ADVANCING DECISION SUPPORT FOR WATER MANAGEMENT OPERATIONS IN THE KLAMATH RIVER BASIN

by

MITCHELL FRISCHMEYER

B.A., University of Nebraska-Lincoln, 2017

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Master of Science

Department of Civil, Environmental &

Architectural Engineering

2020

Committee Members:

Dr. Edith Zagona

Dr. Rajagopalan Balaji

Dr. Joseph Kasprzyk

Abstract

Mitchell Frischmeyer (M.S., Civil, Environmental, and Architectural Engineering)

Advancing Decision Support for Water Management Operations in the Klamath River Basin

Thesis directed by Dr. Edith Zagona

Droughts, highly uncertain forecasts, and competition for limited supply are persistent management challenges within the Klamath River Basin. As a result, the Bureau of Reclamation is subject to a recurring cycle of litigation and reanalysis over stakeholder's seasonal supply allocations. This drives the need for Reclamation to a) improve forecasts' skill and b) develop a more flexible and transparent operations management tool. Our research addresses the need for a better operations management tool by utilizing a widely used hydro policy model software, RiverWare, to build one. The design is based on Reclamation's defined functional requirements. Additionally, features are incorporated that improve the user experience through intuitive run and output management. As for improving forecasts' skill, our research investigates potentially informative climate teleconnections and alternative regression methods to reduce uncertainty. These are sea surface temperature anomaly and 700 hectopascal geopotential height signals and local polynomial and random forest regression respectively. The outcome is a climate informed version of the existing forecast that effectively reduces error at January through March lead times. The practical value of these forecasts is assessed by integrating them into operational projections. Resultant projections of competing agricultural, environmental, and flood control uses of supply have a meaningfully smaller range of error, therefore, advocating for their application.

Acknowledgements

I would like to thank my adviser Dr. Zagona for giving me the amazing opportunity to work on this research effort. Her guidance throughout the study and push to advance water management in the Klamath River Basin through the application of RiverWare were invaluable to the research's completion and positive reception from our collaborators respectively.

I would like to acknowledge David Neumann for his constant availability to answer my model development questions and general RiverWare modeling advice. The array of model features and their quality benefitted greatly from it.

Special thanks to my committee members Dr. Kasprzyk and Dr. Balaji for their suggestions that enhanced the quality of the thesis. I learned a great deal about this research's statistical forecasting techniques and approach to water management for conflicting objectives from Dr. Balaji's and Dr. Kasprzyk's classes respectively.

Bureau of Reclamation for funding the development of the KRMO under the contract 140R8120F185.

Collaboration from the Klamath Basin Area Office and Denver Technical Service Center.

Contents

Tables

Figures

Chapter 1: Introduction

Nestled among the southern pasture and marsh lands of southern Oregon spanning to the coast of northern California, the Klamath River Basin is no stranger to the struggles of water management in the western United States. Each year presents the challenge of properly distributing a dwindling and highly variable supply of water to numerous competing objectives. Transported from mountainous headwaters and trickling into streams passing through the lowlands of the region, a large portion of the spring/summer water supply originates from snowmelt runoff, approximately 75 to 85% [1] [Pasteris & Lea, 2002]. Due to the semiarid climate and the frequently dry summers, demands in March thru September rely heavily on this source. Water for demands is managed through a network pumps, canals, and dams that form the water management project called the Klamath Project. The primary control point in the Project is Upper Klamath Lake which is the largest and most upstream reservoir. The Project operator, the Bureau of Reclamation (Reclamation), controls the storage and distribution of water supply from UKL to provide irrigation water to farmlands, flood control, recreation, and ecosystems needs of Endangered Species Act listed fish [Reclamation, 2019].

Historically, conflicts over irrigation supply and ecosystems needs have been points of frequent contention. Most notable examples of this dispute involve the severe droughts of 2001 and 2002 that reached national headlines. In 2001, low forecasts drove Reclamation to shut off diversions to approximately three fourths of irrigable farmland it supplies, around 150,000 of a total 200,000 acres. In reality, forecasts underestimated supply by a substantial 70 thousand acre-feet (TAF). Experts assess the estimated cost between $27 and $46 million in agricultural losses [Boehlert and Jaeger, 2010]. To prevent similar agricultural losses in 2002, Reclamation allowed higher diversions which gave way to lower flows and increased water temperatures. By late summer, tens of thousands Chinook and coho salmon died from parasitic blooms triggered by the poor stream conditions. While no similar catastrophes have occurred

[1] Based on statistics of snowmelt and supply relationships for basins in the western United States.

since, water stress continues to be an issue. As of 2018, the Upper Klamath Basin has experienced a governor declared drought in 10 of the past 16 years [Burns, 2018]. Concerns about growing drought severity remain too as the year 2015 saw record low observations of snow water equivalent across the majority of the station monitoring network [ORWD, 2018].

All signs indicate basin mangers face increasing planning and distribution challenges. The manifestation of these challenges is perhaps best shown by how frequently operation policies change. Operations policies change through reconsultation with the United States Fisheries and Wildlife Services (USFWS) and National Marine and Fisheries Services (NMFS). Usually, it is initiated by some form of ligation over water allocations. Since 2010, three Biological Opinions resulting in new policies have been implemented with work on a fourth starting in 2020 for 2021 operations. Since Reclamation lacks a robust modeling tool, a new excel sheet-based tool must be created with every reconsultation. For daily operations, it is tedious to operate. The excel sheet-based tool is versioned daily, receives frequent adjustments, and requires manual transfers between data sets. Thus, updates to operations are tedious and hamper their ability to communicate reasoning/projections to stakeholders.

While water stress shorts demands, its effect could be mitigated or even prevented by better planning. Planning water allocations is predicated on seasonal volumetric forecasts. From January thru July, the National Resources and Conservation Service (NRCS) releases forecasts on the first day of each month. When forecasts are highly uncertain, basin managers have difficulty assessing the impacts of their management decisions. Compared to forecast products for other basins in the West, the Klamath River Basin's forecast has one of the highest standard errors. Take, for example, the April 1st forecast standard error. At UKL, it averages around 20%. Whereas, a site such as the northern Rocky Mountains averages around 10% [Risley et al., 2005]. Multiple efforts have been made to address the forecasts' inherent uncertainties. Such efforts include studies of groundwater interactions, sediment transport, ecosystem dynamics, and hydrologic forecasting methods [Wagner and Garnett, 2014; Schenk et al., 2016; Risley et al., 2018]. The latter is of chief interest to this study. In total, three papers have been published on improving forecasts in the Klamath River Basin over the past 15 years. Two focus on streamflow modeling with hydrologic simulation methods and one focuses on water supply modeling with statistical methods [Risley et al., 2005; Hay et al., 2009; Risley, 2019]. These studies investigate the ability of large-scale climate teleconnection indices (El Nino Southern Oscillation and Pacific Decadal Oscillation) as well as groundwater states, air temperatures, etc., to improve skill. Overall, the studies find that generalized indices lack strong correlations with seasonal inflow volumes, whereas the local physical metrics such as groundwater states add some value to the forecasts [Risley et al., 2005; Hay et al., 2009].

1.1. Study Objectives

Frequent reconsultations, cumbersome management tools, lack of adequate forecasts, and growing water stress remain a constant challenge to managing the Klamath Basin. To address these problems, this study has four objectives:

1. Develop a basin RiverWare model with a robust framework that adjusts to accommodate new policy logic associated with different reconsultations and an intuitive design that makes for better data, run, and output management.
2. Identify strong, basin specific climate teleconnections to seasonal volumetric inflow that can inform forecasts at various lead times.
3. Develop climate and local information-based regression models that are more skillful at forecasting seasonal volumetric inflow than the NRCS.
4. Demonstrate meaningful improvement to the quality of operational objective projections when they are based on our climate and local information-based forecasts.

1.2. Study Description

These objectives are accomplished in the following chapters:

Chapter 2 describes the development and testing of the Klamath Basin RiverWare Operations Model (KROM) including the model's requirements, design, and testing as well as stakeholder's feedback from our modeling workshop. Requirements include the extent and features, data, and policy the RiverWare model must represent and the existing daily workflow that it must support. Lastly, we give a brief overview of the stakeholder workshop and their feedback on the RiverWare operations model we presented to them.

The technical approaches and methods used to develop the forecasts are described in Chapter 3. This includes methods utilized in the study's statistical forecasting effort, which include predictive variable identification, regression model development and selection, and operational analysis of improved forecasts. How to find strong climatic teleconnections to seasonal volumetric inflow and convert the teleconnections and NRCS's forecast products into predictive variables are detailed in predictive variable identification. Then we describe, the regression techniques, i.e. local polynomial and random forest, used to create forecast models informed by those predictive variables. Model selection lays out the criteria used to compare and define the forecasts' skill. Operational analysis closes out the chapter by explaining performance metric used to judge operational projections based on our forecasts compared to the NRCS's.

Chapter 4 presents the results, organized according to the statistical forecasting methodology: best predictive variables, regression model skill, and operational performance metrics. Best predictive variables are the representations of the strongest climate teleconnections and adjusted NRCS forecast products. Regression model skill shows the rankings of the regression models' and NRCS's forecasts at each lead time. Operational performance metrics compare the quality of objective projections that are based on our forecasts or the NRCS's.

The final chapter, Conclusions and Future Work, discusses the results of the operations modeling and statistical forecasting findings. The statistical forecasting discussion focuses on why more skillful forecast were or weren't achieved and how the more skillful forecasts are useful for operators and stakeholders. Then, both operations modeling and statistical forecasting accomplishments are recapped together to summarize the accomplishments of our entire study. The chapter ends with some thoughts on future work that could improve or further analyze the results.

Chapter 2: Development and Testing of the Klamath RiverWare Operations Model

To manage operations, the Klamath Basin Area Office computes official releases and projections using a multi-sheet Excel Workbook called the Proposed Action Calculator (PA Calc). The PA Calc is based on the 2019 Basin Biological Opinion, which was developed by the United States Fisheries and Wildlife Service (USFWS). Due to frequent changes in policy from reconsultations, the Excel operations management tool regularly gets replaced by a tool tailored to the new policy. To accommodate policy changes and reduce the recurring model development burden, KBAO wants a more robust and capable tool to manage current and future operations. Additionally, KBAO would like the new operations management tool to have the following features to make daily operations more efficient and outputs more communicable with stakeholders:

1. Multiple operational runs with predeterminable settings that eliminate the trial and error setting configuration process that operators currently use.
2. Scenario operational runs that use hydrology from selected years in the period of record.
3. Custom basin reports and output products that are automatically generated and can be sent stakeholders.

With collaboration from CADSWES and TSC, KBAO decided RiverWare was a viable option for replacing the Excel operations management tool. RiverWare is a generalized river basin modeling tool used to develop and run detailed, site-specific models [Zagona et al., 2001]. Current examples of Reclamation projects that use RiverWare to manage operations include the Truckee Carson Basin and Colorado River System [Rieker et al., 2005]. As an added benefit of familiarity, this effort will leverage a RiverWare model developed collaboratively with Reclamation for a previous research project in which I implemented logic from the last Excel workbook model that reflected the 2013 biological opinion.

Covered in this chapter is the development and testing of the RiverWare model. This encompasses the step taken to lay out the model requirements, form a comprehensive design, and complete rigorous testing to ensure the model can handle day-to-day management of operations. In addition, this chapter explicitly discusses the RiverWare model's added usability features and our RiverWare model presentation at the October 2020 stakeholder workshop. The chapter starts by breaking the requirements into the extent, policy and data that the model needs to run operations and the workflow that operators follow during daily setup and use. Then, it describes our model's design in the categories of model layout, objects and methods, Rpl[2] sets, run and scenario management, output products, and adjusted workflow. The testing work that this chapter covers details the outputs compared between the two management tools (Excel and RiverWare) and the extent to which they are matched. The chapter finishes with the impressions and feedback from stakeholders at the October 2020 workshop. Each step and topic in this chapter is broken down further in the following subsections.

2.1. Model Requirements

The model requirements are the 1) extent and features, 2) data, 3) policy and 4) workflow that the RiverWare model must include to be a viable replacement to the PA Calc. Data is the per run settings, observed hydrology, and hydrology tables. Policy consists of operational logic that sets releases and projects throughout the water year. Workflow capabilities are features such as a dashboard to configure model runs, plot to visualize results, and tables to show the satisfaction of run criteria metrics. These requirements were identified through a thorough analysis of the PA Calc and the operator's daily workflow. We use the requirements later to guide our design for the RiverWare model, which members of the KBAO reviewed and confirmed. A more detailed explanation of each requirement is provided in the sections below.

[2] The coding language specific to RiverWare

2.1.1. Extent and Features

Operations consider the headwaters of the Klamath River basin starting at Chiloquin on the Williamson River all the way to the last controlled release point on the Klamath River at Iron Gate Dam. Major regions/sections on or adjacent to the Klamath River that are actively managed are as follows:

Upper Klamath Lake: The main control points that KBAO uses to release water for agricultural, environmental, and flood control objectives.

Lost River Diversion Canal: Connects the Klamath River to the Lost River System. Allows for water transfer between the two and distribution of water to agricultural area 1.

Agricultural Area 1: Agricultural lands that make up the Klamath Irrigation District (KID).

Agricultural Area 2: Refuge and agricultural lands that make up the Lower Klamath National Wildlife Refuge and Klamath Drainage District (KDD). Together, Agricultural Areas 1 and 2 are referred to as the Klamath Project.

Pacificorp Managed: The section of the Klamath River that runs from Keno Dam to Iron Gate Dam. It is managed by Pacificorp to generate hydropower. In addition to Keno and Iron Gate, there is JC Boyle and Copco 1 that are hydroelectric dams in this section.

2.1.2. Data

Data for operations is classified into three groups, which are per run settings, observed hydrology, and hydrology tables. A more comprehensive explanation of each are as follows:

Per Run Settings: Settings that the operator sets before each model run. They consist of adjustment factors and seasonal supplies for forecasts, switches for climate scenarios, exceedance percentages for table lookups, contribution percentages for accretions to Iron Gate Dam, and operation dates. Many of the settings vary by their period of application. Some apply for a week, others for a month, and some for an entire season. The operator has access to change that period along with their value.

Observed Hydrology: The hydrology that has been measured throughout the basin from the run start date until the operation start date. The operation start date can be thought of as the day at which the system is being operated, often the current day. It could also be a previous day if the system has to be re-operated due to missing data, holidays, etc. The operation start date is used to determine whether the data is in after-the-fact mode (Observed Hydrology) or is in a forecast mode. Thus, if either the run start or operation start date changes, so does the hydrology data in the model. Observed Hydrology takes many forms, it is accretions to or diversion from the Klamath River, inflows at Upper Klamath Lake (UKL), releases at Iron Gate Dam, external volumetric forecasts for the season, or etc. Reclamation stores this data in an external file or database.

Hydrology Tables: Store datasets for several forecast and release computations. Where Per Run Settings vary from one run to the next and Observed Hydrology differs with the operation or run start date, the Hydrology Tables remain constant throughout computations. Tables range from datasets for the minimum/maximum releases at the major reservoirs, flood control pool elevations, daily forecasts at multiple locations, and UKL trajectory correction factors.

2.1.3. Policy

Policy that controls operations is split between pre- and peri- computations. Pre-computations set daily forecasts and account supplies. Their logic is non-competitive and generates the data necessary for simulating operations past the point of observations. Peri-computations determine daily releases. Their logic is competitive and aims to satisfy the basin objectives when possible. An elaboration of main pre- and peri- computations are as follows:

Pre-Computations

Hydrological Forecasts: Generates the daily values for two types of hydrology; 1) the inflow at Upper Klamath Lake and 2) accretions for locations along or adjacent to the Klamath River. As forecasts, the values are computed from the operation start date through the end of the run timestep.

Demand Forecasts: Generates the daily diversion requests for the major canals and pumping stations that supply the agricultural and refuge lands. Additionally, generates daily flows that offset those areas' requests. Since these are also forecasts, the demands and offsets are computed from the operation start date through the end of the run timestep.

Account Supplies: Determines the total volume of water to distribute from Upper Klamath Lake for agricultural and environmental objectives over the Spring/Summer timeframe. This is computed on the first of each month from January through June.

Peri-Computations

Agricultural Release: Fulfills the agricultural areas' diversion requests. Other sources can supply some of these demands, this release supplies the portion of demand that remains.

Environmental Release: Keeps the river stage at a level that maintains a healthy habitat for the species in the Klamath River. It is critical that enough water is sent from UKL to sustain that stage below IGD.

Central Tendency Controlled Release: As a safeguard against excessive pool elevation drawdown at UKL, the Central Tendency Controlled Release rate allows the pool elevation to gradually rise and eventually reach the historical mean. In addition to the reservoir release, the policy imposes limits on the diversion request rates for the canals in agricultural area 2.

Ramping and Minimum Release: At any given timestep, the lowest reservoir release is the Ramping and Minimum Release. The policy defines ramping as the gradual release reduction that prevents sharp dips from day to day. The policy defines minimums as lowest allowable release that sustains the aquatic habitat.

Flood Control Release: Ensures that UKL's pool elevation does not overtop its dam. Commonly, the Flood Control Release persists for a brief period of around 4 to 15 days, which gradually decreases the pool elevation until it is under the flood control threshold elevation.

2.1.4. Workflow

Workflow is the step-by-step process that the operator uses to prepare the data, run the model and generate outputs that are used for operations or sent to stakeholders. At a minimum, the RiverWare model must be able to support the workflow followed by operators when using the PA Calc. If the RiverWare model supports more steps than the PA Calc, that would be an added benefit since that reduces the number of different tools the operator uses. An explanation of each workflow step is as follows:

1. **Acquire, Review, and Archive the Data:** The operator receives the measurements from the USGS/NWIS, Hydromet, and basin reports. Then, they validate and if needed correct the data before transferring to the PA Calc. On the data is reviewed, the operator archives the data. The data can be corrected or updated later, but the data that was used in the operational computations must be known in case of litigation. The operator performs this step in Excel workbooks that serve as the database for the daily data (one for formatting and archiving, another for exporting to PA Calc).
2. **Prepare the Model for Today's Use:** The operator transfers the observed hydrology from the database workbooks to the PA Calc. Then, the operator configures the per run settings such as the operation start date or adjustment factors that control daily hydrology and demand forecasts. Lastly, the operator enters any manual inputs or overrides. The operator performs this step with the export database workbook and PA Calc.
3. **Operate and Plan for the Season based on Observed and Projected Hydrology**: Once the operator's finished model preparation, they run the model and check results. To configure setting properly, the operator usually performs multiple iterative runs. After each iterative run, they check the results to see if hydrology and deliveries match expectations. If satisfied, the operator

sets the operational releases and reports projection to stakeholders. The operator performs this step with the PA Calc.

4. **Run Scenarios and Analyze Possibilities**: After operation are set, the operator makes scenario runs based on historical hydrology. The historical hydrology that is used for the scenarios is the accretions and/or UKL inflow. Also, to estimate the portion of UKL storage committed to agricultural deliveries, the operator runs the model with and without agricultural deliveries. Then, they compare the UKL storage over each run. The operator also performs this step with the PA Calc.
5. **Generate Reports and Output Products:** Every week, the operator creates and delivers reports to stakeholders. Stakeholders include irrigation districts, tribes, and other water users. The weekly reports that the operator generates are as follows:
    a. *USBR Daily Numbers Update*: Observed hydrology from the past week.
    b. *Klamath Project Deliveries and Demands*: Observed deliveries for the current water year.
    c. *Cumulative UKL Inflow*: The cumulative, observed UKL inflow for the current water.
    d. *Smoothed UKL Inflow*: The smoothed, observed UKL inflow for the current water year.
    e. *Wormtrails*: The observed UKL pool elevation for the current water year.

   Report *a*. is in the format of tabular data. Reports *c*. through *e*. are in the format of plots. Report *b*. has a data table and plot. Reports *b*. through *e*. include previous year's hydrology for comparison. Excel workbook of the same name as the reports are used to perform this step.
6. **Archive the Model:** Each day, the operator archives the model with the settings used to compute the official operational releases and projection. This is done to record and preserve the exact data and logic used for assignments. These archives may be need later if litigation requires review of operational decisions.

2.2. Model Design

Model design covers all aspects the of the Klamath RiverWare Operations Model (KROM), which was developed based on the requirements discussed in the prior section. We discuss the design in the order of development, starting with the model workspace and the specifications of the objects on it. The Rpl sets are discussed next, which are either global function sets, initialization rulesets, or simulation rulesets. Usability feature that make the KROM intuitive for the operator to use follows. Then, the output products that the KROM can automatically generate are presented last. Each aspect is explored in depth below.

2.2.1. Model Layout

The model layout is a spatial representation of the Klamath River Basin's physical features on the workspace. Important river reaches, reservoirs, irrigation canals, and etc. are represented as objects on the workspace, which house data and simulate physical processes (i.e. flow routing, reservoir storage, canal diversions, and etc.). The headwaters of the model start at the Williamson River since it is the most upstream location with data in the PA Calc. From the Williamson River, the layout represents locations on or directly adjacent to the Klamath River until IGD, which is where the most downstream releases are assigned. We discuss locations on or directly adjacent to the Klamath River in terms of the major sections described in "Extent and Features" above, i.e., the Lost River System, Klamath Project (Agricultural Areas 1 and 2), and Pacificorp Managed. While the sections much more physically intricate, our proposed layout is simple since it is tailored to run operations rather than comprehensively model every process. The simplified representations of these sections are discussed next, and Figure 1 shows the KROM's workspace afterwards.

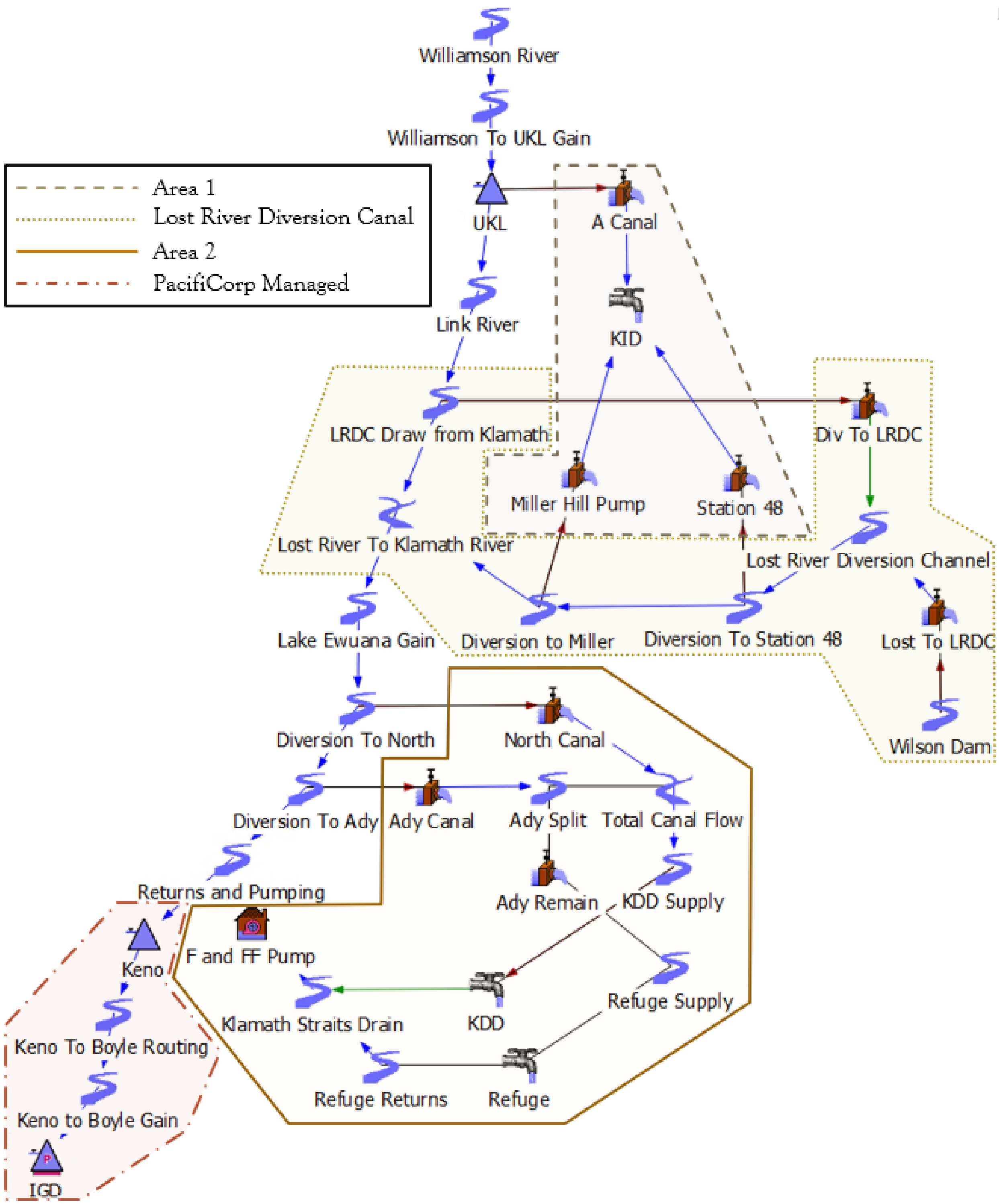


Figure 1. Workspace of the proposed KROM. Arrows signal the direction of flow.

Lost River Diversion Canal: The Lost River Diversion Canal (LRDC) system encompasses the section of the workspace that starts as the flows leaving Wilson Dam and ends at the canal's overflow into the Klamath River. Additionally, it includes the transfer from the Klamath River to the LRDC. We input and forecast the flows from the Lost River to the LRDC. Sections upstream or downstream from Wilson Dam are left unmodeled due to a lack of available logic.

Klamath Project: The Klamath Project covers the network of canals, pumps, and water users on the workspace that make up the basin's agricultural and refuge lands. These lands are classified into two regions, Agricultural Area 1 and 2. Area 1 represents the Klamath Irrigation District (KID) that A Canal, Station 48, and Miller Hill Pump supply with diversions. Area 2 represents the Refuge and Klamath Diversion District that North and Ady Canal supply with diversions. F/FF Pump returns runoff flow back to the Klamath River. Secondary reaches and diversion objects are represented in Area 2 to split flows between the two water users.

Pacificorp Managed: The facilities that PacifiCorp manages include the stretch of objects on the Klamath River that start at Keno Dam, continue through both a lag and gains reach, and end at IGD. The gains and lag represent the aggregate influx/outflux between the reservoirs and total travel time between the workspace's headwaters and tailwaters respectively. In addition to Keno Dam and IGD, PacifiCorp also manages two reservoirs between called JC Boyle and Copco 1. We exclude those from the model since PacifiCorp does not provide the logic nor data to solve for them.

2.2.2. Objects and Methods

The physical features in the Klamath River Basin are represented by seven types of objects on the workspace: Storage Reservoir, Reach, Diversion Object, Water User, Confluence, Data Object, and Inline Pump. Each object represents physical processes and contains physical data that the model uses to route flows, as part of operating policies, and forecast hydrology. Table 1 shows an inventory of the objects in the model organized by their type.

Table 1. Objects on the KROM workspace.

Objects		
Reach		
Williamson River	Ady Split	Diversion To Miller Hill
Williamson to UKL Gain	KDD Supply	Lake Ewuana Gain
Link River	Refuge Supply	Diversion to North
LRDC Draw from Klamath	Refuge Returns	Keno to Boyle Routing
Lost River Diversion Channel	Klamath Straights Drain	Keno to Boyle Gain
Diversion To Station 48	Returns and Pumping	Wilson Dam
Data Object	**Diversion**	
Dashboard	A Canal	Miller Hill Pump
Compliance Metrics	Div to LRDC	North Canal
Reservoir	Lost to LRDC	Ady Canal
UKL	Station 48	Ady Remain
Keno	**Inline Pump**	**Water User**
IGD	F and FF Pump	KID
Confluence		KDD
Lost River to Klamath River	Total Canal Flow	Refuge

Storage Reservoirs: The Storage Reservoirs represent reservoirs with a release and spillways. Since Reclamation does not manage the basin's hydropower operations, which is done by PacifiCorp, we model all the reservoirs as Storage Reservoirs. This means we exclude IGD's energy production. The reservoir's storage is a function of the pool elevation, which the model determines with the Elevation Volume Table. In the PA Calc, that table is available for UKL. For Keno and IGD, we use tables provided by Reclamation from previous modeling studies. The previous model contains physical data associated with the 2013 Biological Opinion, which the PA Calc lacks nor requires for its operation. On the Storage Reservoir, the total Outflow is composed of Spill and Release. In many models, Release represents a lower flow control structure, while Spill represents either controlled gates or uncontrolled spillway crests. In this model, we will use the Release to represent the entire outflow; spill will remain zero. Thus, we set the “Spill” method to “None”. Additionally, one reservoir (UKL) directly supplies a diversion canal. Its “Diversion from Reservoir” method is set to “Available Flow Based Diversion”.

Reaches: A Reach represents a section of the river that routes water downstream. In addition to downstream passage, Reaches may route accretions or diversions, and sometimes lag the flow. Based on these physical processes, we present the relevant methods and objects.

- *Diversions:* We set the “Diversion from Reach” method to “Available Flow Based Diversion”, which creates a diversion slot that links to the adjacent Diversion Object. The LRDC Draw from Klamath, Diversion to Miller, Diversion to Station 48, Diversion to Ady, and Diversion to North are Reaches with this configuration.
- *Accretions:* If the accretion comes as an influx from the surrounding environment, we set the “Local Inflow and Solution Direction” to “Specify Local Inflow, Solve Outflow” or “Solve Local Inflow”. This creates a slot where we can set or solve the hydrologic gain. The Reaches with “Gain” in their name possess this configuration. If the accretion comes as a return flow from a pump, we link the Reach’s return flow slot to the outflow slot of the pump, and set no method. The Returns and Pumping is the Reach with this configuration.
- *Lag:* While incremental in reality, the model represents it as one lag on the Keno to IGD Lag object. For this Reach, we set the “Routing” method to “Time Lag”. The time lag duration is 3 days, which is taken from the PA Calc, and is set on the time lag slot.

Diversion Objects: A Diversion Object represents the physical structure that diverts water from a reach or reservoir. In the Klamath Basin, this is either a canal or pump, which transports water to the irrigation districts and/or refuge lands. The two physical processes that represent the water exchange are the diversion to and outflow from the object. The Diversion Object’s diversion slot links to the adjacent reach’s or reservoir’s diversion slot. As for the irrigation district and/or refuge lands, its diversion slot links to the Diversion Object’s outflow slot. We set the Diversion Object’s methods as follows:

- “Diversion Object Solution Direction” = “Solve for Outflow”
- “Diversion Request” = “Input Diversion Request”
- “Available Flow” = “Available Flow Diversion”

Water Users: A Water User represents irrigation districts or refuge lands along the Klamath River. On the Water User, the physical processes are the diversion and return flow. The return flow is optional since there is no explicit logic that computes it in the PA Calculator. Rather, the PA Calculator assumes that the return flow contributes some unknown portion to the accretion. The model includes the return flow for Area 2’s Water Users since it can estimate the return flow proportions with algebra between the objects’ known physical processes. The algebra works because two canals divert water from the river and one pump returns it back. As for Area 1, we know the canals that divert water, but lack the location where water returns back to the river. Thus, the Water User in Area 1 does not represent the return flow. We set the methods on the Water Users as follows:

- “Diversion and Depletion Request” = “Input Request”
- “Return Flow” = “Fraction Return Flow”
- “Fraction Return Flow Input” = “Input Fraction”

Inline Pumps: An Inline Pump is RiverWare’s representation of a pumping station. We set the “Inline Pump Solution Direction” to “Solve Downstream”, which prevents looping errors. No energy consumption or head is modeled in this model.

Confluence: A Confluence represents a flow junction with two inflows and one outflow. We set the “Confluence Solution Direction” to “Solve Downstream Only” to prevent looping errors.

Data Objects: A Data Object is an object that holds user-defined data in the model. In this model, we store data on a Data Object if it applies to multiple objects or is not location specific (i.e. climate scenarios and dates respectively). These objects have no methods and do not solve. Thus, there is no physical data they require. The Dashboard data object holds slots that classify as Per Run Settings and

Account Supplies. The Compliance Metrics data objects hold post processed slots, which are discussed near the end of the report.

2.2.3. Data and Management

The data that we defined in the requirements is stored on the objects. Operators have the option to manually input data through a few features or the KROM can automatically import it. First, we will discuss the data management features. This introduces how the operator can interact with the data. Then, we will discuss the format, itemization, storage location, and management of the required data. Our discussion is broken down by the types of data management features; then, by the types of required data listed in the "Data" section above.

Management Features:

Data Management Interface (DMI): A feature that transfers data into or out of the RiverWare model from or to a database respectively. The DMI's in the KROM are connected to Excel databases since KBAO stores their data in Excel workbooks. For timeseries data, the operator specifies the time range that data is imported or exported, i.e., run start timestep through operation start date.

Scripts: A feature that automates tasks that the operator would otherwise manually perform whilst operating the model. Each individual task a script performs corresponds to a script action. Some actions ask for the operator's input and others require no interaction to set up. Each individual action can be turned "On" or "Off". This allows the operator to select the actions the script performs upon execution. Examples of data management actions the KROM's script performs are execute a DMI to import data, clear values from series slots, and set timestep parameters. Script actions that affect policy are discussed in the "Rpl Sets" section below.

System Control Table (SCT): A feature that is a tailored user interface for the operator. For example, the operator can use the SCT to access and edit data, make runs, and open scripts. The SCT has multiple sheets, which contain either series, scalar, or other slot types. For the KROM's SCT's, we organize the slots on each sheet by the operation the data is associated.

Required Data:

Per Run Settings: The operator sets these before each run. Table 2 shows an inventory of the slots that represent the Per Run Settings.

Table 2. Objects and their slots that represent the Per Run Settings.

Per Run Settings		
Object	Slot	Definition
Dashboard	Operation Start Date	Determines if the data is observed or forecasted
	Climate	Represents the climate scenario as Dry, Medium, or Wet
	Agricultural Adj	Adjustment factors for demand forecasts based on the timestep
	NRCS Forecast	External March/Month thru September inflow forecast at UKL
Object	Slot	Definition
UKL	Inflow Adj	Adjustment factors for inflow forecasts based on the timestep
	Exceedance Percentages	Represents the historical scenario the timestep is similar to, varies by timestep
Object	Slot	Definition
LRDC F/FF Pump	Contribution Percentages	Adjusts how much these object contribute to the release at IGD from March thru September
Object	Slot	Definition
LRDC F/FF Pump Lake Ewuana Gain Keno to IGD Gain	Exceedance Percentages	Represents the historical scenario the timestep is similar to, varies by timestep
	Accretion Adj	Adjustment factors for accretion forecasts based on the timestep
Object	Slot	Definition
LRDC F/FF Pump A Canal Station 48 Miller Hill Pump North Canal Ady Canal	Seasonal Supply and Adj	Supply volume and adjustment factor for demand or offset forecasts based on the season
	Distribution Percent	The portion of a demand's or offset's supply volume that is distributed on a timestep

The objects store Per Run Settings' slots as scalars or tables. If the slot is location specific, such as the seasonal volume adjustment factor for Ady Canal, that location's object stores the slot (i.e. Ady Canal.Season Adj). Otherwise, we house Per Run Settings that multiple locations use on data objects, such as the climate scenario switch (i.e. Dashboard.Climate). Apart from access on the objects, the Per Run Settings can be accessed and edited on the "Scalar Slots" and "Other Slots" sheets of the SCT named "KlamathOps2019PA".

Observed Hydrology: Hydrology measurements prior to the operation start data. Table 3 shows an inventory of the slots that represent the Observed Hydrology.

Table 3. Slots that represent the Observed Hydrology.

Observed Hydrology	
Williamson.Inflow	A Canal.Diversion
UKL.Inflow	Station 48.Diversion
LRDC.Inflow	Miller Hill Pump.Diversion
F and FF Pump.Inflow	Ady Canal.Diversion
Keno to IGD.Local Inflow	North Canal.Diversion
Keno.Outflow	Refuge.Diversion
UKL.Outflow	Miller Hill Pump.Spill
IGD.Outflow	UKL.Pool Elevation

These slots are time series since the measurements occur over the run period. The operator can automatically import the data in Table 3 for a pre-specified period by executing the "Import Hydrology" DMI. Additionally, this data can be accessed and input on the slot viewer.

Hydrology Tables: Forecast and release computations use data from these tables as an input. Table 4 shows an inventory of the Hydrology Tables' slots.

Table 4. Objects and their slots that represent the Hydrology Table. EP is short for exceedance percentages, the values in the parenthesis show the exact percentages used.

Hydrology Tables		
Object	Slot	Columns
UKL	Inflow Limit Table	Maximum / Minimum
	Max Release Table	Elevation / Max Release
	Min Release Table	Minimum Release
	Inflow Forecast Table	EP (1%, 3%, 5%, 7%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%)
	Trajectory Table	Low / Central / High / Adj
	Flood Control Table	Dry / Wet
Object	Slot	Columns
IGD	Max Release Table	Lower / Upper / Lower EWA / Upper EWA
	Min Release Table	Minimum 1 / Minimum 2
Object	Slot	Columns
LRDC	Forecast Table	EP (5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%)
F/FF Pump		EP (10%, 20%, 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 95%, 99%)
Lake Ewuana Gain		EP (10%, 20%, 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%)
Keno to IGD Gain		EP (10%, 20%, 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%)

Objects store most Hydrology Tables as periodic table slots since the rows vary by date or month. The exception is UKL's maximum release table, which is elevation based. Since these don't change frequently, the operator accesses the object to get to the relevant table (i.e. the Iron Gate Dam Maximum Release Table is on the Iron Gate Dam object.) No data objects store these since the tables are all location dependent.

2.2.4. Rpl Sets

There are three primary Rpl Sets in the KROM. These are the initialization ruleset, simulation ruleset, and global function set. The initialization ruleset is comprised of rules that are executed during the initialization phase preceding the run. These perform the pre-computations that forecast hydrology and demands as well as assign the supply for each water account. Initialization rules fire once per run. The simulation ruleset is comprised of rules that are executed on each run timestep. These perform the peri-computations that assign releases from UKL and Iron Gate Dam. Simulation rules can fire and re-fire on a timestep if their dependencies change due to another rule's execution. This feature is necessary since the rules often set conflicting values for competing objectives. The peri-computations and their order of priority and execution in the simulation ruleset from lowest to highest are 4) Environmental Release, 3) Central Tendency Controlled Release, 2) Ramping and Minimums Release, and 1) Flood Control Release. See the "Policy" section in the "Model Requirements" for a breakdown of the pre and peri-computations. Lastly, the global function set are comprised of functions that replace commonly used or complicated logic structures to make the rules easier to understand. Some functions have arguments that allow them to be manipulated/tailored for specific computations. We present the initialization and simulation rulesets in the tables below. The global function set is excluded since it focuses on explicit logic statements, which is too detailed for the scope of the paper.

Initialization Ruleset: Rules in the initialization ruleset are sorted by the policy they are associated, which are discussed individually below. While initialization rules do have priorities, we do not have different rules set a slot for the same timestep. Thus, the priorities only signal the order of execution, which is high to low index numbers. The initialization rules have two flags, the R and Z flag. These are tied to priority. For initialization rules with the R flag, the priority of slot values they set is IR. This allows those values to be overwritten by any simulation rule. For initialization rules with the Z flag, the priority of the slot values they set is 0, which prevents those values from being overwritten by any simulation rule. See Figure 2 for the KROM's initialization ruleset.

Policy & Utility Groups | Report Groups

Name	Index	Flag	Priority	On	Type
Area Initialization				✓	Policy Group
UKL Initialization	1	R	IR	✓	Rule
PacifiCorp Initialization	2	Z	0	✓	Rule
Return Flow Initialization	3	R	IR	✓	Rule
Ag Initialization	4	R	IR	✓	Rule
Lost River Initialization	5	R	IR	✓	Rule
Account Supply Tables				✓	Policy Group
Project Supply Table	6	R	IR	✓	Rule
Environmental Water Acc...	7	R	IR	✓	Rule
UKL Supply Table	8	R	IR	✓	Rule
Adj NRCS Forecast Table	9	R	IR	✓	Rule
Demand Forecasts				✓	Policy Group
Agricultural Offset Forec...	10	R	IR	✓	Rule
Agricultural Demand Fore...	11	R	IR	✓	Rule
Refuge Demand Forecasts	12	R	IR	✓	Rule
Hydrological Forecasts				✓	Policy Group
Accretion Forecasts	13	Z	0	✓	Rule
Accretion Override	14	Z	0	✓	Rule
UKL Inflow Forecasts	15	Z	0	✓	Rule
Initial Mass Balances				✓	Policy Group
Set Lake Ewauna Gain	16	Z	0	✓	Rule

Figure 2. Initialization ruleset, broken down by policy group and rules, for the KROM.

Area Initialization: These rules set slot values for timesteps prior to the operation start timestep that allow objects to dispatch or simulation rules to solve. Multiple object or regions (collection of objects) require initialization since by default their value is NaN. NaN values can prevent the objects from solving and flow not to be routed through them.

Account Supply Table: These rules set table slot values for the account supplies in the basin. To set the agricultural and environmental account volumes, the rules need to solve for two inputs beforehand. These inputs are stored in the table slots that initialization rules 8 and 9 set.

Demand Forecasts: These rules set the daily agricultural demand forecasts. This includes forecasts for diversion requests and offsets. The refuge and agricultural diversion requests have separate rules since their logic differs.

Hydrological Forecasts: These rules set the daily hydrological forecasts. This includes forecasts for basin accretions and UKL inflow. Commonly, the operator sets overrides for the accretions. We create a separate rule to assign the overrides; then, the accretion forecast rule checks for and does not set the timesteps with overrides.

Initial Mass Balance: This rule performs a simple mass balance calculation that assigns slot values that other initialization rules use as an input. The slot is the Lake Ewuana Gain.Local Inflow.

Simulation Ruleset: Rules in the simulation ruleset are sorted by the policy they are associated, which are discussed individually below. The priority of simulation rules matters since multiple can set a slot value for the same timestep. In that case, the highest priority rule assigns that value. The order of execution is still low to high priority. The flag each rule sets is R#, where # is the priority index of the rule. Lower #'s override higher #'s. Also, the rules highlighted in blue set a slot that allows an object to dispatch. See Figure 3 for the KROM's initialization ruleset.

Policy & Utility Groups | Report Groups

Name	Priority	On	Type
Flood Release		✓	Policy Group
Set IGD Flood Release	1	✓	Rule
Set UKL Flood Release	2	✓	Rule
Ramping and Minimums Release		✓	Policy Group
Set Ramping and Minimums IG...	3	✓	Rule
Central Tendency Controlled Release		✓	Policy Group
Set UKL Supply Diversion Control	4	✓	Rule
Set Central Tendency IGD Rel...	5	✓	Rule
Set Central Tendency Diversio...	6	✓	Rule
Compute Storage Diff Ratio 5-...	7	✓	Rule
Compute PacifiCorp Balance	8	✓	Rule
Environmental Release		✓	Policy Group
Set Environmental UKL Release	9	✓	Rule
Set Environmental IGD Release	10	✓	Rule
Compute Summer Release	11	✓	Rule
Compute Spring Release	12	✓	Rule
Compute Fill Release	13	✓	Rule
Compute IGD Spawn Release	14	✓	Rule
Compute EWA Remain	15	✓	Rule
Compute EWA Used thru Yest...	16	✓	Rule
Compute Link Release Difference	17	✓	Rule
Compute UKL Credit	18	✓	Rule
Project Supply		✓	Policy Group
Set Klamath to LRDC	19	✓	Rule
Set Project Supply UKL Release	20	✓	Rule
UKL Inflow Processing		✓	Policy Group
Compute Smoothed Inflow	21	✓	Rule

Figure 3. Simulation ruleset, broken down by policy group and rules, for the KROM.

Flood Release: These rules set the flood release at Iron Gate Dam and UKL. They only execute when the UKL pool elevation exceeds the flood curve. The highest priority rules since they prevent dam overtopping.

Ramping and Minimums Release: This rule sets the ramping and minimums release at Iron Gate Dam. There is no rule to set this release at UKL since the minimum release is factored into the UKL environmental release logic and there are no ramping operations at UKL.

Central Tendency Controlled Release: These rules set the central tendency controlled release at Iron Gate Dam and UKL. Additionally, they reduce the diversion requests if UKL's pool elevation is low enough. Rules 8 and 9 compute slot values that are inputs to one or multiple of the other rules. Those values are an elevation difference ratio (between current and central tendency elevation) and daily borrow/payback volume respectively.

Environmental Release: These rules set the environmental release at Iron Gate Dam and UKL. Since the environmental release's objective varies by season, so does the logic. Thus, rules 11-14 separately compute their objective's release. Then, based on the season, rules 9 and 10 assign the appropriate release. Rules 15-18 compute slot values thar are inputs to one or multiple of the other rules. Those values are the remaining and used volume of the EWA, daily release difference from UKL, and extra volume available for release respectively.

Project Supply: These rules set the agricultural release at UKL and the transfer of water from the Klamath River to the LRDC. The transfer occurs when there is not enough water in the LRDC to meet diversion requests at Miller Hill Pump and Station 48. No agricultural release occurs at Iron Gate Dam since the agricultural areas are all upstream of it.

UKL Inflow Processing: This rule computes the smoothed inflow that is as an input for other simulation rules. The observed inflow at UKL is determined from a mass balance. Since there are many unknown fluxes at UKL, the computed inflow can greatly vary from day to day. Thus, some rules need the smoothed value to stop the erroneous variance being carried through the rest of the computations.

2.2.5. Run and Scenario Management

The KROM has two modes for runs, single and multiple. Each uses the same types of data management features for setup. First, a script loads the Rpl sets, opens the associated SCT, sets the timestep parameters, clears data from previous runs, and imports the observed hydrology. Then, a SCT is used to enter any necessary manual input (overrides, reported data) and configure the per run settings. See the Appendix for the names of the script and SCT used for single and multiple run setup.

As the name implies, the single run performs one run. It begins once the operator selects "Start" on the run control. The multiple runs are a few to many runs that occur successively and are based on preset configurations. The operator selects the configuration they want to run from the multiple run manager (MRM) and begins them by selecting "Start". We developed three multiple run configurations for the KROM, which are discussed in detail below. Each are of the iterative MRM mode, which allows per run settings to be set and results stored in the model for each run. Additionally, each configuration needs initialization rules to assign values for iterative runs. These rules are discussed below as well and shown in Figure 4.

Policy & Utility Groups | Report Groups

Name	Index	Flag	Priority	On	Type
Run Trials MRM				✓	Policy Group
Assign Demand Settings	18	Z	0	✓	Rule
Assign Inflow Forecast Settings	19	Z	0	✓	Rule
Assign Accretion Forecast Settings	20	Z	0	✓	Rule
Assign Seasonal Forecasts	21	Z	0	✓	Rule
Run Scenarios MRM				✓	Policy Group
Assign Historical Inflows	22	Z	0	✓	Rule
Assign Historical Accretions	23	Z	0	✓	Rule
Assign Agricultural Adjustments	24	Z	0	✓	Rule

Figure 4. Initialization rules that assign slot values for the multiple run configurations.

Multiple Run Configurations:

Run Trials of Different Settings: The operator defines the number of runs and per run settings for each run. See the "Data and Management" section above for a complete list of the per run settings. The operator uses this configuration to determine which per run setting values compute the official operational releases. This is an iterative process that requires checking results and tuning per run settings as the operator sees fit. Performing this with multiple runs rather than single runs expediates the process greatly. Rules 18-21 assign the per run settings defined for each run; their names correspond to which settings they assign.

Run Historical Hydrology Scenarios: The operator defines the number of runs and type(s) of hydrological scenario. The scenarios are 1) use historical UKL inflows instead of forecasts and 2) use historical accretions instead of forecasts. For this configuration, the operator can elect to perform either scenario or both. The historical hydrology is pulled from a specified year in the period of record, which goes from 1981 through the present. For the number of runs, the operator selects a year for each. Note that the historical hydrology is only set after the operation start date and not before. Rules 22 and 23 assign the daily historical values for UKL inflow and accretions respectively. Rules that forecast the same hydrology do not execute in this configuration, i.e. when historical UKL inflows are assigned, rule 15 does not execute.

Run with and without Agricultural Deliveries: Performs two runs; one with and another without agricultural releases from UKL. All other settings remain the same. In the run without deliveries, the project supply volume remains as UKL storage. The operator compares the two run's UKL storage to estimate the portion UKL holds for agricultural demands throughout the rest of the year. Rule 24 shuts off UKL's agricultural release by setting the agricultural adjustment factors to zero.

2.2.6. Output Products

After the model runs, some of the daily observations and projections are post processed into seasonal or monthly volumes. These are used by the operator to evaluate their management decisions as well as give an overall impression of the current and potential future state of the basin to stakeholders. To make regular and post processed results easier to assess and communicate, we developed some output products that the operator can generate. The output products fall into one of the three categories: Plots, Model Report, or Output Canvas. The specific product in each category are presented below.

Plots: Visualize the pool elevation at UKL and release at IGD along with other associated values. Examples are shown in Figure 5 and Figure 6, respectively.

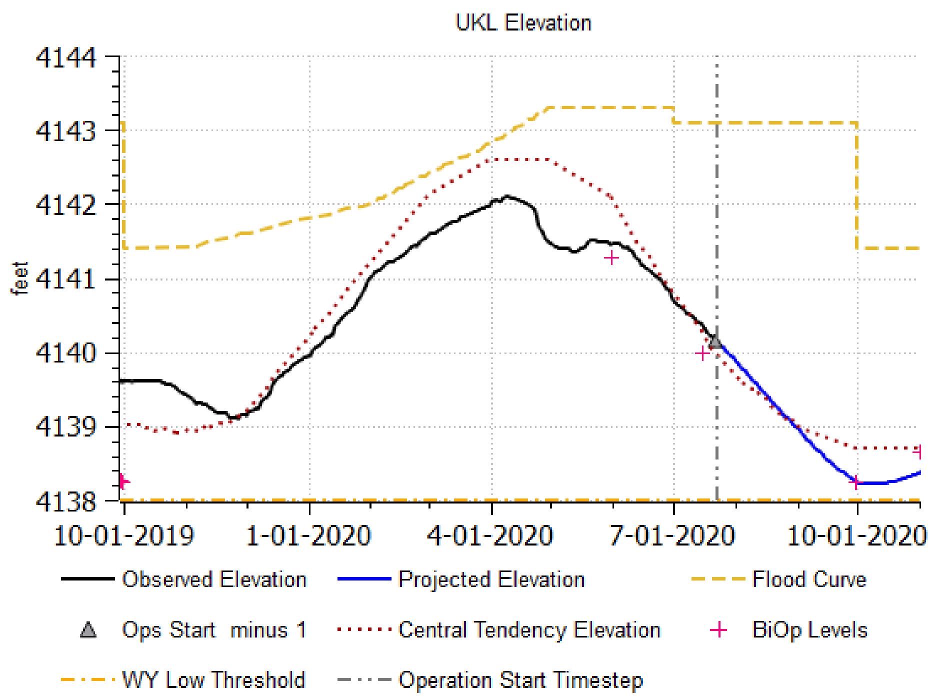


Figure 5. Sample plot of the UKL pool elevation. Includes lines and points for dates, trendlines, and thresholds.

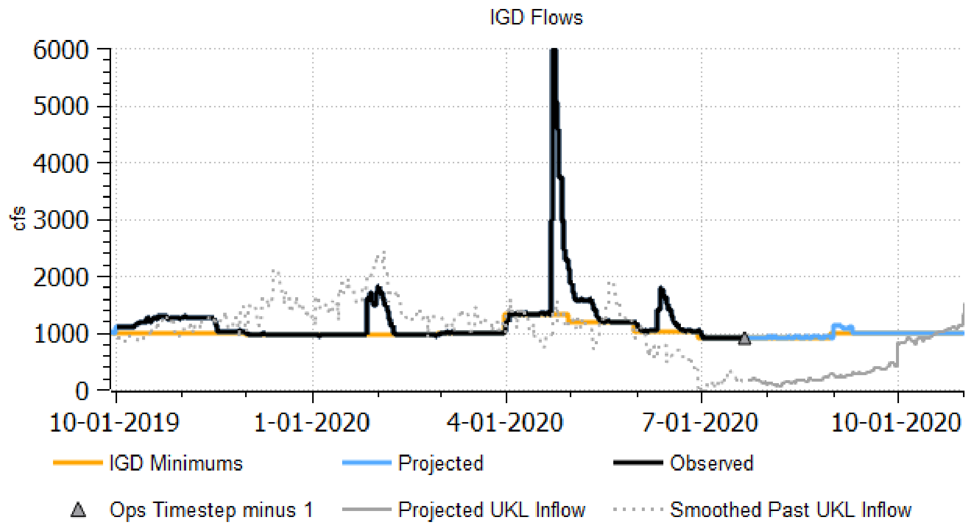


Figure 6. Sample plot of the IGD release. Includes lines and points for dates and UKL inflow.

Each has time on the x-axis, which spans the entire water year. In addition, the operator can zoom in on the UKL pool elevation plot to see more detail. The operator may do this for the month of February if they want to track how UKL fills before the agricultural diversions begin in the spring. The operator can view as well as edit the plots from the plotting feature.

Model Report: Organizes the basin hydrology from the past week into a tabular report that is generated and sent to stakeholders. The report is flexible and can be tailored in content and format by the operator to include other data they want to communicate. See Figure 7 for the model report.

USBR Daily Numbers Update

	UKL Pool Elevation feet	UKL Storage acre-feet	UKL Outflow cfs	Keno Power Canal Flow cfs	A Canal Diversion cfs	Keno Outflow cfs	Iron Gate Outflow cfs	Wilson Dam Outflow cfs	Station 48 Diversion cfs	Miller Hill Pump Diversion cfs	LRDC Flow cfs
01-25-2020	4,140.70	330,724.91	329.00	0.00	0.00	546.00	965.00	0.00	0.00	0.00	210.83
01-26-2020	4,140.76	335,655.28	317.00	0.00	0.00	736.00	964.00	0.00	0.00	0.00	507.55
01-27-2020	4,140.80	338,987.03	286.00	0.00	0.00	1,120.00	1,590.00	0.00	0.00	0.00	750.60
01-28-2020	4,140.85	343,151.72	243.00	0.00	0.00	1,040.00	1,710.00	0.00	0.00	0.00	552.51
01-29-2020	4,140.90	347,359.35	247.00	0.00	0.00	939.00	1,570.00	0.00	0.00	0.00	333.94
01-30-2020	4,140.96	352,421.38	481.00	0.00	0.00	793.00	1,470.00	0.00	0.00	0.00	226.55
01-31-2020	4,141.00	355,818.85	598.00	0.00	0.00	793.00	1,700.00	0.00	0.00	0.00	191.26

Figure 7. Sample model report of the past week's hydrology.

Output Canvas: Presents an organized view of metrics that reflect the status of operations and hydrology in the basin. We developed two output canvases; they are named Run Metrics and Compliance Metrics. Run Metrics present results in tables for the operator to review after each run. Additionally, the operator can generate a html of the Run Metric canvas to send to stakeholders. It has many of the metrics of interest to stakeholders, i.e., supply remaining for agricultural deliveries and projected UKL inflow volume per month. Compliance Metrics presents results that are used to gauge compliance with the policy's criteria. If the metrics fail to satisfy the criteria, the operator must adjust per run settings to generate results that comply. The Run Metrics and Compliance Metrics canvas is shown in Figure 8 and Figure 9 respectively.

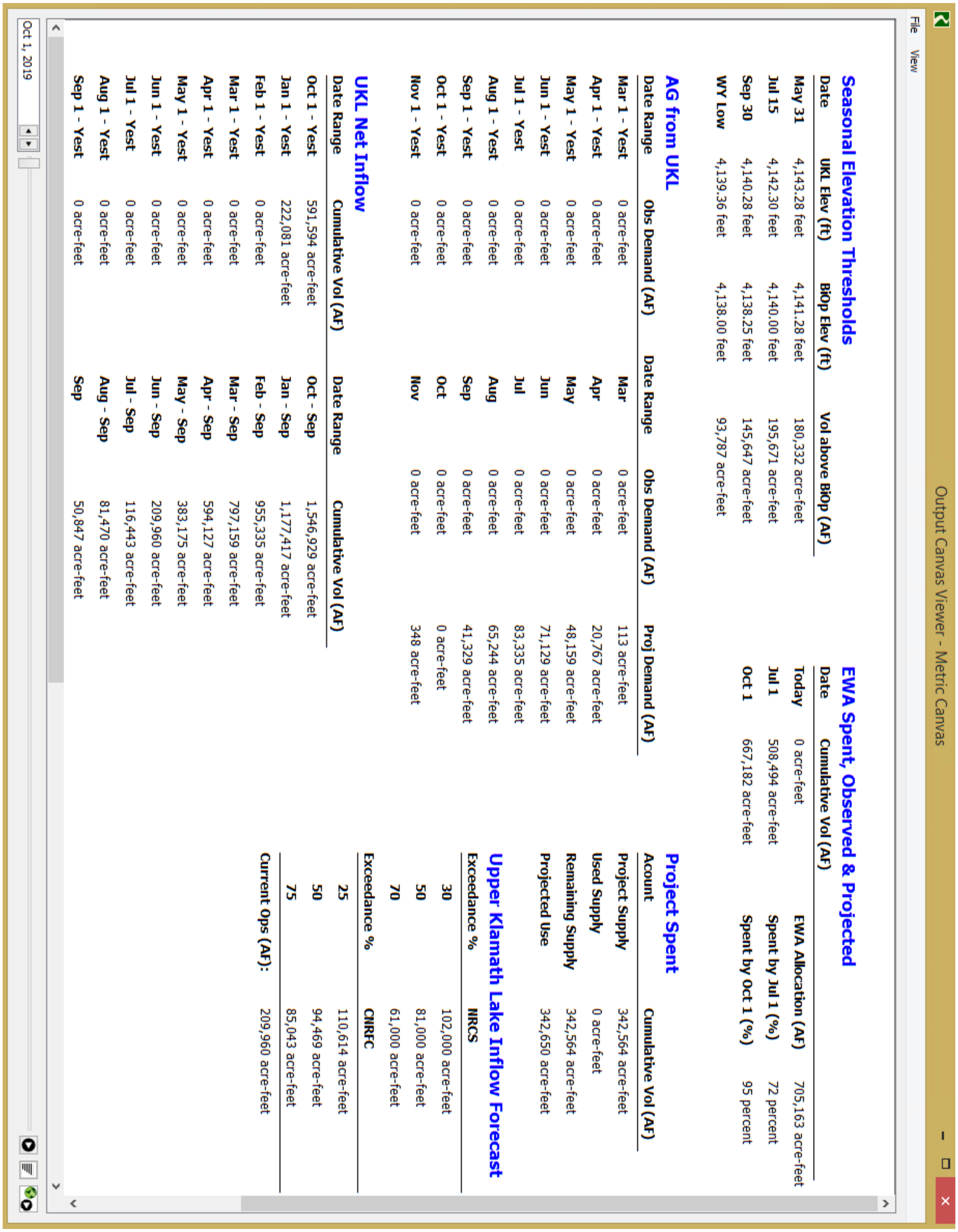


Figure 8. Sample Run Metric canvas.

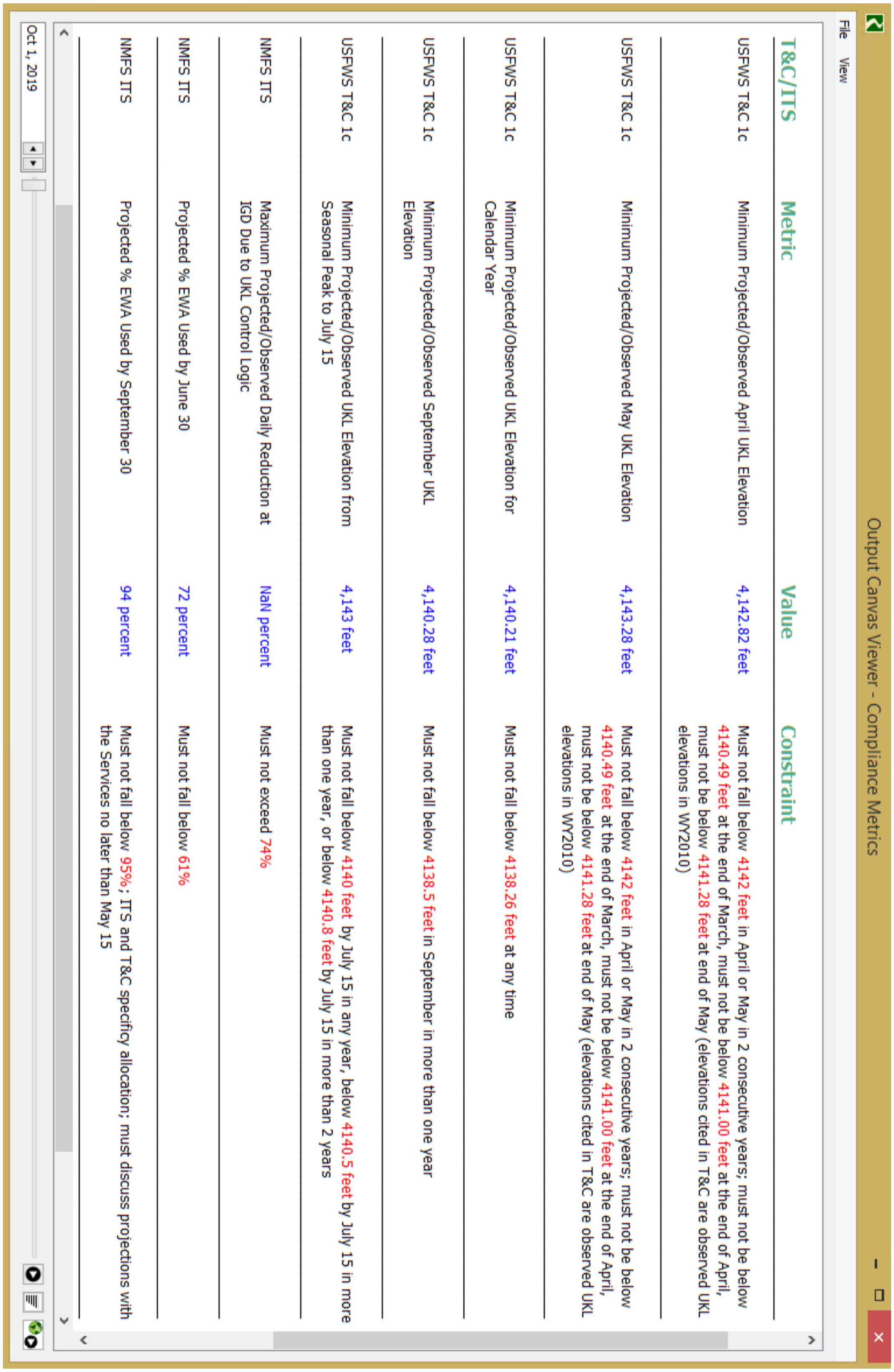


Figure 9. Sample Compliance Metric canvas.

2.2.7. Adjusted Workflow

A step-by-step explanation of the workflow that uses KROM is presented in this section.

1. **Acquire, Review, and Archive the Data:** The operator receives the measurements from the USGS/NWIS, Hydromet, and basin reports. Then, they validate and if needed correct the data. The validated data is linked to a new sheet that formats it for import to the KROM through a DMI. The operator still performs this step in Excel workbooks that serve as the database for the daily data (one for formatting and archiving, another that's linked to the KROM's DMI).
2. **Prepare the Model for Today's Use:** The operator sets up the KROM with a script and SCT. They use the script to load the Rpl sets, open the associated SCT, set the timestep parameters, clear data from previous runs, and import the observed hydrology. Then, they use the SCT to enter any necessary manual inputs (overrides, reported data) and configure the per run settings. The operator performs this step in the KROM.
3. **Operate and Plan for the Season based on Observed and Projected Hydrology**: Once the operator has finished model preparation, they run the KROM and check results. To configure settings properly, the operator uses the multiple run configuration called "Run Trials of Different Settings". After each multiple of runs, they check the results to see if any run's hydrology and deliveries match expectations. If satisfied, the operator makes a single run with the per run settings chosen to compute the operational releases. Then, they report the projections to stakeholders. The operator performs this step with the KROM.
4. **Run Scenarios and Analyze Possibilities**: After operational releases are set, the operator runs scenarios with the multiple run configurations called "Run Historical Hydrology Scenarios" and "Run with and without Agricultural Deliveries". The operator performs this step with the KROM.
5. **Generate Reports and Output Products:** Every week, the operator creates and delivers reports to stakeholders. Stakeholders include irrigation districts, tribes, and other water users. The weekly reports that the operator generates are as follows:

a. *USBR Daily Numbers Update*: Observed hydrology from the past week.

b. *Klamath Project Deliveries and Demands*: Observed deliveries for the current water year.

c. *Cumulative UKL Inflow*: The cumulative, observed UKL inflow for the current water.

d. *Smoothed UKL Inflow*: The smoothed, observed UKL inflow for the current water year.

e. *Wormtrails*: The observed UKL pool elevation for the current water year.

The operator produces report *a*. with the KROM. The operator produces the other reports from Excel workbooks of the same name.

6. **Archive the Model:** Each day, the operator archives the KROM with the per run settings used to compute the official operational releases and projections. Thus, a snapshot of the model, exact data and logic used for assignments is recorded.

2.3. Testing

For the KROM to be viable replacement of the PA Calc, its logic must be able to replicate the PA Calc's results. Our testing plan validates this criterion by comparing a set of the models' results from operational runs starting at various points throughout the water year. This section starts by presenting the set of results we compare. These results are also call comparison variables. The testing methodology is covered next, which goes over model setup, result management, and variable comparison. Results of the comparison are presented last, and we provide commentary on any differences. Each step of the testing plan is discussed in detail in below.

2.3.1. Comparison Variables

A comparison variable is a computed result's timeseries of values, i.e., the Iron Gate Dam release for every timestep in the run range. A result is chosen as a comparison variable for two reasons: 1) they are a hydrologic measure that represents some important flow or delivery or 2) they are an input variable for one of those measures' computation. If either type of variable is off, that difference is likely to propagate to other variables' computed value. And if vice versa, we are confident that the KROM's logic is correct since a difference would most likely be observed in one or more of these variables. Table 5 shows the comparison variables.

Table 5. Comparison variables used to assess the model.

UKL Storage	Iron Gate Dam Central Tendency Release	Keno to IGD Accretion	LRDC Accretion
EWA Used Through Yesterday	Iron Gate Dam Environmental Release	F and FF Pump Accretion	Lake Ewauna Accretion
Iron Gate Dam Outflow	UKL Outflow	UKL Agricultural Delivery	UKL Inflow

2.3.2. Methodology

Our testing plan runs the models for multiple operation start dates, which is representative of the operator performing operations at different days throughout the year. In that way, we ensure that if the model computes releases in the summer or winter, which are based on different logic, they generate the intended results. Essentially, we are trying to have every logical statement execute/compute a value to ensure they were written correctly. We did not vary the per run settings between test runs since the conditions of the tested water year were diverse, i.e., there was flooding early spring and lack of supply mid-summer. That diverse hydrology allows most hydrology-based logic to execute. Whereas, changing the operation start date allows most timestep-based logic to execute. We perform test runs with an operation start date in each month of the year. Each test run follows a simple three step process. First, both models are set up with identical settings. Second, the models are run and the results gathered for comparison. Third and last, variables from each model are compared and we identify any difference in their values throughout the run range. Each step is covered next.

Setup: For the results to match, the data input to each model must be identical. The following tasks are performed to meet this criterion:

- Match the hydrology tables, see Table 4 for a list of them. This is only done for the first run since their values remain the same regardless of the date that we are running operations.
- Match the observed hydrology by setting the same operation start date in each model. Both models pull observed hydrology from the same database based on the operation start date.
- Match the per run settings by setting everything in Table 2 (see "Data" section of "Model Requirements") the same for each model.
- Set the manual inputs/overrides the same for each model. Otherwise, policy will set those values for one model and not the other.

Result Management: After each model is setup and ran, we export the results to databases. Our databases are separate Excel workbooks, one that stores the KROM's results and another that stores the PA Calc's results. Each workbook is linked to its associated model, with a DMI and cell formula respectively. To export results from the KROM, we execute the output DMI named "KROM Outputs". To export results from the PA Calc, we open and save the workbook that stores the PA Calc's results. The databases are formatted identically, i.e. each column represents a variable and each row represent the variable's value on a timestep.

Variable Comparison: We developed an R script that imports, computes the differences, and plots the differences of each variable's values. Prior to running the R script, we input the test run's operation start date so the code can distinguish between the observation- and projection-based values. The goal of testing is to confirm the KROM's result adequately match the PA Calc's. Thus, we define allowable differences for the models' results. The allowable differences for observation- and projection-based values are as follows:

- <u>Based on Observed Data:</u> Any difference must be caused by inconsistencies in processes fundamental to the model, i.e., one model performs a table lookup and the other an interpolation.
- <u>Based on Projected Data:</u> Any difference must be less than 5% of the mean value.

These criteria were priorly approved by KBAO since they must be able to justify to stakeholders that the KROM run operations as intended. The comparison plots the R script creates are automatically formatted into a report. We archive these reports by each test run's operation start date.

2.3.3. Results

The results of the test runs are interpreted through the comparison plots. In this section, we present each variable's comparison plot for one test run. The x axis of each plot is the timestep. The y axis is the unit difference. Additionally, the values computed from observed and projected data are color coded orange and blue respectively. The compared and plotted values are from the 7/10/2020 test run, see the figures below.

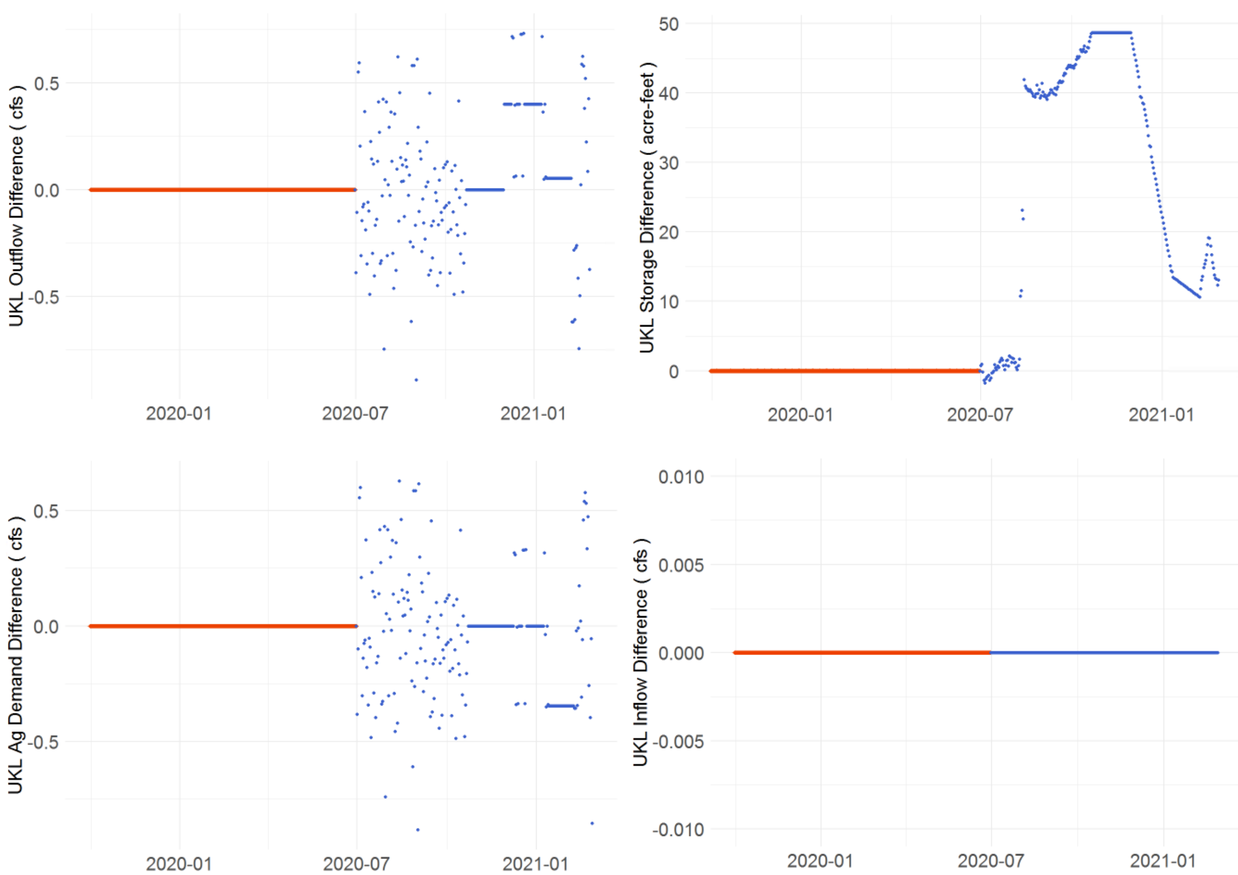


Figure 10. Comparison plots of UKL measures. The outflow (left) and storage (right) are on the top row. The agricultural delivery (left) and inflow (right) are on the bottom row.

For the variables representing UKL hydrology, no differences occur in their observation-based values. The max differences of the projection-based values for outflow (0.08%), storage (0.02%), inflow (0%), and agricultural delivery (0.3%) are all well below 5%. Thus, the differences are allowable. Notice that the agricultural delivery's difference distribution is similar to the outflow's. This occurs since that agricultural delivery is portion of that total release. Thus, we can confidently identify the KROM's agricultural delivery logic as the one source of the difference.

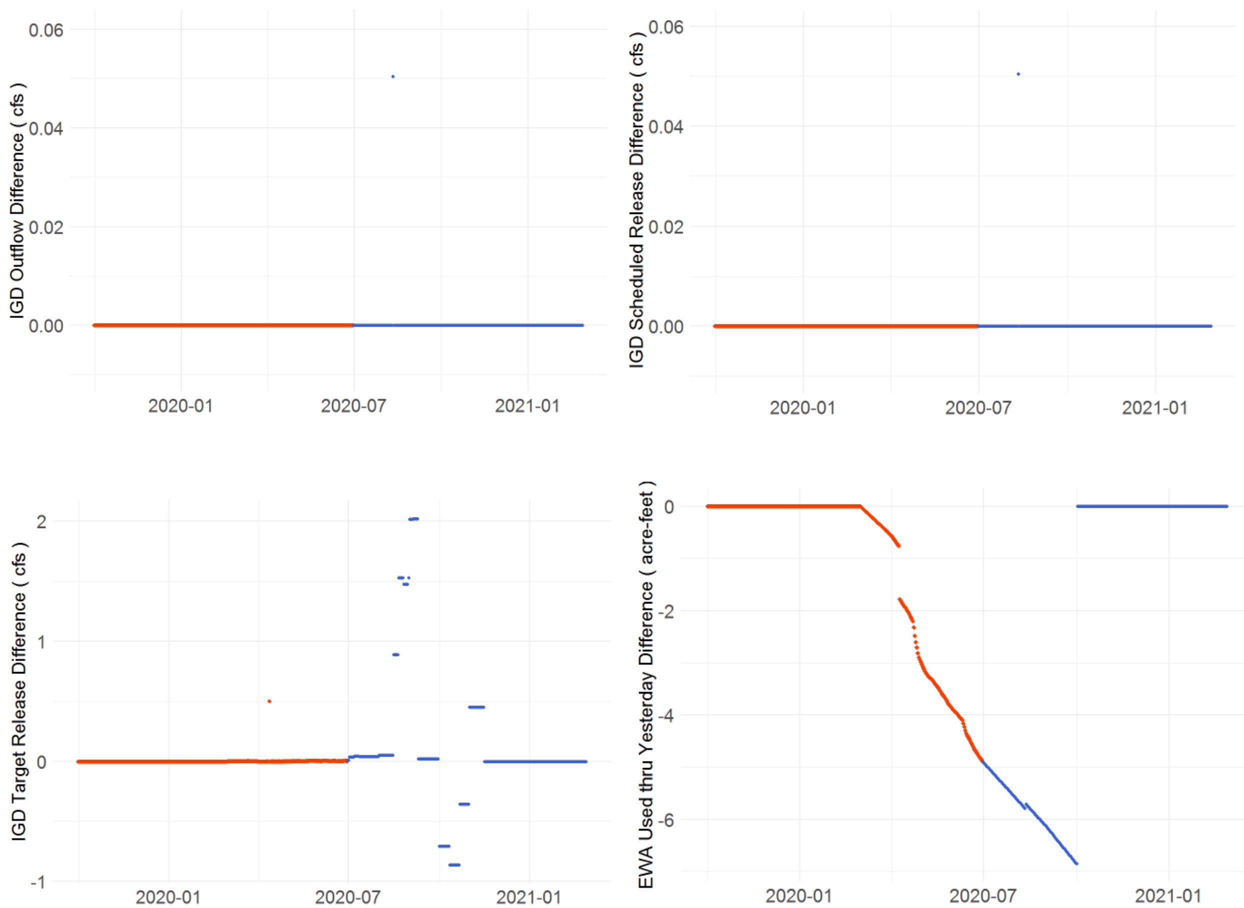


Figure 11. Comparison plot of Iron Gate Dam measures and EWA use. The outflow (left) and central tendency release (right) are on the top row. The environmental release (left) and EWA use (right) are on the bottom row.

The environemntal release's and EWA use's observation-based values exhibit differences. Both differences are caused by an incorrect cell reference error in the PA Calc (on 4/10/2020). Additionally, the EWA use's observation-based values differ since the KROM's flow to volume conversion is more specific (significant digits) than the PA Calc's. Note that the EWA use is a cumulative value, which means the differences compound with time. Since the differences are not logic induced, they are allowable. The max differences of the projection-based values for EWA use (0.02%) and Iron Gate Dam outflow ($4.5*10^{-5}$%), minimum/ramping release ($4.5*10^{-5}$%), and environmental release (0.1%) are all well below 5%. Thus, the differences are allowable. Iron Gate Dam's outflow and ramping/minimum release difference distribution match since the ramping/minimum objective sets the operational release, i.e., had the greatest priority. This is also why differences shown in Iron Gate Dam's environmental release are absent from the outflow, i.e., had lower priority.

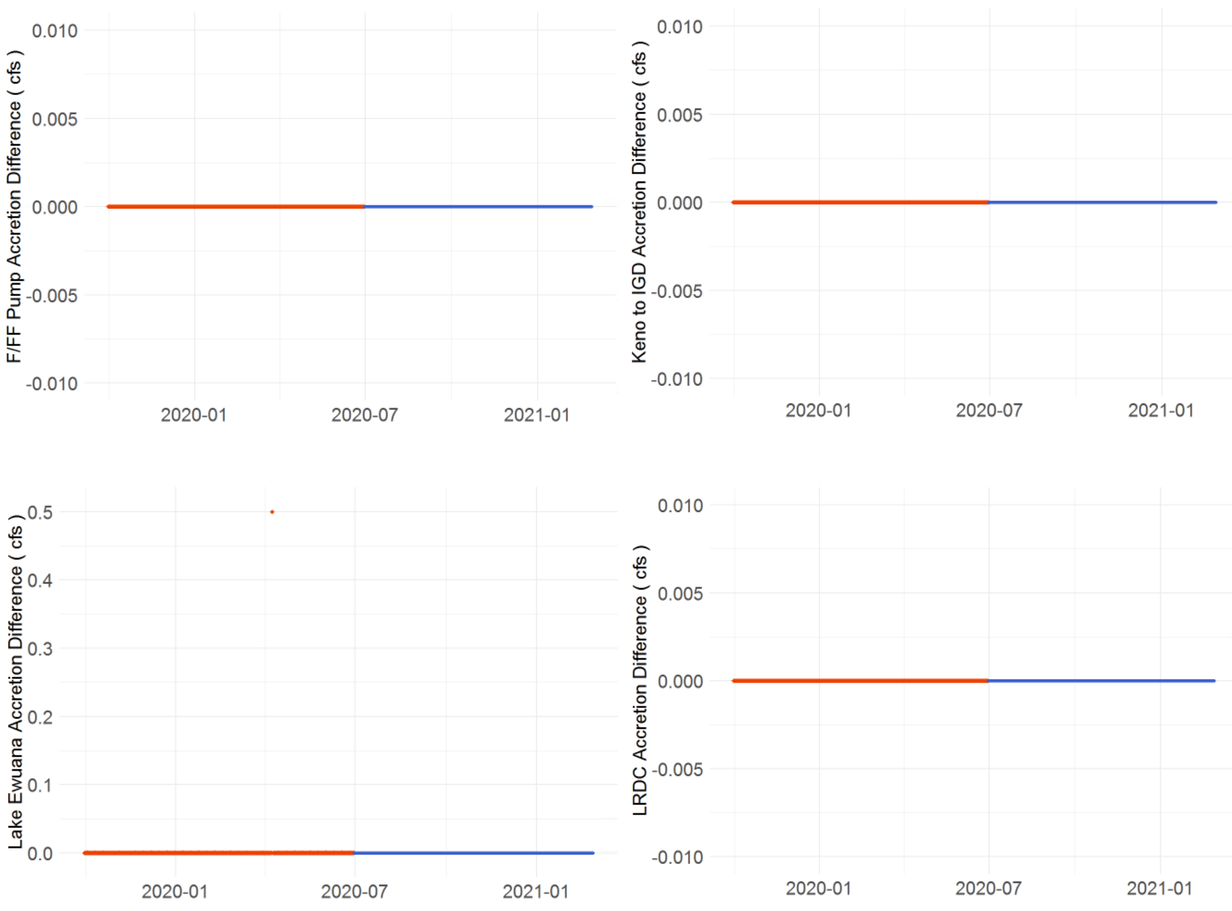


Figure 12. Comparison plots of the accretion measures. The F/FF Pump (left) and Keno to IGD (right) accretion are on the top row. The Lake Ewauna (left) and LRCD (right) accretion are on the bottom row.

There is one observation-based value of Lake Ewauna accretion that differs. It is also caused by the incorrect cell reference error in the PA Calc (on 4/10/2020); the magnitude of difference matches.

Since that difference is not logic induced, it is allowable. No differences occur for the accretion variables' projection-based values.

Based on the comparison plots, the KROM was able to replicate the PA Calc's results within the acceptable range of differences. The observation-based value differences were only caused by model process inconsistencies. The projection-based value differences were well under 5% of each variable's mean. While the difference discussed in this section focus on the 7/01/2020 test run, acceptable standards were also met for the other test runs. Further testing was performed by collaborators KBAO and TSC to validate the KROM's logic performs as intended for even more dates and hydrologic scenarios. These results are pending.

2.4. Stakeholder Workshop

To present the KROM to the basin's stakeholders and receive feedback, we collaborated with KBAO and TSC to hold a workshop. This workshop took place on September 23rd, 2020. Initially, we had planned to hold it in Klamath Falls, Oregon, which is where KBAO is located. But, due to COVID-19, it was held virtually. The participant of the workshop were personnel from both collaborating branches of Reclamation (KBAO and TSC), CADSWES, National Marine Wildlife and Fisheries (NMFS), United States Fisheries and Wildlife Services (USFWS), MBK Engineering, and Confluence Consulting. The personnel from the latter two firms were the modeling consultants who designed the PA Calc. Since not all attendees were familiar with RiverWare, this served as an opportunity to present RiverWare as a software and show other applications where RiverWare is used to manage basin operations. The main topics covered at the workshop were as follows:

- RiverWare Overview
- KROM Tutorial and Demonstration
- RiverWare Capabilities
- Other Basin's RiverWare Operations and Planning Models

- Discussion

The workshop spanned over the entire workday. In this section, we will discuss the feedback from the stakeholders at the end of the day. This will cover impressions of the KROM and using RiverWare as the operations and planning model going forward.

2.4.1. Feedback

Participants' comments on the KROM were quite positive. Most were centered around four key features. 1) Multiple run configurations: this form of run management is regarded as much more powerful and time efficient than the single run process currently used for operations and scenarios. 2) Data transfer: automated input of observed data and output of results is a quality of life improvement over the copy and paste methodology necessary for a few of their data management tasks. 3) Suite of output products: they like that results from multiple runs are presentable on the same plots and that other products are in a ready to distribute form immediately after generation, i.e. html. 4) Script and SCT: the centralized style of data management and quick setup is favored over the per run settings being spread across objective specific worksheets.

Based on the KROM's strong performance and reception at the workshop as well as commitment from KBAO to advance operations' management with RiverWare, there are plans to develop Rpl sets and adapt the KROM for the 2021 operational policy. These tasks would be done alongside the development of the new Excel operations management tool. The reason being, this 1) confirms the KROM can produce the intended results and 2) allows operators to familiarize themselves with managing operations using the new tool. Currently, Reclamation is working with CADSWES to develop a new contract to build an operations/planning model for the 2021 operational policy. Consultants that develop the Excel model are willing to get on board with the switch, but first want to take some time to experiment and get proficient with the RiverWare software. To support this change, CADSWES may host a RiverWare training session for the Klamath River Basin stakeholders.

Chapter 3: Methodology

The Klamath River Basin is no stranger to the water management challenges of the West. Frequent reconsultations, lack of adequate forecasts, and growing water stress persist as constant issues. To address these problems, this study has three objectives:

1. Develop a basin RiverWare model with a robust framework that adjusts to accommodate new policy logic associated with different reconsultations.
2. Develop and test improved forecasting techniques that have been successful in other Western Basins.
3. Determine how much accurate forecasts could improve operational projections.

The previous chapter describes the development and testing of the RiverWare model. This chapter focuses on the steps taken to improve forecasting techniques and operational projections. Our forecasting work starts at identifying basic relationships such as the seasonal inflow volume's connection to climatic and local measurements. Then, it moves to comparing regression methods to identify one that best predicts seasonal inflow volumes based on those relationships. Our operational projections work is an evaluation of the improvement for each of the projected operational objectives. To do this, we run the RiverWare model with our forecast scenarios and analyze the agricultural, environmental, and flood outputs. An elaboration of each step is given in each of the following subsections.

3.1. Predictive Variable Identification

Currently, the NRCS generates forecasts at six different lead times, starting on January 1st and occurring on the first of each month until June 1st. The length of the season depends on the lead time. From January through March lead times, the inflow volume is forecast from March through September. From April through June lead times, the inflow volume is forecast from the month the forecast is made through September. These forecasts are based on short lead, local predictive variables, e.g. first of month snow-water equivalent (SWE), precipitation, and antecedent streamflow [USDA, 2011]. To improve forecasts, we propose to identify climatic antecedent[3] information that has strong connections to inflow during these timeframes, then add that information to the NRCS's local information.

To develop predictive variables from our information, we use three steps:

1) Identify teleconnections that have been useful in other studies in the Western United States.
2) Test the relationship between teleconnections and seasonal inflows in the Upper Klamath Basin and select the strongest signals from the datasets.
3) Consider the teleconnection information and NRCS information as our new set of predictive variables

A more detailed explanation of each action is provided in the sections below.

3.1.1. Potential Climate Teleconnections Identification

Each year, the National Resources and Conservation Service (NRCS) generates the Seasonal Inflow Forecasts for Upper Klamath Lake. These forecasts are based on local predictive variables, e.g. snow-water equivalent (SWE), precipitation, and antecedent streamflow. Noticeably, there is a lack of climate teleconnection variables informing forecasts. In other Western basins, Sea Surface Temperatures (SST) anomalies and 700 hectopascal Geopotential Heights (700 hPa GPH's) have been shown to possess a strong correlation with seasonal inflow volumes [Wang and Tin, 2000; Sagarika, 2015]. Thus, we investigate both climatic data's potential connection, spatially and temporally, with seasonal inflows at Upper Klamath Lake. The definition and scale of the climatic data is as follows:

SST anomalies: The quantity sea temperatures depart from the historical mean at a given location. This study uses 5x5-degree gridded SST anomalies on a monthly temporal scale. "National Center for Atmospheric Research (2020)"

[3] Relative to the forecast's lead time, e.g. the January Regression Model is only fit on info known prior to January.

700 hPa GPH's: The height that the 700 hectopascal pressure surface sits above sea level. This study uses a 2.5x2.5-degree gridded GPH's on a monthly temporal scale. "National Center for Atmospheric Research (2020)"

3.1.2. Strongest Signals Selection

Spatially, each climatic dataset covers the globe. To limit size and improve the quality of the information, only the strongest spatial connections are kept. This prevents overfitting the regression model to the climatic teleconnections' random noise. To develop predictive variables from each climatic dataset, the following step-by-step process is employed:

1. Compile historical Upper Klamath Lake inflows by month. Aggregate into volumes for each seasonal period that is forecasted, e.g. March-September, April-September, May-September, and June-September.
2. Average climatic data over three-month periods. This eliminates the influence of outliers from individual months. The first three-month period is September thru November. The periods step by one month until the February-April average is reached.
3. Compute the spatial correlation for each viable combination of climatic and seasonal inflow datasets. A viable combination is Sep-Oct SST's with Mar-Sep Inflows since the SST's are available before the inflows occur. Whereas, Feb-Apr SST's with Mar-Sep Inflow is an invalid combination since the March and April SST's are unknown before that inflow period begins.
4. Plot each grid point's correlation to determine the regions with the strongest connections.
5. To capture the signal from each region that has strong correlation coefficients, set a physical boundary by visual inspection around the points whose correlation coefficients exceed an absolute value of 0.4. Do this by approximating the max and min latitude and longitude of the box. Then, compute the average of the unit values inside the boundary. The averaged values serve as the predictive variables for the regression models. Each average is denoted by location and period, e.g. Pacific Northwest Sep-Nov SST's.

6. Compile the predictive variables for each forecast date. Since variables vary spatially and temporally, two variables from the same region but different periods are usable for a regression model. To eliminate redundant information, select the variable from the period that shows the strongest correlation.

3.1.3. Incorporation of Existing Forecast Products

To include local information, the NRCS seasonal inflow forecasts are added to the sets. This is done for two reasons: (1) the NRCS forecasts contain highly predictive localized components such as precipitation, SWE, and antecedent flow in their computation, and (2) their inclusion shows if and how much skill the climatic predictors add when our forecasts are compared against them. Using the local variable sets that the NRCS inputs to their model is not a practical option since the exact composition is unknown to non-USDA personnel. To obtain the NRCS seasonal inflow forecasts, we access charts and reports on the USDA's website "United States Department of Agricultural (2020)". The seasonal forecast we need for each lead time are as follows:

Table 6. Seasonal forecast period based on forecast lead time.

Lead Time	Seasonal Forecast
January, February, March	March-September Inflow Volume
April	April-September Inflow Volume
May	May-September Inflow Volume
June	June-September Inflow Volume

The required seasonal inflow forecasts, specifically the periods, are dictated by the operations policy. Throughout the study's analysis period (1982-2019), those exact forecast products are commonly unavailable, except for April and May lead time forecasts. This necessitates the available forecast products to be adjusted accordingly. Before April, the NRCS generates at least two forecast products for each lead time. One is the Month-July forecast and the other is either the March- or April-September forecast (NRCS only produces one of these). To approximate the March-September forecast from the April-September forecast, I estimate the March portion of the Month-July forecast where the "Month" refers to the lead time's month of the seasonal inflow forecast, e.g. if the lead time is January 1, the "Month" is January. Then, that volume is added to the April-September forecast. The March portion is computed as follows.

1. Retrieve observed Upper Klamath Lake inflows from 1981 through 2018.
2. Sum the inflows into the March volume, January-July volume, February-July volume, and March- July volume.
3. For each year, compute the quotient of the March and Month-July inflow volume. Thus, there are three quotients for each year.
4. For each lead time, multiply the NRCS's Month-July forecast by the average quotient from years prior. For example, if the year is 1999 and the lead time is January, use the average of the January quotients from 1981-1998.

For June, half the forecast products are the May-September forecast. Since the May inflows are known by June, the June-September forecast is the difference between the May-September forecast and May inflow volume.

3.2. Regression Model Development and Selection

The regression model development and selection methodology is a two-step process. (1) Develop regression models to forecast seasonal inflows on the first day of January, February, March, April, May, and June; each with its own set of predictors suited to the month. (2) Select the best regression method for each month based on its models' leave one out cross validation root mean square error (RMSE) and Nash Sutcliffe model efficiency coefficient (NSE). Two regression methodologies are employed in this study: local polynomial and random forest.

3.2.1. Local Polynomial Regression

Local polynomial regression is a non-parametric method that invokes the principle of weighted moving averages to smooth/fit the data. Flexible in nature, this method does not require assumptions of normally distributed errors and variables that likely do not hold true in natural systems. As further motivation, local polynomial regression has proven effective in similar hydrologic studies in the western US, e.g. streamflow forecasting on the Colorado River basin [Bracken et al., 2010] and the Truckee and Carson river basins [Grantz et al., 2005]. Each of these prior studies utilized the local weighted polynomial method as laid out by Regonda et al. [2006]. Following this groundwork, this research implements a simplified version to obtain the "best model". The "best model" is classified by the size of the neighborhood, the order of the polynomial, and the subset of predictor variables which are identified by some objective criteria. Allowing the neighborhood and order to be optimized by the "locfit" function developed for R, the only aspect to determine becomes the subset of predictor variables. The subset of predictor variables will be chosen from two sets, (1) climatic and (2) climatic and local. The climatic set lacks the NRCS forecasts, thus eliminating the local information. A purely climatic set is regressed to assess the teleconnection's predictive value relative to the purely local (NRCS Forecasts) and local and climatic information (Combined Set Regression Forecasts).

As a precursor to regression, predictive variables are tested for independence. This is done by developing a correlation matrix of the sets. For regression to work properly, the relationship between each predictive variable and the dependent variable must be isolatable [Frost, 2017; Haque, 2013]. To isolate the relationship, all predictive variables except one are held constant while the unit change in the dependent variable for each 1-unit change in the non-constant predictive variable is observed. That dependent variable's unit change becomes the predictive variable's regression coefficient. When predictive variables aren't independent or are multicollinear, one predictive variable's change correlates to a shift in another. If two predictive variables shift while observing the dependent variable's change, we can't identify the predictive variable that induced the change; thus, both predictive variable's regression coefficients are altered from their true value. Thus, if multicollinearity is strong, a predictive variable's regression coefficient changes depending on other predictive variables in the subset.

Based on the degree of correlation shown in the correlation matrix, we perform Principal Component Analysis (PCA) to make the predictive variables independent. PCA de-correlates predictive variables by finding independent, complex linear combinations of said predictive variables, which are called principal components. To perform PCA, we follow the methodology laid out by Brems [2017], which is as follows:

1. Assemble a matrix of predictive variables (predictors), with each column a predictor and the rows as their time series of data.
2. For each column, subtract the mean from the data so the column's mean is zero.
3. Since predictors do not have the same units, standardize the matrix. If left unstandardized, PCA will emphasize the features of predictors with greater variance. The centered and standardized matrix is called Z.
4. Transpose the matrix Z. Then, multiply the matrix Z by the transposed matrix. The output is the covariance matrix.
5. Perform an eigendecomposition of the covariance matrix to obtain the matrix of eigenvectors and matrix of their corresponding eigenvalues.
6. Sort the eigenvalues from largest to smallest, which is done by rearranging the columns. Arrange the matrix of eigenvectors to match their corresponding eigenvalue column position.
7. Multiply the matrix of sorted eigenvectors by the matrix Z. Now each centered and standardized observation is a linear combination of the raw predictors with the weights set by the eigenvectors. Since the eigenvectors are independent, now the predictors in the product matrix are also independent.

To interpret the contribution of each predictor to individual principal components, we analyze the loadings. The columns are principal components and the row predictors. A high coefficient signals a greater contribution and vice versa.

As a last step before subset selection and regression, the unimportant features of the predictors are eliminated. This is also called feature reduction, which is another application of PCA. Feature reduction is based on a key PCA assumption. It follows that the importance of a principal component corresponds to the fraction of set variance it explains. If the variance explained is large, the more important it is at explaining the dependent variable's behavior and vice versa. Thus, to select the set of principal components, a floor is defined based on the fraction of variance they explain. The floor is at 10% of the set's variance. If a principal component accounts for less than 10%, it is considered noise and left out of the regression set. Whereas, if a principal component represents greater than 10%, it is considered a signal and is valid to be considered in the regression set.

With predictor variables properly prepared for each set (climatic, local and climatic), the simplified local polynomial regression procedure that is based on Regonda et al. [2006] and also used by Bracken et al. [2010] proceeds as follows:

1. Select a subset of predictors
2. Weighted least squares estimation fits a polynomial based on the predictor subset. Use the fit to get an estimate, denoted Ŷ.
3. Compute the objective criteria, generalized cross-validation estimation

$$GCV(K,p) = \frac{\sum_{i=1}^{N} \frac{e_i^2}{N}}{\left(1 - \frac{q}{N}\right)^2},$$

where e_i is the model residual ($Y_i - \hat{Y}$) for the ith data point, N is the number of data points, q is the number of parameters in the local polynomial model. Generalized cross validation (GCV) provides a good estimate of predictive risk of the model, unlike other statistics, which are goodness of fit measures [Craven and Wahba, 1978].

4. Repeat steps 1 and 2 until the GCV values of all the predictor subset combinations are obtained. The goal is to minimize the GCV score, thus the model with the lowest value is ranked the "best model".
5. Retain the top two "best models" as the candidates for the next step in comparing the regression methods since GCV score is not a definitive measure of the best predictive model. This ensures the consideration of a model that may excel at a different set of criteria.

3.2.2. Random Forest Regression

Random forest regression is the other method applied to this data set. It is a non-parametric method that uses an ensemble-learning algorithm to construct decision trees based on the data. The random forest operates by utilizing the decision trees to generate independent forecasts that are assembled into an ensemble. The mean of those forecasts or the ensemble produce the final estimate. Random forest regression is an excellent technique for capturing nonlinear interactions and has been proficient at forecasting reservoir inflows with climatic indices [Kim et al., 2019]. The main parameters to tune for a random forest are (1) the number of trees, (2) the number of variables each node split of the tree considers [Breiman 2001]. For this application, (1) is set at 500 to capture the signal without being too computationally expensive and (2) is set to the default of the number of predictors divided by three to avoid overfitting. Unlike the local weighted polynomial method, no "best model" is found from the random forest method since the fit is on the entire raw predictive variable sets (local and climatic). Thus, no PCA, objective criteria computation, or subset comparison steps are necessary. The entire raw predictive variable sets are suitable for two reasons. (A) Random forest regression does not require linearly independent predictive variables. Thus, PCA is unnecessary. (B) The weight/importance of each predictive variable in a random forest is interpretable, which makes fitting multiple subsets to determine what information is predictively powerful unnecessary. The weight/importance of the predictors are measured by the percent increase mean square error (MSE) and increase of node purity. The definition, formula, and purpose of each importance measure are as follows:

Percent Increase MSE: The squared deviation of the errors when a predictor is excluded from the model. In this study, the errors are the difference between the forecasted and observed inflow volumes. The formula for MSE is

$$MSE_{fo} = \frac{\sum_{i=1}^{N}\left(z_{f_i} - z_{o_i}\right)^2}{N}$$

where N is the number of data points, z_{f_i} is the ith forecasted data point, and z_{o_i} is the ith observed data point. MSE indicates how close the model's forecasts are to the observations. Thus, the greater the percent increase MSE, the more important the variable is since without it the forecasts are much less close to the observations.

Increase of Node Purity: The reduction in sum of squared errors (SSE) when a predictor is selected to split[4]. The formula for SSE is

$$SSE_{fo} = \sum_{i=1}^{N}(z_{f_i} - z_{o_i})^2$$

where N is the number of data points, z_{f_i} is the ith forecasted data point, and z_{o_i} is the ith observed data point. SSE indicates how close the model's forecasts are to the observations. Thus, the greater the SSE reduced, the more important the variable is since when it is selected to split the forecasts are much closer to the observations.

For a comprehensive explanation of the local polynomial or random forest regression methods, refer to Loader [1999] and Brieman [2001] respectively.

[4] Determines which branch of the decision tree to follow at a given node.

3.2.3. Model Comparison and Skill Assessment

From the above steps, there are the following regression models for January thru June: (1) Top local polynomial models, (a) two based on climatic info and (b) two based on local and climatic info. (2) One random forest regression model based on local and climatic information. To compare, the objective criteria are the leave one out cross validation (LOOCV) root mean square error (RMSE) and Nash-Sutcliffe model efficiency coefficient (NSE). The LOOCV test removes 1 data point (p) out of the regression training dataset (n samples). In this study, the left-out point is data from a random year in the analysis period. Then, the model is trained on the altered dataset (n-1 samples) and predicts the removed data point (p). This is repeated until each data point, yearly data in this study, is the (p) removed and subsequently predicted [Shaikh, 2018]. As a result, the set of predictions is entirely based on unknown data[5]. A leave one out cross validation test is a great estimator of model quality in future applications. It analyzes how well a model predicts on new data and avoids overfitting or underfitting the existing data. From the set of predictions, the RMSE and NSE are computed. The definition, formula, and purpose of each objective criterion are as follows:

RMSE: The standard deviation of the errors. The RMSE formula is

$$RMSE_{fo} = \sqrt{\frac{\sum_{i=1}^{N}(z_{f_i} - z_{o_i})^2}{N}}$$

where N is the number of data points, z_{f_i} is the ith forecasted data point, and z_{o_i} is the ith observed data point. RMSE indicates how close the model's forecasts are to the observations. Thus, the lower the RMSE score, the better the predictive skill of the model.

NSE: Goodness-of-fit index, also called the efficiency index, proposed by Nash and Sutcliffe [1970]. The NSE formula is

[5] Data the model has not been trained on, which is what the model will make predictions from in practice.

$$NSE = 1 - \frac{\sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}$$

where N is the number of data points, $\hat{Y}_i$ is the ith forecasted data point, Y_i is the ith observed data point, and $\bar{Y}$ is the mean of the observed data. NSE is widely used by the hydrology community to assess the predictive skill of models. A perfect model achieves NSE score of 1, a model with the same predictive skill as the mean achieves NSE score of 0, and a model with error variance greater than the observations achieves NSE score < 0. Thus, the closer the NSE score is to 1, the better the predictive skill of the model.

The benchmark objective criteria scores are set by the NRCS's model, in each month. Thus, the RMSE and NSE are computed from the NRCS forecast sets. If either regression model beats the NRCS's score, the new predictive variables contain useful information and/or the regression method better represents the system. To visualize this comparison, two plots of the objective criteria scores are made. The Y axis represents the score and the X axis represents the lead time by month. To reduce redundancy, only the best of each local polynomial model (climatic, climatic and local) is plotted. In months the benchmark is beat, the forecast from the best scoring regression model is retained for operation testing. See the results in the "Regression Model Rankings" subsection of chapter 4 for the forecast lead times our models beat the benchmark.

3.3. Operations with Improved Forecasts Analysis

If any of the regression models shows better predictive skill than the NRCS Seasonal Inflow Forecast, the next step is to analyze the accuracy of projected operation's when based on the improved forecasts. Then, those projections are compared against projected operations based on the NRCS forecasts. To do this, we propose a three-step process. (1) Define the performance metrics. (2) For each forecast set, set up the operations model, run, and record results. (3) Compute the values of the performance metrics and plot for comparison.

3.3.1. Performance Metric Definitions

Performance metrics are measures that represent the quality of a product. The product in this research is the seasonal inflow forecasts, which are evaluated for their ability to accurately project operations. The more accurate the projections are, the more useful they are for stakeholder planning. Thus, our performance metrics measure the error from the perfect projection. Error is computed for values that represent Reclamation's three management objectives, which are (1) deliver water for agricultural demands, (2) maintain healthy hydrological conditions for Endangered Species Act (ESA) listed fish through making releases that sustain river stage and keeping adequate storage in Upper Klamath Lake, and (3) control flooding at Upper Klamath Lake. Values for each objective are as follows (number corresponds to objective):

(1) Project Supply Use
(2) Environmental Water Account (EWA) Use and UKL Storage
(3) Cumulative UKL Flood Release

Each value is projected on a daily timestep. All except UKL Storage are a cumulative distribution volume. To summarize their status, one to three assessment dates are selected to analyze the values. Agricultural objectives are the most detailed due to the crop's varying water requirements; it's three assessment periods are March-May, June-July, and August-September. Environmental objectives have two assessment dates, July 1st and August 1st, that correspond to compliance dates defined by the operation policy. Lastly, there is one assessment date, October 1st, for the flood control objective since the timing of flooding is not of interest, but rather the total amount. The difference or error from the perfect value is computed on these dates. Those are the performance metrics, which are 8 in total. See Table 7 for a breakdown of each performance metric.

Table 7. Performance metrics specifications.

Objective	*Agricultural*	*Environmental*	*Flood Control*
Measure	Project Use	EWA Use UKL Storage	Cumulative UKL Flood Release
Assessment Date	June 1st August 1st October 1st	July 1st October 1st	October 1st

3.3.2. Model Setup, Runs, and Results

Every day, the Reclamation Klamath Basin operator will use the Klamath RiverWare Operations Model (KROM) to compute official releases and project the basin's hydrology and demands. One of the main inputs that controls those projections is the seasonal inflow forecast. By configuring the model to the expected supply, which is informed by the forecast, the operator generates projections and informs stakeholders of the potential basin outlook (available supply and how it is distributed).

The seasonal inflow forecast must be disaggregated to daily inflows at UKL. This involves a search and selection from a table of historical daily inflow time series that are disaggregates of historical yearly inflow volumes. They are organized by exceedance percentages, e.g., the yearly volume that exceeds 50% of the historical UKL inflow volumes. The time series is selected for which the total most closely matches the forecasted seasonal volume, and is used for inflows from March thru September. For generating inflows in other months, the operator uses their best judgement of the year's conditions to select the time series of inflows from the table.

The same principles and process the operator follows is performed in our study to generate projections as if we are operating based on our forecasts. Since the KROM is an operations model and not a planning model, each run generates projections for one water year. Thus, each set (NRCS, Best Regression, Perfect) of model forecasts requires 38 individual runs (WY 1982 through 2019). The operator's process that we tailored to setup the model, run, and record results for our study is as follows:

1. Select a water year in the analysis period (WY 1982-2019)

2. For that year, select a forecast on which to base operation projections (NRCS, Best Regression, Perfect)
3. Configure the KROM to model the selected year and forecast
   a. Execute the Historical Hydrology Data Management Interface (DMI) to import that year's hydrology
   b. Set the Operation Start Date to the forecast's lead time
   c. Input the forecast
   d. Adjust the projection table lookup settings to correspond with the expected supply[6]
4. Run the KROM
5. Execute the Record Results DMI to export and record the run's results
6. Repeat for each combination of year and forecast

It should be noted that there are many other adjustable settings or overrides that can affect the outcome of a run. In actual operations, the operator uses these to fine tune projections for periods as small as weeks or days rather than the entire season. In our study, these adjustments are kept constant[7] since such detail is unnecessary. Additionally, anything that adds variability other than step 3's actions should be negated to best capture the forecast's influence.

3.3.3. Performance Metric Computation and Analysis

For each lead time, there are three sets of results. Each set belongs to runs based on the following types of forecasts: (1) Best Regression, (2) NRCS, and (3) Perfect Foresight. For naming, each set will be named after the forecast it is based, e.g. NRCS set. The result sets are composed of eight variables that are laid out in the Performance Metric Definition subsection, see Table 7. The performance metrics represent the error of each variable from perfect foresight. To compute the error or performance metrics for the

---

[6] When adjusted accordingly, (1) the cumulative daily demands equal the project supply and (2) the cumulative UKL inflows match the seasonal inflow forecast, and (3) the accretions lookup settings match the inflow settings.
[7] Overrides are set to "No Value" and adjustment factors are set to one to negate their influence

NRCS and Best Regression sets, we find the difference from the Perfect Foresight set for each variable. To capture the overall variance and average of the errors, performance metrics are shown in boxplots. For each performance metric, there are two boxplots. One represents the NRCS based results and the other the Best Regression based results. From these boxplots, the value of the improved seasonal inflow forecast can be assessed.

Chapter 4: Results

This chapter describes the results beginning with identification of the best predictive variables for each lead time, which are based on either climatic or local information. The climatic variables are formed from strong SST anomaly and 700 hPa GPH teleconnections to seasonal inflow volumes. The local variables are NRCS forecasts, which are adjusted to the seasonal lengths defined by the operations policy. Regression model are fit to these variables or variations of them. The importance of the variables in each model are determined to understand the drivers of our forecasts. Then, the climate-plus-NRCS forecasts are compared with the NRCS forecasts to evaluate when the climatic information can improve forecasting skill. The extent that improved forecasting skill is useful is based on the performance of operational projections. Thus, the last results this section describes are the eight performance metric scores from runs based on the climatic-plus-NRCS and NRCS-only forecasts.

4.1. Best Predictive Variables

To identify the information on which to fit our regression model, the strength of climatic teleconnections to seasonal inflow volume was evaluated. The climatic datasets we worked with were SST anomalies and 700 hPa GPH's. From our investigation, significant regional teleconnections were found across multiple periods. By averaging the values of those regions, we developed our climatic variables. To add local variables to our sets, we utilized the NRCS forecasts. Most of these forecasts weren't for the timeframe the operational policy required. Thus, we estimated from the available NRCS forecasts and historical data the forecasts for the policy specified timeframes. This produced the adjusted NRCS forecasts that serve as our local variables.

4.1.1. Sea Surface Temperatures

SST anomalies from October through December were identified with significant teleconnection strength with the seasonal inflows at all lead times. These SST anomalies are located in the East Atlantic, Pacific Northwest, Central Pacific, and Indian Ocean regions. The connections are ranked by the correlation coefficients. Each region that informs predictive variables possesses a moderate to moderately high correlation coefficient that ranges from 0.4 to 0.6. The spatial correlation plot of the October through December averaged SST anomalies are shown in Figure 13.

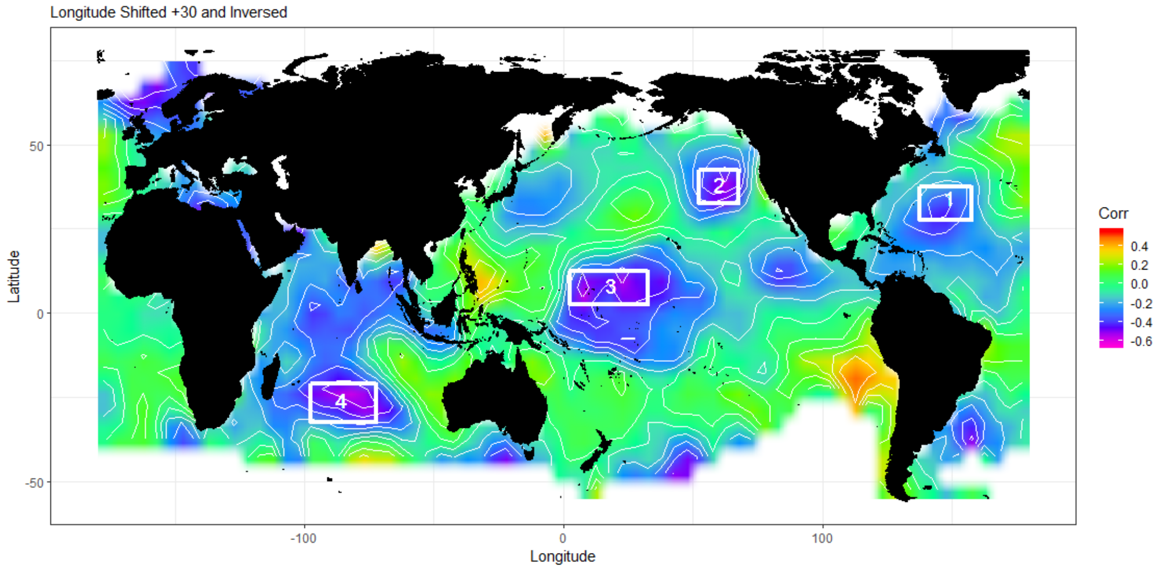
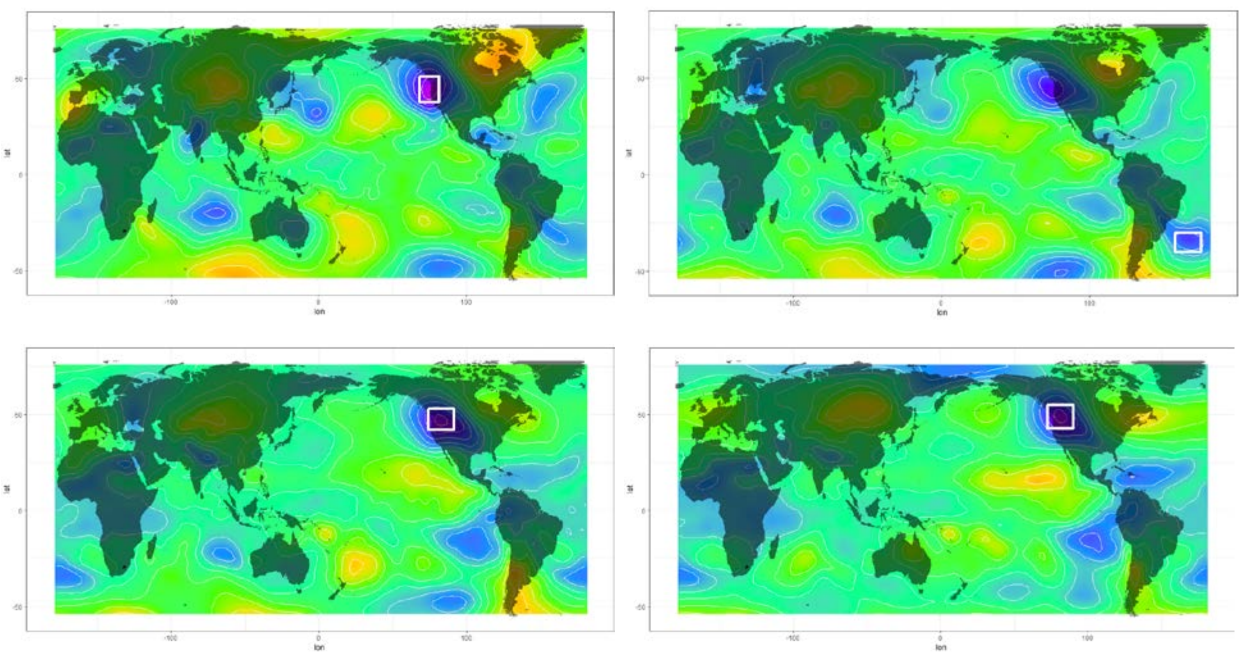


Figure 13. Spatial correlation plot of SST’s with Seasonal UKL Inflow. The area that corresponds to each delineated ocean region is as follows: 1) East Atlantic, 2) Pacific Northwest, 3) Central Pacific, and 4) South Indian.

To form the predictive variables, we delineate boundaries around the specified regions and average the values within. The resulting boundaries are also shown in Figure 13.

4.1.2. 700 Hectopascal Geopotential Heights

Unlike the SST variables, information from multiple periods is found to be significant. The earliest information comes from September thru October and the latest in February thru April. Spatially, the set's information comes from four regions: Pacific Northwest, Southwest Atlantic, Gulf of Mexico and Northwest Atlantic. These produce significant teleconnections during different periods since those regional connections remain strong throughout winter and into spring. Each connection that informs predictive variables possesses a moderate to moderately high correlation coefficient that ranges from 0.4 to 0.6. The spatial correlation plots of the three-month averaged 700 hPa GPH's are shown in Figure 14, and are organized by period.



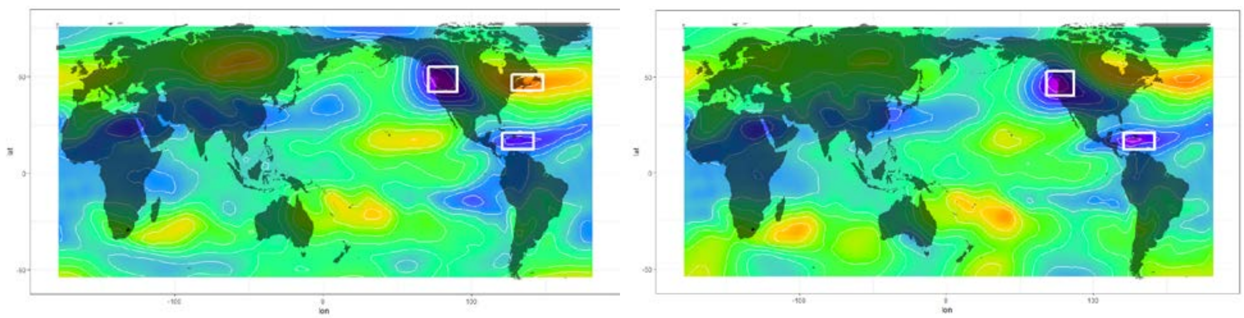


Figure 14. Spatial Plots of 700 hPa Geopotential Heights with Seasonal Upper Klamath Lake Inflows. Each boxed region represents the area averaged value that is a predictive variable. Plots are organized by period. The first row is left) Sep-Oct and right) Oct-Nov average heights, second row is left) Nov-Jan and right) Dec-Feb average heights, and third row is left) Jan-Mar and right) Feb-Apr average heights. Longitude is shifted 30 degrees and inversed.

To form the predictive variables, we delineate boundaries around the specified regions and average the value within. The boundaries are also shown in Figure 14.

For each region, Table 8 shows the periods for which teleconnections are significant for various forecast lead times, providing for the identification of a predictive variable.

Table 8. Significant teleconnections at each Forecast Lead Time.

Region	*Jan 1*	*Feb 1*	*Mar 1*	*Apr 1*	*May 1*	*Jun 1*
Pacific Northwest	Sep-Nov	Nov-Jan	Dec-Feb	Jan-Feb	Feb-Apr	Feb-Apr
Gulf of Mexico	NA	NA	NA	Jan-Mar	Feb-Mar	Feb-Mar
Southwest Atlantic	Oct-Nov	Oct-Nov	Oct-Nov	Oct-Nov	Oct-Nov	Oct-Nov
Northwest Atlantic	NA	NA	NA	Jan-Mar	Jan-Mar	Jan-Mar

Month-Month, i.e., Sep-Nov, specifies the period the regional teleconnections are significant for that forecast lead time. If NA, a significant teleconnection was not found for that region for that forecast lead time. When the regional correlation coefficient is at least 0.4, the most recent period averaged values are used at a that lead time (e.g., April 1st forecast uses the January thru March avg. GPH's).

4.1.3. Adjusted NRCS Forecasts

The NRCS produces seasonal inflow forecasts at six lead times: the first of every month from January through June. Commonly, the seasonal timeframe of the forecast product changes throughout the period of record. Since the operational policy requires a temporally specific forecast, see Table 6, we estimate that forecast from the available NRCS products and historical hydrology. To demonstrate the estimation process and results, we select a year and walk through the steps. For this example, the year 1986 is chosen. First, we retrieve the NRCS Seasonal Inflow Forecasts from the USDA's website, organize them in a table, and denote the forecast season. See Table 9 for color coded breakdown of the forecasts.

Table 9. Original NRCS Seasonal Inflow Forecasts from 1986.

	Upper Klamath Lake Inflow Forecast					
Year	*1-Jan*	*1-Feb*	*1-Mar*	*1-Apr*	*1-May*	*1-Jun*
1986	667	605	700	502	314	184

Key				
	(blue)	Mar-Sep	(green)	Apr-Sep
	(orange)	May-Sep	(purple)	Jun-Sep

Next, we compare the forecasts in Table 9 with Table 6 in the Methodology chapter to determine if the original NRCS forecasts need to be adjusted. The January and February forecasts need to be converted from an April-September to a March-September volume. Also, the June forecast must be converted from a May-September to June-September volume. Since the adjustments differ, they are discussed separately. Starting with the earlier lead times, we estimate the March inflow volume to add to the April-September forecast. The March volume is approximated from two products, the NRCS Inflow Forecast for February-September and the historical fraction of inflow volume that comes in March over the February-September period. See Table 10 for both.

Table 10. The February-September NRCS Inflow Forecasts for January and February lead times and the historical fraction of February-September inflow volume that comes in March.

	February-September Upper Klamath Lake Inflow Forecast		**March Fraction of Feb-Sep Inflow**
Year	*1-Jan*	*1-Feb*	*1981-1985*
1986	815	739	0.221

Now, by taking the product of the Forecasts and Fraction, March volumes are approximated. The March volumes, which are summed with the April-September Inflow Forecasts, and the resultant March-September Forecasts are shown in Table 11.

Table 11. The approximated March Inflow Volume from the NRCS Forecast and the estimated March-September Upper Klamath Lake Inflow Forecast.

	March Inflow Volume		**March - September Upper Klamath Lake Inflow Forecast**	
Year	*1-Jan*	*1-Feb*	*1-Jan*	*1-Feb*
1986	180	163	667	605

For the June lead time, we find the May inflow volume to subtract from the May-September Inflow Forecast. Since the May inflows are known on June 1st, it is a simple subtraction. See Table 12 for May volume and resultant June-September Forecast.

Table 12. The 1986 May Inflow Volume and the computed June-September Upper Klamath Lake Inflow Forecast.

	May Inflow Volume	**June - September Upper Klamath Lake Inflow Forecast**
Year	*1-Jun*	*1-Jun*
1986	136	184

The rest of the adjusted NRCS Seasonal Inflow Forecasts can be found in the appendix. The tables for the other intermediary values such as the March or May Inflow Volume for the relevant lead times are located in appendix as well. These adjusted NRCS forecasts are our local variables.

4.2. Regression Model Skill

For each forecast lead time, regression models (local polynomial and random forest) were fit to the predictive variable set or subset variations (climatic, local & climatic). First, we present the "best model" from each regression method and the most important information in the "best model". For local polynomial models, this is shown with correlation matrices (predictors/predictors, predictors/principal components) and GCV scores. For the random forest model, this is shown with the MSE and node purity importance scores. Second, we compare the predictive skill of the best regression models and NRCS models at each forecast lead time. Predictive skill is measured by the objective criteria scores RMSE and NSE from a LOOCV test. Two plots, one for each criterion, show each model's scores.

4.2.1. Local Polynomial Models

Initially, the information for regression is represented as predictive variables, which were formed from SST's anomalies, 700 hPa GPH's, and adjusted NRCS forecasts in step 1 of the Methodology chapter. Local polynomial regression requires these variables to be independent from each other. Otherwise, the fit will not capture the true relationships between variables and seasonal inflow volume. We tested the independence of our predictors by computing their joint correlation coefficients. The coefficients are displayed in correlation matrices, see in Table 13 for January's. the matrices for the other months are in the Appendix.

Table 13. Correlation matrix of the climate (SST's and GPH's) and NRCS predictors for the January lead time.

	East Atlantic SST	*Pacific NW SST*	*Central Pacific SST*	*South Indian SST*	*Pacific NW GPH*	*SW Atlantic GPH*	*Jan 1st NRCS Forecast*
East Atlantic SST	1.00	0.27	0.20	0.58	0.24	0.29	0.80
Pacific NW SST	0.27	1.00	0.52	0.18	0.48	0.07	0.27
Central Pacific SST	0.20	0.52	1.00	0.46	0.31	0.17	0.21
South Indian SST	0.58	0.18	0.46	1.00	0.26	0.22	0.62
Pacific NW GPH	0.24	0.48	0.31	0.26	1.00	0.29	0.19
SW Atlantic GPH	0.29	0.07	0.17	0.22	0.29	1.00	0.15
Jan 1st NRCS Forecast	0.80	0.27	0.21	0.62	0.19	0.15	1.00

Our predictive variables were unsuitable to fit our local polynomial regression models because there are multiple instances of moderate to moderately strong relationship between variables. For January, the strongest relationship appears between East Atlantic and South Indian SST anomalies, Pacific Northwest and Central Pacific SST anomalies, South Indian SST anomalies and Jan 1st NRCS Adjusted Forecast, and East Atlantic SST anomalies and Jan 1st NRCS Adjusted Forecast. Of course, the diagonal correlation coefficients are equivalent to 1 since each predictive variable is correlated with itself.

Since the predictors are not independent, we need to manipulate them to so they aren't related. Thus, we perform PCA. PCA forms complex linear combinations of the variables, called principal components, that are independent from one another. Some principal components carry greater importance than others. In PCA, the component's importance is tied to the fraction of set variance they explain. The fraction of variance is computed from the principal component's eigenvalues as the quotient of the component's eigenvalue over the set's cumulative eigenvalue. Then, those quotients are plotted on what is called a scree plot. Once again, with January as the example lead time, see Figure 15 for the scree plot of those principal components.

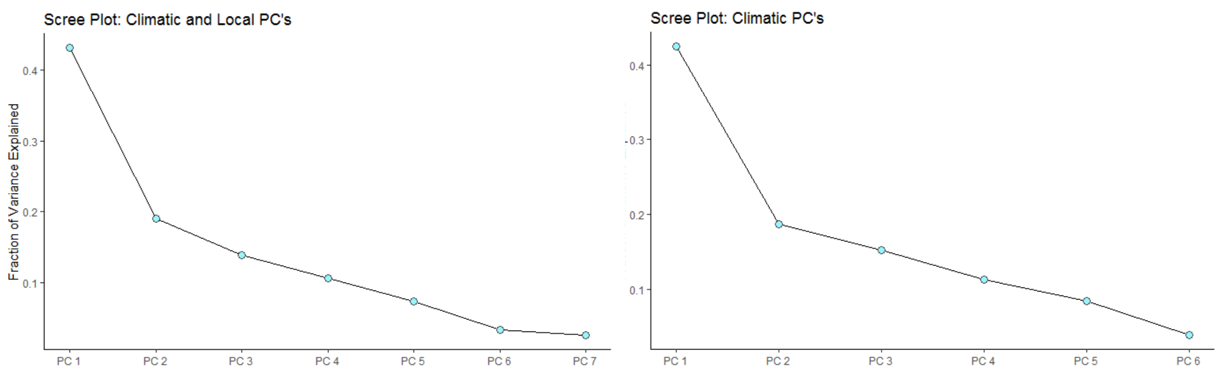


Figure 15. Scree Plots for the principal components formed from the (Left) climate-plus-NRCS and (Right) climate predictor sets for the January lead time.

The other months' scree plots are in the Appendix.

For January, the first principal component carries much more importance than the rest as it explains over 40% of the set's variance. There is a sharp drop-off from PC 1 to PC 2, which explains around 20% of the set's variance. From there, the variance explained by the subsequent principal components drops by about 4-5% per component.

The contribution of each predictive variable to each principal component can be analyzed through the loadings. Loadings are the weight coefficients of the linear combination of the predictors from which the principal components are constructed. The sign of the coefficient denotes the direction of relationship. Thus, if negative, as the predictor increases the component decreases. Coefficients range from a magnitude of 0 to 1. Thus, the closer the coefficient is to 1, the stronger the relationship between the predictor and component and vice versa. Table 14 shows the loading matrices of January's predictor/principal components for both climate-plus-NRCS and climate sets.

Table 14. Loadings matrix of the predictive variables (Top: climate-plus-NRCS, Bottom: climate-only) and principal components used to fit the local polynomial regression model at the January lead time.

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*
East Atlantic SST	0.37	0.48	-0.19	0.27	-0.59	-0.11	0.40
Pacific NW SST	0.38	-0.54	-0.12	-0.03	-0.53	0.24	-0.46
Central Pacific SST	0.37	-0.26	-0.46	-0.55	0.25	0.00	0.48
South Indian SST	0.40	0.40	-0.41	0.10	0.39	-0.15	-0.57
Pacific NW GPH	0.41	-0.29	0.44	0.16	0.11	-0.71	0.06
SW Atlantic GPH	0.26	0.40	0.54	-0.65	-0.12	0.14	-0.15
Jan 1st NRCS Forecast	-0.43	0.05	-0.30	-0.41	-0.35	-0.62	-0.21

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*
East Atlantic SST	0.42	-0.47	0.19	0.50	0.37	-0.44
Pacific NW SST	0.42	0.56	0.02	0.22	0.49	0.47
Central Pacific SST	0.44	0.29	0.27	-0.64	0.03	-0.48
South Indian SST	0.45	-0.38	0.42	-0.10	-0.44	0.53
Pacific NW GPH	0.41	0.28	-0.46	0.35	-0.60	-0.21
SW Atlantic GPH	0.29	-0.40	-0.71	-0.40	0.26	0.16

These are shown in separate tables since component values differ when the variable sets that they are made from differ. See the Appendix for the other months' predictor/principal component matrices.

Based on our loading matrices developed to evaluate the contribution of predictive variables to principal components, we see that the predictors had a similar contribution to the first component with exception of the geopotential heights in the Southwest Atlantic. The other components are not as evenly influenced. Rather, some have no influence such as the Central Pacific SST anomalies to the climate-plus-NRCS's PC 6 (0.00 coefficient) or a clear-cut highest contribution such as the Southwest Atlantic GPH's to the climate-only's PC 3 (-0.71 coefficient). Thus, if either of those components is highly skillful at forecasting seasonal volumetric inflow, we can deduce that there was no signal or a very strong signal from the mentioned regions, respectively.

From each set of principal components, we consider those that explain less than 10% of the set's variance to be noise. To prevent overfitting the model on poor relationships between the predictive variables and seasonal inflow volume, also known as noise or random error, we perform feature reduction. From Figure 15, which is the scree plot for the seasonal inflow forecast at the January lead time, we identified the components to remove. The remaining components from each set that the local polynomial models are fit on at the January lead time are as follows (see the Appendix for the lead time's fit components):

Climate-plus-NRCS: PC 1, PC 2, PC 3, and PC 4

Climate-only: PC 1, PC 2, PC 3, and PC 4

For each set, these four components explain over 80% of the variance, thus, giving us confidence that the strongest relationships are included.

After the local polynomial models are fit on the subsets of the remaining principal components, the GCV scores are computed. The GCV score is the objective criterion used to rank the models; it differs by subsets and determines the "best model." The lower the GCV score the better the model's skill. For each lead time and component set, the variables that fit the two "best models" and the GCV scores of those models are shown in Table 15. We retain two since a model may perform well for one criterion but not another.

Table 15. Best model specifications for the local polynomial regression on the climate-only and climate-plus-NRCS principal components.

Month	*Model Information*	*Independent Variables*	*GCV Scores*
January	Climate-only	a) PC 1, PC 2, PC 3 b) PC 1, PC 2	19, 678 26, 329
	Climate+NRCS	a) PC 1, PC 3 b) PC 1, PC 3, PC 4	21,561 21, 997
February	Climate-only	a) PC 1, PC 2, PC 3 b) PC 1, PC 3	15,822 22,590
	Climate+NRCS	a) PC 1, PC 2 b) PC 1, PC 2, PC 4	17,203 17,301
March	Climate-only	a) PC 1, PC 2 b) PC 1, PC 2, PC 3	18,732 18,895
	Climate+NRCS	a) PC 1, PC 2 b) PC 1, PC 2, PC 3	8,599 10,226
April	Climate-only	b) PC 1, PC 2, PC 3 a) PC 1	13,780 15,443
	Climate+NRCS	a) PC 1, PC 2, PC 3 b) PC 1, PC 2	8,069 9,861
May	Climate-only	a) PC 1, PC 2, PC 3 b) PC 1	5,842 6,116
	Climate+NRCS	a) PC 1, PC 2, PC 3 b) PC 1, PC 2	3,733 3,982
June	Climate-only	a) PC 1 b) PC 1, PC 2	1,813 2020
	Climate+NRCS	a) PC 1, PC 2, PC 3 b) PC 1	1,451 1,661

For the seasonal inflow volume forecasts, the predictive skill as measured by the GCV score improves as the lead time gets later, indicating that the NRCS and climate information is more informative the closer it is to the timeframe that inflow volume is forecasted over. Also note that the seasonal inflow volume forecasted decreases after March since it is over a shorter period. This likely contributes to improving GCV scores at the later lead times since the errors are smaller. At January and February lead times, the best climate-only models outperform the climate-plus-NRCS models. The climate-only models are comparatively much less skillful at March and April lead times when the climate-plus-NRCS models score better by a margin of 4,500 to 9,000. Then, at May and June lead times, the climate-only models are slightly less skillful than the climate-plus-NRCS models. Since the climate information appears most skillful at early and late lead times, we expect the climate-plus-NRCS model to outperform the NRCS model at those lead times for the next round of model comparison, which is based on the LOOCV test results.

4.2.2. Random Forest Models

This regression can use the predictive variable formed from SST anomalies, 700 hPa GPH's, and adjusted NRCS forecast, see the Random Forest Regression section of the Methodology chapter. Random forest regression does not require independence amongst variables. Also, the importance of each variable in the random forest model is determinable. Thus, we exclude no variables from the sets that the model fits, i.e., the model fits all the climate-plus-NRCS variables for the lead time for which it is forecasting seasonal inflow volume. The importance of each variable is measured in two ways. First, we compute the percent increase in mean square error when a variable is excluded from the model. Second, we compute the increase in node purity, which is how much the sum of squared errors decreases when a variable is selected to split. For each, higher values signal greater importance.

We present the measures of importance (MSE and Node Purity) for three of six of seasonal inflow forecast lead times here. The rest can be found in the Appendix. Figure 16 shows each variable's importance in the random forest model that forecasts the seasonal inflow volume on January 1st.

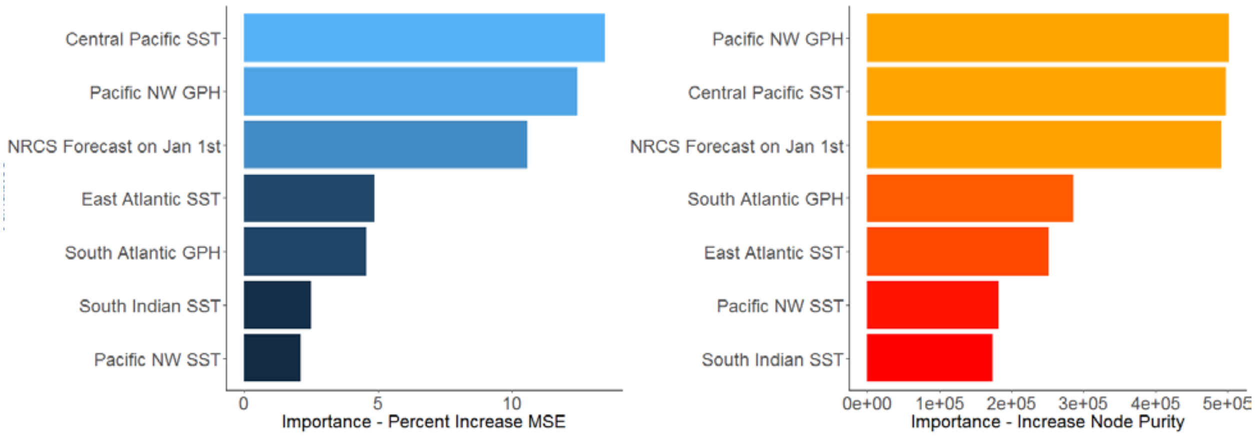


Figure 16. Variable importance scores, percent increase MSE (left) and increased node purity (right), for the seasonal inflow forecast on January 1st.

At this lead time, the climatic information has slightly greater predictive power than the NRCS information. Specifically, the SST anomalies in the Central Pacific and GPH's in the Pacific Northwest are the climatic drivers of the forecasts. This is indicated by a 12-13% MSE uptick when they are excluded from the model and an over $5*10^5$ decrease in the sum of squared errors when they split a node. The other climatic variables are not nearly as important. At most, the MSE increases by 5% per their exclusion and the sum of squares decrease by less than $3*10^5$.

Next, we present the variable importance scores for the random forest model that forecasts seasonal inflow volume on March 1st, see Figure 17.

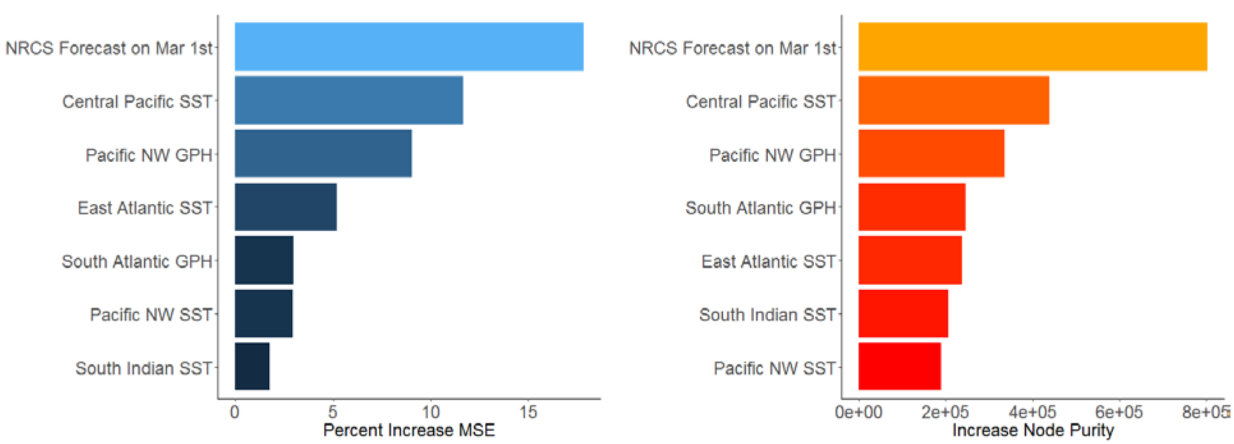


Figure 17. Variable importance scores, percent increase MSE (left) and increased node purity (right), for the seasonal inflow forecast on March 1st.

At this lead time, the NRCS information has overtaken the SST anomalies in the Central Pacific and GPH's in the Pacific Northwest as the drivers of the forecast. Both climatic variables are still powerful predictors as they induce a 10-12% MSE uptick when omitted from the model, but the local variable now causes a 17-18% MSE increase. Also, as compared to January, the sum of squares decreases $8*10^5$ vs $5*10^5$ when the local variable splits the node. The importance of the other climatic variable has gone down marginally as well.

Lastly, we present the variable importance scores for the random forest model that forecasts seasonal inflow volume on June 1st, see Figure 18.

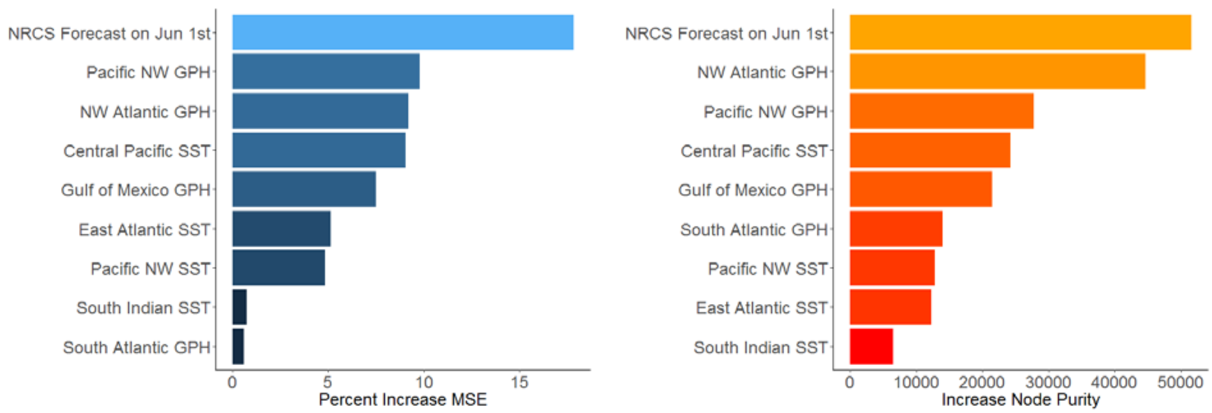


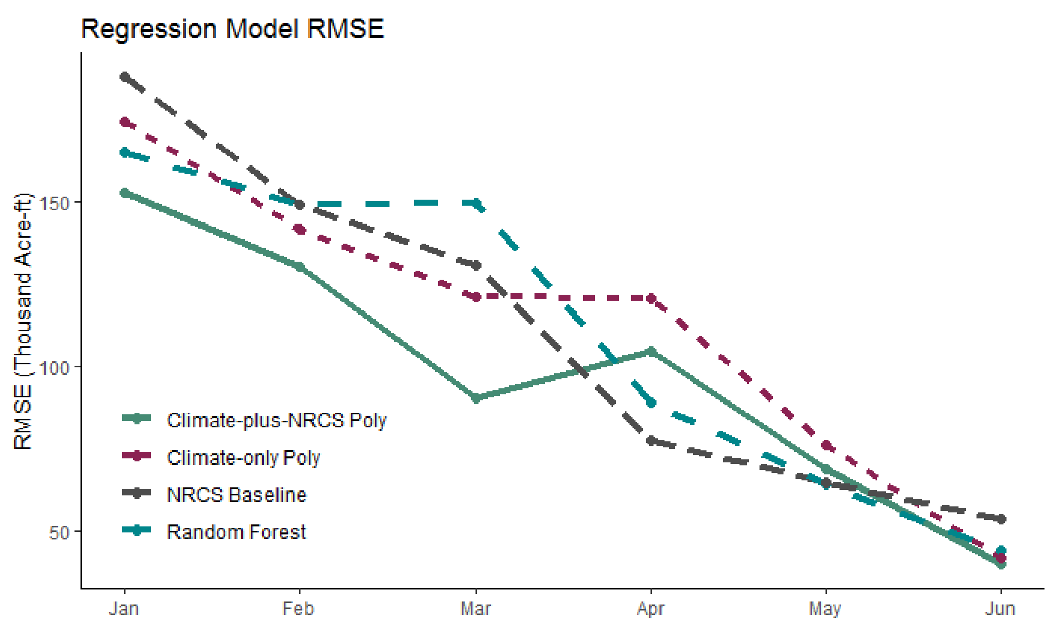
Figure 18. Variable importance scores, percent increase MSE (left) and increased node purity (right), for the seasonal inflow forecast on June 1st.

The NRCS information still possesses the greatest predictive power, which shows with a 17-18% MSE increase when excluded. Unavailable for the models for earlier lead times, GPH's in the Northwest Atlantic ranks in the top three most important variables for both percent MSE increase and increase in node purity. GPH's in the Pacific Northwest and SST anomalies in the Central Pacific remain among the top important climatic predictors with a 7-9% MSE increase when excluded, which is slightly down from March. Noticeably, the node purity scores have dropped from values in 10^5 to 10^4. This is a product of the remaining inflow volume forecast from June through September being much less than the March through September volume.

Overall, the climatic variables importance scores were highest early in the season and waned as the lead time got closer to and into the runoff season. The opposite is the case for the NRCS variables. Although, the node purity scores peaked at the middle lead times. SST anomalies in the Central Pacific and GPH's in the Pacific Northwest were consistently the most important climatic variables, with GPH's in the Pacific Northwest and Gulf of Mexico adding solid predictive skill at later lead times.

4.2.3. Regression Model Rankings

Now that the local polynomial and random forest models have been developed, we want to analyze which have the best predictive skill and if the climate-plus-NRCS models perform better than the current NRCS-only model. For ranking, predictive skill is measured by the RMSE and NSE scores from a LOOCV test. For the prior, the lower the RMSE the better the skill. For the latter, the higher. These scores for all our regression models and the adjusted NRCS forecasts are shown in Figure 19.



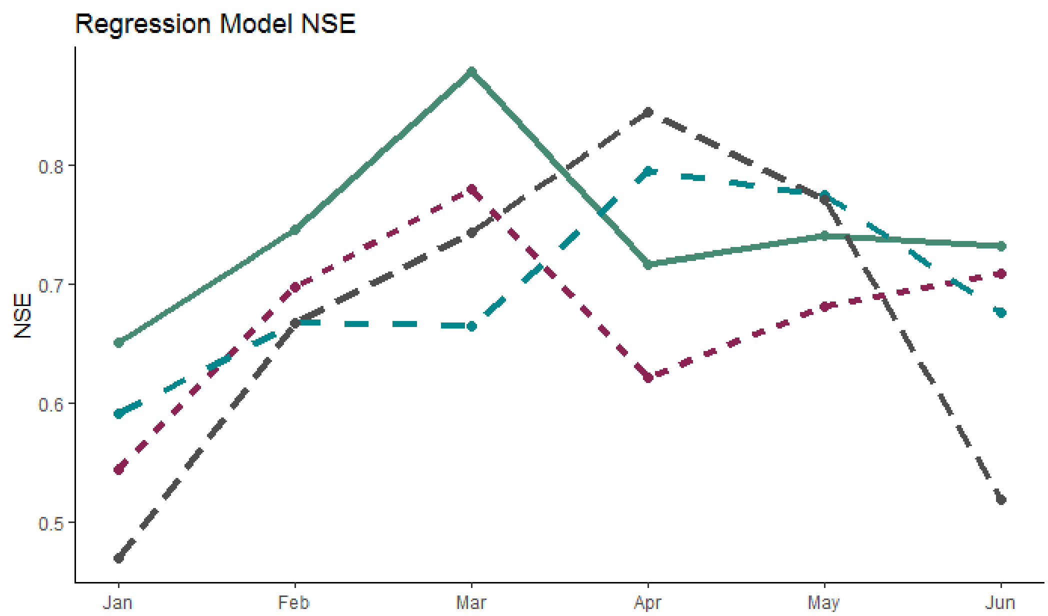


Figure 19. Forecast models' RMSE (top) and NSE (bottom) skill scores at each lead time. The forecast models are the climate-plus-NRCS local polynomial, climate-only local polynomial, random forest, and NRCS.

Of the six lead times that the seasonal inflow volume is forecast, a climate-plus-NRCS model had better predictive skill than the NRCS's model at four. The lead times predictive skill is better for January, February, March, and June 1st. At the January through March and June forecast lead times, our climate-plus-NRCS local polynomial models are the most skillful at forecasting seasonal inflow volumes. Compared to the NRCS models, the respective RMSE and NSE score differences are (-30 TAF, +0.2) for January, (-15 TAF, +0.1) for February, (-35 TAF, +0.2) for March, and (-12 TAF, +0.23) for June lead times. Also, for those aforementioned lead times, the skill scores of the climate-only local polynomial models surpass the NRCS models. This indicates that the teleconnections are predictively powerful and the regression method is not the only factor making the climate-plus-NRCS local polynomial model skillful.

The random forest model exceeds the NRCS model's skill at both January and June lead times - 20 TAF, +0.13 and -7 TAF, +0.17, respectively. However, they do not capture the relationships between our predictive information and seasonal inflow volumes as well as the local polynomial regression method. At April and May lead times, the random forest outperforms the local polynomial models but is worse than or nearly identical to the NRCS models. April and May are the lead times for which the NRCS models have their best NSE score, 0.85 and 0.78 respectively. Their RMSE scores are also low then, but that is partially due to the inflow volume being forecast in April and May being over a shorter period, therefore, lower in overall volume and magnitude of error.

Table 16 shows the best climate-plus-NRCS model for each lead time and specifies if their predictive skill exceeds the NRCS model.

Table 16. Best climate-plus-NRCS model for each lead time.

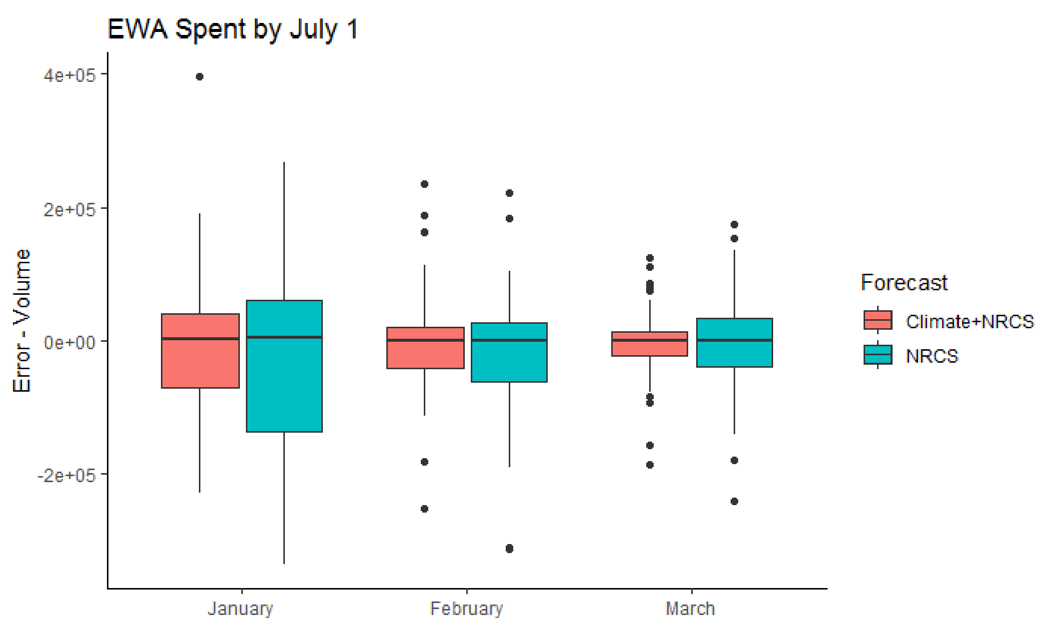
Lead Time	*Regression Method*	*Better than NRCS (Y/N)*
January	Local Polynomial	Y
February	Local Polynomial	Y
March	Local Polynomial	Y
April	Random Forest	N
May	Random Forest	N
June	Local Polynomial	Y

4.3. Performance Metric Scores

We ran the KROM as described in the Model Setup, Runs, and Results section in the Methodology chapter. We compared the operational performance of the climate-plus-NRCS and NRCS-only forecasts for lead times for which the prior exceeded the latter's skill. Thus, we ran operations from the January, February, and March lead times. As an exception, we did not run operations at the June lead time even though predictive skill was better. At June, it is too late to make any significant planning decisions so it is much less relevant than the other lead times.

Operational performance is analyzed based on the major operational objectives. Thus, environmental, agricultural and flood control performances were computed from the runs based on the climate-plus-NRCS and NRCS-only forecasts as error from a perfect forecast run. We present boxplots of these performance metrics below and discuss how to interpret the errors, for example, does the direction of the error matter, how may the operator change operations based on a better projection in a given year, and in what conditions are projection errors most detrimental to operations? The distribution, range, outliers, and abnormal tendencies of the box plotted errors are related to these questions. Boxplots are structured so the middle-shaded region, also called the intra-quartile range, contains the errors between the 25^{th} and 75^{th} percentile. The middle bar of the boxplot is the median of the errors. The lines extending from the intra-quartile range represent the distribution to the minimum and maximum, which are equivalent to the 25^{th} percentile + 1.5 times the intra-quartile range and the 75^{th} percentile – 1.5 times the intra-quartile range respectively. Lastly, the points outside the minimum and maximum limits represent the outliers.

The first performance metric we discuss represents the projected volume of environmental releases from Upper Klamath Lake by July 1^{st} and October 1^{st}. See Figure 20 for the error boxplots.



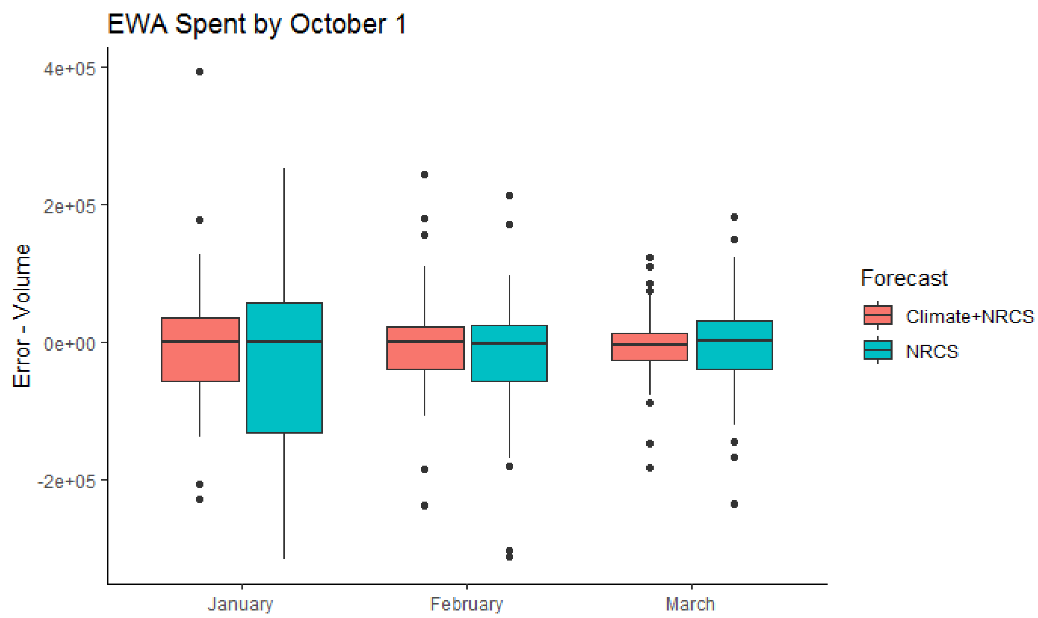


Figure 20. Error of projected EWA spent by June 1 (top) and October 1 (bottom) for lead times January thru March with respect to the perfect forecast results.

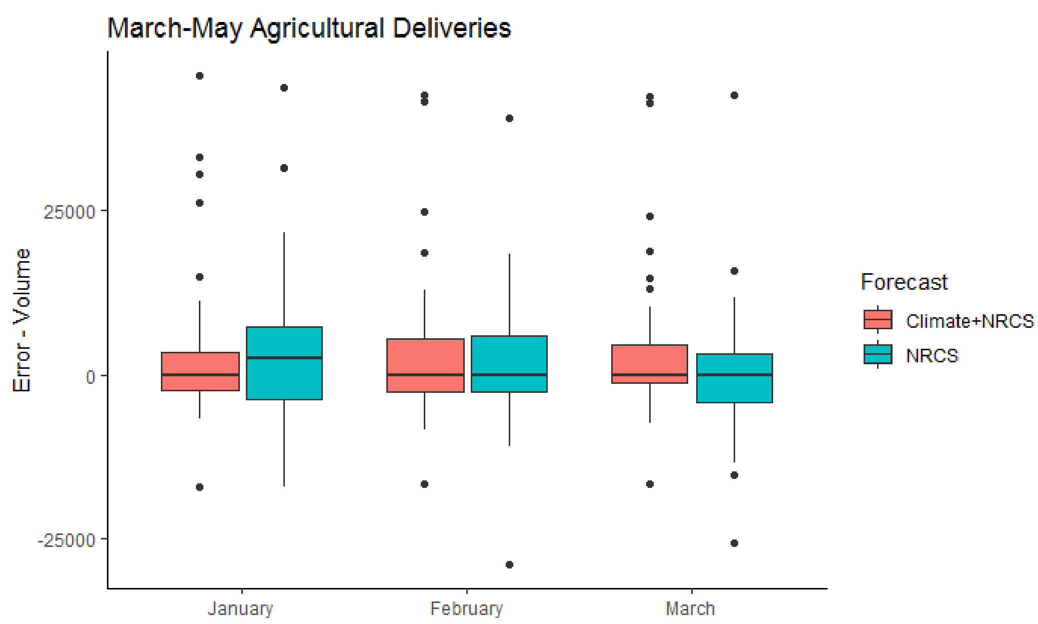
For the climate-plus-NRCS runs, the error range of EWA spent is lower for both analysis dates at each lead time, January through March. Interestingly, the range of errors is nearly identical on both analysis dates even though the spent EWA volume on July 1st is 20-40% less than on October 1st.

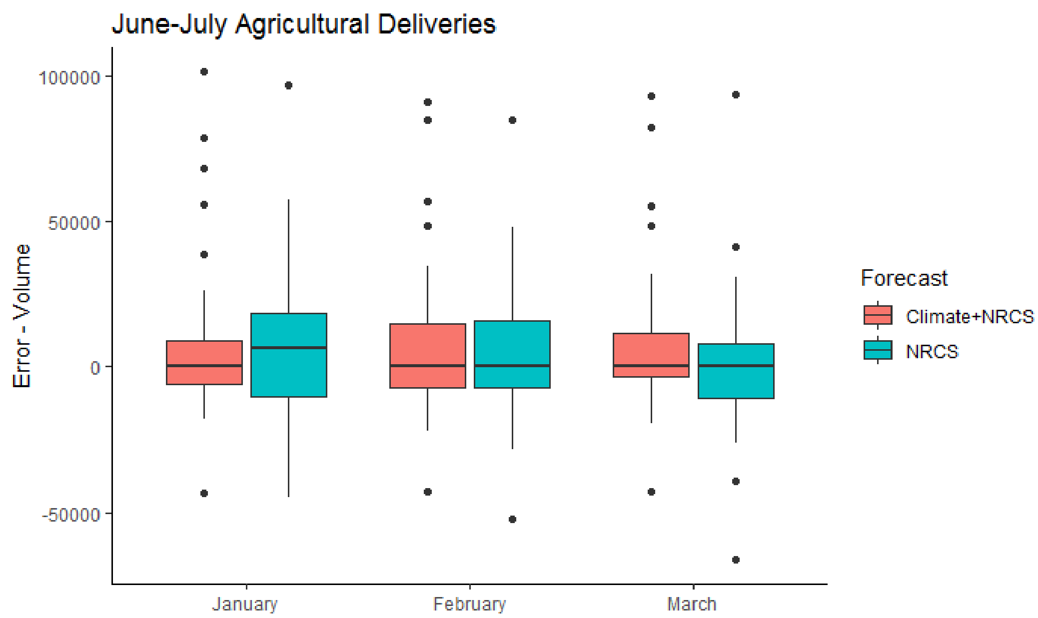
At the analyzed lead times, the forecasts determine the early season distribution of environmental releases. Errors of projected EWA spent, therefore, cause over or under misappropriation of early season supply. Excessive use early can deplete supply necessary for the summer. As a result, the later environmental releases are lower and the stream's water temperature rises more quickly in the hotter conditions. High water temperatures initiate algae and parasitic blooms, which have caused fish die off such as in 2002. Under-use early in the season is harmful to the spring spawning fish that need a higher river stage to migrate up the river.

Of note, excessive under-use early in the season is unlikely since the policy sets a minimum environmental allocation that drives base requirement releases. Thus, forecast overestimates are more problematic in dry years since there are no protections for late base releases if the minimum allocation has been depleted.

From the boxplots, the smaller range of projected EWA spent errors using the climate-plus-NRCS forecast gives confidence that the environmental releases are closer to the optimal amount before and into the spring. Notably, there are more outliers from the climate-plus-NRCS based runs. Upon inspection, these occur in wetter years when the forecast's error variability is greater. Projected EWA spent errors in wet years (1984, 1986, and 2011) are less consequential since water scarcity is not an issue.

The next performance metrics we discuss represent the projected agricultural delivery volume in March-May, June-July, and August-September from Upper Klamath Lake. See Figure 21 for the error boxplots.





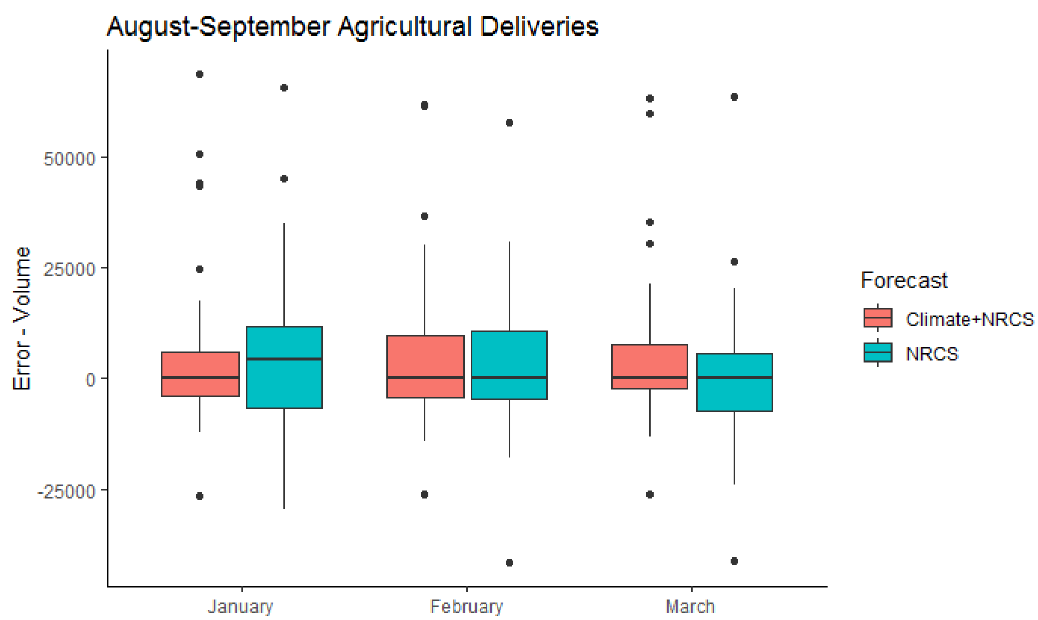


Figure 21. Error of projected agricultural deliveries from March-May (top), June-July (middle), and August-September (bottom) for lead times January thru March.

For each lead time, runs based on the climate-plus-NRCS forecasts have a smaller error range than runs based on the NRCS forecasts. The ranges are greatest from June-July since the largest agricultural deliveries occur then, therefore, causing greater error variability. Of note for each period is the similar error range of the runs based on the climate-plus-NRCS forecasts at the January and March lead times. This is surprising since the skill of the climate-plus-NRCS model at the March lead time is much better than at the January lead time. Unlike the environmental release metrics, the distribution of agricultural delivery errors for both runs, climate-plus-NRCS and NRCS based, skews towards overestimating the delivery volume. There is an exception for runs based on the NRCS forecast at the March lead time whose error distribution skews towards slightly underestimating deliveries. Once again, runs based on the climate-plus-NRCS forecasts have more outlier errors than runs based on the NRCS forecasts. These outlier errors are predominantly delivery overestimates.

Agricultural delivery projections are especially important since the relevant stakeholders, the irrigators, typically adjust their decisions to the forecasts by altering seeding rates or even crop selection based on the forecasted supply. Once these decisions are made, the irrigators rely on the forecasted supply to support what is planted. A similar dry year overuse and underuse dynamic that affects the EWA spent projections exists for the agricultural delivery projections. If the forecast is an overestimate, later distribution will have to be cut back, leading to shortages later that can decrease crop growth or cause crop kill. If the forecast is an underestimate, the lack of distribution early stunts crop growth at a critical growing stage and affects end of season yields. Forecast errors can also be harmful in moderate to moderately wet years. In these years, farmers aim to maximize profit by planting at higher seeding rates, which should boost yield. Thus, it is critical that crops are adequately irrigated or else the farmer will have committed a higher input cost for a standard yield, i.e. lower net gain.

From the boxplots, the error range of the climate-plus-NRCS based runs is smaller and skews towards an overestimate the later the lead time. The smaller range should give irrigators more confidence in their cropping choices, but they should be wary that there may be less water available than the forecasted supply. Some overestimate errors are a product of the policy cutting off agricultural deliveries entirely to satisfy the minimum environmental allocation volume. In reality this would not happen.

Rather, Reclamation and the irrigation districts would come to an agreement on an agricultural allocation volume. Then, the UKL pool elevation would be drawn down more and water would be borrowed from the Pacificorp reservoirs. Since the outlier errors occur in these no delivery years (1991, 1992, and 1994) the error range and skew of the climate-plus-NRCS based runs is better than the boxplots represent.

The following performance metric represents the projected Upper Klamath Lake flood release volume from January, February, and March through September. See Figure 22 for the error boxplots.

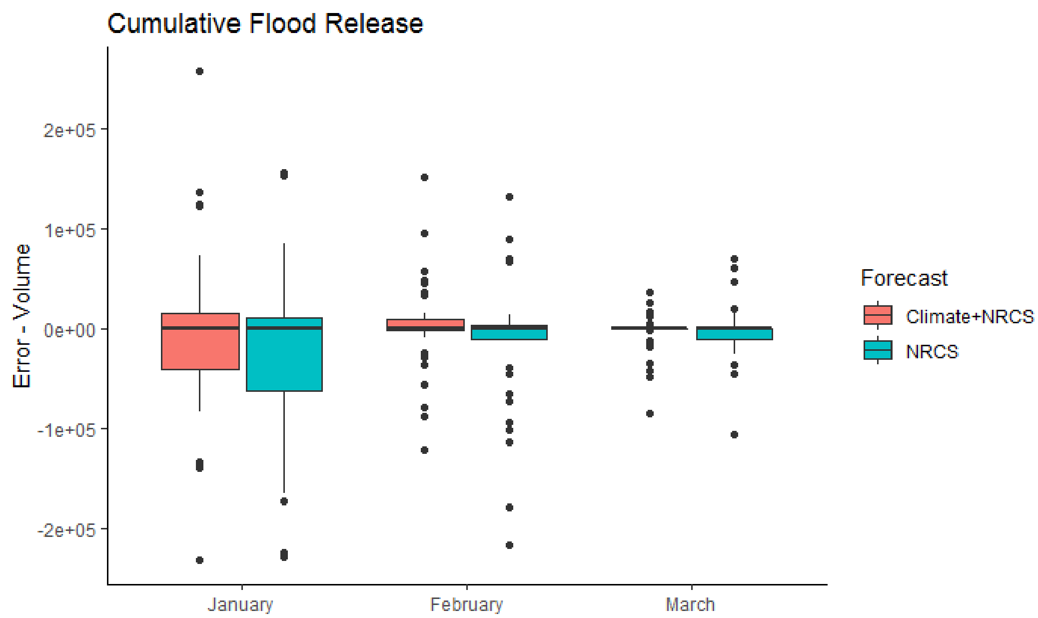


Figure 22. Error of projected flood release for lead times January through March.

The error distributions from runs based on the NRCS forecasts skew towards underestimating the flood release. At the January and February lead times, the error distributions from runs based on the climate-plus-NRCS forecasts skew towards underestimating and overestimating, respectively. At the March lead time, those runs’ errors are normally distributed. As expected, the error range tends to decrease with later lead times since only projected flood releases are tracked. For example, the cumulative flood release at the February lead time does not include flood releases in January and the cumulative at the March lead time does not include flood releases in January or February. There is a considerable amount of outlier errors for the cumulative flood release projected at the February and March lead times. Thus, the intra-quartile range is affected by years with no flood releases. These years are not uncommon since the Klamath River Basin is in a semi-arid region that relies on snow-melt. It is also why the median for each month is at 0.

The consequence of the direction of projected flood release errors is quite different depending on the year. In wet years, it is more favorable to overestimate the forecast. The operator will initiate flood releases earlier in this scenario. While unnecessary, there is little water stress and that released supply will not adversely affect other objectives. Whereas, if the forecast is an underestimate, the pool elevation will rise quicker than expected. As a result, the operator will initiate larger flood releases to reduce overtopping risk. Releases of these magnitudes will partially inundate downstream lands and cause damages. In dry years, it is more favorable to underestimate the forecast. No flood releases are made either way in this scenario. If the forecast overestimates in a dry year, any resultant flood release is wasting critical supply that could have been used for some other objective.

From the boxplots, the most notable aspect of the errors is the outliers compared to the intra-quartile range in February and March. This is due to 22 of 38 years in the analysis period having no flood releases. Thus, by erroring on only 2-6 of these years, most errors are zero and the intra-quartile range is very small. As a result, any marginal error greater than 15,000 acre-ft is an outlier. Even with the outlier errors, the range is meaningfully smaller for the climate-plus-NRCS based runs at each lead time. This is especially useful since flood control decisions are made late-January thru March. Take 2013 for example, the NRCS based run at the January lead time projected 47,000 acre-ft in flood releases when none were required. Comparatively, the climate-plus-NRCS based run at the same lead time projected 5,000 acre-ft in flood releases. While each corrected to no flood releases by the March lead time, a considerable amount of supply for other objectives would have been wasted if informed by the NRCS forecast.

The last performance metric we discuss represents the forecast of Upper Klamath Lake storage on June 1st and October 1st. See Figure 23 for the error boxplots.

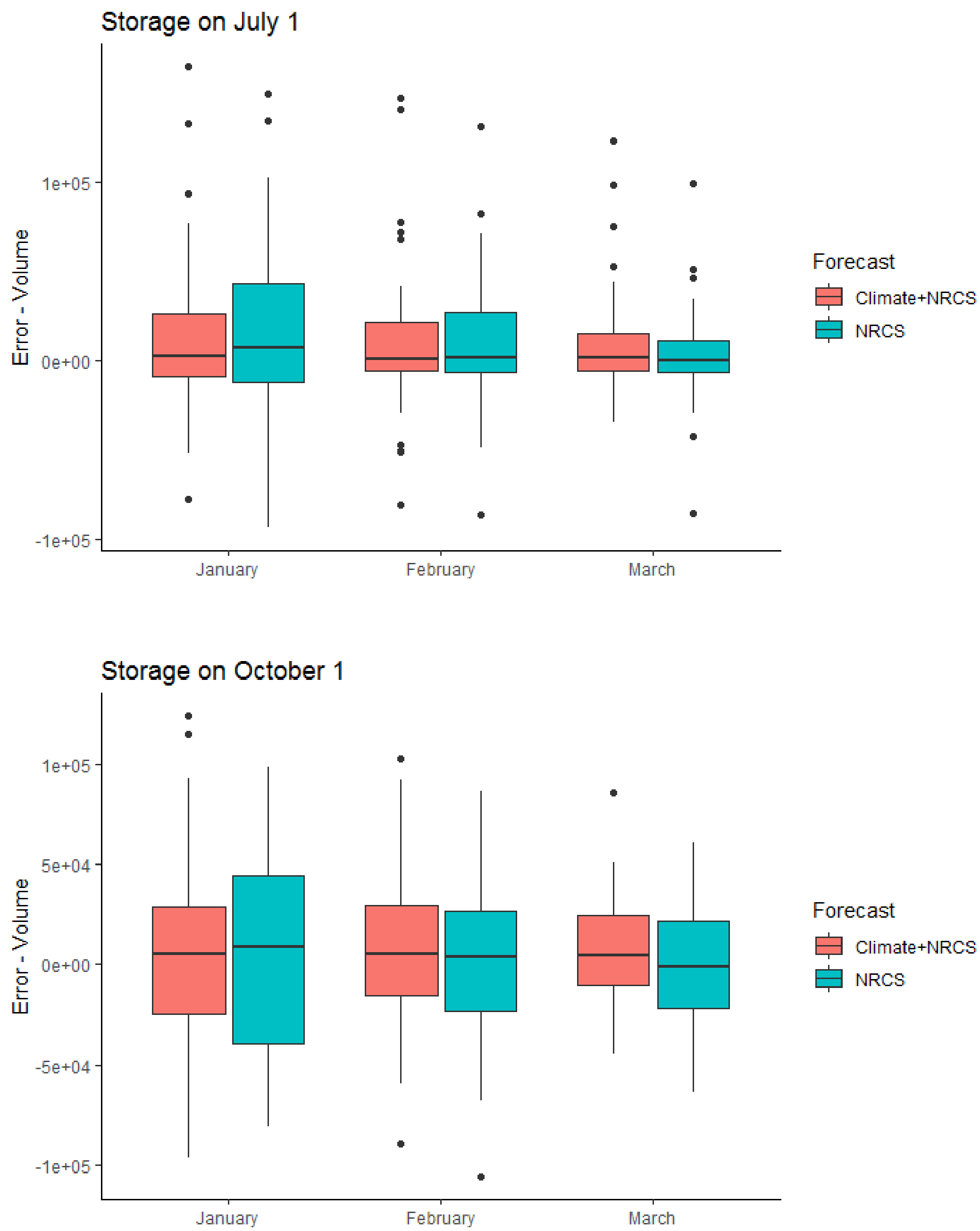


Figure 23. Error of projected Storage on June 1 (top) and October 1 (bottom) for lead times January through March.

The error distribution skews towards an overestimate of the July 1 storage for either forecast run, whereas, the error distribution is relatively normal for either run for the October 1 storage. For both storage dates, the range of errors is smaller for runs based on the climate-plus-NRCS forecasts with one exception. For the July 1 storage, the range of errors for runs based on the NRCS forecasts is smaller at the March lead time. Ranges for both analysis dates are similar since the operating logic tends to correct the storage towards a central tendency. More error outliers occur for the July 1 storage projections, with the climate-plus-NRCS forecast based runs having slightly more than the NRCS forecast based runs.

The projected UKL storage serves two purposes. First, it ensures the storage is at a level that sufficiently supports endangered fish. Second, it shows the storage deficit that will need to be recharged for the next water year. If either is unsatisfactory, the operator will alter their management decisions. The direction of the projected storage error has different effects on these decisions. If the storage is an overestimate, the operator is influenced to increase releases for agricultural or environmental demands. In the process, the storage is either depleted to a critical level or releases are limited in the summer. If the storage is an underestimate, the operator is influenced to be conservative with the supply. Thus, early season demands are shorted.

From the boxplots, the error range of the climate-plus-NRCS based runs are significantly smaller. Thus, decisions affecting storage and its distribution are likely to be better suited to the true seasonal supply. There is an exception, the error range of the July 1st storage is smaller for the NRCS based runs. This is not a concern since all major releases that deplete storage in the climate-plus-NRCS runs had smaller errors.

In this paragraph we present an error characteristic summary of the runs based on the climate-plus-NRCS forecasts compared with the NRCS forecasts for every metric at each lead time. First, the range of errors is noticeably smaller for forecasts that add climate. The smaller the range of errors, the more optimal the supply distribution for the year’s actual conditions. Therefore, early releases are less likely to deplete late season storage or curb early season demands. Second, the distribution of errors is balanced between overestimating and underestimating for most metrics. The one metric that is more likely to overestimate is the agricultural deliveries. This gives pause to irrigators when tailoring cropping decisions to the full forecasted supply. Third, adding climate to the forecast results in more outlier errors. Outlier errors are the most prevalent for the flood control metric, but this is misleading since there are so many years without flood releases that result in errors of 0 acre-ft. Overall, based on having a meaningfully smaller range of errors than runs based on the NRCS forecasts, the climate-plus-NRCS forecasts are effective at improving the operational projections. Thus, managers and stakeholders in the basin can have greater confidence in their planning decisions. But the particular characteristics of the errors for each metric should be considered in understanding how the errors could affect decisions.

Chapter 5: Conclusions and Future Work

This chapter summarizes and provides commentary on our study's findings and lays the framework for future work. From the results chapter, the skill of seasonal inflow volume forecasts is improved at January, February, March, and June first lead times. In the next section we discuss the value of more skillful/accurate forecasts for operations and studies in the Klamath River Basin. Following that, lack of improved skill for forecast at April and May first lead times is addressed; this explains why local and climatic variables are more predictively powerful at certain times of the year. We end the conclusion portion by summarizing what has been accomplished in this study. Then, the final section finishes with proposing further ways the results can be analyzed, how follow up forecasting work can build on our findings, and the next steps for moving towards RiverWare as the Klamath River Basin operations management tool.

5.1. Value of Improved Early Seasonal Forecasts

At January, February, March, and June first lead times, we were able to improve the skill of forecasts based on the RMSE and NSE objective criteria. Comparing the climate-plus-NRCS forecasts to the NRCS's, the respective RMSE and NSE score differences are (-30 TAF, +0.2) for January, (-15 TAF, +0.1) for February, (-35 TAF, +0.2) for March, and (-12 TAF, +0.23) for June lead times. We see these improvements translate to operational projections by reducing the range of uncertainty relative to having perfect foresight for each objective. With less uncertainty, the operators and stakeholders can feel more confident about their plans for the upcoming seasons (spring/summer). Examples of planning decisions affected by early lead time forecasts are as follows:

Agriculture: Based on water availability, farmers must decide how to manage resources for the upcoming season. For farmers with livestock sustained by grazing, they face constraints such as how many acres of pasture can be supplied with the available water and how many cattle can be supported from those acres? In dry years, this may mean planting a dryland mix since it does better with less water. Therefore, it can support more head of cattle than a water intensive forage that dies midseason from drought conditions. For grain farmers, they may have to go with a lower seeding rate or even leave fields to farrow if there is limited water for irrigation. This decision benefits the next crop rotation as less nutrients are depleted from the land. Having the confidence to make these time sensitive, tough decisions based on early seasonal forecasts can be a massive benefit to farmers looking to meet their bottom line in drought years and maximize yield in water surplus years.

Flood Control: Due to the majority of supply coming from snowmelt, operators are challenged early with conserving enough storage to last through the season while not putting UKL in danger of overtopping. Making flood releases in the lead up to a dry season cause diminished releases that persist through mid to late summer due to a lack of supply. Whereas, withholding storage in the lead up to a wet season causes prolonged flood releases that inundate downstream regions due to mitigating overtopping risk at UKL. With a more certain, early outlook of the seasonal supply, the operator can preemptively act to conserve critical storage or reduce flood risk.

Allocation Agreements: Much of the contention and subsequent litigation from stakeholders is caused by supply allocation in years of water stress/drought. Uncertainty plays a role in their frustration as allocations can change drastically as more or less water than expected is available as the season progresses. Stakeholders are realistic, they know the season cannot be perfectly projected and planned for. But, having a tighter range of likely seasonal conditions/scenarios would help bring everyone to the table to prepare contingency plans for either side of forecasted water availability spectrum. Thus, alleviating disagreements later on when the actual supply scenario plays out.

Seasonal Carryover: Because the Klamath River Basin is semi-arid, a full recharge of UKL is not guaranteed over the fall and winter. Therefore, the operator needs to consider how allocations for the current season will adversely affect the supply for the next year. Maybe all the objectives can be met, but if a drought year follows, there will be serious shortages that cause devastating economic and species losses. See the 2001 and 2002 operational challenges and outcomes referenced in the Introduction chapter. A more certain range of end of season storage outcomes will help the operator weight the consequences of early season decisions. Thus, not jeopardizing operations in the year to follow.

To summarize, a more certain range of forecasted seasonal supply maximizes agricultural economic gains, saves extra water or reduces flood risk, lessens litigation, and best prepares for the next year's operations. Both operators and stakeholder stand to benefit greatly from improved early seasonal forecasts that allow for more concentrated planning.

5.2. Explanation for Lack of Improvement in Late Seasonal Forecasts

At April and May first lead times, we were unable to improve the skill of forecasts based on the RMSE and NSE objective criteria. This suggests that the signals the climatic variables represent carries enough noise that the climate-plus-NRCS forecasts are less skillful than the NRCS forecasts. Why is this the case at these lead times and not the others?

To explain, we first reference the source of surface water in the spring/summer period. In the West, approximately 75-85% of surface water originates as snowmelt and the remaining percent as precipitation. This makes established, on-site snowpack measurements an excellent indicator of seasonal volumetric supply [Pagano, 2005]. For a better part of the last century, agencies that forecast water supply have used SWE as their on-site snowpack measurement. Therefore, at the snowpack peak, there is little variability in the supply that the SWE does not indicate. Western snowpack usually starts to accumulate late November, reaches its peak mid-March, and disappears late June [Siegel, 2009]. It is, therefore, expected that the skill of the NRCS forecast, which uses SWE from multiple stations as predictors, follows the same pattern. Thus, the forecast skill is lowest when the snowpack is forming, grows and peaks as the snowpack reaches its maximum, and decreases as the snowpack dissipates. See Figure 19.

Second, we consider the physical processes that connect the predictors with seasonal volumetric inflow. The NRCS forecasts are based on antecedent SWE, precipitation, streamflow, and groundwater well levels [Risley, 2005]. One or two processes connect these ground measures to UKL inflow. For example, precipitation becomes runoff that enters surface water and is transported downstream to UKL. Therefore, the connection is strong for the ground measures that directly precede the response, e.g., precipitation correlates strongly to seasonal volumetric inflow. Our climatic variables represent 700 hPa GPH's and SST's. These measures have been shown to affect airflow dynamics (SST's: circulatory patterns, GPH's: jet stream) that move, induce, or regulate precipitation events in the West [Soumya, 2015]. Compared with the ground measures that the NRCS forecasts are based on, many more processes connect the climate measures to seasonal volumetric inflow. Thus, there are more opportunities for uncertainty to compound and the lag between the signal and response is longer. Our results showed strong connections between GPH's/SST's and seasonal inflow volume up to 6 months prior to the start of the season, see the Best Predictive Variables section.

While there is more uncertainty in the climate connection, their longer lag allows these measures to inform the model of snowpack forming and/or mid to late season precipitation events depending on the forecast lead time. The climate-only forecasts' skill scores support this claim, see Figure 19. The highest scores occur at February and March lead times when variables are connected to both precipitation events. Lower scores occur at other lead times when variables are connected to one precipitation event (January: snowpack forming, April-June: mid to late season rainfall). Unlike the climate measures, the ground measures only inform the model of water that currently exists in the basin. Historical ground measurement patterns/trends are used to inform the model of future supply events, but the relationships are not physically supported.

Based on 1) SWE being a highly skilled predictor when snowpack has recently peaked and 2) climatic predictors representing more uncertainty than the ground measures, but carrying greater insight to snowpack forming and mid to late season precipitation events; we observe the following:

- The climate-plus-NRCS forecasts are the most skilled in January through March lead times when a high to moderate amount of the snowpack has yet to form, and the majority of the supply is yet to come from snowpack forming and mid- to late-season precipitation events.
- The NRCS forecasts are the most skilled in April and May when the snowpack maximum was recently achieved and future precipitation makes up a relatively small portion of the seasonal inflow volume.
- The climate-plus-NRCS forecasts are the most skilled in June when snowpack is nearly dissipated and a moderate to large portion of supply is yet to come from late-season precipitation events.

5.3. Study Accomplishments

Faced with the water management challenges of the West (i.e., more frequent droughts, competing operational objectives, and water rights litigation), our study aims to improve the decision support systems used for operations in the Klamath River Basin. To achieve this improvement, we address four key questions: 1) Can an operational model be developed in RiverWare that replicates the PA Calc's results, but is more robust and capable? 2) Can strong climate teleconnections to seasonal volumetric inflow in the Klamath River Basin be identified? 3) Using the climatic information and alternative regression methods, can the skill of the seasonal volumetric inflow forecast produced by the NRCS be improved upon? 4) Can higher skilled seasonal inflow forecasts improve operational projections to the extent that they are useful to stakeholders and managers?

The name of management tool we developed is the Klamath RiverWare Operations Model. Not only does it fulfill the base operational criteria (i.e., reproduce PA Calc's results and satisfy extent and features, data, policy, and workflow requirements), but it is also a noticeable improvement over the existing management tool, the PA Calc.

The inefficiencies of the PA Calc stem from the setup, runs (operational and scenario), and output products. Starting with setup, the operator begins their daily interaction with the PA Calc by manually transferring observed data from the database. Then, they adjust per run settings by shuffling through multiple worksheets. Configuring per run settings is an involved process. They change daily and are chosen by 1) running the model, 2) checking the results, 3) adjusting per run settings accordingly, and 4) repeating that process until official operational releases and projections are satisfactory. After the operational runs, the operator performs three types of scenario runs that use either A) historical UKL Inflow, B) historical accretions, or C) no agricultural demands. For both operational and scenario runs, the results are automatically post processed into plots and metrics, but they are only viewable for individual runs. Thus, if the operator wants to compare or distribute outputs, they must manually save them after each run.

By addressing those same inefficiencies, the KROM has proved itself a more powerful and intuitive operational model. Starting with setup once again, the operator executes a script that pulls the observed data from the database. Among other actions, it also opens the SCT that serves as the dashboard containing the per run settings. By automating the data import and centralizing the per run settings, it is much simpler to prepare and later alter. After the script executes, the operator generates the operational releases and projection. While setting configuration is still done iteratively, it is now much more streamlined using multiple runs. Using the SCT, the operator defines sets of per run settings. Then, they activate the trials version of the multiple run configuration, which performs as many runs as there are defined per run settings' sets. Not doing this setting configuration task with single runs is an immense time saver. Scenario runs are also prepared and performed with their own multiple run configurations, which allows the operator to run multiple of each scenario type per activation. Since runs are most commonly performed in a multiple format, the output products were designed accordingly. This means the results from each run of the multiple are comparable in plot or tabular data formats. Thus, the operator can either simultaneously evaluate many operational decisions or view a distribution of scenario outcomes. Additionally, each output can be generated into a html that is distributable to stakeholders.

Based on feedback from the workshop where we presented the KROM, the stakeholders were impressed. Specifically, they liked the KROM's added functionality and the proven applications of RiverWare as an operations management tool in other basins. While the switch is not going to be immediate, members of KBAO are committed to developing an updated policy set for the 2021 operating policy. Then, using the KROM in tandem with the new Excel management tool they will validate results and get familiar with the model switch. Eventually, they may wish to develop a planning version of the KROM.

Stepping back from the day to day operations, a glaring issue is the quality of seasonal volumetric inflow forecasts that on which KBAO bases their operational projections. The forecasts carry too much uncertainty, especially before the season begins. Thus, it is difficult for the stakeholders to plan for the season or the operator to evaluate the lasting effects of management decisions since the distribution of possible futures was so wide. The current producer of seasonal volumetric inflow forecasts is the NRCS. They generate forecasts on the first of each month from January through June. Their models exclusively use ground measurements inside the Klamath Basin as predictors.

To address a lack of outside the basin predictors, we analyzed climate measures (SST anomalies and 700 hPa GPH's) that have shown predictive skill in other Western basins. At multiple lead times and in regions, respectively, we found moderate to moderately strong teleconnections with season inflow volume to UKL. Lead times vary from 1 to 6 months ahead of the season. SST anomaly teleconnections are located in the East Atlantic, Pacific Northwest, Central Pacific, and Indian Ocean regions. The 700 hPa GPH's teleconnections are located Pacific Northwest, Southwest Atlantic, Gulf of Mexico and Northwest Atlantic. Serving as the foundation for our statistical modeling work in this study, we incorporated the variable representations of these climate measures with the NRCS forecasts in two types of regression models. The climate teleconnections' findings provide the groundwork for others to utilize this data to create their own forecast models. Additionally, it advocates for the usefulness of climate teleconnections to inform forecasts in other snowmelt driven basins.

The two types of regression models we developed were local polynomial and random forest. Both used a custom set of climate-plus-NRCS predictors for each lead time to generate seasonal volumetric inflow forecasts. To determine if the climatic information was predictively powerful, we compared the skill of the regression models' forecasts against the NRCS's based on LOOCV NSE and RMSE scores. At January, February, March, and June first lead times the climate-plus-NRCS local polynomial model had the best forecasting skill. This validated climatic teleconnections usefulness in early lead time seasonal volumetric inflow forecast, which suggests they have meaningful foresight of snowpack forming and mid- to late-season precipitation events. Additionally, the results advocate for alternative regression methods to be considered for future forecasting efforts.

Improving the forecast skill was encouraging, but the true value of the seasonal volumetric inflow forecasts needed to be judged in an operational context since that is their purpose. Thus, we compared the operational projections of runs based on either the climate-plus-NRCS or NRCS forecast, each relative to perfect foresight. This was done at each of the early season lead times that forecasting skill was improved, i.e., January thru March. From the comparison, the climate-plus-NRCS forecasts effectively reduced the range of errors for the operational objectives that KBAO projects regardless of the lead time. With greater confidence in the potential seasonal conditions, we expect the stakeholders and operators to benefit from better planning and assessment of their management decisions. For example, stakeholders have better information to make difficult cropping decisions, i.e. seeding rate and field mix, to maximize profits and operators have a tighter range of probable outcomes to assess the risk of flood control decisions. Additionally, this effort utilized the operation management tool we developed, i.e. the KROM. This allowed us to showcase its operational potential, but it also stress tested the model to various hydrological conditions to which it performed without issues. Thus, furthering confidence in the KROM's abilities.

To finish the recap of our study, we present the answer to our four key questions. 1) The KROM proves to be a more than suitable alternative to the current operational management tool based on the KBAO's criteria. It is more intuitive to operate, performs daily operational tasks quicker, and has positive feedback from stakeholders. 2) Both SST's and 700 hPa GPH's from multiple regions and lead times have moderate to moderately strong connections to UKL seasonal volumetric inflow that may be useful for forecasting. 3) The climate-plus-NRCS local polynomial forecasts were more skillful than the NRCS's at January, February, March, and June first lead times. 4) The range of operational projection error, relative to perfect foresight, was meaningfully reduced by basing runs on the climate-plus-NRCS forecasts. From the increased confidence in the seasonal conditions, stakeholders and operators stand to be benefit from better planning and decision assessment.

5.4. Future Work

Efforts or analysis outside of this study's scope involve both operations model development and statistical forecasting. Starting with the prior, a scope of work is being drafted for the development of an operations/planning model for the 2021 operational policy. This effort will build on the KROM's framework, e.g., update the layout, data management, and output products for either operations/planning run configurations and create new policy sets for the 2021 policy. Additionally, a testing phase will be included in the scope. It will involve two parts - first, testing the model with various observed hydrology and per run settings scenarios and validating the results, and second, supporting KBAO while they manage daily operations with both the Excel tool and KROM. The second part is to back the transition between management software.

Since recent years have experienced a larger frequency of drought years, we suggest that future studies of improved forecasting may consider non-stationarity. We evaluate our assumption of stationarity for the seasonal volumetric inflow that was used in this study by plotting the seasonal volumetric inflow over the period of record and adding an 8-year moving average to check if there is significant increase or decrease over time. The 8-year moving average is plotted with each year's seasonal value in figure 24.

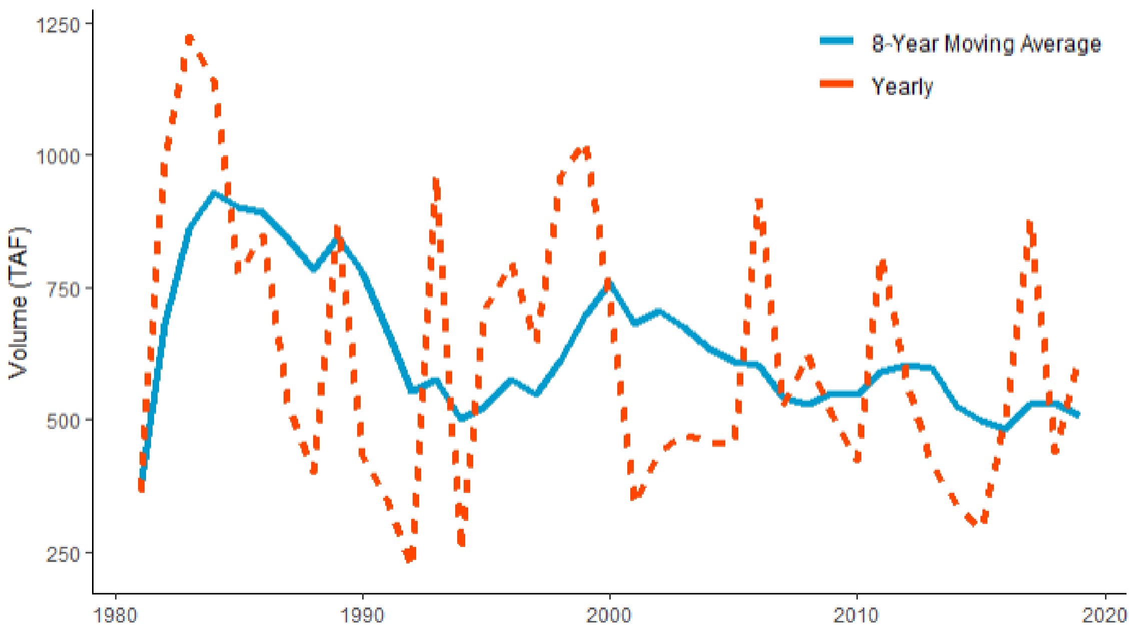


Figure 24. Seasonal volumetric inflow at UKL from 1981 to 2019.

Based on the 8-year moving average, the results suggest that the seasonal volumetric inflow decreases slightly with time. This apparent trend may be misleading since the first available years in the period of record are during a wet period. Another wet period occurs in the late 1990's, which also raises the average near the record's midpoint. A more rigorous trend analysis could be undertaken in future efforts that may benefit from considering a decreasing inflow trend.

As another adjustment to the regression effort, the South Indian Ocean SST anomalies could be excluded from the set of predictors. While there exists a moderately strong relationship between that region and the seasonal volumetric inflow, the physical connection is lacking. The lack of physical connection is due to the distance. The circulation response that the South Indian Ocean's SST's induce does not reach the Klamath River Basin.

To expand the scope of the statistical forecasting research, a deeper analysis into the skill of climate-plus-NRCS forecasts could be performed. Specifically, this could include evaluating and comparing their performance for subsets of years, e.g., dry vs wet or El Nino vs La Nina. Thus, further insight to when climatic teleconnections are most predictively useful may be gained. Additionally, there are other climatic measures not evaluated in our study that may possess meaningful connections to seasonal volumetric inflow in the Klamath River Basin, e.g., longwave radiation or wind vectors [NCEI, 2020]. Such measures could be evaluated with our predictive variable identification methodology.

References

Bracken, C., Rajagopalan, B., and Prairie, J. (2010). A multisite seasonal ensemble streamflow forecasting technique. *Water Resour. Res.*, 46(3), W03532. doi:10.1029/2009WR007965

Brems, M. (2017, April). A One-Stop Shop for Principal Component Analysis. *Medium*. Retrieved from https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c

Boehlert, B.B. and Jaeger, W.K., (2010). Past and future water conflicts in the Upper Klamath Basin: An economic appraisal. *Water Resour. Res.*, 46, W10518. doi:10.1029/2009WR007925

Burns, J. (2018, March). Another Drought Year Promises to Test Klamath Basin. *Oregon Public Broadcasting*. Retrieved from https://www.opb.org/news/article/another-drought-year-promises-to-test-klamath-basin/

Craven, P. and Wahba, G. (1978). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross validation. *Numer. Math.*, 31, 377-403. doi:10.1007/BF01404567

Frost, J. (2017). Overfitting Regression Models: Problems, Detection, and Avoidance. *Statistics by Jim*. Retrieved from https://statisticsbyjim.com/regression/overfitting-regression-models/#comments

Haque, M., Rahman, A., Hagare, D. and Kibria, G. (2013). Principal Component Regression Analysis in Water Demand Forecasting: An Application to the Blue Mountains, NSW, Australia. *Journal of Hydrology and Environment Research*, 1, 49-59. Retrieved from https://www.researchgate.net/publication/268209977_Principal_Component_Regression_Analysis_in_Water_Demand_Forecasting_An_Application_to_the_Blue_Mountains_NSW_Australia

Hay, L.E., McCabe, G.J., Clark, M.P. and Risley, J.C. (2009). Reducing Streamflow Forecast Uncertainty: Application and Qualitative Assessment of the Upper Klamath River Basin, Oregon. *Journal of the American Water Resources Association (JAWRA)*, 45(3), 580-596. doi:10.1111/j.1752-1688.2009.00307.x

Kaplan, Alexy & National Center for Atmospheric Research Staff (Eds). (2020). *The Climate Data Guide: Kaplan Sea Surface Temperature Anomalies*, [1980-2019]. Retrieved from https://climatedataguide.ucar.edu/climate-data/kaplan-sea-surface-temperature-anomalies

Loader, C. (1999). *Local regression and likelihood*. Springer, New York.

Nash, J. E. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*. 10(3): 282–290. doi:10.1016/0022-1694(70)90255-6

National Center for Atmospheric Research (NCAR). (2020). *700 hPa Geopotential Heights*, [1980-2019]. Retrieved from https://climatedataguide.ucar.edu/climate-data/ncep-reanalysis-r2

National Center for Environmental Information (NCEI), (2020). *Teleconnections*. National Oceanic and Atmospheric Administration. Retrieved from https://www.ncdc.noaa.gov/teleconnections/

Oregon Water Resources Department (ORWD), (2018). *Drought annex state of Oregon emergency operations plan: Oregon Water Resources Department*. Retrieved from https://www.oregon.gov/owrd/WRDPublications1/2016ORDroughtAnnex.pdf

Pagano, T. C., and Garen, D. C. (2006). Integration of climate information and forecasts into western US water supply forecasts. In: *Climate Variations, Climate Change, and Water Resources Engineering*, J. D. Garbrecht and T. C. Piechota (eds.), American Society of Civil Engineers, Reston, VA, pp. 86-103

Pasteris, P.A., Vigil, R., and Lea, J. (2002). *Water Supply Forecasting in the Upper Klamath Lake Basin During the 2001 Drought*. 15th Conf. on Biometeorology/Aerobiology and 16th international Congress of Biometeorology, Kansas City, KS, Amer. Meteor. Soc.

Reclamation, (2019). *Biological Opinion on the Effects of Proposed Klamath Project Operations from April 1, 2019, through March 31, 2024, on the Lost River Sucker and the Shortnose Sucker.* U.S. Department of the Interior, Bureau of Reclamation

Regonda, S. K., Rajagopalan, B., Clark, M., and Zagona, E. (2006). A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin. *Water Resour. Res.*, 42, W09404, doi:10.1029/2005WR004653.

Rieker, J.D., Coors, S., Mann, M. and Scott, T. (2005). *Modeling in Support of Water Operations in the Truckee River Basin*. Proceedings of Watershed Management 2005, EWRI and ASCE. Williamsburg, VA.

Risley, J.C. (2019). *Using the precipitation-runoff modeling system to predict seasonal water availability in the upper Klamath River basin, Oregon and California: U.S. Geological Survey Scientific Investigations Report 2019-5044.* USGS, https://doi.org/10.3133/sir20195044

Risley, J.C., Gannett, M.W., Lea, J.K., and Roehl, E.A., Jr. (2005). *An analysis of statistical methods for seasonal flow forecasting in the Upper Klamath River Basin of Oregon and California: U.S. Geological Survey Scientific Investigations Report 2005-5177.* USGS, Retrieved from http://pubs.usgs.gov/sir/2005/5177/

Schenk, L.N., Anderson, C.W., Diaz, P. and Stewart, M.A. (2016). *Evaluating external nutrient and suspended-sediment loads to Upper Klamath Lake, Oregon, using surrogate regressions with real-time turbidity and acoustic backscatter data: U.S. Geological Survey Scientific Investigations Report 2016-5167*. USGS, https://doi.org/10.3133/sir20165167

Siegel, J., (2009). *Examining Monthly Relationships Between Temperature, Precipitation, Snowpack, and Streamflow in the Upper Klamath Basin Over a 26 Year SNOTEL Record.* Department of Earth Science and Geography Master's Thesis, Vassar College, Poughkeepsie, NY

Sagarika, S., Kalra, A., and Ahmad, S. (2015). Pacific Ocean SST and Z_{500} climate variability and western U.S. seasonal streamflow: *International Journal of Climatology*, 36(3), 1515-1533, https://doi.org/10.1002/joc.4442

U.S. Department of Agriculture (USDA), (2011). *Snow Survey and Water Supply Forecasting. National Engineering Handbook Part 622*. Water and Climate Center, Natural Resources Conservation Service

U.S. Department of Agriculture (USDA), (2020). *Water Supply Forecast Chart* [1981-2019]. Retrieved from https://www.nrcs.usda.gov/wps/portal/wcc/home/quicklinks/forecastCharts

Wang H. and Ting M. (2000). Covariabilities of Winter U.S. Precipitation and Pacific Sea Surface Temperatures. *Journal of Climate*, 13, 3711-3719, doi:10.1175/15200442(2000)013<3711:CO@USP>2.0.CO;2

Wagner, B.J. and Gannett, M.W. (2014). *Evaluation of alternative groundwater-management strategies for the Bureau of Reclamation Klamath Project, Oregon and California: U.S. Geological Survey Scientific Investigations Report 2014-5054*, USGS, http://dx.doi.org/10.3133/sir20145054

Zagona, E.A., Fulp, T.J., Shane, R., Magee, T.M. and Goranflo, H.M. (2001). RiverWare: A Generalized Tool for Complex Reservoir System Modeling. *Journal of the American Water Resources Association*, 37(4), 913-929, doi:10.1111/j.1752-1688.2001.tb05522.x

Appendix

A. Predictive Variable Correlation Matrices

Table 17. Correlation matrix of the climate (SST's and GPH's) and NRCS predictors for the February lead time.

	East Atlantic SST	*Pacific NW SST*	*Central Pacific SST*	*South Indian SST*	*Pacific NW GPH*	*SW Atlantic GPH*	*Feb 1st NRCS Forecast*
East Atlantic SST	1.00	0.27	0.20	0.58	0.29	0.29	-0.34
Pacific NW SST	0.27	1.00	0.52	0.18	0.41	0.07	-0.45
Central Pacific SST	0.20	0.52	1.00	0.46	0.09	0.17	-0.36
South Indian SST	0.58	0.18	0.46	1.00	0.26	0.22	-0.49
Pacific NW GPH	0.28	0.41	0.09	0.26	1.00	0.29	-0.75
SW Atlantic GPH	0.29	0.07	0.17	0.22	0.29	1.00	-0.39
Feb 1st NRCS Forecast	-0.34	-0.45	-0.36	-0.49	-0.75	-0.39	1.00

Table 18. Correlation matrix of the climate (SST's and GPH's) and NRCS predictors for the March lead time.

	East Atlantic SST	*Pacific NW SST*	*Central Pacific SST*	*South Indian SST*	*Pacific NW GPH*	*SW Atlantic GPH*	*Mar 1st NRCS Forecast*
East Atlantic SST	1.00	0.27	0.20	0.58	0.27	0.29	-0.25
Pacific NW SST	0.27	1.00	0.52	0.18	0.30	0.07	-0.47
Central Pacific SST	0.20	0.52	1.00	0.46	0.11	0.17	-0.38
South Indian SST	0.58	0.18	0.46	1.00	0.21	0.22	-0.42
Pacific NW GPH	0.27	0.30	0.11	0.21	1.00	0.53	-0.73
SW Atlantic GPH	0.29	0.07	0.17	0.22	0.53	1.00	-0.44
Mar 1st NRCS Forecast	-0.25	-0.47	-0.38	-0.42	-0.73	-0.44	1.00

Table 19. Correlation matrix of the climate (SST's and GPH's) and NRCS predictors for the April lead time.

	East Atlantic SST	*Pacific NW SST*	*Central Pacific SST*	*South Indian SST*	*Pacific NW GPH*	*SW Atlantic GPH*	*Gulf of Mexico GPH*	*NW Atlantic GPH*	*Apr 1st NRCS Forecast*
East Atlantic SST	1.00	0.27	0.20	0.58	0.30	0.29	0.53	-0.01	-0.24
Pacific NW SST	0.27	1.00	0.52	0.18	0.31	0.07	0.39	-0.42	-0.47
Central Pacific SST	0.20	0.52	1.00	0.46	0.34	0.17	0.71	-0.30	-0.45
South Indian SST	0.58	0.18	0.46	1.00	0.24	0.22	0.56	-0.02	-0.33
Pacific NW GPH	0.30	0.31	0.34	0.24	1.00	0.61	0.37	-0.50	-0.69
SW Atlantic GPH	0.29	0.07	0.17	0.22	0.61	1.00	0.21	-0.27	-0.49
Gulf of Mexico GPH	0.53	0.39	0.71	0.56	0.37	0.21	1.00	-0.36	-0.42
NW Atlantic GPH	-0.01	-0.43	-0.30	-0.02	-0.50	-0.27	-0.36	1.00	0.34
Apr 1st NRCS *Forecast*	-0.24	-0.47	-0.45	-0.33	-0.69	-0.49	-0.42	0.34	1.00

Table 20. Correlation matrix of the climate (SST's and GPH's) and NRCS predictors for the May lead time.

	East Atlantic SST	*Pacific NW SST*	*Central Pacific SST*	*South Indian SST*	*Pacific NW GPH*	*SW Atlantic GPH*	*Gulf of Mexico GPH*	*NW Atlantic GPH*	*May 1st NRCS Forecast*
East Atlantic SST	1.00	0.27	0.20	0.58	0.34	0.29	0.66	-0.01	-0.24
Pacific NW SST	0.27	1.00	0.52	0.18	0.29	0.07	0.35	-0.42	-0.44
Central Pacific SST	0.20	0.52	1.00	0.46	0.33	0.17	0.42	-0.30	-0.45
South Indian SST	0.58	0.18	0.46	1.00	0.22	0.22	0.54	-0.02	-0.38
Pacific NW GPH	0.34	0.30	0.33	0.22	1.00	0.57	0.50	-0.36	-0.65
SW Atlantic GPH	0.29	0.07	0.17	0.22	0.57	1.00	0.33	-0.27	-0.45
Gulf of Mexico GPH	0.66	0.35	0.42	0.54	0.50	0.33	1.00	-0.40	-0.44

NW Atlantic GPH	-0.01	-0.43	-0.30	-0.02	-0.36	-0.27	-0.40	1.00	0.35
May 1st NRCS Forecast	-0.24	-0.44	-0.45	-0.38	-0.65	-0.45	-0.44	0.35	1.00

Table 21. Correlation matrix of the climate (SST's and GPH's) and NRCS predictors for the June lead time.

	East Atlantic SST	*Pacific NW SST*	*Central Pacific SST*	*South Indian SST*	*Pacific NW GPH*	*SW Atlantic GPH*	*Gulf of Mexico GPH*	*NW Atlantic GPH*	*Jun 1st NRCS Forecast*
East Atlantic SST	1.00	0.27	0.20	0.58	0.34	0.29	0.66	-0.01	-0.18
Pacific NW SST	0.27	1.00	0.52	0.18	0.29	0.07	0.35	-0.42	-0.37
Central Pacific SST	0.20	0.52	1.00	0.46	0.33	0.17	0.42	-0.30	-0.49
South Indian SST	0.58	0.18	0.46	1.00	0.22	0.22	0.54	-0.02	-0.35
Pacific NW GPH	0.34	0.30	0.33	0.22	1.00	0.57	0.50	-0.36	-0.59
SW Atlantic GPH	0.29	0.07	0.17	0.22	0.57	1.00	0.33	-0.27	-0.50
Gulf of Mexico GPH	0.66	0.35	0.42	0.54	0.50	0.33	1.00	-0.40	-0.40
NW Atlantic GPH	-0.01	-0.43	-0.30	-0.02	-0.36	-0.27	-0.40	1.00	0.41
Jun 1st NRCS Forecast	-0.18	-0.37	-0.49	-0.35	-0.59	-0.50	-0.40	0.41	1.00

B. Scree Plots

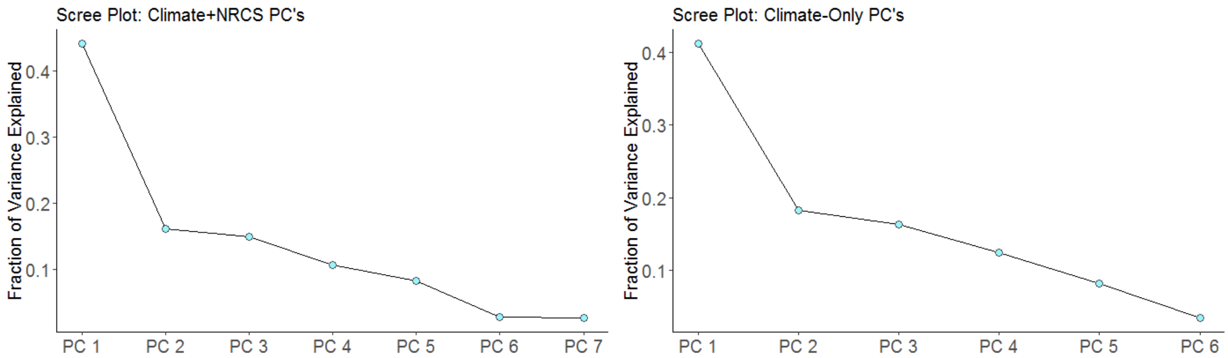


Figure 25. Scree Plots for the principal components formed from the (Left) climate-plus-NRCS and (Right) climate-only predictor sets for the February lead time.

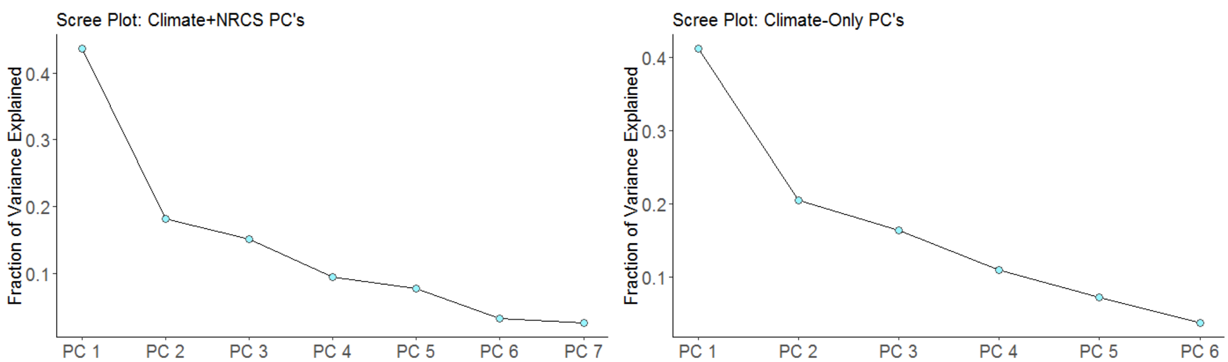


Figure 26. Scree Plots for the principal components formed from the (Left) climate-plus-NRCS and (Right) climate-only predictor sets for the March lead time.

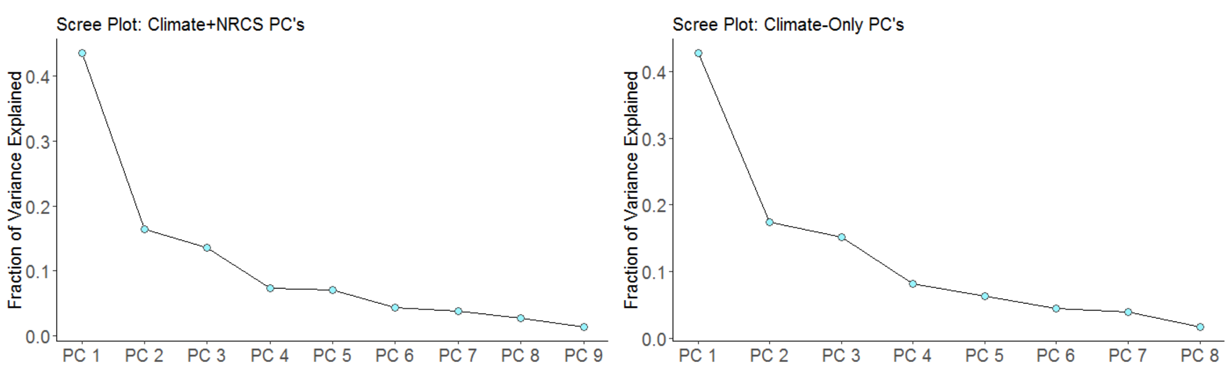


Figure 27. Scree Plots for the principal components formed from the (Left) climate-plus-NRCS and (Right) climate-only predictor sets for the April lead time.

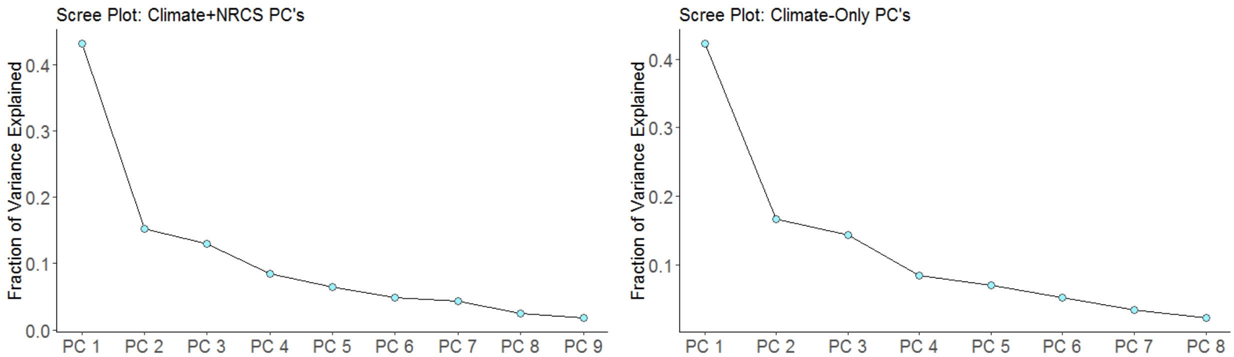


Figure 28. Scree Plots for the principal components formed from the (Left) climate-plus-NRCS and (Right) climate-only predictor sets for the May lead time.

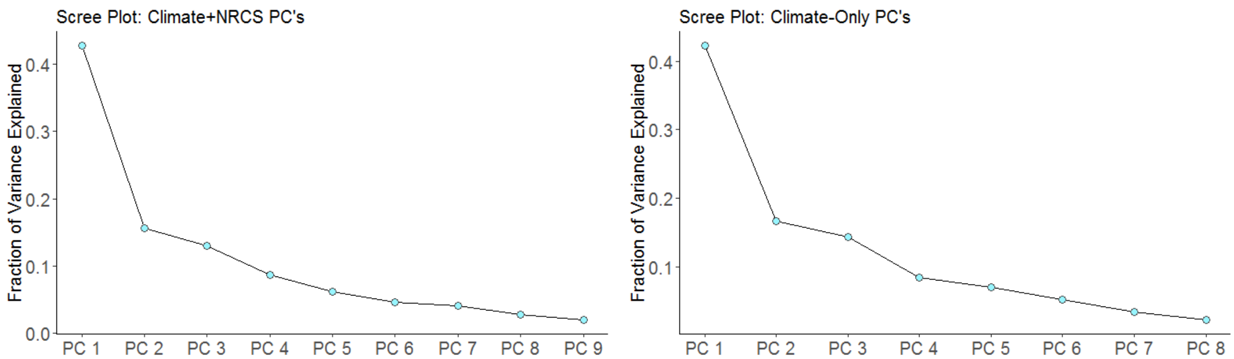


Figure 29. Scree Plots for the principal components formed from the (Left) climate-plus-NRCS and (Right) climate-only predictor sets for the June lead time.

C. Predictor/Principal Component Correlation Matrices

Table 22. Loadings matrix of the predictive variables (Top: climate-plus-NRCS, Bottom: climate-only) and principal components used to fit the local polynomial regression model at the February lead time.

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*
East Atlantic SST	0.36	-0.04	-0.51	-0.38	0.56	0.35	0.16
Pacific NW SST	0.35	-0.32	0.53	-0.07	0.50	-0.49	-0.03
Central Pacific SST	0.33	-0.61	0.09	0.45	-0.17	0.52	-0.10
South Indian SST	0.40	-0.24	-0.49	-0.17	-0.44	-0.49	-0.29
Pacific NW GPH	0.40	0.47	0.33	-0.25	-0.12	0.30	-0.59
SW Atlantic GPH	0.29	0.44	-0.25	0.74	0.27	-0.20	-0.06

Feb 1st NRCS Forecast	-0.48	-0.23	-0.20	0.07	0.37	-0.01	-0.73

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*
East Atlantic SST	0.45	-0.26	-0.33	0.42	-0.56	0.35
Pacific NW SST	0.41	0.49	0.45	0.14	-0.37	-0.48
Central Pacific SST	0.41	0.54	-0.15	-0.47	0.14	0.52
South Indian SST	0.47	-0.05	-0.53	0.12	0.49	-0.49
Pacific NW GPH	0.38	-0.28	0.61	0.30	0.48	0.30
SW Atlantic GPH	0.31	-0.56	0.15	-0.69	-0.21	-0.20

Table 23. Loadings matrix of the predictive variables (Top: climate-plus-NRCS, Bottom: climate-only) and principal components used to fit the local polynomial regression model at the March lead time.

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*
East Atlantic SST	0.34	0.19	-0.57	0.50	-0.32	0.25	0.34
Pacific NW SST	0.34	0.31	0.53	0.42	-0.33	-0.36	-0.29
Central Pacific SST	0.34	0.49	0.25	-0.55	-0.14	0.49	0.12
South Indian SST	0.38	0.33	-0.47	-0.17	0.46	-0.31	-0.44
Pacific NW GPH	0.41	-0.50	0.13	0.20	0.21	0.54	-0.43
SW Atlantic GPH	0.34	-0.47	-0.20	-0.45	-0.58	-0.31	-0.07
Mar 1st NRCS Forecast	-0.47	0.21	-0.24	-0.01	-0.43	0.28	-0.64

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*
East Atlantic SST	-0.45	0.02	-0.49	-0.50	0.40	0.39
Pacific NW SST	-0.38	-0.37	0.56	-0.43	0.18	-0.45
Central Pacific SST	-0.41	-0.48	0.22	0.55	0.04	0.50
South Indian SST	-0.46	-0.18	-0.53	0.18	-0.47	-0.48

Pacific NW GPH	-0.38	0.51	0.34	-0.21	-0.58	0.30
SW Atlantic GPH	-0.36	0.58	0.09	0.44	0.50	-0.29

Table 24. Loading matrix of the predictive variables (Top: climate-plus-NRCS, Bottom: climate-only) and principal components used to fit the local polynomial regression model at the April lead time.

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*	*PC 8*	*PC 9*
East Atlantic SST	0.28	0.42	-0.33	0.61	-0.04	-0.14	0.20	0.04	-0.44
Pacific NW SST	0.31	-0.05	0.48	0.41	-0.51	-0.17	-0.31	0.06	0.33
Central Pacific SST	0.37	0.17	0.37	-0.51	0.03	-0.34	-0.03	0.35	-0.45
South Indian SST	0.31	0.49	-0.19	-0.16	0.07	0.53	-0.53	0.08	0.16
Pacific NW GPH	0.38	-0.37	-0.25	0.02	0.00	0.22	0.36	0.65	0.25
SW Atlantic GPH	0.28	-0.30	-0.55	-0.11	0.09	-0.55	-0.42	-0.16	0.11
Gulf of Mexico GPH	0.40	0.30	0.16	-0.08	0.37	-0.17	0.43	-0.37	0.49
NW Atlantic GPH	-0.27	0.42	-0.31	-0.31	-0.59	-0.25	0.22	0.14	0.28
Apr 1st NRCS Forecast	-0.39	0.24	0.09	0.25	0.49	-0.33	-0.21	0.51	0.27

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*	*PC 8*
East Atlantic SST	0.33	0.40	-0.30	0.58	-0.16	0.28	-0.01	0.44
Pacific NW SST	0.33	-0.14	0.46	0.57	0.46	-0.17	0.14	-0.28
Central Pacific SST	0.40	0.08	0.37	-0.48	0.36	0.13	-0.04	0.56
South Indian SST	0.35	0.48	-0.14	-0.18	-0.12	-0.75	0.07	-0.13
Pacific NW GPH	0.37	-0.39	-0.31	0.01	0.08	-0.09	-0.77	-0.08
SW Atlantic GPH	0.27	-0.29	-0.61	-0.15	0.33	0.11	0.56	-0.09
Gulf of Mexico GPH	0.45	0.21	0.16	-0.22	-0.32	0.51	0.01	-0.57
NW Atlantic GPH	-0.28	0.55	-0.21	-0.05	0.63	0.19	-0.26	-0.25

Table 25. Loading matrix of the predictive variables (Top: climate-plus-NRCS, Bottom: climate-only) and principal components used to fit the local polynomial regression model at the May lead time.

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*	*PC 8*	*PC 9*
East Atlantic SST	0.31	0.53	-0.04	0.36	-0.32	0.18	-0.01	0.11	-0.58
Pacific NW SST	0.30	-0.26	0.49	0.10	-0.54	0.35	0.22	0.04	0.36
Central Pacific SST	0.33	-0.08	0.43	-0.43	0.29	0.24	-0.54	-0.07	-0.27
South Indian SST	0.31	0.50	0.15	-0.26	0.36	-0.09	0.40	0.41	0.30
Pacific NW GPH	0.37	-0.19	-0.38	-0.07	-0.28	-0.31	-0.45	0.52	0.18
SW Atlantic GPH	0.29	-0.08	-0.62	-0.10	0.17	0.66	0.12	-0.17	0.14
Gulf of Mexico GPH	0.41	0.21	0.03	0.41	0.16	-0.28	-0.24	-0.56	0.38
NW Atlantic GPH	-0.26	0.52	-0.09	-0.51	-0.47	0.05	-0.22	-0.26	0.26
May 1st NRCS Forecast	-0.39	0.20	0.11	0.41	0.20	0.42	-0.43	0.36	0.32

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*	*PC 8*
East Atlantic SST	0.37	-0.48	0.01	-0.41	0.25	0.15	-0.16	-0.59
Pacific NW SST	0.32	0.35	-0.44	-0.16	0.56	0.35	0.17	0.30
Central Pacific SST	0.36	0.15	-0.41	0.62	-0.05	-0.17	-0.43	-0.29
South Indian SST	0.35	-0.47	-0.21	0.27	-0.35	0.17	0.61	0.13
Pacific NW GPH	0.38	0.17	0.42	0.11	0.34	-0.60	0.39	-0.08
SW Atlantic GPH	0.29	0.04	0.65	0.32	0.04	0.56	-0.22	0.15
Gulf of Mexico GPH	0.46	-0.14	-0.01	-0.36	-0.27	-0.31	-0.40	0.56
NW Atlantic GPH	-0.27	-0.59	-0.03	0.31	0.56	-0.15	-0.17	0.34

Table 26. Loading matrix of the predictive variables (Top: climate-plus-NRCS, Bottom: climate-only) and principal components used to fit the local polynomial regression model at the June lead time.

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*	*PC 8*	*PC 9*
East Atlantic SST	0.31	0.54	-0.06	0.35	-0.25	0.14	-0.06	-0.13	-0.62

Pacific NW SST	0.30	-0.21	0.52	0.19	-0.56	0.35	-0.13	0.11	0.32
Central Pacific SST	0.34	-0.06	0.42	-0.48	0.03	-0.14	0.60	-0.16	-0.27
South Indian SST	0.31	0.49	0.11	-0.33	0.34	0.16	-0.22	0.55	0.20
Pacific NW GPH	0.37	-0.18	-0.35	0.06	-0.35	-0.59	0.07	0.48	-0.01
SW Atlantic GPH	0.29	-0.12	-0.61	-0.10	-0.04	0.58	0.37	-0.07	0.18
Gulf of Mexico GPH	0.41	0.23	0.03	0.37	0.27	-0.31	0.09	-0.45	0.51
NW Atlantic GPH	-0.27	0.50	-0.13	-0.43	-0.56	-0.15	0.02	-0.24	0.30
Jun 1st NRCS Forecast	-0.38	0.27	0.13	0.41	-0.01	0.04	0.65	0.40	0.13

	PC 1	*PC 2*	*PC 3*	*PC 4*	*PC 5*	*PC 6*	*PC 7*	*PC 8*
East Atlantic SST	0.37	-0.48	0.01	-0.41	0.25	0.15	-0.16	-0.59
Pacific NW SST	0.32	0.35	-0.44	-0.16	0.56	0.35	0.17	0.30
Central Pacific SST	0.36	0.15	-0.41	0.62	-0.05	-0.17	-0.43	-0.29
South Indian SST	0.35	-0.47	-0.21	0.27	-0.35	0.17	0.61	0.13
Pacific NW GPH	0.38	0.17	0.42	0.11	0.34	-0.60	0.39	-0.08
SW Atlantic GPH	0.29	0.04	0.65	0.32	0.04	0.56	-0.22	0.15
Gulf of Mexico GPH	0.46	-0.14	-0.01	-0.36	-0.27	-0.31	-0.40	0.56
NW Atlantic GPH	-0.27	-0.59	-0.03	0.31	0.56	-0.15	-0.17	0.34

D. Variable Importance Scores

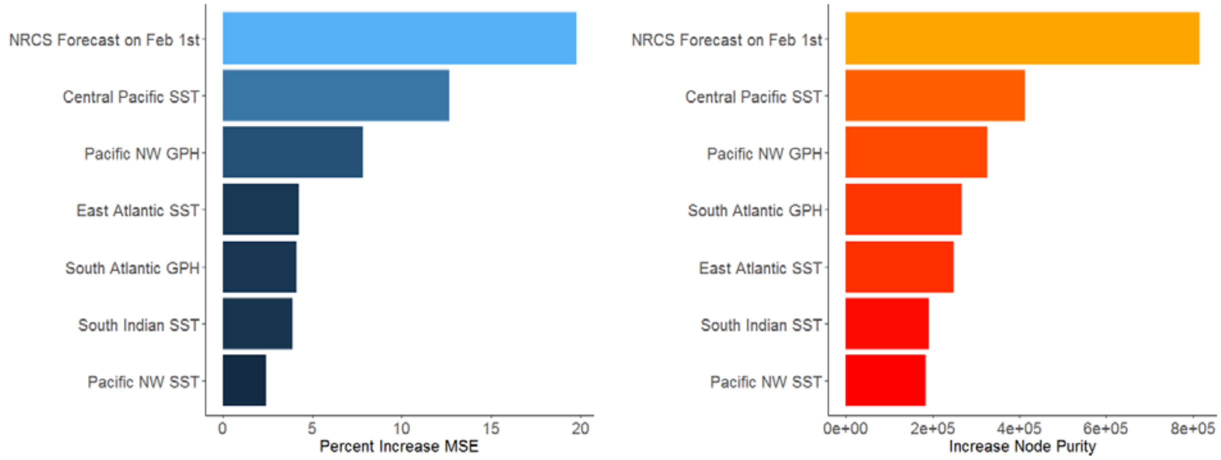


Figure 30. Variable importance scores, percent increase MSE (left) and increased node purity (right), for the seasonal inflow forecast on February 1st.

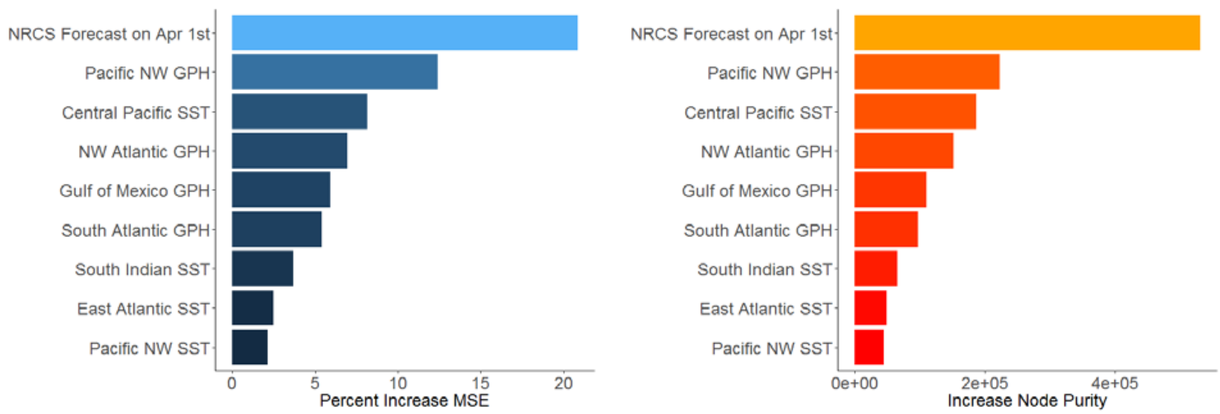


Figure 31. Variable importance scores, percent increase MSE (left) and increased node purity (right), for the seasonal inflow forecast on April 1st.

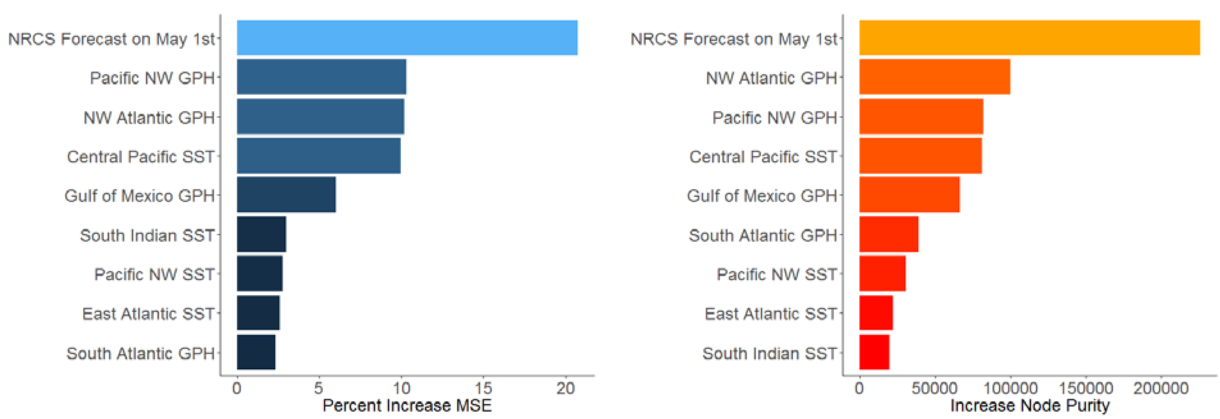


Figure 32. Variable importance scores, percent increase MSE (left) and increased node purity (right), for the seasonal inflow forecast on May 1st.