EXAMINING THE RELATIONSHIP BETWEEN SNOW TELEMETRY AND INTERACTIVE MULTISENSOR SNOW TO PREDICT SPRING STREAMFLOW IN THE COLUMBIA RIVER BASIN

by

KENNETH TURNER

(Under the Direction of Lynne Seymour)

ABSTRACT

In the Northwestern United States, meltwater from snow accumulated on mountains serves as the dominant water supply for many communities. The efficient distribution and use of this renewable, yet temporally and spatially variable resource, relies on the accurate forecasting of Spring streamflow. Here, we examine the utility of adding a specifically defined Snow Telemetry (SNOTEL) variable to already existing satellite-derived snow cover models to predict Spring discharge in the Columbia River Basin. We examined six subbasins of interest in the Columbia River Basin: the Yakima, the Deschutes, the John Day, the Clearwater, the Pend Oreille, and the Kootenai. Within these six subbasins, we propose 144 models; the majority of which contain statistically significant predictive value in forecasting Spring streamflow in the Columbia River Basin. We also discuss remediation for multicollinearity by Principal Component Analysis for the models in which satellite-derived data and snow telemetry data are highly correlated.

INDEX WORDS: Columbia River Basin, streamflow, discharge, remote sensing, snow telemetry, snow cover, principal component analysis

EXAMINING THE RELATIONSHIP BETWEEN SNOW TELEMETRY AND INTERACTIVE MULTISENSOR SNOW TO PREDICT SPRING STREAMFLOW IN THE COLUMBIA RIVER BASIN

by

KENNETH TURNER

BS, University of South Carolina, 2016

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2020

EXAMINING THE RELATIONSHIP BETWEEN SNOW TELEMETRY AND INTERACTIVE MULTISENSOR SNOW TO PREDICT SPRING STREAMFLOW IN THE COLUMBIA RIVER BASIN

by

KENNETH TURNER

Major Professor: Lynne Seymour
Committee: Nicole Lazar
Tharuvai Sriram

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
December 2020

DEDICATION

I would like to thank my wife, Caroline, and my two children, David and Delilah, for the love and support that they provided throughout my graduate study. For me, there was no greater motivating force than making the sacrifices necessary to provide a better life for them.

ACKNOWLEDGMENTS

First, I would like to thank the researchers that laid the foundation for this project. Thomas Estilow, from Rutgers University, was responsible for providing the processed satellite data from the Ice Mapping System (IMS) to be used in this analysis. Thomas Mote, from the University of Georgia, provided the shapefiles for the Columbia River Basin and its respective subbasins. Benjamin Washington, from the University of Georgia, was my predecessor on this project; he laid the groundwork by developing the methodology to detect a signal between IMS data and streamflow discharge in the Columbia River Basin.

Secondly, I would like to thank all of the professors at the University of Georgia – Department of Statistics and the University of South Carolina – Department of Statistics for exposing me to the beauty that is the world of Statistics; without their guidance and patience, I would not be the researcher that I am today.

Lastly, I would like to thank my thesis advisor, Dr. Lynne Seymour. She was my toughest professor, and I knew that her expectations of me would drive me to greatness. She was a light in a dark tunnel for the duration of this thesis, and her contribution to this project cannot be overstated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

As is true in much of the Pacific Northwest, snowpacks constitute a persistent concern for people living in the Columbia River Basin and surrounding areas. Meltwaters from these snowpacks provide the source water for the streams and rivers that drain throughout the basin, supplying the majority of the water upon which the entire population relies (McCabe and Clark, 2005; Serreze et al., 1999). While these snowpacks are essential resources, they also contribute potential threats. In periods of little snow accumulation, low-flow conditions can fall short of the region's water needs. On the other hand, periods of much snow accumulation, or during high rates of snowmelt, high river flows can spur dangerous surges, flooding population centers, threatening lives, and destroying land and property (O'Connor and Costa, 2000; Historical high water events, 2014). Impacts of these extremes and more common subtle inter-annual variations are of importance to public safety, potable water, hydroelectric generation, irrigation, and the economy of the region (Mote, 2003).

Snow cover plays a vital role in the hydrologic cycle by acting as the frozen storage term in the water balance equation (Derksen and LeDrew, 2000). Thus, snow ablation is a significant contributor to streamflow, soil moisture, and potable water supplies. Every river basin has its own snow-ablation characteristics and melt-discharge relationships, with the Columbia River Basin and its subbasins being no exception.

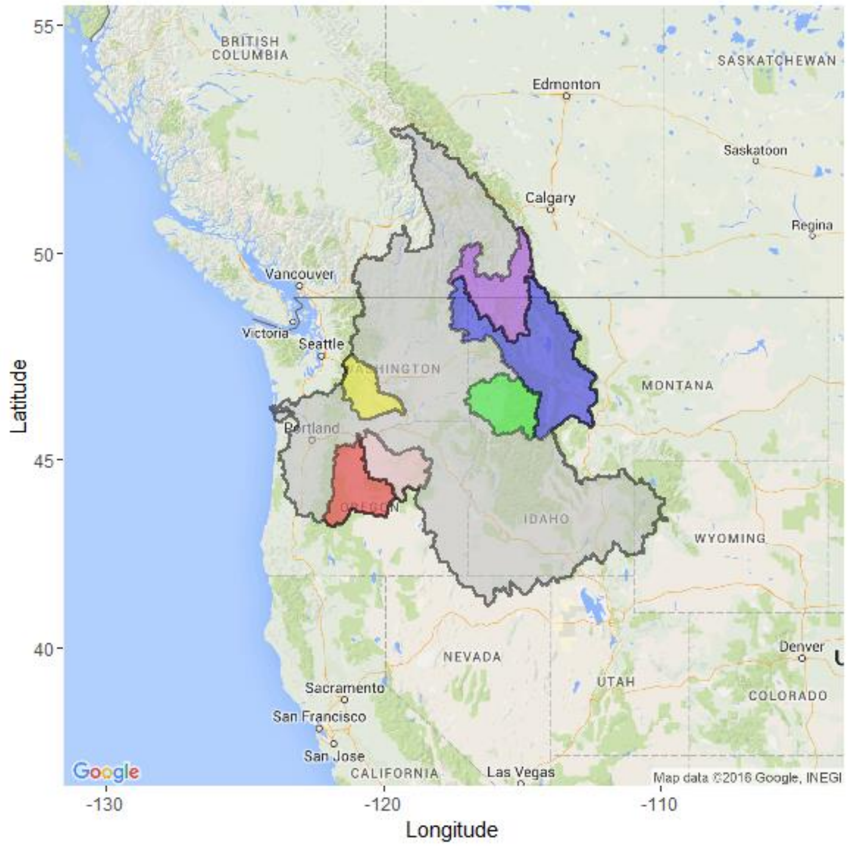

Figure 1. The subbasins within the Columbia River Basin system that were examined in this study. They include the Yakima subbasin (yellow), the Deschutes subbasin (red), the John Day subbasin (pink), the Pend Oreille subbasin (green), the Clearwater subbasin (blue), and the Kootenai subbasin(purple).

Figure 1 depicts the expansiveness of this region and highlights the subbasins of interest in this paper. In this region, where snow cover can be deep and water-laden in the mountains, intra-seasonally fleeting elsewhere, and where discharge exhibits notable inter-annual variability due to melt timing and snow-water content, forecasting of water release from the snowpack presents a daunting challenge. Such variability may lead to harmful environmental and societal consequences, including snowmelt-induced floods, transport of pollutants or excess nutrients in rapid snowmelt events, and lack of adequate streamflow for irrigation, human consumption, and power generation. Thus, accurate forecasting of water release from snowpack is an essential component of any reliable streamflow model.

The motivation behind this project is built upon our perceived insufficiency of models that only use snow telemetry data as a predictor variable. The data points that are retrieved from these stations are not telling the full story of what is going on in these subbasins. These stations are expensive to build and maintain, and thus are in areas where snow is expected. Thus, there is little variation of snowpack in these areas, so these data points introduce an inherent bias since not all areas of a given subbasin have as much snowpack at these remote stations at these high elevations. We will introduce a satellite-derived snow metric, which will tell another part of the story in these subbasins as this metric runs complimentary to snow telemetry data. When reading these introductory sections introducing our variables for this project, take special note of where snow telemetry stations actually reside, and observe how our IMS sampling locations capture a much clearer picture of snowpack. Both of these variables, snow telemetry data and IMS data, both hold important information, and the goal of this project is to get these two variables together in a model to predict Spring streamflow in these select subbasins and examine the results.

CHAPTER 2

Background

Exploratory work in identifying and extracting a seasonal streamflow signal from remotely sensed snow cover in the Columbia River Basin was investigated by Washington et al. (2018). They explore the use of satellite-derived Ice Mapping System (IMS) data in 4-kilometer and 24-kilometer grid format to explain the inter-annual variation in streamflow in the Columbia River Basin. The subbasins they examined were the Kootenai, Pend Oreille, John Day, Deschutes, Clearwater, and Yakima. Their analyses focused on the critical spring discharge season, April to July, when snowmelt is often at its highest, and snowpack distribution and morphology is most challenging to characterize. They proposed a metric, Percent Snow Cover (PSC), to describe the percentage of days under snow cover for a given time window. They then averaged all PSC values, Spatial PSC, for a given region to examine its relationship with its respective streamflow volumes in a respective subbasin and recorded the results for each of the subbasins they considered. This “full” PSC included all the IMS locations in each subbasin.

Washington et al. (2018) found further improvement in the signal by reducing the model. They did this by selecting points in the satellite image that are highly correlated with spring discharge, as well as high inter-annual snow cover variability, by finding the correlation of each IMS location’s February to March snow cover with April to July streamflow discharge. They then calculated each location’s inter-annual variance in February to March and selected only the points that were in the top 50^{th} percentile in both correlation and variance. This “reduced” PSC included only the IMS locations in a given subbasin that met both of these reduction criteria.

The motivation behind the February to March prediction period is that the months immediately preceding the streamflow season should be the most instructive regarding the upcoming spring discharge. They noted improvement in the model in each of the subbasins they examined due to the model reduction. They also noted more inter-annual median variances in an only March prediction period with respect to the reduced PSC.

Table 1

Comparison of the February – March reduced PSC to that of the March reduced PSC linear regression models in the all six subbasins from Washington et al. (2018).

	Feb-March Snow Signal		**March Snow Signal**		
	R^2	**P-value**	**R^2**	**P-value**	**Degrees of Freedom**
Yakima Full	0.35	0.0124	0.42	0.0047	15
Yakima Red.	0.52	0.0012	0.56	0.0006	15
Deschutes Full	0.52	0.0017	0.64	0.0002	14
Deschutes Red.	0.63	0.0002	0.69	0.0001	14
John Day Full	0.30	0.0232	0.29	0.0245	15
John Day Red.	0.34	0.0146	0.39	0.0075	15
Clearwater Full	0.31	0.0204	0.35	0.0118	15
Clearwater Red.	0.35	0.0119	0.41	0.0053	15
Pend Oreille Full	0.24	0.0462	0.33	0.0158	15
Pend Oreille Red.	0.25	0.0412	0.34	0.0138	15
Kootenai Full	0.16	0.1097	0.17	0.0997	15
Kootenai Red.	0.17	0.0963	0.19	0.0810	15

Table 1 are the recorded results of the February – March snow cover reduced PSC compared to that of the March snow cover reduced PSC from their research.

This paper carries forward the “full” and “reduced” PSC methodology for reducing the IMS data for the models that we propose, and then add snow telemetry data as a second predictor to Spring streamflow discharge.   Exploring how these two predictors behave in a multiple regression setting, assessing the value of adding a second predictor for each respective subbasin, determining the predictive utility of having both in a model, and discussing what conclusions can be drawn by the use of both will be the focal point of this analysis.

CHAPTER 3

Data Sources

Ice Mapping System (IMS)

The primary predictor in our analysis is the satellite-derived imagery from the Interactive Multisensor Snow and Ice Mapping System (IMS), which is used to detect the presence or absence of snow cover in the United States, provided by the National Snow and Ice Data Center. IMS data are collected by satellites that can detect a variety of signals synoptically from Earth's surface in regular spatial intervals over large areas every day. These maps are generated at the 24-kilometer resolution and the 4-kilometer resolution daily; approximately 1,500 grid points lie in the Columbia River Basin for the 24-kilometer resolution, and approximately 50,000 grid points lie in the Columbia River Basin for the 4-kilometer resolution.

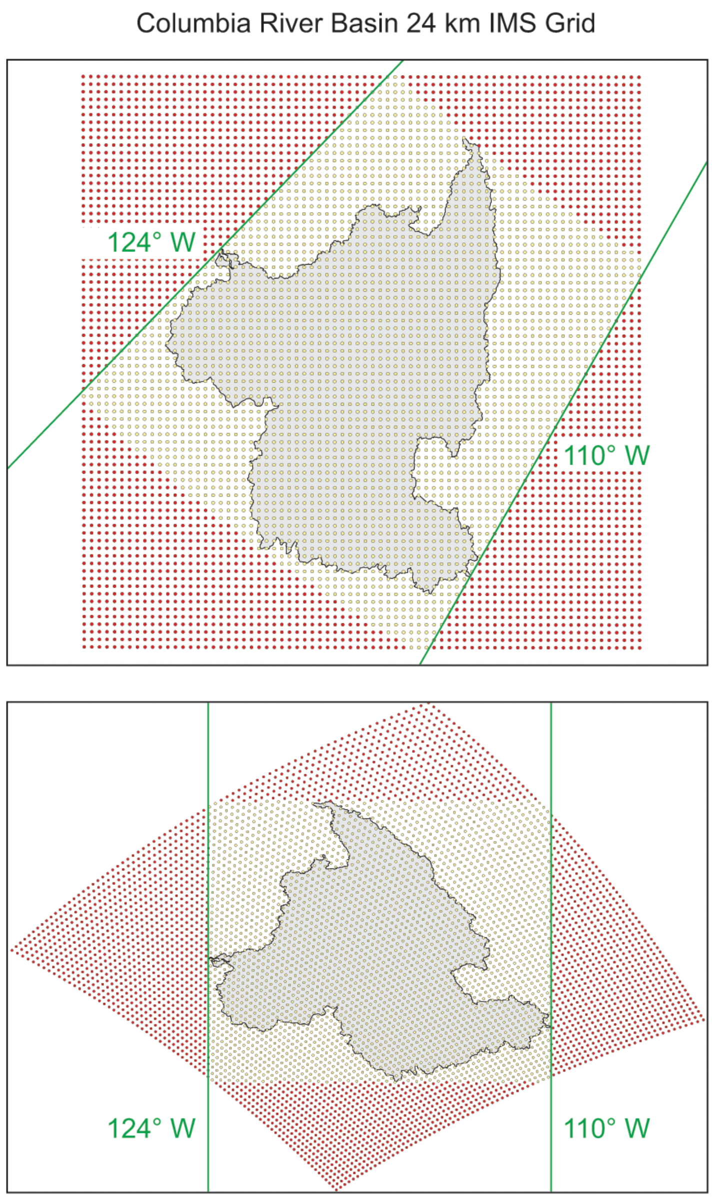

Figure 2. The 24-kilometer resolution IMS grid coordinates in polar stereographic projection (top) and equirectangular projection (below) over the Columbia River Basin.

To better understand the format of the IMS data, it is best to think of a square grid overlaid on top of a polar stereographic projection of the entire Northern Hemisphere, as seen in Figure 2. Geographic shapefiles are used to then remove grid points outside of the Columbia

River Basin; the same type of reduction can be done to remove grid points outside of subbasins of interest contained in the Columbia River Basin.

The IMS methodology facilitates the incorporation of data from multiple satellite and in situ sources and includes interactive image analysis to better recognize regions covered by snow and ice from those lacking such covers (Helfrich et al. 2012). Despite the development of more imagery and better analytical tools, the product has consistently relied upon trained analysts primarily evaluating visible imagery to generate snow cover extent (SCE) maps (Washington et al. 2018). Both the 24-kilometer and the 4-kilometer IMS resolutions used in this analysis can be accessed through the National Snow and Ice Data Center (National Ice Center, 2012).

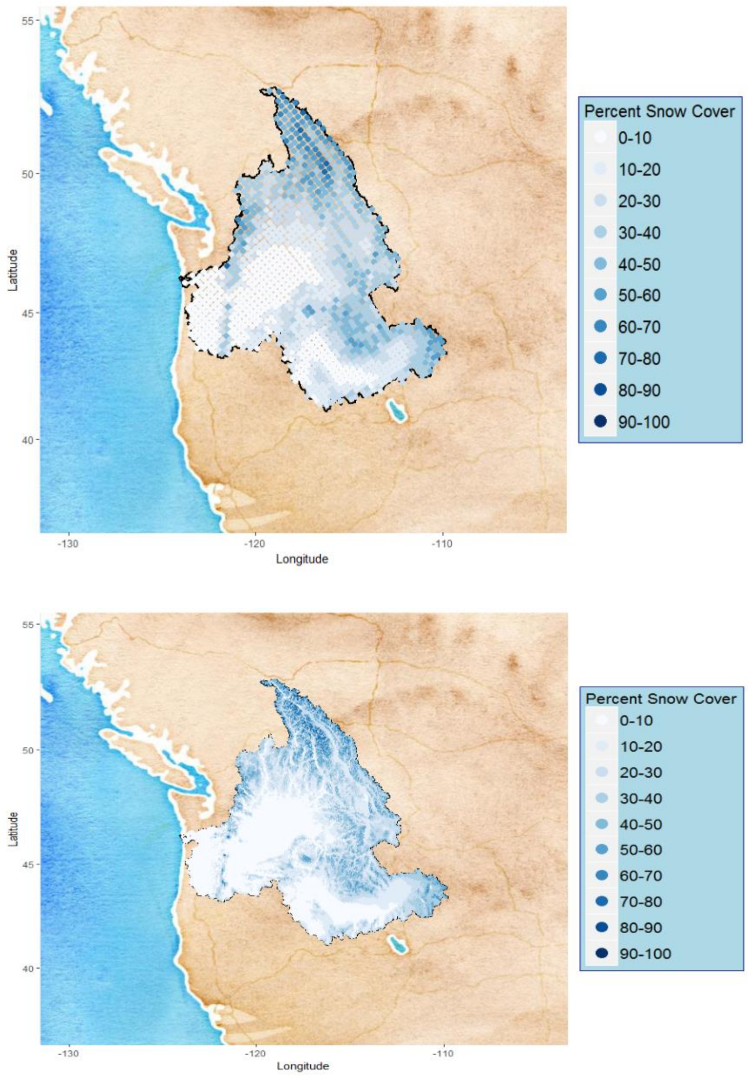

Figure 3. The percent of days with snow cover in the Columbia River Basin based on 24-kilometer resolution data (upper: 1999 - 2018) and 4-kilometer resolution data (lower: 2005 – 2018).

As noted earlier, this gridded satellite product exists in two formats, 24-kilometer and 4-kilometer resolutions. Upon first inspection, when looking at Figure 3, it seems as if the higher resolution would provide a more powerful predictive tool than the lower resolution. However, there is a possibility that predictive power is lost in the higher resolution, since it only dates back to March 2004, whereas the 24-km resolution dates back to January 1999. We believe that the

24-km product provides the spatial and temporal resolution appropriate for extracting a meaningful metric to aid in streamflow prediction, while by the same token not sacrificing five years of predictive power. However, we will consider both in the scope of our analysis, and their respective predictive value will be compared using our model selection criteria.

Snow Telemetry (SNOTEL) Data

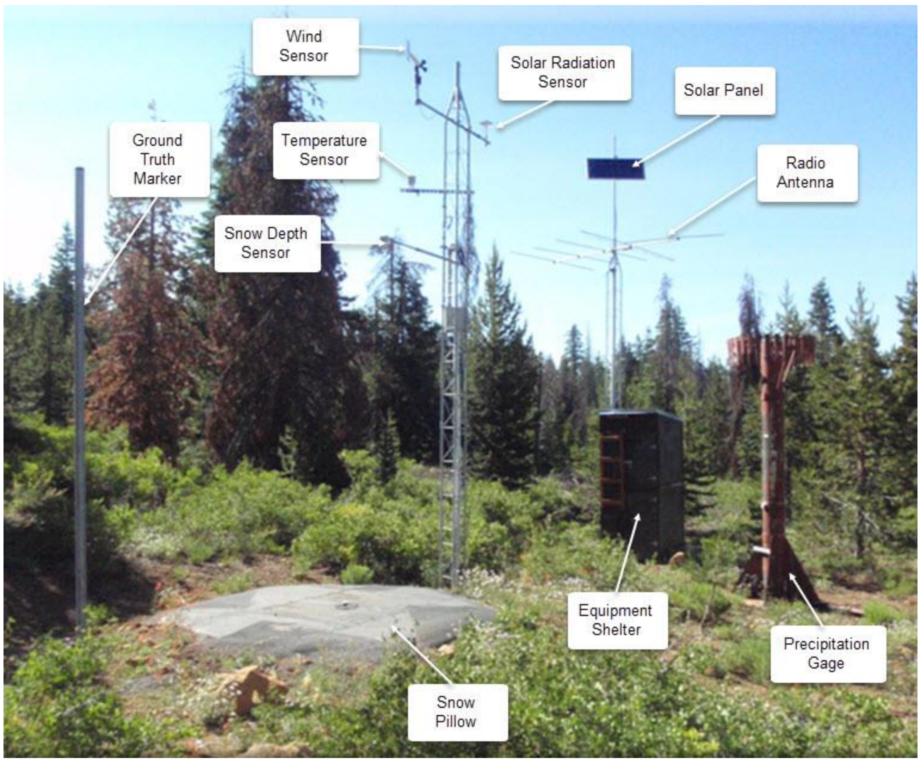

Figure 4. The parts and mechanisms of a Snow Telemetry (SNOTEL) station (NRCS, 2016).

The second predictor in our analysis is snow water equivalent (SWE) data retrieved from snow telemetry (SNOTEL) stations belonging to the Columbia River Basin. These stations use pressure-sensing snow pillows to measure the snow water equivalent (SWE), temperature, precipitation, and change in snow depth, shown in Figure 4. Because SNOTEL stations are expensive to upkeep and maintain, they are often located in remote locations where snow cover is expected; one caveat being that there is little variation in snow cover at the higher elevations where snow telemetry stations are present. Figure 5 is a map of the current snow telemetry sites in the Columbia River Basin.

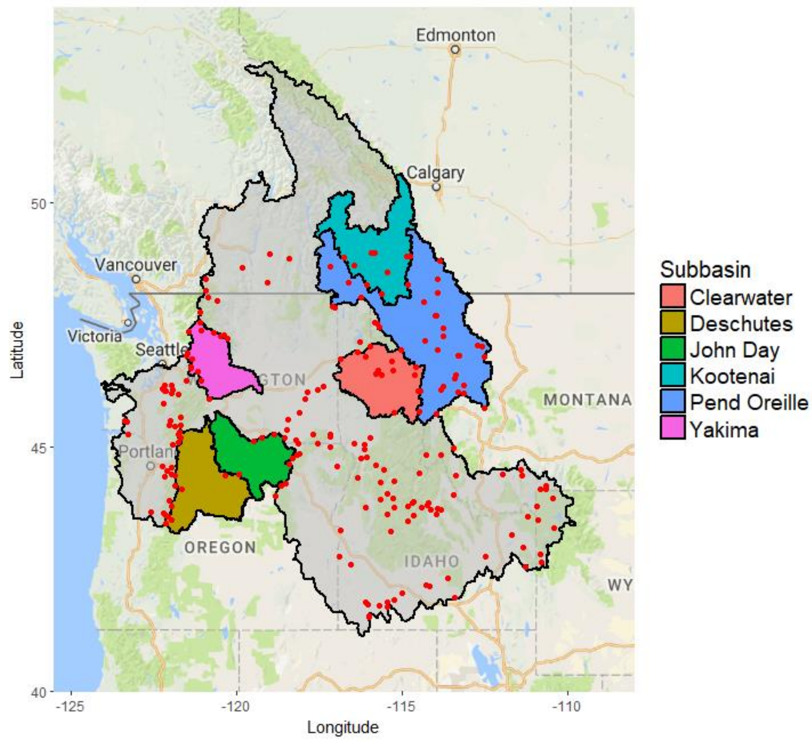

Figure 5. The locations of US SNOTEL stations (red dots) superimposed on a map of the subbasins of interest. Note that the SNOTEL stations commonly are situated along subbasin boundaries, typically at or near geographic apices, where snow cover commonly is thicker and/or more persistent than lower elevations.

While SNOTEL data are temporally detailed, they are spatially sparse at the watershed scale and not intended to represent total snow across the basin (Gleason et al., 2016). Furthermore, using only SNOTEL data to forecast streamflow inherently introduces bias, since not all locations in the Columbia River Basin have as much snow as these stations at high elevations.

Additionally, Nolan and Brown (2008) found that SNOTEL sites in the Willamette River Basin did not sample 50% of the elevation range that is typically covered with snow. Similar concerns with the basin-wide representativeness of SNOTEL locations have been explored at the headwaters of the Rio Grande River Basin (Molotch and Bales, 2006). Within the Columbia River Basin, SNOTEL sites range from 128 meters to 2902 meters above sea level. We will keep this in mind, along with inspecting this range at the subbasin level when we present our own analyses with respect to these subbasins.

Our variable of interest, average daily SWE, is recorded at SNOTEL stations at daily intervals in the Columbia River Basin. This variable will be presented in the model in addition to our IMS data. Current streamflow models use a program called Visual Interactive Prediction and Estimation Routines (VIPER) to make seasonal streamflow forecasts. These models use only SWE for their forecasts. We will be using snow telemetry as a complement to our model, with IMS data as the primary predictor to predict Spring streamflow.

Streamflow Gage Discharge Data

The response in our analysis is measured streamflow discharge. The United States Geological Survey streamgaging network is a multipurpose network that comprises more than 10,000 streamflow gages across the United States, one of the largest in the world (US Geological Survey 2019). We will calculate this output for monthly mean discharge from April to July, a period identified as a seasonal melting period for the Columbia River Basin. More specifically, we will study the relationship between the sum of the monthly means in cubic feet per second from April to July and February and March snowpack data from IMS and SNOTEL within each subbasin, respectively. Figure 6 is an example of a streamflow gage.

Figure 6. A streamflow gage in the Columbia River Basin (Weiser, 2017).

With the selection of our streamflow gages for each respective subbasin, we considered two factors; elevation of the streamflow gage, and how regulated the streamflow gage is. The goal in mind was to find a streamflow gage as low in elevation as possible, with respect to flowing to the Columbia River, without the signal between our predictors being compromised by a regulated streamflow gage. Reservoir storage and dam operations have the potential to affect streamflow characteristics substantially. Documentation of methods of classification of the regulation status of gaging stations is important (Sando, McCarthy, and Dutton 2016). Based on the USGS regulation-classification criteria, a gaging station is regulated if the cumulative drainage area of all upstream dams exceeds 20 percent of the drainage area of the gaging station. The classification goes on to describe possible regulation as major and minor, leading us to believe that the issue is not as simple as classifying a streamflow gage as regulated or unregulated, but rather the extent of the regulation. Examples of major regulations are dams, large diversion canals, and things of this nature. An example of a minor regulation would be smaller diversions for irrigation. All but one of the streamflow gages selected for our study were not regulated at all. The only case of minor regulation was the Service Creek station (14046500) in the John Day subbasin. It should be noted that the majority of the streamflow gages in the John Day subbasin are regulated or missing too many data points to be reliable; in this instance, we deferred to the streamflow gage that the NRCS uses for their streamflow prediction model. Table 2 contains the information for all of the streamflow gages used in this study.

Table 2. A list of the streamflow gages used.

Subbasin	Streamflow Gage	Gage Identification Number
Yakima Subbasin	Station Above Ahtanum Creek at Union Gap	12500450
Deschutes Subbasin	Station at Moody	14103000
John Day Subbasin	Service Creek Station	14046500
Clearwater Subbasin	Orofino Station	13340000
Pend Oreille Subbasin	Clark Fork Station	12392155
Kootenai Subbasin	Yaak River near Troy MT Station	12304500

For the Orofino streamflow station, the month of May 2017 was missing from the data set obtained from the USGS. We used linear regression to impute May 2017, using uninterrupted streamflow data from 1965 – 2016. We simply treated the year 2017 as a new data point, and using the other months of 2017 that we had knowledge of, we regressed on May 2017. The model, model summary, and residual analysis to ensure that linear model assumptions are met can be found in Table 3.

Table 3

For the Orofino streamflow gage in the month of May 2017 was missing from the USGS data set, so that month was imputed using the step-wise Reduced Model (bottom) from the following linear regressions. The global validation of linear model assumptions is on the right-hand side for both models.

Orofino Imputation Model

Full Orifino Imputation Model

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.138e+04	3.204e+03	3.550	0.000887 ***
March	2.912e-01	2.443e-01	1.192	0.239217
April	3.853e-01	1.815e-01	2.123	0.039019 *
June	2.523e-01	1.346e-01	1.875	0.067032 .
July	3.681e-01	4.103e-01	0.897	0.374275

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5478 on 47 degrees of freedom

Multiple R-squared: 0.4636, Adjusted R-squared: 0.418

Predicted R-squared: 0.3212

F-statistic: 10.16 on 4 and 47 DF, p-value: 5.222e-06

VIFs: March = 1.17, April = 1.14, June = 3.92, July = 3.90

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS

USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:

Level of Significance = 0.05

Call:

gvlma(lm(formula = May ~ March + April + June + July))

	Value	p-value	Decision
Global Stat	5.2971	0.2582	Assumptions acceptable.
Skewness	3.0044	0.083	Assumptions acceptable.
Kurtosis	0.0661	0.7972	Assumptions acceptable.
Link Function	0.3636	0.2429	Assumptions acceptable.
Heteroscedasticity	0.863	0.3529	Assumptions acceptable.

Reduced (Stepwise) Orifino Imputation Model

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.249e+04	3.120e+03	4.002	0.000212 ***
April	4.410e-01	1.706e-01	2.585	0.012779 *
June	3.719e-01	6.827e-02	5.448	1.64e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5495 on 49 degrees of freedom

Multiple R-squared: 0.4372, Adjusted R-squared: 0.4142

Predicted R-squared: 0.3561

F-statistic: 19.03 on 2 and 49 DF, p-value: 7.657e-07

VIF: 1.0003

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS

USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:

Level of Significance = 0.05

Call:

gvlma(x = lm(May ~ April + June))

	Value	p-value	Decision
Global Stat	4.8127	0.3071	Assumptions acceptable.
Skewness	1.8824	0.1701	Assumptions acceptable.
Kurtosis	0.0272	0.869	Assumptions acceptable.
Link Function	2.2496	0.1336	Assumptions acceptable.
Heteroscedasticity	0.6534	0.4189	Assumptions acceptable.

The full model was comprised of March, April, June, and July as predictors, but a VIF calculation showed the presence of multicollinearity in the model. We chose to reduce the model in a step-wise fashion to minimize the mean squared error for the model, as well as combat the multicollinearity in the model. The reduced model contained only the months of June and April as predictors for May streamflow. May 2017 streamflow was imputed using June 2017 streamflow and April 2017 streamflow as predictors.

CHAPTER 4

Methodology

Previous Analyses

The purpose of the previous analysis in Washington et al. (2018), as discussed in Chapter 2, was to determine whether IMS data would be a useful predictor for predicting the inter-annual variability in streamflow in the Columbia River Basin. They developed a measure for determining the percentage of cumulative snow cover for a given month using IMS grid data points at the 24-kilometer resolution, as well as the 4-kilometer resolution, and determined the utility of using that metric in a model for predicting seasonal Spring streamflow. Now we are interested in adding another predictor into that model, one that would determine how much snow has accumulated in a subbasin via snow telemetry data. We propose a model in which we use the predictors of if the presence/absence of snow (IMS data) and how much snow there is (SNOTEL/SWE data).

We adopted the methodology of Washington et al. (2018) for using the IMS grid points and reducing them to fit their model. We replicated the beginning stages of their analysis to set a baseline for the performance of our model, since the streamflow gages we chose to use are different from the ones used for their analysis in some cases. We produced a full PSC model, containing all the IMS grid locations for our six respective subbasins. Then that model was reduced based on the criteria that Washington et al. (2018) developed, containing only the IMS grid locations that meet their specific reduction criteria. Once we reduced our IMS variable, we added our snow telemetry variable to our model, which is simply the average snow water equivalent (SWE) for the month of March for all snow telemetry stations that fall in the latitude and longitude specifications of our shapefiles for each respective subbasin. At every step, we reported our results and compared the utility of adding the March SWE data to our model. Looking back further may be of interest to stakeholders, so we examined the predictive ability of our same model structure with February IMS data and February SWE data for predicting April – July Spring streamflow.

Model Comparison

Our criteria for comparing these models will be a calculated value, Predicted R^2. This value is derived from the PRESS statistic, a value that comes from the cross-validation process of a model. The predicted residual error sum of squares (PRESS) statistic is a cross-validated measure used in regression analysis to provide a summary measure of the fit of a model to a sample of observations that were not themselves used to estimate the model. It is calculated as the sums of squares of the prediction residuals for those removed observations. With a fitted model having been produced, each observation is in turn removed and the model is refitted using the remaining observations. The out-of-sample predicted value is calculated for the omitted observation in each case, and the PRESS statistic is calculated as the sum of the squares of all the resulting prediction errors. Given this procedure, the PRESS statistic can be calculated for all of our candidate models with the same data set, with the lowest values of PRESS indicating the best model structures.

Since our sample sizes for our models are small, we will be using a jackknifing approach for our cross-validation, also known as Leave-One-Out cross-validation. A normal R^2 calculation would not be sufficient for the comparison of these models since it is a value that can be artificially inflated by several factors, one of them being multicollinearity. Since multicollinearity is a prevalent issue that needs to be addressed in most of our models, we do not want to chase a value that is artificially inflated for their comparison against the models that do not contain issues of multicollinearity. Also worth noting, a normal R^2 wouldn't provide us any way of assessing the actual predictive power of these models across a consistent baseline for comparison. In the simple linear regression setting (only one predictor), the reason that our Adjusted R^2 would be smaller than our Multiple R^2 is that Adjusted R^2 is intended to be an unbiased estimate of population variance explained using the population regression equation. In this case, as sample size decreases, the difference between Multiple R^2 and Adjusted R^2 increases; in the reverse, as sample size increases, the difference between the two converges to zero. Since our sample sizes are small throughout our analysis, we will continue to see this discrepancy between Multiple R^2 and Adjusted R^2. When we use our jackknifing method of cross-validation, we get our Predicted R^2, which tells us the percentage of the variance in the out-of-sample prediction that is explained by the model.

Our cross-validation method removes a point from the data set, calculates the regression equation with that data point missing, evaluates how well the model predicts the data point that we removed, and then repeats this process for all of the data points in the data set. Predicted R^2 is also going to penalize us overfitting the model. An overfit model is one where too many terms are included to predict the response, and the model starts to fit the random noise in our sample. A Predicted R^2 that is distinctly lower than our Multiple R^2 and Adjusted R^2 will be a sign that we have overfit our model, which is something we will want to keep in mind when evaluating the utility of adding snow telemetry to our existing models. In the multiple regression case, it is important to note that the difference between Multiple R^2 and Adjusted R^2 could be due to sample size, but it could also be due to the model being penalized for adding a predictor that doesn't contribute enough extra sum of squares to warrant it staying in the model.

Testing for Linear Model Assumptions

When using linear regression, we are required to make sure that the four assumptions of linear regression hold for our results to be valid. These four assumptions associated with the linear regression model are:

1. Independence: Observations are independent of one another.
2. Normality: For any fixed value of X, Y is normally distributed.
3. Linearity: The relationship between X and the mean of Y is linear.
4. Homoscedasticity: The variance of residuals is the same for any value of X.

We assess whether these four assumptions hold by performing residual diagnostics. With the number of models that we are proposing in this paper, it would be quite cumbersome to do residual diagnostics for each one. Global Validation of Linear Model Assumptions (gvlma), provides us with a way to evaluate models quickly, and only examine models where there are violations of linear model assumptions (Pena and Slate 2006). The global statistic they discuss, is a test on whether or not all of the assumptions hold. That is, if at least one assumption does not hold, the global statistic will flag that an assumption of the linear model has not been met. Given a particular model, we test globally if all four linear model assumptions are met:

$$H_0: Assumptions\ (1) - (4)\ hold$$

$$H_a: At\ least\ one\ Assumption\ (1) - (4)\ does\ not\ hold$$

That is, the null hypothesis for the global test is that all linear model assumptions hold. The alternative hypothesis is that at least one assumption does not hold. This test is denoted by the name 'Global Stat' in our output. Assumption 1 is denoted as 'Skewness'; this is a measurement of the distribution to see if it is skewed positively or negatively. Rejection of the null for Assumption 1 would indicate that a transformation would likely be needed to remediate the violation of the assumption. Assumption 2 is denoted as 'Kurtosis'; this is a measurement of the assumption of normality. Rejection of the null for Assumption 2 would indicate that our distribution is highly peaked or very shallowly peaked, necessitating a transformation to remediate the violation of the assumption. Assumption 3 is denoted as 'Link Function'; this determines if the model is appropriate. Is there a linear relationship between the response and the predictor(s)? Rejection of the null for Assumption 3 would indicate that our model is incorrectly specified and that we would need to use an alternative form of the generalized linear model to remediate the violation of the assumption. Assumption 4 is denoted as 'Heteroscedasticity'; this checks to make sure the variance of our model residuals are constant across the range of X, or homoscedasticity. Rejection of the null for Assumption 4 would indicate that our residuals are not homoscedastic, and thus non-constant across the range of our predictors. It is important to note that these metrics are mathematical derivations from multiple statistical analysis methods, rather than visually checking the residual diagnostic plots. We will check the residual diagnostic plots as a precaution only if there is a violation of assumptions.

Multicollinearity & Remediation

Remediation will also require some discussion, as IMS data and snow telemetry are highly correlated amongst the respective subbasins. For remediation measures, we will use principal component analysis for each subbasin. The variance inflation factor (VIF) determines how much multicollinearity there is in a given model, and depending on how stringent a researcher is about a VIF threshold, it varies greatly. Due to multicollinearity, we will be using principal component analysis for each respective subbasin, for each respective resolution, for each respective month in addition to our multiple regressions for each case. Readers will have an array of 144 models for each month for prediction. Setting it up this way allows the reader to examine the tradeoff between the complexity of the model and the actual predictive ability of the model.

As mentioned above, we will need to explore the multicollinearity that shows up in a fair amount of our models. This section will discuss what multicollinearity is, the consequences of multicollinearity, how we choose to remediate it with principal component analysis; and then will examine the output from our twelve principal component models (six for 24-kilometer resolution, and six for 4-kilometer resolution).

Is it helpful to first review multiple regression, and what the matrix notation looks like for a multiple regression model. The general linear regression model is defined as:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{i,p} + \epsilon_i$$

where:

- $\beta_0, \beta_1, \ldots, \beta_p$ are parameters;
- $X_{i1}, \ldots, X_{i,p}$ are known constants;
- ϵ_i are independent N(0,σ^2);
- $i = 1, \ldots, n$

which can be converted to the following matrix notation:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \mathbf{X} = \begin{bmatrix} 1 & X_{1,1} & X_{1,2} & \dots & X_{1,p} \\ 1 & X_{2,1} & X_{2,2} & \dots & X_{2,p} \\ \vdots & \vdots & . \vdots & & \vdots \\ 1 & X_{n,1} & X_{n,2} & \dots & X_{n,p} \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

So, in matrix terms, the model is $\boldsymbol{Y} = \boldsymbol{X}\beta + \epsilon$, where:

- **Y** is a vector of responses;
- β is a vector of parameters;
- **X** is a matrix of constants;
- ϵ is a vector of independent normal random variables.

Multicollinearity occurs when one predictor variable in a multiple regression model can be linearly predicted from another predictor variable, for example, when two predictor variables are highly correlated. In this situation, the estimated parameters of the multiple regression may change erratically in response to small changes in the model or the data. Multicollinearity does not reduce the predictive power or reliability of the model as a whole, at least within the sample data set; it only affects calculations regarding individual predictors. That is, a multiple regression model with collinearity among the predictors can indicate how well the set of predictors predicts the outcome variable, but it may not give valid results about any individual predictor while holding other predictors constant. Multicollinearity can have effects on the extra sums of squares, fitted values and predictions, regression coefficients, and several other parts of multiple linear regression.

In a case when we have perfect multicollinearity (the correlation between our predictor variables being 1), our design matrix has less than full rank, and therefore, the moment matrix

$X^T X$ is not invertible. Under these circumstances, for the general linear model, $\boldsymbol{Y} = \boldsymbol{X}\beta + \epsilon$, the ordinary least squares estimator for our parameters, $\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y$ does not exist. The cases in our paper are not instances of perfect multicollinearity, so our $X^T X$ matrix may be invertible, but we may not be able to calculate an approximate inverse, and if we are able to calculate an approximate inverse, it may be highly sensitive to slight variations in the data and may be very inaccurate or very sample-dependent. Another consequence of multicollinearity for our paper is that the estimate of one variable's impact on the response variable Y while controlling for the others will be less precise than if the predictors were uncorrelated with one another. The normal interpretation of a regression coefficient is that it provides an estimate of the effect of a one-unit change in a predictor variable, X_1, while holding the other variable constant. If X_1 is highly correlated with another predictor variable, X_2, in the given data set, then we have a set of observations for which X_1 and X_2 have a linear relationship. We do not have a set of observations for which all changes in X_1 are independent of changes in X_2, so we have an imprecise estimate of the effect of independent changes in X_1 . Thus, making the interpretation of our multiple regression models difficult with the presence of multicollinearity. It is important to note that multicollinearity does not actually bias results; it just produces large standard errors in the related independent variables.

Our measure of remediation to combat the multicollinearity in our models is called Principal Component Analysis. Principal Component Analysis is a dimension reduction tool used to reduce a set of correlated predictor variables to a smaller, less correlated set, called principal components, that still contains most of the information of the larger set. The first principal component contains as much of the variability in the data as possible, and the principal components that follow the first, account for the remaining variability as much as they possibly can. The principal components for a set of predictors are a set of linear combinations of the predictors, chosen so that this captures the most information in a smaller subset of predictors. In our case, we have two predictors that are very linearly correlated. At most, we can have two principal components, the first being a reduced dimension set of both predictors, and the second component being the full dimension set of both predictors. In other words, if we are using principal component analysis of our Reduced March IMS Model with Snow Telemetry data added as a second predictor, including both principal components from that model would just return the regression parameters of that very same model.

A large part of determining principal components is manipulating the X matrix we discussed earlier. In our case, we start with the X matrix of data and two predictor variables, IMS data and SNOTEL data. We assume that the first X matrix has n rows and k+1 columns, and the first column is the 1-vector with the next k = 2 columns are IMS data and SNOTEL data. We also need the columns to be centered and have mean zero. After we center our data, we have

$$X^T X = \begin{pmatrix} n & 0 \\ 0 & A \end{pmatrix} .$$

Taking the eigenvalues and eigenvectors of A, we let V be the diagonal matrix (λ_1,λ_2), and U have u_1, u_2 as columns. Using the spectral decomposition of A, we get $U^T \bar{X}^T \bar{X} U = A$. Finally, we let $P = \bar{X}U$, where each column of P, p_i, is a linear combination of the columns of X, and we call the p_i's principal components of the predictor variables. Since these principal components are linear combinations of the covariates, they are in the column space spanned by the covariates.

The package we use in R to perform our Principal Component Analysis approaches it by singular value decomposition rather than eigenvalue decomposition. The reason that we decided to go with this package is that after the principal components are calculated, it uses an algorithm to convert our principal components back into our least squares regression parameters for each interpretation. Singular value decomposition is a different method of dimensional reduction, so we will describe how we are arriving at our principal components via singular value decomposition, and how it is related to eigenvalue decomposition.

Let **X** be an *n* x *p* design matrix, where *n* is the number of samples and *p* is the number of predictor variables. Let us assume, like before, that the data matrix X is centered. Then, the *p* x *p* covariance matrix **C** is given by $\boldsymbol{C} = \frac{\boldsymbol{X^T X}}{n-1}$. **C** is a symmetric matrix, so it can be diagonalized as $\boldsymbol{C} = \boldsymbol{VLV^T}$, where **V** is a matrix of eigenvectors (each column is an eigenvector) and **L** is a diagonal matrix with eigenvalues λ_i in the decreasing order on the diagonal. The eigenvectors are called principal directions of the data. Projections of the data on the principal directions are called principal components; these can be seen as new, transformed, variables. The j^{th} principal component is given by j^{th} column of **XV**. The coordinates of the i^{th} data point in the new principal component space are given by the i^{th} row of **XV**. If we perform singular value decomposition on our **X**, we obtain the decomposition $\boldsymbol{X} = \boldsymbol{USV^T}$, where **U** is a unitary matrix and **S** is the diagonal matrix of the singular values s_i. From here, we can see that

$$\boldsymbol{C} = \frac{\boldsymbol{VSU^T USV^T}}{n-1} = \boldsymbol{V}(\frac{\boldsymbol{S^2}}{n-1})\boldsymbol{V^T}$$

meaning that the right singular vectors **V** are principal directions and that singular values are related to the eigenvalues of the covariance matrix via $\lambda_i = \frac{s_i^2}{n-1}$, and principal components are given by $\boldsymbol{XV} = \boldsymbol{USV^T V} = \boldsymbol{US}$. We will look at the results of our principal component analysis for the 24-kilometer resolution and the 4-kilometer resolution in the next section.

CHAPTER 5

Analysis

Yakima Subbasin

The Yakima subbasin is located in south-central Washington and contains a diverse landscape of rivers, ridges, and mountains totaling just over 6,100 square miles. Along the western portion of the basin, the glaciated peaks and deep valleys of the Cascade Mountains exceed 8,000 feet. At the 24-kilometer resolution, the Yakima subbasin contains 33 IMS sampling locations for the Full PSC metric and 8 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure 7.

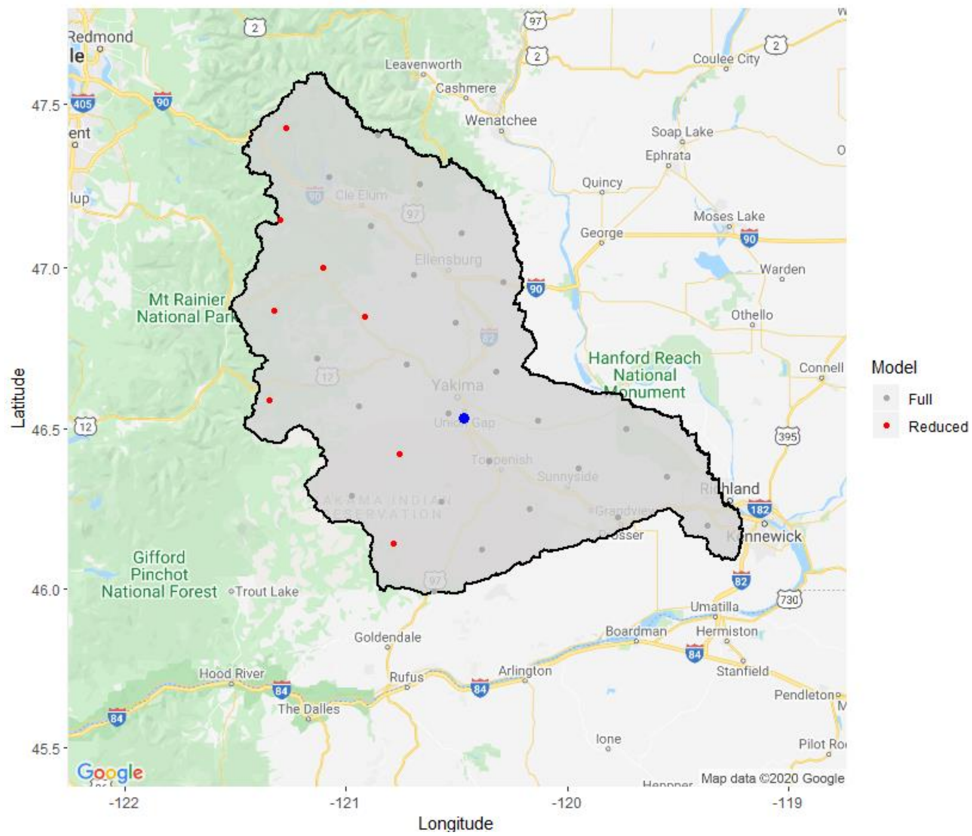

Figure 7. The 24-kilometer resolution IMS sample locations of the Yakima subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 1999 - 2018. The blue dot indicates the location of the streamflow gage for the station Above Ahtanum Creek at Union Gap (12500450).

At the 4-kilometer resolution, the Yakima subbasin contains 1,161 IMS sampling locations for the Full PSC metric and 412 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure 8.

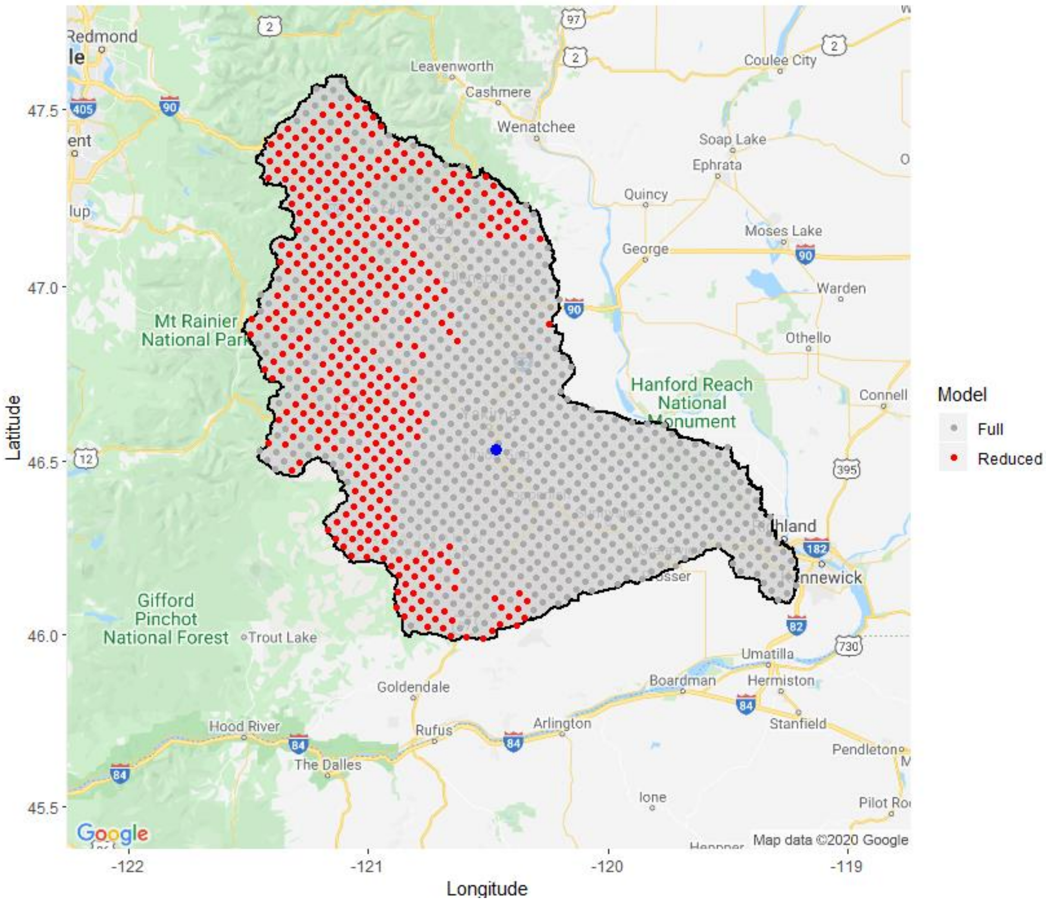

Figure 8. The 4-kilometer resolution IMS sample locations of the Yakima subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 2004 - 2018. The blue dot indicates the location of the streamflow gage for the station Above Ahtanum Creek at Union Gap (12500450).

Our IMS variable for our models in this section is derived from taking the average of these values for the month of March from 1999 to 2018 in the 24-kilometer resolution case, and from 2004-2018 in the 4-kilometer resolution case. The Yakima subbasin is home to 9 snow telemetry stations, ranging from 3,430 feet to 5,920 feet in elevation. Our SNOTEL variable for our models in this section is derived from taking the average daily SWE across all the snow telemetry stations contained in the subbasin.

i. 24-kilometer Resolution

For the case of the 24-kilometer resolution, Table A1 (p. 68) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (24-kilometer resolution), we have a baseline at 0.3922 Multiple R^2. Our value for our Adjusted R^2 is 0.3565, which informally means that approximately 36% of the variation in Spring streamflow in the Yakima subbasin can be explained by the variable comprised of the Full PSC Metric. When we use our jackknifing method of cross-validation, we get a Predicted R^2 of 0.2911, which tells us that approximately 29% of the variance in the out-of-sample prediction is explained by the model. Linear model assumptions were satisfied for the Full March IMS Model (24-kilometer resolution).

For the Reduced March IMS Model (24-kilometer resolution), we have a baseline at 0.4942 Multiple R^2, an Adjusted R^2 of 0.4644, and a Predicted R^2 of 0.402. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.5618 Multiple R^2, an Adjusted R^2 of 0.507, and a Predicted R^2 of 0.433, with no violations of linear model assumptions. For this model, we also need to consider the Variance Inflation Factor (VIF), since it is a multiple regression, and the two predictor variables are highly correlated as shown in Figure A1 (p. 69). The VIF for the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor is 2.28; nothing alarming, but we will consider how highly correlated the two predictors are when we discuss remediation. For this subbasin, at this resolution, we show improvement in our model through its stages of reduction and then adding a second predictor when considering the measurements of model predictive ability.

For the Yakima subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table A2 (p. 70) of the Appendix. Table A1 is the model in the multiple regression setting for comparison. The VIF from the original model was 2.28, which would not be considered high by most standards, but the correlation between the two predictor variables was 0.7208. From the output of Table A2, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing two principal components. Thus, the dimensional reduction from the principal component analysis is not justified, since we arrive back at the regression coefficients, or parameters, that we had with the model before we performed the dimensional reduction. We also have the same Multiple R^2 and Predicted R^2.

ii. 4-kilometer Resolution

For the case of the 4-kilometer resolution, Table A3 (p. 71) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (4-kilometer resolution), we have a baseline at 0.6579 Multiple R^2. Our Adjusted R^2 is 0.6294, which informally means that approximately 63% of the variation in Spring streamflow in the Yakima subbasin can be explained by the variable comprised of the Full PSC Metric. When we use our jackknifing method of cross-validation, we get a Predicted R^2 of 0.5355. Linear model assumptions were satisfied for the Full March IMS Model (4-kilometer resolution).

For the Reduced March IMS Model (4-kilometer resolution), we have a baseline at 0.7323 Multiple R^2, an Adjusted R^2 of 0.71, and a Predicted R^2 of 0.6556. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.7328 Multiple R^2, an Adjusted R^2 of 0.683, and a Predicted R^2 of 0.6055, with no violations of linear model assumptions. For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we need to check the VIF, since it is a multiple regression. The VIF for the March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor is 3.74, which is high, but no surprise as the predictor variables are highly correlated, as shown in Figure A2 (p. 72). Possible remediation will be discussed later on. For this subbasin, at this resolution, we did see improvement from the Full Model to the Reduced Model. However, we did not see improvement from the Reduced Model to the multiple regression with the SNOTEL variable added as a second predictor. The two predictor variables are highly correlated, so it is possible that the second predictor did not add any new information (extra sum of squares) to the model while in the presence of the first predictor. This will be discussed in more depth, after possible remediation measures and model selection criteria.

For the Yakima subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table A4 (p. 73) of the Appendix. Table A3 is the model in the multiple regression setting for comparison. The VIF from the original model was 3.74, which would be considered high by some standards, and the correlation between the two predictor variables was 0.7757. From the output of Table A4, with the results of the cross-validation, we can see that the Root Mean Squared Error for the

Prediction (RMSEP) is minimized for the model containing only one principal component. Thus, the dimensional reduction from the principal component analysis is justified, and it can be concluded that the multicollinearity in the original model had enough of an effect to warrant the dimensional reduction.

Deschutes Subbasin

The Deschutes subbasin stretches over 10,700 square miles of land in central Oregon. Covering eleven percent of Oregon's land area, the Deschutes subbasin is the second-largest watershed in Oregon behind the Willamette. At the 24-kilometer resolution, the Deschutes subbasin contains 58 IMS sampling locations for the Full PSC metric and 14 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure B1 (p. 74). At the 4-kilometer resolution, the Deschutes subbasin contains 2,098 IMS sampling locations for the Full PSC metric, and 536 IMS sampling locations for the Reduced PSC metric as can be seen in Figure B3 (p. 78). Our IMS variable for our models in this section is derived from taking the average of these values for the month of March from 1999 to 2018 in the 24-kilometer resolution case, and from 2004-2018 in the 4-kilometer resolution case. The Deschutes subbasin is home to 8 snow telemetry stations, ranging from 3,810 feet to 5,850 feet in elevation. Our SNOTEL variable for our models in this section is derived from taking the average daily SWE across all the snow telemetry stations contained in the subbasin.

i. 24-kilometer Resolution

For the case of the 24-kilometer resolution, Table B1 (p. 75) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (24-kilometer resolution), we have a baseline at 0.7622 Multiple R^2. Our value for our Adjusted R^2 is 0.7482, which informally means that approximately 75% of the variation in Spring streamflow in the Deschutes subbasin can be explained by the variable comprised of the Full PSC Metric. When we use our jackknifing method of cross-validation, we get a Predicted R^2 of 0.7083, which tells us that approximately 71% of the variance in the out-of-sample prediction is explained by the model. Linear model assumptions were satisfied for the Full March IMS Model (24-kilometer resolution).

For the Reduced March IMS Model (24-kilometer resolution), we have a baseline at 0.806 Multiple R^2, an Adjusted R^2 of 0.7946, and a Predicted R^2 of 0.7651. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.8062 Multiple R^2, an Adjusted R^2 of 0.782, and a Predicted R^2 of 0.7494, with no violations of linear model assumptions. For this model, we also need to consider the Variance Inflation Factor (VIF), since it is a multiple regression, and the two predictor variables are highly correlated as shown in Figure B2 (p. 76). The VIF for the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor is 5.16, which is high by some standards; we will consider this when discussing remediation. For this subbasin, at this resolution, we show improvement in our model through its stages of reduction and then adding a second predictor when considering the measurements of model predictive ability.

For the Deschutes subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table B2 (p. 77) of the Appendix. Table B1 is the model in the multiple regression setting for comparison. The VIF from the original model was 5.16, which would be considered high by some standards, and the correlation between the two predictor variables was 0.7287. From the output of Table B2, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing only one principal components. Thus, the dimensional reduction from the principal component analysis is justified, and it can be concluded that the multicollinearity in the original model had enough of an effect to warrant the dimensional reduction.

ii. 4-kilometer Resolution

For the case of the 4-kilometer resolution, Table B3 (p. 79) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (4-kilometer resolution), we have a baseline at 0.7733 Multiple R^2. Our Adjusted R^2 is 0.7544, which informally means that approximately 75% of the variation in Spring streamflow in the Deschutes subbasin can be explained by the variable comprised of the Full PSC Metric. Our Predicted R^2 is 0.7199. Linear model assumptions were satisfied for the Full March IMS Model (4-kilometer resolution).

For the Reduced March IMS Model (4-kilometer resolution), we have a baseline at 0.7845 Multiple R^2, an Adjusted R^2 of 0.7665, and a Predicted R^2 of 0.7199; our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.8014 Multiple R^2, an Adjusted R^2 of 0.7653, and a Predicted R^2 of 0.6468, with no violations of linear model assumptions. For the March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we need to check the VIF, since it is a multiple regression. The VIF for the March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor is 5.04, which is high, but no surprise as the predictor variables are highly correlated as shown in Figure B4 (p. 80). For this subbasin, at this resolution, we did see improvement, with regards to predictive value, from the Full Model to the Reduced Model. However, we did not see improvement from the Reduced Model to the one with the SNOTEL variable added as a second predictor. The two predictor variables are highly correlated, so it is possible that the second predictor did not add any new information to the model (extra sum of squares) while in the presence of the first predictor. This will be discussed in more depth later, after possible remediation measures and model selection criteria.

For the Deschutes subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table B4 (p. 81) of the Appendix. Table B3 is the model in the multiple regression setting for comparison. The VIF from the original model was 5.04, which would be considered high by some standards, and the correlation between the two predictor variables was 0.8227. From the output of Table B4, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing only one principal component. Thus, the dimensional reduction from the principal component analysis is justified, and it can be concluded that the multicollinearity in the original model had enough of an effect to warrant the dimensional reduction.

John Day Subbasin

The John Day subbasin stretches over 8,000 square miles of land in northeastern Oregon. Through its tributaries, the John Day river drains much of the western side of the Blue Mountains, flowing across the lightly populated arid part of the state east of the Cascade Range in a northwest zigzag, then entering the Columbia upstream from the Columbia River Gorge. At the 24-kilometer resolution, the John Day subbasin contains 47 IMS sampling locations for the Full PSC metric and 10 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure C1 (p. 82). At the 4-kilometer resolution, the John Day subbasin contains 1,538 IMS sampling locations for the Full PSC metric and 564 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure C7 (p. 90). Our IMS variable for our models in this section is derived from taking the average of these values for the month of March from 1999 to 2018 in the 24-kilometer resolution case, and from 2004-2018 in the 4-kilometer resolution case. The John Day subbasin is home to 4 snow telemetry stations, ranging from 5,150 feet to 5,870 feet in elevation. Our SNOTEL variable for our models in this section is derived from taking the average daily SWE across all the snow telemetry stations contained in the subbasin.

i. 24-kilometer Resolution

For the case of the 24-kilometer resolution, Table C (p. 83) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (24-kilometer resolution), we have a baseline at 0.3234 Multiple R^2. Our value for our Adjusted R^2 is 0.2836, which informally means that approximately 28% of the variation in Spring streamflow in the John Day subbasin can be explained by the variable comprised of the Full PSC Metric. When we use our jackknifing method of cross-validation, we get a Predicted R^2 of 0.1709, which tells us that approximately 17% of the variance in the out-of-sample prediction is explained by the model. Linear model assumptions were not satisfied for the Full March IMS Model (24-kilometer resolution), specifically the assumptions for measuring the distribution of the model, kurtosis and skewness. From the test we are using, these are not necessarily visually perceived violations from the assumptions, because our tests are derived from statistical analysis methods. However, if you look closely at the output for these violations of the assumptions in Figure C3 (p. 85) of the Appendix, it can be clearly seen that the histogram of the Standardized residuals is not a bell-shaped curve (approximately normal). Also worth noting that the normal probability plot of the standardized residuals (also known as a quantile-quantile plot) has a very heavy right tail, indicating a possible outlier.

For the Reduced March IMS Model (24-kilometer resolution), we have a baseline at 0.4126 Multiple R^2, an Adjusted R^2 of 0.3781, and a Predicted R^2 of 0.2567; our global test of linear model assumptions found several violations for this model. Assumptions for Kurtosis, Skewness, and Link Function were not satisfied, as can be shown in Figure C4 (p. 86) in the Appendix. The link function assumption usually points to an incorrectly specified model; in the simple linear regression case, this means there is not a sufficient linear relationship between the response variable March Spring streamflow volume and Reduced March IMS (24-kilometer resolution), per the global test.

For the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.4332 Multiple R^2, an Adjusted R^2 of 0.3623, and a Predicted R^2 of 0.2354, with violations of linear model assumptions with respect to skewness and kurtosis as can be seen visually in Figure C6 (p. 88) in the Appendix. In this multiple regression case, it is important to note that the difference between Multiple R^2 and Adjusted R^2 could be due to sample size, but it could also be due to the model being penalized for adding a predictor that doesn't contribute enough extra sums of squares to warrant it staying in the model. For the last model, we also need to consider the Variance Inflation Factor (VIF), since it is a multiple regression, and the two predictor variables are highly correlated as shown in Figure C2 (p. 84). The VIF for the March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor is 1.76, which would not be considered close to most thresholds for identifying multicollinearity. We will explore remediation anyway, given that the two predictors are highly correlated. For this subbasin, at this resolution, we did see improvement, with regards to predictive value, from the Full Model to the Reduced Model, however, we did not see improvement from the Reduced Model to the one with the SNOTEL variable added as a second predictor.

For the John Day subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table C2 (p. 89) of the Appendix. Table C1 is the model in the multiple regression setting for comparison. The VIF from the original model was 1.76, which would not be considered high by any standards, but the correlation between the two predictor variables was 0.7896. From the output of Table C2, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing only one principal component. Thus, the dimensional reduction from the principal component analysis is justified, and it can be concluded that the multicollinearity in the original model had enough of an effect to warrant the dimensional reduction.

ii. 4-kilometer Resolution

For the case of the 4-kilometer resolution, Table C3 (p. 91) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (4-kilometer resolution), we have a baseline at 0.4424 Multiple R^2 , which informally means that approximately 44% of the variation in Spring streamflow in the John Day subbasin can be explained by the variable comprised of the Full PSC Metric. Our Adjusted R^2 is 0.396, and our Predicted R^2 is 0.2271. Linear model assumptions were not all satisfied for the Full March IMS Model (4-kilometer resolution). If you look closely at the table, on the right-hand side for this specified model, you will notice that the global stat test did not reject its null hypothesis, but the individual test for skewness did reject its null hypothesis. The global stat test is designed to reject its null hypothesis whenever any one of the four assumptions are violated. The authors of the Global Validation of Linear Model Assumptions recommend that this be ignored, but we will do an individual test with a Bonferroni correction to verify that this is an incorrect flag of this assumption being violated. With the new p-value cut off for the individual test being 0.0125, alpha = 0.05 divided by 4 (the number of comparisons/tests), we see that we do not reject the individual test with the Bonferroni correction. The residual diagnostic plots from this model can be found in Figure C9 (p. 93) of the Appendix. The histogram of the standardized residuals does not look skewed, but the quantile-quantile plot of the standardized residuals does show a possible outlier. Going forward, we will consider this model not to have any linear model assumptions violations.

For the Reduced March IMS Model (4-kilometer resolution), we have a baseline at 0.4604 Multiple R^2, an Adjusted R^2 of 0.4155, and a Predicted R^2 of 0.2734. We have the same issue with the Full Model when looking at the Global Validation of Linear Model Assumptions. The global stat test did not reject its null hypothesis, but the individual test for skewness did. We approach this issue in the same manner as above and perform a Bonferroni correction for the individual test. With the new p-value cutoff being 0.0125, we do not reject the null hypothesis of the skewness test. The residual diagnostic plots from this model can be found in Figure C10 (p. 94) in the Appendix. The histogram of the standardized residuals looks only slightly skewed, and there is a possible outlier shown in the top right corner of the quantile-quantile plot. Going forward, we will not consider this model to have any linear model assumptions violations.

For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.4702 Multiple R^2, an Adjusted R^2 of 0.3739, and a Predicted R^2 of 0.2218, with violations on the skewness test and the kurtosis test on the Global Validation of Linear Model Assumptions, shown in Figure C12 (p. 96) of the Appendix. We can see the heavy right tail of the quantile-quantile plot of the standardized residuals, and also a slight skew to the histogram of the standardized residuals. For the March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we need to check the VIF. The VIF for the March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a data added as a second predictor is 1.89, which would not be considered close to most thresholds for identifying multicollinearity. We will explore remediation anyway, given that the two predictors are highly correlated, as shown in Figure C8 (p. 92) of the Appendix. For this subbasin, at this resolution, we did see improvement, with regards to predictive ability, from the Full Model to the Reduced Model. However, we did not see improvement from the Reduced Model to the one with SNOTEL added as a second predictor. The two predictor variables are highly correlated, so it is possible that the second predictor did not add any new information to the model (extra sum of squares) while in the presence of the first predictor. This will be discussed in more depth later, after possible remediation measures and model selection criteria.

For the John Day subbasin, the Principal Component Analysis for the Reduced March IMS Model with March SNOTEL added as a second predictor can be found in Table C4 (p. 97) of the Appendix. Table C3 is the model in the multiple regression setting for comparison. The VIF from the original model was 1.89, which would not be considered high by any standards, but the correlation between the two predictor variables was 0.8696. From the output of Table C4, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing only one principal component. Thus, the dimensional reduction from the principal component analysis is justified, and it can be concluded that the multicollinearity in the original model had enough of an effect to warrant the dimensional reduction.

Clearwater Subbasin

The Clearwater River subbasin is located in north-central Idaho, east of Lewiston, with much of the subbasin lying within the Nez Perce Reservation. At the 24-kilometer resolution, the Clearwater subbasin contains 47 IMS sampling locations for the Full PSC metric and 15 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure D1 (p. 98). At the 4-kilometer resolution, the Clearwater subbasin contains 1,779 IMS sampling locations for the Full PSC metric and 599 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure D4 (p. 103). Our IMS variable for our models in this section is derived from taking the average of these values for the month of March from 1999 to 2018 in the 24-kilometer resolution case, and from 2004-2018 in the 4-kilometer resolution case. The Clearwater subbasin is home to 9 snow telemetry stations, ranging from 3,080 feet to 6,360 feet in elevation. Our SNOTEL variable for our models in this section is derived from taking the average daily SWE across all the snow telemetry stations contained in the subbasin.

i. 24-kilometer Resolution

For the case of the 24-kilometer resolution, Table D1 (p. 99) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (24-kilometer resolution), we have a baseline at 0.3474 Multiple R^2. Our value for our Adjusted R^2 is 0.309, which informally means that approximately 31% of the variation in Spring streamflow in the Clearwater subbasin can be explained by the variable comprised of the Full PSC Metric. When we use our jackknifing method of cross-validation, we get a Predicted R^2 of 0.2245, which tells us that approximately 22% of the variance in the out-of-sample prediction is explained by the model. Linear model assumptions were satisfied for the Full March IMS Model (24-kilometer resolution).

For the Reduced March IMS Model (24-kilometer resolution), we have a baseline at 0.3866 Multiple R^2, an Adjusted R^2 of 0.3505, and a Predicted R^2 of 0.2585. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.5433 Multiple R^2, an Adjusted R^2 of 0.4862, and a Predicted R^2 of 0.3675, with a violation on the linear model assumption on the skewness test. If you look closely at the table, on the right-hand side for this specified model, you will notice that the global stat test did not reject its null hypothesis, but the individual test for skewness did reject its null hypothesis. The global stat test is designed to reject its null hypothesis whenever any one of the four assumptions are violated. The authors of the Global Validation of Linear Model Assumptions recommend that this be ignored, but we will do an individual test with a Bonferroni correction to verify that this is an incorrect flag of this assumption being violated. With the new p-value cut off for the individual test being 0.0125, alpha = 0.05 divided by 4 (the number of comparisons/tests), we see that we do not reject the individual test with the Bonferroni correction. The residual diagnostic plots from this model can be found in Figure D3 (p. 101) of the Appendix. The histogram of the standardized residuals does show a slight skew, but the quantile-quantile plot of the standardized residuals seems to follow the normal distribution line very closely. Going forward, we will consider this model not to have any linear model assumptions violations. For this model, we also need to consider the Variance Inflation Factor (VIF), since it is a multiple regression, and the two predictor variables are highly correlated as shown in Figure D2 (p. 100). The VIF for the Reduced March IMS

Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor is 2.19, which is not high, but we will consider how highly correlated the two predictors are when we discuss remediation. For this subbasin, at this resolution, we show improvement in our model through its stages of reduction and then adding a second predictor when considering the measurements of model predictive ability.

For the Clearwater subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table D2 (p. 102) of the Appendix. Table D1 is the model in the multiple regression setting for comparison. The VIF from the original model was 2.19, which would not be considered high by most standards, but the correlation between the two predictor variables was 0.6517. From the output of Table D2, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing two principal components. Thus, the dimensional reduction from the principal component analysis is not justified, since we arrive back at the regression coefficients, or parameters, that we had with the model before we performed the dimensional reduction. We also have the same Multiple R^2 and Predicted R^2.

ii. 4-kilometer Resolution

For the case of the 4-kilometer resolution, Table D3 (p. 104) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (4-kilometer resolution), we have a baseline at 0.475 Multiple R^2. Our Adjusted R^2 is 0.4313, which informally means that approximately 43% of the variation in

Spring streamflow in the Clearwater subbasin can be explained by the variable comprised of the Full PSC Metric. Our Predicted R^2 is 0.3027, which tells us that approximately 30% of the variance in the out-of-sample prediction is explained by the model. Linear model assumptions were satisfied for the Full March IMS Model (4-kilometer resolution).

For the Reduced March IMS Model (4-kilometer resolution), we have a baseline at 0.551 Multiple R^2, an Adjusted R^2 of 0.5136, and a Predicted R^2 of 0.4166. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.617 Multiple R^2, an Adjusted R^2 of 0.5473, and a Predicted R^2 of 0.44, with no violations of the linear model assumptions. For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we need to check the VIF, since it is a multiple regression. The VIF for the March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor is 2.61, which is not high, but the predictor variables are highly correlated, as shown in Figure 5 (p. 105). We will consider how highly correlated the two variables are when we discuss possible remediation. For this subbasin, at this resolution, we show improvement in our model through its stages of reduction and then adding a second predictor when considering the measurements of model predictive ability.

For the Clearwater subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table D4 (p. 106) of the Appendix. Table D3 is the model in the multiple regression setting for comparison. The VIF from the original model was 2.61, which would not be considered high by most standards, but the correlation between the two predictor variables was 0.8249. From the output of Table D4, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing two principal components. Thus, the dimensional reduction from the principal component analysis is not justified, since we arrive back at the regression coefficients, or parameters, that we had with the model before we performed the dimensional reduction. We also have the same Multiple R^2 and Predicted R^2.

Pend Oreille Subbasin

The Pend Oreille subbasin is located in northern Idaho, a northwestern Montana, and northeastern Washington. The Pend Oreille subbasin extends from Cabinet Gorge Dam downstream to the United States (U.S.) - Canadian border. At the 24-kilometer resolution, the Pend Oreille subbasin contains 140 IMS sampling locations for the Full PSC metric and 47 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure E1 (p. 107). At the 4-kilometer resolution, the Pend Oreille subbasin contains 4,884 IMS sampling locations for the Full PSC metric and 1,588 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure E3 (p. 111). Our IMS variable for our models in this section is derived from taking the average of these values for the month of March from 1999 to 2018 in the 24-kilometer resolution case, and from 2004-2018 in the 4-kilometer resolution case. The Pend Oreille subbasin is home to 33 snow telemetry stations, ranging from 4,350 feet to 8,250 feet in elevation. Our SNOTEL variable for our models in this section is derived from taking the average daily SWE across all the snow telemetry stations contained in the subbasin.

i. 24-kilometer Resolution

For the case of the 24-kilometer resolution, Table E1 (p. 108) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (24-kilometer resolution), we have a baseline at 0.3888 Multiple R^2. Our value for our Adjusted R^2 is 0.3528, which informally means that approximately 35% of the variation in Spring streamflow in the Pend Oreille subbasin can be explained by the variable comprised of the Full PSC Metric. When we use our jackknifing method of cross-validation, we get a Predicted R^2 of 0.1925, which tells us that approximately 29% of the variance in the out-of-sample prediction is explained by the model. Linear model assumptions were satisfied for the Full March IMS Model (24-kilometer resolution).

For the Reduced March IMS Model (24-kilometer resolution), we have a baseline at 0.4137 Multiple R^2, an Adjusted R^2 of 0.3792, and a Predicted R^2 of 0.2272. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.5541 Multiple R^2, an Adjusted R^2 of 0.4984, and a Predicted R^2 of 0.2988, with no violations of the linear model assumptions. For the last model, we also need to consider the Variance Inflation Factor (VIF), since it is a multiple regression, and the two predictor variables are highly correlated as shown in Figure E2 (p. 109). The VIF for the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor is 2.24, which is not high, but we will consider how highly correlated the two predictors are when we discuss remediation. For this subbasin, at this resolution, we show improvement in our model through its stages of reduction and then adding a second predictor when considering the measurements of model predictive ability.

For the Pend Oreille subbasin, the Principal Component Analysis for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor can be found in Table E2 (p. 110) of the Appendix. Table E1 is the model in the multiple regression setting for comparison. The VIF from the original model was 2.24, which would not be considered high by most standards, but the correlation between the two predictor variables was 0.6775. From the output of Table E2, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing two principal components. Thus, the dimensional reduction from the principal component analysis is not justified, since we arrive back at the regression coefficients, or parameters, that we had with the model before we performed the dimensional reduction. We also have the same Multiple R^2 and Predicted R^2.

ii. 4-kilometer Resolution

For the case of the 4-kilometer resolution, Table E3 (p. 112) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (4-kilometer resolution), we have a baseline at 0.532 Multiple R^2. Our Adjusted R^2 is 0.493, which informally means that approximately 49% of the variation in Spring streamflow in the Pend Oreille subbasin can be explained by the variable comprised of the Full PSC Metric. Our Predicted R^2 is 0.2973. Linear model assumptions were satisfied for the Full March IMS Model (4-kilometer resolution).

For the Reduced March IMS Model (4-kilometer resolution), we have a baseline at 0.5803 Multiple R^2, an Adjusted R^2 of 0.5454, and a Predicted R^2 of 0.3813. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.6044 Multiple R^2, an Adjusted R^2 of 0.5324, and a Predicted R^2 of 0.3044, with no violations of the linear model assumptions. For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we need to check the VIF, since it is a multiple regression. The VIF for the March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor is 2.53, which is not high, but the predictor variables are highly correlated as shown in Figure E4 (p. 113). For this subbasin, at this resolution, we did see improvement, with regards to predictive ability, from the Full Model to the Reduced Model. However, we did not see improvement from the Reduced Model to the one with the SNOTEL variable added as a second predictor. The two predictor variables are highly correlated, so it is possible that the second predictor did not add any new information to the model (extra sum of squares) while in the presence of the first predictor.

For the Pend Oreille subbasin, the Principal Component Analysis for the Reduced March IMS Model with March SNOTEL added as a second predictor can be found in Table E4 (p. 114) of the Appendix. Table E3 is the model in the multiple regression setting for comparison. The VIF from the original model was 2.53, which would not be considered high by most standards, but the correlation between the two predictor variables was 0.8323. From the output of Table

E4, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing only one principal component. Thus, the dimensional reduction from the principal component analysis is justified, and it can be concluded that the multicollinearity in the original model had enough of an effect to warrant the dimensional reduction.

Kootenai Subbasin

The Kootenai River Subbasin is an international watershed that encompasses parts of British Columbia (B.C.), Montana, and Idaho. The headwaters of the Kootenai River originate in Kootenay National Park, B.C. The river flows south within the Rocky Mountain Trench into the reservoir created by Libby Dam, which is located near Libby, Montana. From the reservoir, the river turns west, passes through a gap between the Purcell and Cabinet Mountains, enters Idaho, and then loops north where it flows into Kootenay Lake, B.C. The waters leave the lake's West Arm and flow south to join the Columbia River at Castlegar, B.C. In terms of runoff volume, the Kootenai is the second largest Columbia River tributary. At the 24-kilometer resolution, the Kootenai subbasin contains 58 IMS sampling locations for the Full PSC metric and 22 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure F1 (p. 115). At the 4-kilometer resolution, the Kootenai subbasin contains 2,109 IMS sampling locations for the Full PSC metric and 851 IMS sampling locations for the Reduced PSC metric, as can be seen in Figure F3 (p. 119). Our IMS variable for our models in this section is derived from taking the average of these values for the month of March from 1999 to 2018 in the 24-kilometer resolution case, and from 2004-2018 in the 4-kilometer resolution case. The Kootenai subbasin is home to 10 snow telemetry stations, ranging from 3,520 feet to 6,903 feet in elevation. Our SNOTEL variable for our models in this section is derived from taking the average daily SWE across all the snow telemetry stations contained in the subbasin.

i. 24-kilometer Resolution

For the case of the 24-kilometer resolution, Table F1 (p. 116) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (24-kilometer resolution), we have a baseline at 0.232 Multiple R^2. Our value for our Adjusted R^2 is 0.1868, which informally means that approximately 19% of the variation in Spring streamflow in the Kootenai subbasin can be explained by the variable comprised of the Full PSC Metric. When we use our jackknifing method of cross-validation, we get a Predicted R^2 of -0.0164; which tells us that approximately 0% of the variance in the out-of-sample prediction is explained by the model. This negative number for our Predicted R^2 is very troubling; we would have more predictive value in just a random sample of the values from the distribution of the streamflow gage values for the months of April – July for past years. Linear model assumptions were satisfied for the Full March IMS Model (24-kilometer resolution).

For the Reduced March IMS Model (24-kilometer resolution), we have a baseline at 0.2578 Multiple R^2, an Adjusted R^2 of 0.2142, and a Predicted R^2 of 0.1114. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.7064 Multiple R^2, an Adjusted R^2

of 0.6697, and a Predicted R^2 of 0.6088, with no violations of the linear model assumptions. For this model, we also need to consider the Variance Inflation Factor (VIF), since it is a multiple regression. The VIF for the Reduced March IMS Model (24-kilometer resolution) with the March SNOTEL variable added as a second predictor is 3.41, which may be concerning depending on how strict of a threshold for VIF you are using. A look at the correlation between the two variables is not high at 0.5934 as seen in Figure F2 (p. 117). We will consider the high VIF when we discuss remediation. For this subbasin, at this resolution, we show improvement in our model through its stages of reduction and then adding a second predictor when considering the measurements of model predictive ability.

For the Kootenai subbasin, the Principal Component Analysis for the Reduced March IMS Model with March SNOTEL added as a second predictor can be found in Table F2 (p. 118) of the Appendix. Table F1 is the model in the multiple regression setting for comparison. The VIF from the original model was 3.41, which would be considered high by some standards, but the correlation between the two predictor variables was 0.5934. From the output of Table F2, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing two principal components. Thus, the dimensional reduction from the principal component analysis is not justified, since we arrive back at the regression coefficients, or parameters, that we had with the model before we performed the dimensional reduction. We also have the same Multiple R^2 and Predicted R^2.

ii. 4-kilometer Resolution

For the case of the 4-kilometer resolution, Table F3 (p. 120) in the Appendix contains the information of three of the four models we will propose for this subbasin at this resolution. For the Full March IMS Model (4-kilometer resolution), we have a baseline at 0.4339 Multiple R^2. which informally means that approximately 43% of the variation in Spring streamflow in the Kootenai subbasin can be explained by the variable comprised of the Full PSC Metric. Our Adjusted R^2 is 0.3867, which informally means that approximately 43% of the variation in Spring streamflow in the Kootenai subbasin can be explained by the variable comprised of the Full PSC Metric. Our Predicted R^2 is 0.2109. Linear model assumptions were satisfied for the Full March IMS Model (4-kilometer resolution).

For the Reduced March IMS Model (4-kilometer resolution), we have a baseline at 0.4751 Multiple R^2, an Adjusted R^2 of 0.4274, and a Predicted R^2 of 0.2943. Our global test of linear model assumptions was acceptable for this reduced model.

For the Reduced March IMS Model (4-kilometer resolution) with the March SNOTEL variable added as a second predictor, we have a baseline at 0.6044 Multiple R^2, an Adjusted R^2 of 0.5324, and a Predicted R^2 of 0.3044, with a very serious violation of the link function test. Even if we were to do a Bonferroni correction on the individual test, the p-value for the test would still fall far beneath the Bonferroni corrected p-value. If we review the residual diagnostic plots in Figure F6 (p. 123), the problem becomes apparent. The histogram of the standardized residuals shows a departure from the normality assumption of linear models. In this case, a linear model is no longer an appropriate model for the multiple regression. For this subbasin, at this resolution, we show improvement in our model through its stages of reduction and then adding a second predictor when considering the measurements of model predictive ability.

However, there was a serious departure from our normality assumption for the Reduced March IMS Model with the March SNOTEL variable added as a second predictor. Even though the model has predictive value, we would not be able to interpret it the way we would normally interpret the results of a linear model. In Figure F5 (p. 123) in the Appendix, the same type of departure from the normality assumption of linear models can be found in the Full March IMS Model with the March SNOTEL variable added as a second predictor as well. Remediation for these models will not be considered in this paper, since they are the only departures from the linear model assumptions of its kind, and there are many other suitable models from which to choose.

For the Kootenai subbasin, the Principal Component Analysis for the Reduced March IMS Model with March SNOTEL added as a second predictor can be found in Table F4 (p. 124) of the Appendix. Table F3 is the model in the multiple regression setting for comparison. The VIF from the original model was 3.29, which would be considered high by some standards, and the correlation between the two predictor variables was 0.7884. From the output of Table F4, with the results of the cross-validation, we can see that the Root Mean Squared Error for the Prediction (RMSEP) is minimized for the model containing two principal components. Thus, the dimensional reduction from the principal component analysis is not justified, since we arrive back at the regression coefficients, or parameters, that we had with the model before we performed the dimensional reduction. We also have the same Multiple R^2 and Predicted R^2. However, back in the chapter for the Analysis, we found that there was a serious violation of a linear model assumption regarding the original form of this model. With this being the case, we would not find this principal component model to be suitable either since it returns the same multiple regression parameters for a model with a serious violation of a linear model assumption.

CHAPTER 6

Results & Discussion

Summary

i. 24-kilometer Resolution

Table G1 (p. 125) of the Appendix holds all of the model results for the models proposed for the month of March at the 24-kilometer resolution. This table contains the models for March IMS Snow Signal on the left-hand side for the Full and Reduced PSC Metric for each subbasin. It also contains the models for March IMS Snow Signal with the addition of the March SNOTEL variable on the right-hand side for the Full and Reduced PSC Metric for each subbasin. We did not review the Full March IMS Model with the addition of March SNOTEL data in Chapter 5 in any detail, but the results for those models can be found in this table as well.

For the Yakima subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model with the addition of the March SNOTEL variable as a second predictor in the multiple regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

For the Deschutes subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model in the simple linear regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

For the John Day subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model in the simple linear regression case. However, none of the John Day models for the 24-kilometer resolution rejected the null hypothesis that all assumptions tests were found to be acceptable.

For the Clearwater subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model with the addition of the March SNOTEL variable as a second predictor in the multiple regression case. This was a special case regarding the global validation of linear model assumptions, where it failed to reject the null hypothesis of the overall test, but the individual test regarding the assumption of the skewness of the distribution rejected the null hypothesis that it was acceptable. We performed a Bonferroni correction of the individual test to gain some insight into the problem, and we did not find a violation of this assumption after all. Therefore, we did not have any evidence to reject the null hypothesis of the global test.

For the Pend Oreille subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model with the addition of the March SNOTEL variable as a second predictor in the multiple regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

For the Kootenai subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model with the addition of the March SNOTEL variable as a second predictor in the multiple regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

ii. 4-kilometer Resolution

Table G2 (p. 126) of the Appendix holds all of the model results for the models proposed for the month of March at the 4-kilometer resolution. This contains the models for March IMS Snow Signal model on the left-hand side for the Full and Reduced PSC Metric for each subbasin.

It also contains the models for March IMS Snow Signal model with the addition of the March SNOTEL variable on the right-hand side for the Full and Reduced PSC Metric for each subbasin. We did not review the Full March IMS Model with the addition of the March SNOTEL variable in Chapter 5 in any detail, but the results for those models can be found in this table as well.

For the Yakima subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model in the simple linear regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

For the Deschutes subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model in the simple linear regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

For the Clearwater subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model with the addition of the March SNOTEL variable as a second predictor in the multiple regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

For the Pend Oreille subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal model in the simple linear regression case. The global validation of linear model assumptions did not reject the null hypothesis that all assumptions tests were found to be acceptable.

For the Kootenai subbasin, our highest Predicted R^2 was found in the Reduced March IMS Snow Signal with the addition of the March SNOTEL variable as a second predictor in the multiple regression case. Our second highest Predicted R^2 was found in the Full March IMS

Snow Signal with the addition of the March SNOTEL variable as a second predictor in the multiple regression case. However, both of these models rejected the link function individual test, with the regression diagnostics confirming. The highest Predicted R^2 without rejecting a null hypothesis on an assumption test was the Reduced March IMS Snow Signal model in the simple linear regression case.

Model Selection Criteria

The model selection criteria we chose is actually quite simple. We used leave-one-out cross-validation to test the predictive ability of every one of our proposed models, which returned us a PRESS statistic, which was then used to calculate our Predicted R^2. We supply an array of all of the models proposed for the month of March predictors in Table G3 (p. 127) in the Appendix. Setting it up this way allows the reader to examine the tradeoff between the complexity of the model and the actual predictive ability of the model. Model assumption violations are noted, and highest Predicted R^2 are highlighted for each subbasin such that determinations of most appropriate model can be made easily. For example, let's look at the John Day subbasin; we have assumptions not met for the Full March IMS Model (24-kilometer), the Reduced March IMS Model (24-kilometer), the Full March IMS Model with the addition of SNOTEL variable (24-kilometer and 4-kilometer), and the Reduced March IMS Model with the addition of the SNOTEL variable (24-kilometer and 4-kilometer). The highest Predicted R^2 belongs to the one component Principal Component Analysis of the Reduced March IMS Model with the addition of SNOTEL (4-kilometer), but this is a complex model to interpret without the algorithm simulating the regression parameters. If we look closely, we see that another Predicted R^2 comes close with the Reduced March IMS Model in the 4-kilometer resolution. This is a simple linear regression case with nearly the same predictive ability. With all of these factors in consideration, one might decide that the simplest model with a competitive predictive ability would be the best choice.

CHAPTER 7

Concluding Thoughts

The focal point of this paper was to introduce snow telemetry data, as we defined it as a variable, into already existing models of Interactive Multisensor Snow data. We wanted to explore the relationships between these two predictor variables with regards to the response, Spring streamflow in the Columbia River Basin. That included the correlation between the two predictors, how they behaved in a model together, and what remediations would be needed in models where they were too related to one another. The methodology we developed delivered an array of models of varying complexities for which a researcher could implement to find an appropriate model, at an appropriate resolution for a specific subbasin, whether that subbasin is in the Columbia River Basin or elsewhere where snow accumulates.

There were several important overall results of the Analysis portion of this paper. The first is that as the IMS resolution increased, so did the correlation of the snow telemetry variable with the IMS variable. Along with this result, we also see the 4-kilometer resolution IMS variable contributing more sum of squares than the models that used the 24-kilometer resolution IMS variable, and thus diminishing the value of adding the snow telemetry variable in most cases. There was one instance where the snow telemetry variable made a great contribution to an IMS model, the Kootenai subbasin. If we think about the characteristics of the Kootenai subbasin, it makes sense; there is almost no variability in snow cover, so obviously an indicator of whether or not there is snow is not as powerful. In the case of the Kootenai subbasin, the snow telemetry variable picks up where the IMS variable lacks. We think that this scenario speaks to the versatility of these models, and how they interact in these models together. This also gives credence to the thought that both of these variables have important information that we do not want to lose in a model reduction. Lastly, in the introduction, we were concerned with the 4-kilometer resolution having enough predictive power, given that there were five years less data than the 24-kilometer resolution. We do not see anything in our analysis that suggests that there is a drop off of predictive power from the 24-kilometer resolution to the 4-kilometer resolution, but rather we see the opposite in most cases.

Of course, now the question becomes, is there any predictive value in February predictors of Snow Telemetry and IMS data? Waiting for March predictors to be vetted and made available for an April – July streamflow prediction could be problematic for some researchers, so we evaluated models of the very same structure for the month of February. The final results for the purposes of model selection can be found in Table H27 (p. 175) of the Appendix. All of the tables and figures used to present our methodology for the March analysis were replicated with February predictor variables, and all can be found in Appendix H for Additional Materials.

When examining Table H27 for the model selection criteria for the month of February, we see greatly diminished returns for all of the subbasins except for the John Day subbasin. In fact, the best performing March and February models for the John Day subbasin are comparable in predictive value; the only issue being that the John Day subbasin had the poorest performing models in the March selection. It could be concluded that there is little drop off in the John Day subbasin from March predictors to February predictors, in the sense that they both performed poorly. The February prediction models for the Yakima and Deschutes subbasin perform okay, but they pale in comparison to their March counterparts. Again, with the array of models across two months of predictors, research can weigh which model best suits their needs for the level of complexity and predictive value that they desire.

The analysis we performed was just another step in the direction of exploring better predicting models for the Columbia River Basin, but there is so much more to accomplish. There is an extraordinary amount of data provided by snow telemetry sites and Interactive Multisensor Snow data that went unutilized in these models: daily temperature, humidity, elevation, slope, etc. One could look at the impact of rain in these months of melt, and how that variable interacts with the already existing variables. There could also be explorations of the model methodology we proposed in other basins where snow accumulates around the world. The opportunities in this subbasin and others for predicting Spring streamflow seem equal parts endless and exciting.

References

Bergeron, J., A. Royer, R. Turcotte, and A. Roy (2014), Snow cover estimation using blended MODIS and AMSR-E data for improved watershed-scale Spring streamflow simulation in Quebec, Canada. *Hydrological Processes*, 28: 4626-4639. doi: 10.1002/hyp.10123.

Biggs, T. W., and T. M. Whitaker (2012), Critical elevation zones of snowmelt during peak discharges in a mountain river basin, *J. Hydrol.*, 438-439, 52–65, doi: 10.1016/j.jhydrol.2012.02.048.

Derksen, C., and E. LeDrew, 2000: Variability and change in terrestrial snow cover: data acquisition and links to the atmosphere. *Progress Physical Geog*, 24: 469–498.

Estilow, T.W., A. H. Young and D.A. Robinson (2015) A long-term Northern Hemisphere snow cover extent data record for climate studies and monitoring. *Earth System Science Data*, 7: 137-142, doi: 10.5194/essd-7-137-2015.

Gleason, K. E., A. W. Nolin and T.R. Roth, 2016. Developing a representative snow monitoring network in a forested mountain watershed. *Hydrol. Earth Syst. Sci. Discuss*., doi: 10.5194/hess-2016-317.

Hall, D.K., J.L. Foster, N.E. DiGirolamo and G.A. Riggs (2012). Snow cover, snowmelt timing and stream power in the Wind River Range, Wyoming. *Geomorphology*, 137: 87–93. doi: 10.1016/j.geomorph.2010.11.011.

Helfrich, S. R., D. McNamara, B. H. Ramsay, T. Baldwin and T. Kasheta, (2007). Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrol. Process.*, 21: 1576–1586. doi: 10.1002/hyp.6720.

Helfrich, S. R., M. Li and C. Kongoli. 2012. Interactive Multisensor Snow and Ice Mapping System Version 3 (IMS V3) Algorithm theoretical basis document version 2.0 Draft 4.1. NOAA NESDIS Center for Satellite Applications and Research (STAR). 59 pp.

Historical High Water Events. 2014. http://orsolutions.org/wp-content/uploads/2014/10/Historic-High-Water-Portland-Fact-Sheet.pdf (accessed 8 July 2019).

Kumar, S., C. Peters-Lidard, D. Mocko and Y. Tian, 2013: Multiscale Evaluation of the Improvements in Surface Snow Simulation through Terrain Adjustments to Radiation. J. Hydrometeor., 14, 220-232, doi: 10.1175/JHM-D-12-046.1.

Liu, Y., C. D. Peters-Lidard, S. V. Kumar, K. R. Arsenault, *and* D. M. Mocko (2015), Blending satellite-based snow depth products with in situ observations for streamflow predictions in the Upper Colorado River Basin, *Water Resources,* 51: 1182-1201, doi: 10.1002/2014WR016606.

McCabe, G. J., and M. P. Clark, 2005: Trends and variability in snowmelt runoff in the western United States. *J. Hydrometeor*., 6: 476–482.

Molotch, N.P. and R.C. Bales. 2006. SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence. *Hydrological Processes,* 20: 723-739. doi: 10.1002/hyp.6128.

Mote, P.W., 2003. Trends in snow water equivalent in the Pacific Northwest and their climatic causes. *Geophysical Research Letters,* 30. doi:10.1029/2003GL017258.

National Ice Center. 2008, updated daily. IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions, Version 1. 1999–2015. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi: 10.7265/N52R3PMC. (accessed 7 April 2019).

NRCS. “SNOTEL automated data collection site.” National Resources and Conservation Service and National Water Climate Center. https://www.wcc.nrcs.usda.gov/about/mon_automate.html. (accessed 4 October 2020).

Nolin, A. W. and A. Brown, 2008. Assessing the representativeness of the NRCS SNOTEL system at a basin-wide scale. 76th Annual Western Snow Conference, April 2008, Hood River, OR.

O’Connor, J. E., and J. E. Costa, 2000: Large floods in the United States: Where they happen and why. U.S. Geological Survey circular 1245, 19pp.

Pena, Edsel and Elizabeth Slate. (2006). Global Validation of Linear Model Assumptions. Journal of the American Statistical Association. 101. 341. doi: 10.1198/016214505000000637.

Ramsay, B. H. 1998. The interactive multisensor snow and ice mapping system. *Hydrol. Process.,* 12: 1537–1546. doi: 10.1002/(SICI)1099-1085(199808/09)12:10/11<1537::AID-HYP679>3.0.CO;2-A.

Rango, A., V. Salomonson, and J. Foster, 1977. Seasonal streamflow estimation in the Himalayan region employing meteorological satellite snow cover observations. *Water Resour. Res*. 13: 109–112.

Rango, A. and J. Martinec, 1979. Application of a snowmelt-runoff model using Landsat data. *Nordic Hydrology*, 10: 225–238.

Rango, A. 1993. Snow hydrology processes and remote sensing. *Hydrol. Proc.* 7: 121–138. doi: 10.1002/hyp.3360070204.

Robinson, D.A., K. F. Dewey and R. Heim, Jr. (1993) Global snow cover monitoring: an update. Bulletin of the American Meteorological Society, 74, 1689-1696. http://dx.doi.org/10.1175/1520-0477(1993)074%3C1689:GSCMAU%3E2.0.CO;2

Sando, S.K., P.M. McCarthy and D.M. Dutton, 2016, Peak-flow frequency analyses and results based on data through water year 2011 for selected streamflow-gaging stations in or near Montana: U.S. Geological Survey Scientific Investigations Report 2015–5019–C, 27: 3-4, doi: 10.3133/sir20155019C.

Serreze, M.C., M.P. Clark and R.L. Armstrong. (1999). Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resources Research*, 35(7), 2145-2160.

U.S. Geological Survey, 2012. National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), accessed 6 May 2019, at URL http://waterdata.usgs.gov/nwis/.

Washington, Benjamin, Lynne Seymour, Thomas Mote, David Robinson and Thomas Estilow. (2018). Identifying and extracting a seasonal streamflow signal from remotely sensed snow cover in the Columbia River Basin. *Remote Sensing Applications: Society and Environment*. doi: 10.1016/j.rsase.2018.03.003.

Weiser, Matt. (2017). Funding Woes for Stream Gages Put Crucial Water Data at Risk. The New Humanitarian. https://deeply.thenewhumanitarian.org/water/community/2017/06/20/how-colorado-plans-to-future-proof-its-water-supply. (accessed on 4 October 2020).

Yang, D., D. A. Robinson, Y. Zhao, T. Estilow, and B. Ye, 2003: Streamflow response to seasonal snow cover extent changes in large Siberian watersheds. *J. Geophys. Res.*, 108: 2156–2202, doi: 10.1029/202JD003149.

Yang, D., Y. Zhao, R. Armstrong and D. A. Robinson, 2009: Yukon River streamflow response to seasonal snow cover changes. *Hydrol. Processes*, 23: 109–121, doi: 10.1002/hyp.7216.

Zhou, X., H. Xie, J.M.H. Hendrickx, 2005. Statistical evaluation of remotely sensed snow-cover products with constraints from streamflow and SNOTEL measurements. *Remote Sensing of Environment*, 94: 214–231.

APPENDIX A

Yakima Subbasin

Table A1

The Full PSC March IMS Model for the 24-kilometer resolution (top), the Reduced PSC March IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 24-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Yakima subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

YAKIMA SUBBASIN

Full March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	7670.44	4182.25	1.834	0.08421 .
yakima.imsfull	287.31	86.75	3.312	0.00412 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5615 on 17 degrees of freedom
Multiple R-squared: 0.3922, Adjusted R-squared: 0.3565
Predicted R-squared: 0.2911
F-statistic: 10.97 on 1 and 17 DF, p-value: 0.004121

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(yakima.stream ~ yakima.imsfull))

	Value	p-value	Decision
Global Stat	2.1258	0.7126	Assumptions acceptable.
Skewness	0.6731	0.412	Assumptions acceptable.
Kurtosis	0.1198	0.7292	Assumptions acceptable.
Link Function	0.003	0.9563	Assumptions acceptable.
Heteroscedasticity	1.3298	0.2488	Assumptions acceptable.

Reduced March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8410.61	3270.55	2.572	0.019808 *
yakima.imsred	172.14	42.24	4.075	0.000788 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5122 on 17 degrees of freedom
Multiple R-squared: 0.4942, Adjusted R-squared: 0.4644
Predicted R-squared: 0.402
F-statistic: 16.61 on 1 and 17 DF, p-value: 0.0007877

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(yakima.stream ~ yakima.imsred))

	Value	p-value	Decision
Global Stat	2.3763	0.6669	Assumptions acceptable.
Skewness	0.2443	0.6211	Assumptions acceptable.
Kurtosis	0.1322	0.7162	Assumptions acceptable.
Link Function	1.2127	0.2708	Assumptions acceptable.
Heteroscedasticity	0.7872	0.3749	Assumptions acceptable.

Reduced March IMS Model (24-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6088.17	3468.31	1.755	0.0983 .
yakima.imsred	105.91	58.47	1.812	0.0889 .
yakima.snotel	251.13	159.80	1.572	0.1356

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4914 on 16 degrees of freedom
Multiple R-squared: 0.5618, Adjusted R-squared: 0.507
Predicted R-squared: 0.433
F-statistic: 10.26 on 2 and 16 DF, p-value: 0.001359
VIF: 2.28

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(yakima.stream ~ yakima.imsred + yakima.snotel))

	Value	p-value	Decision
Global Stat	0.3755	0.9844	Assumptions acceptable.
Skewness	0.1615	0.6878	Assumptions acceptable.
Kurtosis	0.0022	0.9622	Assumptions acceptable.
Link Function	0.1926	0.6607	Assumptions acceptable.
Heteroscedasticity	0.0191	0.8901	Assumptions acceptable.

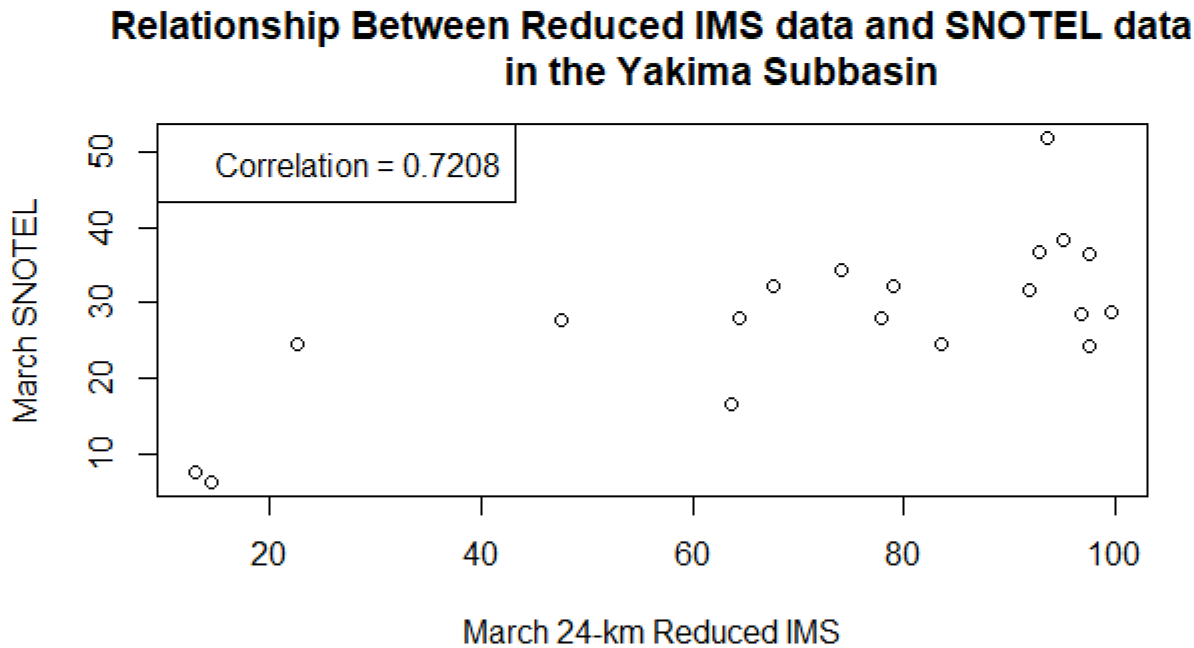

Figure A1. This correlation plot shows the correlation between the March 24-kilometer resolution Reduced PSC variable and the March Snow Telemetry (SNOTEL) variable for the Yakima subbasin.

Table A2

Principal Component Analysis performed for the Yakima subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (24-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Yakima Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	7191	5144	5129
adjCV	7191	5133	5113

TRAINING: % variance explained

	1 comps	2 comps
X	94.72	100.00
yakima.stream	51.64	56.18

Coefficients:

1 component -

(Intercept)	7737.309
yakima.imsred	163.49453
yakima.snotel	45.86402

Multiple R-squared: 0.5164
Predicted R-squared: 0.4298

2 components -

(Intercept)	6088.168
yakima.imsred	105.9121
yakima.snotel	251.1320

Multiple R-squared: 0.5618
Predicted R-squared: 0.4331

Table A3

The Full PSC March IMS Model for the 4-kilometer resolution (top), the Reduced PSC March IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 4-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Yakima subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

YAKIMA SUBBASIN	
Full March IMS Model (4-km) Coefficients: Estimate / Std. Error / t value / Pr(>|t|) (Intercept) 5366.3 3501.5 1.533 0.15131 yakima.imsfull 461.2 96.0 4.804 0.00043 *** --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 4403 on 12 degrees of freedom Multiple R-squared: 0.6579, Adjusted R-squared: 0.6294 Predicted R-squared: 0.5355 F-statistic: 23.08 on 1 and 12 DF, p-value: 0.0004305	ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM: Level of Significance = 0.05 Call: gvlma(x = lm(yakima.stream ~ yakima.imsfull)) Value / p-value / Decision Global Stat 0.6796 0.9538 Assumptions acceptable. Skewness 0.0075 0.9308 Assumptions acceptable. Kurtosis 0.018 0.8932 Assumptions acceptable. Link Function 0.3148 0.5748 Assumptions acceptable. Heteroscedasticity 0.3392 0.5603 Assumptions acceptable.
Reduced March IMS Model (4-km) Coefficients: Estimate / Std. Error / t value / Pr(>|t|) (Intercept) 6122.5 2831.7 2.162 0.0515 . yakima.imsred 250.9 43.8 5.729 9.47e-05 *** --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 3895 on 12 degrees of freedom Multiple R-squared: 0.7323, Adjusted R-squared: 0.71 Predicted R-squared: 0.6556 F-statistic: 32.82 on 1 and 12 DF, p-value: 9.47e-05	ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM: Level of Significance = 0.05 Call: gvlma(x = lm(yakima.stream ~ yakima.imsred)) Value / p-value / Decision Global Stat 1.6329 0.8029 Assumptions acceptable. Skewness 0.002 0.9643 Assumptions acceptable. Kurtosis 0.0981 0.7542 Assumptions acceptable. Link Function 0.3554 0.5511 Assumptions acceptable. Heteroscedasticity 1.1775 0.2779 Assumptions acceptable.
Reduced March IMS Model (4-km) w/ March SNOTEL Coefficients: Estimate / Std. Error / t value / Pr(>|t|) (Intercept) 6378.24 3390.62 1.881 0.08667 . yakima.imsred 259.55 72.41 3.584 0.00429 ** yakima.snotel -28.63 186.23 -0.154 0.88062 --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 4064 on 11 degrees of freedom Multiple R-squared: 0.7328, Adjusted R-squared: 0.6843 Predicted R-squared: 0.6055 F-statistic: 15.09 on 2 and 11 DF, p-value: 0.0007033 VIF: 3.74	ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM: Level of Significance = 0.05 Call: gvlma(x = lm(yakima.stream ~ yakima.imsred + yakima.snotel)) Value / p-value / Descision Global Stat 1.7187 0.7873 Assumptions acceptable. Skewness 0.0008 0.977 Assumptions acceptable. Kurtosis 0.1344 0.7139 Assumptions acceptable. Link Function 0.5741 0.4486 Assumptions acceptable. Heteroscedasticity 1.0094 0.315 Assumptions acceptable.

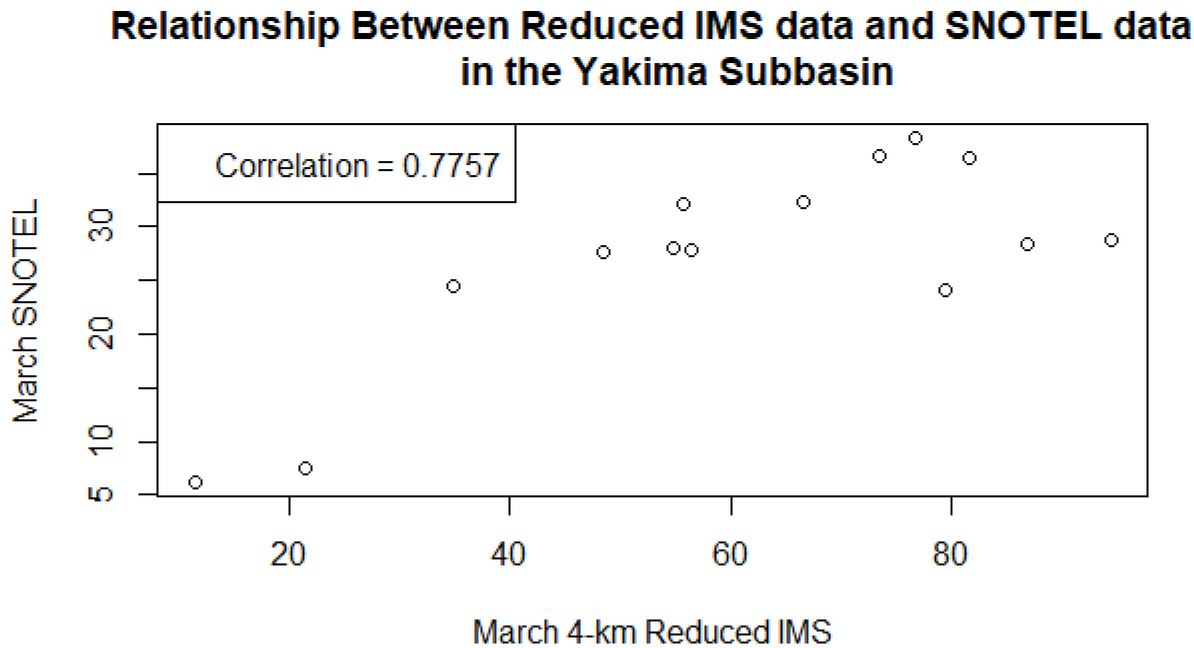

Figure A2. This correlation plot shows the correlation between the March 4-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Yakima subbasin.

Table A4

Principal Component Analysis performed for the Yakima subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (4-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Yakima Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 14 2
Y dimension: 14 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 14 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	7506	4132	4378
adjCV	7506	4115	4349

TRAINING: % variance explained

	1 comps	2 comps
X	95.23	100.00
yakima.stream	72.56	73.28

Coefficients:

1 component -

(Intercept)	5581.22
yakima.imsred	227.26045
yakima.snotel	72.53312

Multiple R-squared: 0.7256
Predicted R-squared: 0.6485

2 components -

(Intercept)	6378.236
yakima.imsred	259.54639
yakima.snotel	-28.62506

Multiple R-squared: 0.7328
Predicted R-squared: 0.6055

APPENDIX B

Deschutes Subbasin

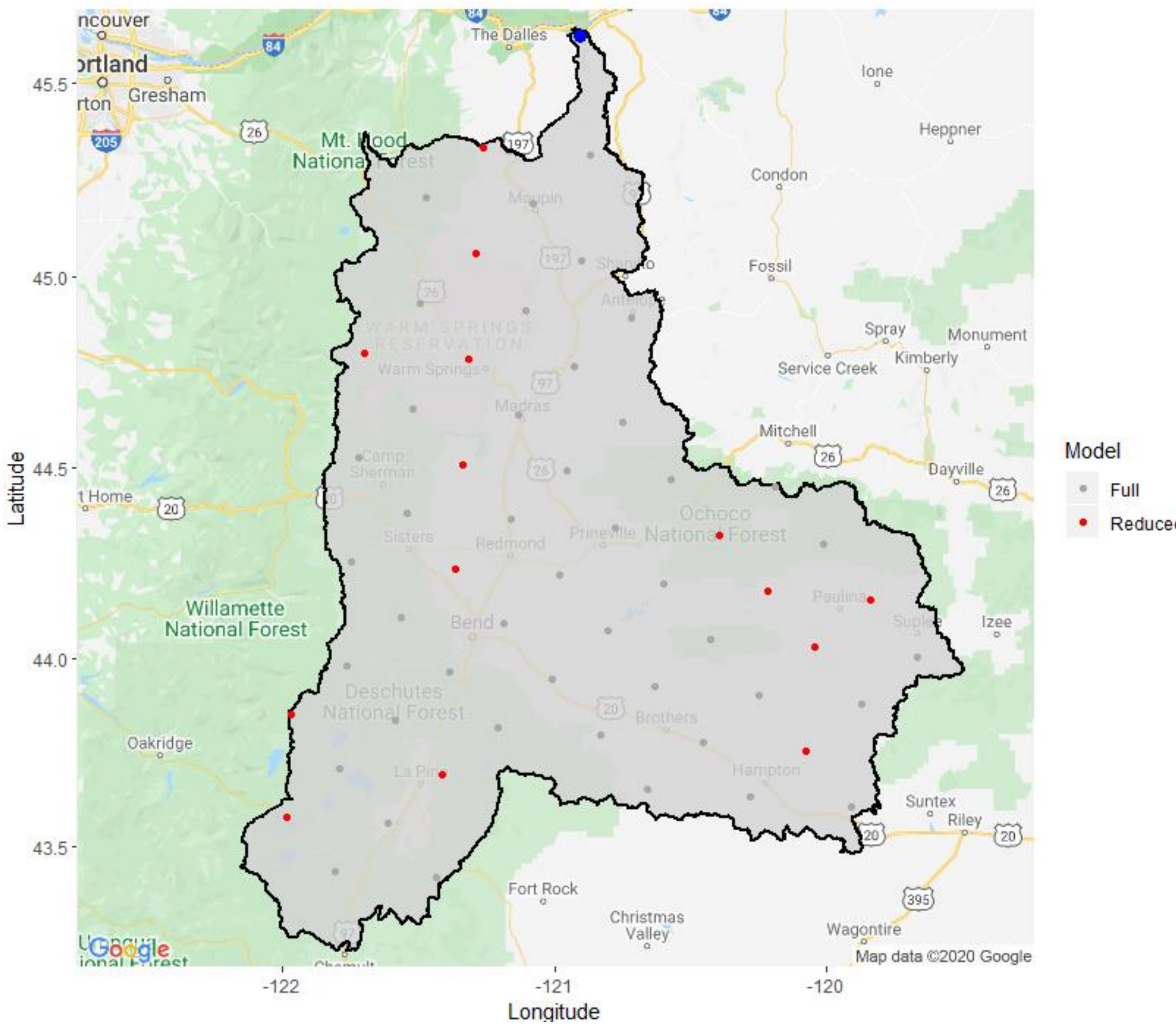

Figure B1. The 24-kilometer resolution IMS sample locations of the Deschutes subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 1999 - 2018. The blue dot indicates the location of the streamflow gage for the station at Moody (14103000).

Table B1

The Full PSC March IMS Model for the 24-kilometer resolution (top), the Reduced PSC March IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 24-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Deschutes subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

DESCHUTES SUBBASIN

Full March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>| t|)
(Intercept)	16772.84	961.89	17.437	2.78e-12 ***
deschutes.imsfull	183.14	24.81	7.382	1.07e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1723 on 17 degrees of freedom
Multiple R-squared: 0.7622, Adjusted R-squared: 0.7482
Predicted R-squared: 0.7083
F-statistic: 54.49 on 1 and 17 DF, p-value: 1.072e-06

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsfull))

	Value	p-value	Decision
Global Stat	3.339	0.503	Assumptions acceptable.
Skewness	0.322	0.57	Assumptions acceptable.
Kurtosis	0.07	0.792	Assumptions acceptable.
Link Function	2.744	0.098	Assumptions acceptable.
Heteroscedasticity	0.203	0.652	Assumptions acceptable.

Reduced March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>| t|)
(Intercept)	19001.0	618.5	30.724	2.45e-16 ***
deschutes.imsred	111.8	13.3	8.405	1.85e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1556 on 17 degrees of freedom
Multiple R-squared: 0.806, Adjusted R-squared: 0.7946
Predicted R-squared: 0.7651
F-statistic: 70.65 on 1 and 17 DF, p-value: 1.85e-07

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred))

	Value	p-value	Decision
Global Stat	2.371	0.668	Assumptions acceptable.
Skewness	0.76	0.383	Assumptions acceptable.
Kurtosis	0.02	0.887	Assumptions acceptable.
Link Function	0.508	0.476	Assumptions acceptable.
Heteroscedasticity	1.082	0.298	Assumptions acceptable.

Reduced March IMS Model (24-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>| t|)
(Intercept)	19110.8	1076.205	17.758	5.94e-12 ***
deschutes.imsred	113.977	22.183	5.138	9.92e-05 ***
deschutes.snotel	-9.866	77.962	-0.127	0.901

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1603 on 16 degrees of freedom
Multiple R-squared: 0.8062, Adjusted R-squared: 0.782
Predicted R-squared: 0.7494
F-statistic: 33.29 on 2 and 16 DF, p-value: 1.986e-06
VIF: 5.16

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred + deschutes.snotel))

	Value	p-value	Descision
Global Stat	2.326	0.676	Assumptions acceptable.
Skewness	0.708	0.4	Assumptions acceptable.
Kurtosis	0.027	0.87	Assumptions acceptable.
Link Function	0.568	0.451	Assumptions acceptable.
Heteroscedasticity	1.024	0.312	Assumptions acceptable.

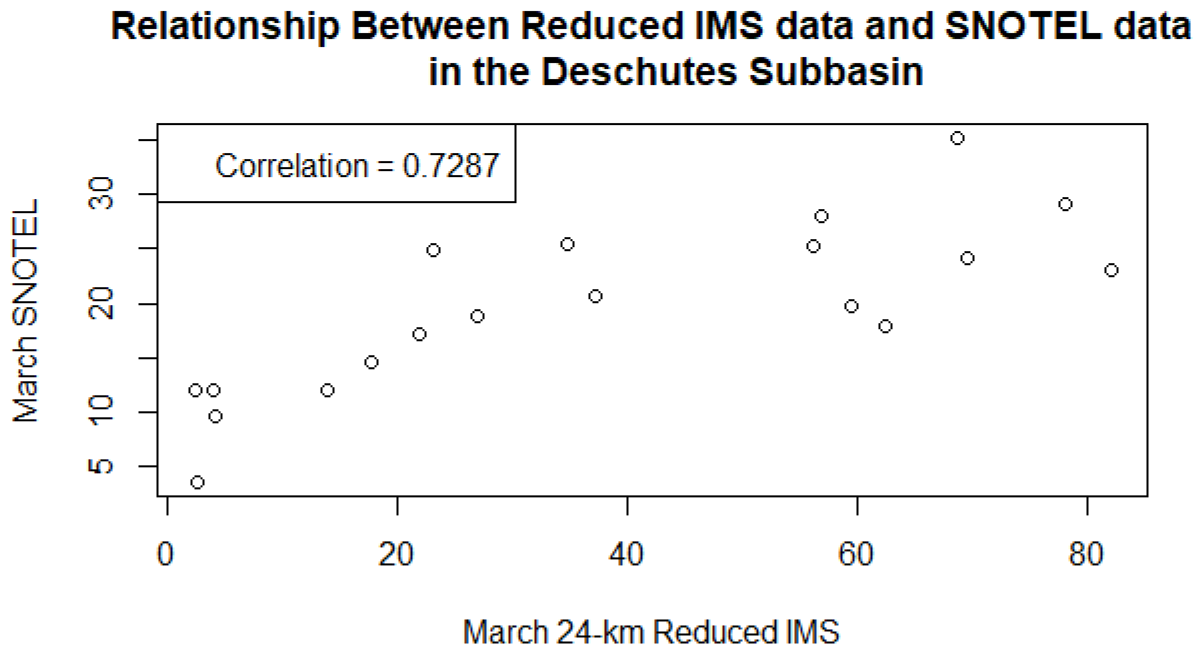

Figure B2. This correlation plot shows the correlation between the March 24-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Deschutes subbasin.

Table B2

Principal Component Analysis performed for the Deschutes subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (24-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Deschutes Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	3527	1628	1673
adjCV	3527	1624	1667

TRAINING: % variance explained

	1 comps	2 comps
X	97.28	100.00
deschutes.stream	80.39	80.62

Coefficients:

1 component -

(Intercept)	18737.83
deschutes.imsred	106.06422
deschutes.snotel	24.45366

Multiple R-squared: 0.8039
Predicted R-squared: 0.7625

2 components -

(Intercept)	19110.8
deschutes.imsred	113.976827
deschutes.snotel	-9.866125

Multiple R-squared: 0.8062
Predicted R-squared: 0.7494

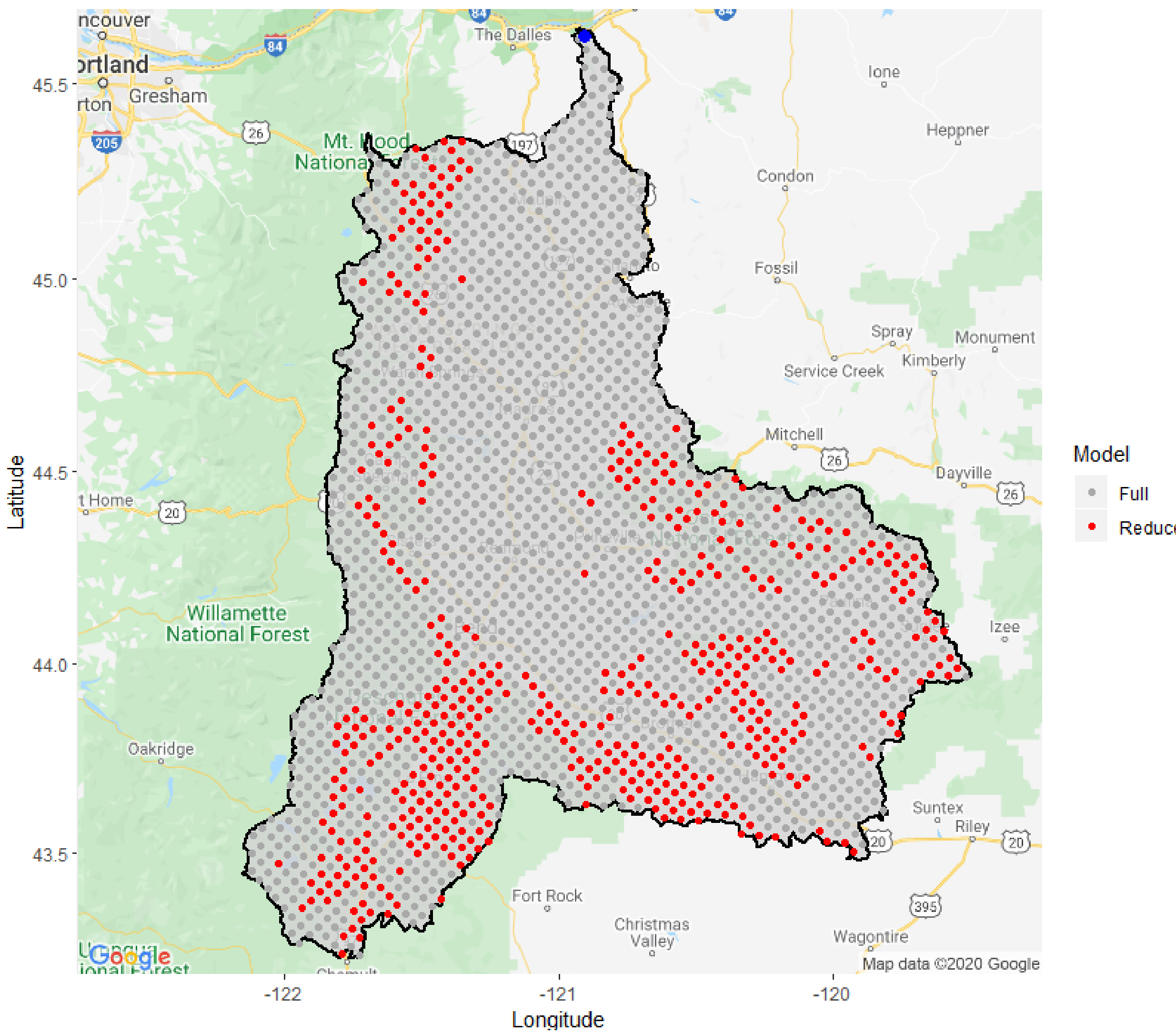

Figure B3. The 4-kilometer resolution IMS sample locations of the Deschutes subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 2004 - 2018. The blue dot indicates the location of the streamflow gage for the station at Moody (14103000).

Table B3

The Full PSC March IMS Model for the 4-kilometer resolution (top), the Reduced PSC March IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 4-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Deschutes subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

DESCHUTES SUBBASIN

Full March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	17165.09	1093.15	15.702	2.30e-09 ***
deschutes.imsfull	197.15	30.82	6.398	3.41e-05 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1661 on 12 degrees of freedom
Multiple R-squared: 0.7733, Adjusted R-squared: 0.7544
Predicted R-squared: 0.7103
F-statistic: 40.93 on 1 and 12 DF, p-value: 3.415e-05

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsfull))

	Value	p-value	Decision
Global Stat	2.8144	0.5893	Assumptions acceptable.
Skewness	0.2569	0.6122	Assumptions acceptable.
Kurtosis	0.4245	0.5147	Assumptions acceptable.
Link Function	2.1321	0.1442	Assumptions acceptable.
Heteroscedasticity	0.001	0.9754	Assumptions acceptable.

Reduced March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	19030.25	810.10	23.491	2.12e-11 ***
deschutes.imsred	120.37	18.21	6.609	2.51e-05 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1619 on 12 degrees of freedom
Multiple R-squared: 0.7845, Adjusted R-squared: 0.7665
Predicted R-squared: 0.7199
F-statistic: 43.68 on 1 and 12 DF, p-value: 2.505e-05

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred))

	Value	p-value	Decision
Global Stat	1.539	0.8197	Assumptions accepted.
Skewness	0.604	0.437	Assumptions accepted.
Kurtosis	0.5497	0.4584	Assumptions accepted.
Link Function	0.3761	0.5397	Assumptions accepted.
Heteroscedasticity	0.0091	0.924	Assumptions accepted.

Reduced March IMS Model (4-km) w/ SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	20063.01	1340.84	14.963	1.17e-08 ***
deschutes.imsred	145.96	32.12	4.544	0.000839 ***
deschutes.snotel	-105.73	109.22	-0.968	0.353813

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1624 on 11 degrees of freedom
Multiple R-squared: 0.8014, Adjusted R-squared: 0.7653
Predicted R-squared: 0.6468
F-statistic: 22.19 on 2 and 11 DF, p-value: 0.0001377
VIF: 5.04

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred + deschutes.snotel))

	Value	p-value	Decision
Global Stat	1.8416	0.7649	Assumptions accepted.
Skewness	0.6353	0.4254	Assumptions accepted.
Kurtosis	0.4308	0.5116	Assumptions accepted.
Link Function	0.2498	0.6172	Assumptions accepted.
Heteroscedasticity	0.5257	0.4584	Assumptions accepted.

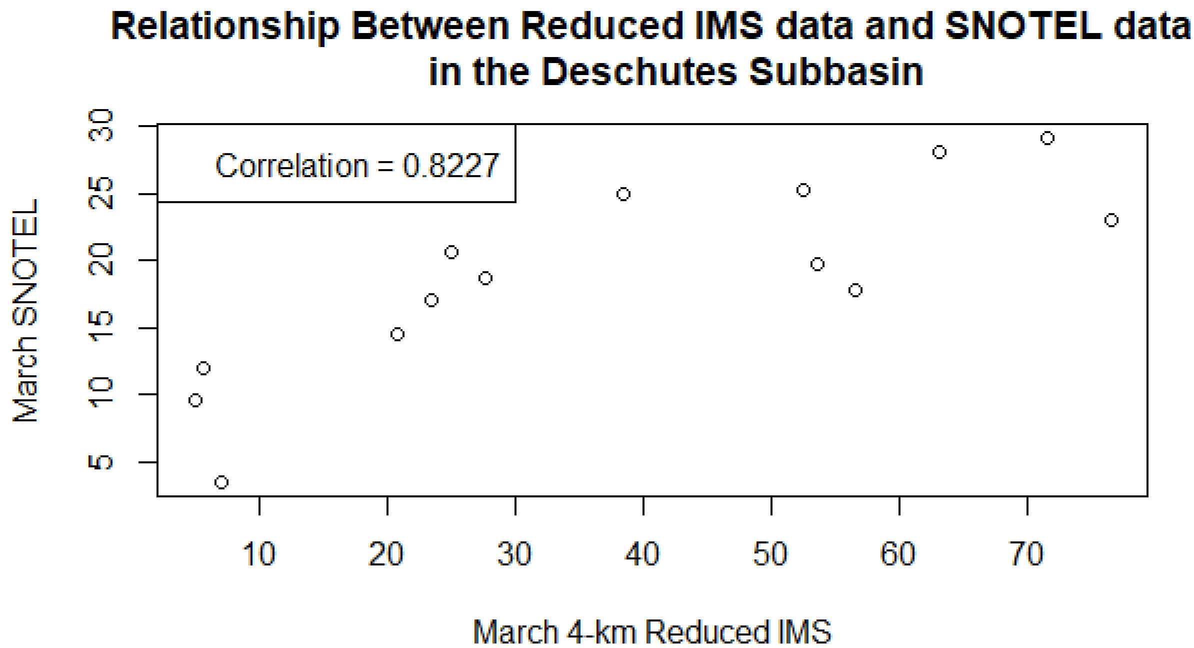

Figure B4. This correlation plot shows the correlation between the March 4-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Deschutes subbasin.

Table B4

Principal Component Analysis performed for the Deschutes subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (4-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Deschutes Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 14 2
Y dimension: 14 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 14 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	3478	1747	1919
adjCV	3478	1739	1901

TRAINING: % variance explained

	1 comps	2 comps
X	97.57	100.00
deschutes.stream	77.43	80.14

Coefficients:

1 component -

(Intercept)	18789.6
deschutes.imsred	122.7167
deschutes.snotel	28.012

Multiple R-squared: 0.7743
Predicted R-squared: 0.7073

2 components -

(Intercept)	20063.01
deschutes.imsred	145.955
deschutes.snotel	-105.7348

Multiple R-squared: 0.8014
Predicted R-squared: 0.6468

APPENDIX C

John Day Subbasin

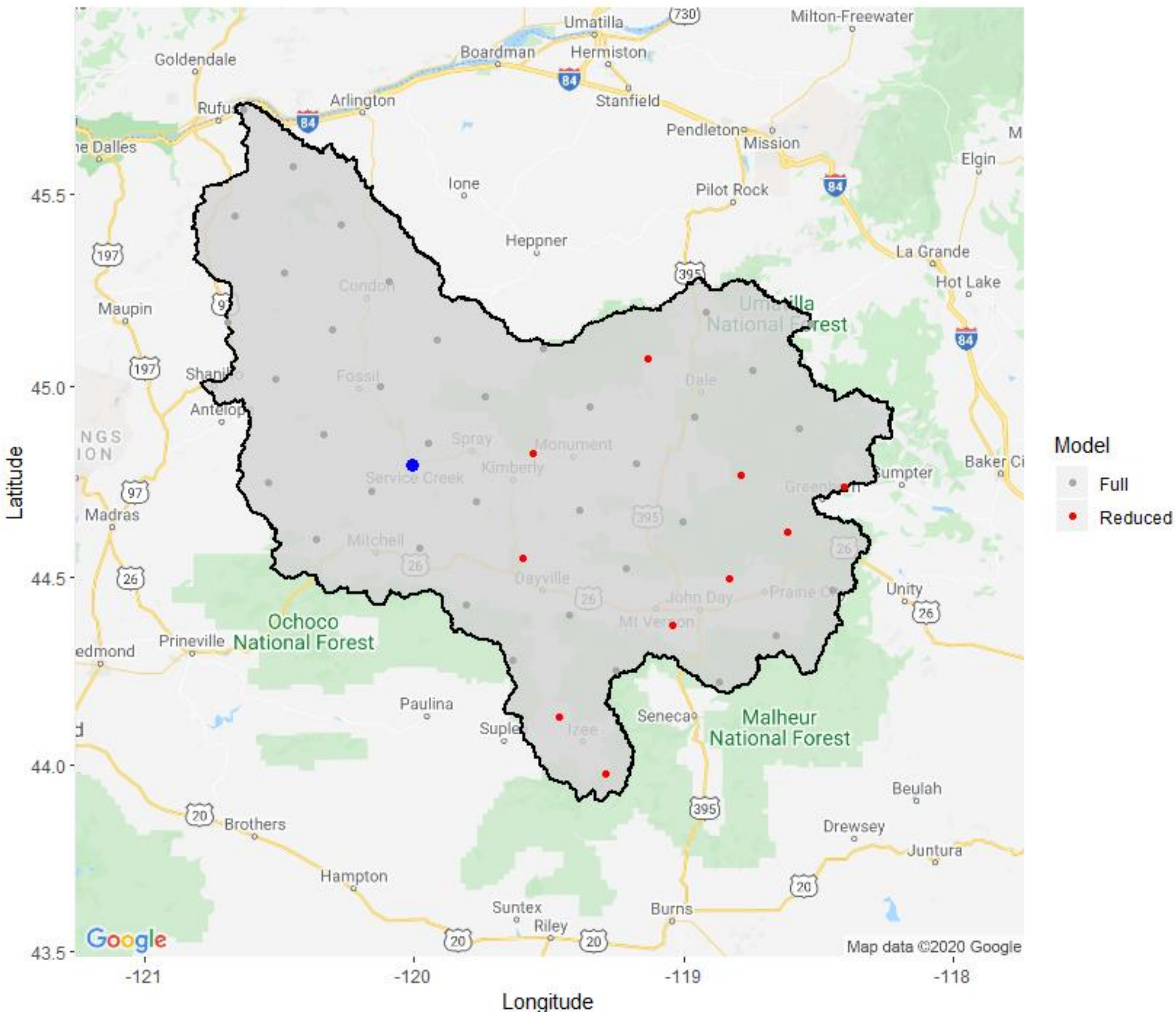

Figure C1. The 24-kilometer resolution IMS sample locations of the John Day subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 1999 - 2018. The blue dot indicates the location of the streamflow gage for the station at Service Creek (14046500).

Table C1

The Full PSC March IMS Model for the 24-kilometer resolution (top), the Reduced PSC March IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 24-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the John Day subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

JOHN DAY SUBBASIN

Full March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6552.25	2515.63	2.605	0.0185 *
johnday.imsfull	228.58	80.19	2.851	0.0111 *

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5217 on 17 degrees of freedom
Multiple R-squared: 0.3234, Adjusted R-squared: 0.2836
Predicted R-squared: 0.1709
F-statistic: 8.126 on 1 and 17 DF, p-value: 0.01106

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsfull))

	Value	p-value	Decision
Global Stat	11.14	0.025	Assumptions NOT satisifed!
Skewness	4.9958	0.0254	Assumptions NOT satisifed!
Kurtosis	5.0753	0.0243	Assumptions NOT satisifed!
Link Function	0.9619	0.3267	Assumptions acceptable.
Heteroscedasticity	0.1069	0.7437	Assumptions acceptable.

Reduced March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5719.63	2347.77	2.436	0.02614 *
johnday.imsred	145.03	41.97	3.456	0.00302 **

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4861 on 17 degrees of freedom
Multiple R-squared: 0.4126, Adjusted R-squared: 0.3781
Predicted R-squared: 0.2567
F-statistic: 11.94 on 1 and 17 DF, p-value: 0.003019

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred))

	Value	p-value	Descision
Global Stat	14.672	0.0054	Assumptions NOT satisfied!
Skewness	5.328	0.021	Assumptions NOT satisfied!
Kurtosis	4.525	0.0334	Assumptions NOT satisfied!
Link Function	4.508	0.0337	Assumptions NOT satisfied!
Heteroscedasticity	0.312	0.5765	Assumptions acceptable.

Reduced March IMS Model (24-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3981.87	3294.76	1.209	0.244
johnday.imsred	103.37	69.26	1.493	0.155
johnday.snotel	357.32	469.06	0.762	0.457

Residual standard error: 4922 on 16 degrees of freedom
Multiple R-squared: 0.4332, Adjusted R-squared: 0.3623
Predicted R-squared: 0.2354
F-statistic: 6.114 on 2 and 16 DF, p-value: 0.01065
VIF: 1.76

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred + johnday.snotel))

	Value	p-value	Descision
Global Stat	21.625	0.0002	Assumptions NOT satisfied!
Skewness	8.6795	0.0032	Assumptions NOT satisfied!
Kurtosis	9.1619	0.0657	Assumptions NOT satisfied!
Link Function	3.3864	0.0657	Assumptions acceptable.
Heteroscedasticity	0.3968	0.5288	Assumptions acceptable.

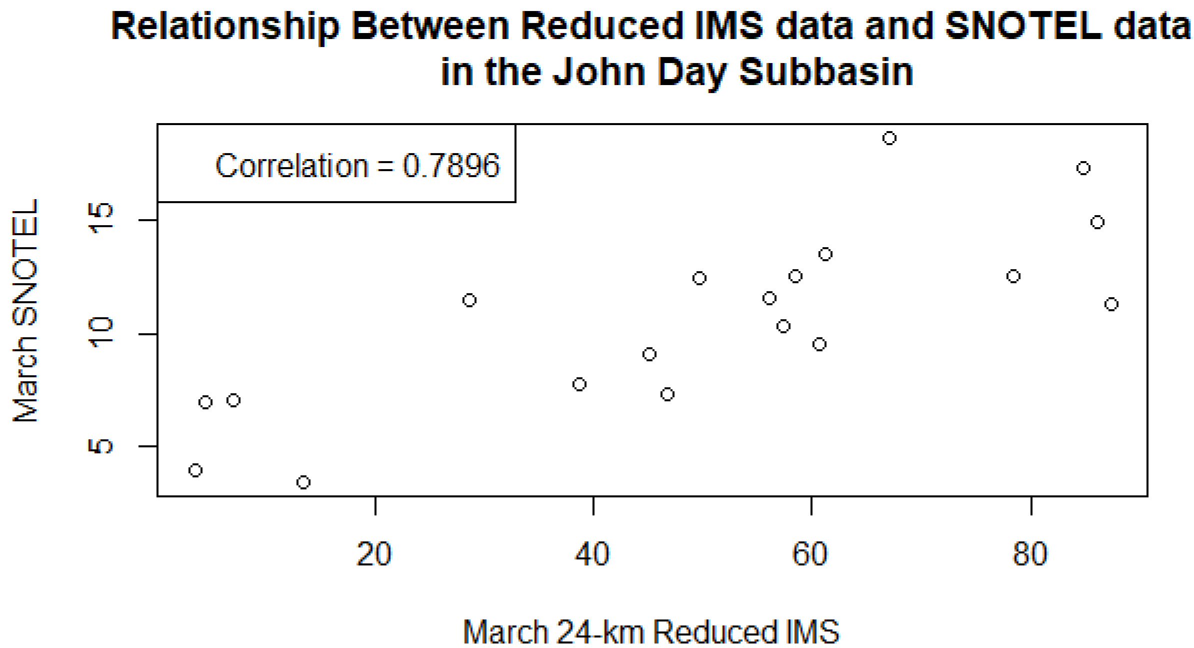

Figure C2. This correlation plot shows the correlation between the March 24-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the John Day subbasin.

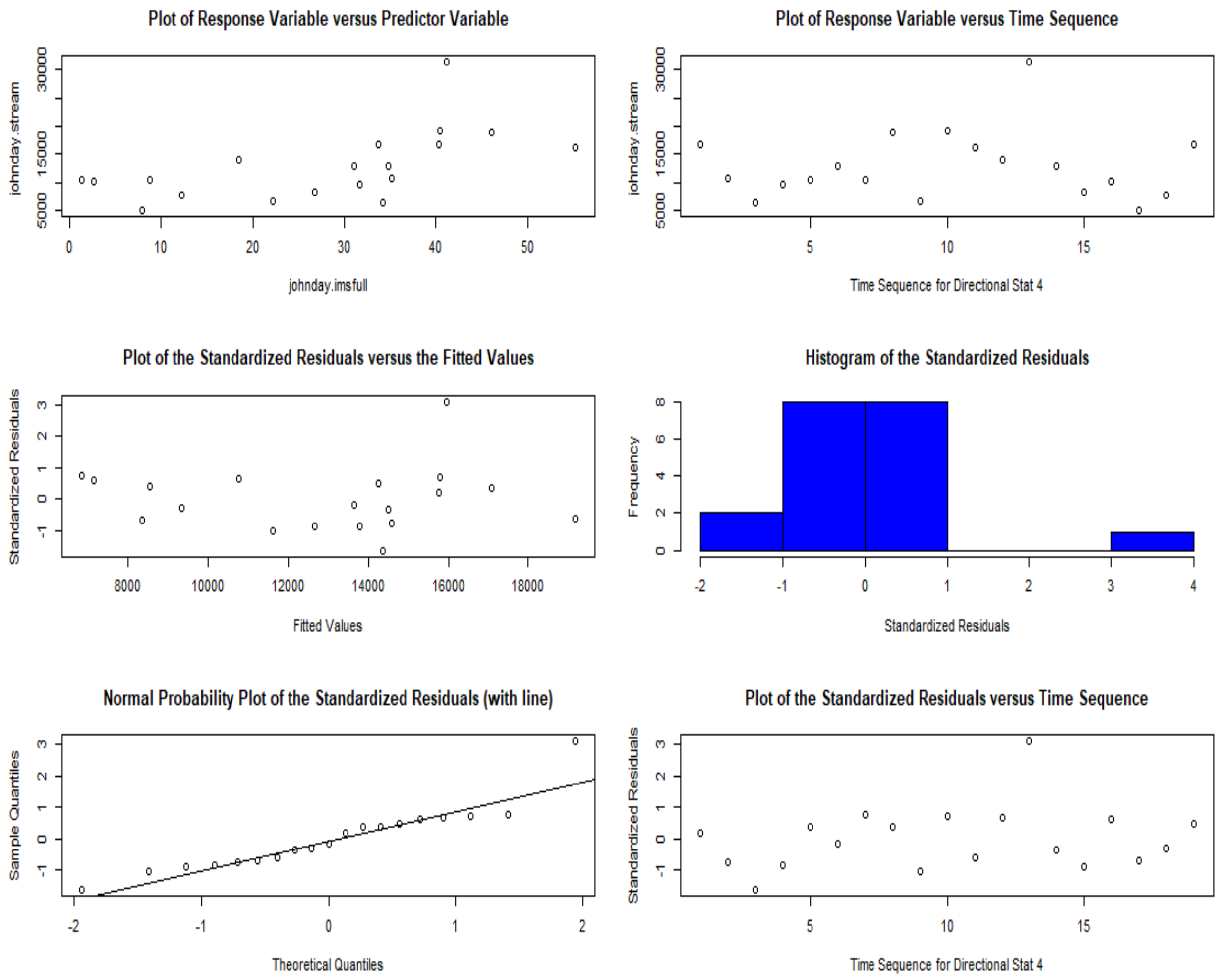

Figure C3. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Full March IMS Model (24-km).

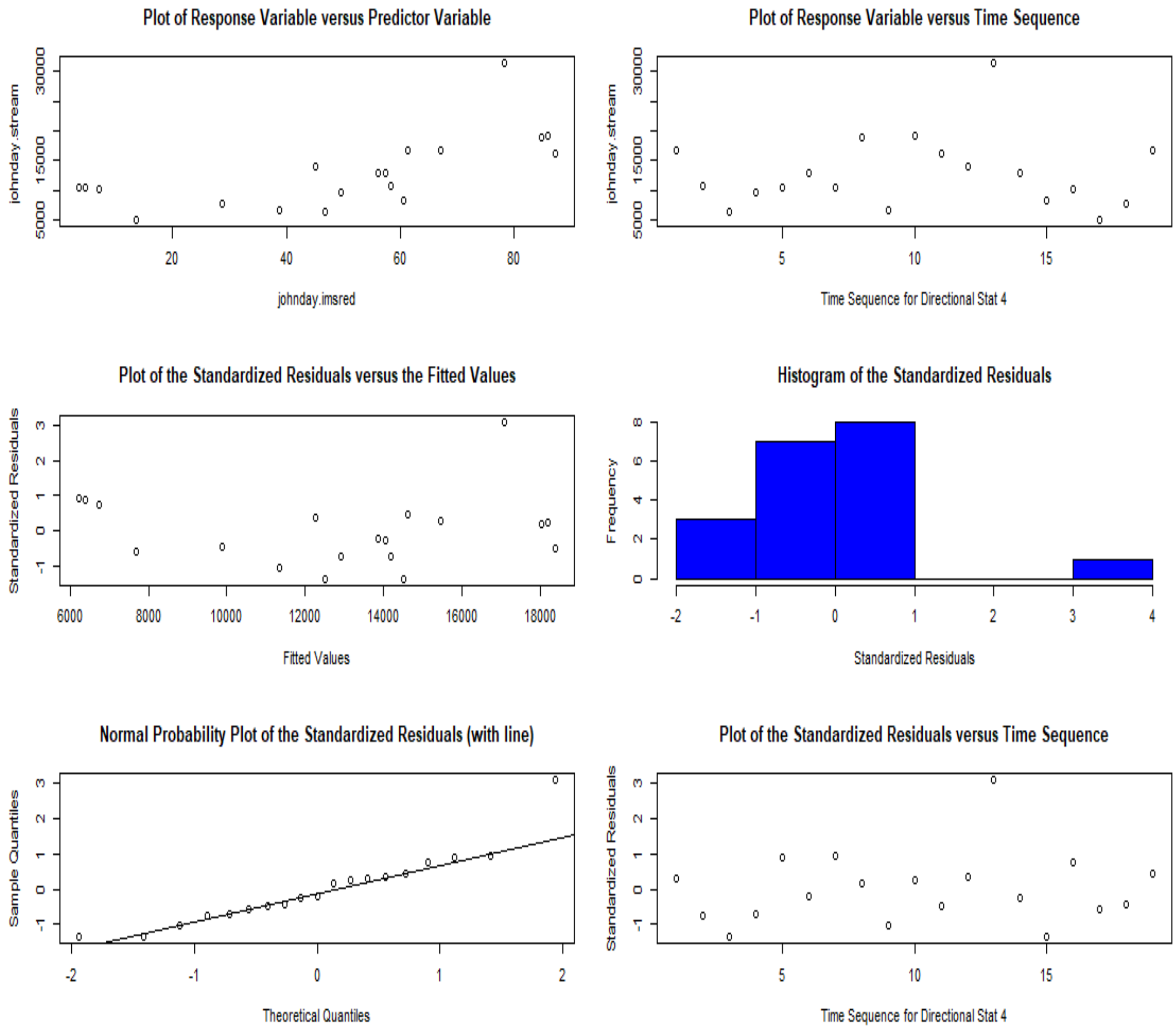

Figure C4. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day Subbasin Reduced March IMS Model (24-km).

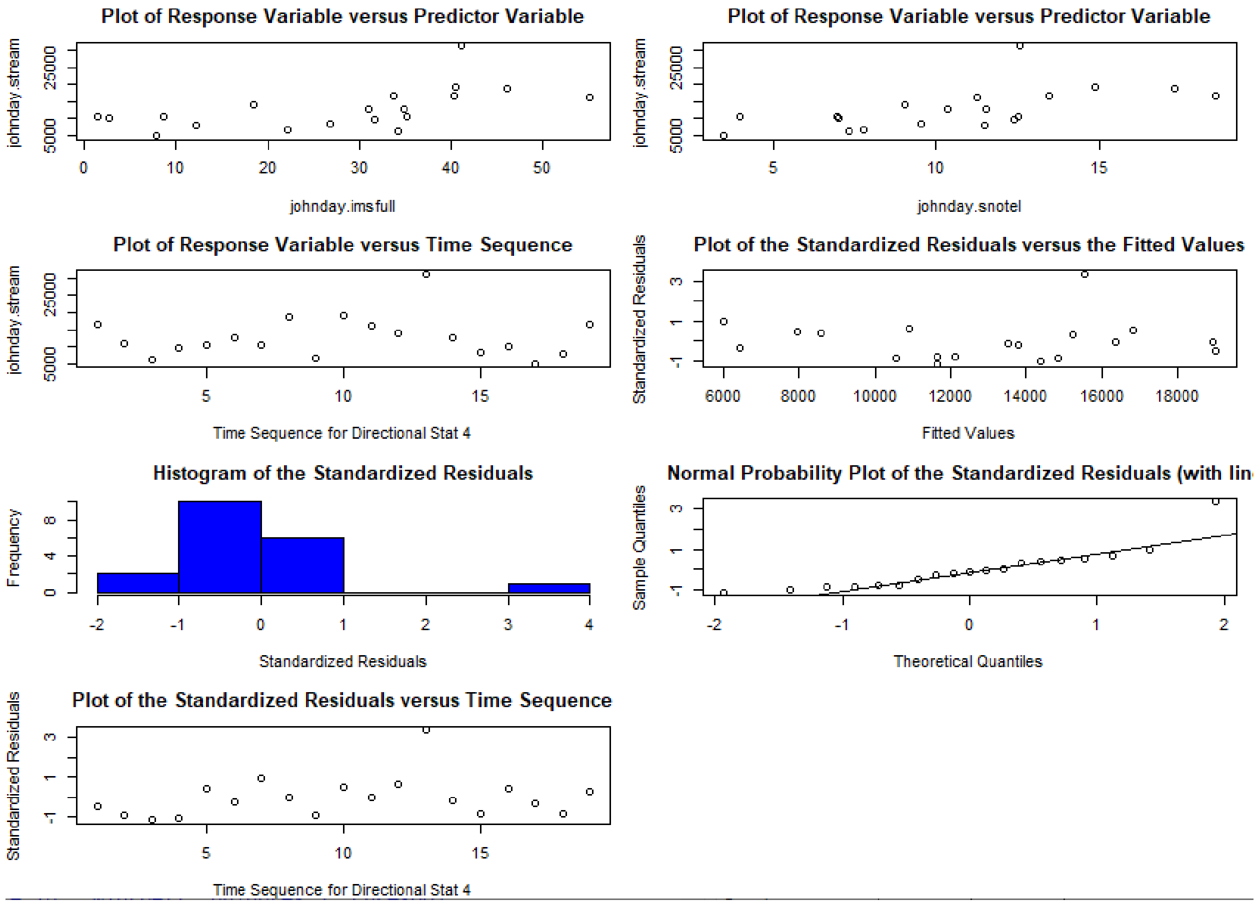

Figure C5. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Full March IMS Model (24-km) with March Snow Telemetry data (SNOTEL) added as a second predictor.

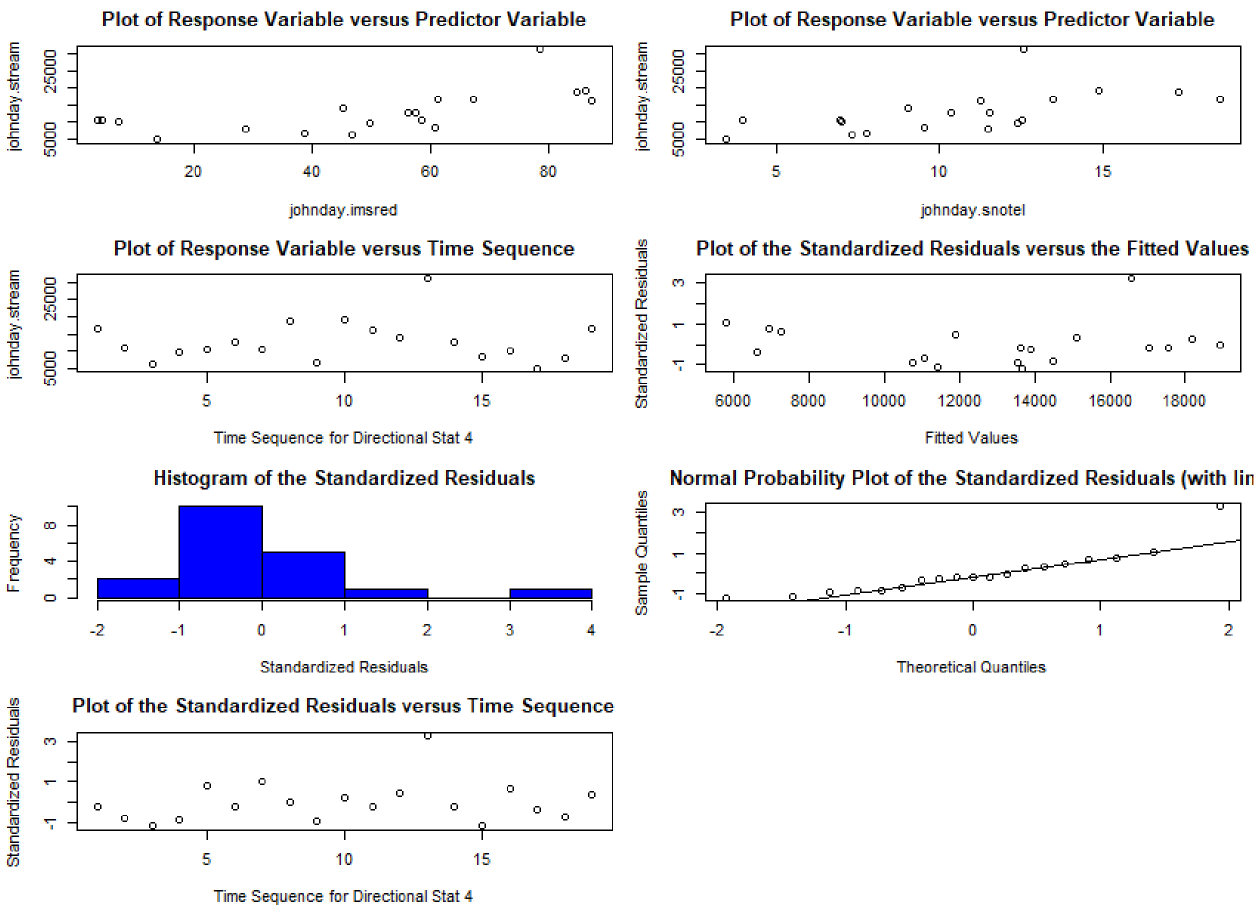

Figure C6. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Reduced March IMS Model (24-km) with March Snow Telemetry data (SNOTEL) added as a second predictor.

Table C2

Principal Component Analysis performed for the John Day subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (24-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

John Day Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	6332	5162	5255
adjCV	6332	5147	5236

TRAINING: % variance explained

	1 comps	2 comps
X	99.21	100.00
johnday.stream	41.45	43.32

Coefficients:

1 component -

(Intercept)	5621.712
johnday.imsred	143.39088
johnday.snotel	16.85442

Multiple R-squared: 0.4145
Predicted R-squared: 0.2596

2 components -

(Intercept)	3981.866
johnday.imsred	103.3721
johnday.snotel	357.3190

Multiple R-squared: 0.4332
Predicted R-squared: 0.2325

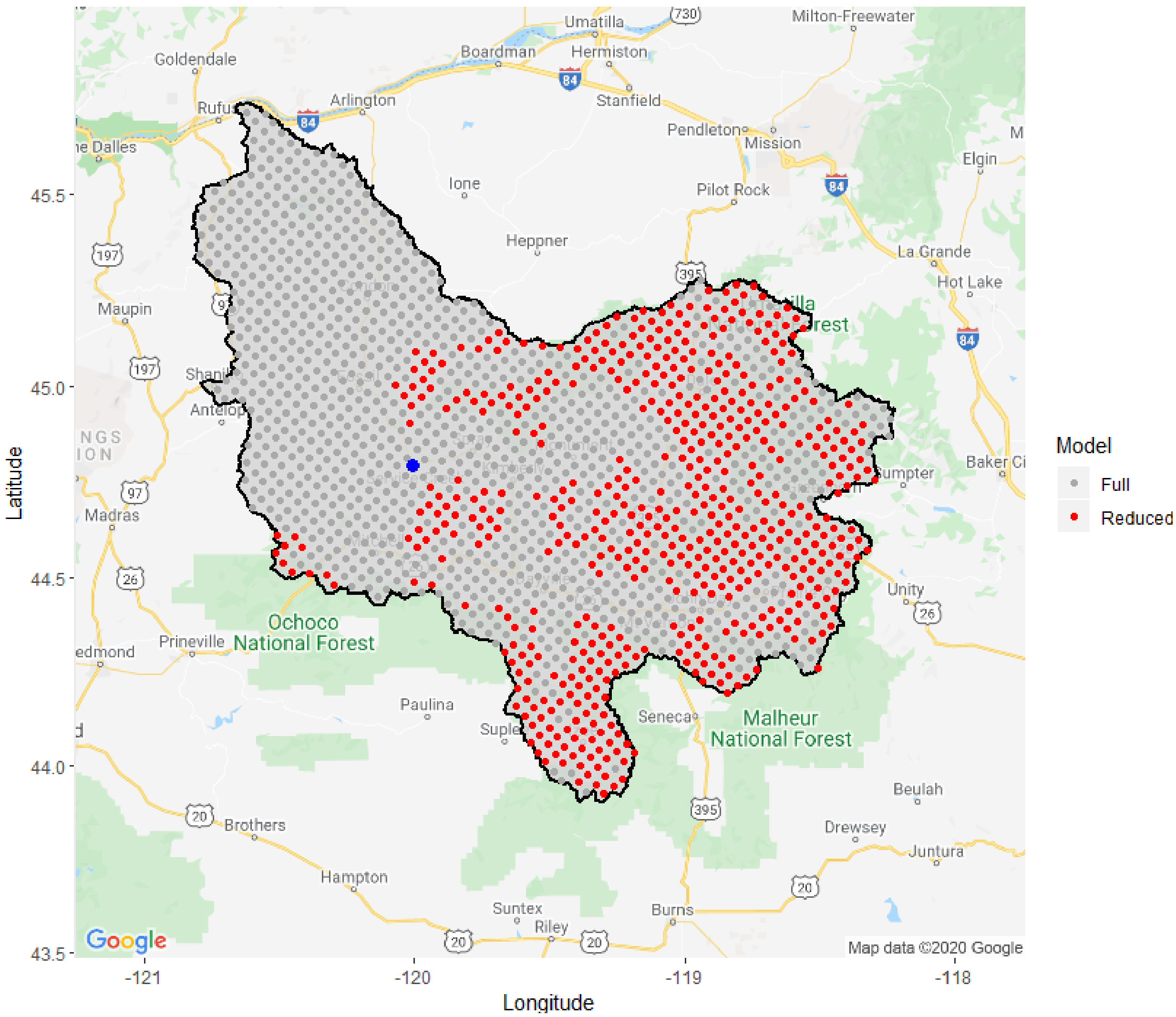

Figure C7. The 4-kilometer resolution IMS sample locations of the John Day subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 2004 - 2018. The blue dot indicates the location of the streamflow gage for the station at Service Creek (14046500).

Table C3

The Full PSC March IMS Model for the 4-kilometer resolution (top), the Reduced PSC March IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 4-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the John Day subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

JOHN DAY SUBBASIN

Full March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	4302.57	3326.36	1.293	0.22019	
johnday.imsfull	275.99	89.44	3.086	0.00944	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5269 on 12 degrees of freedom
Multiple R-squared: 0.4424, Adjusted R-squared: 0.396
Predicted R-squared: 0.2271
F-statistic: 9.522 on 1 and 12 DF, p-value: 0.009436

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsfull))

	Value	p-value	Decision
Global Stat	6.3361	0.1754	Assumptions acceptable.
Skewness	4.1835	0.0408	Assumptions NOT satisfied!
Kurtosis	1.9844	0.1589	Assumptions acceptable.
Link Function	0.1641	0.6854	Assumptions acceptable.
Heteroscedasticity	0.0042	0.9484	Assumptions acceptable.

Reduced March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	4748.10	3094.27	1.534	0.15084	
johnday.imsred	172.31	53.85	3.200	0.00763	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5184 on 12 degrees of freedom
Multiple R-squared: 0.4604, Adjusted R-squared: 0.4155
Predicted R-squared: 0.2734
F-statistic: 10.24 on 1 and 12 DF, p-value: 0.007634

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred))

	Value	p-value	Decision
Global Stat	8.8443	0.0651	Assumptions acceptable.
Skewness	4.7815	0.0288	Assumptions NOT satisfied!
Kurtosis	2.7257	0.0988	Assumptions acceptable.
Link Function	1.2694	0.2599	Assumptions acceptable.
Heteroscedasticity	0.0678	0.7945	Assumptions acceptable.

Reduced March IMS Model (4-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3430.6	4337.8	0.791	0.446
johnday.imsred	128.1	112.9	1.135	0.281
johnday.snotel	349.8	776.9	0.450	0.661

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5365 on 11 degrees of freedom
Multiple R-squared: 0.4702, Adjusted R-squared: 0.3739
Predicted R-squared: 0.2218
F-statistic: 4.881 on 2 and 11 DF, p-value: 0.03039
VIF: 1.89

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred + johnday.snotel))

	Value	p-value	Decision
Global Stat	11.352	0.0229	Assumptions NOT satisfied!
Skewness	6.0095	0.0142	Assumptions NOT satisfied!
Kurtosis	4.0247	0.0448	Assumptions NOT satisfied!
Link Function	1.2583	0.262	Assumptions acceptable.
Heteroscedasticity	0.0597	0.807	Assumptions acceptable.

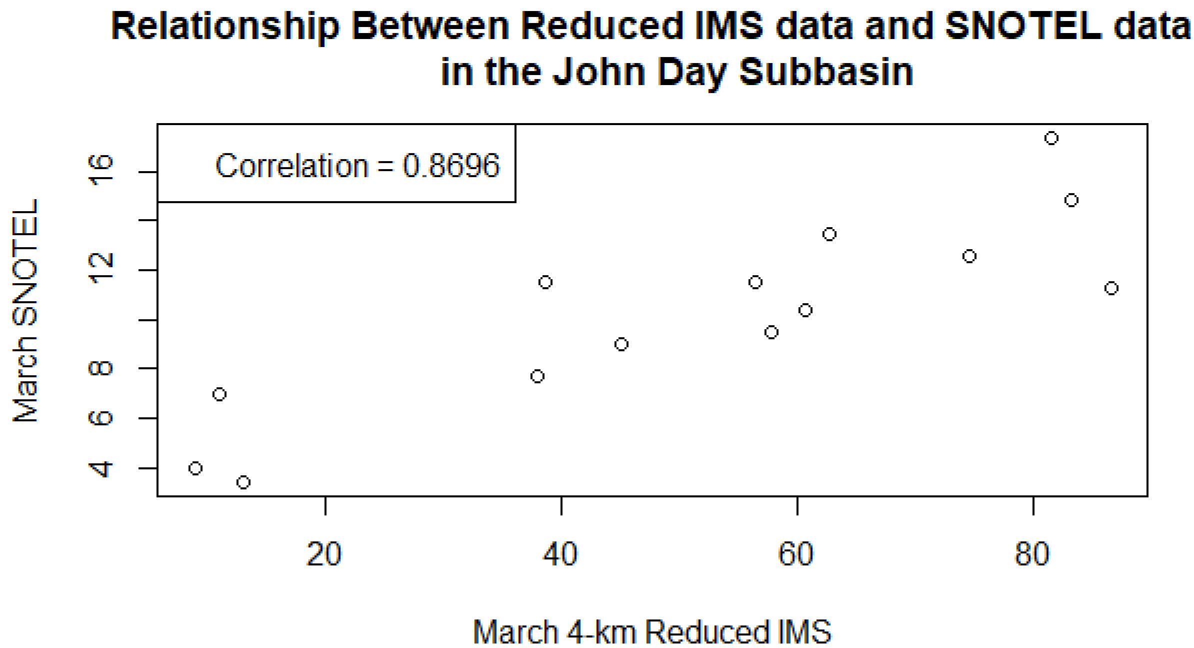

Figure C8. This correlation plot shows the correlation between the March 4-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the John Day subbasin.

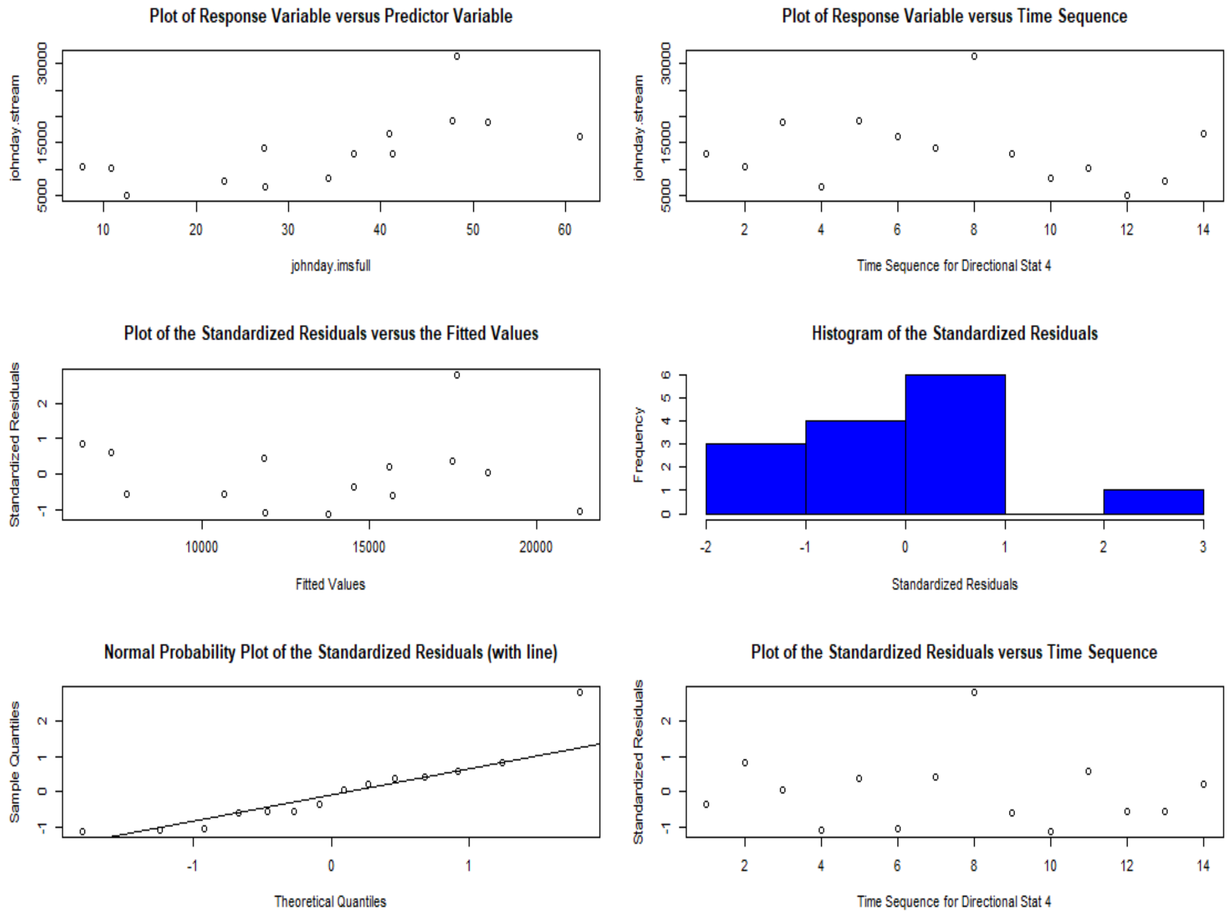

Figure C9. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Full March IMS Model (4-km).

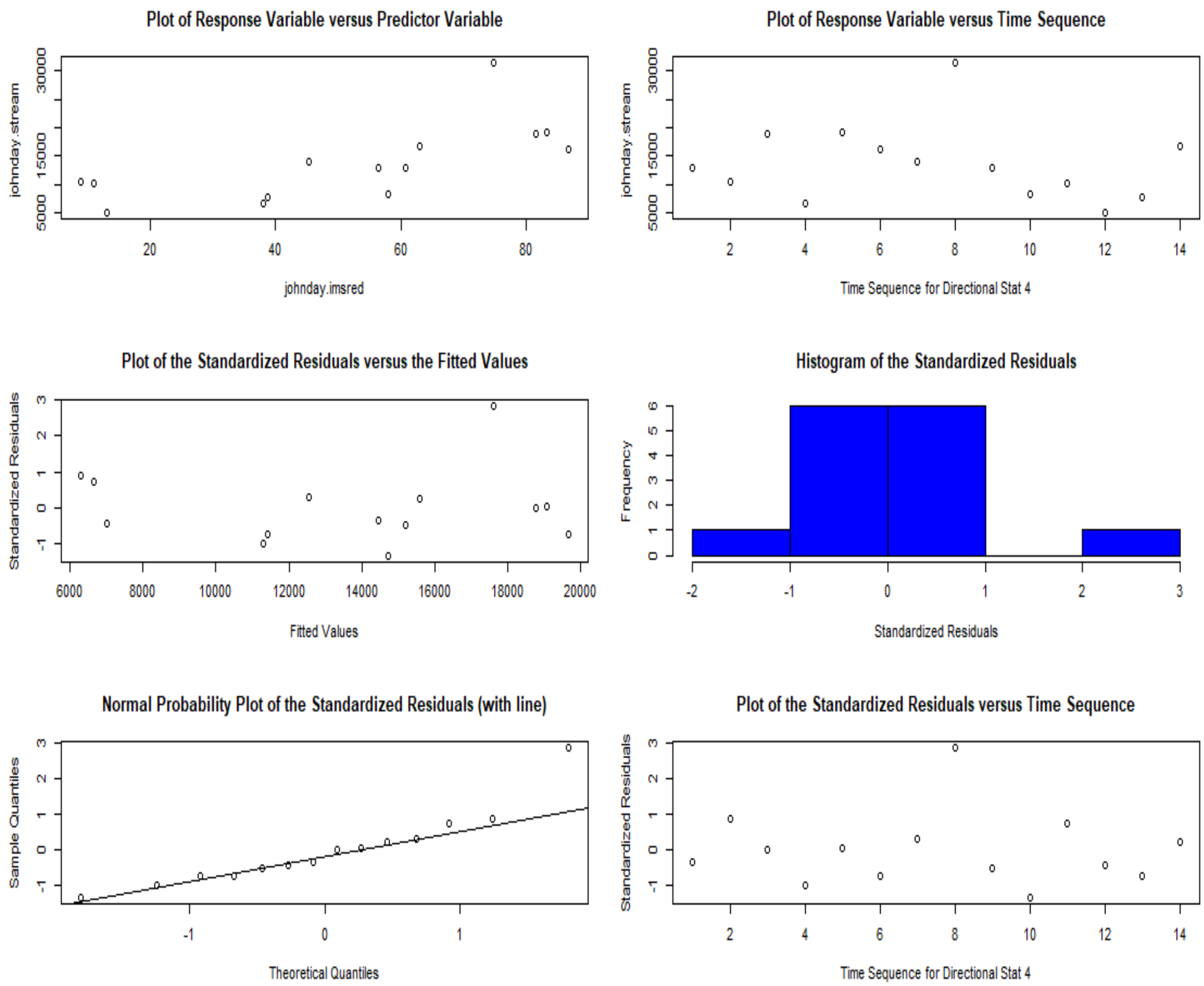

Figure C10. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Reduced March IMS Model (4-km).

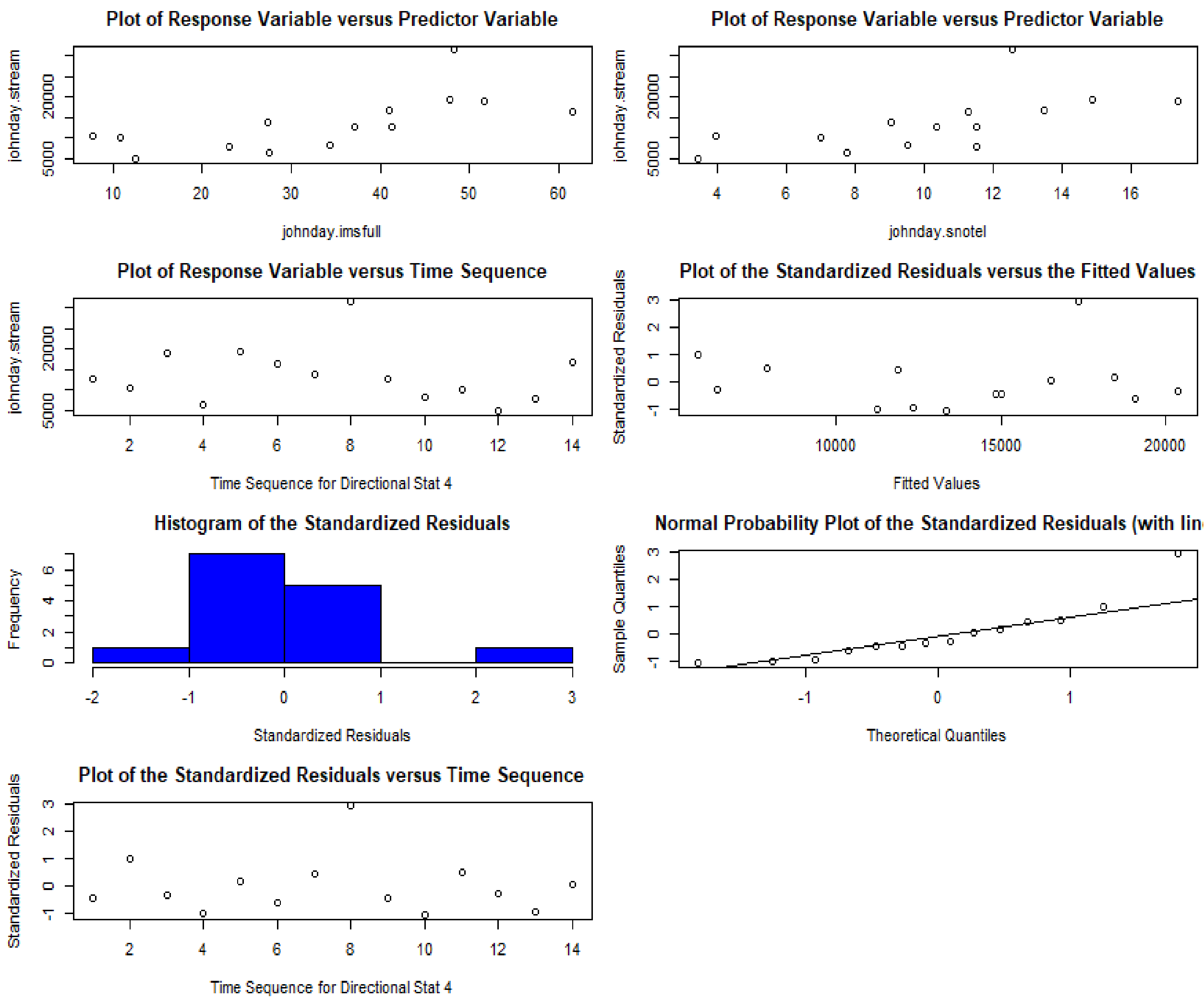

Figure C11. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Full March IMS Model (4-km) with March Snow Telemetry data (SNOTEL) added as a second predictor.

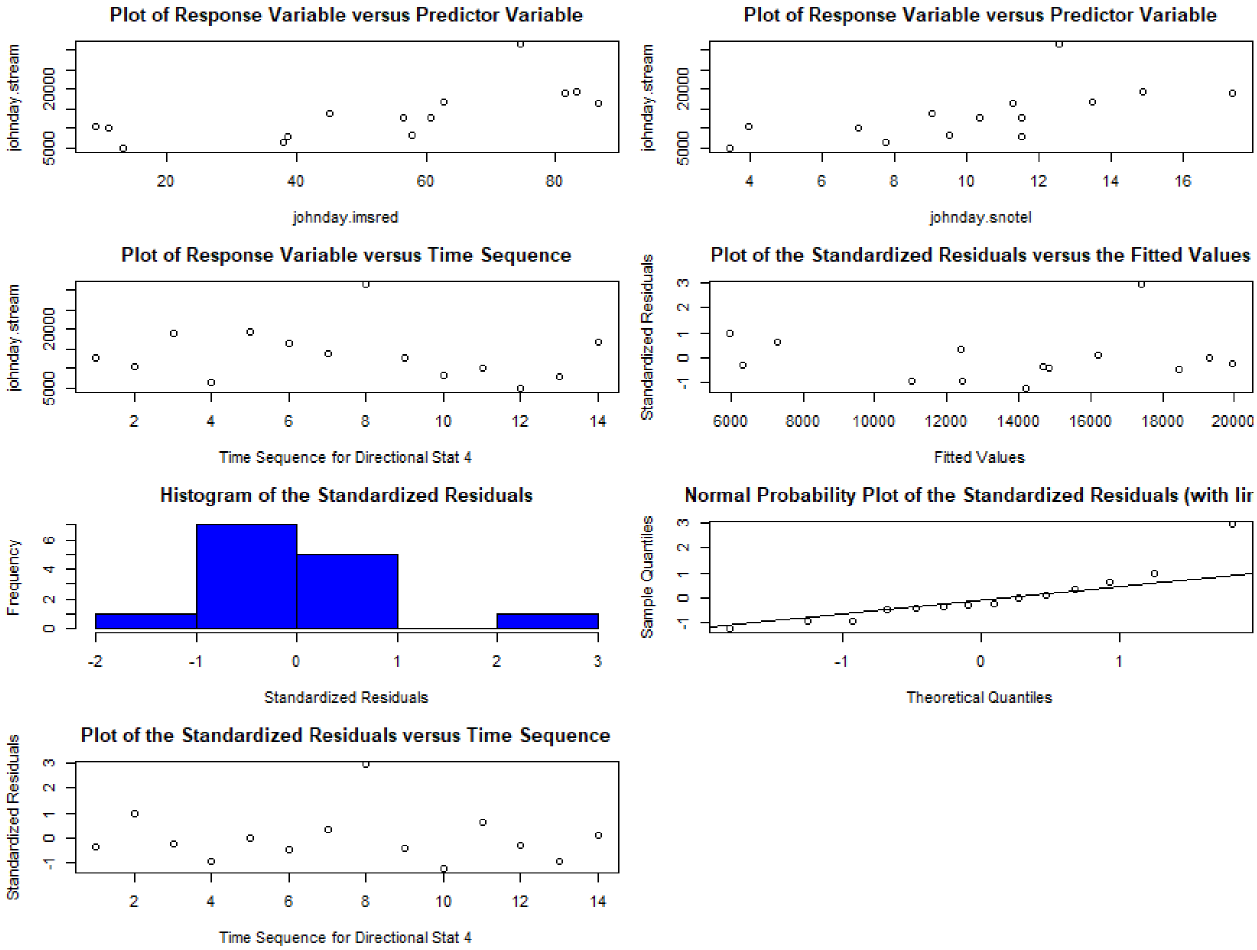

Figure C12. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Reduced March IMS Model (4-km) with March Snow Telemetry data (SNOTEL) added as a second predictor.

Table C4

Principal Component Analysis performed for the John Day subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (4-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

John Day Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 14 2
Y dimension: 14 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 14 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	7036	5559	5764
adjCV	7036	5531	5727

TRAINING: % variance explained

	1 comps	2 comps
X	99.50	100.00
johnday.stream	46.16	47.02

Coefficients:

1 component -

(Intercept)	4656.036
johnday.imsred	169.7982
johnday.snotel	21.5638

Multiple R-squared: 0.4616
Predicted R-squared: 0.2761

2 components -

(Intercept)	3430.588
johnday.imsred	128.1115
johnday.snotel	349.8138

Multiple R-squared: 0.4702
Predicted R-squared: 0.2218

APPENDIX D

Clearwater Subbasin

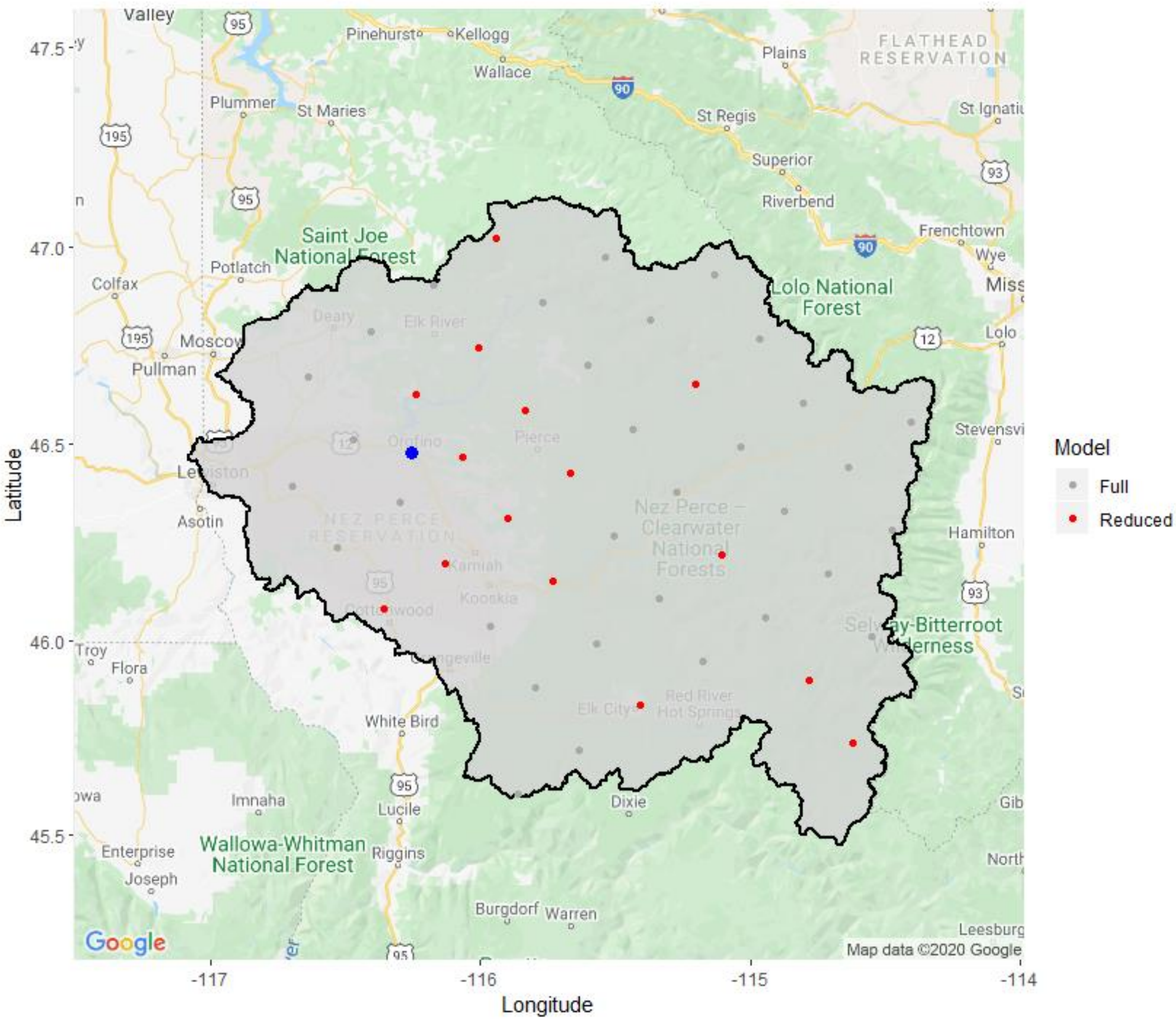

Figure D1. The 24-kilometer resolution IMS sample locations of the Clearwater subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 1999 - 2018. The blue dot indicates the location of the streamflow gage for the station at the Orofino station (13340000).

Table D1

The Full PSC March IMS Model for the 24-kilometer resolution (top), the Reduced PSC March IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 24-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Clearwater subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

CLEARWATER SUBBASIN

Full March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	41369.7	10535.0	3.927	0.00109 **
clearwater.imsfull	477.0	158.6	3.008	0.00791 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15210 on 17 degrees of freedom

Multiple R-squared: 0.3474, Adjusted R-squared: 0.309

Predicted R-squared: 0.2245

F-statistic: 9.051 on 1 and 17 DF, p-value: 0.007911

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS

USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:

Level of Significance = 0.05

Call:

gvlma(x = lm(clearwater.stream ~ clearwater.imsfull))

	Value	p-value	Decision
Global Stat	1.2327	0.8727	Assumptions acceptable.
Skewness	0.8967	0.3437	Assumptions acceptable.
Kurtosis	0.1951	0.6587	Assumptions acceptable.
Link Function	0.001	0.9753	Assumptions acceptable.
Heteroscedasticity	0.14	0.7083	Assumptions acceptable.

Reduced March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	49847.9	7368.5	6.765	3.3e-06 ***
clearwater.imsred	386.1	117.9	3.273	0.00448 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14750 on 17 degrees of freedom

Multiple R-squared: 0.3866, Adjusted R-squared: 0.3505

Predicted R-squared: 0.2585

F-statistic: 10.71 on 1 and 17 DF, p-value: 0.00448

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS

USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:

Level of Significance = 0.05

Call:

gvlma(x = lm(clearwater.stream ~ clearwater.imsred))

	Value	p-value	Decision
Global Stat	1.3171	0.8585	Assumptions acceptable.
Skewness	0.9609	0.327	Assumptions acceptable.
Kurtosis	0.2133	0.6442	Assumptions acceptable.
Link Function	0.045	0.832	Assumptions acceptable.
Heteroscedasticity	0.098	0.7543	Assumptions acceptable.

Reduced March IMS Model (24-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	27703.7	11503.0	2.408	0.0284 *
clearwater.imsred	174.9	138.3	1.264	0.2242
clearwater.snotel	1116.8	476.8	2.343	0.0324 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13120 on 16 degrees of freedom

Multiple R-squared: 0.5433, Adjusted R-squared: 0.4862

Predicted R-squared: 0.3675

F-statistic: 9.515 on 2 and 16 DF, p-value: 0.001894

VIF: 2.19

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS

USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:

Level of Significance = 0.05

Call:

gvlma(x = lm(clearwater.stream ~ clearwater.imsred + clearwater.snotel))

	Value	p-value	Decision
Global Stat	6.1312	0.1896	Assumptions acceptable.
Skewness	4.0783	0.0434	Assumptions NOT satisifed!
Kurtosis	1.2319	0.267	Assumptions acceptable.
Link Function	0.3776	0.5389	Assumptions acceptable.
Heteroscedasticity	0.4434	0.5055	Assumptions acceptable.

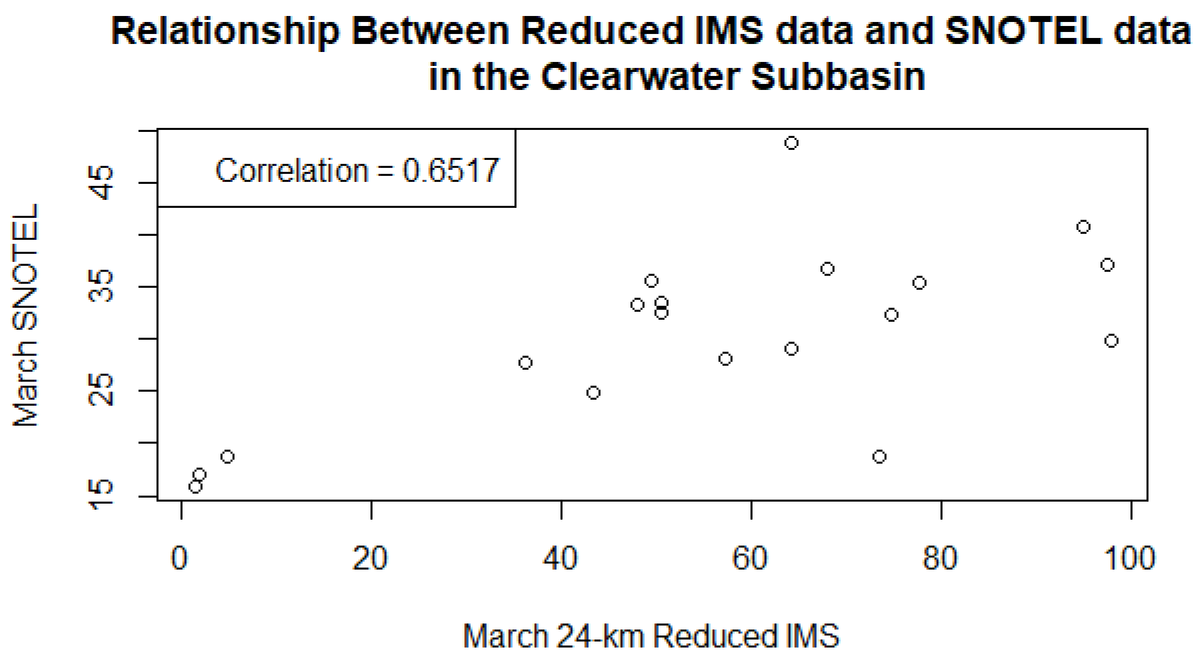

Figure D2. This correlation plot shows the correlation between the March 24-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Clearwater subbasin.

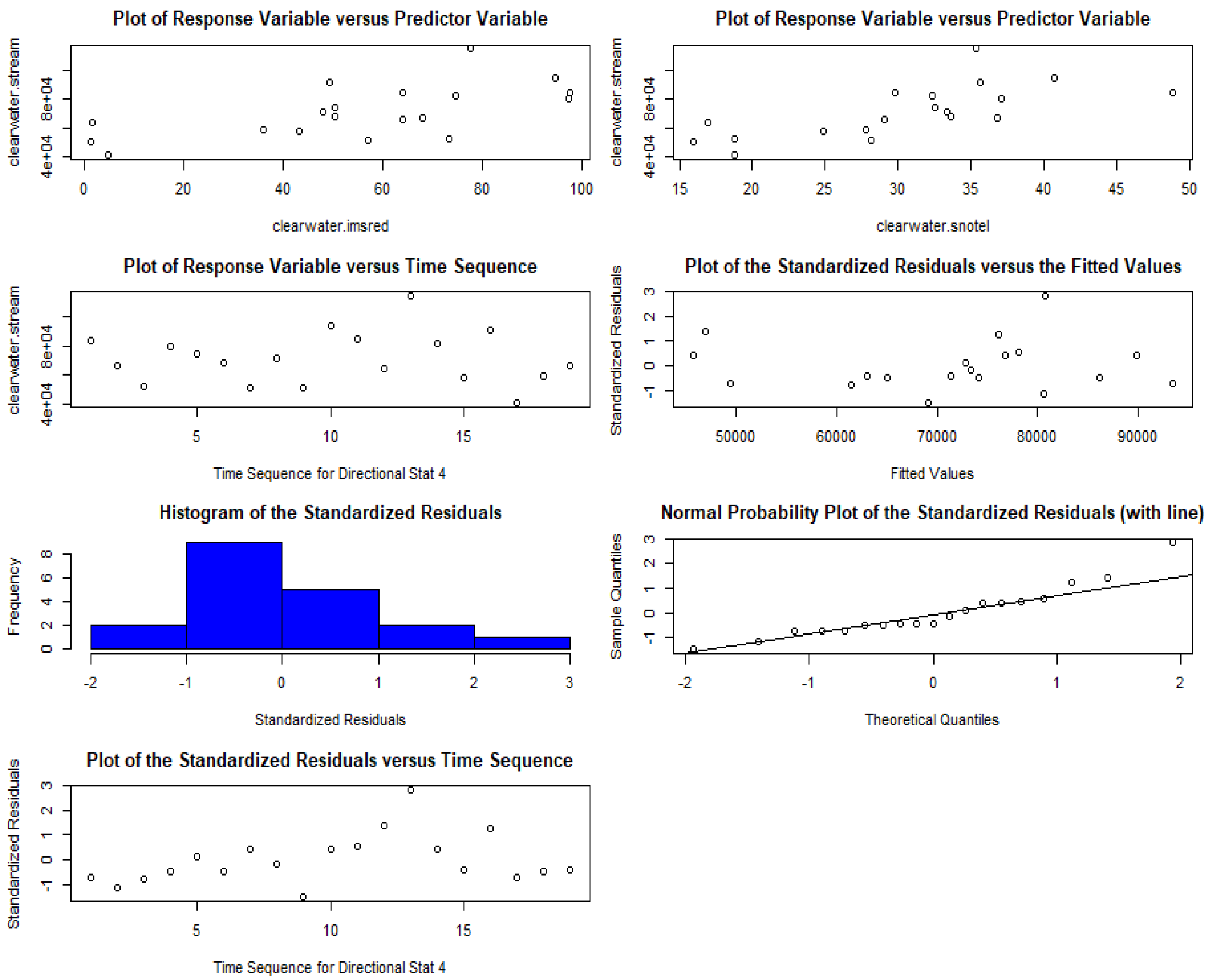

Figure D3. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the Clearwater subbasin Reduced March IMS Model (4-km) with March Snow Telemetry data (SNOTEL) added as a second predictor.

Table D2

Principal Component Analysis performed for the Clearwater subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (24-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Clearwater Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	18804	15090	14168
adjCV	18804	15052	14109

TRAINING: % variance explained

	1 comps	2 comps
X	95.70	100.00
clearwater.stream	40.69	54.33

Coefficients:

1 component -

(Intercept)	47813.19
clearwater.imsred	381.39532
clearwater.snotel	75.64407

Multiple R-squared: 0.4069
Predicted R-squared: 0.2825

2 components -

(Intercept)	27703.65
clearwater.imsred	174.8926
clearwater.snotel	1116.8250

Multiple R-squared: 0.5433
Predicted R-squared: 0.3675

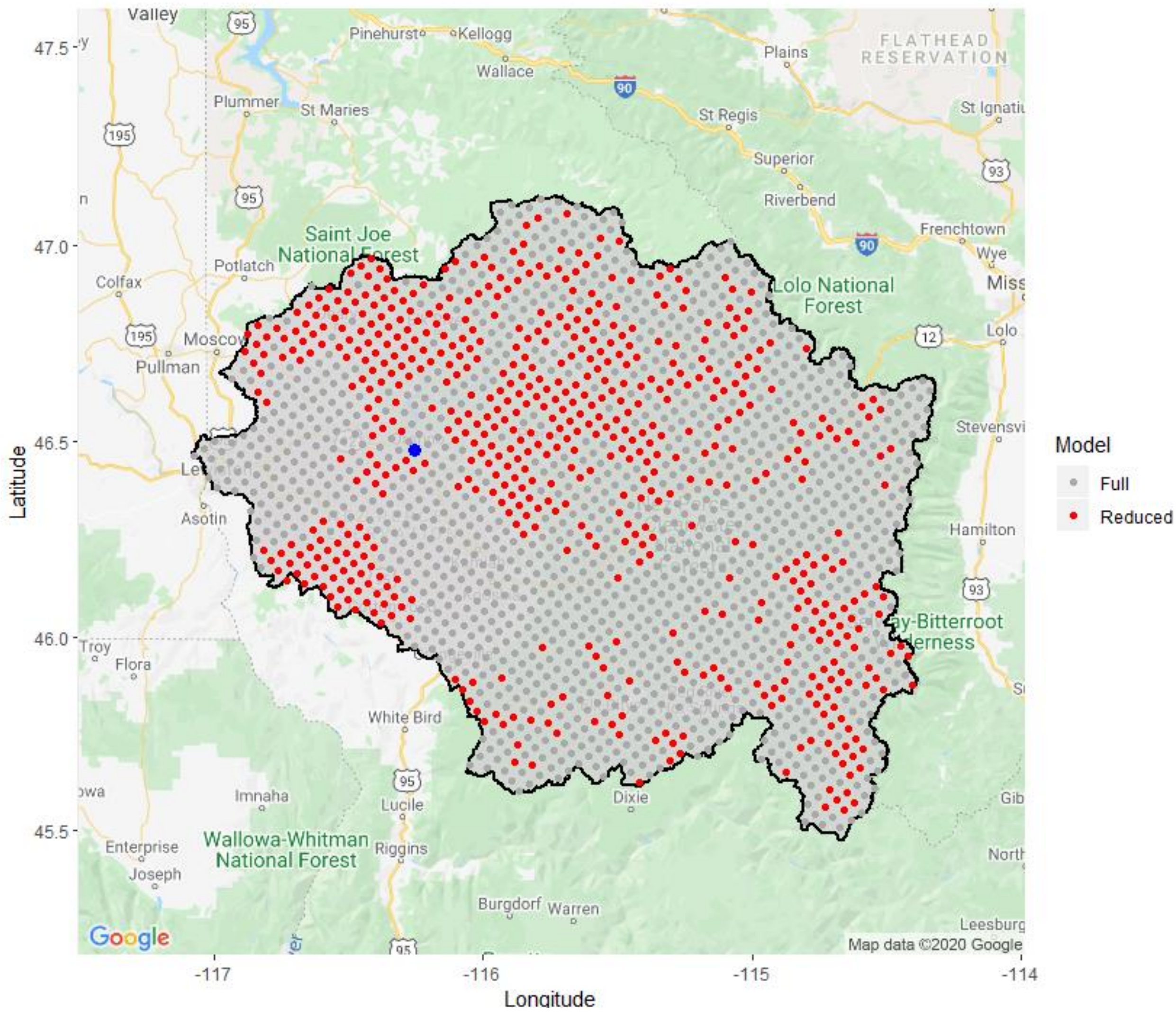

Figure D4. The 4-kilometer resolution IMS sample locations of the Clearwater subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 2004 - 2018. The blue dot indicates the location of the streamflow gage for the station at the Orofino station (13340000).

Table D3

The Full PSC March IMS Model for the 4-kilometer resolution (top), the Reduced PSC March IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 4-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Clearwater subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

CLEARWATER SUBBASIN

Full March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	32471.7	12444.8	2.609	0.0228 *
clearwater.imsfull	677.4	205.6	3.295	0.0064 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15350 on 12 degrees of freedom
Multiple R-squared: 0.475, Adjusted R-squared: 0.4313
Predicted R-squared: 0.3027
F-statistic: 10.86 on 1 and 12 DF, p-value: 0.006399

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsfull))

	Value	p-value	Decision
Global Stat	1.3338	0.8889	Assumptions acceptable.
Skewness	0.8048	0.3697	Assumptions acceptable.
Kurtosis	0.0426	0.8365	Assumptions acceptable.
Link Function	0.0901	0.764	Assumptions acceptable.
Heteroscedasticity	0.1963	0.6577	Assumptions acceptable.

Reduced March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	44219.0	7985.5	5.537	0.000128 ***
clearwater.imsred	512.8	133.6	3.838	0.002362 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14190 on 12 degrees of freedom
Multiple R-squared: 0.551, Adjusted R-squared: 0.5136
Predicted R-squared: 0.4166
F-statistic: 14.73 on 1 and 12 DF, p-value: 0.002362

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsred))

	Value	p-value	Decision
Global Stat	1.6895	0.7926	Assumptions acceptable.
Skewness	1.3007	0.2541	Assumptions acceptable.
Kurtosis	0.0026	0.9593	Assumptions acceptable.
Link Function	0.1732	0.6773	Assumptions acceptable.
Heteroscedasticity	0.213	0.6445	Assumptions acceptable.

Reduced March IMS Model (4-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	21976.0	17908.3	1.227	0.245
clearwater.imsred	254.0	228.0	1.114	0.289
clearwater.snotel	1248.0	907.1	1.376	0.196

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13690 on 11 degrees of freedom
Multiple R-squared: 0.617, Adjusted R-squared: 0.5473
Predicted R-squared: 0.440
F-statistic: 8.859 on 2 and 11 DF, p-value: 0.005103
VIF: 2.61

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsred + clearwater.snotel))

	Value	p-value	Decision
Global Stat	4.3002	0.3669	Assumptions acceptable.
Skewness	1.5764	0.2093	Assumptions acceptable.
Kurtosis	0.0271	0.8692	Assumptions acceptable.
Link Function	2.6882	0.1011	Assumptions acceptable.
Heteroscedasticity	0.0086	0.9263	Assumptions acceptable.

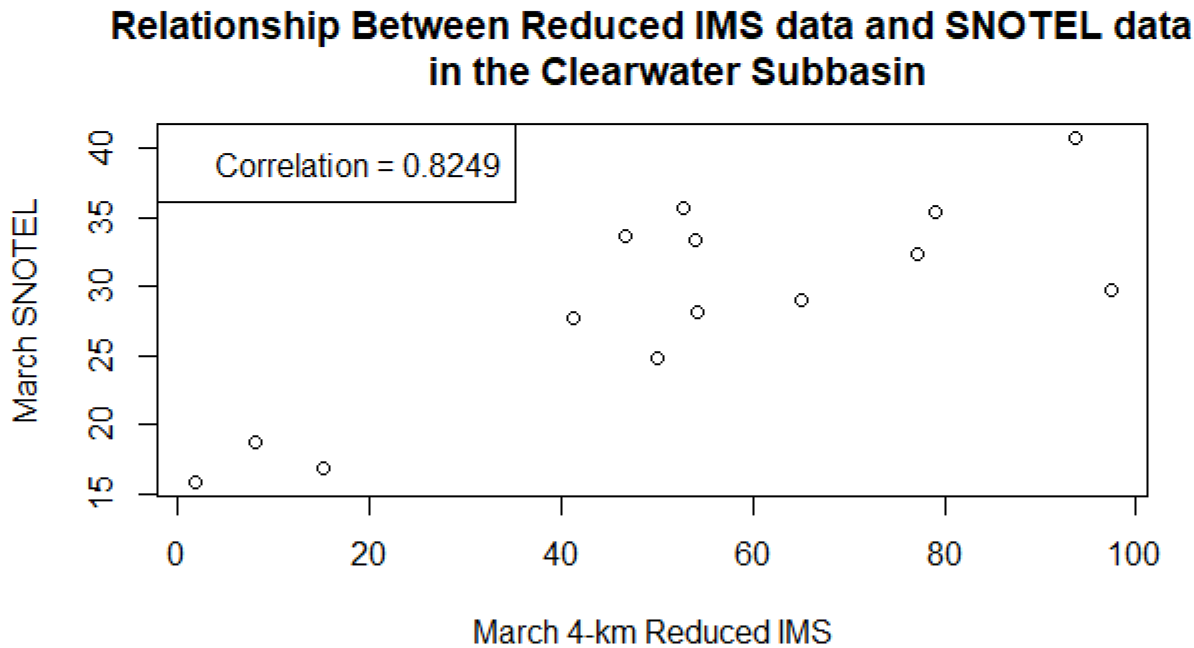

Figure D5. This correlation plot shows the correlation between the March 4-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Clearwater subbasin.

Table D4

Principal Component Analysis for the Clearwater subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (4-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Clearwater Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 14 2
Y dimension: 14 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 14 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	21118	14801	14675
adjCV	21118	14735	14582

TRAINING: % variance explained

	1 comps	2 comps
X	98.18	100.00
clearwater.stream	56.16	61.7

Coefficients:

1 component -

(Intercept)	42104.75
clearwater.imsred	495.7406
clearwater.snotel	104.828

Multiple R-squared: 0.5616
Predicted R-squared: 0.4303

2 components -

(Intercept)	21976.03
clearwater.imsred	254.0051
clearwater.snotel	1248.0159

Multiple R-squared: 0.617
Predicted R-squared: 0.440

APPENDIX E

Pend Oreille Subbasin

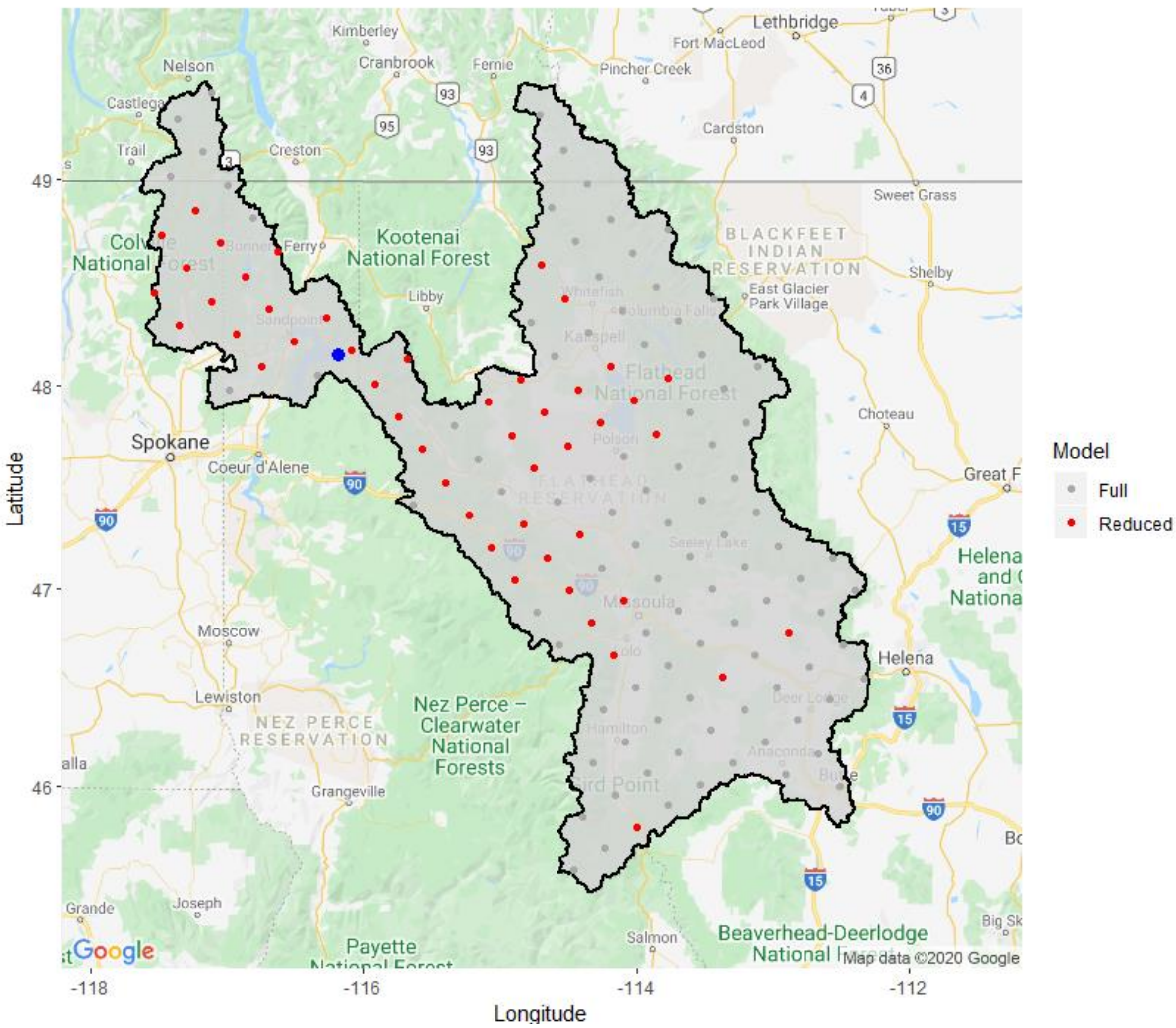

Figure E1. The 24-kilometer resolution IMS sample locations of the Pend Oreille subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 1999 - 2018. The blue dot indicates the location of the streamflow gage for the station at the Clark Fork station (12392155).

Table E1

The Full PSC March IMS Model for the 24-kilometer resolution (top), the Reduced PSC March IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 24-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Pend Oreille subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

PEND OREILLE SUBBASIN

Full March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	791.24	906.85	0.873	0.39509	
pendoreille.imsfull	37.76	11.48	3.288	0.00434	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 909.7 on 17 degrees of freedom
Multiple R-squared: 0.3888, Adjusted R-squared: 0.3528
Predicted R-squared: 0.1925
F-statistic: 10.81 on 1 and 17 DF, p-value: 0.004337

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsfull))

	Value	p-value	Decision
Global Stat	0.174	0.996	Assumptions acceptable.
Skewness	0.116	0.734	Assumptions acceptable.
Kurtosis	0.036	0.85	Assumptions acceptable.
Link Function	0.02	0.889	Assumptions acceptable.
Heteroscedasticity	0.003	0.958	Assumptions acceptable.

Reduced March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	2070.05	511.31	4.049	0.000835	***
pendoreille.imsred	23.93	6.91	3.464	0.002969	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 891 on 17 degrees of freedom
Multiple R-squared: 0.4137, Adjusted R-squared: 0.3792
Predicted R-squared: 0.2272
F-statistic: 12 on 1 and 17 DF, p-value: 0.002969

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred))

	Value	p-value	Decision
Global Stat	0.388	0.984	Assumptions acceptable.
Skewness	0.024	0.876	Assumptions acceptable.
Kurtosis	2E-04	0.987	Assumptions acceptable.
Link Function	0.361	0.548	Assumptions acceptable.
Heteroscedasticity	0.003	0.96	Assumptions acceptable.

Reduced March IMS Model (24-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	354.565	891.933	0.398	0.6962	
pendoreille.imsred	11.093	8.446	1.313	0.2076	
pendoreille.snotel	126.000	56.143	2.244	0.0393	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 800.9 on 16 degrees of freedom
Multiple R-squared: 0.5541, Adjusted R-squared: 0.4984
Predicted R-squared: 0.2988
F-statistic: 9.941 on 2 and 16 DF, p-value: 0.001563
VIF: 2.24

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred + pendoreille.snotel))

	Value	p-value	Decision
Global Stat	1.949	0.745	Assumptions acceptable.
Skewness	0.677	0.411	Assumptions acceptable.
Kurtosis	0.013	0.91	Assumptions acceptable.
Link Function	0.163	0.686	Assumptions acceptable.
Heteroscedasticity	1.096	0.295	Assumptions acceptable.

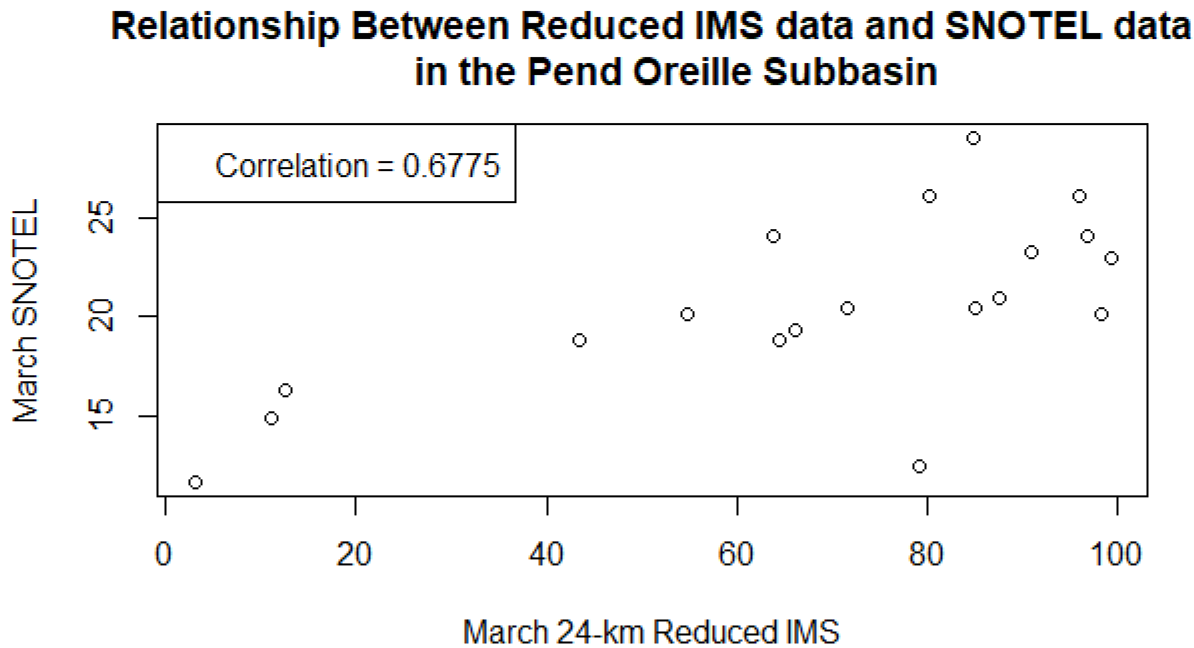

Figure E2. This correlation plot shows the correlation between the March 24-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Pend Oreille subbasin.

Table E2

Principal Component Analysis performed for the Pend Oreille Subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (24-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Pend Oreille Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	1162	963.2	921.7
adjCV	1162	959.8	916.5

TRAINING: % variance explained

	1 comps	2 comps
X	98.82	100.00
pendoreille.stream	41.91	55.41

Coefficients:

1 component -

(Intercept)	2026.089
pendoreille.imsred	23.83854
pendoreille.snotel	2.45949

Multiple R-squared: 0.4191
Predicted R-squared: 0.2342

2 components -

(Intercept)	354.5648
pendoreille.imsred	11.09253
pendoreille.snotel	125.99991

Multiple R-squared: 0.5541
Predicted R-squared: 0.2988

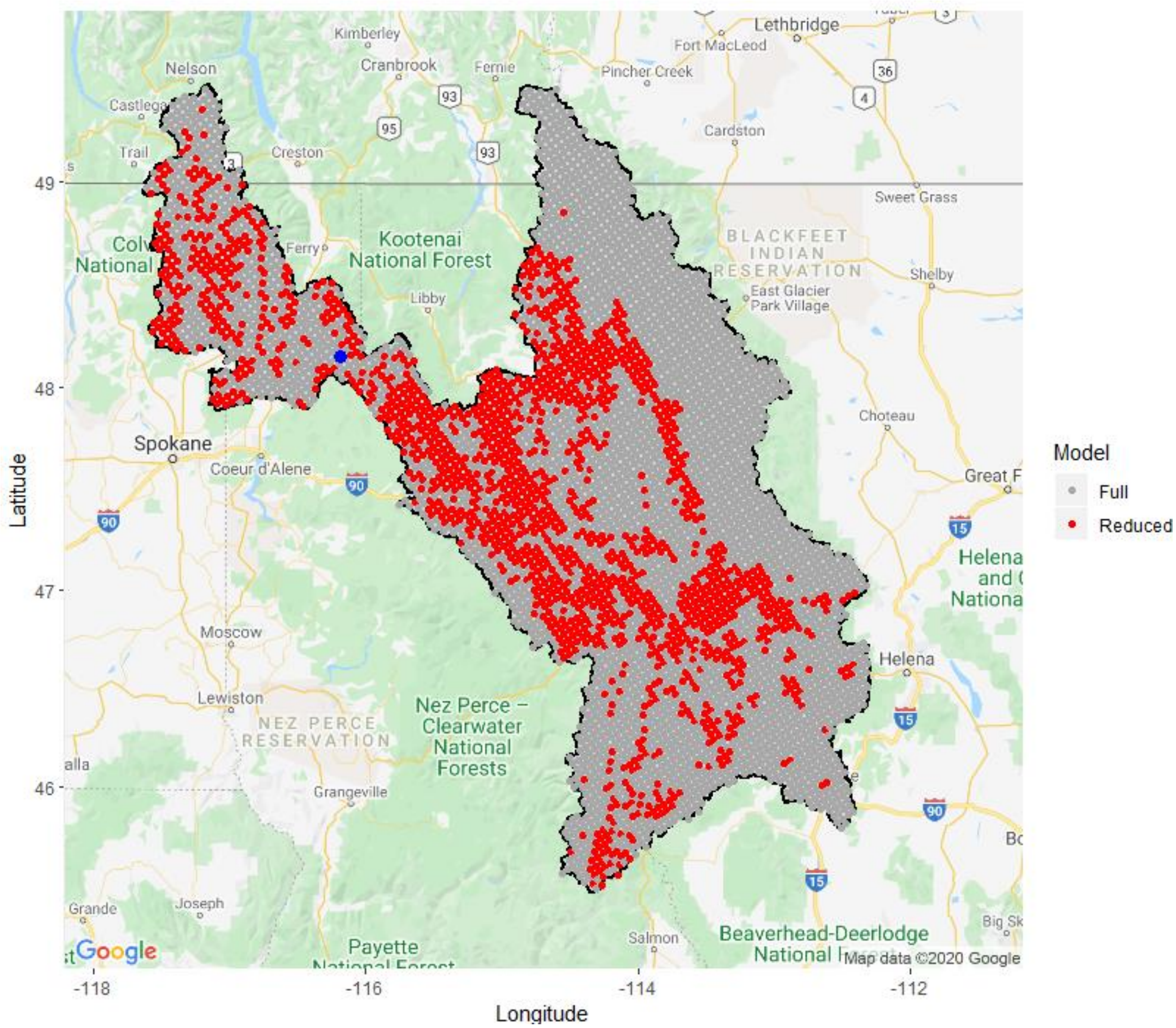

Figure E3. The 4-kilometer resolution IMS sample locations of the Pend Oreille subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 2004 - 2018. The blue dot indicates the location of the streamflow gage for the station at the Clark Fork station (12392155).

Table E3

The Full PSC March IMS Model for the 4-kilometer resolution (top), the Reduced PSC March IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 4-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Pend Oreille subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

PEND OREILLE SUBBASIN

Full March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	595.39	881.07	0.676	0.51201
pendoreille.imsfull	45.29	12.26	3.694	0.00307 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 850.1 on 12 degrees of freedom
Multiple R-squared: 0.532, Adjusted R-squared: 0.493
Predicted R-squared: 0.2973
F-statistic: 13.64 on 1 and 12 DF, p-value: 0.003072

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsfull))

	Value	p-value	Decision
Global Stat	0.797	0.939	Assumptions acceptable.
Skewness	0.016	0.9	Assumptions acceptable.
Kurtosis	0.031	0.861	Assumptions acceptable.
Link Function	0.448	0.503	Assumptions acceptable.
Heteroscedasticit	0.303	0.582	Assumptions acceptable.

Reduced March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2040.098	469.413	4.346	0.000951 ***
pendoreille.imsred	29.098	7.143	4.074	0.001544 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 805 on 12 degrees of freedom
Multiple R-squared: 0.5803, Adjusted R-squared: 0.5454
Predicted R-squared: 0.3813
F-statistic: 16.59 on 1 and 12 DF, p-value: 0.001544

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred))

	Value	p-value	Decision
Global Stat	0.529	0.971	Assumptions acceptable.
Skewness	0.012	0.912	Assumptions acceptable.
Kurtosis	0.013	0.91	Assumptions acceptable.
Link Function	0.245	0.621	Assumptions acceptable.
Heteroscedasticit	0.26	0.61	Assumptions acceptable.

Reduced March IMS Model (4-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	953.87	1411.86	0.676	0.513
pendoreille.imsred	20.21	13.07	1.546	0.150
pendoreille.snotel	79.10	96.79	0.817	0.431

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 816.4 on 11 degrees of freedom
Multiple R-squared: 0.6044, Adjusted R-squared: 0.5324
Predicted R-squared: 0.3044
F-statistic: 8.401 on 2 and 11 DF, p-value: 0.006098
VIF: 2.53

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred + pendoreille.snotel))

	Value	p-value	Decision
Global Stat	1.03	0.905	Assumptions acceptable.
Skewness	0.215	0.643	Assumptions acceptable.
Kurtosis	2E-04	0.99	Assumptions acceptable.
Link Function	0.153	0.696	Assumptions acceptable.
Heteroscedasticit	0.661	0.416	Assumptions acceptable.

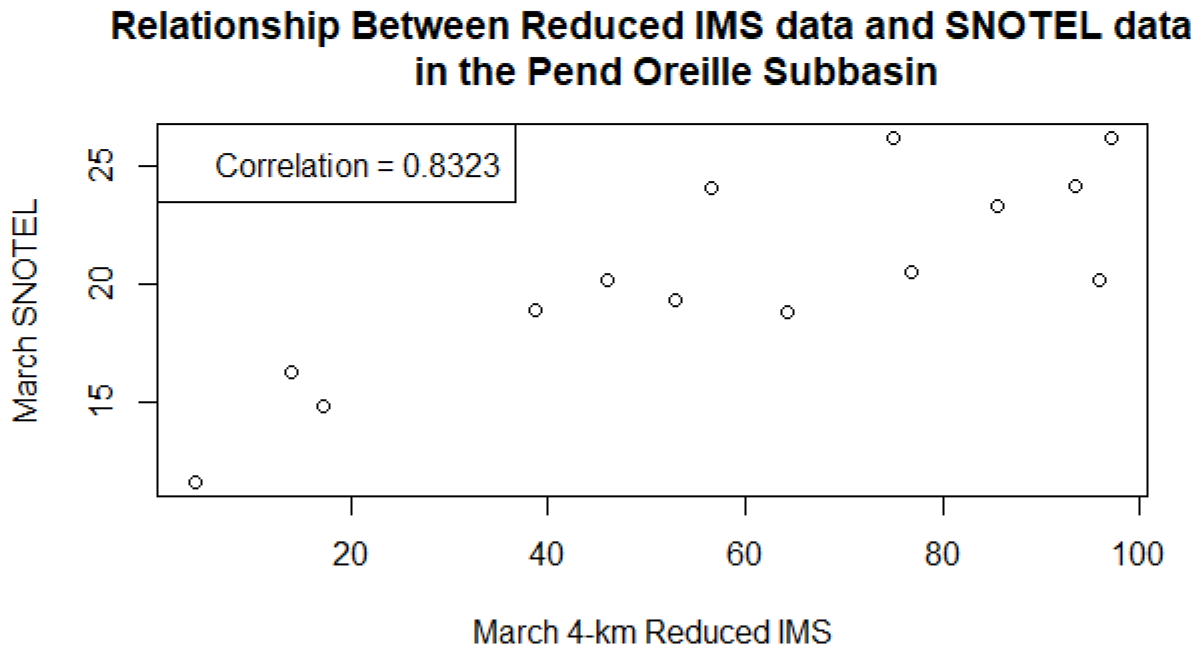

Figure E4. This correlation plot shows the correlation between the March 4-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Pend Oreille subbasin.

Table E4

Principal Component Analysis performed for the Pend Oreille subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (4-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Pend Oreille Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 14 2
Y dimension: 14 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 14 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	1239	901.9	959.6
adjCV	1239	896.2	950.8

TRAINING: % variance explained

	1 comps	2 comps
X	99.46	100.00
pendoreille.stream	58.23	60.44

Coefficients:

1 component -

(Intercept)	1992.666
pendoreille.imsred	28.779582
pendoreille.snotel	3.252273

Multiple R-squared: 0.5823
Predicted R-squared: 0.3855

2 components -

(Intercept)	953.8729
pendoreille.imsred	20.20858
pendoreille.snotel	79.09761

Multiple R-squared: 0.6044
Predicted R-squared: 0.3044

APPENDIX F

Kootenai Subbasin

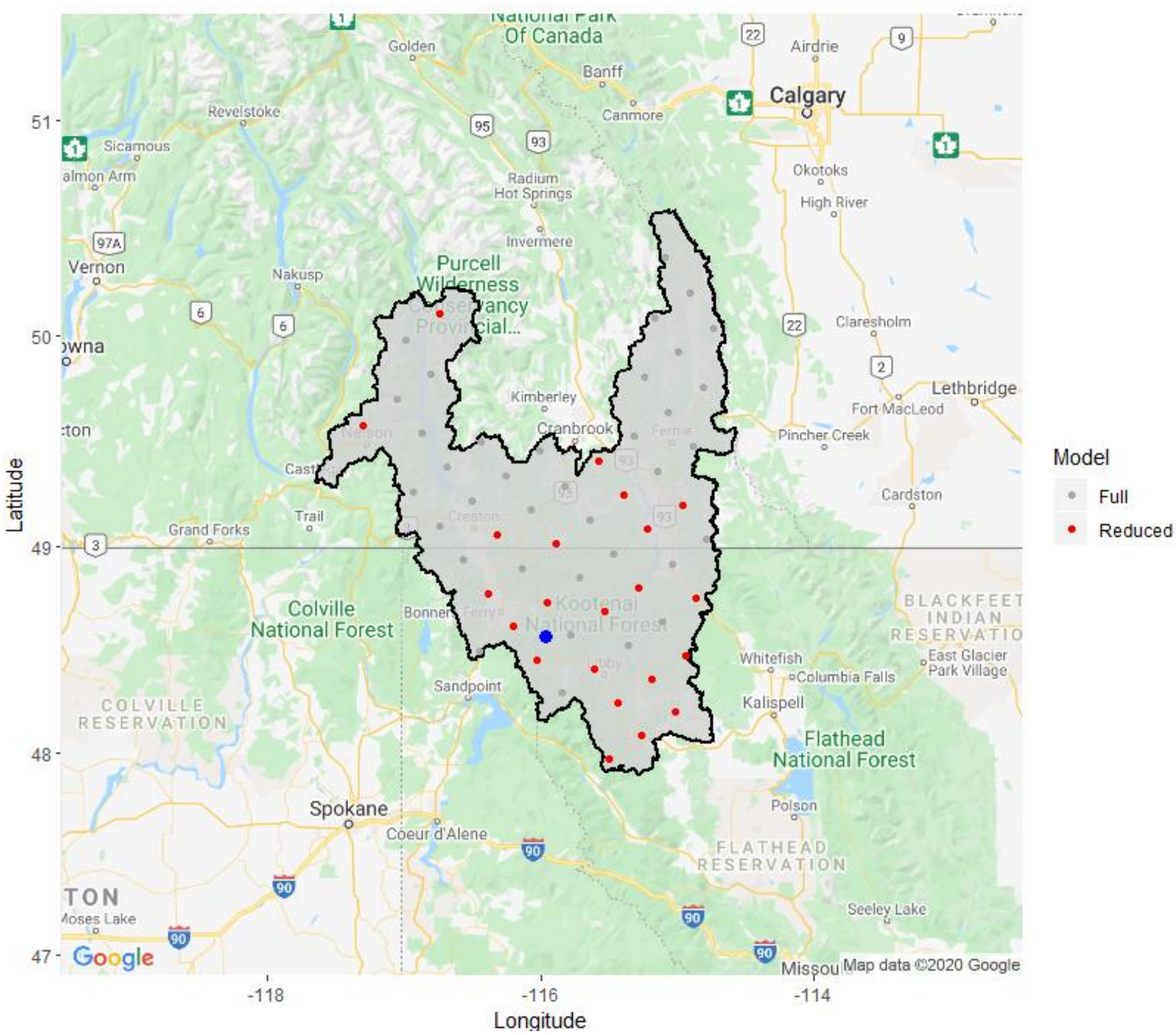

Figure F1. The 24-kilometer resolution IMS sample locations of the Kootenai subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 1999 - 2018. The blue dot indicates the location of the streamflow gage for the station at the Yaak River near Troy MT (12304500). The horizontal line at 49 degrees is the Canadian border.

Table F1

The Full PSC March IMS Model for the 24-kilometer resolution (top), the Reduced PSC March IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 24-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Kootenai subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

KOOTENAI SUBBASIN

Full March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-252.39	3135.33	-0.080	0.9368
kootenai.imsfull	78.71	34.73	2.266	0.0368 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2366 on 17 degrees of freedom
Multiple R-squared: 0.232,   Adjusted R-squared: 0.1868
Predicted R-squared: -0.0164
F-statistic: 5.135 on 1 and 17 DF, p-value: 0.03679

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsfull))

	Value	p-value	Decision
Global Stat	2.975	0.562	Assumptions acceptable.
Skewness	0.072	0.788	Assumptions acceptable.
Kurtosis	0.073	0.787	Assumptions acceptable.
Link Function	2.636	0.104	Assumptions acceptable.
Heteroscedastic	0.194	0.66	Assumptions acceptable.

Reduced March IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3315.29	1509.05	2.197	0.0422 *
kootenai.imsred	43.91	18.07	2.430	0.0265 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2326 on 17 degrees of freedom
Multiple R-squared: 0.2578,   Adjusted R-squared: 0.2142
Predicted R-squared: 0.1114
F-statistic: 5.905 on 1 and 17 DF, p-value: 0.02646

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred))

	Value	p-value	Decision
Global Stat	1.77	0.778	Assumptions acceptable.
Skewness	0.002	0.962	Assumptions acceptable.
Kurtosis	0.017	0.895	Assumptions acceptable.
Link Function	1.379	0.24	Assumptions acceptable.
Heteroscedastic	0.371	0.542	Assumptions acceptable.

Reduced March IMS Model (24-km) w/ March SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-2462.388	1524.068	-1.616	0.125710
kootenai.imsred	1.208	14.555	0.083	0.934874
kootenai.snotel	407.504	82.422	4.944	0.000146 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1508 on 16 degrees of freedom
Multiple R-squared: 0.7064,   Adjusted R-squared: 0.6697
Predicted R-squared: 0.6088
F-statistic: 19.25 on 2 and 16 DF, p-value: 5.523e-05
VIF: 3.41

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred + kootenai.snotel))

	Value	p-value	Decision
Global Stat	2.449	0.654	Assumptions acceptable.
Skewness	0.567	0.451	Assumptions acceptable.
Kurtosis	0.258	0.611	Assumptions acceptable.
Link Function	1.229	0.268	Assumptions acceptable.
Heteroscedastic	0.395	0.53	Assumptions acceptable.

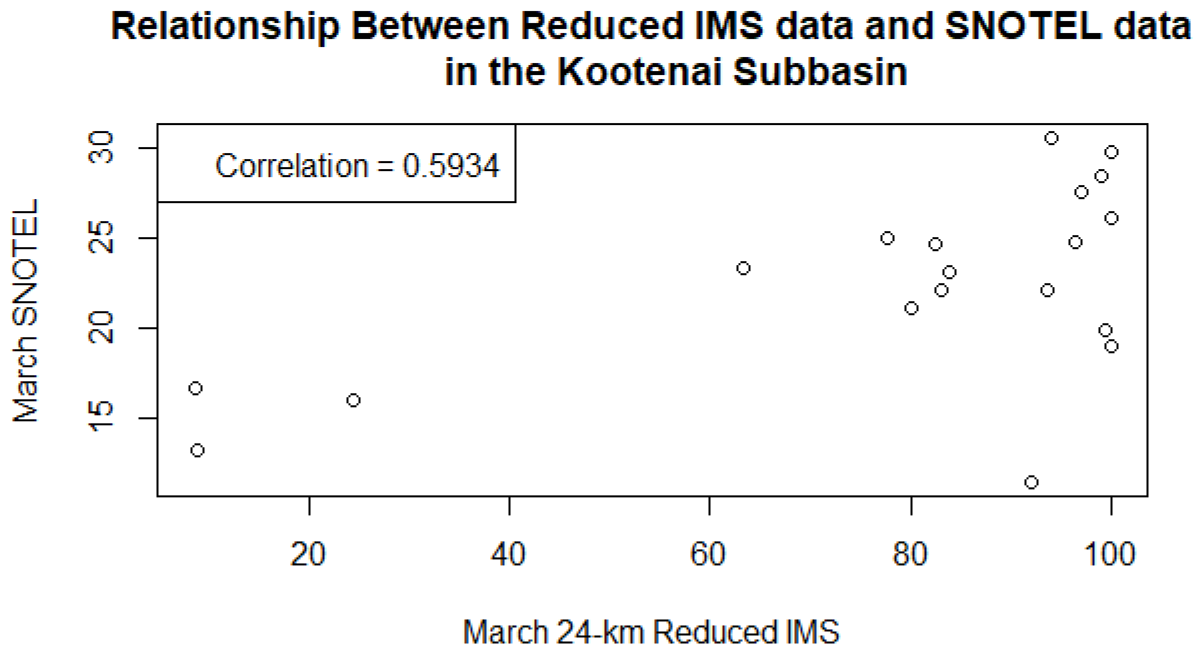

Figure F2. This correlation plot shows the correlation between the March 24-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Kootenai subbasin.

Table F2

Principal Component Analysis performed for the Kootenai subbasin at the 24-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (24-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Kootenai Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	2696	2393	1597
adjCV	2696	2387	1592

TRAINING: % variance explained

	1 comps	2 comps
X	98.06	100.00
kootenai.stream	26.81	70.64

Coefficients:

1 component -

(Intercept)	3180.941
kootenai.imsred	44.276478
kootenai.snotel	4.734493

Multiple R-squared: 0.2681
Predicted R-squared: 0.1222

2 components -

(Intercept)	-2462.388
kootenai.imsred	1.208214
kootenai.snotel	407.504361

Multiple R-squared: 0.7064
Predicted R-squared: 0.6088

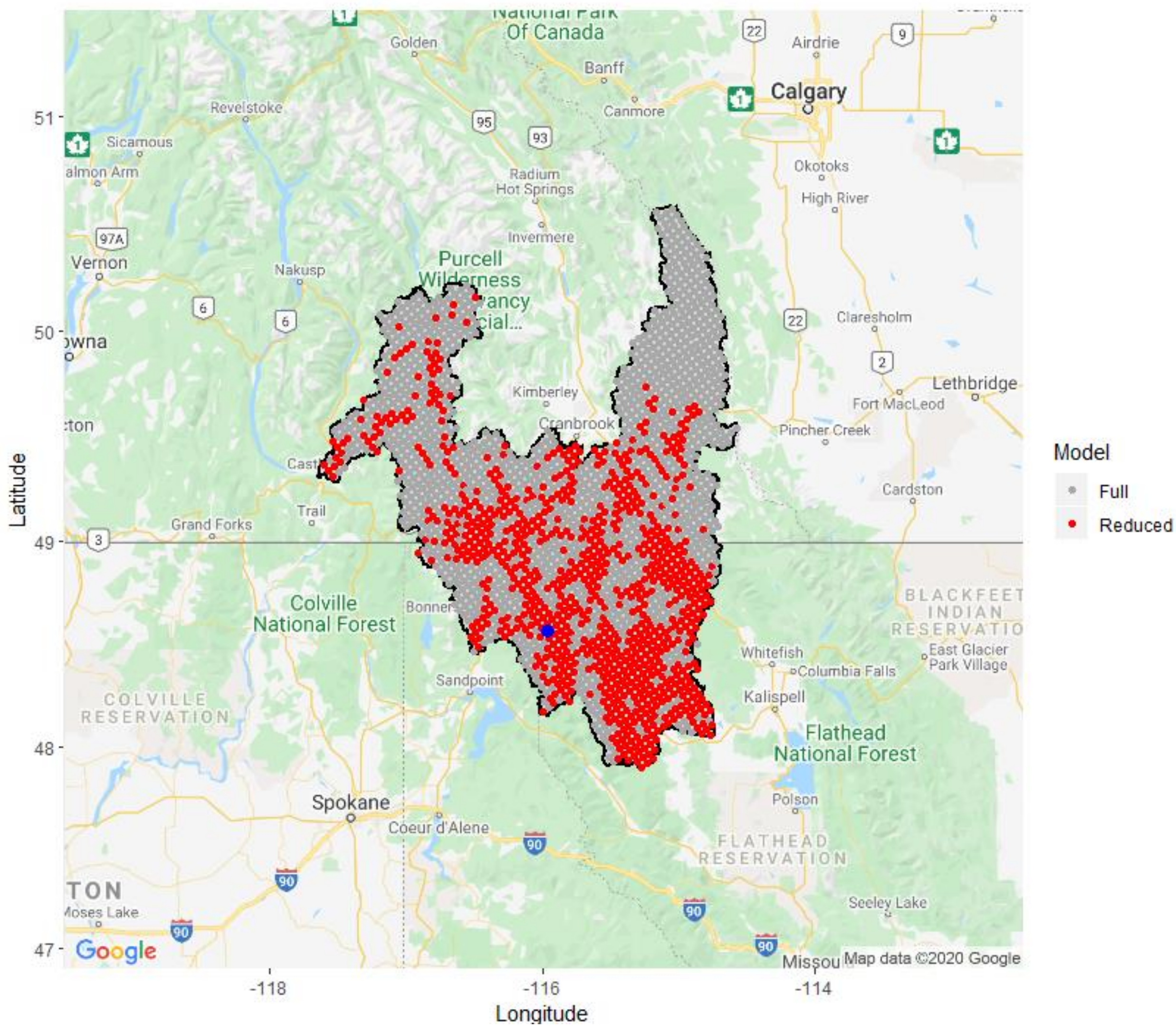

Figure F3. The 24-kilometer resolution IMS sample locations of the Kootenai subbasin for the Full PSC metric (grey and red dots) as well as the Reduced PSC metric (only red dots) for the month of March, 2004 - 2018. The blue dot indicates the location of the streamflow gage for the station at the Yaak River near Troy MT (12304500). The horizontal line at 49 degrees is the Canadian border.

Table F3

The Full PSC March IMS Model for the 4-kilometer resolution (top), the Reduced PSC March IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC March IMS Model for the 4-kilometer resolution with March Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Kootenai subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

KOOTENAI SUBBASIN

Full March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-89.90	2346.55	-0.038	0.9701
kootenai.imsfull	88.83	29.29	3.033	0.0104 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2055 on 12 degrees of freedom
Multiple R-squared: 0.4339, Adjusted R-squared: 0.3867
Predicted R-squared: 0.2109
F-statistic: 9.197 on 1 and 12 DF, p-value: 0.01042

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsfull))

	Value	p-value	Decision
Global Stat	2.0134	0.7333	Assumptions acceptable.
Skewness	0.5677	0.4512	Assumptions acceptable.
Kurtosis	0.2248	0.6354	Assumptions acceptable.
Link Function	1.1073	0.2927	Assumptions acceptable.
Heteroscedasticity	0.1137	0.736	Assumptions acceptable.

Reduced March IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3626.06	1113.42	3.257	0.00687 **
kootenai.imsred	52.01	15.90	3.272	0.00668 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1986 on 12 degrees of freedom
Multiple R-squared: 0.4715, Adjusted R-squared: 0.4274
Predicted R-squared: 0.2943
F-statistic: 10.71 on 1 and 12 DF, p-value: 0.00668

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred))

	Value	p-value	Decision
Global Stat	1.038	0.904	Assumptions acceptable.
Skewness	0.3354	0.5625	Assumptions acceptable.
Kurtosis	0.195	0.6588	Assumptions acceptable.
Link Function	0.3246	0.5688	Assumptions acceptable.
Heteroscedasticity	0.183	0.6688	Assumptions acceptable.

Reduced March IMS Model w/ March SNOTEL (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-3426.193	2627.905	-1.304	0.2189
kootenai.imsred	6.031	20.472	0.295	0.7738
kootenai.snotel	444.706	156.096	2.849	0.0158 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1573 on 11 degrees of freedom
Multiple R-squared: 0.6959, Adjusted R-squared: 0.6406
Predicted R-squared: 0.4992
F-statistic: 12.59 on 2 and 11 DF, p-value: 0.001435
VIF: 3.29

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred + kootenai.snotel))

	Value	p-value	Decision
Global Stat	8.3622	0.0792	Assumptions acceptable.
Skewness	0.1317	0.7167	Assumptions acceptable.
Kurtosis	0.7191	0.3964	Assumptions acceptable.
Link Function	7.4143	0.0065	Assumptions NOT satisifed!
Heteroscedasticity	0.0971	0.7554	Assumptions acceptable.

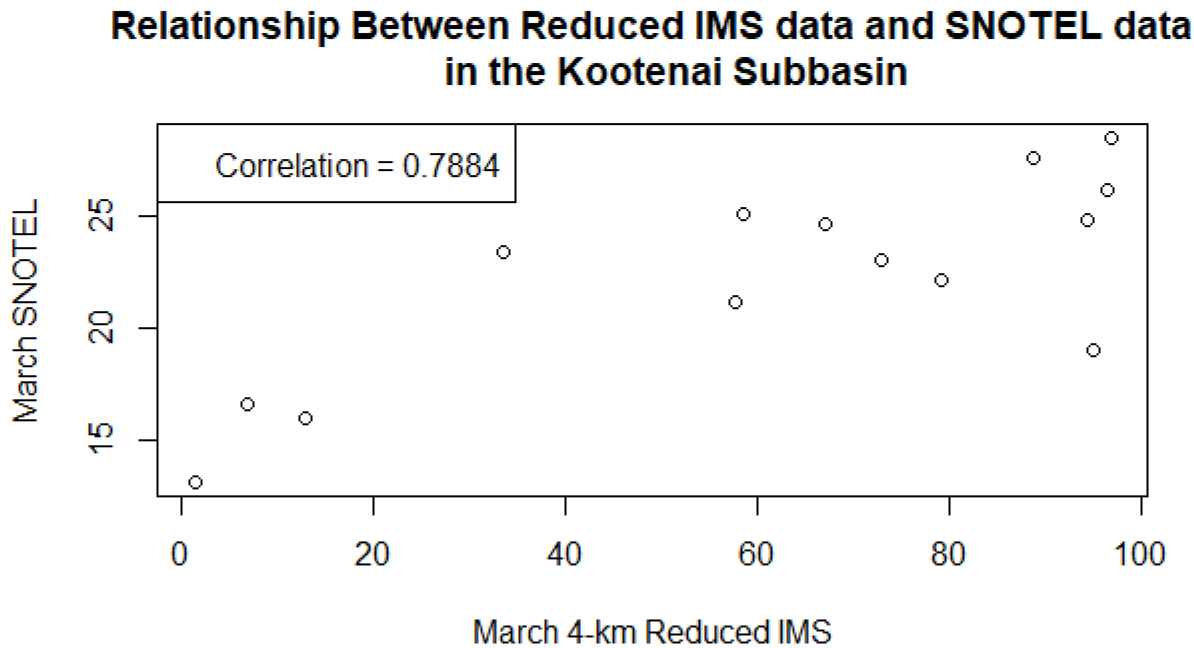

Figure F4. This correlation plot shows the correlation between the March 4-kilometer resolution Reduced PSC IMS variable and the March Snow Telemetry (SNOTEL) variable for the Kootenai subbasin.

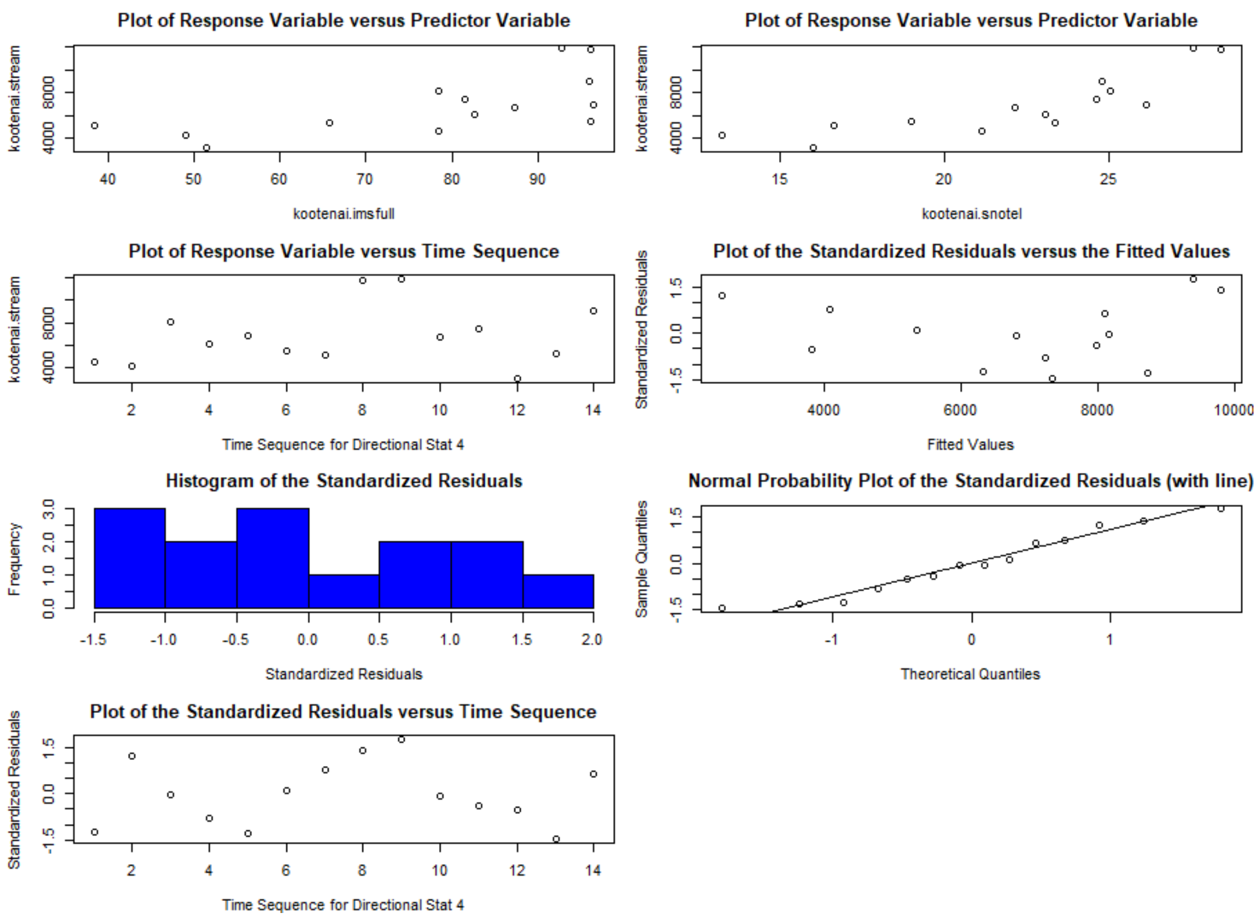

Figure F5. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the Kootenai subbasin Full March IMS Model (4-km) with March Snow Telemetry data (SNOTEL) added as a second predictor.

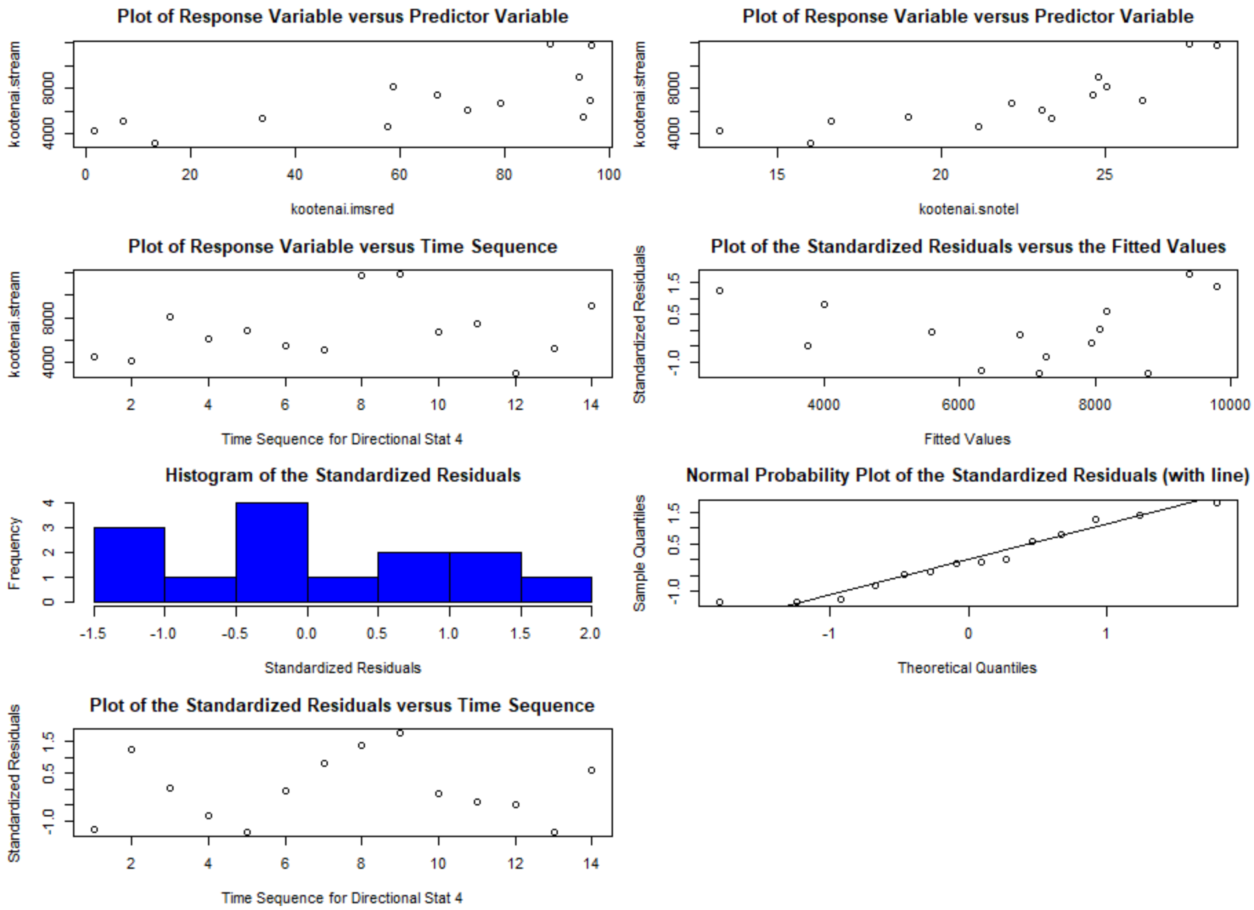

Figure F6. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the Kootenai subbasin Reduced March IMS Model (4-km) with March Snow Telemetry data (SNOTEL) added as a second predictor.

Table F4

Principal Component Analysis performed for the Kootenai subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced March IMS Model (4-km) with the March Snow Telemetry (SNOTEL) added as a second predictor.

Kootenai Subbasin

Principal Component Regression of March IMS + March SNOTEL

Data: X dimension: 14 2
Y dimension: 14 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 14 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	2723	2115	1790
adjCV	2723	2104	1775

TRAINING: % variance explained

	1 comps	2 comps
X	99.37	100.00
kootenai.stream	47.69	69.59

Coefficients:

1 component -

(Intercept)	3522.522
kootenai.imsred	51.749912
kootenai.snotel	5.385424

Multiple R-squared: 0.4769
Predicted R-squared: 0.3007

2 components -

(Intercept)	-3426.193
kootenai.imsred	6.031462
kootenai.snotel	444.705665

Multiple R-squared: 0.6959
Predicted R-squared: 0.4992

APPENDIX G

Summary Tables

Table G1

The Summary Table for all the models proposed for the month of March at the 24-kilometer resolution.

Subbasin	March IMS Snow Signal (24-km)				March IMS Snow Signal (24-km) + March SNOTEL			
	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom
Yakima (Full)	0.3565	0.2911	0.004121	17	0.4568	0.3287	0.002956	16
Yakima (Reduced)	0.4644	0.402	0.0007877	17	0.507	0.4331	0.001359	16
Deschutes (Full)	0.7482	0.7083	0.000001	17	0.7384	0.697	0.000008	16
Deschutes (Reduced)	0.7946	0.7651	0.0000001	17	0.782	0.7494	0.0000001	16
John Day (Full)	0.2836	0.1709	0.01106	17	0.3122	0.2003	0.01951	16
John Day (Reduced)	0.3781	0.2567	0.003019	17	0.3623	0.2325	0.01065	16
Clearwater (Full)	0.309	0.2245	0.007911	17	0.4693	0.357	0.002454	16
Clearwater (Reduced)	0.3505	0.2585	0.00488	17	0.4862	0.3675	0.001894	16
Pend Oreille (Full)	0.3528	0.1925	0.004337	17	0.4956	0.2898	0.001632	16
Pend Oreille (Reduced)	0.3792	0.2272	0.002969	17	0.4984	0.2988	0.001563	16
Kootenai (Full)	0.1868	-0.0164	0.03679	17	0.6696	0.6049	0.000005	16
Kootenai (Reduced)	0.2142	0.1114	0.02646	17	0.6697	0.6088	0.000005	16

Table G2

The Summary Table for all the models proposed for the month of March at the 4-kilometer resolution.

Subbasin	March IMS Snow Signal (4-km)				March IMS Snow Signal (4-km) + March SNOTEL			
	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom
Yakima (Full)	0.6294	0.5355	0.0004305	12	0.6138	0.5046	0.002129	11
Yakima (Reduced)	0.71	0.6556	0.0000947	12	0.6843	0.6055	0.0007033	11
Deschutes (Full)	0.7544	0.7013	3.415E-05	12	0.7502	0.679	0.0001941	11
Deschutes (Reduced)	0.7665	0.7199	2.505E-05	12	0.7653	0.6468	0.0001377	11
John Day (Full)	0.396	0.2271	0.009436	12	0.3737	0.1836	0.030402	11
John Day (Reduced)	0.4155	0.2734	0.007634	12	0.3739	0.2218	0.030309	11
Clearwater (Full)	0.4313	0.3027	0.006399	12	0.5283	0.4059	0.006398	11
Clearwater (Reduced)	0.5136	0.4166	0.002362	12	0.5473	0.434	0.005103	11
Pend Oreille (Full)	0.493	0.2973	0.003072	12	0.5023	0.2681	0.008598	11
Pend Oreille (Reduced)	0.5454	0.3813	0.001544	12	0.5324	0.3044	0.006098	11
Kootenai (Full)	0.3867	0.2109	0.01042	12	0.6379	0.4976	0.001494	11
Kootenai (Reduced)	0.4274	0.2934	0.00668	12	0.6406	0.4992	0.001435	11

Table G3

These are all the models proposed for the month of March prediction period, by Predicted R^2. The models with the highest $Predicted\ R^2$ are highlighted in yellow.

Model Selection Based on Predicted R^2 - MARCH												
Subbasin	March IMS (Full)		March IMS (Reduced)		March IMS (Full) + March Snotel		March IMS (Reduced) + March Snotel		1-Component PCA		2-Component PCA	
	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution
Yakima	0.5355	0.2911	0.6556	0.402	0.5046	0.3287	0.6055	0.4331	0.6485	0.4298	0.6055	0.4331
Deschutes	0.7013	0.7083	0.7199	0.7651	0.679	0.697	0.6468	0.7494	0.7073	0.7625	0.6468	0.7494
John Day	0.2271	0.1709**	0.2734	0.2567***	0.1836**	0.2003**	0.2218**	0.2325**	0.2761	0.2596	0.2218	0.2325
Clearwater	0.3072	0.2245	0.4166	0.2585	0.4059	0.357	0.44	0.3675	0.4303	0.2825	0.44	0.3675
Pend Oreille	0.2973	0.1925	0.3813	0.2272	0.2681	0.2898	0.3044	0.2988	0.3855	0.2342	0.3044	0.2988
Kootenai	0.2109	-0.0164	0.2934	0.1114	0.4976*	0.6049	0.4992*	0.6088	0.3007	0.1222	0.4992	0.6088

*One linear model assumption violated

**Two linear model assumptions violated

***Three or more linear model assumptions violated

APPENDIX H

Additional Materials

Table H1

The Full PSC February IMS Model for the 24-kilometer resolution (top), the Reduced PSC February IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 24-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Yakima subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

YAKIMA SUBBASIN

Full February IMS Model (24-km)

```
Coefficients:
                  Estimate    Std. Error   t value   Pr(>|t|)
(Intercept)       12125.97    5139.27      2.359     0.0305 *
yakima.imsfull    143.89      80.99        1.777     0.0935 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6614 on 17 degrees of freedom
Multiple R-squared:  0.1566,    Adjusted R-squared:  0.107
Predicted R-squared:  -0.0121
F-statistic: 3.156 on 1 and 17 DF,  p-value: 0.09354
```

```
ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance =  0.05

Call:
 gvlma(x = lm(yakima.stream ~ yakima.imsfull))

                     Value    p-value    Decision
Global Stat          2.3878   0.6648     Assumptions acceptable.
Skewness             0.1633   0.6862     Assumptions acceptable.
Kurtosis             0.0528   0.8182     Assumptions acceptable.
Link Function        1.7947   0.1804     Assumptions acceptable.
Heteroscedasticity   0.377    0.5392     Assumptions acceptable.
```

Reduced February IMS Model (24-km)

```
Coefficients:
                  Estimate    Std. Error   t value   Pr(>|t|)
(Intercept)       10721.15    3513.15      3.052     0.00721 **
yakima.imsred     144.78      46.55        3.110     0.00636 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5750 on 17 degrees of freedom
Multiple R-squared:  0.3627,    Adjusted R-squared:  0.3252
Predicted R-squared:  0.244
F-statistic: 9.675 on 1 and 17 DF,  p-value: 0.006361
```

```
ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance =  0.05

Call:
 gvlma(x = lm(yakima.stream ~ yakima.imsred))

                     Value    p-value    Decision
Global Stat          0.8609   0.9301     Assumptions acceptable.
Skewness             0.0032   0.9546     Assumptions acceptable.
Kurtosis             0.2421   0.6227     Assumptions acceptable.
Link Function        0.0256   0.8729     Assumptions acceptable.
Heteroscedasticity   0.59     0.4424     Assumptions acceptable.
```

Reduced February IMS Model (24-km) w/ February SNOTEL

```
Coefficients:
                  Estimate    Std. Error   t value   Pr(>|t|)
(Intercept)       8704.42     4112.67      2.116     0.0503 .
yakima.imsred     111.09      58.61        1.895     0.0763 .
yakima.snotel     194.53      204.65       0.951     0.3560
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5766 on 16 degrees of freedom
Multiple R-squared:  0.3968,    Adjusted R-squared:  0.3213
Predicted R-squared:  0.2372
F-statistic: 5.262 on 2 and 16 DF,  p-value: 0.01754
VIF: 1.58
```

```
ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance =  0.05

Call:
 gvlma(x = lm(yakima.stream ~ yakima.imsred + yakima.snotel))

                     Value    p-value    Decision
Global Stat          1.6422   0.8012     Assumptions acceptable.
Skewness             0.4626   0.4964     Assumptions acceptable.
Kurtosis             0.9127   0.3394     Assumptions acceptable.
Link Function        0.238    0.6257     Assumptions acceptable.
Heteroscedasticity   0.0289   0.8651     Assumptions acceptable.
```

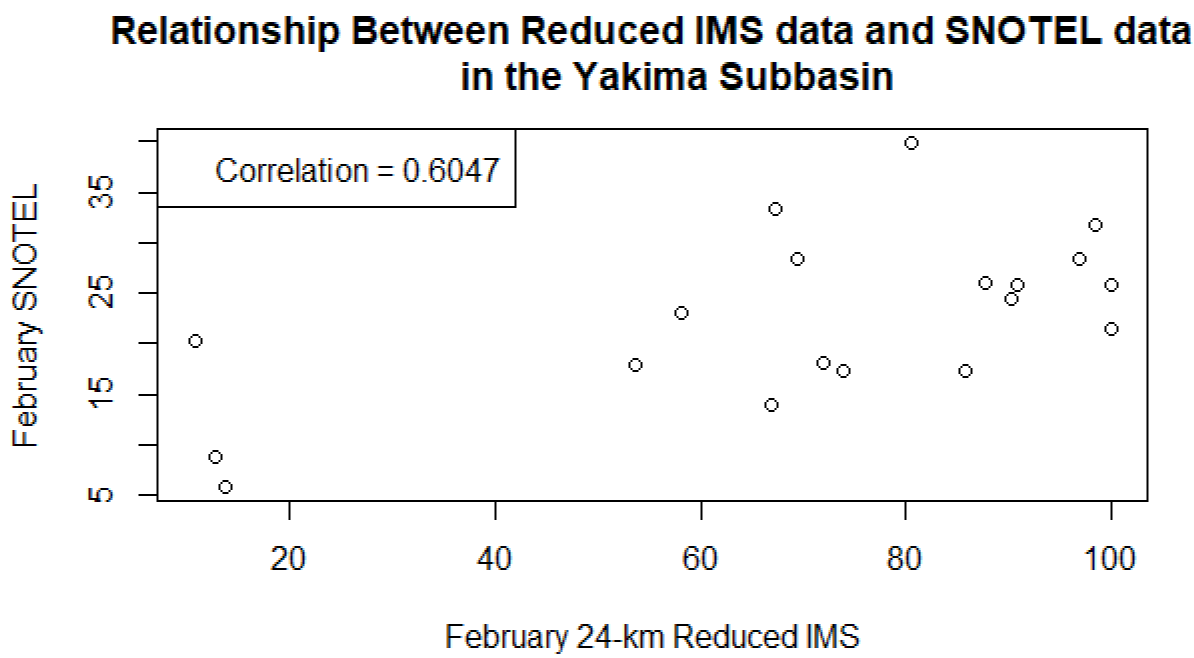

Figure H1. This correlation plot shows the correlation between the February 24-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Yakima subbasin.

Table H2

Principal Component Analysis performed for the Yakima subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (24-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Yakima Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	7191	5877	5950
adjCV	7191	5864	5932

TRAINING: % variance explained

	1 comps	2 comps
X	95.34	100.00
yakima.stream	37.11	39.68

Coefficients:

1 component -

(Intercept)	10344.38
yakima.imsred	141.854
yakima.snotel	25.87319

Multiple R-squared: 0.3711
Predicted R-squared: 0.2557

2 components -

(Intercept)	8704.425
yakima.imsred	111.0929
yakima.snotel	194.5257

Multiple R-squared: 0.3968
Predicted R-squared: 0.2372

Table H3

The Full PSC February IMS Model for the 4-kilometer resolution (top), the Reduced PSC February IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 4-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Yakima subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

YAKIMA SUBBASIN

Full February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12829.76	4211.34	3.046	0.0111 *
yakima.imsfull	192.99	82.73	2.333	0.0397 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6092 on 11 degrees of freedom
Multiple R-squared: 0.331, Adjusted R-squared: 0.2702
Predicted R-squared: 0.0677
F-statistic: 5.443 on 1 and 11 DF, p-value: 0.03966

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(yakima.stream ~ yakima.imsfull))

	Value	p-value	Decision
Global Stat	5.4893	0.2407	Assumptions acceptable.
Skewness	2.0646	0.1508	Assumptions acceptable.
Kurtosis	0.0276	0.8681	Assumptions acceptable.
Link Function	3.375	0.0662	Assumptions acceptable.
Heteroscedasticity	0.0222	0.8816	Assumptions acceptable.

Reduced February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	11302.80	3143.18	3.596	0.00420 **
yakima.imsred	188.00	50.49	3.724	0.00336 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4954 on 11 degrees of freedom
Multiple R-squared: 0.5576, Adjusted R-squared: 0.5174
Predicted R-squared: 0.4186
F-statistic: 13.86 on 1 and 11 DF, p-value: 0.003361

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(yakima.stream ~ yakima.imsred))

	Value	p-value	Decision
Global Stat	5.7025	0.2225	Assumptions acceptable.
Skewness	2.7987	0.0943	Assumptions acceptable.
Kurtosis	0.8587	0.3541	Assumptions acceptable.
Link Function	2.0321	0.154	Assumptions acceptable.
Heteroscedasticity	0.0131	0.909	Assumptions acceptable.

Reduced February IMS Model (4-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	11005.51	4275.78	2.574	0.0277 *
yakima.imsred	183.87	65.09	2.825	0.0180 *
yakima.snotel	24.81	227.45	0.109	0.9153

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5193 on 10 degrees of freedom
Multiple R-squared: 0.5581, Adjusted R-squared: 0.4698
Predicted R-squared: 0.3579
F-statistic: 6.316 on 2 and 10 DF, p-value: 0.01684
VIF: 1.51

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(yakima.stream ~ yakima.imsred + yakima.snotel))

	Value	p-value	Decision
Global Stat	6.6939	0.153	Assumptions acceptable.
Skewness	2.9239	0.0873	Assumptions acceptable.
Kurtosis	1.067	0.3016	Assumptions acceptable.
Link Function	2.6837	0.1014	Assumptions acceptable.
Heteroscedasticity	0.0193	0.8896	Assumptions acceptable.

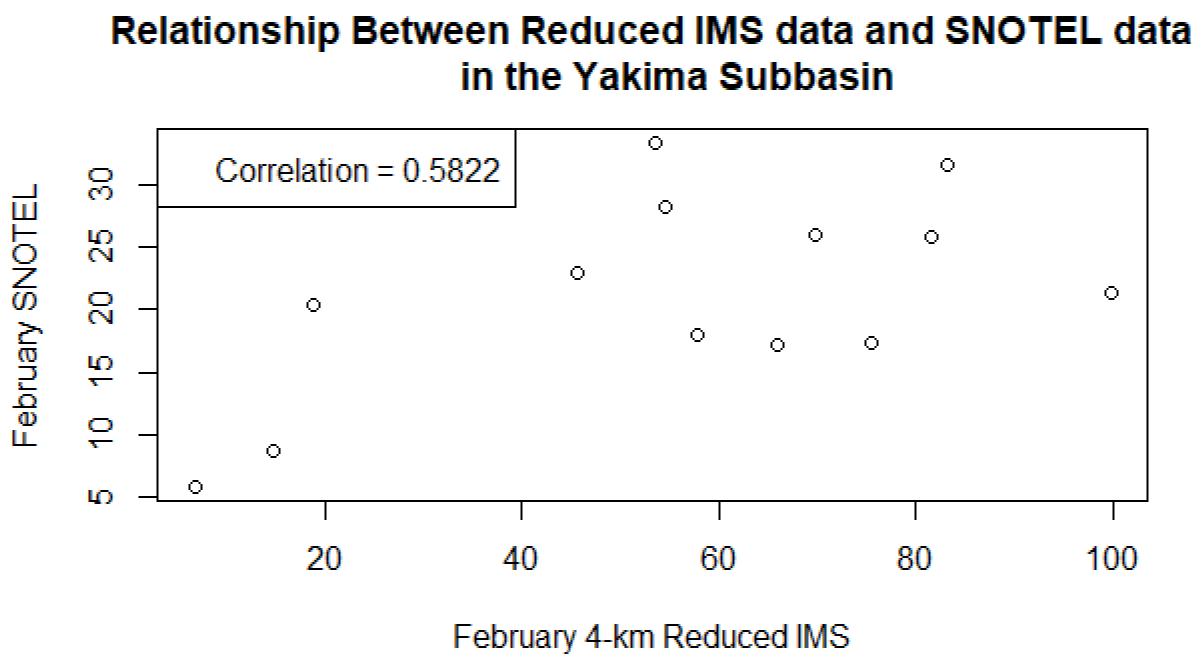

Figure H2. This correlation plot shows the correlation between the February 4-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Yakima subbasin.

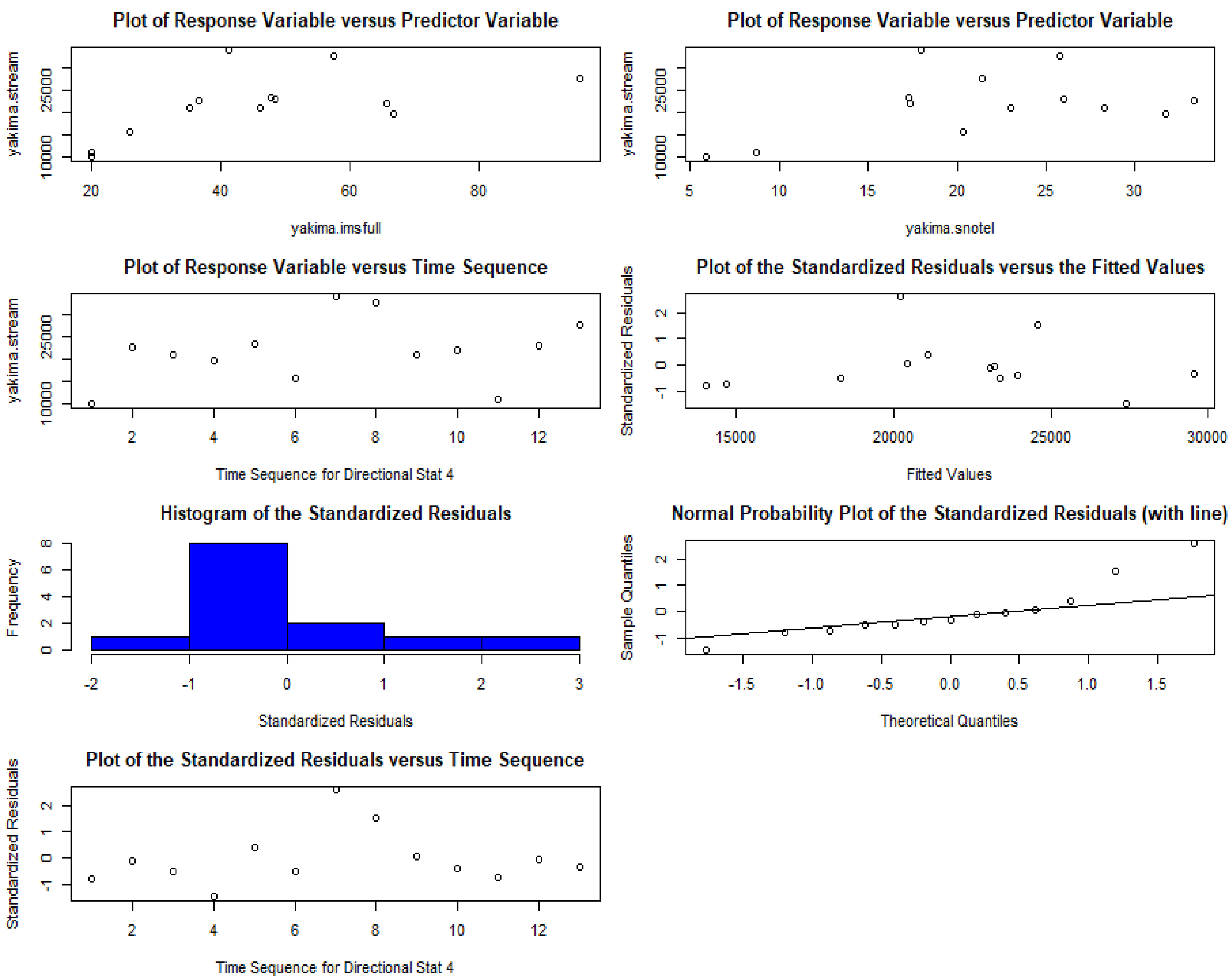

Figure H3. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the Yakima subbasin Reduced February IMS Model (4-km) with February Snow Telemetry data (SNOTEL) added as a second predictor.

Table H4

Principal Component Analysis performed for the Yakima subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (4-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Yakima Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 13 2
Y dimension: 13 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 13 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	7422	5211	5490
adjCV	7422	5185	5453

TRAINING: % variance explained

	1 comps	2 comps
X	95.14	100.00
yakima.stream	55.81	55.81

Coefficients:

1 component -

(Intercept)	10921.86
yakima.imsred	182.58247
yakima.snotel	32.10814

Multiple R-squared: 0.5581
Predicted R-squared: 0.4215

2 components -

(Intercept)	11005.51
yakima.imsred	183.86562
yakima.snotel	24.81154

Multiple R-squared: 0.5581
Predicted R-squared: 0.3579

Table H5

The Full PSC February IMS Model for the 24-kilometer resolution (top), the Reduced PSC February IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 24-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Deschutes subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

DESCHUTES SUBBASIN

Full February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	19055.46	1797.88	10.599	6.56e-09 ***
deschutes.imsfull	79.63	31.53	2.525	0.0218 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3012 on 17 degrees of freedom
Multiple R-squared: 0.2728, Adjusted R-squared: 0.23
Predicted R-squared: 0.1428
F-statistic: 6.376 on 1 and 17 DF, p-value: 0.0218

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsfull))

	Value	p-value	Decision
Global Stat	0.9328	0.9198	Assumptions acceptable.
Skewness	0.0304	0.8616	Assumptions acceptable.
Kurtosis	0.024	0.8769	Assumptions acceptable.
Link Function	0.4859	0.4858	Assumptions acceptable.
Heteroscedasticity	0.3923	0.531	Assumptions acceptable.

Reduced February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18994.22	1445.83	13.137	2.49e-10 ***
deschutes.imsred	75.89	23.18	3.273	0.00448 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2767 on 17 degrees of freedom
Multiple R-squared: 0.3866, Adjusted R-squared: 0.3505
Predicted R-squared: 0.2978
F-statistic: 10.71 on 1 and 17 DF, p-value: 0.004481

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred))

	Value	p-value	Decision
Global Stat	0.3896	0.9833	Assumptions acceptable.
Skewness	0.0001	0.9905	Assumptions acceptable.
Kurtosis	0.0692	0.7925	Assumptions acceptable.
Link Function	0.0463	0.8296	Assumptions acceptable.
Heteroscedasticity	0.2739	0.6007	Assumptions acceptable.

Reduced February IMS Model (24-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18304.39	1786.26	10.247	1.95e-08 ***
deschutes.imsred	56.75	36.73	1.545	0.142
deschutes.snotel	108.60	159.92	0.679	0.507

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2811 on 16 degrees of freedom
Multiple R-squared: 0.4038, Adjusted R-squared: 0.3293
Predicted R-squared: 0.1998
F-statistic: 5.418 on 2 and 16 DF, p-value: 0.01597
VIF: 2.43

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred + deschutes.snotel))

	Value	p-value	Decision
Global Stat	0.5312	0.9704	Assumptions acceptable.
Skewness	0.2996	0.5841	Assumptions acceptable.
Kurtosis	0.0712	0.7897	Assumptions acceptable.
Link Function	0.1452	0.7031	Assumptions acceptable.
Heteroscedasticity	0.0153	0.9017	Assumptions acceptable.

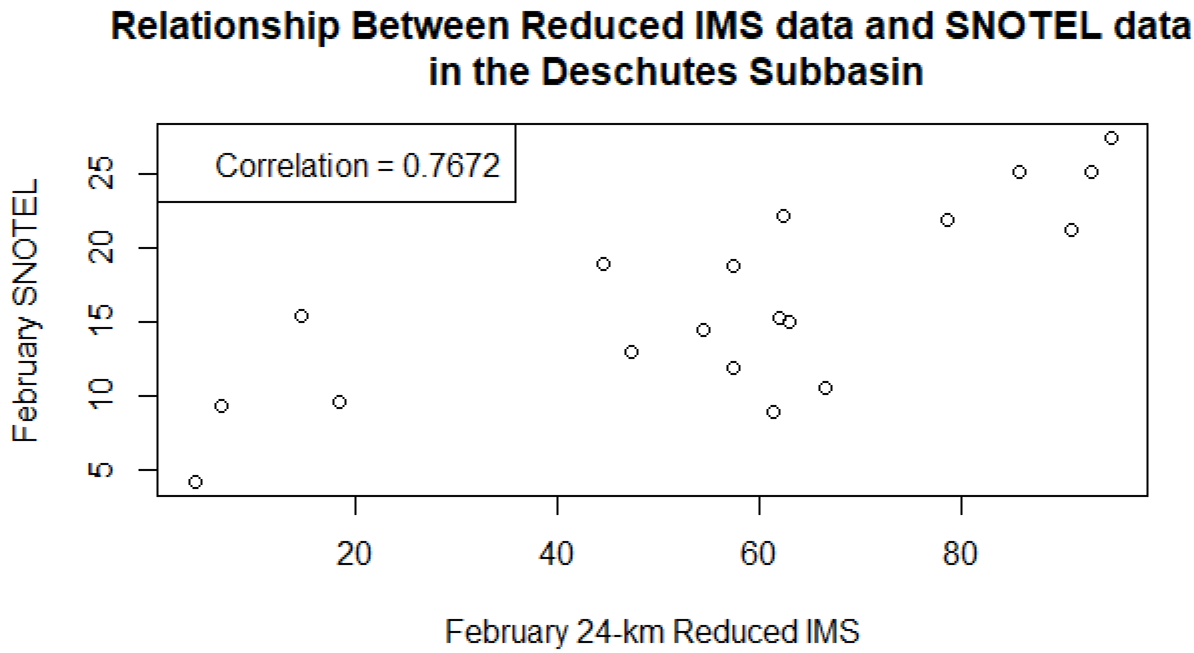

Figure H4. This correlation plot shows the correlation between the February 24-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Deschutes subbasin.

Table H6

Principal Component Analysis performed for the Deschutes subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (24-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Deschutes Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	3527	2788	2989
adjCV	3527	2783	2978

TRAINING: % variance explained

	1 comps	2 comps
X	98.00	100.00
deschutes.stream	39.05	40.38

Coefficients:

1 component -

(Intercept)	18889.48
deschutes.imsred	73.90307
deschutes.snotel	13.30435

Multiple R-squared: 0.3095
Predicted R-squared: 0.3036

2 components -

(Intercept)	18304.39
deschutes.imsred	56.74807
deschutes.snotel	108.59697

Multiple R-squared: 0.4038
Predicted R-squared: 0.1998

Table H7

The Full PSC February IMS Model for the 4-kilometer resolution (top), the Reduced PSC February IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 4-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Deschutes subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

DESCHUTES SUBBASIN

Full February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18879.05	1938.22	9.740	9.61e-07 ***
deschutes.imsfull	102.72	37.89	2.711	0.0202 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2784 on 11 degrees of freedom
Multiple R-squared: 0.4006, Adjusted R-squared: 0.3461
Predicted R-squared: 0.2277
F-statistic: 7.351 on 1 and 11 DF, p-value: 0.02025

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsfull))

	Value	p-value	Decision
Global Stat	2.6965	0.6098	Assumptions acceptable.
Skewness	2.0357	0.1536	Assumptions acceptable.
Kurtosis	0.004	0.9498	Assumptions acceptable.
Link Function	0.4653	0.4951	Assumptions acceptable.
Heteroscedasticity	0.1915	0.6617	Assumptions acceptable.

Reduced February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18497.95	1683.81	10.986	2.87e-07 ***
deschutes.imsred	86.65	25.53	3.393	0.006 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2514 on 11 degrees of freedom
Multiple R-squared: 0.5114, Adjusted R-squared: 0.467
Predicted R-squared: 0.4073
F-statistic: 11.51 on 1 and 11 DF, p-value: 0.005999

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred))

	Value	p-value	Decision
Global Stat	1.8863	0.7567	Assumptions acceptable.
Skewness	0.6275	0.4283	Assumptions acceptable.
Kurtosis	0.0087	0.9258	Assumptions acceptable.
Link Function	0.2584	0.6112	Assumptions acceptable.
Heteroscedasticity	0.9918	0.3193	Assumptions acceptable.

Reduced February IMS Model (4-km) w/ SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18861.59	2002.64	9.418	2.75e-06 ***
deschutes.imsred	101.32	47.23	2.145	0.0575 .
deschutes.snotel	-81.76	217.44	-0.376	0.7148

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2618 on 10 degrees of freedom
Multiple R-squared: 0.5182, Adjusted R-squared: 0.4219
Predicted R-squared: 0.3272
F-statistic: 5.379 on 2 and 10 DF, p-value: 0.02595
VIF: 3.15

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(deschutes.stream ~ deschutes.imsred + deschutes.snotel))

	Value	p-value	Decision
Global Stat	1.5966	0.8094	Assumptions acceptable.
Skewness	0.3767	0.5394	Assumptions acceptable.
Kurtosis	0.054	0.8163	Assumptions acceptable.
Link Function	0.1934	0.6601	Assumptions acceptable.
Heteroscedasticity	0.9726	0.324	Assumptions acceptable.

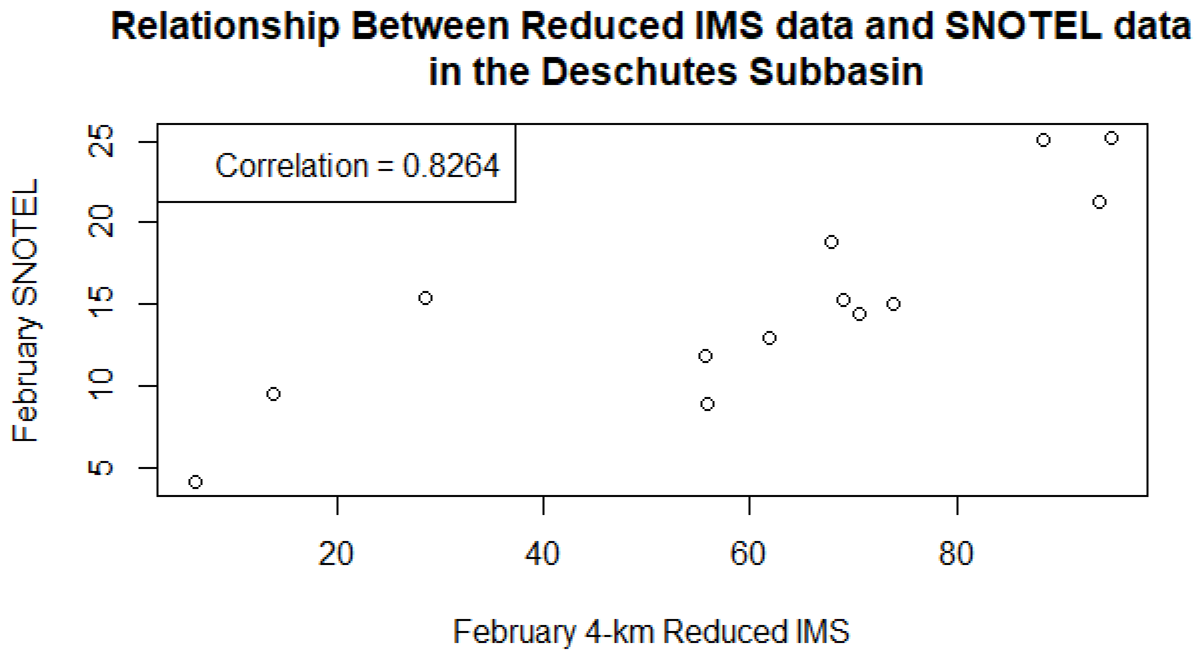

Figure H5. This correlation plot shows the correlation between the February 4-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Deschutes subbasin.

Table H8

Principal Component Analysis performed for the Deschutes subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (4-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Deschutes Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 13 2
Y dimension: 13 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 13 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	3584	2556	2714
adjCV	3584	2546	2697

TRAINING: % variance explained

	1 comps	2 comps
X	98.62	100.00
deschutes.stream	50.87	51.82

Coefficients:

1 component -

(Intercept)	18445.54
deschutes.imsred	83.65527
deschutes.snotel	15.23809

Multiple R-squared: 0.5087
Predicted R-squared: 0.4033

2 components -

(Intercept)	18861.59
deschutes.imsred	101.32387
deschutes.snotel	-81.76042

Multiple R-squared: 0.5182
Predicted R-squared: 0.3272

Table H9

The Full PSC February IMS Model for the 24-kilometer resolution (top), the Reduced PSC February IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 24-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the John Day subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

JOHN DAY SUBBASIN

Full February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6701.58	2953.39	2.269	0.0366 *
johnday.imsfull	135.67	58.75	2.309	0.0338 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5533 on 17 degrees of freedom
Multiple R-squared: 0.2388, Adjusted R-squared: 0.194
Predicted R-squared: 0.0614
F-statistic: 5.333 on 1 and 17 DF, p-value: 0.03375

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsfull))

	Value	p-value	Decision
Global Stat	12.637	0.0132	Assumptions NOT satisfied!
Skewness	4.3276	0.0375	Assumptions NOT satisfied!
Kurtosis	7.8409	0.0051	Assumptions NOT satisfied!
Link Function	0.4344	0.5099	Assumptions acceptable.
Heteroscedasticity	0.0342	0.8533	Assumptions acceptable.

Reduced February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6426.54	2829.43	2.271	0.0364 *
johnday.imsred	103.19	40.79	2.530	0.0216 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5406 on 17 degrees of freedom
Multiple R-squared: 0.2735, Adjusted R-squared: 0.2307
Predicted R-squared: 0.1057
F-statistic: 6.399 on 1 and 17 DF, p-value: 0.0216

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred))

	Value	p-value	Decision
Global Stat	12.477	0.0141	Assumptions NOT satisfied!
Skewness	4.1044	0.0428	Assumptions NOT satisfied!
Kurtosis	8.0225	0.0046	Assumptions NOT satisfied!
Link Function	0.098	0.7543	Assumptions acceptable.
Heteroscedasticity	0.2521	0.6156	Assumptions acceptable.

Reduced February IMS Model (24-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4170.84	4201.98	0.993	0.336
johnday.imsred	78.01	53.71	1.452	0.166
johnday.snotel	400.38	545.00	0.735	0.473

Residual standard error: 5481 on 16 degrees of freedom
Multiple R-squared: 0.2972, Adjusted R-squared: 0.2093
Predicted R-squared: 0.01
F-statistic: 3.383 on 2 and 16 DF, p-value: 0.05954
VIF: 1.69

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred + johnday.snotel))

	Value	p-value	Decision
Global Stat	23.696	9E-05	Assumptions NOT satisfied!
Skewness	8.3025	0.004	Assumptions NOT satisfied!
Kurtosis	14.475	0.0001	Assumptions NOT satisfied!
Link Function	0.2738	0.6008	Assumptions acceptable.
Heteroscedasticity	0.6443	0.4222	Assumptions acceptable.

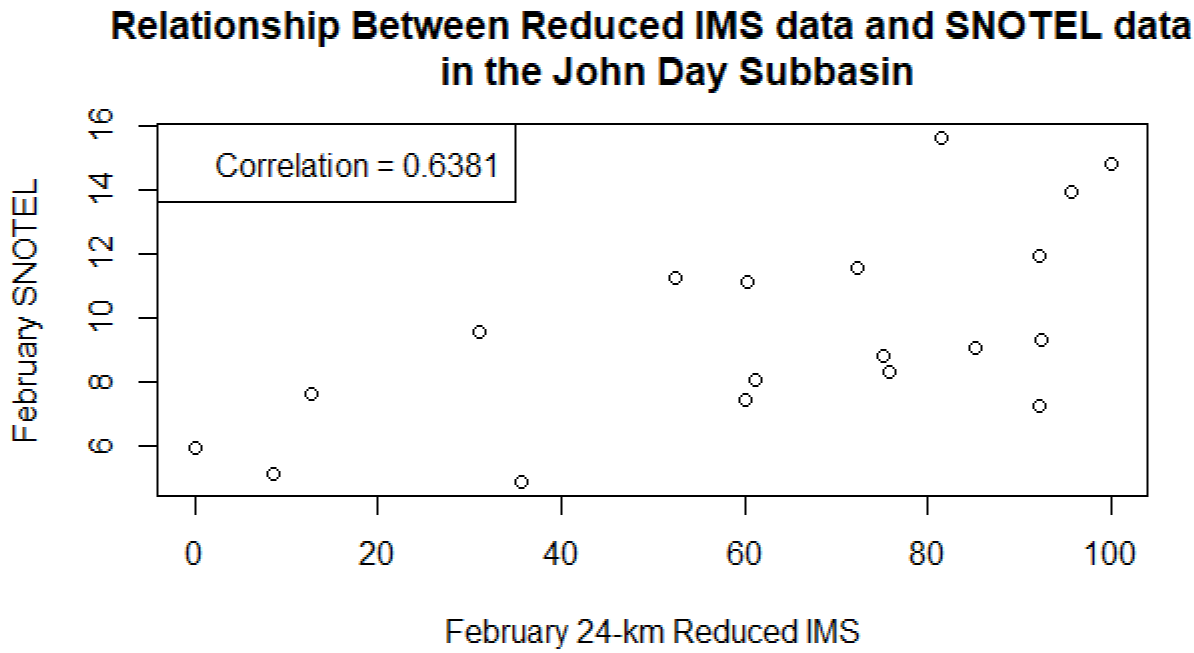

Figure H6. This correlation plot shows the correlation between the February 24-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the John Day subbasin.

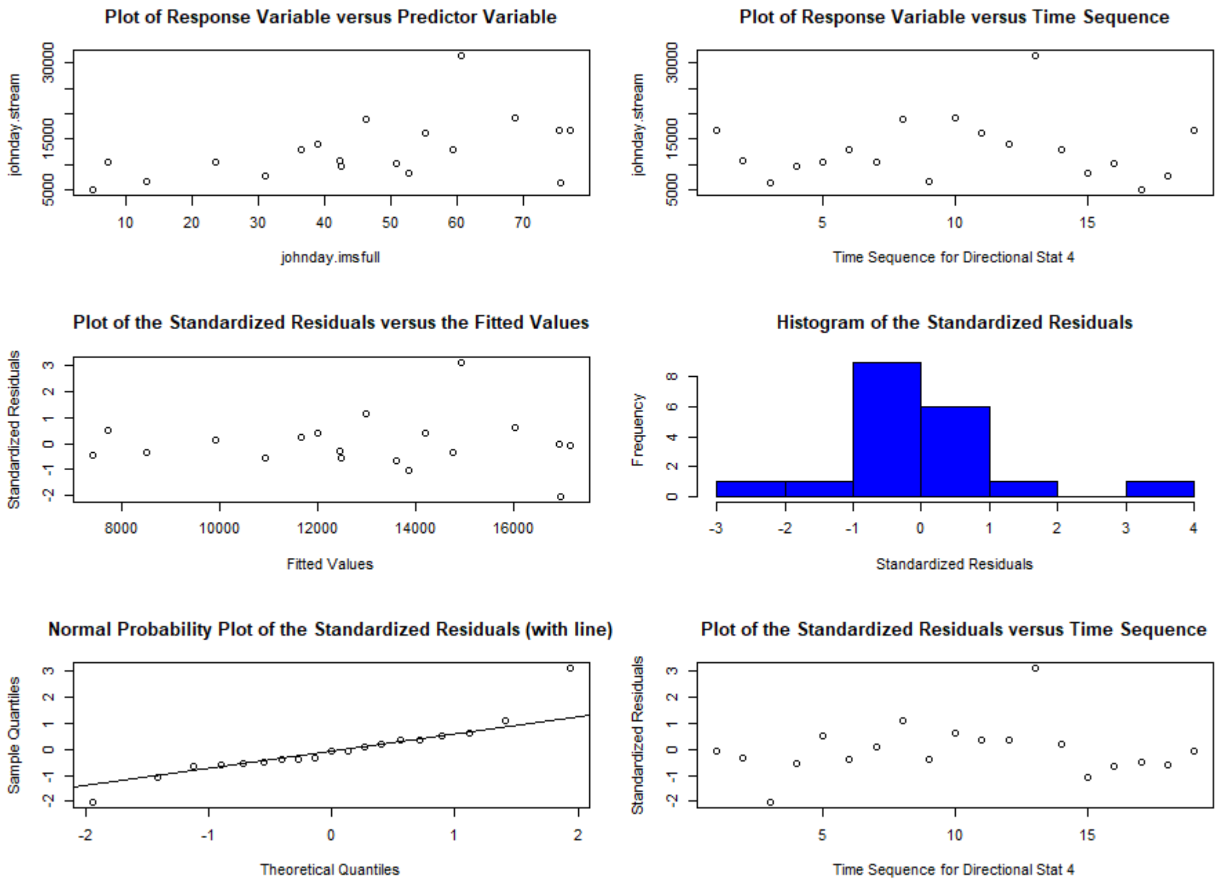

Figure H7. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Full February IMS Model (24-km).

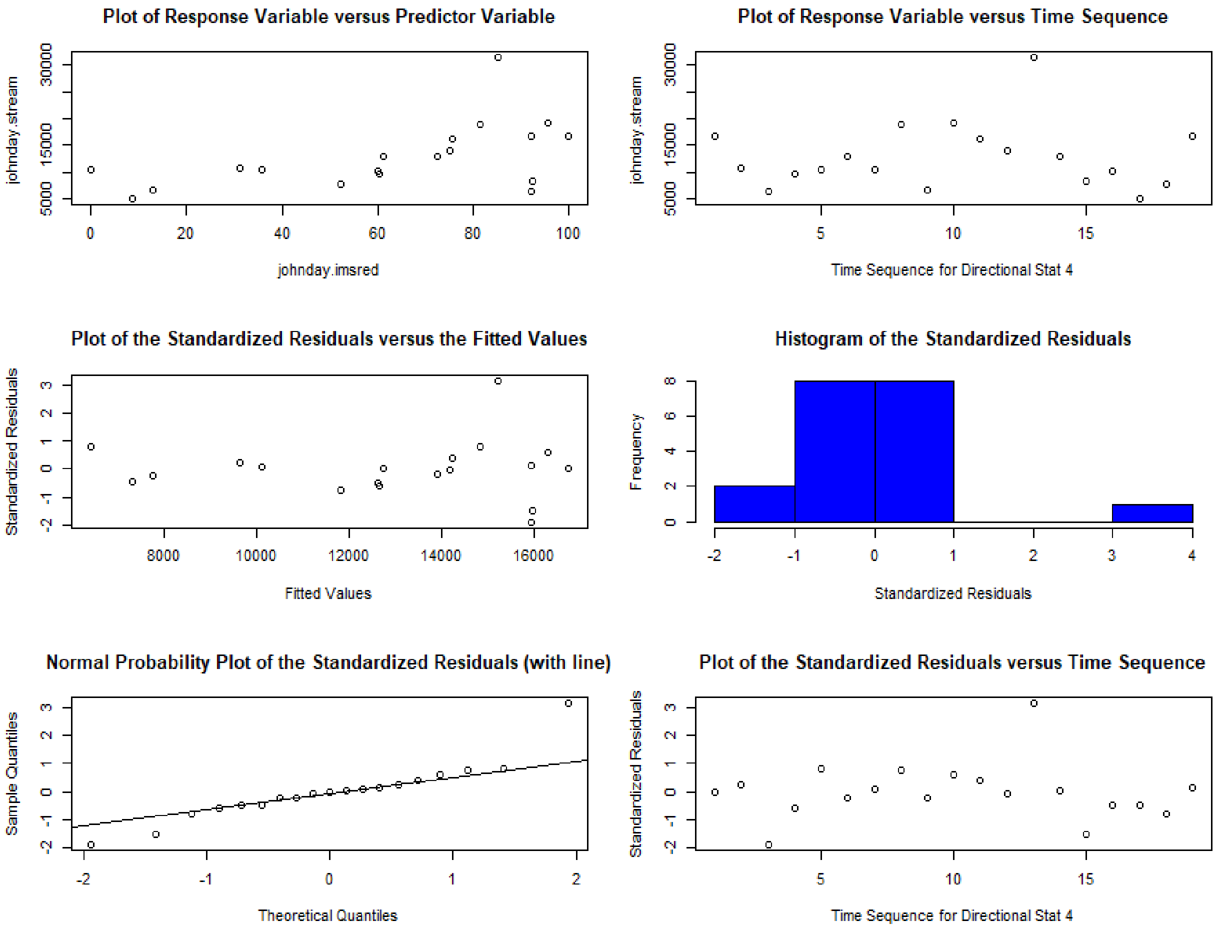

Figure H8. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Reduced February IMS Model (24-km).

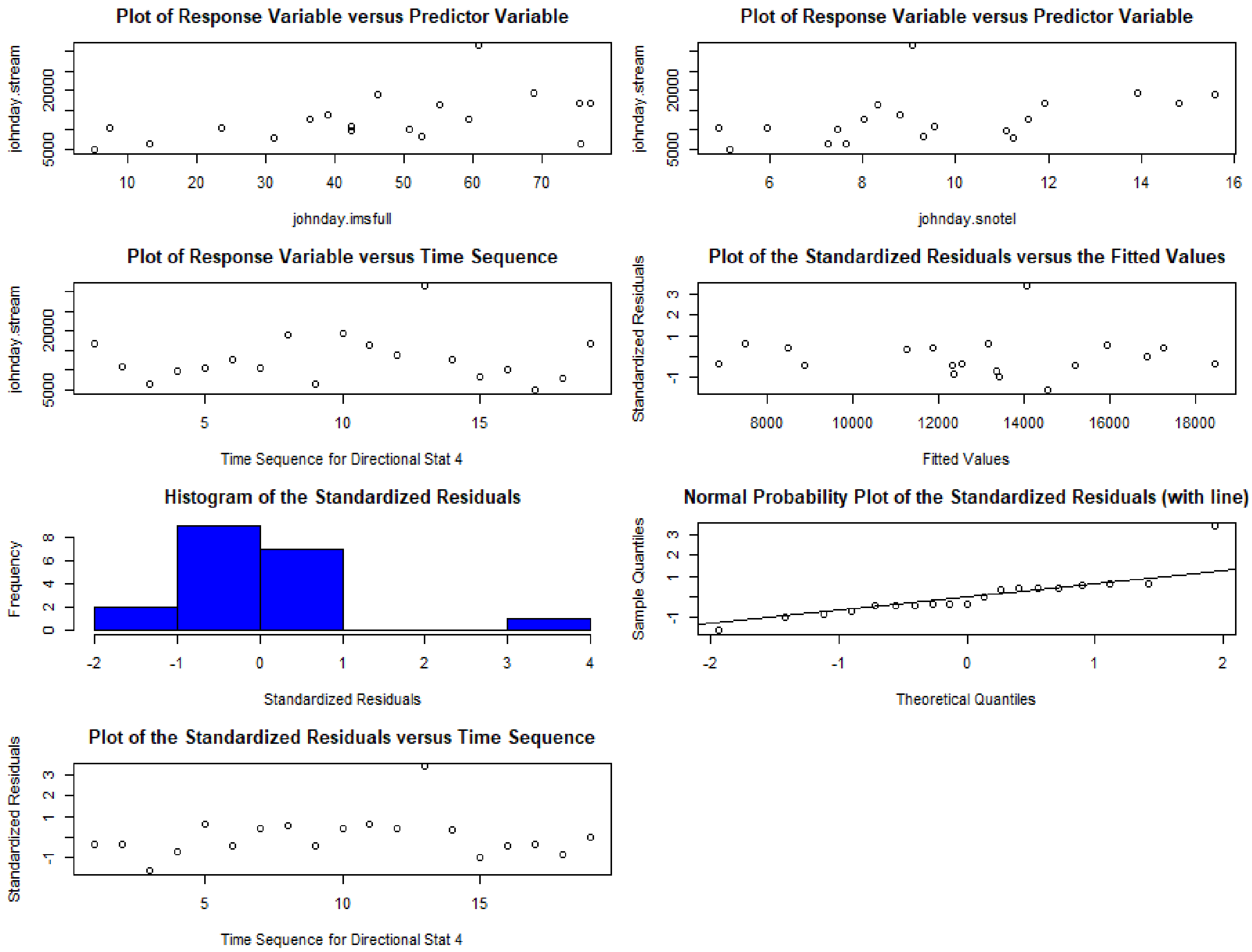

Figure H9. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Full February IMS Model (24-km) with February Snow Telemetry data (SNOTEL) added as a second predictor.

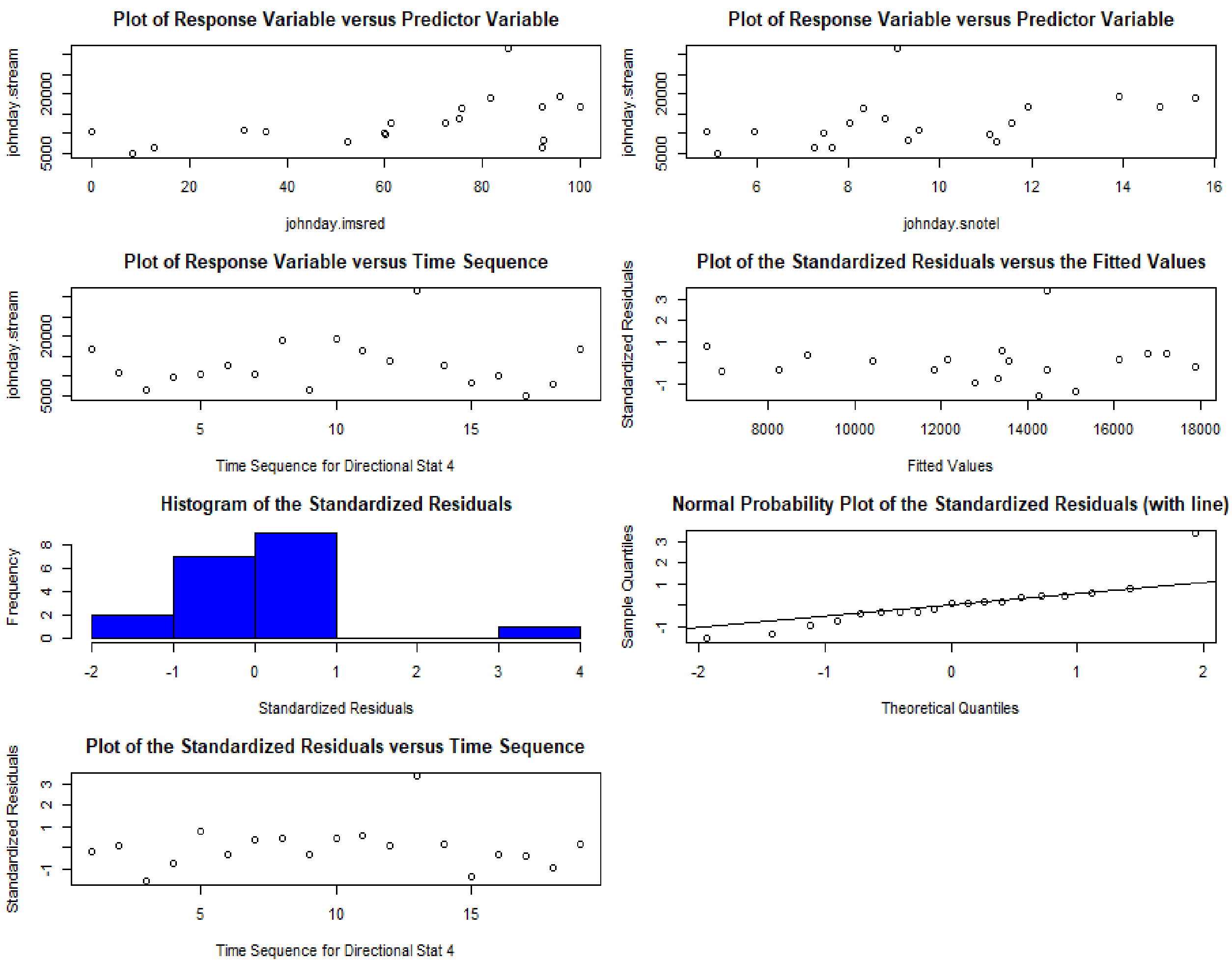

Figure H10. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Reduced February IMS Model (24-km) with February Snow Telemetry data (SNOTEL) added as a second predictor.

Table H10

Principal Component Analysis performed for the John Day subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (24-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

John Day Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	6332	5668	5969
adjCV	6332	5653	5944

TRAINING: % variance explained

	1 comps	2 comps
X	99.43	100.00
johnday.stream	27.42	29.72

Coefficients:

1 component -

(Intercept)	6380.956
johnday.imsred	102.922302
johnday.snotel	6.510439

Multiple R-squared: 0.2742
Predicted R-squared: 0.1075

2 components -

(Intercept)	4170.843
johnday.imsred	78.00799
johnday.snotel	400.37612

Multiple R-squared: 0.2972
Predicted R-squared: 0.01

Table H11

The Full PSC February IMS Model for the 4-kilometer resolution (top), the Reduced PSC February IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 4-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the John Day subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

JOHN DAY SUBBASIN

Full February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3131.51	4182.50	0.749	0.47
johnday.imsfull	219.77	80.83	2.719	0.02 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5697 on 11 degrees of freedom
Multiple R-squared: 0.4019, Adjusted R-squared: 0.3475
Predicted R-squared: 0.2262
F-statistic: 7.392 on 1 and 11 DF, p-value: 0.01997

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsfull))

	Value	p-value	Decision
Global Stat	5.5713	0.2335	Assumptions acceptable.
Skewness	3.0071	0.0829	Assumptions acceptable.
Kurtosis	2.2311	0.1353	Assumptions acceptable.
Link Function	0.117	0.7323	Assumptions acceptable.
Heteroscedasticity	0.2161	0.642	Assumptions acceptable.

Reduced February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3266.6	3781.6	0.864	0.4061
johnday.imsred	168.9	56.3	3.000	0.0121 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5464 on 11 degrees of freedom
Multiple R-squared: 0.45, Adjusted R-squared: 0.4
Predicted R-squared: 0.2694
F-statistic: 8.999 on 1 and 11 DF, p-value: 0.01208

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred))

	Value	p-value	Decision
Global Stat	7.5415	0.1099	Assumptions acceptable.
Skewness	2.4049	0.121	Assumptions acceptable.
Kurtosis	4.7395	0.0295	Assumptions NOT satisfied!
Link Function	0.0223	0.8814	Assumptions acceptable.
Heteroscedasticity	0.3749	0.5404	Assumptions acceptable.

Reduced February IMS Model (4-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3947.5	5173.3	0.763	0.4631
johnday.imsred	181.1	84.0	2.156	0.0565 .
johnday.snotel	-153.6	751.6	-0.204	0.8421

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5718 on 10 degrees of freedom
Multiple R-squared: 0.4523, Adjusted R-squared: 0.3427
Predicted R-squared: 0.1168
F-statistic: 4.128 on 2 and 10 DF, p-value: 0.0493
VIF: 2.03

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(johnday.stream ~ johnday.imsred + johnday.snotel))

	Value	p-value	Decision
Global Stat	6.3072	0.1774	Assumptions acceptable.
Skewness	1.7565	0.1851	Assumptions acceptable.
Kurtosis	4.1643	0.0413	Assumptions NOT satisfied!
Link Function	0.0268	0.8699	Assumptions acceptable.
Heteroscedasticity	0.3596	0.5487	Assumptions acceptable.

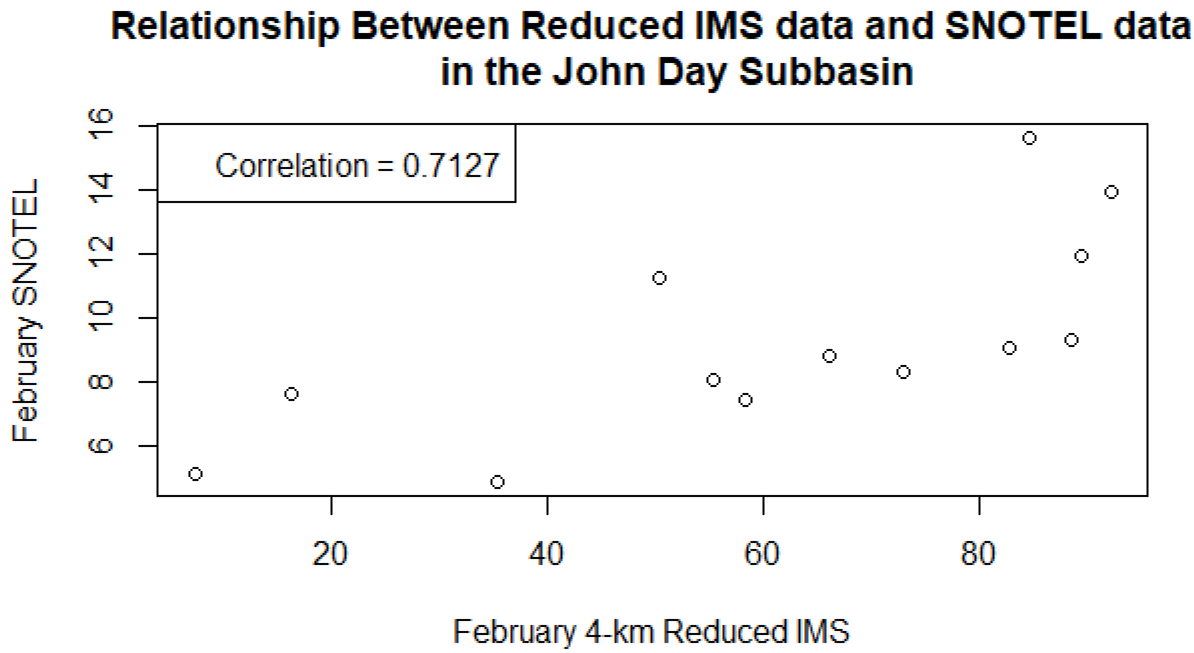

Figure H11. This correlation plot shows the correlation between the February 4-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the John Day subbasin.

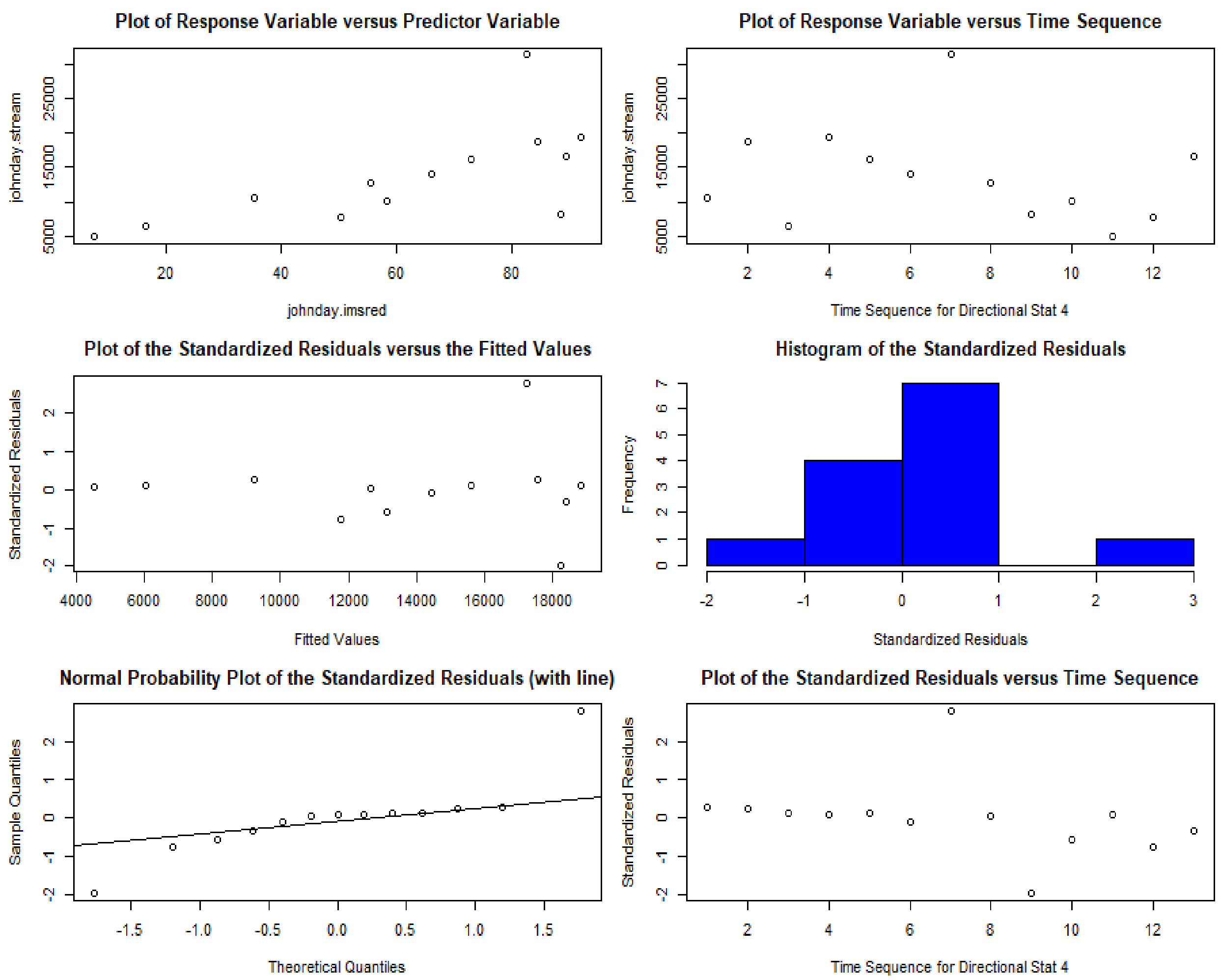

Figure H12. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Reduced February IMS Model (4-km).

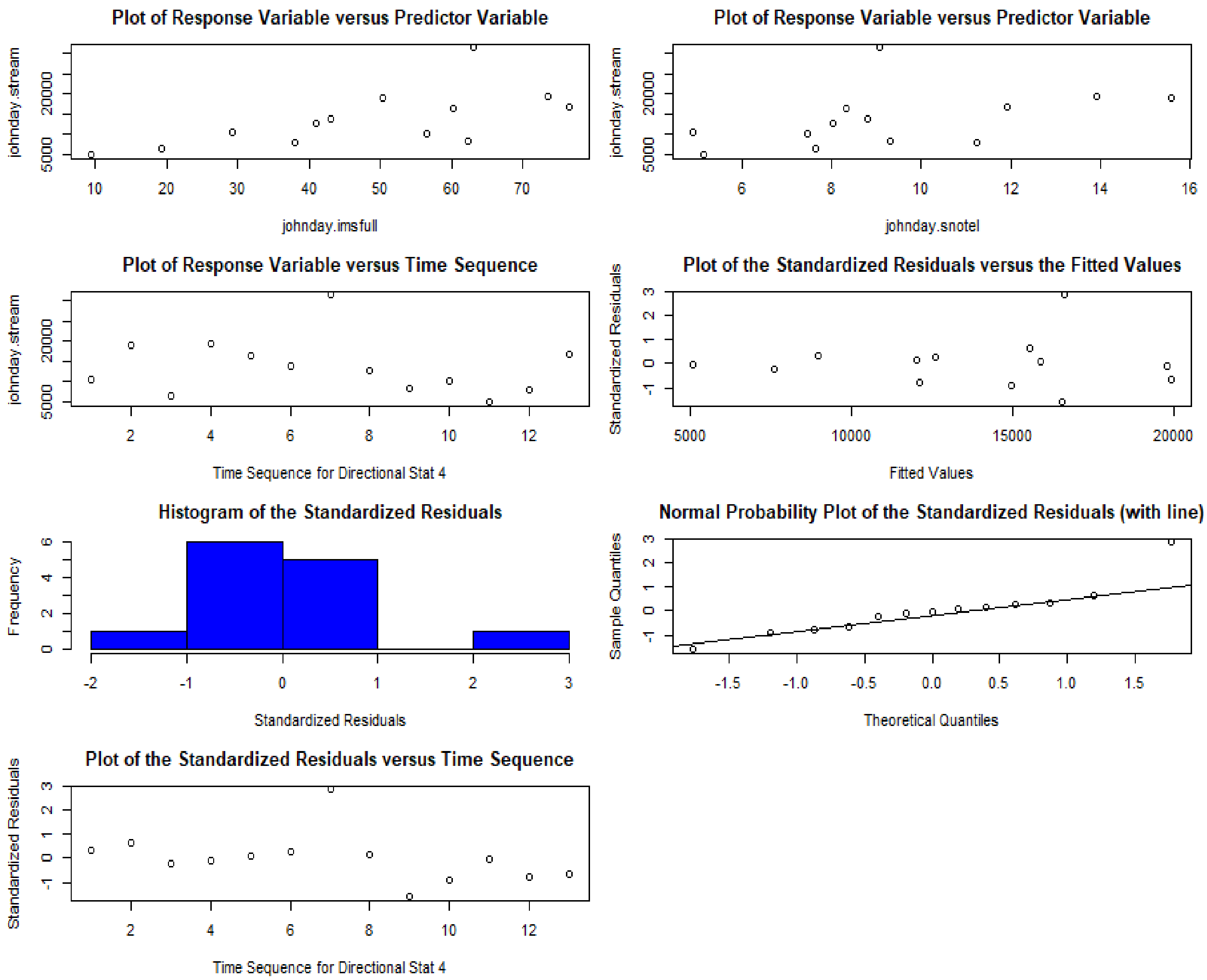

Figure H13. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Full February IMS Model (4-km) with February Snow Telemetry data (SNOTEL) added as a second predictor.

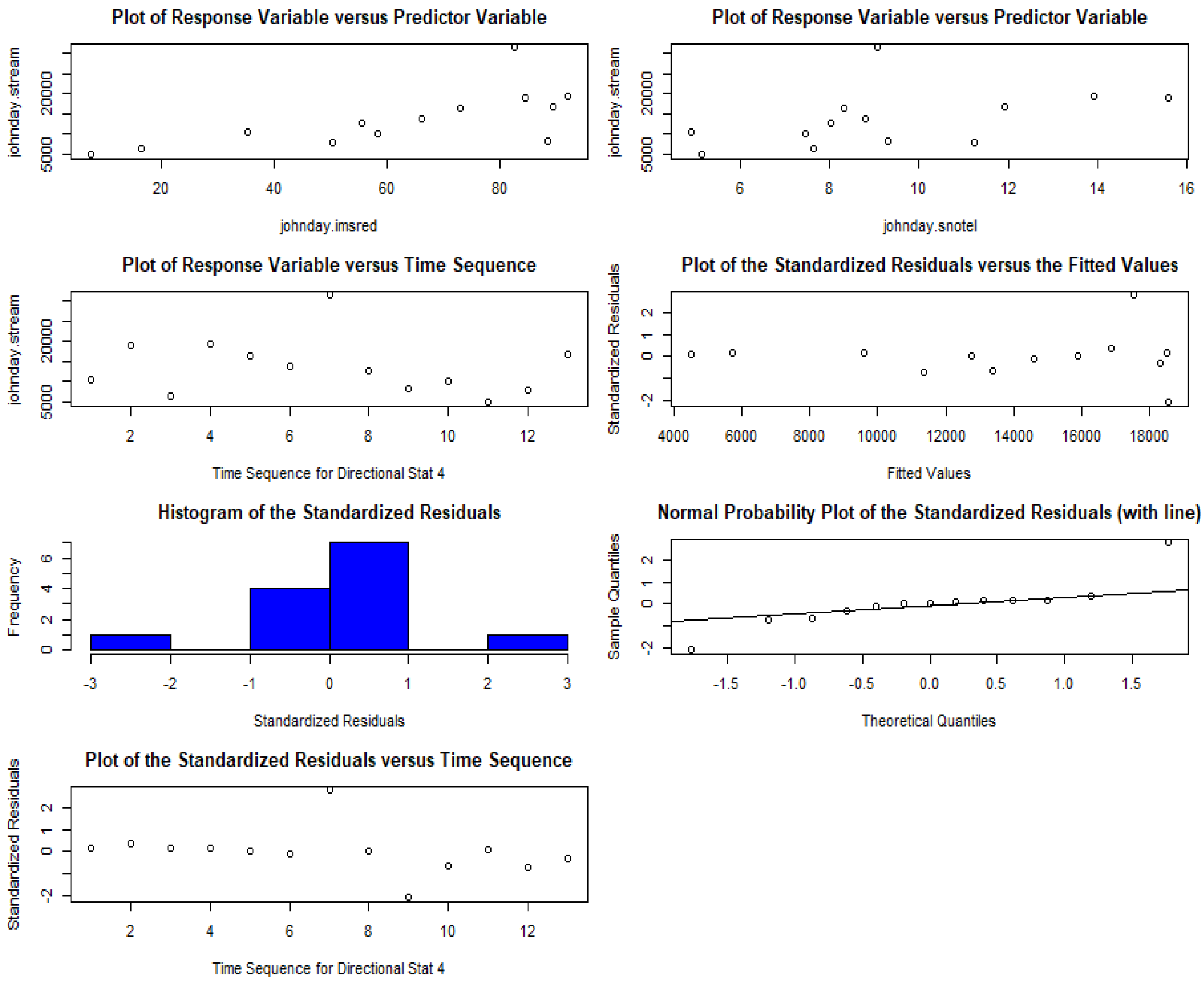

Figure H14. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the John Day subbasin Reduced February IMS Model (4-km) with February Snow Telemetry data (SNOTEL) added as a second predictor.

Table H12

Principal Component Analysis performed for the John Day subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (4-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

John Day Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 13 2
Y dimension: 13 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 13 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	7341	5788	6369
adjCV	7341	5759	6316

TRAINING: % variance explained

	1 comps	2 comps
X	99.40	100.00
johnday.stream	44.96	45.23

Coefficients:

1 component -

(Intercept)	3212.033
johnday.imsred	167.73688
johnday.snotel	13.44257

Multiple R-squared: 0.4496
Predicted R-squared: 0.2704

2 components -

(Intercept)	3947.506
johnday.imsred	181.1264
johnday.snotel	-153.6324

Multiple R-squared: 0.4523
Predicted R-squared: 0.1168

Table H13

The Full PSC February IMS Model for the 24-kilometer resolution (top), the Reduced PSC February IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 24-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Clearwater subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

CLEARWATER SUBBASIN

Full February IMS Model (24-km)

```
Coefficients:
                    Estimate   Std. Error   t value   Pr(>|t|)
(Intercept)         37897.9    16361.4      2.316     0.0333 *
clearwater.imsfull  407.0      193.9        2.099     0.0511 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16780 on 17 degrees of freedom
Multiple R-squared:  0.2058,    Adjusted R-squared:  0.1591
Predicted R-squared:  0.0631
F-statistic: 4.405 on 1 and 17 DF,  p-value: 0.05107
```

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsfull))

	Value	p-value	Decision
Global Stat	2.317	0.678	Assumptions acceptable.
Skewness	1.339	0.247	Assumptions acceptable.
Kurtosis	0.274	0.601	Assumptions acceptable.
Link Function	0.326	0.568	Assumptions acceptable.
Heteroscedasticity	0.38	0.538	Assumptions acceptable.

Reduced February IMS Model (24-km)

```
Coefficients:
                    Estimate   Std. Error   t value   Pr(>|t|)
(Intercept)         50840.6    10224.1      4.973     0.000116 ***
clearwater.imsred   272.1      126.3        2.155     0.045758 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16690 on 17 degrees of freedom
Multiple R-squared:  0.2146,    Adjusted R-squared:  0.1684
Predicted R-squared:  0.0692
F-statistic: 4.646 on 1 and 17 DF,  p-value: 0.04576
```

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsred))

	Value	p-value	Decision
Global Stat	3.768	0.438	Assumptions acceptable.
Skewness	2.026	0.155	Assumptions acceptable.
Kurtosis	0.822	0.365	Assumptions acceptable.
Link Function	0.448	0.503	Assumptions acceptable.
Heteroscedasticity	0.473	0.492	Assumptions acceptable.

Reduced February IMS Model (24-km) w/ February SNOTEL

```
Coefficients:
                    Estimate   Std. Error   t value   Pr(>|t|)
(Intercept)         29786.60   14599.62     2.040     0.0582 .
clearwater.imsred   99.64      148.46       0.671     0.5117
clearwater.snotel   1411.71    742.31       1.902     0.0754 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15540 on 16 degrees of freedom
Multiple R-squared:  0.3594,    Adjusted R-squared:  0.2794
Predicted R-squared:  0.158
F-statistic: 4.489 on 2 and 16 DF,  p-value: 0.02835
VIF: 1.6
```

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsred + clearwater.snotel))

	Value	p-value	Decision
Global Stat	4.528	0.339	Assumptions acceptable.
Skewness	2.843	0.092	Assumptions acceptable.
Kurtosis	0.455	0.5	Assumptions acceptable.
Link Function	0.415	0.519	Assumptions acceptable.
Heteroscedasticity	0.815	0.367	Assumptions acceptable.

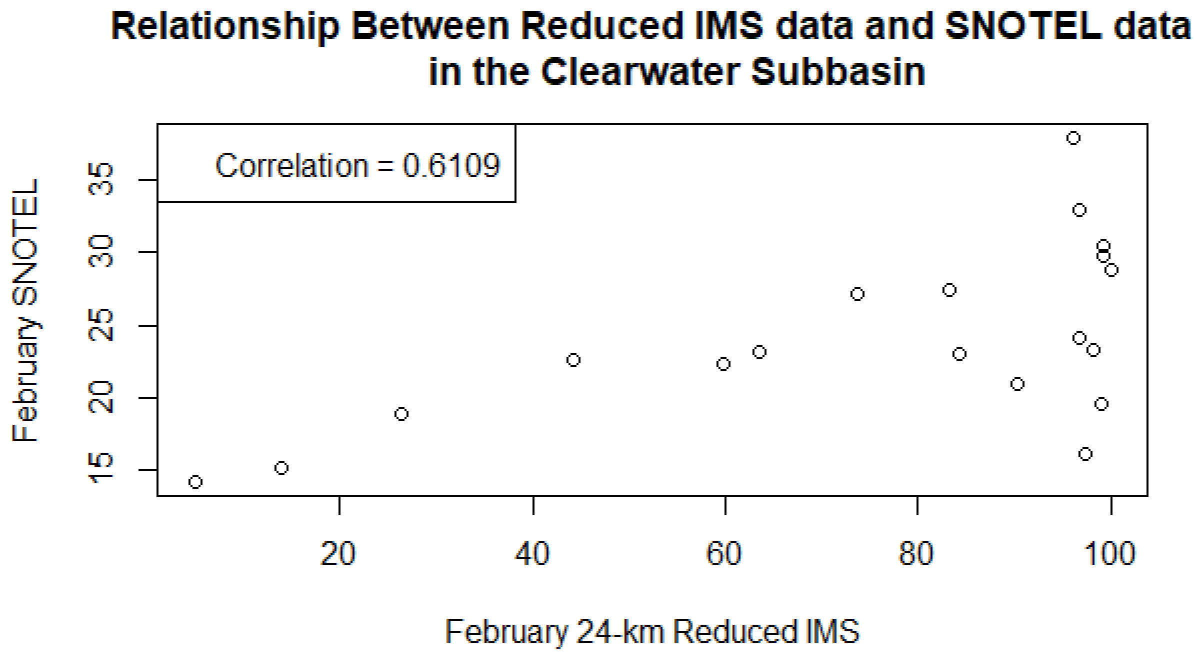

Figure H15. This correlation plot shows the correlation between the February 24-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Clearwater subbasin.

Table H14

Principal Component Analysis performed for the Clearwater subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (24-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Clearwater Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	18804	17100	16347
adjCV	18804	17063	16290

TRAINING: % variance explained

	1 comps	2 comps
X	97.63	100.00
clearwater.stream	22.15	35.94

Coefficients:

1 component -

(Intercept)	50012.11
clearwater.imsred	272.21899
clearwater.snotel	34.10129

Multiple R-squared: 0.2215
Predicted R-squared: 0.0786

2 components -

(Intercept)	29786.6
clearwater.imsred	99.64382
clearwater.snotel	1411.71022

Multiple R-squared: 0.3594
Predicted R-squared: 0.158

Table H15

The Full PSC February IMS Model for the 4-kilometer resolution (top), the Reduced PSC February IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 4-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Clearwater subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

CLEARWATER SUBBASIN

Full February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	27298.2	19417.1	1.406	0.1874	
clearwater.imsfull	590.7	251.3	2.351	0.0384	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 18030 on 11 degrees of freedom
Multiple R-squared: 0.3344, Adjusted R-squared: 0.2739
Predicted R-squared: 0.1314
F-statistic: 5.526 on 1 and 11 DF, p-value: 0.03844

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsfull))

	Value	p-value	Decision
Global Stat	0.7291	0.9477	Assumptions acceptable.
Skewness	0.5772	0.4474	Assumptions acceptable.
Kurtosis	0.0024	0.961	Assumptions acceptable.
Link Function	0.0281	0.8668	Assumptions acceptable.
Heteroscedasticity	0.1213	0.7276	Assumptions acceptable.

Reduced February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	46601.2	11160.1	4.176	0.00155	**
clearwater.imsred	394.3	159.3	2.475	0.03085	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17720 on 11 degrees of freedom
Multiple R-squared: 0.3577, Adjusted R-squared: 0.2993
Predicted R-squared: 0.1602
F-statistic: 6.126 on 1 and 11 DF, p-value: 0.03085

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsred))

	Value	p-value	Decision
Global Stat	1.1123	0.8923	Assumptions acceptable.
Skewness	0.8028	0.3703	Assumptions acceptable.
Kurtosis	0.0954	0.7574	Assumptions acceptable.
Link Function	0.0029	0.9574	Assumptions acceptable.
Heteroscedasticity	0.2112	0.6458	Assumptions acceptable.

Reduced February IMS Model (4-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	13708.1	22442.6	0.611	0.555
clearwater.imsred	173.2	199.6	0.868	0.406
clearwater.snotel	2082.1	1259.7	1.653	0.129

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16470 on 10 degrees of freedom
Multiple R-squared: 0.4955, Adjusted R-squared: 0.3946
Predicted R-squared: 0.2058
F-statistic: 4.911 on 2 and 10 DF, p-value: 0.03268
VIF: 1.82

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(clearwater.stream ~ clearwater.imsred + clearwater.snotel))

	Value	p-value	Decision
Global Stat	1.8471	0.7639	Assumptions acceptable.
Skewness	0.6273	0.4284	Assumptions acceptable.
Kurtosis	0.1093	0.741	Assumptions acceptable.
Link Function	1.0969	0.2949	Assumptions acceptable.
Heteroscedasticity	0.0136	0.907	Assumptions acceptable.

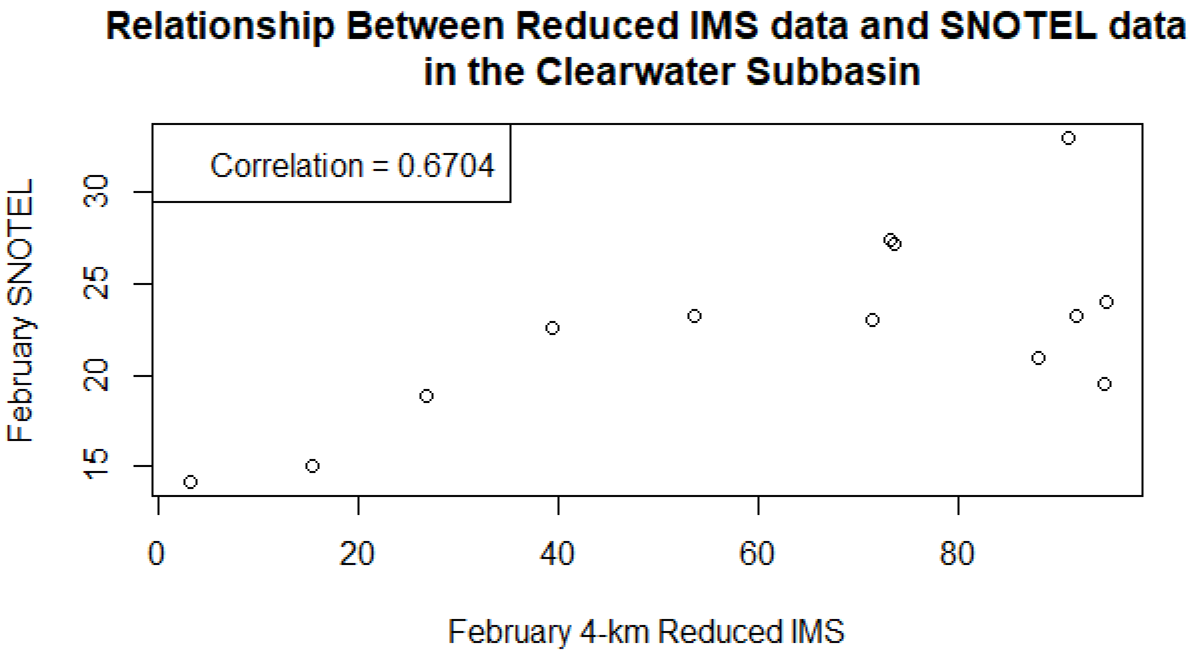

Figure H16. This correlation plot shows the correlation between the February 4-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Clearwater subbasin.

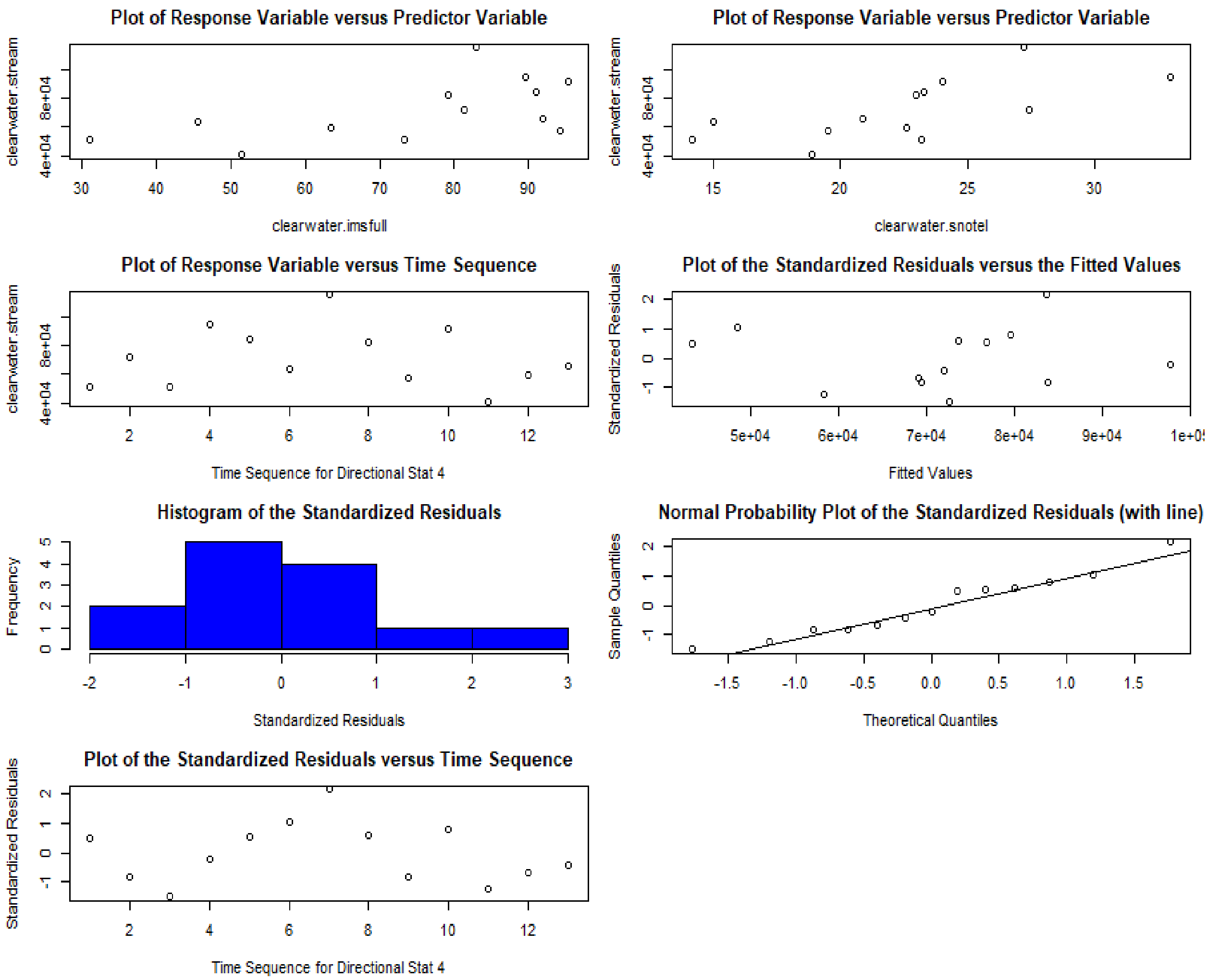

Figure H17. Plots from the residual diagnostics performed through the Global Validation of Linear Model Assumptions package in R. These plots are for the Clearwater subbasin Full February IMS Model (4-km) with February Snow Telemetry data (SNOTEL) added as a second predictor.

Table H16

Principal Component Analysis performed for the Clearwater subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (4-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Clearwater Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 13 2
Y dimension: 13 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 13 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	22028	18547	18121
adjCV	22028	18456	17975

TRAINING: % variance explained

	1 comps	2 comps
X	98.67	100.00
clearwater.stream	36.32	49.55

Coefficients:

1 component -

(Intercept)	45744.15
clearwater.imsred	392.81404
clearwater.snotel	42.29907

Multiple R-squared: 0.3632
Predicted R-squared: 0.1680

2 components -

(Intercept)	13708.09
clearwater.imsred	173.1692
clearwater.snotel	2082.0506

Multiple R-squared: 0.4955
Predicted R-squared: 0.2058

Table H17

The Full PSC February IMS Model for the 24-kilometer resolution (top), the Reduced PSC February IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 24-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Pend Oreille subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

PEND OREILLE SUBBASIN

Full February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-214.17	1395.18	-0.154	0.8798	
pendoreille.imsfull	42.48	14.98	2.836	0.0114	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 958.7 on 17 degrees of freedom
Multiple R-squared: 0.3212, Adjusted R-squared: 0.2813
Predicted R-squared: -0.4566
F-statistic: 8.044 on 1 and 17 DF, p-value: 0.0114

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsfull))

	Value	p-value	Decision
Global Stat	3.888	0.421	Assumptions acceptable.
Skewness	0.367	0.545	Assumptions acceptable.
Kurtosis	0.174	0.676	Assumptions acceptable.
Link Function	3.135	0.077	Assumptions acceptable.
Heteroscedasticity	0.212	0.645	Assumptions acceptable.

Reduced February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1195.681	777.638	1.538	0.14255	
pendoreille.imsred	29.099	8.731	3.333	0.00394	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 904.9 on 17 degrees of freedom
Multiple R-squared: 0.3952, Adjusted R-squared: 0.3596
Predicted R-squared: 0.0347
F-statistic: 11.11 on 1 and 17 DF, p-value: 0.003941

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred))

	Value	p-value	Decision
Global Stat	2.267	0.687	Assumptions acceptable.
Skewness	0.255	0.614	Assumptions acceptable.
Kurtosis	0.241	0.623	Assumptions acceptable.
Link Function	1.77	0.183	Assumptions acceptable.
Heteroscedasticity	0.001	0.974	Assumptions acceptable.

Reduced February IMS Model (24-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-98.94	1028.33	-0.096	0.9245	
pendoreille.imsred	20.34	9.56	2.127	0.0493	*
pendoreille.snotel	124.98	69.77	1.791	0.0922	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 851.3 on 16 degrees of freedom
Multiple R-squared: 0.4962, Adjusted R-squared: 0.4333
Predicted R-squared: -0.0042
F-statistic: 7.88 on 2 and 16 DF, p-value: 0.004149
VIF: 1.35

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred + pendoreille.snotel))

	Value	p-value	Decision
Global Stat	3.353	0.501	Assumptions acceptable.
Skewness	0.036	0.849	Assumptions acceptable.
Kurtosis	0.377	0.539	Assumptions acceptable.
Link Function	1.9	0.168	Assumptions acceptable.
Heteroscedasticity	1.04	0.308	Assumptions acceptable.

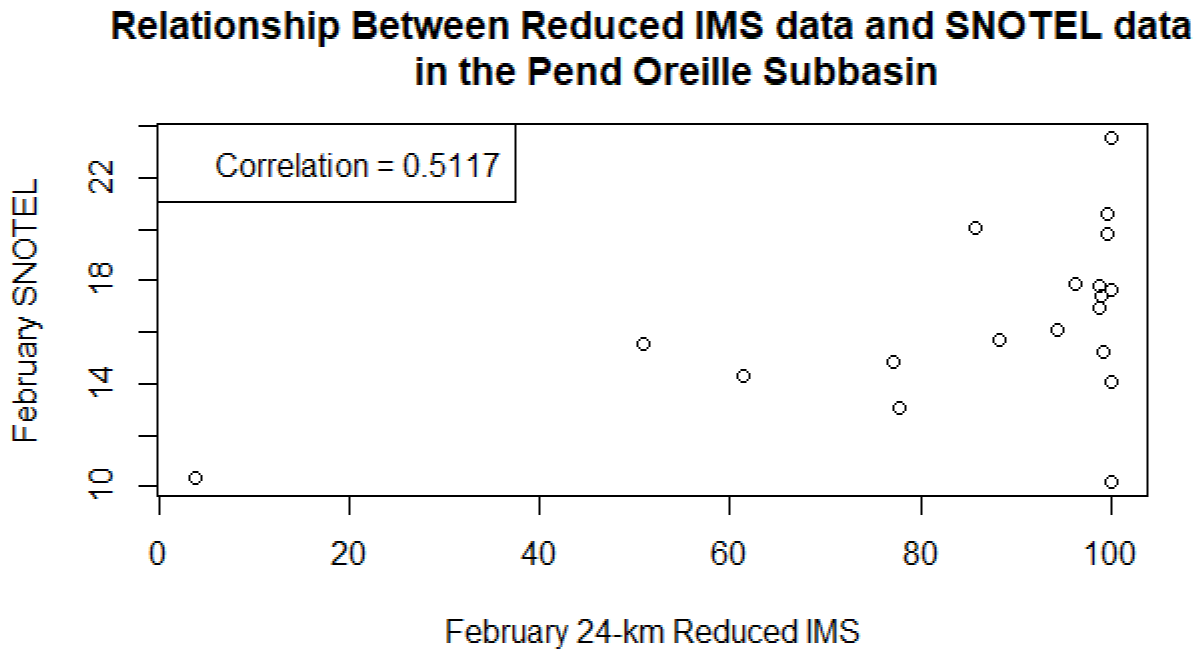

Figure H18. This correlation plot shows the correlation between the February 24-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Pend Oreille subbasin.

Table H18

Principal Component Analysis performed for the Pend Oreille subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (24-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Pend Oreille Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	1162	1082	1103
adjCV	1162	1075	1093

TRAINING: % variance explained

	1 comps	2 comps
X	98.65	100.00
pendoreille.stream	39.85	49.62

Coefficients:

1 component -

(Intercept)	1163.922
pendoreille.imsred	29.07465
pendoreille.snotel	2.06726

Multiple R-squared: 0.3985
Predicted R-squared: 0.0335

2 components -

(Intercept)	-98.9356
pendoreille.imsred	20.33534
pendoreille.snotel	124.97979

Multiple R-squared: 0.4962
Predicted R-squared: -0.0042

Table H19

The Full PSC February IMS Model for the 4-kilometer resolution (top), the Reduced PSC February IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 4-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Pend Oreille subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

PEND OREILLE SUBBASIN

Full February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-31.27	1377.99	-0.023	0.982
pendoreille.imsfull	43.61	15.53	2.808	0.017 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 988.2 on 11 degrees of freedom
Multiple R-squared: 0.4176, Adjusted R-squared: 0.3647
Predicted R-squared: -0.8245
F-statistic: 7.888 on 1 and 11 DF, p-value: 0.01701

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsfull))

	Value	p-value	Decision
Global Stat	3.82347	0.4304	Assumptions acceptable.
Skewness	0.1377	0.7106	Assumptions acceptable.
Kurtosis	0.02915	0.8644	Assumptions acceptable.
Link Function	3.25687	0.0711	Assumptions acceptable.
Heteroscedasticity	0.39974	0.5272	Assumptions acceptable.

Reduced February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1157.715	758.075	1.527	0.15495
pendoreille.imsred	32.677	9.014	3.625	0.00399 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 874.1 on 11 degrees of freedom
Multiple R-squared: 0.5443, Adjusted R-squared: 0.5029
Predicted R-squared: 0.0608
F-statistic: 13.14 on 1 and 11 DF, p-value: 0.003991

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred))

	Value	p-value	Decision
Global Stat	1.8709	0.7595	Assumptions acceptable.
Skewness	0.00761	0.9305	Assumptions acceptable.
Kurtosis	0.03749	0.8465	Assumptions acceptable.
Link Function	1.80708	0.1789	Assumptions acceptable.
Heteroscedasticity	0.01873	0.8911	Assumptions acceptable.

Reduced February IMS Model (4-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	416.86	1475.26	0.283	0.7833
pendoreille.imsred	28.30	11.87	2.384	0.0384 *
pendoreille.snotel	67.44	113.90	0.592	0.5669

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 901.1 on 10 degrees of freedom
Multiple R-squared: 0.5598, Adjusted R-squared: 0.4717
Predicted R-squared: -0.0631
F-statistic: 6.358 on 2 and 10 DF, p-value: 0.01653
VIF: 1.63

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(pendoreille.stream ~ pendoreille.imsred + pendoreille.snotel))

	Value	p-value	Decision
Global Stat	3.12E+00	0.5373	Assumptions acceptable.
Skewness	7.59E-05	0.9931	Assumptions acceptable.
Kurtosis	1.88E-02	0.891	Assumptions acceptable.
Link Function	3.05E+00	0.0807	Assumptions acceptable.
Heteroscedasticity	5.45E-02	0.8155	Assumptions acceptable.

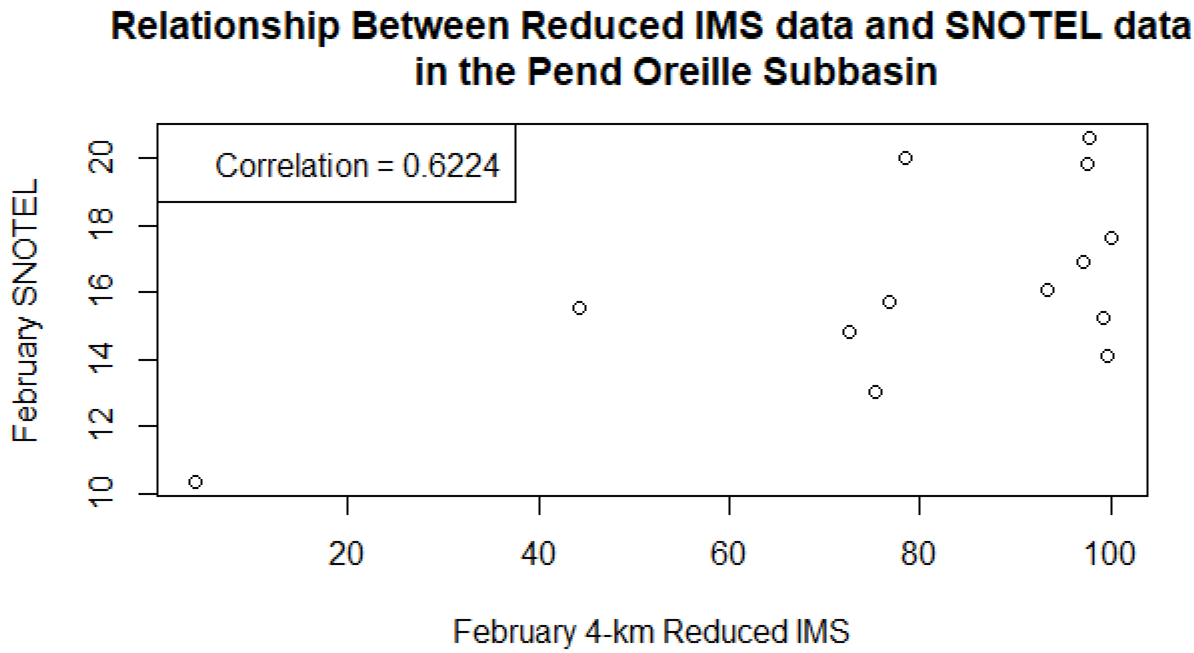

Figure H19. This correlation plot shows the correlation between the February 4-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Pend Oreille subbasin.

Table H20

Principal Component Analysis performed for the Pend Oreille subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (4-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Pend Oreille Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 13 2
Y dimension: 13 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 13 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	1290	1156	1228
adjCV	1290	1139	1209

TRAINING: % variance explained

	1 comps	2 comps
X	99.34	100.00
pendoreille.stream	54.53	55.98

Coefficients:

1 component -

(Intercept)	1132.096
pendoreille.imsred	32.567064
pendoreille.snotel	2.127072

Multiple R-squared: 0.5453
Predicted R-squared: 0.0588

2 components -

(Intercept)	416.8646
pendoreille.imsred	28.30125
pendoreille.snotel	67.43991

Multiple R-squared: 0.5598
Predicted R-squared: -0.0631

Table H21

The Full PSC February IMS Model for the 24-kilometer resolution (top), the Reduced PSC February IMS Model for the 24-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 24-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Kootenai subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

KOOTENAI SUBBASIN

Full February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-4369.36	6895.52	-0.634	0.535
kootenai.imsfull	115.11	71.16	1.618	0.124

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2513 on 17 degrees of freedom
Multiple R-squared: 0.1334, Adjusted R-squared: 0.0824
Predicted R-squared: -0.6587
F-statistic: 2.616 on 1 and 17 DF, p-value: 0.1242

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsfull))

	Value	p-value	Decision
Global Stat	2.3636	0.6874	Assumptions acceptable.
Skewness	0.3096	0.5779	Assumptions acceptable.
Kurtosis	0.1385	0.7098	Assumptions acceptable.
Link Function	1.8134	0.1781	Assumptions acceptable.
Heteroscedasticity	0.0021	0.9631	Assumptions acceptable.

Reduced February IMS Model (24-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2711.72	2334.64	1.162	0.2615
kootenai.imsred	44.84	25.17	1.781	0.0927 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2479 on 17 degrees of freedom
Multiple R-squared: 0.1573, Adjusted R-squared: 0.1077
Predicted R-squared: -0.4383
F-statistic: 3.173 on 1 and 17 DF, p-value: 0.09272

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred))

	Value	p-value	Decision
Global Stat	2.288	0.6829	Assumptions acceptable.
Skewness	0.3041	0.5814	Assumptions acceptable.
Kurtosis	0.1074	0.7432	Assumptions acceptable.
Link Function	1.8634	0.1722	Assumptions acceptable.
Heteroscedasticity	0.0133	0.9083	Assumptions acceptable.

Reduced February IMS Model (24-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-2437.10	2418.18	-1.008	0.32854
kootenai.imsred	15.14	21.90	0.691	0.49924
kootenai.snotel	440.37	132.88	3.314	0.00439 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1967 on 16 degrees of freedom
Multiple R-squared: 0.5003, Adjusted R-squared: 0.4379
Predicted R-squared: 0.0698
F-statistic: 8.01 on 2 and 16 DF, p-value: 0.003886
VIF: 1.2

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred + kootenai.snotel))

	Value	p-value	Decision
Global Stat	3.6281	0.4587	Assumptions acceptable.
Skewness	1.8412	0.1748	Assumptions acceptable.
Kurtosis	0.0003	0.9862	Assumptions acceptable.
Link Function	0.2223	0.6373	Assumptions acceptable.
Heteroscedasticity	1.5643	0.211	Assumptions acceptable.

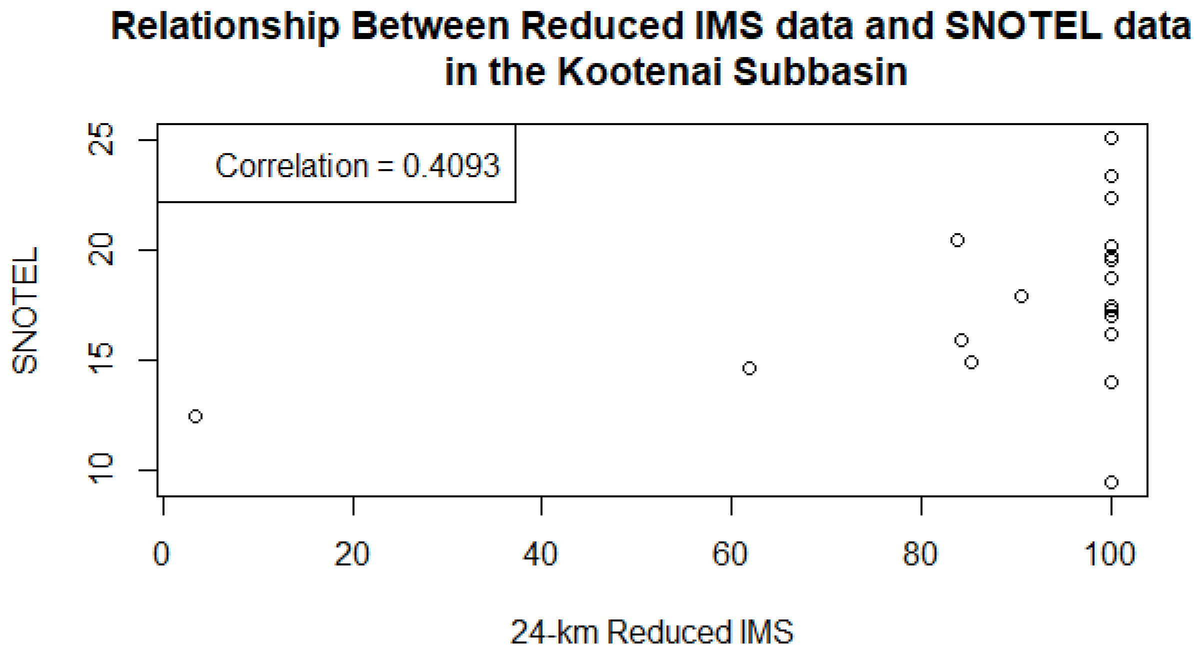

Figure H20. This correlation plot shows the correlation between the February 24-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Kootenai subbasin.

Table H22

Principal Component Analysis performed for the Kootenai subbasin at 24-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (24-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Kootenai Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 19 2
Y dimension: 19 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 19 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	2696	3117	2463
adjCV	2696	3089	2440

TRAINING: % variance explained

	1 comps	2 comps
X	97.81	100.00
kootenai.stream	16.21	50.03

Coefficients:

1 component -

(Intercept)	2614.083
kootenai.imsred	44.312136
kootenai.snotel	3.126725

Multiple R-squared: 0.1621
Predicted R-squared: -0.4896

2 components -

(Intercept)	-2437.101
kootenai.imsred	15.14031
kootenai.snotel	440.37343

Multiple R-squared: 0.5003
Predicted R-squared: 0.0698

Table H23

The Full PSC February IMS Model for the 4-kilometer resolution (top), the Reduced PSC February IMS Model for the 4-kilometer resolution (middle), and the Reduced PSC February IMS Model for the 4-kilometer resolution with February Snow Telemetry data (SNOTEL) added as a second predictor (bottom) for the Kootenai subbasin. Their respective global validation of linear model assumptions can be found on the right-hand side.

KOOTENAI SUBBASIN

Full February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-1737.84	4341.28	-0.400	0.6966	
kootenai.imsfull	97.48	47.87	2.036	0.0665	.

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2355 on 11 degrees of freedom
Multiple R-squared: 0.2738, Adjusted R-squared: 0.2078
Predicted R-squared: -0.4157
F-statistic: 4.147 on 1 and 11 DF, p-value: 0.06651

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsfull))

	Value	p-value	Decision
Global Stat	3.142	0.5344	Assumptions acceptable.
Skewness	1.3325	0.2484	Assumptions acceptable.
Kurtosis	0.1828	0.669	Assumptions acceptable.
Link Function	1.5077	0.2195	Assumptions acceptable.
Heteroscedasticity	0.119	0.7301	Assumptions acceptable.

Reduced February IMS Model (4-km)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	2549.55	2070.16	1.232	0.2438	
kootenai.imsred	53.56	23.71	2.259	0.0451	*

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2284 on 11 degrees of freedom
Multiple R-squared: 0.317, Adjusted R-squared: 0.2549
Predicted R-squared: -0.1863
F-statistic: 5.105 on 1 and 11 DF, p-value: 0.04514

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred))

	Value	p-value	Decision
Global Stat	3.1112	0.5394	Assumptions acceptable.
Skewness	1.4831	0.2233	Assumptions acceptable.
Kurtosis	0.2015	0.6535	Assumptions acceptable.
Link Function	1.392	0.2381	Assumptions acceptable.
Heteroscedasticity	0.0346	0.8524	Assumptions acceptable.

Reduced February IMS Model (4-km) w/ February SNOTEL

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-2491.46	3895.25	-0.640	0.537
kootenai.imsred	24.24	29.80	0.814	0.435
kootenai.snotel	425.50	284.03	1.498	0.165

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2165 on 10 degrees of freedom
Multiple R-squared: 0.4422, Adjusted R-squared: 0.3306
Predicted R-squared: -0.0965
F-statistic: 3.963 on 2 and 10 DF, p-value: 0.05402
VIF: 1.76

ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance = 0.05

Call:
gvlma(x = lm(kootenai.stream ~ kootenai.imsred + kootenai.snotel))

	Value	p-value	Decision
Global Stat	1.1542	0.8856	Assumptions acceptable.
Skewness	0.5523	0.4574	Assumptions acceptable.
Kurtosis	0.3357	0.5623	Assumptions acceptable.
Link Function	0.2254	0.635	Assumptions acceptable.
Heteroscedasticity	0.0409	0.8397	Assumptions acceptable.

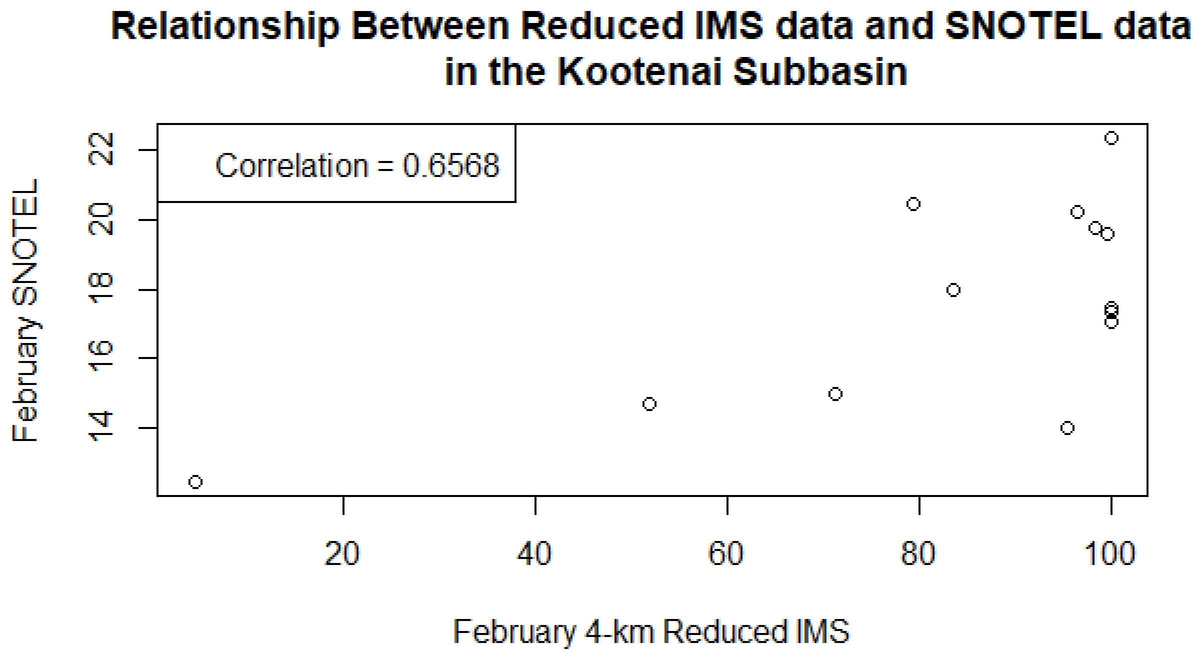

Figure H21. This correlation plot shows the correlation between the February 4-kilometer resolution Reduced PSC IMS variable and the February Snow Telemetry (SNOTEL) variable for the Kootenai subbasin.

Table H24

Principal Component Analysis performed for the Kootenai subbasin at 4-kilometer resolution. The model we performed this analysis on was the Reduced February IMS Model (4-km) with the February Snow Telemetry (SNOTEL) added as a second predictor.

Kootenai Subbasin

Principal Component Regression of Feb IMS + Feb SNOTEL

Data: X dimension: 13 2
Y dimension: 13 1
Fit method: Singular Value Decomposition Principal Components
Number of components considered: 2

VALIDATION: RMSEP

Cross-validated using 13 leave-one-out segments.

	(Intercept)	1 comps	2 comps
CV	2754	2773	2662
adjCV	2754	2741	2628

TRAINING: % variance explained

	1 comps	2 comps
X	99.37	100.00
kootenai.stream	31.91	44.22

Coefficients:

1 component -

(Intercept)	2490.456
kootenai.imsred	53.49102
kootenai.snotel	3.70914

Multiple R-squared: 0.3191
Predicted R-squared: -0.1899

2 components -

(Intercept)	-2491.461
kootenai.imsred	24.24365
kootenai.snotel	425.49723

Multiple R-squared: 0.4422
Predicted R-squared: -0.0965

Table H25

The Summary Table for all the models proposed for the month of February at the 24-kilometer resolution.

Subbasin	February Snow Signal (IMS)				February Snow Signal + February SNOTEL			
	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom
Yakima (Full)	0.107	-0.0121	0.09354	17	0.2046	0.0961	0.06246	16
Yakima (Reduced)	0.3252	0.244	0.006361	17	0.3213	0.2372	0.01754	16
Deschutes (Full)	0.23	0.1428	0.0218	17	0.2581	0.097	0.03578	16
Deschutes (Reduced)	0.3505	0.2978	0.004481	17	0.3293	0.1998	0.01597	16
John Day (Full)	0.194	0.0614	0.03375	17	0.1882	-0.0554	0.07349	16
John Day (Reduced)	0.2307	0.1057	0.0216	17	0.2093	0.01	0.05954	16
Clearwater (Full)	0.1591	0.0631	0.05107	17	0.2799	0.1702	0.02818	16
Clearwater (Reduced)	0.1684	0.0692	0.04576	17	0.2794	0.158	0.02835	16
Pend Oreille (Full)	0.2813	-0.4566	0.0114	17	0.3793	-0.3544	0.008585	16
Pend Oreille (Reduced)	0.3596	0.0347	0.003941	17	0.4333	-0.0042	0.004149	16
Kootenai (Full)	0.0824	-0.6587	0.1242	17	0.4324	-0.0644	0.004201	16
Kootenai (Reduced)	0.1077	-0.4383	0.09272	17	0.4379	0.0698	0.003886	16

Table H26

The Summary Table for all the models proposed for the month of February at the 4-kilometer resolution.

Subbasin	February Snow Signal (IMS)				February Snow Signal + February SNOTEL			
	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom	Adjusted R^2	Predicted R^2	P-Value	Degrees of Freedom
Yakima (Full)	0.2702	0.0677	0.03966	11	0.2676	0.1101	0.08469	10
Yakima (Reduced)	0.5174	0.4186	0.003361	11	0.4698	0.3579	0.01684	10
Deschutes (Full)	0.3461	0.2277	0.02025	11	0.2999	0.1403	0.06761	10
Deschutes (Reduced)	0.467	0.4073	0.005999	11	0.4219	0.3272	0.02595	10
John Day (Full)	0.3475	0.2262	0.01997	11	0.2897	0.0815	0.07264	10
John Day (Reduced)	0.4	0.2694	0.01208	11	0.3427	0.1168	0.0493	10
Clearwater (Full)	0.2739	0.1314	0.03844	11	0.3821	0.211	0.03621	10
Clearwater (Reduced)	0.2993	0.1602	0.03085	11	0.3946	0.2058	0.03268	10
Pend Oreille (Full)	0.3647	-0.8245	0.01701	11	0.3444	-0.8963	0.04867	10
Pend Oreille (Reduced)	0.5029	0.0608	0.003991	11	0.4717	-0.0631	0.01653	10
Kootenai (Full)	0.2078	-0.4157	0.06651	11	0.3138	-0.2292	0.06113	10
Kootenai (Reduced)	0.2549	-0.1863	0.04514	11	0.3306	-0.0965	0.05402	10

Table H27

These are all the models proposed for the month of March prediction period, by Predicted R^2. The models with the highest Predicted R^2 are highlighted in yellow.

Model Selection Based on Predicted R^2 - February												
Subbasin	February IMS (Full)		February IMS (Reduced)		Feb. IMS (Full) + Feb. Snotel		Feb. IMS (Reduced) + Feb. Snotel		1-Component PCA		2-Component PCA	
	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution	4-km Resolution	24-km Resolution
Yakima	0.0677	-0.0121	0.4186	0.244	0.1101*	0.0961	0.3579	0.2372	0.4215	0.2557	0.3579	0.2372
Deschutes	0.2277	0.1428	0.4073	0.2978	0.1403	0.097	0.3272	0.1998	0.4033	0.3036	0.3272	0.1998
John Day	0.2262	0.0614**	0.2694	0.1057**	0.0815	-0.0554**	0.1168	0.01001**	0.2704	0.1075	0.1168	0.01001
Clearwater	0.1314	0.0631	0.1602	0.0692	0.211*	0.1702	0.2058	0.158	0.168	0.0786	0.2058	0.158
Pend Oreille	-0.8245	-0.4566	0.0608	0.0347	-0.8963	-0.3544	-0.0631	-0.0042	0.0588	0.0335	-0.0631	-0.0042
Kootenai	-0.4157	-0.6587	-0.1863	-0.4383	-0.2292	-0.0645	-0.0965	0.0698	-0.1899	-0.4896	-0.0965	0.0698

*One linear model assumption violated

**Two linear model assumptions violated

***Three or more linear model assumptions violated