (7/1/2015; -0.48 ℃) when *Gagediff* was very close to 0 ft. (0.005 ft.; Figure 13).

Mean accuracy by *Gagediff* offers discharge ranges and their corresponding mean LTO. The smallest *Gagediff* metric ±0.1 ft., produced a mean LTO of 0.8 °C (n=20), the next range ±0.2 ft. produced a mean LTO of 2.38 °C (n=40), and a range of ±0.5 ft. produced the full mean of 4.56 °C (n=60; Figure 14).

Results suggest models with stream gage data performed relatively strongly in explaining LTO; however, air temperature data near stream temperature prediction sites may be more readily available. To investigate the performance of air temperature data in isolation of stream gage data, models were run without *Gagediff* to assess the use of *Airdaily* as a covariate. A GAM model run with LTO ~ *Rosclass + f(Airdaily)* (R^2_{adj} = 0.858, RMSE= 1.87℃) preformed quite well, with RMSE only slightly higher than the two top models (top model: LTO ~ *Rosclass + f(Gagediff)*, second top model: LTO ~ *Rosclass + f(Gagediff) + f(Airdaily)*). A GAM model run with LTO ~ *f(Airdaily)* (R^2_{adj} = 0.788, RMSE= 2.35℃), performed significantly less well than both top models or the model run with LTO ~ *Rosclass + f(Airdaily).*

Possible Field Errors

In-situ temperature loggers need to be wetted and free of debris to provide accurate stream temperature data. During the field study, loggers were periodically checked; however, it is possible that debris such as salmon carcasses or woody debris may have contacted sensors, or that extreme changing water levels may have rendered some sensors dry for short periods of time. This would potentially bias in-situ readings and contribute to perceived LTO error if such sensor anomalies occurred during a Landsat pass. That notwithstanding, such errors were

unlikely during this study; during a total of 9 field visits to all sites, no equipment errors were observed apart from the July 26 2015 (7/26/2015) date when a salmon carcass was found above the in-situ sensor hung on the embedded rebar.

DISCUSSION

The goal of this study was to identify which stream environmental variables most strongly influence the accuracy of Landsat 8 TIR-derived stream temperature and to determine whether GAM modeling framework could correct Landsat 8 TIR into temperature data representative of in-situ stream temperatures in smaller order systems. Model selection analysis identified *Rosclass* and *Gagediff* as the top covariates that influence Landsat 8 TIR-derived stream temperatures, with weaker support for *Airdaily* as an influential explanatory covariate. The AICc top model appears to suggest the accuracy of TIR-derived stream temperature data are driven by morphological (*Rosclass*) and hydrological (*Gagediff*) processes. Results indicated that the relationship between Landsat 8 TIR-derived stream temperature and in-situ measured temperature may depend on flow conditions in targeted stream reaches.

One plausible mechanism expounding the relationship between stream flow conditions and LTO is thermal stratification, which can lead TIR based information at the surface of a waterbody to not accurately reflect temperature conditions in the remainder of the water column. Other work on remotely sensed temperature products have implicated thermal stratification as decreasing the accuracy of TIR-derived stream temperatures (Torgersen et al. 2001; Handcock et al. 2006). Thermal stratification can occur on a micro-level, causing a thin formation of warm surface water created by air temperature anomalies; this is known as the thermal boundary effect.

In most rivers, the thermal boundary effect is limited to pools, off-channel habitat, and areas with extremely low gradient, but if turbulent stream mixing is absent due to low discharge, a thermal boundary can form in normally lotic stream reaches if air temperature abnormalities occur concurrently (Torgersen et al. 2001).

During the study period in this effort, peak air temperature (28.8℃) occurred on June 16 2015, one day after Landsat 8 TIR was collected from the same location (air temperature reached 25.8℃ on this date). This high temperature is uncharacteristic for a subarctic climate such as the Anchor River basin, which has a mean historic air temperature of 14.33℃ for June 15 (NWS 2016). A closer look at the week preceding the June 15 2015 Landsat 8 TIR collection, a rapid decline in stream discharge (*Gagediff*= -0.28 ft.) and a rapid increase in air temperature occurred concurrently (Figure 15). These factors may have caused a thermal boundary effect, decreasing vertical mixing. The mean LTO for June 15 2015 was the highest of any day (13.92 ℃), the same day included the highest overall LTO for the whole study (18.55 ℃).

When air temperature was consistent with regional norms and stream discharge was stable, LTO was always smaller. Just 15 days following the June 15 2015 TIR collection, on July 1 2015, air temperature was back within average range (5-20 ℃) and stream discharge was extremely stable (*Gagediff* = 0.005 ft.; NWS 2016; Figure 16). The LTO for this sample event was the lowest for the entire field season (0.64℃) and produced the lowest overall LTO for the whole study (-0.13 ℃).

Stream discharge fluctuations followed seasonal trends and produced a parabolic distribution over the study period. When stream discharge increased significantly, such as on August 18 2015 (*Gagediff* = 0.37 ft.), Landsat TIR accuracy decreased (mean LTO = 3.87℃).

Increased stream discharge can increase vertical mixing; however, it can decrease overall accuracy by increasing wetted width, saturating vegetation and creating stratified off channel pools (Handcock et al. 2006). Temporal effects can all be linked to levels of discharge and display the influence of seasonal changes (Figure 17).

Air temperature's (*Airdaily*) influence appears to be exclusively relevant during heat anomalies. When all data associated with June 15 2015 are removed from the global model, the relative importance of *Airdaily* drops from 52% to 43%. Similarly, removing data associated with June 15 2015 leads to an increase in *Gagediff* relative variable importance from 91% to 100%. Without this anomalous sample event, *Rosclass* enjoys similar AICc relative variable importance support (RI = 94%), underscoring the interactions between discharge and morphology and their effects on the temporal variability of LTO.

The magnitude of discharge, coupled with the underlying stream morphology, can significantly affect vertical mixing (Arntzen et al. 2006). Channel morphology, as captured by *Rosclass*, largely dictates the speed and depth of discharge. Inference can be made from each *Rosclass* stream reach as to the degree of stratification and the mechanisms causing it. *Rosclass* E reaches are slow moving, highly sinuous, low gradient reaches with little turbidity. These reaches are more susceptible to atmospheric influences due to the low stream velocity. *Rosclass* B and C reaches have similarities in their channel morphology. B reaches are slightly more sinuous than C reaches; however, both reach types host riffles and have moderate entrenchment. The slight advantage of improved accuracy when predicting in-situ stream temperature in C type reaches may result from this habitat type's wider wetted-width. B and C reaches often have input from smaller tributaries, thus increasing their discharge. E type reaches are often spring or

groundwater fed (Rosgen 1996). Mean LTO across all sites and dates reflects this notion; E type sites have higher overall LTO than B and C sites, which have similar levels of overall accuracy.

Mulvihill et al. (2009) examined 18 models for bank-full discharge to drainage area relations using *Rosclass* as a stratification-predicting variable. The authors found that while it is difficult to quantify differences in discharge between B and C type reaches, E type reaches had lower mean discharge due to the smaller wetted width, smaller drainage area, and lower gradient and, thus, have significantly higher stratification rates.

Another plausible mechanism for discrepancies between TIR and in-situ stream temperatures is rates of surface heating and cooling. Solar radiation warms both land surfaces and water. Land surface absorbs much more solar radiation than water, which reflects most solar radiation back to the atmosphere, so land retains more heat and warms more quickly than water (Forman 2014). Hence, land and water heat and cool at different rates. Land temperature might change faster than water, for example, if it's anomalously hot (such as on June 15 2015), the land may be warming rapidly, but because water has high thermal mass, it changes temperature more slowly and, thus, results in a bigger LTO. The same processes that drive stratification (i.e. temperature anomalies and discharge flux) also may cause water and land to heat and cool at different rates. This mechanism would vary depending on channel characteristics (explaining why *Rosclass* is so significant), and water velocity and flow volume (i.e. *Gagediff*; Torgersen 2001). The implications of this reiterates the importance of understanding the relative temperature differences between land and water.

Accuracy Assessment

The accuracy of Landsat TIR derived stream temperature predictions varied depending on geomorphic (*Rosclass*) and hydrological (*Gagediff*) covariates. Without correcting for stream covariates through regression, the mean raw LTO across all field sites and dates was 4.56 °C (n=60), with a maximum LTO of 18.54 °C and a most negative offset of -5.67 °C. Using the EPA best practices margin of error of ± 0.5 °C as the gold standard of "point source" stream monitoring, 10% of all LTO fell within this range (McCullough 1999). Fifty-one percent of total raw LTO fell within the ± 4 °C range (a widely accepted TIR error range).

Regression modeling of LTO identified stream conditions that may lead to high accuracy in predicting in-situ stream temperature for smaller streams using Landsat TIR products (Figure 18). The range with the lowest mean LTO, *Rosclass* B type field sites within the *Gagediff* range of ±0.1 ft., had a mean LTO of -0.22 °C (n=10), both *Rosclass* B and C type field sites within the *Gagediff* range of ±0.1 ft. had a mean LTO of -0.63 °C (n=16), *Rosclass* C type within the same *Gagediff* range had a mean LTO of -1.3 °C (n=6).

A larger *Gagediff* range lowers the accuracy of mean LTO slightly, but increases the sample size. *Rosclass* C type field sites within the *Gagediff* range of ±0.2 ft. had a mean LTO of 1.07 °C (n=12). *Rosclass* B type field sites within the *Gagediff* range of ±0.2 ft. had a mean LTO of 1.95 °C (n=20). Rosclass B and C type field sites within the *Gagediff* range of ±0.2 ft. had a mean LTO of 1.63 °C (n=32).

These covariate ranges allow relative confidence in using raw TIR-derived stream temperatures if used during periods of stream discharge stability and at field sites with channel morphology consistent with *Rosclass* B or C type characteristics. The mean LTO of all *Rosclass* stream types (B, C, and E) within the *Gagediff* range of ±0.1 ft. was 0.8 °C (n=20); however, only *Rosclass* E type field sites within the *Gagediff* range of ±0.1 ft. had a mean LTO of 2.92 °C

(n=4). The high mean LTO of E type field sites under such ideal discharge circumstances underscores the importance of stream morphology and the power of the top model including both covariates.

Comparison to Previous Studies

Much like the Handcock et al. (2006) study, these findings show that Landsat TIR appears to over-predict in-situ temperatures. However, model-produced predictions can correct for fluctuations in discharge, air temperature anomalies, and E type steam reaches, producing accurate in-situ stream temperature predictions for lower order streams. Handcock et al. (2006) found that field sites with a "fraction of a pixel across the stream" had a mean LTO of 7.6 °C. The authors proposed that high LTO is caused by thermal stratification between warm surface water and cooler water deeper in the stream where the loggers were located and is more prevalent in these smaller streams due to lower discharge.

Thermal stratification is not common in well-mixed streams, such as the Anchor River, but it may occur during sunny, warm conditions or slow stream flows when solar heating of the surface layer is not compensated by vertical mixing (Wunderlich 1969). Handcock et al. (2006) used study regions in the inland Pacific Northwest, the "fraction of a pixel across the stream" sites came from the Green River in Washington state. The Green River has a dam-controlled water discharge and passes through largely urban areas. The decreased accuracy of Landsat TIR-derived stream temperature could be a result of any one of these factors. In comparison, the Anchor River has no dam and no urban infrastructure. Dam-controlled stream systems are more prone to vertical stratification due to warming of upstream stagnant water in lakes and reservoirs (Baxter 1977). Urban infrastructure, such as adjacent roads and buildings, can influence TIR as

well by becoming thermal heat sinks. Thus, results from the Green River may not directly reflect dynamics in the Anchor River due to its absence of dams and urban infrastructure.

Handcock et al. (2006) found high LTO variability in smaller streams ranging in size from 5 to 20 m across. In the same study, results found the minimum LTO was 0.2 °C and the maximum 22.2 °C (n=22); however, due to the absence of environmental data, no conclusion as to the cause of variability can be made besides the possibility of difference in geomorphology and hydrologic discharge. The authors postulated the cause of such wide disagreement between Landsat TIR-derived stream temperature and in-situ temperatures in smaller stream systems as: 1) radiometric error due to atmospheric uncertainty, 2) vertical stratification occurring, and 3) the influence and percent of stream bank material (ground cover).

Reconstructing In-situ Temperatures Using Remotely-sensed TIR Products

The narrow range in conditions over which raw Landsat 8 temperature agreed with in-stream water temperatures suggests that without further correction to accommodate stream conditions, it is of limited value for predicting in-situ water temperatures in small order subarctic streams. However, this thesis demonstrates that relatively simple GAM models can employ readily available environmental variables, such as Rosgen Stream Classification type, stream stage, and air temperature data, to accurately correct for LTO. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as:

$$LST_C = LST_{Raw} - LTO_{\hat{y}}$$

Corrected Landsat temperatures were calculated for each site and date for the duration of the study period. The performance of the LST_c model (LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~

Rosclass + f(Airdaily)) was examined on a sub-basin scale by predicting temperatures for the North Fork Anchor River, South Fork Anchor River, and Chakok River basins for comparison with in-situ temperatures (Figures 19-24). This was also done for the whole Anchor River basin (Figure 25-26). Corrected Landsat temperatures across all sub-basins and watershed-wide, had mean absolute error (MAE) and root mean squared error (RMSE) between 1.1 °C and 1.93 °C (Table 5).

Handcock et al. (2006) reported that uncorrected Landsat temperatures did not characterize actual stream temperatures for small streams. This thesis verifies this finding and outlines a modeling framework for predicting corrected Landsat temperatures that performs efficiently enough for landscape-scale stream temperature monitoring.

Applications of Landsat-derived Temperature Information for Small Order Stream Ecology

Landsat 8 TIR can be a useful tool in assessing watershed-scale stream temperature heterogeneity amidst concurrent trends in salmonid size-at-age and abundance in remote watersheds in Alaska. While Landsat 8 TIR cannot provide absolute temperature, or offer fine-scale thermal maps suitable for detecting cold water refugia or ground water influence, it can offer temperatures representative of temporal shifts in broad-scale trends. Recent studies (i.e. Schindler et al. 2015) have suggested that juvenile salmonids utilize larger portions of a watershed than previously thought (i.e. high use rearing areas change yearly), and that shifts in landscape-scale water temperature heterogeneity are representative of processes affecting juvenile salmonids' life histories. Addressing such trends is useful for understanding the effects of climate change on freshwater-stage salmonids and predicting the trajectory of such trends.

Additionally, Landsat TIR is available from the Landsat 5, 6, and 7 satellites, offering the possibility of obtaining stream temperatures from 1982 on, with 16-day frequency across Alaska with an error range under 2℃. Alaska Department of Fish and Game's Anadromous Waters Catalogue has salmonid size and age data along the same time scale, offering the possibility of viewing stream temperature heterogeneity and juvenile salmon size and age information over a 35-year time-period (ADF&G 2015).

Further Studies

In this effort, analysis protocol and field design provided an opportunity to incorporate insight provided by Handcock et al. (2006), and addressed three primary proposed causes of inaccurate Landsat TIR-derived stream temperatures: 1) images were given an atmospheric correction, 2) vertical stratification was assessed using *Rosclass* and *Gagediff,* and 3) ground cover was classified and found to be an insignificant influence on LTO. Overall, efforts here were successful in improving insight about the margin of error of Landsat TIR-derived stream temperatures on the Anchor River watershed and improved understanding about the influence of in-stream and riparian habitat conditions on LTO in high latitude small stream systems. Ultimately, this study sought to investigate whether Landsat TIR derived predictions of in-situ stream temperature are useful for assessing relative temperature trends over large spatial and temporal scales.

Thermal stratification fluctuates seasonally, with clear delineations between layers during the summer, narrower layers in winter, and a "turnover" in the spring and fall when temperature is uniform throughout the water column (Boehrer and Schultze 2008). Expanding this study to view the effects of these covariates on LTO year-round may increase understanding of these interactions. Conducting a "paired study", where one in-situ logger sits on the stream bed and

one on the stream surface would clarify the effects of stratification and the seasonal trends in vertical mixing.

Hydrologic models, such as the Soil, Water Assessment Tool (SWAT), exist that can quantify stratification and even predict it based on stream discharge and air temperature (Santhi et al. 2001; Mulvihill et al. 2009). Use of these models in unison with the methods described here may offer insight into field site and date selection, to further eliminate the influence of thermal stratification.

To assess how rates of warming affect terrestrial areas versus wetted areas would provide understanding on how rates of warming and cooling affect LTO. To test this, a temperature sensor could be placed on dry land adjacent to an in-situ logger to test rates of warming and cooling of land vs. water over different stream reaches with varying morphology.

Applying these methods to other, similar Alaskan watersheds will validate this study as a useful fisheries management tool. As it currently sits, these methods can only speak to relative trends in the Anchor River watershed. Validating these methods on additional systems would make progress toward filling the stream temperature data gap and allow fisheries managers to draw useful correlations between stream temperature flux and juvenile salmonid populations in remote watersheds.

CONCLUSION

This thesis investigated the environmental variables that influence the accuracy of Landsat 8 TIR-derived stream temperatures in the Anchor River basin. Multimodel based inference supported Rosgen Stream Classification (*Rosclass*) and 7 day trends in stream

discharge (*Gagediff*) as being strongly associated with Landsat 8 thermal offset (LTO). Stability in stream discharge (*Gagediff* = ±0.2 ft.) in the week preceding a Landsat TIR data collection event, coupled with stream morphology characteristics consistent with *Rosclass* B and C reaches, were associated with the highest accuracy of Landsat TIR-derived stream temperatures. Across all sites and dates, GAM regressions accurately translated raw Landsat TIR data into temperatures reflecting in-situ water temperatures in small order streams. While regression analyses supported *Rosclass* and *Gagediff* habitat covariates as being influential in explaining LTO, the underlying mechanisms that these variables reflect remains an open question. This study hypothesizes that these variables may provide indication of conditions that lead to stream stratification, or alternatively, may be reflecting other environmental processes leading to stronger temperature differentials between temperature in the stream and that on the surrounding riparian land. Future studies exploring the degree of stratification associated with in-situ stream temperatures and coupled riparian land- and water-based in-situ validation efforts may be useful for further clarifying conditions that lead to accurate Landsat TIR-based stream temperature predictions and improve development of regression-based correction models to reduce LTO. Finally, continued validation efforts in other high latitude watersheds over longer time periods will be important for testing the robustness of methods to use Landsat TIR-based stream temperature predictions, progressing towards a goal of providing a broad-scale tool for fisheries and watershed managers to assess the degree of, and potential impacts from, climate change on Pacific salmon resources in Alaska.

REFERENCES

Alaska Department of Fish and Game (ADF&G). 2013. Chinook salmon stock assessment and research plan, 2013. *Alaska Department of Fish and Game, Special Publication* No. 13-01, Anchorage, Alaska.

Alaska Department of Fish and Game (ADF&G). 2015. Anadromous waters catalog. Alaska Department of Fish and Game, Juneau, Alaska.

Armour, C. L. 1991. Guidance for evaluating and recommending temperature regimes to protect fish. U.S. Fish & Wildlife Service, Fort Collins, CO.

Arntzen, E.V., Geist, D.R. and Dresel, P.E. 2006. Effects of fluctuating river flow on groundwater/surface water mixing in the hyporheic zone of a regulated, large cobble bed river. *River Research and Applications*, *22*(8), 937-946.

Arora, M., Kumar, R., Malhotra, J. and Kumar, N. 2014. Correlations of Stream Flow and Climatic Variables for a Large Glacierized Himalayan Basin. *Journal of Water Resource and Protection*, 6(1), 1326-1334.

Barton, K. 2015. MuMIn for R. R News 1(2):20-25

Baxter, R.M. 1977. Environmental effects of dams and impoundments. *Annual Review of Ecology and Systematics*, *8*(1), 255-283.

Bio, A.M.F., Alkemade, R. and Barendregt, A. 1998. Determining alternative models for vegetation response analysis: a non-parametric approach. *Journal of Vegetation Science* 9, 5–16.

Bjornn, T. C. and Reiser, D. W. 1991. Habitat requirements of salmonids in streams. Pages 83–138 in W. R. Meehan, editor. *Influences of forest and rangeland management on salmonid fishes and their habitats*. American Fisheries Society, Special Publication 19, Bethesda, Maryland.

Blodgett, R.H. and Stanley, K.O. 1980. Stratification, bedforms, and discharge relations of the Platte braided river system, *Nebraska. Journal of Sedimentary Research*, 50(1).

Boehrer, B. and Schultze, M. 2008. Stratification of Lakes. Helmhotz Centre for Environmental Research -UFZ, Magdelburg. *Reviews of Geophysics*, Vol 46.

Boggs, K., Boucher, T.V., Kuo, T.T., Fehringer, D., and Guyer, S. 2014. Vegetation map and classification: Northern, Western and Interior Alaska, 2014 Update. Alaska Natural Heritage Program, University of Alaska Anchorage, Anchorage, Alaska. 80

Brabets, T.P., Wang, B. and Meade, R.H. 2000. Environmental and hydrologic overview of the Yukon River Basin, Alaska and Canada (No. 99-4204). US Dept. of the Interior, US Geological Survey; Branch of Information Services

Budy, P., Baker, M., and Dahle, S.K. 2011. Predicting fish growth potential and identifying water quality constraints: a spatially-explicit bioenergetics approach. *Environmental Management* 48, 691-709.

Bunn, S.E. and Arthington, A.H. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental management*, *30*(4), 492-507.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multi-modal inference: a practical information–theoretic approach. Springer-Verlag, New York, New York, USA.

Calabretta, C.J. and Oviatt, C.A. 2008. The response of benthic macrofauna to anthropogenic stress in Narragansett Bay, Rhode Island: a review of human stressors and assessment of community conditions. *Marine Pollution Bulletin* 56, 1680 -1695.

Chapin, F. S., III, S. F. Trainor, P. Cochran, H. Huntington, C. Markon, M. McCammon, A. D. McGuire, and Serreze, M. 2014. Ch. 22: Alaska. Climate Change Impacts in the United States: The Third National Climate Assessment, J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., *U.S. Global Change Research Program*, 514-536. doi:10.7930/J00Z7150.

Cristea, N. C. and Burges, S. J. 2009. Use of thermal infrared imagery to complement monitoring and modeling of spatial stream temperatures. *Journal of Hydrologic Engineering*, 14, 1080–1090.

Dey, S. 2014. Fluvial Hydrodynamics, GeoPlanet: Earth and Planetary Sciences, DOI: 10.1007/978-3-642-19062-9_9, Springer-Verlag Berlin Heidelberg 2014.

ESRI. 2011. ArcGIS Desktop: Release 10.3.2. Redlands, CA: Environmental Systems Research Institute.

Evenson, D.F. and Carlile, J.K. 2014. Chinook Salmon Abundance and Productivity Trends. *ADF&G Chinook News* 1 (2014): 2. Print.

Forman, R.T. 2014. *Land Mosaics: The Ecology of Landscapes and Regions (1995)* (p. 217). Island Press.

GPS Pathfinder Office. 2014. Trimble Corporation. Version 2. Sunnyvale, CA

Geist, M., Trammell, E.J. and Bogan, D. 2014. AKOATS –Alaska Online Aquatic Temperature Site, an Inventory of Continuous Stream and Lake Temperature Monitoring Stations across Alaska, Summary Report, Alaska Natural Heritage Program, UAA, Anchorage, AK.

Griffiths, J.R. and Schindler, D.E. 2012. Consequences of changing climate and geomorphology for bioenergetics of juvenile sockeye salmon in a shallow Alaskan lake. *Ecology of Freshwater Fish* 21,349-362.

Gu, R., Montgomery, S., and Austin, T.A. 1998. Quantifying the effects of stream discharge on summer river temperature. *Hydrological Sciences Journal*, 43(6), 885-904.

Handcock, R.N., Gillespie, A.R., Cherkauer, K.A, Kay, J.E., Burges, S.J., and Kampf, S.K. 2006. Accuracy and uncertainty of thermal- infrared remote sensing of stream temperatures at multiple spatial scales. *Remote Sensing of Environment* 100 (2006), 427 – 440.

Handcock, R.N., Torgersen, C.E., Cherkauer, K.A., Gillespie, A.R., Tockner, K., Faux, R.N., Tan, J., and Carbonneau, P.E. 2012. Thermal infrared remote sensing of water temperature in riverine landscapes. *Fluvial remote sensing for science and management*, 85-113.

Hoboware. 2014. Hoboware Pro for desktop. Version 5.1. Onset Corporation 2001

Keys, R. 1981. Cubic convolution interpolation for digital image processing. *IEEE transactions on acoustics, speech, and signal processing*, *29*(6), 1153-1160.

Kyle, R.E. and Brabets, T.P. 2001. Water temperature of streams in the Cook Inlet basin, Alaska, and implications of climate change. US Department of the Interior, US Geological Survey.

Lewis, B., Grant W.S., Brenner R.E., and Hamazaki T. 2015. Changes in Size and Age of Chinook Salmon (Oncorhynchus tshawytscha) Returning to Alaska. *PLoSONE, 10* (6): e0130184 doi:10.1371.

Mantua, N.J., Crozier, L.G., Reed, T.E., Schindler, D.E. and Waples, R.S. 2015. Response of Chinook Salmon to climate change. *Nature Climate Change, 5*(7),613-615.

Mauger, S., Shaftel, E.J., Trammell, R., Geist, M., and Bogan, D. 2014. Stream temperature data collection standards and protocol for Alaska: minimum standards to generate data useful for regional-scale analyses. Cook Inletkeeper, Homer, AK and Alaska Natural Heritage Program, UAA, Anchorage, AK. 53.

Mazerolle, M.J. 2006. Improving data analysis in herpetology: using Akaike's Information Criterion (AIC) to assess the strength of biological hypotheses. *Amphibia-Reptilia* 27:169-180

McCullough, D.A. 1999. *A review and synthesis of effects of alterations to the water temperature regime on freshwater life stages of salmonids, with special reference to Chinook salmon* (p. 279). US Environmental Protection Agency, Region 10.

McPhee, M. 2015. Causes and Consequences of Changing Age at Maturity of Female Chinook Salmon in Western Alaska. In *145th Annual Meeting of the American Fisheries Society*. AFS.

Mulvihill, C.I., Baldigo, B.P., Miller, S.J., DeKoskie, D., and DuBois, J. 2009. Bankfull discharge and channel characteristics of streams in New York State: *U.S. Geological Survey Scientific Investigations Report* 2009–5144, 51 p.

National Aeronautics and Space Administration (NASA). 2011. Landsat 7 Science Data Users Handbook. http://landsathandbook.gsfc.nasa.gov/.

Natural Resource Conservation Service (NRCS). 2015. SNOTEL Site Information and Reports: Site 1062. https://wcc.sc.egov.usda.gov/nwcc/site?sitenum=1062

National Weather Service (NWS). 2015. Archival Daily Air Temperature: Anchor Point, AK. http://w2.weather.gov/climate/index.php?wfo=pafc.

Quigley, J.T. and Hinch, S.G. 2006. Effects of rapid experimental temperature increases on acute physiological stress and behavior of stream dwelling juvenile Chinook salmon. *Journal of Thermal Biology* 31: 429-441.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rosgen, D. L. 1996. Applied River Morphology. Wildland Hydrology Books, 1481 Stevens Lake Road, Pagosa Springs, Co. 81147.

Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R. and Hauck, L.M. 2001. Validation of the swat model on a large river basin with point and nonpoint sources. *JAWRA Journal of the American Water Resources Association*, *37*(5), 1169-1188.

Schindler, D.E., Armstrong, J.B., and Reed, T.E. 2015. The portfolio concept in ecology and evolution. *Frontiers in Ecology and the Environment*, *13*(5), 257-263.

Schlosser, I. J. 1991. Stream fish ecology: A landscape perspective. *BioScience* 41: 704–712.

Stephenson, D.B. 2003. Statistical concepts in environmental science. University of Reading, Department of Meteorology.

Symonds, M. R., and Moussalli, A. 2011. A brief guide to model selection, multimodel inference and model averaging in behavioral ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, 65: 13–21.

Torgersen, C.E., Price, D.M., Li, H.W., and McIntosh, B.A. 1999. Multiscale thermal refugia and stream habitat associations of Chinook salmon in northeastern Oregon. *Ecological Applications,* 9:301-319.

Torgersen, C.E., Faux R.N., McIntosh B.A., Poage, N.J., and Norton, D.J. 2001. Airborne thermal remote sensing for water temperature assessment in rivers and streams. *Remote Sensing of the Environment*, 76: 386-398.

U.S. Environmental Protection Agency (EPA). 2003. EPA Region 10 Guidance for Pacific Northwest State and Tribal Temperature Water Quality Standards. EPA 910-B-03-002. Region 10 Office of Water, Seattle, WA.

United States Geological Survey (USGS). 2013. Using the Landsat 8 Product. Online Document, US Department of the Interior Washington D.C. [available at: http://landsat.usgs.gov/Landsat8_Using_Product.php].

United States Geological Survey (USGS). 2015. Landsat 7 Project Website. http://landsat.usgs.gov/

United States Geological Survey (USGS). 2016. National Water Information System. USGS 15239900 Anchor River Gage. https://waterdata.usgs.gov/nwis/uv?site_no=15239900

Vannote, R. R., Minshall, G.W., Cummins, K.W., Sedell, J.R. and Cushing, C.E. 1980. The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences, 37*: 130-137.

Viereck, L.A., Dyrness, C.T., Batten, A.R., and Wenzlick, K.J. 1992. The Alaska vegetation classification. Gen. Tech. Rep. PNW-GTR-286. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 278.

Wood, S.N. 2001. mgcv: GAMs and Generalized Ridge Regression for R. R News 1(2):20-25

Wood, S.N. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J. Amer. Statist. Ass*. 99: 673-686.

TABLES

Table 1. Habitat delineations from the ten 30 m × 30 m in-situ data collection field sites. *Light, Medium* and *Dark* refer to ground cover classifications combining vegetation of similar thermal output. *Water* indicates the percent riverine water within the site. Boggs refers to the dominate vegetation class as outlined by Boggs Indices. PWC is Percent Water Class, which represents the range of percentage of water within the pixel. *Rosclass* refers to the Rosgen Stream Classification System, a method of classifying stream morphology based on width-to-depth ratio, sinuosity, gradient, deposition, entrenchment and channel pattern. Three Rosgen Stream Classifications were used (B, C, and E). Type B sites have a moderately steep channel slope, moderately entrenched, have a cross-section width/depth ratio, low sinuosity, and pool-riffle dominated bed morphology. Type C sites feature a well-developed floodplain, moderate channel slope, slightly entrenched with sequencing riffles and pools, moderately sinuous and characteristic "point bars" within the active channel. Type E sites have very low elevational relief, are slightly entrenched, very low channel width/depth ratios, very high sinuosity and high number of pools.

Field Site	% Light	% Medium	% Dark	% Water	Boggs Indices	PWC	Rosclass
NF-1	44.69	0.00	0.00	55.31	Black spruce-deciduous canopy	3	C
NF-2	0.00	58.83	0.00	41.17	Black spruce-deciduous canopy	2	B
NF-3	0.00	27.14	37.53	35.33	Black spruce-deciduous canopy	2	B
NF-4	0.00	0.00	91.39	8.61	Black spruce-deciduous canopy	1	B
NF-5	0.00	0.00	80.61	19.39	Deciduous forest with low scrub	1	B
CR-1	0.00	72.86	0.00	27.14	Herbaceous	2	E
CR-2	0.00	87.22	0.00	12.78	Herbaceous	1	E
SF-1	0.00	25.55	2.60	71.85	Deciduous forest	3	C
SF-2	0.00	0.00	0.00	100.00	Deciduous forest	4	C
SF-3	36.64	2.00	0.00	61.36	Deciduous forest and herbaceous	3	B

Table 2. Stream discharge and air temperature data from the United States Geologic Survey (USGS) stream gage site located on the South Fork Anchor River (Gage Number: 15239900; Latitude 59.4450, Longitude -151.4511) and the Natural Resource Conservation Service (NRCS) Snotel Site at the Anchor River Divide (Site Number**:** 1062; Latitude 59. 8597, Longitude -151. 3111) from each Landsat TIR data collection event. *Gagedaily* is the daily average gage height (ft.), and *Gageweekly* is the average gage height (ft.) from the seven days preceding the Landsat TIR image. *Gagediff* is the difference between *Gagedaily* and *Gageweekly*. *Airdaily* is daily average air temperature (ºC). *Airweekly* is the average air temperature (ºC) from the seven days preceding the Landsat TIR pass over. *Airdiff* is the difference between *Airdaily* and *Airweekly*.

Date (mm/dd/yyyy)	***Gagedaily* (ft.)**	***Gageweekly* (ft.)**	***Gagediff* (ft.)**	***Airdaily* (°C)**	***Airweekly* (°C)**	***Airdiff* (°C)**
5/30/2015	1.01	1.17	-0.16	15.24	6.93	8.31
6/15/2015	0.94	1.22	-0.28	19.80	8.90	10.90
7/1/2015	0.73	0.73	0.01	8.53	10.89	-2.36
8/11/2015	0.59	0.69	-0.10	10.85	14.11	-3.26
8/18/2015	0.95	0.58	0.37	13.20	10.76	2.44
9/12/2015	0.98	0.96	0.03	6.09	7.83	-1.74

Table 3. Summary of the GAM fit as ranked by the AICc

AICc Variable Importance	**RI**	**N Containing Models**			
Rosclass	0.93	109			
Gagediff*	0.91	120			
Airdaily*	0.52	72			
*Gagedaily**	0.39	63			
*Airdiff**	0.36	61			
Dark	0.3	62			
Medium	0.27	66			
Water	0.24	57			
Light	0.24	52			
AICc Top 10 models	**DF**	**AICc**	**ΔAICc**	**AICc Weight**	**adj. R^2**
LTO~*Rosclass+ f(Gagediff)*	6.77	253.33	0	0.07	0.876
LTO~*Rosclass+ f(Gagediff)+f(Airdaily)*	7.83	253.64	0.3	0.06	0.879
LTO~*Dark + Rosclass + f(Gagediff)*	7.77	254.91	1.58	0.03	0.876
LTO~ *f(Airdaily)+f(Gagedaily)+Rosclass+f(Airdiff)*	8.61	255.19	1.85	0.03	0.878
LTO~ *f(Airdaily)+ Dark + Rosclass + f(Gagediff)*	8.83	255.28	1.95	0.03	0.879
LTO~*f(Gagedaily)+Rosclass+ f(Gagediff)*	7.74	255.44	2.10	0.03	0.878
LTO~*Medium+ Rosclass + f(Gagediff)*	7.76	255.44	2.11	0.03	0.875
LTO~ *Rosclass + f(Gagediff)+Water*	7.76	255.47	2.13	0.03	0.875
LTO~ *f(Gagedaily)+Rosclass+ f(Airdiff)+ f(Gagediff)*	8.81	255.53	2.20	0.02	0.878
LTO~ *f(Airdaily) + f(Gagedaily)+Rosclass+ f(Gagediff)*	8.83	255.58	2.24	0.02	0.878

* *f*(Gagediff) and *f*(Airdaily) indicate covariates were modeled as cubic splines with 4 knots; all other covariates entered GAM models in level form.

** Assumes asymptotically normally distributed coefficient estimates

Table 4. Model formula and variable terms for the top two models.

Model	**Coefficient**	**SE or F[a]**	**P**
LTO ~ *Rosclass* + *f*(*Gagediff*)	.	.	.
Rosclass B	4.44	0.34	<0.001
Rosclass C	-1.38	0.55	0.015
Rosclass E	2.63	0.63	0.001
f(*Gagediff*)	.	127.0	<0.001
LTO ~ *Rosclass* + *f*(*Gagediff*) + *f*(*Airdaily*)	.	.	.
Rosclass B	4.44	0.33	<0.001
Rosclass C	-1.38	0.54	0.014
Rosclass E	2.62	0.62	<0.001
f(*Gagediff*)	.	13.19	<0.001
f(*Airdaily*)	.	1.36	0.249

* *f*(Gagediff) and *f*(Airdaily) indicate covariates were modeled as cubic splines with 4 knots; all other covariates entered GAM models in level form.
** Assumes asymptotically normally distributed coefficient estimates
*** A Standard errors (SE) are reported for level-variables, F-statistic values for smooth terms. Coefficient values are reported only for level terms.

Table 5. Corrected Landsat temperature values error evaluation for each site and date for the duration of the study period. The performance of the LST_c model (LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*)) was examined on a sub-basin scale by predicting temperatures for the North Fork Anchor River, South Fork Anchor River and Chakok River basins for comparison with in-situ temperatures. RMSE is Root Mean Squared Error, which is a measure of the difference between values predicted by a model and the values observed. MAE is Mean Absolute Error, which is a measure of how close predictions are to the eventual outcomes.

	LTO~*Rosclass* + *f*(*Gagediff*)		LTO~*Rosclass* + *f*(*Airdaily*)	
Basin	**MAE**	**RMSE**	**MAE**	**RMSE**
Anchor River	1.38 °C	1.62 °C	1.40 °C	1.56 °C
Sub-Basin				
North Fork Anchor River	1.46 °C	1.82 °C	1.41 °C	1.78 °C
South Fork Anchor River	1.20 °C	1.56 °C	1.10 °C	1.43 °C
Chakok River	1.45 °C	1.78 °C	1.58 °C	1.93 °C

* *f*(Gagediff) and *f*(Airdaily) indicate covariates were modeled as cubic splines with 4 knots; all other covariates entered GAM models in level form.

FIGURES

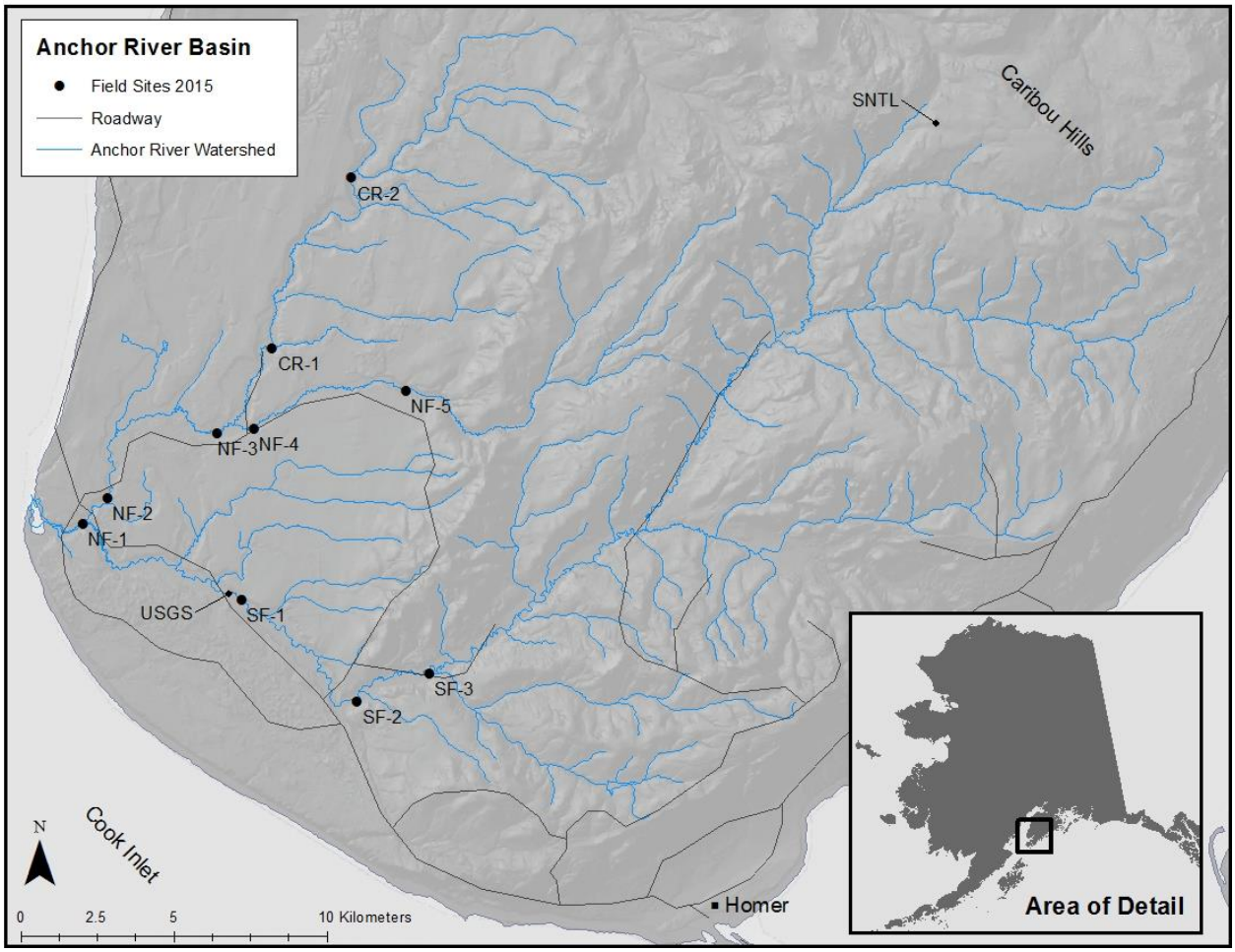


Figure 1. The Anchor River watershed and locations of in-situ temperature collection field sites. U.S. Geological Survey stream discharge gage is annotated as USGS (Gage Number: 15239900; Latitude 59.4450, Longitude -151.4511), Natural Resource Conservation Service Snotel (Site Number**:** 1062; Latitude 59. 8597, Longitude -151. 3111) air temperature data collection site is annotated as SNTL. Field sites labeled SF are on the South Fork Anchor River, NF are on the North Fork Anchor River and CR are on the Chakok River.

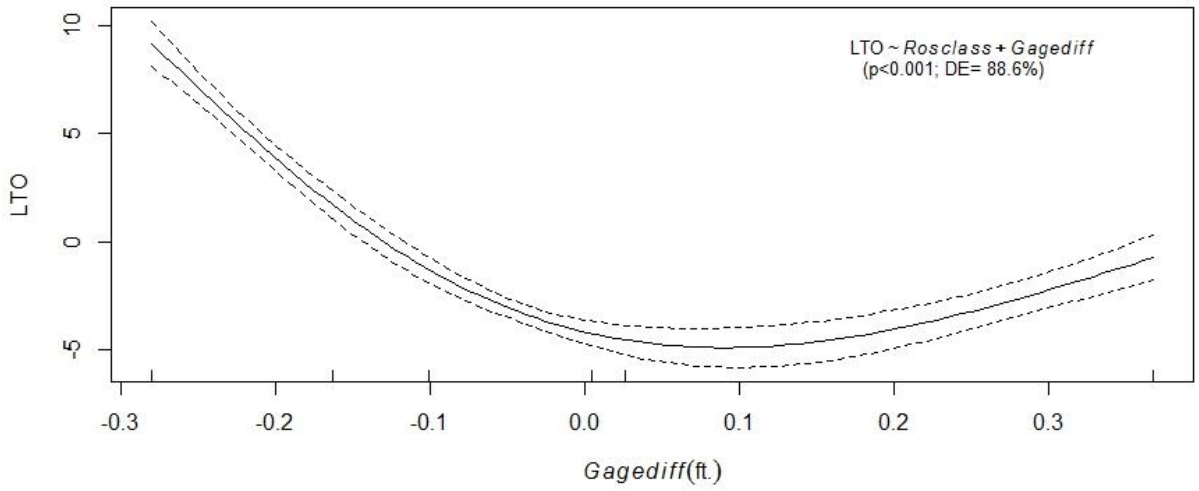


Figure 2. The AICc top-ranked GAM (LTO ~ *Rosclass* + *f*(*Gagediff*)) displaying the relationship between LTO and *Gagediff*. *f*(Gagediff) indicates Gagediff entered the model as a cubic spline with four knots. DE refers to percent deviance explained, a measure of goodness of fit of a model.

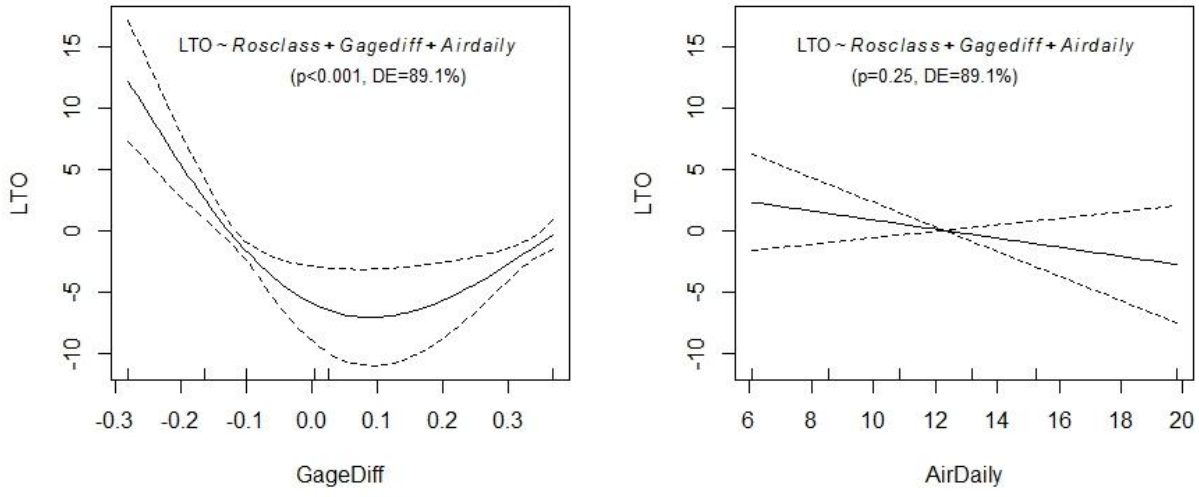


Figure 3. The AICc second-ranked GAM (LTO ~ *Rosclass* + *f(Gagediff)* + *f(Airdaily)*) displaying the *Gagediff* and *Airdaily* curves, both run as smoothing functions with 4 knots. DE refers to percent deviance explained, a measure of goodness of fit of a model.

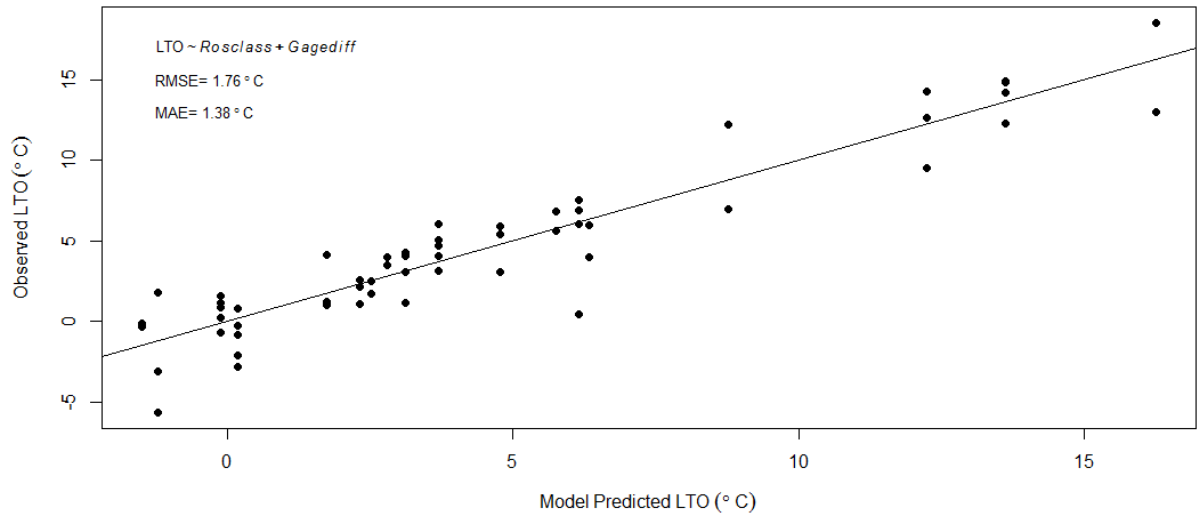


Figure 4. GAM-produced LTO predictions from the top-ranked model (LTO ~ *Rosclass* + *f(Gagediff)*) compared to observed LTO across all dates and field sites. RMSE is Root Mean Squared Error, which is a measure of the difference between values predicted by a model and the values observed. MAE is Mean Absolute Error, which is a measure of how close predictions are to the eventual outcomes.

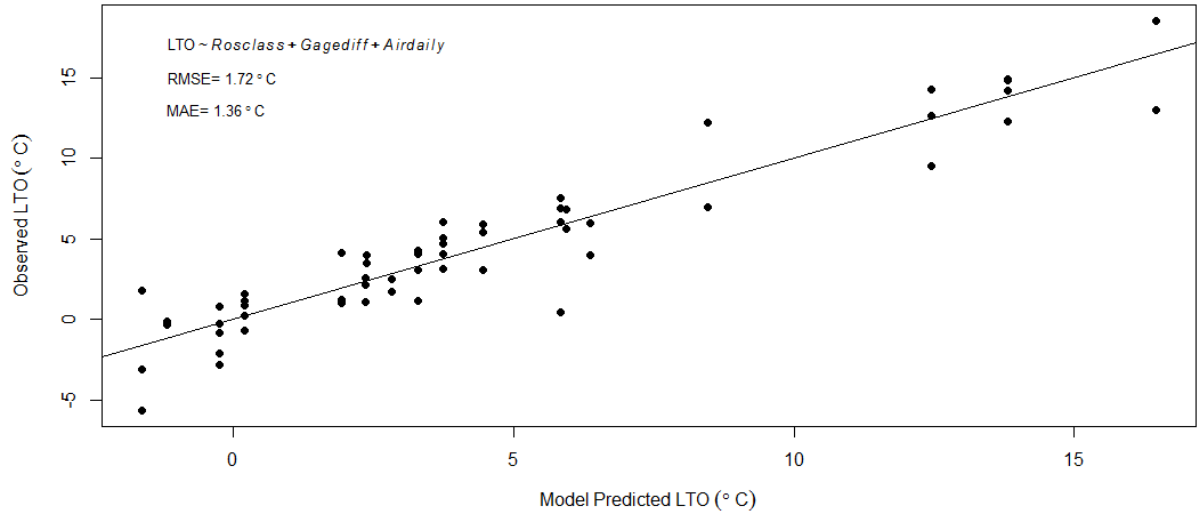


Figure 5. GAM-produced LTO predictions from the second-ranked model (LTO ~ *Rosclass* + *f(Gagediff)* + *f(Airdaily)*) compared to observed LTO across all dates and field sites. RMSE is Root Mean Squared Error, which is a measure of the difference between values predicted by a model and the values observed. MAE is Mean Absolute Error, which is a measure of how close predictions are to the eventual outcomes.

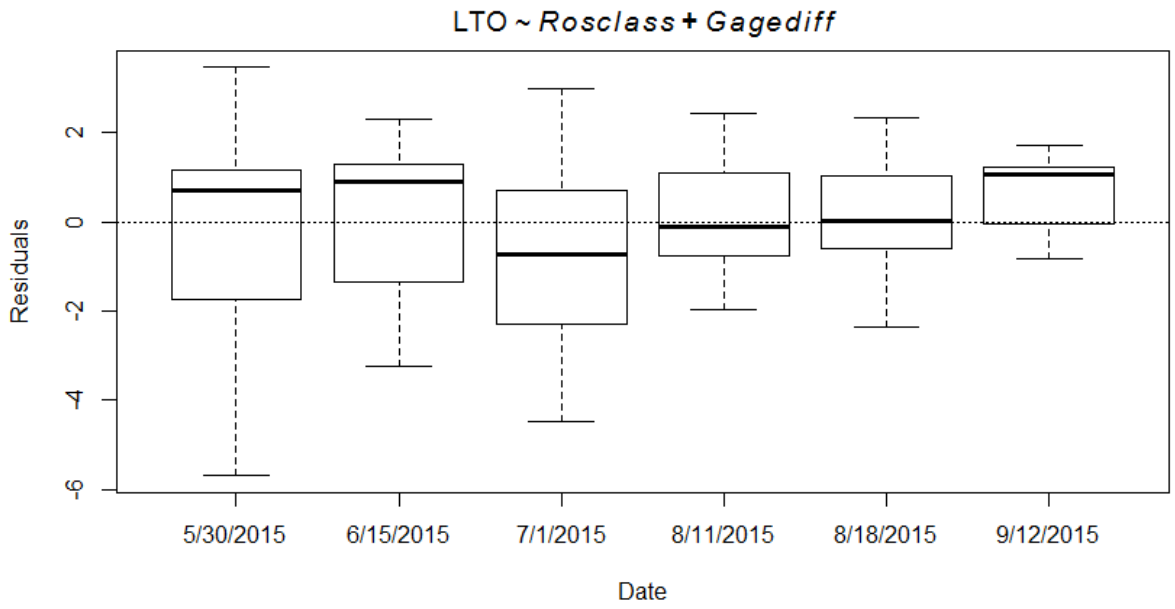


Figure 6. Residuals from the top-ranked GAM (LTO ~ *Rosclass* + *f(Gagediff)*) by Landsat 8 TIR collection date.

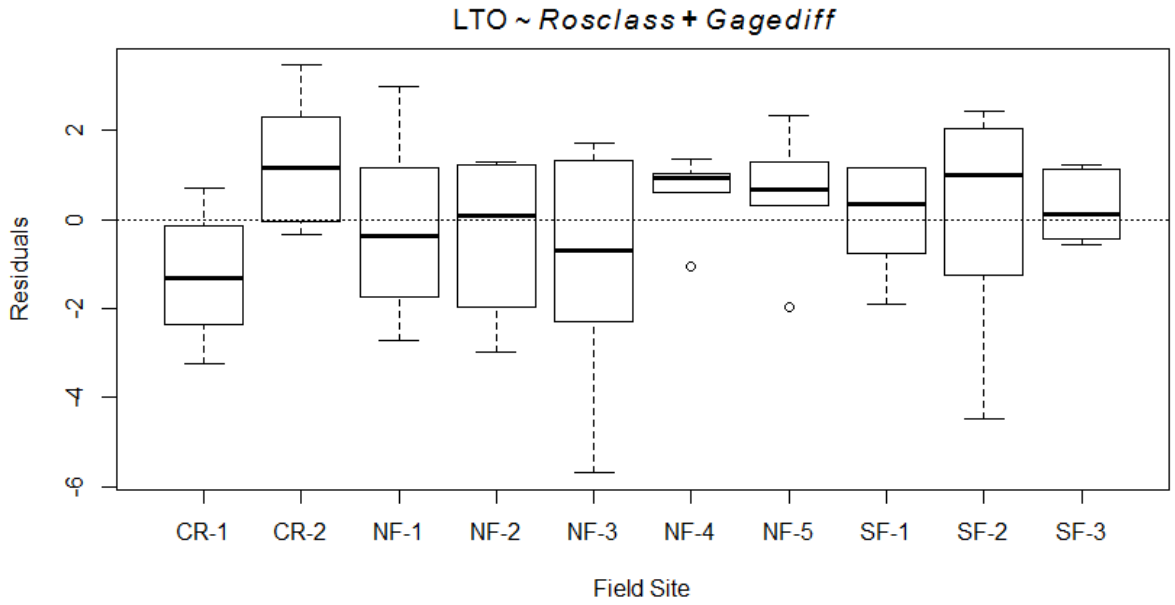


Figure 7. Residuals from the top-ranked GAM (LTO ~ *Rosclass* + *f(Gagediff)*) by field site. CR=Chakok River, NF=North Fork and SF=South Fork. Sites are displayed from downstream to upstream.

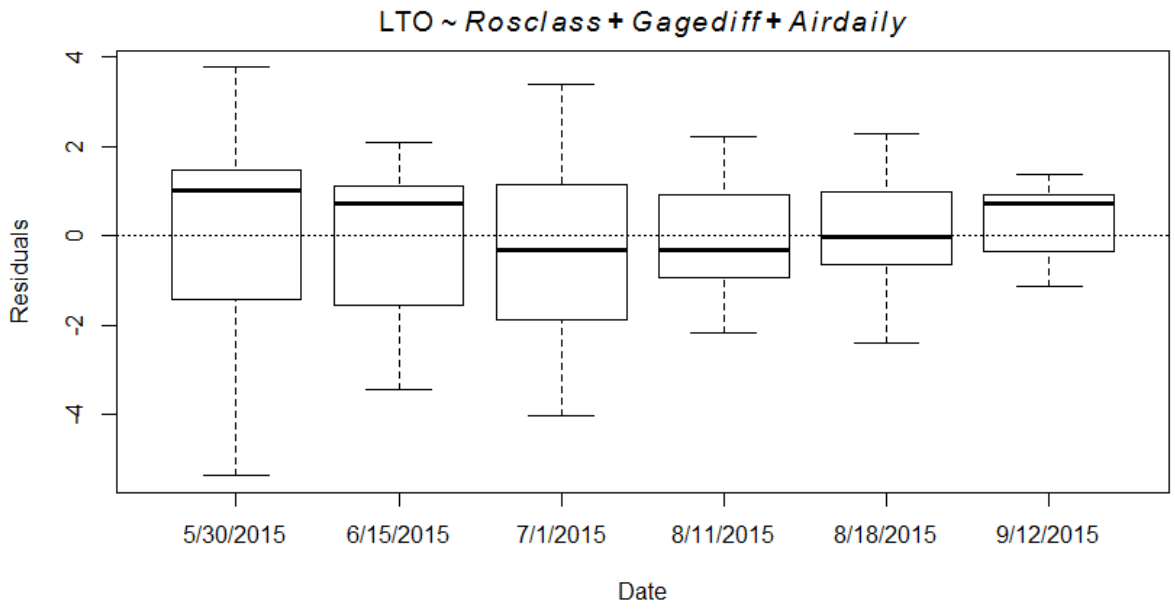


Figure 8. Residuals from the second-ranked GAM (LTO ~ *Rosclass* + *f(Gagediff)* + *f(Airdaily)*) by Landsat TIR collection date.

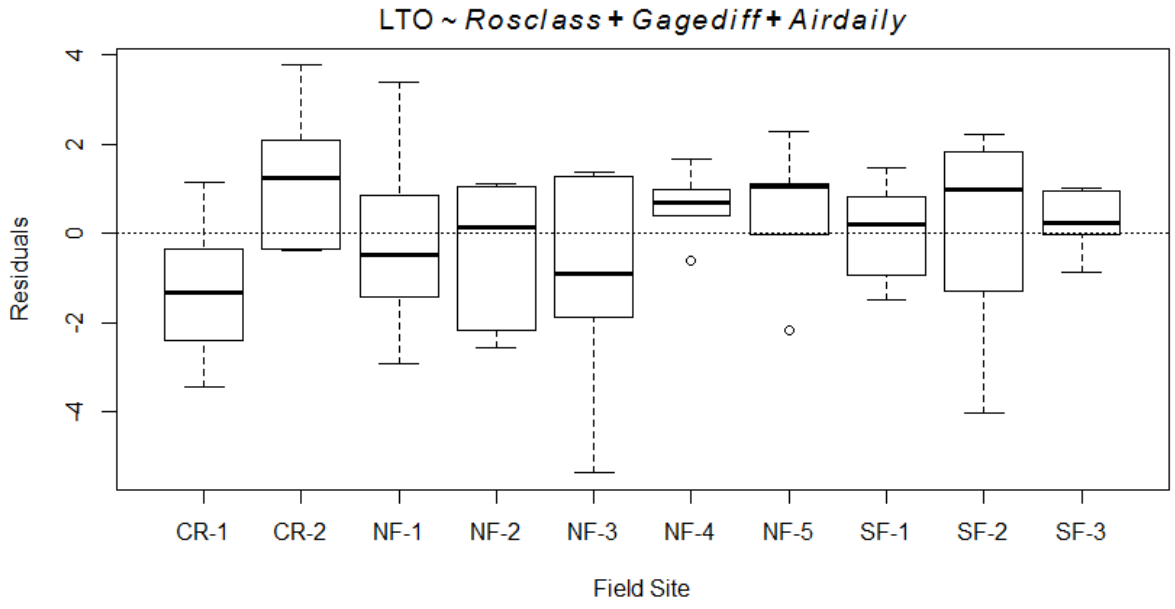


Figure 9. Residuals from the second-ranked GAM (LTO ~ *Rosclass* + *f(Gagediff)* + *f(Airdaily)*) by field site. CR=Chakok River, NF=North Fork and SF=South Fork. Sites are displayed from downstream to upstream.

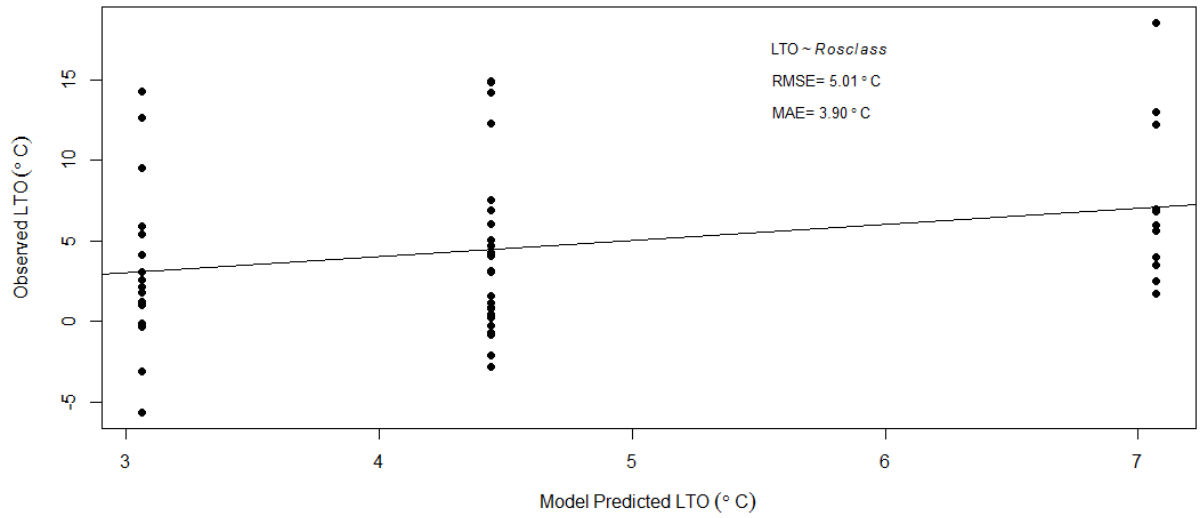


Figure 10. GAM-produced LTO predictions from running the mode with Rosclass (LTO ~ *Rosclass)* as the only variable compared to observed LTO across all dates and field sites. RMSE is Root Mean Squared Error, which is a measure of the difference between values predicted by a model and the values observed. MAE is Mean Absolute Error, which is a measure of how close predictions are to the eventual outcomes.

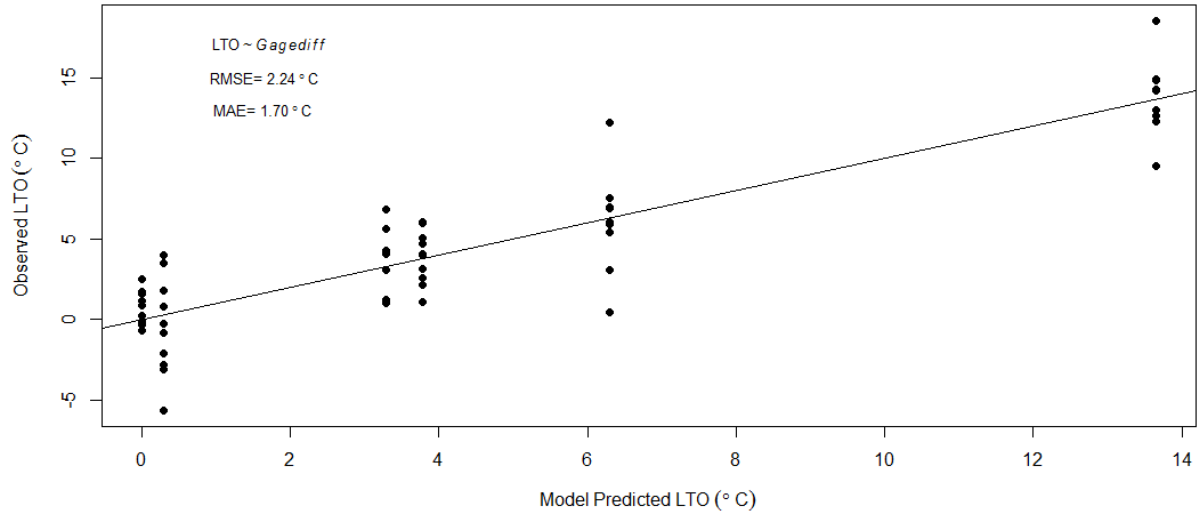


Figure 11. GAM-produced LTO predictions from running the mode with *Gagediff* (LTO ~ *f(Gagediff)*) as the only variable compared to observed LTO across all dates and field sites. RMSE is Root Mean Squared Error, which is a measure of the difference between values predicted by a model and the values observed. MAE is Mean Absolute Error, which is a measure of how close predictions are to the eventual outcomes.

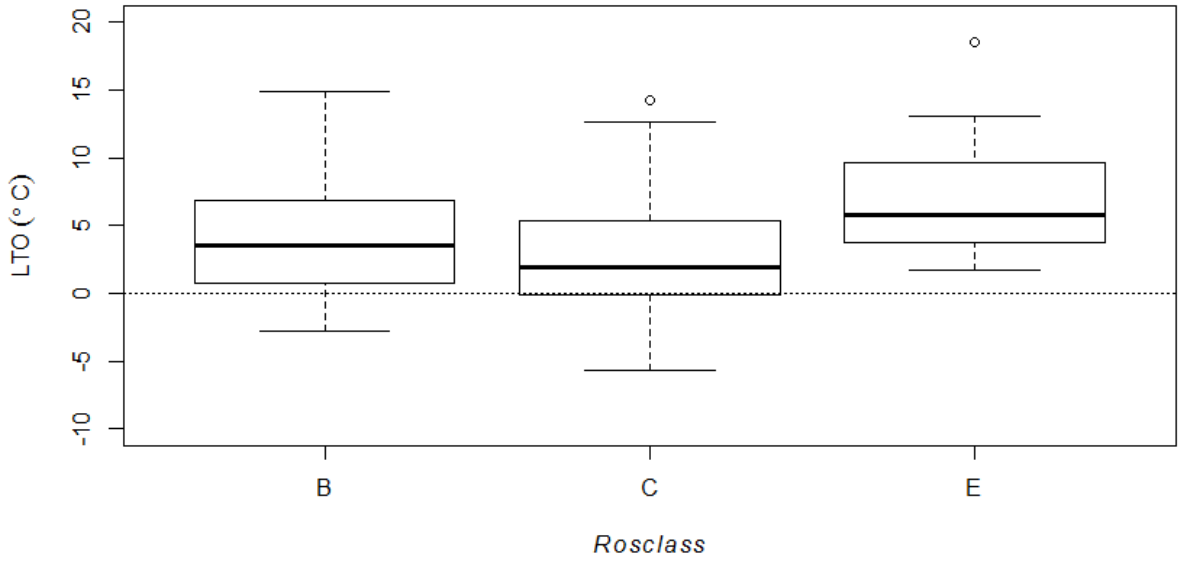


Figure 12. Raw LTO distributions across all dates and field sites by *Rosclass.*

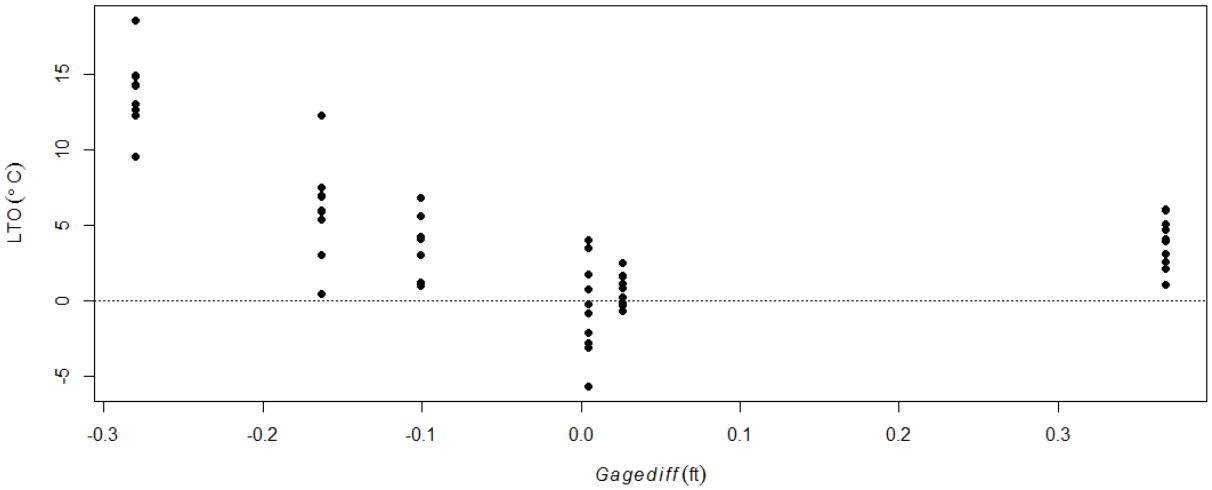


Figure 13. Raw LTO distribution across all dates and sites by *Gagediff.*

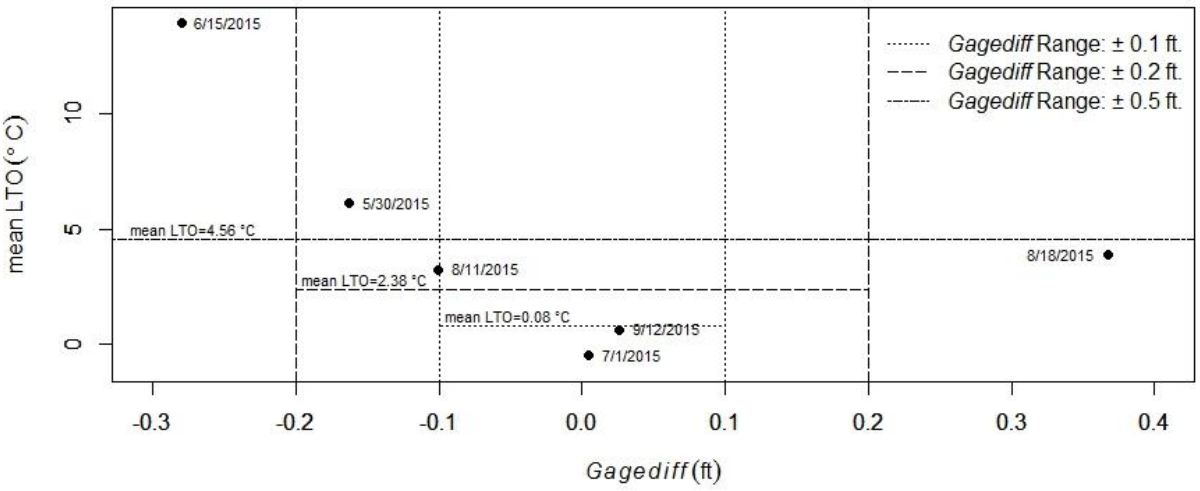


Figure 14. Mean LTO by date and *Gagediff.* Each mean is labeled by Landsat 8 TIR collection date. *Gagediff* ranges and their associated mean LTO temperature values are included. As the *Gagediff* ranges grows, mean LTO grows concurrently.

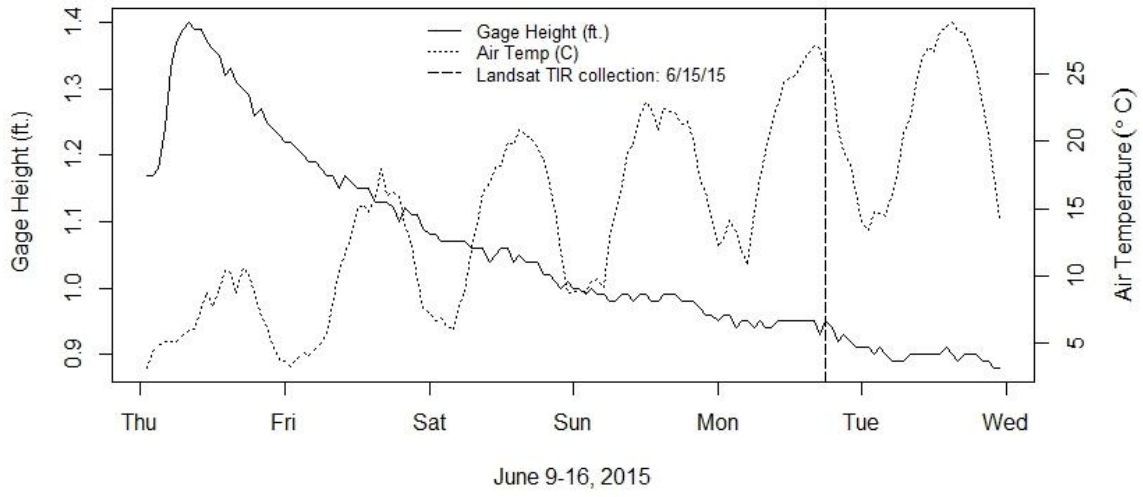


Figure 15. Stream discharge and air temperature in the week preceding the 6/15/2015 Landsat 8 TIR collection. This week experienced peak air temperature observed during the study (28.8 ℃) and declining stream discharge (*Gagediff*= -0.28 ft.).

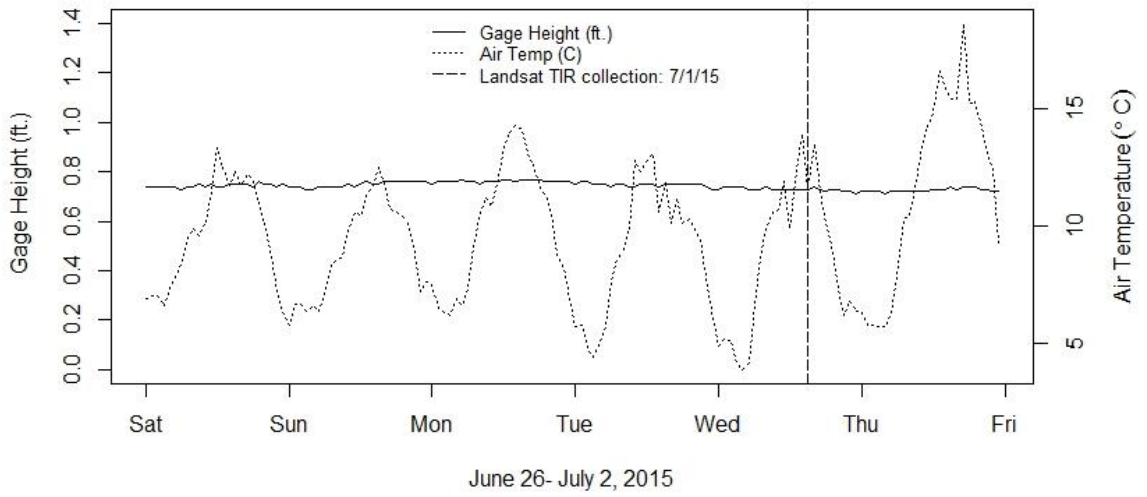


Figure 16. Stream discharge and air temperature in the week preceding the 7/1/2015 Landsat 8 TIR collection**.** Air temperature was within average range (5-20 ℃) and stream discharge was extremely stable (*Gagediff* = 0.005 ft).

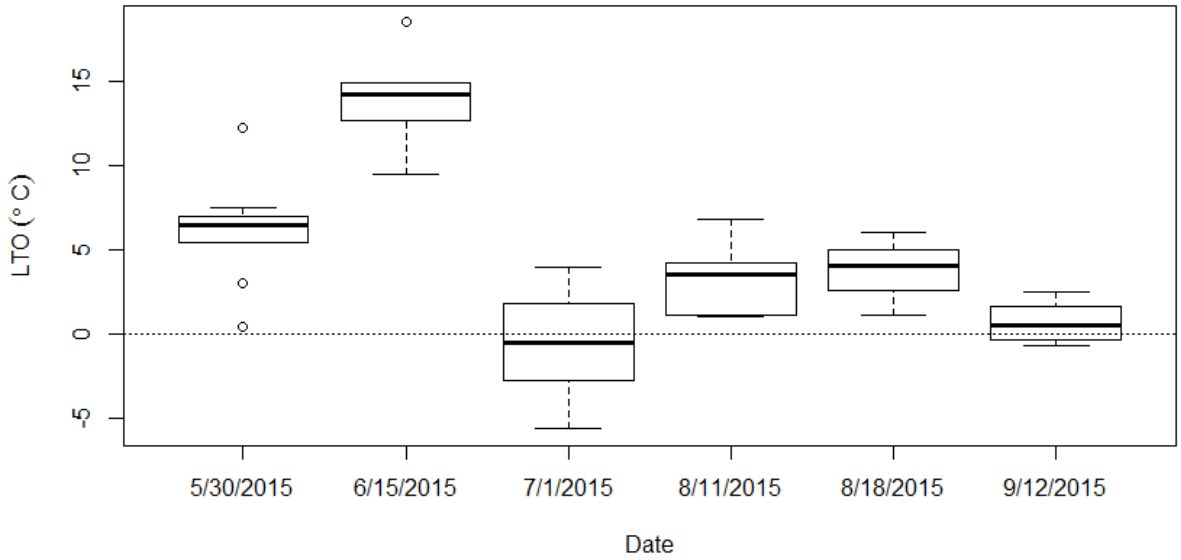


Figure 17. Raw LTO across Landsat 8 TIR collection dates at all field sites**.**

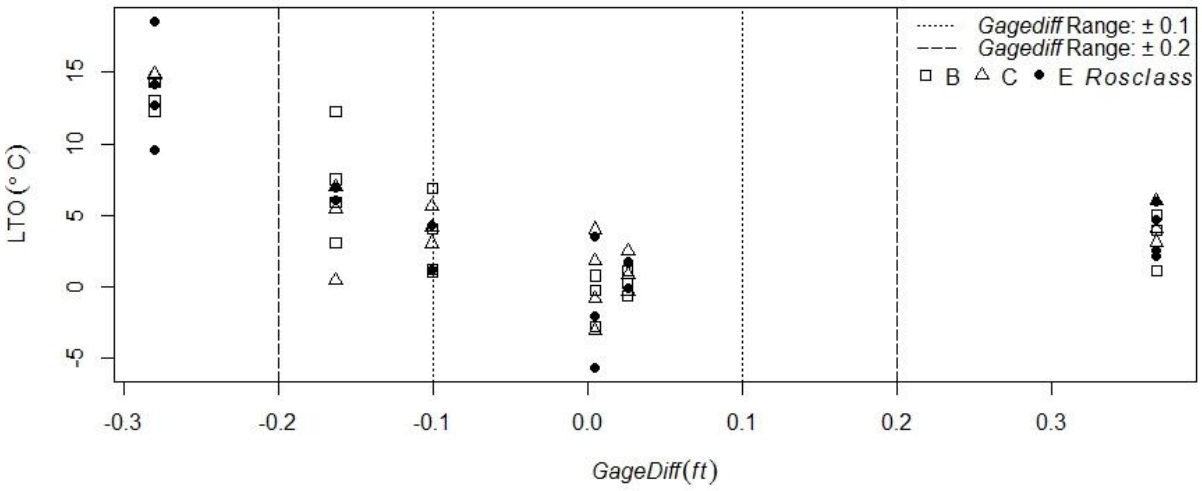


Figure 18. Raw LTO across all dates and field sites with *Gagediff* (x-axis) and *Rosclass* (plot symbols). The *Gagediff* ranges of ±0.1 ft. and ±0.2 ft. are marked as dashed lines.

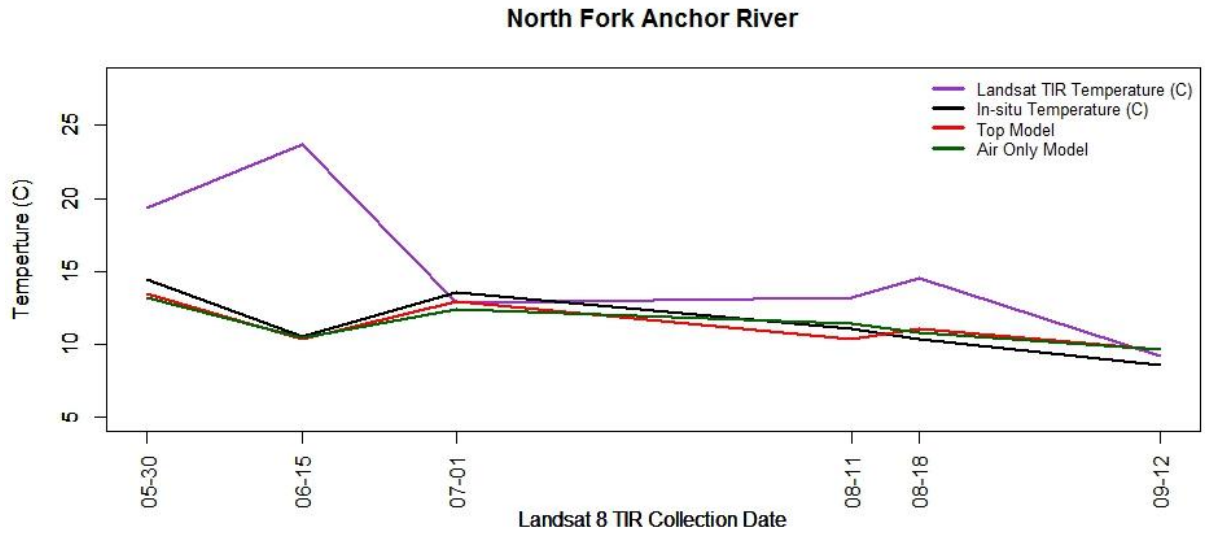


Figure 19. North Fork Anchor River model-produced TIR corrections. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

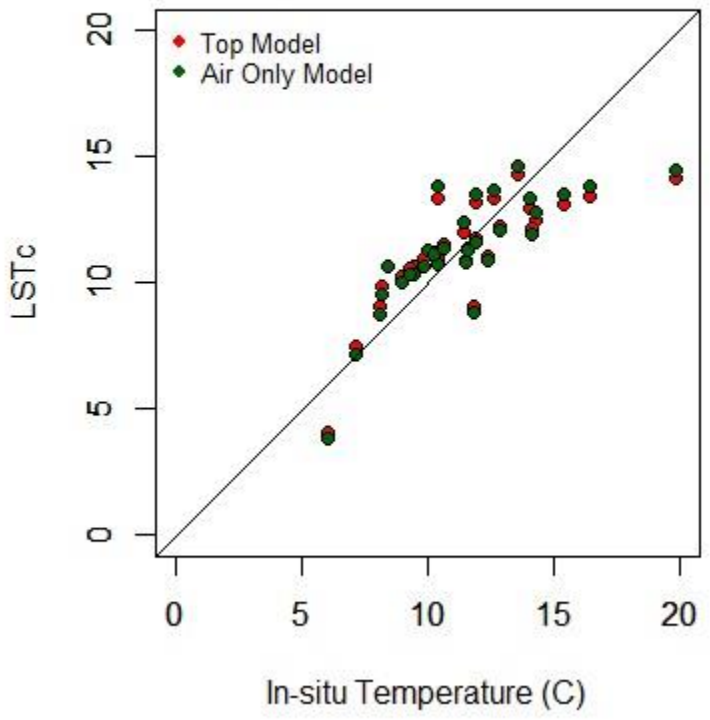


Figure 20. North Fork Anchor River model-produced TIR corrections and in-situ temperature along the 1:1 line. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

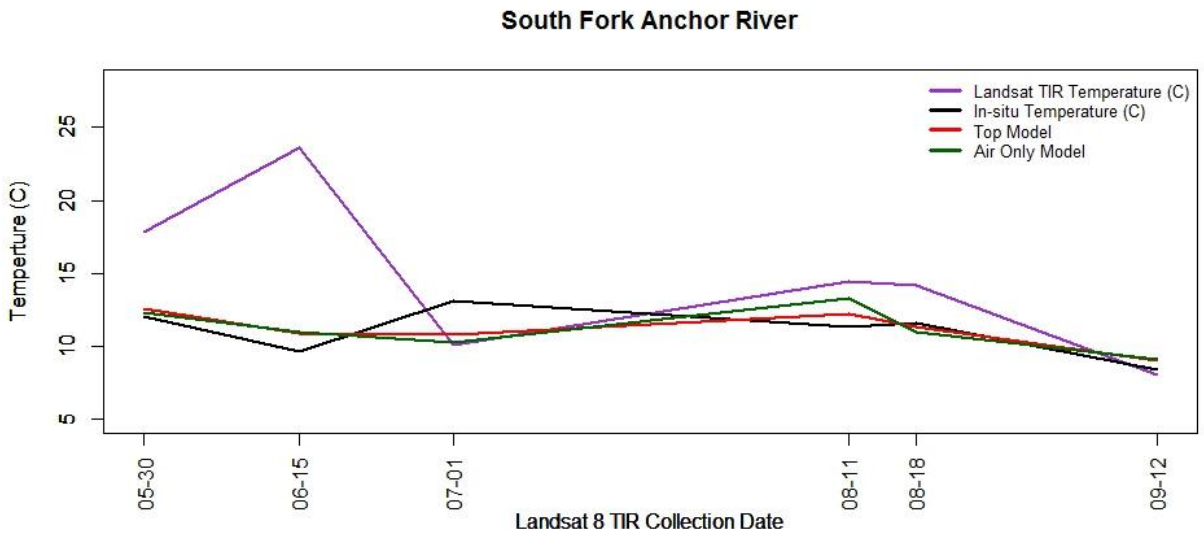


Figure 21. South Fork Anchor River model-produced TIR corrections. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

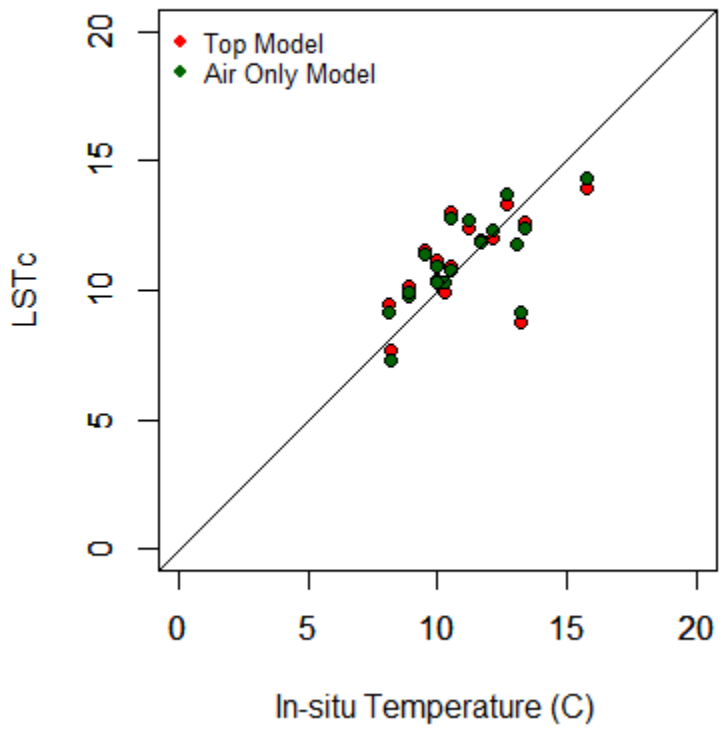


Figure 22. South Fork Anchor River model-produced TIR corrections and in-situ temperature along the 1:1 line. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

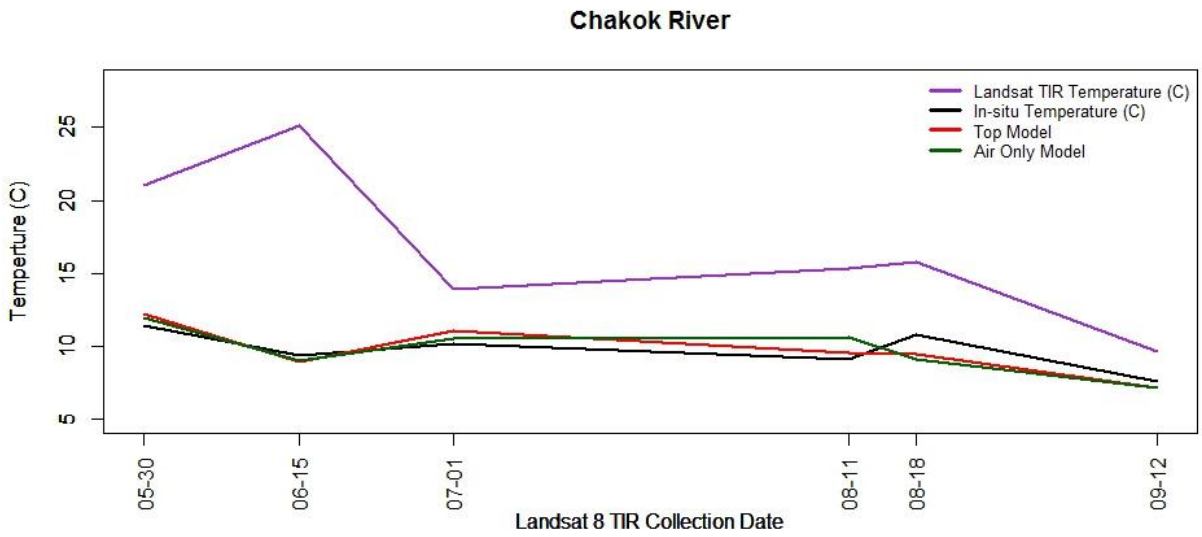


Figure 23. Chakok River model-produced TIR corrections. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

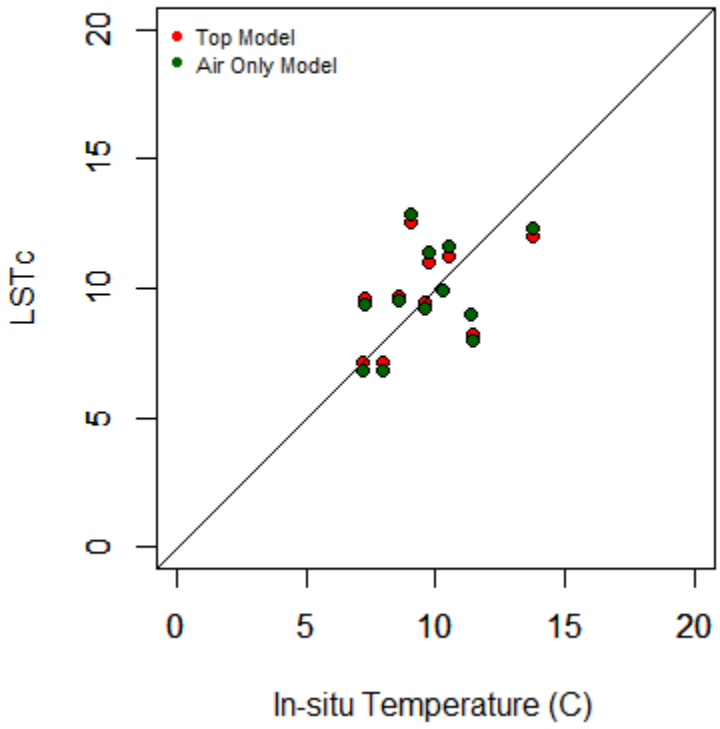


Figure 24. Chakok River model-produced TIR corrections and in-situ temperature along the 1:1 line. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

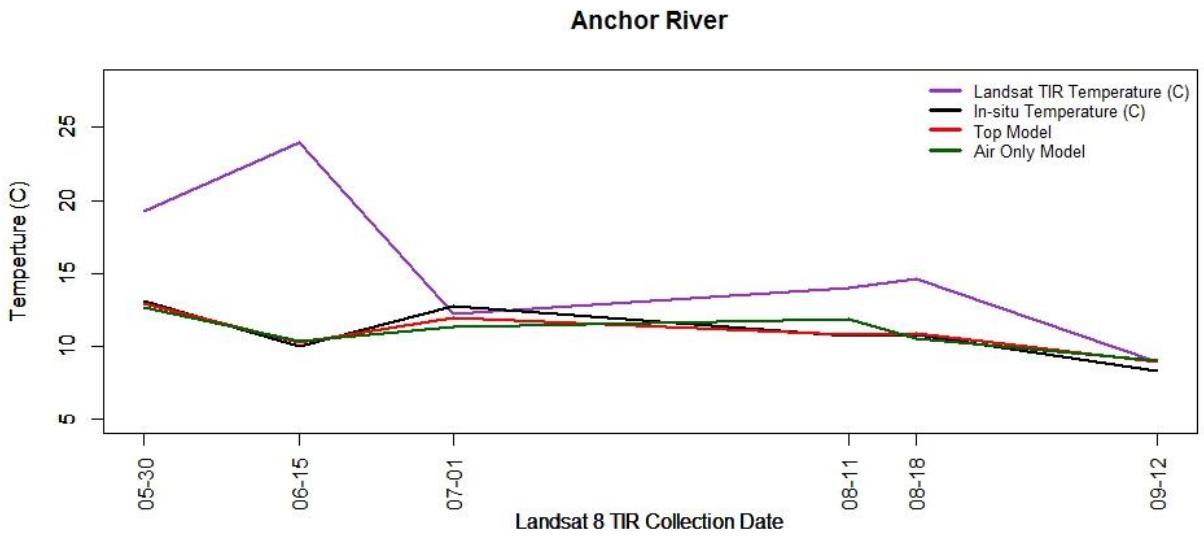


Figure 25. Anchor River model-produced TIR corrections. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

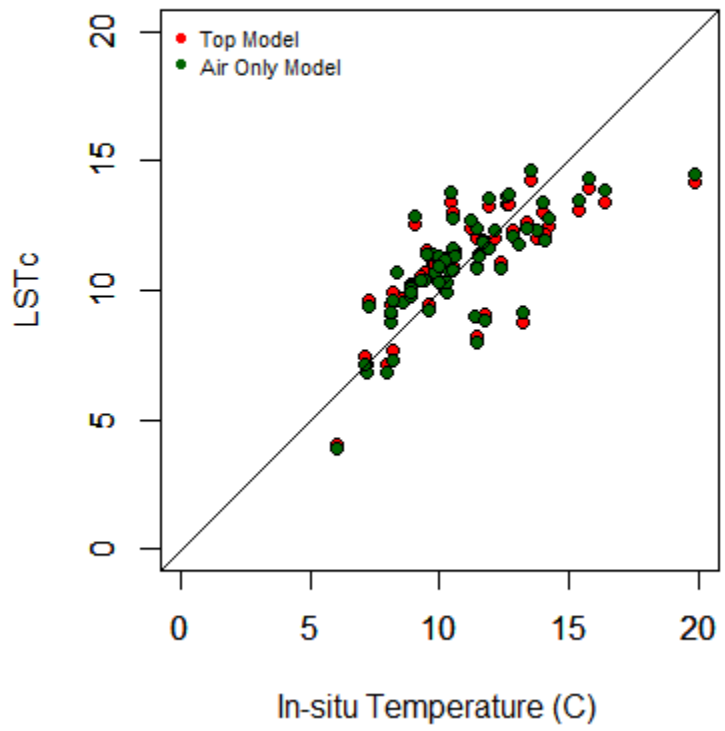


Figure 26. Anchor River model-produced TIR corrections and in-situ temperature along the 1:1 line. A corrected Landsat temperature value (LST_C) can be calculated by subtracting the model-produced LTO predictions ($LTO_{\hat{y}}$) from the corresponding raw Landsat temperature value (LST_{Raw}) as: $LST_C = LST_{Raw} - LTO_{\hat{y}}$. Corrected Landsat temperatures were calculated for each site and date for the duration of the study period using predication from two LST_c models:(LTO ~ *Rosclass* + *f*(*Gagediff*) and LTO ~ *Rosclass* + *f*(*Airdaily*).

APPENDIX

Appendix 1. Quantized and calibrated scaled Digital Numbers (DN) representing the multispectral image data (USGS 2013).

$$L_\lambda = M_L Q_{cal} + A_L$$

Where:

L_λ = TOA spectral radiance (Watts/(m2 * srad * μm))
M_L = Band-specific multiplicative rescaling factor from the metadata (RADIANCE_MULT_BAND_x, where x is the band number)
A_L = Band-specific additive rescaling factor from the metadata (RADIANCE_ADD_BAND_x, where x is the band number)

Q_{cal} = Quantized and calibrated standard product pixel values (DN)

$$T = \frac{In((K1/L\lambda)+1))}{K_2}$$

Where:

T = At-satellite brightness temperature (K)
L_λ = TOA spectral radiance (Watts/(m2 * srad * μm))
K_1 = Band-specific thermal conversion constant from the metadata (K1_CONSTANT_BAND_x, where x is the band number, 10 or 11)
K_2 = Band-specific thermal conversion constant from the metadata (K2_CONSTANT_BAND_x, where x is the band number, 10 or 11)

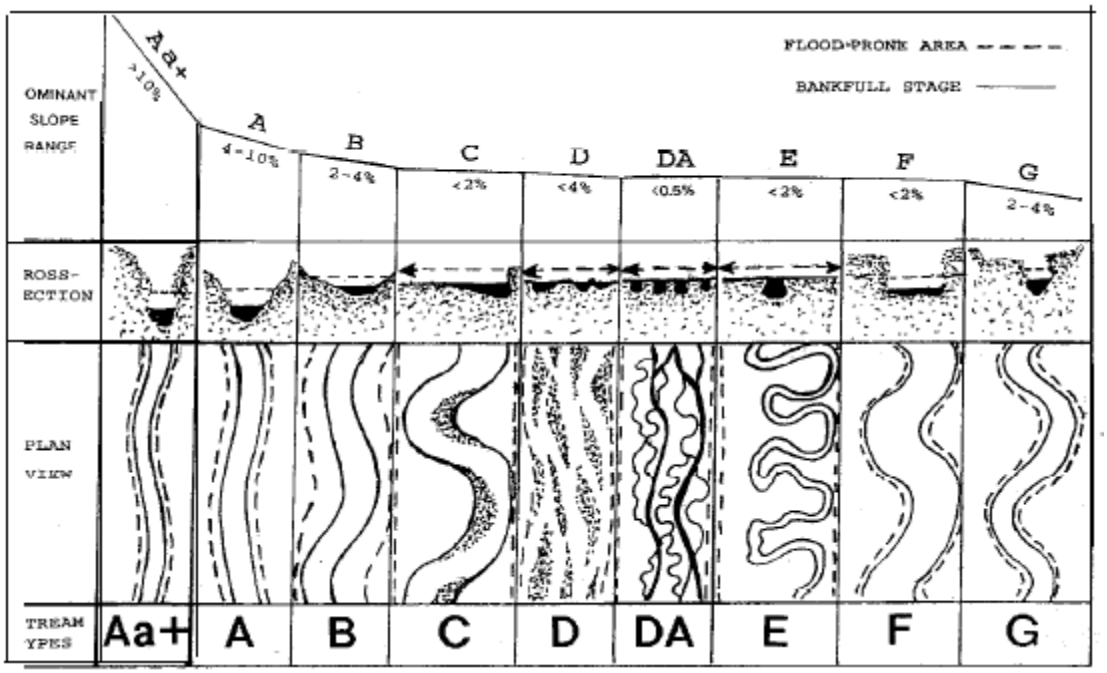
Appendix 2. The atmospheric correction to account for Earth's emitted radiation (Handcock et al. 2012).

$$Lg(\lambda) = \frac{Ls(\lambda) - Lp(\lambda)}{\tau(\lambda)}$$

Where:

Lg= land-leaving spectral radiance at a particular wavelength (Wm−2μm−1sr−1)
Ls= sensor spectral radiance (Wm−2μm−1sr−1)
Lp= path spectral radiance (Wm−2μm−1sr−1)
τ= transmissivity
λ= is the wavelength of the sensor

Appendix 3. The Rosgen Stream Classification field guide (reproduced from Rosgen 1996). This guide graphically describes each level 1 classification by slope, cross-section and valley view.



Appendix 4. Generalized Additive Model data and R script for model calls.

R script to fit the top two GAM models using data provided in Table A4.1 is as follows:

```
# R version 3.3.3 (2017-03-06)
# mgcv version 1.8-17

# Read in data after setting relevant working directory
   setwd("your_directory")
   dat <- read.csv ("data_from_tableA4.1.csv",header=T)
# load mgcv
   library(mgcv)
# Fit AICc top model:
   mod1 <- gam(LTO~RosClass+ s(GageDiff,k=4), data=dat)
   summary(mod1)
# Fit second best AICc model:
   mod2 <- gam(LTO~RosClass+ s(GageDiff,k=4) + s(AirDaily,k=4), data=dat)
   summary(mod2)
# Fit the model with RosClass and AirDaily only
   mod3 <- gam(LTO~RosClass+ s(AirDaily,k=4), data=dat)
   summary(mod3)
```

Table A4.1 Cleaned and processed data for Generalized Additive Modeling[a]

SiteCode	Water	Light	Medium	Dark	LTO	DateTime	RosClass	PWC	GageDaily	GageWeekly	GageDiff	AirDaily	AirWeekly	AirDiff
NF-1	55.31148	44.68852	0	0	3.0233	5/30/2015	C	3	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
NF-2	41.17157	0	58.82843	0	6.895	5/30/2015	B	2	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
NF-3	35.33333	0	27.13889	37.52778	0.4516	5/30/2015	B	2	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
NF-4	8.611111	0	0	91.38889	7.4984	5/30/2015	B	1	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
NF-5	19.38889	0	0	80.61111	6.895	5/30/2015	B	1	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
CR-1	27.13889	0	72.86111	0	6.9691	5/30/2015	E	2	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
CR-2	12.77778	0	87.22222	0	12.2372	5/30/2015	E	1	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
SF-3	71.85126	0	25.54978	2.59896	6.002	5/30/2015	B	3	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
SF-2	100	0	0	0	5.4063	5/30/2015	C	4	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
SF-1	61.36	36.64	2	0	5.9175	5/30/2015	C	3	1.012083	1.17494048	-0.16286	15.2375	6.931548	8.305952
NF-1	55.31148	44.68852	0	0	9.5342	6/15/2015	C	3	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
NF-2	41.17157	0	58.82843	0	14.9246	6/15/2015	B	2	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
NF-3	35.33333	0	27.13889	37.52778	12.2863	6/15/2015	B	2	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
NF-4	8.611111	0	0	91.38889	14.2114	6/15/2015	B	1	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
NF-5	19.38889	0	0	80.61111	14.9246	6/15/2015	B	1	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
CR-1	27.13889	0	72.86111	0	13.015	6/15/2015	E	2	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
CR-2	12.77778	0	87.22222	0	18.5543	6/15/2015	E	1	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
SF-3	71.85126	0	25.54978	2.59896	14.858	6/15/2015	B	3	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
SF-2	100	0	0	0	14.2823	6/15/2015	C	4	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
SF-1	61.36	36.64	2	0	12.6547	6/15/2015	C	3	0.942917	1.22267857	-0.27976	19.8	8.90119	10.89881
NF-1	55.31148	44.68852	0	0	1.7695	7/1/2015	C	3	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
NF-2	41.17157	0	58.82843	0	-2.8157	7/1/2015	B	2	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
NF-3	35.33333	0	27.13889	37.52778	-2.114	7/1/2015	B	2	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
NF-4	8.611111	0	0	91.38889	-0.8698	7/1/2015	B	1	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
NF-5	19.38889	0	0	80.61111	0.7621	7/1/2015	B	1	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
CR-1	27.13889	0	72.86111	0	3.5133	7/1/2015	E	2	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
CR-2	12.77778	0	87.22222	0	3.9852	7/1/2015	E	1	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
SF-3	71.85126	0	25.54978	2.59896	-0.2702	7/1/2015	B	3	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
SF-2	100	0	0	0	-5.67162	7/1/2015	C	4	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
SF-1	61.36	36.64	2	0	-3.0995	7/1/2015	C	3	0.730417	0.72541667	0.005	8.529167	10.8869	-2.35774
NF-1	55.31148	44.68852	0	0	1.2103	8/11/2015	C	3	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
NF-2	41.17157	0	58.82843	0	1.14007	8/11/2015	B	2	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
NF-3	35.33333	0	27.13889	37.52778	3.0348	8/11/2015	B	2	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
NF-4	8.611111	0	0	91.38889	4.0559	8/11/2015	B	1	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
NF-5	19.38889	0	0	80.61111	1.14007	8/11/2015	B	1	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
CR-1	27.13889	0	72.86111	0	5.589	8/11/2015	E	2	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
CR-2	12.77778	0	87.22222	0	6.8461	8/11/2015	E	1	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
SF-3	71.85126	0	25.54978	2.59896	4.2579	8/11/2015	B	3	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
SF-2	100	0	0	0	4.1566	8/11/2015	C	4	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
SF-1	61.36	36.64	2	0	0.9915	8/11/2015	C	3	0.587083	0.68757576	-0.10049	10.85	14.10909	-3.25909
NF-1	55.31148	44.68852	0	0	2.114686	8/18/2015	C	3	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
NF-2	41.17157	0	58.82843	0	3.112142	8/18/2015	B	2	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
NF-3	35.33333	0	27.13889	37.52778	5.033159	8/18/2015	B	2	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
NF-4	8.611111	0	0	91.38889	4.716448	8/18/2015	B	1	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
NF-5	19.38889	0	0	80.61111	6.026026	8/18/2015	B	1	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
CR-1	27.13889	0	72.86111	0	3.966416	8/18/2015	E	2	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
CR-2	12.77778	0	87.22222	0	5.986457	8/18/2015	E	1	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
SF-3	71.85126	0	25.54978	2.59896	4.063837	8/18/2015	B	3	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
SF-2	100	0	0	0	1.077316	8/18/2015	C	4	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
SF-1	61.36	36.64	2	0	2.565439	8/18/2015	C	3	0.95125	0.58375	0.3675	13.2	10.76012	2.439881
NF-1	55.31148	44.68852	0	0	-0.33228	9/12/2015	C	3	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
NF-2	41.17157	0	58.82843	0	1.113781	9/12/2015	B	2	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
NF-3	35.33333	0	27.13889	37.52778	1.589899	9/12/2015	B	2	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
NF-4	8.611111	0	0	91.38889	0.835816	9/12/2015	B	1	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
NF-5	19.38889	0	0	80.61111	0.193905	9/12/2015	B	1	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
CR-1	27.13889	0	72.86111	0	1.694735	9/12/2015	E	2	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
CR-2	12.77778	0	87.22222	0	2.470946	9/12/2015	E	1	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
SF-3	71.85126	0	25.54978	2.59896	-0.6744	9/12/2015	B	3	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
SF-2	100	0	0	0	-0.13913	9/12/2015	C	4	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512
SF-1	61.36	36.64	2	0	-0.34406	9/12/2015	C	3	0.98125	0.95535714	0.025893	6.091667	7.826786	-1.73512

[a]Field codes:

SiteCode: In-situ and Landsat temperature pixel sites.

Water: Percent water in each pixel.

Light: Percent light groundcover in each pixel.

Medium: Percent medium groundcover in each pixel.

Dark: Percent dark groundcover in each pixel.

LTO: Landsat 8 TIR image-derived temperature – in situ stream temperature.

Rosclass: Rosgen Stream Classification level I class type (A-F).

PWC: Percent water class.

Gagedaily: Daily stream stage average from Landsat image acquisition date and time.

Gageweekly: Weekly stream stage average from Landsat image acquisition date and time.

Gagediff: Daily average minus weekly average, this provides context for fluctuation.

Airdaily: Daily average air temperature from Landsat image acquisition date and time.

Airweekly: Weekly average air temperature from Landsat image acquisition date and time.

Airdiff: Daily average minus weekly average, this provides context for fluctuation.