Eastern Washington Nomograph Project Report

Submitted to:
Washington Department of Natural Resources,
Forest Practices Division
1111 Washington Street SE
Olympia, WA 98504

Submitted by:
Domoni Glass
Glass Environmental Consulting, Inc.
15208 Goodrich Dr NW
Gig Harbor, WA 98329

January 7, 2005

Contents

List of Figures

List of Tables

This page intentionally blank

EXECUTIVE SUMMARY

In the Pacific Northwest (PNW), stream temperature is an important determinant of water quality and thus the health of aquatic ecosystems. Consequently, both the natural variations in stream temperature and the human activities that change the thermal conditions of surface waters are significant concerns for natural resource managers. These concerns have prompted research on the natural variability of stream temperatures across landscapes.

The State of Washington has a set of nomographs inserted into its Forest Practices rules and regulations. These nomographs, which are based on previous studies, provide guidance regarding the amount of canopy closure (shade) that needs to be left to meet state water quality standards as a function of site elevation. The original nomographs developed for eastern Washington were based on a small dataset. Since they were published, stream temperature data has become available at many more locations in eastern Washington, providing an opportunity to update the existing nomographs.

This study uses the available existing data to address the effect that canopy closure and other physical parameters have on stream temperature in eastern Washington. Developed relationships between those variables can be used to guide the amount of canopy closure required to meet temperature targets.

Existing stream temperature data, as well as measures of canopy closure, elevation, basin size, distance from divide, precipitation, air temperature, stream width and depth, stream flow, channel aspect, and other characteristics for the sample sites were assembled into a database to support the analysis. Data from 305 sample sites was assembled.

The existing nomographs were tested for accuracy against the larger database. Regression analyses were then used to develop new relationships. The database was stratified using three different classification systems: region, ecoregion, and lithology. There were four regions, size ecoregions, and seven lithology groups in the database. Regression analyses were conducted for all the strata and for eastern Washington as a whole. Two sets of regressions were developed. The first set related canopy closure and elevation to stream temperature. This set of analyses produced equations that were similar to the existing nomographs. The second set of analyses evaluated all available independent variables with potential to affect stream temperature.

The analysis found that the existing nomograph underestimated the amount of canopy closure required to meet the 16^{o}C and 18^{o}C temperature targets 10.5 and 9.2 percent of the time, respectively. Both the 16^{o}C and 18^{o}C nomographs overestimate the amount of shade needed more often than they underestimate shade. Graphically, there is little difference between the existing nomographs and the regression lines incorporating elevation and canopy closure as the independent variables developed by this project. The equations developed in this study that incorporate other variables could not be compared with the existing nomographs.

Problems with data gaps were encountered with all of the individual strata evaluated. Some had extremely small sample sizes. The data in others was concentrated within one set of physical conditions (such as very small basin or high canopy closure). Hence, the dataset was not well represented across the range of conditions in the strata. The small sample sizes hindered the power of the analyses and, in some cases, were too small to support the development of regression equations. The poor representation of conditions across the strata biased the results.

The primary variables found to be significant in the regression analyses were canopy closure, elevation, average annual precipitation, distance from divide, and basin size. Collectively, these variables are indirect measurements of direct solar radiation, ambient air temperature (convective exchange between the water and the air), and water volume. Evaporation rates, longwave radiation, conductive heat transfer between the stream and the streambed, and groundwater inputs are not accounted for in the equations. Variations in these factors contribute to the unexplained variance in the data. The imprecision of the indirect measures of ambient air temperature and water volume also contribute to the unexplained variability.

Canopy closure was identified as a significant variable in nine of the potential 17 strata and for the entire eastern Washington database. Where canopy closure was not included in the regression equations, the objective of providing guidance regarding the amount of canopy closure that needed to be retained could not be met.

Stratification on region, ecoregion, or lithology is not recommended for the purposes of developing new nomographs. The exception to this recommendation is the Blue Mountain area of Washington. The analyses suggest that the amount of canopy required to meet specified temperature standards may be higher in that area. However, the data used to develop the Blue Mountain regressions came primarily from the Oregon Blue Mountains. Additional data is needed to confidently assess the applicability of Oregon data to the Washington Blue Mountains.

The regression equation using the entire database and including only canopy closure and elevations as independent variables was most similar to the existing nomographs. That equation was highly significant and explained 37 percent of the variance in the dataset. A five variable model was developed for the dataset when all possible independent variables were evaluated. The variables in this model were canopy closure, elevation, latitude, average annual precipitation, and distance from divide. This model was highly significant and explained 54 percent of the variability in the model. Recognizing that a 5-parameter model would be difficult to implement, the two parameters that contributed the least to the 5-parameter model were removed to create a 3-parameter model. The remaining variables were elevation, canopy closure, and distance from divide. This last model was also highly significant and explained 51 percent of the variability in the data. Hence, the three-parameter model is substantially more accurate than the model that incorporated only canopy closure and elevation. However, there was little difference in the relative accuracy of the 3-parameter and 5-parameter models.

The selection of the preferred models for implementation (if any) is a policy decision and is beyond the scope of this assessment.
Information that would have the greatest effect on improvement of the analyses conducted in this study includes 1) site specific information on stream size and 2) average July air temperature data with sufficient coverage to support the development of more precise air temperature maps. With the exception of the situation in the Blue Mountains mentioned above, this data would likely have a greater effect on improving the models than collecting data to fill gaps in the strata.

This page is intentionally blank.

1.0 INTRODUCTION

In the Pacific Northwest (PNW), stream temperature is an important determinant of water quality and thus the health of aquatic ecosystems. Consequently, both the natural variations in stream temperature and the human activities that change the thermal conditions of surface waters are significant concerns for natural resource managers. These concerns have prompted research on the natural variability of stream temperatures across landscapes (e.g. Sullivan et al. 1990; Lewis et al. 2000), the influence of land use activities on the thermal properties of streams (Brown and Krygier 1970; Beschta et al.1987; Adams and Sullivan 1990; Sullivan et al. 1990; Lewis et al. 2000), and the effect of temperature on native salmonids (NMFS 1998; Berman 1998; Sullivan et al. 2000).

This study uses existing data to address the effect that canopy closure and other physical parameters have on stream temperature and develops relationships between those variables that can be used to guide the amount of canopy closure required to meet temperature targets in eastern Washington. The development of an updated nomograph (or multiple nomographs should stratification of the data be found to be beneficial) using a larger dataset can help to ensure that the proper amount of canopy closure will be maintained to protect fish and water quality. The accuracy of nomograph(s) developed in this study is variable and was affected in many instances by sample size and/or the range of conditions represented by the dataset.

1.1 PURPOSE AND OBJECTIVES

The purpose of this project was to create an improved temperature nomograph for eastern Washington. The nomograph is based upon modeling of stream temperature as a function of a number of variables. Separate nomographs were developed for each region of eastern Washington, for each of the Bailey's ecoregions (Bailey 1995), and for areas of different lithology.

The specific objectives of the study included the following:

1. Compile all available existing temperature data and physical data associated with temperature dataset for forested lands in eastern Washington. Stratify the data by region, ecoregion, and lithology.

2. Evaluate and report upon the accuracy of the existing temperature prediction nomographs compared to the available temperature data. Evaluate the accuracy of the newly developed nomographs for all of eastern Washington and for the stratified subsets and compare accuracy with existing nomographs.

3. Evaluate additional variables besides elevation and canopy cover that can improve predictive accuracy of temperature prediction models for eastern Washington.

4. Make recommendations for improving the accuracy of the nomographs used in the State of Washington's Forest Practices rules.

1.2 BACKGROUND

In 1992, the Washington State Forest Practices Board (FPB) adopted new forest practices regulations in an effort to bring forest practices rules into compliance with State Water Quality Standards. Included in these regulations were temperature prediction nomographs that were designed to assist landowners, other interested parties, and federal, state, and tribal agencies to determine proper canopy closure levels that needed to be maintained along streams to meet water quality standards. Further monitoring and data analysis in eastern Washington found problems with the initial statewide approach. In 1993, different temperature prediction nomographs were developed for eastern Washington conditions (Figures 1 and 2). The temperature prediction nomographs that were developed in the 1990s were specific to the State Water Quality Standards of 16.0 ^{o}C for Class AA (extraordinary water quality) streams and 18.0 ^{o}C for Class A (excellent water quality) streams.

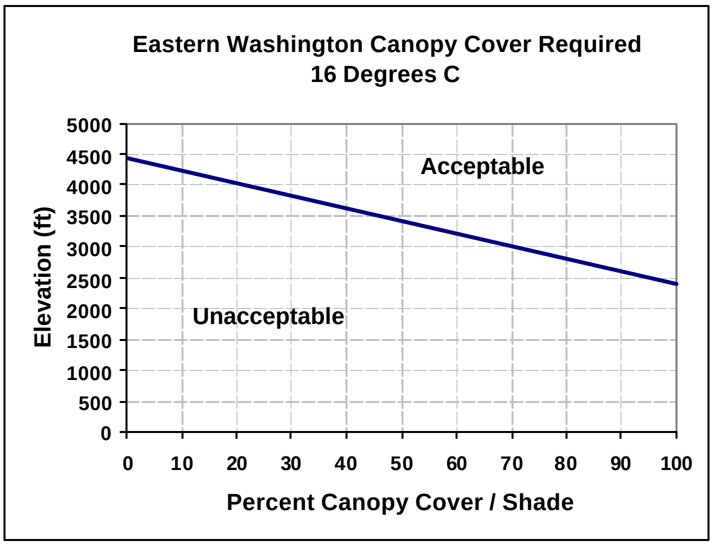


Figure 1. Existing nomograph defining the amount of canopy cover required to be left at various elevations to meet a stream temperature criterion of 16 degrees C.

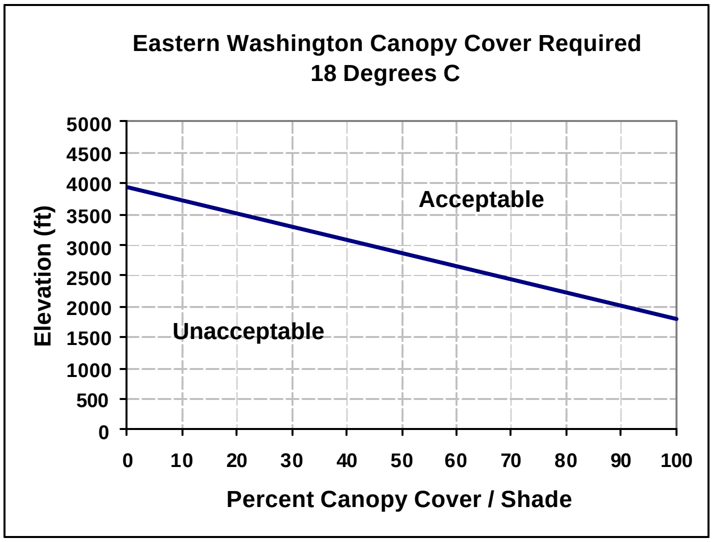


Figure 2. Existing nomograph defining the amount of canopy cover required to be left at various elevations to meet a stream temperature criterion of 18 degrees C.

Effective on July 1, 2001, Washington State's Forest Practices Board adopted a new set of permanent rules. The new rules included requirements for adaptive management aimed at developing implementation tools, evaluating the effectiveness of the new rules, conducting studies to provide better understanding regarding the relationship between forest management activities and fish habitat needs, and monitoring of trends. The Cooperative Monitoring, Evaluation, and Review Committee (CMER) was given the charge to advance the science needed to support the adaptive management process. In recognition of the fact that the original eastern Washington nomographs were developed using relatively small datasets, CMER has identified the need to evaluate and improve the accuracy of the eastern Washington temperature prediction nomographs as a priority project.

Water quality standards have been set for all surface waters in the State of Washington (WAC 173-201A). These standards define the criteria that are used to determine whether the water body is meeting acceptable conditions. There are two sets of water quality standards that are currently of interest. The water quality standards that are currently applicable within the State of Washington are those standards that were adopted in 1997. This set of standards has been under review and revision. As of June 30, 2003, a new set of water quality standards was legally adopted by the state and forwarded to EPA for approval. This new set of water quality standards applies for all actions within the state. The adopted changes cannot be used for federal Clean Water Act actions until the EPA has approved them. Hence, Ecology has indicated "the 1997 standards and criteria should be used as a basis for decision-making until approval is received" (www.ecy.wa.gov/programs/wq/swqs/rev_rule.html). Further information regarding the rules, the adoption process, the applicable standards and other related information can be found at: www.ecy.wa.gov/programs/wq/swqs.

The new water quality standards are "use-based" standards. The applicable criteria for various water segments are defined by the use of that water. Each stream segment is given a "designated use" based upon the use of the stream by various life stages of key fish species, as well as use for water withdrawals and primary contact activities. Standards for water temperature are specified for each designated use. Table 1 summarizes the 2003 temperature standards. Under these new standards, any of seven temperature criteria may apply to a stream segment.

Table 1. 2003 Washington State Water Quality Criteria

Designated Use	Temperature [1]
Char	12°C (53.6°F); 9°C (48.2°F) during spawning and incubation
Salmon and trout spawning, **core** rearing, and migration	16°C (60.8°F); 13°C (55.4°F) during spawning and incubation
Salmon and trout spawning, **noncore** rearing, and migration	17.5°C (63.5°F)
Salmon and trout rearing and migration **only**	17.5°C (63.5°F)
Non-anadromous interior redband trout	18°C (64.4°F)
Indigenous warm water species	20°C (68°F)

(1) Temperature is measured by the 7-day average of the daily maximum temperatures
(2) Lowest one-day minimum

The standards recognize that some locations will naturally exceed the specified criterion. The full text of the rules addresses this issue and limit increases in temperature associated with land use to 0.3 degrees C over natural background. The standards presented in Table 1 provide only the state criteria and do not address natural exceedance of those criteria.

1.3 POTENTIAL INFLUENCE OF TIMBER MANAGEMENT ON STREAM TEMPERATURE

Logging activities such as tree removal, road building, instream modifications, and other operations have been shown to alter the thermal regime of adjacent streams (Brown and Krygier 1970; Feller 1981; Beschta and Taylor 1988; Hatten and Conrad 1995). Tree thinning and clear-cutting within the riparian corridor reduces both stream canopy closure and thermal insulation. This exposes the stream to increased thermal loading, particularly during the summer months when solar radiation is highest. Vegetation thinning and removal generally lead to increased temperature maxima (Beschta and Taylor 1988). Conversely, riparian canopy removal can increase heat loss during the winter and subsequently decrease winter water temperature minima (Beschta et al. 1987).

Sullivan et al (2000), Brown (1983), Sullivan and Adams (1990), Brown and Krygier (1970), Burns (1972), Holtby and Newcombe (1982), Levno and Rothacher (1967), Sullivan et al (1990), and others have presented evidence of the effect of land use practices on thermal conditions in the Pacific Northwest. Studies conducted in

Washington indicate that ambient temperatures have been increased by canopy removal in many areas (Sullivan et al 1990; Caldwell et al 1991; Levno and Rothacher 1969, Levno and Rothacher 1967, Holtby and Newcombe 1982). Reported increases in temperatures have ranged from 0 to 15.8 ^{o}C (Meehan 1970, Holtby and Newcombe 1982, Levno and Tothacher 1967, Levno and Rothacher 1969, Brown and Krygier 1970, Brown and Krygier 1971, Brown 1971). The majority of these studies were conducted to evaluate temperature increases associated with clearcuts of riparian areas. No systematic evaluation of the effect that occurs under the current or previous forest practices and/or the areal extent of these impacts across the state has been completed for forested lands in the State of Washington. A few studies addressing the effects under existing rules are currently underway.

Increases in stream temperature can have consequences for the growth, survival, and migratory behavior of anadromous salmonids. All salmonid species exhibit the greatest growth and survival within a range of temperatures typically referred to as the "preferred" temperature range. These temperature preferences have been the target of many studies (e.g. Grett 1971, Brett 1995, Iverson 1972, Brungs and Jones 1977, Wurtsbaugh 1973), which have been summarized by Coutant (1977), Jobling (1981), and Sullivan et al (2000). Growth rates increase up to a "preferred" maximum temperature and then plummet at temperatures in excess of this preferred maximum. The preferred maximum increases to a limit with increasing prey availability (Everson 1973, Adams and Breck 1990, Brett 1995). Lethal water temperatures are reported in the range from 21.5 ^{o}C to 25 ^{o}C and vary significantly with acclimation temperature (Beschta et al 1987, Sullivan et al 2000). At lower temperatures, water temperatures in excess of the preferred temperature range can influence behavior, triggering concentration of fish in cooler waters (Brett 1952, Mantelman 1960). Upstream movement of adult salmonids can be curtailed by warm temperature (Lantz 1971) and the rate of downstream migration of smolts can be increased at warmer temperatures (Quinn and Adams 1996, Keenleyside and Hoar 1954). Hence, warming of waters to temperatures at or below the preferred temperatures for salmonids will tend to increase growth and survival and warming of waters to temperatures in excess of the preferred temperature range can have significant effects on fish survival, behavior, and migration.

2.0 METHODS

The Nomograph Project was conducted between July 2002 and December 2004. The project consisted of three primary tasks including development of the temperature database, GIS mapping, and statistical analyses.

2.1 PROJECT SETTING

The project study area includes all forested lands in the State of Washington east of the Cascade Mountain divide. Within this area, there are four sets of mountainous ranges including the eastern flanks of the Cascade Mountains, the Okanogan highlands and the

Selkirk Mountains in Northeast Washington, and the Blue Mountains in Southeast Washington.

The east slope of the Cascade Mountains decreases in elevation from the summit of the range to approximately 1,500 feet above sea level. Precipitation tends to decrease as elevation decreases (Western Regional Climate Center, www.wrcc.dri.edu). Annual precipitation in the higher elevations often exceeds 90 inches per year, while precipitation at lower elevations may be less than 20 inches per year. Maximum summer temperatures at lower elevations often reach 100°F or higher and 80°F is usually recorded in the higher elevations.

Elevations in northeastern Washington range from roughly 2000 feet to 6000 feet along the higher ridges. Average annual precipitation increases in a northeasterly direction and ranges from 17 inches in the Spokane area to greater than 60 inches in the northeastern corner of the region. Maximum summer air temperatures reach 100°C on at least a few days each summer. Air temperatures in the mountains decrease three to five degrees F with each 1000 feet increase in elevation (Western Regional Climate Center, www.wrcc.dri.edu).

The Blue Mountain area of the state ranges in elevation from 1000 feet near Walla Walla to 6000 feet in the higher mountains. Annual precipitation is between 10 to 20 inches per year at lower elevations, increasing to over 60 inches at higher elevations. Average maximum July air temperature is in the upper 80s and exceeds 100°F on a few days of each year (Western Regional Climate Center, www.wrcc.dri.edu).

2.2 STUDY VARIABLES

The temperature of a stream reflects the concentration of heat energy in the water. This is determined by the initial temperature at the headwaters and the amount of heat energy either lost or gained along the length of the stream. The water temperature recorded at any location along a stream thus reflects a balance between heat input and heat loss from the system. The major source of heat input, solar radiation, is inherently dynamic. It changes over daily and seasonal time scales, and it can vary spatially with biological and physical shading (Brown 1969). Rates of heat input and heat loss also depend on local factors, both within the stream and along the stream corridor, including microclimate, canopy cover, stream size, and groundwater inflow.

To explain spatial variability in stream temperature, several factors that affect heat inputs and heat losses to the natural stream system were considered within this study. A total of thirteen parameters were examined including:

- Site elevation,
- Stream flow,
- Bankfull width and depth,
- Wetted width and depth,
- Canopy closure,
- Drainage basin area,

- Annual precipitation,
- Hill slope gradient,
- Distance to watershed divide, and
- Channel aspect.

These parameters were selected because they have specific relevance to the project and have been shown to influence stream temperature conditions (Adams and Sullivan 1989; National Academy of Sciences/National Academy of Engineering, 1972; Beschta et al. 1987, Brown 1983, Brown 1969, Edinger et al.1968, Hatten and Conrad 1995). Some of these variables can only be obtained through data collection at site-specific locations, while other variables can be obtained using data available on maps and GIS layers at a given scale. Descriptions of these parameters and information on their relevance in stream systems are provided below.

Site elevation can have an influence on stream temperature in that air temperature has been shown to decrease with increasing elevation (Sullivan et al. 1990). Air temperature, in turn, influences water temperature as is discussed in detail below.

Air temperature refers to the local atmospheric temperature at the location being monitored. Air temperature, like solar radiation, is dynamic and changes with meteorological shifts on both a daily and seasonal basis. Air temperature also tends to decrease with elevation. Many studies have examined the relationship between air temperature and water temperature (Smith and Lavis 1975, Webb and Walling 1992, Mohseni and Stefan 1999). These studies have shown that, as expected, water temperature correlates closely with air temperature. However, Crittenden (1978) found that for streams subjected to intense solar radiation, the thermal properties of the streambed, the time of year, the depth of the stream, the amount of vegetative shading, and wind speed were more important contributors to stream temperature than air temperature.

Stream flow can have an effect on stream temperature in terms of volume. A larger mass of water will take longer to heat and cool through the water surface to air interchange. The larger mass of water can also assimilate more heat load for the same rise in temperature as a smaller mass, thus modifying the heat effect. Higher flow velocities decreases the time and proportion water segments are exposed to solar radiation (Ecology 2000).

Bank full width and depth are measurements that describe the relative size of a stream within a given basin. The depth and width of a stream are the most important size-related factors that influence stream temperature. Small streams have less capacity for heat storage than larger streams (Brown 1969), and consequently they show greater fluctuations in daily temperature. For example, Brown (1969) showed that diurnal temperature shifts in small streams (< 1 cms) could be as much as $20^{o}C$ while large streams (> 142 cms) fluctuated only $2^{o}C$ (Brown 1969). The width-to-depth relationship of a stream also influences thermal conditions. Shallow streams heat quickly in response to solar radiation, and, as such, display greater diurnal fluctuations. Conversely, deeper streams respond more slowly to increases in solar radiation. Accordingly, streams with different depth profiles will have different temperature regimes (Sullivan et al. 1990).

Stream shading or canopy closure refers to the amount of riparian vegetation shielding the stream from direct and indirect solar radiation. Removal of the canopy over a stream generally results in higher stream temperatures and greater diurnal fluctuations (Brown and Krygier 1970, Holtby and Newcombe 1982, Hatten and Conrad 1995). Of the various aspects of the thermal regime, daily temperature maxima of small streams increase the most in response to vegetation removal (Sullivan et al. 1990).

Drainage basin area relates to the watershed position of a given monitoring site. The distance between the site and the watershed divide is a measure that refers to the location of a given point in the context of the larger drainage area. Hynes (1970) summarized findings from European streams in which summer stream temperatures increased proportionally with the logarithmic distance of the site position from the watershed divide. Lewis et al. (2000) also found that streams in northern California become warmer as they extend away from the watershed divide with the exception of streams that are cooled by fog and marine air along the coast of the Pacific Ocean. Conversely, research by McIntosh et al. (1996) examined longitudinal temperature patterns in streams in the Pacific Northwest and found significant variability in stream temperatures within a watershed. Using Forward-Looking Infrared (FLIR), they found that stream temperatures can be highly variable at the watershed scale, alternating between cool and warm reaches longitudinally. In some cases, this variability was attributed to point sources such as dams or tributary streams; however, this did not explain the variability in all cases. In general, streams tend to decrease in elevation and increase in size as they move away from the watershed divide. Therefore, correlations between drainage basin area and stream temperature at least partially reflect the differences in air temperature and stream size, which were discussed previously.

Hill slope gradient on either side of the channel has the potential to affect the amount of topographic shading on a stream independent of canopy closure. Topographic shading can affect the amount of direct solar radiation reaching a stream.

Precipitation shifts dramatically from the crest of the Cascade Mountains eastward to the floodplain of the Columbia River (Western Regional Climate Center, www.wrcc.dri.edu). High precipitation correlates with the generally cool temperatures of high-elevation streams. Conversely, low precipitation correlates with warmer stream temperatures at lower elevations. Precipitation serves as the source of water for surface flow and groundwater recharge, both of which affect stream flow and, hence, the volume of water in a stream to be heated and cooled.

Channel aspect refers to the orientation of the channel with respect to the position of the sun. Generally, streams flowing in a north-to-south direction or a south-to-north direction have a relatively shorter period of exposure to direct overhead solar radiation than do east-to-west or west-to-east flowing streams (Sullivan et al. 1990). However, there is some debate over the impact of riparian canopy closure in relation to channel orientation. The path of the sun from east to west as well as the peak angle of the sun in the sky during the summer month suggests that both riparian vegetation and topography may provide more canopy closure to north-south oriented streams than to east-west flowing streams (Sullivan et al. 1990). However, Lewis et al. (2000) found no significant relationship between channel orientation and stream temperature and suggest that it plays a relatively minor role.

2.3 TEMPERATURE DATA COLLECTION

2.3.1 Minimum Acceptable Standards for Data

To reduce error and improve consistency between data sets, all data used for the project had to meet several requirements. These requirements extended to both the specific method of collection of stream and air temperature data as well as the specific site location where those data were collected.

Data must have been collected at sites defined as thermal reaches (Sullivan et al. 1990). A thermal reach has relatively homogenous riparian and channel characteristics for a distance upstream from the monitoring station to allow the stream temperature to reach equilibrium with those conditions. The distance upstream will vary with stream width, depth, and flow (as well as other parameters). As a general rule of thumb, roughly 2000 feet of relatively uniform conditions (this can be uniformly low canopy closure, uniformly high canopy closure, etc.) are needed upstream of a monitoring site on smaller streams.

The temperature sampling sites were preferred that met the following target criteria (data were accepted for inclusion in this study's database if items 2, 5, or 6 had not been checked when the data were collected):

1. Relatively uniform riparian conditions upstream of the monitoring site. Consistent riparian conditions for at least 1000 feet upstream of the site were preferred.
2. Less than 20% of the total flow at the site is input by tributaries within 2000 feet upstream of the monitoring site.
3. The site is not located at the confluence of two streams.
4. There are no large wetlands, lakes, or dams within 2000 feet upstream of the monitoring site. If dams or lakes are present and the stream is large, the distance from the dam or lake should be increased to 3000 feet.
5. There are no obvious large volumes of groundwater inputs (springs) upstream of the site that may be inputting more than 20% of the total flow at the monitoring site.
6. The channel geomorphology is relatively uniform for 2000 feet upstream of the monitoring site.

The temperature data must meet the following criteria:

1. Data must be collected using continuous monitoring devices, with temperature recorded at least hourly.
2. Recording devices must have been in place in the stream from at least July 7 through September 1 of any given year.

In addition to stream temperature, canopy closure measurements were required. Preferably, canopy closure measurements would have been taken at several points upstream of the monitoring site, but one measurement at the site was acceptable if the riparian conditions upstream of the site were relatively consistent upstream. Consistency upstream of the site was confirmed with each of the parties providing data. In most cases, parties collecting data had collected reach averaged canopy closure measurements. One of the parties contributing data confirmed consistency in shade levels through use of aerial photographs rather than verifying consistency in the field. Twenty-seven (27) sites in Northeast Washington and five (5) sites in the northern Cascades region had only point measurements of canopy closure. These data were included in the database because sample sizes in both areas were small and the parties contributing data confirmed that consistency of canopy closure had been checked visually.

Additional required information for each site:

1. Year data were collected.
2. Persons or entity collecting the data (with contact information).
3. Site of data collection (site legal description, longitude and latitude, or GIS coordinates). Several data points were provided by one data contributor that preferred to treat the sample location as proprietary data. This contributor provided the GIS information we needed for each location; hence, the need to know the specific location was significantly diminished and the data was accepted for inclusion in the database.
4. Type of data collection instrument used and accuracy of that instrument (if known).
5. If data has been published, citation of publication.

The above requirements and specifications for site selection and data collection are consistent with the stream temperature survey methods described by Schuett-Hames, et al. (1999).

2.3.2 Assistance with Additional Data Collection by Contributing Parties

A lengthy list of potential collaborators was considered for this project. Following extensive contacts and discussion, the following organizations have contributed stream temperature data, and, where available, associated channel characteristic information:

- Wenatchee/ Okanogan National Forest
- Okanogan Conservation District
- Yakama Indian Nation
- Boise Cascade Company

- Plum Creek Timber Company
- Kalispel Tribe
- Upper Columbia United Tribes
- Colville Confederated Tribes
- Entiat Ranger District
- Washington State Department of Ecology
- Ferry County Conservation District
- Foster County Conservation District
- Washington Department of Fish and Wildlife
- Campbell Group
- Cooperative Monitoring, Evaluation and Research (CMER) original data used to develop nomographs
- U.S. Fish and Wildlife Service
- Washington Department of Natural Resources

An agreement was been made with some of the supporting and collaborating entities to maintain confidentiality data.

2.4 CONTACT OF PARTIES

In August 2002, a memo soliciting data were sent to entities in eastern Washington who may have data relevant to this project. Data requested included at a minimum, stream temperature, site elevation and canopy closure. The specific quality assurance/quality control of the data was also included in this letter in an attempt to maintain a high level of accuracy and consistency between the datasets received.

2.5 DATA COLLECTION QA/QC

Once received, data underwent quality assurance/quality control (QA/QC) screening to ensure data accuracy and consistency. In addition, consistency among stream name, monitoring site number, and site description was verified. Finally, a QA/QC assessment sheet was filled out for each entity.

2.6 DATA PROCESSING

2.6.1 Raw Data

A few datasets were received as raw temperature data. These data required both QA/QC screening and the development of summary information before they could be incorporated into the database. Summaries included the reduction of hourly stream temperature data to daily values. Reduction was conducted using a macro developed using MS-DOS. Temperature summaries included maximum, mean, and minimum values, as well as a 7-day rolling average of the maximum temperature.

2.6.2 Pre-Processed Data

Most datasets received contained only a summary of the data collected at each monitoring site. These data included only the maximum summer stream temperature and, in some cases, the 7-day average maximum temperature (calculated as the rolling sum of 7 days maximum temperature divided by 7). These datasets required basic screening and no additional calculations were needed before they were included in the database.

2.7 DATABASE DEVELOPMENT

The database was compiled using Microsoft Excel.

2.7.1 Field Collected Data

All field collected data were entered into the database as received unless conversions from standard to metric measures had to be made or other conversions were necessary to develop a database with consistent units of measure.

2.7.2 GIS Data

Temperature site location data were received in a variety of formats including ArcInfo coverages, shapefiles, text files, map images, and Excel spreadsheets. Each format and dataset required a different process to bring it into the GIS. Coverages and shapefiles were imported directly and projected to the project coordinate system. When latitude and longitude coordinates were provided in a tabular format, the data were formatted, then imported using ArcInfo Generate processes. Points provided graphically on maps were calibrated to GIS base layers and digitized manually. Points received without spatial coordinates or graphic map location information were used if they could be correlated from a combination of legal description, stream name, and point elevation. Point data were linked to attribute data tables via GISID.

Base GIS data including hydrography, transportation, Public Land Survey, 1:100,000 Scale Digital Geology, and 10 meter digital elevation models (DEM), which were provided by Washington Department of Natural Resources (WDNR). The WDNR data were received in the Washington State Plane, South Zone, NAD27 coordinate system. This data and coordinate system was used as a basis for all GIS processing. Additional background spatial data were collected and imported into the GIS. Ecoregion data (Bailey 1995) was acquired from the USDA Forest Service (http://www.fs.fed.us/institute/ecoregions/ecoreg1_home.html) (Figure 3). Regions were defined using the maps presented in the *Geology of Washington* located at http://www.dnr.wa.gov/geology/geolofwa.htm. Precipitation data were acquired from the Water and Climate Center of the Natural Resources Conservation Service (www.ocs.orst.edu/prism) (Figure 4). Air temperature data were acquired from the National Climate Data Center (www.nndc.noaa.gov/cgi-bin/climaps/climaps.pl) (Figure 5).

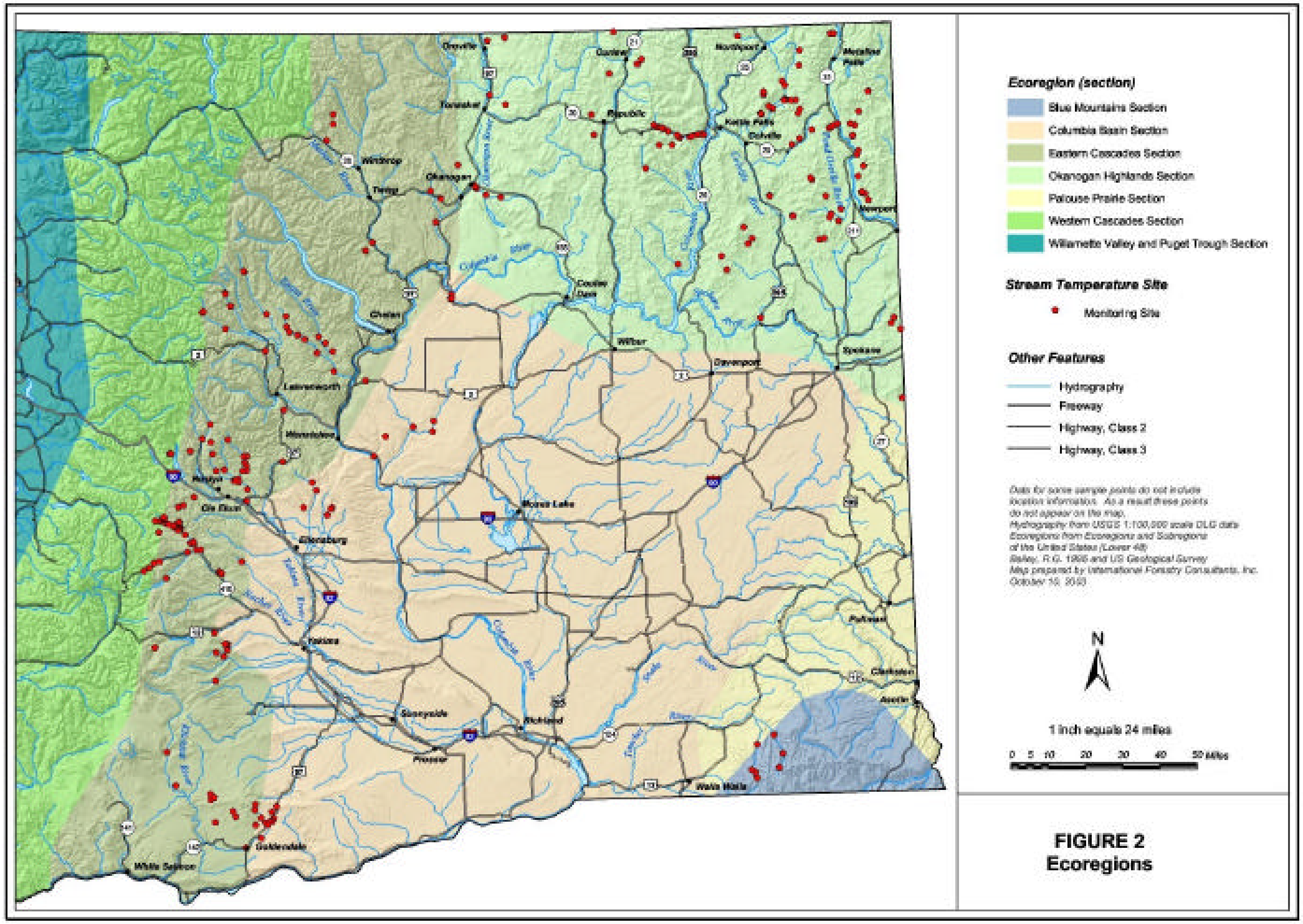


Figure 3. Bailey's ecoregions in eastern Washington.

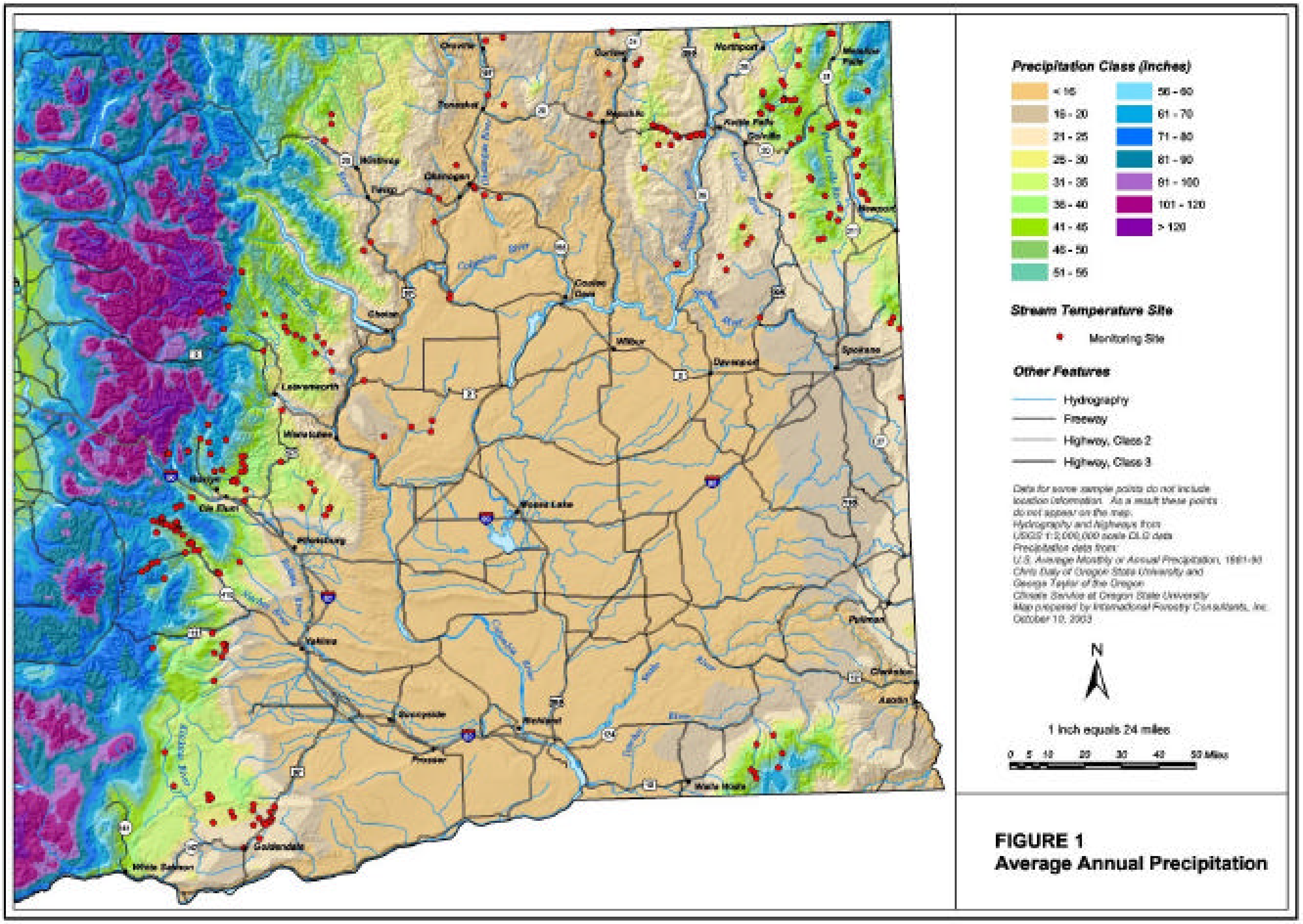


Figure 4. Average annual precipitation in eastern Washington.

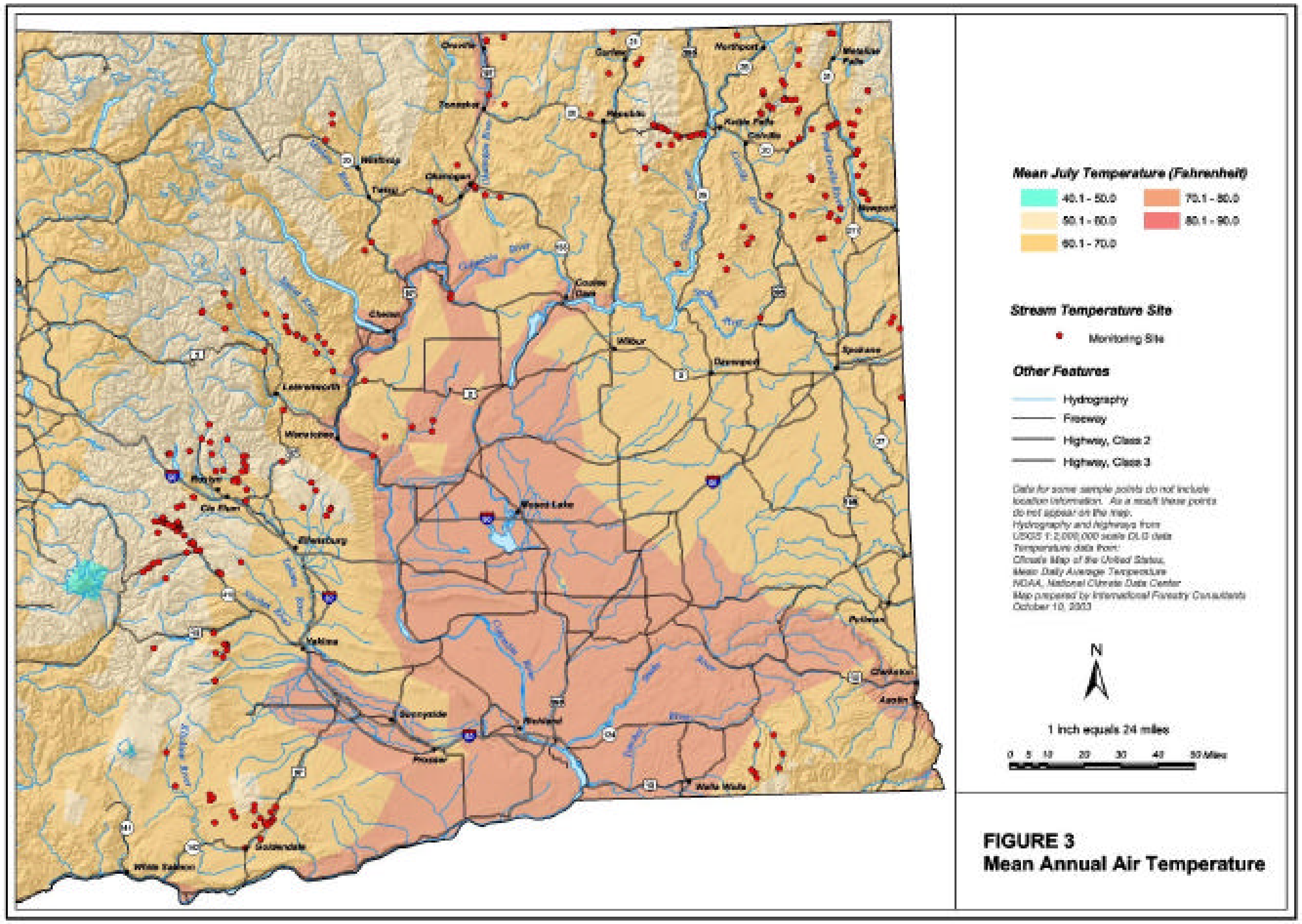


Figure 5. Mean July air temperature in eastern Washington.

Drainage basins above each sample point were delineated using a combination of automated ArcInfo Grid watershed processing and manual digitizing. Manual digitizing was performed on-screen, based on 60 foot contours, DNR hydro and Watershed Administrative Unit (WAU) coverages, and a shaded relief image. The contours and shaded relief were generated from the DEMs. Automated watershed processing was thoroughly edited using manual methods.

GIS analysis was performed using ArcInfo software. This included spatial feature overlays, proximity analyses, and measurements. For a brief description for each of these processes, please refer to GIS Metadata (Appendix E). All analysis results were copied to, and are stored in the point attribute table in the master temperature point GIS coverage.

2.7.3 Concatenating Variables Where Needed

In some cases, different measurements had to be combined to create one variable. This occurred with canopy closure. In some cases, contributors provided only canopy closure measured at the site or canopy closure averaged over the 1000 feet upstream of the site. In other cases, both measures were provided. A new variable was created (condensed shade) that used the average canopy closure measured over the 1000 feet upstream if available and used the site measurement if not. This choice of variables ensured that the most accurate of the provided variables was always used.

Elevation was also concatenated in some cases. In most cases, elevation was not provided and the GIS database was used to determine elevation. In some cases, no specific location information was provided, so the elevation measure reported by the contributor of the information was used.

2.8 DATA ANALYSIS

2.8.1 Descriptive Statistics

Descriptive statistics were used to describe the basic features of the data in the study. Summary statistics were calculated for all the study variables including the number of cases, and minimum, maximum, and summary values. Univariate analyses were also performed including the distribution, the central tendency (mean), and the dispersion (range and standard deviation). All descriptive statistics were developed using SYSTAT 10.2 for Windows.

2.8.2 Transformations

The only variable that was transformed was channel aspect. Channel aspect is a variable that is measured in degrees. It was transformed into a linear variable reflecting its tendency to run north-south to east-west. The transformation used was the absolute value of the sine of the channel aspect. This resulted in a variable that approached zero for streams that trend towards a north to south orientation and approached one (1) for streams that trend towards an east to west orientation. Throughout this document, this transformed variable is referred to as Abs(sine aspect).

2.8.3 Correlations

Correlation matrices were developed using SYSTAT 10.2 for Windows to evaluate the expected degree of correlation between dependent and independent variables and to investigate potential interactions between variables. Eight variables were examined including:

- Site elevation
- Canopy closure
- Drainage basin area
- Annual precipitation
- Distance to watershed divide
- Average July air temperature
- Average August air temperature
- Average annual air temperature

Other variables in the dataset were not included due to small sample sizes.

2.8.4 Tests of the Goodness of Fit of the Existing Nomograph Rules

There are several difficulties with statistically testing the goodness of fit of the existing nomographs to the dataset that was assembled for this project. The nomographs are set up such that if certain minimum conditions (canopy closure and elevation) are met or exceeded, the resulting stream temperature will be equal to or lower than 16.0ºC for Class AA streams and 18.0ºC for Class A streams. Conversely, it is assumed that if those conditions are not met, the target stream temperatures will also not be met.

There are several considerations regarding the existing nomograph that must be taken into consideration as the appropriate statistical tests are identified. These include:

- The nomographs were apparently based on predictive equations; however, the information needed to statistically compare those regression lines with those developed in this study is not available. An attempt was made to recreate the original regression equations, but the resulting equations were different from those published in the Forest Practices Manual. This may have occurred if the dataset assumed to represent the original data was incorrect or if assumptions regarding outliers were different. Therefore, the original nomograph regression cannot be compared statistically with those equations developed in this study using standard techniques to test for differences between two regression lines.

- No information on standard deviation or other measures of accuracy is available for the existing graphs. Hence, data means and distributions cannot be tested to determine if they fall within the same population.

- It cannot be assumed that the data in the database generated in Task 1 are representative of a random sample. For instance, WDOE data is likely biased toward sites where violations of water quality standards are suspected. Additionally, some data collection efforts were focused on collecting data at sites that represent the range of canopy closure and elevation conditions rather than a random sample. Therefore, means and standard deviations calculated from these sites are not necessarily representative of the entire population.

Given these problems, most statistical methods that might otherwise be employed to test the fit of the nomograph to the data cannot be used. However, a chi-square approach (Steele and Torrie, 1980) can, and was, used to test the goodness of fit of the nomograph.

2.8.5 Regression Analysis

Regression analyses using maximum annual stream temperature as the dependent variables and the various other parameters previously discussed as the independent variables were completed using SYSTAT 10.2 for Windows. Analyses were conducted for the entire data set and for subsets of the data stratified on region, ecoregion, and lithology. One set of regression analyses were completed using canopy closures and elevation as the only independent variables. This set of analyses provided results that were most comparable to the existing nomographs. A second set of regression completed using all possible input variables. Variables that were found to be significant at the p=0.05 level were retained. Other variables were excluded from each equation. In some cases, equations that included variables that were almost but not quite significant at the p=0.05 level were also reported. This occurred most frequently when the inclusion of the marginal variable resulted in an equation that was consistent with others developed within a strata (region, ecoregion, or lithology) or where the marginal variable was canopy closure.

For the purposes of the regression analyses, a second dataset representing data points collected in the Blue Mountains in Oregon was used. The Washington Blue Mountain dataset was very small (n=7). Inclusion of the Oregon data allowed for the development of a larger database that could be used to explore possible relationships between stream temperature, elevation and canopy closure. The dataset is a compilation of data collected by the US Forest Service and the State of Oregon in the Blue Mountains area. There are potential problems with using an Oregon dataset to evaluate temperature relationships in Washington. These are described in the “Discussion” section of this report.

For each analysis, outliers identified in the SYSTAT 10.2 output were evaluated to determine if they should be included in the dataset. The determinations to include or exclude outliers was based on the relative contribution to the coefficient of determination (r^2), the effect of inclusion of the case had on the overall F statistic for the regression analysis, and the influence the case had on the resulting equations. Specific discussion regarding the determination to include or exclude outliers is provided in detail in the results section of this report.

Equations developed within each strata (region, ecoregion, and lithology) were compared using methods for comparing multiple regression equations described by Zar (1974).

3.0 RESULTS

3.1 SAMPLE SIZE AND DISTRIBUTION

The final database includes 305 cases from the State of Washington with at least elevation, canopy closure, and maximum stream temperature. By region, the sample size varied from 7 (Blue Mountains) to 178 (Southern Cascades) (Table 2). The samples covered a wide range of elevation and canopy closure situations in eastern Washington (Figure 6).

Most of the samples from NE Washington were from areas with higher canopy closures (Figure 6). Lower elevation samples from the North Cascades region tended to have lower canopy closure than those at higher elevation. The samples from the Southern Cascades region were well distributed across the range of elevation and canopy closure. There were only seven samples for the Blue Mountains from the State of Washington. The physical locations of the samples are depicted in Figure 7.

Table 2. Sample sizes for the complete dataset and for each region for the variables included in the database.

Variable	All Data	NE Washington	Southern Cascades	Northern Cascade	Blue Mountains
Max. Water Temperature	305	83	178	37	7
Canopy Closure	305	83	178	37	7
Elevation (some GIS)	305	83	178	37	7
Bank full Width	66	17	30	28	0
Bank full Depth	36	36	7	0	0
Stream Gradient	21	30	0	0	0
Wetted Width	2	2	0	0	0
Wetted Depth	21	21	0	0	0
Left Side Hill Gradient	8	8	0	0	0
Right Side Hill Gradient	8	8	0	0	0
Road Density	1	1	0	0	0
Lat and Lon (some GIS)	267	83	146	37	7
Basin Size (GIS)	267	83	146	37	7
Ecoregion (GIS)	267	83	146	37	7
Lithology (GIS)	267	83	146	37	7
Average Annual Precipitation (GIS)	267	83	146	37	7
Stream Gradient (GIS)	267	83	146	37	7
Hill Slope Gradient	267	83	146	37	7

Variable	All Data	NE Washington	Southern Cascades	Northern Cascade	Blue Mountains
(GIS)					
Distance from Divide (GIS)	267	83	146	37	7
Road Density (GIS)	263	83	142	37	7
Average Annual Air Temperature (GIS)	267	83	146	37	7
Average July Air Temperature (GIS)	267	83	146	37	7
Average August Air Temperature (GIS)	267	83	146	37	7
Aspect (GIS)	267	83	146	37	7

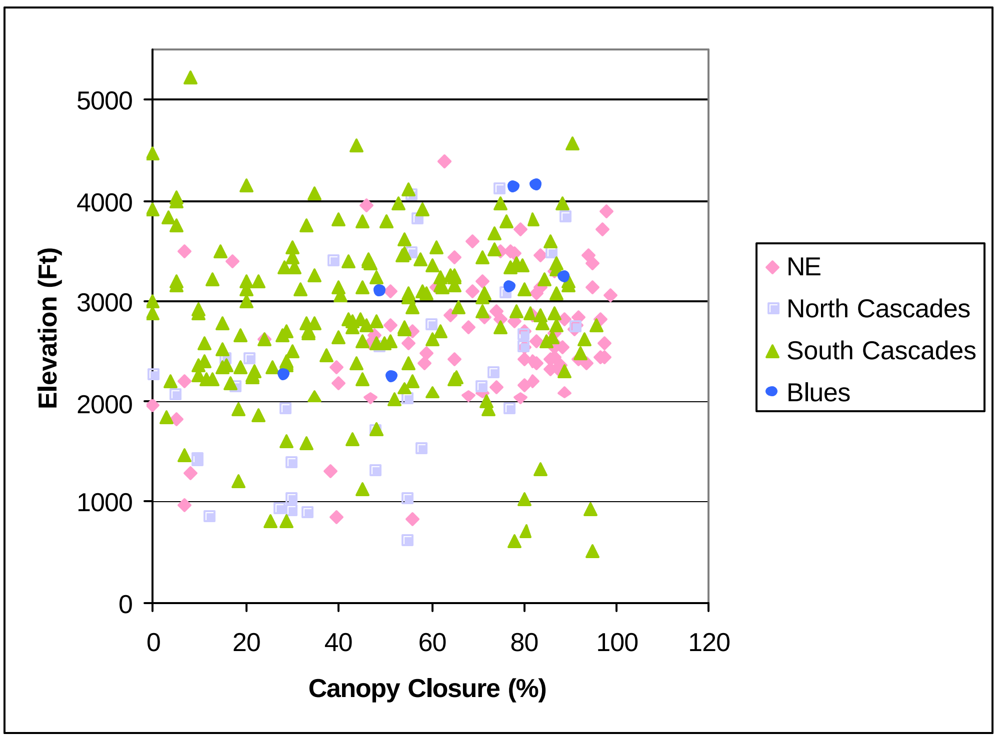


Figure 6. Distribution of samples in the database across elevation and canopy closure by region. Note, the majority of the samples from the Blue Mountains are from Oregon.

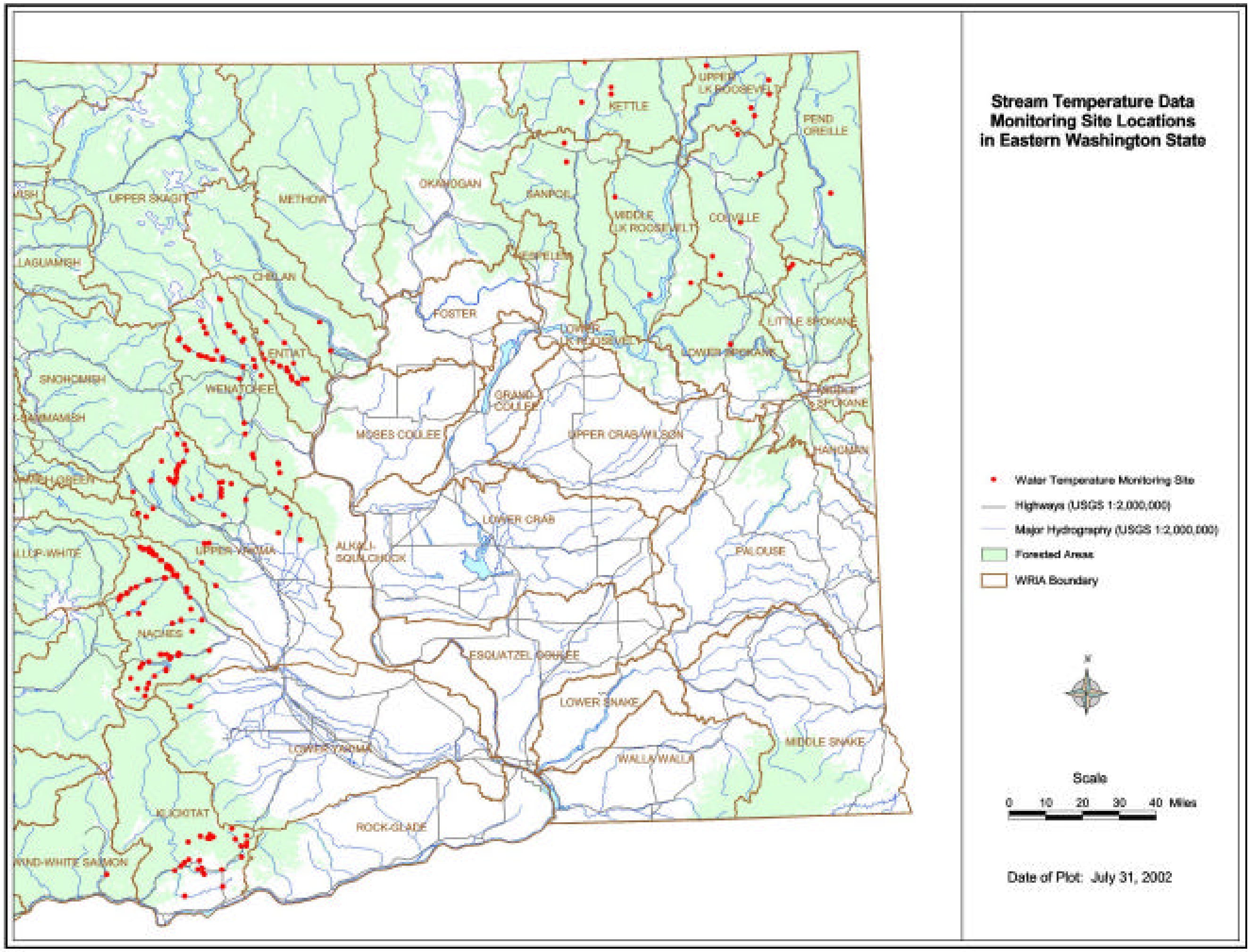


Figure 7. Data collection sites represented in the stream temperature dataset. Not shown on the map are 29 sites in the eastern Cascade Mountain area whose location is proprietary and 19 sites in NE Washington that were received too late to be included in the mapping efforts.

The sample size of some of the variables was significantly lower than that for the primary variables (Table 2). This was particularly true for field measured parameters (other than canopy closure) that were provided by the contributors of the data. These parameters included stream flow, wetted width, wetted depth, bank full width, bank full depth, stream flow, maximum air temperature measured in the field, stream gradient and hill slope gradient measured in the field, channel type, valley width, dominant tree species in the riparian area, and channel gradient measured in the field.

In a smaller number of cases, information was provided to us without specific locations. In these cases, the contributors provided much of the needed information to fill the database; however, some of the information that was generated using GIS could not be generated for these cases. As a result, the sample size for some of the GIS generated variables is smaller than the total number of cases in the database. Greater detail regarding the descriptive statistics (minimum, maximum, range, standard deviation, variance) for each of the variables is provided in Appendix A.

3.2 CORRELATIONS

To determine whether the environmental variables considered in this study are independent, a correlation analysis was run to identify relationships among the variables (Table 3). Eight variables (site elevation, canopy closure, drainage basin area, annual precipitation, distance to watershed divide, July air temperature, August air temperature, and annual air temperature) were considered. A complete table of correlation analysis results including both the correlations reported for the entire study area and those identified for each study region is provided in Appendix B.

There are several highly correlated variables in the database (Table 3). Elevation was highly correlated with basin size, precipitation, distance from divide, and average July, August, and annual air temperatures. These correlations are all as expected. With decreasing elevation, precipitation tends to decrease and air temperature tends to increase. Distance to divide was negatively correlated with annual precipitation and site elevation, and positively correlated with basin size, and mean and maximum air temperatures. Basin size was positively correlated with distance from divide and negatively correlated to both precipitation and site elevation. Average July, August and annual air temperatures were positively correlated.

3.3 GOODNESS OF FIT OF THE EXISTING NOMOGRAPHS

To assist in the evaluation of the relative goodness of fit of the data to the existing nomographs, the data set was divided into 4 possible situations:

1. The canopy closure at the site met or exceeded the canopy closure required by the nomograph and the measured stream temperature equaled or was cooler than the target temperature (the nomograph correctly predicted the stream would be cool).

Table 3. Correlation matrix of the relationships among selected environmental variables showing the pairs that had higher correlation coefficients only.

	Site Elevation	**Basin Size**	**Annual Precip.**	**Dist. to divide**	**July air temp**	**August air temp**	**Annual air temp**	**Canopy closure**
Site Elevation	1.00							
Basin Size	r= -0.48 p <0.01 n = 264	1.00						
Annual Precip.	r= 0.45 p <0.01 n = 264		1.00					
Distance from divide	r= -0.55 p <0.01 n = 264	r= 0.83 p <0.01 n = 267		1.00				
July air temp.	r= -0.54 p <0.01 n = 264		r= -0.59 p <0.01 n = 267		1.00			
August air temp.	r= -0.55 p <0.01 n = 264		r= -0.55 p <0.01 n = 267		r= 0.85 p <0.01 n = 267	1.00		
Annual air temp.	r= -0.66 p <0.01 n = 264				r= 0.60 p <0.01 n = 267	r= 0.64 p <0.01 n = 297	1.00	
Canopy closure		r= -0.42 p <0.01 n = 267		r= -0.53 p <0.01 n = 267				1.00

2. The canopy closure at the site was less than the canopy closure required by the nomograph and the measured stream temperature warmer than the target temperature (the nomograph correctly predicted the stream would be warm).

3. The canopy closure at the site met or exceeded the canopy closure required by the nomograph, however, the measured stream temperature was warmer than the target temperature (the canopy closure required using the existing nomograph was not sufficient to meet the target temperature).

4. The canopy closure at the site was less than the canopy closure required by the nomograph, however the measured stream temperature was equal or cooler than the target temperature (the canopy closure required using the existing nomograph was greater than needed to meet the target temperature).

In the first two situations, the existing nomograph correctly predicted the outcome. In the last two situations, the existing nomograph required a canopy closure that was either insufficient or greater than needed to meet the target temperature.

The number of cases in each situation is summarized in Table 4 and the relative distribution of cases over these outcomes is depicted in Figures 8 and 9. Overall, 63.6 and 61.3 percent of the cases fell into one of the correctly predicted categories for the 16 and 18 degree C nomographs, respectively. The balance fell into one of the categories where canopy closure either was over or underestimated. The existing nomograph underestimated the canopy closure needed to meet the temperature target in 10.5 percent of the cases. Broken down by region, the portion that fell into one of the categories where the nomograph correctly predicted the result ranged from 56.2% (North Cascades region, 16 degrees) to 67.5% (North Cascades, 18 degrees). The portion that fell into the situation where the nomograph underestimated the amount of canopy closure required ranged from 6% (Northeast Washington region, 16 degrees) 28.6% (Blue Mountains region, 16 degrees).

The first null hypothesis tested was that the nomographs correctly predicted the canopy closure needed to meet the target 50 percent of the time. Essentially, this is a test to determine whether the nomographs are more accurate than a random assignment of cases into the categories where the nomograph correctly predicted the result and those where the nomograph was incorrect in that it under or overestimated the amount of canopy closure required. This hypothesis was rejected for the complete dataset for both nomographs (Tables 5 and 6) at the probability (p) level of 0.05. This indicates that, overall, the nomographs have better accuracy than a random assignment of canopy closures. The results broken down by region were variable. The hypothesis was rejected for the South Cascades and Northeast Washington regions in the test against both the 16 and 18-degree rules, but was accepted for both rules as applied in the North Cascades and Blue Mountains regions, implying that the nomographs were no more accurate than a random assignment of cases into the various categories. The Blue Mountain region was affected by a small sample size.

A second approach to testing the fit of the existing nomographs to the data is to test the distribution of the errors around the nomograph lines. If the lines were developed through a regression of data, the errors should be evenly distributed on either side of the line. To test this, the number of cases in the "incorrect" situations was used. If the existing nomographs were a good fit, the number of cases where the needed canopy closure is overestimated would be roughly equivalent to the number of cases where the needed canopy closure is underestimated. Hence, the null hypothesis is that the errors are evenly balanced on each side of the nomograph lines.

A chi-square test was used to test the null hypothesis (Steele and Torrie 1980) using the $p=0.05$ test statistic. The null hypothesis of evenly balanced errors was rejected when using the entire dataset to test both the $16^{o}C$ and $18^{o}C$ nomograph lines (Tables 7 and 8). The null hypothesis was also rejected for the NE Washington and Southern Cascades regions in the test of both the nomograph lines. The null hypothesis was accepted for the Blue Mountain and Northern Cascades regions. On average, the existing nomographs overestimated the canopy closure required to meet the target temperature in all cases except the 16-degree nomograph tested against the data for the Blue Mountains (Figures 10 and 11).

Table 4. Number and percent of cases falling into each possible scenario regarding the goodness of fit of the existing nomographs.

		Existing Nomograph Correctly Predicts Outcome		Existing Nomograph Over or Underestimates Required Canopy Closure		
Target Temperature	**Region**	**Canopy Closure target and stream temperature targets met**	**Neither Canopy Closure target nor stream temperature target met**	**Canopy Closure target met but stream temperature exceeds target**	**Canopy closure target not met but stream temperature less than or equal to target**	**Total**
16	All Data	47 (15.4%)	147 (48.2%)	32 (10.5%)	79 (28.9%)	305
	NE Washington	15 (19.3%)	30 (44.6%)	5 (6.0%)	25 (30.1%)	83
	Northern Cascades	4 (15.7%)	20 (40.5%)	5 (6.8%)	8 (32.4%)	37
	Southern Cascades	27 (15.2%)	86 (48.3%)	20 (11.2%)	45 (25.3%)	178
	Blue Mountains	0 (0.0%)	4 (57.1%)	2 (28.6%)	1 (14.3%)	7
18	All Data	109 (35.7%)	78 (25.6%)	28 (9.2%)	90 (29.5%)	305
	NE Washington	34 (41.0%)	18 (21.7%)	6 (7.2%)	25 (30.1%)	83
	Northern Cascades	13 (35.1%)	12 (32.4%)	4 (10.8%)	8 (21.6%)	37
	Southern Cascades	61 (34.3%)	45 (26.4%)	17 (9.6%)	55 (30.9%)	178
	Blue Mountains	1 (14.3%)	3 (42.9%)	1 (14.3%)	2 (28.6%)	7

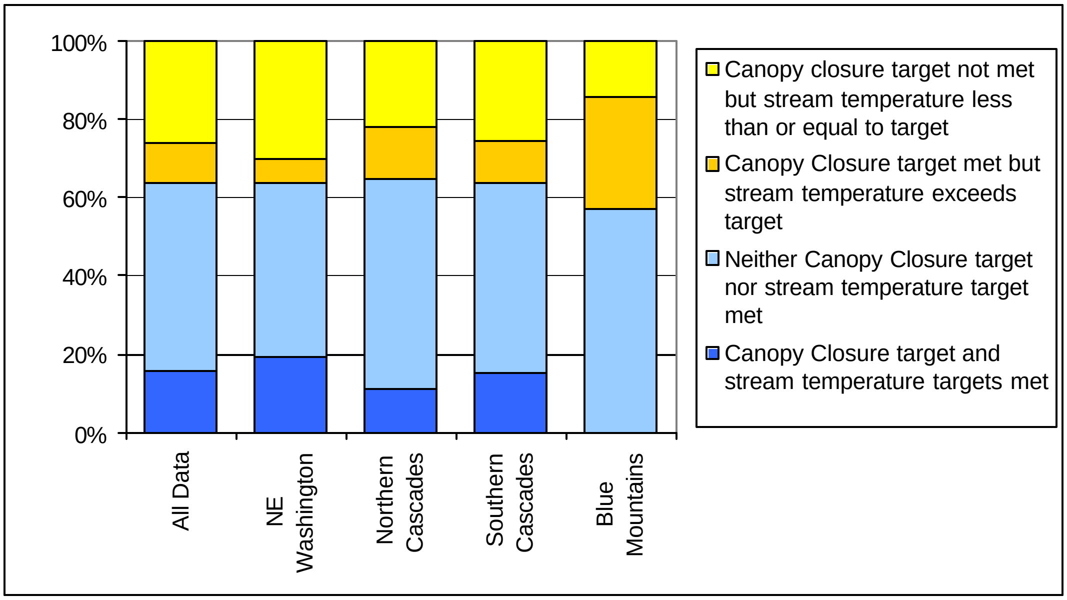


Figure 8. Graphical depiction of the four possible scenarios regarding the ability of the existing 16-degree nomograph to represent actual outcomes. Blue shades represent cases where the nomograph was "correct" and warmer shades represent cases where it over or under predicted the canopy closure required (see text). The darker gold color is represents the number of cases where the existing nomograph did not require sufficient canopy closure.

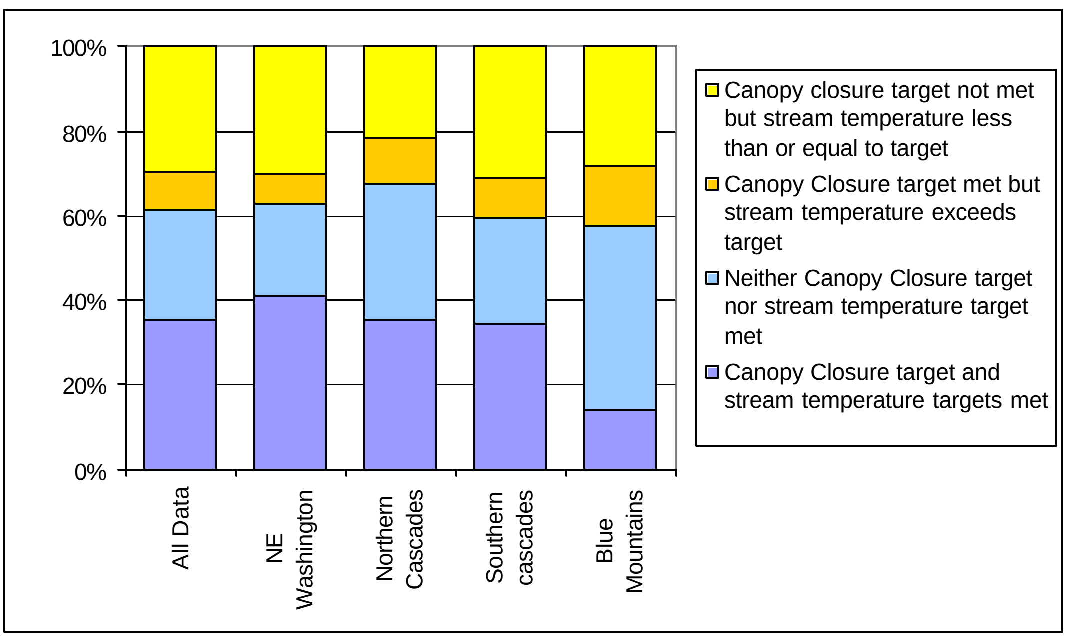


Figure 9. Graphical depiction of the four possible scenarios regarding the ability of the existing 16-degree nomograph to represent actual outcomes. Blue shades represent cases where the nomograph was "correct" and warmer shades represent cases where it over or under estimated the canopy closure required (see text). The darker gold color is represents the number of cases where the existing nomograph did not require sufficient canopy closure.

Table 5. Chi-square table (upper table), calculations (lower table), and results (lower table) of test of the null hypothesis that the existing 16-degree nomograph correctly represents actual situations 50 percent of the time.

	Result was as predicted by nomograph		**Nomograph over or underestimated the required canopy closure**		**Total**
REGION	**Observed**	**Expected**	**Observed**	**Expected**	
NE WA	53	62.25	30	20.75	83
S Cascades	113	133.5	65	44.5	178
N Cascades	24	27.75	13	9.25	37
Blue Mtns	4	5.25	3	1.75	7
TOTAL	194	228.75	111	76.25	305

	Result as predicted by nomograph	**Nomograph over or underestimated required canopy closure**	**Total**	**Degrees of freedom**	**probability**	
NE WA	1.37	4.12	5.50	1	0.01<p<0.025	Reject
S Cascades	3.15	9.44	12.59	1	p<<0.005	Reject
N Cascades	0.51	1.52	2.03	1	0.1<p<0.25	Accept
Blue Mtns	0.30	0.89	1.19	1	0.25<p<0.5	Accept
TOTAL	5.28	15.84	21.12	3	p<0.005	Reject

Table 6. Chi-square table (upper table), calculations (lower table), and results (lower table) of test of the null hypothesis that the existing 18-degree nomograph correctly represents actual situations 50 percent of the time.

	Result was as predicted by nomograph		**Nomograph over or underestimated the required canopy closure**		**Total**
REGION	**Observed**	**Expected**	**Observed**	**Expected**	
NE WA	52	62.25	31	20.75	83
S Cascades	106	133.5	72	44.5	178
N Cascades	25	27.75	12	9.25	37
Blue Mtns	4	5.25	3	1.75	7
TOTAL	187	228.75	118	76.25	305

	Result as predicted by nomograph	**Nomograph over or underestimated required canopy closure**	**Total**	**Degrees of freedom**	**Probability**	
NE WA	1.69	5.06	6.75	1	0.005<p<0.01	Reject
S Cascades	5.66	16.99	22.66	1	p<<0.005	Reject
N Cascades	0.27	0.82	1.09	1	0.25<p<0.5	Accept
Blue Mtns	0.30	0.89	1.19	1	0.25<p<0.5	Accept
TOTAL	7.62	22.86	30.48	3	p<0.005	Reject

Table 7. Chi-square table and results of test of hypothesis that the errors are evenly balanced on each side of the nomograph line for 16°C temperature target.

	Under-estimates Canopy Closure Needed to Meet 16°C Target		Over-estimates Canopy Closure Needed to Meet 16°C Target		
REGION	**Observed**	**Expected**	**Observed**	**Expected**	**Total**
NE WA	5	15	25	15	30
S Cascades	20	32.5	45	32.5	65
N Cascades	5	6.5	8	6.5	13
Blue Mtns	2	1.5	1	1.5	3
TOTAL	32	55.5	79	55.5	111

	Under-estimates Canopy Closure Needed to Meet 16°C Target	**Over-estimates Canopy Closure Needed to Meet 16°C Target**	**Total**	**Degrees of freedom**	**Probability**	**Result**
NE WA	4.00	7.50	11.50	1	p<<0.005	Reject
S Cascades	3.47	16.25	19.72	1	p<<0.005	Reject
N Cascades	0.28	3.25	3.53	1	0.05<p<0.10	Accept
Blue Mtns	0.25	0.75	1.00	1	0.25<0<0.50	Accept
TOTAL	6.99	27.75	34.74	3	p<<0.005	Reject

Table 8. Chi-square table and results of test of hypothesis that the errors are evenly balanced on each side of the nomograph line for 18°C temperature target.

	Under-estimates Canopy Closure Needed to Meet 18°C Target		Over-estimates Canopy Closure Needed to Meet 18°C Target		
REGION	**Observed**	**Expected**	**Observed**	**Expected**	**Total**
NE WA	6	15.5	25	15.5	31
S Cascades	17	36	55	36	72
N Cascades	4	6	8	6	12
Blue Mtns	1	1.5	2	1.5	3
TOTAL	28	59	90	59	118

	Under-estimates Canopy Closure Needed to Meet 18°C Target	**Over-estimates Canopy Closure Needed to Meet 18°C Target**	**Total**	**df**	**p**	**Result**
NE WA	3.61	7.75	11.36	1	p<0.005	Reject
S Cascades	6.56	18.00	24.56	1	p<<0.005	Reject
N Cascades	0.50	3.00	3.50	1	0.05<0<0.10	Accept
Blue Mtns	0.13	0.75	0.88	1	0.25<0<0.50	Accept
TOTAL	10.68	29.50	40.18	3	p<<0.005	Reject

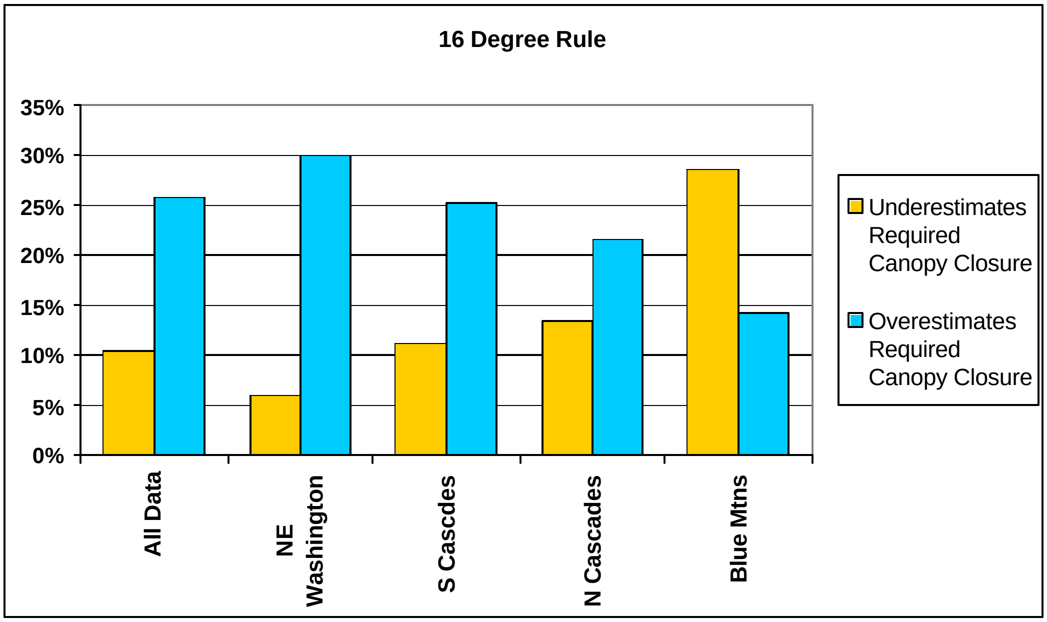


Figure 10. Frequency that the existing 16°C nomograph overestimates and underestimates the amount of canopy closure needed to meet the target temperature.

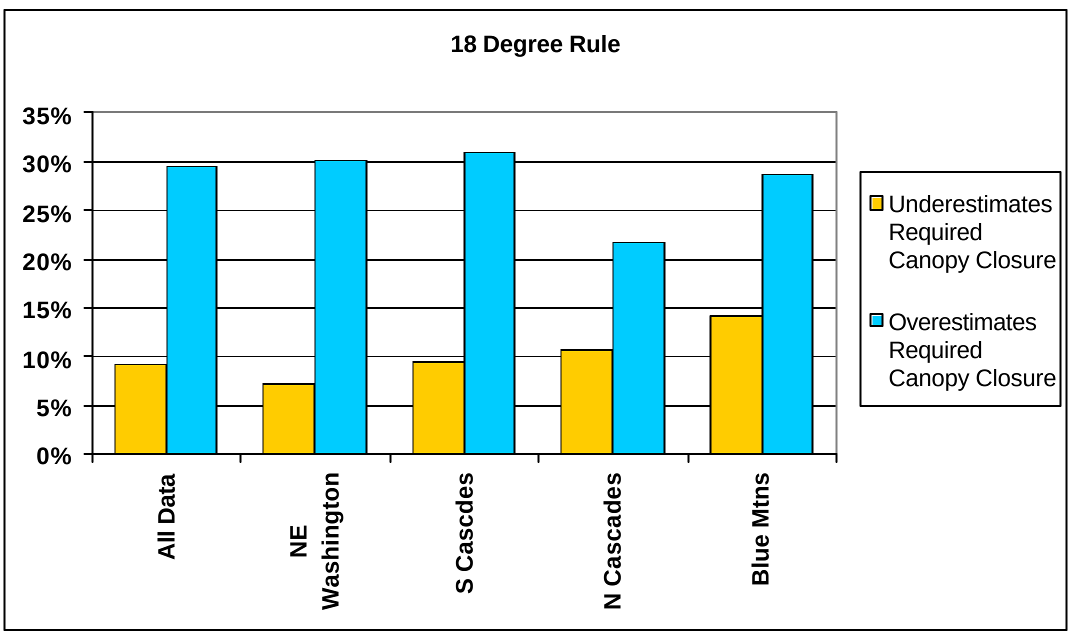


Figure 11. Frequency the existing 18°C nomograph overestimates and underestimates the amount of canopy closure needed to meet the target temperature.

3.4 REGRESSION MODELS WITH CANOPY CLOSURE AND ELEVATION AS ONLY INDEPENDENT VARIABLES

The discussion below includes the results of regression analyses with all variables included in the dataset and results for each of the various strata depicting region, ecoregions, and lithology. Throughout this document, there are references to a squared multiple r. This is the squared correlation coefficient, often written as r^2. The term multiple refers to the fact that the regressions are all multivariate analyses.

Note that the coefficients are reported in table form along with the confidence intervals and the statistical significance of the parameter included in the model. To develop predictive equations using the tabled information, equations should be written as follows:

Temperature (^oC) = constant coefficient + (Elevation * coefficient for elevation) + (canopy closure * coefficient for canopy closure),

where the coefficients are taken from the tables presented in each section, and the elevation and canopy closure are variables whose values are to be specified. Hence, if the coefficient for the constant were 2.3, the coefficient for elevation were 10.5 and the coefficient for canopy closure were 22.2, then the equation would be:

Temperature = 2.3 + 10.5 * elevation + 22.2 * canopy closure
(example only, does not represent an actual equation developed in this analysis).

3.4.1 Stratified by Region

NE Washington

The Regression analysis with all of the data points from Northeast Washington included in the equation identified two cases that tested as outliers. The regression was rerun without the identified outliers and the resulting equations were compared to determine if the removal of the outliers had a substantial effect on the resulting equation. The results of these two analyses are provided in Table 9. Both regressions tested as significant with a 2-tail probability of 0.000.

Removal of the outliers resulted in a substantial improvement in the portion of the variability in the data that is explained by the regression equation. A comparison of plots of the two equations depicts this difference (Figure 12). Based on this analysis, the model without the outliers was selected as the best fit. The relationship between canopy closure and elevation for a wide range of temperature targets is depicted in Figure 13.

The sample points used to develop the regression models are also displayed on Figure 13. The majority of the sample points were collected at elevations between 2000 and 4000 feet and in areas with more than 50% canopy closure. Low canopy closure situations are poorly represented in the database. As a result, the dataset is biased. A larger number of samples taken from sites with lower canopy closure levels could result in a substantial change in the predictive equations.

There are several figures in this text similar to Figure 13. Some explanation regarding the interpretations may be useful. For Figure 13 and all other similar figures, the expected stream temperature for any specific combination of canopy closure and elevation can be determined by finding the temperature nearest the intersection of a vertical line drawn from the elevation of interest and a horizontal line drawn from the canopy closure of interest. Alternatively, the canopy closure required at a particular elevation can be determined by following a vertical line up from the elevation of interest until it intersects the desired temperature target. Then follow a horizontal line from the point of intersection on the temperature target line to the canopy closure axis. The value on the canopy closure axis is the amount of canopy closure that would be required to meet the selected temperature target.

Table 9. Results of regression analyses for Northeast Washington with and without the identified outliers removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	28.868872	25.325134	32.412609	0.00000
Elevation	-0.00294	-0.004172	-0.001709	0.00000
Canopy Closure	-0.06541	-0.09793	-0.032889	0.00014
n				86
Multiple R				0.627947
Squared Multiple R				0.394318
Adjusted squared multiple R				0.670021
2 Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0
Constant	30.258771	27.766957	32.750584	0.07679
Elevation	-0.00347	-0.004337	-0.002603	0.01464
Canopy Closure	-0.070824	-0.093296	-0.048351	0
n				84
Multiple R				0.796461
Squared Multiple R				0.63435
Adjusted squared multiple R				0.625322

A plot of the residuals from the regression without the outliers (Figure 14) depicts the deviation of actual measured values from values predicted by the model. For northeast Washington, the deviation ranges from -6.23 to 7.08 degrees. The pattern in the residuals indicates no need for a transformation.

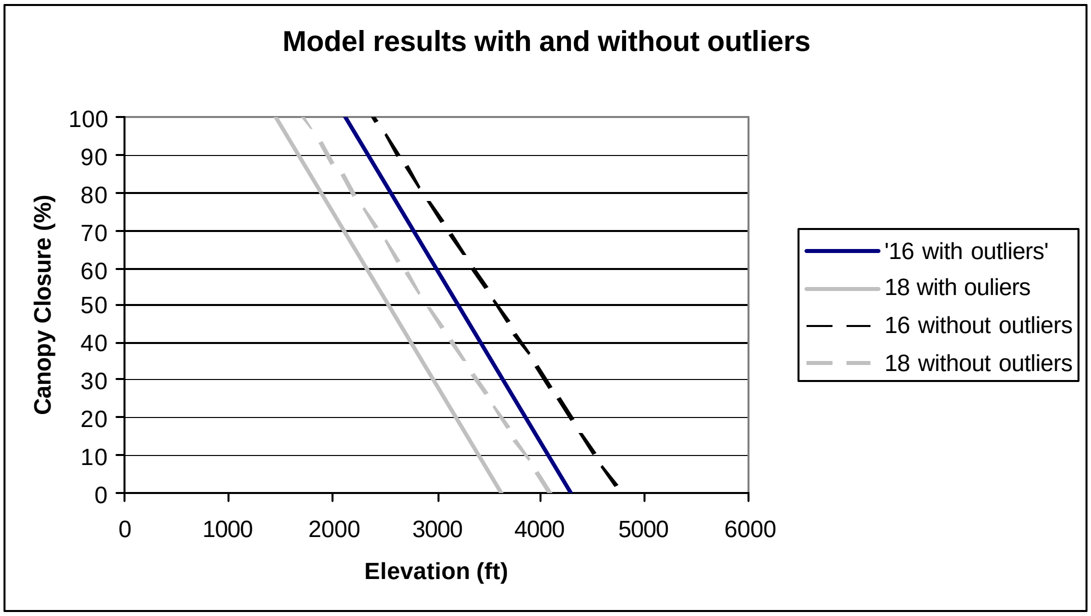


Figure 12. Comparison of the lines predicting 16 and 18 degrees centigrade resulting from regression analyses using NE Washington data with and without the identified outliers.

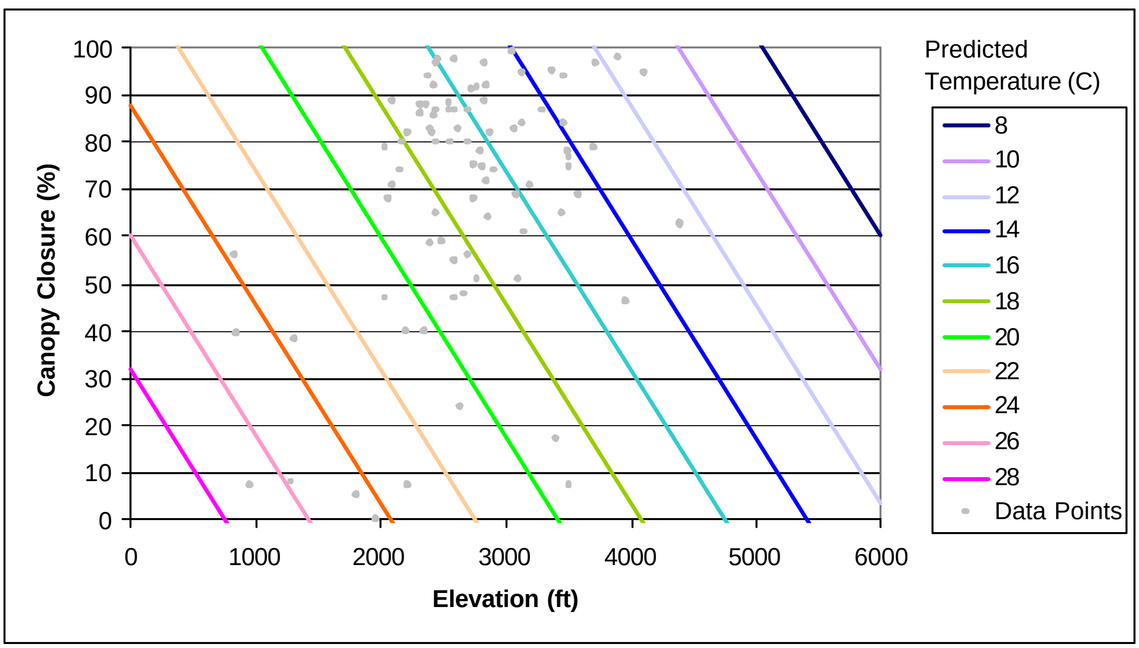


Figure 13. Regression model for NE Washington incorporating only canopy closure and elevation as independent variables. Each line corresponds to a different target temperature. The dots represent actual data points used in the analyses.

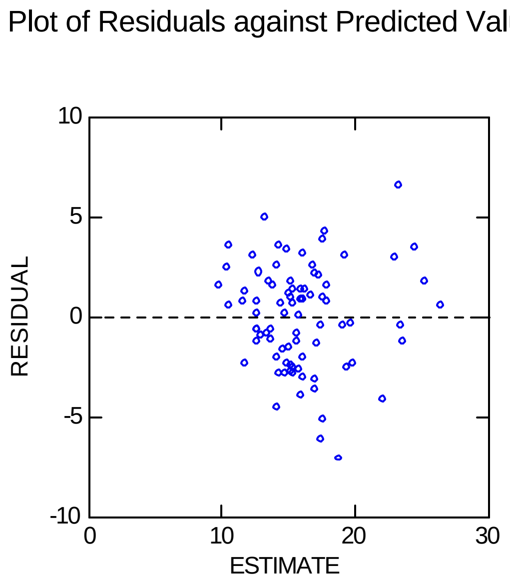


Figure 14. Plot of residuals (predicted minus actual value) for the regression using northeast Washington data without the outliers indicates a random distribution of errors around the regression line.

Northern Cascades

The regression analysis with data points from the Northern Cascades region of Washington did not identify any outliers. The results of the analysis are summarized in Table 10. The regression was highly significant (p=0.007). Canopy Closure, however, did not contribute significantly to the model. The probability levels calculated for the inclusion of elevation and canopy closure in the model were 0.032 and 0.248, respectively.

The relationship between canopy closure and elevation for a wide range of temperature targets is depicted in Figure 15. The sample points used to develop the regression models are also displayed on Figure 15. This figure was included to meet the objective of developing nomographs that aid in the identification of the appropriate amount of canopy closure to be left at harvest sites. The reader should keep in mind that canopy closure was far from significant in the regression equation.

Situations with low canopy closure at high elevation and high canopy closure at low elevation are poorly represented in the database. As a result, the dataset is biased. Additionally, all of the data points in this dataset come from basins with less than 30 inches of average annual precipitation (discussed in detail later; see Figure 70). Hence, the dataset represents only the driest portions of the region. The wetter areas at higher elevations in the north Cascades Mountains are not represented. A larger number of samples taken from under-represented conditions could result in a substantial change in the predictive equations and may result in the inclusion of canopy closure as a significant variable in the model.

Table 10. Results of regression analyses for the North Cascades region (no outliers identified). Canopy closure was not statistically significant at the p = 0.05 probability level (indicated in italics).

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00667
Constant	24.923794	21.39403	28.45355	0.00000
Elevation	-0.001767	-0.00337	-0.00016	0.03187
Canopy Closure	*-0.033944*	*-0.09266*	*0.024772*	*0.24822*
n				37
Multiple R				0.50524
Squared Multiple R				0.255267
Adjusted squared multiple R				0.211459

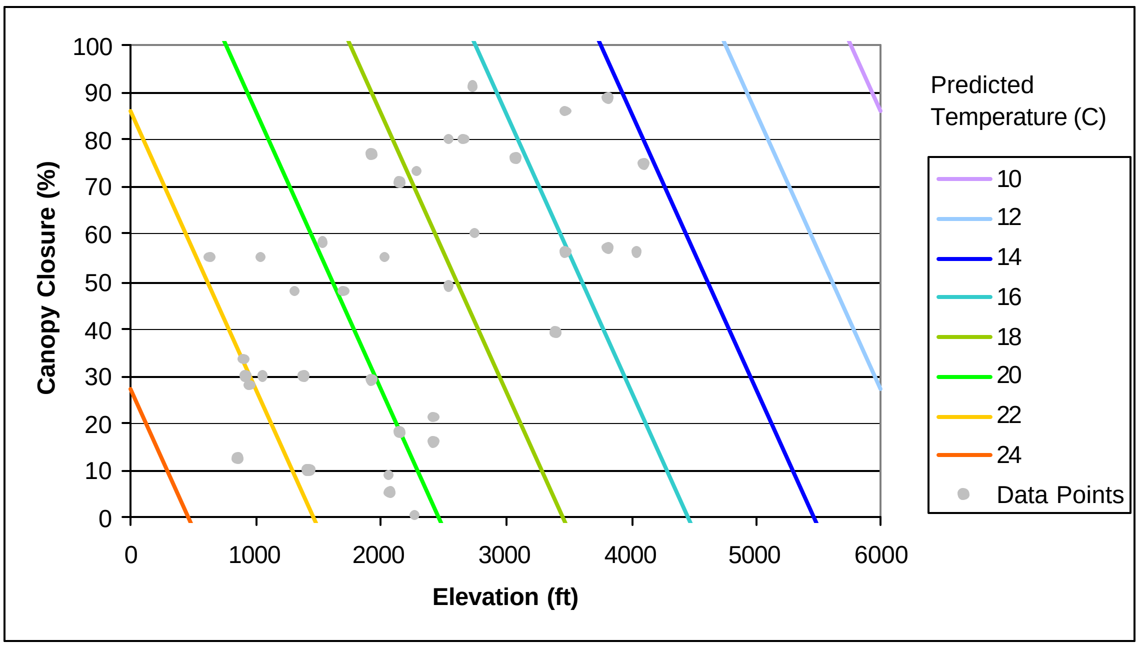


Figure 15. Regression model for North Cascades area of Washington incorporating only canopy closure and elevation as independent variables. Each line corresponds to a different target temperature. The dots represent the data used in the analyses. Note, canopy closure, though included in this model, did not test as a significant variable. Also note, for consistency sake, the depicted regression line was extended beyond the limits of the data. The accuracy of the line beyond the extent of the data is unknown.

A regression analysis was also run incorporating only elevation as the independent variable. The results of this analysis are summarized in Table 11 and depicted in Figure 16. Plots of the residuals from each of the regression analyses (Figure 17) depict the deviation of actual measured values from values predicted by the model. Statistically, this is the best-fit model.

Table 11. Results of the regression analysis for the North Cascades region incorporating only elevation as the independent variable.

Elevation As Only Independent Variable				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.003014
Constant	24.338613	20.94236	27.734865	0
Elevation	-0.002216	-0.003627	-0.000805	0.00301
n				37
Multiple R				0.474377
Squared Multiple R				0.225034
Adjusted squared multiple R				0.202892

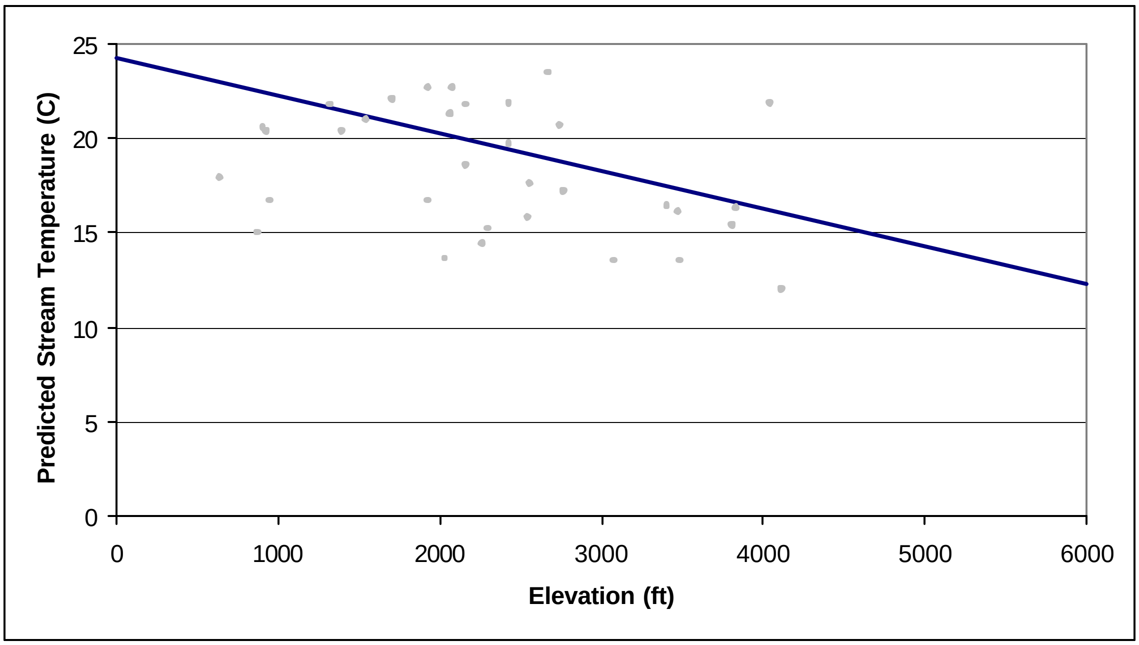


Figure 16. Regression model for the North Cascades area of Washington with elevation only. Canopy Closure was not included because it did not test as a significant variable. Data points are depicted as gray dots and reflect actual rather than predicted temperature. Note, for consistency sake, the depicted regression line was extended considerably beyond the limits of the data. The accuracy of the line beyond the extent of the data is unknown.

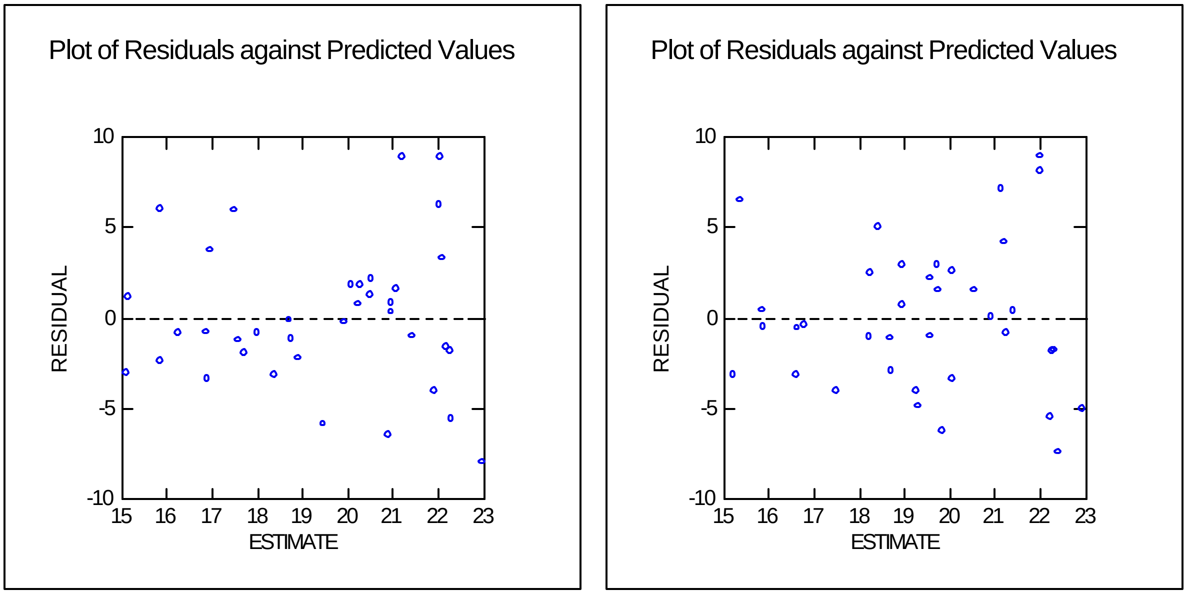


Figure 17. Plots of residuals for the regression analyses conducted with the dataset from the North Cascades area of Washington. The plot on the left is the residual plot for the regression equation that incorporated both elevation and canopy closure. The plot on the right is for the regression equation incorporating only elevation. Both indicate a random distribu tion of errors around the regression line.

Southern Cascades

Several outliers were identified during the regression analyses conducted for the southern Cascades region of eastern Washington. Regressions were completed eliminating the outliers as they were identified. The values for the equations, the sample size, number of outliers removed, and correlation coefficients derived through the regression analyses are summarized in Table 12.

Removal of four of the outliers resulted in a substantial improvement in the portion of the variability in the data that is explained by the regression equation. The fifth outlier had little effect on the overall model. A comparison of plots depicting the predictive equations of the various models is presented in Figure 18 and Figure 19. Based on this analysis, the model without the first four outliers was selected as the best-fit model.

The relationship between canopy closure and elevation for a wide range of temperature targets using the model with four outliers removed is depicted in Figure 20. The sample points used to develop the regression models are also displayed on Figure 20. The majority of the sample points were collected at elevations between 2000 and 4000 feet, however other elevations are reasonably represented. The full range of canopy closures is also reasonably represented. The difference between predicted stream temperatures and actual measured temperatures is depicted in the plot of residuals for the model with four outliers removed (Figure 21).

Table 12. Results of regression analyses for the Southern Cascades region with all data included in the analysis and with various numbers of identified outliers excluded.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	25.571806	23.292549	27.851063	0.00000
Elevation	-0.001985	-0.00268	-0.00129	0.00000
Canopy Closure	-0.058467	-0.079354	-0.03758	0.00000
n				175
Multiple R				0.51440
Squared Multiple R				0.26461
Adjusted squared multiple R				0.25606
2 Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	25.505331	23.381164	27.629499	0.00000
Elevation	-0.00193	-0.002574	-0.001286	0.00000
Canopy Closure	-0.060308	-0.079926	-0.04069	0.00000
n				173
Multiple R				0.5434
Squared Multiple R				0.295284
Adjusted squared multiple R				0.286993
4 Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	26.283985	24.284368	28.283603	0.00000
Elevation	-0.002188	-0.002804	-0.001572	0.00000
Canopy Closure	-0.061235	-0.079795	-0.042674	0.00000
n				171
Multiple R				0.597823
Squared Multiple R				0.357392
Adjusted squared multiple R				0.349742
5 Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	26.235624	24.310164	28.161084	0.00000
Elevation	-0.002223	-0.002816	-0.00163	0.00000
Canopy Closure	-0.059596	-0.077487	-0.041705	0.00000
n				170
Multiple R				0.61208
Squared Multiple R				0.374642
Adjusted squared multiple R				0.367153

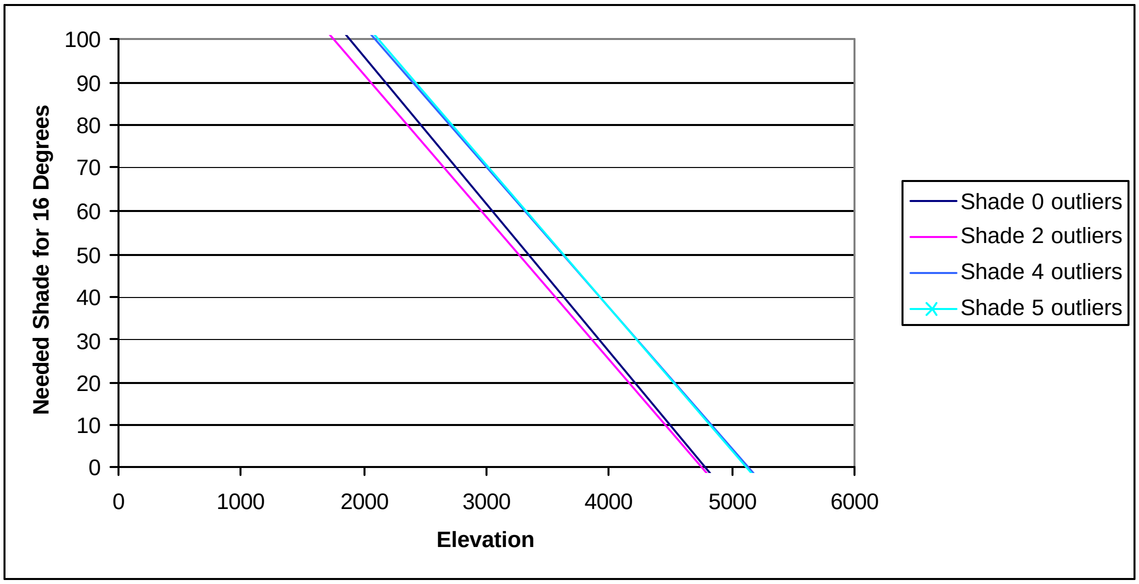


Figure 18. Canopy closure needed to meet a 16 degrees C temperature target in the southern Cascades area of eastern Washington based on 4 regression models with varying number of outliers removed.

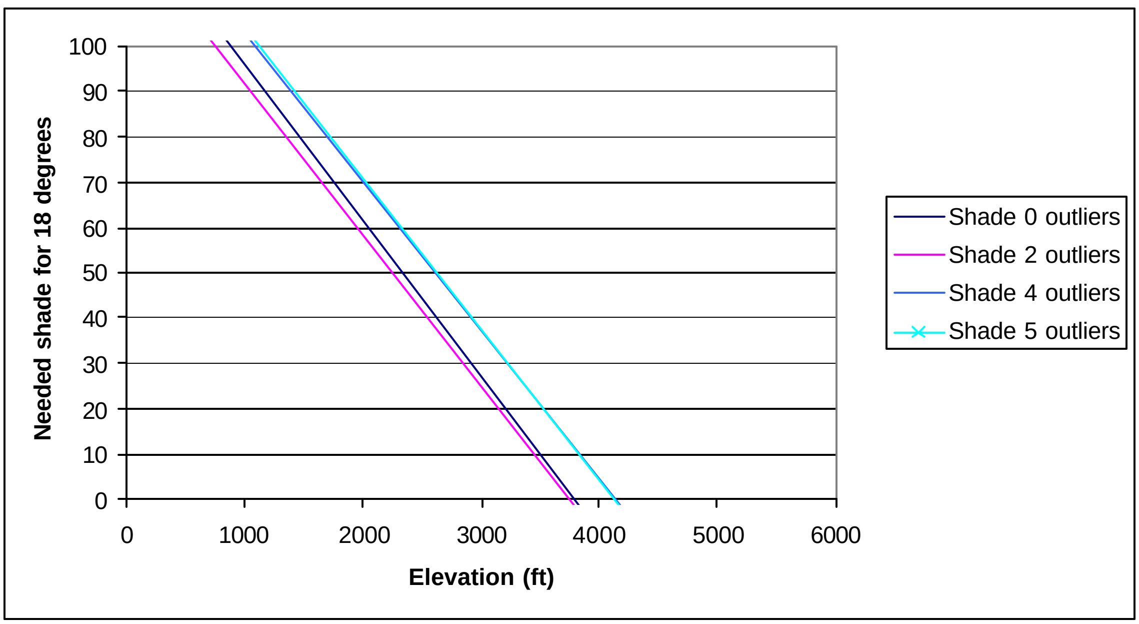


Figure 19. Canopy closure needed to meet a 18 degrees C temperature target in the southern Cascades area of eastern Washington based on 4 regression models with varying number of outliers removed.

Blue Mountains

There were only seven data points representing the Washington portion of the Blue Mountains in the data set. The regression predicting stream temperature as a function of canopy closure and elevation was not significant at the p=0.05 level. Outliers were identified in the regression analysis. When these outliers were removed, more outliers were identified. This continued until all but 2 points were removed. Given the small sample size, the regression equations with outliers removed are not reported. The results of the analysis using only Washington data with all data points in the analysis are presented in Table 13. The regression was not significant at the 0.05 probability level and canopy closure did not significantly contribute to the equation (p = 0.895). The poor performance of the regression analysis may be related to the extremely small sample size for this dataset.

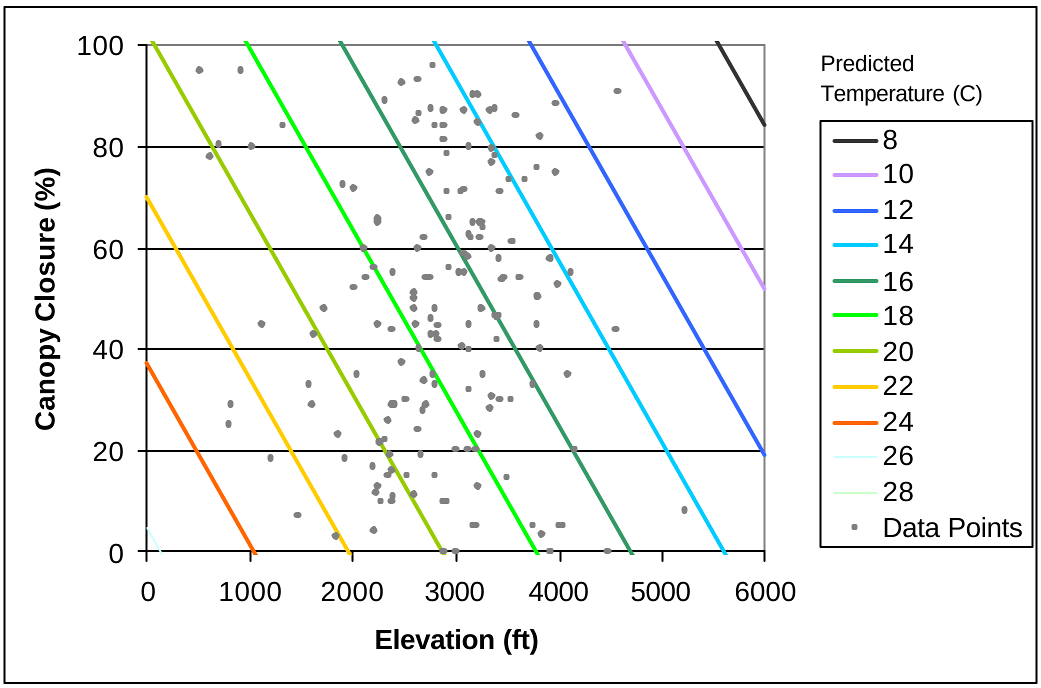


Figure 20. Regression model for the southern Cascades area of eastern Washington incorporating only canopy closure and elevation as independent variables. Four outliers were removed. Each line corresponds to a different target temperature (legend). The dots represent the data used in the analyses.

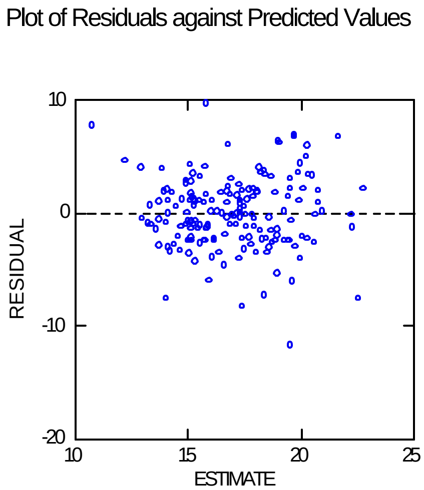


Figure 21. Plot of residuals for the regression analyses conducted with the dataset from the southern Cascades area of eastern Washington with four outliers removed. The plot indicates a relatively random distribution of data around the mean (balanced errors).

Table 13. Results of regression analysis for the Blue Mountain data of Washington using only canopy closure and elevation as independent variables.

All Washington State Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				*0.09439*
Constant	34.083987	12.988313	55.179662	0.01094
Canopy Closure	*-0.007256*	*-0.15076*	*0.13625*	*0.89514*
Elevation	-0.005029	-0.009855	-0.000203	0.04443
n				7
Multiple R				0.832329
Squared Multiple R				0.692771
Adjusted squared multiple R				0.539156

An additional dataset was available from the Oregon side of the Blue Mountains. This dataset was evaluated to see if the two sets of data were comparable. Since the model developed using only Washington data were not significant, a test of two models to determine if they were significantly different is not recommended. Therefore, a model using only the Oregon data was developed (Table 14). The regression and both independent variables were significant at the p = 0.05 probability level. The residuals of the Oregon data were plotted (Figure 22).

The predicted stream temperature for the Washington data points was estimated using the Oregon model. The resulting residuals were also plotted to determine if those residuals fell within the range of the residuals from the Oregon data (Figure 22). The plot of the residuals of the Oregon regression model with the residuals of the Washington data points added indicates that all but one of the Washington residuals lies within the distribution of the Oregon residuals (Figure 22). Based on this evaluation, the Washington and Oregon datasets were combined to develop a new model. One outlier was identified during the regression analysis. The results of the regression analyses with and without the outlier removed are provided in Table 15.

Table 14. Results of regression analysis for the Blue Mountain data of Oregon using only canopy closure and elevation as independent variables.

All Oregon State Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000013
Constant	32.982786	26.717387	39.248184	0.00000
Canopy Closure	-0.083652	-0.126081	-0.041223	0.00028
Elevation	-0.002107	-0.003524	-0.000689	0.00461
n				42
Multiple R				0.662218
Squared Multiple R				0.438532
Adjusted squared multiple R				0.409739

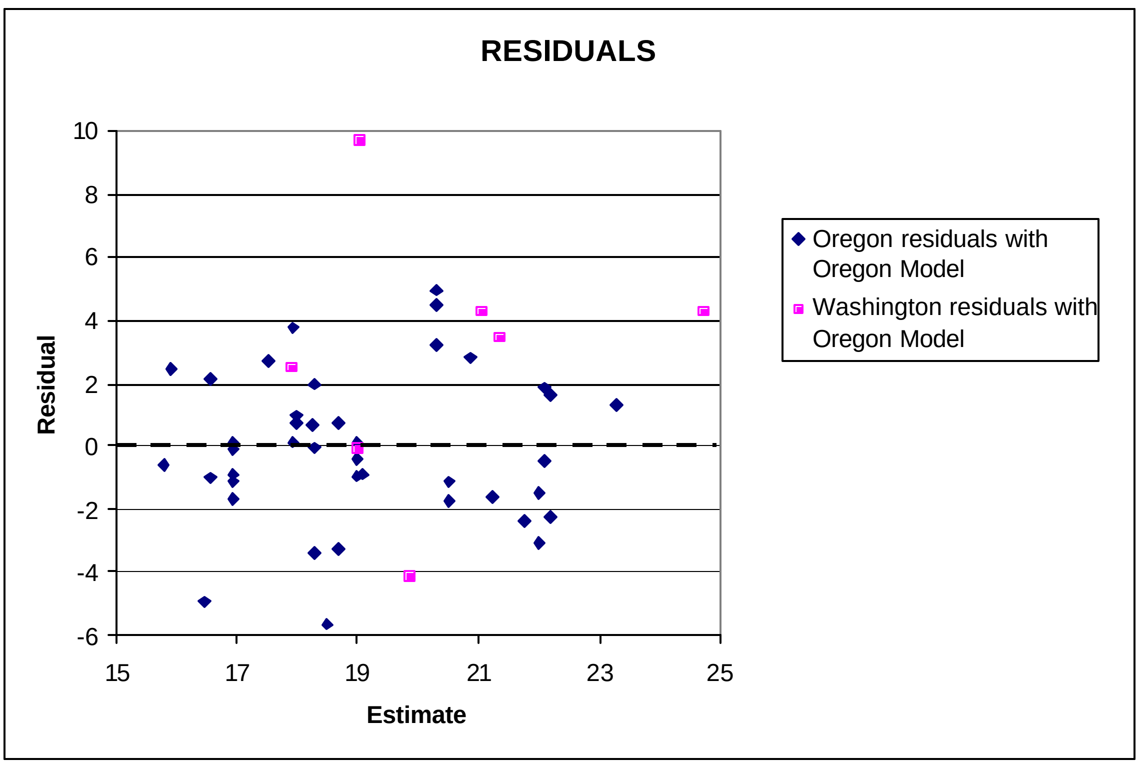


Figure 22. Residuals of the model developed using only Oregon data with canopy closure and elevation as the only independent variables. The Washington data points were developed using the Oregon model with the appropriate canopy closure and elevation for the Washington points. The plot indicates a relatively random distribution of the Oregon data around the mean (balanced errors). Six of the seven Washington data points lie within the range of variability of the Oregon data points. The seventh point tested as an outlier.

Table 15. Results of the regression analysis for the combined Oregon and Washington Blue Mountain using only canopy closure and elevation as independent variables.

Washington and Oregon State Data Combined				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00027
Constant	30.58616	24.804193	36.368127	0.00000
Canopy Closure	-0.072658	-0.115074	-0.030243	0.00122
Elevation	-0.001788	-0.002969	-0.000607	0.00381
n				49
Multiple R				0.547683
Squared Multiple R				0.299956
Adjusted squared multiple R				0.26952
Washington and Oregon State Data Combined, One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000055
Constant	30.744273	25.613875	35.874671	0.000000
Canopy Closure	-0.070856	-0.1085	-0.033213	0.000440
Elevation	-0.001805	-0.002853	-0.000757	0.001160
n				41
Multiple R				0.594335
Squared Multiple R				0.353234
Adjusted squared multiple R				0.324489

The relationship between canopy closure and elevation for a wide range of temperature targets is depicted in Figure 23. The majority of the sample points were collected at elevations above 3000 feet (Figure 23), which roughly corresponds to the elevations where timber is found in that area. Canopy closure is reasonably well represented in the dataset across the range of elevations present in the dataset. The deviations between the predicted stream temperature and the temperatures measured in the field are depicted in Figure 24.

Tests of Significant Differences between Equations for Each Region

The equations that were developed for each region appear to be similar for all regions except the Blue Mountains (Figures 25 and 26). The equations generated for each region and for all regions as a whole using elevation and canopy closure as independent variables were tested to determine if the various datasets were from one population (Zar 1974). All populations were found to be significantly different from each other at the 0.5 probability level (Table 16).

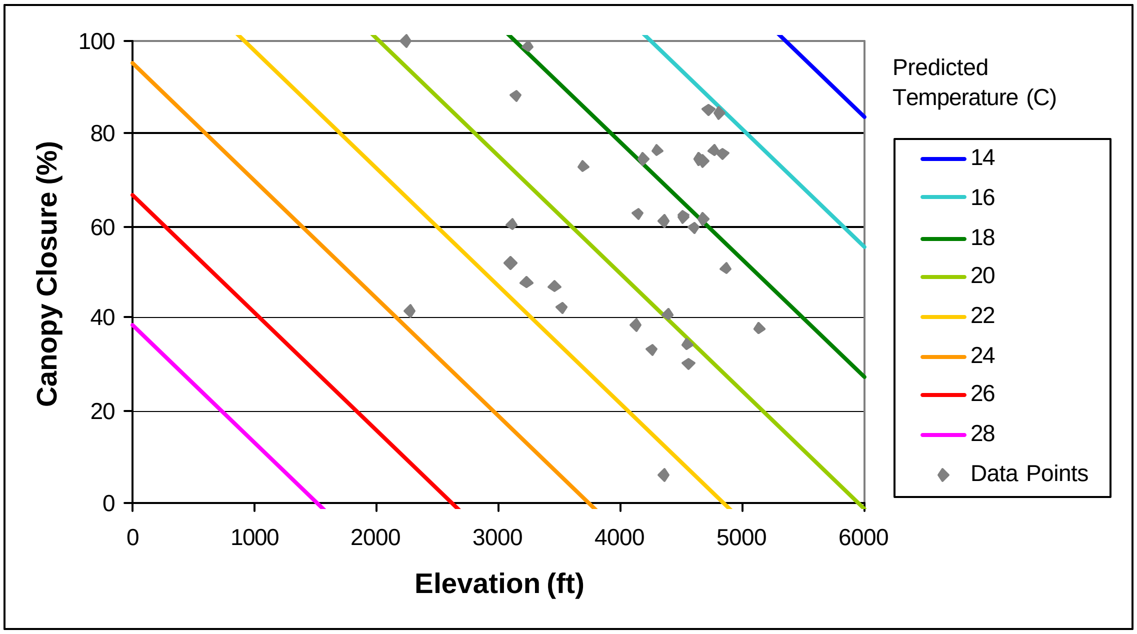


Figure 23. Regression model for the Blue Mountains (Washington and Oregon data combined) incorporating only canopy closure and elevation as independent variables. One outlier was removed. Each line corresponds to a different target temperature. The dots represent the actual data in the dataset. The lines extend beyond the range of the data and the accuracy of the lines beyond the range of data is unknown.

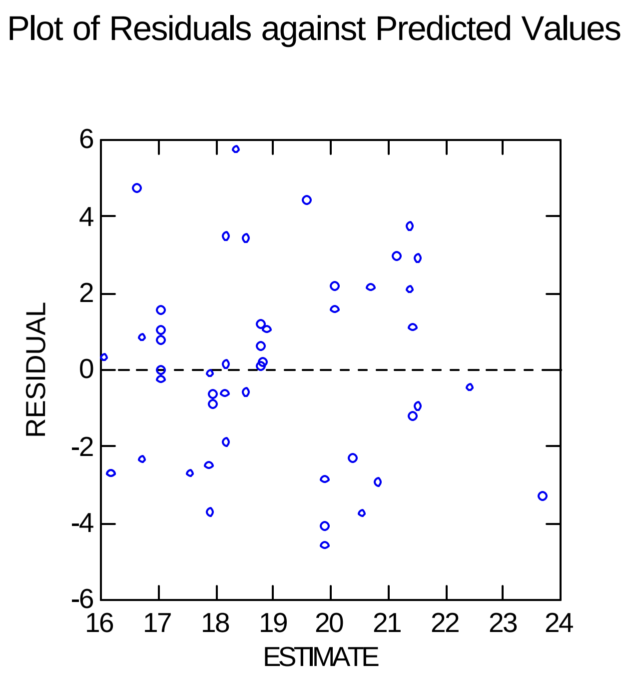


Figure 24. Plot of residuals for the regression analyses conducted with the combined Oregon and Washington dataset for the Blue Mountains with one outlier removed. The plot suggests a random distribution of errors around the regression line.

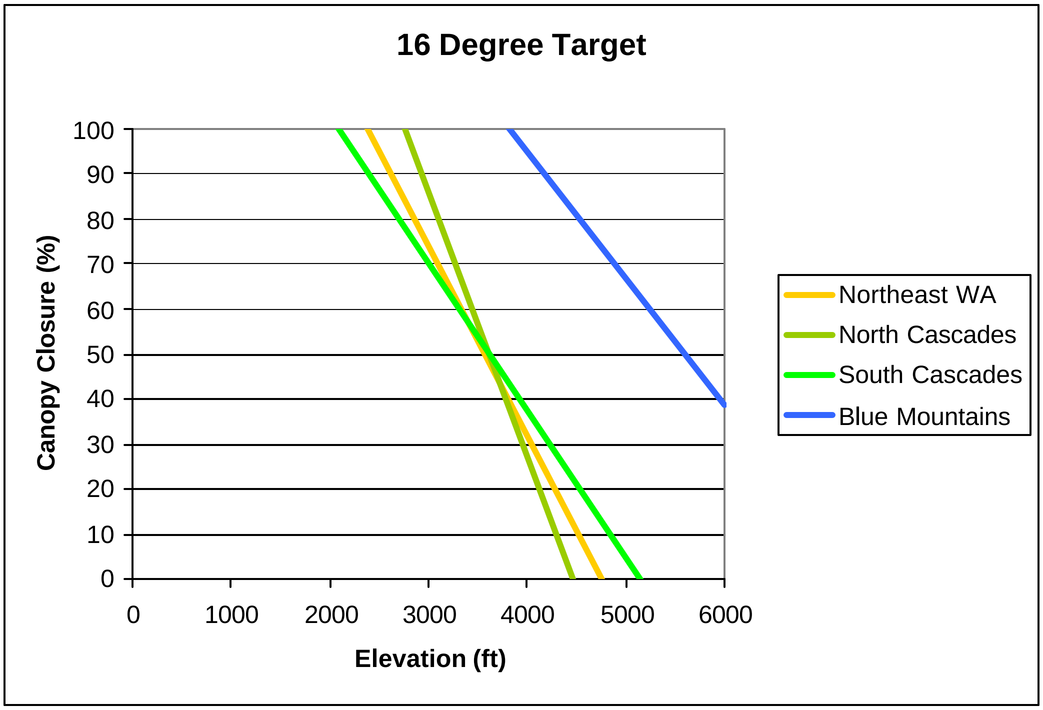


Figure 25. Canopy closure required to attain a 16 degree C temperature target for each region based on results of regression analyses using canopy closure and elevation as independent variables.

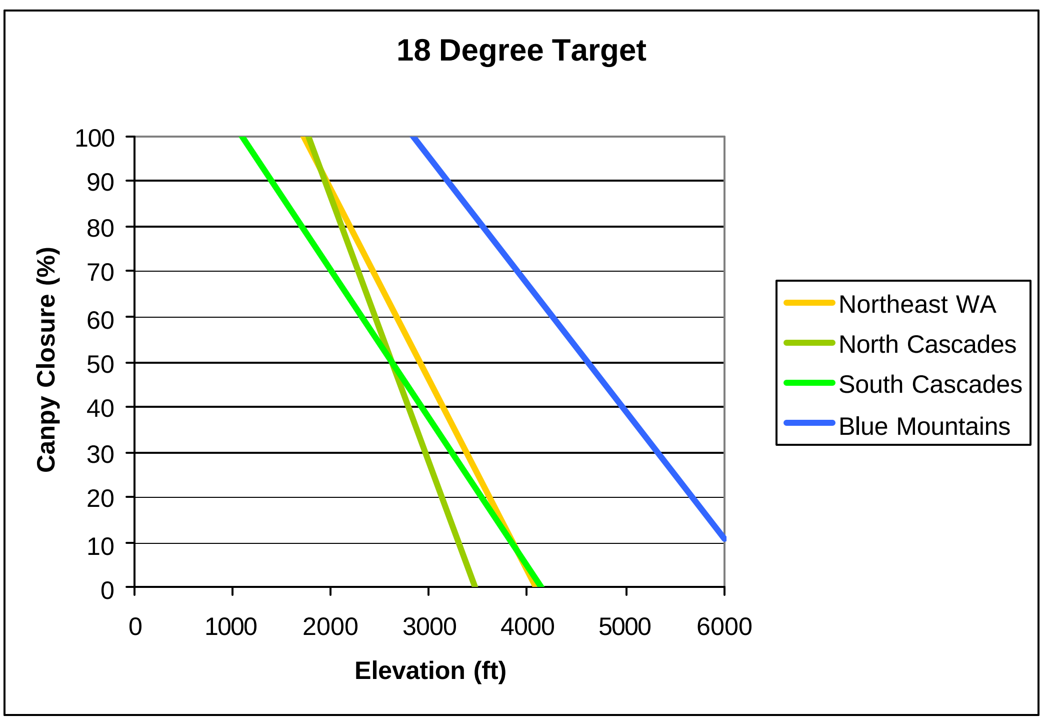


Figure 26. Canopy closure required to attain an 18 degree C temperature target for each region based on results of regression analyses using canopy closure and elevation as independent variables.

Table 16. F statistics and estimated probability to reject the null hypothesis that two populations are the same (in parentheses) for pairs of datasets representing the various regions. In all cases, the null hypothesis was rejected and populations were found to be significantly different.

REGION	South Cascades	NE Washington	Blue Mountains
North Cascades	14.738 (p<.0005)	4.127 (.0025<p<.005)	7.239 (p<0.0005)
South Cascades		3.186 (0.005<p<0.01)	14.988 (p<0.0005)
NE Washington			9.748 (p<0.0005)

3.4.2 Stratified by Ecoregion

Regression analyses were conducted with the data stratified by ecoregion. There are six ecoregions represented in the data (Table 17). Ecoregion -331A (Palouse Prairie Section) only had one data point. Therefore, no regression could be completed for this region. No ecoregion was available for 29 of the data points. Most of the data points with these missing values were from the area of the southern Cascade Mountains.

M333A (Okanogan Highlands Section)

Two outliers were identified during the regression analyses conducted for the M333A ecoregions (Okanogan Highlands Section). Regressions were completed eliminating the outliers as they were identified. Both regression equations and all the variables included in the regressions tested as significant with a 2-tail probability of 0.000. The values for the equations, the sample size, number of outliers removed, and correlation coefficients derived through the regression analyses are provided in Table 18.

Removal of the outliers resulted in a substantial improvement in the portion of the variability in the data that is explained by the regression equation and substantially modified the resulting predictive equation (Figure 27). Based on this analysis, the model without the outliers was selected as the best fit. The relationship between canopy closure and elevation for a wide range of temperature targets using the model with two outliers removed is depicted in Figure 28. The difference between predicted stream temperatures and actual measured temperatures is depicted in the plot of residuals for the model with two outliers removed (Figure 29).

There is a slight amount of bias in the dataset. Situations with high canopy closure and low elevation and situations with low canopy closure and high elevation are under-represented. Furthermore, low elevation data points are sparse relative to the number of points from elevations greater than 2500 feet. Additional data representing those situations that are currently under represented in the database may affect the regression results.

Table 17. Description of the domain, division, province, and section represented by each ecocode per Bailey (1995).

ECOCODE	DOMAIN	DIVISION	PROVINCE	SECTION	SAMPLE SIZE
M333A	Dry Domain	Temperate Steppe Regime Mountains	Northern Rocky Mountain Forest-Steppe-Coniferous Forest-Alpine Meadow Province	Okanogan Highlands Section	113
M242B	Humid Temperate Domain	Marine Regime Mountains Redwood Forest Province	Cascade Mixed Forest-Coniferous Forest-Alpine Meadow Province	Western Cascades Section	11
M242C	Humid Temperate Domain	Marine Regime Mountains Redwood Forest Province	Cascade Mixed Forest-Coniferous Forest-Alpine Meadow Province	Eastern Cascades Section	120
-342I	Dry Domain	Temperate Desert Division	Intermountain Semi-Desert Province	Columbia Basin Section	25
-331A	Dry Domain	Temperate Steppe Division	Great Plains-Palouse Dry Steppe Province	Palouse Prairie Section	1
M332G	Dry Domain	Temperate Steppe Regime Mountains	Middle Rocky Mountain Steppe-Coniferous Forest-Alpine Meadow Province	Blue Mountains Section	6

Table 18. Results of regression analyses with and without outliers removed for Ecoregion M333A (Okanogan Highlands Section).

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	25.571806	23.292549	27.851063	0.00000
Elevation	26.904328	24.231005	29.577651	0.00000
Canopy Closure	-0.002476	-0.003439	-0.001514	0.00006
n	-0.054858	-0.08088	-0.028835	113
Multiple R				0.607394
Squared Multiple R				0.368927
Adjusted squared multiple R				0.357453
2 Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	27.734495	25.628922	29.840068	0.00000
Elevation	-0.002771	-0.003529	-0.002013	0.00000
Canopy Closure	-0.06022	-0.080506	-0.039934	0.00000
n				111
Multiple R				0.734725
Squared Multiple R				0.53982
Adjusted squared multiple R				0.531299

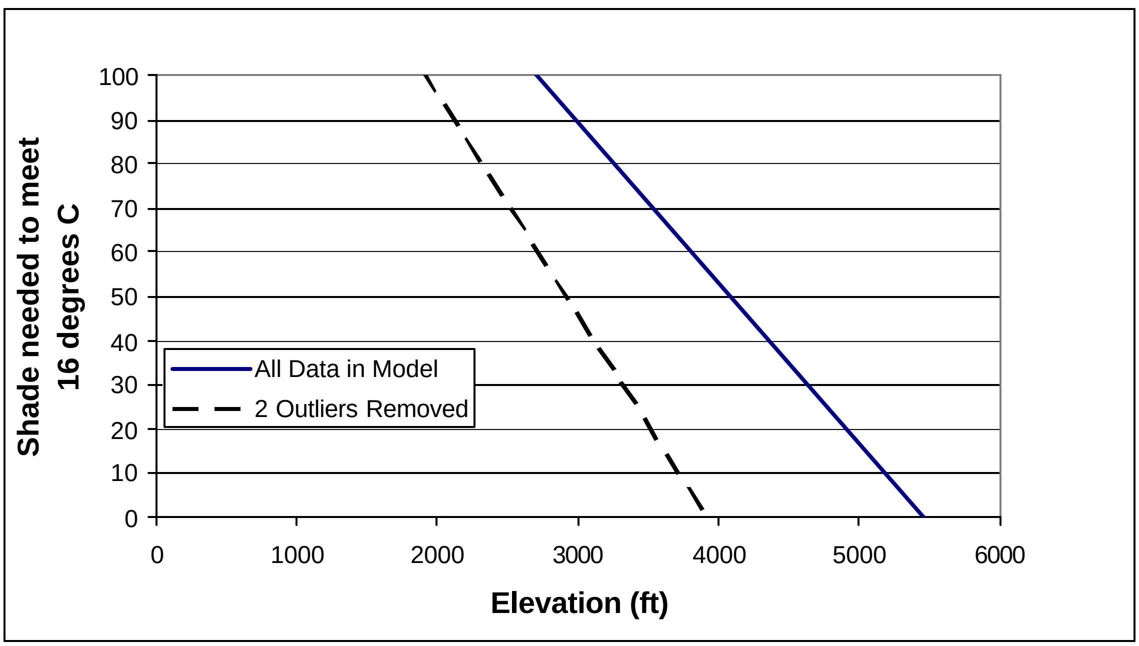


Figure 27. Comparison of the Predictive models with and without the outliers included. The figure shows the differences in the percent canopy closure required to meet a 16 degree C temperature target for both models.

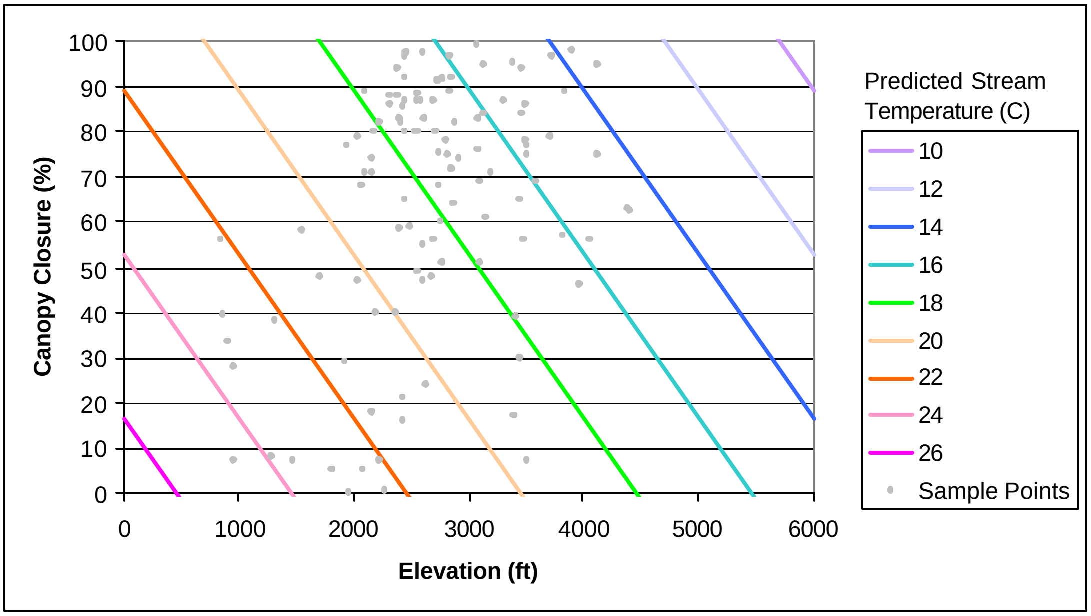


Figure 28. Regression model for the M333A Ecoregion (Okanogan Highlands Section) incorporating canopy closure and elevation as independent variables. One outlier was removed. Each line corresponds to a different target temperature. The dots correspond to actual sample points in the dataset.

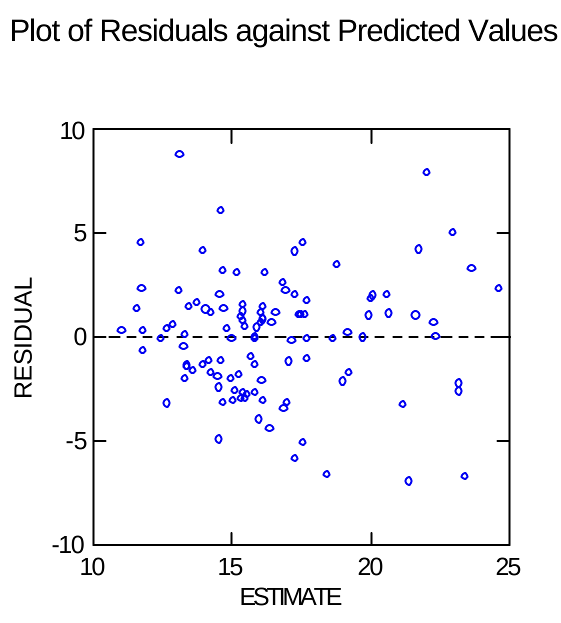


Figure 29. Plot of residuals for the regression analyses conducted with the dataset for the M333A ecoregions (Okanogan Highlands Section) with one outlier removed. The plot suggests a random distribution of errors around the regression line.

M242B (Western Cascades Section)

All of the data points in Ecoregion M242B are in the southern portion of the Eastern Cascade Mountains. Two outliers were identified in the regression analysis using elevation and canopy closure as independent variables. The effect of these outliers is depicted in Figure 30.

Without the removal of the first outlier, elevation is not a significant variable for inclusion in the model. With the removal of the first and second outliers identified, canopy closure no longer is found to contribute significantly to the model (Table 19).

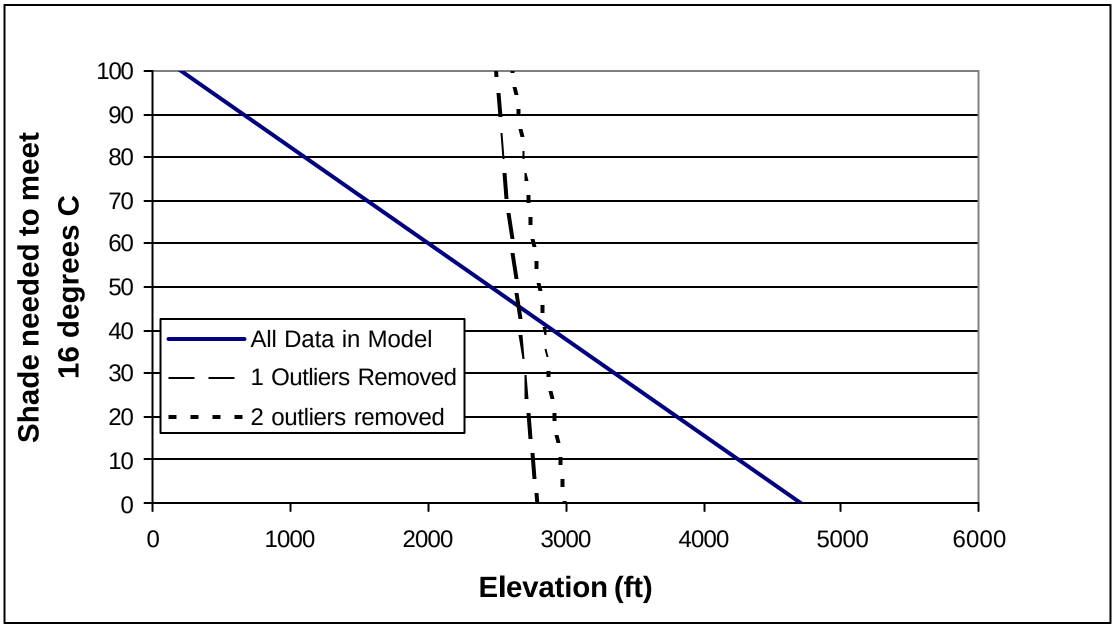


Figure 30. Effect of removing outliers on the prediction of the amount of canopy closure required to meet a temperature standard of 16 degrees C.

The results of the regression models incorporating both canopy closure and elevation were further explored to determine which outliers, if any, should be removed. The difference in the regression lines between the model with all data included and the models with one or two outliers removed is substantial (Figure 30). The removal of the first outlier appears to be justified. The difference in the regression lines between that developed with one outlier removed and that with two outliers removed is not substantial. Therefore, the removal of the second outlier may not be justified. The best-fit model with both canopy closure and elevation included as independent variables is therefore assumed to be the model with one outlier removed, although canopy closure is not significant in this equation at the 0.05 probability level. The resulting model developed with one outlier removed is depicted in Figure 31. The data points used in that model are also depicted in the figure.

Table 19. Results of regression analyses for Ecoregion M242B (western Cascades section) with all data in the analysis and with one and two outliers removed. Italic cells indicate variables that are not significant at p≤0.05.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.012018
Constant	25.417429	16.394731	34.440127	0.00019
Elevation	*-0.002204*	*-0.004778*	*0.000369*	*0.08370*
Canopy Closure	-0.090132	-0.151253	-0.02901	0.00935
n				11
Multiple R				0.817864
Squared Multiple R				0.668902
Adjusted squared multiple R				0.586127
One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00158
Constant	27.162268	21.83042	32.49412	0.00001
Elevation	-0.003917	-0.00569	-0.00214	0.00123
Canopy Closure	*-0.012172*	*-0.06836*	*0.044017*	*0.62425*
n				10
Multiple R				0.917377
Squared Multiple R				0.841581
Adjusted squared multiple R				0.796319
Two Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000455
Constant	27.97327	23.755998	32.190541	0.00000
Elevation	-0.004178	-0.005581	-0.002776	0.00034
Canopy Closure	*-0.01484*	*-0.058585*	*0.028905*	*0.43827*
n				9
Multiple R				0.960765
Squared Multiple R				0.923069
Adjusted squared multiple R				0.897426

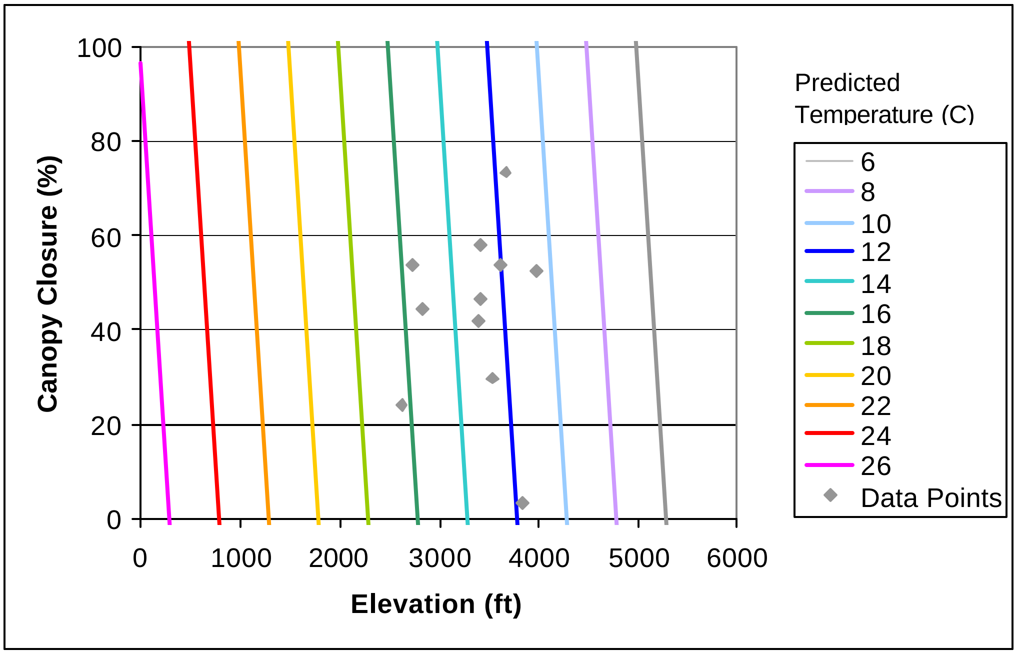


Figure 31. Depiction of the model results from the regression for the M424B ecoregion (Western Cascades Section) with one outlier removed. The gray dots represent the actual data used in the analysis. The lines correspond to the predicted temperatures indicated in the legend. Canopy closure was not significant in this equation. Note the lack of representation of high and low elevation situations in the dataset. The accuracy of the model beyond the range of the data is unknown.

The elevation range represented by the data is rather narrow and most of the data points are clustered around a small range of canopy closure. The narrow range of the input data has affected the regression results. A different outcome would be expected with a larger database drawn from a wider range of conditions.

Statistically, the best-fit model is one that does not include canopy closure (Table 20). In the regression using only elevation as an independent variable, one outlier was identified that made elevation non-significant. With that outlier removed, one more was identified. The models with one and two outliers removed are depicted in Figure 32. There is little difference in the plots of the two lines. The squared multiple r increases with the second outlier removed, but given the narrow range of samples and the marginal difference between the plots, the removal of the second outlier is not justified. Therefore, the model with one outlier removed and only elevation included as an independent variable must be considered the best-fit model for this dataset. The residuals for that model are depicted in Figure 33. The fit to the data in this model is quite tight, however the range of elevations represented by the data is narrow. Additional data representing a wider range of elevations would likely have a substantial effect on the analysis results.

Overall, we have little confidence in the results of any models developed for this ecoregions and do not recommend their use. Additional data for this ecoregions would improve the confidence in the results.

Table 20. Results of regression analyses for Ecoregion M242B using only elevation as an independent variable with and without the removal of outliers.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.179774
Constant	22.091255	9.455953	34.726558	0.003330
Elevation	-0.00239	-0.006108	0.001327	0.179770
n				11
Multiple R				0.436265
Squared Multiple R				0.190327
Adjusted squared multiple R				0.100364
One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000214
Constant	27.146481	22.192776	32.100186	0.000000
Elevation	-0.004088	-0.005566	-0.00261	0.000210
n				10
Multiple R				0.914135
Squared Multiple R				0.835643
Adjusted squared multiple R				0.815098
Two Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000056
Constant	27.938418	23.955707	31.921128	0.000000
Elevation	-0.00438	-0.005579	-0.003181	0.000060
n				9
Multiple R				0.956156
Squared Multiple R				0.914235
Adjusted squared multiple R				0.901983

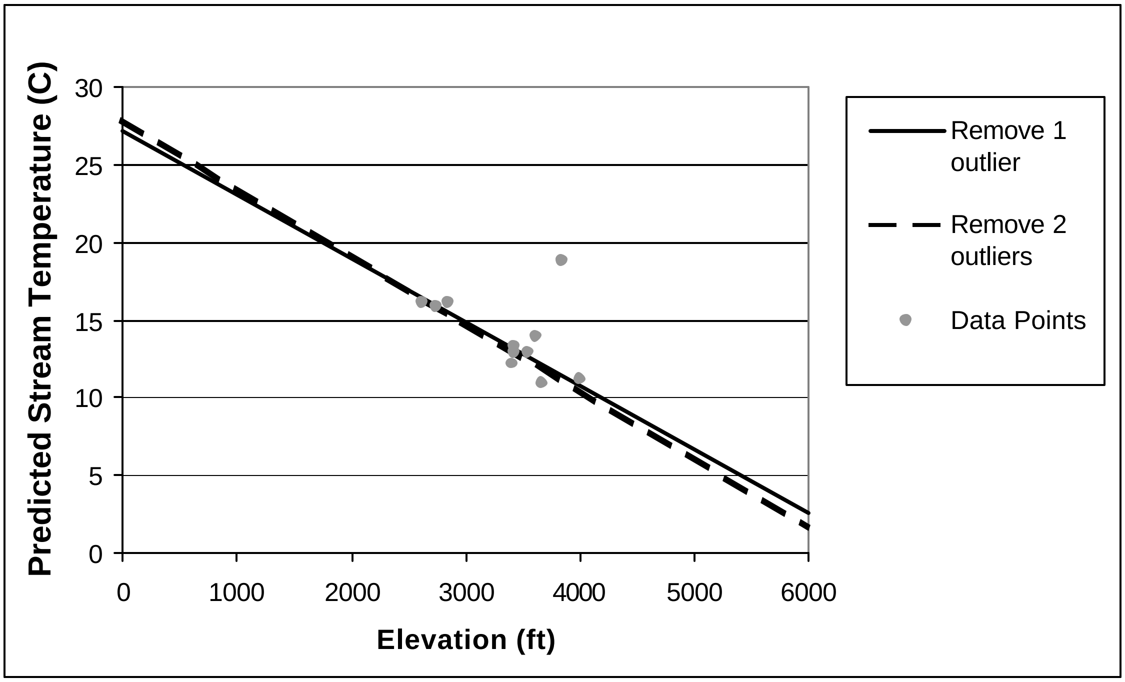


Figure 32. Representation of the regression equations developed for Ecoregion M242B (Western Cascades Section) using only elevation as the independent variable with one and two outliers removed. The gray dots correspond to the actual data used in the analysis. The dots are plotted on actual rather than predicted temperature. Note the narrow range of the data on the elevation scale. The accuracy of the plot beyond the range of the data is unknown.

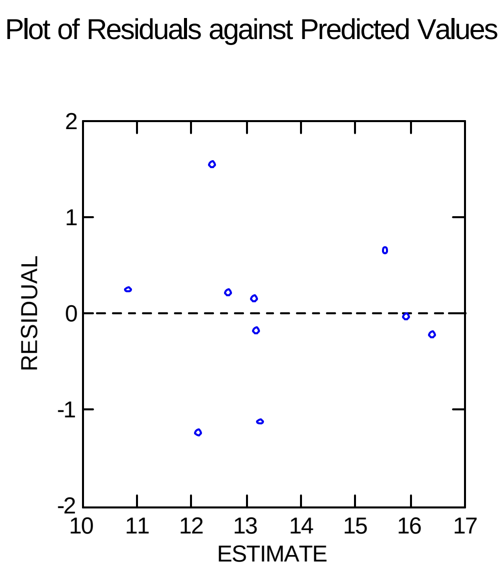


Figure 33. Residuals for the regression model developed for ecoregion M242B (Western Cascades Section). This model includes only elevation. Two outliers were removed from the model. The plot suggests a random distribution of errors around the regression line, although the sample is small. The plot also indicates that the two points at the bottom of the plot may have had a large effect on the outcome.

M242C (Eastern Cascades Section)

The data points in Ecoregion M242C include 7 points that were in the North Cascades region and 113 points that lie in the southern portion of the Cascade Mountains. One outlier was identified in the regression analysis. The models with and without the outlier are both significant at the $p \leq 0.05$ level as are the variables included in the models (Table 21). The removal of the outlier did not have a substantial effect on adjusted squared multiple R square, the model parameters, or the predictive relationships (Figure 34). Therefore, the identified outlier was not removed. The results of the model with the outlier left in the dataset are depicted in Figure 35 along with the scatter of data points used to derive the model. The range of canopy closure and elevation are well represented in the dataset for this ecoregion. The differences between measured and predicted values are represented in Figure 36.

Table 21. Results of regression analyses for Ecoregion M242C (East Cascades Section) with and without the outlier removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	24.587646	22.317537	26.857756	0.00000
Elevation	-0.000995	-0.001715	-0.000275	0.00718
Canopy Closure	-0.081823	-0.10544	-0.058207	0.00000
n				120
Multiple R				0.577902
Squared Multiple R				0.333971
Adjusted squared multiple R				0.322586
One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	24.190182	22.025754	26.354611	0.00000
Elevation	-0.000991	-0.001674	-0.000308	0.00485
Canopy Closure	-0.076329	-0.098928	-0.053731	0.00000
n				119
Multiple R				0.574377
Squared Multiple R				0.329909
Adjusted squared multiple R				0.318356

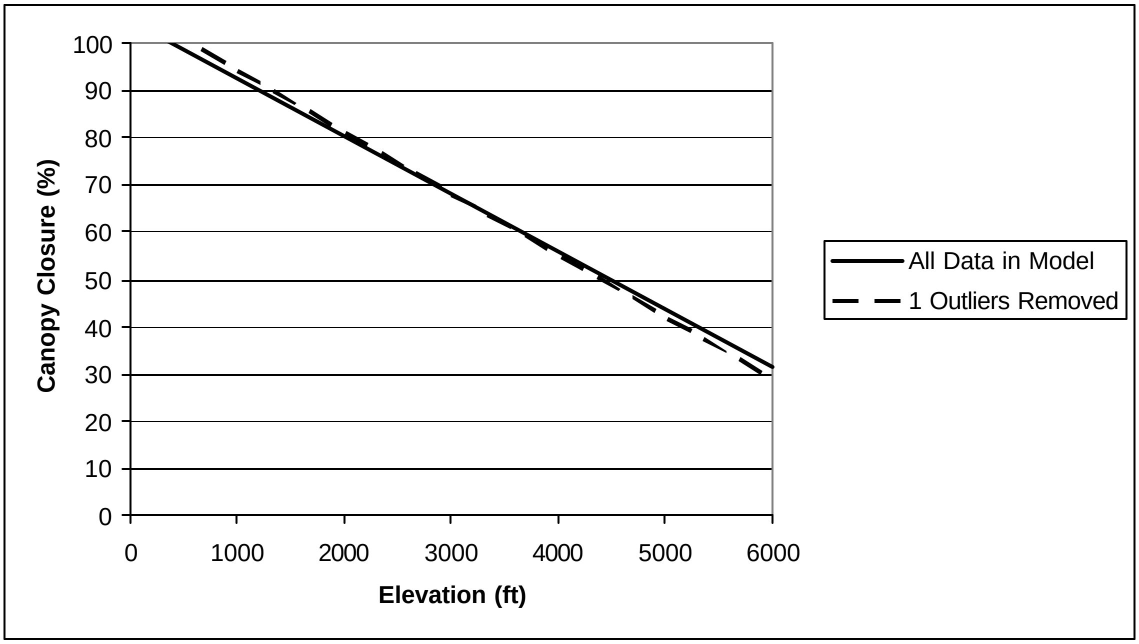


Figure 34. Comparison of two models developed through the regression analyses with and without the identified outlier. The lines represent the predicted canopy closure needed to meet a 16 degree C temperature target.

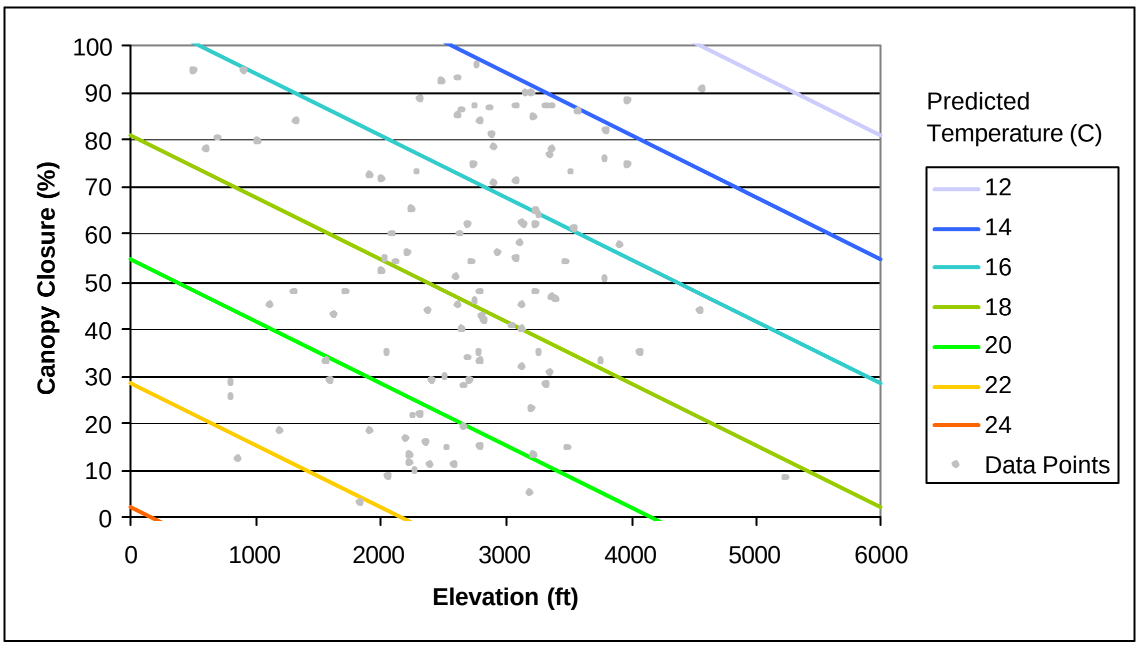


Figure 35. Regression model for the M242C Ecoregion (Eastern Cascades Section) incorporating only canopy closure and elevation as independent variables. No outliers were removed. Each line corresponds to a different target temperature.

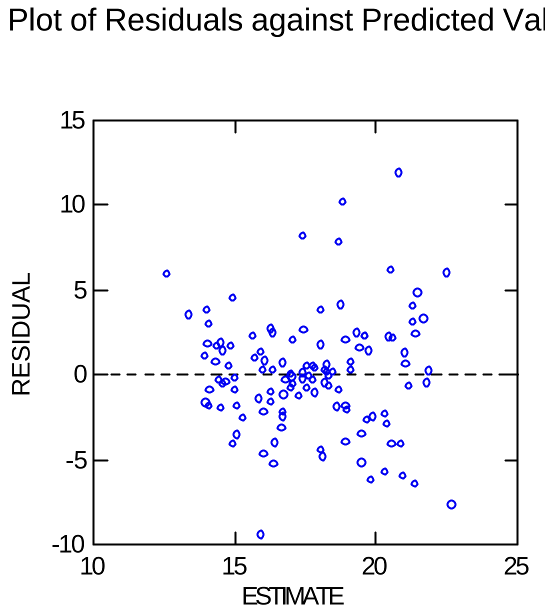


Figure 36. Distribution of residuals for the regression model for ecoregion M242C (Eastern Cascades Section) with no outliers removed. The plot suggests a random distribution of errors around the regression line.

-342I (Columbia Basin Section)

Eight (8) of the data points in Ecoregion -342I (Columbia Basin Section) are in the North Cascades region of eastern Washington and 13 of the data point lie in the southern portion of the eastern Cascade Mountains. No outliers were identified in the regression analysis with both canopy closure and elevation included as independent variables. The 2-tail p level for inclusion of canopy closure in the model was 0.081 (Table 22). Statistically, this parameter should not be included in the model since it does not meet a $p \leq 0.05$ criteria. The 2-tail p level for inclusion of elevation in the model was 0.08. However, the test statistic is close to the 0.05 probability level and therefore, should be given consideration.

Because canopy closure was not statistically significant in the regression model incorporating both elevation and canopy closure as independent variables, a regression analysis was completed using only elevation in the model. Outliers were identified for this model. Additional models were developed eliminating the outliers as they were identified to allow comparison of results and to determine if removal of the outlier(s) was justified. The results of the regression analyses without canopy closure in the model are summarized in Table 23.

Table 22. Results of regression analysis for Ecoregion –342I using canopy closure and elevation as independent variables. Italics indicate variables that were not significant at the p = 0.05 probability level.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000379
Constant	30.419477	25.958168	34.880787	0
Elevation	-0.00284	-0.004864	-0.000815	0.00813
Canopy Closure	*-0.064892*	*-0.138544*	*0.008759*	*0.08126*
n				25
Multiple R				0.715115
Squared Multiple R				0.511389
Adjusted squared multiple R				0.46697

Table 23. Results of regression analyses for Ecoregion –342I using only elevation as an independent variable. Table includes results with all data from the ecoregions included and results with one, two, and three outliers removed.

Canopy Closure Not In Equation				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00032
Constant	29.644234	25.065106	34.223362	0.00000
Elevation	-0.003739	-0.005569	-0.001909	0.00032
n				25
Multiple R				0.661238
Squared Multiple R				0.437235
Adjusted squared multiple R				0.412767
Canopy Closure Not In Equation, One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000008
Constant	32.482164	28.149308	36.81502	0.000000
Elevation	-0.004754	-0.006453	-0.003056	0.000010
n				24
Multiple R				0.777782
Squared Multiple R				0.604944
Adjusted squared multiple R				0.586987
Canopy Closure Not In Equation, Two Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000000
Constant	34.886974	30.771411	39.002537	0.000000
Elevation	-0.005592	-0.007175	-0.004008	0.000000
n				23
Multiple R				0.848289
Squared Multiple R				0.719594
Adjusted squared multiple R				0.706241

Table 23 Continued.

Canopy Closure Not In Equation, Three Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	36.660422	32.877897	40.442946	0
Elevation	-0.006174	-0.007607	-0.004742	0
n				22
Multiple R				0.89531
Squared Multiple R				0.801581
Adjusted squared multiple R				0.79166

Since the test statistic for inclusion of canopy closure was not far from significant, plots depicting the results of the models with and without canopy closure were developed (Figure 37 and Figure 38). The plots of residuals are found in Figure 39. The data points in the models are sparse. The model would likely be improved and possibly substantially modified with the inclusion of additional data from this ecoregion.

For the models with elevation as the only independent variable, the exclusion of the outliers resulted in a substantial change in the adjusted squared multiple r and the predictive relationship (Figure 38). Therefore, use of the model with all the outliers removed is recommended over the models that include one or more of the outliers. Given that the test statistic for inclusion of canopy closure was reasonably low (p = 0.08) and that the goal of this study was to develop nomographs that guided the amount of canopy closure to leave in harvest units, the model incorporating both canopy closure and elevation is recommended.

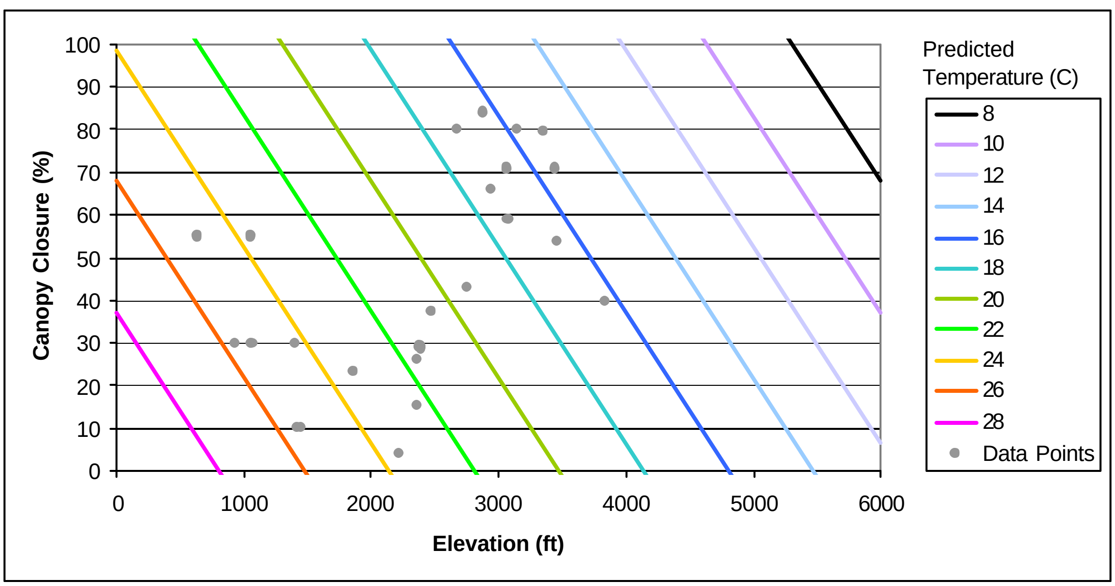


Figure 37. Regression model for the -342I ecoregion (Columbia Basin Section) incorporating only canopy closure and elevation as independent variables. No outliers were removed. Canopy closure was not significant at the p ≤ 0.05 level. The 2-tailed p level for inclusion of canopy closure in the model was 0.08. Each line corresponds to a different target temperature.

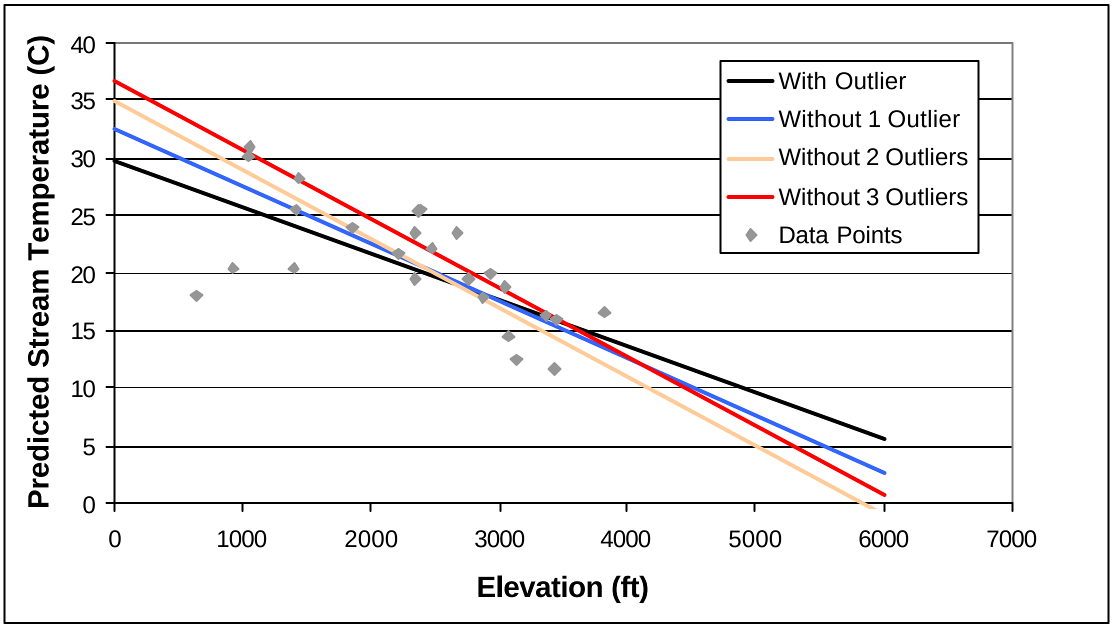


Figure 38. Depiction of predicted stream temperature (C) as a function of elevation only for regression analyses including and excluding the identified outliers for the –342I ecoregion. The gray points correspond to the input data for the analysis. The dots are plotted on actual rather than predicted temperature. The data points at around 17 to 20 degrees and 800 to 1500 feet elevation are the identified outliers.

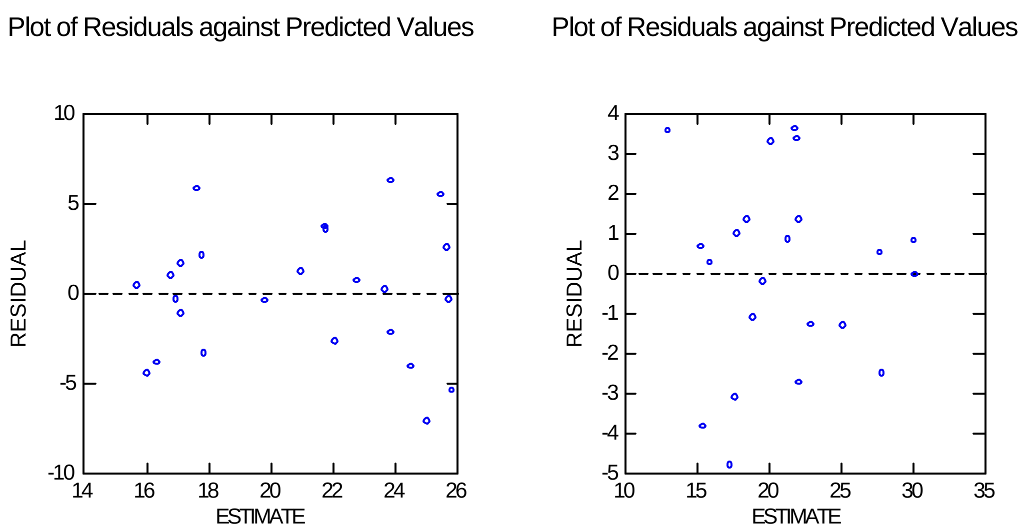


Figure 39. Residual plots for the regression analyses conducted for ecoregion -342I (Columbia Basin Section). The left plot shows residuals for the model that includes both canopy closure and elevation as independent variables. The right plot depicts residuals for the model with elevation included as the only independent variable with three outliers removed. The plots suggest a random distribution of errors around the regression lines.

-331A (Palouse Prairie Section)

There was only one data point from this ecoregion. This data point is located in the Blue Mountains. No regression was developed.

M332G (Blue Mountain Section)

There are six data points in this ecoregion. All are located in the Blue Mountains. This data set includes all but one of the Blue Mountain region data. The number of data points is very low. The results presented earlier for the Blue Mountain region are applicable for this ecoregion. Hence, no additional analysis was conducted for ecoregion M332G.

Tests of Significant Differences between Equations

Tests of significant difference between equations were developed for analyses discussed above incorporating both canopy closure and elevation as independent variables. It should be noted, however, that many of the ecoregions did not have equations where both the independent variables were significant and analyses could not be completed for two of the ecoregions. In addition, the data representing ecoregion M242B (Western Cascades Section) was poorly distributed, resulting in low confidence in the regression results. Ecoregions M333A (Okanogan Highlands Section) and M242C (Eastern Cascades Section) were the only ecoregions where both independent variables were significant and confidence in the models was reasonably high. The analyses for Ecoregion –342I (Columbia Basin Section) was reasonably good; canopy closure, however, was significant at the 0.08 probability level and not at 0.05.

The plots of temperature and canopy closure combinations that are predicted to result in 16 and 18 degree temperatures appear to be very scattered and inconsistent (Figures 40 and 41). Much of the variation between lines is due to small sample sizes, poor representation of the range of independent samples in the data set, and/or independent variables that did not test as significant variables in the prediction of stream temperature.

The two lines representing ecoregions where the analyses were robust (M333A and M242C) are substantially different (Figures 40 and 41). The first represents primarily north central and northeastern Washington and the second represents the east flanks of the Cascades Mountains. The lines for these two ecoregions would indicate that the east flanks of the Cascade Mountains require less canopy closure to meet a temperature target at low elevations than the region in north and northeast Washington. Conversely, greater canopy closure is apparently needed on the east flanks of the Cascade Mountains at higher elevation. Weather patterns and air temperature changes substantially with elevation on the east flanks of the Cascade Mountains. If these changes are more pronounced than in north and northwest Washington, the differences in the position of the two lines may be logical. The equation for Ecoregion 0342I (Columbia Basin Section) is similar to that developed for M333A (East Cascades Section) (Figures 40 and 41).

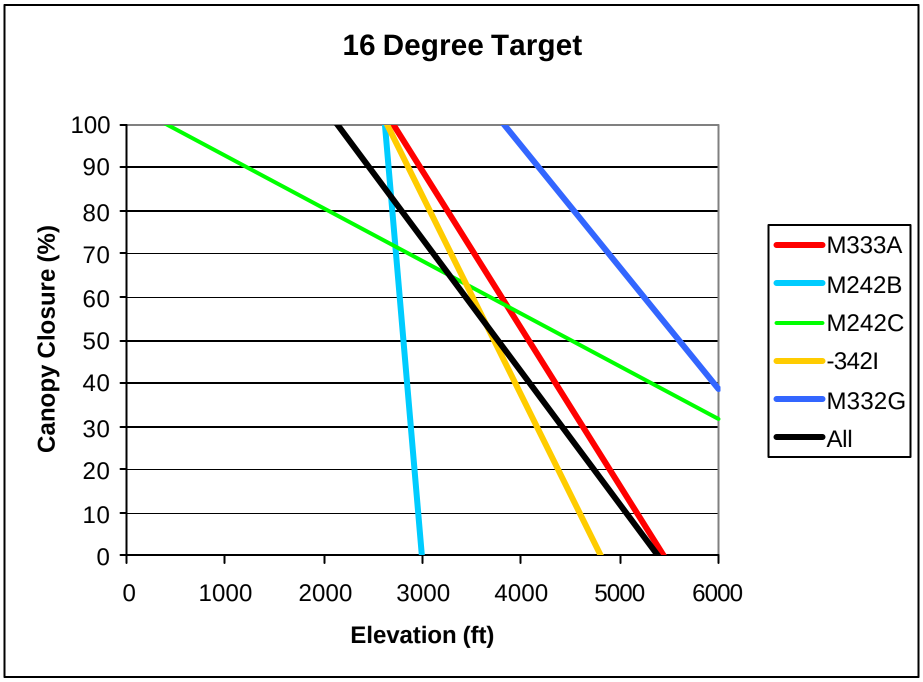


Figure 40. Plots of the combinations of canopy closure and elevation by ecoregion that are predicted to result in a 16 degree C stream temperature. Note, Ecoregions M333A and M242C were the only ecoregions where both independent variables were significant and confidence in the models was reasonably high. Ecoregion –342I came close to meeting the 0.05 probability level.

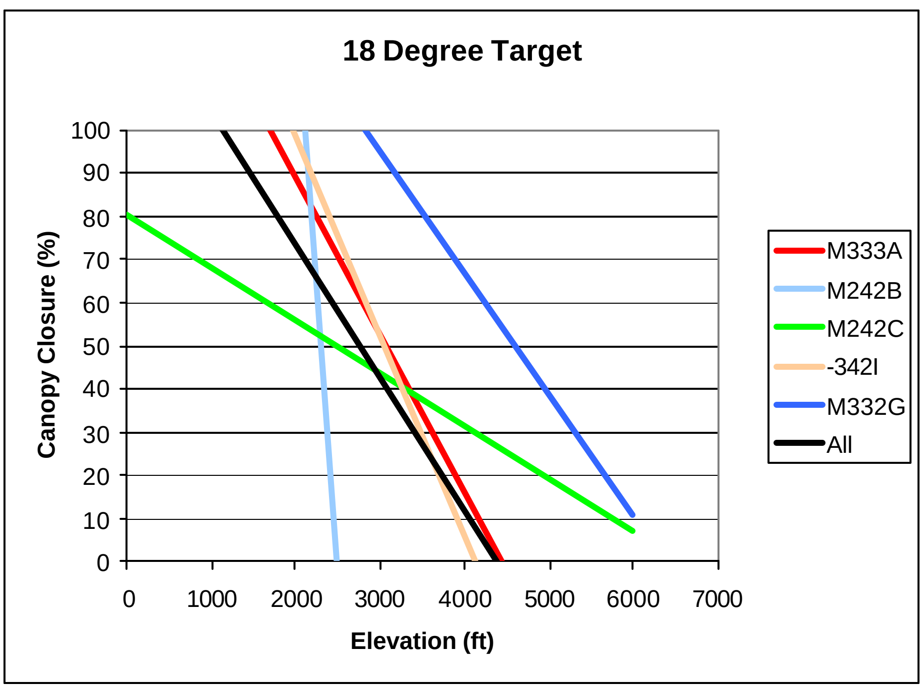


Figure 41. Plots of the combinations of canopy closure and elevation by ecoregion that are predicted to result in a 18 degree C stream temperature. Note, Ecoregions M333A and M242C were the only ecoregions where both independent variables were significant and confidence in the models was reasonably high. Ecoregion –342I came close to meeting the 0.05 probability level.

The equations generated for each ecoregion elevation and canopy closure as independent variables were tested to determine if the populations of the various combinations of datasets were from one population (Zar 1974). M332G (Blue Mountains Section) was excluded from this test because the data represented the region rather than the ecoregion.

The test to determine if all ecoregion datasets are from the same population indicated that they are not all the same ($p<0.05$). Data representing pairs of ecoregions were then tested to see if they were significantly different. Interestingly, M333A (Okanogan Highlands Section) and M242B (Western Cascades Section) were not significantly different (Table 24). These were the only two ecoregions where confidence in the regression results was reasonably high. All other pairs of populations were found to be significantly different from each other at the 0.05 probability level (Table 24). Given that ecoregions M333A and M242B are not statistically different, and given the difficulties with the datasets in other ecoregions, the use of ecoregions as a stratification parameter is not recommended.

Table 24. F statistics and estimated probability to reject the null hypothesis that two populations are the same (in parentheses) for pairs of datasets representing the various regions.

ECOREGION	M242B	M242C	-342I
M333A	2.086 (.05<p<.10)	3.960 (p<0.0025)	12.099 (p<0.0005)
M242B		4.356 (p<0.0025)	20.968 (P<0.0005)
M242C			9.911 (P<0.0005)

3.4.3 Stratified by Lithology Group

Regression analyses were conducted with the data stratified by lithology. No lithology was available for 57 of the data points. Most of the data points with these missing values were from the area of the southern portion of the Cascade Mountains.

There are 33 lithologies represented in the dataset. Sample sizes were very low for most of these (Table 25). Due to the small sample sizes, the lithologies were combined into seven lithology groups made up of similar lithologic types (Table 26). Regression analyses were run for each of these lithology groups.

Alluvium

There are 74 data points in the dataset representing the alluvium lithology group. Five of these data points are from the Blue Mountain area, 7 are from the North Cascades area, 54 are from the southern portion of the Cascade Mountains, and 8 are from northeast Washington.

Table 25. Sample sizes by lithology from the WDNR GIS database.

Lithology	Sample Size
2 mica granite	7
acidic intrusive rocks	2
alluvium	73
alluvium, older	1
alpine glacial drift, Fraser-age	5
alpine glacial till, pre-Fraser	1
andesite flows	6
argillite	2
banded gneiss	2
basalt flows	38
continental glacial drift, Fraser-age	37
continental glacial outwash, Fraser-age	3
continental glacial till, Fraser-age	2
continental sedimentary deposits or rocks	1
continental sedimentary deposits or rocks, conglomerate	1
gabbro	1
glaciolacustrine deposits, Fraser-age	7
granodiorite	1
heterogeneous metamorphic rocks, chert bearing	4
intermediate intrusive rocks	1
intrusive andesite and dacite	1
marine sedimentary rocks	1
mass wasting deposits	2
mass-wasting deposits	5
mass-wasting deposits, mostly landslides	2
metacarbonate	2
metavolcanic rocks	2
orthogneiss	1
outburst flood deposits, sand and silt, late Wisconsin	1
phyllite low grade	1
quartzite low grade	4
schist low grade	1
sedimentary rocks or deposits	11
tonalite	3
volcaniclastic deposits or rocks	16
Grand Total	**248**

Table 26. Lithologies included in each of the lithology groups.

Condensed Lithology Group						
Alluvium (n=74)	**Basalt (n=47)**	**Glacial Drift (n=48)**	**Granites (n=13)**	**Metamorphic (n=19)**	**Other Deposits (n=33)**	**Sedimentary (n=14)**
• Alluvium • Alluvium, older	• Andesite Flows • Basalt Flows • Gabbro • Intermediate Intrusive Rocks • Intrusive Andesite and Dacite	• Alpine glacial drift, Fraser-age • Alpine glacial till, pre-Frasier • Continental glacial drift, Fraser-age • Continental glacial outwash, Fraser-age • Continental glacial till, Fraser-age	• 2 mica granite • Ccidic intrusive rocks • Granodiorite • Tonalite	• Argillite • Banded gneiss • Heterogeneous metamorphic rocks chert bearing • Metavolcanic rocks • Orthogneiss • Phyllite low grade • Quartzite low grade • Schist low grade	• Glaciolacustrine deposits Fraser-age • Mass-wasting deposits • Mass-wasting deposits, mostly landslides • Outburst flood deposits, sand and silt, late Wisconsin • Volcaniclastic deposits or rocks	• Continental sedimentary deposits or rocks • Continental sedimentary deposits or rocks, conglomerate • Marine sedimentary rocks • Sedimentary rocks or deposits

No outliers were identified for this group in the regression analysis. The results of the regression analysis is summarized in Table 27 and depicted in Figure 42. The data set is well distributed across elevation and canopy closure for all elevations less than 4000 feet (Figure 42). The residuals for the analysis are found in Figure 43.

The residuals of the analysis are somewhat skewed, but no transformation was performed since normality is apparent in most of the other regression analysis conducted with this nomograph data set.

Table 27. Results of regression analyses for the dataset within the alluvium lithology group.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000006
Constant	26.677486	23.809062	29.545909	0.00000
Elevation	-0.002149	-0.003282	-0.001016	0.00032
Canopy Closure	-0.051573	-0.094111	-0.009036	0.01820
n				74
Multiple R				0.536731
Squared Multiple R				0.288081
Adjusted squared multiple R				0.268026

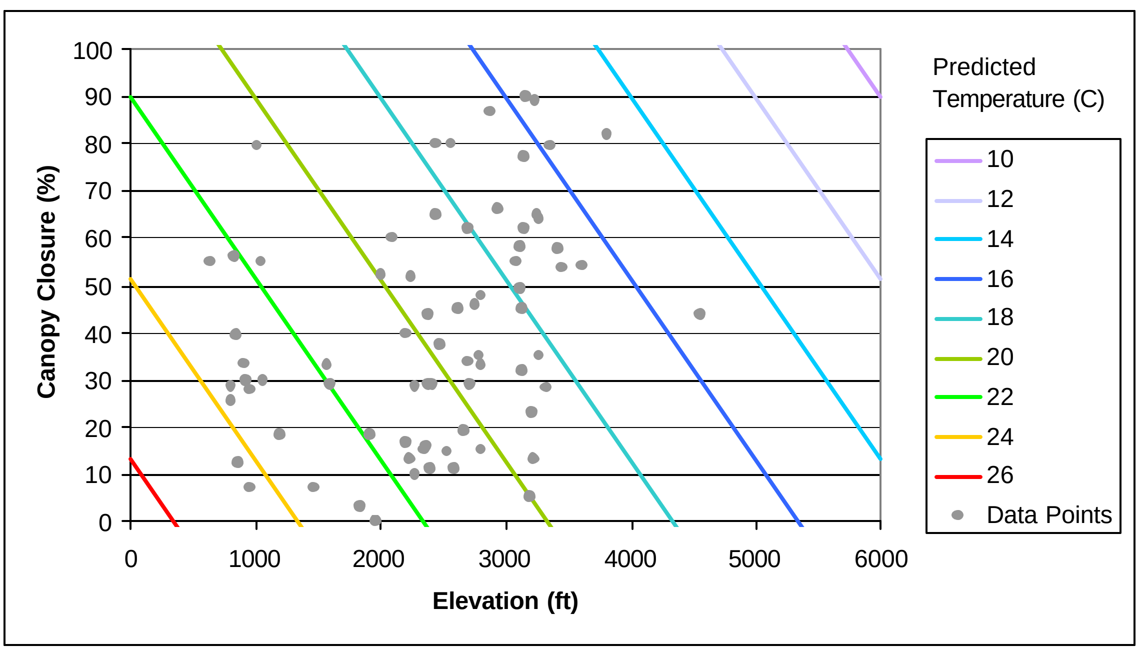


Figure 42. Regression model for the alluvium lithology group incorporating only canopy closure and elevation as independent variables. No outliers were removed. Each line corresponds to a different target temperature.

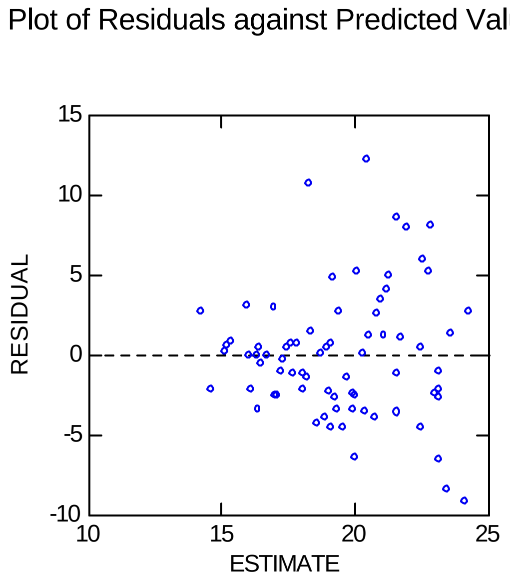


Figure 43. Residual plots for the regression analyses conducted for the alluvium lithology group. The residuals are somewhat skewed (tighter on the left and wider on the right of the graph), suggesting a possible need for a transformation. No transformation was performed since normality is apparent in most of the other regression analyses conducted with this nomograph data set.

Basalt

There are 47 data points in the database representing the basalt lithology group. Two of these data points were located in the Blue Mountains, three were located in the North Cascades region, and the rest were located in the southern portion of the Cascade Mountains.

No outliers were identified for this group in the regression analysis. The results of the regression analysis is summarized in Table 28 and depicted in Figure 44. The data set is reasonably well distributed although low canopy closure at high elevation is under-represented as is all elevations below 1000 feet (Figure 44). The residuals for the analysis are found in Figure 45.

Glacial Drift

There are 48 data points in the dataset representing the glacial drift lithology group. Five of these data points are from the North Cascades area, seven are from the southern portion of the Cascade Mountains, and 36 are from northeast Washington.

One outlier was identified for this group in the regression analysis. Regression analyses were conducted with the outlier removed as well as for the situation with the outlier in the model. The results of the regression analysis are summarized in Table 29.

Table 28. Results of regression analysis for the basalt lithology. No outliers were identified.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	28.23372	25.410705	31.056735	0.00000
Elevation	-0.001985	-0.002924	-0.001047	0.00011
Canopy Closure	-0.090752	-0.11965	-0.061854	0.00000
n				47
Multiple R				0.790884
Squared Multiple R				0.625497
Adjusted squared multiple R				0.608474

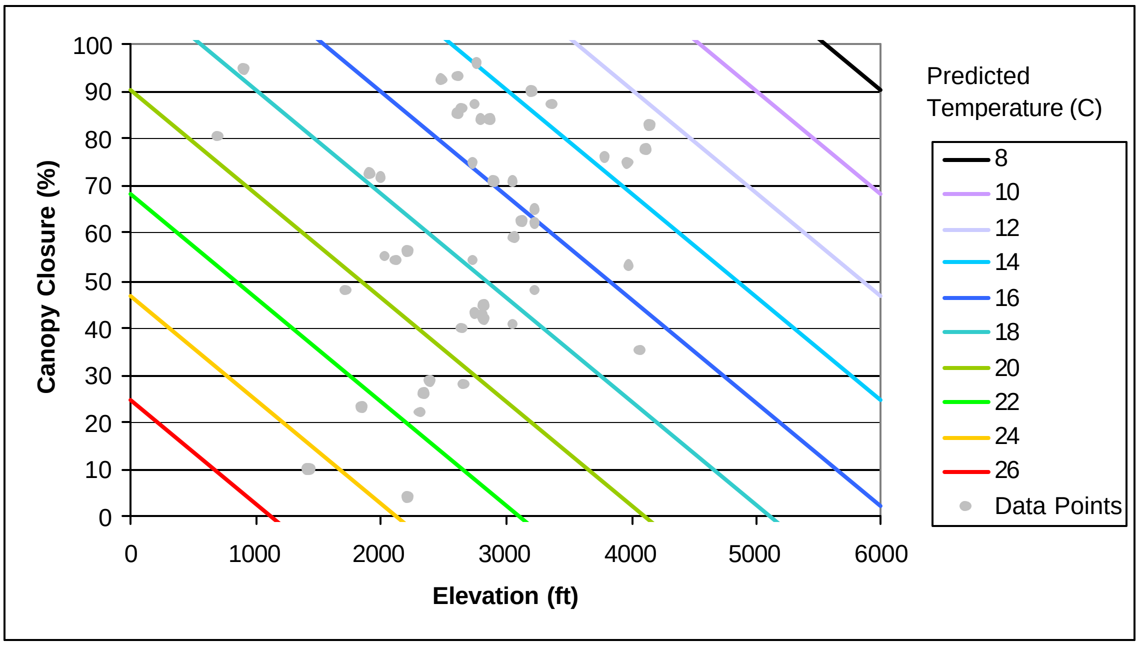


Figure 44. Regression model for the basalt lithology group incorporating only canopy closure and elevation as independent variables. No outliers were removed. Each line corresponds to a different target temperature.

Plot of Residuals against Predicted Values

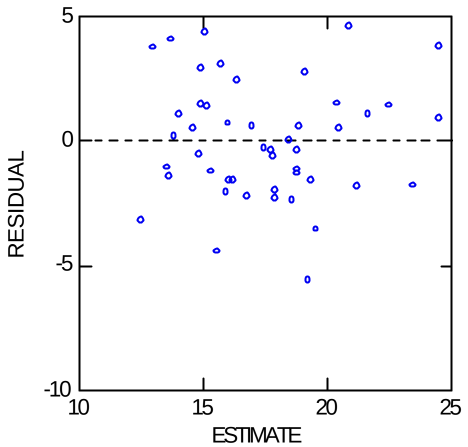


Figure 45. Residual plots for the regression analyses conducted for the basalt lithology group. The residuals indicate a relatively random distribution of errors around the regression line.

Table 29. Results of regression analyses for the dataset within the glacial drift lithology group with and without the identified outlier removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000001
Constant	26.380208	22.119725	30.640691	0.000000
Elevation	-0.001808	-0.003062	-0.000554	0.005710
Canopy Closure	-0.089862	-0.121827	-0.057896	0.000000
n				48
Multiple R				0.68741
Squared Multiple R				0.472533
Adjusted squared multiple R				0.44909
One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	28.380374	24.664646	32.096101	0.00000
Elevation	-0.003013	-0.00421	-0.001817	0.00001
Canopy Closure	-0.071546	-0.099841	-0.04325	0.00001
n				47
Multiple R				0.758841
Squared Multiple R				0.57584
Adjusted squared multiple R				0.55656

The canopy closure required to meet a temperature standard of 18 degrees C was estimated for both regression models for the glacial drift lithology group and plotted to evaluate the effect that removing the outlier had on the resulting equation (Figure 46). Both the adjusted squared multiple r and the model results changed substantially with the removal of the outlier. Therefore, the model with the outlier removed is the better model.

The predicted canopy closure required to meet a range of stream temperatures based on the regression model without the outlier is depicted in Figure 47. The scatter of the data points used to develop the model is also shown on this figure. The majority of the data points are from situations with higher canopy closure and elevation (Figure 47). Lower canopy closure and lower elevations are under-represented. Additional data representing these situations may affect the model results. The residuals for the analysis are found in Figure 48.

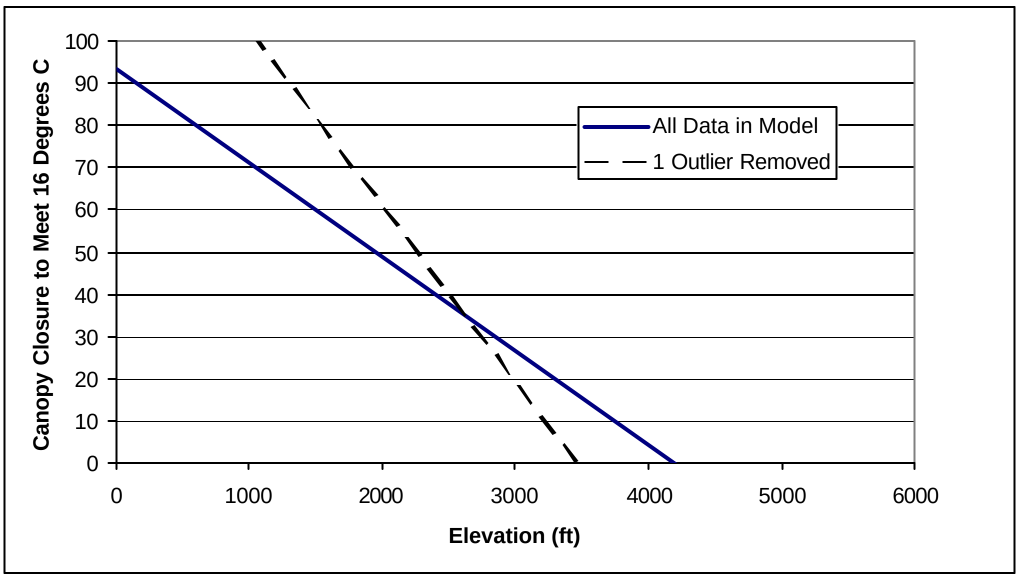


Figure 46. Comparison of the estimated canopy closure required to meet an 18 C degree temperature standard based on the regression models for the glacial drift dataset with and without the identified outlier.

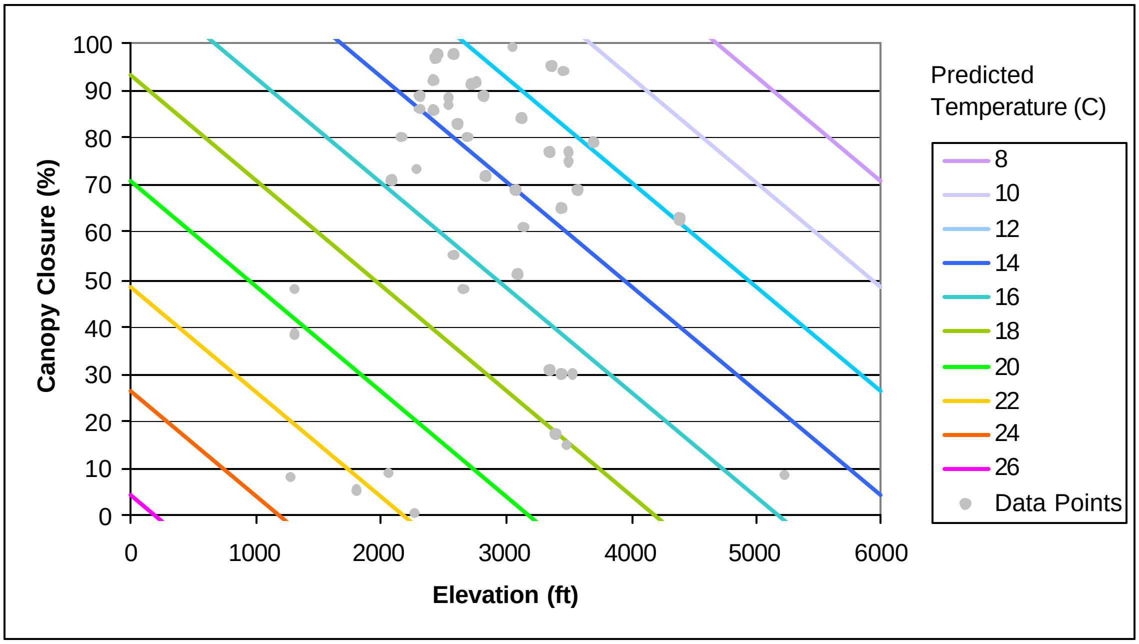


Figure 47. Regression model for the glacial drift lithology group incorporating only canopy closure and elevation as independent variables. One outlier was removed. Each line corresponds to a different target temperature.

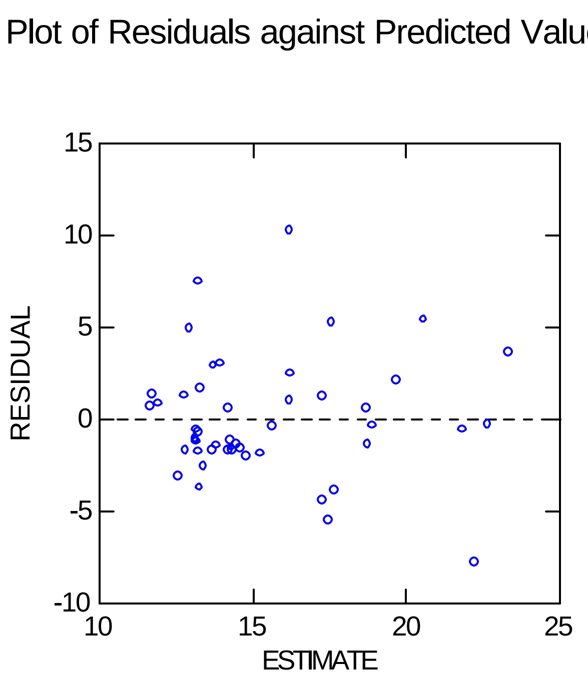


Figure 48. Residual plots for the regression analyses conducted for the glacial drift lithology group. The residuals indicate a relatively random distribution of errors around the regression line.

Granites

There are 13 data points representing the granite lithology group. Three of these come from the southern portion of the eastern Cascade Mountains. The balance is located in northeast Washington.

The regression analyses using canopy closure and elevation as independent variables with all data points from the lithology group found no significant relationships (Table 30). The t-tailed test probability values for inclusion of canopy closure and elevation into the model were 0.657 and 0.439, respectively. Neither of these values is near the p value of 0.05 used to determine whether a variable should be included in the model.

One outlier was identified in the initial analysis. When this outlier was removed, a statistically significant ($P<0.05$) equation was developed. In this second analysis, canopy closure was found to be a significant variable, but elevation was not. An outlier was also identified in the second analysis of this dataset. When the second outlier was removed, the overall equation was still significant, canopy closure remained in the equation, and elevation approached the probability level of 0.05 ($p = 0.06467$ for elevation). Once again, the regression identified an outlier. When this last outlier was removed, the equation was significant and elevation very nearly met the probability level of 0.05 for inclusion in the model ($p = 0.051670$).

The equations that were developed for each of these regressions are depicted in Figure 49. Note that the three lines that were statistically significant (those with outliers removed) all indicate a need for increasing canopy closure with increasing elevation. This does not make physical sense. All equations were therefore assumed spurious. Hence, no model for the granite lithology was given further consideration.

Table 30. Results of regression analyses for the granite lithology group with all data in the analysis and with various numbers of outliers removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.639624
Constant	20.953752	12.867537	29.039968	0.000180
Elevation	-0.000871	-0.00328	0.001537	0.438790
Canopy Closure	-0.014531	-0.085256	0.056194	0.656900
n				13
Multiple R				0.292399
Squared Multiple R				0.085497
Adjusted squared multiple R				0.000000

Table 30 Continued.

One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.02299
Constant	23.459585	18.186454	28.732717	0.00000
Elevation	0.00037	-0.001292	0.002032	0.62673
Canopy Closure	-0.091137	-0.152362	-0.029912	0.00829
n				12
Multiple R				0.753385
Squared Multiple R				0.56759
Adjusted squared multiple R				0.471498
Two Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.003414
Constant	23.358037	19.09544	27.620634	0.00000
Elevation	0.001665	-0.000128	0.003458	0.06467
Canopy Closure	-0.132872	-0.195435	-0.07031	0.00120
n				11
Multiple R				0.870787
Squared Multiple R				0.758271
Adjusted squared multiple R				0.697838
Three Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000355
Constant	25.551033	22.082012	29.020054	0.000000
Elevation	0.001296	-0.000012	0.002604	0.051670
Canopy Closure	-0.144806	-0.190306	-0.099306	0.000130
n				10
Multiple R				0.946928
Squared Multiple R				0.896673
Adjusted squared multiple R				0.556560

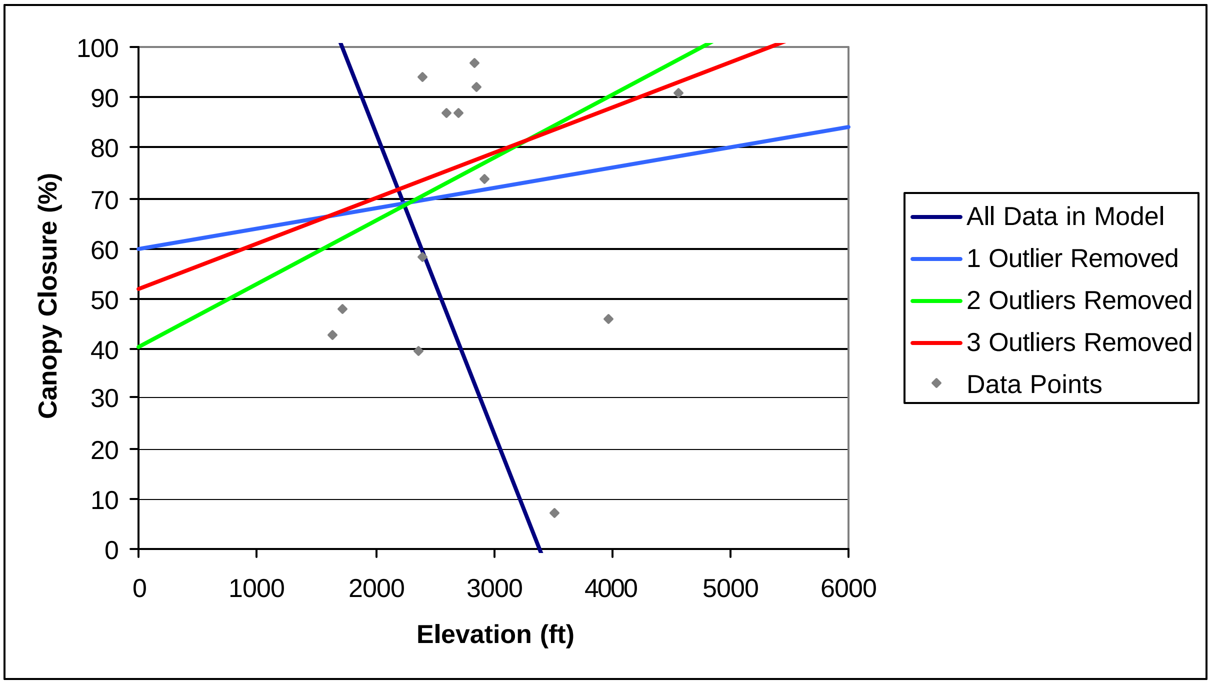


Figure 49. Depiction of the four regression equations developed using the dataset for the granite lithology group. The data points are depicted as dots on the graph. Only the lines with outliers removed were statistically significant; however all lines with outliers removed indicate an increasing requirement for canopy closure with increasing elevation. This trend does not make physical sense.

Metamorphic

There are 19 data points representing the metamorphic lithology group. Nine of these are points located in the southern portion of the eastern Cascade Mountains and the balance is located in northeast Washington.

The regression analyses using canopy closure and elevation as independent variables were completed. The analyses identified four outliers in the dataset. Estimates of canopy closure needed to meet an 18 degree C temperature standard were developed for each of four models with zero, one, three, and four outliers removed. The parameters associated with each of these four models are summarized in Table 31.

The model developed with all the data points for the metamorphic lithology group was not significant (Table 31). The first analysis of this data identified one outlier. With this outlier removed, the regression was significant at the $p = 0.05$ probability level; elevation, however, did not significantly contribute to the equation. This second analysis identified two more outliers. These were removed for the third regression analysis of the dataset and one more outlier identified in the third regression analysis was removed in the fourth analysis. The significance of the equation became progressively higher with each outlier removed (Table 31). The probability that elevation contributes significantly to the equation approached the 0.05 probability level with the four outliers removed (p for elevation with 4 outliers removed was 0.0782).

Table 31. Results of regression analyses conducted with the database for the metamorphic lithology group. Table shows results of the analysis with all the data points used and with outliers removed as they were identified.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.317774
Constant	21.914623	11.145236	32.68401	0.000540
Elevation	0.000143	-0.002848	0.003134	0.920480
Canopy Closure	-0.082819	-0.197831	0.032193	0.146400
n				19
Multiple R				0.365387
Squared Multiple R				0.133507
Adjusted squared multiple R				0.025196
One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.018539
Constant	28.380374	24.664646	32.096101	0.000000
Elevation	-0.003013	-0.00421	-0.001817	0.303240
Canopy Closure	-0.071546	-0.099841	-0.04325	0.015620
n				18
Multiple R				0.642188
Squared Multiple R				0.412405
Adjusted squared multiple R				0.334059
Three Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00215
Constant	26.899328	21.121897	32.676759	0.00000
Elevation	-0.000753	-0.002635	0.001129	0.40312
Canopy Closure	-0.122499	-0.210279	-0.034718	0.00995
n				16
Multiple R				0.781883
Squared Multiple R				0.611341
Adjusted squared multiple R				0.551548
Four Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000291
Constant	28.033136	23.072885	32.993386	0.000000
Elevation	-0.001504	-0.003212	0.000204	0.079200
Canopy Closure	-0.113141	-0.187506	-0.038776	0.00617
n				15
Multiple R				0.861704
Squared Multiple R				0.742534
Adjusted squared multiple R				0.699623

The model parameters change dramatically with each outlier removed (Figure 50). These changes in the models with each removal of an outlier suggest that any model based on the existing dataset, with or without the outliers included, may be spurious. In most situations, removal of outliers simply refines the accuracy of the model. In this case, the model is clearly unstable and highly affected by individual data points. Therefore, we have little confidence in the model developed for the metamorphic lithology group. An increase in sample size may help to stabilize the model and result in a significant relationship between stream temperature and the independent variables.

Because elevation was not significant in any of the regression analyses conducted using both elevation and canopy closure, a regression analysis without elevation was conducted. One outlier was identified in the first analysis conducted. When this outlier was removed, two more outliers were identified. A third analysis was conducted without these last two outliers. The results of the analyses are summarized in Table 32.

The regression without the outliers removed was not significant at the 0.05 probability level (Table 32). The regression lines changed substantially with each removal of identified outliers (Figure 51). The overall significance of the equations and the squared multiple R increased with each removal of the outliers. Hence, the last equation developed is likely the best of the various equations. A plot of the residuals for that equation is provided in Figure 52.Confidence in the equation is low. Additional samples from the metamorphic lithology group would likely result in substantial changes in the regression results.

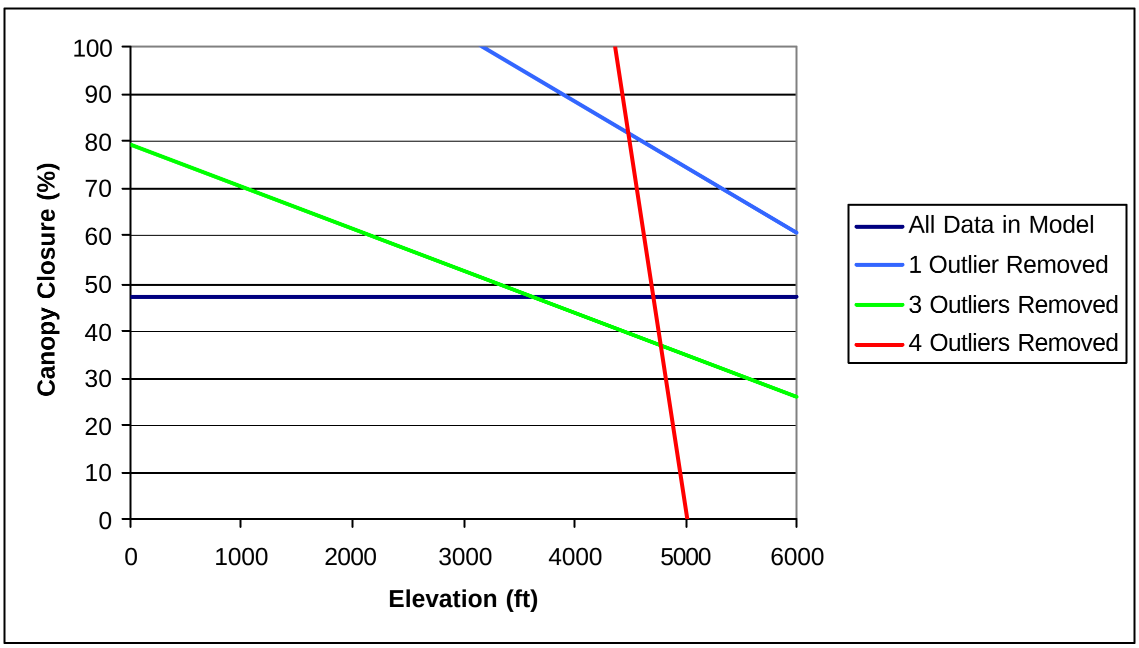


Figure 50. Depiction of the estimated canopy closure needed to meet and 18 degree C temperature standard for each of four models for the metamorphic lithology group.

Table 32. Results of the regression analyses for the metamorphic lithology group using only canopy closure as the independent variable with and without the removal of identified outliers.

Canopy Closure Only in Equation				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.124818
Constant	22.225313	13.92772	30.522906	0.00003
Canopy Closure	-0.081189	-0.187283	0.024906	0.12482
n				19
Multiple R				0.364624
Squared Multiple R				0.13295
Adjusted squared multiple R				0.081947
Canopy Closure Only in Equation, One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.007615
Constant	24.206905	17.973928	30.439883	0.000000
Canopy Closure	-0.116265	-0.197035	-0.035495	0.007610
n				18
Multiple R				0.606533
Squared Multiple R				0.367883
Adjusted squared multiple R				0.328375
Canopy Closure Only in Equation, Three Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000519
Constant	26.393375	20.847556	31.939194	0.000000
Canopy Closure	-0.143715	-0.212528	-0.074901	0.000520
n				16
Multiple R				0.767471
Squared Multiple R				0.589011
Adjusted squared multiple R				0.559655

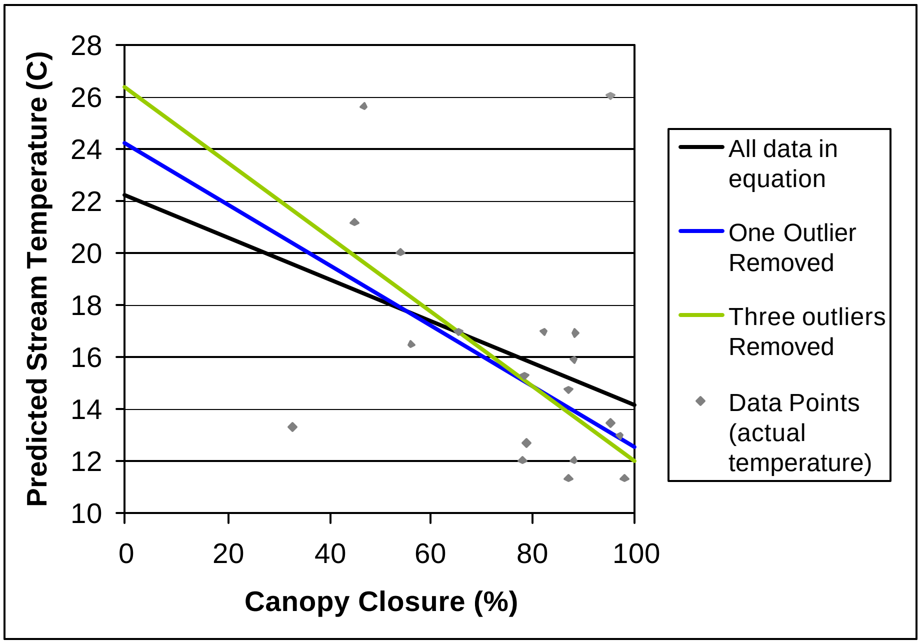


Figure 51. Depiction of the equations developed through the regression analyses of the metamorphic lithology dataset using only canopy closure as the independent variable. Equations represent analysis results with all data for the lithology in the equation and with outliers removed. Note, the equation with all the data included was not significant at the 0.05 probability level.

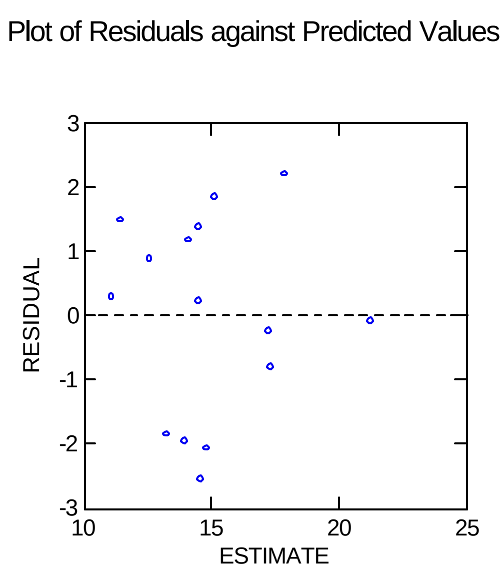


Figure 52. Plot of the residuals of the regression equation for the metamorphic lithology group using canopy closure as the independent variable with three outliers removed. No transformation is suggested by the residuals and the errors appear to be well balanced.

Other Deposits

There are 33 cases in the lithology group termed "other deposits". Two of these cases are located in the North Cascades region of eastern Washington, 15 are located in eastern Washington, and 16 are located in the southern portion of the Cascade Mountains. This group included several deposits. The similarity between the lithologies included in this group is that all are or were unconsolidated deposits of material resulting from glacial dam breaks, landslides, and other events that deposit large volumes of material.

In the regression analysis for this group, canopy closure was not identified as a significant variable for inclusion in the model (Table 33). The 2-tail probability value for the inclusion of canopy closure was 0.125. The best-fit model with both elevation and canopy closure included as independent variables is depicted in Figure 53.

Table 33. Results of regression analyses conducted with the database for the various other deposits. Canopy closure was not significant at the 0.05 probability level (indicated in italics). The results of the analysis without canopy closure as an independent variable are also shown in the table.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.014158
Constant	24.617141	18.811549	30.422733	0
Elevation	-0.002236	-0.003735	-0.000737	0.00494
Canopy Closure	*-0.038784*	*-0.089081*	*0.011513*	*0.12526*
n				30
Multiple R				0.520078
Squared Multiple R				0.270482
Adjusted squared multiple R				0.216443
Canopy Closure Not In Equation				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.01251
Constant	21.412268	17.25841	25.566126	0.00000
Elevation	-0.00193	-0.003411	-0.000449	0.01251
n				30
Multiple R				0.450386
Squared Multiple R				0.202847
Adjusted squared multiple R				0.174378

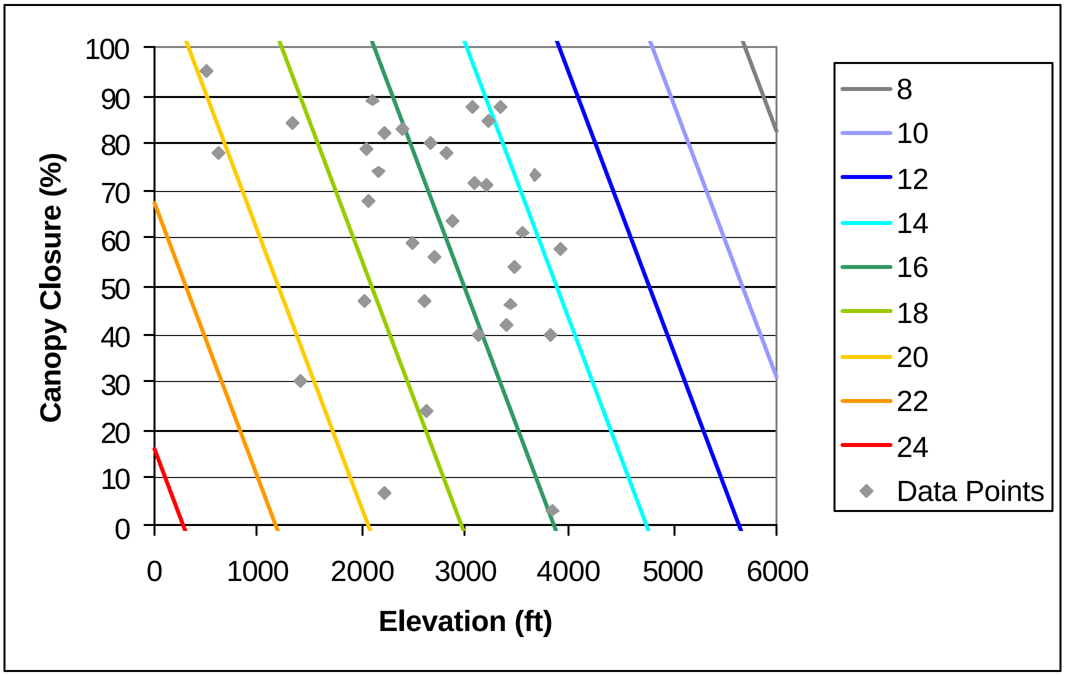


Figure 53. Depiction of the regression line developed for the metamorphic lithology group using canopy closure and elevation as independent variables with three outliers removed. Note, canopy closure in the best-fit equation was not significant at the 0.05 probability level. The actual data points are depicted as gray dots on the plots. Each of the colored lines corresponds to the temperature indicated in the legend.

Additional regression analyses were conducted for this lithology group using only elevation as the independent variable (Table 33, Figure 54). This equation was significant, but does not function as an aid in the identification of the amount of canopy closure needed to attain target temperatures.

Situations at low elevation with low shade were not represented in the database. Additional data covering this missing sector would likely affect the regression results.

The residuals depicting the scatter of the predicted data from actual observations are provided for the regression analysis incorporating both canopy closure and elevation as independent variables and for the regression analysis that did not include canopy closure are depicted in Figure 55. The residuals look good in both plots.

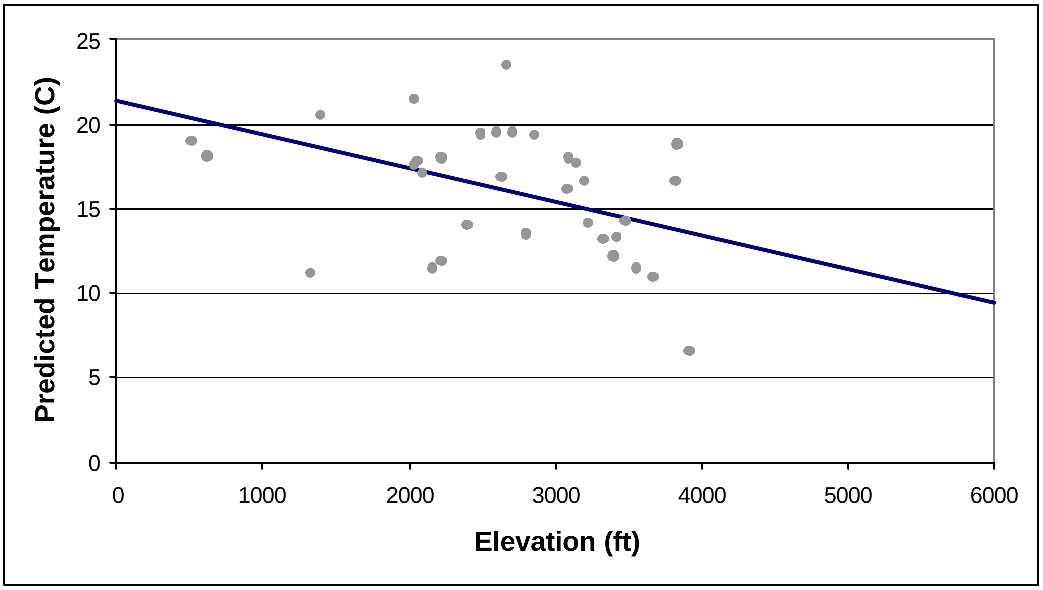
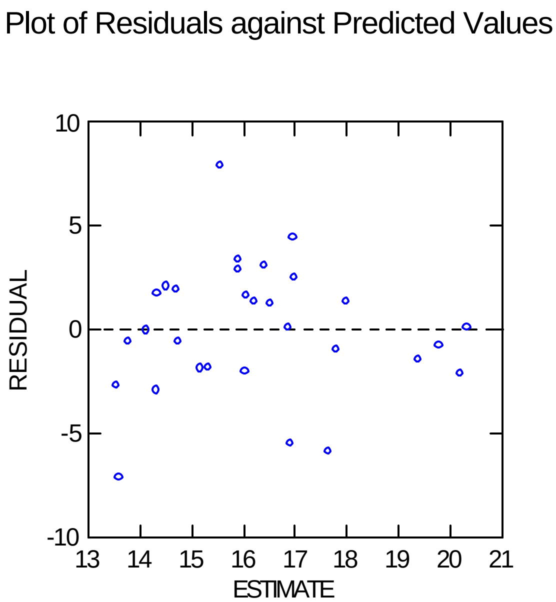


Figure 54. Depiction of the relationship between elevation and predicted stream temperature based on the regression analysis conducted for the lithology group termed "other deposits". Dots depict actual temperature of sample points used in the model.



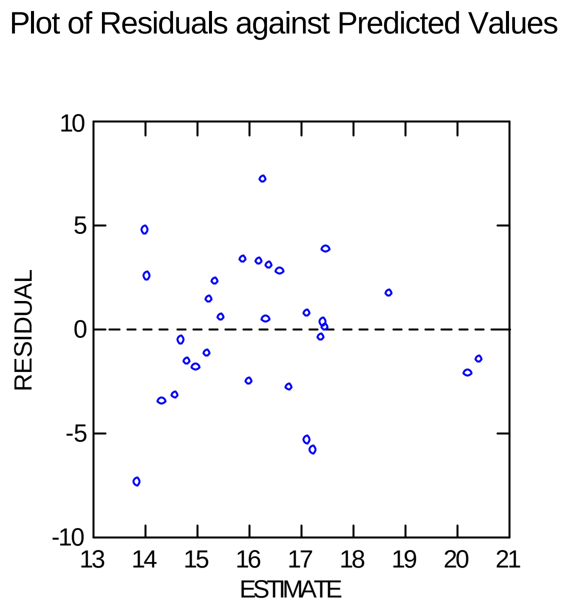


Figure 55. Residuals of the regression models developed for the other deposits lithology group. The plot on the left depicts the residuals of the best-fit model incorporating both canopy closure and elevation as independent variables. The plot on the right depicts the residuals of the best-fit model that included only elevation as an independent variable. The residuals indicate a relatively random distribution of errors around the regression line and do not suggest the need for a transformation.

Sedimentary

There are 14 cases in the sedimentary lithology group. All of these data points come from locations in the southern portion of the Cascade Mountains.

In the regression analysis for this group, elevation was not identified as a significant variable for inclusion in the model (Table 34). The 2-tail probability value for the inclusion of elevation was 0.199. The model with both independent variables is depicted in Figure 56. The small sample size may have affected this analysis. In addition, most of the data points for this group lie within a 1500-foot range of elevations and there is little variation in elevation at each canopy closure level (Figure 56). This poor representation of the range of canopy closure and elevation conditions likely has affected the analyses and leads to low confidence in the model results. A larger, more representative dataset may produce different results.

A second regression analysis was conducted using only canopy closure as an independent variable (Table 34, Figure 57). Given that elevation did not contribute significantly to the original model, the second model is probably the best choice at this time, but confidence is low in either equation due to the poor distribution of data.

The residuals depicting the scatter of the predicted data from actual observations are provided for both equations in Figure 58. The residuals appear to be well balanced and do not suggest a need for a transformation of the data.

Table 34. Results of the regression analyses conducted for the sedimentary lithology group. No outliers were identified. The table includes the analysis where both elevation and canopy closure were included as independent and the analysis with only canopy closure in the equation.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00067
Constant	30.30058	21.205326	39.395833	0.00001
Elevation	*-0.00246*	*-0.006416*	*0.001499*	*0.19883*
Canopy Closure	-0.123438	-0.215077	-0.031799	0.01287
n				14
Multiple R				0.857431
Squared Multiple R				0.735189
Adjusted squared multiple R				0.687041
Elevation Not In Equation				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00023
Constant	25.163457	21.281499	29.045416	0.00000
Canopy Closure	-0.162429	-0.230878	-0.093981	0.00023
n				14
Multiple R				0.830773
Squared Multiple R				0.690184
Adjusted squared multiple R				0.664366

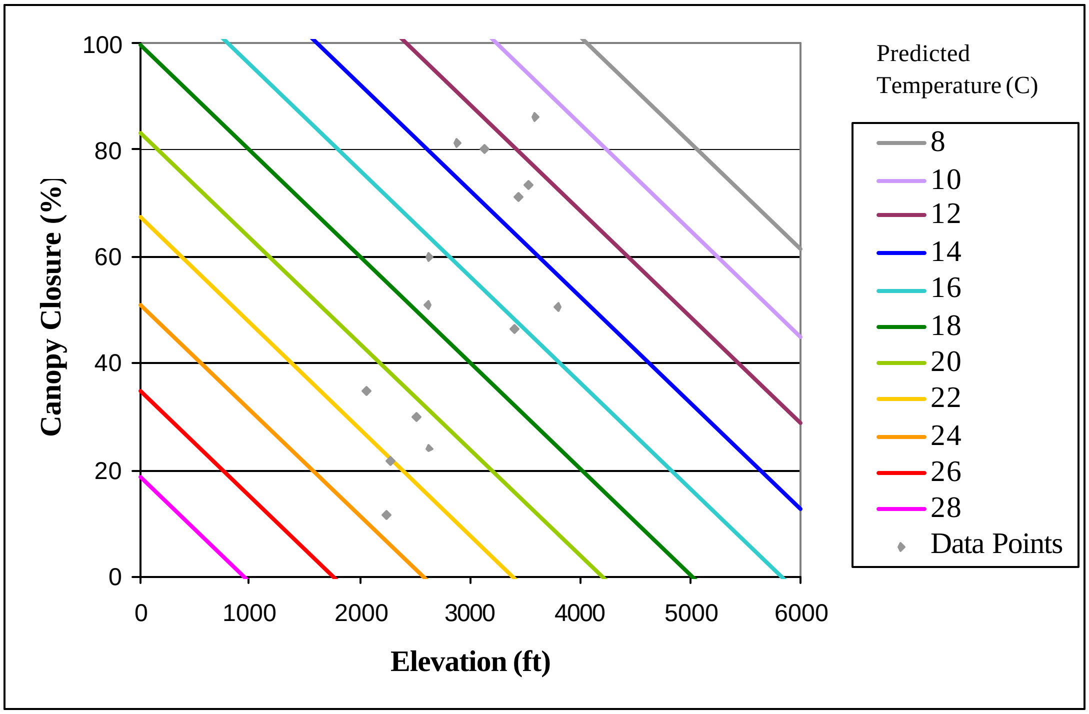


Figure 56. Depiction of the regression model developed for the sedimentary lithology group using elevation and canopy closure as independent variables. Note that elevation was not significant in this model at the 0.05 probability level. The gray points represent the actual data points used in the analysis. The colored lines represent various target temperatures as indicated in the legend.

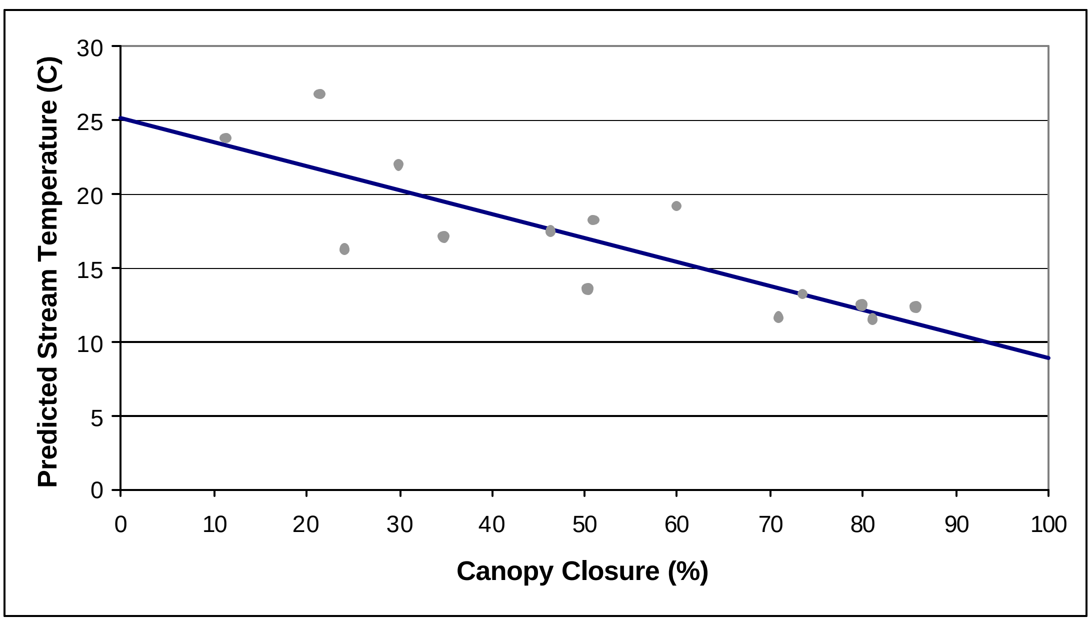


Figure 57. Depiction of the relationship between elevation and predicted stream temperature based on the regression analysis conducted for the lithology group termed "other deposits". Dots depict actual temperature of sample points used in the model.

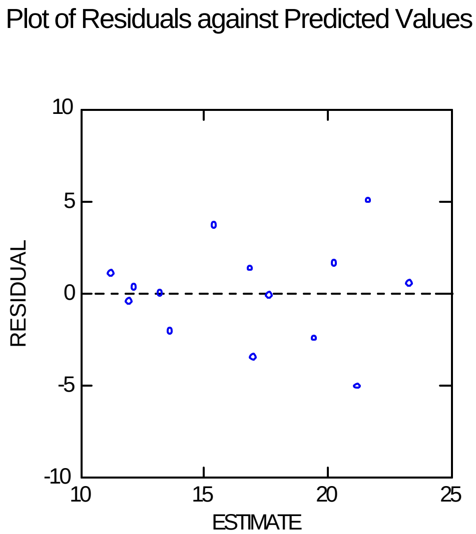


Figure 58. Residuals for the regression equation predicting stream temperature as a function canopy closure for the sedimentary lithology group. The residuals indicate a relatively random distribution of errors around the regression line.

Tests for Significant Differences between Populations

The relationships between canopy closure and elevation that are predicted to result in 16 and 18 degree C stream temperatures vary quite a bit between lithology groups (Figures 59 and 60). Two of the lithology groups represent a bedrock lithology (basalt and sedimentary) while the other three are lithologies defined as coarse deposits of material (alluvium, glacial drift, and other deposits). The two lines representing harder bedrock material tend to be relatively similar. Likewise, the lines representing deposits have somewhat similar slope, although the intercepts are dissimilar.

The equations generated for each lithology group and for all lithology groups as a whole using elevation and canopy closure as independent variables were tested to determine if the populations of the various combinations of datasets were from one population (Zar 1974). The test to determine if all ecoregion datasets are from the same population indicated that they are not all the same ($p<0.05$). Data representing pairs of lithology groups were then test to see if they were significantly different (Table 35). The alluvium, glacial drift, and basalt lithology groups are not from significantly distinct populations at the $p<0.05$ probability level. Likewise, basalt, other deposits, and sedimentary lithology groups were found to be statistically from the same populations.

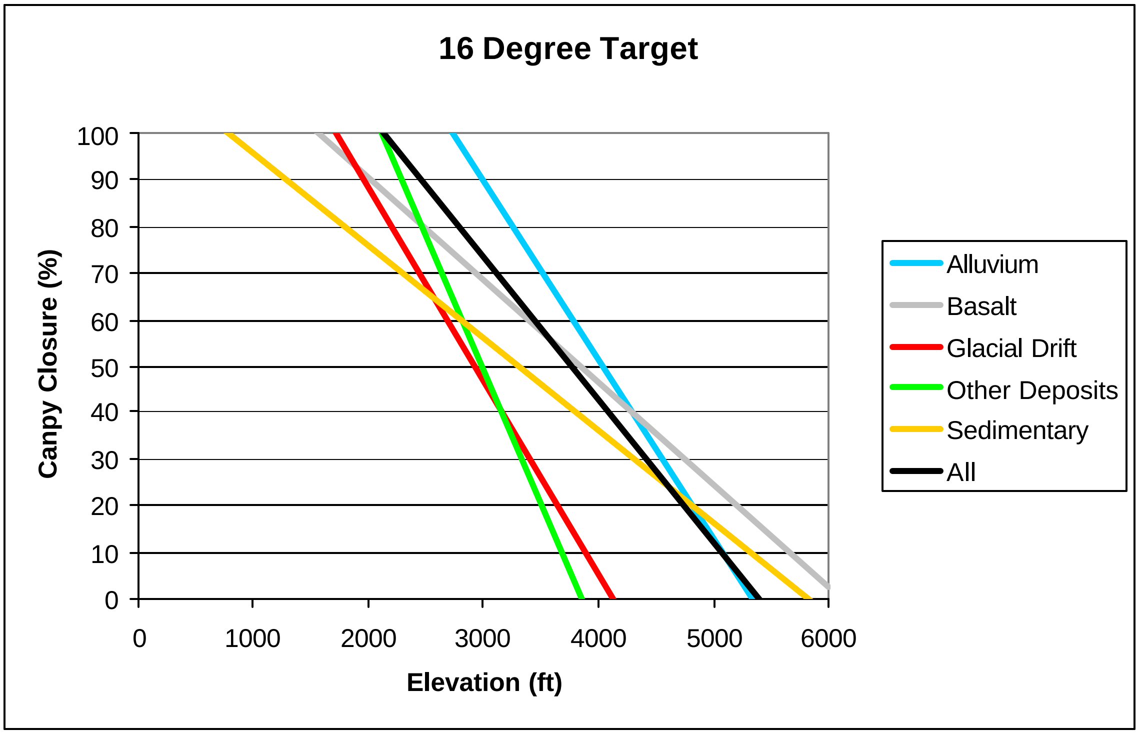


Figure 59. Lines depicting the combinations of canopy closure and elevation that are predicted to result in stream temperatures of 16 degrees C for the various lithology strata.

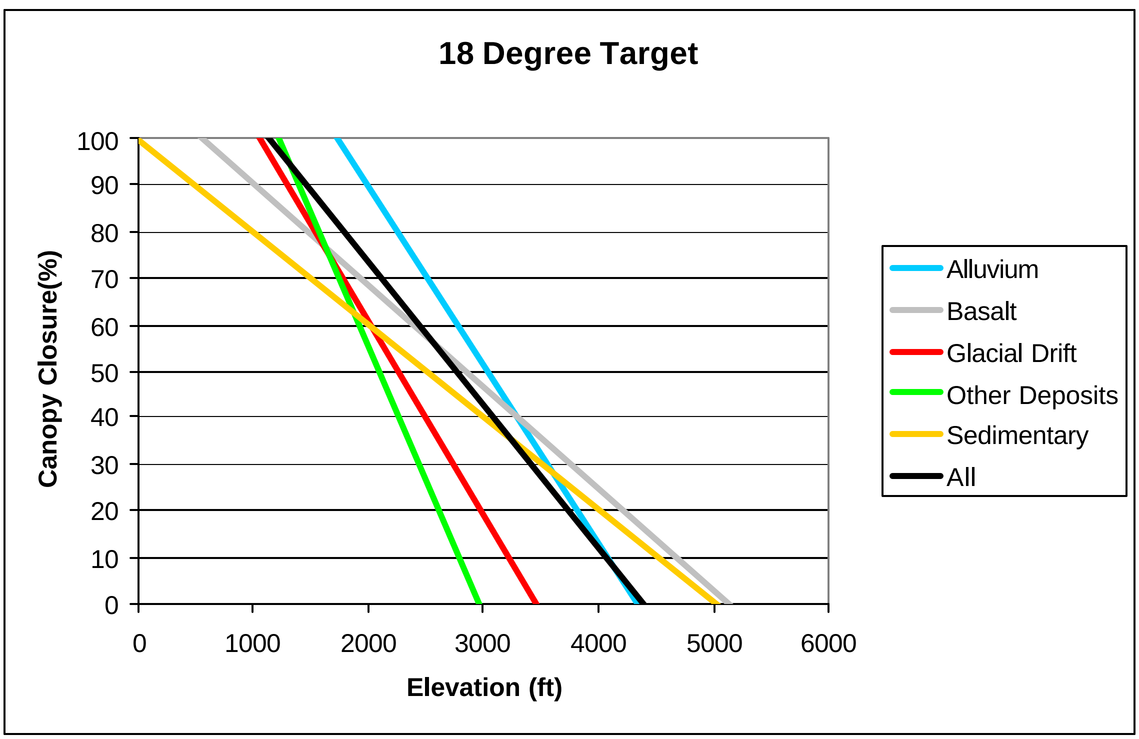


Figure 60. Lines depicting the combinations of canopy closure and elevation that are predicted to result in stream temperatures of 16 degrees C for the various lithology strata.

Table 35. F statistics and estimated probability to reject the null hypothesis that two populations are the same (in parentheses) for pairs of datasets representing the various lithology groups. The hypothesis was accepted (the datasets are statistically the same) for the pairs indicated with light gray shading. All other pairs were found to be from significantly different populations.

Lithology	Basalt	Glacial Drift	Other Deposits	Sedimentary
Alluvium	0.658 (p>0.25)	1.660 (0.1<P<0.25)	5.510 (P<0.0005)	0.711 (P>0.25)
Basalt		9.333 (p<0.0005)	1.190 (P>0.25)	0.410 (P>0.25)
Glacial Drift			5.911 (P<0.0005)	6.366 (P<0.0005)
Other Deposits				0.020 (P>0.25)

3.4.4 All Data Unstratified

There are 305 points in the dataset. Regression of temperature against canopy closure and elevation identified three outliers. Regression models were developed to evaluate whether these outliers should be excluded. The results of the three models are summarized in Table 36.

The model parameters exhibited the greatest amount of change when the first two outliers were removed. To further evaluate whether to include the outliers, the estimated canopy closure needed to meet a temperature standard of 16 degrees C was estimated for each model and plotted (Figure 61). The elimination of the first two outliers from the dataset substantially changed the model; however, elimination of the third outlier had a negligible effect. Therefore, the best model is that which excludes the first two outliers.

The resulting model with two outliers removed is depicted in Figure 62. The scatter of the data points used to develop this model is also depicted in the figure. The residuals from the analysis are presented in Figure 63.

Table 36. Results of regression analyses using elevation and canopy closure as independent variables and incorporating the entire database. Results with all data included in the analysis and with two and three outliers removed are presented.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	26.330075	24.724363	27.935787	0.0000
Elevation	-0.002199	-0.002723	-0.001675	0.0000
Canopy Closure	-0.057412	-0.072637	-0.042188	0.0000
n				305
Multiple R				0.566938
Squared Multiple R				0.321419
Adjusted squared multiple R				0.316925
Two Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	26.649863	25.144098	28.155628	0.00000
Elevation	-0.00228	-0.002771	-0.001789	0.00000
Canopy Closure	-0.061332	-0.07559	-0.047073	0.00000
n				303
Multiple R				0.610987
Squared Multiple R				0.373305
Adjusted squared multiple R				0.369127
Three Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	26.778834	25.309357	28.248311	0.00000
Elevation	-0.002246	-0.002725	-0.001767	0.00000
Canopy Closure	-0.064598	-0.07859	-0.050606	0.00000
n				302
Multiple R				0.627202
Squared Multiple R				0.393383
Adjusted squared multiple R				0.389325

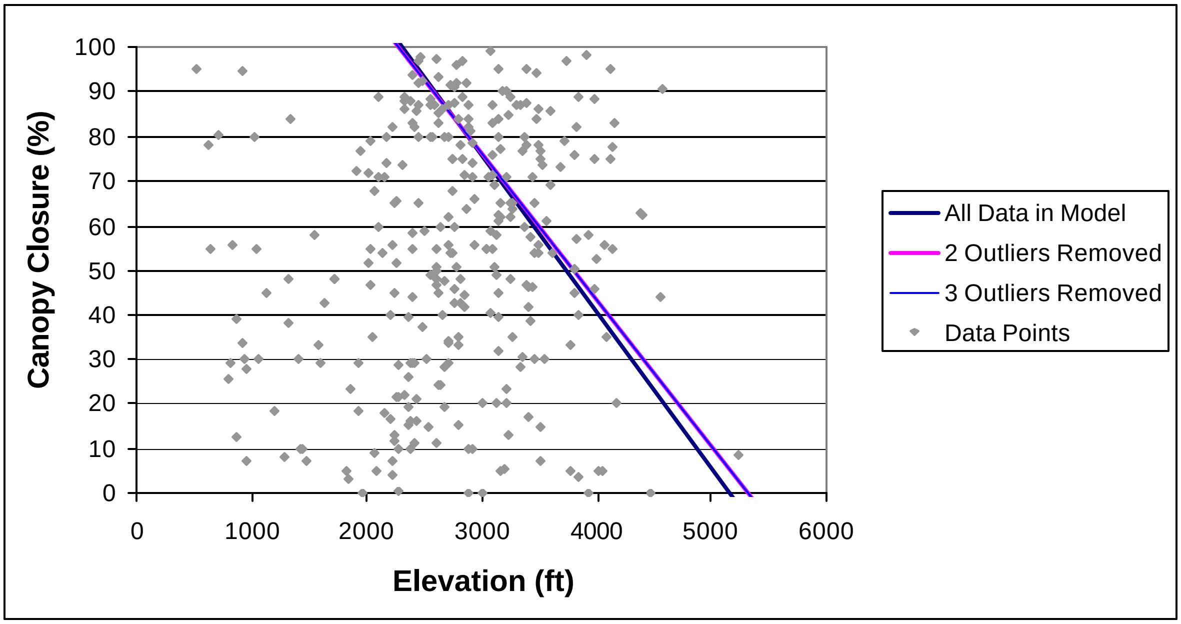


Figure 61. Depiction of the estimated canopy closure required throughout the range of elevation based on three models developed from the full unstratified data set. Each model has a different number of outliers excluded. Note, the lines with two outliers excluded and three outliers excluded are virtually identical.

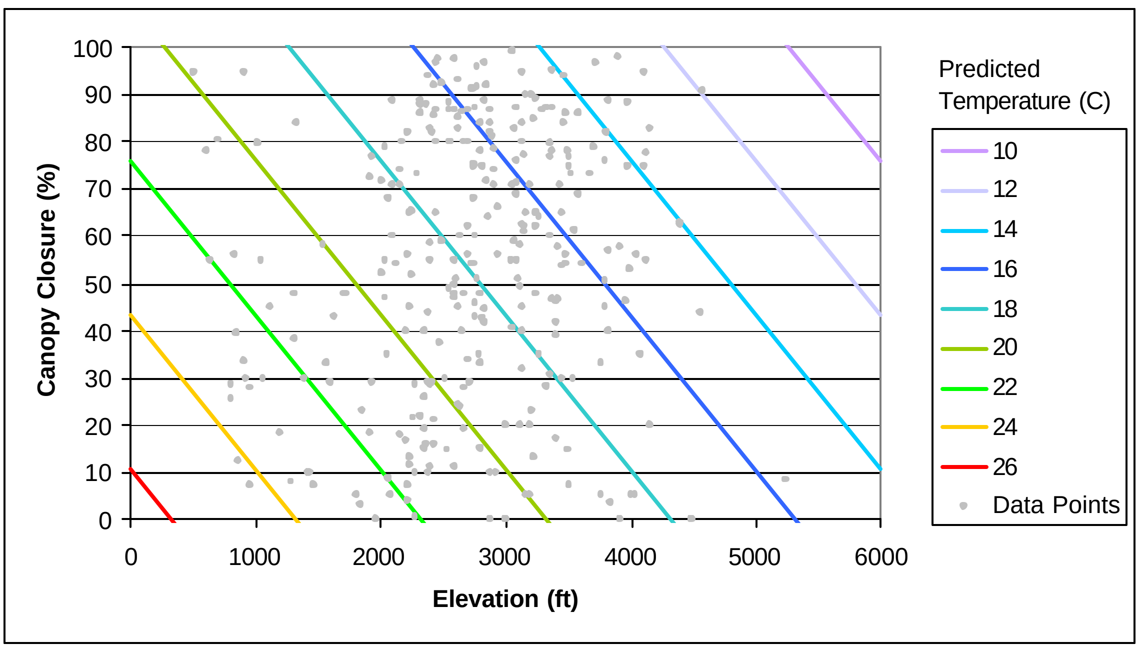


Figure 62. Regression model for the entire data set incorporating only canopy closure and elevation as independent variables. Two outliers were removed. Each line corresponds to a different target temperature.

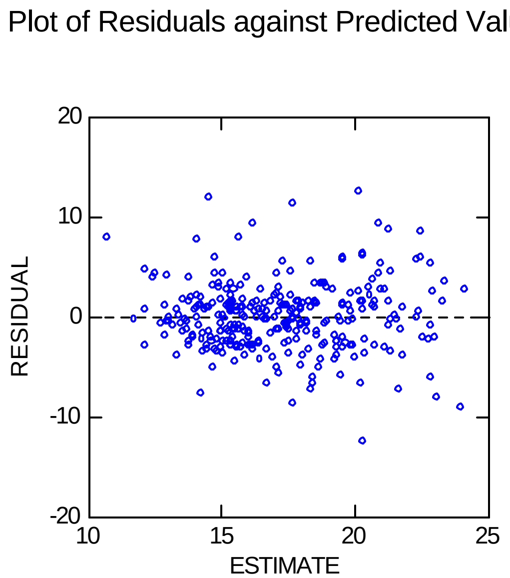


Figure 63. Residuals for the regression model based on the entire data set with two outliers removed. The residuals indicate a relatively random distribution of errors around the regression line.

3.5 REGRESSION MODELS WITH ALL INDEPENDENT VARIABLES INCLUDED

Regression models using the full set of independent variables were developed stratified by region, ecoregion, and lithology as well as for the entire dataset with no stratification. For each stratum, the best-fit model was identified. Where more than two independent variables were present in the best-fit model, the models were simplified by dropping the variable(s) that contributed the least to the regression. This was done to investigate the relative contribution of the additional variables in defining the amount of canopy closure that should be provided and to facilitate the development of models that are simpler to implement operationally.

The results of the regression analyses with all the independent variables are reported in tabular form. Tables contain the coefficients for the constant and each of the independent variables as well as the coefficients for the confidence intervals and the statistical significance of the parameter included in the model. To develop predictive equations using the tabled information, equations should be written as follows:

Temperature (^{o}C) = constant coefficient + (Variable A * coefficient for Variable A) + (Variable B * coefficient for Variable B + …),

where the variables in the equation are taken from the first column in the table and the coefficients for each variable are taken from the second column in the table. Hence, if canopy closure and basin size were the only variables in the equation, the coefficient for the constant were 2.3, the coefficient for basin size were 10.5 and the coefficient for canopy closure were 22.2, then the equation would be:

Temperature = 2.3 + 10.5 * basin size + 22.2 * canopy closure.

The full equations are provided in the Discussion section along with equations that are rearranged to identified canopy closure needed to meet target temperatures for various locations.

3.5.1 By Region

Southern Cascade Mountains

The dataset for the southern region of the Cascade Mountains did not include the site descriptive variables that were measured in the field (wetted and bank full width and depth, stream flow, valley width, and valley type); hence, the independent variables included in the regression analyses were the variables developed through GIS coverages, elevation, and canopy closure.

The regression analysis was started using a stepwise backward elimination regression. Two cases were identified as outliers. The stepwise regression was rerun without these outliers. The results of these regression analyses are summarized in Table 37. The variables that remained in the model in all three cases were canopy closure, basin size, elevation, and annual precipitation. The portion of the variance in the data explained by the model as measured by the squared multiple r increased from 46 to 51 percent with the removal of the outliers.

To test the effect that removal of the outliers had on the ultimate model predictions, temperature was estimated for roughly the maximum and minimum values of the other parameters in the model. Two estimates of predicted temperature were developed for each model using elevation set at 0 and 6000 feet, canopy closure set at 0 and 100 percent, basin size set at 500 and 250,000 acres, and annual precipitation set at 10 and 100 inches. The removal of the first outlier had a substantial effect on the predicted temperature (Table 37). The removal of the second outlier had a much smaller effect. This would suggest that the removal of the second outlier was not supported; however, the increase in the percent of the variance in the data explained by the removal of the outlier was approximately 3 percent. Therefore, both outliers were removed. The regression results for the third model in Table 38 are therefore the results of the best-fit model. The residuals for this model are provided in Figure 64. Inclusion of both elevation and basin size in the model raises questions regarding independence of the variables, but in this case, the correlation between those two variables was reasonably low (-0.279890).

Table 37. Results of the regression analyses for the region in the southern portion of the Cascade Mountains showing model parameters with and without identified outliers removed.

All Data In Model				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2- Tailed P***
Regression				0.00000
Constant	25.771197	23.010395	28.531999	0.00000
Elevation	-0.001264	-0.001922	-0.000605	0.00022
Canopy Closure	-0.066215	0.000011	0.000051	0.00000
Basin size	0.000031	-0.088722	-0.043708	0.00295
Annual Precip	-0.492998	-0.071564	-0.014955	0.00299
n				146
Multiple R				0.679283
Squared Multiple R				0.461426
Adjusted squared multiple R				0.446147
Remove 1 Outlier				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2- Tailed P***
Regression				0.00000
Constant	25.719304	23.072159	28.36645	0.00000
Elevation	-0.001305	-0.001937	-0.000674	0.00007
Basin size	0.000031	0.000012	0.000051	0.00171
Canopy Closure	-0.063674	-0.085297	-0.042052	0.00000
Annual Precip	-0.044274	-0.071418	-0.017131	0.00157
n				145
Multiple R				0.693934
Squared Multiple R				0.481544
Adjusted squared multiple R				0.466731
Remove 2 Outliers				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2- Tailed P***
Regression				0.00000
Constant	26.062722	23.521021	28.604422	0.00000
Elevation	-0.001597	-0.002222	-0.000973	0.00000
Basin size	0.000033	0.000014	0.000051	0.00063
Canopy Closure	-0.056943	-0.077961	-0.035926	0.00000
Annual Precip	-0.043643	-0.069639	-0.017646	0.00115
n				144
Multiple R				0.715175
Squared Multiple R				0.511475
Adjusted squared multiple R				0.497416

Table 38. Results of test of effect of removal of outliers for the south Cascade Mountains region on predicted temperature.

Model Tested	Elevation (ft)	Canopy Closure (%)	Basin Size (Acres)	Annual Precipitation (in.)	Predicted Temperature (C)
All Data	0	0	500	10	20.85672
	6000	100	250000	100	-29.9841
1 Outlier Removed	0	0	500	10	25.29206
	6000	100	250000	100	14.84450
2 Outliers Removed	0	0	500	10	25.64279
	6000	100	250000	100	14.67212

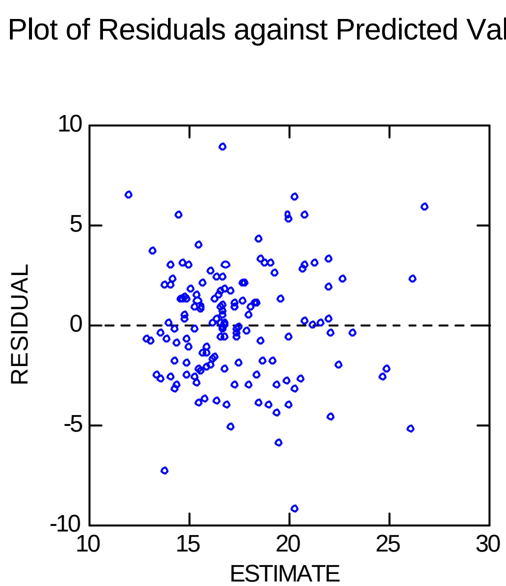


Figure 64. Residuals for the regression model using data from the south Cascade Mountains region with two outliers removed. The residuals indicate a relatively random distribution of errors around the regression line.

The results of the regression analyses were further explored by developing a model that excluded annual precipitation, which was the variable with the lowest probability for inclusion in the model and was the last variable entered in a stepwise regression. The results of the regression analysis excluding annual precipitation are summarized in Table 39. The portion of the variability in the data explained by the model decreased by roughly 4 percent with the exclusion of this variable. With the exclusion of average annual precipitation, the model predicts temperatures that are roughly 1.5^{o}C cooler at low elevations with little canopy closure and 4 ^{o}C warmer at high elevations with high canopy closure relative to the model with precipitation included (Table 40). The residuals for the model excluding precipitation are presented in Figure 65.

Table 39. Results of regression analysis for the south Cascade Mountains region excluding annual precipitation, the variable with the lowest probability of inclusion in the overall model.

Variable	*Coefficient*	*Lower 95%*	*Upper 95%*	*2- Tailed P*
Regression				0.00000
Constant	23.9702320	21.677393	26.26307	0.00000
Elevation	-0.0016200	-0.002271	-0.000978	0.00000
Basin size	0.0000390	0.000021	0.000058	0.00005
Canopy Closure	-0.0534190	-0.075065	-0.031772	0.00000
n				144
Multiple R				0.68757
Squared Multiple R				0.472752
Adjusted squared multiple R				0.461454

Table 40. Comparison of predicted temperature for similar parameters resulting from the best-fit model and the same model excluding average annual precipitation.

Model Tested	**Elevation (ft)**	**Canopy Closure (%)**	**Basin Size (Acres)**	**Annual Precipitation (in.)**	**Predicted Temperature (C)**
2 Outliers Removed (best fit model)	0	0	500	10	25.64279
	6000	100	250000	100	14.67212
Best-fit model with average annual precipitation removed	0	0	500	n/a	23.97023
	6000	100	250000	n/a	18.65833

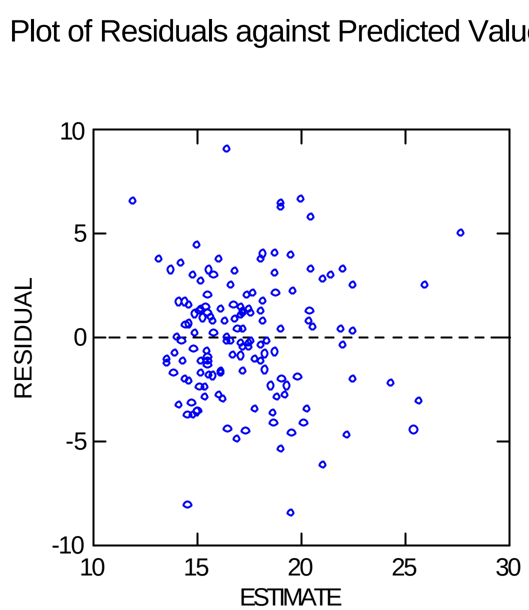


Figure 65. Residuals of the regression model for the south Cascades Mountain region with basin size, elevation, and canopy closure included in the model and average annual precipitation excluded. Two outliers were excluded in this model. The residuals indicate a relatively random distribution of errors around the regression line.

To complete the exploration of the models for the south Cascade Mountains region, the variables elevation, canopy closure, and basin size, were paired and included in regressions with two independent variables. The results of these analyses are provided in Table 41. Note that the precise value of the coefficients for the model with canopy closure and elevation vary somewhat from the model previously discussed. This is due to the removal of an outlier that was not removed from the original analysis of canopy closure and elevation. The paired model with the highest squared multiple r was the model that included basin size and elevation as the only independent variables. The model with the second highest squared multiple r was the model that included canopy closure and basin size. Depictions of the results of the models with two independent variables are provided in Figures 66, 67, and 68. The residuals from these models are provided in Figure 69.

These model results make very good physical sense. Basin size and elevation are the two variables that explain the most variance. Basin size affects stream size and air temperature tends to decrease with elevation. This model result therefore indicates that small high elevation streams tend to have cooler temperatures than large low elevation streams. The addition of canopy closure as the third most significant variable indicates that canopy closure is a modifier on stream temperature at any given elevation (air temperature) and basin (stream) size. Finally, the addition of average annual precipitation would tend to further modify the relationships by affecting stream size.

Table 41. Results of the regression analyses conducted on pairs of independent variables included in the best-fit model, excluding average annual precipitation. Two outliers were removed from the models.

Canopy Closure and Basin Size				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2- Tailed P***
Regression				0.00000
Constant	19.329356	17.861195	20.797517	0.00000
Basin size	0.000051	0.000031	0.00007	0.00000
Canopy Closure	-0.054232	-0.077621	-0.030843	0.00001
n				144
Multiple R				0.616289
Squared Multiple R				0.379813
Adjusted squared multiple R				0.371016

Canopy Closure and Elevation				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2- Tailed P***
Regression				0.00000
Constant	26.047389	23.879281	28.215496	0.00000
Canopy Closure	-0.053022	-0.072853	-0.03319	0.00000
Elevation	-0.002292	-0.002964	-0.001619	0.00000
n				173
Multiple R				0.551999
Squared Multiple R				0.304703
Adjusted squared multiple R				0.296523

Elevation and Basin Size				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2- Tailed P***
Regression				0.00000
Constant	20.979393	18.88166	23.077125	0.00000
Elevation	-0.001648	-0.002345	-0.000952	0.00001
Basin Size	0.000059	0.000041	0.000078	0.00000
n				144
Multiple R				0.618953
Squared Multiple R				0.383103
Adjusted squared multiple R				0.374352

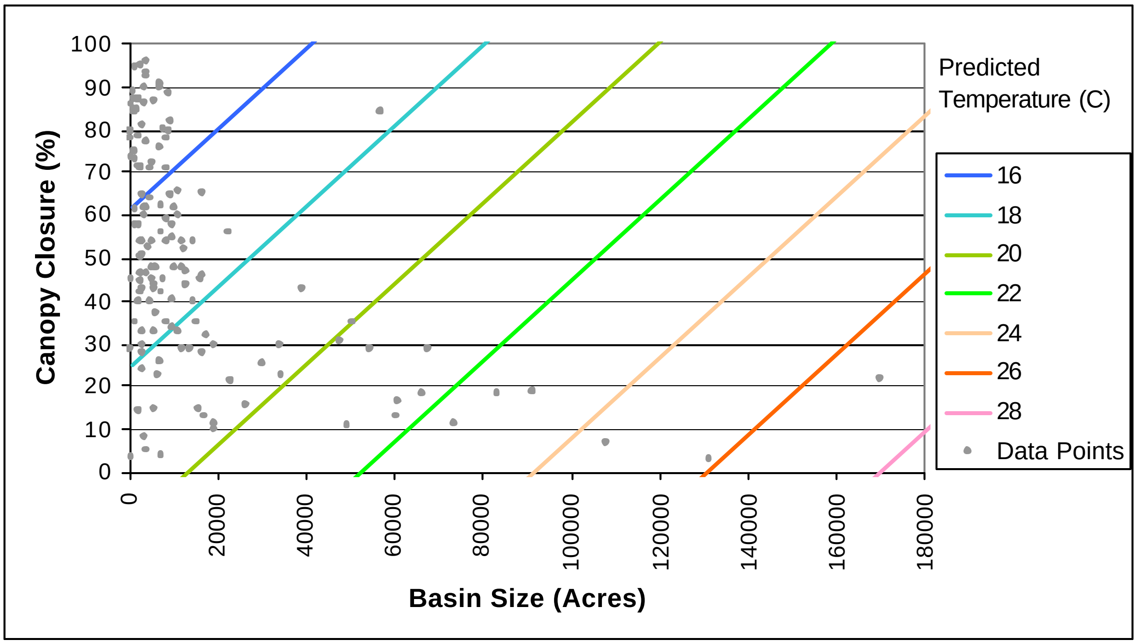


Figure 66. Depiction of the results of the model for the south Cascades Mountain region, which includes canopy closure and basin size as independent variables. Two outliers were removed from this model. The colored lines correspond to different predicted stream temperatures. The gray dots correspond to the data points used in the analysis.

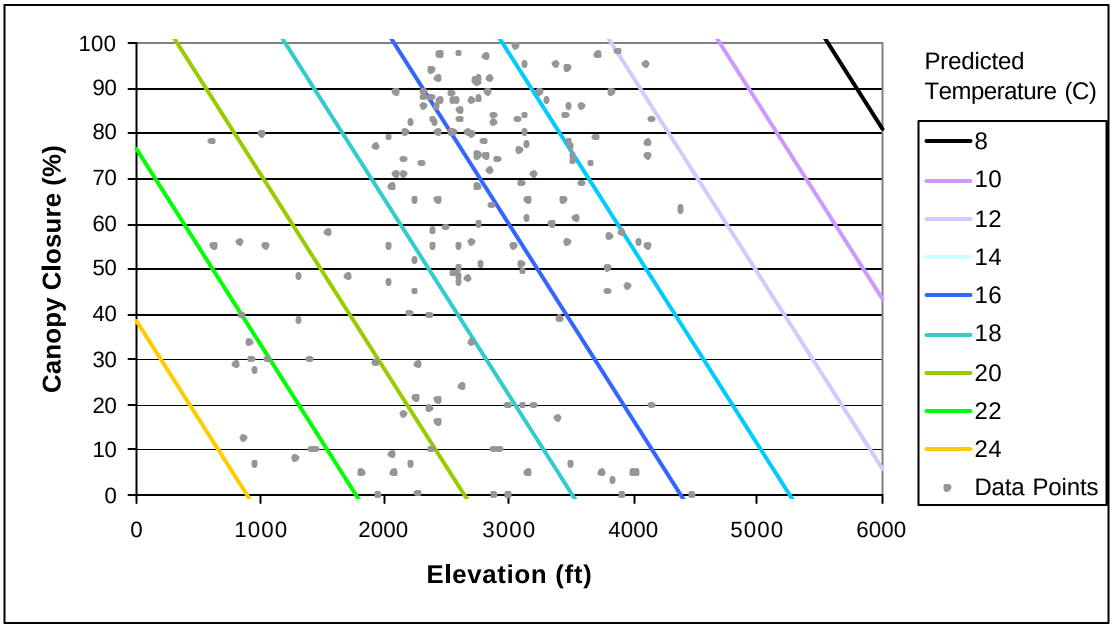


Figure 67. Depiction of the results of the model for the south Cascades Mountain region, which includes canopy closure and elevation as independent variables. Two outliers were removed from this model. The colored lines correspond to different predicted stream temperatures. The gray points represent the data points used in the analysis.

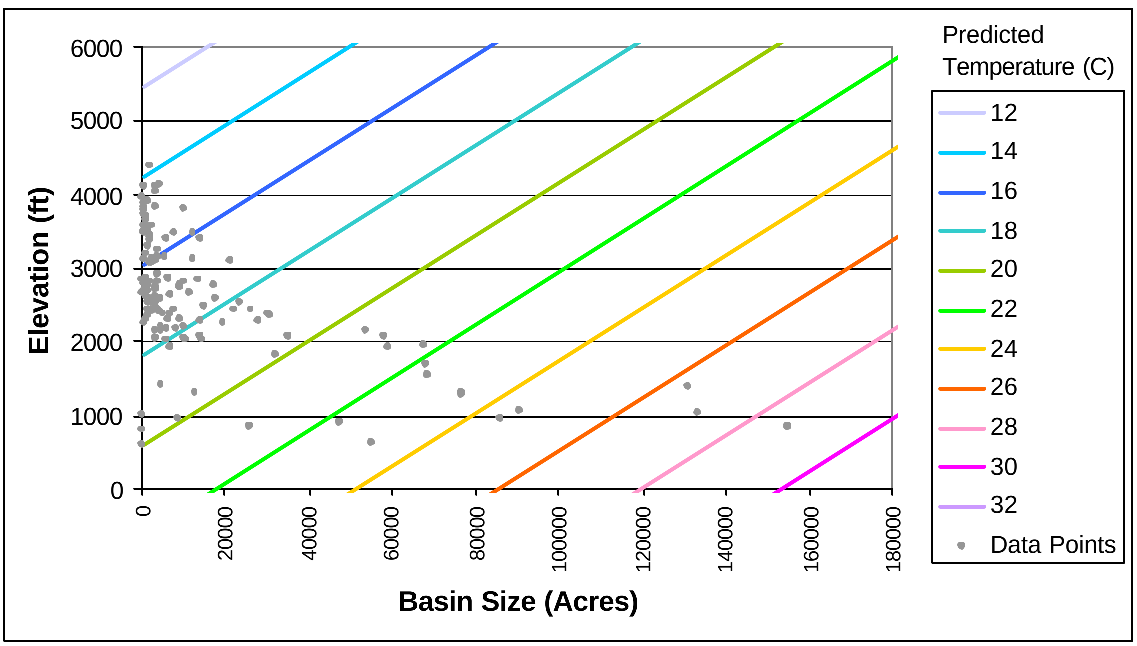


Figure 68. Depiction of the results of the model for the south Cascades Mountain region, which includes basin size and elevation as independent variables. Two outliers were remove d from this model. The colored lines correspond to different predicted stream temperatures. The dots represent the actual data points in the dataset.

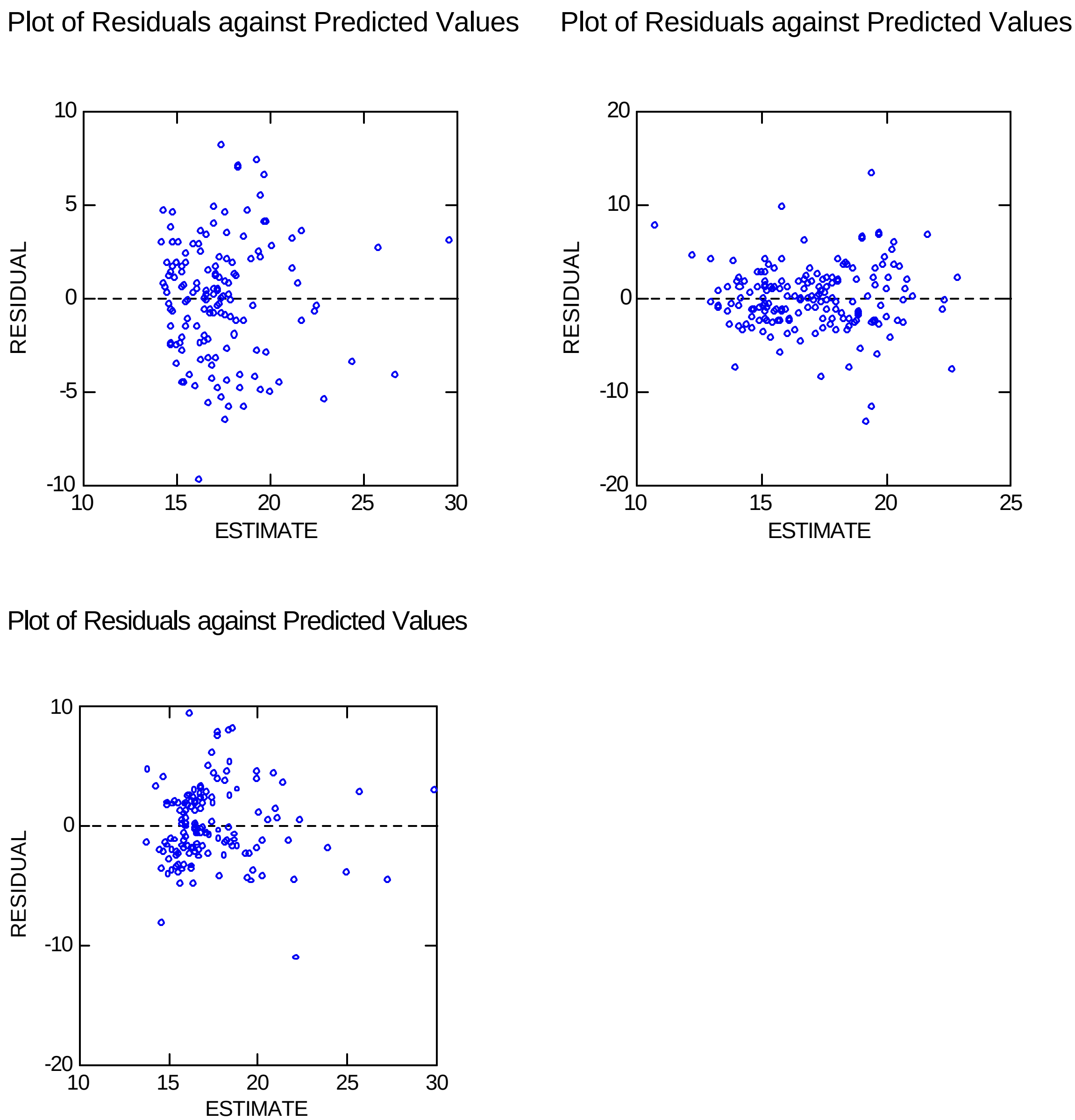


Figure 69. Residuals from the three regression models for the south Cascade Mountains region using two independent variables. The top left plot contains the residuals for the model that included canopy closure and basin size. The top right plot contains the residuals for the model that included canopy closure and elevation. The bottom plot contains the residuals for the model that included basin size and elevation. The residuals indicate a relatively random distribution of errors around the regression lines.

North Cascades

The North Cascades region has 37 cases in the full data set. Twelve of these cases included measures of air temperature and 28 of the cases included measures of bank full width. Bank full depth, wetted width, wetted depth, and stream flow were not included in any of the cases.

Basin size and stream gradient were causing an unexplained singularity in a matrix when the regression analysis was run with all the variables included. The regression parameters could not be calculated. Also notable was the high correlation between annual precipitation and elevation (0.8995) and the high correlation between basin size and distance from divide (0.880). One of each of these highly correlated pairs should be excluded from the analysis. Stream gradient was dropped from the analysis, which addressed the issue with the singularity.

An initial regression analysis was run which included all of the GIS generated independent variables and bank full width. No outliers were found. Bank full width and average annual precipitation were the only variables that were found to be significant in the model (Table 43). The regression explained 57.3 percent of the variance in the database. The residuals look well balanced.

The range of average annual precipitation in the region was not well represented by the database. All of the sample sites came from areas with an average of less than 30 inches per year (Figure 72). Hence, the wetter, higher elevation locations of the northern Cascades Mountains are not represented in the regression results.

Table 42. Results of regression based on data from the North Cascades region including all the GIS generated variables and bank full width as independent variables. No outliers were found in the dataset.

All GIS variables and bank full width				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2-Tailed P***
Regression				0.00016
Constant	18.767835	14.546148	22.989523	0.00000
Annual precipitation	-0.204538	-0.386878	-0.022199	0.02941
Bank full width	0.665919	0.356556	0.975282	0.00016
n				28
Multiple R				0.757248
Squared Multiple R				0.573425
Adjusted squared multiple R				0.563146

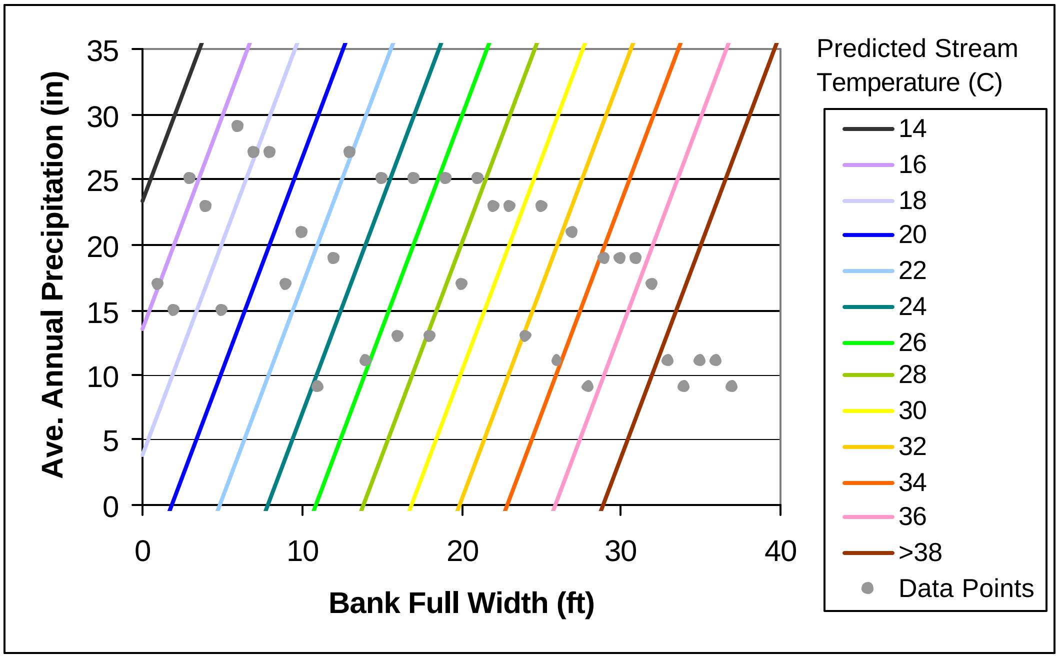


Figure 70. Depiction of the results of the model for the north Cascades Mountain region, which includes average annual precipitation and bank full width as independent variables. Note the narrow range of precipitation represented by the data. The colored lines correspond to different predicted stream temperatures. The dots represent the actual data points in the dataset.

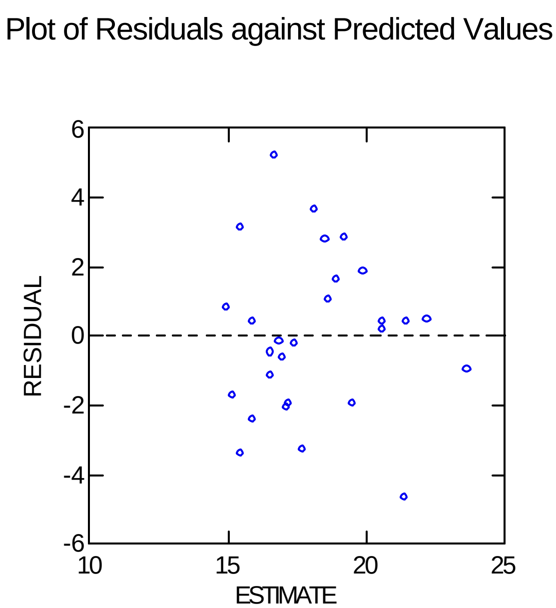


Figure 71. Residuals for the regression analysis conducted using the North Cascades region dataset and including all the GIS generated variables and bank full width as independent variables. The residuals indicate a relatively random distribu tion of errors around the regression lines.

A backward stepwise regression was then run with all the GIS variables, excluding bank full width. This provided a means of including the entire data set in a model. The only significant variable found in this analysis was distance from divide. The model explained 30.2% of the variance in the data (Table 43). Predicted temperature as a function of distance from divide is depicted in Figure 73 and the residuals from this analysis are provided in Figure 74.

Table 43. Results of regression analysis conducted using the North Cascades region data set with only the GIS generated independent variables included. No outliers were identified.

All GIS variables included				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00043
Constant	15.95284	12.434674	16.368023	0.00000
Distance to Divide	0.390652	0.582735	0.833921	0.00043
n				37
Multiple R				0.549158
Squared Multiple R				0.301575
Adjusted squared multiple R				0.281620

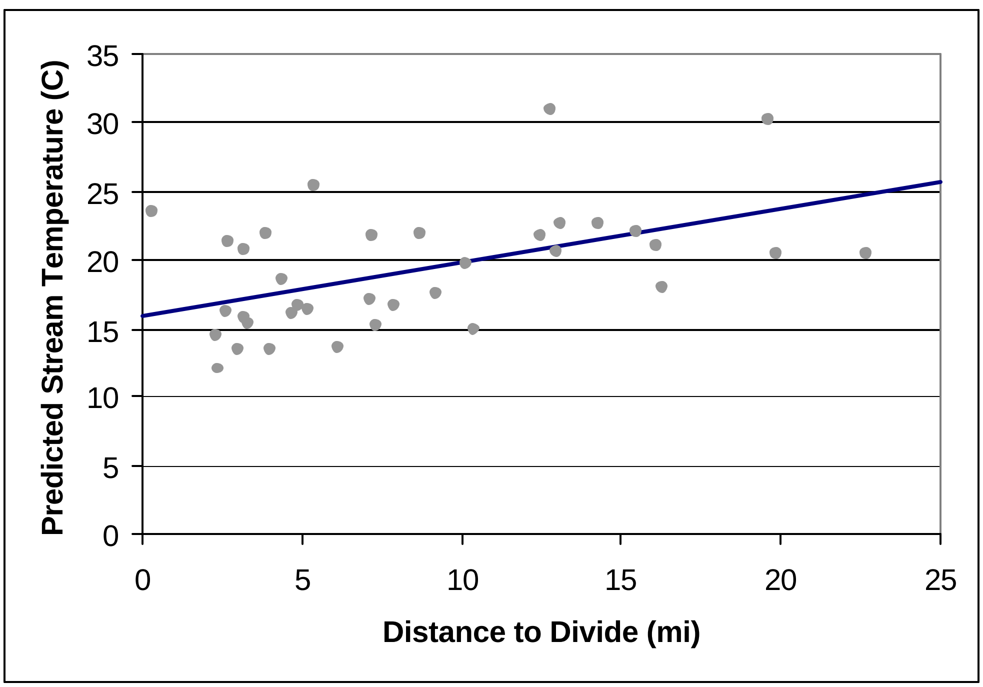


Figure 72. Depiction of the results of the regression analysis with the data from the North Cascades region incorporating all the GIS independent variables. Only distance from divide was found to be significant. The dots represent the actual data points in the dataset.

Plot of Residuals against Predicted Values

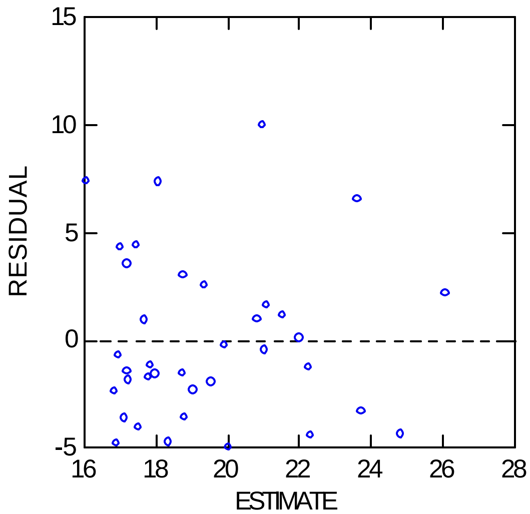


Figure 73. Plot of the residuals from the regression analysis for the North Cascades region incorporating all GIS generated independent variables. The top four data points have high leverage on the resulting regression equation, but did not test as outliers.

Given that the purpose of this study was to provide information that could be used to develop nomographs that guide the amount of canopy closure to leave during timber harvest operations, an additional regression analysis was completed forcing canopy closure into the model with distance from divide. This regression model explained 32.7 percent of the variance in the data set (Table 44). The addition of canopy closure resulted in a very small increase in the squared multiple r over that calculated with only distance from divide in the equation. The resulting equation is depicted in Figure 74.

Table 44. Results of regression analysis with canopy closure forced into equation along with distance to divide using the dataset for the North Cascades region.

Distance to divide and canopy closure				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2-Tailed P***
Regression				0.001193
Constant	17.84225	13.785705	21.898796	0.00000
Canopy Closure	*-0.030218*	*-0.084431*	*0.023995*	*0.26524*
Distance to Divide	0.336171	0.110508	0.561834	0.00468
n				37
Multiple R				0.571817
Squared Multiple R				0.326974
Adjusted squared multiple R				0.287385

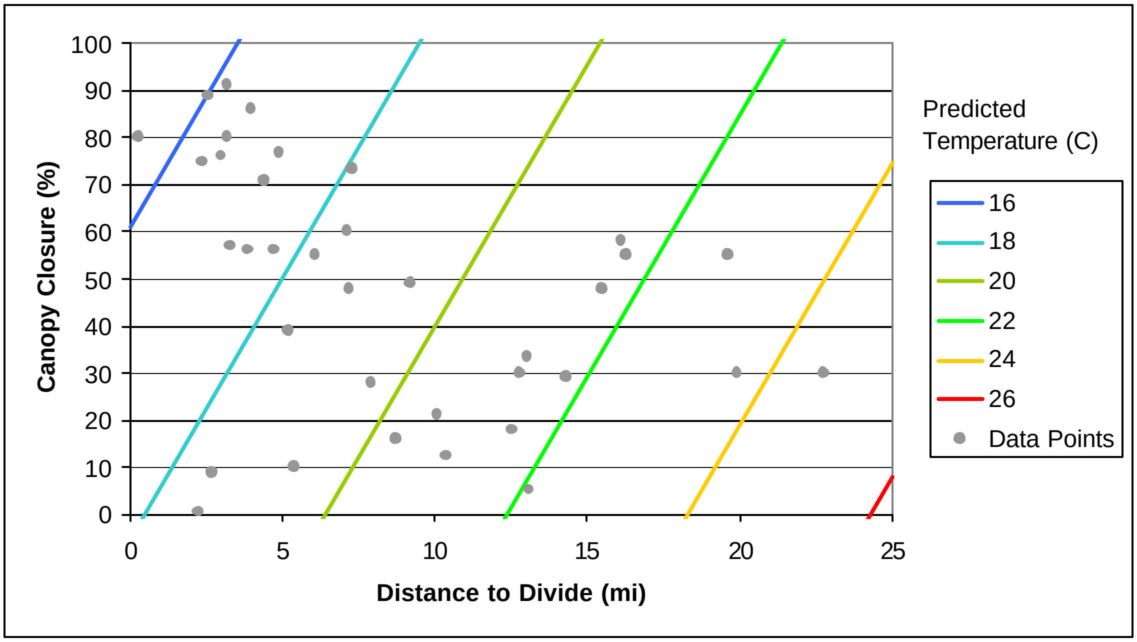


Figure 74. Depiction of the results of the model for the north Cascades Mountain region, which includes distance to divide and canopy closure as independent variables. Note, canopy closure was far from significant in this equation. The colored lines correspond to different predicted stream temperatures. The dots represent the actual data points in the dataset.

The range of canopy closures represented at each value of distance to divide tends to be rather narrow. In basins closer to the divide, most of the sample sites had higher canopy closures. Sample sites with moderate distances to divide tended to have low canopy closure, and those at the largest distance from divide tended to have moderate canopy closures. The poor representation of canopy closure at various distances from divide likely has affected the regression results and contributed to the non-significance of canopy closure in the analysis.

Northeast Washington

There are 86 data points in the Northeast Washington dataset. Of these, only 15 data points have measurements for wetted depth, bank full depth, and stream flow. Eleven of these data points have canopy closure levels in excess of 80 percent and 10 of these points are within an elevation range of 2400 to 3000 feet. None of the data points has measurements for wetted width or bank full width. As a result, the set of samples present in the data base with these physical measurements is not only small, but is also poorly representative of the range of canopy closure and elevation conditions within the region. For these reasons, the variables wetted depth, bank full depth, and stream flow should be excluded from the analyses. Nevertheless, a backward stepwise regression was run on these 15 data points with all available independent variables included. None of the variables was eliminated. This left 12 variables in a model driven by 15 data points. As the number of variables in a model approaches the number of data points, the squared multiple r naturally approaches 1.0, regardless of whether the model makes physical sense. In this case, the model with 12 variables and 15 data points had a squared multiple r of 0.999897. The results must be considered spurious. Wetted depth, bank full depth, and stream flow were eliminated from all other analyses in this region.

A backward stepwise regression was run for all 86 cases for Northeast Washington excluding the variables discussed above. Two outliers were identified. Regressions were run with and without the outliers to compare the results. In the case with all the data included, the variables that remained in the model were hill slope gradient and distance to divide. The squared multiple r was 0.573425 (Table 45). In the case where the outliers were removed, canopy closure and distance to divide remained in the model. The squared multiple r with the outliers removed was 0.744811 (Table 45). The variables in the model changed with the removal of the outliers and the portion of the variance in the data increased substantially. Therefore, it was assumed that removal of the outliers was appropriate.

The results of the regression model with the outliers removed are depicted in Figure 75 and the residuals from that regression analysis are displayed in Figure 76. The dataset is not well distributed across the range of canopy closure and distance from divide. Most of the data were collected in areas with high canopy closure levels located within 7 miles of the divide (Figure 75). Low canopy closure situations are poorly represented and areas with larger (>10 miles) distances from divide with high canopy closure levels are entirely missing from the dataset. The bias in the dataset is also apparent in the residuals (Figure 76). Additional data points filling these poorly represented situations would likely affect the model outputs.

Table 45. Results of regressions run on the NE Washington dataset with and without the outliers removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2-tailed P***
Regression				0.00000
Constant	10.761427	9.046098	12.476756	0.00000
Hill Slope Gradient	6.377788	0.984199	11.771377	0.02105
Distance to Divide	0.809122	0.655918	0.962326	0.00000
n				86
Multiple R				0.757248
Squared Multiple R				0.573425
Adjusted squared multiple R				0.563146
Outliers Removed (cases 204 and 208)				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	14.401348	12.434674	16.368023	0.00000
Canopy Closure	-0.028686	-0.050275	-0.007098	0.00984
Distance to Divide	0.708328	0.582735	0.833921	0.00000
n				84
Multiple R				0.863024

Squared Multiple R	0.744811
Adjusted squared multiple R	0.738510

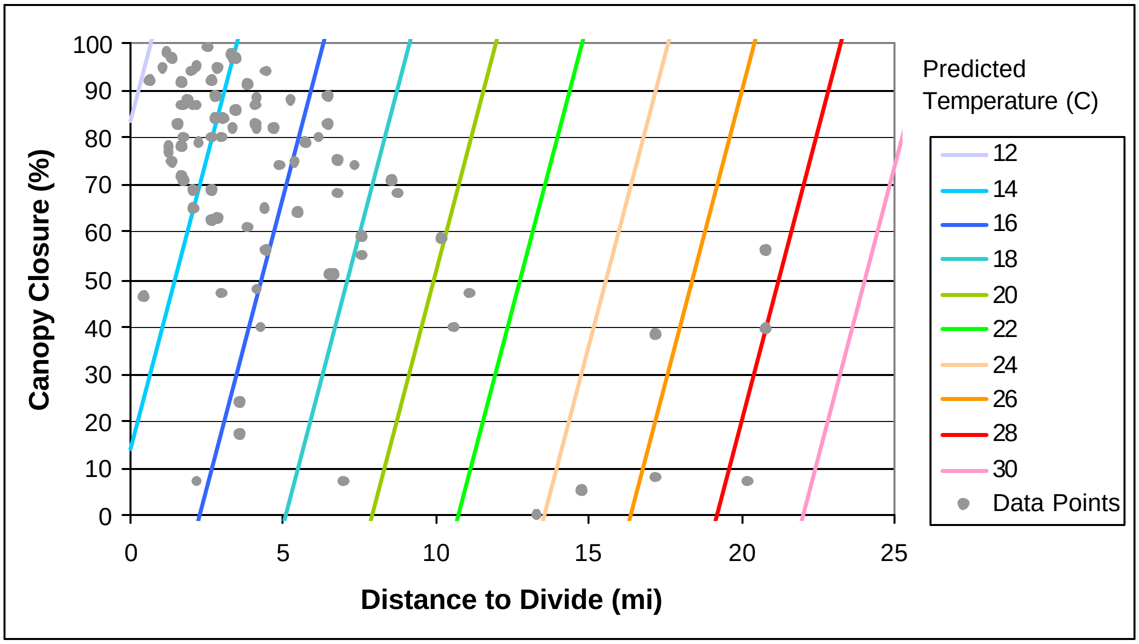


Figure 75. Depiction of the results of the regression model developed for NE Washington. Two outliers were removed. The points on the graph represent the scatter of the data used to develop the model. The lines represent a range of temperature targets.

Plot of Residuals against Predicted Values

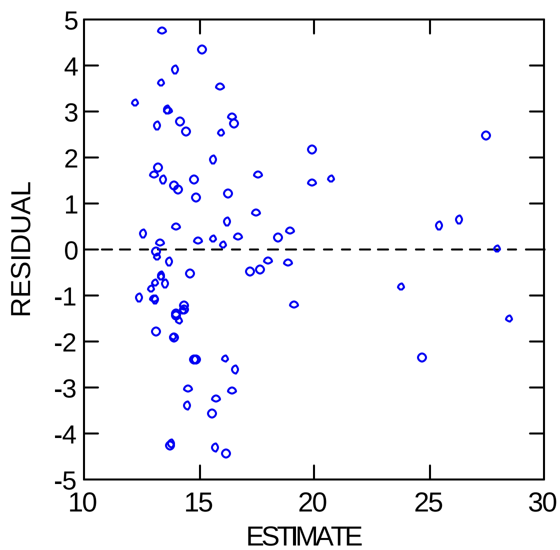


Figure 76. Residuals from the regression analysis conducted with NE Washington data with two outliers removed. The residuals indicate a relatively random distribution of errors around the regression line.

Blue Mountains

The Blue Mountain dataset for Washington State is small. There are only seven points in the dataset. In the analyses related to models developed using only elevation and canopy closure, a second dataset from Oregon was used to increase the sample size. This second data set does not include basin size or distance from divide, which are variables that were found to be significant in other regions. Therefore, no additional analyses were conducted in this region. The reader is referred to the discussion of canopy closure and elevation models to review the best-fit model for the Blue Mountain region.

3.5.2 By Ecoregion

The datasets for Ecoregions M332G (Blue Mountains Section) and -331A (Palouse Prairie Section) were too small to allow the development of regression equations. The regression results for the other regions are summarized below.

M333A (Okanogan Highlands Section)

An initial regression analysis was run which included all of the GIS generated independent variables. Two outliers were identified. Regressions were run with and without the outliers to compare the results. In the case with all the data included, the variables that remained in the model were hill slope gradient, distance to divide, and abs(sine aspect). The squared multiple r was 0.587279 (Table 46). In the case where the outliers were removed, canopy closure and distance to divide remained in the model, the abs (sine of aspect) was removed, and precipitation was added. The squared multiple r with the outliers removed was 0.744811 (Table 46). Hence, the portion of the variance in the data that was explained by the model increased by 15.7 percent with the removal of the outliers. The variables in the model also changed with the removal of the outliers. Therefore, it was assumed that removal of the outliers was appropriate. The residuals from the regression with the outliers removed are displayed in Figure 77.

The range of predicted temperature as a function of the three significant variables in the model (distance from divide, canopy closure, and annual precipitation) is depicted in Table 47. Distance from divide has the greatest effect on predicted temperature and was also the most significant variable. Annual precipitation had the second greatest effect on predicted temperature, followed by canopy closure.

To complete the exploration of the models for Ecoregion M333A (Okanogan Highlands Section), the variable distance from divide, which was the most significant variable, was paired with annual precipitation and with canopy closure in regressions with two independent variables. The effectiveness of distance from divide as the only independent variable was also evaluated in a regression model. The results of these analyses are provided in Table 48. The portion of the variance in the data explained by these two models as measured by the squared multiple r were almost identical. Depictions of the results of the models with two independent variables and with distance from divide alone are provided in Figures 78, 79, and 80.

Table 46. Results of regressions run on the Ecoregion M333A (Okanogan Highlands Section) dataset with and without the outliers removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2-tailed P***
Regression				0.00000
Constant	10.338899	8.517647	12.160151	0.00000
Hill Slope Gradient	5.106027	0.719127	9.492926	0.02295
Distance from Divide	0.732298	0.611868	0.852729	0.00000
Abs (Sine of Aspect)	1.781783	0.045429	3.518138	0.0444
n				113
Multiple R				0.758024
Squared Multiple R				0.574601
Adjusted squared multiple R				0.562893
Outliers Removed (cases 204 and 208)				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	16.420597	13.727502	19.113692	0.00000
Canopy Closure	-0.020744	-0.039434	-0.002055	0.02993
Annual Precipitation	-0.066021	-0.125413	-0.006628	0.02969
Distance from Divide	0.573397	0.442904	0.70389	0.00000
n				111
Multiple R				0.842201
Squared Multiple R				0.709303
Adjusted squared multiple R				0.701152

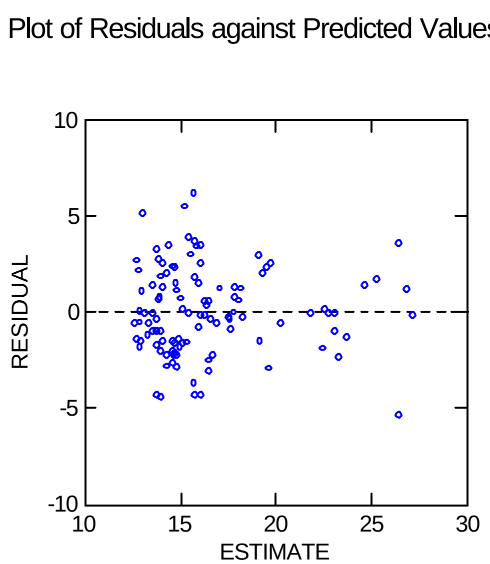


Figure 77. Residuals of the regression model for Ecoregion (Okanogan Highlands Section covering predominately the NE region of Washington and the northern Cascade Mountains) canopy closure,

annual precipitation, and distance from divide included in the model. Two outliers were excluded in this model.

Table 47. Predicted temperature at a range of distance from divide, canopy closure, and annual precipitation for the regression model on data from Ecoregion M333A (Okanogan Highlands Section) with two outliers removed.

Distance from Divide (mi)	Canopy closure (%)	Annual Precipitation (in)	Predicted Temperature (C)	Lower 95%	Upper 95%
25	0	10	30.09531	23.54597	36.64466
25	0	50	27.45447	18.52945	36.37954
25	0	100	24.15342	12.2588	36.04814
25	50	10	29.05811	21.57427	36.54191
25	50	50	26.41727	16.55775	36.27679
25	50	100	23.11622	10.2871	35.94539
25	100	10	28.02091	19.60257	36.43916
25	100	50	25.38007	14.58605	36.17404
25	100	100	22.07902	8.315402	35.84264
12	0	10	22.64115	17.78822	27.49409
12	0	50	20.00031	12.7717	27.22897
12	0	100	16.69926	6.50105	26.89757
12	50	10	21.60395	15.81652	27.39134
12	50	50	18.96311	10.8	27.12622
12	50	100	15.66206	4.52935	26.79482
12	100	10	20.56675	13.84482	27.28859
12	100	50	17.92591	8.8283	27.02347
12	100	100	14.62486	2.55765	26.69207
0	0	10	15.76039	12.47337	19.04741
0	0	50	13.11955	7.456852	18.78229
0	0	100	9.818497	1.186202	18.45089
0	50	10	14.72319	10.50167	18.94466
0	50	10	14.72319	10.50167	18.94466
0	50	100	8.781297	-0.7855	18.34814
0	100	10	13.68599	8.529972	18.84191
0	100	50	11.04515	3.513452	18.57679
0	100	100	7.744097	-2.7572	18.24539

Table 48. Results of regression runs for the Ecoregion M333A (Okanogan Highlands Section) dataset using distance from divide paired with canopy closure, paired with annual precipitation, and alone as the independent variables.

Distance from Divide and Canopy Closure				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2-tailed P***
Regression				0.00000
Constant	14.078066	12.371708	15.784425	0.00000
Canopy Closure	-0.020255	-0.039268	-0.001243	0.03702
Distance From Divide	0.652548	0.541271	0.763825	0.00000
n				111
Multiple R				0.834332
Squared Multiple R				0.696110
Adjusted squared multiple R				0.690482
Distance from Divide and Annual Precipitation				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2-tailed P***
Regression				0.00000
Constant	14.656068	12.44413	16.868006	0.00000
Distance from Divide	0.644227	0.528406	0.760047	0.00000
Annual Precipitation	-0.064468	-0.124884	-0.004052	0.03672
n				111
Multiple R				0.834355
Squared Multiple R				0.696149
Adjusted squared multiple R				0.690522
Distance from Divide Only				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***2-tailed P***
Regression				0.00000
Constant	12.408009	11.723452	13.092567	0.00000
Distance from Divide	0.719928	0.62694	0.812916	0.00000
n				111
Multiple R				0.826778
Squared Multiple R				0.683562
Adjusted squared multiple R				0.680659

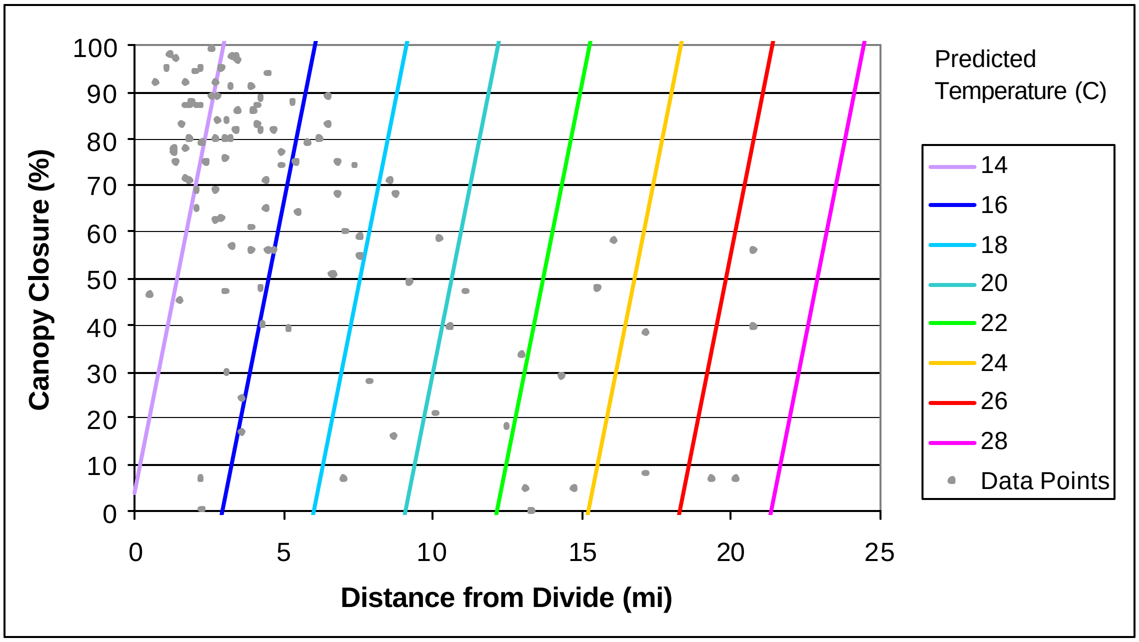


Figure 78. Depiction of the results of the regression analysis using the data from Ecoregion M333A (Okanogan Highlands Section) incorporating distance from divide and canopy closure.

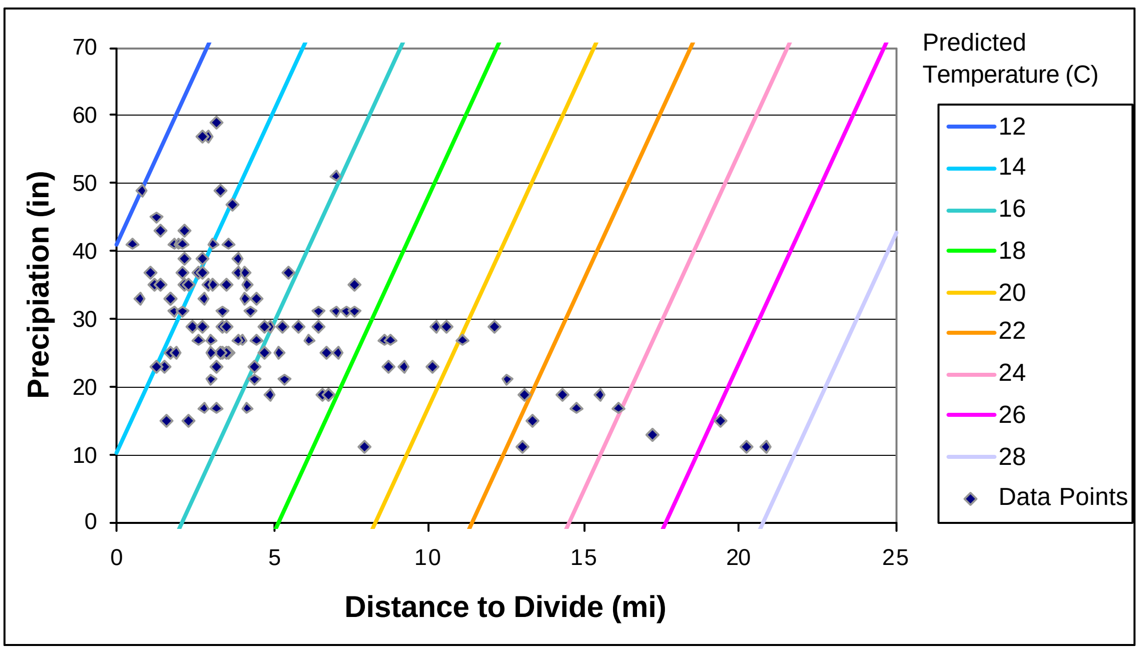


Figure 79. Depiction of the results of the regression analysis with the data from Ecoregion M333A (Okanogan Hi ghlands Section) incorporating distance from divide and annual precipitation.

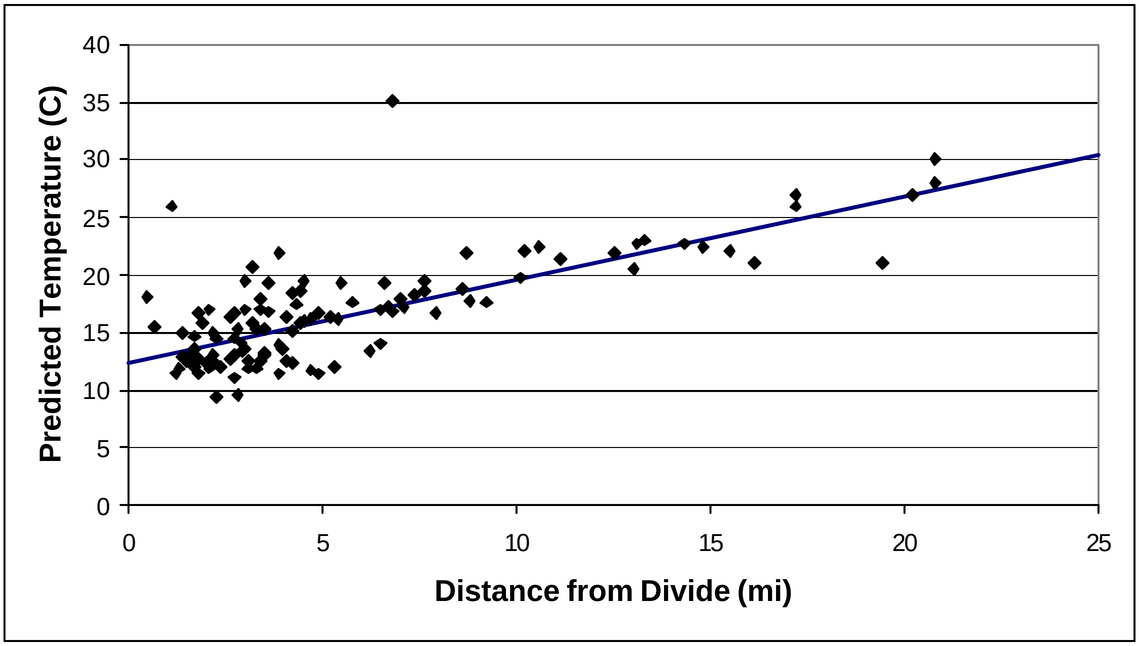


Figure 80. Depiction of the predicted temperature (C) for Ecoregion M333A (Okanogan Highlands Section) using distance from divide (mi) as the only independent variable. Dots represent actual data (distance from divide and actual summer maximum temperature).

Data represented in this dataset is somewhat biased. The majority of the data comes from locations that are less than 7 miles from the divide with greater than 50 percent canopy closure (Figure 78). The range of precipitation was well represented at sample sites that were less than 7 miles from the divide; however sample locations that were greater than 12 miles from the divide tended to be very dry sites with less than 20 inches of annual precipitation (Figure 79).

The plots of the residuals from the three regression models discussed above are presented in Figure 81. All three sets of the residuals are very similar. This reflects the dominant effect of distance from divide in the equations.

M242B (Western Cascades Section)

There are only 11 samples from Ecoregion M242B (Western Cascades Section). These are located primarily in the southern portion of the eastern Cascade Mountains. The regression equation with all possible independent variables identified three significant variables. These were canopy closure, distance from divide, and annual precipitation.

Two outliers were identified and removed. The resulting model had a high squared multiple r (0.976716) (Table 49), however it has low degrees of freedom. This results in wide confidence intervals around the various coefficients (Table 49) and the predicted stream temperatures (Table 50). The variables included in the regression were the same as those that were included in the model were canopy closure, average annual precipitation, and distance from divide.

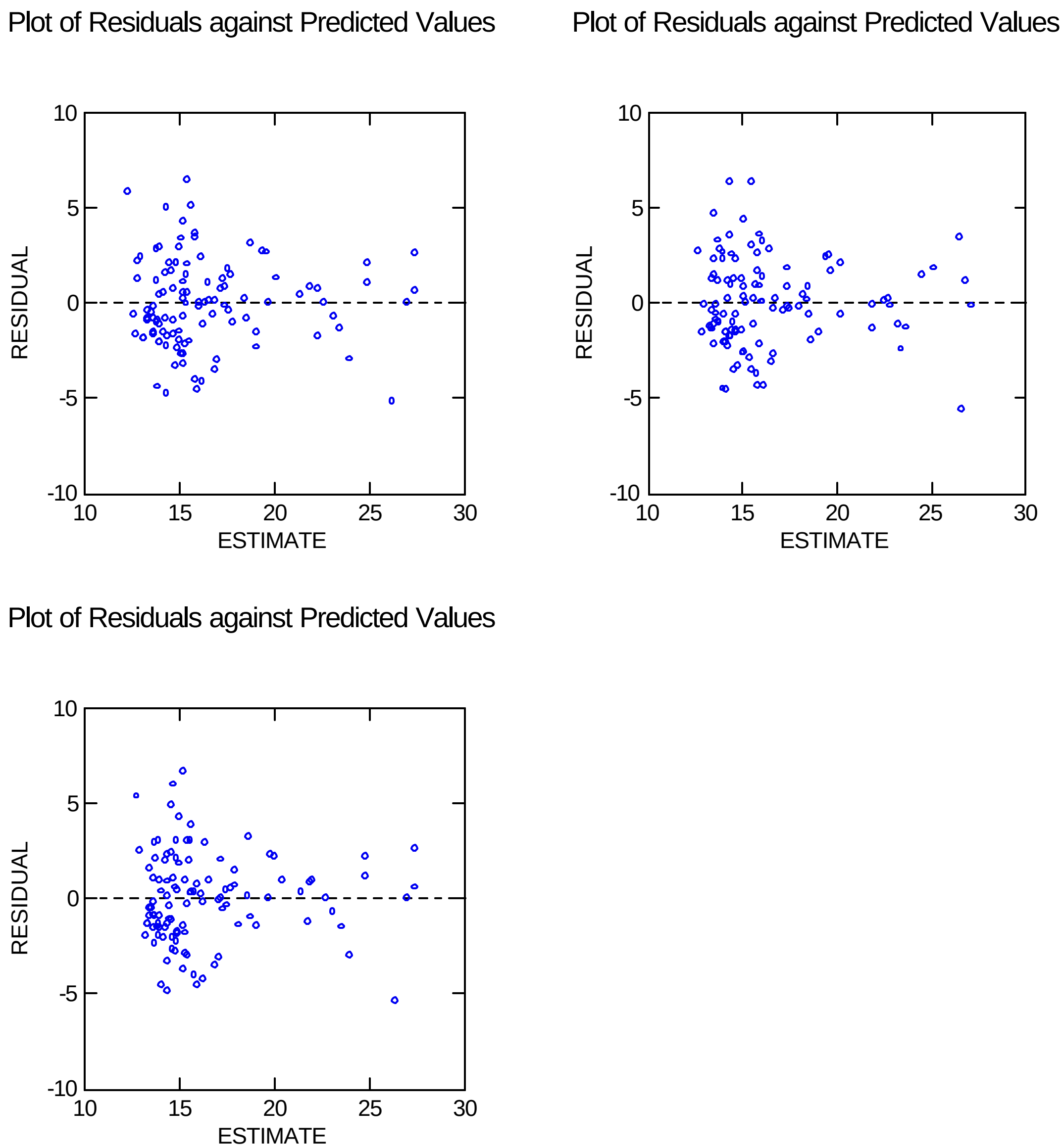


Figure 81. Residuals from three regression analyses for Ecoregion (Okanogan Highlands Section). The top right is from the regression including distance from divide and canopy closure, the top left is from the regression including distance from divide and annual precipitation, and the bottom is from the regression using only distance from divide as the independent variable.

An additional model was developed that included canopy closure and precipitation as the only dependent variables. These variables explained the highest portion of the variability in the data. Canopy closure was not significant in this equation and the 0.05 probability level (p=0.14742). The squared multiple r for this model was lower than that for the 3-parameter model described above (Table 49).

The dataset for Ecoregion M242B (Western Cascades Section) represents high precipitation areas (Figure 82). Data points with canopy closures greater than 80 percent were missing from the database. Distance from divide ranged from 1.0 to 8.9 miles. Residuals for both equations are provided in Figure 83.

Confidence in any of the regressions developed for this ecoregion is low. With two or three outliers removed, the dataset is reduced to 8 or 9 points. Use of any of these equations is not recommended. Should stratification on ecoregion be pursued, we recommend additional data be collected in this ecoregion.

Table 49. Results of regression analyses using dataset from Ecoregion M242B (Western Cascades Section). Rows in italics are not significant at the 0.05 probability level.

All Data (2 outliers removed)				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000997
Constant	13.933518	10.611971	17.255066	0.00012
Canopy Closure	-0.104542	-0.138509	-0.070576	0.00052
Annual Precipitation	0.089094	0.048059	0.130128	0.00255
Distance from Divide	-0.363069	-0.619318	-0.106819	0.01487
n				9
Multiple R				0.975841
Squared Multiple R				0.952265
Adjusted squared multiple R				0.923624
Canopy Closure and Annual Precipitation Only (3 outliers removed)				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.010335
Constant	8.902862	3.830046	13.975678	0.00633
Canopy Closure	*-0.047569*	*-0.118958*	*0.023821*	*0.14742*
Annual precipitation	0.099019	0.046404	0.151634	0.00472
n				8
Multiple R				0.916192
Squared Multiple R				0.839408
Adjusted squared multiple R				0.775171

Table 50. Predicted stream temperature with confidence limits for various combinations of canopy closure, annual precipitation, and distance from divide.

Canopy Closure (%)	Annual Precipitation (in)	Distance from Divide (mi)	Predicted Temperature (C)	Lower 95%	Upper 95%
25	0	10	7.689278	0.956066	15.49067
25	0	50	-6.83348	-23.8167	15.49067
25	0	100	-24.9869	-54.7826	15.49067
25	50	10	12.14398	0.956066	21.99707
25	50	50	-2.37878	-23.8167	21.99707
25	50	100	-20.5322	-54.7826	21.99707
25	100	10	16.59868	0.956066	28.50347
25	100	50	2.075918	-23.8167	28.50347
25	100	100	-16.0775	-54.7826	28.50347
12	0	10	9.048324	2.756683	16.40815
12	0	50	-5.47444	-22.016	16.40815
12	0	100	-23.6279	-52.9819	16.40815
12	50	10	13.50302	2.756683	22.91455
12	50	50	-1.01974	-22.016	22.91455
12	50	100	-19.1732	-52.9819	22.91455
12	100	10	17.95772	2.756683	29.42095
12	100	50	3.434964	-22.016	29.42095
12	100	100	-14.7185	-52.9819	29.42095
0	0	10	10.30283	4.418791	17.25507
0	0	50	-4.21993	-20.3539	17.25507
0	0	100	-22.3734	-51.3198	17.25507
0	50	10	14.75753	4.418791	23.76147
0	50	10	14.75753	4.418791	23.76147
0	50	100	-17.9187	-51.3198	23.76147
0	100	10	19.21223	4.418791	30.26787
0	100	50	4.689468	-20.3539	30.26787
0	100	100	-13.464	-51.3198	30.26787

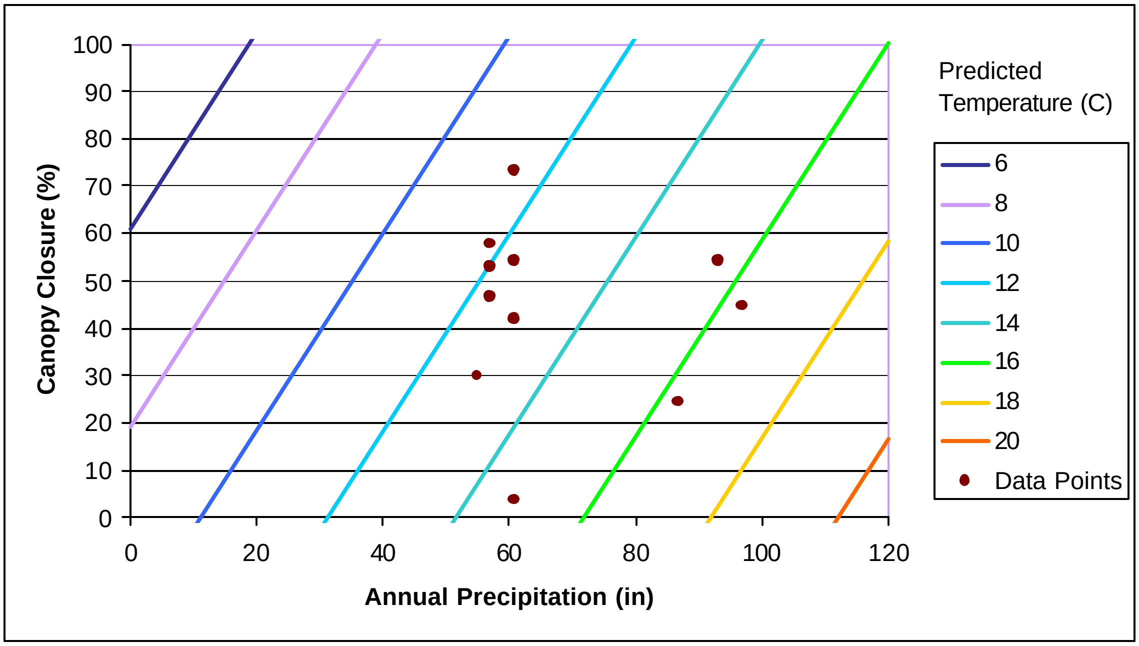
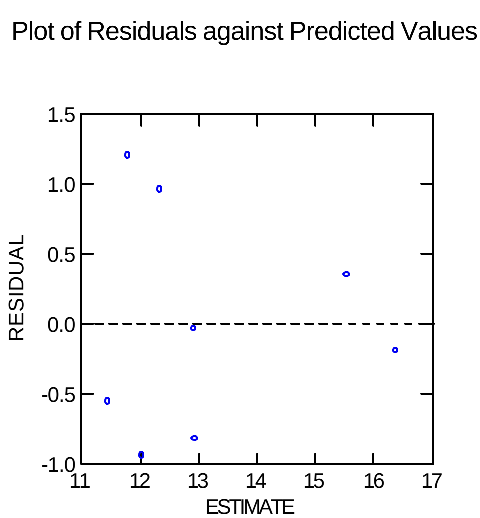


Figure 82. Depiction of predicted stream temperature using regression model developed with canopy closure and annual precipitation as independent variables. Actual data used in the analysis are depicted with black dots. Note, for consistency sake, the depicted regression line was extended beyond the limits of the data. The accuracy of the line beyond the extent of the data is unknown.



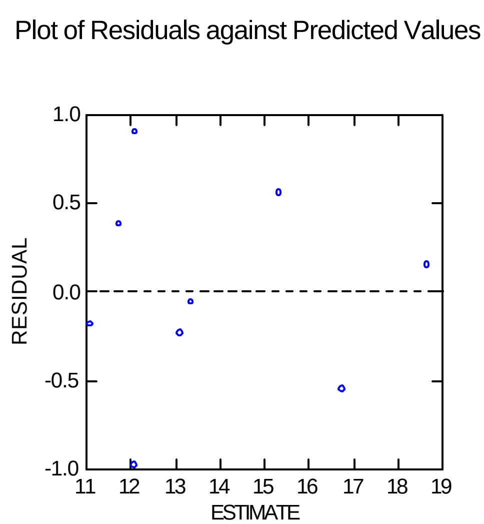


Figure 83. Graphs of residuals for regressions developed for Ecoregion M2442B. The left graph is for the regression that included canopy closure, annual precipitation, and distance from divide. The right graph is for the regression that included canopy closure and annual precipitation. The dataset is too small to make any determinations regarding the distribution of data around the regression line.

M242C(Eastern Cascades Section)

There are 120 data points in Ecoregion M242C. All of these sample locations are in the southern portion of the eastern Cascade Mountains. The regression analysis that included all the GIS generated independent variables identified one outlier. The regression identified canopy closure, basin size, and August air temperature as significant independent variables. A second analysis was conducted without this outlier. The same independent variables were included in the model. The model results of these two regressions are presented in Table 51.

The removal of this outlier did not appreciably change the squared multiple r (Table 51). Correlation coefficients for the two regressions are 0.67 and 0.69 for the models with and without the identified outlier, respectively. Additionally, the predicted temperature with and without this outlier changed by less than 0.5 degrees C (Table 52). Since there was no evidence that the identified sample was substantially different from the rest of the dataset for the ecoregion, the identified outlier was not removed. The residual graph for the regression without the outlier removed is depicted in Figure 82.

Regression analysis were also completed using only canopy closure and basin size, which were the two variables explaining the most variance in the dataset. The results of this analysis are also found on Table 51.

The predicted stream temperatures at various basin sizes and canopy closures are depicted in Figure 83. Predicted stream temperatures range from approximately 14 degrees C to 21 degrees throughout the range of data. The majority of the data points come from basins less than 10,000 acres in size with greater than 30% canopy closure. Therefore, the regression equations developed with this dataset do not represent areas where basin size is greater than 10,000 acres.

August air temperatures represented in the database include only two temperature bands: 50-60 degrees and 60-70 degrees. The effect of the addition of August air temperature in the model is depicted in Figure 84. The additional variable effectively increases the canopy closure required by roughly 30 percent for any given basin size.

Table 51. Summary of results of regression analysis conducted on the dataset for Ecoregion M242C (Eastern Cascades Section) with and without the identified outlier removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	8.434834	1.305402	15.564267	0.02082
Canopy Closure	-0.060807	-0.085151	-0.036463	0.00000
Basin Size	0.000045	0.000025	0.000066	0.00002
August Air Temperature	0.189795	0.07434	0.30525	0.00148
n				120
Multiple R				0.667914
Squared Multiple R				0.446109
Adjusted squared multiple R				0.431784
1 Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	6.694231	-0.093416	13.481878	0.05318
Canopy Closure	-0.057478	-0.08052	-0.034437	0.00000
Basin Size	0.000047	0.000028	0.000066	0.00000
August Air Temperature	0.213539	0.10389	0.323188	0.00019
n				119
Multiple R				0.690884
Squared Multiple R				0.477321
Adjusted squared multiple R				0.463686
No outliers removed, canopy closure and basin size only				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	19.593384	18.023712	21.163057	0.00000
Canopy Closure	-0.055139	-0.079485	-0.030793	0.00002
Basin Size	0.000049	0.000029	0.000069	0.00000
n				120
Multiple R				0.640068
Squared Multiple R				0.409687
Adjusted squared multiple R				0.399509

Table 52. Predicted stream temperature with upper and lower confidence levels computed for models with and without the outlier removed for a range of input variables.

			All Data			With Outlier Removed			
Canopy Closure	***Basin Size***	***Aug Air Temp***	***Predicted Temp.***	***Lower 95%***	***Upper 95%***	***Predicted Temp.***	***Lower 95%***	***Upper 95%***	***Difference in Predicted Temp***
0	0	55	18.87356	5.394102	32.35302	18.43888	5.620534	31.25722	0.434683
100	0	55	12.79286	-3.121	28.70672	12.69108	-2.43147	27.81352	0.101783
0	32000	55	20.31356	6.194102	34.46502	19.94288	6.516534	33.36922	0.370683
100	32000	55	14.23286	-2.321	30.81872	14.19508	-1.53547	29.92552	0.037783
0	0	65	20.77151	6.137502	35.40552	20.57427	6.659434	34.4891	0.197243
100	0	65	14.69081	-2.3776	31.75922	14.82647	-1.39257	31.0454	-0.13566
0	32000	65	22.21151	6.937502	37.51752	22.07827	7.555434	36.6011	0.133243
100	32000	65	16.13081	-1.5776	33.87122	16.33047	-0.49657	33.1574	-0.19966
0	0	55	18.87356	5.394102	32.35302	18.43888	5.620534	31.25722	0.434683
100	0	55	12.79286	-3.121	28.70672	12.69108	-2.43147	27.81352	0.101783
0	32000	55	20.31356	6.194102	34.46502	19.94288	6.516534	33.36922	0.370683
100	32000	55	14.23286	-2.321	30.81872	14.19508	-1.53547	29.92552	0.037783
0	0	65	20.77151	6.137502	35.40552	20.57427	6.659434	34.4891	0.197243
100	0	65	14.69081	-2.3776	31.75922	14.82647	-1.39257	31.0454	-0.13566
0	32000	65	22.21151	6.937502	37.51752	22.07827	7.555434	36.6011	0.133243
100	32000	65	16.13081	-1.5776	33.87122	16.33047	-0.49657	33.1574	-0.19966

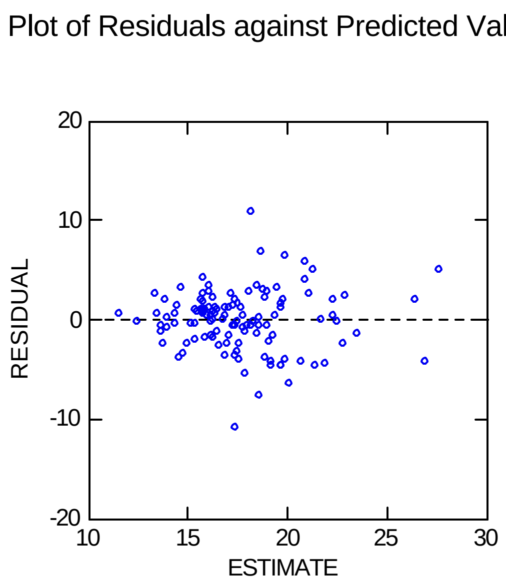


Figure 84. Residuals for the regression developed for Ecoregion M242C (Eastern Cascades Section) including all data. No outliers were removed. The data appears to be well balanced on each side of the regression line.

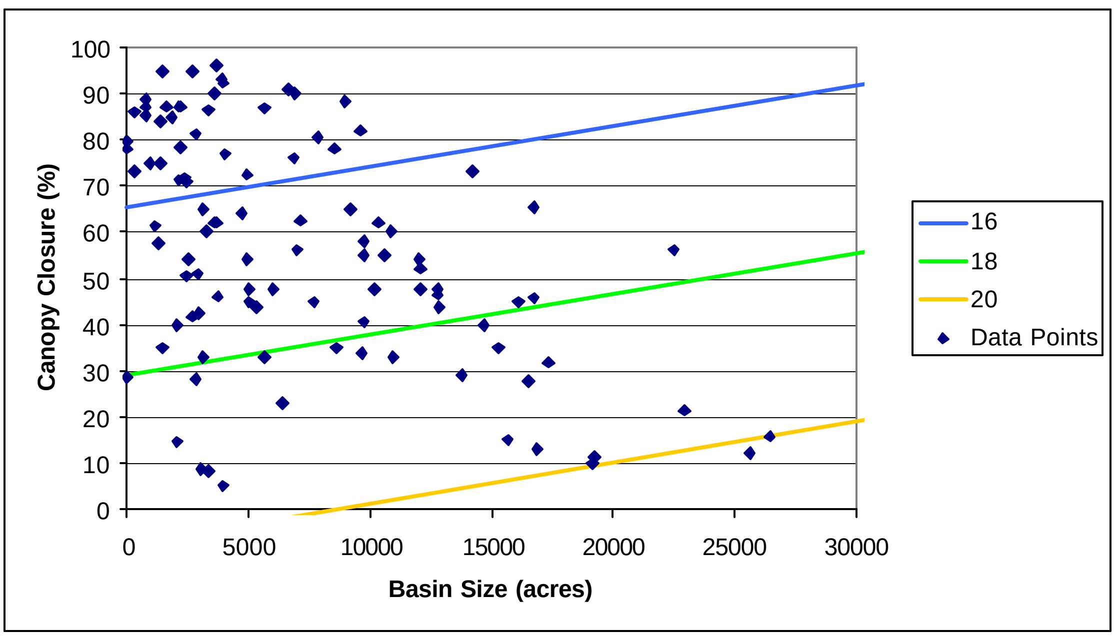


Figure 85. Depiction of predicted stream temperature as a function of canopy closure and basin size for Ecoregion M242C (Eastern Cascades Section).

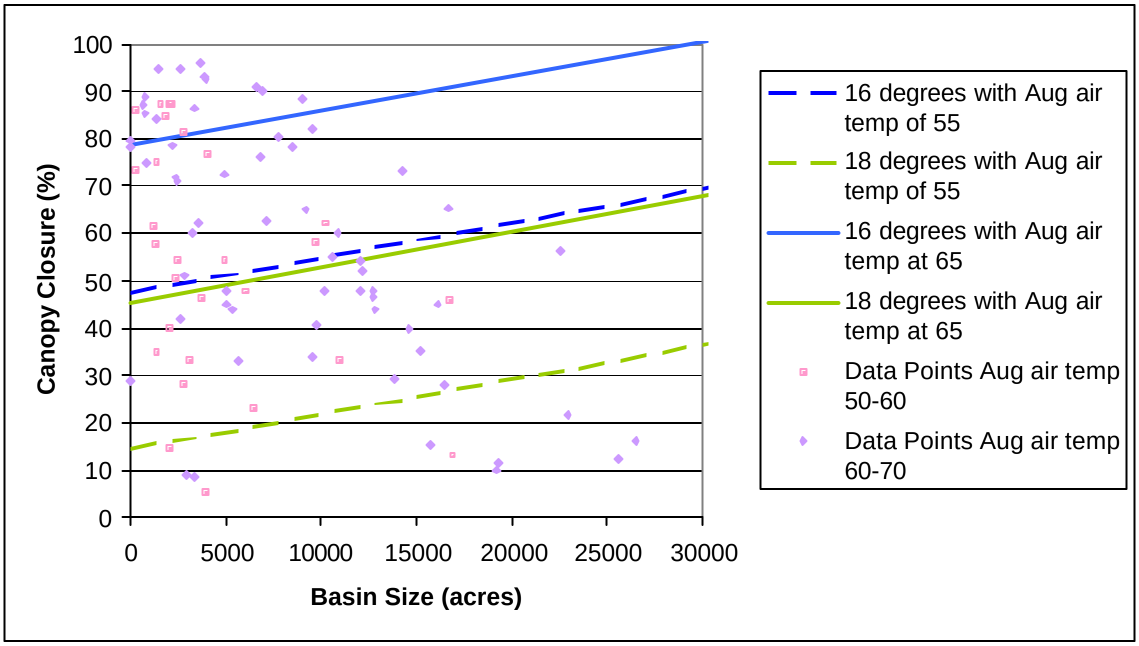


Figure 86. Depiction of the effect of the inclusion of August air temperature in the equation predicting stream temperature for Ecoregion M242C (Eastern Cascades Section).

-342I (Columbia Basin Section)

Ecoregion -342I (Columbia Basin Section) has 25 data points. The regression equation that resulting from a forward stepwise analysis resulted in the inclusion of only one independent variable, average annual air temperature (Table 53). To further explore the relationships in this ecoregion, a forward stepwise regression analysis was conducted that included all variables except average annual air temperature. In this latter analysis, canopy closure and average August air temperature were identified as the only two significant variables (Table 53). The appearance of a second air temperature measure when average annual air temperature was excluded from the analysis suggests climatic patterns are important in this ecoregion.

Two data points have a strong influence on the relationships developed in both analyses (Figures 87 and 88). The stream temperature at these sample sites is unusually warm. We have no basis for excluding this data from the analyses; however, the data does appear to be somewhat questionable. In the absence of these data points, August air temperature may not have been significant in the equation.

At present, there are only 25 points in this dataset. Substantially different results of the analyses may be found for this ecoregion if the sample size was increased. Residuals for the two regression equations are provided in Figure 89.

Table 53. Results of regression analyses for Ecoregion -342I (Columbia Basin Section).

All Variables Included				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000054
Constant	-16.585662	-32.338764	-0.832559	0.03991
Ave Annual Air Temp	0.842753	0.490028	1.195479	0.00005
n				25
Multiple R				0.717679
Squared Multiple R				0.515064
Adjusted squared multiple R				0.49398
All Variables Included Except Average Annual Air Temperature				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.000166
Constant	*-0.574219*	*-17.648225*	*16.499787*	*0.94503*
August Air Temperature	0.375167	0.138853	0.611482	0.00332
Canopy Closure	-0.072665	-0.139983	-0.005347	0.03562
n				25
Multiple R				0.739424
Squared Multiple R				0.546748
Adjusted squared multiple R				0.505543

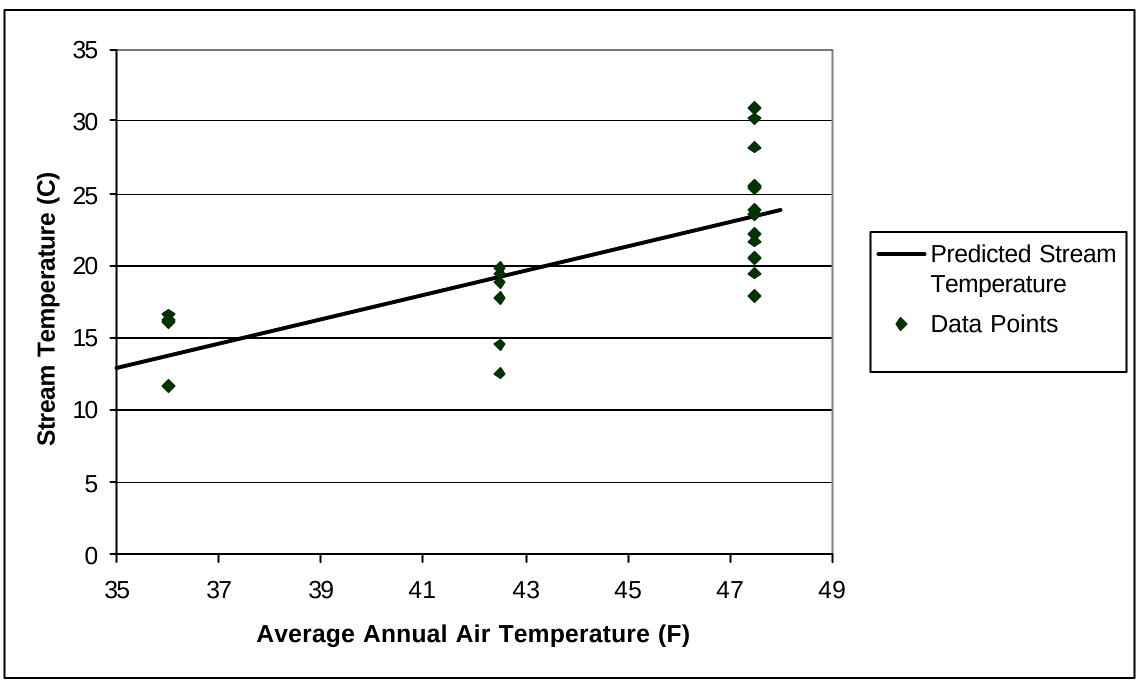


Figure 87. Depiction of regression equation developed for Ecoregion -342I (Columbia Basin Section). Average annual air temperature was the only significant independent variable included in equation. Dots correspond to data points representing the ecoregion.

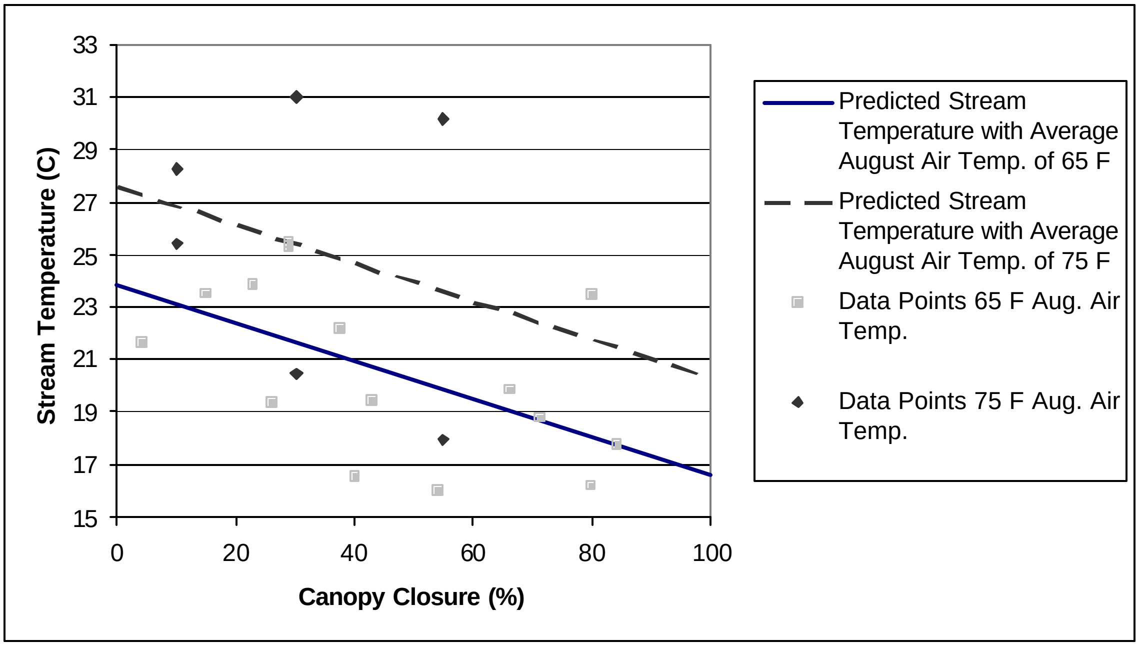


Figure 88. Depiction of regression equation developed for Ecoregion -342I (Columbia Basin Section) using all variables except average annual air temperature. Canopy closure and average August air temperature were the only significant variables found in this analysis. There are only two values for average August air temperature in the dataset for this ecoregion. Lines on graph represent each of these two values. Dot correspond to the data points representing ecoregion.

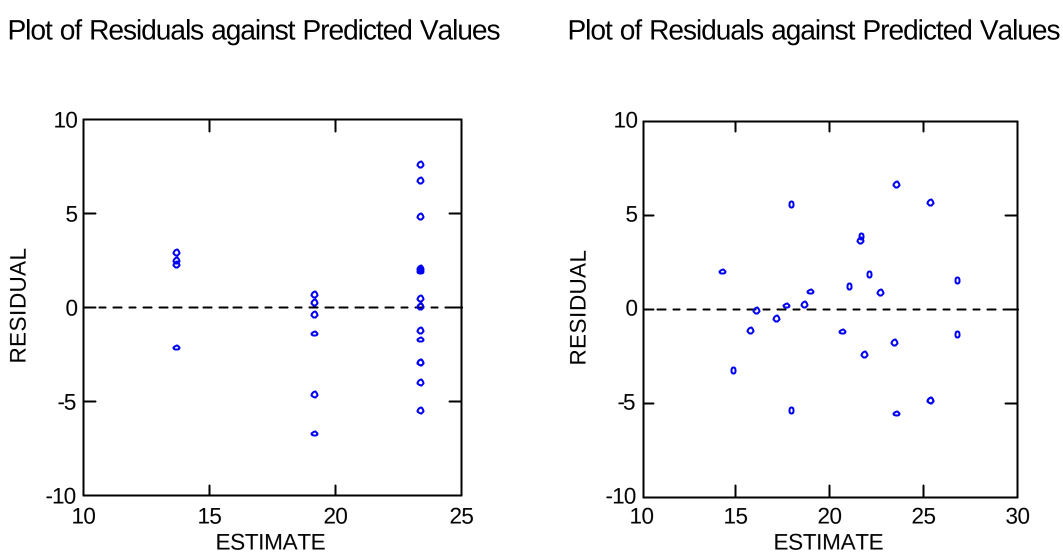


Figure 89. Plots of residuals for the regressions developed for Ecoregion -342I (Columbia Basin Section). The plot on the left corresponds to the regression that included only average annual air temperature. The plot on the right corresponds to the regression that included average August air temperature and canopy closure. The residuals do not indicate any need for data transformation.

3.5.3 By Lithology

Alluvium

The alluvium dataset includes points that are primarily from the southern portion of the eastern Cascade Mountains, but also includes points from all other areas of eastern Washington.

Two outliers were identified during the regression analysis of all GIS data points for the alluvium lithology group. These data points were removed and the resulting equations compared to determine if the removal of outliers was appropriate. The squared multiple r increased from 0.508483 to 0.669431 with the removal of the two identified data points (Table 54). The removal of the outliers increased the portion of the variance in the data explained by the regression equation by roughly 16%. This increase was deemed sufficient to support the removal of the two cases.

Canopy closure was not included as a significant variable in either equation. Within the alluvium lithology group data set, canopy closure ranges from 5 to 90%. Therefore, the exclusion of canopy closure from the equations was not due to the absence of variation within the variable in the dataset. The variables that were found to be significant in the equation with the outliers removed include basin size (acres), latitude, and average July air temperature (F).

There are only three values for average July air temperature in the data set. Figures depicting the regression model with the outliers removed were developed for each of these three air temperature values (Figures 90, 91, and 92). Interestingly, the input data with average July air temperatures of 55 degrees F were primarily found around latitude 47. Average July air temperatures of 65 degrees F were found primarily at latitudes 46 and 48 and Average July air temperatures of 75 were found exclusively at latitude 49. Given this segregation of data within the variables included in the regression, a separate regression was developed excluding latitude (Table 55). The regression that included only basin size and average July air temperature was nearly as robust as the one that also included latitude. A depiction of this equation is provided in Figure 93. Residuals graphs for regression summarized in Table 54 (with the outliers removed) and the regression summarized in Table 55 are provided in Figure 94.

Table 54. Results of regression analyses for the alluvium lithology group.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	91.61093	39.001748	144.220101	0.00089
Basin Size	0.000056	0.000038	0.000075	0
Latitude	-1.51413	-2.615506	-0.412759	0.00775
Annual Precip.	-0.06513	-0.113426	-0.01683	0.00894
n				74
Multiple R				0.71308
Squared Multiple R				0.508483
Adjusted squared multiple R				0.487418
2 Outliers Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	60.14713	19.490305	100.803949	0.00433
Basin Size	0.000076	0.00006	0.000091	0
Latitude	-1.18726	-2.046687	-0.327826	0.00749
Ave. Jul. Air Temp.	0.210184	0.100235	0.320133	0.0003
n				72
Multiple R				0.818188
Squared Multiple R				0.669431
Adjusted squared multiple R				0.654847

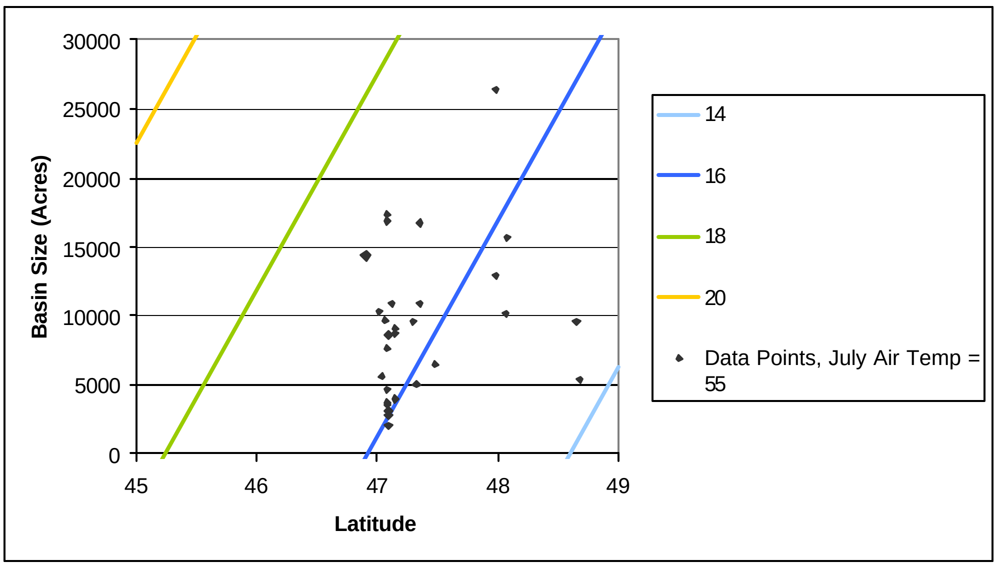


Figure 90. Depiction of the results of the regression equation developed with the data set represented by the alluvium lithology group. Lines represent the equation solved with average July air temperature held at 55 degrees F. Note, for consistency sake, the depicted regression line was extended beyond the limits of the data. The accuracy of the line beyond the extent of the data is unknown.

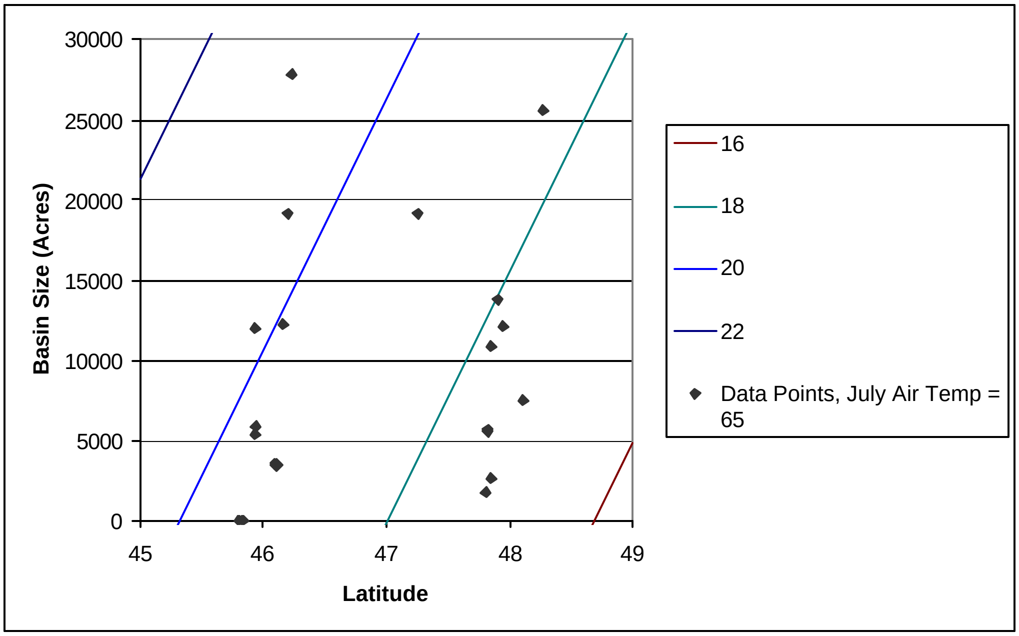


Figure 91. Depiction of the results of the regression equation developed with the data set represented by the alluvium lithology group. Lines represent the equation solved with average July air temperature held at 65 degrees F.

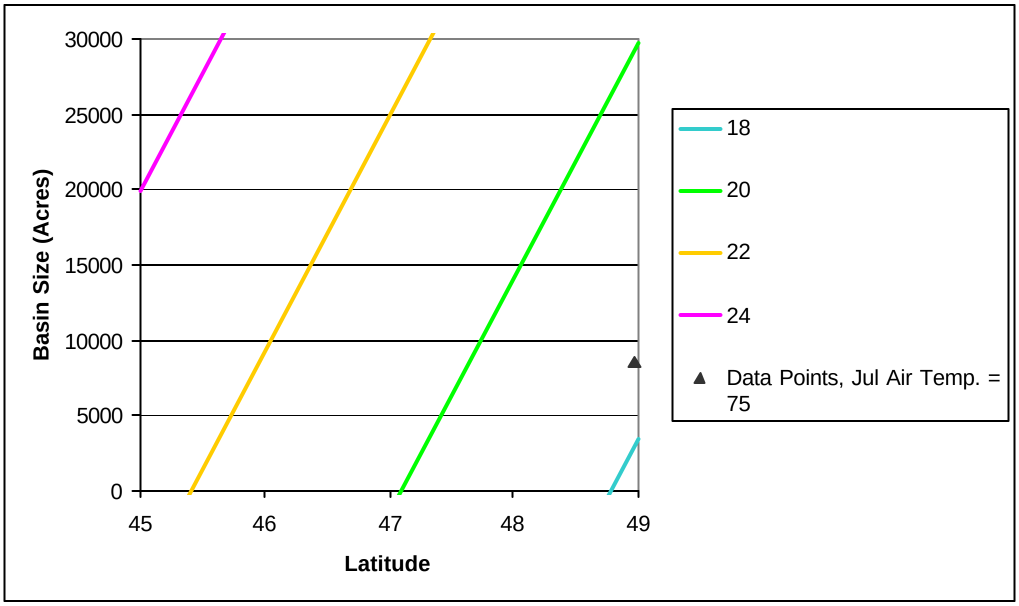


Figure 92. Depiction of the results of the regression equation developed with the data set represented by the alluvium lithology group. Lines represent the equation solved with average July air temperature held at 75 degrees F.

Table 55. Results of regression analysis for the alluvium lithology group with latitude excluded and outliers removed.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	4.726126	-2.176758	100.803949	0.17642
Basin Size	0.000072	0.000056	0.000091	0
Ave. Jul. Air Temp.	0.200204	0.085396	0.320133	0.00088
n				72
Multiple R				0.795292
Squared Multiple R				0.63249
Adjusted squared multiple R				0.621837

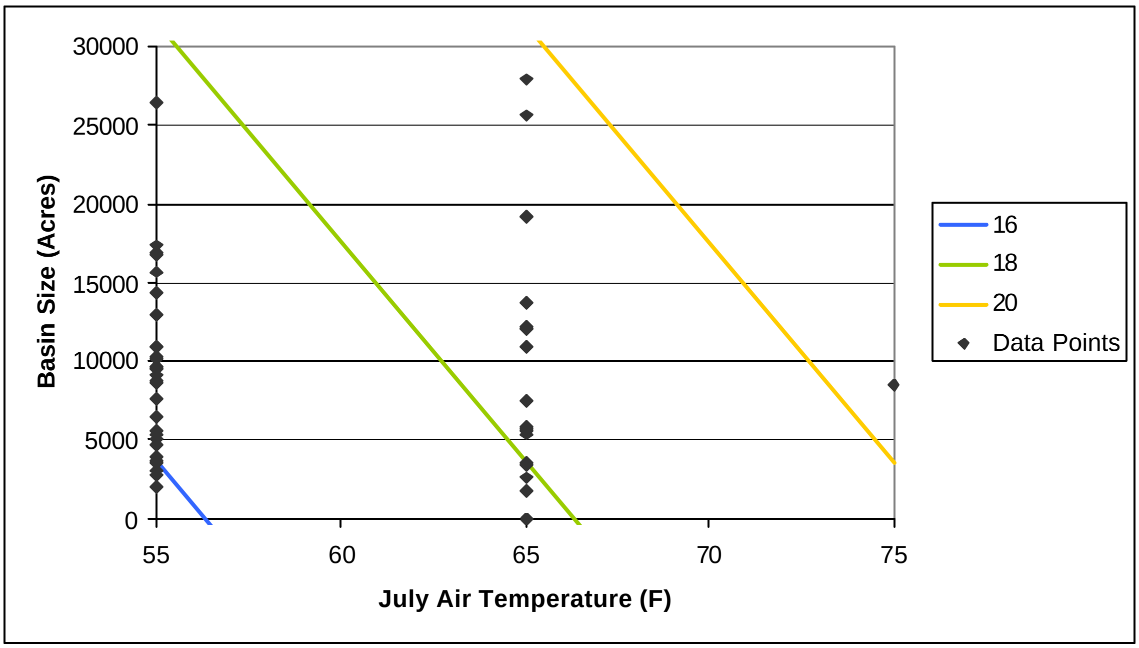


Figure 93. Depiction of the results of the regression analysis for the alluvium lithology group excluding outliers and the independent variable, latitude.

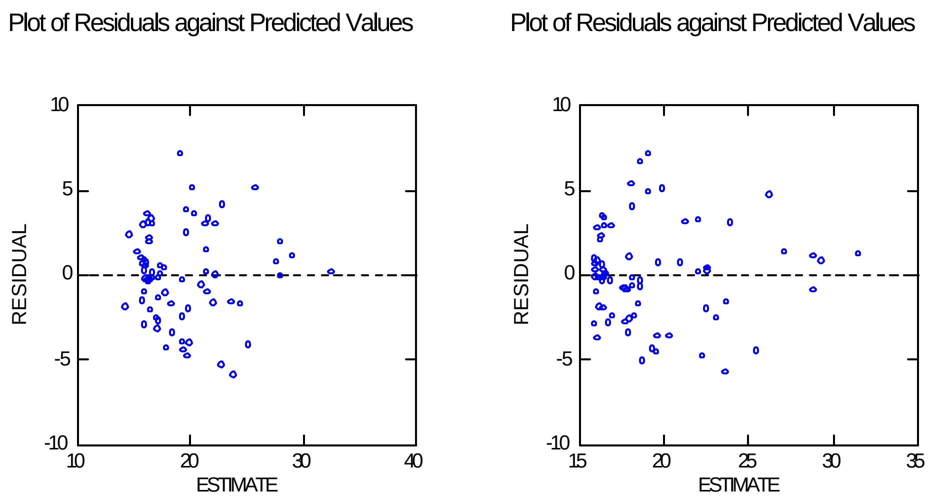


Figure 94. Residuals for regressions using database for the alluvium lithology group. The left plot is for the regression summarized on Table 54 with the outliers removed. The right plot is for the regression summarized on Table 55. The data appears to be well balanced on either side of the regression line.

Basalt

All but six of the 47 cases in the basalt database are from the southern portion of the eastern Cascade Mountains. The regression analysis including all possible variables found no outliers. The significant variables included in the equation were canopy closure, elevation, and average August air temperature (Table 56).

The relationship between canopy closure and elevation was discussed in an earlier section. The addition of average August air temperature adds a new dimension to the relationship. As average August air temperature increases, the amount of canopy closure needed to meet specific temperature targets increases (Figures 95, 96, and 97). The residuals for the equation are provided in Figure 98.

Table 56. Results of regression analysis for the basalt lithology group.

All Data				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	10.72126	-1.203616	22.646138	0.07679
Elevation	-0.00125	-0.002241	-0.000259	0.01464
Canopy Closure	-0.09041	-0.116953	-0.063857	0.00000
Average August Air Temperature	0.24414	0.081875	0.406405	0.00408
n				47
Multiple R				0.83159
Squared Multiple R				0.691541
Adjusted squared multiple R				0.670021

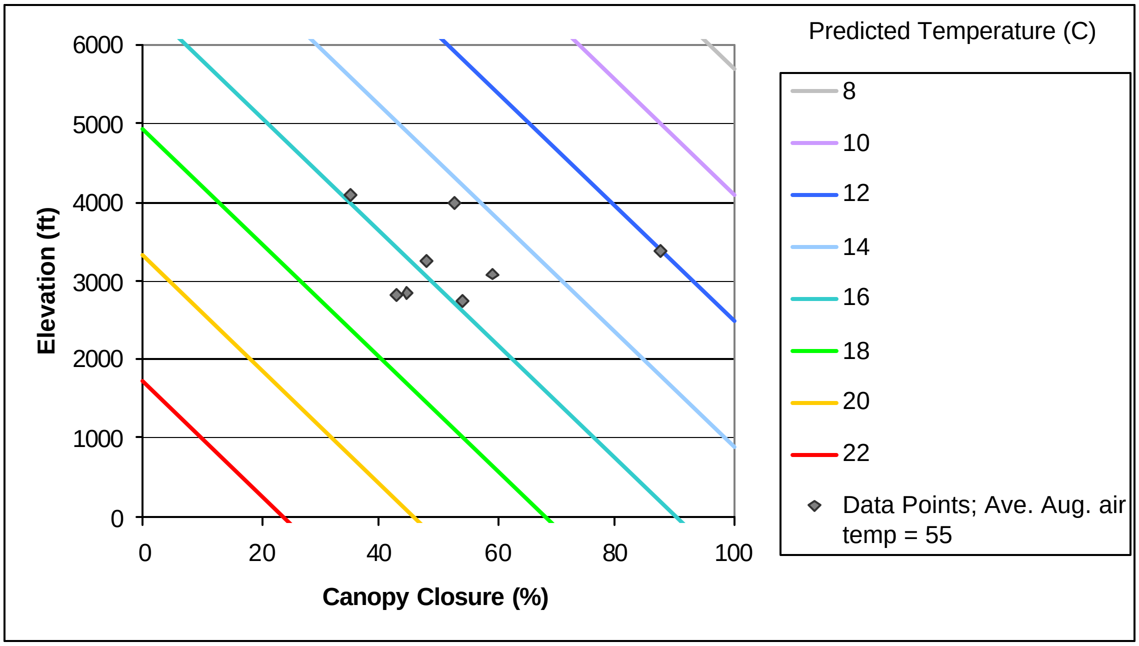


Figure 95. Depiction of the relationship between canopy closure and elevation developed using the basalt lithology group dataset for areas with average August air temperature of 55. Note, for consistency sake, the depicted regression line was extended beyond the limits of the data. The accuracy of the line beyond the extent of the data is unknown.

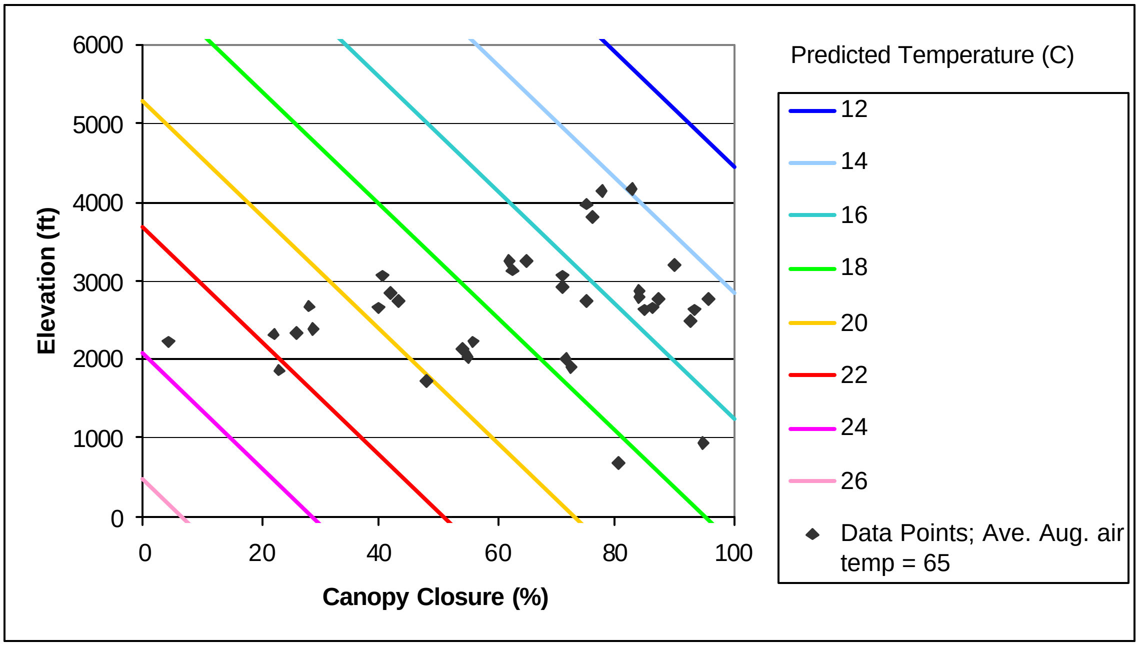


Figure 96. Depiction of the relationship between canopy closure and elevation developed using the basalt lithology group dataset for areas with average August air temperature of 65.

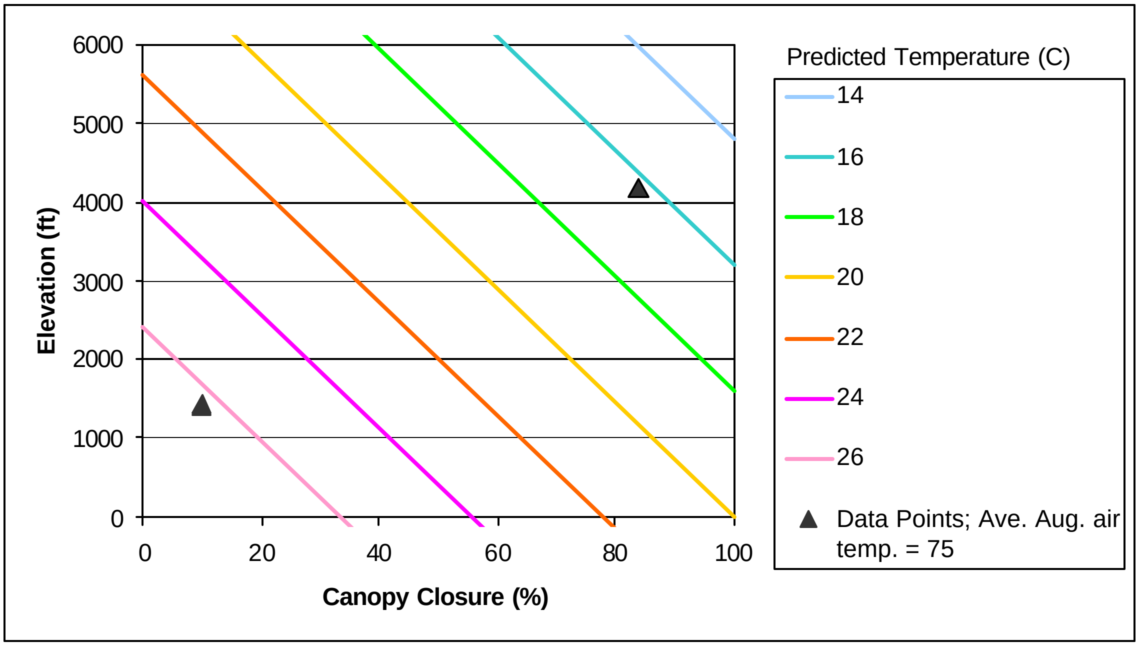


Figure 97. Depiction of the relationship between canopy closure and elevation developed using the basalt lithology group dataset for areas with average August air temperature of 65.

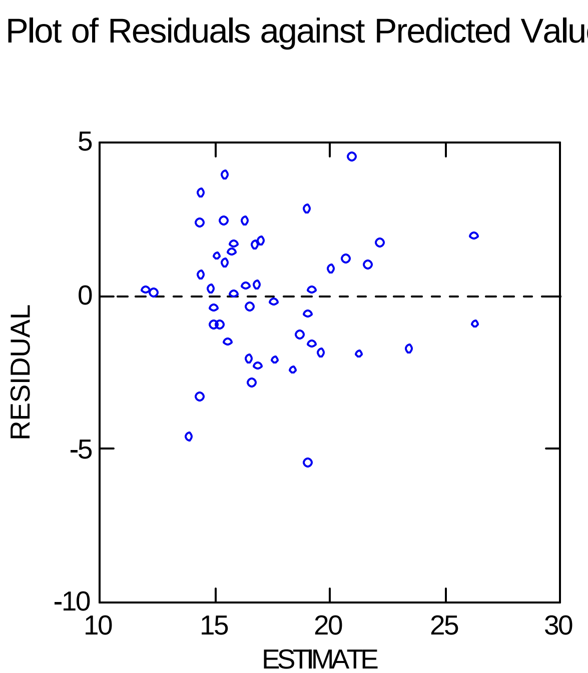


Figure 98. Graph of the residuals for the regression equation developed for the basalt lithology group. No patterns of concern are apparent in the residuals.

Glacial Drift

The regression with all GIS generated variables included for the dataset represented by the glacial drift lithology group identified five significant independent variables. These are canopy closure, basin size, average annual precipitation, average July air temperature, and average August air temperature (Table 57). The predicted temperatures for various combinations of the five independent variables are presented in Table 58. The inclusion of both average July air temperature and average August air temperature was somewhat problematic. The two variables tend to be highly correlated; hence, they violate assumptions of independence. Additionally, there was very little variability in the data within these two variables (Table 59). Hence, the removal of one of the two air temperature variables was warranted. As a further complication, close inspection of Table 58 reveals that predicted stream temperature tends to decrease with increasing air temperature. This does not make physical sense. Air temperature affects stream temperature in a positive direction. Therefore, the five variable model is spurious.

Regression models were developed using four independent variables, excluding either average July air temperature or August air temperature. In each case, predicted stream temperature decreased with increasing air temperature. Again, this does not make physical sense. These equations were also considered spurious.

Once the air temperature variables were excluded from the model, a model with the three remaining independent variables (canopy closure, basin size, and annual precipitation) was developed (Table 60). One outlier was identified. The predicted stream temperature with and without this outlier was compared (Table 61). The predicted stream temperature varied by 1 to 2 degrees with and without the outlier. The residuals of the regressions were also inspected. The outlier lies substantially outside the range of the rest of the dataset (Figure 99). Therefore, the model without the outlier was determined to be the preferred model.

The predicted stream temperature estimated using the 3 variable model with the outlier removed was plotted for 16 and 18 degree C target temperatures and a range of average annual precipitation values (Figure 100). All three of the variables in the model have a substantial effect on predicted temperature.

Table 57. Results of regression analyses for the glacial drift lithology group.

All Significant Variables				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	38.97484	13.589756	64.359915	0.00346
Canopy Closure	-0.04988	-0.078618	-0.021141	0.00111
Basin Size	0.000099	0.000045	0.000154	0.00067
Annual Precipitation	-0.1412	-0.240142	-0.042263	0.00623
Ave. July Air Temperature	-0.63702	-1.026294	-0.247737	0.00196
Ave. August Air Temperature	0.380275	0.112206	0.648344	0.00652
n				48
Multiple R				0.848257
Squared Multiple R				0.71954
Adjusted squared multiple R				0.686152
4 Variable Model				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	48.28456	21.808295	74.760828	0.00065
Canopy Closure	-0.05279	-0.083744	-0.021845	0.0013
Basin Size	0.00009	0.000031	0.000148	0.00344
Ave. July Air Temperature	-0.3897	-0.765496	-0.013909	0.04244
Annual Precipitation	-0.1629	-0.268443	-0.057365	0.00329
n				48
Multiple R				0.815361
Squared Multiple R				0.664813
Adjusted squared multiple R				0.633633

Table 58. Predicted temperature for a range of values for the five independent variables included in the original regression model for the glacial drift lithology group.

Canopy Closure (%)	Basin Size (acres)	Annual Precipitation (in)	Ave. July Air Temperature (F)	Ave. August Air Temperature (F)	Predicted Stream Temperature (C)	Lower 95%	Upper 95%
0	0	10	55	55	23.44211	-39.0865	85.97067
0	0	10	55	65	27.24486	-37.9644	92.45411
0	0	10	65	55	17.07196	-49.3494	83.4933
0	0	10	65	65	20.87471	-48.2274	89.97674
0	0	80	55	55	13.5579	-55.8964	83.01226
0	0	80	55	65	17.36065	-54.7744	89.4957
0	0	80	65	55	7.187745	-66.1594	80.53489
0	0	80	65	65	10.9905	-65.0373	87.01833
0	36000	10	55	55	27.00611	-37.4665	91.51467
0	36000	10	55	65	30.80886	-36.3444	97.99811
0	36000	10	65	55	20.63596	-47.7294	89.0373
0	36000	10	65	65	24.43871	-46.6074	95.52074
0	36000	80	55	55	17.1219	-54.2764	88.55626
0	36000	80	55	65	20.92465	-53.1544	95.0397
0	36000	80	65	55	10.75175	-64.5394	86.07889
0	36000	80	65	65	14.5545	-63.4173	92.56233
100	0	10	55	55	18.45421	-46.9483	83.85657
100	0	10	55	65	22.25696	-45.8262	90.34001
100	0	10	65	55	12.08406	-57.2112	81.3792
100	0	10	65	65	15.88681	-56.0892	87.86264
100	0	80	55	55	8.569995	-63.7582	80.89816
100	0	80	55	65	12.37275	-62.6362	87.3816
100	0	80	65	55	2.199845	-74.0212	78.42079
100	0	80	65	65	6.002595	-72.8991	84.90423
100	36000	10	55	55	22.01821	-45.3283	89.40057
100	36000	10	55	65	25.82096	-44.2062	95.88401
100	36000	10	65	55	15.64806	-55.5912	86.9232
100	36000	10	65	65	19.45081	-54.4692	93.40664
100	36000	80	55	55	12.134	-62.1382	86.44216
100	36000	80	55	65	15.93675	-61.0162	92.9256
100	36000	80	65	55	5.763845	-72.4012	83.96479
100	36000	80	65	65	9.566595	-71.2791	90.44823

Table 59. Sample size of July and August average air temperatures for the glacial drift lithology group

	July Air Temperature		
August Air Temperature	**55**	**65**	**Total**
55	6	3	**9**
65	1	38	**39**
Total	**7**	**41**	**48**

Table 60. Results of regression models with three independent variables with and without the outlier; glacial drift lithology group.

3 Variable Model				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	21.07855	17.380608	24.776486	0
Canopy Closure	-0.06304	-0.093452	-0.032623	0.00014
Basin Size	0.000108	0.000051	0.000166	0.00047
Annual Precipitation	-0.07752	-0.145983	-0.009057	0.02738
n				48
Multiple R				0.794179
Squared Multiple R				0.63072
Adjusted squared multiple R				0.605542
3 Variable Model Without Outlier				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	19.96385	16.750934	23.176755	0
Canopy Closure	-0.04568	-0.073067	-0.018295	0.00162
Basin Size	0.000124	0.000074	0.000174	0.00001
Annual Precipitation	-0.08803	-0.146881	-0.029183	0.00428
n				47
Multiple R				0.833746
Squared Multiple R				0.695132

Table 61. Comparison of predicted stream temperatures for the glacial drift lithology estimated using the three independent variable models with and without the outlier removed.

			Outlier in	***Outlier Removed***		
Canopy Closure	**Basin Size**	**Annual Precipitation**	**Predicted Temperature**	**Predicted Temperature**	**Lower 95%**	**Upper 95%**
0	1000	15	20.02375	18.76737	-0.26375	22.91301
0	30000	15	23.15575	22.36337	4.782255	27.95901
100	1000	15	13.71995	14.19927	-42.6656	22.91301
100	30000	15	16.85195	17.79527	-37.6196	27.95901
0	1000	70	15.76015	13.92561	-1.86881	21.30795
0	30000	70	18.89215	17.52161	3.17719	26.35395
100	1000	70	9.456347	9.357505	-44.2707	21.30795
100	30000	70	12.58835	12.95351	-39.2247	26.35395

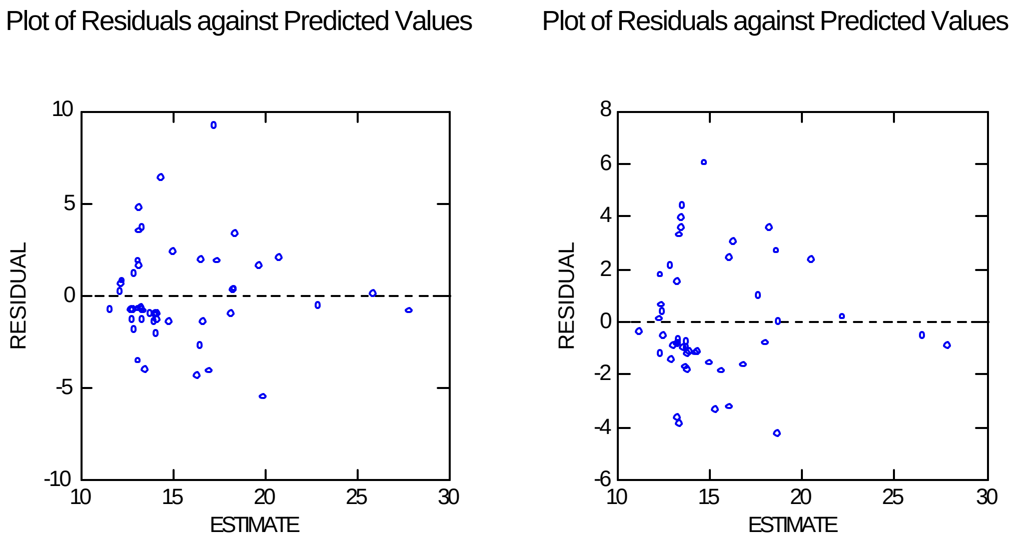


Figure 99. Residuals of models developed using the database for the glacial drift lithology group and including three independent variables. The plot on the left is the residuals for the model with the identified outlier left in and the plot on the right depicts the residuals for the same model without the outlier. No patterns of concern are apparent in the residuals; however the three data points on the right may have greater influence on the resulting equations than the other points in the dataset.

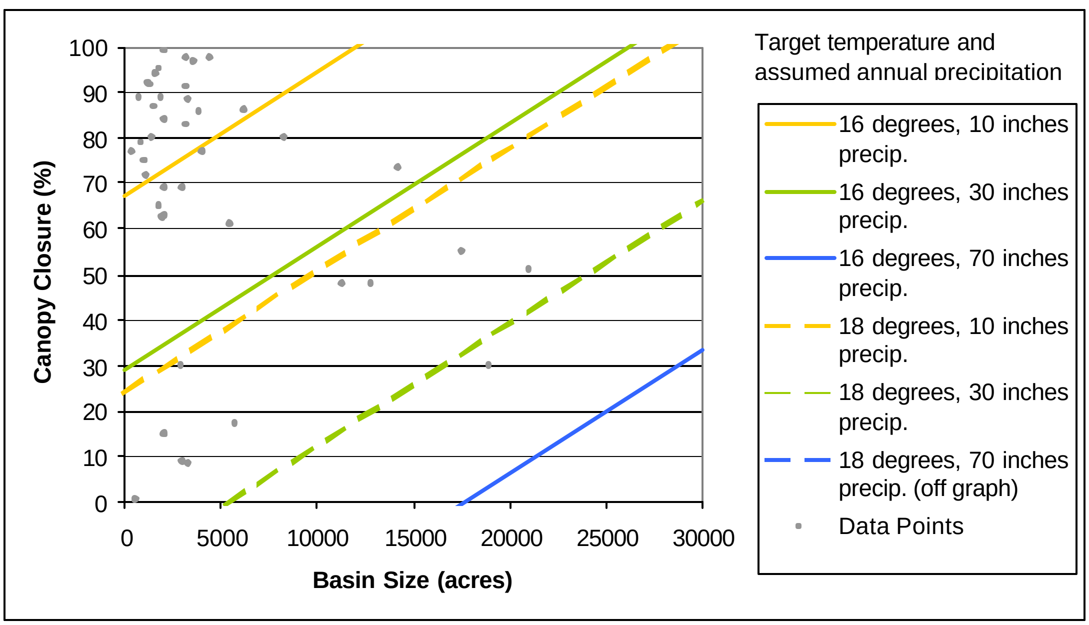


Figure 100. Depiction of the relationship between average annual precipitation, basin size, and canopy closure developed for targeted stream temperature of 16 and 18 degrees C. The solid lines represent a 16-degree target and the dotted lines represent an 18-degree target. Colors are constant for both temperature targets with the same precipitation levels. Gray dots represent the input data set. The accuracy of the model beyond the range of the data is unknown.

Granites

The sample site for the granite lithology group is very small (n=13). The regression analysis that included all potential independent variables resulted in a four variable model that included basin size, stream gradient, hill slope gradient, and distance from divide (Table 62). An outlier was identified. The database is too small to reliably determine if that identified case was truly an outlier, so the case was not removed from the analyses. Two of the four parameters in the model did not make physical sense. These are hill slope gradient and stream gradient. In both cases, the model indicated that stream temperature increases with gradient. The reverse should be true. Therefore, the inclusion of these two parameters was assumed spurious.

Table 62. Results of regression analyses for the granite lithology group.

Full model with all independent variables included, no outliers removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00006
Constant	7.31023	3.728593	10.891875	0.00153
Basin Size	0.00015	0.0001	0.000209	0.00018
Stream Gradient	35.6745	18.51423	52.834827	0.00137
Hill slope Gradient	5.76574	1.343084	10.188396	0.01691
Distance from Divide	0.86144	0.516923	1.205965	0.00042
n				13
Multiple R				0.96987
Squared Multiple R				0.940648
Adjusted squared multiple R				0.910972
Model with stream gradient and hill slope gradient excluded, no outliers removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00047
Constant	15.6385	14.287216	16.989759	0
Basin Size	0.00018	0.0001	0.000261	0.00047
n				13
Multiple R				0.828567
Squared Multiple R				0.686524
Adjusted squared multiple R				0.658026

The second analysis completed included all possible independent variables except hill slope and stream gradient. The resulting model contained only one independent variable that was significant at the 0.05 probability level. This was basin size. Canopy closure was forced into the model to determine if it improved that model. It did not result in any significant improvement. Therefore, the best-fit model was determined to be the model that predicted stream temperature as a function of basin size alone (Table 62). The modeled relationship between stream temperature and basin size is depicted in Figure 101 and the residuals of the model are depicted in Figure 102.

Note that the entire dataset for the granite lithology group was collected in relatively small basins (Figure101). Also note that most of the data for this lithology group comes from NE Washington. This is an area where low levels of canopy closure and low elevation sites were also poorly represented. With an expanded dataset representing a larger range of basin sizes, canopy closures, and other conditions, the relationships that are presented here may be found to be spurious.

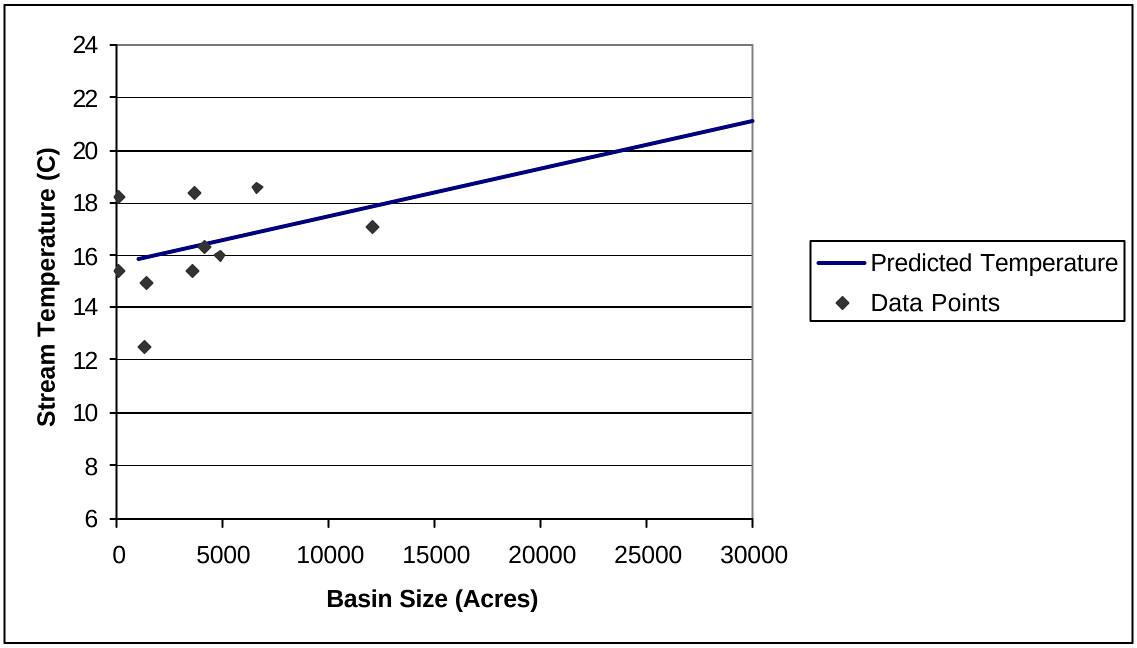


Figure 101. Depiction of the relationship between basin size and predicted stream temperature developed through regression analysis for the granite lithology group. The dots represent the data points used in the analysis. The accuracy of the model beyond the range of the data is unknown.

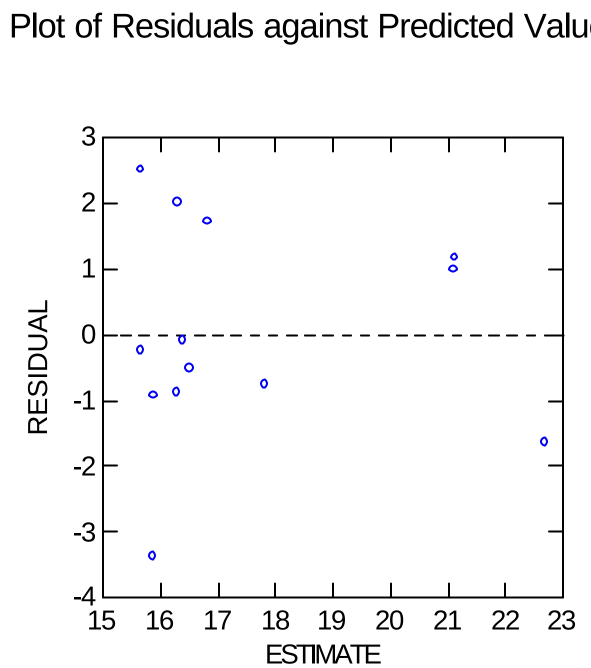


Figure 102. Residuals for the regression analysis incorporating only basin size as the independent variable. Analysis was completed using the granite lithology group database. No patterns of concern are apparent in the residuals. Note, the sample size is small.

Metamorphic

The database for the metamorphic lithology group has only 19 sample points. These come from NE Washington and the southern portion of the eastern Cascade Mountains.

A regression analysis was conducted that included all the GIS generated independent variables. Distance to divide was the only significant variable identified (Table 63). The squared multiple r was low (0.198054) suggesting a poor fit to the data. One outlier was identified. This outlier was removed to determine if the model fit would improve. With the outlier removed, a model with four independent variables was developed (Table 63). These variables included elevation, canopy closure, distance to divide, and average annual air temperature. The database was inspected to identify the range of these variables represented in the data. Canopy closure, distance to divide, and elevation were reasonably represented; however, average annual air temperature was poorly distributed. Fourteen of the eighteen samples had the same average annual air temperature (Table 64). Given the poor distribution of samples across this variable and the small overall sample size, this variable was eliminated from the equation.

A regression analysis was completed with the remaining variables. In this analysis, canopy closure and elevation were no longer significant at the 0.05 probability level. These, too, were eliminated, leaving only distance from divide, which was also the only variable in the original equation. The elimination of the outlier, however, improved the fit to the data. The relationship between distance to divide and predicted stream temperature is depicted in Figure 103 and the residuals for that analysis are depicted in Figure 104.

Table 63. Results of regression analyses for the metamorphic lithology group.

Model with all Independent Variables Considered				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.05622
Constant	12.97096	9.222863	16.719051	0.00000
Distance to Divide	0.79381	-0.023559	1.611179	*0.05622*
n				19
Multiple R				0.445032
Squared Multiple R				0.198054
Adjusted squared multiple R				0.15088
Model with all Independent Variables Considered, One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00001
Constant	-33.1774	-55.596767	-10.758067	0.00701
Elevation	0.002316	0.000574	0.004058	0.01308
Canopy Closure	-0.13205	-0.188386	-0.075716	0.00022
Distance to Divide	0.8488	0.483722	1.213877	0.00023
Ave. Annual Air Temperature	1.115073	0.646096	1.58405	0.00019
n				18
Multiple R				0.940988
Squared Multiple R				0.885458
Adjusted squared multiple R				0.850214
Model with Annual Air Temperature Eliminated From Equation, One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00159
Constant	18.70717	10.144157	27.270182	0.00035
Canopy Closure	-0.05344	-0.129387	0.022515	***0.15354***
Distance to Divide	0.882362	0.2749	1.489823	0.0076
Elevation	-0.00093	-0.002732	0.000863	***0.28375***
n				18
Multiple R				0.808072
Squared Multiple R				0.652980
Adjusted squared multiple R				0.578619
Model with Distance From Divide, One Outlier Removed				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00044
Constant	10.84941	8.248078	13.450741	0.00000
Distance to Divide	1.149515	0.596519	1.70251	0.00044
n				18
Multiple R				0.740446
Squared Multiple R				0.548261
Adjusted squared multiple R				0.520027

Table 64. Sample sizes for average annual air temperature in the database represented by the metamorphic lithology group.

Average Annual Air Temperature (F)	Number of Samples
42.5	14
47.5	3
36	1

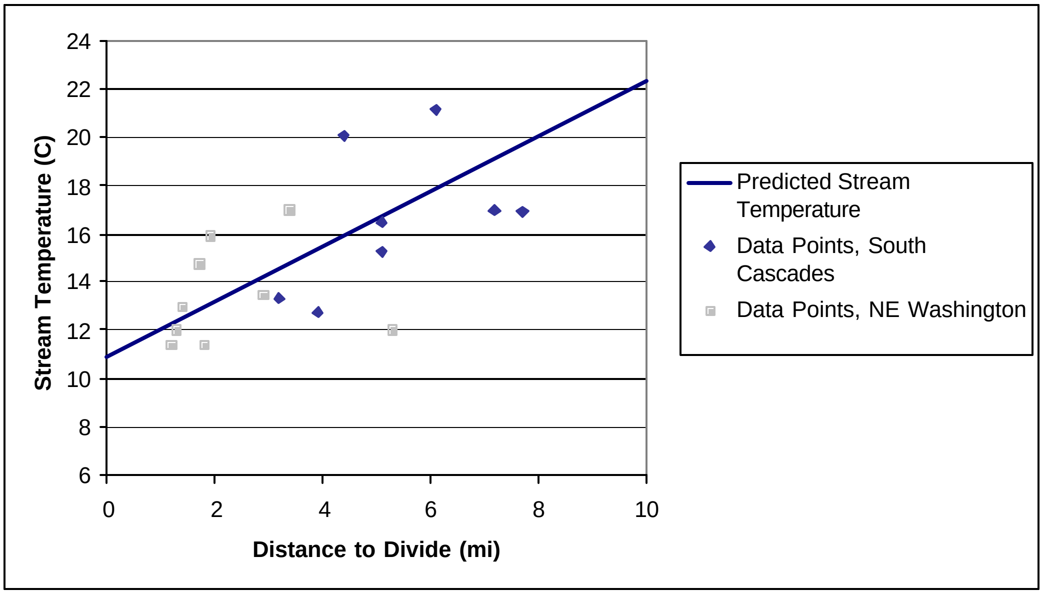


Figure 103. Depiction of the relationship between distance to divide and predicted stream temperature arising from the regression model developed for the metamorphic lithology group.

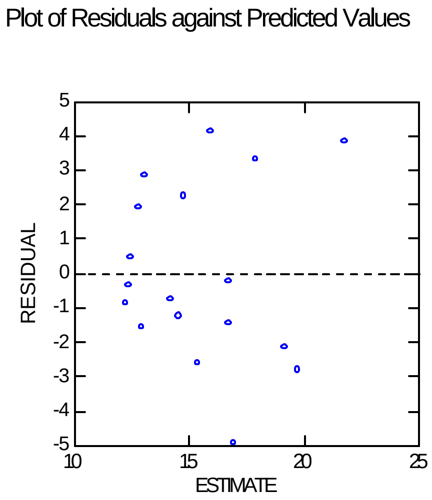


Figure 104. Plot of residuals for the regression including only distance from divide with one outlier removed based upon the database for the metamorphic lithology group. No patterns of concern are apparent in the residuals.

Other Deposits

A regression analysis was conducted for the database representing other deposits. All GIS generated variables were evaluated. Average July air temperature was the only significant independent variable identified (Table 65). The relationship between average July air temperature and predicted stream temperature is depicted in Figure 105. The residuals for the analysis are depicted in Figure 106.

Table 65. Results of regression analyses for the other deposits lithology group.

Model with All Significant Independent Variables				
Variable	*Coefficient*	*Lower 95%*	*Upper 95%*	*P*
Regression				0.00352
Constant	-5.21308	-19.062374	8.63622	0.44714
Ave. July Air Temperature	0.348077	0.124342	0.571812	0.00352
n				30
Multiple R				0.515914
Squared Multiple R				0.266167
Adjusted squared multiple R				0.239959

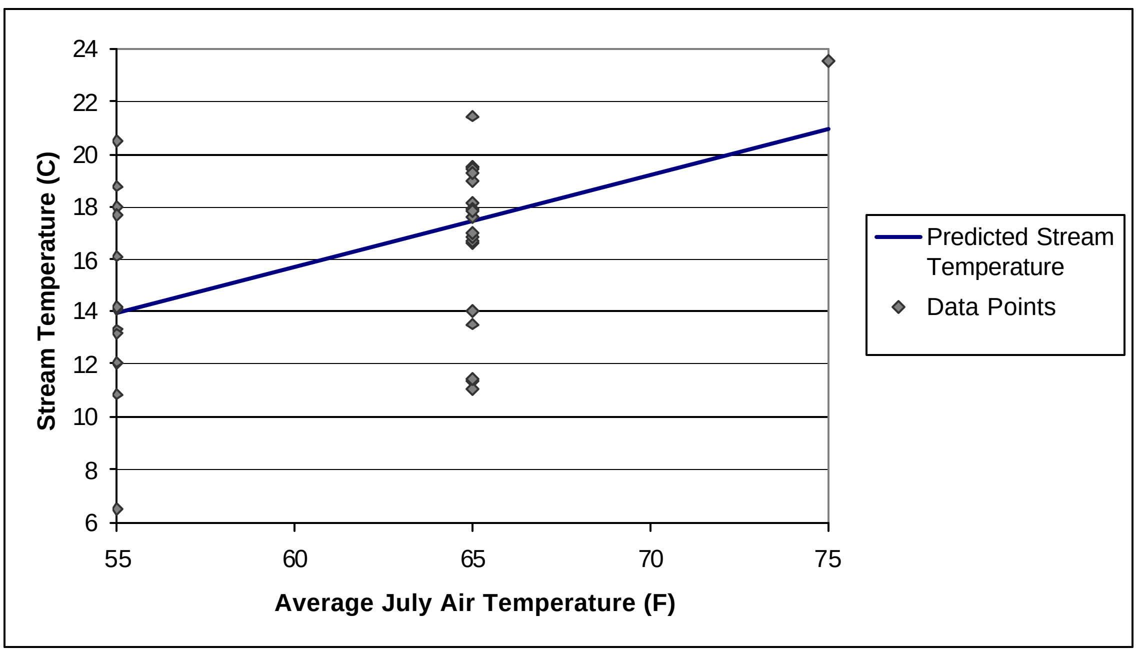


Figure 105. Depiction of the relationship between average July air temperature and stream temperature developed through regression on the other deposit database.

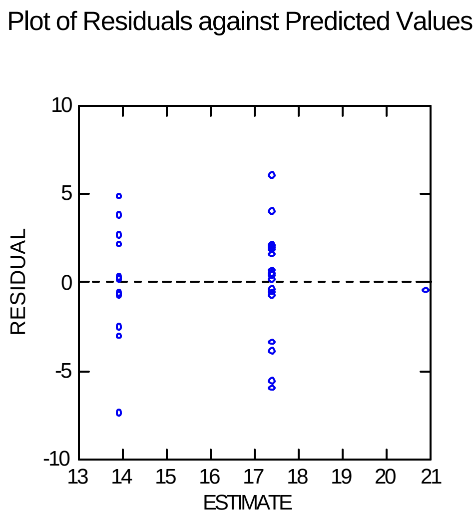


Figure 106. Depiction of the residuals for the regression for the other deposits lithology group regression. The clumping of the residuals reflects the categorical nature of the air temperature data.

Sedimentary

The regression analysis for the sedimentary lithology group identified stream gradient and canopy closure as significant variables explaining stream temperature. The model predicts that stream temperature would increase as gradient decreases, which makes intuitive sense. A depiction of the results of the model is provided in Figure 106 and the residuals are depicted in Figure 107. Note the sample size is very small for this lithology group. Different results would be expected with a larger sample size.

Table 66. Results of the regression analysis for the sedimentary lithology group.

Model with All Significant Independent Variables				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression Constant	25.71307	22.439738	28.986409	0.00012 0.00000
Canopy Closure	-0.10274	-0.179437	-0.026045	0.01325
Stream Gradient	-68.7048	-127.62377	-9.785813	0.02621
n				14
Multiple R				0.897900
Squared Multiple R				0.806224
Adjusted squared multiple R				0.770991

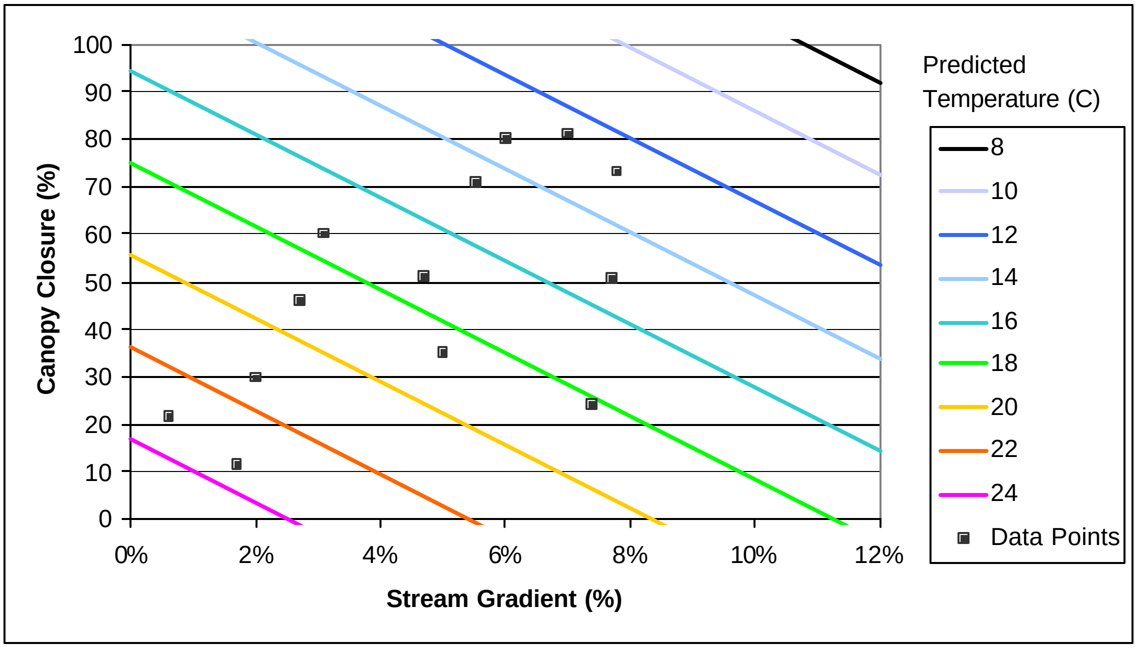


Figure 107. Depiction of the results of the regression equation developed with the database representing the sedimentary lithology group. Note, the depicted regression line was extended beyond the limits of the data. The accuracy of the line beyond the extent of the data is unknown.

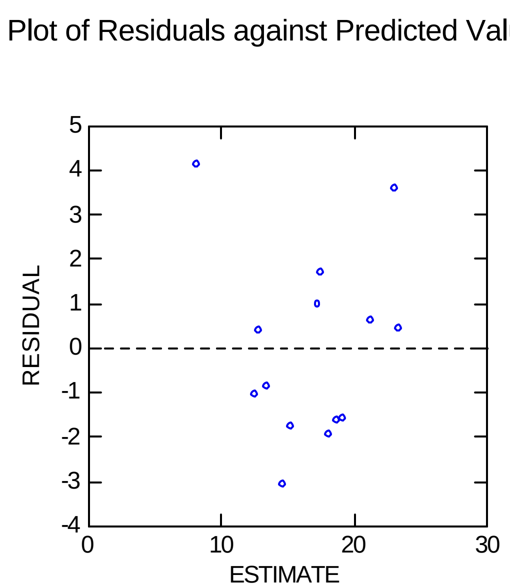


Figure 108. Depiction of residuals for the regression developed from the sedimentary lithology group database. No patterns of concern are apparent in the residuals.

3.5.4 All Data Combined

The regression of all variables on maximum stream temperature produced a model with six variables (Table 67). The variables that were included in the model were elevation, canopy closure, latitude, basin size, average annual precipitation, and distance from divide. In a forward stepwise regression, the variables entered at each step and resulting r-square were:

1. Distance from divide r^2=0.401
2. Canopy closure r^2=0.470
3. Elevation r^2=0.510
4. Latitude r^2=0.520
5. Average annual precipitation r^2=0.537
6. Basin size r^2=0.546

The variables basin size and distance from divide are strongly correlated (r^2=.83), hence the assumption of independence of these two variables is violated. Distance from divide and elevation are also highly correlated in this dataset (r^2=0.49). Basin size was the last variable to be entered into the model; therefore, elimination of this variable from the regression model is justified.

This leaves us with a five variable model (Table 67). The five variable model may be very justifiable, but it would be difficult to implement in the field. The first three variables (canopy closure, elevation, and distance from divide) explained the majority of the variance in the database. Therefore, a model containing these three variables was also developed (Table 67, Figures 109 and 110). Residuals of the models are depicted in Figure 111.

Table 67. Results of the regression analysis of all physical variables on maximum stream temperature including the entire database, without stratification.

Model with All Significant Independent Variables				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0.00000
Constant	55.299181	35.922913	74.675449	0.00000
Elevation	-0.001007	-0.001519	-0.000494	0.00014
Canopy Closure	-0.052496	-0.068146	-0.036846	0.00000
Latitude	-0.677399	-1.08056	-0.274238	0.00107
Basin Size	0.00002	0.000003	0.000038	0.02423
Average Annual Precipitation	-0.038821	-0.062672	-0.014971	0.00152
Distance from Divide	0.15687	0.010398	0.303341	0.03590
n				274
Multiple R				0.738866
Squared Multiple R				0.545923
Adjusted squared multiple R				0.535719

Table 67 Continued.

All Data, Basin Size Excluded (5 Variable Model)				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0
Constant	54.420375	34.910655	73.930094	0
Canopy Closure	-0.052013	-0.067778	-0.036249	0
Elevation	-0.001049	-0.001565	-0.000534	0.00008
Latitude	-0.776391	-1.071827	-0.259578	0.00141
Average Annual Precipitation	-0.665702	-0.062315	-0.014258	0.0019
Distance from Divide	-0.038287	0.179868	0.380844	0
n				274
Multiple R				0.732931
Squared Multiple R				0.537188
Adjusted squared multiple R				0.528553
All Data, 3 Most Significant Variables				
Variable	***Coefficient***	***Lower 95%***	***Upper 95%***	***P***
Regression				0
Constant	21.570568	19.418719	23.722416	0
Canopy Closure	-0.051059	-0.066855	-0.035262	0
Elevation	-0.001238	-0.001756	-0.000721	0
Distance from Divide	0.315486	0.217173	0.413798	0
n				274
Multiple R				0.714211
Squared Multiple R				0.510098
Adjusted squared multiple R				0.0504654

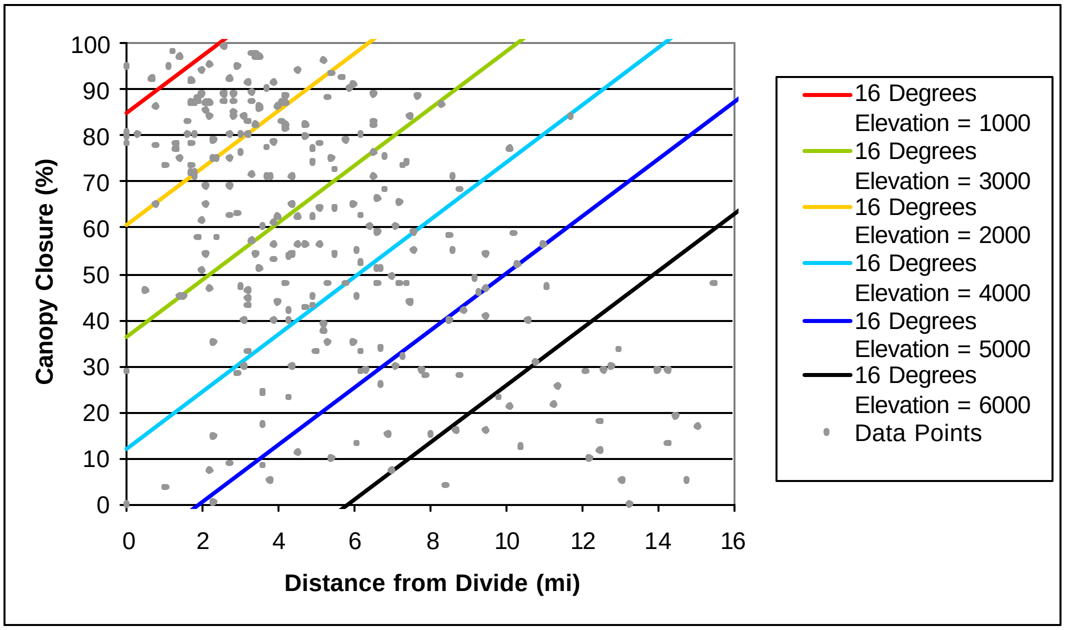


Figure 109. Depiction of the canopy closure needed to meet 16°C based on the model developed using all data in the database and the three most significant variables (distance from divide, canopy closure, and elevation in that order). Gray points represent data points used in analysis.

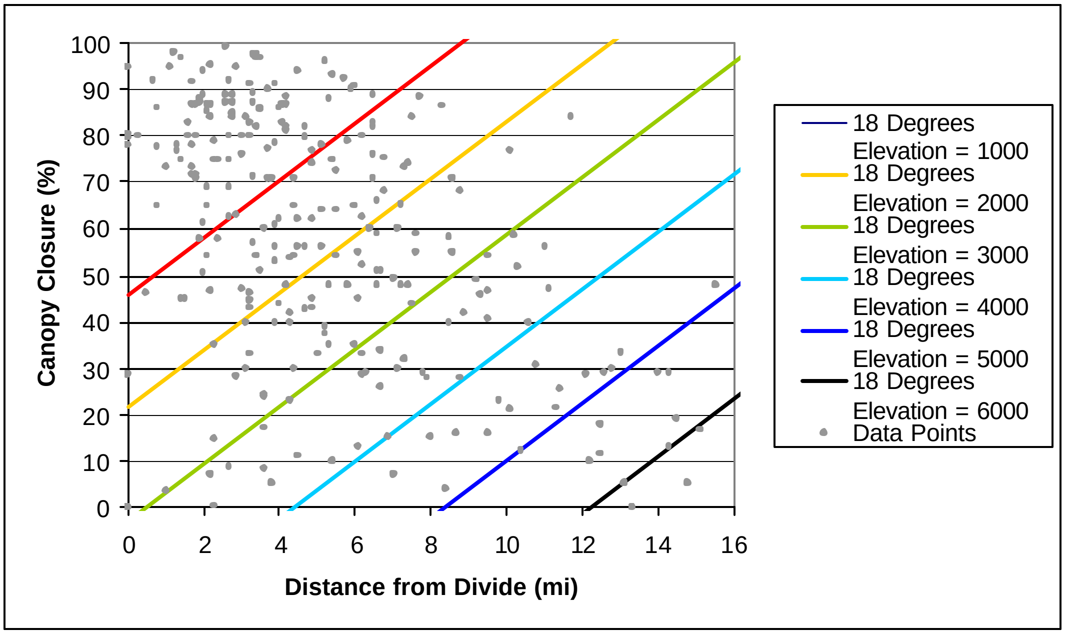


Figure 110. Depiction of the canopy closure needed to meet 18°C based on the model developed using all data in the database and the three most significant variables (distance from divide, canopy closure, and elevation in that order). Gray points represent data used in the analyses.

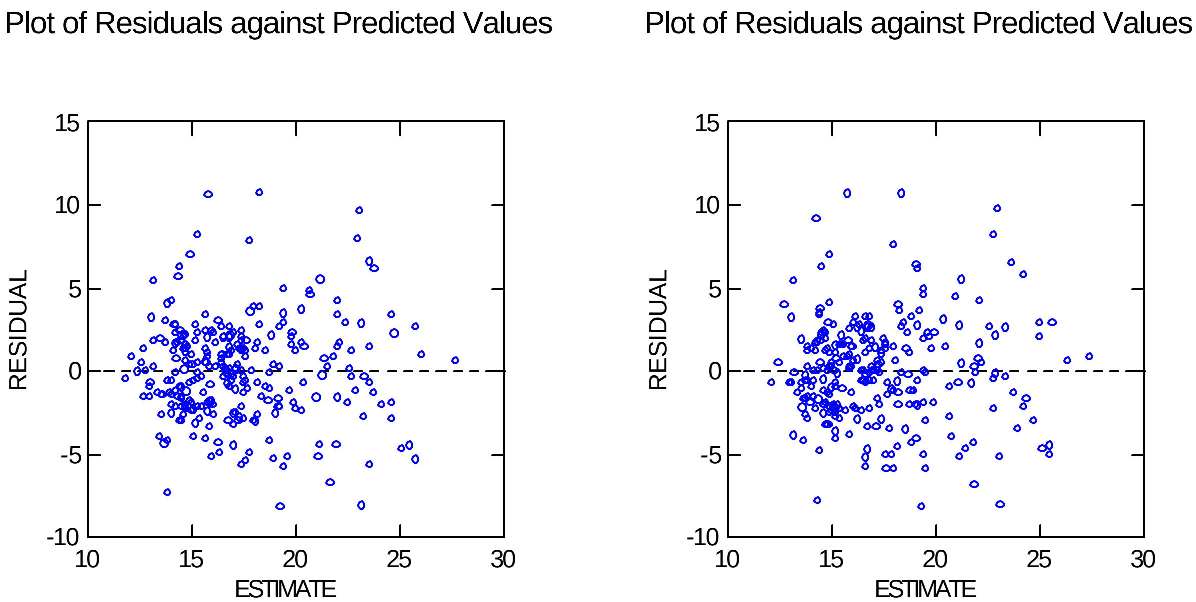


Figure 111. Residuals for the five variable model (left) using the entire database and the three variable model (right). No patterns of concern are apparent in the residuals.

4.0 DISCUSSION

The analyses presented in this document cover two basic regression analysis approaches across three stratifications. The two basic analysis approaches were 1) regressions that included only canopy closure and elevation as independent variable, which duplicates the variables included in the existing nomographs and b) regressions that included a wider range of physical parameters that could potentially affect stream temperature. Stratifications included 1) across geographical region, 2) across Bailey's ecoregion, and 3) across major lithology groups. The two analysis approaches were also presented incorporating the entire dataset without stratification. Additionally, analyses have been presented that addressed the relative fit of the existing nomographs to the larger database that was used here to conduct the various analyses.

4.1 REGRESSIONS USING CANOPY CLOSURE AND ELEVATION AS INDEPENDENT VARIABLES

4.1.1 By Region

Northeast Washington: Northeast Washington results affected by biased sample with low elevation situations poorly represented. The regression equation and both independent variables contributed significantly to the regression equation (Table 68). Additional samples representing the under-represented conditions could have a substantial effect on the results of the analyses.

North Cascades Region: In the Northern Cascades region, canopy closure did not contribute significantly to the regression equation (p=0.248; Table 68). Situations with low canopy closure at high elevation and high canopy closure at low elevation are poorly represented in the database. As a result, the dataset is biased. A larger number of samples taken from under-represented conditions could result in a substantial change in the predictive equations and may result in the inclusion of canopy closure as a significant variable in the model.

Southern Cascades Region: In the Southern Cascades region, the majority of the sample points were collected at elevations between 2000 and 4000 feet, however other elevations are reasonably represented. The full range of canopy closures is also reasonably represented. However, over 40 percent of the samples came from basins less than 5000 acres in size and over 60 percent of the samples came from basins less than 10,000 acres in size. Hence, the database tends to represent only smaller basins. This is discussed in more detail in section 4.3. The regression equation and both the independent variables were highly significant (Table 68); however, the squared multiple r was low for this dataset (0.35739) implying that the equation explained only 36 percent of the variability in the data.

Blue Mountains: There were only seven data points representing the Washington portion of the Blue Mountains in the data set and the regression with these seven data points was far from significant. Available data from Oregon was combined with the Washington data in order to increase the sample size. The distributions of the Oregon and Washington datasets appear to be similar; however, assumptions regarding this similarity should be tested. The regression equation as a whole and the independent variables were statistically significant, however the squared multiple r was low for this dataset (0.35323; Table 68) implying that the equation explained only 35 percent of the variability in the data.

Comparison of Regressions for all Regions: The regression lines for the Northern Cascades region, the Southern Cascades region, and the Northeast Washington region appear to be similar (Figure 112 and 113), however the differences between the lines are statistically significant. The regression line for the Blue Mountains is very different from the lines for the other regions. The analysis results indicate a need for higher canopy closures to meet target stream temperatures. Once again, the dataset for the Washington Blue Mountains was very small and the regression line for that region is based primarily on Oregon data. Additional data from the State of Washington would allow for a more rigorous evaluation of the applicability of the Oregon data to the Washington Blue Mountains.

Table 68. Best-fit models developed for each region. Table includes the original regression equation, the same equation rearranged to allow for predictions of required canopy closure at a specified target temperature and elevation. Also included are notes on the significance of the equations, the squared multiple r (R^2), and comments on major sources of bias or error in the analysis.

Strata	Best-Fit Equation[1]	Regression and All Variables Significant at P<0.05?	R^2	Major Sources of bias or error
By Region				
Northeast Washington	Temp = 30.258771-0.00347*Elev-0.070824*Canopy Canopy = (Target Temp -30.258771+ 0.00347*Elev)/(-0.070824)	Yes	0.63435	Some conditions under-represented
Northern Cascades	Temp = 24.338613-0.002216*Elev *With canopy forced into equation*: Temp = 24.923794-0.001767*Elev-0.033944*Canopy Canopy = (Target Temp -24.923794 + 0.001767*Elev)/(-0.033944)	Canopy not significant (p=0.248)	0.22503 0.25527	Dataset includes only drier sites. Other conditions also under-represented.
Southern Cascades	Temp = 26.283985-0.002188*Elev-0.061235*Canopy Canopy = (Target Temp -26.283985+ 0.002188*Elev)/(-0.061235)	Yes	0.35739	80% of the data comes from basins less than 10,000 acres in size
Blue Mountains	Temp = 30.744273-0.001805*Elev-0.070856*Canopy Canopy = (Target Temp -30.744273+ 0.001805*Elev)/(-0.070856)	Yes	0.35323	Wash. sample size small, most of data from Oregon

[1] Temp is Stream temperature in degrees C, Canopy is canopy closure measured in percent, Elev is site elevation in feet.

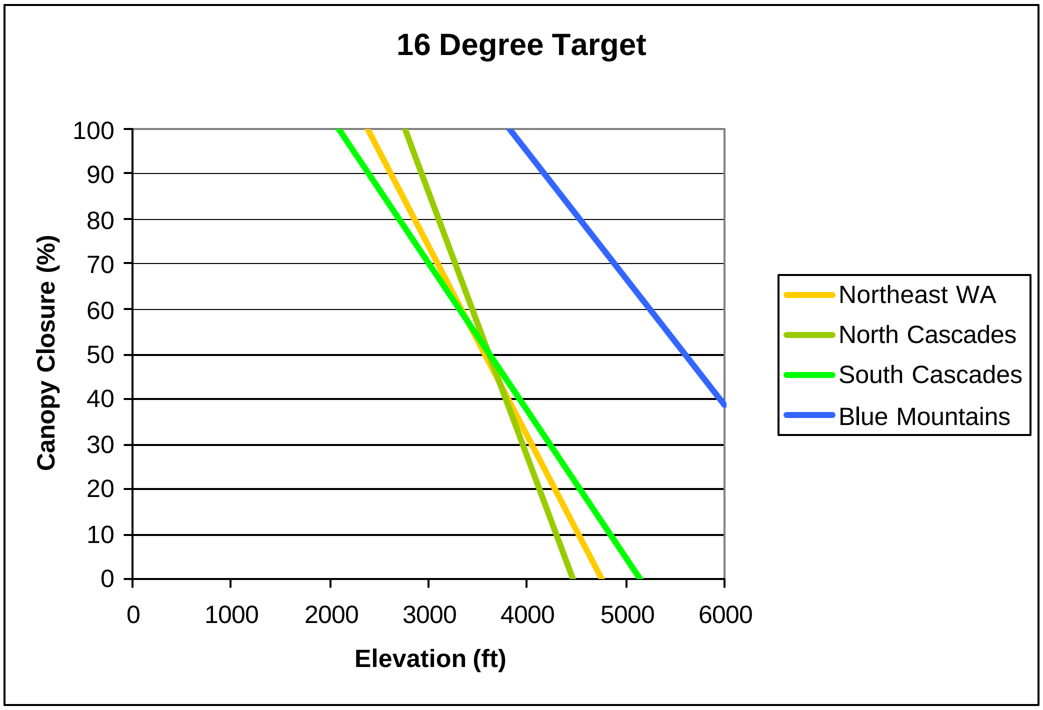


Figure 112. Canopy closure required to attain a 16 degree C temperature target for each region based on results of regression analyses using canopy closure and elevation as independent variables.

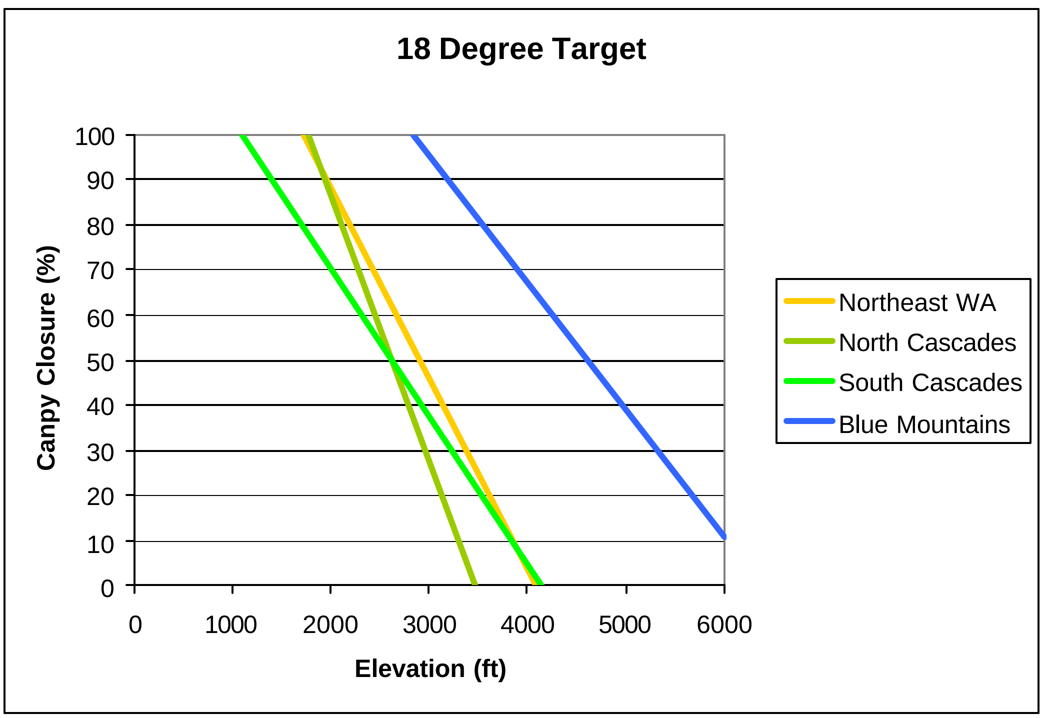


Figure 113. Canopy closure required to attain an 18 degree C temperature target for each region based on results of regression analyses using canopy closure and elevation as independent variables.

4.1.2 By Ecoregion

M333A: Ecoregion M333A (Okanogan Highlands Section) is represented by a large sample size. The regression as a whole and the individual independent variables were statistically significant (Table 69). There is a slight amount of bias in the dataset. Situations with high canopy closure and low elevation and situations with low canopy closure and high elevation are under-represented. Furthermore, low elevation data points are sparse relative to the number of points from elevations greater than 2500 feet. Additional data representing those situations that are currently under represented in the database may affect the regression results.

M242B: The elevation range represented by the data for Ecoregion M242B (Western Cascades Section) is rather narrow and most of the data points are clustered around a small range of canopy closure. This narrow distribution of the dataset has had a large effect on the analyses. A different outcome would be expected with a larger database drawn from a wider range of conditions. Statistically, the best-fit model is one that does not include canopy closure (Table 69); however, confidence is low in the results due to the limited range of conditions represented by the data. The best-fit equation developed for this ecoregion with canopy closure forced into the equation is also presented in Table 69. Overall, we have little confidence in the results of any models developed for this ecoregion and do not recommend their use. Additional data for this ecoregions would improve the confidence in the results.

M242C: The range of elevation and canopy closure conditions for Ecoregion M242C (Eastern Cascades Section) are well represented in the database. The regression overall and both canopy closure and elevation were statistically significant at the 0.05 probability level. The portion of the variance in the data explained by the regression was, however, low (r^2=0.33397). Nevertheless, confidence in the developed equations (Table 69) is good.

-342I: Canopy closure did not contribute significantly to the regression equation developed for the –342I ecoregion (Columbia Basin Section) at the 0.05 probability level. The 2-tail p level for inclusion of elevation in the model was 0.081 (Table 69). The test statistic is close to the 0.05 probability level. Given that the goal of the project was to develop relationships that would aid in the identification of the appropriate amount of canopy closure to be retained during harvests and given that canopy closure is significant at the 0.08 probability level, the model that incorporates canopy closure is recommended over one that does not. High canopy closure situations at low elevations are not well represented in the data. Improvements in the regression equations might be expected with additional samples from these physical locations. The magnitude of effect additional samples would have on the predictive equations is unknown.

-331A: Ecoregion –331A (Palouse Prairie Section) had only one data point. Therefore, no regression analyses were completed for this ecoregion.

M332G: Only one sample site differentiated the dataset for the M332G ecoregion (Blue Mountain Section) and the Blue Mountain region. The ecoregion had only six data points. No analyses were conducted for this ecoregion because the results are virtually identical to those of the Blue Mountain region and are subject to the same difficulties with the small dataset.

Table 69. Best-fit models developed for each ecoregion. Table includes the original regression equation, the same equation rearranged to allow for predictions of required canopy closure at a specified target temperature and elevation. Also included are notes on the significance of the equations, the squared multiple r (R^2), and comments on major sources of bias or error in the analysis.

Strata	Best-Fit Equation[1]	Regression and All Variables Significant at P<0.05?	R^2	Major Sources of bias or error
By Ecoregion				
M333A (Okanogan Highlands)	Temp = 27.734495-0.002771*Elev-0.06022*Canopy Canopy = (Target Temp -27.734495+ 0.002771*Elev)/(-0.06022)	Yes	0.53982	Some conditions slightly under-represented
M242B (Western Cascades)	No recommended equation developed due to small sample size and narrow range of represented conditions.	No, canopy not significant (p=0.62425)		Data is very small (n=11). Use of equations not recommended.
M242C (Eastern Cascades)	Temp = 24.587646-0.000995*Elev-0.081823*Canopy Canopy = (Target Temp -24.587646+ 0.000995*Elev)/(-0.081823)	Yes	0.33397	
-342I (Columbia Basin)	Temp = 30.419477-0.00284*Elev-0.064892*Canopy Canopy = (Target Temp -30.419477+ 0.00284*Elev)/(-0.064892)	No, canopy closure is significant at p=0.08.	0.51139	Conditions at low elevation, high canopy closure are under-represented.
-331A (Palouse Prairie)	N/A	N/A	N/A	Only one data point in ecoregion.
M332G (Blue Mountain)	See results for the Blue Mountain region.			See results for the Blue Mountain region.

[1] Temp is Stream temperature in degrees C, Canopy is canopy closure measured in percent, Elev is site elevation in feet.

Comparison and Discussion of Regressions for all Ecoregions: The two lines representing ecoregions where the analyses were robust (M333A and M242C) are substantially different The first represents primarily north central and northeastern Washington and the second represents the east flanks of the Cascades Mountains. The lines for these two ecoregions would indicate that the east flanks of the Cascade Mountains require less canopy closure to meet a temperature target at low elevations than the region in north and northeast Washington. Conversely, greater canopy closure is apparently needed on the east flanks of the Cascade Mountains at higher elevation. Weather patterns and air temperature changes substantially with elevation on the east flanks of the Cascade Mountains. If these changes are more pronounced than in north and northwest Washington, the differences in the position of the two lines may be logical. The equation for Ecoregion 0342I (Columbia Basin Section) is similar to that developed for M333A (East Cascades Section).

The equations for Ecoregion M333A (Okanogan Highlands Section) and Ecoregion M242B (Western Cascades Section) did not test as significantly different. All other pairs of populations were found to be significantly different from each other at the 0.05 probability level. Given that ecoregions M333A and M242B are not statistically different although the differences in their regression equations were large, and given the difficulties with the datasets in other ecoregions, the use of ecoregions as a stratification parameter is not recommended.

4.1.3 By Lithology

Alluvium: The range of canopy closure and elevation conditions was well represented in the dataset for the alluvium lithology. The regression was significant at the 0.05 probability level, as were the inclusions of canopy closure and elevation as independent variables. The squared multiple r, however, was low (0.288081) indicating that less than 29 percent (Table 70) of the variability in the data was explained by the two independent variables. Sample size was relatively good, so additional sampling is not likely to result in a substantial improvement in the model.

Basalt: The data set representing the basalt lithology group is reasonably well distributed although situations with low canopy closure at high elevation and situations at elevations below 1000 feet are under-represented. Sample size is reasonably good (n=47). The regression was significant at the 0.05 probability level, as were the inclusions of canopy closure and elevation as independent variables (Table 70).

Glacial Drift: The regression for the glacial drift lithology was significant at the 0.05 probability level, as were the inclusions of canopy closure and elevation as independent variables (Table 70). The majority of the data points are from situations with higher canopy closure and elevations ranging from 2000 to 4000 feet. Lower canopy closure (<50%) and both low and high elevations are under-represented. Additional data representing these situations may affect the model results.

Granites: The sample size for the granites lithology is very small (n=13). Three of the data points tested as outliers. The regression equations that were developed all indicate that increasing amounts of canopy closure are required with increasing elevation. Since air tends to cool at increasing elevation, the results did not make physical sense. Hence, all equations developed are assumed spurious. No equation is recommended. A larger sample size is likely to produce different results.

Table 70. Best-fit models developed for each lithology group. Table includes the original regression equation, the same equation rearranged to allow for predictions of required canopy closure at a specified target temperature and elevation. Also included are notes on the significance of the equations, the squared multiple r (R^2), and comments on major sources of bias or error in the analysis.

Strata	Best-Fit Equation[1]	Regression and All Variables Significant at P<0.05?	R^2	Major Sources of bias or error
By Lithology				
Alluvium	Temp = 26.677486-0.002149*Elev-0.090752*Canopy Canopy = (Target Temp -26.677486+ 0.002149*Elev)/(-0.05157)	Yes	0.28808	
Basalt	Temp = 28.23372-0.001985*Elev-0.012172*Canopy Canopy = (Target Temp -28.23372+ 0.001985*Elev)/(-0.090752)	Yes	0.62550	Low canopy closure (<30%) at high elevation and very low elevation poorly represented
Glacial Drift	Temp = 28.380374 - 0.003013*Elev - 0.071546*Canopy Canopy = (Target Temp -28.380374+ 0.003013*Elev)/(0.07155)	Yes	0.57584	<50% canopy cover, low and high elevations are under-represented.
Granites	None of the regression results are physically logical and all are assumed spurious. No equation recommended.			Small sample size (13).
Meta-morphic	N/A Model unstable and highly affected by individual data points. All results likely spurious.	N/A	N/A	Small sample size (19). Unstable results.
Other Deposits	Temp = 21.412268-0.00193*Elev *With canopy forced into equation*: Temp = 24.617141-0.002236*Elev-0.038784*Canopy Canopy = (Target Temp -24.617141+ 0.002236*Elev)/(-0.03878)	No, canopy not significant (p=0.125)	0.20285 0.27048	Low elevation with low canopy closure not represented.
Sedimentary	Equations were developed but are not recommended for use.			Small sample size, poor distribution of the data

[1] Temp is Stream temperature in degrees C, Canopy is canopy closure measured in percent, Elev is site elevation in feet.

<u>Metamorphic</u>: The sample size is small (19) for this lithology group. Several of the samples tested as outliers, further reducing sample size. The model parameters change dramatically with the removal of each outlier suggesting that any model based on the existing dataset, with or without the outliers included, may be spurious. The model is clearly unstable and highly affected by individual data points. Therefore, we have little confidence in the model developed for the metamorphic lithology group. An increase in sample size may help to stabilize the model

and result in a significant relationship between stream temperature and the independent variables.

Other Deposits: In the regression analysis for this lithology group, canopy closure was not identified as a significant variable for inclusion in the model. The 2-tail probability value for the inclusion of canopy closure was 0.125 (Table 70). The regression using only elevation as the independent variable is statistically the better model (Table 70). Both regression equations explained only a small (<28%) portion of the variability in the data. The results of the modeling effort were likely affected by a moderately low sample size and by the distribution of the data across the range of elevation and canopy closure. Situations at low elevation with low shade were not represented in the database. Additional data covering this missing sector would likely affect the regression results. Because canopy closure is significant at the 0.125 level, the model incorporating canopy closure could be employed in the development of eastern Washington nomographs. Confidence in the results are, however, low.

Sedimentary: In the regression analysis for this group, elevation was not identified as a significant variable for inclusion in the model. The 2-tail probability value for the inclusion of elevation was 0.199. The small sample size (14) may have had an affect on this regression analysis. In addition, most of the data points for the sedimentary lithology group lie within a 1500-foot range of elevations and there is little variation in elevation at each canopy closure level. This distribution of the data has also affected the analysis. A wider representation of elevation may produce a different result. A second regression analysis was conducted using only canopy closure as an independent variable (Table 70). Given that elevation did not contribute significantly to the original model, the second model is probably the best choice at this time, but additional data would likely produce different results. Confidence is extremely low regarding the analyses conducted for this lithology group; hence, we do not recommend either equation be used.

Comparison and Discussion of Regressions for all Lithology Groups: There are three lithology groups with regressions that are statistically significant at the 0.05 probability level and have good datasets. One of these three groups, the alluvium lithology group, has a poor squared multiple r. Therefore, only two lithology groups have good regressions that explain over 50% of the variability in the database. These two groups are basalt and glacial drift. These two lithology groups represent only 38% of the entire database.

The two lithology groups with robust analyses (basalt and glacial drift) did test as statistically different. The alluvium lithology group (with the low squared multiple r) did not test as statistically different from either the basalt or glacial drift lithology groups. This is likely due to the high level of variability in that dataset.

The differences between the basalt and glacial drift datasets may be related to the differences in the spatial distribution of the two sets. The majority of the data for the basalt lithology group comes from the south Cascades (Figure 114) and the majority of the data for the glacial drift lithology group comes from northeast Washington. Therefore, apparent difference in response between lithology groups may be reflective of differences in the climate of the two regions.

Given the potential that apparent differences in lithology may in fact reflect differences in location, the small portion of the entire dataset encompassed by these two groups, the small sample sizes in many of the other groups, and the poor performance of the regression analyses for the other groups, we do not recommend stratifying by lithology when developing eastern Washington nomographs.

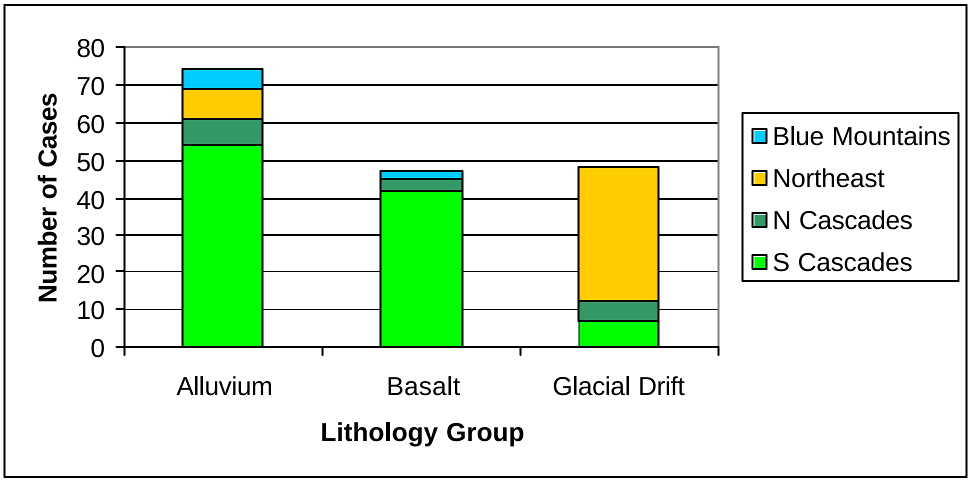


Figure 114. Distribution of the cases in the three lithology groups with significant regressions and well distributed datasets across region.

4.1.4 All Data Combined – No Stratification

Regressions were conducted on the full dataset without stratification. The data in the full dataset were well distributed across all variables. Both canopy closure and elevation were significant in the equation; however, the squared multiple r was relatively small (0.37331; Table 71).

Table 71. Best-fit model developed for the entire dataset without stratification. Table includes the original regression equation and the same equation rearranged to allow for predictions of required canopy closure at a specified target temperature and elevation.

Best-Fit Equation[1]	Regression and All Variables Significant at P<0.05?	R^2	Major Sources of bias or error
Temp = 26.649863-0.00228*Elev-0.061332*Canopy Canopy = (Target Temp -26.649863+ 0.00228*Elev)/(-0.061332)	Yes	0.37331	

[1] Temp is Stream temperature in degrees C, Canopy is canopy closure measured in percent, Elev is site elevation in feet.

4.2 REGRESSIONS WITH ALL OF THE AVAILABLE INDEPENDENT VARIABLES

4.2.1 By Region

Northeast Washington: The most robust regression for Northeast Washington included canopy closure and distance from divide as independent variables (Table 72). Distance from divide may be representative of stream size and/or air temperature.

The dataset is not well distributed across the range of canopy closure and distance from divide. Most of the data were collected in areas with high canopy closure levels located within 7 miles of the divide. Low canopy closure situations are poorly represented (Figure 115). Only 19 percent of the samples had canopy closures that were less than 50 percent. Sites that were more 10 miles from the divide are entirely missing from the dataset. Additional data points filling these poorly represented situations would likely affect the model results.

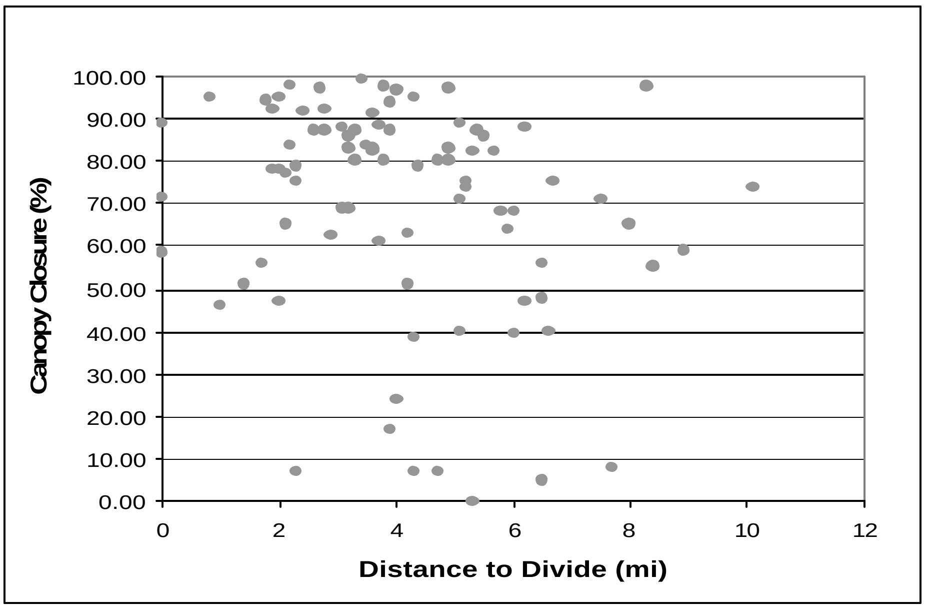


Figure 115. Distribution of sample sites in northeast Washington across canopy closure (%) and distance to divide (mi).

Table 72. Best-fit models developed for each region. Table includes the original regression equation, the same equation rearranged to allow for predictions of required canopy closure at a specified target temperature and elevation. Also included are notes on the significance of the equations, the squared multiple r (R^2), and comments on major sources of bias or error in the analysis.

Strata	Best-Fit Equation[1]	Regression and All Variables Significant at P<0.05?	R^2	Major Sources of bias or error
By Region				
Northeast Washington	Temp = 14.401348 + 0.708328*Divide - 0.028686*Canopy Canopy = (Target Temp - 14.401348 - 0.708328*Divide)/ (-0.028686)	Yes	0.74481	Lower canopy closure situations under-represented.
Northern Cascades	Temp = 18.767835 - 0.204538*Precip - 0.665919*Bank full *With canopy forced into equation*: Temp = 17.84225 + 0.336171*Divide - 0.030218*Canopy Canopy = (Target Temp - 17.84225 - 0.336171*Divide) / (- 0.030218)	Canopy not significant (p=0.265)	0.57343 0.32697	Dataset includes only drier sites. Other conditions also under-represented.
Southern Cascades	*3-parameter model:* Temp = 23.9702320 - 0.0016200*Elev - 0.0534190*Canopy + 0.0000390*Basin Size Canopy = (Target Temp - 23.9702320 + 0.0016200*Elev - 0.0000390*Basin Size.)/(-0.0534190) *2-parameter model:*[2] Temp = 19.329356-0.000051*Basin Size - 0.054232*Canopy Canopy = (Target Temp -19.329356+ 0.000051*Basin Size)/ (-0.054232) *Best 2-parameter model without Basin Size:* See regression using elevation and canopy closure reported in the previous section.	Yes	0.47275 0.37981 0.35739	Basin sizes biased towards very small streams.
Blue Mountains	See results of regressions including only canopy closure and elevation.	Yes	0.35323	Wash. sample size small, most of data from Oregon

1 Temp is Stream temperature in degrees C, Canopy is canopy closure measured in percent, Elev is site elevation in feet, basin size is in acres, precip is average annual precipitation in inches, bank full is bank full width in feet, divide is distance to divide in miles.

2 The coefficients for the southern Cascades model with canopy closure and elevation only are slightly different than earlier reported. This is due to a difference of one outlier that was removed in this equation but retained in the earlier analysis.

Northern Cascades Mountains: The North Cascades region has 37 cases in the full data set. Twelve of these cases included measures of air temperature and 28 of the cases included measures of bank full width (Table 72). As a result, this region provided an opportunity to investigate the potential effects of air temperature and bank full width on the regressions. The best-fit model contained average annual precipitation and bank full width, but no canopy closure. The inclusion of the precipitation and bank full width variables would suggest that stream size is a primary driver of stream temperature at the range of sites represented by the data. When bank full width was excluded from the analyses (to increase sample size), distance from divide was found to be the only significant variable. Distance from divide can represent stream size and/or air temperature. Canopy closure was forced into the equation with distance from divide to provide an equation that may be useful in the development of a nomograph to guide the amount of canopy closure needed to be left during harvest. Canopy closure far from significant in this equation (p=0.265).

Two prominent biases in the data were identified. Average annual precipitation at all locations was less than 30 inches. Hence, the database represents only the drier portions of the region. No sample sites came from the wetter locations of the northern Cascades Mountains. Hence, the regression equations are only applicable to locations with low average annual precipitation. Canopy closure was also poorly represented across the range of values for distance from divide. The higher canopy closures (greater than 60 percent) were only found in the basins closest to the divide and those that were furthest from the divide. The poor representation of the range of canopy closure situations across the range of distances from divide may have contributed to the lack of significance of canopy closure in the analyses.

Southern Cascades Mountains: Site descriptive variables that were measured in the field (wetted and bank full width and depth, stream flow, valley width, and valley type) were not available for any of the sites in the Southern Cascades Mountains region. Hence, the independent variables included in the regression analyses were limited to those that were developed through GIS, elevation, and canopy closure.

The model with the highest correlation coefficient was a model with four independent variables in it including elevation, canopy closure, average annual precipitation, and basin size. Operationally, the four-parameter model would be difficult to implement, hence the relative contribution of the four variables was evaluated. The addition of average annual precipitation only had a small effect on the overall predictive capability of the model; hence, a three-parameter model is recommended over the four-parameter model (Table 72). Two-parameter models were also explored. The model with the highest squared multiple r was the model that included elevation and basin size. The squared multiple r for the two-parameter modeled was roughly 20 percent lower than that for three-parameter model (Table 72). Therefore, the predictive power and accuracy of the model is reduced somewhat by dropping basin size from the model, but there may be operational reasons to want to simplify the nomograph to two parameters.

The dataset is not evenly distributed across the range of basin sizes. Over 40 percent of the samples came from basins less than 5000 acres in size and over 60 percent of the samples came from basins less than 10,000 acres in size (Figure 116). In eastern Washington, streams that drain less than 5000 acres tend to be small headwater streams. These small streams tend to warm to higher temperatures in day and cool at night to lower temperatures due to the small amount of water subject to heating and cooling processes. Therefore, the database is dominated by a narrow range of stream sizes with pronounced responses to factors influencing temperature. Preferably, the samples would be representative of the entire range of the data. With the distribution of basin sizes in this region, the left side of the regression line is driven by the abundance of small streams and the right side is strongly affected by the small number of larger basins in the database. Therefore, the dataset is rather biased. The regression results would likely be different with a dataset that better represents the range of basin. Therefore, caution is urged regarding the use of the regressions developed for this region that incorporate basin size.

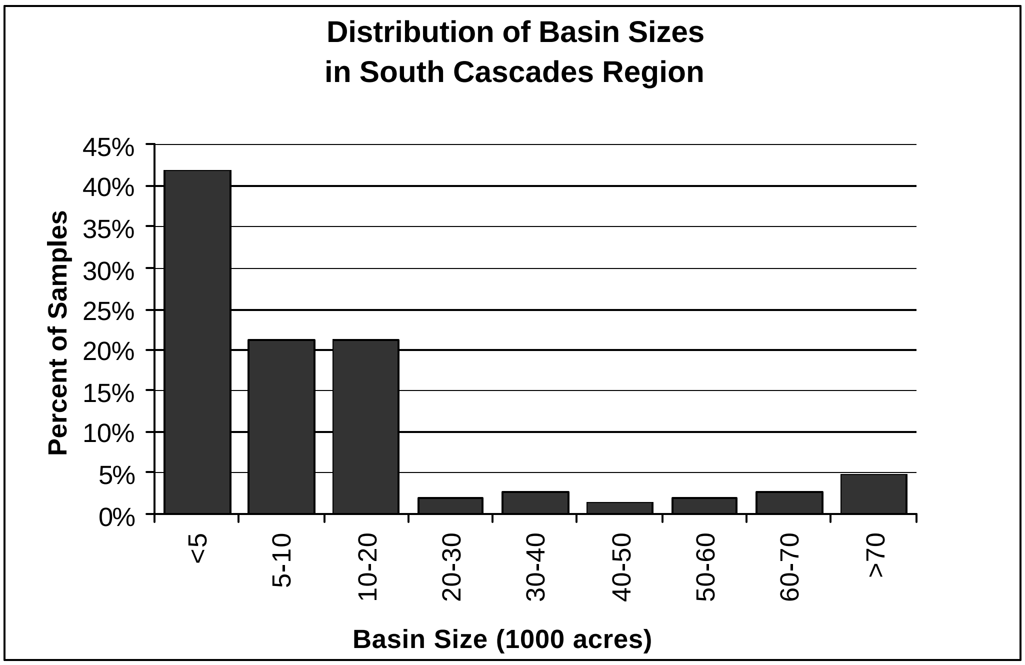


Figure 116. Distribution of basin sizes in the South Cascades region.

The most robust equation developed for the southern Cascades Mountain region was the equation reported in the previous section incorporating canopy closure and elevation. The difference in the squared multiple r values is small (Table 72). The dataset, however, is still affected by the high percentage of very small streams.

Some would suggest that the distribution of basin size in the dataset would suggest a transformation, however, in regression analysis, transformations are only indicated if the distribution of a variable around a particular value of the dependent variable (stream temperature) is not normal (Steele and Torrie 1980). In the case of these analyses, there was no reason to assume that the distribution of basin sizes around a particular temperature was not normal; hence, a transformation was not appropriate. Limiting the analysis to a smaller range of basin sizes or increasing the number of larger basin sizes represented in the dataset may have significant effects on the regression results.

Blue Mountains: There were only seven data points available in the Washington dataset for the Blue Mountains. The only physical parameters available for the Oregon dataset were elevation and canopy closure. Therefore, no additional data analyses were conducted. The results of analyses incorporating elevation and canopy closure were previously discussed.

4.2.2 By Ecoregion

M333A (Okanogan Highlands Section): This ecoregion has a large sample size. The analyses produced an equation with canopy closure, distance from divide, and average annual precipitation in the equation (Table 73). Distance from divide explained the largest portion of the variability in the dataset. The list of variables included would suggest that stream size is correlated with stream temperature. In this area, however, average annual precipitation and distance from divide tend to be inversely correlated. Streams farther from the divide receive less precipitation. Hence, streams that drain solely low elevation terrain may tend to be small.

The three-parameter model may be difficult to implement on the ground. Therefore, two and one parameter models were explored. Analyses that included distance from divide with canopy closure, distance from divide with average annual precipitation, and distance from divide alone were all completed. Interestingly, all three models were similarly robust. The regressions in all three models were highly significant, the coefficients for the constant and distance from divide changed very little, and the squared multiple r values were within 0.0126 of each other, indicating the amount of the variability explained by the three models and the fit to the data were very similar. The significance of the regressions and the squared multiple r were also very similar to those in the three-parameter model. Since there are four models with almost identical performance, the simplest of the four would normally be recommended. The one-parameter model, however, does not include canopy closure. Therefore, the two-parameter model incorporating distance from divide and canopy closure (Table 73) is the recommended model for this ecoregion.

Data represented in this dataset is somewhat biased. The majority of the data comes from locations that are less than 7 miles from the divide with greater than 50 percent canopy closure (Figure 119). The range of precipitation was well represented at sample sites that were less than 7 miles from the divide; however sample locations that were further from the divide tended to be very dry sites with less than 20 inches of annual precipitation (Figure 117). There are no sites with high canopy closure at larger distances from the divide in the dataset. The low number (or absence) of samples representing some conditions likely affected the results of the regression analyses.

Table 73. Best-fit models developed for each ecoregion using all the independent variables in the analyses. Table includes the original regression equation, the same equation rearranged to allow for predictions of required canopy closure at a specified target temperature and elevation. Also included are notes on the significance of the equations, the squared multiple r (R^2), and comments on major sources of bias or error in the analysis.

Strata	Best-Fit Equation[1]	Regression and All Variables Significant at P<0.05?	R^2	Major Sources of bias or error
By Ecoregion				
M333A (Okanogan Highlands)	*3-parameter model:* Temp = 16.420597+ 0.573397*Divide - 0.066021*Precip - 0.020744*Canopy Canopy = (Target Temp - 16.420597 - 0.573397*Divide + 0.066021*Precip) / (-0.020744) *2-parameter model with canopy closure*: Temp = 14.078066+ 0.652548*Divide - 0.020255*Canopy Canopy = (Target Temp -14.078066 - 0.652548*Divide)/(-0.020255)	Yes Yes	0.70930 0.69611	Most of the data from small basins with high shade. Sites at larger distances from divide also tend to be dry sites. Distance from divide alone performs as well as these equations.
M242B (Western Cascades)	Use of developed equations not recommended due to small sample size.			Dataset is very small (n=11). Use of equations not recommended.
M242C (Eastern Cascades)	*3-parameter model:* Temp = 8.434834 + 0.000045*Basin + 0.189795*Aug Air - 0.060807*Canopy Canopy = (Target Temp - 8.434834 - 0.000045*Basin - 0.189795*Aug Air) / (-0.060807) *2-parameter model with canopy closure*: Temp = 19.593384 + 0.000049*Basin - 0.055139*Canopy Canopy = (Target Temp -19.593384 - 0.000049*Basin)/(-0.055139)	Yes Yes	0.44611 0.40969	Dataset includes only basins less than 10,000 acres; equations should not be applied in basins much larger than that.
-342I (Columbia Basin)	Temp = -16.585662 + 0.842753*average air	Yes	0.54675	Conditions at low elevation, high canopy closure are under-represented. Small sample size. Unusual data points affecting results.

1 Temp is Stream temperature in degrees C, Canopy is canopy closure measured in percent, Elev is site elevation in feet, basin is basin size in acres, precip is average annual precipitation in inches, bank full is bank full width in feet, divide is distance to divide in miles, Aug air is average August air temperature (F), average air is average annual air temperature (F).

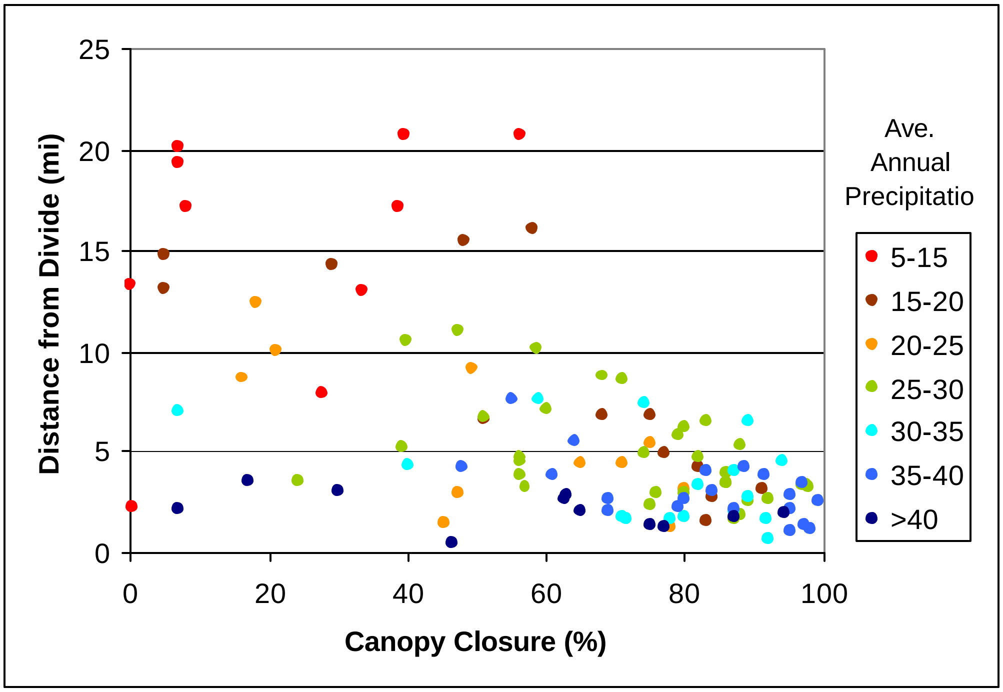


Figure 117. Distribution of samples in the M333A (Okanogan Highlands Section) ecoregion across distance from divide, canopy closure, and average annual precipitation.

<u>M242B (Western Cascades Section)</u>: There were only 11 data points in the database representing Ecoregion M242B. A three-parameter model was developed that included canopy closure, distance from divide, and average annual precipitation. Two outliers had to be removed, leaving only 9 points in the analysis. A two-parameter model was also developed, but three outliers had to be removed from that analysis; hence, the two-parameter model was only represented by 8 data points. Confidence in any of the regressions developed for this ecoregion is low due to the low sample size. Use of any of the equations developed is not recommended.

<u>M242C (Eastern Cascades Section)</u>: A 3-parameter model was developed using basin size, canopy closure, and August air temperature as the independent variables (Table 73). Recognizing that a 3-parameter model may be difficult to implement, 2-parameter models were explored. The best-fit 2-parameter model incorporated canopy closure and basin size (Table 73). The differences in the squared multiple r values for the two models was small, indicating that the two models explained a similar amount of variance in the data are were similar in overall accuracy. The inclusion of August air temperature in the model increased the canopy closure required by roughly 30 percent for any given basin size. Two August air temperatures were represented in the dataset. The predicted stream temperatures for these two air temperatures differed by approximately 2 degrees C.

The majority of the data points for this ecoregion came from basins that were less than 10,000 acres in size with greater than 30% canopy closure. Therefore, the regression equations developed with this dataset do not represent areas where basin size is greater than 10,000 acres. A of these models is not recommended for basins much larger than 10,000 acres.

-342I (Columbia Basin Section): Ecoregion -342I (Columbia Basin Section) has only 25 data points. The most robust equation developed for this ecoregion included only average annual air temperature as an independent variable (Table 73). Average annual air temperature was experimentally excluded from the analysis. The subsequent model included August air temperature and canopy closure. The appearance of a measure of air temperature in both models suggests climatic patterns are important in this ecoregion.

Two data points have a strong influence on the relationships developed in both sets of analyses. The stream temperature at these sample sites is unusually warm. We have no basis for excluding this data from the analyses; however, the data does appear to be somewhat questionable. In the absence of these data points, August air temperature may not have been significant in the second model. The influence of these data points reduces confidence in the second model. The first model, however, does not include canopy closure and is not useful as a guide to the amount of canopy closure that needs to be left during harvest activities. Given the small sample size and the influence of the two data points, we are hesitant to recommend the use of the second model. Substantially different results may be found for this ecoregion if the sample size was increased.

-331A (Palouse Prairie Section): Ecoregion –331A had only one data point. Therefore, no regression analyses were completed for this ecoregion.

M332G (Blue Mountain Section): Only one sample site differentiated the dataset for the M332G ecoregion and the Blue Mountain region. The ecoregion had only six data points. No analyses were conducted for this ecoregion because the results are virtually identical to those of the Blue Mountain region and are subject to the same difficulties with the small dataset.

4.2.3 By Lithology

Alluvium: The variables that were found to be significant in the equation with the outliers removed include basin size (acres), latitude, and average July air temperature (F) (Table 74). Canopy closure was not included as a significant variable. Numerous runs were made to try to force canopy into the equation. With basin size or elevation in the equation, canopy closure was always far from significant. The combination of canopy closure, elevation, and precipitation was significant, however the squared multiple r for that equation was only 0.394. The squared multiple r for the regression that included basin size, latitude and July air temperature had a squared multiple r of 0.669. Therefore, none of the attempted models that incorporated canopy closure was included in this assessment.

The datasets had physical conditions that were poorly represented or missing in the database. High canopy closure situations were missing from all but the smallest basins. Precipitation was well represented for only basins that were less then 7,000 acres. The regression analyses would be much improved and would possibly have different outcomes if these under-represented (or absent) situations were better represented in the database.

Table 74. Best-fit models developed for each lithology group. Table includes the original regression equation, the same equation rearranged to allow for predictions of required canopy closure at a specified target temperature and elevation. Also included are notes on the significance of the equations, the squared multiple r (R^2), and comments on major sources of bias or error in the analysis.

Strata	Best-Fit Equation[1]	Regression and All Variables Significant at P<0.05?	R^2	Major Sources of bias or error
By Lithology				
Alluvium	Temp = 60.14713 + 0.000076*Basin -1.18726*Lat + 0.210184*Jul Air	Yes	0.66943	Under-represented conditions.
Basalt	Temp = 10.72126 - 0.00125*Elev - 0.09041*Canopy + 0.24414* Jul Air Canopy = (Target Temp - 10.72126 + 0.00125*Elev - 0.24414*Jul Air)/(-0.09041)	Yes	0.69154	Low canopy closure (<30%) at high elevations and sites <1000 feet poorly represented
Glacial Drift	Temp = 19.96385 - 0.08803*Precip - 0.04568*Canopy + 0.000124*Basin Canopy = (Target Temp - 19.96385 + 0.08803*Precip - 0.000124*Basin)/(-0.04568) *Recommend not using the equation for basins that are much more than 10,000 acres.*	Yes	0.69513	<50% canopy cover, low and high elevations are under-represented. Few basins >5000 acres in the dataset.
Granites	Temp = 15.6385 + 0.00018*Basin	Canopy not significant	0.68652	Small sample size (13). Large basins very poorly represented. High canopy at low elevation not well represented.
Meta-morphic	Temp = 10.84941+ 1.149515*Divide	Canopy not significant	0.54826	Small sample size (19). Somewhat unstable results.
Other Deposits	Temp = -5.21308 + 0.348077*Jul Air	Canopy not significant	0.26617	Low elevation with low canopy closure not represented.
Sedimentary	Temp = 25.71307- 0.10274*Canopy - 68.7048*stream gradient Canopy = (Target Temp - 25.71307+ 68.7048*stream gradient)/(- 0.10274)	Yes	0.80622	Small sample size, poor distribution of the data

[1] Temp is Stream temperature in degrees C, Canopy is canopy closure measured in percent, Elev is site elevation in feet, basin is basin size in acres, precip is average annual precipitation in inches, bank full is bank full width in feet, divide is distance to divide in miles, Aug air is average August air temperature (F), average air is average annual air temperature (F), Lat is latitude, Jul air is average July air temperature (F).

Basalt: The majority of the basalt lithology group dataset comes from the southern portion of the eastern Cascade Mountains. A significant regression analysis was developed that included canopy closure, elevation, and average August air temperature (Table 74).

As was discussed earlier, situations with low canopy closure at high elevation and sites at elevations less than 1000 feet are under-represented in the database. Additional data for these situations may have an effect on the outcome of the analyses.

Glacial Drift: The most robust regression for the glacial drift lithology group was a three-parameter model that included canopy closure, basin size, and average annual precipitation (Table 74). As was discussed previously, the majority of the data points representing this lithology group are from situations with higher canopy closure and elevations ranging from 2000 to 4000 feet. Lower canopy closure (<50%) and both low and high elevations are under-represented. Additionally, 81 percent of the samples are from basins that are less than 5000 acres in size. Given the distribution of the data, the model results should be extended to only those locations that are no more than 10,000 acres in size. Additional data representing these situations would likely affect the model results.

Granites: The best-fit model for the granite lithology group was a model that predicted stream temperature as a function of basin size alone (Table 74). Canopy closure was not significant at the 0.05 probability level. When forced into the equation, canopy closure did not results in substantive changes in the predictive ability of the equation. The entire dataset for the granite lithology group was collected in relatively small basins. Most of the data for this lithology group comes from NE Washington, where situations with high canopy closure at low elevation are also poorly represented. With an expanded dataset representing a larger range of basin sizes, canopy closures, and other conditions, better predictive equations might have been developed.

Metamorphic: The regression with all possible independent variables identified only distance from divide as significantly explaining stream temperature (Table 74). Canopy closure was not significant at the 0.05 probability level. The sample size is very small for this strata and the results may be improved with an increased sample size.

Other Deposits: Average July air temperature was the only significant independent variable identified (Table 74). Sample size was relatively small, which may have affected this analysis. In addition, some canopy situations are not well represented in the database.

Sedimentary: The regression analysis for the sedimentary lithology group identified stream gradient and canopy closure as significant variables explaining stream temperature (Table 74). The sample size for this lithology group was only 14 points. With additional data representing this group, difference regression results would be expected.

4.2.4 All Data Combined – No Stratification

A five variable model was developed through a regression analysis using the full dataset and all the potential independent variables. The included variables were canopy closure, elevation, distance from divide, latitude, and average annual precipitation (Table 75). The five variable model was robust and may be very justifiable, but it would be difficult to implement in the field. Hence, a smaller set of variables was evaluated. The first three variables that entered into the forward stepwise regression were (canopy closure, elevation, and distance from divide) explained the majority of the variance in the database. Therefore, a model containing these three variables was developed (Table 75). The elimination of average annual precipitation and latitude resulted in a decrease in the squared multiple r of only 0.027. Hence, the predictive capability of the model was only marginally compromised. The distribution of the data across all variables was reasonably good, which supports a high degree of confidence in the model.

Table 75. Best-fit equations developed using the entire database without stratification and all possible independent variables. Equations are given in two forms. One describing the equation to estimated predicted stream temperature and one rearranged to calculate the needed canopy closure to meet a target temperature.

Best-Fit Equation[1]	Regression and All Variables Significant at $P<0.05$?	R^2	Major Sources of bias or error
5-parameter model Temp = 54.420375 - 0.052013*canopy - 0.001049*elev - 0.776391*lat - 0.665702*precip - 0.038287*divide Canopy = (Target temp - 54.420375 + 0.001049*elev + 0.776391*lat + 0.665702*precip + 0.038287*divide) /(- 0.052013)	Yes	0.53719	None.
3-parameter model Temp = 21.570568 - 0.001238*elev + 0.315486*divide - 0.051059*canopy Canopy = (Target temp - 21.570568 + 0.001238*elev - 0.315486*divide) / (- 0.051059)	Yes	0.51010	None.

[1] Temp is stream temperature (C), canopy is canopy closure (%), divide is distance from divide (mi), precip is average annual precipitation (in), elev is elevation in feet, and lat is latitude.

4.3 CHOICE OF STRATIFICATION APPROACH

There were four regions, six ecoregions, and seven lithology groups in the three stratification aproaches tested. In total, 17 strata were evaluated. There were poorly represented situations that affected the analyses in every region, every ecoregion, and every lithology (Table 1). Seven of the 17 strata had small sample sizes. All of the possible strata with sample sizes large enough to support the development of regressions had physical situations that were poorly represented or

absent in the database. Therefore, the data for all strata were biased to some degree due to sample size, poorly distributed data, or both.
Statistically significant models that included canopy closure were developed for three out of four regions (Table 76). A statistically significant model was developed for the north Cascades Mountain region, but it did not include canopy closure. Statistically significant models were developed for three out of six ecoregions, but only two of these included canopy closure. Statistically significant models were developed for all of the lithology groups, but three out of seven of the models did not include canopy closure. Overall, canopy closure was identified as a variable that significantly explained part of the variance in the database in 10 out of 17 strata.

Because substantial problems regarding sample size and bias associated with the distribution of variables has been identified for all strata, we do not recommend that stratification of the dataset be considered for the purposes of developing new nomographs to be incorporated into the Washington State Forest Practices regulations.

One possible exception to this recommendation is the Blue Mountain region. The analyses appear to indicate that the canopy closure required to meet a specified temperature criterion in the Blue Mountains may be substantially different from other areas of the state. These analyses, however, are based primarily on data from Oregon State. We recommend additional data collection in the Blue Mountain region of Washington State to improve on the assessment regarding the applicability of the Oregon State data. If the Oregon State data is found to be unrepresentative or even questionable, we recommend that the additional sampling effort be sufficient to allow for the development of region specific regression models. Any models that are developed should be assessed relative to nomographs for the rest of eastern Washington to determine if there are statistical differences in the canopy closure needed to meet temperature targets.

Stratification could potentially increase the predictive capabilities of the nomographs if additional data were available that addresses some of the shortcomings in the current database. It is difficult to estimate how much improvement might be gained. In some the strata with larger sample sizes reported here, the squared multiple r value is substantially larger than that calculated for the regressions that combined all the data. These higher values may be indicative of the kind of improvement that might be attained through stratification of a larger dataset that better represents the range of conditions in each stratum. On the other hand, the missing data regarding the range of conditions in a stratum tends to reduce the variability in a dataset. Therefore, the higher squared multiple r values may actually be the result of an incomplete database. Given these considerations, the improvement in the predictive capabilities of the equations with the stratification of an improved database cannot be estimated.

Table 76. Summary of sources of bias or under-represented data, identification of models incorporating canopy closure, and best squared multiple r for each of the strata evaluated.

Strata	Sources of Bias or Under-representation in the database	Statistically significant model developed?	Statistically significant model developed with canopy closure included?	Variables in best equation
Region				
Northeast WA	Low elevation situations poorly represented.	Yes	Yes, best r^2 =0.74	Canopy closure Distance from divide
North Cascades	Situations with low canopy closure at high elevation and high canopy closure at low elevation are poorly represented in the database.	Yes	No	
South Cascades	Over 40 percent of the samples came from basins less than 5000 acres in size and over 60 percent of the samples came from basins less than 10,000 acres in size.	Yes	Yes, best r^2 =0.47	Elevation Canopy closure Basin Size
Blue Mountains	Only 7 data points from the State of Washington. Regressions largely based on Oregon data. Applicability of the Oregon data appears to be good, but needs to be verified. Physical variables besides elevation and canopy closure are not available.	Yes	Yes, best r^2 =0.35	Canopy closure Elevation
Ecoregion				
M333A	Situations with high canopy closure and low elevation and situations with low canopy closure and high elevation are under-represented. Low elevation data points are sparse relative to the number of points from elevations greater than 2500 feet.	Yes	Yes, best r^2 =0.71 2^{nd} best r^2 =0.70	Distance from divide Ave Annual Precip. Canopy Closure Distance from divide Canopy closure
M242B	Eleven points in the database. The elevation range represented by the data for Ecoregion M242B is rather narrow and most of the data points are clustered around a small range of canopy closure.	No	No	

Strata	Sources of Bias or Under-representation in the database	Statistically significant model developed?	Statistically significant model developed with canopy closure included?	Variables in best equation
M242C	The majority of the data points for this ecoregion came from basins that were less than 10,000 acres in size with greater than 30% canopy closure	Yes	Yes, best r^2 =0.45 2nd best r^2 =0.41	Basin size Ave. August air temp. Canopy closure Basin size Canopy closure
-342I	High canopy closure situations at low elevations are not well represented in the data.	Yes	No	
-331A	Only one point in the dataset.	No	No	
M332G	Six data points in the dataset.	No	No	
Lithology				
Alluvium	High canopy closure situations were missing from all but the smallest basins. Precipitation was well represented for only basins that were less then 7,000 acres.	Yes	Yes, best r^2 =0.29	Canopy closure Elevation
Basalt	Situations with low canopy closure at high elevation and situations at elevations below 1000 feet are under-represented.	Yes	Yes, best r^2 =0.69	Canopy closure Elevation Ave. July air temp.
Glacial Drift	The majority of the data points are from situations with higher canopy closure and elevations ranging from 2000 to 4000 feet. Lower canopy closure (<50%) and both low and high elevations are under-represented. Additionally, 81 percent of the samples are from basins that are less than 5000 acres in size.	Yes	Yes, best r^2 =0.70	Canopy closure Ave. annual precip. Basin size
Granite	The sample size is small (n=13). The entire dataset for the granite lithology group was collected in relatively small basins. Most of the data comes from NE Washington, where situations with high canopy closure at low elevation are also poorly represented.	Yes	No	
Metamorphic	The sample size is small (19).	Yes	No	
Other Deposits	Situations at low elevation with low shade were not represented in the database.	Yes	No	

Strata	Sources of Bias or Under-representation in the database	Statistically significant model developed?	Statistically significant model developed with canopy closure included?	Variables in best equation
Sedimentary	Small sample size (14). Most of the data points lie within a 1500-foot range of elevations and there is little variation in elevation at each canopy closure level.	Yes	Yes, best r^2 =0.81	Canopy closure Stream gradient
All Data Combined	None.	Yes	Yes, best r^2 =0.54 2[nd] best r^2 =0.51	Canopy closure Elevation Ave. annual precip. Distance from divide Canopy closure Elevation Distance from divide

The most valuable stratification approach would be one that stratifies areas of eastern Washington such that the strata reflect variability in the fundamental process of heating and cooling. Stratifications that reflect fundamental differences in plant communities of riparian areas (e.g. conifers, deciduous trees, brush), air temperature, stream flow, precipitation, summer wind patterns (affect cooling), or groundwater inputs may be useful. The overwhelming majority of the riparian communities in the forested areas of eastern Washington are dominated by conifer stands. There are some areas where oaks or cottonwoods are the dominant species along streams, but these are found primarily in the marginal areas between forested and desert communities. Hence, stratification on dominant plant form is unlikely to be useful. Air temperature, stream flow, and precipitation can all be treated as continuous variables in regression analyses; hence, stratification using those parameters is unnecessary. Major differences in groundwater inputs are hard to map on a statewide scale. Groundwater inputs tend to be driven by localized geological processes. Therefore, stratification on groundwater inputs is probably not possible.

The stratifications used in the analysis included region, ecoregion, and lithology. Conceptually, ecoregion may be expected to reflect some of the fundamental differences in climate, geology, and biotic communities across eastern Washington. In practice, the ecoregions as defined reflect primarily differences in soils (Table 77). The primary eastern Washington ecoregions that support forests are the East Cascades Section, the Okanogan Highlands Section, and the Blue Mountain Section. The range of air temperatures in these three ecoregions is identical (Table 77). Average annual precipitation in all the three ecoregions ranges from very dry to very wet (although some have higher extremes than others) (Table 77). Hence, ecoregion does not separate eastern Washington into strata where there are fundamental differences in the processes affecting stream temperature.

Table 77. Primary geologic processes affecting soils, the range of average annual precipitation, and the range of average July air temperatures in the dominant Bailey's ecoregions of eastern Washington.

Ecoregion	Forested	Primary geologic processes affecting soils	Range of average annual precipitation	Range of average July air temperatures
Eastern Cascades Section	Yes	Volcanic	<10 to >120 inches	50 to 80 ^{o}F
Okanogan Highlands Section	Yes	Diverse underlying geology	<10 to 70 inches	50 to 80 ^{o}F
Blue Mountains Section	Yes	Lava flows and ash deposits	<10 to 60 inches	50 to 80 ^{o}F
Palouse Section	Minor areas	Wind driven soil deposits	<10 to 25 inches	60 to 80 ^{o}F
Columbia Basin Section	Minor areas	Large ice age floods	<10 to 25 inches	60 to 80 ^{o}F

Region, as was defined in this study, is very similar to ecoregion. The regions used in this study include the Blue Mountain region, which overlaps almost precisely with the Blue Mountain Ecoregion, Northeast Washington, which covers an area that is very similar to the Okanogan Highlands Ecoregion, and the North and South Cascades Mountains region, which are located primarily in the Eastern Cascades Ecoregion. These regions have the same range of precipitation and air temperature that is present in the ecoregions. Region, therefore, does not differentiate on fundamental differences in heating and cooling processes.

Lithology was used as a stratification approach to try to address fundamental differences in flow as it is affected by groundwater inputs and/or subduction (alluvium). The geologies in the eastern portion of the state are highly diverse and stratification on geology turned out to be impractical. The larger lithologies that were used likely include a range of groundwater inputs that reflect local scale variations in geology.

With additional information, significant differences between the strata in any of the classification systems used here may be found. In many cases, the regression equations developed in the analyses completed in this study were statistically different. In practice, however, there was little difference between the plotted lines. The major exception to this was the Blue Mountain region, which appeared to require more shade than the other areas to attain a target temperature. The data used in the Blue Mountain region came primarily from Oregon. The Oregon data was collected at a more southern latitude, which may have effected the outcome. The applicability of the Oregon data in the Washington Blue Mountains should be further tested with a larger sample size from Washington.

4.4 MOST SIGNIFICANT VARIABLES

The most common independent variables that were found to significantly contribute to the regressions were canopy closure, elevation, average annual precipitation, basin size, and distance from divide (Figure 118). Four of these parameters (canopy closure, elevation, average annual precipitation, and distance from divide) were included in the analyses of the entire database without stratification. The significance levels of all the most common independent variables listed above were high in all equations in which they appeared (with the exception of those cases where canopy closure was forced into the equation – those situations are not reflected in Figure 120).

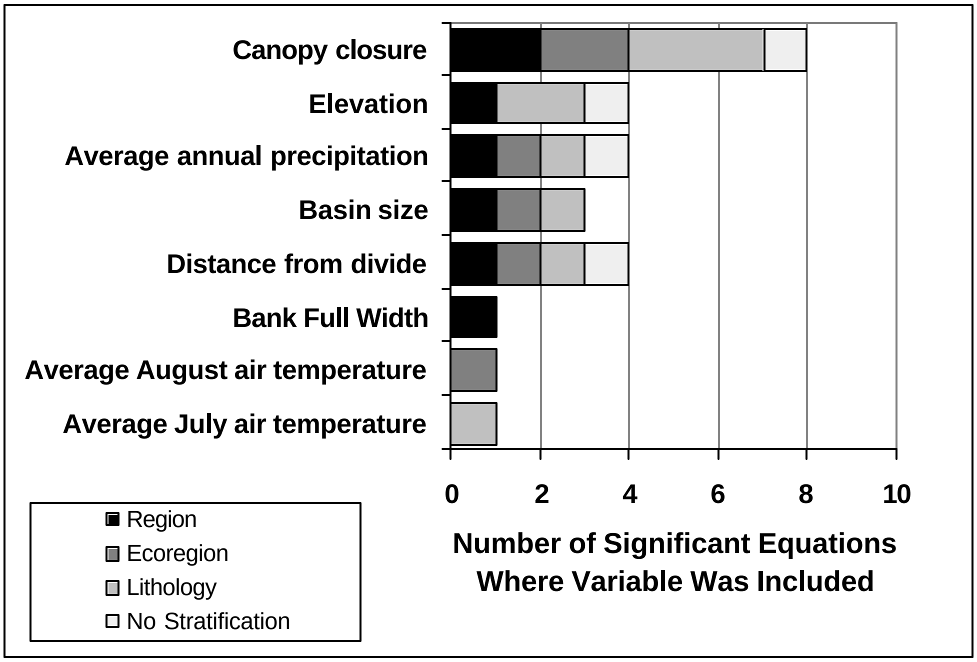


Figure 118. Frequency that each independent variable appeared in significant equations for strata with reasonably large sample sizes (Northeast WA, North Cascades, South Cascades, M333A, M242C, alluvium, glacial drift, basalt, and all data combined).

Basin size and distance from divide may reflect similar physical conditions. Basin size is the size of the drainage area upstream of the sampling point. Distance from divide is the distance from the sampling point to the watershed divide. Both should reflect a) the amount of drainage area upstream of the sample site, b) stream flow as it is affected by drainage area, c) a decrease in elevation relative to upstream points, and d) a change in air temperature also relative to a fixed upstream point. Hence, both variables arguably represent the same characteristics of a sample site. The difference between the two variables is that basin size is probably the better measure of drainage area, and subsequently stream flow. Given the similarities between these two variables, the total number of equations where they are represented may be summed, in which case some measure of drainage basin size was the second most common variable in the equations.

Three of the primary physical processes that affect cooling and heating of streams are reflected in the five most common variables and in the four variables included in the regression using all of the available data. Elevation reflects air temperature; average annual precipitation and basin size reflect stream size and possibly air temperature; and canopy closure reflects the amount of direct solar radiation. Hence, these analyses would suggest that direct solar radiation, convective exchange of heat between the air and the stream, and the volume of water are the primary factors affecting stream temperature. Other factors not reflected in the variables used in the regressions are known to be important. In particular, wind driven evaporation, the volume of groundwater inputs, and conductive heat exchange between the stream and the streambed can have a significant effect on stream temperature (Sullivan and Adams 1990). The modifying effect of these and other heating and cooling processes contribute to the unexplained variability in the equations. Variability is also introduced through the imprecise representation of water volume and air temperature by elevation, basin size, and distance from divide.

The analyses conducted in this study would have been more robust if more accurate measures of ambient air temperature and stream size were available. Unfortunately, the number of samples with measures of stream size and air temperature was very small. Therefore, the effect of stream size and air temperature could not be directly evaluated. It is noteworthy, that bank full width was found to be a significant variable in the one region where that information was available for a large number of cases. Should additional sampling be conducted in the future, it is recommended that some measure of stream size, preferably summer stream size, and air temperature be incorporated into the sampling scheme.

4.5 FIT OF DATA TO EXISTING NOMOGRAPH

The analysis found that the nomograph underestimated the amount of canopy closure required to meet the 16^{o}C and 18^{o}C temperature targets 10.5 and 9.2 percent of the time, respectively. Hence, the existing nomographs required at least as much shade as was needed to meet the targets 89.5 to 90.8 percent of the time. The region with the highest percentage of cases where the nomograph under predicted the canopy closure needed was the Blue Mountains region. This region, however, had only seven cases in the analysis. Different results may by found with a larger sample size.

The analyses comparing the data in the database to the existing nomographs suggests that both the 16^{o}C and 18^{o}C nomographs overestimate the amount of shade needed more often than they underestimate shade (Figures 119 and 120). The difference between the number of cases that were overestimated versus those that were underestimated was statistically significant at the p=0.05 level in the Northeast Washington and Southern Cascades regions and for the entire dataset as a whole.

Although the statistical tests indicate that the errors are not balanced on each side of the lines representing the existing nomographs, the graphics indicate there is little difference between the existing nomographs and the newly developed regression lines (Figures 119 and 120). In most cases, the differences in the estimated required shade are not large; however, minor modifications in the lines representing the nomographs can result in as much as a ±20% difference in required canopy closure.

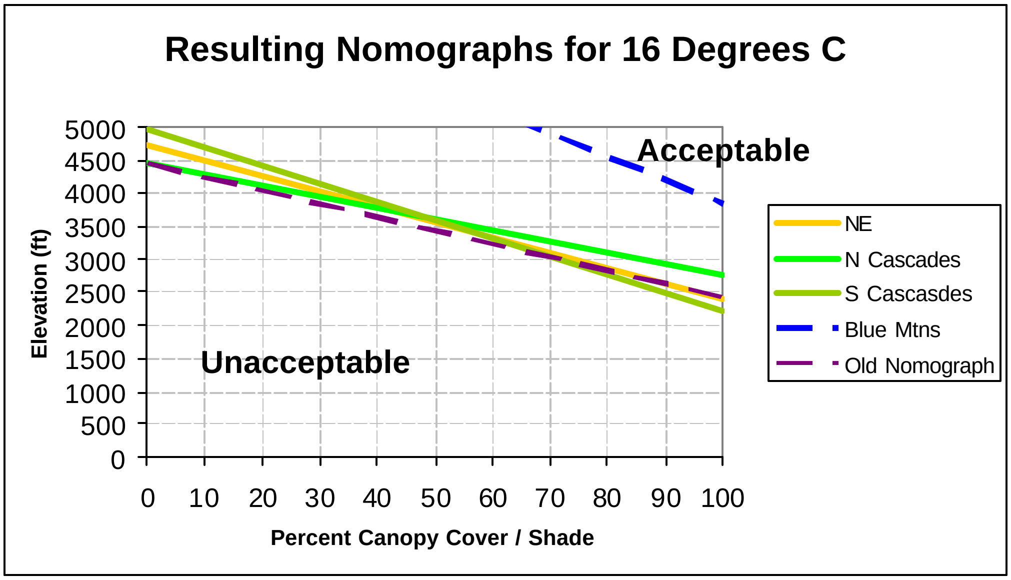


Figure 119. Predicted canopy closure and elevation combinations needed to meet a 16 degree C temperature standard for each region compared to the existing nomograph.

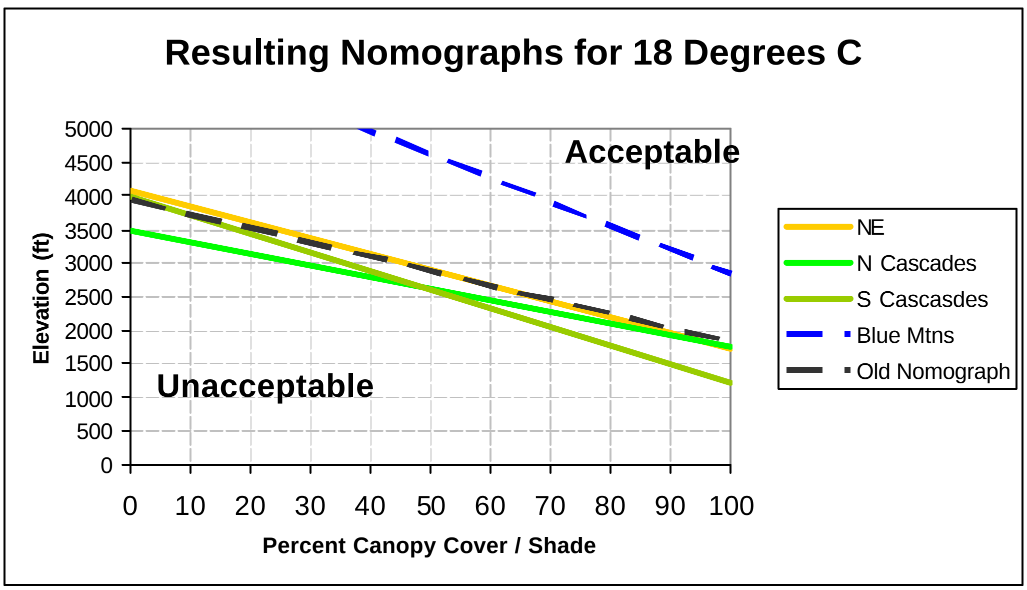


Figure 120. Predicted canopy closure and elevation combinations needed to meet a 18 degree C temperature standard for each region compared to the existing nomograph.

The reader should note that the tests to check for balanced errors tended to indicate that on average, the canopy closure required to meet the temperature standards was overestimated in most cases; but the predicted regression relationships suggest the opposite in many cases. The differences lie in the relative magnitude of difference that drives the mean values at any given elevation and canopy closure situation.

Given the difference in variables used in the existing nomographs and the variables used in the best fit models developed with the full suite of physical parameters, the predicted models with the full suite of parameters could not be compared to the existing nomographs. The Predicted equations for the Blue Mountains are clearly different from the rest of the regions. The regression equations suggest that substantial more shade is required in the Blue Mountains to meet temperature targets than is needed in other regions. The reader is reminded that the data representing the Blue Mountains was collected primarily in Oregon. Only seven data points are from the State of Washington. The results of the regressions could change substantially if more Washington data were available and the Oregon data could be eliminated from the analyses.

4.6 OPERATIONAL CONSIDERATIONS

Any regression model that accounts for direct solar radiation, ambient air temperature, and stream size must include at least three independent variables. The current nomographs published in the Washington State Forest Practices Rules incorporate only two variables and address two water temperature criteria. These variables in the current nomographs account for air temperature and solar radiation, but not stream size. With two variables and two temperature criteria, the nomographs are relatively easy to use. The two variables are plotted on two plots, one for each temperature criteria. The user identifies which criteria applies, selects the appropriate graph, and looks up the target canopy closure based using the known elevation of a site.

The new Washington State water quality standards include seven temperature criteria. If new two-variable nomographs are created to address all these criteria, seven plots will have to be developed or all seven nomgraph lines will need to be plotted on the same graph. This adds a level of complexity to the timber manager. A plot containing nomographs for all seven temperature criteria based on model with canopy closure and elevation that was developed with all the data without stratification is provided in Figure 121.

Many of the best-fit equations developed in this study incorporate three or more independent variables. Where four or more variables were found to be significant, the variables that contributed the least to the equations were identified and removed to develop regression equations with no more than three variables.

Three variable nomographs will be complex to implement. The simplest method of implementation would be to provide equations that can be used to solve for required canopy closure based upon the target temperature criterion and local measures of two physical variables.

If look-up plots are desired, the situation becomes more complex. Plots on three axes could be developed for each of the potentially applicable temperature criterion. Three-dimensional plots would have to categorize canopy closure (e.g. 60-70 percent shade) in order to be useful. A set of these plots for the 3-parameter model developed for all the data without stratification is provided in Figures 122, 123, and 124. No plot was developed for the temperature criterion of 9^{o}C because 100 percent canopy closure (or more) is needed at all locations.

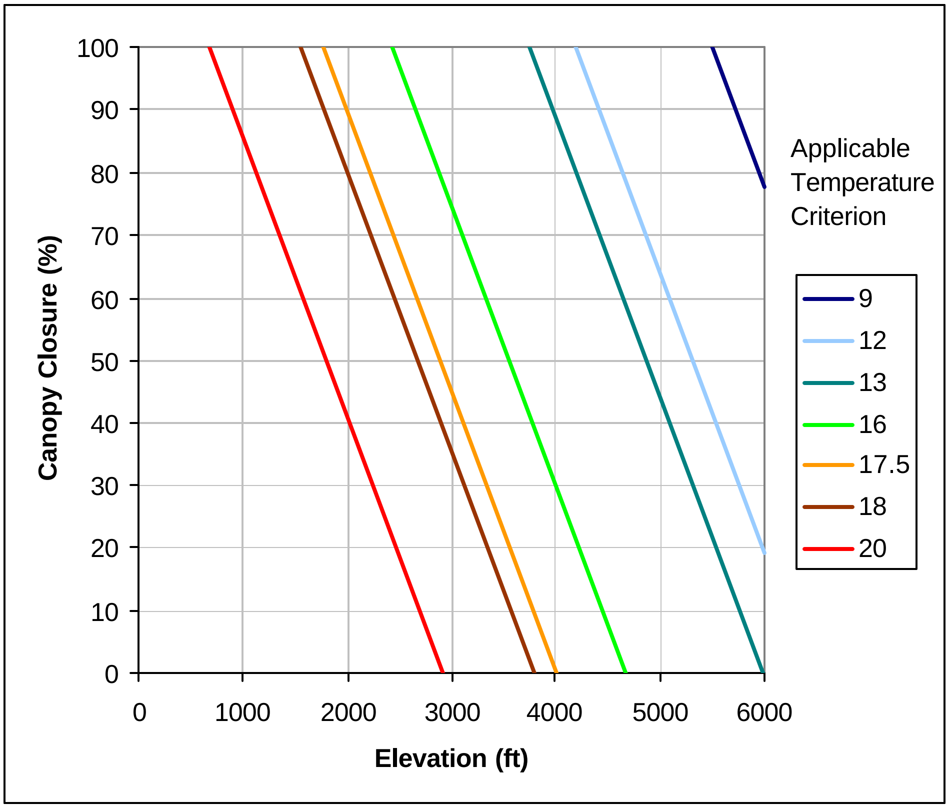


Figure 121. Example of a nomograph plot containing all seven of the new State of Washington temperature criteria. The lines on the graph were derived using the results of the regression analyses for all data combined incorporating only canopy closure and elevation as independent variables.

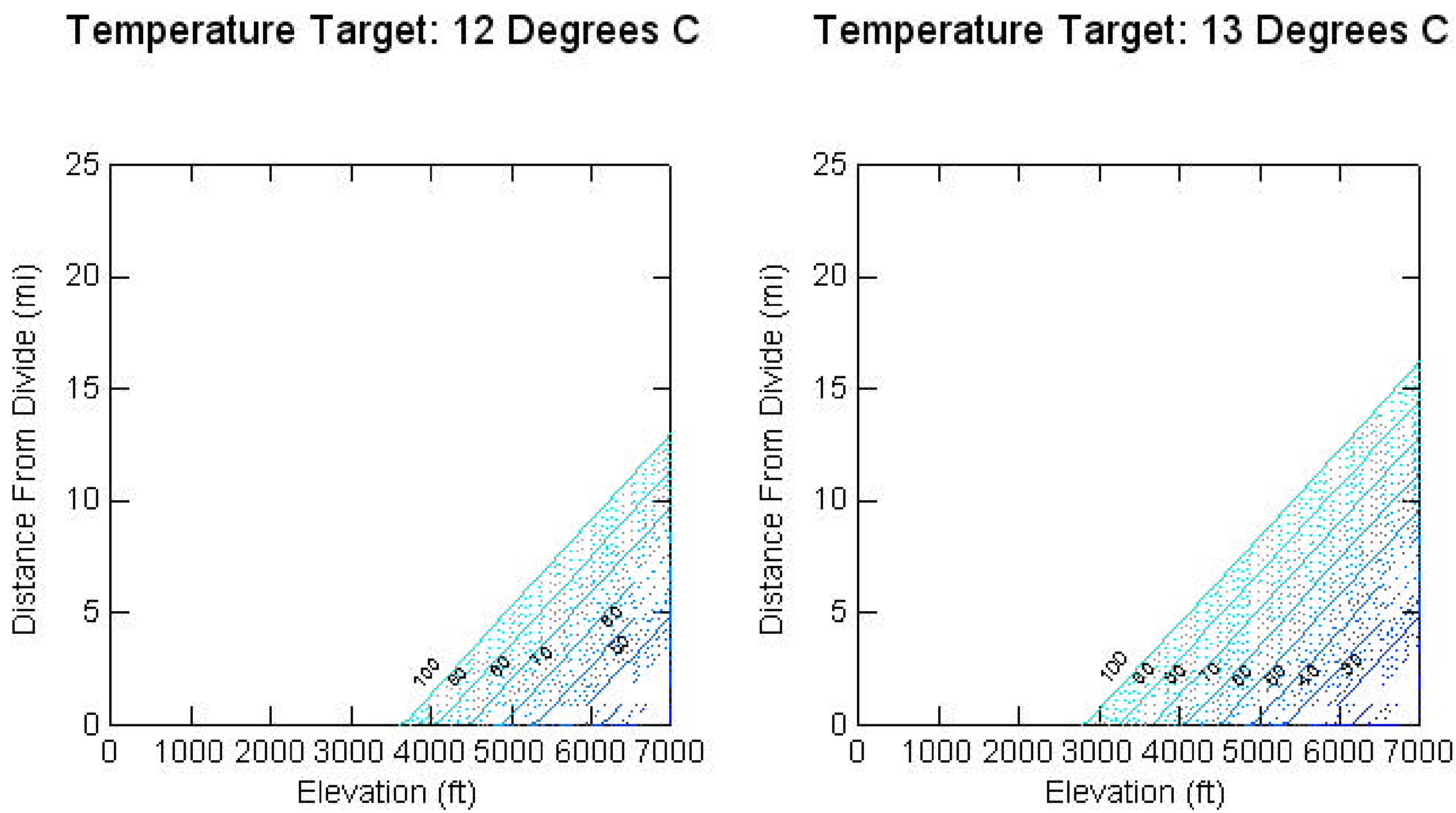


Figure 122. Contours of estimated canopy closure (%) needed to attain target temperatures of 12 and 13 degrees C. Contours based on regression results incorporating the entire database without stratification and all potential independent variables. The 3-parameter model that was developed is depicted here.

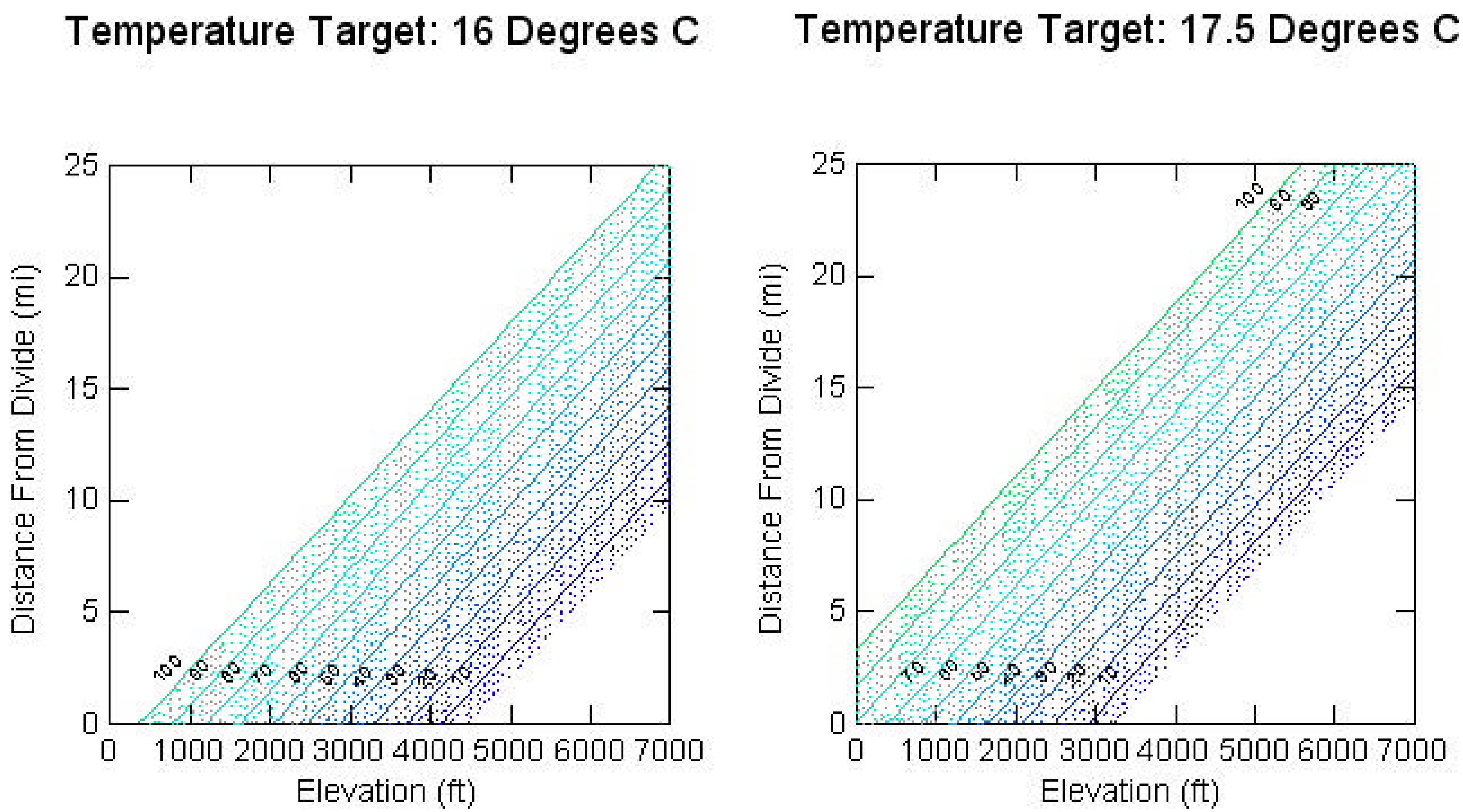


Figure 123. Contours of estimated canopy closure (%) needed to attain target temperatures of 16 and 17.5 degrees C. Contours based on regression results incorporating the entire database without stratification and all potential independent variables. The 3-parameter model that was developed is depicted here.

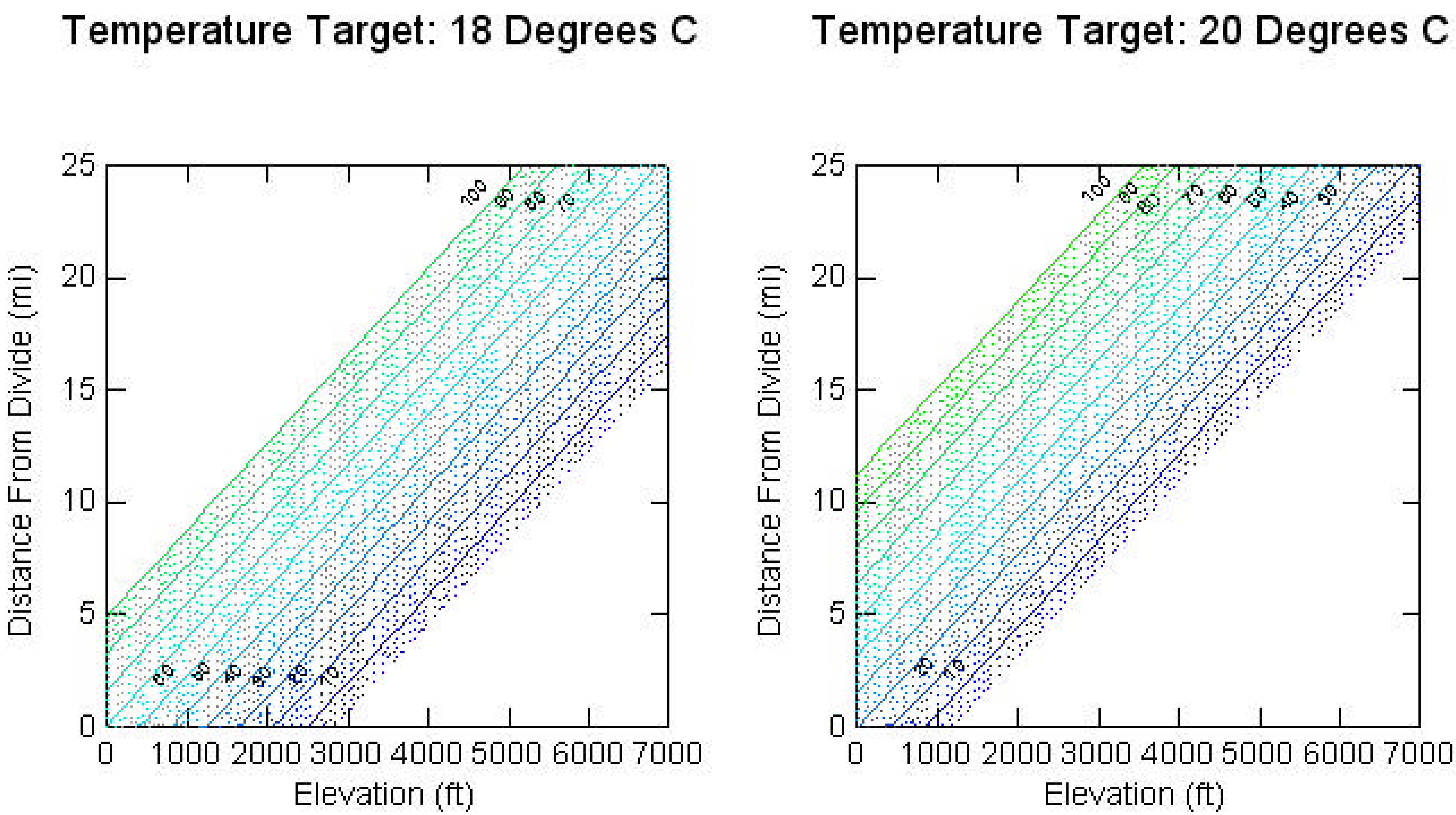


Figure 124. Contours of estimated canopy closure (%) needed to attain target temperatures of 18 and 20 degrees C. Contours based on regression results incorporating the entire database without stratification and all potential independent variables. The 3-parameter model that was developed is depicted here.

Another option is to categorize one of the physical variables other than shade and develop a separate plot for each category. This would result in a large number of plots that are potentially applicable at any one site. If one of the variables, for instance elevation, was categorized into 6 intervals, then 42 plots would have to be developed (7 plots representing each potentially applicable temperature criterion for each of 6 categories of the physical variable). Regardless of the preferred method for depicting target canopy closure, the implementation of the three variable models would be substantially more complicated than the current nomograph. Implementation of 4 or 5 parameter models would require exponentially more graphs and would probably be better handled by applying the equation directly.

If the state determines that stratification across locations is desired, then the process described above would also have to be repeated for each stratum.

4.7 SOURCES OF ERROR

4.7.1 Database

In many, if not most, of the cases where data were contributed to the database, we had to rely upon the assurances of the contributor that the data met the minimum standards for our dataset. No means of field-verifying this information was available. There is some possibility, though unlikely, that contributors misinterpreted our description of minimum standards and submitted data that did not meet those standards. Such data would tend to increase the range of variability in the database. As such, it may have affected the correlation coefficients and the confidence intervals, but we doubt that these errors would have been biased in one direction. Therefore, we do not believe that such errors would affect any estimates of mean tendency, including overall regression equations. There is some possibility that data that did not meet the minimum standards would have resulted in the elimination of marginal variables from inclusion as significant variables in regression equations.

The number of cases where we had measured data for hill slope gradient and stream gradient was sparse. A GIS estimate of these parameters was used where the data were missing. GIS estimates are acceptable in a larger scale, but at the site scale are inherently inaccurate. Therefore, use of the GIS estimates for these parameters likely introduced error into the analyses. Neither variable was included in any significant equation.

Mean annual, July, and August temperature were based on a broad scale coverage. Local variability in temperature was not reflected in the data. Precise measurements of air temperature would have improved the accuracy of the data and may have affected the results of the regression analyses.

Elevation generated by the GIS analyses was based on 10 meter DEMs. Some error is likely associated with this data due to the coarseness of the data. Given the error associated with the regression equations, the additional measurement error associated with the use of 10 meter DEMs is likely small. Likewise, the estimates of drainage area and distance form divide were likely affected by the use of the 10 meter DEMs. Again, this error is likely small relative to the error associated with the regression equations.

Several of the variables used in the regressions that incorporated all of the physical parameters in the database were auto correlated to some extent. Minimal effort was put into addressing this autocorrelation. Attempts to address autocorrelation were limited to extreme situations. In no case were interaction variables introduced to deal with the autocorrelation found to be significant.

Several of the physical parameters included in the database were poorly represented. These parameters include stream width, stream depth, stream gradient, hill slope gradient, flow, and measured air temperature at the site of data collection.

Canopy closure upstream of the data collection site should be relatively consistent for a distance upstream. Preferably, canopy closure would be represented as the average of measurements taken along the reach upstream of the site. In all but 27 cases (8.8 percent of the total number of samples), the canopy closure was correctly measured as the average of the reach. Consistency of canopy closure of the upstream reach was verified visually in the remaining 27 cases. Twenty-two (22) of these cases were located in Northeast Washington and the other five (5) were located in the northern Cascades Mountains. These cases represented roughly 1/3 of the total database for Northeast Washington and 14 percent of the cases for the North Cascades. In 47 cases in NE Washington, we had both the reach averaged canopy closure measurement and the canopy closure measurement at the point where the thermograph was deployed. The differences between the two measurements ranged from 0 to 50 and the standard deviation of the differences between the two measurements was 19.25. The differences between the two measurements indicate some error was introduced into the models through the inclusion of the data points that did not have reach averaged canopy closure measurements. The actual degree of error is unknown and dependent upon the care the collectors of the data took to find reaches with consistent shade. Both regions affected by these point measurements had small sample sizes; therefore, the data were not excluded. Larger sample sizes are recommended to both regions to improve confidence in the models. With larger sample sizes, these data point should be excluded from the analyses to reduce measurement error.

The database includes data collected over several years (1987-2003). The may be considerable variability in air temperature and/or flow between years. This variability likely contributes to some of the variability seen in the analysis results.

In many of the strata evaluated, the range of physical conditions within that stratum was not well represented. Specific information regarding which conditions are under-represented (or absent) was discussed earlier. Datasets that do not represent the entire range of conditions or are heavily weighted to a small set of conditions will tend to bias results.

The Blue Mountains have only seven data points from the Washington portion of the region. Oregon data were used to increase the sample size, but this data were collected up to 150 miles south of the Washington/Oregon border. Hence, the results of the analyses for the Blue Mountains may be affected by the difference in latitude between the Washington area of the Blue Mountains and the Oregon portion of the same area. The difference in latitude may have introduced error related to the change in air temperature with latitude. We would ask the reader to be cautious in their interpretation of the information reported herein regarding the Blue Mountains. Significant differences in the equations reported herein and equations developed using only Washington data could arise.

Small sample sizes affected the ability to develop regression equations and/or affected the confidence in the outcomes in numerous situations.

Virtually all statistical references caution against extrapolating data beyond the range of the input data. Errors associated with such extrapolation cannot be estimated. The data distribution for all tests in this document was included in the accompanying plots. In those cases where the data distribution did not cover the range of results that are predicted, the errors associated with extrapolation may be occurring.

4.7.2 Likelihood of Type I Errors

Each hypothesis that was tested was compared to a F or T statistic at the P=0.05 level. This test statistic indicates that the probability of rejecting a null hypothesis when it should have been accepted is 5 in 100. This is known as a Type I error. There are number of tests reported in this document. We haven't counted them. Assuming there are approximately 100 tests reported where the null hypothesis was rejected, we would expect that five of these were rejected in error. There is no way to identify which these may be, but it is reasonable to assume that tests where the test statistic was close to the p=0.05 level have a higher probability of being error than those that had a very small probability level for accepting the null hypothesis.

5.0 PRIMARY FINDINGS

1. The analysis found that the nomograph underestimated the amount of canopy closure required to meet the 16^{o}C and 18^{o}C temperature targets 10.5 and 9.2 percent of the time, respectively. Hence, the existing nomographs required at least as much shade as was needed to meet the targets 89.5 to 90.8 percent of the time. Both the 16^{o}C and 18^{o}C nomographs overestimate the amount of shade needed more often than they underestimate shade. The regression lines developed in this study balance the errors (are equally likely to over and under-estimate canopy closure). Graphically, there is little difference between the existing nomographs and the regression lines incorporating elevation and canopy closure as the independent variables developed by this project (Figures 119 and 120). The equations developed in this study that incorporate other variables could not be compared with the existing nomographs.

2. The primary variables that were found to be significant in the regression analyses were canopy closure, elevation, average annual precipitation, distance from divide, and basin size. Collectively, these variables are indirect measurements of direct solar radiation, ambient air temperature (convective exchange between the water and the air), and water volume. Evaporation rates, longwave radiation, conductive heat transfer between the stream and the streambed, and groundwater inputs are not accounted for in the equations. Variations in these factors contribute to the unexplained variance in the data. The imprecision of the indirect measures of ambient air temperature and water volume also contribute substantially to the unexplained variability.

3. Stratification on region, ecoregion, or lithology is not support with the existing database. All of the strata tested had small sample size or had substantial sets of physical conditions that were under-represented or absent in the database. Additional data that fills the missing information may improve the analyses. The one exception to this is the Blue Mountain area. The analyses suggest that the amount of canopy required to meet specified temperature standards may be higher in that area. The data used to develop the Blue Mountain regressions comes primarily from the Oregon Blue Mountains. Additional data is needed to confidently assess the applicability of Oregon data to the Washington Blue Mountains.

4. The regression equation using the entire database and including only canopy closure and elevations as independent variables was most similar to the existing nomographs. That equation was highly significant and explained 37 percent of the variance in the dataset. A five variable model was developed for the dataset when all possible independent variables were evaluated. The variables in this model were canopy closure, elevation, latitude, average annual precipitation, and distance from divide. This model was highly significant and explained 54 percent of the variability in the model. The two parameters that contributed the least to the 5-parameter model were removed to create a 3-parameter model. The remaining variables were elevation, canopy closure, and distance from divide. This last model was also highly significant and explained 51 percent of the variability in the data. Hence, the three-parameter model is substantially more accurate than the model that incorporated only canopy closure and elevation. However, there was little difference in the relative accuracy of the 3-parameter and 5-parameter models.

5. Information that would have the greatest effect on improvement of the analyses conducted in this study includes 1) site specific information on stream size and 2) average July air temperature data with sufficient coverage to support the development of more precise air temperature maps. With the exception of the situation in the Blue Mountains discussed above, this data would likely have a greater effect on improving the models than collecting data to fill gaps in the various strata.

6.0 RECOMMENDATIONS

1. The adoption of new temperature standards by the State of Washington raises policy questions regarding implementation of the Forest Practices rules with regards to those standards. We suggest that the implications of the change in standards be addressed by policy representatives. In recognition that the temperature criteria were under review and subject to change, the equations for all regression analyses completed in this study have been provided. With these equations, nomographs for any desired temperature standard can be created.

2. If stratification by either ecoregion or lithology group is pursued, this should be done with the awareness that some of the ecoregions and lithology groups are poorly represented in the database. Prior to adoption of either of these stratification approaches, we would recommend that additional data be collected for strata that are currently poorly represented.

3. Stratification by region, ecoregion, or lithology is not recommended at this time. All of the various strata had problems with sample size and/or data gaps that biased the results. The regressions that were conducted using the entire dataset had a large number of samples distributed widely across the range of physical conditions represented in the database.

4. Future data collection efforts should emphasize the benefits of following the TFW criteria for site selection and data collection. We are aware that a large number of samples were not submitted because the sites did not meet our minimum criteria.

5. We encourage the collection of information related to summer stream size and average July air temperatures in future stream temperature monitoring projects. Over time (years), sufficient information may be collected to allow for the development of a more precise set of maps depicting average July air temperatures and may allow for an analysis that uses direct measures of stream size.

6. The missing data that was discussed earlier in each of the strata evaluated by this project likely affected the outcome of the analyses. Additional data collection to fill these gaps may improve the regressions for the various strata. However, we feel that the data collection efforts described in item number 5 above would provide more information regarding stream heating and cooling and would be of greater benefit to any future evaluations of nomographs.

7.0 ACKNOWLEDGEMENTS

We would like to thank the Wenatchee/ Okanogan National Forest, the Okanogan Conservation District, the Yakama Indian Nation, Boise Cascade Company, Plum Creek Timber Company, the Kalispel Tribe, the Upper Columbia United Tribes, the Colville Confederated Tribes, the Entiat Ranger District, Washington State Department of Ecology, Ferry County Conservation District, Foster County Conservation District, Washington Department of Fish and Wildlife, the Campbell Group, the U.S. Fish and Wildlife Service, Washington Department of Natural Resources, and all members of the Cooperative Monitoring, Evaluation and Research (CMER) committee for their contributions of data, reviews of this document, and general support of the project.

8.0 REFERENCES

Adams, S.M. and J.E. Breck. 1990. Bioenergetics. *In*: Shreck, C.B. and P.B. Moyle, eds. Methods for fish biology. American Fisheries Society. Bethesda, MD.

Adams, T.A., and Sullivan, K. 1989. The Physics of forest stream heating: a simple model. Timber-Fish-Wildlife Report No. TFW-WQ3-90-007, Washington Department of Natural Resources, Olympia, Washington.

Bailey, R. G.; Avers, P. E.; King, T.; McNab, W. H., eds. 1994. Ecoregions and subregions of the United States (map). Washington, DC: USDA Forest Service. 1:7,500,000. With supplementary table of map unit descriptions, compiled and edited by W. H. McNab and R. G. Bailey

Bartholow, J. 2003. Modeling uncertainty: quicksand for water temperature modeling. Hydrological Science and Technology 19: 221-232.

Bartholow, J.M. 1989. Stream temperature investigations: field and analytic methods. Instream Flow Information Paper No. 13. U.S. Fish Wildl. Serv. Biol. Rep. 89(17). 139 pp.

Berman, C.H. 1998. Oregon Temperature Standard Review, U.S. Environmental Protection Agency, Region 10, Seattle, WA, 63 pp.

Beschta, R.L. and R.L. Taylor. 1988. Stream temperature increases and land use in a forested Oregon watershed. Water Res. Bull., 24(1):19-26.

Beschta, R.L., R.E. Bilby, G.W. Brown, L.B.Holtby, and T.D. Hofstra. 1987. Stream Temperature and Aquatic Habitat: Fisheries and Forestry Interactions. Chapter six in E.O. Salo, T.W. Cundy (eds). 1987. Streamside Management: Forestry and Fishery Interactions. Univ. of Wash. Inst. of Forest Resources, Contribution No. 57.

Brown, G. W. 1969. Predicting temperature of small streams. Water Res. Research. 5(1):68-75.

Brett, J.R. Energetics. Pp. 3-68. *In*: C. Groot, L. Margolis, and W.C. Clarke (eds). Physiological Ecology of Pacific Salmon UBC Press, Vancouver, B.C.

Brown, G.W. 1983. Forestry and water quality. Oregon State University Bookstore, Inc., Corvallis, OR. 142 pp.

Brown, G.W. and J.T. Kryier. 1970. Effects of clearcutting on stream temperature. Water Resour. Res., 6(4):1133-1139.

Burns, J.W. 1972. Some effects of logging and associated road construction on northern California streams. Trans. Am. Fish. Soc. 101:1-17.

Caldwell, J.E., K. Doughty, and K. Sullivan. 1991. Evaluation of downstream temperature effects of type 4/5 waters. Timber, Fish and Wildlife Report. TFW-WQ5-91-004.

Coutant, C.C. 1977. Compilation of temperature preference data. J. Fish. Res. Board Can. 34:739-745.

Crittenden, R.N. 1978. Sensitivity analysis of a theoretical energy balance model for water temperatures in small streams. Ecol. Model. 5:207-224.

Edinger, J.E., Duttweiler, D.W. and Geyer, J.C., 1968. The response of water temperatures to meteorological conditions. Water Resources Research, Vol. 4, No. 5.

Everson, L.B. 1973. Growth and food consumption of juvenile coho salmon exposed to natural and elevated fluctuating temperatures. M.S. Thesis. Oregon State unive3rsity, Corvallis, OR. 68 pp.

Feller, M.C. 1981. Effects of clearcutting and slashburning on stream temperature in southwestern British Columbia. Water Res. Bull. 17(5):863-867.

Hatten, J.R. and R. Conrad. 1995. A comparison of summer stream temperatures in unmanaged and managed sub-basins of Washington's Western Olympic Peninsula. Northwest Fishery Resource Bulletin. Project Report Series No.4. 52p.

Holtby, L.B. and C.P. Newcombe. 1982. A preliminary analysis of logging-related temperature changes. pgs 81-99 In G.F. Hartman (ed) *Proceedings of the Carnation Creek Workshop, A 10 year Review*. Malaspina College. Nanaimo, B.C. 404 pp.

Hynes, H.B.N. (1970) The ecology of running waters. Toronto, Canada: University of Toronto Press.

Isaak, D. J., and W. A. Hubert. 2001. A hypothesis about factors that affect maximum summer stream temperatures across montane landscapes. Journal of the American Water Resources Association 37:351-366.

Jobling, M. 1981. Temperature tolerance and the final preferendum: rapid methods for the assessment of optimum growth temperatures. J. Fish. Biol. 19:439-455.

Keenleyside, M.H.A. and W.S. Hoar. 1954. Effects of temperature on the response of young salmon to water currents. Behavior 7:77-87.

Lantz, R.L. 1971. Influence of water temperature on fish survival, growth, and behavior. In J.T. Krygier and J. D. Hall (eds) Forest land uses and stream environment Proceedings of a symposium, p. 182-193. Oregon State University, Corvallis, OR.

Levno, A. and J. Rothacher. 1967. Increases in maximum stream temperatures after logging old growth Douglas-fir watersheds. USDA For. Serv. Res. Note. PNW-65. Pac. Northwest For. And Range Exp. Stn., Portland, Oregon. 12 pp.

Levno, A. and J. Rothacher. 1969. Increases in maximum stream temperature after slash burning in a small experimental watershed. USDA For. Serv. Res. Note PNW-110. Pac. Northwest For. And Range Exp. Stan., Portland, OR

Lewis, T., D. Lamphear, D. McCanne, A. Webb, J. Krieter, and W. Conroy. 2000. Regional assessment of stream temperatures across Northern California and their relationship to various landscape-level and site-specific attributes. Pages 420. Forest Science Project, Humboldt State University Foundation, Arcata, CA.

McIntosh, B.A., D. M. Price, C.E. Torgerson, and H.W. Li. 1996. Distribution, habitat utilization, movement patterns, and the use of thermal refugia by spring chinook in the Grande Ronde, Imnaha, and John Day basins. Progress Report. Portland, OR: Bonneville Power Administration, Project No. 93-700. 16 pp.

Mohseni O. and Stefan, H. G., 1999. Stream Temperature/Air Temperature Relationship: A Physical Interpretation. Journal of Hydrology, 218(3-4):128-141.

National academy of Sciences/National Academy of engineering. 1972. Water quality criteria. A report of the committee on Water Quality Criteria. National Acad. Sci/Nat’l Acad. End. Us Gov’t Printing Office, WA D.C. 594 pp.

Pilgrim, J.M., X. Fang, and H.G. Stefan, 1998: Stream temperature correlations with air temperatures in Minnesota: implications for climate warming. Journal of the American Water Resources Association, 34, 1109–1121.

Quinn, TP, DJ Adams. 1996. Environmental changes affecting the migratory timing of American shad and sockeye salmon. Ecology 77:1151-1162

Sinokrot, B.A., and Stefan, H.G. 1994. Stream water-temperature sensitivity to weather and bed parameters. J. Hydraulic Eng. 120(6):722-735.

Smith, K., and M.E. Lavis. 1975. Environmental influences on the temperature of a small upland stream. Oikos 26(2):228-236.

Steele, R.G.D., and J.H. Torrie. 1980. Principles and Procedures of Statistics. A Biometrical Approach. McGraw-Hill Book Company. New York.

Sullivan, K. and T.A. Adams. 1990. The physics of stream heating: An analysis of temperature patterns in streams environments based on physical principles and stream data. Weyerhaeuser Technical Report. 044-5002.90/2. Technology Center, Tacoma, WA 98477.

Sullivan K., J. Tooley, K. Doughty, J.E. Caldwell, and P. Knudsen. 1990. Evaluation of prediction models and characterization of stream temperature regimes in Washington. Timber/Fish/Wildlife Rep. No. TFW-WQ3-90-006. Washington Dept. Nat. Resources, Olympia, Washington. 224 pp.

Sullivan, K., D. Martin, R. Cardwell, J. Toll, and S. Duke. 2000. An analysis of the effects of temperature on salmonids of the Pacific Northwest with implications for selecting temperature criteria. Pages 192. Sustainable Ecosystems Institute, Portland, OR.

Walling, D.E. and B.W. Webb. 1992. Water quality. I. Physical characteristics. Pages 48-72 in P. Calow and G.E. Petts (eds). The River Handbook. Volume 1.

Zar, J.H. 1974. Biostatistical Analysis. Prentice-Hall Biological Sciences Series. W.D. McElroy and C.Pl Swanson (eds.). Prentice-Hall, Inc. Englewood Cliffs, NJ