

THE IMPACT OF DATA GRANULARITY AND STREAM CLASSIFICATION ON
TEMPERATURE GRADIENT MODELING IN MICHIGAN'S STREAMS

By

HALIL IBRAHIM DERTLI

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife—Master of Science

2021

ABSTRACT

THE IMPACT OF DATA GRANULARITY AND STREAM CLASSIFICATION ON TEMPERATURE GRADIENT MODELING IN MICHIGAN'S STREAMS

By

HALIL IBRAHIM DERTLI

Stream temperature is an important parameter of water quality and developing models capable of reliable predictions are critically important in stream management. In addition to the structure of these models (e.g., predictive variables), there are other factors that may influence model performance such as the selection of data granularity (i.e., level of temporal aggregation) and seasonal coverage of data collection. Data granularity and seasonal extent of data collection vary widely in the literature and have often been arbitrarily selected in stream temperature modeling studies in the past, but the consequences of these choices have not been explored. I applied different data granularity and time period selections to regression models, which were developed by Andrews (2019) to predict temperature gradient (i.e., stream temperature change) in Michigan's streams. Applying higher data granularity increased overall model performances and changed model selection results, however applying different time periods did not have a substantial effect on model performances. Using higher data granularity also changed model parameter estimates by increasing the multicollinearity in best-fitting models. In addition to temporal data granularity, data may be pooled spatially across streams within a thermal class to reduce the costs of data collection. I examined the impact of stream classification on model performance by applying data pooling within stream classes. Stream-Specific Models had better performance compared to Class-Based Models. Additional analyses suggested that classifying streams based on temperature gradient instead of stream temperature may result in better Class-Based Model performance.

Copyright by
HALIL IBRAHIM DERTLI
2021

This thesis is dedicated to my beloved sister.

ACKNOWLEDGEMENTS

My research was funded by Republic of Turkey Ministry of Agriculture and Forestry and Michigan State University. Data collection was funded by Michigan Department of Natural Resources.

I thank my academic advisor Dr. Daniel Hayes for giving me the opportunity to pursue my master's degree. Without his endless patience and encouragement, this thesis would have not been possible. I sincerely thank my committee members Dr. Troy Zorn and Dr. Scott Peacor for their advice and contribution to my research. I also thank Ryan Andrews for the great work in his research and on data collection that inspired and shaped my study.

I greatly appreciate my colleagues Trey McClinton, Tyler Firkus, Katie Kierczynski, Samantha Strandmark, Nicole Watson, Paige Filice, Janice Albers, Alexandria Walus, Samantha Tank and Rob Hunter for their precious friendship and endless support. I am also indebted to all members of a great family, Michigan State University – Department of Fisheries and Wildlife.

I greatly thank the faculty members of Middle East Technical University – Department of Biological Sciences: Dr. Musa Doğan, Dr. Can Bilgin and Dr. Meryem Beklioğlu, who encouraged and inspired me to be a part of the field of ecology and environmental sciences. I am deeply grateful to my parents: Gönül Dertli and Mahmut Dertli and my family, who endlessly supported me in my academic journey. Also, special thanks to Selena Lovell, who patiently supported me against the challenges of graduate school and research, for her invaluable partnership.

TABLE OF CONTENTS

LIST OF TABLES..... viii

LIST OF FIGURES..... x

CHAPTER 1: THE IMPACT OF DATA GRANULARITY ON TEMPERATURE GRADIENT MODELING IN MICHIGAN’S STREAMS..... 1

INTRODUCTION..... 1

Physics Behind Temperature Gradient in Streams..... 4

Modeling Stream Temperatures..... 7

A Challenge for Modeling: Selection of Time Period and Data Granularity.....10

Purpose of the Study..... 15

METHODS..... 15

Study Site..... 15

Data Collection..... 16

Revising the Data, Applying Data Granularity, and Testing Linearity..... 18

Comparisons of Goodness of Fit for Each Data Granularity Scenario..... 19

Multicollinearity Diagnosis and Response of Parameter Estimates to Data Granularity. 20

Evaluating Model Performances by Using July-Restricted and June-October Data..... 21

RESULTS..... 21

Data Granularity Influenced Model Predictive Power and Model Weight..... 21

Data Granularity Leads to Instability of Parameter Estimates in Best Fitting Model..... 25

Data Granularity Increased Multicollinearity in Raw Data..... 31

Using July-Restricted Data Did Not Improve Model Prediction Power..... 35

DISCUSSION..... 38

How Does Model Performance and Choice Vary with Data Granularity?..... 39

What are the Possible Reasons for Model Performance and Selection Changes with Data Granularity?..... 43

How Do Models Perform with July-Restricted Data?..... 46

CONCLUSIONS AND IMPLICATIONS..... 48

APPENDICES..... 51

APPENDIX A: Tables..... 52

APPENDIX B: Figures..... 63

APPENDIX C: Model Parameter Calculation..... 78

APPENDIX D: RStudio Codes..... 80

BIBLIOGRAPHY..... 81

CHAPTER 2: THE EFFECT OF STREAM THERMAL CLASSIFICATION AND DATA POOLING ON TEMPERATURE GRADIENT MODELING IN MICHIGAN’S STREAMS... 88

INTRODUCTION..... 88

A Challenge in Stream Management: Limited Data..... 88

Recent History of Stream Classification..... 89

Data Pooling, Model Generalization and Stream Management Practices..... 91

METHODS..... 93

Study Site and Data Collection..... 93

Stream Classification and Model Performance..... 93

Obtaining and Evaluating Models..... 94

RESULTS..... 96

Pooling Data Changed Model Dynamics and Model Outcomes..... 96

Classifying Streams Reduced Overall Model Performance with July-Restricted Data... 104

DISCUSSION..... 107

What is the Effect of Data Pooling on Model Dynamics?..... 108

Does Stream Classification Improve Model Performance?..... 108

Do Models Work Better for Warm or Cold Streams?..... 110

Does Using July-Restricted Data Change Model Performance?..... 113

CONCLUSIONS AND IMPLICATIONS..... 114

APPENDIX..... 117

BIBLIOGRAPHY..... 126

LIST OF TABLES

Table 1.1. Mean adjusted correlation (R^2) values of each model by data granularity across all streams with June-October data.....	22
Table 1.2. Percentage (%) of streams where each model had the highest model weight (w) across levels of data granularity. June-October data were used in models.....	24
Table 1.3. Mean adjusted correlation (R^2) values of each model by data granularity across all streams with July 2016 data.....	35
Table 1.4. Stream temperature models (Magnusson et. al. 2012; Andrews 2019).....	52
Table 1.5. Intercept and parameter estimate values of Model 10 across different data granularity. Tobacco River June – October 2016 data was used.....	53
Table 1.6. Streams and rivers with their regions (SLP: Southern Lower Peninsula; NLP: Northern Lower Peninsula; UP: Upper Peninsula), thermal classes, upstream latitudes, upstream longitude, downstream latitude, and downstream longitude (Zorn et. al. 2008; Andrews 2019).....	54
Table 1.7. Starting and ending day of year for sampling in each stream for 2016.....	55
Table 1.8. Correlation matrix of β values with hourly data. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.....	56
Table 1.9. Correlation matrix of β values with 2-hour data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.....	57
Table 1.10. Correlation matrix of β values with 6-hour data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.....	58
Table 1.11. Correlation matrix of β values with 12-hour data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.....	59
Table 1.12. Correlation matrix of β values with daily data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.....	60
Table 1.13. Correlation matrix of β values with weekly data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.....	61
Table 1.14. Mean adjusted correlation (R^2) values from July-restricted and June-October across models. Student t-test was used to find p-values.....	62

Table 2.1. Intercepts and parameter estimates from Stream-Specific models (SSMs) applied to each stream for June – October hydrological data..... 97

Table 2.2. Parameter estimates of Class-Based and Global Based models. June-October 2016 data were used.....98

Table 2.3. Bias values (B) and their average (\bar{B}) of Stream-Specific models (SSM), Class-Based models (CBM) and Global-Based model (GBM) predictions. June- October 2016 data were used.....99

Table 2.4. Mean Pearson’s correlation coefficient values of SSM, CBM and GBM across stream classes. June-October 2016 data were used.....100

Table 2.5. Mean Pearson’s correlation coefficient (r) values of SSM, CBM and GBM across stream classes. July 2016 data were used.....104

Table 2.6. Mean observed and predicted temperature gradient values, absolute bias values of Stream-Specific models (SSM), Class-Based models (CBM) and Global-Based model (GBM) predictions. June- October 2016 data were used.....118

Table 2.7. Mean downstream temperatures of streams with June-October and July-restricted data for year 2016. The stream classes are based on Zorn et. al. (2008), cold (C): July Mean Temperature (JMT) ≤ 17.5 °C; Cold-Transitional: 17.5 °C < JMT ≤ 19.5 °C; Warm-transitional: 19.5 °C < JMT ≤ 21.0 °C; Warm (W) = JMT > 21.0 °C. Streams were assigned to their classes based on their mean JMT values from 30-years of data (Zorn et. al. 2008).....119

LIST OF FIGURES

Figure 1.1. Components of heat energy budget in streams.....	5
Figure 1.2. Components of water budget in streams. Heat energy budget forms the downstream discharge.....	6
Figure 1.3. The locations of 16 streams that were selected for this study.....	16
Figure 1.4. Mean adjusted correlation (R^2) values of all streams (June-October 2016) based on different data granularity scenarios and different models.....	22
Figure 1.5. The percentage of the models having the highest model weight at least one stream for each data granularity with June-October 2016 data.....	24
Figure 1.6. Response of $\bar{\beta}$ to data granularity. Model 10 was used with June-October 2016 data. β values of all streams were averaged. Q_{Up} : upstream discharge; $Q_{Down} - Q_{Up}$: difference between downstream and upstream discharge; $T_{Air} - T_{Up}$: difference between air temperature and upstream temperature.....	26
Figure 1.7. Response of $\bar{\beta}$ to data granularity. Model 10 was used with June-October 2016 data. β values of all streams were averaged. S : day length; α : altitude angle; ΔT_{Up} : upstream heat flow; ΔT_{Base} : baseflow heat flow; ΔT_{Over} : Overflow heat flow.....	27
Figure 1.8. Correlation (r) between β values across all streams with hourly data granularity. Model 10 was used with June-October 2016 data. The color and the size of circles indicate the sign and the numerical value of correlation.....	28
Figure 1.9. Correlation (r) between β values across all streams with 2-hour data granularity. Model 10 was used with June-October 2016 data.....	29
Figure 1.10. Correlation (r) between β values across all streams with 6-hour data granularity. Model 10 was used with June-October 2016 data.....	29
Figure 1.11. Correlation (r) between β values across all streams with 12-hour data granularity. Model 10 was used with June-October 2016 data.....	30
Figure 1.12. Correlation (r) between β values across all streams with daily data granularity. Model 10 was used with June-October 2016 data.....	30
Figure 1.13. Correlation (r) between β values across all streams with weekly data granularity. Model 10 was used with June-October 2016 data.....	31
Figure 1.14. The amount of correlation (r) between environmental variables in Tobacco River with hourly (a) and 2-hour (b) June-October 016 data shown in correlogram.....	32

Figure 1.15. The amount of correlation (r) between environmental variables in Tobacco River with 6-hour (a) and 12-hour (b) data granularity. Model 10 with June-October 2016 data was used.....	33
Figure 1.16. The amount of correlation (r) between environmental variables in Tobacco River with daily (a) and weekly (b) data granularity. Model 10 with June-October 2016 data was used.....	34
Figure 1.17. Mean adjusted correlation (R^2) values of each data granularity scenarios based on all regression models. Whiskers represent standard errors of sample.....	36
Figure 1.18. Mean adjusted correlation (R^2) values of models based on averaging all data granularity scenarios. Lines represent mean adjusted correlation values obtained by using July restricted data (blue) June-October data (orange).....	37
Figure 1.19. Mean adjusted correlation (R^2) values of models across all streams with July 2016 data across data granularity scenarios.....	38
Figure 1.20: Air temperature – downstream temperature ($T_a - T_w$) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....	63
Figure 1.21. Upstream discharge (Q_{Up}) (cubic meters per second – CMS) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....	64
Figure 1.22. Upstream discharge – downstream discharge ($Q_{Up} - Q_{Down}$) (cubic meters per second – CMS) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....	65
Figure 1.23. Day length (S) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....	66
Figure 1.24. Altitude angle (α) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....	67
Figure 1.25. Upstream heat flow (ΔT_{Up}) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....	68
Figure 1.26. Baseflow heat flow (ΔT_{Base}) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....	69

Figure 1.27. Overflow heat flow (ΔT_{Over}) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.....70

Figure 1.28. Observed and predicted temperature gradient ($^{\circ}\text{C}$) of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016. Predictions were obtained from Model 10.....71

Figure 1.29. Mean adjusted correlation (R^2) values of models based on averaging all data granularity scenarios with June-October 2016 data.....72

Figure 1.30. Mean adjusted correlation (R^2) values of each data granularity scenarios based on all regression models with June-October data.....73

Figure 1.31. Parameter estimate (β) values of some predictor variables across streams with hourly June-October 2016 data. Model 10 was used to obtain β values for each stream.....74

Figure 1.32. Parameter estimate (β) values of some predictor variables across streams with weekly June-October 2016 data. Model 10 was used to obtain β values for each stream.....75

Figure 1.33. Air temperature across time with hourly (a), daily (b) and weekly (c) data granularity. Tobacco River July 2016 (July 1 – July 31) data were used.....76

Figure 1.34. Adjusted correlation (R^2) values across data granularity. Model 10 was used with June-October 2016 data.....77

Figure 2.1. Classification of streams and rivers at a national scale based on annual stream regimes (from Maheu et. al. 2016).....90

Figure 2.2. The absolute value of biases averaged for each stream class. The higher the mean absolute bias, the higher the overall mean temperature gradient prediction deviates from the overall mean observed temperature gradient.....100

Figure 2.3. Pearson’s correlation coefficient (r) values of SSM, CBM. GBM across mean downstream temperatures from June-October 2016.....102

Figure 2.4. Bias (B) versus mean downstream temperature. June – October 2016 data were used.....102

Figure 2.5. Pearson correlation coefficient (r) values of SSM, CBM. GBM across mean observed temperature gradient from June-October 2016.....103

Figure 2.6. Bias (B) versus mean observed temperature gradient. June – October 2016 data were used.....104

Figure 2.7. Mean Pearson’s correlation coefficient (r) of Class-Based Models with June- October 2016 data and July 2016 data.....105

Figure 2.8. Pearson correlation coefficient (r) values of SSM, CBM. GBM across mean downstream temperatures from July 2016.....106

Figure 2.9. Pearson correlation coefficient (r) values of SSM, CBM, GBM across mean observed temperature gradient from July 2016.....107

Figure 2.10. Observed and predicted temperature gradient ($^{\circ}\text{C}$) from Stream-Specific, Class-Based, and Global-Based models. Cedar Creek (cold) June-October 2016 data were used.....120

Figure 2.11. Observed and predicted temperature gradient ($^{\circ}\text{C}$) from Stream-Specific, Class-Based, and Global-Based models. Tobacco River (cold-transitional) June-October 2016 data were used.....121

Figure 2.12. Observed and predicted temperature gradient ($^{\circ}\text{C}$) from Stream-Specific, Class-Based, and Global-Based models. Escanaba River (warm-transitional) June-October 2016 data were used.....122

Figure 2.13. Observed and predicted temperature gradient ($^{\circ}\text{C}$) from Stream-Specific, Class-Based, and Global-Based models. Prairie River (warm) June-October 2016 data were used.....123

Figure 2.14. Average Pearson correlation coefficient (r) based on stream classes. June – October data were used.....124

Figure 2.15. Mean Pearson correlation coefficient (r) values were averaged based on stream classes. July 2016 data were used.....125

CHAPTER 1: THE IMPACT OF DATA GRANULARITY ON TEMPERATURE GRADIENT MODELING IN MICHIGAN'S STREAMS

INTRODUCTION

Freshwater ecosystems are a priority of conservation efforts since they are more prone to lose their biodiversity compared to terrestrial ecosystems (Sala et. al. 2000). In addition to their ecological importance, freshwater resources are very important for humans as they constitute only 0.01% of the total water budget in the world (Dudgeon et. al. 2006). It is known that these critical water systems and their biodiversity show regional differences in their reactions to environmental changes based on their unique environmental conditions. For example, Carpenter et al. (1992) predicted that the biodiversity in high altitude and latitude streams is more susceptible to decline when compared to biodiversity in tropical and temperate streams due to alterations in stream temperature patterns, mostly based on climate change and changes in land cover (e.g., Woltemade and Hawkins 2016). In addition to climate change and land cover changes, an important driver of stream temperature is the amount of groundwater input (e.g., Woltemade and Hawkins 2016), a factor that is vulnerable to human alteration by groundwater withdrawal. Climate change, land cover change, and groundwater withdrawal occur across the globe, but manifest themselves in changes to water temperature as a local scale. This is exemplified with the statement: “Water resources are a global problem with local roots” by Raymond Nace from U.S. Geological Survey (Nace 1967).

Stream temperature is one of the most important aspect of riverine systems as all freshwater organisms and their life cycles are affected by it. Therefore, the effect of water temperature has been well studied with a long history of investigation. For example, the effect of stream temperature on aquatic plants and their photosynthesis rates is well explained by Iversen

(1971) and Sand-Jensen (1989). They showed that while light availability is the main driving factor of photosynthesis, stream temperature can change the structure of the primary producer community especially in pools and slow-flowing streams in addition to littoral zones because of the lack of vertical mixing. In addition to the direct effect on the growth rate of primary producers by changing the rate of photosynthesis, water temperature can also change the chemistry of water by changing the solubility of water chemicals (Wetzel 1960).

In addition to primary producers, there have been numerous studies on aquatic invertebrates, with documented changes to drift behavior (e.g., Wojtalik and Waters 1970; Jackson et. al. 2007), and production (e.g., Galbraith and Vaughn 2009). Patrick et. al. (2019), for example, revealed the relationship between stream invertebrate production and hydrological characteristics of streams in a global scale. They used estimates of secondary production of stream invertebrates from 164 sites distributed globally. Secondary production is particularly important because it is considered as a main determinant of dynamics in higher trophic levels. By using their metamodel, they concluded that stream temperature had the highest overall effect on annual community secondary production among other environmental covariates (e.g., latitude, elevation, forest cover, monthly discharge). Although the streams may have unique hydrological characteristics and biota, this study posed an overall picture of how stream temperature affects invertebrate biomass in streams from a global perspective.

Fish have also been a focus of many studies, and the effect of water temperature on fish distribution, productivity and survival is well-understood. For example, the effect of water temperature on fish physiology is well explained by Ficke et. al. (2007) who described the relationship between fish metabolic rate and water temperature. They also emphasize that the effect of water temperature occurs even at the cellular level as the stability of proteins varies with

temperature. Since fish physiology responds strongly to water temperature, it can be concluded that water temperature directly affects fish reproduction and survival. In addition, fish community structure can also change with water temperature. In a recent study, Morales-Marin et. al. (2019) modelled the distribution of Athabasca Rainbow Trout, *Oncorhynchus mykiss*, which is considered as a species at risk, by using predicted future stream temperatures in Athabasca River basin, AB, Canada. Using the rainbow trout water temperature tolerance ranges and predicted distribution of water temperature in the basin, they concluded that the changing temperatures would constrain the Rainbow Trout to the Northern parts of the basin and this can potentially change the fish community structure by opening new niche areas for other fish species.

The effect of stream temperature and water withdrawal on fish distribution and growth in Michigan has been observed in several recent studies (Zorn et. al. 2004; Wehrly et. al. 2007; Nuhfer et. al. 2017). For example, Nuhfer et. al. (2017) observed that reductions in discharge did not have a significant effect on brook trout density, but spring-to-fall growth of fish declined significantly under 75% or more discharge reductions. They also observed that warming rates increased with increased water withdrawal, but the change in temperature was relatively small because the reach was quite short (602 m). However, they predicted that the increase in water temperature that would be caused by 90% flow reduction would have eliminated over 80% of habitable areas for brook trout in the whole river system.

As stream temperature is critical for riverine systems, it is important to understand the physical processes that drive and affect stream temperatures. Therefore, the following section is devoted to describing those processes and environmental variables.

Physics Behind Temperature Gradient in Streams

The change in water temperature between two points in a stream (which I will refer to hereafter as temperature gradient) is determined by several environmental factors or processes. Four of the main processes influencing temperature gradient are radiative energy exchange, conduction, evaporation, and direct changes due to input or loss of water to the stream (Figure 1.1). Radiative energy exchange occurs via incoming solar radiation (i.e., shortwave radiation), longwave radiation that is mainly emitted by the water body, and back radiation that includes reflected solar radiation by the water body (Cheng and Wiley 2016). Heat transfer via conduction occurs between the river base and the water body and between the water body and the atmosphere. Evaporative heat loss can occur in streams but is generally thought to be a minor component in the overall heat budget (Cheng and Wiley 2016). Finally, the heat energy contained in incoming surface water and groundwater contribute to temperature gradient by directly adding water with a potentially different temperature signature than the stream itself.

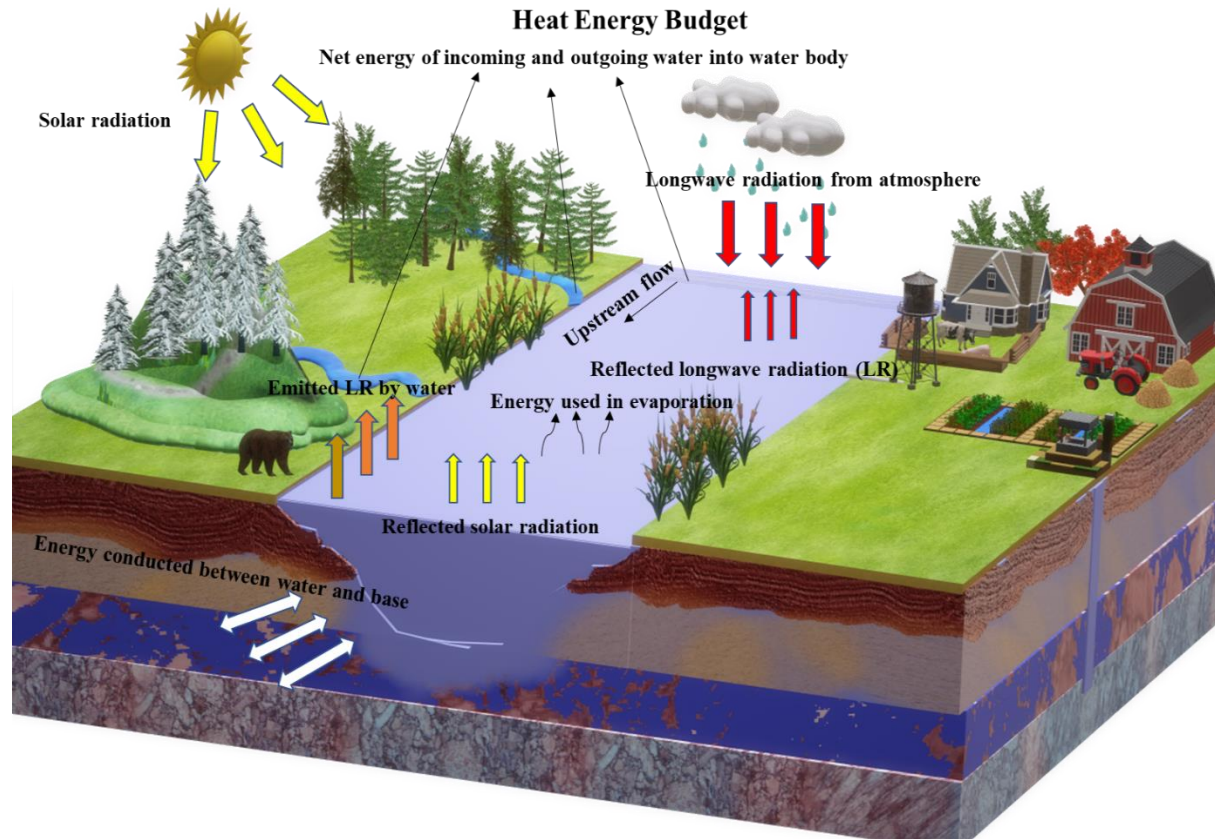


Figure 1.1. Components of heat energy budget in streams.

As the thermal signature of runoff and groundwater contributions influence temperature gradient, it is important to consider the water budget within a stream. The discharge at a point in a river is based on upstream discharge and the net effects of evaporation, transpiration, incoming-outgoing surface water runoff (mostly determined by amount of precipitation) and incoming-outgoing groundwater (Figure 1.2).

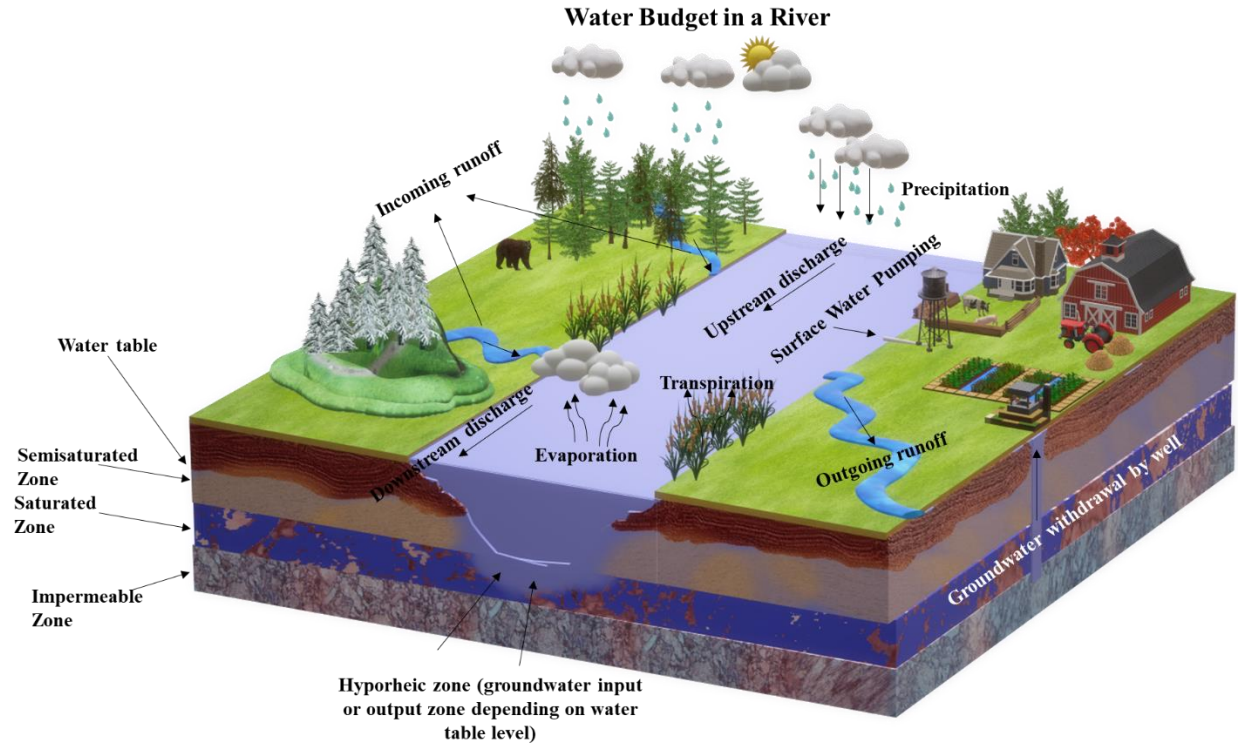


Figure 1.2. Components of water budget in streams. Heat energy budget forms the downstream discharge.

A simple equation for downstream discharge can be written as follows:

$$Q_{down} = Q_{up} + (R_{in} - R_{out}) + (G_{in} - G_{out})$$

where Q_{down} stands for the downstream discharge, Q_{up} stands for upstream discharge, R_{in} stands for incoming runoff, R_{out} stands for outgoing runoff, G_{in} and G_{out} stands for input and outflow of groundwater, respectively. Groundwater inputs occur as water moves from the water table through hyporheic zone into a stream (Vogt et. al. 2010) and they are vulnerable to groundwater withdrawal. If the water table is equal or higher than the surface water, groundwater input occurs (i.e., gaining reach) (Storey et. al. 2003). However, if the water table is lower than surface water level, the stream loses water to the aquifer, which can be viewed as reducing in-stream discharge (Ruehl et. al. 2006). Precipitation is included in incoming runoff because the majority of precipitation joins the stream from the landscape instead of directly falling on the

stream. Although outgoing runoff is conceptually possible, it does not have a substantial influence on the downstream discharge. In this equation, evaporation and transpiration are not represented as these are typically minor quantities in streams (Cheng and Wiley 2016). As indicated in the above equation, the amount of groundwater contribution is especially important in smaller streams where groundwater flow plays a large role in the water budget, and consequently in the amount of temperature gradient along a river.

Modeling Stream Temperatures

There are many models for representing stream temperature dynamics. Stream temperature models can be divided into two main groups: deterministic and statistical/stochastic models. Both have different features, strengths, and weaknesses under different circumstances. Therefore, selection of the model type is important to make reliable representations of stream temperatures.

Deterministic models use mathematical expressions and equations based on physical laws (such as laws of thermodynamics, fluid mechanics, etc.) that govern the interactions between the stream and its surroundings (Benyahya et. al. 2007). Since they use an energy budget approach, they generally require large amounts of detailed data for driving variables such as air temperature, solar radiation, wind, humidity, depth of water, velocity and so on (Morin and Couillard 1990; Sinokrot and Stefan 1993; St-Hilaire et. al. 2000; Benyahya et. al. 2007; Cheng and Wiley 2016). Deterministic models have been successfully used in a variety of situations and can be effective and appropriate to use because the heat budget equations can be modified based on different purposes such as analyzing and comparing the impacts of environmental changes (St-Hilaire et. al. 2000; Benyahya et. al. 2007).

Because they are typically complicated and costly to implement due to intensive data requirements, practitioners have sought to simplify deterministic models without losing their robustness. Cheng and Wiley (2016), for example, addressed some challenges of building and using physically based heat balance models, such as scarcity and unreliability of data for parameter values especially for large watersheds (Edinger et. al. 1974; Crittenden 1978), using a steady-state solution that assumes that the parameters do not change temporally or spatially (Bartholow 2000a; Borman and Larson 2003; Bartholow et. al. 2004), and region-specific relationships between stream temperatures and stream flows.

Statistical models are alternatives for deterministic models. One of the main differences between deterministic and statistical models is that the latter tend to be more simplistic and require less data, which can be advantageous in such cases that data collection may cost workforce, time, and money (Benyahya et. al. 2007). Benyahya et al. (2007) classified statistical models into two groups: parametric and non-parametric models. The structure of non-parametric models depends on the data and do not use conventional mathematical functions; instead, they adopt a set of relations between parameters and the output variable (e.g., Artificial Neural Networks; Benyahya et. al. 2007). Parametric models, on the other hand, adopt mathematical functions and they are very useful explaining the variation in some environmental variables (e.g., water temperature) by using the variation in other variables (e.g., air temperature; Benyahya et. al. 2007).

Benyahya et. al. (2007) classified linear regression models, which are the focus of my study, as parametric models. Linear regression models have been used to simulate stream temperatures as a function of one (e.g., air temperature) or more independent variables (e.g., air

temperature, vegetation cover, groundwater recharge; Benyahya et. al. 2007). Although simple regression models use the structure:

$$T_w(t) = a_0 + a_1 T_a(t) + \varepsilon(t),$$

where $T_w(t)$ is modelled water temperature for a given time period; $T_a(t)$ is air temperature for the same time period; a_0 and a_1 are regression coefficients and $\varepsilon(t)$ is the error term for given time, the model can be modified to a multiple regression equation by adding other independent variables such as amount of flow (Webb et. al. 2003; Benyahya et. al. 2007; Andrews 2019).

Andrews (2019) developed a suite of regression models to simulate the temperature gradient in 21 streams in Michigan. He collected hydrological and meteorological data from 15 streams in 2015 (July to early November) and 21 streams in 2016 (May through October) at 15-minute intervals. He built 11 regression models (Table 1.4) that included different independent variables, and one model that was a deterministic model based on a previous study (Magnusson et. al. 2012). He compared those models based on their fit and parsimony by using Akaike's Information Criterion (AIC) (Akaike 1973), and root mean square errors (RMSE) (Janssen and Heuberger 1995) between observed and predicted values. In addition to model accuracy and correlation with observed data, he used partial regression analysis to determine the strength of the impacts of different parameters in the best model that was selected by AIC. Finally, he evaluated the implications of baseflow reductions by using the most highly selected model.

One of the findings from his analysis was that two models received the highest weight of evidence across the majority of streams, with an average model weight ($\omega_{Avg.}$) of 0.74 for the highest ranked model, Model 10 (*Eqn. 1*). This model also had the highest correlation with

observed data in 76% of the 21 streams, with an average correlation (r) for one-year and two-year data sets of 0.66 and 0.58, respectively.

Eqn. 1.

$$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_5(Q_{down} - Q_{up}) + \beta_6(S) + \beta_7(\alpha) + \beta_8(\Delta T_{up}) \\ + \beta_9(\Delta T_{base}) + \beta_{10}(\Delta T_{over})$$

where T_a is air temperature ($^{\circ}\text{C}$), T_w is water temperature, Q_{up} is the upstream flow (m^3/sec), Q_{Down} is the downstream flow (m^3/sec), S is the day length (hours), α is the altitude angle, ΔT_{up} is the upstream heat gradient ($^{\circ}\text{C}$), ΔT_{base} is the baseflow heat gradient ($^{\circ}\text{C}$), T_{over} is the overland flow heat gradient ($^{\circ}\text{C}$) (Andrews 2019).

Although Andrews (2019) successfully applied these regression models, which provided a number of insights into drivers of stream temperature gradient, several questions remain considering the possible challenges that might be encountered in other hydrological modeling studies. I will address these potential challenges in following section.

A Challenge for Modeling: Selection of Time Period and Data Granularity

Data collection and modeling serve a variety of purposes for ecological and stream conservation. Because of the variety of uses, the time period across which data are collected and the level of data aggregation in time varies widely. For example, if the long-term effects of some environmental parameter change are the main focus, researchers tend to use yearly periods or all seasons when predicting the response variable. Studies that focus on the effects of global climate change are good examples for selection of annual periods (Sinokrot et. al. 1995; Isaak et. al. 2012; Anderson and Konrad 2019). On the other hand, a narrower time period is often used to predict the effects of environmental parameters that can change seasonally such as vegetative cover, soil temperature, concentration of nitrates and phosphates (St-Hilaire et. al. 2000; Álvarez

—Cabria et. al. 2016). Shorter time periods (e.g., monthly) may be used when the focus is on periods of ecological stress; for example, Zorn et. al. (2004) modelled the distribution of fish populations based on predicted July mean temperature under different baseflow reduction scenarios.

Although the time period for data collection is generally selected based on the purpose of study, and not based on model success, the reliability of model outputs is still important for explaining the variation in response variables with predictor variables. Therefore, it is crucial to understand and interpret the response of model success to use of different time periods (e.g., seasonal, and monthly). Moreover, understanding the response of model reliability with different time periods can give researchers a clue how model reliability varies as the ecological relevance of the time period selection varies.

Selecting the level of time aggregation, which I will refer to as data granularity in this study, is an important decision-making step in modeling. The term “data granularity” has been used in the field of business (e.g., Kim et al. 2019) and energy production and distribution (e.g., Kools and Phillipson 2016), but to my knowledge it has not been used in the hydrological literature. With current technology and data collection tools, researchers can collect environmental data at very fine time intervals such as every minute or 15 minutes and use the data with various data granularity levels by taking averages at broader time intervals (e.g., hourly time interval). In the literature, different studies have used a variety of data granularity ranging from hourly (Caissie et. al. 2001) and daily (Cheng and Wiley 2016), to weekly (Stefan and Preud’homme 1993) and monthly averages (Zorn et. al. 2004). In some of these studies, the purpose of the study was shaped by the ecological relevance of the selected data granularity. For example, Zorn et. al. (2004) used July averages to model fish distribution based on the close

relationship observed between July mean temperatures cold water fish populations. However, selecting the level of data granularity may not be entirely dependent on the purpose of study or ecological relevance of data granularity.

Data granularity may be selected for a variety of reasons such as the features of data collection tools (e.g., data collection devices may have variety of sampling interval) (Johnson et. al. 2005) or the modeler's arbitrary preference. Based on my reading of the hydrological modeling literature, the reason for selecting a level of data granularity is not stated often or explained in detail in the majority of studies. This implies that data granularity may be selected arbitrarily in most cases. However, arbitrary selection of data granularity may cause biases in model evaluation and selection processes (Kirchner 2006). This may eventually affect the decision-making processes and evaluation of hydrological and ecological implications. Therefore, selection of data granularity poses a considerable challenge for researchers and managers as the conclusions may depend on arbitrary choices. Some studies in the past examining the consequences of using different data granularity on model success have already supported my perspective on this issue. For example, Stefan and Preud'homme (1993) found that water and air temperatures were more correlated, and their relationship was less scattered, as the time averaging of data increased from two hours to weekly averages. Pilgrim et. al. (1998) also found that the slope of the regression line increased with increasing data granularity (daily, weekly, and monthly). Webb et. al. (2003) obtained similar results when they used hourly, daily, and weekly temperature mean values of different streams in Devon River System, that is, the correlation coefficient (r^2) between air temperature - stream temperature increased from hourly mean temperature values to weekly mean values in all streams.

Considering the magnitude of the problem, the number of studies in the literature is still limited and the issue needs to be addressed for recent hydrological studies. For example, although his models were useful in representing the dynamics of temperature gradient in Michigan streams, Andrews (2019) only used a single data granularity (i.e., hourly). Therefore, evaluating the response of his models to different data granularity levels would lead a better understanding of these models. Although I address the effect of data granularity on model success in this study, the focus of my study is not to define what is the most appropriate or relevant time period or data granularity for a particular problem, but rather to determine the modeling implications or consequences of changing either of these factors.

Using different data granularity can alter the model dynamics (i.e., the influence of predictor variables) and affect the results of model evaluation methods. Change in parameter estimates of models with different data granularity can be responsible for differences in perceived system dynamics and model predictive power. For example, the best fitting model with hourly data (Andrews 2019) may have different parameter estimates with different data granularity and this may potentially change conclusions based on predictive powers of models. In addition to effects on model predictive power, using different data granularity may also change model selection. Model selection results with hourly data showed that Andrews' Model 10 (2019) had the best model fit-complexity balance (i.e., model weight), however, it is unknown whether using coarser data granularity (e.g., daily) would still lead to Model 10 having the highest model weight, and the best option for temperature gradient prediction. If not, which regression model would give the best model fit-complexity balance with daily data? Although my questions are related to the specific cases from Andrews (2019), they are relevant in many other ecological and hydrological modeling studies. Therefore, finding answers is important for

future studies and environmental implications of regression models because it would reveal which environmental factors are specifically important with different data granularity selection. For these reasons, it was necessary to provide more information and a better perspective on data granularity - model reliability relationship.

Considering the potential effects of data granularity on model reliability, preliminary findings suggested that parameter estimates were not stable across different levels of data granularity (Table 1.5). Although there are many potential causes of parameter instability in regression models, a common source for this problem is multicollinearity in the independent variables. Multicollinearity is defined as the dependency of two or more predictor variables in a regression model. The primary effect of multicollinearity is an increase of the standard error of parameter estimates. The biased standard errors of parameter estimates affect the significance of parameter estimates potentially leading to biases in model selection processes that can make selecting an appropriate model hard for decision makers and may cause failure in ecological and environmental implementations (Daoud 2018). The problems related with multicollinearity in regression models have been addressed in various studies (Farrar and Glauber 1967; Haitovsky 1969; Daoud 2018).

The multicollinearity problem is pervasive in hydrological modeling because many environmental variables in topography, geology, geo-morphology, and meteorology are naturally correlated (Kroll and Song 2013). Moreover, Kroll and Song (2013) also concluded that the sample size (e.g., number of sampled streams) also might affect the amount of correlation between variables. Similarly, Mason and Perreault (1991) found that smaller sample sizes exaggerated the effect of multicollinearity on model success. This is particularly important for my research because higher data granularity naturally leads to lower sample sizes. Therefore,

there was a need for addressing multicollinearity issues in regression models that I used in my research. This would help researchers to have a better perspective of the influence of data granularity on model success and selection.

Purpose of the Study

In my research, I address the consequences of using different data granularity and time periods by using Andrews' regression models (2019). Although my study primarily will focus on these models, I believe that my findings will be a guide for many other modeling approaches since modelers have common challenges. In response to these challenges, the main objectives of my study are:

- 1) To compare the performance of regression models across different levels of data granularity by evaluating their goodness of fit and model weights,
- 2) To observe the effect of data granularity on parameter estimates and to seek possible explanations for the changing model dynamics with changing data granularity,
- 3) To analyze multicollinearity of independent variables with different data granularity to have a better insight of parameter estimate instability with changing data granularity,
- 4) To determine the relative performance of models developed for a broad time frame (June-October) compared to models developed for a narrow time frame (July) that represents a critical ecological period for cold water fishes to observe whether model performance (i.e., model prediction reliability) varies with data window choice based on the ecological relevance of data selection.

METHODS

Study Site

The choice of study streams was based on sites modelled in Zorn et. al. (2008) and Andrews (2019). For Andrews' study (2019), 24 streams were chosen that are known to be close

to groundwater extraction points in different regions of Michigan based on the different thermal classifications that are explained in Zorn et. al. (2008). I chose 16 of the 24 streams Andrews sampled based on data requirements that are explained in following sections (Table 1.6; Figure 1.3).

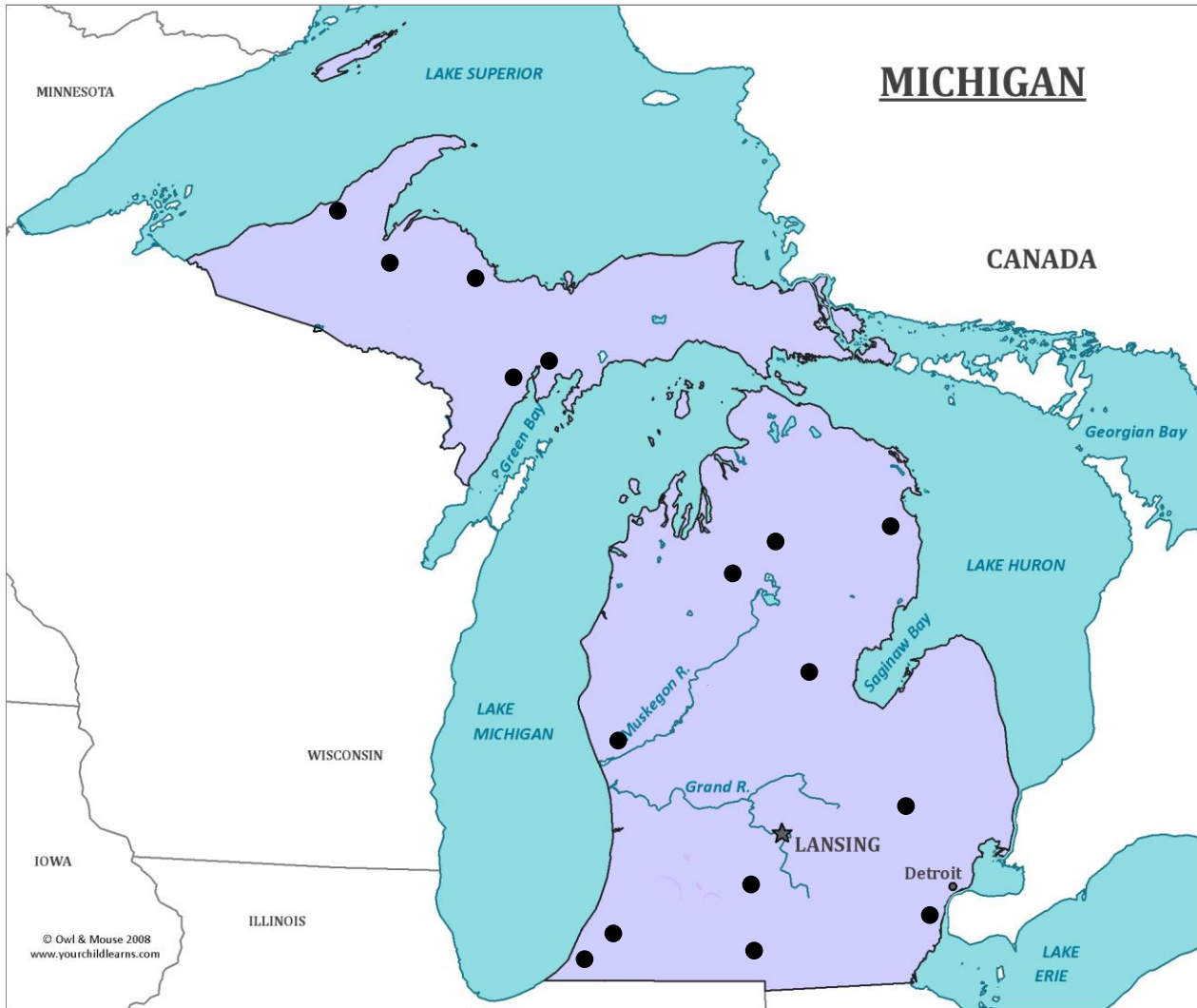


Figure 1.3. The locations of 16 streams that were selected for this study.

Data Collection

Andrews (2019) collected hydrological and meteorological data in 2015 from 15 streams with different time periods for each stream but generally ranging from July to early November

and from 21 streams in 2016, generally ranging from May to October. He placed stream gauges by using PVC pipes that were stabilized by attaching them to a fence post fixed in the streambed. To obtain water stage data, he integrated staff rulers to gauges and he used HOB0® U20 Water Level Loggers to gauges to obtain water temperature data for every 15 minutes after calibrating the loggers by placing them into ice bath (0 °C) and then letting them reach room temperature slowly. The temperatures that were obtained from all loggers were consistent but were adjusted to the same temperature. Air temperature data was collected using Monarch® Track-It data loggers with 15-minute intervals, and all water and air temperature data were averaged into hourly temperatures. To obtain stream discharge levels, he used both staff rulers and SonTek Flowtracker®. He collected barometric pressure readings from SonTek Flowtracker® to subtract them from total pressure and find water pressure.

The equation that Andrews (2019) used for the discharge calculation (*Eqn. 2*) was:

$$\text{Eqn. 2} \quad Q = aG^b,$$

where Q stands for the stream discharge (m³/sec), G stands for the reading on the gauge (inches), and a and b are parameter estimates that were obtained by using a power function while building stage-discharge curve. He derived other constants (or parameters), c , e , f , h , i , and j , from the power function to calculate other hydrological variables (*Eqn. 3*, *Eqn. 4*, *Eqn. 5*):

$$\text{Eqn. 3.} \quad w = cQ^e,$$

$$\text{Eqn. 4.} \quad d = fQ^h,$$

$$\text{Eqn. 5.} \quad V = iQ^j,$$

where w stands for the width (m) of the stream, d stands for the depth of the stream (m) and V stand for the water velocity (m/sec).

Revising the Data, Applying Data Granularity, and Testing Linearity

Although data were available from 2015 and 2016, I chose to use only the streams and rivers that were sampled in 2016 as these had data that covered the longest and most consistent time interval (i.e., June to October; Table 1.7). Data were trimmed so that the data started from 1 June 2016 to 31 October 2016 for each stream. Before modeling, I evaluated residual plots for each stream, removing outliers when necessary and removing some data frames based on unrealistic discharge changes. I also plotted the relationship between dependent and each independent variable as well as between observed and predicted temperature gradient to evaluate whether a linear model appeared to be appropriate constant (Poole and O'Farrell 1971). Example results from the Tobacco River, which had the best goodness of fit, between June-October 2016 are presented as an example (Figure 1.20 to Figure 1.28).

I also changed the usage of some parameters: upstream heat flow (ΔT_{up}), baseflow heat flow (ΔT_{base}), overflow heat flow (ΔT_{over}), and total heat flow (ΔT_{flow}) (Table 1.4) to better reflect the dynamics of stream discharge. In his study, Andrews (2019) equalized all these parameters to zero when the downstream discharge was lower than the upstream discharge because he suggested that if downstream discharge were lower than upstream discharge, the contribution of upstream flow, baseflow and overflow on downstream discharge and temperature gradient would be ignorable. Another reason was that these parameters tend to have negative values in that case. On the contrary, I directly used the values of these parameters although their values were negative because I suggested that the discharge loss might be a result of natural processes (i.e., downwelling) or anthropogenic process (i.e., groundwater or surface water withdrawal), and those parameters might have had an effect on discharge and temperature

gradient. Moreover, temporal changes of those parameters might have had explanatory power on temperature gradient even if they had negative values.

After these refinements and revisions, I took hourly, 2-hour, 6-hour, 12-hour, daily (24-hour) and weekly (168-hour) averages of the thermal gradient data and of the environmental data from each stream to create data of increasing granularity.

Comparisons of Goodness of Fit for Each Data Granularity Scenario

To achieve the first goal in my study, I applied 11 regression models based on Andrews (2019) (Table 1.4) to June-October 2016 data with different data granularity scenarios. I fit each model to each stream and determined the best-fitting models by using two measures of goodness of fit for each data granularity scenario. The first measure was adjusted correlation coefficient (R^2), which is commonly used as a measure of a regression's explanatory power in many stream temperature modeling studies (Ahmadi - Nedushan et. al. 2007; Mayer 2012; Hill et. al. 2013). Adjusted correlation coefficient (R^2) was used to explain the variation of a variable (e.g., predicted temperature gradient) across other variable (e.g., observed temperature gradient). Based on the nature of the equation, value of R^2 is always between 0 and 1, and as the value approaches to 1, model predictive power becomes greater. To obtain R^2 , I used *Eqn. 6*:

Eqn. 6.
$$R^2 = 1 - \left(\frac{n-1}{n-p} \right) * \frac{SSE}{SST}$$

where n is the number of observations, p is the number of parameters, SSE is the sum of squared residuals and SST , and SST is the total sum of squares.

As the second goodness of fit measure, I applied Akaike's Information Criterion (AIC) (Akaike 1973) based on the principle of parsimony. AIC is defined as *Eqn. 7*:

Eqn. 7.
$$AIC = 2k - 2 \ln \ln (L(data)); \text{ and } L(data) = -\binom{n}{2} \cdot \log_e(SSE),$$

where L stands for the likelihood, k stands for the number of unknown parameters, and n stands for the sample size (Seber and Wild 1989). I prioritized the models by determining their weight of evidence using the formula:

Eqn. 8.
$$\omega_i = \frac{\exp\left(-\frac{\Delta_i}{2}\right)}{\sum_m^M \exp\left(-\frac{\Delta_i}{2}\right)}$$

where M is the total number of models, m is the model number, and Δ_i is the difference of AIC values of that model from the AIC value of the best-fitting model. By using model weights, I was able to order the models from the best-fitting to poorest-fitting model while balancing model complexity (Andrews 2019).

Multicollinearity Diagnosis and Response of Parameter Estimates to Data Granularity

I adopted two approaches to evaluate the implications of multicollinearity among predictor variables. First, I obtained Pearson's correlation coefficients (r) (Eqn. 9) between the parameter estimates of predictor variables that were used in Model 10, the best performing model, across streams to understand how parameter estimates covary. I obtained r values by using:

Eqn. 9.
$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}}$$

where x and y are variables, and \bar{x} and \bar{y} represent means of variables. I obtained and used correlation diagrams and correlograms to visualize the change of correlation (r) between parameter estimates for each data granularity scenario. To obtain correlation diagrams, “writexl” package was used in RStudio Version 0.98.1103 (Appendix D: RStudio Codes). To obtain correlograms, “corrplot” package was used (Appendix D: RStudio Codes). The purpose of this approach was to observe the response of mean β values to increasing data granularity, leading to a better understanding on the insight of the best predicting model.

As multicollinearity in the input data has long been known to influence the stability of parameter estimates, I calculated the degree of multicollinearity between variables in the raw data by using an example stream to evaluate whether the level of collinearity in the data could be driving the instability of parameter estimates. I used Tobacco River June-October 2016 data as an example since model predictive power was highest based on my preliminary results. Correlation matrices and correlograms were obtained for each time aggregation to analyze the effect of data granularity on the level of correlation in the raw data.

Evaluating Model Performances by Using July-Restricted and June-October Data

As another purpose in my research, I fit the linear regression models to July 2016 restricted datasets for each data granularity scenario and I found adjusted correlation coefficient (R^2) to evaluate model predictive power across data granularity scenarios. Then, I compared model predictive power of July-restricted model and June-October model. The variations between model predictive powers indicated the importance of selecting seasonal or monthly dataset on accuracy of the best fitting models.

RESULTS

Data Granularity Influenced Model Predictive Power and Model Weight

The relationship between data granularity and the predictive power of linear regression models as measured by the adjusted correlation coefficient (R^2) showed three major patterns. Firstly, overall model prediction powers of all models increased with data granularity (Table 1.1). The second major pattern is that Model 10 had the highest mean adjusted correlation for each of the levels of data granularity (Figure 1.4).

Table 1.1. Mean adjusted correlation (R^2) values of each model by data granularity across all streams with June-October data.

Model	Data Granularity (hour)						Average
	1	2	6	12	24	168	
1	0.139	0.142	0.149	0.198	0.315	0.498	0.240
2	0.094	0.098	0.108	0.133	0.202	0.415	0.175
3	0.188	0.209	0.207	0.226	0.311	0.499	0.273
4	0.205	0.209	0.225	0.253	0.340	0.571	0.301
5	0.278	0.284	0.309	0.368	0.502	0.732	0.412
6	0.253	0.257	0.279	0.360	0.485	0.737	0.395
7	0.329	0.336	0.367	0.502	0.515	0.754	0.467
8	0.258	0.375	0.391	0.453	0.591	0.812	0.480
9	0.332	0.336	0.358	0.45	0.587	0.823	0.481
10	0.418	0.423	0.447	0.563	0.598	0.842	0.548
11	0.312	0.320	0.342	0.419	0.536	0.793	0.454
Average	0.255	0.272	0.289	0.357	0.453	0.680	

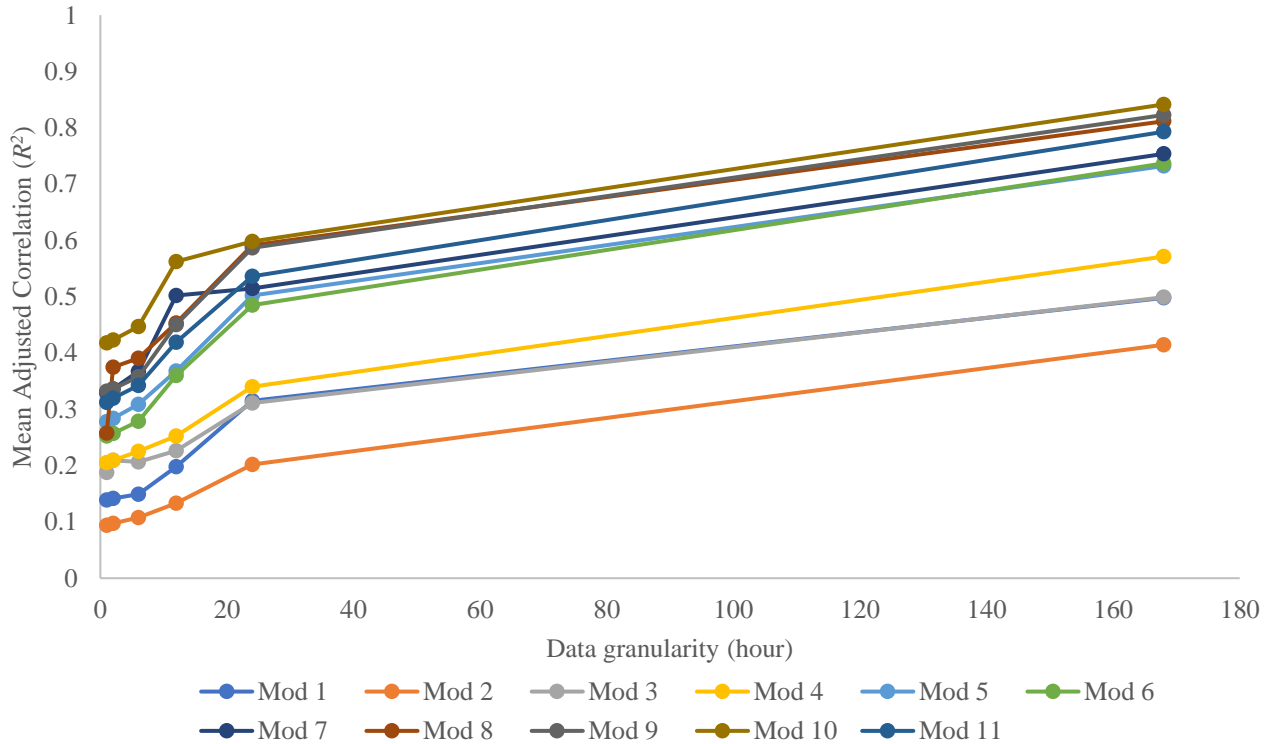


Figure 1.4. Mean adjusted correlation (R^2) values of all streams (June-October 2016) based on different data granularity scenarios and different models.

When averaged across all levels of data granularity, Model 10 had an average R^2 of 0.548 (Table 1.1; Figure 1.29). Models 8 and 9 followed closely behind Model 10 in their predictive capacity, with a mean R^2 value of 0.480 and 0.481, respectively (Table 1.1). Models 7 and 11 were generally close in their predictive capacity, with a mean R^2 value of 0.467 and 0.454, respectively (Table 1.1). Models 1 through 4 showed distinctly lower predictive power than the other models (Figure 1.4). These models lacked parameters representing solar insolation, such as altitude angle and day length, indicating that these parameters were of large importance in explaining patterns of temperature gradient across all levels of data granularity. The last major pattern was that the mean correlation generally increased for all models as data granularity was increased from hourly to weekly time scales (Table 1.1; Figure 1.30). When averaged across all models, the mean R^2 value increased from 0.255 for hourly data granularity to 0.680 for weekly data granularity (Table 1.1). While these patterns were quite consistent for the mean response of adjusted correlation coefficients to data granularity, preliminary analysis suggested that the trends of model predictive power across data granularity varied among streams.

Overall, Model 10 received the highest weight of evidence in the majority of data granularity scenarios (Table 1.2; Figure 1.5). However, the same results showed that the level of data granularity changed the outcome of model selection substantially, where increasing data granularity (i.e., reducing the number of data points) led to reduced weights for the most complex models, and broadened the support for less complex models (Table 1.2; Figure 1.5).

Table 1.2. Percentage (%) of streams where each model had the highest model weight (w) across levels of data granularity. June-October data were used in models.

Data granularity (hour)	Models											Total
	1	2	3	4	5	6	7	8	9	10	11	
1	0	0	0	0	0	0	0	6.25	6.25	62.50	25.00	100
2	0	0	0	0	0	0	0	6.25	0	62.50	31.25	100
6	0	0	0	0	0	0	0	12.50	0	50.00	37.50	100
12	0	0	0	0	0	0	18.75	6.25	6.25	43.75	25.00	100
24	0	0	0	0	6.25	6.25	0	25.00	12.50	18.75	31.25	100
168	6.25	0	0	0	6.25	0	6.25	25.00	0	31.25	25.00	100

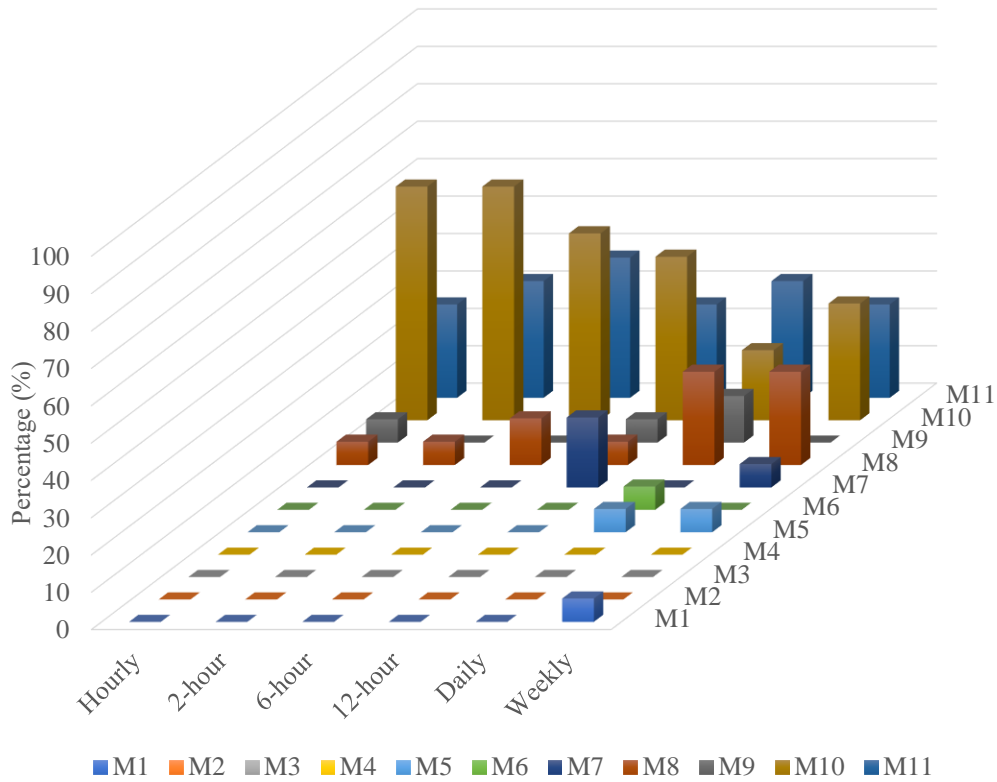


Figure 1.5. The percentage of the models having the highest model weight at least one stream for each data granularity with June-October 2016 data.

The effect of data granularity on model selection was clearly noticeable as model weights changed across data granularity scenarios (Table 1.2). For hourly data granularity, Model 10 had the highest model weight for more than 60% of streams. Model 10 continued to receive the

highest weight for the most streams across all levels of data granularity except for a daily data granularity, in which Model 11 had the highest percentage. The percentage of other models that were most highly chosen increased with data granularity. For data granularity scenarios of 1-, 2-, and 6-hours, Models 8, 9, 10, and 11 were the only models to be selected as the top models. As data granularity increased to higher levels (i.e., 12-hour, daily and weekly), less complex models, such as Model 1, 5, 6 and 7, emerged as the most highly selected model in some streams.

Data Granularity Leads to Instability of Parameter Estimates in Best Fitting Model

Parameter estimates ($\bar{\beta}$) for Model 10 averaged across all streams showed instability with increasing data granularity. In most cases, $\bar{\beta}$ for predictor variables showed consistent trends with higher data granularity (Figure 1.6 and 1.7). The $\bar{\beta}$ value associated with upstream discharge (Q_{Up}) showed a strong increasing trend with greater data granularity, and even showed a change in the sign of the parameter estimate (Figure 1.6). In contrast, $\bar{\beta}$ value for day length (S) started with a positive sign and ended up with a negative sign with weekly data (Figure 1.7). Furthermore, the general picture indicated that the trends of mean parameter estimate values across data granularity influenced each other. For example, upstream heat flow (ΔT_{Up}) and overflow heat flow (ΔT_{Over}) increased from hourly to 12-hour data granularity and a decrease for greater granularity scenarios, whereas baseflow heat flow (ΔT_{Base}) showed a decrease from 2-hour to daily data granularity but increased in weekly granularity (Figure 1.7).

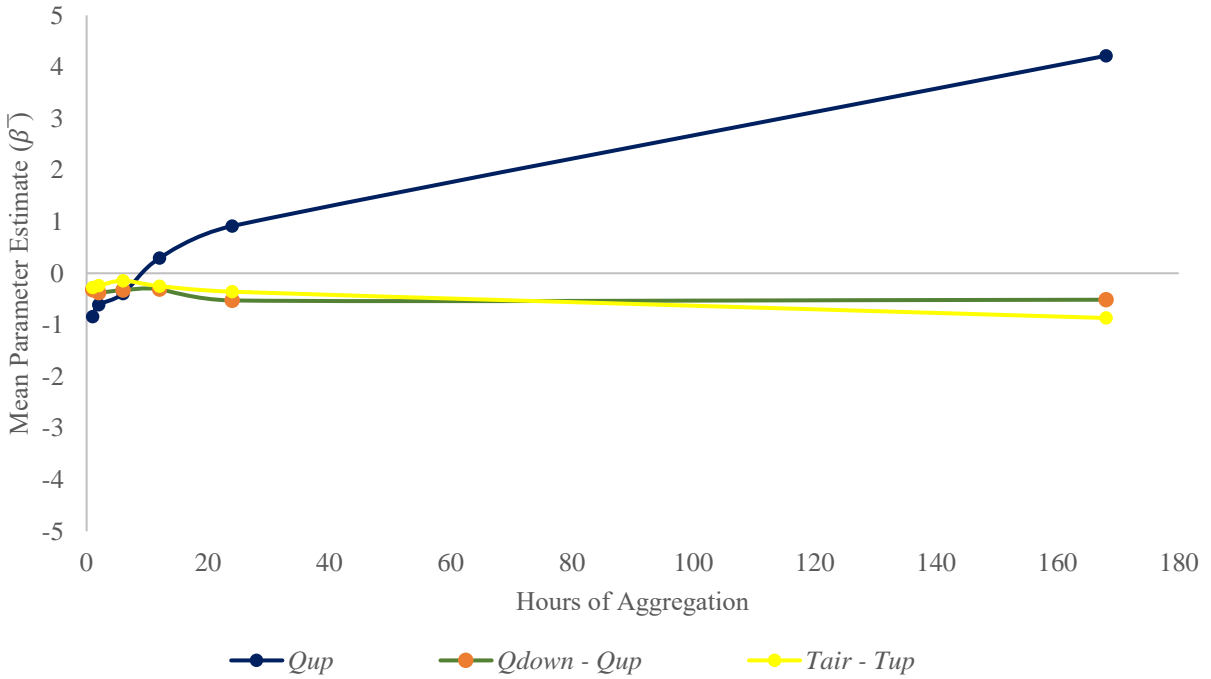


Figure 1.6. Response of $\bar{\beta}$ to data granularity. Model 10 was used with June-October 2016 data. β values of all streams were averaged. Q_{Up} : upstream discharge; $Q_{Down} - Q_{Up}$: difference between downstream and upstream discharge; $T_{Air} - T_{Up}$: difference between air temperature and upstream temperature.

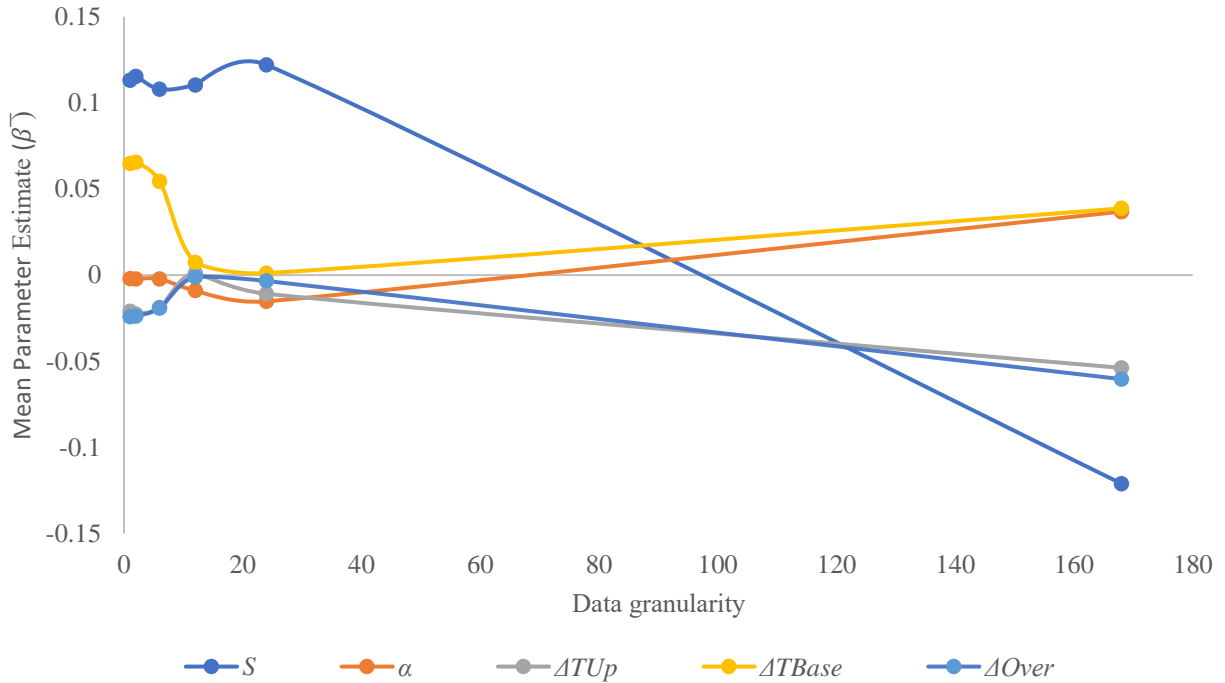


Figure 1.7. Response of $\bar{\beta}$ to data granularity. Model 10 was used with June-October 2016 data. β values of all streams were averaged. S : day length; α : altitude angle; ΔT_{Up} : upstream heat flow; ΔT_{Base} : baseflow heat flow; ΔT_{Over} : Overflow heat flow.

The potential interaction between mean parameter estimates ($\bar{\beta}$) led me to evaluate multicollinearity for parameter estimates (β) of Model 10 across streams since the interaction between $\bar{\beta}$ values might have been explained by high correlation between β values. Preliminary analysis on ΔT_{Up} , ΔT_{Base} and ΔT_{Over} revealed the interaction between these variables (Figure 1.31 and 1.32). These figures showed that if β values of ΔT_{Up} and ΔT_{Over} are high on a stream, β value of ΔT_{Base} tended to be low for that stream, or vice versa. To evaluate the interactions between all parameter estimates, multicollinearity between β values was tested by observing coefficient of correlation (r). Results showed that increasing data granularity resulted in a change of overall correlation between the β values of explanatory variables (Table 1.8 to 1.13). Correlograms clearly showed this change, as the number of darker and bigger circles varied across data granularity (Figure 1.8 to 1.13). In addition, some of the β values were highly

correlated in all data granularity scenarios. Baseflow heat flow (ΔT_{Base}) and overflow heat flow (ΔT_{Over}) had the highest negative correlation across scenarios. Moreover, Overflow heat flow (ΔT_{Over}) and upstream heat flow (ΔT_{Up}) had the highest positive correlation in all scenarios.

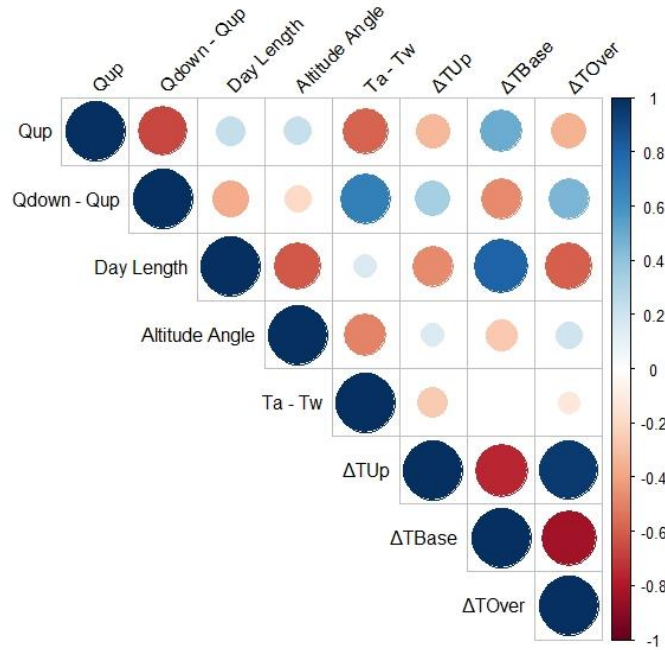


Figure 1.8. Correlation (r) between β values across all streams with hourly data granularity. Model 10 was used with June-October 2016 data. The color and the size of circles indicate the sign and the numerical value of correlation.

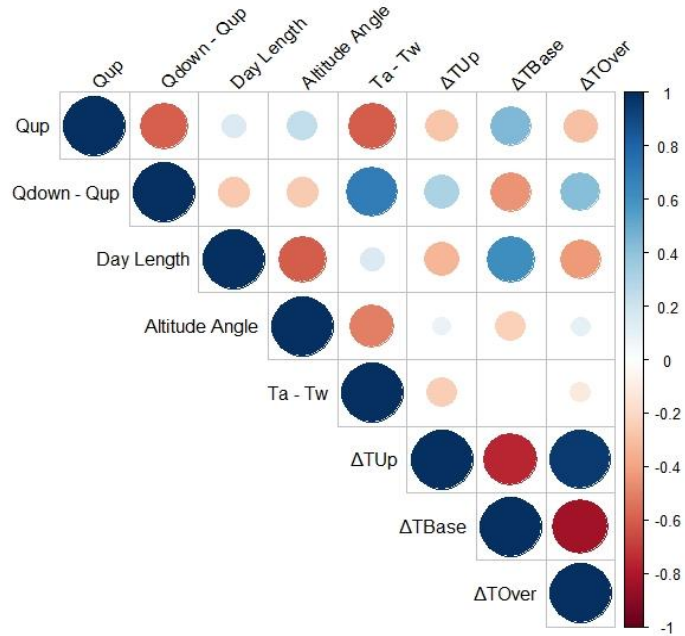


Figure 1.9. Correlation (r) between β values across all streams with 2-hour data granularity. Model 10 was used with June-October 2016 data.

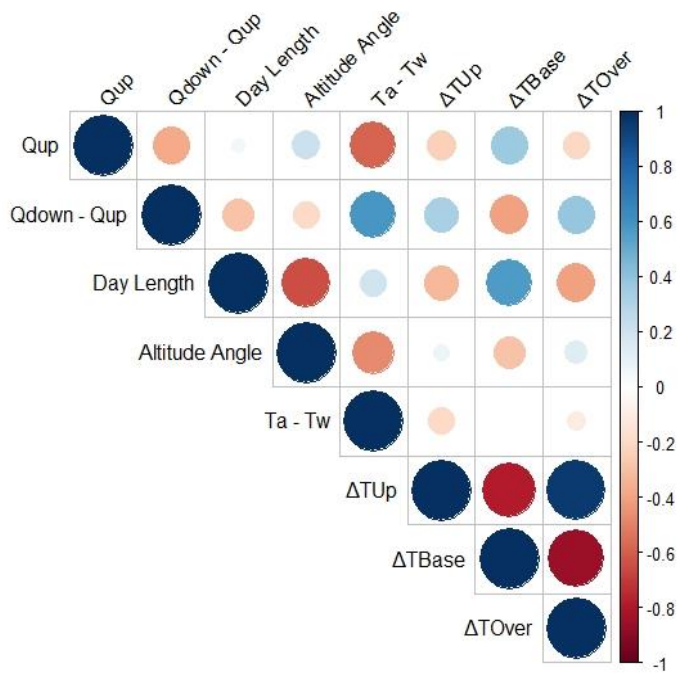


Figure 1.10. Correlation (r) between β values across all streams with 6-hour data granularity. Model 10 was used with June-October 2016 data.

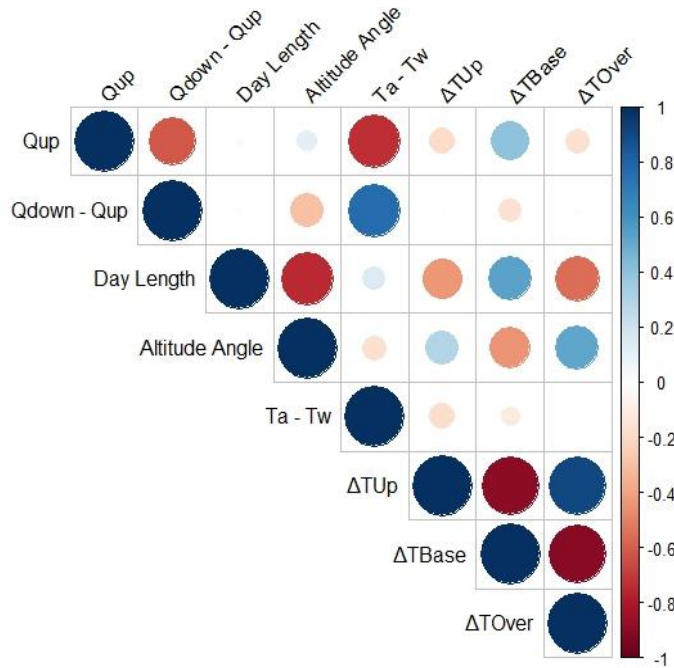


Figure 1.11. Correlation (r) between β values across all streams with 12-hour data granularity. Model 10 was used with June-October 2016 data.

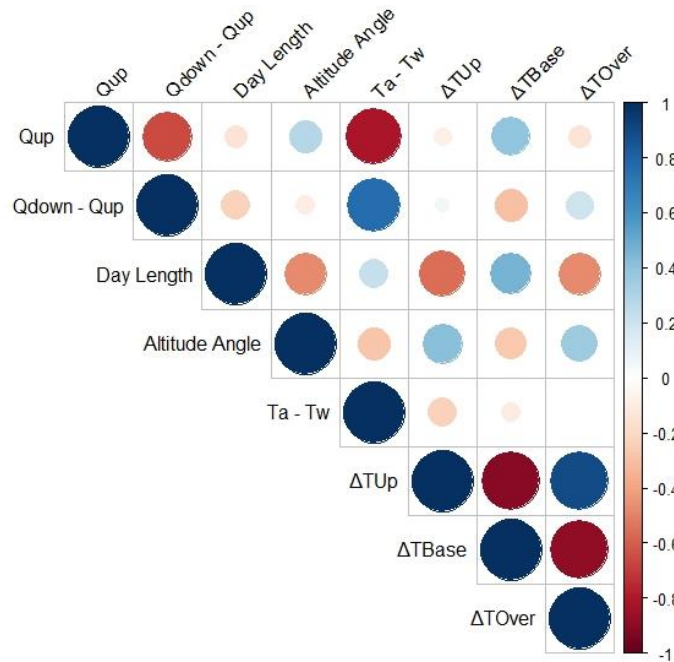


Figure 1.12. Correlation (r) between β values across all streams with daily data granularity. Model 10 was used with June-October 2016 data.

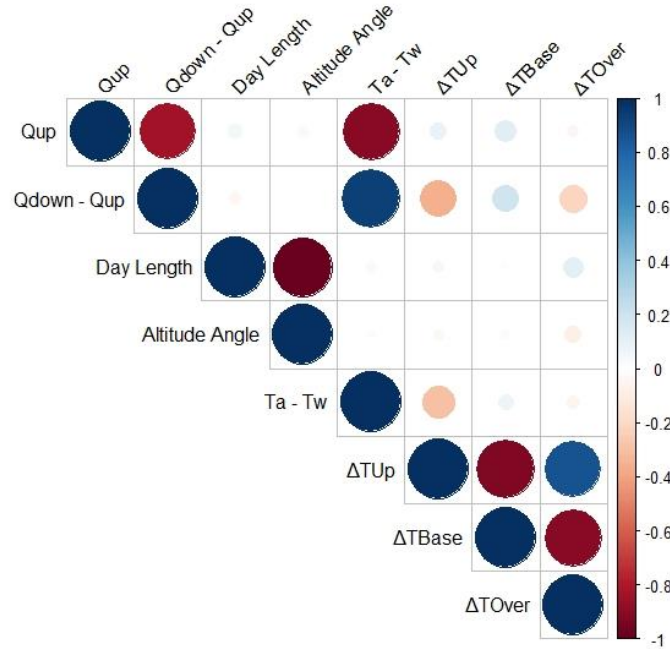


Figure 1.13. Correlation (r) between β values across all streams with weekly data granularity. Model 10 was used with June-October 2016 data.

Data Granularity Increased Multicollinearity in Raw Data

A potential cause of parameter instability and multicollinearity between parameter estimates (β) might have been the intrinsic multicollinearity between environmental variables in the raw data. Multicollinearity between environmental variables was tested by using Tobacco River data (June-October 2016). Correlation (r) between environmental variables showed two major patterns. First, an increase of r between environmental variables was observed (Figure 1.14 to 1.16). Moreover, although the magnitude of correlation varied with increasing data granularity, the sign of r values did not change with data granularity. Some of the variables (e.g., Q_{up} and $Q_{Down} - Q_{up}$ versus ΔT_{Up}) were consistently negatively correlated, whereas some parameters (e.g., altitude angle versus day length) were positively correlated. Second, at hourly data granularity, several variables showed high correlation. Both Q_{up} and $Q_{Down} - Q_{up}$ values and ΔT_{Up} had the highest correlation in all scenarios. In addition, Q_{up} and $Q_{Down} - Q_{up}$ were other variables that had high correlation in all scenarios (Figure 1.14 to 1.16).



Figure 1.14. The amount of correlation (r) between environmental variables in Tobacco River with hourly (a) and 2-hour (b) June-October 016 data shown in correlogram.



Figure 1.15. The amount of correlation (r) between environmental variables in Tobacco River with 6-hour (a) and 12-hour (b) data granularity. Model 10 with June-October 2016 data was used.

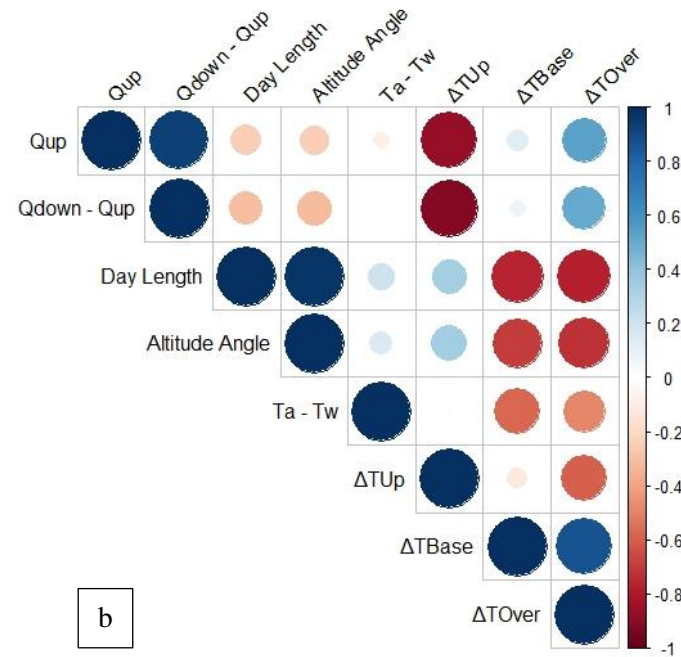
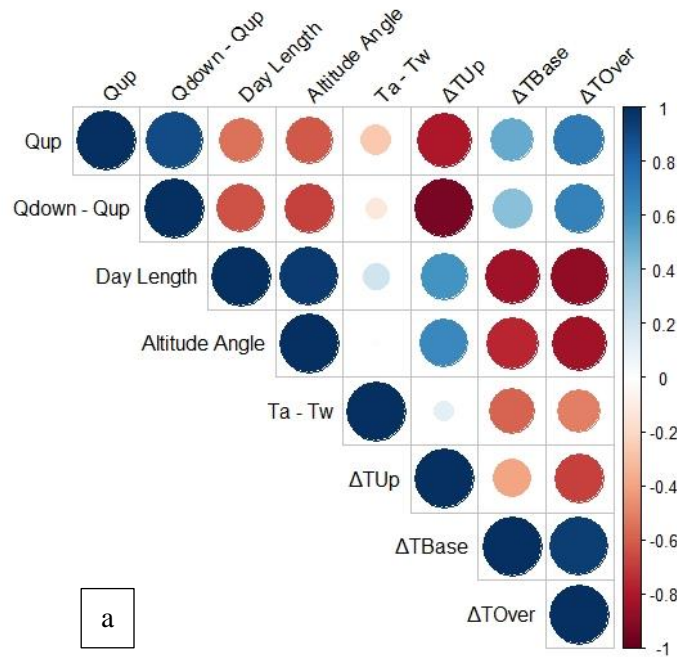


Figure 1.16. The amount of correlation (r) between environmental variables in Tobacco River with daily (a) and weekly (b) data granularity. Model 10 with June-October 2016 data was used.

Using July-Restricted Data Did Not Improve Model Prediction Power

Although overall R^2 increased with greater data granularity for the July-restricted models, it was less apparent than for June-October models (Table 1.3; Figure 1.17). This observation was supported by the fact that, in all data granularity scenarios, the p-values were greater than $p=0.05$ (1-hour: $p=0.1681$; 2-hour: $p=0.2869$; 6-hour: $p=0.3859$; 12-hour: $p=0.7024$; 24-hour: $p=0.2581$), that is, I failed to conclude that the mean R^2 values of July-restricted models and June-October models within the same aggregation were significantly different (Table 1.14). In other words, using July restricted dataset did not cause a significant difference between overall predictive power of models.

Table 1.3. Mean adjusted correlation (R^2) values of each model by data granularity across all streams with July 2016 data.

Model	Data Granularity (hour)						Average
	1	2	6	12	24	168	
1	0.144	0.143	0.130	0.163	0.116	0.144	0.139
2	0.136	0.139	0.145	0.120	0.181	0.136	0.144
3	0.252	0.257	0.274	0.290	0.356	0.252	0.286
4	0.261	0.265	0.282	0.298	0.377	0.261	0.297
5	0.275	0.278	0.290	0.280	0.409	0.275	0.306
6	0.341	0.346	0.366	0.400	0.407	0.341	0.372
7	0.355	0.358	0.375	0.421	0.444	0.355	0.391
8	0.394	0.398	0.401	0.399	0.497	0.394	0.418
9	0.448	0.452	0.461	0.443	0.494	0.448	0.460
10	0.472	0.476	0.486	0.463	0.519	0.472	0.483
11	0.438	0.432	0.441	0.417	0.455	0.438	0.437
Average	0.320	0.322	0.332	0.336	0.387	0.320	

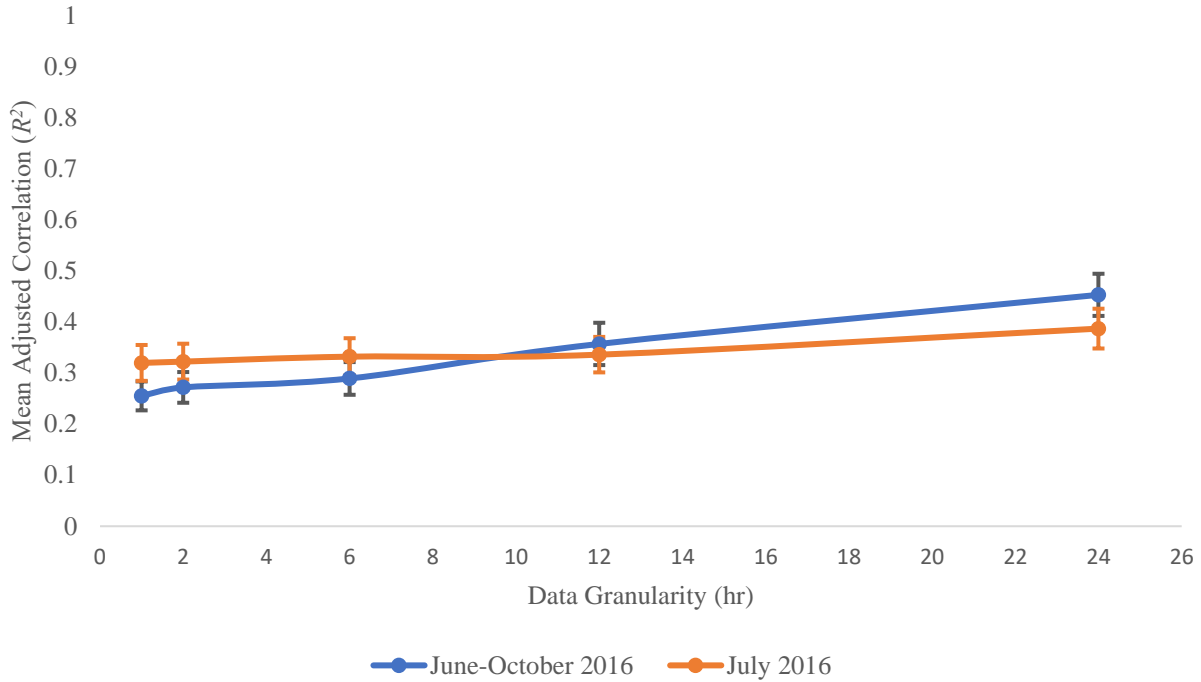


Figure 1.17. Mean adjusted correlation (R^2) values of each data granularity scenarios based on all regression models. Whiskers represent standard errors of sample.

In addition, mean R^2 values showed little relation to data granularity for July restricted data (Figure 1.17). As observed for June-October data, Model 10 had the highest mean correlation coefficient (0.483) in all data granularity scenarios for models applied to July 2016 data (Figure 1.18). Moreover, Model 3 and higher models were grouped together based on their predictive power (Figure 1.19) when July-restricted data were used, but this grouping pattern was different since Model 5 and higher models were grouped when June-October data were used (Figure 1.4). This conclusion suggests that the influence of parameters (i.e., day length and altitude angle) used in models differs between June - October and July- restricted data.

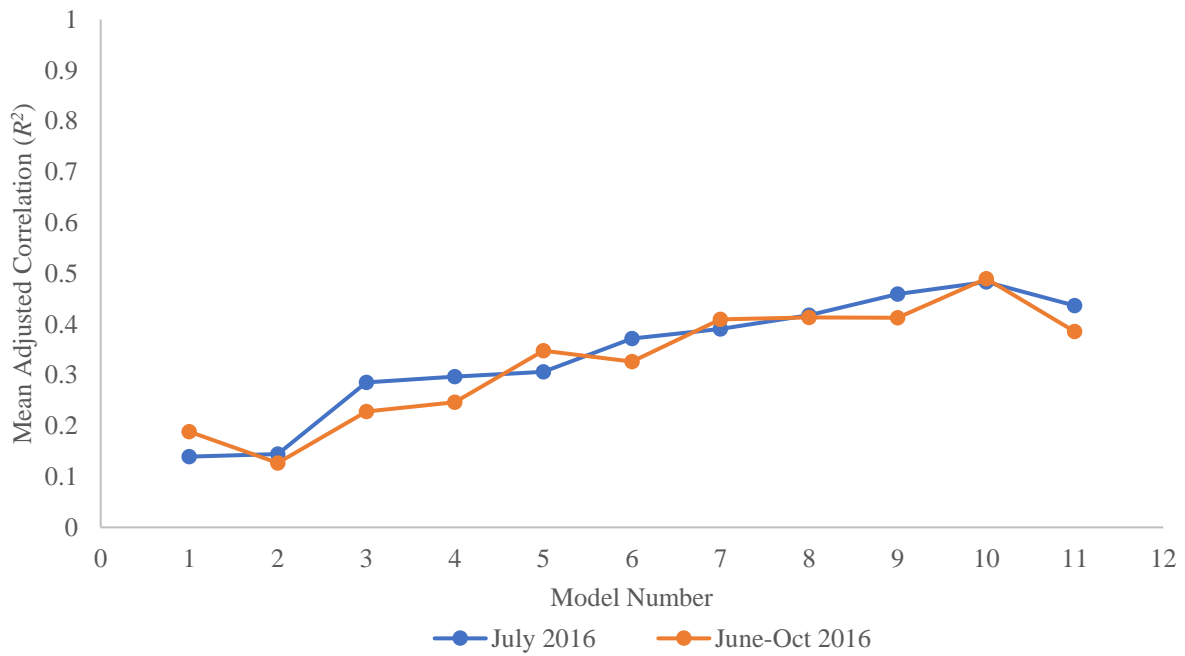


Figure 1.18. Mean adjusted correlation (R^2) values of models based on averaging all data granularity scenarios. Lines represent mean adjusted correlation values obtained by using July restricted data (blue) June-October data (orange).

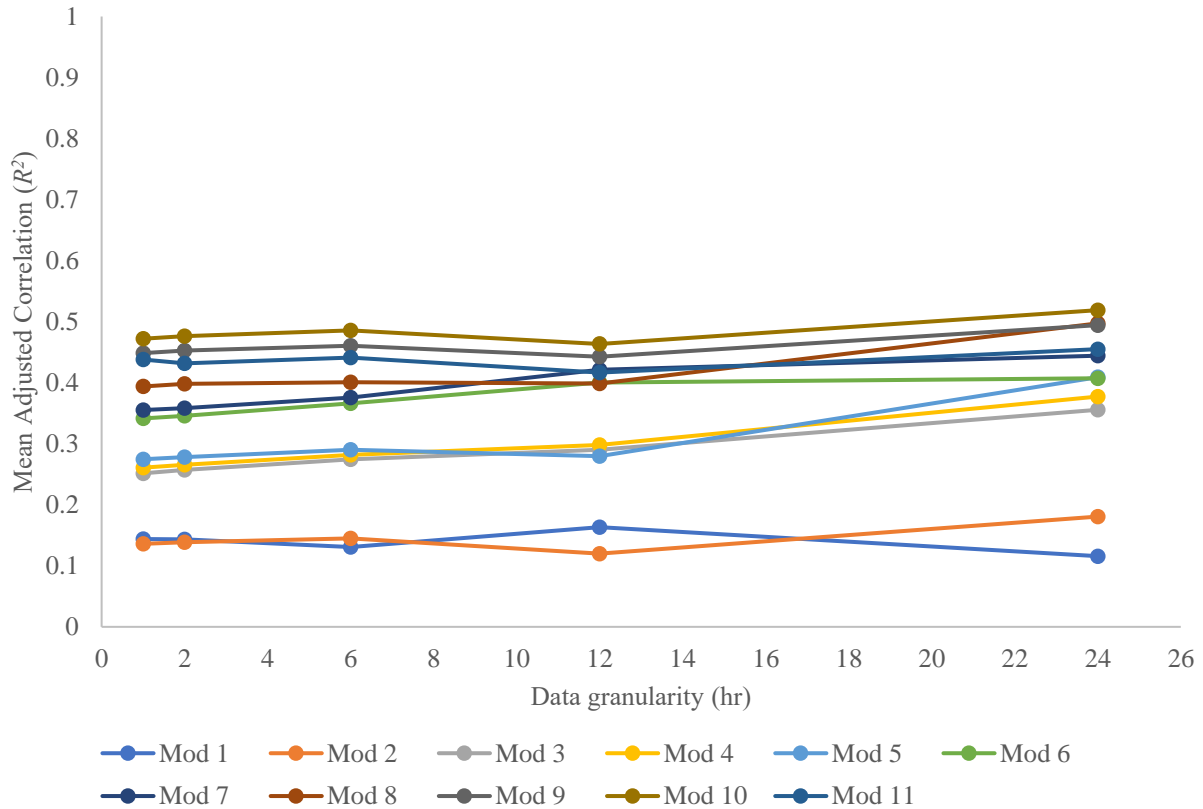


Figure 1.19. Mean adjusted correlation (R^2) values of models across all streams with July 2016 data across data granularity scenarios.

DISCUSSION

My findings address the gaps in previous modeling studies I identified by answering three main questions: “How does model performance and selection vary with data granularity?”, “What are the possible reasons for model performance and selection changes with data granularity?” and “How do models perform with July-restricted data?”. My results provide a clear picture of how model performance varied with different data granularity scenarios, as well as the possible reasons for variable model performances by observing the changes of model dynamics with data granularity. Revealing the changes in model dynamics by referring multicollinearity has given a better insight into the regression models that can be used when implementing these models in future studies.

How Does Model Performance and Choice Vary with Data Granularity?

Selection of different data granularity, or time aggregation, scenarios had substantial effects on model performances and model outcomes. Higher data granularity increased overall prediction power of regression models with July-October data (Table 1.1; Figure 1.30). Data granularity did not only change overall model prediction power but also changed decisions in model selections by influencing the model weights. For example, depending on the ecological perspective and purpose, Model 10 can be selected and used to make more accurate temperature gradient predictions with hourly data, whereas Model 11 can be selected for the same purpose with daily data, since model weight of Model 11 was the highest for most of the streams with daily data (Table 1.2; Figure 1.5).

If data granularity changes the model performances and selection, then what is the “best” data granularity to be used in environmental studies? My study cannot directly answer this question since the term “best” is highly depended on the purpose of the study and the ecological relevance of selected data granularity. To illustrate, predicting monthly stream temperature or temperature gradient averages to evaluate and simulate the habitable streams for certain fish species (e.g., Zorn et. al. 2004; Zorn et.al. 2008) would be more plausible than using hourly stream temperature predictions since many fish species can tolerate the hourly variations in stream temperature. Therefore, using monthly average would give a better perspective to simulate the fish distribution. On the other hand, if the daily change of stream discharge based on daily or weekly groundwater withdrawal is under focus, using greater data granularity scenarios (e.g., daily, weekly) would be the best decision (Fleury et. al. 2009). Selection of the “best” data granularity also depends on the expectations from the model performances. In the literature, there is no agreement on what the range of adjusted correlation coefficient (R^2) should be used as

an indicator of good model performance (Prairie 1996). For example, the value of 0.8 for R^2 may be considered as a low model performance in a study, while the value of 0.7 may be considered as high model performance in another study depending on that field of research. Moreover, the methods that researchers employ to evaluate model success can affect the “best” data granularity selection. As a clear example, my study showed that the model prediction power (i.e., R^2) of Model 10 with weekly data granularity was the highest, whereas with the same data granularity, Model 11 had the highest model weight which is used in model selection processes based on the balance between model complexity - model prediction power (Akaike 1973).

Although my study could not come up with the “best” data granularity selection, it proposed answers for many other questions. For example, if the question was changed to “What is the best data granularity selection based on the purposes and the methods used in this study?”, the answer would be: “the higher data granularity, the better model performances”. Since applying greater data granularity reduces the number data points, it consequently reduces the number of sharp variations within environmental variables. As an example, my preliminary results based on July 2016 Tobacco River data showed that the air temperature tended to fluctuate during the day when hourly data is used. However, applying daily or higher data granularity (e.g., daily and weekly) reduces these fluctuations (Figure 1.33). As a result, overall model predictive power tended to increase with hourly data to weekly data granularity.

My study also revealed the best-fitting model based on model predictive power and model weights. Model 10 had the highest overall model predictive power on temperature gradient among all regression models based on adjusted correlation coefficient (R^2) considering all data granularity scenarios (Figure 1.29). Model 10 is the most complex model having 8 independent variables. As expected, model complexity increased the model predictive power.

Compared to models Model 8 and Model 9, Model 10 contains both altitude angle (α) and day length (S) as driving variables, leading to higher predictive power. Since altitude angle, which is directly related with the amount of solar radiation reaching the stream, is very important especially during the summer season, variations in altitude angle helped Model 10 to have better prediction power. Moreover, day length is also an important predictive parameter since it determines the amount of time that the stream is exposed to solar radiation. Therefore, using both parameters in the same model was critical. Additional evidence for the importance of these parameters is the model grouping based on model predictive power (Figure 1.4). The models that included at least one of these parameters (e.g., Model 5 through Model 10) were grouped based on their distinctly higher predictive power, whereas the models that do not include these parameters were grouped based on their lower predictive power.

Model 10 also had the highest model weight in the majority on data granularity scenarios (Figure 1.5). In other words, the trade-off between goodness of fit components and the model complexity was the lowest in Model 10. Therefore, I concluded that Model 10 should be selected as a linear regression model to make more reliable temperature gradient predictions, until another model is developed that can have better reliability-complexity balance. However, it was clearly observed that the model weight of Model 10 decreased with data granularity (Figure 1.5). For example, Model 11 and Model 8 with daily data granularity had higher percentage of being the best model selection across streams. This was not the only conclusion. The results also showed that the less complex models began showing up as the best models for some streams as data granularity increased (Table 1.2). This was an important finding, because it implied that higher model complexity may be a disadvantage, since less complex models can make temperature gradient predictions as good as the complex models as data granularity increases. This was a

consequence of decrease in the number of predicted temperature gradient (i.e., data points) as greater data granularity was used, making the model complexity less important, yet making the explanatory power of parameters in the model more important. Therefore, researchers should consider the data granularity when building the models, since the complexity may reduce the model efficiency (i.e., predictive power - model complexity balance). Moreover, designing less complex but more reliable models may save resources such as time, work force and finances during the data collection.

The biggest picture that my findings posed was that the arbitrary selection of data granularity may have serious consequences, such as biases in model evaluations and model selection processes. More importantly, my literature readings showed that arbitrary selection of data granularity has not been a big concern for many researchers and managers, and the reasons for data granularity selection were not detailly explained in many studies that deal with riverine systems modeling. However, if the reasoning for data granularity selection is not purely based on ecological relevance (i.e., data granularity is selected arbitrarily), researchers can easily come up with conclusions on success of their models by using arbitrary data granularity selection, which may not be realistic when other data granularity scenarios are considered. The solution, in these cases, might be to define the ecological relevance of a particular level of data granularity in the first place, then evaluate the models based on their performances. By doing so, researchers may have a better understanding of the weaknesses of their models with ecologically relevant data granularity, and design models that do not only have higher performance, but also have ecological relevance with their purposes.

What are the Possible Reasons for Model Performance and Selection Changes with Data Granularity?

As above, I propose that one of the reasons for overall model performance increase with higher data granularity was the lower number of data points in the data. However, my preliminary results showed that the model prediction power may decrease with higher data granularity for some streams. For example, adjusted correlation value of Tobacco River was lower for 12-hour data granularity when compared to the same value for 12-hour and daily data granularities (Figure 1.34). Likewise, the value of adjusted correlation decreased from 12-hour to daily data granularity for Butterfield Creek, Carp River and Prairie River (Figure 1.34). Therefore, lower number of data points could not be the only reason for prediction power change with data granularity. Changes in model dynamics, which are caused by variation between parameter estimate (β) values (i.e., parameter instability) across granularity scenarios, is likely a more plausible reason for model prediction power changes as well as the changes in model weights, since parameter estimate (β) values indicate the weight (or influence) of each predictor variable on temperature gradient predictions. The simplest way to show the change of model dynamics was to observe the trends of mean parameter estimate ($\bar{\beta}$) values across data granularity scenarios. The instability of $\bar{\beta}$ values, leading in some cases to a change in the sign of the $\bar{\beta}$ value (Figure 1.7), suggests that changes in data granularity changed the structure of data resulting in the instability of $\bar{\beta}$ values.

Mean parameter estimate instability was not the only critical finding. Increasing and decreasing trends of $\bar{\beta}$ values across data granularity showed a potential interaction between $\bar{\beta}$ values (Figure 1.6 and 1.7). This was interpreted as a clear sign of interaction between parameter estimates (β). Observing the β values of ΔT_{Up} , ΔT_{Base} and ΔT_{Over} parameters across streams also supported this interpretation and revealed a sign of potential multicollinearity between β values

(Figure 1.31 and 1.32). Indeed, correlograms showed a clear increase of multicollinearity between β values, meaning that the overall independency of predictor variables decreased with data granularity (Figure 1.8 to 1.13). The correlation between these values across streams showed whether the weight of a predictor variable on predictions was changed with the weight of other predictor variables or not. Therefore, the higher the correlation, the higher the influences of predictor variables on each other. From the modeling perspective, if the correlation between environmental variables is high, those environmental variables cannot be considered as independent from each other, which violates one of the assumptions of linear regression models, that is, independence of model variables. Especially some β values (e.g., ΔT_{Up} , ΔT_{Base} and ΔT_{Over}) were found to be highly correlated across streams in all data granularity scenarios. This finding revealed that some predictor variables were correlated in the majority of streams.

The first main conclusion was that data granularity increased the overall correlation between environmental variables in the raw data (Figure 1.14 to 1.16) and that this increase in multicollinearity likely contributes to the instability of parameter estimates across levels of granularity. A potential reason for higher overall correlation was that the nature of some environmental variables, such as the altitude angle (α) varied systematically with data granularity. During the nighttime, the value of altitude angle was calculated as “0” in datasets based on the altitude angle equation. This caused substantial variation between the values during the daytime and nighttime. Daily and weekly data reduced this variation increasing the correlation between altitude angle and day length (S). The correlation between altitude angle versus baseflow heat flow (ΔT_{Base}) and overflow heat flow (ΔT_{Over}) increased with data granularity for the same reason. The signs of r were also helpful to better understand the relationship between variables. As expected, the level of data granularity did not change the

negative and positive sign of correlations because increasing time granularity should not have any substantial effect on increasing and decreasing trends of environmental variables.

The second main conclusion was that some of the environmental variables used in Model 10 were naturally correlated. Both discharge variables (i.e., Q_{Up} and $Q_{Down} - Q_{Up}$) were highly correlated with ΔT_{Up} . This was an expected result, considering the equation for ΔT_{Up} that includes ratio of upstream discharge (Q_{up}) and downstream discharge (Q_{Down}) (Appendix C: Model Parameter Calculation). In addition, obtaining a high r between Q_{Up} and $Q_{Down} - Q_{Up}$ was also an expected outcome since the value of $Q_{Down} - Q_{Up}$ was highly dependent on Q_{Up} . Furthermore, some environmental variables were found negatively or positively correlated. Negative correlation between ΔT_{Up} versus both discharge variables (i.e., Q_{Up} and $Q_{Down} - Q_{Up}$) was observed. This was a consequence of nature of the equation of ΔT_{Up} that includes upstream discharge (Q_{up}) as numerator (Appendix C: Model Parameter Calculation). In other words, as upstream discharge (Q_{up}) increased, ΔT_{Up} decreased. Moreover, day length and altitude angle were found positively correlated. This result matched with the natural processes since both variables mostly decrease between June and October in the Northern Hemisphere.

All these observations supported my conclusion that data granularity affects the multicollinearity between environmental variables in raw data, consequently affecting parameter estimates (β) and outcome of regression models. Therefore, the change in multicollinearity would certainly be of concern for decision-makers on environmental issues since multicollinearity affects model designing and selection processes. For example, increased multicollinearity makes it hard to separate the individual effects of each environmental variable (Alin 2010), as a result, making it hard to resolve the influence of driving environmental variables. Therefore, data granularity selection and potential multicollinearity between

environmental variables should be considered together while designing models. As an example, including altitude angle in the model may be redundant if the day length is also included if daily or weekly data will be used. For the same reason, there may be utility in avoiding the inclusion of naturally correlated environmental variables such as Q_{Up} and $Q_{Down} - Q_{up}$. Also, including correlated environmental variables may magnify the effect of a certain parameter (e.g., upstream discharge) on response variable and may cause uncertainties on evaluation of the effect of environmental variables. Another advantage of eliminating redundant environmental variables is that it may significantly reduce the effort for collecting environmental data and effort for modeling applications. A downside of this approach, however, is that overall predictive power may be lost due to the removal of variables (O'Brien 2017). Put another way, the cost of increasing predictive power in situations where “natural experiments” are conducted and the predictor variables may be correlated is that the parameter estimates are unstable, and as such, difficult to interpret.

How Do Models Perform with July-Restricted Data?

Overall model predictive power across data granularity did not substantially change with July-restricted data (Figure 1.17). In addition, I found that no significant difference between model predictive power for July-restricted and June-October data within the same data granularity scenario (Table 1.14). Although the sample size, (i.e., period of data, the number of data granularity scenarios or the number of streams) may not be enough to conclude that the effect of data granularity was significantly changed, one can expect that using longer time period (i.e., June-October) may cause lower prediction powers (Tian et. al. 2017), since there would be larger variations within the same environmental variable. For example, the variation in day length and altitude angle during June-October would be higher than July-restricted data, which

may lead low fit between observed and predicted temperature gradient. However, it should also be considered that limiting time period may increase multicollinearity between environmental variables and change the model outputs and performance (Cropper 1984). Therefore, without a multicollinearity analysis, it was hard to come up with a conclusion on the exact reasons for insignificant effects of using July-restricted data on model predictive powers.

My results also revealed the fact that model selection, based on model prediction power, can be affected by time period selection. Using July-restricted data reduced the effect of model complexity since Model 1 and Model 2 were grouped as the least-fitting models (Figure 1.19), whereas Model 1 through 4 were grouped as the least-fitting models when June-October data were used (Figure 1.4). This was a clear sign for the effect of time period on the importance of environmental variables. As the data were restricted to July, the importance of day length (S) and altitude angle (α) parameters, which appeared in Model 5 and upper models, was reduced, therefore, Model 3 and 4 were grouped with best-fitting models, even though they lacked these parameters. These results emphasize the importance of time period selection when interpreting the output of models.

Although my results showed no significant differences on model performances, using larger time periods may increase or decrease biases between predicted and observed values (Jetten et. al. 1999; Tian et. al. 2017). In both cases, decision-makers need to decide between the model performance and the purpose of their study. As I explained, the purpose of the study naturally overrides the model performance expectations in most cases, that is, decision-makers favor ecological relevance over model performance. The perspective that my study brought to this issue is that the models can be designed or re-adjusted by using different time periods (e.g., my study showed that using day length and altitude angle may not be necessary for July-

restricted data). This approach will reveal the critical environmental variables that should be used in their models or point out the redundant parameters, eventually resulting to better model predictions. By understanding the effect of data granularity and different time periods on their model performances, researchers can optimize and use their models without losing the ecological relevance of their data and without reducing their expectations from model reliability.

CONCLUSIONS AND IMPLICATIONS

Although this research provides a variety of insights into hydrological modeling, the following conclusions are of most importance:

- 1) Selection of data granularity is a significant factor in modeling applications as it directly affects parameter estimates, model selection, and goodness of fit measures. Therefore, arbitrary selection of data granularity may lead to conflicting insights across studies where none exist. If the selection of data granularity does not include a strong ecological relevance, then the model performance should be one of the biggest concerns while deciding on data granularity. Another concern should be the effect of data granularity on multicollinearity between predictor variables since multicollinearity may influence the model dynamics and performance. Because of different responses of models and streams to data granularity, it was impossible to propose a “best” selection of data granularity. However, my study clearly showed that model performance changes with the type of data granularity, giving a better perspective to researchers on possible consequences of arbitrary data granularity selections. More research on this topic is needed in ecological and environmental sciences, considering the lack of studies enlightening the remaining unknowns at this topic. A better understanding on the implications of data granularity

will help researchers to design models that work best for their purposes and this will lead to more accurate decisions on ecological implications.

- 2) The best-fitting model among the regression models was Model 10, however multicollinearity analyses showed that some of the parameters in Model 10 were dependent, which violates one of the assumptions of linear regression models. Thus, I suggest that additional work could be done to improve this model. More analysis, such as Variance Inflation Factors, on multicollinearity can be done to have a better understanding on which parameters are mostly causing the multicollinearity. Modifying Model 10, such as discarding and adding parameters, based on my findings may decrease the dependency of predictor variables to each other and this may lead a better understanding of which environmental variables have a greater effect on temperature gradient. Improvements to this model will help to improve predictions relevant to environmental applications, such as predicting the fish distributions based on temperature gradient predictions, the effect of variations in climate and the impact of groundwater withdrawal on stream temperature changes (Carlson et. al. 2020).
- 3) Variation between the same environmental variables across different streams showed that the characteristics of streams influence model dynamics and reliability. Because it is hard to design models that are specific to each stream, classifying streams based on some characteristics may help to find a generalized model for each stream class. Finding generalized models may reduce the costs of data collection and improve the model performances that result to more robust predictions that will help decision makers a better perspective in natural resource management. Because the effect of stream classification

on model performances and the selection of time aggregation is not well studied, I explore this issue in the next chapter of my thesis.

- 4) Using July-restricted data did not substantially influence overall model performances.

Different time periods can either reduce or increase the influence of environmental variables on temperature gradient predictions. However, my findings are insufficient to conclude whether restricting data improves model performance or not. In fact, time period selection is critically dependent on the purpose of a study and ecological relevance of time period. Therefore, selection of the model and time period is study specific. However, optimizing the models by using different time periods can be helpful to maximize model performance within an ecologically relevant time period.

APPENDICES

APPENDIX A: Tables

Table 1.4. Stream temperature models (Magnusson et. al. 2012; Andrews 2019).

Model 1	$\Delta T = \beta_0 + \beta_1(T_a - T_w)$
Model 2	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_2 \left(\frac{Q_{up}}{Q_{down}} \right)$
Model 3	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_4(\Delta T_{flow})$
Model 4	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_4(\Delta T_{flow}) + \beta_5(Q_{down} - Q_{up})$
Model 5	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_4(\Delta T_{flow}) + \beta_5(Q_{down} - Q_{up}) + \beta_6(S)$
Model 6	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_4(\Delta T_{flow}) + \beta_5(Q_{down} - Q_{up}) + \beta_7(\alpha)$
Model 7	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_4(\Delta T_{flow}) + \beta_5(Q_{down} - Q_{up}) + \beta_6(S) + \beta_7(\alpha)$
Model 8	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_5(Q_{down} - Q_{up}) + \beta_6(S) + \beta_8(\Delta T_{up}) + \beta_9(\Delta T_{base}) + \beta_{10}(\Delta T_{over})$
Model 9	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_5(Q_{down} - Q_{up}) + \beta_7(\alpha) + \beta_8(\Delta T_{up}) + \beta_9(\Delta T_{base}) + \beta_{10}(\Delta T_{over})$
Model 10	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_3(Q_{up}) + \beta_5(Q_{down} - Q_{up}) + \beta_6(S) + \beta_7(\alpha) + \beta_8(\Delta T_{up}) + \beta_9(\Delta T_{base}) + \beta_{10}(\Delta T_{over})$
Model 11	$\Delta T = \beta_0 + \beta_1(T_a - T_w) + \beta_9(\Delta T_{base}) + \beta_{10}(\Delta T_{over}) + \beta_{11} \left(\frac{1}{Q_{up}} * \left[(T_a + 273.16)^4 + (T_w + 273.16)^4 \right] \right) + \beta_{12} \left[\frac{1}{Q_{up}} * (e^{T_w} - e^{T_a}) \right] + \beta_{13} \left(\frac{1}{Q_{up}} * \alpha \right)$

Table 1.5. Intercept and parameter estimate values of Model 10 across different data granularity. Tobacco River June – October 2016 data was used.

	Data Granularity (hours)					
	1	2	6	12	24	168
Intercept	0.627	0.629	0.868	1.337	1.092	0.780
$T_a - T_w$	0.023	0.023	0.014	0.001	-0.023	-0.030
Q_{Up}	-1.166	-1.169	-1.430	-2.045	-2.630	-1.383
$Q_{Down} - Q_{Up}$	-0.579	-0.583	-0.522	0.242	0.366	-0.436
S	-0.006	-0.006	-0.023	-0.051	0.000	-0.056
α	0.019	0.019	0.022	0.023	0.001	0.000
ΔT_{Up}	-0.223	-0.225	-0.221	-0.055	-0.088	-0.071
ΔT_{Base}	0.138	0.139	0.125	-0.001	0.008	-0.016
ΔT_{Over}	-0.144	-0.145	-0.138	-0.031	-0.060	-0.067

Table 1.6. Streams and rivers with their regions (SLP: Southern Lower Peninsula; NLP: Northern Lower Peninsula; UP: Upper Peninsula), thermal classes, upstream latitudes, upstream longitude, downstream latitude, and downstream longitude (Zorn et. al. 2008; Andrews 2019).

Stream	Region	Thermal Class	Upstream Latitude	Upstream Longitude	Downstream Latitude	Downstream Longitude
Pokagon Creek	SLP	C	41.89517	-86.162632	41.915803	-86.175679
Pigeon River	SLP	CT	42.932887	-86.081828	42.91636	-86.146075
Nottawa Creek	SLP	WT	42.192564	-85.060415	42.195998	-85.104618
Middle Branch Tobacco River	SLP	WT	43.909194	-84.697312	43.929905	-84.666327
Hasler Creek	SLP	W	43.042332	-83.423206	43.083594	-83.442947
Prairie River	SLP	W	41.801832	-85.116614	41.832568	-85.165065
Swan Creek	SLP	W	41.90477	-85.297885	41.921249	-85.312047
Cedar Creek	NLP	C	44.375846	-85.972647	44.369588	-85.999598
Cedar River	NLP	C	44.956875	-85.132748	44.968664	-85.138993
East Branch Black River	NLP	C	45.070651	-84.283728	45.089439	-84.284929
Butterfield Creek	NLP	CT	44.273249	-85.094087	44.256377	-85.03362
Morgan Creek	UP	C	46.519698	-87.504502	46.521351	-87.494782
Spring Creek	UP	CT	46.512909	-90.156133	46.513418	-90.177011
Carp River	UP	CT	46.509131	-87.418924	46.510534	-87.388497
Middle Branch Escanaba River	UP	WT	46.420206	-87.797962	46.398398	-87.770883
Squaw Creek	UP	W	46.057035	-87.18974	45.985396	-87.140559

Table 1.7. Starting and ending day of year for sampling in each stream for 2016.

Stream Name	Data Start Date	Data End Date
Black River	177	284
Butterfield Creek	144	296
Carp River	151	285
Cedar Creek	144	296
Cedar River	143	296
Escanaba River	151	284
Hasler Creek	160	315
Morgan Creek	151	285
Nottawa Creek	138	289
Pigeon River	137	307
Pokagon Creek	137	307
Prairie River	138	289
Spring Creek	151	284
Squaw Creek	152	285
Swan Creek	139	289
Tobacco River	144	287

Table 1.8. Correlation matrix of β values with hourly data. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.

Parameter	Q_{up}	$Q_{Down} - Q_{up}$	S	α	$T_a - T_w$	ΔT_{Up}	ΔT_{Base}	ΔT_{Over}
Q_{up}	1.000	-0.665	0.233	0.224	-0.585	-0.326	0.493	-0.35
$Q_{Down} - Q_{up}$		1.000	-0.361	-0.194	0.686	0.339	-0.476	0.453
S			1.000	-0.615	0.152	-0.475	0.810	-0.600
α				1.000	-0.500	0.156	-0.269	0.195
$T_a - T_w$					1.000	-0.257	0.003	-0.130
ΔT_{Up}						1.000	-0.759	0.957
ΔT_{Base}							1.000	-0.832
ΔT_{Over}								1.000

Table 1.9. Correlation matrix of β values with 2-hour data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.

Parameter	Q_{up}	$Q_{Down} - Q_{up}$	S	α	$T_a - T_w$	ΔT_{Up}	ΔT_{Base}	ΔT_{Over}
Q_{up}	1.000	-0.599	0.159	0.248	-0.603	-0.273	0.440	-0.295
$Q_{Down} - Q_{up}$		1.000	-0.267	-0.256	0.695	0.318	-0.444	0.420
S			1.000	-0.605	0.155	-0.332	0.616	-0.429
α				1.000	-0.505	0.086	-0.239	0.108
$T_a - T_w$					1.000	-0.245	-0.003	-0.119
ΔT_{Up}						1.000	-0.758	0.959
ΔT_{Base}							1.000	-0.839
ΔT_{Over}								1.000

Table 1.10. Correlation matrix of β values with 6-hour data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.

Parameter	Q_{up}	$Q_{Down} - Q_{up}$	S	α	$T_a - T_w$	ΔT_{Up}	ΔT_{Base}	ΔT_{Over}
Q_{up}	1.000	-0.371	0.053	0.219	-0.582	-0.239	0.368	-0.200
$Q_{Down} - Q_{up}$		1.000	-0.289	-0.198	0.583	0.320	-0.403	0.382
S			1.000	-0.647	0.191	-0.321	0.566	-0.410
α				1.000	-0.473	0.077	-0.287	0.137
$T_a - T_w$					1.000	-0.197	-0.008	-0.106
ΔT_{Up}						1.000	-0.783	0.959
ΔT_{Base}							1.000	-0.858
ΔT_{Over}								1.000

Table 1.11. Correlation matrix of β values with 12-hour data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.

Parameter	Q_{up}	$Q_{Down} - Q_{up}$	S	α	$T_a - T_w$	ΔT_{Up}	ΔT_{Base}	ΔT_{Over}
Q_{up}	1.000	-0.614	-0.022	0.117	-0.739	-0.187	0.406	-0.158
$Q_{Down} - Q_{up}$		1.000	-0.016	-0.299	0.760	-0.018	-0.151	0.010
S			1.000	-0.749	0.150	-0.432	0.531	-0.554
α				1.000	-0.169	0.293	-0.444	0.523
$T_a - T_w$					1.000	-0.170	-0.106	-0.005
ΔT_{Up}						1.000	-0.895	0.907
ΔT_{Base}							1.000	-0.903
ΔT_{Over}								1.000

Table 1.12. Correlation matrix of β values with daily data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.

Parameter	Q_{up}	$Q_{Down} - Q_{up}$	S	α	$T_a - T_w$	ΔT_{Up}	ΔT_{Base}	ΔT_{Over}
Q_{up}	1.000	-0.654	-0.145	0.288	-0.811	-0.084	0.392	-0.141
$Q_{Down} - Q_{up}$		1.000	-0.226	-0.096	0.764	0.057	-0.293	0.204
S			1.000	-0.474	0.228	-0.560	0.461	-0.479
α				1.000	-0.277	0.416	-0.268	0.352
$T_a - T_w$					1.000	-0.227	-0.104	0.003
ΔT_{Up}						1.000	-0.909	0.896
ΔT_{Base}							1.000	-0.888
ΔT_{Over}								1.000

Table 1.13. Correlation matrix of β values with weekly data granularity. Model 10 with seasonal data was used to obtain β values. Correlation between variables were obtained across streams.

Parameter	Q_{up}	$Q_{Down} - Q_{up}$	S	α	$T_a - T_w$	ΔT_{Up}	ΔT_{Base}	ΔT_{Over}
Q_{up}	1.000	-0.838	0.055	-0.036	-0.904	0.081	0.127	-0.040
$Q_{Down} - Q_{up}$		1.000	-0.041	-0.006	0.930	-0.355	0.197	-0.212
S			1.000	-0.985	-0.034	0.044	0.015	0.115
α				1.000	0.017	-0.033	-0.023	-0.080
$T_a - T_w$					1.000	-0.298	0.067	-0.052
ΔT_{Up}						1.000	-0.925	0.861
ΔT_{Base}							1.000	-0.900
ΔT_{Over}								1.000

Table 1.14. Mean adjusted correlation (R^2) values from July-restricted and June-October across models. Student t-test was used to find p-values.

Data Granularity (hour)	Mean Adjusted Correlation		
	July 2016	June-October 2016	p-value
1	0.320	0.255	0.168
2	0.322	0.272	0.287
6	0.332	0.289	0.386
12	0.336	0.357	0.702
24	0.387	0.453	0.258

APPENDIX B: Figures

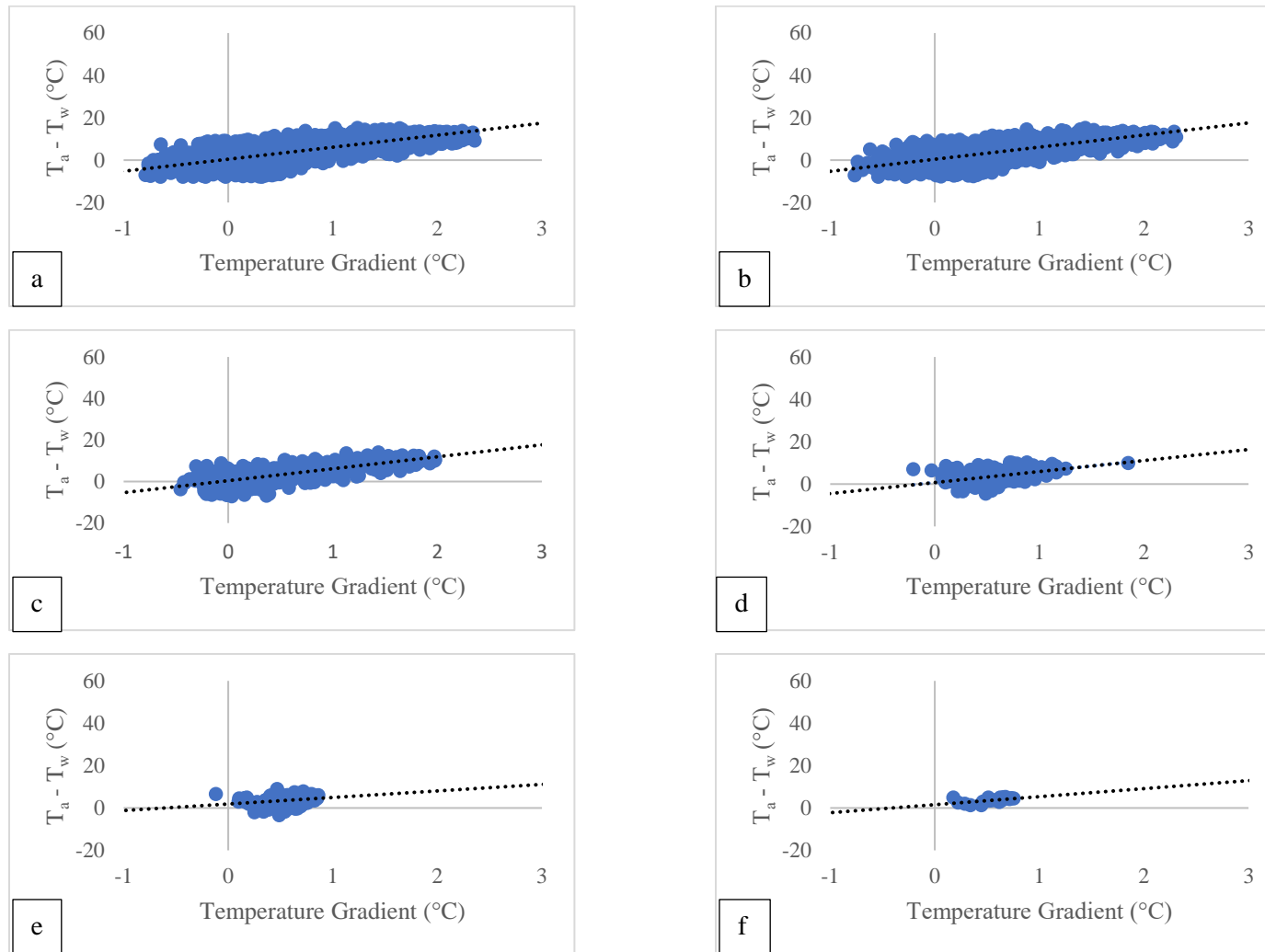


Figure 1.20. Air temperature – downstream temperature ($T_a - T_w$) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

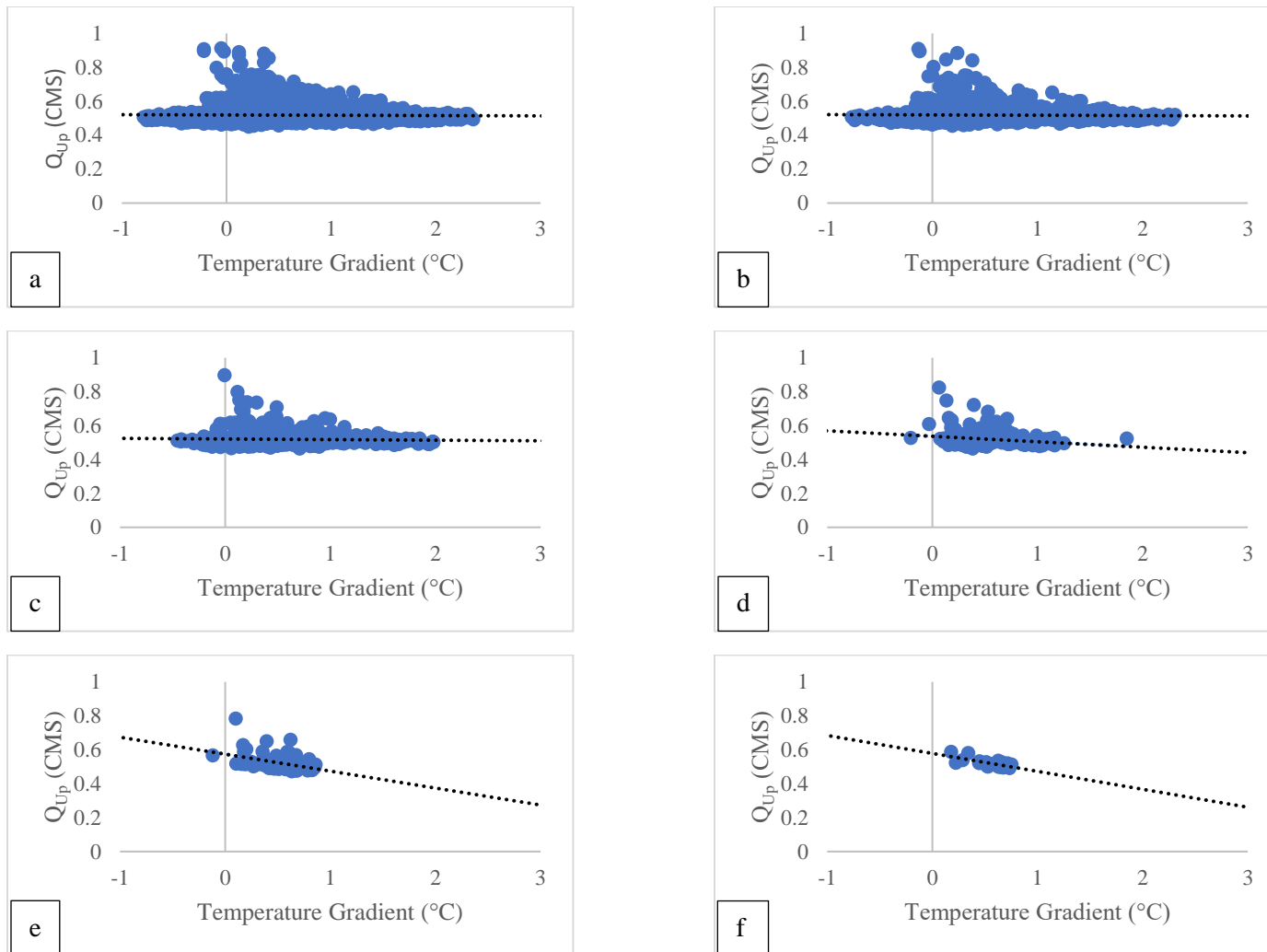


Figure 1.21. Upstream discharge (Q_{up}) (cubic meters per second – CMS) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

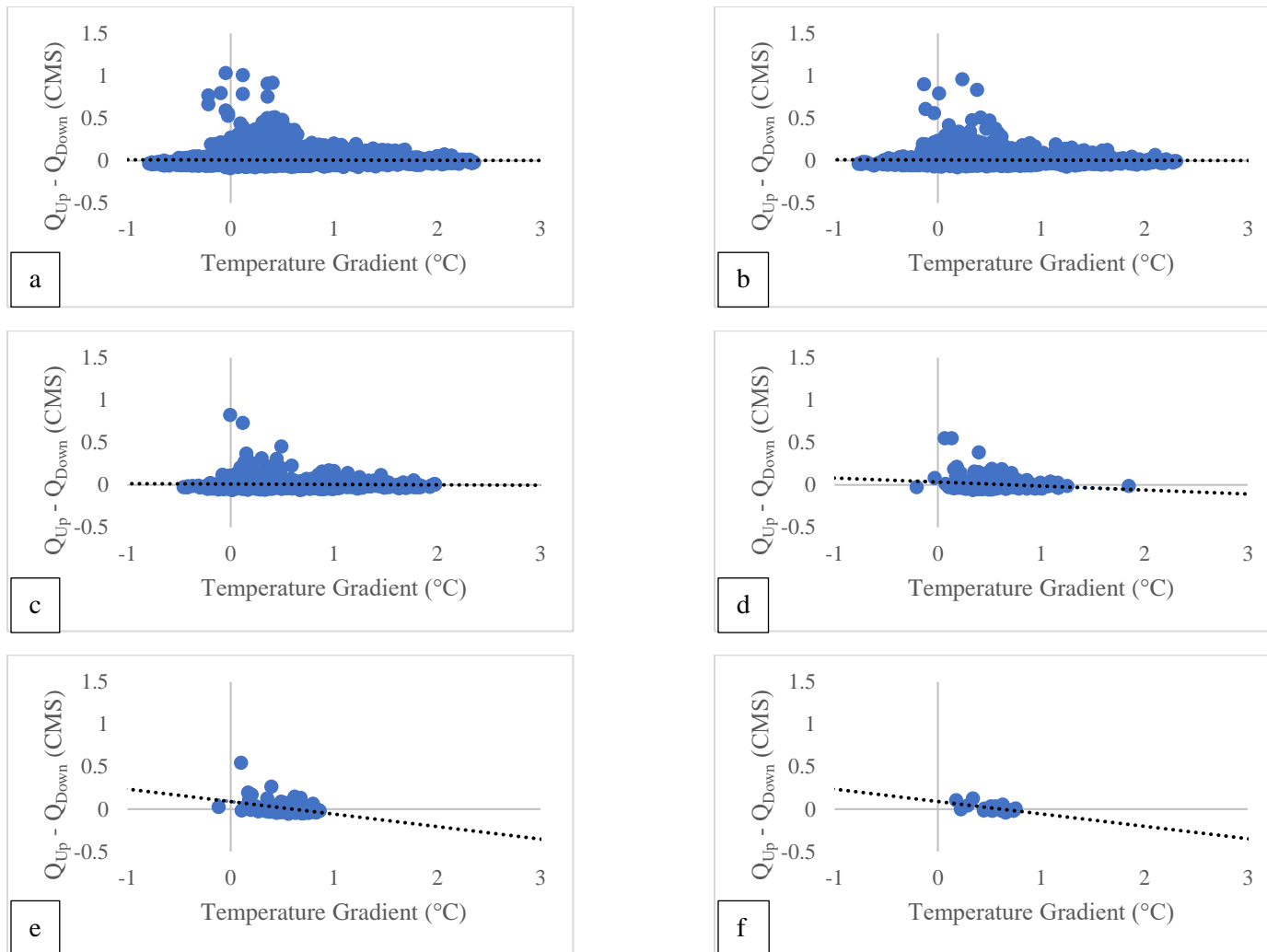


Figure 1.22. Upstream discharge – downstream discharge ($Q_{Up} - Q_{Down}$) (cubic meters per second – CMS) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

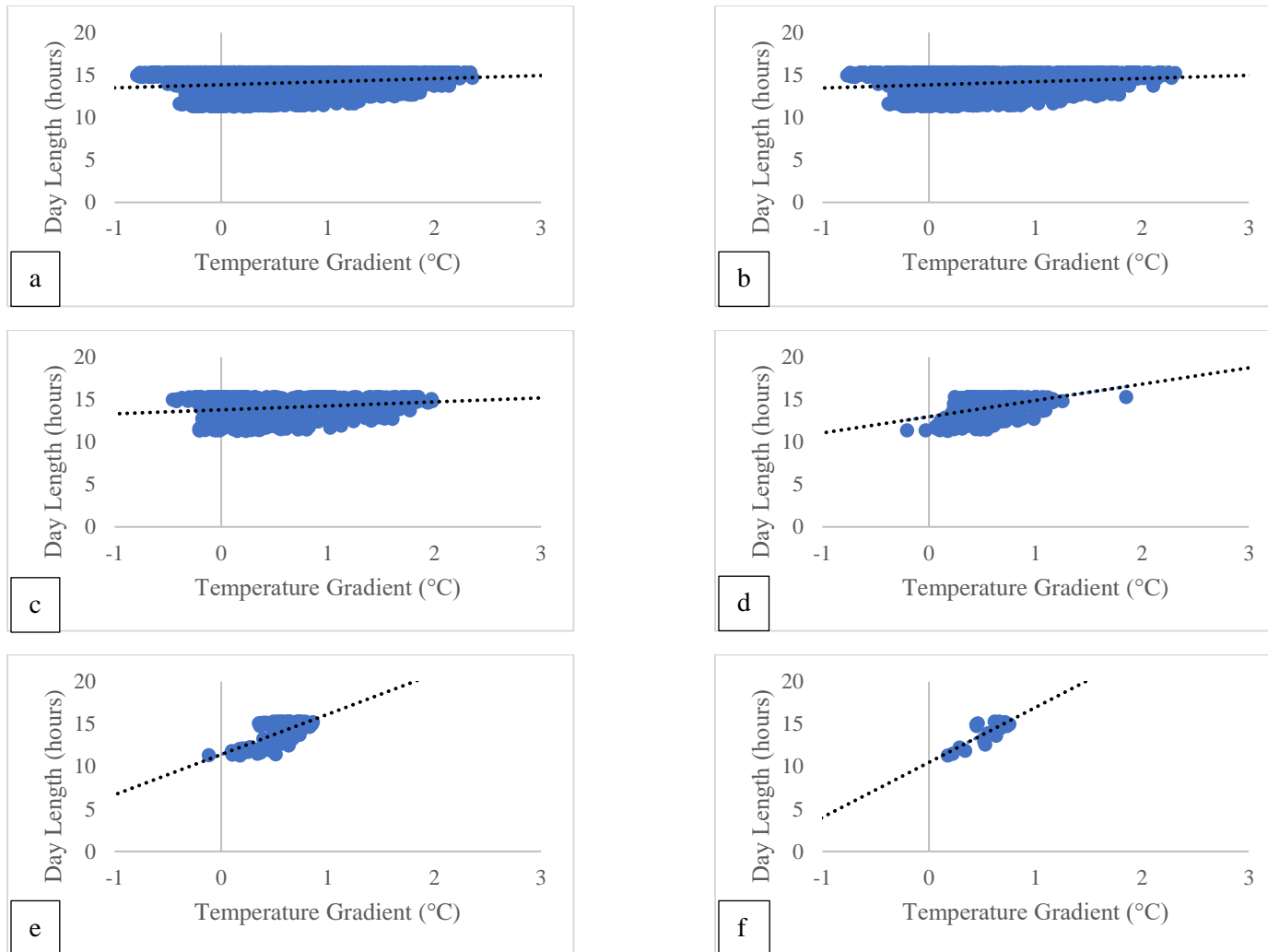


Figure 1.23. Day length (S) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

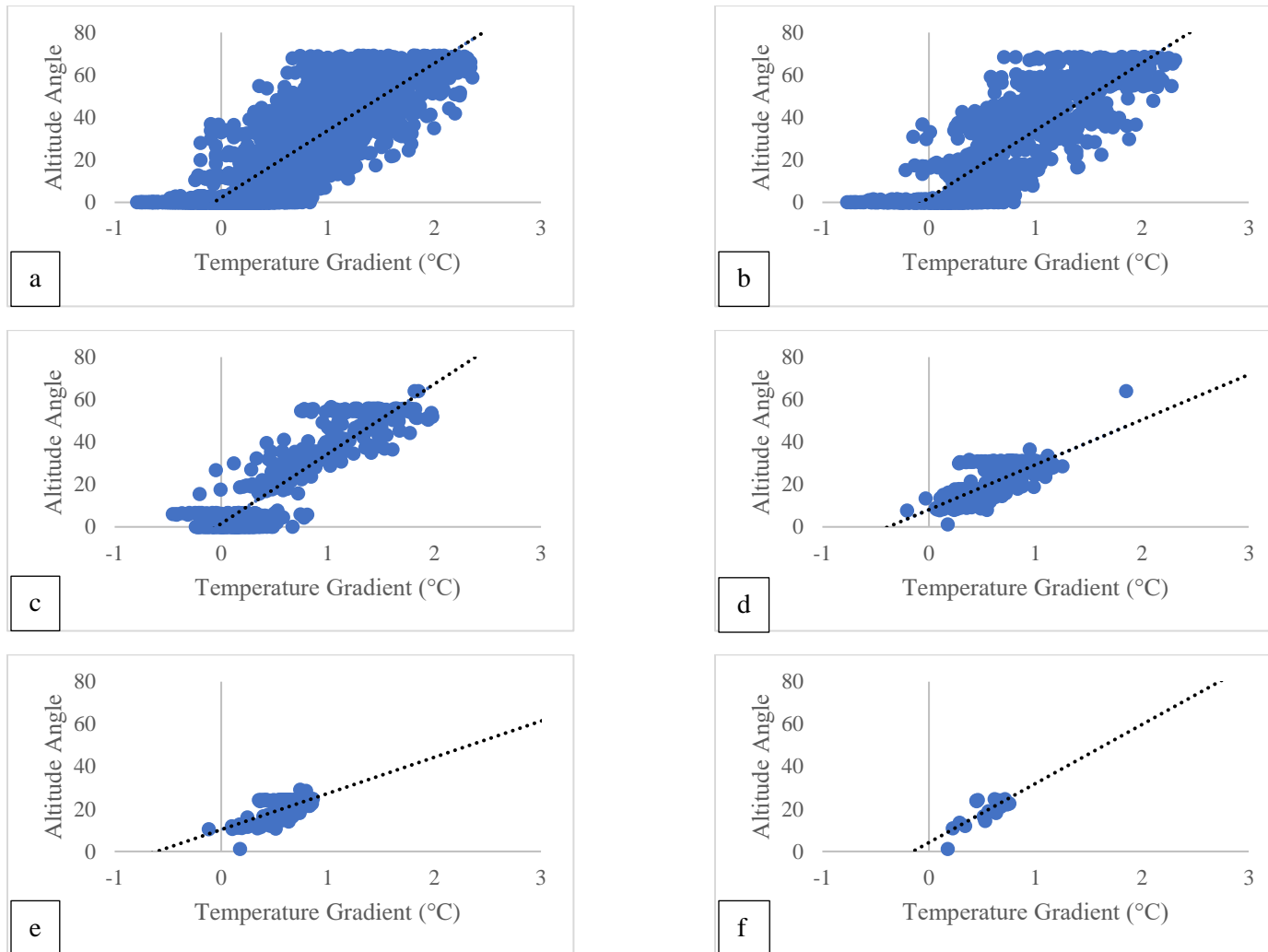


Figure 1.24. Altitude angle (α) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12-hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

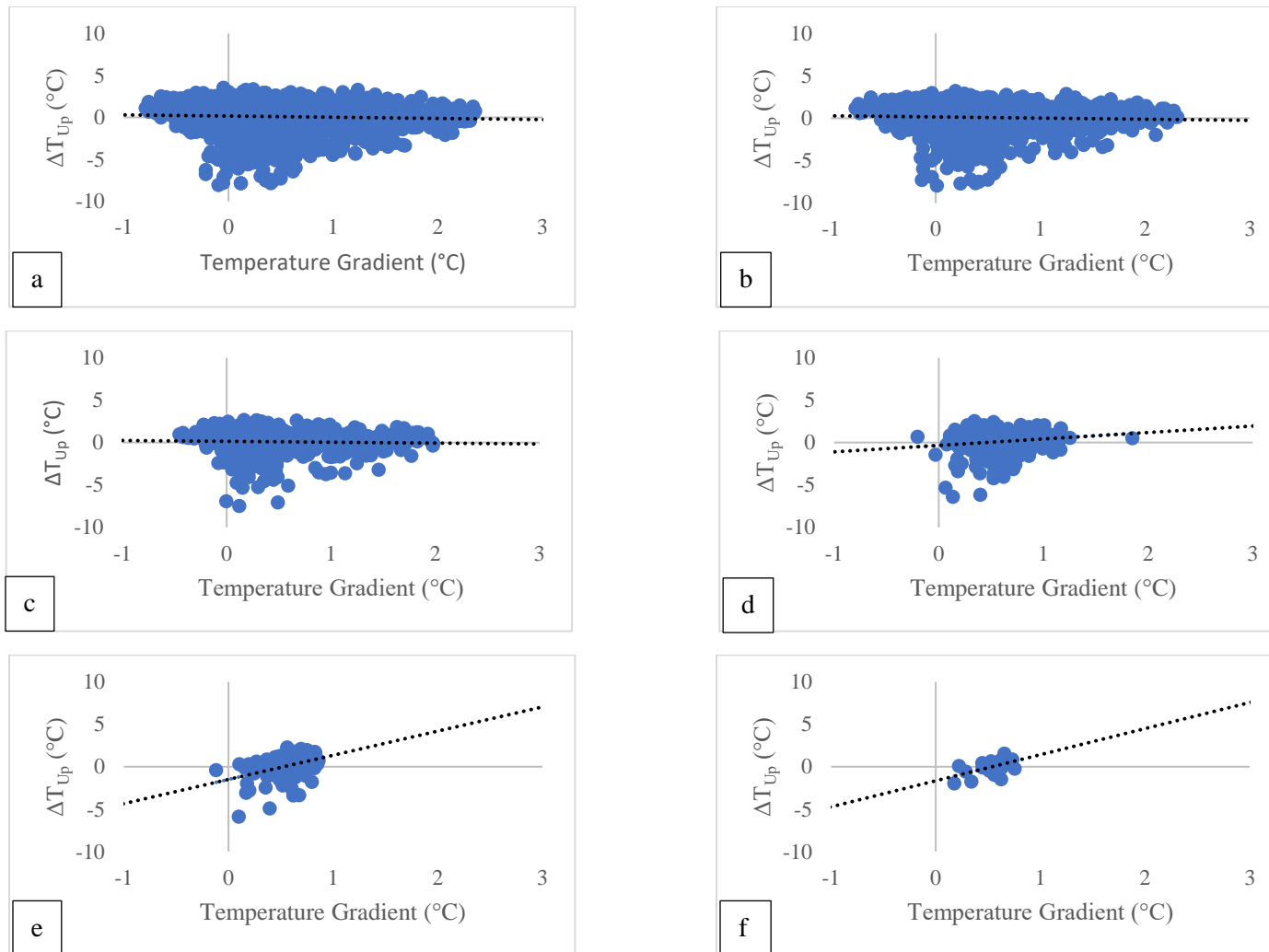


Figure 1.25. Upstream heat flow (ΔT_{Up}) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12-hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

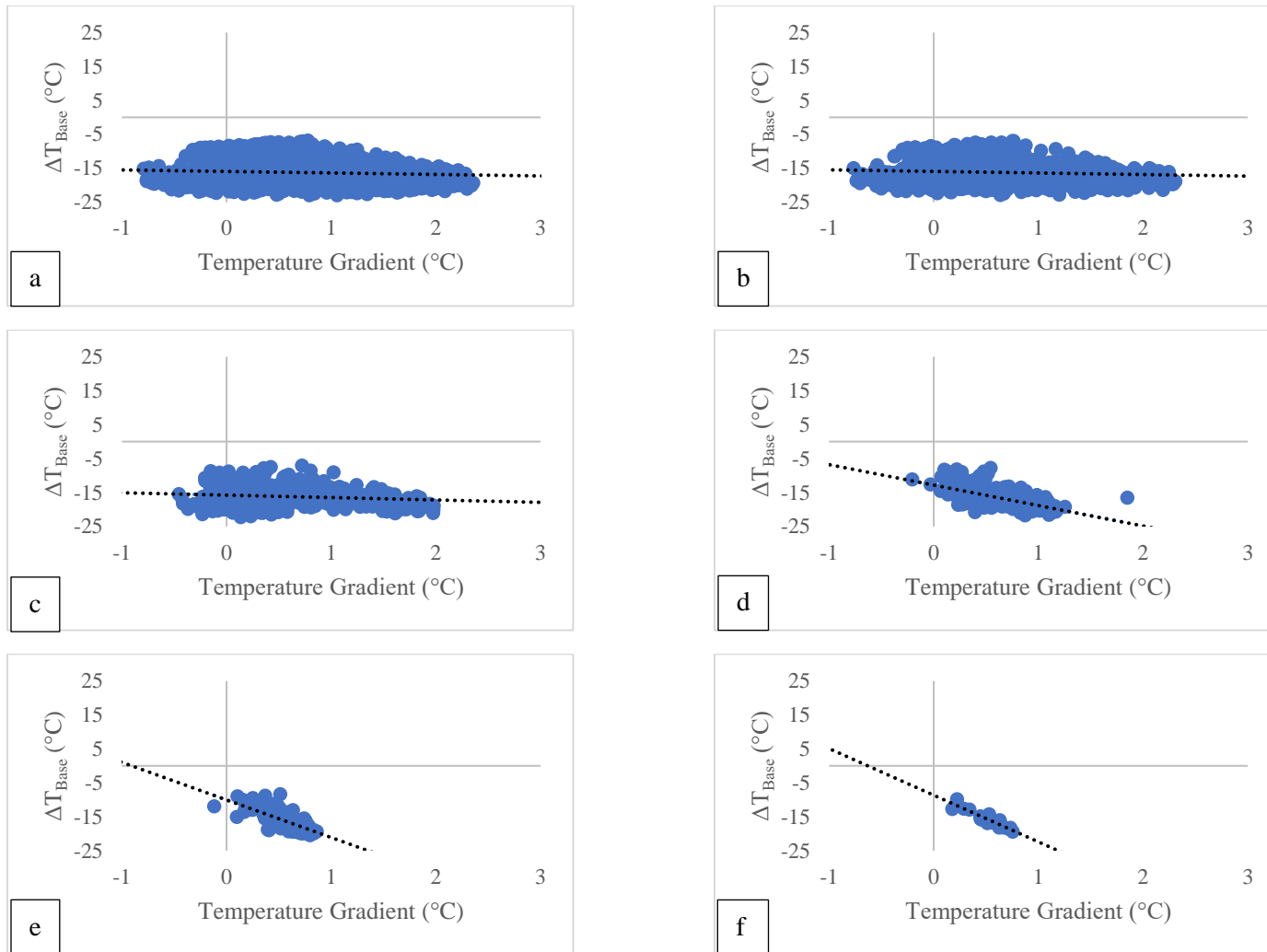


Figure 1.26. Baseflow heat flow (ΔT_{Base}) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12-hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

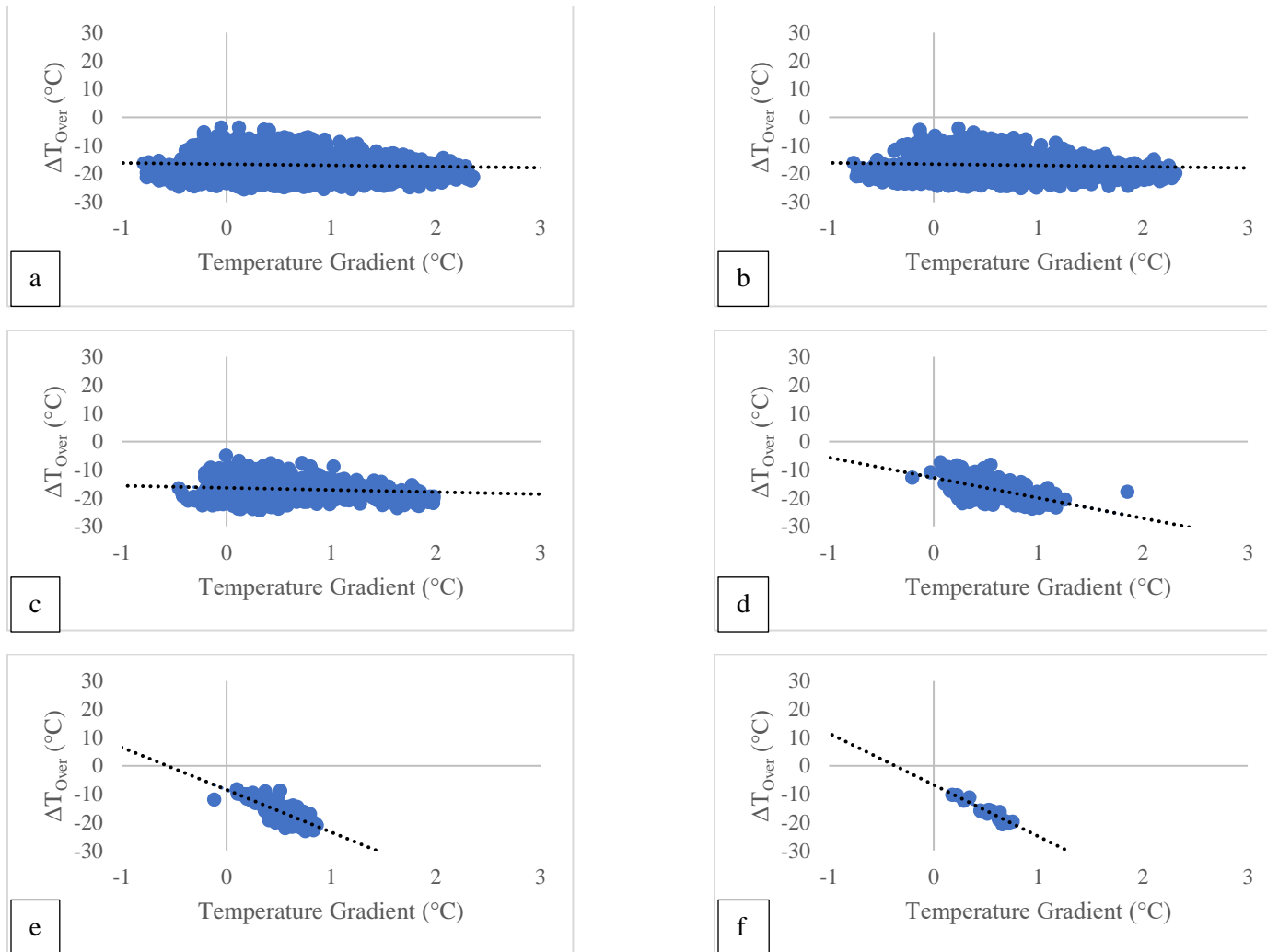


Figure 1.27. Overflow heat flow (ΔT_{Over}) across observed temperature gradient of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12-hour (d), daily (e), and weekly (f) data granularity between June-October 2016.

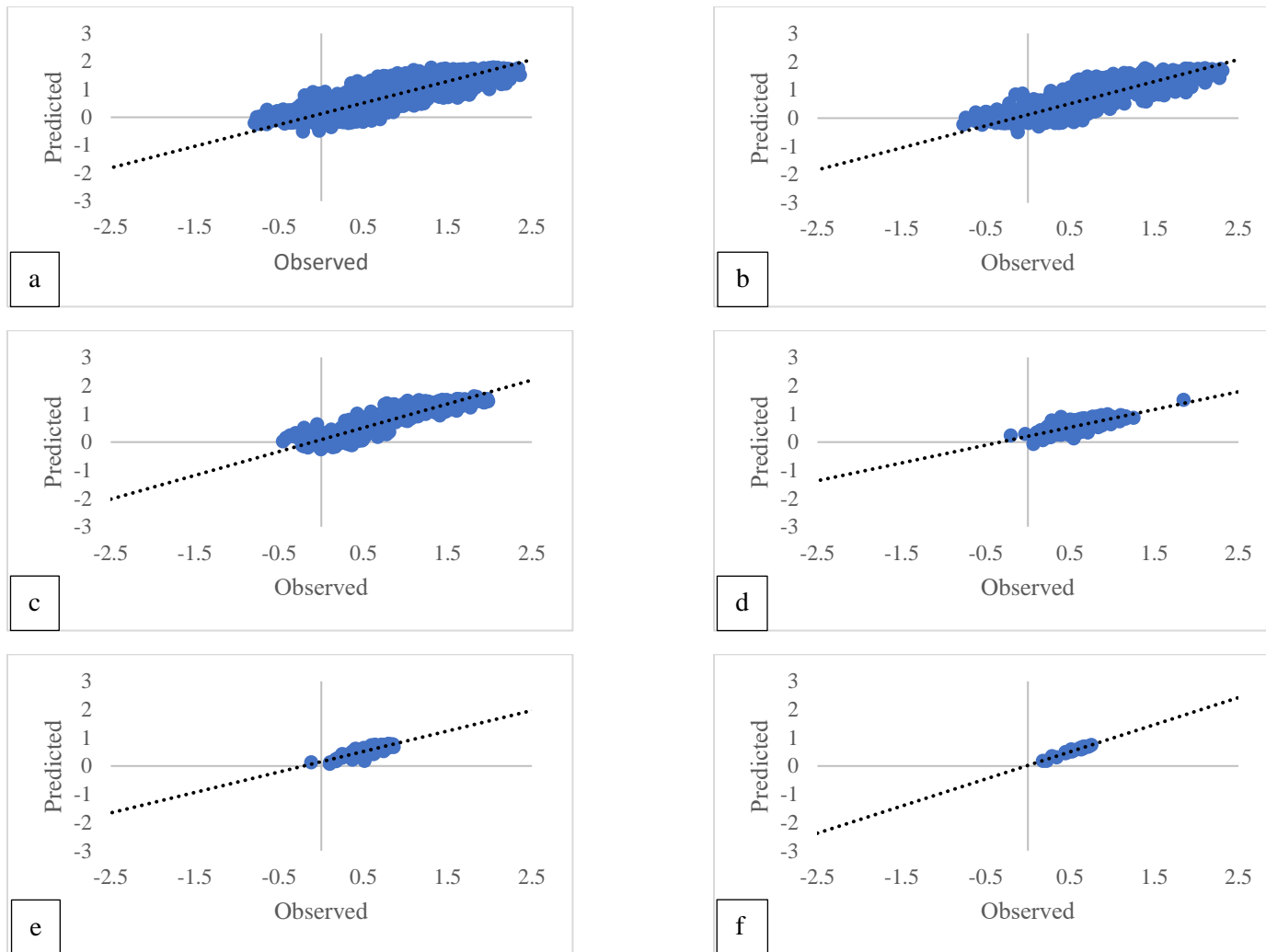


Figure 1.28. Observed and predicted temperature gradient ($^{\circ}\text{C}$) of Tobacco River with hourly (a), 2-hour (b), 6-hour (c), 12- hour (d), daily (e), and weekly (f) data granularity between June-October 2016. Predictions were obtained from Model 10.

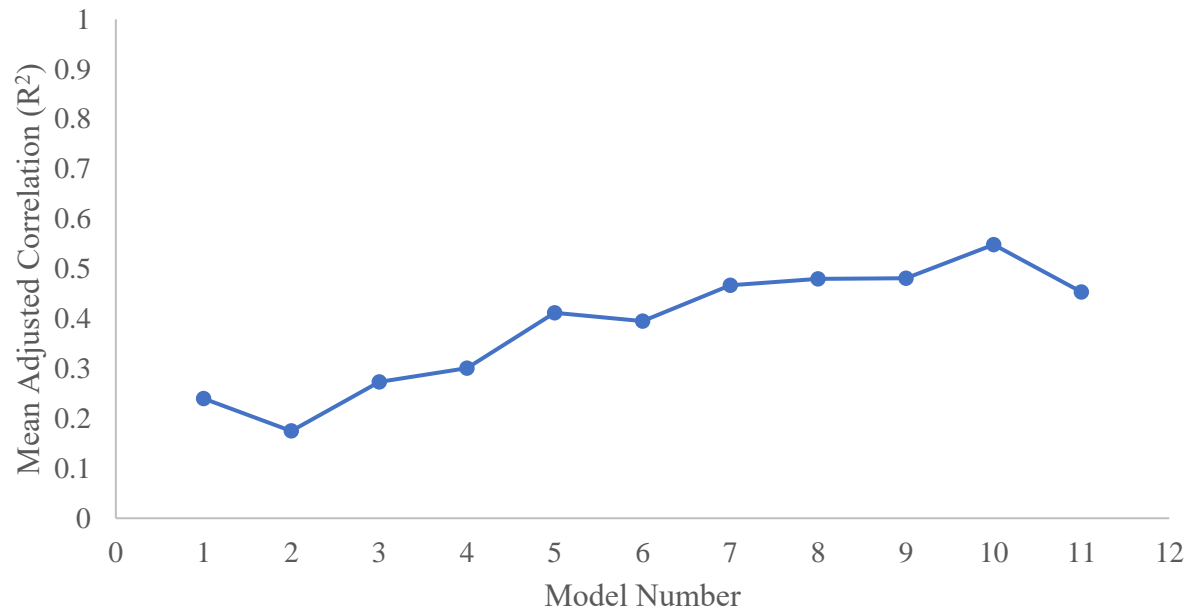


Figure 1.29. Mean adjusted correlation (R^2) values of models based on averaging all data granularity scenarios with June-October 2016 data.

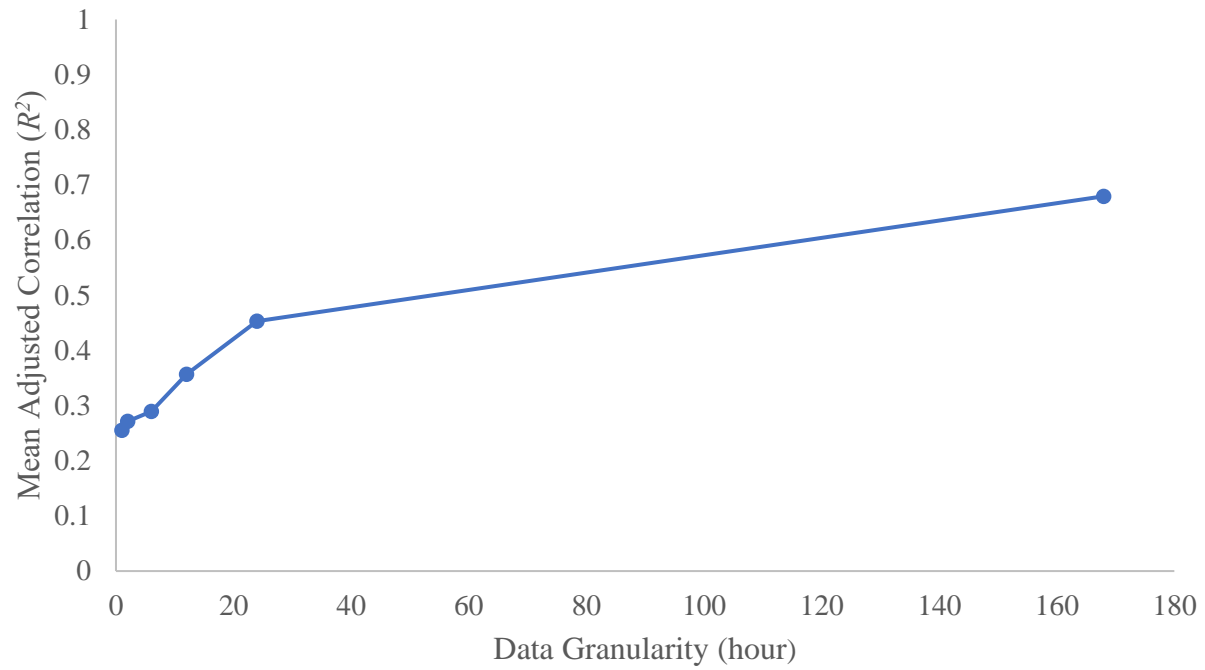


Figure 1.30. Mean adjusted correlation (R^2) values of each data granularity scenarios based on all regression models with June-October data.

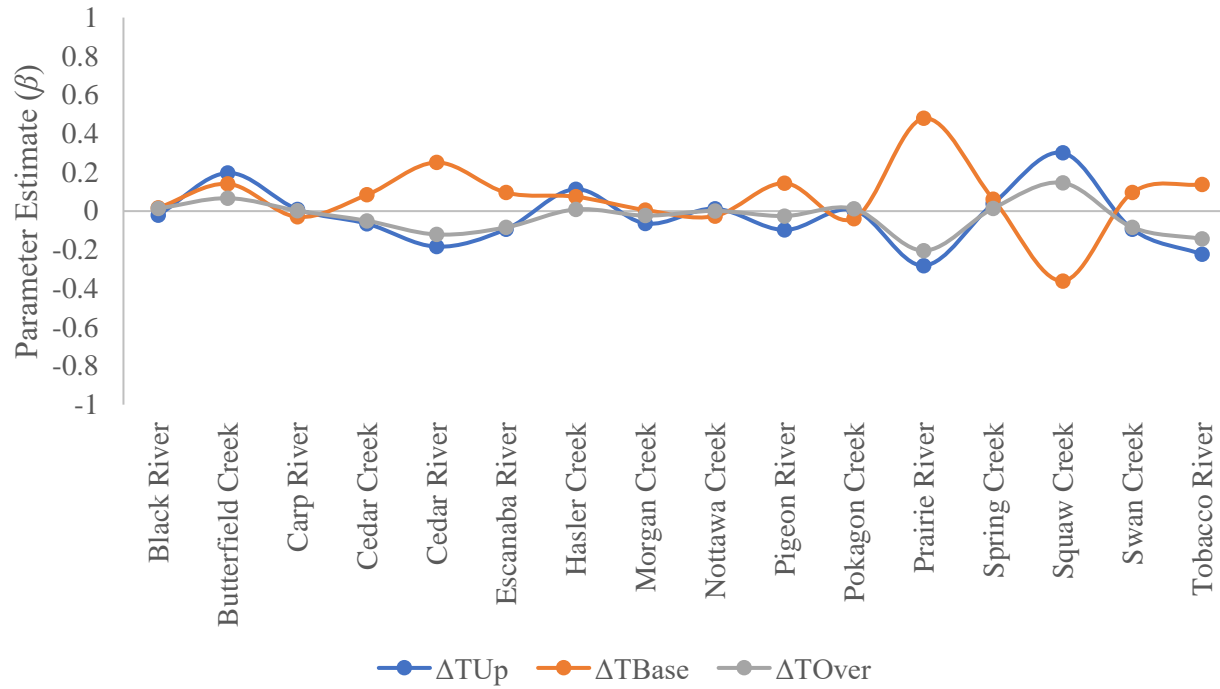


Figure 1.31. Parameter estimate (β) values of some predictor variables across streams with hourly June-October 2016 data. Model 10 was used to obtain β values for each stream.

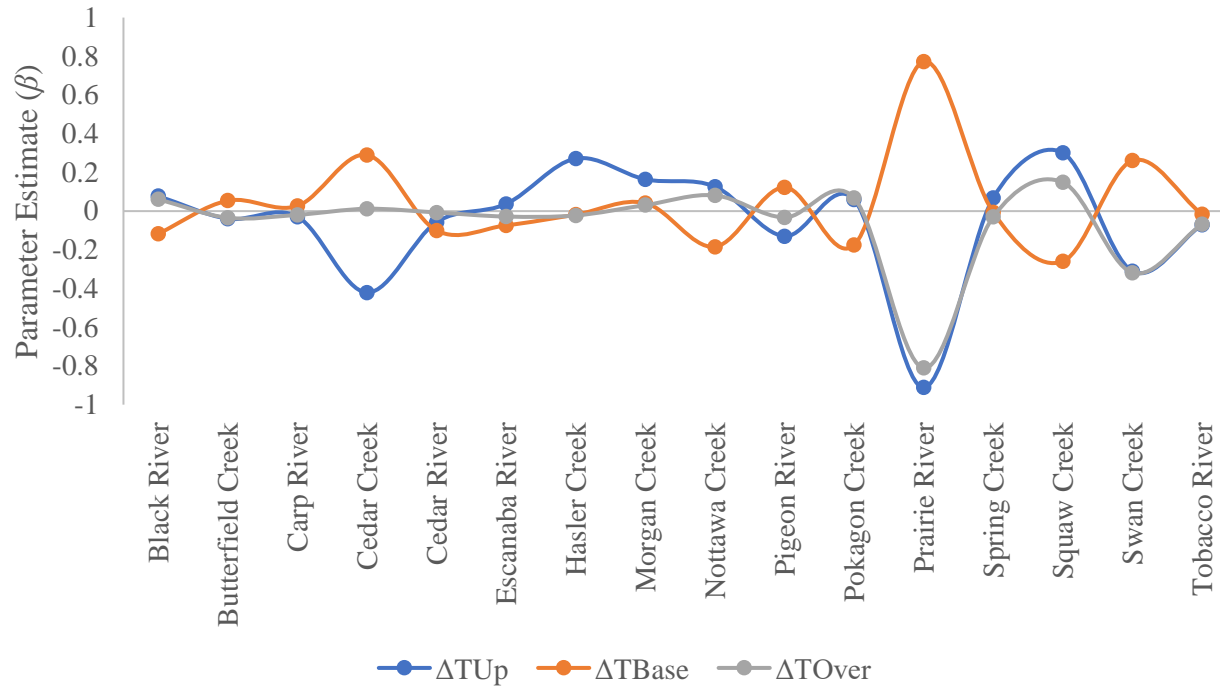


Figure 1.32. Parameter estimate (β) values of some predictor variables across streams with weekly June-October 2016 data. Model 10 was used to obtain β values for each stream.

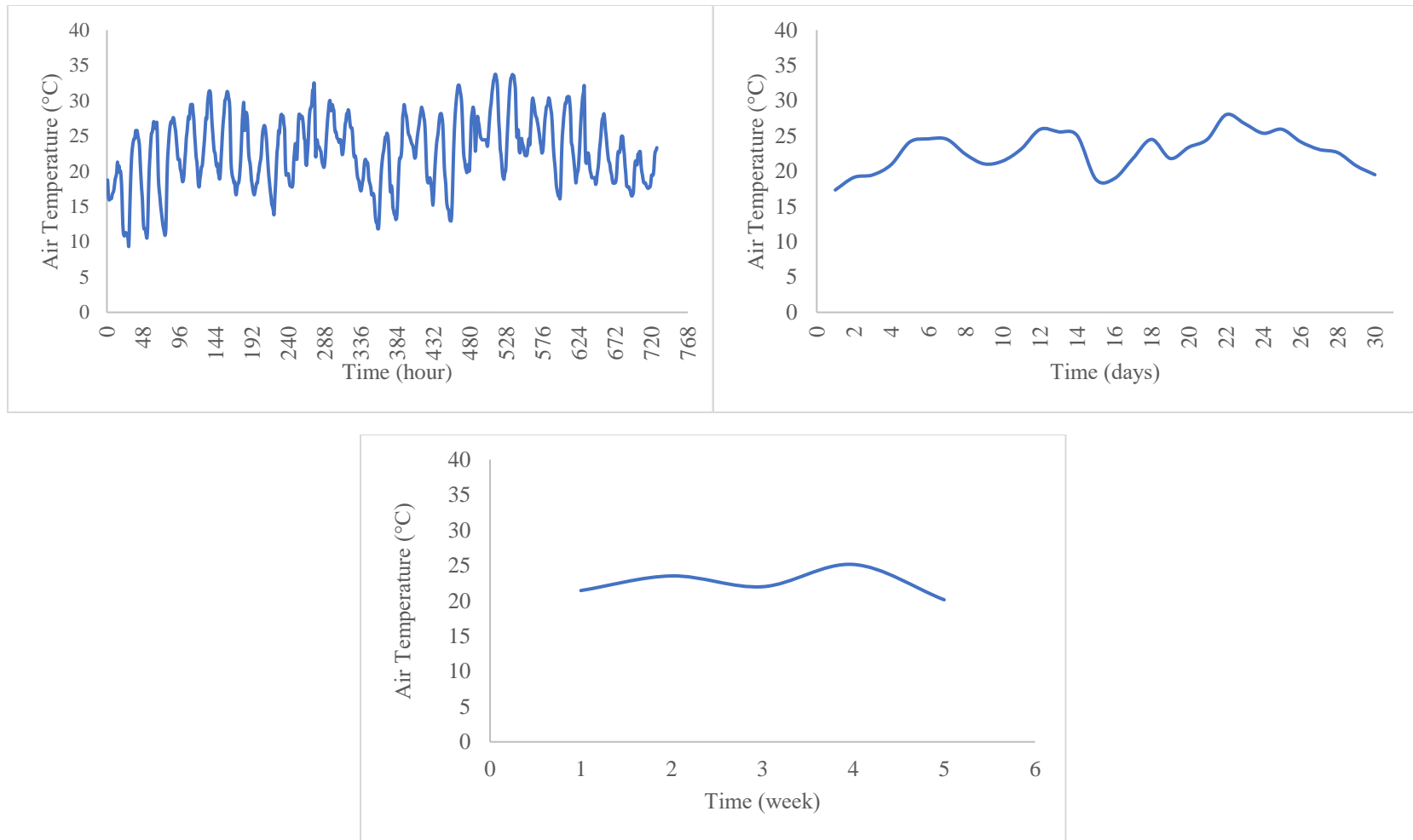


Figure 1.33. Air temperature across time with hourly (a), daily (b) and weekly (c) data granularity. Tobacco River July 2016 (July 1 – July 31) data were used.

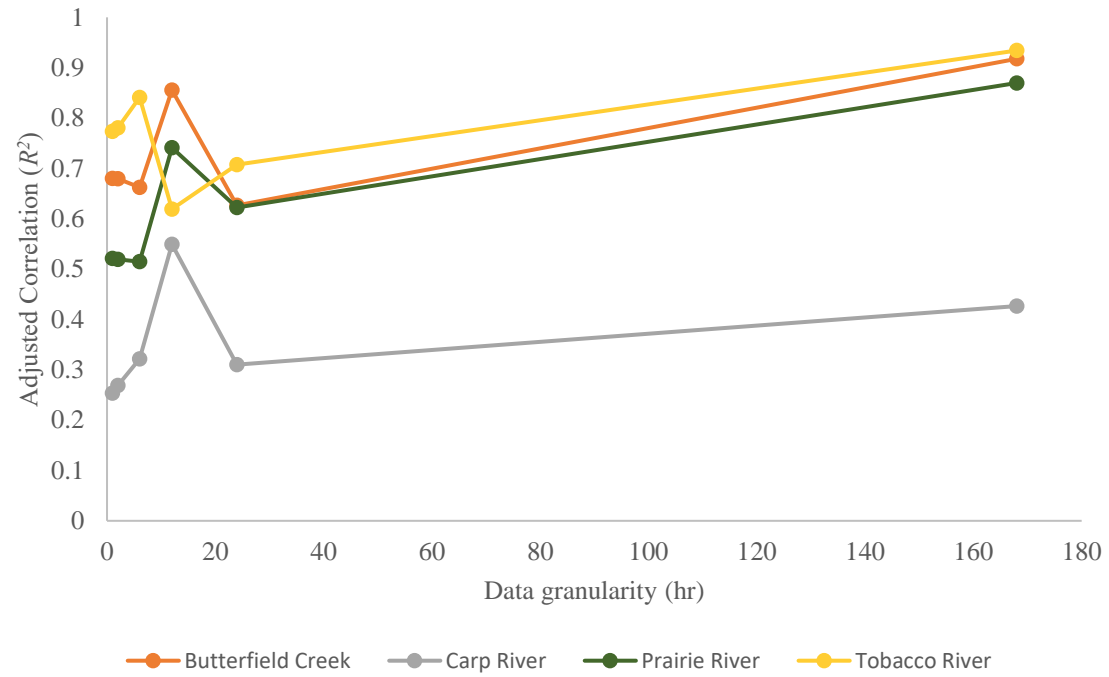


Figure 1.34. Adjusted correlation (R^2) values across data granularity. Model 10 was used with June-October 2016 data.

APPENDIX C: Model Parameter Calculation

Eqn. 10. ΔT_{flow} (Andrews 2019).

$$\Delta T_{flow} = T_{up} - \frac{[(Q_{up} \cdot T_{up}) + (Q_{base} \cdot T_{gw}) + (Q_{down} - Q_{up} - Q_{base}) \cdot \overline{T_{air}}]}{Q_{down}},$$

where Q_{up} : upstream discharge (cms); Q_{down} : downstream discharge (cms); Q_{Base} : baseflow discharge (cms); T_{Up} : upstream temperature ($^{\circ}\text{C}$), T_{gw} : groundwater temperature; $\overline{T_{air}}$: average air temperature of every 12-hour.

Eqn. 11. ΔT_{up} (Andrews 2019).

$$\Delta T_{up} = T_{up} - \frac{(Q_{up} \cdot T_{up})}{Q_{down}}$$

Eqn. 12. ΔT_{base} (Andrews 2019).

$$\Delta T_{base} = T_{up} - \frac{(Q_{base} \cdot T_{base})}{Q_{down}},$$

where T_{base} : baseflow temperature (cms).

Eqn. 13. ΔT_{over} (Andrews 2019).

$$\Delta T_{over} = T_{up} - \frac{[(Q_{down} - Q_{up} - Q_{base}) \cdot \overline{T_{air}}]}{Q_{down}}$$

Eqn. 14. day length (S) (Andrews 2019).

$$S = \frac{\tau}{15} \cdot 2,$$

$$\tau = \arccos [-\tan(lat) * \tan(\Delta)],$$

where lat : latitude and Δ : declination angle of the Sun:

$$\Delta = 23.5 \cdot \sin \left[\left(\frac{x}{365} \right) \cdot 360 \right],$$

where x : the number of days since the vernal equinox (March 21).

Eqn. 15. altitude angle (α) (Andrews 2019).

$$\alpha = \arcsin [\cos(lat) \cdot \cos(\Delta) \cdot \cos(h) + \sin(lat) \cdot \sin(\Delta)],$$

$$h = \frac{(No.of\ minutes\ past\ midnight,AST) - 720\ min}{4\ min/deg},$$

$$AST = LST + \left(4 \frac{min}{deg} \right) (LSTM - long) + ET$$

where LST : local standard time, $long$: longitude.

Eqn. 15. (cont'd)

$$LSTM = 15^\circ \cdot \frac{long}{15^\circ},$$

$$ET = 9.87 \cdot \sin(2D) - 7.53 \cdot \cos(D) - 1.5 \cdot \sin(D),$$

$$D = 360^\circ \cdot \left(\frac{N - 81}{365} \right),$$

where N : day of the year.

APPENDIX D: RStudio Codes

```
Output data=aggregate (Raw data, by=list (Raw data $ Day_of_Year, Raw data$X2hours),  
FUN=mean) #obtaining 2-hour data granularity.
```

```
Output data=aggregate (Raw data, by=list (Raw data $ Day_of_Year, Raw data$X6hours),  
FUN=mean) #obtaining 6-hour data granularity.
```

```
Output data=aggregate (Raw data, by=list (Raw data $ Day_of_Year, Raw data$X12hours),  
FUN=mean) #obtaining 12-hour data granularity.
```

```
Output data=aggregate (Raw data, by=list (Raw data $ Day_of_Year, Raw data$daily),  
FUN=mean) #obtaining daily data granularity.
```

```
Output data=aggregate (Raw data, by=list (Raw data $ Day_of_Year, Raw data$weekly),  
FUN=mean) #obtaining weekly data granularity.
```

```
Model10<- lm
```

```
(down_up_delta_tempc~air_tempc_minus_up_tempc+up_dischargecms+down_up_delta_discharge+day_length+altitude_angle+up_heat_load+base_heat_load+over_heat_load, data=Output data) #simulation of Model 10.
```

```
summary (Model10) # obtaining adjusted correlation ( $R^2$ ) and parameter estimates ( $\beta$ ).
```

```
AIC<-AIC (Model11, Model2, Model3, Model4, Model5, Model6, Model7, Model8, Model9, Model10, Model11, k=2) #obtaining AIC results for each granularity scenario and stream.
```

```
summary (AIC) #summarizing AIC results for each granularity scenario and stream.
```

```
Correlation matrix<-cor(Parameter estimate data) #obtaining correlation matrix for each data granularity scenario based on parameter estimates of Model 10.
```

```
round (Correlation matrix,2) # rounding the numbers in Correlation matrix.
```

```
install.packages("writexl") #installing "writexl" package.
```

```
library("writexl") #extracting "writexl" package.
```

```
write_xlsx(Correlation matrix,"File destination/Correlation matrix table.xlsx")
```

```
install.packages("corrplot") # installing "corrplot" package.
```

```
colnames(Correlation matrix)<-c("Qup", "Qdown - Qup", "Day Length", "Altitude Angle", "Ta - Tw", " $\Delta T_{up}$ ", " $\Delta T_{base}$ ", " $\Delta T_{over}$ ") #setting column names of correlogram.
```

```
rownames(Correlation matrix)<-c("Qup", "Qdown - Qup", "Day Length", "Altitude Angle", "Ta - Tw", " $\Delta T_{up}$ ", " $\Delta T_{base}$ ", " $\Delta T_{over}$ ") #setting row names of correlogram.
```

```
library(corrplot) # extracting "corrplot" package.
```

```
corrplot(Correlation matrix, type = "upper", order = "alphabet",
```

```
tl.col = "black", tl.srt = 45) # obtaining correlogram for each data granularity scenario.
```

BIBLIOGRAPHY

BIBLIOGRAPHY

- Ahmadi-Nedushan, B., St.-Hilaire, A., Ouarda, T. B. M. J., Bilodeau, L., Robichaud, É., Thiémonge, N., & Bobée, B. (2007). Predicting river water temperatures using stochastic models: Case study of the Moisie River (Québec, Canada). *Hydrological Processes*. <https://doi.org/10.1002/hyp.6353>
- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*. <https://doi.org/10.1093/biomet/60.2.255>
- Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*. <https://doi.org/10.1002/wics.84>
- Álvarez-Cabria, M., Barquín, J., & Peñas, F. J. (2016). Modeling the spatial and seasonal variability of water quality for entire river networks: Relationships with natural and anthropogenic factors. *Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2015.12.109>.
- Anderson, S. W., & Konrad, C. P. (2019). Downstream-Propagating Channel Responses to Decadal-Scale Climate Variability in a Glaciated River Basin. *Journal of Geophysical Research: Earth Surface*. <https://doi.org/10.1029/2018JF004734>.
- Andrews, R. (2019). Effects of flow reduction on thermal dynamics of streams: improving an important link in Michigan's water withdrawal assessment tool. M.S. Thesis, Michigan State University, East Lansing, MI.
- Bartholow, J. (2000a). The Stream Segment and Stream Network Temperature Models. Technical Report. U.S. Department of Interior. U.S. Geological Survey.
- Bartholow, J. M., Campbell, S. G., & Flug, M. (2004). Predicting the thermal effects of dam removal on the Klamath river. *Environmental Management*. <https://doi.org/10.1007/s00267-004-0269-5>
- Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T. B. M. ., & Bobée, B. (2007). A Review of Statistical Water Temperature Models. *Canadian Water Resources Journal*. <https://doi.org/10.4296/cwrj3203179>
- Borman, M. M., & Larson, L. L. (2003). A case study of river temperature response to agricultural land use and environmental thermal patterns. *Journal of Soil and Water Conservation*.

- Caissie, D., El-Jabi, N., & Satish, M. G. (2001). Modeling of maximum daily water temperatures in a small stream using air temperatures. *Journal of Hydrology*. [https://doi.org/10.1016/S0022-1694\(01\)00427-9](https://doi.org/10.1016/S0022-1694(01)00427-9)
- Carlson, A. K., Taylor, W. W., & Infante, D. M. (2020). Modeling effects of climate change on Michigan brown trout and rainbow trout: Precipitation and groundwater as key predictors. *Ecology of Freshwater Fish*. <https://doi.org/10.1111/eff.12525>
- Carpenter, S. R., Fisher, S. G., Grimm, N. B., & Kitchell, J. F. (1992). Global Change and Freshwater Ecosystems. *Annual Review of Ecology and Systematics*. <https://doi.org/10.1146/annurev.es.23.110192.001003>
- Cheng, S. T., & Wiley, M. J. (2016). A Reduced Parameter Stream Temperature Model (RPSTM) for basin-wide simulations. *Environmental Modeling and Software*. <https://doi.org/10.1016/j.envsoft.2016.04.015>
- Crittenden, R. N. (1978). Sensitivity analysis of a theoretical energy balance model for water temperatures in small streams, *Ecological Modeling*, Volume 5, Issue 3, 1978, Pages 207-224, ISSN 0304-3800, [https://doi.org/10.1016/0304-3800\(78\)90021-2](https://doi.org/10.1016/0304-3800(78)90021-2).
- Cropper, J. (1984). Multicollinearity within selected western north American temperature and precipitation data sets. *Tree Ring Bulletin*.
- Daoud, J. I. (2018). Multicollinearity and Regression Analysis. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/949/1/012009>
- Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z. I., Knowler, D. J., Lévêque, C., ... Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society*. <https://doi.org/10.1017/S1464793105006950>
- Edinger, John & Brady, D.K. & Geyer, J.C. (1974). Heat Exchange and Transport in the Environment.
- Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in Regression Analysis: The Problem Revisited. *The Review of Economics and Statistics*. <https://doi.org/10.2307/1937887>
- Ficke, A. D., Myrick, C. A., & Hansen, L. J. (2007). Potential impacts of global climate change on freshwater fisheries. *Reviews in Fish Biology and Fisheries*. <https://doi.org/10.1007/s11160-007-9059-5>.
- Fleury, P., Ladouche, B., Conroux, Y., Jourde, H., & Dörfliger, N. (2009). Modeling the hydrologic functions of a karst aquifer under active water management - The Lez spring. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2008.11.037>

- Galbraith, H. S., & Vaughn, C. C. (2009). Temperature and food interact to influence gamete development in freshwater mussels. *Hydrobiologia*. <https://doi.org/10.1007/s10750-009-9933-3>
- Haitovsky, Y. (1969). Multicollinearity in Regression Analysis: Comment. *The Review of Economics and Statistics*. <https://doi.org/10.2307/1926450>
- Hill, R. A., Hawkins, C. P., & Carlisle, D. M. (2013). Predicting thermal reference conditions for USA streams and rivers. *Freshwater Science*. <https://doi.org/10.1899/12-009.1>
- Isaak, D. J., Wollrab, S., Horan, D., & Chandler, G. (2012). Climate change effects on stream and river temperatures across the northwest U.S. from 1980-2009 and implications for salmonid fishes. *Climatic Change*. <https://doi.org/10.1007/s10584-011-0326-z>.
- Iversen, T.M. (1971). The ecology of a mosquito population (*Aedes communis*) in a temporary pool in a Danish beech wood. *Arch. Hydrobiol.*, 69: 309-332.
- Jackson, H. M., Gibbins, C. N., & Soulsby, C. (2007). Role of discharge and temperature variation in determining invertebrate community structure in a regulated river. *River Research and Applications*. <https://doi.org/10.1002/rra.1006>
- Janssen, P. H. M., & Heuberger, P. S. C. (1995). Calibration of process-oriented models. *Ecological Modeling*. [https://doi.org/10.1016/0304-3800\(95\)00084-9](https://doi.org/10.1016/0304-3800(95)00084-9)
- Jetten, V., De Roo, A., & Favis-Mortlock, D. (1999). Evaluation of field-scale and catchment-scale soil erosion models. *Catena*. [https://doi.org/10.1016/S0341-8162\(99\)00037-5](https://doi.org/10.1016/S0341-8162(99)00037-5)
- Johnson, A. N., Boer, B. R., Woessner, W. W., Stanford, J. A., Poole, G. C., Thomas, S. A., & O'Daniel, S. J. (2005). Evaluation of an inexpensive small-diameter temperature logger for documenting ground water-river interactions. *Ground Water Monitoring and Remediation*. <https://doi.org/10.1111/j.1745-6592.2005.00049.x>
- Kim, M. Bradlow, E. and Iyengar, R. (2019) Selecting Data Granularity Using the Power Likelihood. Available at SSRN: <https://ssrn.com/abstract=3453170> or <http://dx.doi.org/10.2139/ssrn.3453170>
- Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*. <https://doi.org/10.1029/2005WR004362>
- Kools, L., & Phillipson, F. (2016). Data granularity and the optimal planning of distributed generation. *Energy*. <https://doi.org/10.1016/j.energy.2016.06.089>.
- Kroll, C. N., & Song, P. (2013). Impact of multicollinearity on small sample hydrologic regression models. *Water Resources Research*. <https://doi.org/10.1002/wrcr.20315>

- Magnusson, J., Jonas, T., & Kirchner, J. W. (2012). Temperature dynamics of a proglacial stream: Identifying dominant energy balance components and inferring spatially integrated hydraulic geometry. *Water Resources Research*. <https://doi.org/10.1029/2011WR011378>
- Mason, C. H., & Perreault, W. D. (1991). Collinearity, Power, and Interpretation of Multiple Regression Analysis. *Journal of Marketing Research*. <https://doi.org/10.2307/3172863>
- Mayer, T. D. (2012). Controls of summer stream temperature in the Pacific Northwest. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2012.10.012>
- Morales-Marín, L. A., Rokaya, P., Sanyal, P. R., Sereda, J., & Lindenschmidt, K. E. (2019). Changes in streamflow and water temperature affect fish habitat in the Athabasca River basin in the context of climate change. *Ecological Modeling*. <https://doi.org/10.1016/j.ecolmodel.2019.108718>.
- Morin, G., and Couillard, D. (1990). Predicting river temperatures with a hydrological model. In Encyclopedia of fluid mechanic, surface and groundwater flow phenomena. Edited by N.P. Cheremisinoff. Gulf Publishing Company, Huston, Tex., Vol. 10, Chap. 5. pp. 171–209.
- Nace, R. L. (1967). Water resources: A global problem with local roots. *Environmental Science and Technology*. 1. No. 7. July 1967.
- Nuhfer, A. J., Zorn, T. G., & Wills, T. C. (2017). Effects of reduced summer flows on the brook trout population and temperatures of a groundwater-influenced stream. *Ecology of Freshwater Fish*. <https://doi.org/10.1111/eff.12259>
- O'Brien, R. M. (2017). Dropping Highly Collinear Variables from a Model: Why it Typically is Not a Good Idea*. *Social Science Quarterly*. <https://doi.org/10.1111/ssqu.12273>
- Patrick, C. J., McGarvey, D. J., Larson, J. H., Cross, W. F., Allen, D. C., Benke, A. C., ... Woodward, G. (2019). Precipitation and temperature drive continental-scale patterns in stream invertebrate production. *Science Advances*. <https://doi.org/10.1126/sciadv.aav2348>
- Pilgrim, J. M., Fang, X., & Stefan, H. G. (1998). Stream temperature correlations with air temperatures in Minnesota: Implications for climate warming. *Journal of the American Water Resources Association*. <https://doi.org/10.1111/j.1752->
- Poole, M. A., & O'Farrell, P. N. (1971). The Assumptions of the Linear Regression Model. *Transactions of the Institute of British Geographers*. <https://doi.org/10.2307/621706>
- Prairie, Y. T. (1996). Evaluating the predictive power of regression models. *Canadian Journal of Fisheries and Aquatic Sciences*. <https://doi.org/10.1139/cjfas-53-3-490>
- Ruehl, C., Fisher, A. T., Hatch, C., Huertos, M. L., Stemler, G., & Shennan, C. (2006). Differential gauging and tracer tests resolve seepage fluxes in a strongly-losing stream. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2006.03.025>

- Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., ... Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. *Science*. <https://doi.org/10.1126/science.287.5459.1770>
- Sand-Jensen, K. (1989). Environmental variables and their effect on photosynthesis of aquatic plant communities. *Aquatic Botany*. [https://doi.org/10.1016/0304-3770\(89\)90048-X](https://doi.org/10.1016/0304-3770(89)90048-X).
- Seber, G. A. F., & Wild, C. J. (1989). Autocorrelated Errors. In *Nonlinear Regression*. <https://doi.org/10.1002/0471725315.ch6>
- Sinokrot, B. A., & Stefan, H. G. (1993). Stream temperature dynamics: Measurements and modeling. *Water Resources Research*. <https://doi.org/10.1029/93WR00540>
- Sinokrot, B. A., Stefan, H. G., McCormick, J. H., & Eaton, J. G. (1995). Modeling of climate change effects on stream temperatures and fish habitats below dams and near groundwater inputs. *Climatic Change*. <https://doi.org/10.1007/BF01091841>.
- Stefan, H. G., & Preud'homme, E. B. (1993). STREAM TEMPERATURE ESTIMATION FROM AIR TEMPERATURE. *JAWRA Journal of the American Water Resources Association*. <https://doi.org/10.1111/j.1752-1688.1993.tb01502.x>
- St-Hilaire, A., Morin, G., El-Jabi, N., & Caissie, D. (2000). Water temperature modeling in a small forested stream: Implication of forest canopy and soil temperature. *Canadian Journal of Civil Engineering*. <https://doi.org/10.1139/100-021>.
- Storey, R. G., Howard, K. W. F., & Williams, D. D. (2003). Factors controlling riffle-scale hyporheic exchange flows and their seasonal changes in a gaining stream: A three-dimensional groundwater flow model. *Water Resources Research*. <https://doi.org/10.1029/2002WR001367>
- Tian, Y., Xiao, Z., Wang, H., Peng, X., Guan, L., Huangfu, Y., ... Feng, Y. (2017). Influence of the sampling period and time resolution on the PM source apportionment: Study based on the high time-resolution data and long-term daily data. *Atmospheric Environment*. <https://doi.org/10.1016/j.atmosenv.2017.07.003>
- Vogt, T., Schneider, P., Hahn-Woernle, L., and Cirpka, O. A. (2010). Estimation of seepage rates in a losing stream by means of fiber-optic high-resolution vertical temperature profiling. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2009.10.033>
- Webb, B. W., Clack, P. D., & Walling, D. E. (2003). Water-air temperature relationships in a Devon river system and the role of flow. *Hydrological Processes*. <https://doi.org/10.1002/hyp.1280>

- Wehrly, K. E., Wang, L., & Mitro, M. (2007). Field-Based Estimates of Thermal Tolerance Limits for Trout: Incorporating Exposure Time and Temperature Fluctuation. *Transactions of the American Fisheries Society*. <https://doi.org/10.1577/t06-163.1>
- Wetzel, R.G. (1960). Marl encrustation on hydrophytes in several Michigan lakes. *Oikos*, 11: 223- 236.
- Wojtalik, T. A., & Waters, T. F. (1970). Some Effects of Heated Water on the Drift of Two Species of Stream Invertebrates. *Transactions of the American Fisheries Society*. [https://doi.org/10.1577/1548-8659\(1970\)99<782:seohwo>2.0.co;2](https://doi.org/10.1577/1548-8659(1970)99<782:seohwo>2.0.co;2).
- Woltemade, C.J., and Hawkins, T.W. (2016). Stream Temperature Impacts Because Of Changes In Air Temperature, Land Cover And Stream Discharge: Navarro River Watershed, California, USA. *River Research and Applications* 32:2020-2031. DOI: 10.1002/rra.3043
- Zorn, T. G. (2008). A Regional-Scale Habitat Suitability Model to Assess the Effects of Flow Reduction on Fish Assemblages in Michigan Streams¹, (October 2017). <https://doi.org/10.1111/j.1752-1688.2012.00656.x>
- Zorn, T.G., Seelbach, P.W., and Wiley, M.J. (2004). Utility of Species-Specific, Multiple Linear Regression Models for Prediction of Fish Assemblages in Rivers of Michigan's Lower Peninsula. Michigan Department of Natural Resources, Fisheries Research Report 2072, Ann Arbor, Michigan.
<http://www.michigandnr.com/PUBLICATIONS/PDFS/ifr/ifrlibra/Research/reports/2072rr.pdf>

CHAPTER 2: THE EFFECT OF STREAM THERMAL CLASSIFICATION AND DATA POOLING ON TEMPERATURE GRADIENT MODELING IN MICHIGAN'S STREAMS

INTRODUCTION

A Challenge in Stream Management: Limited Data

Data availability is critically important for environmental studies. Availability and integrity of environmental data determines the outcomes of environmental studies, and eventually influence the decisions for environmental problems. Data limitation is a global problem and might be a consequence of many factors, such as limited time, intensive labor need and high costs (Niemczynowicz 1999; Tavares Wahren et. al. 2016). Although the reason for data limitation varies case by case, the need for making environmental predictions with limited data is a common problem. In some cases, reducing the number of data collection sites by determining reference data collection sites (e.g., McManamay et. al. 2018) can be a reasonable solution to reduce the expenses of data collection procedures. Data collection sites are usually determined by a set of key environmental characteristics that vary between environments and are commonly used for classifying these environments. For example, different hydrological (e.g., thermal) and ecological (e.g., species diversity) characteristics of streams are used for stream classification, and they help identifying reference data collection sites that represent a broader group of streams (Zorn et. al. 2008; Leathwick et. al. 2011; Maheu et. al. 2016). Therefore, stream classification has been an effective tool to reduce the costs of data collection and has been an important topic in environmental sciences. Moreover, key stream characteristics (e.g., discharge change) help researchers gain deeper insight and make better predictions of other environmental variables (e.g., groundwater inflow or outflow) for which data collection might be challenging.

Because of its importance to data collection practices and needs, I will primarily focus on stream classification in this study and its use for reducing the need for extensive data collection. However, the consequences of those applications will also be under focus. A detailed analysis and interpretation of the outcomes of stream classification and its applications on linear regression models will be the main theme as there is no such study that was dedicated to this issue for Michigan's streams. Although Zorn et. al. (2004, 2008) and Andrews (2019) considered the consequences of stream classification for stream temperature and temperature gradient modeling, some concepts related to these issues remained unknown. For example, linear regression models have not been generalized and applied based on stream classes. Before explaining possible applications of stream classes to linear regression models and its possible results, I will touch on some applications of stream classification in the United States and in Michigan.

Recent History of Stream Classification

Classification of streams has been a useful tool in stream management in many aspects (Tadaki et. al. 2014) and many different approaches have been adopted. Classification of streams has been based on various characteristics of streams. The U.S. Environmental Protection Agency (EPA), for example, uses average water depth, surface area, water velocity and sediment type to classify streams and rivers (ROSGEN stream classification) in the United States (Rosgen 1994,1996). Stream temperature has been considered as another classification criteria since water temperature is an important water quality criterion and can help decision-makers to monitor anthropogenic effects. For example, Maheu et. al. (2016) characterized the thermal regime of streams by describing the patterns in water temperature variability at a national scale. They used annual mean stream temperatures that were obtained from daily mean stream temperatures at 79

sites. They also included annual and diel water temperature variability in their classification by using other environmental variables such as air temperature. Based on these inputs, they developed six stream thermal classes: highly variable cool, variable cold, variable cool, variable warm, stable cold and stable cool. Based on their findings, they mapped streams nation-wide based on their stream classes (Figure 2.1). In addition to such wide-scale classification, researchers have classified streams at smaller scales since local environmental variables can also be critical.

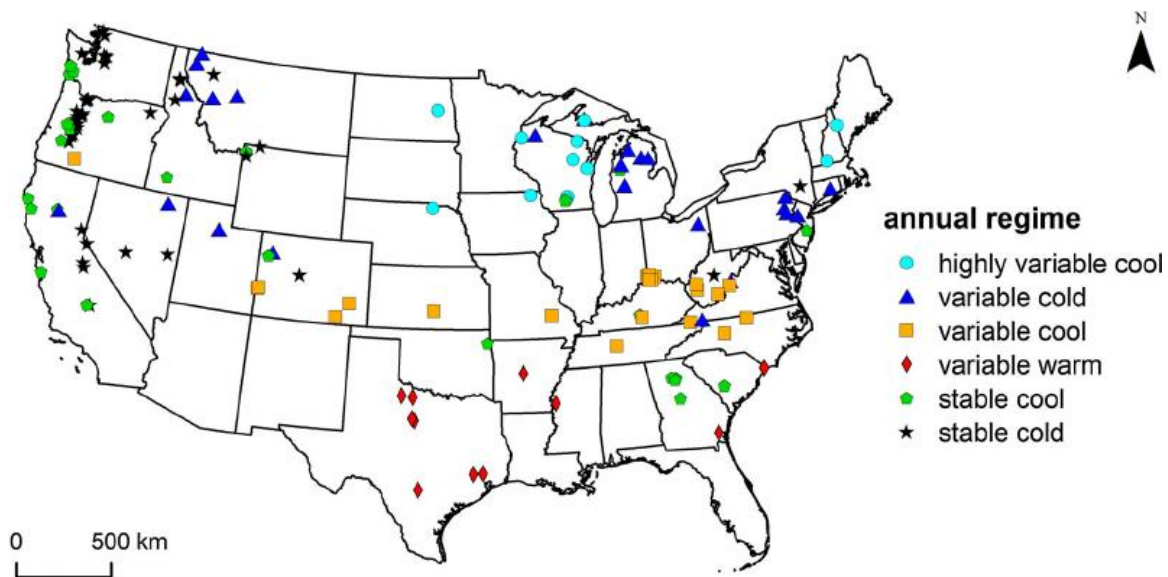


Figure 2.1. Classification of streams and rivers at a national scale based on annual stream regimes (from Maheu et. al. 2016).

In addition to nation-wide efforts, stream classification approaches have been implemented at a state-wide scale (Kendy et. al. 2012). Michigan is one of the states where stream classification is well-studied topic, going back to the late 1990s. Seelbach et al. (1997) developed and used a landscape-based classification model to classify river valley segments in lower Michigan based on their ecological features, such as catchment size, water temperature, hydrology and fish assemblages. Several years later, Brenden et. al. (2008) further refined the initial classification system. In addition, considering stream temperatures as one of the main

factors for fish habitat preference, Wehrly et al. (2003) classified streams into three classes (cold, cool and warm) by using July mean temperature (JMT) data from 171 sites in Michigan. By referencing the classification approaches in previous studies (Seelbach et. al (1997); Zorn et. al (2002); Wang et. al. (2003); Wehrly et. al. (2003); Baker (2006); Seelbach et. al. (2006); Brenden et. al. (2008)), Zorn et al. (2008) developed a model to evaluate the effect of flow reduction on stream fish assemblages in Michigan. In this study, stream thermal classes were developed based on July mean temperatures: (cold (C) = $JMT \leq 17.5 \text{ }^{\circ}\text{C}$ ($63.5 \text{ }^{\circ}\text{F}$), cold-transitional (CT) = $17.5 \text{ }^{\circ}\text{C}$ ($63.5 \text{ }^{\circ}\text{F}$) < $JMT \leq 19.5 \text{ }^{\circ}\text{C}$ ($67 \text{ }^{\circ}\text{F}$), warm-transitional (WT) = $19.5 \text{ }^{\circ}\text{C}$ ($67 \text{ }^{\circ}\text{F}$) < $JMT \leq 21.0 \text{ }^{\circ}\text{C}$ ($70 \text{ }^{\circ}\text{F}$), warm (W) = $JMT > 21.0 \text{ }^{\circ}\text{C}$ ($70 \text{ }^{\circ}\text{F}$)) and were applied to make predictions by using the Water Withdrawal Assessment Tool (WWAT). These categories are the current basis for classification under current Michigan legislation.

In previous research, Andrews (2019) developed a suite of regression models to predict thermodynamics in streams. However, those regression models have not been evaluated within a stream classification framework. In this study, I adopted the best performing linear regression model among Andrews (2019) models and applied data pooling to determine if these models could be generalized across thermal stream classes. My study is important in many aspects since my findings can lead to new perspectives in stream classification – stream temperature modeling and can be implemented in state-wide stream management processes.

Data Pooling, Model Generalization and Stream Management Practices

Andrews' (2019) linear regression models were designed to be used to provide stream-specific temperature gradient (ΔT) predictions. In other words, the model dynamics changed from stream to stream since data from individual streams were used in parameter estimation. Hypothetically, pooling the data from streams within the same thermal class could result in more

generalized models. If these generalized class-based models (e.g., Cold stream class model) were applied to predict temperature gradient for an individual stream, the predictions would reflect the temperature gradient predictions based on the overall stream characteristics of that stream class (e.g., Cold stream class). If class-based temperature gradient predictions are realistic, these generalized models would be useful for numerous management purposes.

If generalized models work well, the most practical use of those models would be to reduce the need for extensive data from individual streams. As such, the class-based models based on a set of representative stream data could be used to make temperature gradient predictions with limited data for other streams. For example, reliable temperature gradient predictions can be made based on common behavioral characteristics of the streams within that stream class (Tadaki et. al. 2014). Also, real-time predictions of response variables can be achieved without collecting individual stream data beforehand, but by retrieving instantaneous data on predictor variables from various data sources (e.g., GIS and weather station data). Future predictions of response variables can also be made by applying hypothetical data for different scenarios. For example, future fish population distributions based on stream temperature changes can be predicted by using hypothetical data that reflect different climate change scenarios (e.g., Lyons et. al. 2010).

In this chapter, I refer to “model performance” as the indicator of the fit between model predictions and observed temperature gradient values, as well as the consistency of trends of observed and predicted thermal gradient across time. Naturally, potential uses of generalized regression models depend on their model performances, particularly on their precision and potential bias. Therefore, I evaluated overall performance of class-specific models, as well as a “global” model that was based on all the streams in my study to evaluate the performance of the

most general model. Additionally, evaluating model performances across stream classes gives valuable information on which stream classes can be most accurately modeled. Furthermore, I evaluated the performance of generalized models when July-restricted data were used to develop those models since time period selection was an important factor affecting model performances (see Chapter 1). All these considerations shaped the main goals of my study, which are:

- 1) To apply data pooling (with June-October 2016 data) based on stream thermal classes (C, CT, WT, W) to obtain generalized models;
- 2) To investigate the changes of model dynamics across data pooling;
- 3) To evaluate overall model performances of stream-specific and generalized models and evaluate their success across stream classes;
- 4) To evaluate overall model performances of stream-specific and generalized models by applying July-restricted data.

METHODS

Study Site and Data Collection

The same study streams and data collection methods in Chapter 1 were used for this chapter. Moreover, the same refined and revised datasets of streams and regression models that were defined in Chapter 1 were used.

Stream Classification and Model Performance

Streams were classified based on July Mean Temperatures (JMT) predictions as described in Zorn et. al. (2008). I decided to use daily data granularity because using daily data granularity resulted in generally high model predictive power for Model 10 (see Chapter 1), and because daily data granularity was used in the WWAT (Zorn et. al. 2008). Although overall model predictive power was highest with weekly data granularity, I did not use it for my

applications in this chapter to avoid overfitting problem especially with July-restricted datasets (see Chapter 1). June-October (starting from June to October 2016) and July-restricted (July 2016) time periods were used to evaluate the effect of stream classification on model performances for each period. To evaluate model performances, model prediction reliability and model prediction powers were observed. Model prediction reliability were evaluated based on bias (B) for individual streams and mean bias (\bar{B}) values for the class as a whole. Pearson correlation coefficient (r) between observed and predicted temperature gradient was used to evaluate the consistency between observed and predicted values.

Obtaining and Evaluating Models

I chose Model 10 as the “base” model to obtain Stream-Specific (SSM), Class-Based (CBM) and Global-Based GBM) models. SSMs were obtained by applying the base model to data from individual streams, as it was done in Chapter 1. CBMs for each stream class were obtained by pooling the data of streams within the same stream class and running the base regression model for the pooled data. Hypothetically, the dynamics (i.e., the intercepts and parameter estimates) in the base model for each stream class would be expected to vary since each class had different environmental characteristics and data, therefore, the outputs from CBMs for each stream class were expected to be different. When compared to SSMs, CBMs were more generalized models since the dynamics of the base model were determined by the sets of streams that were in the same stream classes. The datasets of all streams were limited to the span between day of the year 177 to 270 to ensure all streams were equally represented.

The Global-Based Model was obtained by pooling the data from all streams and applying the base regression model to the pooled data. Like SSMs and CBMs, the GBM was expected to have unique values of intercept and parameter estimates. Temperature gradient predictions were

obtained using the GBM for each stream. Since the data of all streams were pooled, GBM was the most generalized model. As the CBMs and the GBM are more broadly applicable than the SSMs, I will refer to these models as “generalized models”.

After obtaining the temperature gradient predictions from each model by using June-October data, I obtained the Pearson Correlation (r) between observed temperature gradient and predicted temperature gradient for each stream. Moreover, to find the amount of bias between observed and predicted temperature gradient for each model, I obtained the mean observed and mean predicted temperature gradient for each stream and used the equation:

Eqn. 16.
$$B = \overline{\Delta T_{Pred}} - \overline{\Delta T_{Obs}}$$

where B stands for bias, $\overline{\Delta T_{Pred}}$ (°C) stands for mean predicted temperature gradient, and $\overline{\Delta T_{Obs}}$ (°C) stands for mean observed temperature gradient. The overall bias between observed and predicted temperature gradient values would be expected to be zero as the sum of residuals (which are the difference between observed and predicted values) is zero in linear regression. The bias, B , calculated here indicates the magnitude of deviation that occurs for subsets of data, which is not guaranteed to be zero for linear regression with subgroups. I calculated the mean absolute value of the stream-specific bias for each stream class as:

Eqn. 17.
$$\bar{B} = \frac{\sum |\Delta T_{Pred} - \Delta T_{Obs}|}{n},$$

where \bar{B} is the mean absolute value of bias and n is the number of the streams in the thermal class. The absolute difference between mean predicted temperature gradient and mean observed head flux were found to observe the magnitude of deviation between these values.

In addition to evaluating model performances across stream classes, I also explored how model performance varied with mean observed downstream temperature and mean observed temperature gradient within each stream. The effect of downstream temperature was explored

because this provided a direct measure of the stream's thermal conditions in the year of study, in contrast to the stream thermal classification, which is based on predictions of the 30-year mean July mean temperature for a stream from Brenden et al. (2008). I also explored model performance as a function of a stream's mean observed temperature gradient to determine if generalized models performed equally across the range of temperature gradients observed.

RESULTS

Pooling Data Changed Model Dynamics and Model Outcomes

Stream-Specific models (SSMs) were obtained by applying Model 10 on the individual dataset of each stream. Substantial variation between the values of intercepts and parameter estimates (β) of the SSM for each stream was observed across stream-specific models (Table 2.1). As an example, the value of intercept in SSM for Black River was 0.004, yet the same value in the model for Hasler Creek was 2.638.

Table 2.1. Intercepts and parameter estimates from Stream-Specific models (SSMs) applied to each stream for June – October hydrological data.

Streams	<i>Intercept</i>	<i>T_a-T_w</i>	<i>Q_{up}</i>	<i>Q_{down} - Q_{up}</i>	<i>S</i>	<i>α</i>	<i>ΔT_{up}</i>	<i>ΔT_{base}</i>	<i>ΔT_{over}</i>
Black River	0.004	-0.380	-0.872	0.002	0.004	-0.037	-0.060	-0.013	-0.026
Cedar River	0.982	-0.665	2.810	-0.154	0.005	-0.011	0.097	-0.029	0.015
Cedar Creek	-3.800	-0.044	0.102	-0.041	-0.001	0.009	0.004	-0.012	0.001
Morgan C.	-0.258	7.516	-0.225	0.155	-0.012	-0.015	-0.231	0.186	-0.027
Pokagon C.	-2.326	-1.105	-1.405	0.072	-0.014	0.027	0.043	-0.105	0.048
Butterfield C.	1.690	-0.323	-0.343	0.309	0.053	-0.016	0.020	-0.006	-0.003
Carp River	0.500	28.338	-9.301	-0.038	0.018	-5.671	0.133	0.056	-0.004
Pigeon River	-3.953	1.841	5.079	0.009	-0.020	0.018	-0.038	0.015	0.013
Spring Creek	1.284	2.412	1.146	0.147	-0.020	0.002	0.060	-0.094	0.038
Escanaba R.	-5.148	1.626	-1.804	0.352	-0.066	0.078	-0.130	0.123	-0.033
Nottawa C.	-4.156	-0.436	-0.050	0.245	-0.060	-0.017	0.064	-0.120	0.068
Tobacco R.	1.092	-1.513	-2.864	0.584	-0.077	-0.018	-0.472	0.370	-0.340
Hasler C.	2.638	0.121	-0.061	-0.037	-0.022	-0.041	0.041	0.004	0.008
Prairie River	-5.764	-18.751	1.618	0.082	-0.038	0.019	0.244	-0.256	0.117
Squaw Creek	0.246	-1.359	-2.618	0.265	0.008	-0.019	0.140	-0.108	0.130
Swan Creek	-2.335	-2.630	0.366	0.000	0.001	-0.023	-0.088	0.008	-0.060
Average	-1.207	0.916	-0.526	0.122	-0.015	-0.357	-0.011	0.001	0.000

In addition, the intercept and parameter estimate values of CBMs and GMB were unique to each class-specific model and the global model (Table 2.2). To illustrate, the intercept value in cold CBM was 0.479, and the value across stream classes and GBM varied. Also, parameter estimates of the same environmental variable (e.g., Q_{up}) changed sign across class-specific models (Table 2.2). For example, Q_{up} had a positive sign in cold CBM (0.236), whereas its value was negative in warm-transitional CBM (-0.400). These variations between parameter estimates indicated potential conflicts in interpretations of how environmental factors influence model predictions as well as the amount of variance explained.

Table 2.2. Parameter estimates of Class-Based and Global Based models. June-October 2016 data were used.

<i>Stream Class</i>	<i>Intercept</i>	T_a-T_w	Q_{up}	$Q_{down}-Q_{up}$	S	α	ΔT_{up}	ΔT_{base}	ΔT_{over}
<i>C</i>	0.479	0.030	0.236	-0.122	0.005	0.002	-0.010	-0.038	0.008
<i>CT</i>	-0.042	-0.004	-0.015	-0.39	0.067	-0.019	0.069	-0.036	0.056
<i>WT</i>	-2.622	0.032	-0.400	-0.875	0.230	-0.023	0.002	-0.020	-0.008
<i>W</i>	0.606	0.027	-2.713	-2.527	0.052	0.050	0.111	-0.051	0.105
<i>Global</i>	-2.096	0.072	0.101	-0.166	0.178	-0.041	0.012	-0.015	-0.002

Naturally, changes in model parameter estimates with data pooling resulted in changes of model predictions. Observed and predicted temperature gradient values showed that the congruence between observed and predicted temperature gradient varied among streams within a class (Figure 2.10 to 2.13). Cedar Creek, Pigeon River, Escanaba River and Prairie River were selected as example streams from each stream class as they had the overall highest mean r values of models (0.6482, 0.5436, 0.5961, 0.5729 respectively) among all streams. The fit of SSMs was generally higher than the fit for CBMs and GBMs. For example, the predicted temperature

gradient from the SSM of Cedar Creek displayed a more similar trend across time to the observed temperature gradient compared to predictions from CBM and GBM (Figure 2.10).

Generalized models generally showed lower overall accuracy of temperature gradient predictions (Table 2.3) compared to SSMs. The mean bias (\bar{B}) values of SSMs were generally the lowest for all stream classes, and mean biases of GBMs were the highest for all classes (Figure 2.2). Moreover, overall mean bias values of GBMs were higher than mean bias values of CBMs. For example, mean bias value of GBM for Warm stream class (0.794) was almost five times greater than the same value of CBM (0.160) for the same stream class.

Table 2.3. Bias values (B) and their average (\bar{B}) of Stream-Specific models (SSM), Class-Based models (CBM) and Global-Based model (GBM) predictions. June- October 2016 data were used.

Stream Class	Stream	Bias (B) (SSM)	Bias (B) (CBM)	Bias (B) (GBM)	Mean Bias (\bar{B}) SSM	Mean Bias (\bar{B}) CBM	Mean Bias (\bar{B}) GBM
C	Black River	0.000	-0.118	-0.066	0.008	0.073	0.181
	Cedar River	-0.008	0.083	0.117			
	Cedar Creek	0.004	-0.034	0.201			
	Morgan C.	-0.002	-0.031	-0.179			
	Pokagon C.	-0.028	0.100	0.342			
CT	Butterfield C.	-0.017	0.507	-0.887	0.021	0.160	0.388
	Carp River	-0.006	0.044	-0.224			
	Pigeon River	0.004	0.074	-0.406			
	Spring Creek	-0.057	0.202	0.037			
WT	Escanaba R.	0.051	-0.037	0.296	0.028	0.025	0.275
	Nottawa C.	0.017	0.007	-0.168			
	Tobacco R.	-0.015	0.030	0.361			
W	Hasler Creek	-0.021	-0.225	-1.223	0.037	0.160	0.794
	Prairie River	0.027	-0.091	-0.489			
	Squaw Creek	-0.031	0.213	1.170			
	Swan Creek	0.070	0.110	0.293			

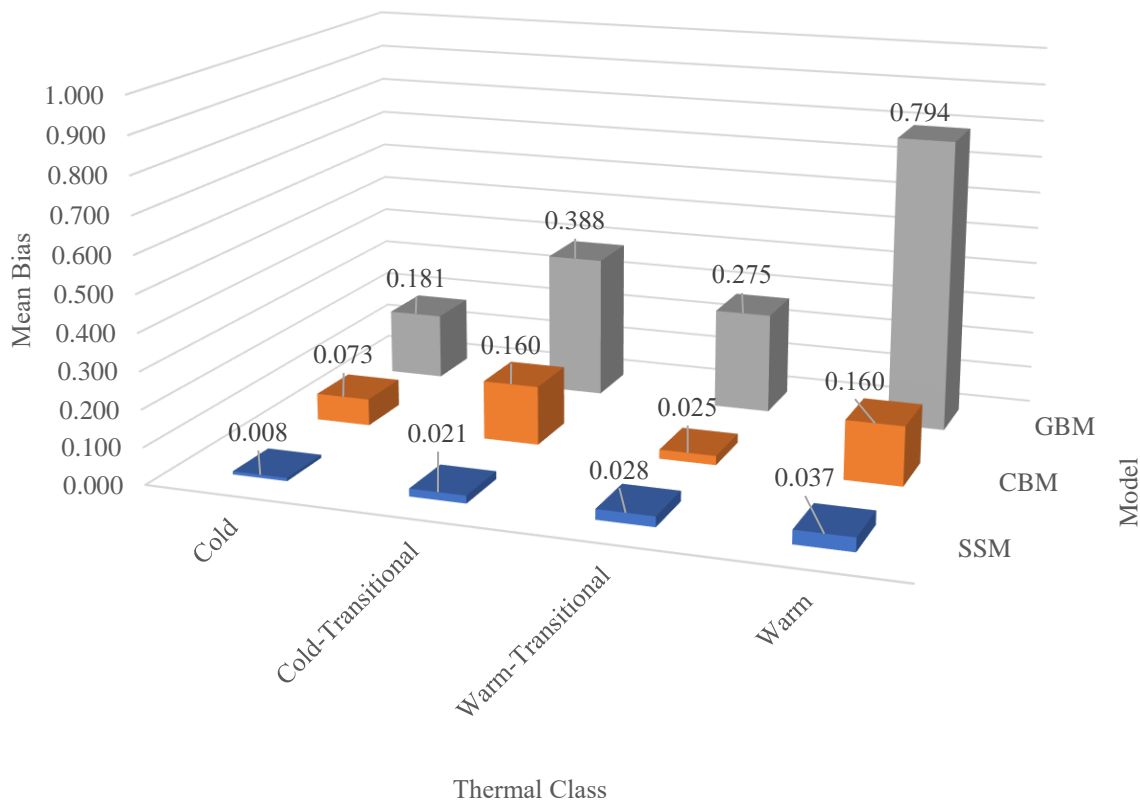


Figure 2.2. The absolute value of biases averaged for each stream class. The higher the mean absolute bias, the higher the overall mean temperature gradient prediction deviates from the overall mean observed temperature gradient.

Based on mean r values, SSMs had distinctively higher model predictive power compared to CBMs and GBMs (Table 2.4; Figure 2.14). Moreover, CBMs had higher model predictive power compared to GBMs, supporting the conclusion that model prediction reliability decreases as generalization of models increases (i.e., SSMs to GBMs).

Table 2.4. Mean Pearson’s correlation coefficient values of SSM, CBM and GBM across stream classes. June-October 2016 data were used.

Stream Class	SSM	CBM	GBM
Cold	0.691	0.212	0.333
Cold-Transitional	0.618	0.106	0.059
Warm-Transitional	0.699	0.584	0.163
Warm	0.796	0.472	0.191

Cold stream class generally had the lowest mean biases for all models, whereas Warm stream classes generally had the highest mean biases (Table 2.3; Figure 2.2). However, streams in Cold-Transitional stream class showed the lowest mean r values in all models (Figure 2.14). In contrast, warmer streams (i.e., Warm and Warm-Transitional classes) posed higher mean r values in majority of models.

Stream classifications used to this point were based on (Zorn et. al. 2008), which uses model-based predictions for each stream's thermal classification. As such, there is a potential mismatch between predicted stream class membership and the observed mean stream temperatures for my study streams between June- October in a single year: 2016. These differences were apparent for several streams (Table 2.7), which lead me to evaluate model performances as a function of mean downstream temperature. Model prediction power values showed no clear relation to mean downstream temperatures (Figure 2.3). In other words, model prediction power did not substantially change with increasing or decreasing stream temperatures. Likewise, bias (B) did not show a trend across mean downstream temperatures (Figure 2.4).

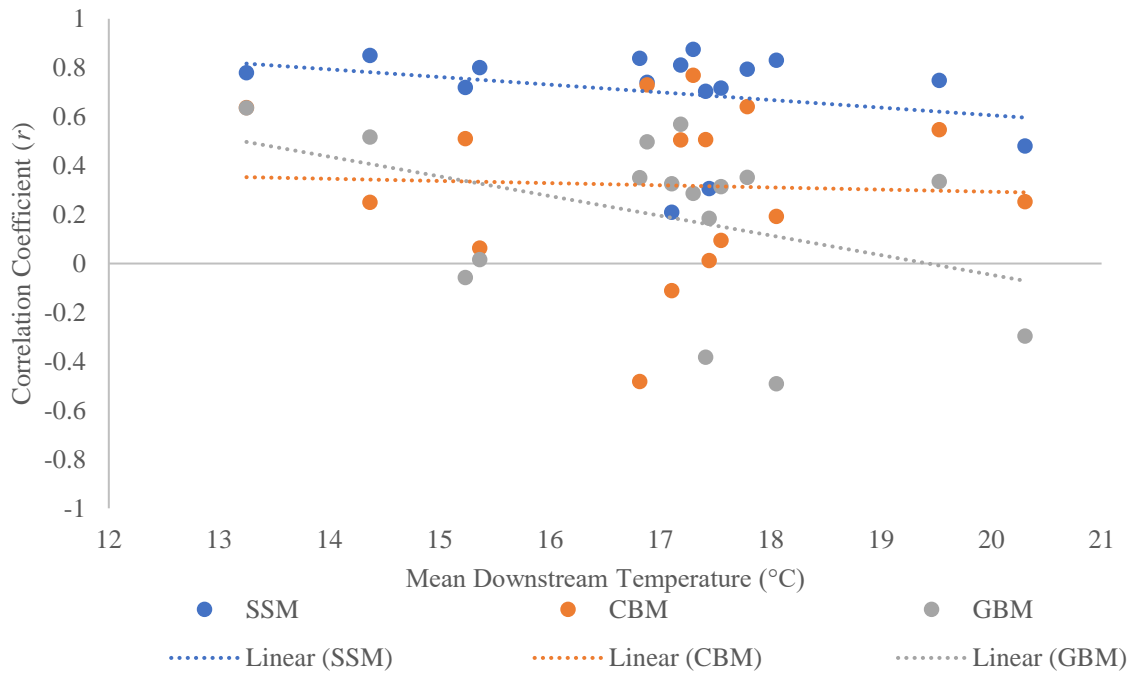


Figure 2.3. Pearson's correlation coefficient (r) values of SSM, CBM, GBM across mean downstream temperatures from June-October 2016.

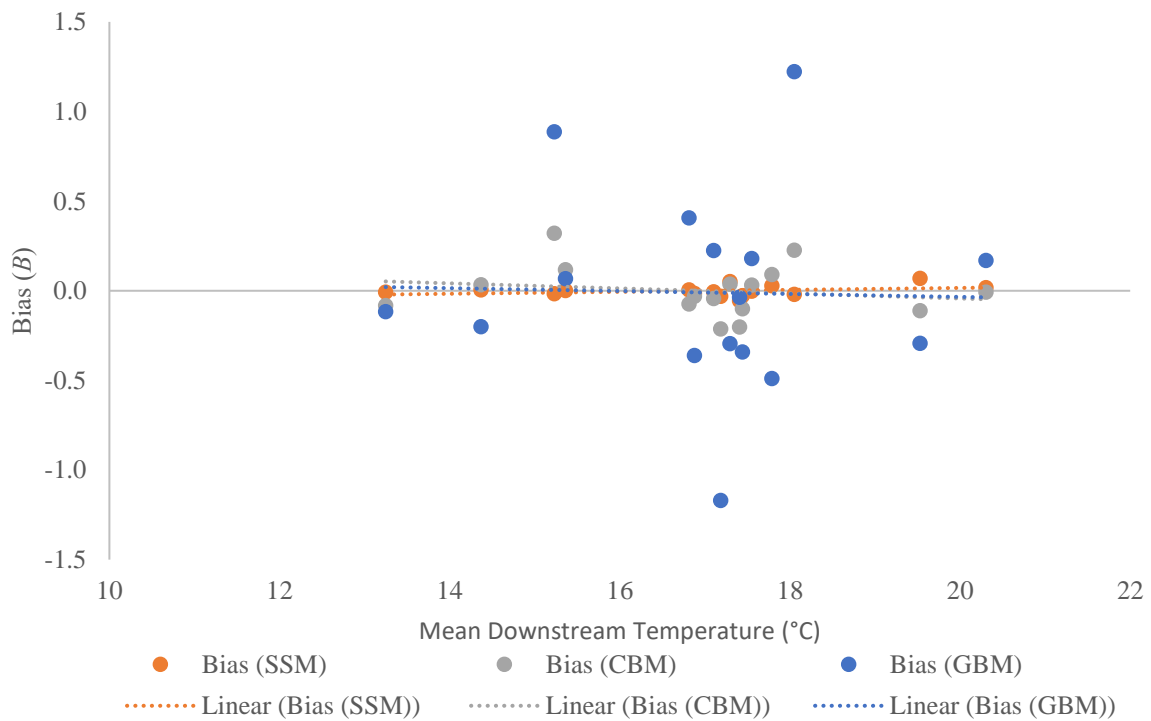


Figure 2.4. Bias (B) versus mean downstream temperature. June – October 2016 data were used.

I also evaluated the relationship of model performances to mean observed temperature gradient to determine if streams that show more or less warming are modeled more accurately. Correlation (r) between generalized model predictions and observed temperature gradient increased with mean observed temperature gradient (Figure 2.5). Model predictive power of GBMs especially showed a considerable increase (from negative values of r to values of 0.6), indicating that generalized models predicted the trends of temperature gradient more accurately for warming stream reaches. On the other hand, the highest bias values for generalized models were observed at the high and low ends of the range of temperature gradient values observed (Figure 2.6). In other words, high temperature changes between upstream and downstream resulted in greater inaccuracies in model predictions.

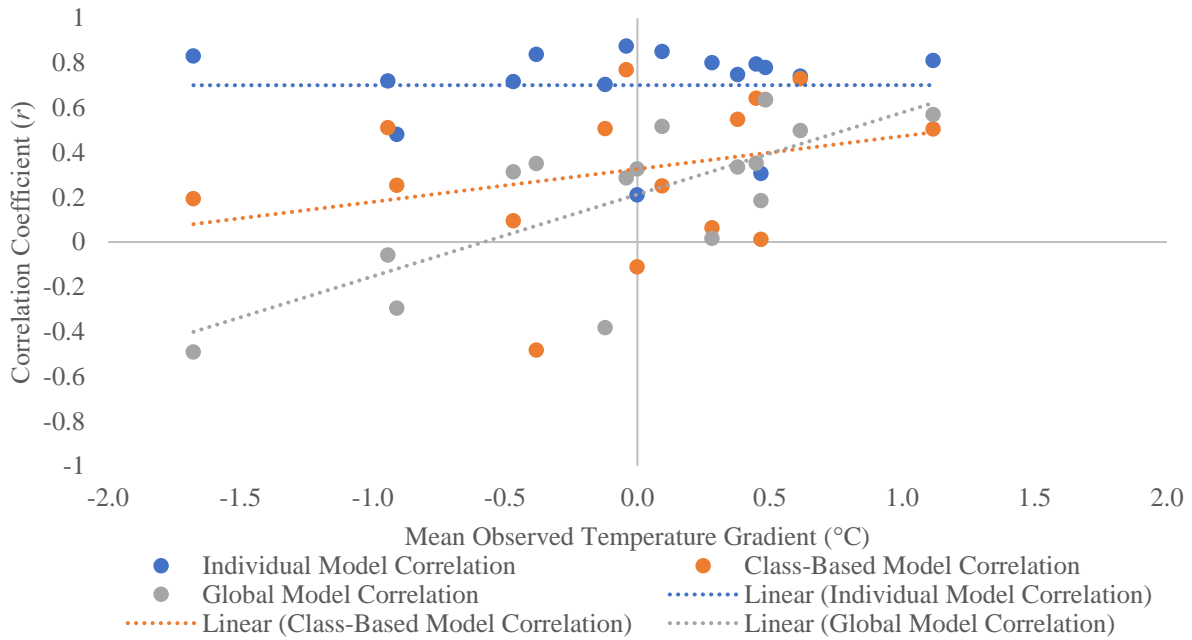


Figure 2.5. Pearson correlation coefficient (r) values of SSM, CBM, GBM across mean observed temperature gradient from June-October 2016.

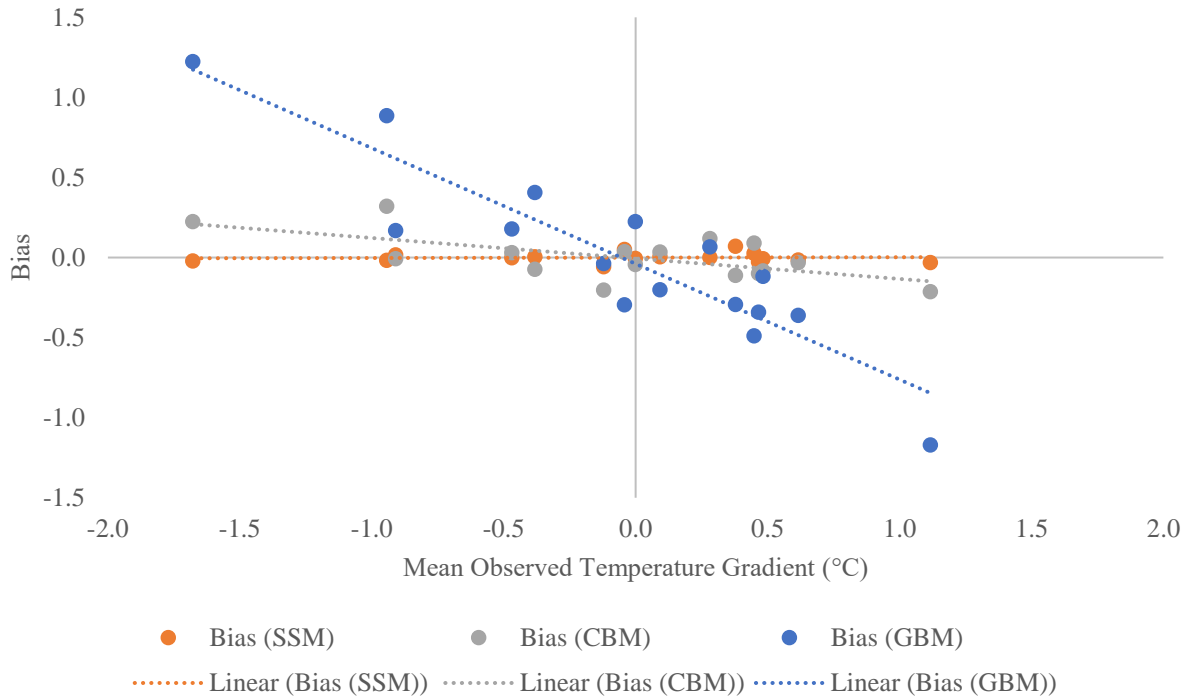


Figure 2.6. Bias (B) versus mean observed temperature gradient. June – October 2016 data were used.

Classifying Streams Reduced Overall Model Performance with July-Restricted Data

With July-restricted data, the predictive power of SSMs was still higher compared to CBMs and GBMs (Table 2.5; Figure 2.15). Although the model predictive power of SSM was substantially higher for all stream classes, neither CBM nor GBM were found to have distinctly higher model predictive power for any particular stream classes when July-restricted data were used. Thus, using July-restricted data did not increase model predictive power over data from the full summer season for either of the generalized models.

Table 2.5. Mean Pearson’s correlation coefficient (r) values of SSM, CBM and GBM across stream classes. July 2016 data were used.

Stream Class	SSM	CBM	GBM
Cold	0.796	0.014	0.165
Cold-Transitional	0.831	0.252	0.330
Warm-Transitional	0.734	0.340	-0.031
Warm	0.861	0.324	0.226

Mean model prediction power of Class-Based models with June-October and July-restricted data were compared to understand whether using July restricted data would make Class-Based models work better or not. Surprisingly, CBMs performed better when June-October data were used in most cases, except for the Cold-Transitional stream class (Figure 2.7). Moreover, results showed that there was no substantial change in model predictive power of models across July mean downstream temperature (Figure 2.8). Nevertheless, the model prediction power of CBMs and GBMs increased with higher mean temperature gradient values (Figure 2.9).

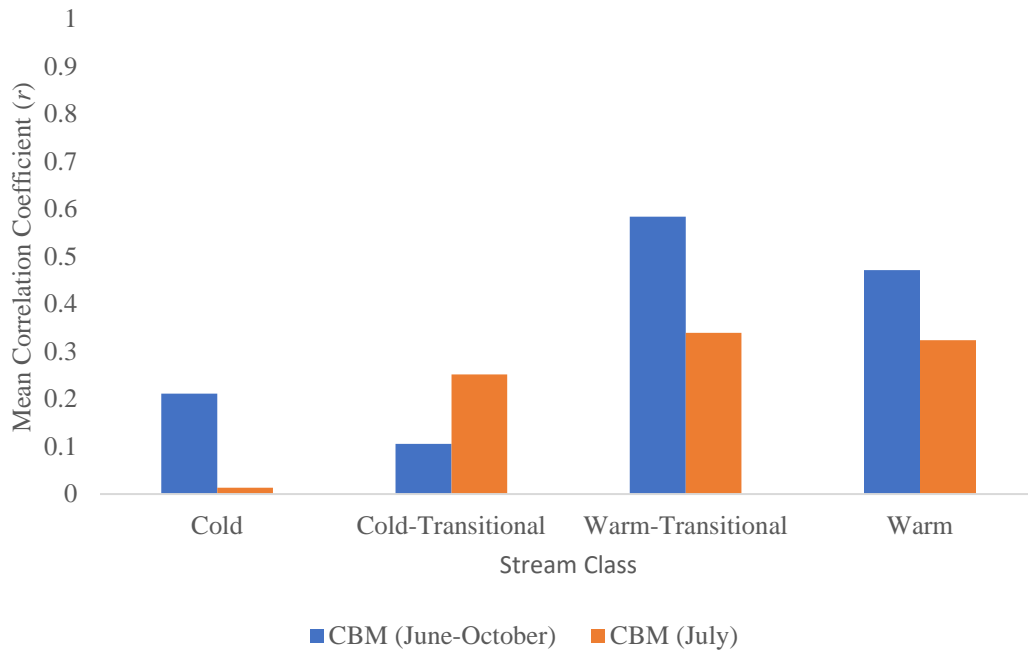


Figure 2.7. Mean Pearson's correlation coefficient (r) of Class-Based Models with June- October 2016 data and July 2016 data.

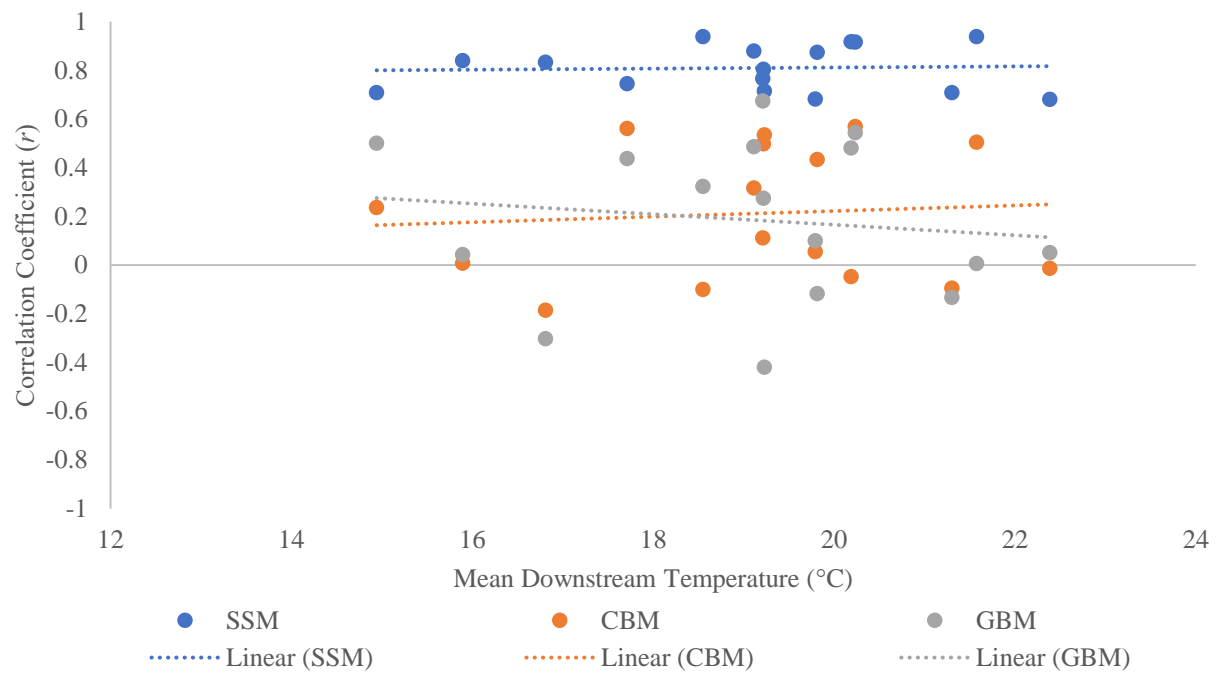


Figure 2.8. Pearson correlation coefficient (r) values of SSM, CBM, GBM across mean downstream temperatures from July 2016.

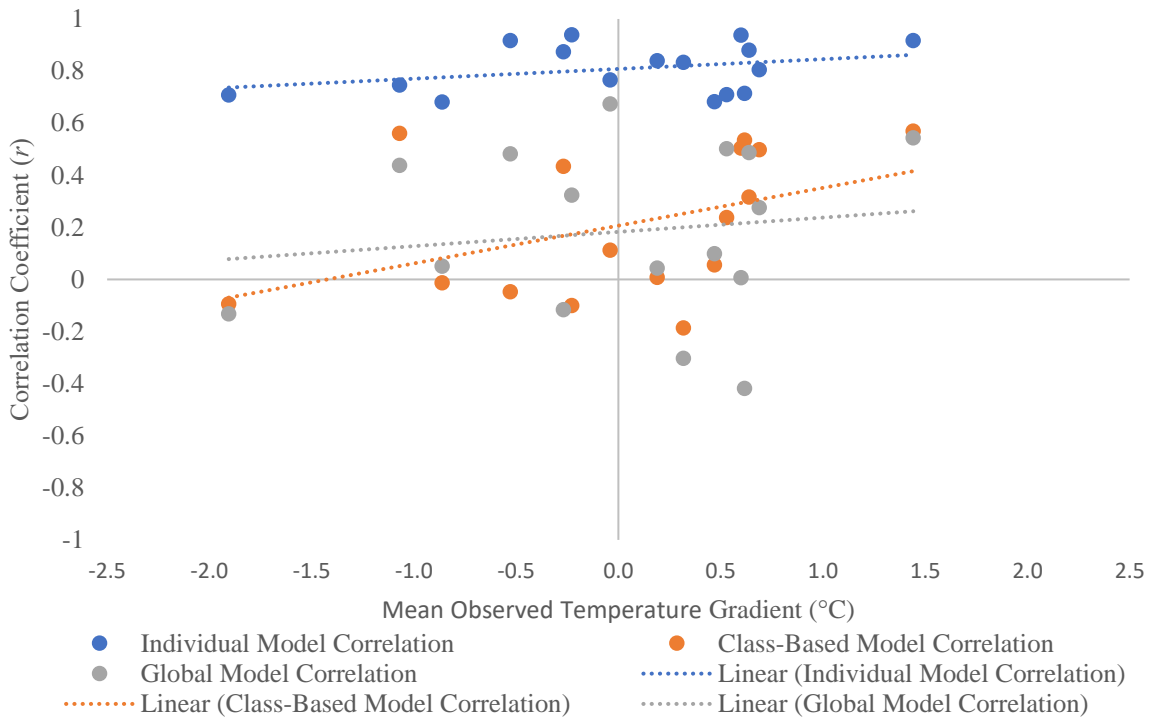


Figure 2.9. Pearson correlation coefficient (r) values of SSM, CBM. GBM across mean observed temperature gradient from July 2016.

DISCUSSION

Although the Michigan Department of Natural Resources has applied stream classification with physical models for many years (Zorn et. al. 2008), the effects of applying data pooling on linear regression models has not been tested. Therefore, applying data pooling to the regression models that were designed by Andrews (2019) provided insight into my four main research questions: “What is the effect of data pooling on model dynamics?” “Does stream classification improve the model performance?”, “Do models work better for warm or cold streams?”, and “Does using July-restricted data change model performance? The answers to these questions are intended to help guide researchers and managers select proper models for their particular needs and to determine if adequate better model predictions can be made without collecting extensive and expensive stream data (Carlson et. al. 2017).

What is the Effect of Data Pooling on Model Dynamics?

The results showed that data pooling based on stream classes resulted in different parameter estimates of Model 10, indicating substantial changes in model dynamics or highly variable parameter estimates as data were broken into subsets (Table 2.1 & Table 2.2). Therefore, data pooling resulted in significant variation in temperature gradient predictions, leading to substantial differences in model performance (i.e., biases and model predictive powers) between models. Basically, using generalized regression models (i.e., CBMs and GBMs) reduces the explanatory power of models especially for the streams that have unique environmental conditions (Carlson et. al. 2017). Therefore, I hypothesized that generalizing the regression model by using the global stream data (across Michigan) may increase the magnitude of biases in model predictions and reduce overall model predictive powers. This hypothesis was tested by evaluating performances of stream-specific and generalized models, and results are discussed in the following section.

Does Stream Classification Improve Model Performance?

I used two approaches to evaluate model performances: the mean bias (\bar{B}) between observed and predicted temperature gradient (*Eqn. 16*) and Pearson's correlation coefficient (r) as an indicator of model prediction power. One of the key findings was that generalized models had higher bias values when compared to SSMs (Figure 2.2). Thus, Class-Based and Global-Based models had lower model prediction reliability. Higher \bar{B} values of CBMs and GBMs supported my hypothesis that data pooling would result in less accurate predictions, especially for the streams that had distinct environmental characteristics. It is important to note that although CBMs did not produce predictions that were as reliable as SSM predictions, they were generally more reliable than GBM predictions.

Model predictive power based on correlation values (r) was another indicator of model performance. Overall model predictive power of SSMs was distinctly higher for the majority of streams and stream classes (Table 2.4). This finding matched with bias findings as bias values (\bar{B}) of SSMs were the lowest for most of streams. Therefore, using SSMs would give better temperature gradient predictions and better estimation of temperature gradient trends. In general, predictions from CBMs had higher correlation with observed data than GBMs. This may be caused by higher similarity between the environmental conditions of the streams that are grouped in the same class. Therefore, the temperature gradient trends were predicted better by CBMs. However, the low sample size of streams within each stream class might be a constraint to developing reliable CBMs, and consequently, may be partly a cause of lower model predictive powers for these models. Thus, if the number of streams that are used to obtain CBMs is increased, model predictive power may be improved.

Both mean bias (\bar{B}) and correlation (r) revealed limitations of implementing CBMs and GBMs, but how can these limitations be considered from the perspective of stream management? The bias results implied that although CBMs had less model prediction reliability when compared to SSMs, they still have the potential to be used. The mean bias of CBMs ranged from 0.025 to 0.160 (Table 2.3) across stream classes. In other words, the difference between average predicted and observed temperature gradient was less than 0.2 °C across stream classes. From an ecological perspective, such difference may be negligible since some salmonid species (e.g., brown trout: *Salmo trutta*) have the ability to acclimate to a temperature of 27 to 30 °C within 24 hours (Brett 1956; Sullivan et. al. 2000). In addition, daily water temperature changes up to 13.5 °C did not substantially affect the survival and growth of salmonids, unless lethal temperature levels were reached (Thomas et. al. 1986). Considering these tolerance ranges, the mean bias

values of CBMs may be acceptable, depending on the focus of the study (e.g., the characteristics of fish species) and as well as the availability of physical and financial resources for stream data collection. For example, if the study's goal focuses on temperature gradient predictions for multiple streams that are distributed within a small spatial range, adopting stream-specific models may be most appropriate as the range of temperature gradient values may be quite narrow. On the other hand, generalized models may be more useful in studies that require modeling for the streams within very large spatial range (e.g., state-wide, e.g., Steward et. al. 2015) that have a wider range of conditions and where the impact of bias would be less. At this point, the efficiency of using generalized models must be evaluated by researchers and decision-makers based on their purpose, the range of bias that is acceptable, and their resource availability needed for data collection.

Do Models Work Better for Warm or Cold Streams?

Cold stream class had the lowest overall bias, indicating model prediction reliability was relatively higher for the Cold stream class. Moreover, the Warm stream class had the highest overall bias thereby yielding models with the lowest model prediction reliability. Interestingly, the correlation between observed and predicted temperature gradient of CBMs was highest for warmer stream classes (Warm and Warm-Transitional). This apparent conflict highlights the difference between predictions that correlate to temporal trends in temperature gradient, and predictions that are offset from the observed data, leading to bias.

A potential limitation for making conclusions based on model performance across stream thermal classes is that the observed stream temperatures for the time period I used (June-October 2016) did not always match a priori stream classes. For example, Morgan Creek should have been included in Warm-Transitional stream class based on its observed mean July downstream

temperature (17.550; Table 2.7). Likewise, Butterfield, Spring, Hasler, Squaw, Swan creeks and Carp, Pigeon, Escanaba, Tobacco, Prairie rivers would fall into different stream classes based on their mean July downstream temperatures. Therefore, to cross-validate my findings on model performance and stream class and make more reliable conclusions on model performance versus stream temperatures, I tested model performances across mean downstream temperatures as another criterion. The distribution of r values of models showed no clear relation to mean downstream temperature (Figure 2.3). In addition, bias also did not show a substantial increase or decrease with increasing mean downstream temperature (Figure 2.4). Therefore, it appears that these models are equally applicable to cooler or warmer streams. This conclusion should be tempered, however, by the narrow range in mean temperature (15 °C-18 °C) among my study streams. Given the low diversity of thermal characteristics of streams studied, it is unknown whether model generalization approaches would work better across a broader range of thermal characteristics.

Response of bias to temperature gradient, which was another thermal criterion, varied between models (Figure 2.6). Generalized models (CBMs and GBMs) had higher biases when compared to SSMs as for streams that showed the highest and lowest mean temperature gradient values. A potential reason was that generalized models can result in biases, especially when a stream has unique hydrological characteristics, such as having complex groundwater-surface water interactions. For example, using a generalized model for a stream section with a high degree of groundwater loss (e.g., positive temperature gradient) or gain (e.g., negative temperature gradient) may result in overestimation of temperature gradient for gaining streams and underestimation for losing streams. In addition to B values, I used Pearson's correlation coefficient (r) to observe the model predictive power across mean temperature gradient. The

results showed that model predictive power of generalized models substantially increased with mean observed temperature gradient, that is, generalized worked better for warming streams. (Figure 2.5). This result was important because it may be a sign for reduced performance based on the amount of groundwater input in the system. As mentioned before, cooling streams may be considered as groundwater gaining streams. As such, the predictive power for warming streams may be better because they lack complex groundwater-surface water interactions. Evaluating model performance across temperature gradient also indicated that other environmental processes (e.g., stream shading, discharge, groundwater) that lead to heat gain or loss of the streams may be more important considerations beyond the observed temperature at a point in the stream (Webb and Zhang 1997; Dugdale et. al. 2018).

Although I did not observe a clear relationship between model performance and stream class or mean downstream temperature, it appears that generalized models perform more poorly for streams with high temperature gradient. Considering that extreme temperature gradients tend to be observed in streams that are highly altered by human activity (e.g., surface or groundwater withdrawal; Xin and Kinouchi 2013) or observed in the streams that might have complex groundwater and surface water dynamics (Westhoff et. al. 2007), managers are recommended to use Stream-Specific models instead of generalized models to obtain reliable temperature gradient predictions. Cooling streams should be of particular concern since the cooling trend generally indicates a groundwater-driven stream, for which models had lower performances. Poor decisions on groundwater withdrawal based on poor model predictions could severely affect dynamics in groundwater-driven streams as well as its biota (Boulton et. al. 2010; Carlson et. al. 2019).

Does Using July-Restricted Data Change Model Performance?

Predictive powers of CBMs and GBMs with July-restricted data were lower when compared to model predictive powers with June-October data (Table 2.5; Figure 2.15). Evaluation of model predictive power of CBMs with July-restricted and June-October data validated this conclusion except prediction power increased for Cold-Transitional stream class (Figure 2.7). Using shorter time periods, such as July, may increase temporal and spatial variation of hydrological events (e.g., groundwater flow, precipitation, snowmelt) across streams. For example, average monthly precipitation is typically the highest in June and July in the Great Lakes basin (Norton et. al. 2019). High spatial variation of rainfall during July may cause larger variations between physical characteristics of streams, consequently reducing the performance of generalized models.

The response of model prediction power across July mean observed temperature gradient matched previous results, that is, as mean observed temperature gradient increased, the predictive power of generalized models increased (Figure 2.9). Thus, restricting data to the warmest part of the year, which may be ecologically the most relevant, does not appear to improve model fits, particularly for sections of streams that show longitudinal cooling streams and that potentially have complex groundwater-surface water dynamics. Using July-restricted data did not significantly change the relation of model predictive power to July mean downstream temperature (Figure 2.8), however, implying that these models work equally well across observed mean downstream temperatures.

Based on these results, my main conclusion was that using shorter time period made generalized models even more disadvantageous than SSMs. Because the ecological relevance of the time period selected for the purpose of the study should come first (as explained in Chapter

1) and other time period options are not applicable in most cases, improving the class-based models appears to be the most effective way to reduce the costs of data sampling and making better predictions. The ways to improve class-based models for July-restricted time period is the same as for the full June-October period: increasing the number of streams used to develop the model and using more representative streams. Certainly, the optimum number of streams varies depending on various cases, however the number of cooling (i.e., groundwater-driven) streams should be carefully chosen to obtain generalized model due to the high complexity and low predictability in these streams. In addition, Model 10, which was the base for CBMs, can potentially be improved by adding new parameters or modifying the existing parameters so that the model can deal with complex groundwater- surface water dynamics and can be less sensitive to extreme temperature gradient values, as well as it can deal with variations between streams within the shorter time period.

CONCLUSIONS AND IMPLICATIONS

- 1) Stream classification is a useful approach to group streams based on their characteristics for many purposes, but especially important to decrease the need for extensive data collection. Data pooling is an effective practice to create class-specific and global models. Class-specific and global regression models had unique model dynamics, therefore, they resulted in different outcomes and showed different performances. Generalized models have the potential to make accurate predictions on response variables without the need of data from streams, as well as to predict future effects of an environmental change (e.g., groundwater withdrawal) on ecological characteristics of streams and stream classes.

- 2) Predictions from the Global-based model showed the highest degree of bias and were not highly correlated to temporal trends in temperature gradient in individual stream.

Thermal class-based models performed better than the Global-based model, but had poorer performance compared to Stream-Specific models. Even though the streams were classified in the same thermal class, some showed distinct physical characteristics, thus class-specific models did not work well for those streams. Another reason for lower performance of class-specific models was the low number of representative streams that were used to create these models. Using larger number of representative streams to create these models might increase model performance.
- 3) My study did not reveal any relationship between model performance across stream thermal classes or mean downstream temperature. The performance of generalized models increased as temperature gradient increased, however, implying better predictive capacity in streams with less groundwater contribution. Therefore, I suggest that modifying the base model and data inputs to better represent groundwater-surface water interactions would be a starting point to develop better generalized models that can explain the influence of groundwater on thermal dynamics of streams and stream classes.
- 4) Restricting the time period to July decreased the overall model performances of generalized models. The reason for this is unclear, but high temporal and spatial variations between environmental phenomena in July (e.g., precipitation) could have increased the distinct physical features of streams, resulting in lower model performances of generalized models for those streams. Therefore, using Stream-Specific models may be more useful in management practices. Nevertheless, although class-specific and global models had lower performances with July-restricted data, ecological relevance of time

period selection may be more important. Thus, improving generalized models would be more effective than using a time period that has lower ecological relevance with the purpose of the study.

APPENDIX

Table 2.6. Mean observed and predicted temperature gradient values, absolute bias values of Stream-Specific models (SSM), Class-Based models (CBM) and Global-Based model (GBM) predictions. June- October 2016 data were used.

Stream class	Stream	Mean Observed ΔT (°C)	SSM Mean Predicted ΔT (°C)	CBM Mean Predicted ΔT (°C)	GBM Mean Predicted ΔT (°C)
<i>C</i>	Black River	0.282	0.282	0.400	0.348
	Cedar River	0.484	0.476	0.401	0.366
	Cedar Creek	0.093	0.097	0.127	-0.108
	Morgan C.	-0.469	-0.471	-0.437	-0.290
	Pokagon C.	0.467	0.439	0.367	0.125
<i>CT</i>	Butterfield C.	-0.943	-0.960	-0.622	-0.056
	Carp River	0.000	-0.007	-0.045	0.224
	Pigeon River	-0.381	-0.377	-0.456	0.025
	Spring Creek	-0.121	-0.178	-0.323	-0.158
<i>WT</i>	Escanaba R.	-0.042	0.009	-0.005	-0.338
	Nottawa C.	-0.909	-0.891	-0.915	-0.740
	Tobacco R.	0.616	0.601	0.586	0.255
<i>W</i>	Hasler Creek	-1.678	-1.698	-1.452	-0.455
	Prairie River	0.449	0.477	0.540	-0.040
	Squaw Creek	1.117	1.086	0.904	-0.053
	Swan Creek	0.379	0.449	0.269	0.086
	<i>Average</i>	-0.041	-0.042	-0.041	-0.051

Table 2.7. Mean downstream temperatures of streams with June-October and July-restricted data for year 2016. The stream classes are based on Zorn et. al. (2008), cold (C): July Mean Temperature (JMT) ≤ 17.5 °C; Cold-Transitional: 17.5 °C $<$ JMT ≤ 19.5 °C; Warm-transitional: 19.5 °C $<$ JMT ≤ 21.0 °C; Warm (W) = JMT $>$ 21.0 °C. Streams were assigned to their classes based on their mean JMT values from 30-years of data (Zorn et. al. 2008).

Stream Class	Stream	Mean Downstream Temperature June-October 2016 (°C)	Mean Downstream Temperature July 2016 (°C)
Cold	Black river	15.362	16.813
	Cedar river	13.248	14.943
	Cedar Creek	14.369	15.896
	Morgan Creek	17.550	20.191
	Pokagon Creek	17.441	19.793
Cold-Transitional	Butterfield Creek	15.231	17.712
	Carp river	17.103	19.213
	Pigeon river	16.814	18.553
	Spring Creek	17.410	19.814
Warm-Transitional	Escanaba River	17.296	19.230
	Nottawa Creek	20.306	22.388
	Tobacco River	16.878	19.221
Warm	Hasler Creek	18.051	21.307
	Prairie River	17.788	19.116
	Squaw Creek	17.186	20.236
	Swan Creek	19.529	21.578

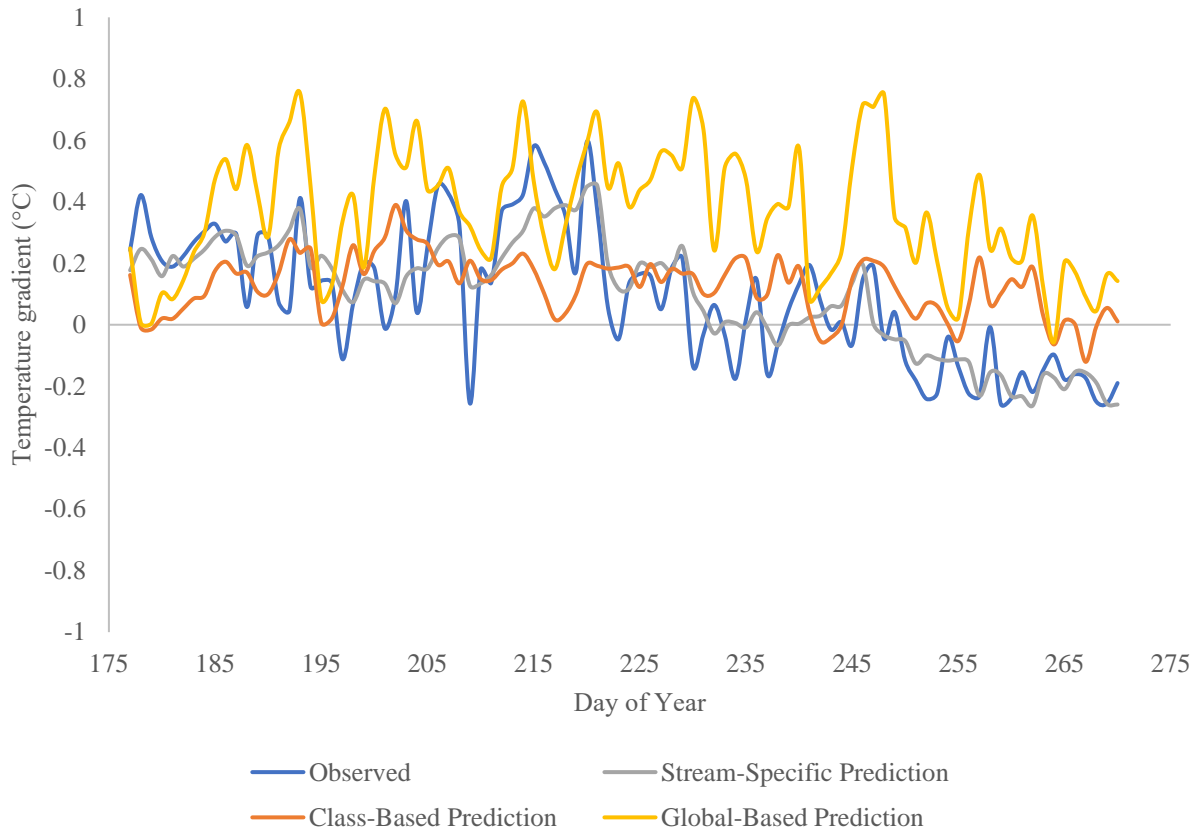


Figure 2.10. Observed and predicted temperature gradient (°C) from Stream-Specific, Class-Based, and Global-Based models. Cedar Creek (cold) June-October 2016 data were used.

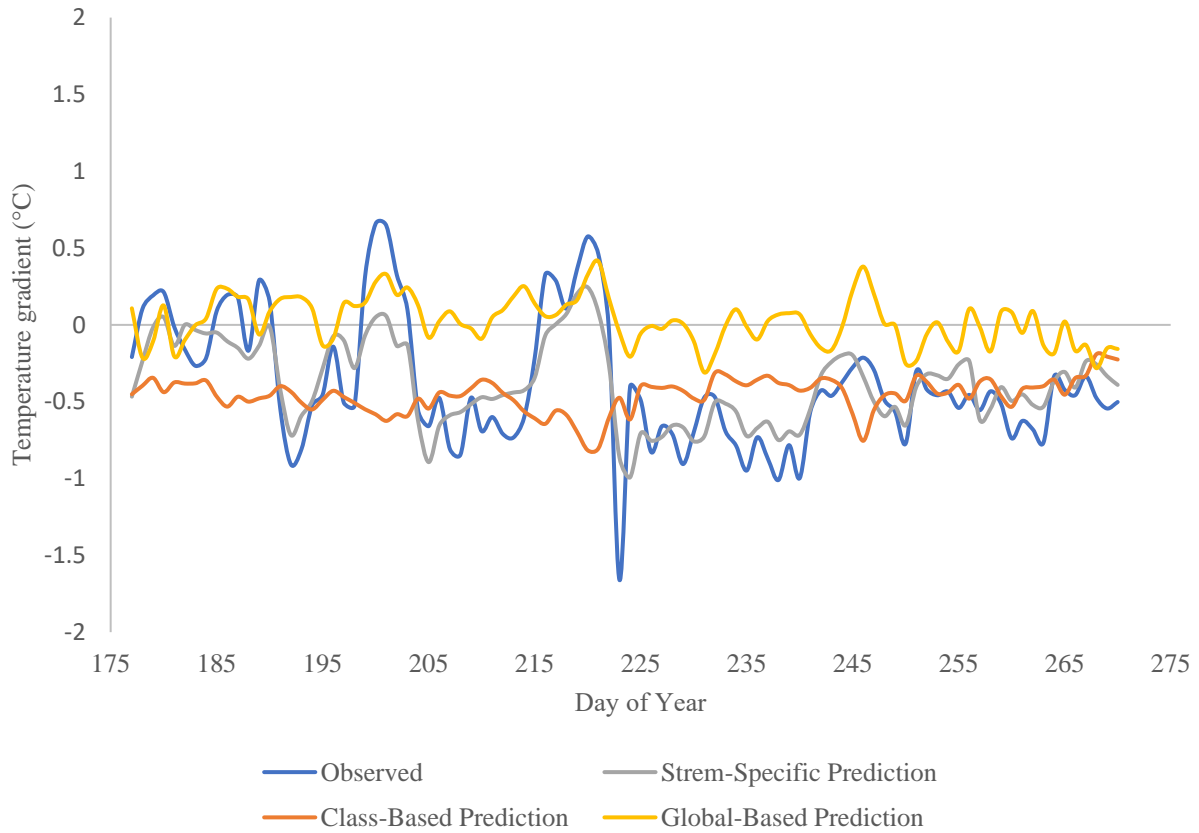


Figure 2.11. Observed and predicted temperature gradient (°C) from Stream-Specific, Class-Based, and Global-Based models. Tobacco River (cold-transitional) June-October 2016 data were used.

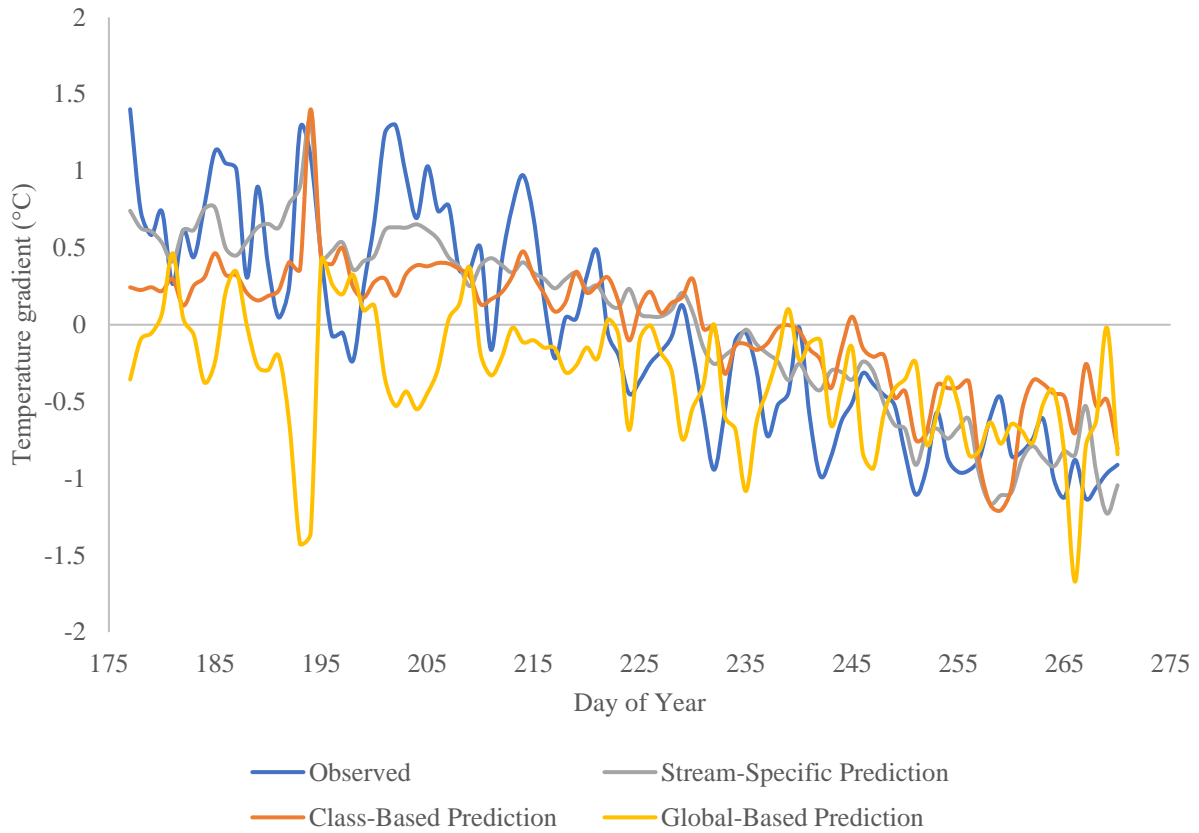


Figure 2.12. Observed and predicted temperature gradient (°C) from Stream-Specific, Class-Based, and Global-Based models. Escanaba River (warm-transitional) June-October 2016 data were used.

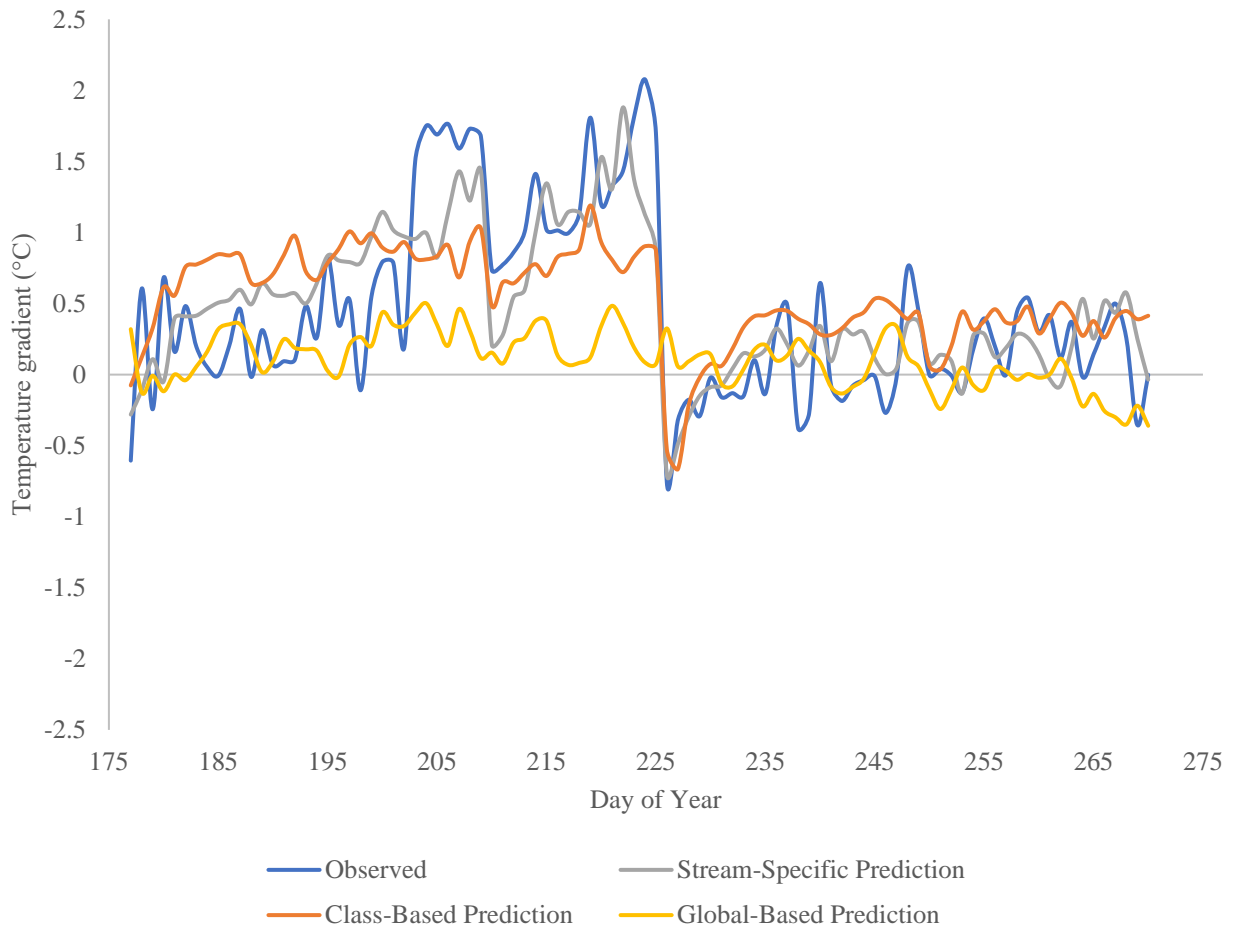


Figure 2.13. Observed and predicted temperature gradient (°C) from Stream-Specific, Class-Based, and Global-Based models. Prairie River (warm) June-October 2016 data were used.

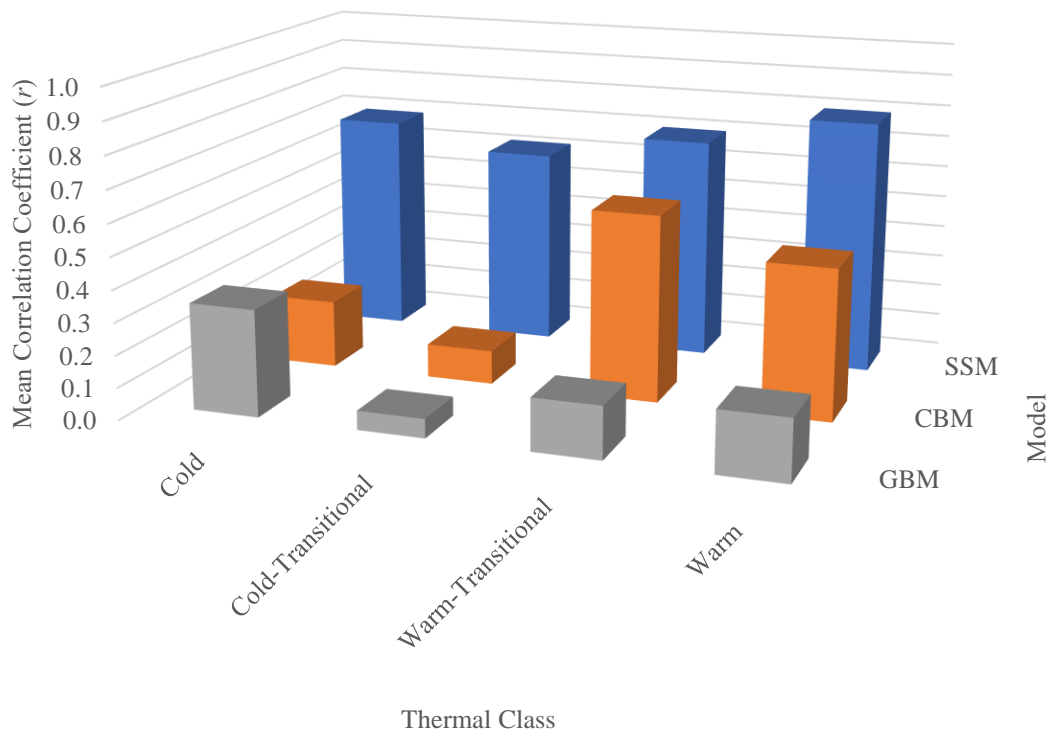


Figure 2.14. Average Pearson correlation coefficient (r) based on stream classes. June – October data were used.

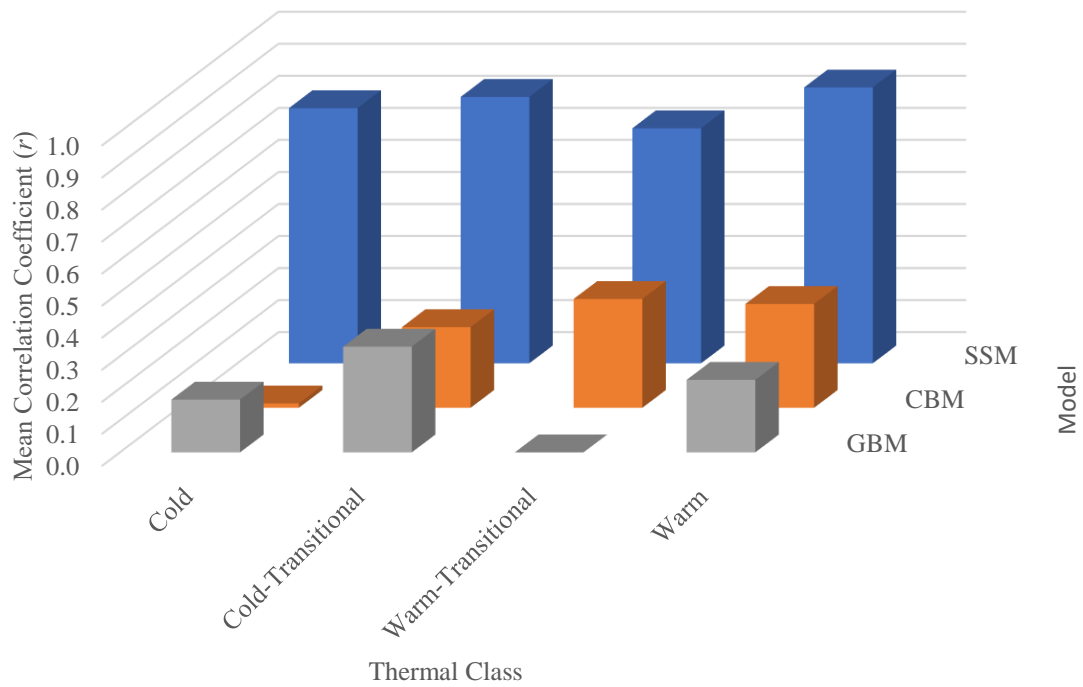


Figure 2.15. Mean Pearson correlation coefficient (r) values were averaged based on stream classes. July 2016 data were used.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Andrews, R. 2019. Effects of flow reduction on stream dynamics of streams: improving an important link in Michigan's water withdrawal assessment tool (-0-0. M.S. Thesis, Michigan State University, East Lansing, MI.
- Baker, E. A. (2006). A landscape-based ecological classification system for river valley segments in Michigan's Upper Peninsula. Michigan Department of Natural Resources, Fisheries Research Report 2085, Ann Arbor.
- Boulton, A. J., Datry, T., Kasahara, T., Mutz, M., & Stanford, J. A. (2010). Ecology and management of the hyporheic zone: Stream-groundwater interactions of running waters and their floodplains. *Journal of the North American Benthological Society*.
<https://doi.org/10.1899/08-017.1>
- Brenden, T. O., Wang, L., & Seelbach, P. W. (2008). A River Valley Segment Classification of Michigan Streams Based on Fish and Physical Attributes. *Transactions of the American Fisheries Society*. <https://doi.org/10.1577/T07-166.1>
- Brett, J. R. (1956). Some Principles in the Thermal Requirements of Fishes. *The Quarterly Review of Biology*. <https://doi.org/10.1086/401257>
- Carlson, A. K., Taylor, W. W., & Infante, D. M. (2019). Developing precipitation- and groundwater-corrected stream temperature models to improve brook charr management amid climate change. *Hydrobiologia*. <https://doi.org/10.1007/s10750-019-03989-1>
- Carlson, A. K., Taylor, W. W., Hartikainen, K. M., Infante, D. M., Beard, T. D., & Lynch, A. J. (2017). Comparing stream-specific to generalized temperature models to guide salmonid management in a changing climate. *Reviews in Fish Biology and Fisheries*.
<https://doi.org/10.1007/s11160-017-9467-0>
- Dugdale, S. J., Malcolm, I. A., Kantola, K., & Hannah, D. M. (2018). Stream temperature under contrasting riparian forest cover: Understanding thermal dynamics and heat exchange processes. *Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2017.08.198>
- Kendy, E., Apse, C., Blann, K., & Richardson, A. (2012). a Practical Guide To Environmental Flows for Policy and Planning. *Nat Conserv*.
- Leathwick, J. R., Snelder, T., Chadderton, W. L., Elith, J., Julian, K., & Ferrier, S. (2011). Use of generalised dissimilarity modeling to improve the biological discrimination of river and stream classifications. *Freshwater Biology*. <https://doi.org/10.1111/j.1365-2427.2010.02414.x>

- Lyons, J., Stewart, J. S., & Mitro, M. (2010). Predicted effects of climate warming on the distribution of 50 stream fishes in Wisconsin, U.S.A. *Journal of Fish Biology*. <https://doi.org/10.1111/j.1095-8649.2010.02763.x>
- Maheu, A., Poff, N. L., & St-Hilaire, A. (2016). A Classification of Stream Water Temperature Regimes in the Conterminous USA. *River Research and Applications*. <https://doi.org/10.1002/rra.2906>
- McManamay, R. A., Smith, J. G., Jett, R. T., Mathews, T. J., & Peterson, M. J. (2018). Identifying non-reference sites to guide stream restoration and long-term monitoring. *Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2017.10.107>
- Niemczynowicz, J. (1999). Urban hydrology and water management - present and future challenges. *Urban Water*. [https://doi.org/10.1016/s1462-0758\(99\)00009-6](https://doi.org/10.1016/s1462-0758(99)00009-6)
- Norton, P.A., Driscoll, D.G., and Carter, J.M. (2019). Climate, streamflow, and lake-level trends in the Great Lakes Basin of the United States and Canada, water years 1960–2015: Scientific Investigations Report 2019–5003, 47 p., <https://doi.org/10.3133/sir20195003>.
- Rosgen, D. L. (1994). A classification of natural rivers. *Catena*. [https://doi.org/10.1016/0341-8162\(94\)90001-9](https://doi.org/10.1016/0341-8162(94)90001-9)
- Rosgen, D.L., (1996). *Applied River Morphology (Second Edition)*. Wildland Hydrology, Pagosa Springs, Colorado.
- Seelbach, P. W., Wiley, M. J., Baker, M. E. and Wehrly K. E. (2006). Initial classification of river valley segments across Michigan's Lower Peninsula. Pages 25–48 in R. Hughes, L. Wang, and P. W. Seelbach, editors. *Landscape influences on stream habitats and biological communities*. American Fisheries Society, Symposium 48, Bethesda, Maryland.
- Seelbach, P.W. & Wiley, Michael & Kotanchik, J.C. & Baker, Matthew. (1997). *A Landscape-Based Ecological Classification for River Valley Segments in Lower Michigan*.
- Stewart, J.S., Westenbroek, S.M., Mitro, M.G., Lyons, J.D., Kammel, L.E., and Buchwald, C.A. (2015). A model for evaluating stream temperature response to climate change in Wisconsin: U.S. Geological Survey Scientific Investigations Report 2014–5186, 64 p., <http://dx.doi.org/10.3133/sir20145186>.
- Sullivan, K., D.J. Martin, R.D. Cardwell, J. E. Toll, and Duke, S. (2000). An analysis of the effects of temperature on salmonids of the Pacific Northwest with implications for selecting temperature criteria. Sustainable Ecosystems Institute, Portland Oregon.
- Tadaki, M., Brierley, G., & Cullum, C. (2014). River classification: theory, practice, politics. *Wiley Interdisciplinary Reviews: Water*. <https://doi.org/10.1002/wat2.1026>

- Tavares Wahren, F., Julich, S., Nunes, J. P., Gonzalez-Pelayo, O., Hawtree, D., Feger, K. H., & Keizer, J. J. (2016). Combining digital soil mapping and hydrological modeling in a data scarce watershed in north-central Portugal. *Geoderma*. <https://doi.org/10.1016/j.geoderma.2015.08.023>
- Thomas, R. E., Gharrett, J. A., Carls, M. G., Rice, S. D., Moles, A., & Korn, S. (1986). Effects of Fluctuating Temperature on Mortality, Stress, and Energy Reserves of Juvenile Coho Salmon. *Transactions of the American Fisheries Society*. [https://doi.org/10.1577/1548-8659\(1986\)115<52:eoftom>2.0.co;2](https://doi.org/10.1577/1548-8659(1986)115<52:eoftom>2.0.co;2)
- Wang, L., Lyons, J., Rasmussen, P., Seelbach, P., Simon, T., Wiley, M., ... Stewart, P. M. (2003). Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, U.S.A. *Canadian Journal of Fisheries and Aquatic Sciences*. <https://doi.org/10.1139/f03-043>
- Webb, B. W., & Zhang, Y. (1997). Spatial and seasonal variability in the components of the river heat budget. *Hydrological Processes*. [https://doi.org/10.1002/\(sici\)1099-1085\(199701\)11:1<79::aid-hyp404>3.0.co;2-n](https://doi.org/10.1002/(sici)1099-1085(199701)11:1<79::aid-hyp404>3.0.co;2-n)
- Wehrly, K. E., Wiley, M. J., & Seelbach, P. W. (2003). Classifying Regional Variation in Stream Regime Based on Stream Fish Community Patterns. *Transactions of the American Fisheries Society*. [https://doi.org/10.1577/1548-8659\(2003\)132<0018:CRVITR>2.0.CO;2](https://doi.org/10.1577/1548-8659(2003)132<0018:CRVITR>2.0.CO;2)
- Westhoff, M. C., Savenije, H. H. G., Luxemburg, W. M. J. ., Stelling, G. S., van de Giesen, N. C., Selker, J. S., ... Uhlenbrook, S. (2007). A distributed stream temperature model using high resolution temperature observations. *Hydrology and Earth System Sciences Discussions*. <https://doi.org/10.5194/hessd-4-125-2007>
- Xin, Z., & Kinouchi, T. (2013). Analysis of stream temperature and heat budget in an urban river under strong anthropogenic influences. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2013.02.048>
- Zorn, T. G., Seelbach, P. W., & Wiley, M. J. (2002). Distributions of Stream Fishes and their Relationship to Stream Size and Hydrology in Michigan's Lower Peninsula. *Transactions of the American Fisheries Society*. [https://doi.org/10.1577/1548-8659\(2002\)131<0070:DOSFAT>2.0.CO;2](https://doi.org/10.1577/1548-8659(2002)131<0070:DOSFAT>2.0.CO;2)
- Zorn, T., Wiley, M. J., Rutherford, E. S. (2008). A Regional-Scale Habitat Suitability Model to Assess the Effects of Flow Reduction on Fish Assemblages in Michigan Streams1, (October 2017). <https://doi.org/10.1111/j.1752-1688.2012.00656.x>
- Zorn, T.G., Seelbach, P.W., and Wiley, M.J. (2004). Utility of Species-Specific, Multiple Linear Regression Models for Prediction of Fish Assemblages in Rivers of Michigan's Lower Peninsula. Michigan Department of Natural Resources, Fisheries Research Report 2072, Ann Arbor, Michigan.

<http://www.michigandnr.com/PUBLICATIONS/PDFS/ifr/ifrlibra/Research/reports/2072rr.pdf>