Enhancing Municipal Water System Planning and Operations Through Climate-Sensitive Demand Estimates

Ryan C. Johnson[1], Steven J. Burian [1,2], Carlos A. Oroza [3], James Halgren [1,4], Trevor Irons [3,5] Emily Baur [3] Danyal Aziz[2], Daniyal Hassan [3], Jiada Li[6], Tracie Kirkham[7], Jessie Stewart[7], Laura Briefer[7]

[1]Alabama Water Institute, University of Alabama, Tuscaloosa, Alabama, USA
[2]Civil, Construction and Environmental Engineering, University of Alabama, Tuscaloosa, Alabama, USA
[3]Civil and Environmental Engineering, University of Utah, Salt Lake City, Utah, USA
[4]Lynker Technologies, Leesburg, Virginia, USA
[5]Montana Technical University, Butte, Montana, USA
[6]Department of Civil and Environmental Engineering, College of Engineering, Colorado State University, Fort Collins, Colorado, USA
[7]Salt Lake City Department of Public Utilities, Salt Lake City, Utah, USA

Corresponding author: Ryan C. Johnson, rjohnson18@ua.edu

Abstract

High seasonality and interannual climate patterns drive the western U.S.'s water supply and demand variability. While the mean and variance of supply and demand drivers are changing with climate and urbanization, the metrics of reliability, resilience, and vulnerability (RRV) that guide urban water systems (UWS) seasonal management and operations tend to be built on assumptions of stationarity. In this research, we use documented performance of a real-world UWS as a testbed to investigate how RRV metrics – and therefore UWS planning and operations guidance – change in response to demands modeled with and without assumptions of stationarity during dry, average, and wet hydroclimate conditions. The results indicate an assumption of stationary demands leads to large differences between simulated and observed RRV metrics for all supply scenarios, and especially in supply-limiting conditions when the peak severity is 129% from the observed. The management implications of relying on stationary demands are severe: if seasonal operational decisions were made on these model results, managers might over-estimate seasonal out-of-district water requests by 50%. In contrast, when using non-stationary demands, one can expect system performance error reduction between 30% to 60% for average and dry climate conditions, respectively, and accurate RRV metrics. Our results further indicate that this UWS is more sensitive to percent changes in per-capita demand relative to percent changes in supply, but because the supply variability is so much greater (158% vs. demand range of 28%), we suggest further work to examine the combined (and coupled) influence of both factors in overall system performance.

Key Points:

- Machine learning water demand model driven by hydroclimate phenomena reduce overall error in seasonal water system assessment.
- Water demand uncertainty characterization enhances water system decision making confidence during supply limiting conditions.
- Water systems can exhibit significant performance sensitivity to seasonal demand projection accuracy.

1 Introduction

Climatic drivers of water supply and demand determine a snowpack-dominated municipal water system's ability to deliver clean and reliable water supplies; current climatic trends are negatively impacting supply in the western U.S. For example, in northern Utah, a seasonally disproportionate amount of precipitation occurs in the winter (over 70% on average), with projections estimating up to a 10% decrease by mid-century (Khatri & Strong, 2020). In the same region, complex Great Basin topography and larger global climate oscillations cause additional interannual climate variability, for instance, annual snow-water equivalent (SWE) accumulations with standard deviations of approximately 200 mm/yr (S.-Y. Wang et al., 2010; Smith et al., 2015). Both the high seasonality and the strong interannual climate variability influence the hydrologic system in myriad ways, including: snowpack accumulation, melt rate, spring runoff timing, and overall annual runoff volume. All of these factors impact the volume and timing of surface water available for domestic uses (Schewe et al., 2014; Scalzitti et al., 2016; Brooks et al., 2021). Looking to the future, climate change in the western US exhibits non-stationary characteristics with respect to historical records – with earlier spring snowmelt runoff and late season low-flow volumes (Muir et al., 2018).

Compounding surface water supply conditions require novel approaches when evaluating reliability, resilience, and vulnerability (RRV) metrics for urban water systems (UWS) (Goharian & Burian, 2018; Makropoulos et al., 2018; Nikolopoulos et al., 2019). A comprehensive UWS RRV analysis integrates streamflow forecasts, reservoir storage, demand projections, and other system performance drivers (i.e. groundwater withdrawal) into a systems framework(Goharian et al., 2016, 2017). This methodology is routinely supply-centric, characterizing system performance and operational decisions in response to the timing and duration of surface water peak runoff and lowflows (Finnessey et al., 2016). When anticipating hydrological drought, management searches for ways to extend supplies. This includes groundwater extraction, acquisition of out-of-district water, and the use of reservoir storage to supplement reduced surface water availability and manage system RRV (Wei & Gnauck, 2007; Finnessey et al., 2016). While supply availability is a critical determinant of system performance, such analyses often only recognize part of the system variability, leaving demand as a static, per-capita estimate independent of climate drivers. (Milly et al., 2008; Donkor et al., 2014; Zhao et al., 2018).

Because existing industry methods relying on historical mean per-capita demands do not capture the observed variability or external influences on water demand, system performance forecasts informing strategic and operational decisions likewise ignore that variability (Billings & Jones, 2011). For clarity in our discussion, we refer to variability as the fluctuation due to the random or chaotic behavior of the climate around a given mean or central tendency across seasons or in general through time (Grayman, 2005; Vose, 2008). Stationarity and non-stationarity refer to the stasis or trending drift, respectively, of that central tendency across many cycles of variation (Koutsoyiannis, 2006; Westra et al., 2014). As we discuss demand variability and non-stationarity in UWS RRV assessments, we note here that for our purposes, these may be referred to either interchangeably or always together. Noting these definitions, Johnson et al. (2022) found traditional per-capita demand forecasting methods, with embedded assumptions of stationarity and no variability from the historical mean, exhibit a significant increase in error and in comparison to machine learning (ML) models integrating driver-demand dynamics (e.g., air temperature, snowpack, surface water availability, precipitation, population density). Johnson et al. (2022) further demonstrate that such industry-standard static demand forecasting methods can overestimate municipal monthly water use by 90% and seasonal water use by 40% during hydrological drought. As a result of the demand forecasting error, downstream decision making process are confounded (Brown et al., 2012).

The recognition of non-stationarity within the UWS supply and demand drivers can lead to more comprehensive water resources planning and management analyses. Zhao et al. (2018) applied stochastic population projections, downscaled climate model supply outputs (Taylor et al., 2012), and spatially distributed hydrology to investigate water system resilience to long-term non-stationary total demand and supply processes. While the results indicate that future climate conditions impose greater uncertainty than urbanization-driven demand dynamics, per-capita demand estimates in their study retained assumptions of stationarity and were disconnected from exogenous drivers. In a set of drought scenario simulations to prepare a municipality for an anticipated drought in Northern Utah, Johnson et al. (2021) found that modeling demand responses to exogenous influences, rather than unchanging industry methods, can result in a 42% reduction in system vulnerability. In a separate but similar study, K. Wang and Davies (2018) used Calgary, Alberta's demand dynamics driven by exogenous influences to inform long-term water resource planning and management to potentially large changes to both seasonal and non-seasonal water system performance, identifying a need to enhance historical water management policies with new policies such as xeriscaping and greywater reuse to achieve water management goals. These studies demonstrate both random variation and climactic mean shifts impacts on water system performance; yet additional work is required to specifically separate the value of modeling variability and stationarity as independent influences on effectiveness of UWS RRV analysis. Specifically, the operational decision-making components driven by water system response to seasonal hydroclimate phenomena, exhibiting characteristics of non-stationarity, influences on supply and demand. Addressing this need, the current research adds a quantitative benchmark for the degree of improvement that may be expected from consideration of demand variability and non-stationarity in UWS performance planning.

As climate change progresses, urbanization continues, and new resource development becomes impractical, water resource planning and management must advance to maintain reliable UWS operations. This includes the compounding influences of non-stationarity within supply and demand processes that have the potential to introduce large model errors and increase system performance uncertainties, having far reaching effects that can misinform critical operational decisions. The non-stationarities we refer to are trends and/or a new normal of demands exceeding the bounds of historical observations in response to changes in land cover (development) and climate patterns (precipitation and temperature deviations in response to climate change). The principle uncertainties include *a priori* estimates of the difference between forecasted demand with respect to dynamic supply input, and observations. A related metric, error, refers to the *a posteriori* difference between the model and reality. Simultaneous non-stationarity within both supply and demand processes confounds the *a priori* estimation of RRV metrics, particularly the key metric of difference between forecast demand and supply. To address the gap of few studies recognizing variable and non-stationary demand processes in UWS assessments, this research examines the following questions:

- What level of error reduction in predicted UWS performance assessment may be achieved by using externally influenced seasonal demand forecasting instead of traditional stationary demand forecasting methods during episodes of average to extreme hydroclimate conditions?
- What additional information in terms of uncertainty quantification can seasonal demand forecasting provide for UWS performance forecasts?
- To which source of variability do UWS RRV metrics exhibit greater sensitivity: supply or demand?

This study addresses these questions by adopting a systems modeling framework to replicate the Salt Lake City, Utah UWS. We use the modeled UWS to investigate changes in RRV metrics in response to historical dry, average, and wet hydroclimate conditions and demand forecasts with and without embedded stationarity assumptions. By characterizing the impacts of variable and non-stationary climate on demand forecasting applied to UWS assessment, this research has the potential to advance the state of the practice towards the integration of non-stationary demand processes to enhance water resource planning and management.

2 Methods

This research considers the Salt Lake City Department of Public Utilities (SLCDPU) as a generalizable snowpack driven UWS. To represent supply inputs, we use dry, average, and wet hydroclimate conditions and the respective influence on surface water availability and demand to form the foundation of the RRV analysis. To forecast water demands, we use two methods; 1) an industry-standard static monthly per-capita methods based on the historical mean and 2) a dynamic forcing with exogenous hydroclimate and other variable inputs through the Climate Supply Development Water Demand Model (CSD-WDM) as demonstrated in Johnson et al. (2022). These supply and demand inputs drive the Salt Lake City Water Systems Model (SLC-WSM) to determine the volume, timing, and duration of out-of-district water requests, the indicator to gauge UWS RRV. The following subsections describe the study area, scenarios, water demand model, water systems model, and RRV methods.

2.1 Study Area

Dependence on winter snowpack, characteristics of high seasonality and interannual climate variability, and extensive data archives all make the SLCDPU a useful and generalizable mountainous western study area (Collins & Associates, 2019). This municipal water district currently serves approximately 350,000 people within the northern Utah's Salt Lake Valley, see Figure 1. The region's cold semi-arid (BSk) to cold desert climate (BWk) has four distinct seasons that influence water demands (Peel et al., 2007). Increases in temperature during spring and the quantity of precipitation strongly influence the beginning of the growing season; a hot, dry summer with temperatures exceeding 35.0 °C drives high evapotranspiration leading to high outdoor water use; and decreasing fall temperatures coupled with the return of precipitation end the growing season and the strong hydroclimate connection to outdoor municipal water use. From April to October, outdoor water use for landscaping irrigation can exceed 1000 mm per person, contributing to Utah being routinely ranked as the 2nd or 3rd highest per-capita water use state in the country (UDNR, 2010, 2014).

`Figures/StudyAreaV5.pdf`

Figure 1. Salt Lake City, Utah, depends on winter snowpack in adjacent Wasatch mountains to supply its four major surface water supplies, to fill the Dell reservoir storage system, and to replenish valley groundwater aquifers (Johnson et al., 2021)

The utility reports its monthly water treatment facility releases (in acre-feet) into the distribution system, including leakage and unaccounted system losses, to the Utah Division of Water Rights with near-continuous records since 1980 (UDWR, 2021). These data include residential, institutional, and commercial sectors covering the total volume of treated water delivered to the service area. From these records, monthly water use indicates significant year-to-year variability, with a minimum of 428 liters per-capita day (*lpcd*) in April, 2017; a maximum approximately six times greater of 2,635 *lpcd* in July, 1991; and an overall standard deviation of $\sigma = 13.0\text{x}10^6$ m^3 or +/-25% of the historical mean as illustrated in Table 1. Further demonstrating demand variability, monthly water use can vary by +/-45% of the respective months historical mean.

Table 1. The SLCDPU per-capita water use exhibits high April to October water use seasonality, with high variability observed from year-to-year.

Month	Minimum	Mean	Maximum	σ
Apr*	428	719	1,011	140
May*	609	1,105	522	246
Jun*	1,090	1,722	2,180	280
Jul*	1,465	2,074	2,635	276
Aug*	1,279	1,931	2,392	280
Sep*	1,030	1,442	1,839	208
Oct*	598	874	1,226	159
Season*	1,060	1,408	1,685	174
Season**	79.1	105.1	125.7	13.0

*units in lpcd
**units in m^3 (x10^6)

To meet these demands, the SLCDPU uses surface water, groundwater, and out-of-district water contracts. Surface water sources include City Creek (CC), Parley's Creek (PC), Big Cottonwood Creek (BCC), and Little Cottonwood Creek (LCC) that flow west from the adjacent Wasatch Mountains to on average supply 60% of the municipality's water. Sustainable groundwater withdrawal is up to 22.2x10^6 m^3 per year via 27 deep groundwater wells. Extraction from these wells tends to occur in summer months when surface water supplies cannot satisfy high outdoor water use. During periods of high water use and low surface water supplies, contracts with the Central Utah Project (CUP) permit SLCDPU to withdrawal up to 61.0x10^6 m^3 per year from the Deer Creek reservoir.

2.2 Simulation Scenarios

Annual hydroclimate variability is high in the Intermountain West, and to examine UWS response to extremes and averages, this study selects water years (October—September) that correlate with hydrological drought, average, and above average surface water supply conditions. These supply conditions demonstrate a direct connection to annual snowpack, the driving force behind groundwater recharge, peak runoff timing and volume, and annual water yield (Brooks et al., 2021). Connecting hydroclimate to UWS operations, we leverage the long-term snowfall record (1945-present) provided by the Alta Guard station at the headwaters of LCC (as a proxy for the region) and the percent of normal snowpack metric employed by the Natural Resources Conservation Service (NRCS) to identify the most recent dry (2015), average (2017), and wet (2008) hydroclimate years (NOAA, 2021). A Log-Pearson Type III analysis from these scenarios indicates the dry year demonstrates an exceedance probability greater than 200 years, and the wet year an exceedance probability of 50 years. The daily streamflow at the canyon mouths from these scenarios form the surface water supply inputs, as this is where in stream diversions supply water to treatment facilities.

The two water demand forecasting methods focus on April to October outdoor water use, featuring unchanging traditional industry methods embedded in stationarity and dynamic demands from the CSD-WDM algorithm capturing climate variabilities and non-stationarities. This research focuses on outdoor demand variability as indoor demands remain relatively consistent throughout the year, a function of showers, dishwashing, laundry, bathroom usage, etc., that do not substantially vary intra or interannually compared to outdoor use. In response to these observations, the indoor demands remain fixed throughout the simulation at the historical mean of 500 *lpcd*. For the industry methods, outdoor demands are climate independent and a function of each month's historical mean (Billings & Jones, 2011). Equation 1 displays the formula to calculate each month's demand

$$\overline{lpcd_m} = \frac{\sum_{i=1}^{30} lpcd_{m,i}}{30 \text{ yrs}} \tag{1}$$

where m is the month of interest and i represents a year in the training data. For the non-stationary dynamic demands, each scenario's hydroclimate and service area conditions are input into the CSD-WDM to estimate monthly mean per-capita demands. The Dynamic Water Demand Modeling section explains the inputs, architecture, and prediction error of this model. The observed per-capita demands that align with each hydroclimate scenario establish a baseline to investigate RRV errors and sensitivity to demands modeled with and without stationarity.

All monthly mean demand values require downscaling to match the SLC-WSM's daily time step. To downscale the demand data, this research develops an iterative python-based cubic spline interpolation program to create a continuous daily resolution demand time series. This approach reduces the residual difference between each month's mean value from interpolated daily demands and the original monthly-scale mean demand. This results in each month's mean daily demands equaling the observed or predicted mean monthly per-capita demand (*lpcd*) value.

2.3 Dynamic Water Demand Modeling

The CSD-WDM is a python-based (v3.8.5) ML optimization algorithm taking in exogenous service area variables to predict a municipality's mean monthly per-capita produced water demand (Johnson et al., 2022). These features include air temperature and precipitation data, conservation goals, surface water supplies, supply source snowfall, and service area (population, land-use, density) dynamics further discussed in the supplementary materials. The model uses a hierarchical framework, where each outdoor irrigation month (e.g., April to October) has a unique set of variable inputs to drive an OLS regression model built in the Statsmodels v0.13.0 package. During model calibration, the model evaluates feature correlation with the per-capita water use (*lpcd*), checks for feature colinearity, removes the lesser demand correlated colinear feature, and performs recursive feature elimination to identify key demand drivers to minimize model forecasting error. Related to error, the CSD-WDM communicates internal modeling error through the Statsmodels v0.13.1 python package by calculating the amount of variation in each demand driver coefficient and the corresponding standard error at a 95% confidence interval within the training data (Davidson et al., 2004; Seabold & Perktold, 2010; Montgomery et al., 2021; Johnson et al., 2022). Together, the framework enhances model interpretability, communicating both driver-demand interaction coefficients and corresponding internal model uncertainty in predictions. This research uses thirty years of data between 1980-2017 to calibrate the CSD-WDM, and three years (e.g., 2015 (dry), 2017 (average), and 2008 (wet) to form the validation scenarios. The calibration data omits the validation scenarios, which test model prediction for the case of hydroclimate conditions exceeding the bounds of stationarity (dry, wet). Model performance on the validation data is as follows; R^2 = 0.98, mean absolute error = 62.8 *lpcd*, and mean absolute percent error = 8.4%.

2.4 Water Systems Model

The SLC-WSM was designed to support SLCDPU decision-making regarding internal and external factors impacting reservoir performance (Goharian et al., 2016;

Goharian & Burian, 2018). The model operates within the GoldSim software environment, coupling submodels and linear programming to replicate the utility's interconnections between different water system components at a daily time step, including reservoir operations, water transfer infrastructure, water treatment systems, wells, withdrawal limitations, and more (Goldsim, 2013).

Water demand initiates the system operations. This includes six subregions (north, central, and southern Salt Lake City, Millcreek, Cottonwood Heights, and Holladay) indoor and outdoor per-capita demands (*lpcd*) and populations to determine daily demand requests (m^3/day). Each subregion uses the same daily per-capita demands, only varying depending on the demand scenario (stationary traditional, non-stationary CSD-WDM, or observed).

An essential component of SLC-WSM architecture is source selection and prioritization. Each subregion has a unique set of sources as a result of SLCDPU's gravity-centric distribution system. For example, the northern Salt Lake region has access to all sources due to its geographic location having the lowest elevation. Cottonwood Heights has the highest elevation and only access to LCC and BCC surface water supplies, a select number of wells, and selected out-of-district water. A critical aspect of system operations is source prioritization, which is as follows: surface water sources (CC, PC, LCC, and BCC), groundwater, and then out-of-district Deer Creek reservoir water. If surface water supplies cannot satisfy demands, then groundwater withdrawal initiates. If surface water supplies and groundwater withdrawal (e.g., limited by the number of wells, extraction rates, and annual withdrawal limitations) cannot satisfy demands, then out-of-district water is requested. This order of prioritization minimizes costs attributed to pumping, treatment, and transfers.

2.5 System Performance Assessment

This system performance assessment determines municipal water system performance by using the simulation time series of a specific parameter or indicator to represent the system's status (e.g., out-of-district supply requests)

$$X_t; \quad t = 1, 2, ..., T \tag{2}$$

where X_t is the system performance at timestep t; and T is the time period of the analysis (e.g., 213 days from April 1^{st} to October 30^{th}). Using this indicator, the calculation of the system performance index (*SPI*) is

$$SPI = f(X_t); \quad t = 1, 2, ..., T \tag{3}$$

Equation 3 forms the foundation of the system performance assessment and we use daily and monthly temporal resolutions (t) for evaluation. This analysis investigates both temporal resolutions to examine the peak intensities and timing (daily), along with the larger system informing water volumes (monthly). Integrating both temporal scales supports a comprehensive seasonal water systems assessment.

The *SPI* can be made more meaningful by connecting the index with an indicator state or threshold at each time step ($t(Z_t)$). This establishes a measure of comparison to define and differentiate satisfactory (S) and unsatisfactory (U) system states. By integrating these measures, the calculation of the *SPI* is

$$SPI = f(Z_t) \quad t = 1, 2, ..., T \quad and \quad \begin{cases} Z_t = 1 & X_t \in S \\ Z_t = 0 & X_t \in U \end{cases} \tag{4}$$

For this assessment, the *SPI* calculation uses the volume of out-of-district Deer Creek reservoir water (X_t) above the simulated historical mean amount (Z_t). Water from this source comes at an increased operational cost, supporting its usage as an indicator to investigate UWS performance. The historical mean Deer Creek use is a function of the observed water demand, supply, and systems operations at a daily time step using simulations spanning from 2000-2020. Thus, this threshold defines unsatisfactory (e.g., >historical mean) and satisfactory (e.g., <historical mean) Deer Creek reservoir requests.

Using the Deer Creek SPI, this study modifies the RRV metrics originally presented in Goharian and Burian (2018). The reliability metric describes the relative frequency of the system operating in a satisfactory state compared to the total simulation length.

$$\alpha = \frac{\sum_{t=1}^{T} Z_t}{T} = 1 - \left(\frac{n_f}{T}\right) \tag{5}$$

where α is the reliability estimate, and n_f is the number of unsatisfactory days out of the period of interest (T). The calculation of reliability is at a respective temporal resolution for each simulation. Values closer to 1 indicate high levels of reliability, and values close to 0 indicate low levels.

Resilience measures the average speed that the system can rebound from an unsatisfactory to a satisfactory state

$$W_t = \begin{cases} 1 \text{ if } X_t \in U \text{ and } X_{t+1} \in S \\ 0, \text{ otherwise} \end{cases} \tag{6}$$

where W_t is an indicator capturing the transition from unsatisfactory to satisfactory states. Using this indicator, the calculation of resilience (RS) is

$$RS = \frac{\sum_{t=1}^{T} W_t}{T - \sum_{t=1}^{T} Z_t} \tag{7}$$

Using this formula, resilience accounts for the number of rebounds (e.g., transition from unsatisfactory to satisfactory states) as a percentage of the total number of unsatisfactory states. From this metric, the inverse of resilience ($1/RS$) is the duration that the system remains in an unsatisfactory state and is the preference for expressing resilience in a water system (Asefa et al., 2014).

Since reliability and resilience cannot fully describe UWS behavior, this research uses vulnerability to capture the severity of unsatisfactory conditions and corresponding system response at both a daily and monthly resolution. Using this framework, exposure and severity further define vulnerability

$$\text{Vulnerability} = f(\text{exposure}, \text{severity}) \tag{8}$$

Exposure is the occurrence of unsatisfactory conditions in Deer Creek reservoir water use because of limited surface water supplies from the respective hydroclimate scenario.

$$WRI_S = 1 - \frac{WR_S}{WR_H} \tag{9}$$

where the out-of-district Deer Creek water requests index to snowpack (WRI_S) is the ratio of water requests due to snowpack (WR_S) and historical water requests (WR_H). The WRI_S varies from 0 to 1, with values closer to 1 representing increased vulnerability and 0 displaying no change from historical conditions. We use the 2000-2020 simulation period and a unique WR_S for each hydroclimate and demand scenario to calculate the WR_H.

Severity characterizes the magnitude of impact that unsatisfactory conditions have on the system. The calculation of this metric is as follows

$$S = \sum s_t e_t \quad X_t \in U \tag{10}$$

where s_t quantifies the severity of unsatisfactory conditions at time t, and e_t is the occurrence probability of X_t (in the form of s_t), as the most severe result from a set of unsatisfactory states. Using both exposure and severity, this study calculates average system vulnerability by

$$\text{Vulnerability} = WRI_S\beta_{WR} + S\beta_S \tag{11}$$

where the application of β_{WR} and β_S weights is because of each variable's different degree of *subjective* importance. Goharian et al. (2016) analyzed the perceived relative importance of these factors based on judgment, stakeholder surveys, management, and sensitivity analysis to determine that equal weighting is appropriate for this system. Thus, we assign equal weights (0.5) to exposure and severity metrics (0-1) to determine vulnerability.

In addition to average system vulnerability, this study calculates peak severity and reports it at daily and monthly time steps. Rather than taking the average severity throughout the simulation using Equation 10, the maximum s_t during each simulation determines the peak severity of unsatisfactory conditions.

This study uses five categories to illustrate different levels of vulnerability and peak severity for the daily and monthly time scales. With values ranging from (0,1), the vulnerability and peak severity analyses leverage the Jenks optimization technique to identify the natural metric breaks within the historical simulations (2000-2020) (Jenks, 1967). This methodology minimizes each class's average deviation from the class mean while maximizing each class's deviation from the means of other classes. This creates five categories ranging from Category 1 (Low) with the lowest vulnerability/peak severity to Category 5 (Extreme) with the greatest. Category 5's Extreme rating is for system performance exceeding the bounds of stationarity, e.g., the historical record. Table 2 displays the vulnerability levels and their ranges.

Table 2. Jenk's classification of system vulnerability and maximum severity leverages historical values to determine categorical means and their distributions. A classification of extreme indicates a level unseen in the historical record.

Metric	Scale	Low	Medium	High	Very High	Extreme
Vulnerability	Day	0-0.08	0.08-0.31	0.31-0.46	0.46-0.59	0.59-
	Month	0-0.20	0.20-0.50	0.50-0.75	0.75-1.0	1.0-
Severity	Day	0-0.12	0.12-0.28	0.28-0.46	0.46-0.66	0.66-
	Month	0-0.08	0.08-0.38	0.38-0.67	0.67-1.0	1.0-

2.6 Water System Sensitivity

The range of daily and monthly RRV metric values form the baseline to investigate UWS sensitivity to supply and demand inputs. System sensitivity is a function of the maximum metric difference among each forecast category and the supply or demand variability ratio.

$$S_m = \frac{m_{max} - m_{min}}{R_y} \tag{12}$$

where S is the system sensitivity for metric m, m_{max} is the largest value and m_{min} is the least, and R_y is the range in the seasonal supply or demand as a ratio of the historical average. For example, using Equation 12 to determine the system vulnerability sensitivity to supply, within each demand type (observed, stationary traditional, non-stationary CSD-WDM), we calculate the range in system vulnerability for wet, average, and dry climate conditions and then divide by the range in streamflow (e.g., above average ratio to the historical average (2.11) minus the below average ratio to the historical average (0.53)). The maximum S_m from observed, stationary traditional, and non-stationary CSD-WDM demands is the system vulnerability to supply. Tables **??** and 3 provide the foundation to calculate the RRV metrics' sensitivity to supply and demand.

In addition to these calculations, we present the total volume of out-of-district water requests in response to percent differences (10%) in average supply and demand. This supplementary system sensitivity analysis responses to the greater variability in historical demands than present in the three hydroclimate driven testing scenarios. We vary the supply by +/-50% of average and demand by +/-40% of average to represent the observed historical variability. While the range of observed supply exhibits a range exceeding 150%, the lower bounds are of greater significance water resources management and we capture the upper bound within the aforementioned wet hydroclimate scenario.

2.7 Model Error and Uncertainty

Quantifying internal model error and prediction uncertainty is a critical component of operational water resources management as it establishes a foundation for informed decision making (Brown et al., 2012). In this research, error refers to the *a posteriori* difference between each simulations water system performance (with and without assumptions of stationarity) to the observed. We calculate total system error as the percent difference in out-of-district Deer Creek reservoir requests from the observed for each hydroclimate and demand simulation. This research acknowledges other sources of error are present (i.e., differences in operations, system interactions, service and maintenance, etc.), but focus on the modeling errors related to demand estimation in this analysis.

In this research, we define prediction uncertainties as the *a priori* estimates in the range of predictions. While the stationary/traditional demands do not support the characterization of prediction uncertainty, the non-stationary CSD-WDM leverages the Statsmodels v0.13.1 python package to calculate the amount of variation in each demand driver coefficient and the corresponding standard error at a 95% confidence interval (Davidson et al., 2004; Seabold & Perktold, 2010; Montgomery et al., 2021; Johnson et al., 2022). In addition to the predicted values, this allows for the estimation of high and low bounds for total municipal demand as a function of internal demand modeling errors. We determine the range in water system performance (volume of out-of-district Deer Creek reservoir water requests) uncertainty by running the low, predicted, and high non-stationary dynamic demand simulations for each hydroclimate scenario. The range in system performance (in response to demand uncertainty) characterizes the prediction uncertainty in the RRV analysis. We exhibit these values for each hydroclimate simulation, with the upper and lower bounds of the non-stationary dynamic demands complementing the predicted at a 95% confidence interval. This novel approach to water system evaluation enhances system performance prediction confidence–especially compared to the deterministic results produced using the stationary traditional methods.

3 Results

The results section first begins by comparing water system performance errors between the stationary and non-stationary demand estimates with the simulated observed for each hydroclimate scenario. We use the simulated out-of-district Deer Creek reservoir requests to calculate the RRV metrics at both a daily and monthly temporal resolution for all simulations (including the observed) and classify the vulnerability level with the Jenks classification algorithm, establishing a baseline for comparison. In this analysis, we determine stationary and non-stationary demand simulation RRV percent errors from the observed to further exemplify the methodological differences. The second part of this results section investigates water system sensitivity to the variability observed in supply and demand. Using the average hydroclimate condition and non-stationary dynamic demands, 99 simulations varying supply and demand percentages from the mean support the evaluation of water system sensitivity to these drivers.

3.1 Reducing Water System Performance Error and Uncertainty

This research calculates the water system RRV and peak severity for all supply and demand scenarios. These values are in Table 3, with Figure 2 illustrating the percent differences from the observed and the range of uncertainty (CSD-WDM simulations). While all temporal resolutions provide essential water system performance information, the tables and figures presented focus on the monthly resolution for its significance in operational water system performance. We discuss the daily resolution results but include these in the supplementary materials (Table **??**). Similarly, we present the all hydroclimate conditions simulation results for the stationary and non-stationary demand estimates in this section, but place the average and wet hydroclimate conditions figures in the supplementary materials (Figures **??**-**??**). Figure 3 and 4 illustrate water system performance during the dry hydroclimate conditions for the stationary and non-stationary demand estimates, respectively. These figures present the range of prediction uncertainty within the CSD-WDM results, a missing component of the stationary traditional demand forecasting method. The final component of this section highlights the categorical and percentage difference from the observed, indicating each forecasts' system performance error and how it varies depending on climate.

There is little difference in daily RRV values among demand models in a high snowfall year (wet). The best measure of system performance differences in this scenario is categorizing vulnerability and peak severity. Classification of both metrics is Low for each demand model, even though there is a large percentage difference in vulnerability (400%). Evaluating the system RRV at the monthly scale indicates a greater difference in reliability (-14%) and vulnerability (250%) from the observed when using stationary traditional demand forecasts. This increase is in response to four days in June with out-of-district requests ranging between 3,300-4,200 m^3/d above average and results in one category higher in vulnerability than the observed (Medium vs. Low). By reducing the demand forecasting error (non-stationary CSD-WDM), the wet climate scenario's system performance mirrors the observed RRV. While there is a 50% increase in vulnerability, this value is only 0.02 greater than the observed and remains in the same Jenk's category, a negligible difference. Furthermore, these results indicate high forecasting confidence with small uncertainties. For example, the uncertainty in daily UWS reliability, resilience, and peak severity are 0, and internal model error demonstrates a small range of vulnerability (e.g., 0-0.12) encompassing the observed (0.01).

In an average snowpack year, the SLCDPU's RRV exhibits a greater change in performance depending on the demand forecasting error. At a daily resolution, the stationary traditional demand forecast exhibits 22% less reliability and 32%, 26%, 15% greater resilience, vulnerability, and peak severity than the observed, respectively. The classification of vulnerability and peak severity are Very High, one level greater than the observed. By integrating more accurate demand estimates (i.e. non-stationary CSD-

WDM), UWS performance reflects the observed conditions for all metrics but resilience, where the simulations suggest a 44% increase. At a monthly resolution, the impact of demand forecasting error on system performance becomes more significant. The stationary traditional demands suggest a 25% reduction in reliability and a 53% and 33% increase in vulnerability and peak severity from the observed. This categorizes the system as entering the greatest observed vulnerability and peak severity state in the historical record (Very High), one level greater than the observed. The non-stationary CSD-WDM simulation closely predicts all RRV metrics in the average climate scenario, with the range prediction uncertainties encompassing the observed system states.

Table 3. By relying on stationary demands, the monthly water system RRV metrics demonstrate the incorrect classification of extreme vulnerability and peak severity (along with no prediction uncertainty characterization) during dry climate conditions which could incorrectly trigger unnecessarily aggressive operational and management actions. By using the non-stationary (CSD-WDM) demand forecast, the predictions exhibit reduced forecast error (value in parenthesis) and characterize the range of uncertainty in response to internal model errors.

Metric	Climate Scenario (snowpack)	Observed Demands	Stationary Demands	Non-Stationary Demands	Non-Stationary Uncertainty (Lo/Hi)
Reliability	Dry	0.29	0.0 (-100%)	0.29 (0%)	0.29
	Average	0.57	0.43 (-25%)	0.57 (0%)	0.43-0.71
	Wet	1.0	0.86 (-14%)	1.0 (0%)	0.76-1.0
Resilience*	Dry	6	8 (-33%)	6 (0%)	6
	Average	2	2 (0%)	2 (0%)	2-5
	Wet	1	1 (0%)	1 (0%)	1
Vulnerability	Dry	0.49	0.68 (39%)	0.48 (-2%)	0.36 -.59
	Average	0.34	0.52 (53%)	0.35 (3%)	0.19-0.51
	Wet	0.04	0.14 (250%)	0.06 (50%)	0.01-0.17
Peak Severity	Dry	0.55	1.28 (133%)	0.66 (20%)	0.33-1.0
	Average	0.58	0.77 (33%)	0.62 (7%)	0.28-1.0
	Wet	0.0	0.01 (INF)	0.0 (0%)	0.0
Vulnerability Level	Dry	Very High	Extreme	Very High	High-Very High
	Average	High	Very High	High	Medium-Very High
	Wet	Low	Medium	Low	Low-Medium
Peak Severity Level	Dry	High	Extreme	High	Medium-Very High
	Average	High	Very High	High	Medium-Extreme
	Wet	Low	Low	Low	Low

*units in months

Figures/MonthRRV.png

Figure 2. Monthly RRV for observed (OBSD), stationary traditional (TD), and non-stationary dynamic (CSD-WDM) water demand simulations. The non-stationary CSD-WDM simulations mirror the observed results and communicate prediction uncertainty estimates to a 95% confidence interval, while the traditional methods indicate reduced reliability, increased vulnerability, and no communication of error.

The most significant differences in UWS performance appear in the dry hydro-climate scenario where an approximate 50% decrease in surface water supply occurs in the 200 year drought event. The results of this scenario are also the most critical to decision-making. When using stationary traditional demands, there is a 15% reduction in daily reliability and a 42% increase in resilience compared to the observed. The differences are more severe for vulnerability and peak severity, 39% and 129%, respectfully. The peak severity value of 1.19 is significant as it exceeds the bounds of stationarity, indicating the UWS is entering a state exceeding all of those in the historical record. The vulnerability and peak severity categories also capture this with the extreme rating, two levels greater than the observed. At a monthly resolution, these demands result in the vulnerability and peak severity being 39% and 133% greater than the observed, and again the classification of Extreme. By using the non-stationary CSD-WDM forecasted demands, the UWS RRV resembles the observed except for daily resilience (+50%), daily peak severity (+21%), and monthly peak severity (+20%). Even with the mean prediction value exhibiting little error, the model's 95% prediction confidence interval completely encompasses the observed. Overall, the reduced forecasting error correctly classifies the system's vulnerability and peak severity at daily and monthly resolutions.

Figures/2021Dry Supply and CSD_WDM Demands_Unc.pdf

Figure 4. The forecasted SLCDPU performance during dry supply conditions for observed and non-stationary CSD-WDM demands. The figure illustrates the similarities between the two with respect to the magnitude and timing of demands and Deer Creek water request and the respective seasonal hydrographs. The CSD-WDM-generated prediction confidence intervals provide a sense for the range of potential prediction uncertainty.

In all climate scenarios, the results indicate demand forecasting error decreases directly translate into more representative daily and monthly system RRV estimates. Comparing the demand forecasting methods, these error decreases are significant with the mean percent reduction in error for the non-stationary demand forecasts being 31% and 59% for average and dry climate conditions, respectively.

3.2 Water System Supply and Demand Sensitivity

Comparing the three hydroclimate and demand simulations, and using the historical mean as a baseline, surface water supply exhibits a greater percentage variability than demand. For example, the dry climate scenario yields 53% of normal seasonal streamflow yield while the wet conditions delivered 211%, producing a 158% range in seasonal supply yield, see Supplementary materials Table **??**. The greatest range in seasonal demands varies by 28%, observed in the dry climate scenario where demands were 131% (Traditional) and 103% (Observed) of the seasonal historical mean.

Applying Equation 12 to the values in Tables 3 provides a measure to gauge system sensitivity to supply and demand variability. Table 4 displays the supply and demand system sensitivity values for each RRV metric at daily and monthly temporal resolutions. In both temporal resolutions, the SLCDPU's RRV demonstrates two to three times greater sensitivity to demand than supply. Although demand demonstrates a greater percentage wise influence on water system performance, the greater range in supply influences system performance to a greater extent with these ranges.

While the wet and dry hydroclimate scenarios capture the variability in supply availability, the municipality's historically observed demand variability differs to a much greater percent than observed in these simulations. Over the past 40 years, the municipality's per-capita water use exhibits a monthly range of demand by +/- 45% from the historical mean.

Table 4. The observed range in daily and monthly SLCDPU water system metrics as a function of supply and demand variability. The larger demand values indicate the water system is more sensitive to percent changes in water demand than supply.

	Demand*	Supply*	Demand**	Supply**
Reliability	0.77	0.37	1.0	0.54
Resilience	98	35	7	4
Vulnerability	0.63	0.36	0.70	0.34
Peak Severity	2.32	0.75	2.55	0.80

* Daily.

**Monthly.

To characterize the system performance influences attributed to these ranges, we perform a sensitivity assessment varying supply and demands in 10% intervals from the mean to the historically observed variability (+/-40% for demand, +/-50% for supply). Figure 5 illustrates the water system performance (as a function of the total seasonal volume of Deer Creek reservoir requests) sensitivity difference between supply and demand. For example, considering the influence of demand variability on water system performance, with the streamflow scenario (50% seasonal reduction) constant and evaluating the full +/-40% range in demands, we observe greater than an $80.0x10^6$ m^3 range in the volume of out-of-district water use. Considering the influence of supply variability on water system performance (+/-50%) and holding demands constant (+40%), the results demonstrate a range of under $40.0x10^6$ m^3 of out-of-district water requests. Similar to the analysis of hydroclimate influenced water system sensitivity, even with the lesser percentage range in variability, municipal demand demonstrated a two- to three-fold greater influence on water system performance than supply. These results illustrate the need to complement supply-focused water system assessments with representative demand estimates.

Figures/DC_GW_Heatmap.png

Figure 5. Using Deer Creek Reservoir water request as a system performance indicator, the SLCDPU water system demonstrates greater sensitivity to percent changes in demand than streamflow. This system response suggests that advances in demand forecasting error reduction will greatly reduce errors in seasonal UWS performance assessments.

4 Discussion

In this section we discuss the management and operational impacts that assumptions of stationarity in water demand impose on a seasonal water systems assessment. First, we expand on demand forecasting errors and the compounding impacts they have on water system performance. This sections connects operational decision making with model simulation results, discussing how the financial and source acquisition actions needed to mitigate supply limitations differ between simulations of demands modeled with and without assumptions of stationarity. The second section discusses the water system performance sensitivity to supply and demand, describing how both components substantially influence water system performance and highlights future research needs to respond to these findings. The final section discusses the impacts of variability and non-stationarity in the water system, providing a high-level overview of how future water resources research (likely involving climate change) can benefit from the realization and modeling of non-stationary processes in these systems to improve future prediction and increase overall resilience.

4.1 Demand Forecasting Error and Water Resource Management

Urban water system simulations need to provide management with criteria for evaluating system performance that can inform operational decisions. An evaluation of the simulation results (see Table 5) and RRV assessment with the observed demands indicates that assumptions of demand stationarity profoundly impact system performance forecasting efficacy. For example, the vulnerability and peak severity levels match each supply scenario when using the lower error CSD-WDM demand forecast, capable of capturing demand variabilities and non-stationarities. In contrast, assumptions of demand stationarity suggest increasing differences from the observed in vulnerability and peak severity levels, especially in the dry climate scenario with the Extreme system state. From the observed system performance, relying on stationary traditional demand forecasting methods suggests a daily average 20% reduction in daily reliability, a 37% increase in resilience, a 33% increase in vulnerability, and a 72% increase in peak severity during average and dry climate conditions.

From a management perspective, the actions needed to mitigate supply limitations are different among demand forecasts with and without assumptions of embedded stationarities. During average snowpack conditions, the stationary traditional demands simulation suggests a 72% increase in seasonal out-of-district water requests and Very High vulnerability classifications. To management, this would trigger alarm and likely initiate a supply limited contingency plan such as water rationing and money spent on conservation awareness (Inman & Jeffrey, 2006; Liu et al., 2015). While the non-stationary CSD-WDM simulation indicates high levels of system vulnerability, it suggests an average seasonal volume of out-of-district requests within the bounds of the 95% confidence interval capturing the observed. This would lead management to closely monitor physical system performance but not require critical operational decisions prior to increased levels of municipal indoor-outdoor use beginning in April or May.

As surface water supply becomes limited, management actions are likely necessary regardless of demand forecasting error. The difference in action requirements (e.g. requested vs. mandatory water use reductions) is driven by the severity of forecasted system performance. In a region dominated by prior appropriations, reductions in total water use are challenging. As an example, the stationary traditional demand simulation suggests the SLCDPU water system entering a non-stationary vulnerability state during dry conditions. This results in a suggested 200% increase in out-of-district water requests, which could prompt aggressive and mandatory Stage II management actions for water rationing (Salt Lake City Department of Public Utilities, 2021). Management solutions require an aggressive conservation plan approaching a 35% reduction in combined indoor/outdoor water use to achieve average historical system performance. Aggressive demand-sided management activities supporting water conservation awareness and mandatory irrigation schedules may achieve this significant reduction (Inman & Jeffrey, 2006; Liu et al., 2015). However, a significant short-term reduction of this magnitude may lead to severe economic consequences for end-users (DeOreo, 2006).

Table 5. With assumptions of demand stationarity, the SLC-WSM overestimates the volume of out-of-district water requests and does not accurately capture the timing of these requests during dry climate conditions. These results further demonstrate the advantages of modeling for demand variabilities and non-stationarities and the need to characterize internal model error and resulting prediction uncertainties. The non-stationary CSD-WDM forecasts indicate a range of metrics values (95% confidence interval values in parenthesis) to communicate uncertainty surrounding the prediction.

Metric	Observed Demands	Stationary (Traditional) Demand	Non-Stationary (CSD-WDM) Demand
Peak Daily System Demand*	57	73	56 (51-63)
Peak Deer Creek Request*	27	40	27 (19-36)
Peak Demand Date	Aug-26	Aug-03	Sep 11 (Sep 6-12)
Deer Creek Request Duration**	111	127	111 (51-123)
Peak Monthly System Demand*	1,750	2,250	1,690 (430-1,930)
Peak Monthly Deer Creek Request*	750	1,130	720 (510-972)
Peak Deer Creek Request Month	Sep	Jul	Sep (Aug-Sep)
Deer Creek Request Duration***	3	4	3 (1-4)
Seasonal Demand*	8,300	10,600	8,300 (7,200-9,400)
Seasonal Deer Creek Request*	2,080	3,040	2,030 (1,200-2,900)
Seasonal Streamflow Supply*		5,000	
Percent of Average Seasonal Streamflow Supply		-47%	

* in x10^4 m^3/d

** units in days

*** units in months

Still examining the dry hydroclimate scenario, the non-stationary CSD-WDM simulations capture the observed voluntary actions to 'survive the drought' and naturally reduce the magnitude of peak system demands by nearly 25%. However, the simulation suggests a 75% increase in out-of-district requests, which would require a 13% mean reduction in outdoor water use to maintain historical system performance. While this number is nearly three times less than the stationary traditional demand simulation (35%), it likely accounts for modified irrigation schedules and the implementation of conservation strategies, making achieving further reductions difficult due to demand hardening (Howe & Goemans, 2007). A key metric to guide management is the seasonal timing and volume of peak out-of-district requests. As a result of an extended period of indoor-outdoor water use, the model suggests high irrigation rates through September, which leads to above average out-of-district requests. While the observed and non-stationary CSD-WDM climate-demand scenarios present significant operational challenges, an approach recognizing demand responses to external factors provides a more comprehensive RRV assessment to guide operational decisions.

For the SLCDPU and other utilities in the western US, a utility is one of many supply requests in large reservoir systems. This emphasizes the seasonal forecasting error of the timing and volume of these requests where reservoir storage-release operations, storage agreements with other utilities, and minimum release requirements for aquatic ecosystems challenge reservoir operations in supply limiting conditions. Again using the dry climate conditions as an example, the stationary traditional per-capita demand forecasting scenario suggests 127 days of unsatisfactory conditions compared to the observed and non-stationary CSD-WDM demand forecasts of 111 days. A similar trend extends to the daily, monthly, and seasonal peak volumes where the stationary traditional demand modeling methods overestimate out-of-district requests by $\sim 50\%$. Table 5 further illustrates the differences in the physical timing, duration, and magnitude of out-of-district water requests.

The inferior system performance and high error resulting from the stationary demand forecasts does not capture the demand response to climate dynamics that influence the magnitude and intensity of April to October indoor-outdoor water use, especially during supply limiting conditions. Thus, responding to the first research question, integrating non-stationarity driven demand estimates has significant impacts on total water system performance, where we demonstrate a 31% and 59% reduction in system forecasting error for average and dry climate conditions, respectively. Responding to the third research question, integrating demand uncertainty measures provide system operations with increased confidence in seasonal system operations, up to a 95% confidence level in these cases.

4.2 Water System Performance Sensitivity to Supply and Demand

The results indicate that this snowpack driven UWS's RRV and peak severity are more sensitive to changes in demand than supply. However, the hydroclimate driven simulations present much greater variability in supply (158%) than demand (28%). While the system may be more sensitive to changes in demands, for these scenarios the greater range in supply availability has a stronger influence on overall system performance. This aligns with the long-term reservoir operations analysis performed by Zhao et al. (2018), demonstrating that while water demand has a substantial influence on reliability, there is greater uncertainty in reliability attributed to supply availability than demand variability. While our analysis did not focus on streamflow forecasting uncertainty, the results do indicate that reductions in demand forecasting error and corresponding prediction uncertainty will enhance confidence in water system performance forecasts.

Recognizing the three hydroclimate scenario's demand variability did not represent the full range historical demands, this study evaluates system performance in response the historical range of supply and demand variability to serve as a preliminary system sensitivity analysis. The municipality's historical demand indicated +/- 40% deviations from the mean, yielding an approximate 80% range that is much greater than that observed in response to hydroclimate variability. Running the systems analysis on the greater range in demand produced similar system response to the smaller hydroclimate driven demands, a two- to three-fold greater influence on water system performance compared to supply availability. The difference is that the water system analysis suggested an overall greater influence on system performance from demands compared to supply. Thus, responding to the second research question, these simulation suggest water system performance exhibits greater sensitivity to demand compared to supply. While these results indicate a significant water system performance response to possible errors and uncertainty in demand prediction, there is a need for further research characterizing system response to both supply and demand forecasting accuracy and error influences on system performance. For example, this can include a more comprehensive sensitivity analysis varying supply and demands by smaller percentages and evaluating over additional hydroclimate conditions. Characterizing these water system performance responses would identify supply and demand forecasting error and uncertainty goals to enhance water resources management and operations.

4.3 Non-stationarity in the Water System

This analysis indicates the assumption of stationarity introduces error when evaluating UWS performance. This is apparent in supply, where average hydroclimate conditions (2017) produced a seasonal surface water yield of 62% of the historical average. While this scenario is exemplary of an average snowfall year, the average snowpack does not correlate to an average April to October surface water yield. This is the result of complex hydrological processes governing Wasatch streamflow yields (Brooks et al., 2021). However, this change in snowpack-water yield aligns with Muir et al. (2018) anticipating a reduction in summer flows for the same winter precipitation amounts as climate change progresses. To assume an average surface water yield from April to October, an above average snowpack will likely be necessary.

With respect to demand non-stationarity, the results indicate that even with reductions in per-capita demands, total system demands will continue to increase due to population growth. For example, even with significant reductions in per-capita water use ($\sim$ 25%) from the dry climate scenario, the results indicate an increase in total water demand (+3%, observed). The total observed system demands are 6% greater than the historical average during an average snowpack and average per-capita demands. As populations continue to increase, total demands will exceed the bounds of the stationarity regardless of hydroclimate conditions (Milly et al., 2008; Zhao et al., 2018).

In this analysis, the observed and non-stationary CSD-WDM demand simulations never exceed the bounds of historical RRV with the mean prediction. However, internal model errors communicating prediction uncertainty connect water system performance during supply limiting conditions to an Extreme vulnerability and peak severity state. This characterization of demand prediction uncertainties (to a 95% confidence interval) is important and novel to water system operations, communicating critical information to water system managers relevant to maintaining water system performance as surface supplies become limiting. While no immediate action is necessary, managers are explicitly informed of possible system compromising conditions.

By integrating exogenous drivers into demand models to reduce prediction error the resulting forecast reduces water system RRV errors and characterizes the associated uncertainties. This will improve water resource management, especially as climate change progresses and supply availability continues to depart from the range of historical observations.

5 Conclusion

This research is part of an ongoing and comprehensive research program to address existing knowledge gaps in municipal water demand forecasting and systems modeling literature. Research activities described here included a seamless coupling between predictions from a non-stationary demand forecasting model (CSD-WDM) and a dynamic systems models (SLC-WSM). We have prepared a comprehensive RRV assessment utilizing Jenk's classification to segregate dry, average, and wet climate scenarios to allow comparison of water system performance among simulations of hydroclimate phenomena.

Using Salt Lake City, Utah as a case study, this research uses recent dry, average, and wet hydroclimate regimes and their respective observed demands to determine the implications of considering (or not considering) demand variability and nonstationarity when predicting UWS performance. This research takes advantage of novel non-stationary demand forecasting methods (e.g., CSD-WDM) to demonstrate significant error reduction and uncertainty characterization of RRV for a snowpack driven UWS, as compared to the same analysis under traditional demand forecast assumptions of stationarity. The results indicate that these demand forecasting methods introduce high errors in UWS performance estimates for all supply scenarios, with maximum errors of -15%, 42% 39%, and 129% for out-of-district (Deer Creek Reservoir) water request RRV and peak severity, respectively. These system differences extend to the timing and magnitude of peak severity and the duration of unsatisfactory conditions.

By integrating novel ML demand models, this research demonstrates that applying advanced demand forecasting methods which capture hydroclimate-influenced service area demand can enhance UWS performance assessment through error reductions in all climate scenarios. Building on the UWS performance improvements, a key contribution to water systems modeling is the realization that integrating demand prediction uncertainties supports the characterization of downstream water system performance. This research demonstrated that in many cases (e.g., supply limiting conditions) reductions in demand forecasting error and integrating uncertainty estimates profoundly impacts overall simulation confidence, supporting enhanced decision making. The need to advance demand forecasting performance and characterizing underlying uncertainties were made more profound by this UWS exhibiting greater sensitivity to demand vs. surface water supply variability. Complementing this finding, the results indicate that this UWS is more sensitive to percent changes in per-capita demand relative to percent changes in supply, but because the supply variability is so much greater (158% vs. demand range of 28%), we suggest further work to examine the combined (and coupled) influence of both factors in overall system performance to cope with hydrological droughts and variable climate conditions.

6 Open Research

This research uses open-source python v3.8.5 software for all ML applications and the GoldSim software environment for the SLCDPU systems model. We provide access to all python-base models at the following github link: `https://github.com/whitelightning450/Water-Demand-Forecasting`. This repository contains all data to train and run the CSD-WDM. The SLC-WSM is not provided for review due to security reasons specified by SLCDPU. Permission for this model require direct consent from SLCDPU. We do provide access to simulation results and analysis tools in an open source data repository” `https://github.com/whitelightning450/SLC_Water_Systems_Analysis`

Acknowledgments

This study is made possible by collaborative dedication to integrating climate resilience into water resource management by the Salt Lake City Department of Public Utilities, the University of Utah Climate Vulnerability Group, and the Alabama Water Institute.

References

Asefa, T., Clayton, J., Adams, A., & Anderson, D. (2014, 01). Performance evaluation of a water resources system under varying climatic conditions: Reliability, resilience, vulnerability and beyond. *Journal of Hydrology*, *508*, 53–65. doi: 10.1016/j.jhydrol.2013.10.043

Billings, B., & Jones, C. (2011). *Forecasting urban water demand* (3rd ed.). Denver, CO: American Waterworks Association.

Brooks, P. D., Gelderloos, A., Wolf, M. A., Jamison, L. R., Strong, C., Solomon, D. K., … others (2021). Groundwater-mediated memory of past climate controls water yield in snowmelt-dominated catchments. *Water Resources Research*, *57*(10), e2021WR030605.

Brown, C., Ghile, Y., Laverty, M., & Li, K. (2012). Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector. *Water Resources Research*, *48*(9). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR011212` doi: 10.1029/2011WR011212

Collins, B., & Associates. (2019). *2018 supply and demand master plan* (Tech. Rep.). Salt Lake City, Utah: BCA.

Davidson, R., MacKinnon, J. G., et al. (2004). *Econometric theory and methods* (Vol. 5). Oxford University Press New York.

DeOreo, W. (2006, 02). The role of water conservation in a long-range drought plan. *Journal American Water Works Association*, *98*, 94-101. doi: 10.1002/j.1551-8833.2006.tb07591.x

Donkor, E., Mazzuchi, T., Soyer, R., & Roberson, A. (2014, 02). Urban water demand forecasting: Review of methods and models. *Journal of Water Resources Planning and Management*, *140*, 146-159. doi: 10.1061/(ASCE)WR.1943-5452.0000314

Finnessey, T., Hayes, M., Lukas, J., & Svoboda, M. (2016). Using climate information for drought planning. *Climate Research*, *70*, 251-263. doi: 10.3354/cr01406

Goharian, E., & Burian, S. (2018, 02). Developing an integrated framework to build a decision support tool for urban water management. *Journal of Hydroinformatics*, *20*, jh2018088. doi: 10.2166/hydro.2018.088

Goharian, E., Burian, S., Lillywhite, J., & Hile, R. (2016, 11). Vulnerability assessment to support integrated water resources management of metropolitan water supply systems. *Journal of Water Resources Planning and Management*, *143*. doi: 10.1061/(ASCE)WR.1943-5452.0000738

Goharian, E., Burian, S. J., Lillywhite, J., & Hile, R. (2017). Vulnerability assessment to support integrated water resources management of metropolitan water supply systems. *Journal of Water Resources Planning and Management*, *143*(3). doi: 10.1061/(asce)wr.1943-5452.0000738

Goldsim. (2013). *Goldsim probabilistic simulation environment.* Issaquah, Washington: GoldSim Technological Group LLC.

Grayman, W. M. (2005). *Incorporating uncertainty and variability in engineering analysis* (Vol. 131) (No. 3). American Society of Civil Engineers.

Howe, C., & Goemans, C. (2007, 10). The simple analytics of demand hardening. *Journal / American Water Works Association*, *99*, 24-25. doi: 10.1002/j.1551-8833.2007.tb08052.x

Inman, D., & Jeffrey, P. (2006, 09). A review of residential water conservation tool performance and influences on implementation effectiveness. *Urban Water Journal - URBAN Water J*, *3*, 127-143. doi: 10.1080/15730620600961288

Jenks, G. (1967). The data model concept in statistical mapping. *Int. Yearbook Cartography*, 186-190.

Johnson, R. C., Burian, S., Oroza, A., Carlos, Stewart, J., & Kirkham, T. (2022). Water demand is not stationary: A supply-based machine learning approach to forecast seasonal demands in variable climate conditions. *Journal of Water Resources Planning and Management*, *in submission*.

Johnson, R. C., Wolf, M., Jamison, L., Burian, S., Oroza, A., Carlos, Brooks, D., Paul, . . . Kirkham, T. (2021). Drought in the west: Embedded water demand stationarity compromises system vulnerability analysis. *Open Water Journal*, *7*.

Khatri, K. B., & Strong, C. (2020). *Climate change, water resources, and potential adaptation strategies in utah.* Salt Lake City, UT: Division of Water Resources, Utah Department of Natural . . . .

Koutsoyiannis, D. (2006). Nonstationarity versus scaling in hydrology. *Journal of Hydrology*, *324*(1-4), 239–254.

Liu, A., Giurco, D., & Mukheibir, P. (2015, 10). Urban water conservation through customised water and end-use information. *Journal of Cleaner Production*, *112*. doi: 10.1016/j.jclepro.2015.10.002

Makropoulos, C., Nikolopoulos, D., Palmen, L., Kools, S., Segrave, A., Vries, D., . . . others (2018). A resilience assessment method for urban water systems. *Urban Water Journal*, *15*(4), 316–328.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science*, *319*, 573 - 574.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis.* John Wiley & Sons.

Muir, M., Luce, C., Gurrieri, J., Matyjasik, M., Bruggink, J., Weems, S., . . . Leahy, S. (2018). *Effects of climate change on hydrology, water resources, and soil.* Washington, DC: USDA Forest Service.

Nikolopoulos, D., van Alphen, H.-J., Vries, D., Palmen, L., Koop, S., van Thienen, P., . . . Makropoulos, C. (2019). Tackling the "new normal": A resilience assessment method applied to real-world urban water systems. *Water*, *11*(2), 330.

NOAA. (2021). *Mesowest alta guard current conditions.* https://www.wrh.noaa.gov/mesowest/getobext.php.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the kÃppen-geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. Retrieved from `https://www.hydrol-earth-syst-sci.net/11/1633/2007/` doi: 10.5194/hess-11-1633-2007

Salt Lake City Department of Public Utilities. (2021). *Continuing dry conditions trigger a stage 2 water shortage response and a stronger call for conservation from mayor mendenhall.* https://www.slc.gov/blog/2021/05/27/continuing-dry-conditions-trigger-a-stage-2-Water-shortage-response-and-a-stronger-call-for-conservation-from-mayor-mendenhall/.

Scalzitti, J., Strong, C., & Kochanski, A. K. (2016). A 26 year high-resolution dynamical downscaling over the wasatch mountains: Synoptic effects on winter precipitation performance. *Journal of Geophysical Research: Atmospheres*, *121*(7), 3224-3240. doi: 10.1002/2015JD024497

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N., Clark, D., . . . Kabat, P. (2014, 03). Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 3245-3250. doi: 10.1073/pnas.1222460110

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In *Proceedings of the 9th python in science conference* (Vol. 57, p. 61).

Smith, K., Strong, C., & Wang, S.-Y. (2015, 05). Connectivity between historical great basin precipitation and pacific ocean variability: A cmip5 model evaluation. *Journal of Climate*, *28*, 150520110052002. doi: 10.1175/JCLI-D-14-00488.1

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of cmip5 and the experiment design. *Bulletin of the American meteorological Society*, *93*(4), 485–498.

UDNR. (2010). *Jordan river basin planning for the future* (Tech. Rep.). Utah Division of Water Resources. Retrieved from `https://Water.utah.gov/wp-content/uploads/2019/SWP/JordanRiver/Jordan-River-Basin-Final2010.pdf`

UDNR. (2014). *State of utah municipal and industrial water supply and use study summary 2010* (Tech. Rep.). Utah Department of Natural Resources. Retrieved from `$https://Water.utah.gov/wp-content/uploads/2019/03/2010-M_I-Statewide-SummaryCH.pdf$`

UDWR. (2021). *Water records/use information* (Tech. Rep.). Salt Lake City, Utah: Utah Division of Water Rights. Retrieved from `https://Waterrights.utah.gov/Wateruse/WaterUseList.asp`

Vose, D. (2008). *Risk analysis: a quantitative guide*. John Wiley & Sons.

Wang, K., & Davies, E. (2018, 03). Municipal water planning and management with an end-use based simulation model. *Environmental Modelling and Software*, *101*, 204-217. doi: 10.1016/j.envsoft.2017.12.024

Wang, S.-Y., Gillies, R. R., Jin, J., & Hipps, L. E. (2010). Coherence between the great salt lake level and the pacific quasi-decadal oscillation. *Journal of Climate*, *23*(8), 2161-2177. doi: 10.1175/2009jcli2979.1

Wei, S., & Gnauck, A. (2007). *Simulating water conflicts using game theoretical models for water resources management in ecosystems and sustainable development* (4th ed.). WIT Press.

Westra, S., Thyer, M., Leonard, M., Kavetski, D., & Lambert, M. (2014). A strategy for diagnosing and interpreting hydrological model nonstationarity. *Water Resources Research*, *50*(6), 5090–5113.

Zhao, G., Gao, H., Kao, S., Voisin, N., & Naz, B. (2018, 8). A modeling framework for evaluating drought resilience of a surface water supply system under non-stationarity. *Journal of Hydrology*, *563*, 22-32. doi: doi.org/10.1016/j.jhydrol.2018.05.037

Figures/2021Dry Supply and TD Demands.pdf

Figure 3. The forecasted SLCDPU performance during dry supply conditions for observed and stationary traditional demands. The traditional demand estimate is a poor forecast of true demand during the dry climate simulation and produces SLCDPU forecasted performance in terms of Deer Creek water request significantly different from observed. Also, traditional methods do not provide any estimate of the prediction confidence (e.g. range of prediction uncertainty)

Figure 3.

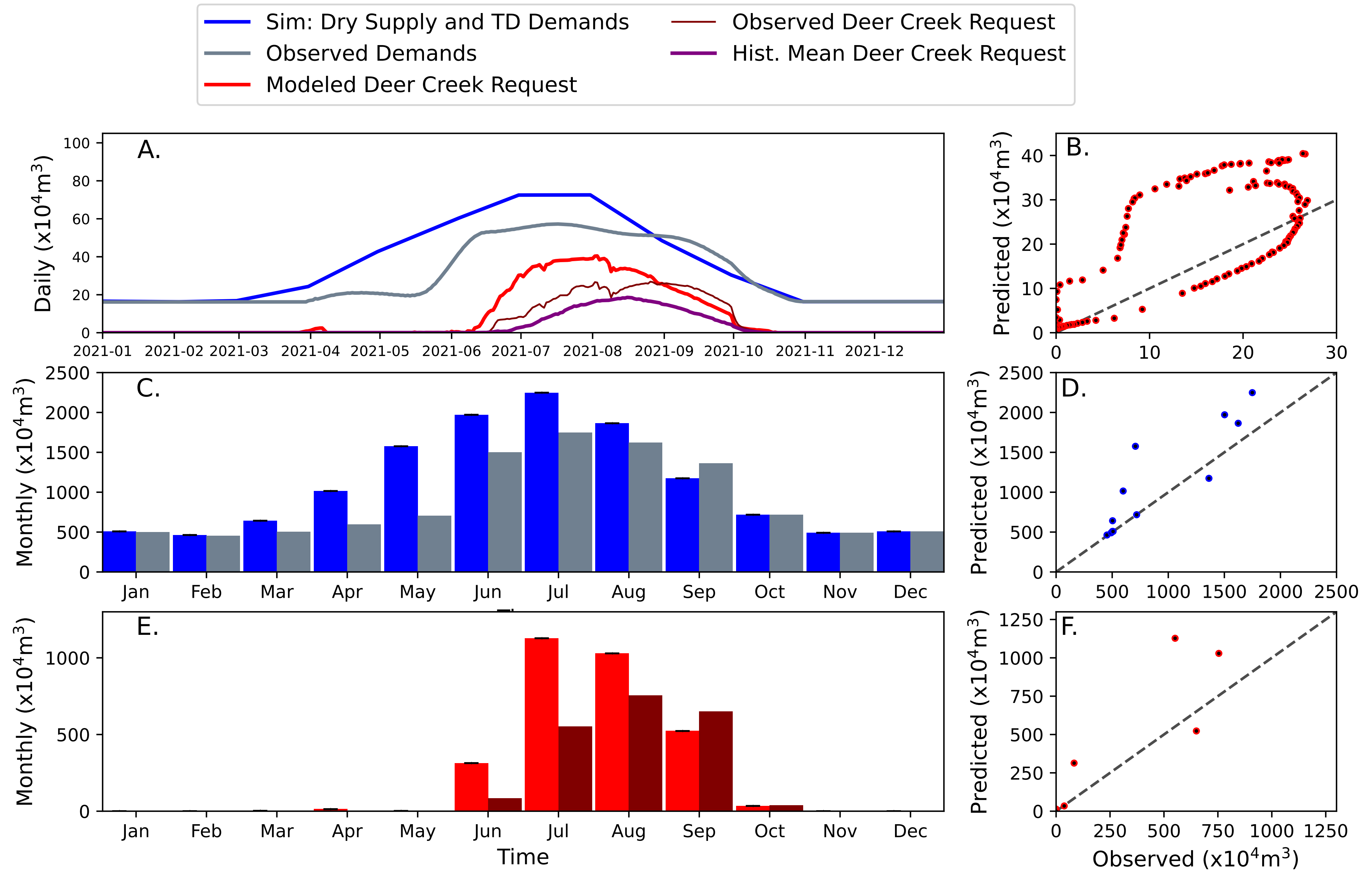
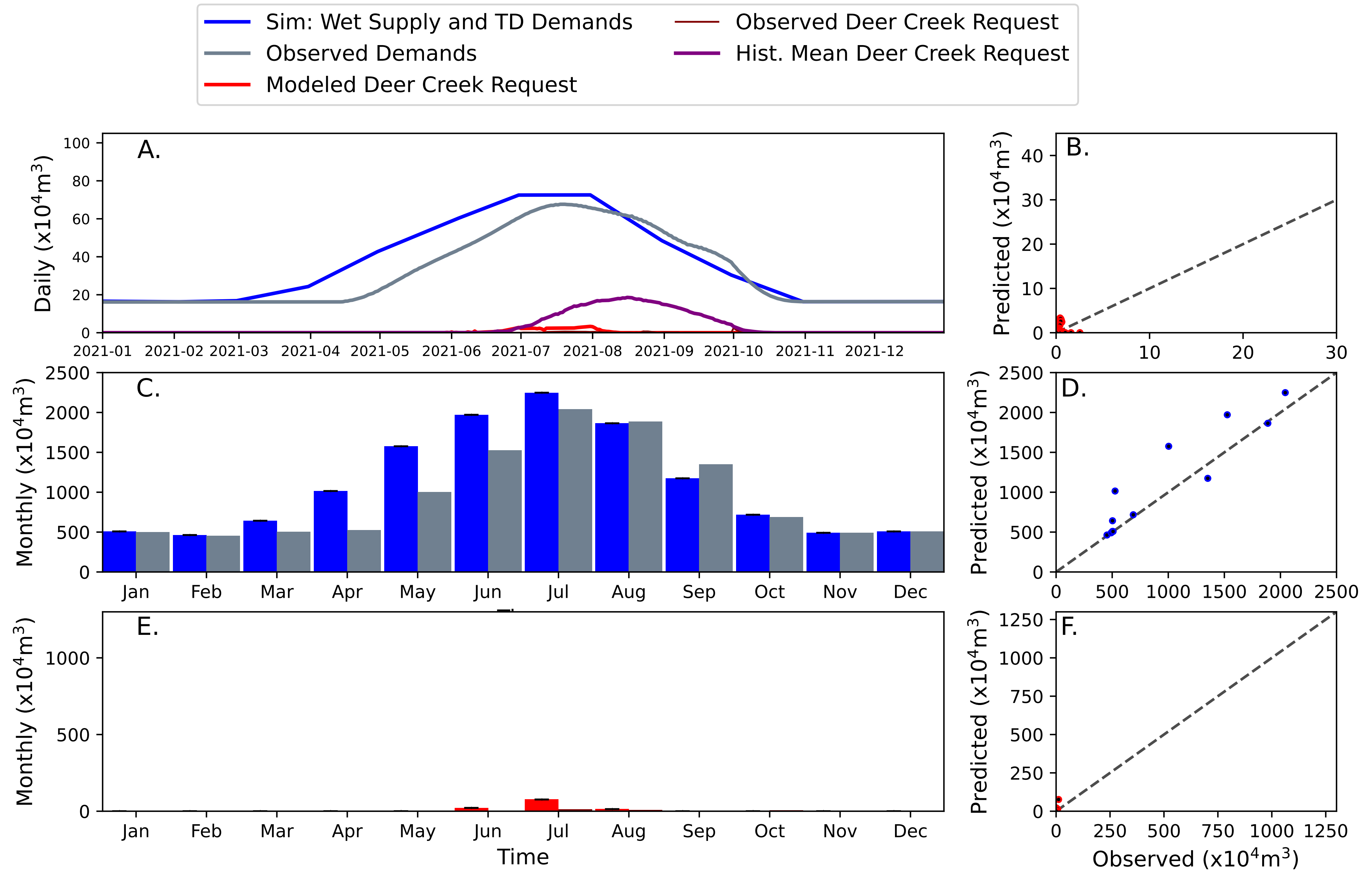
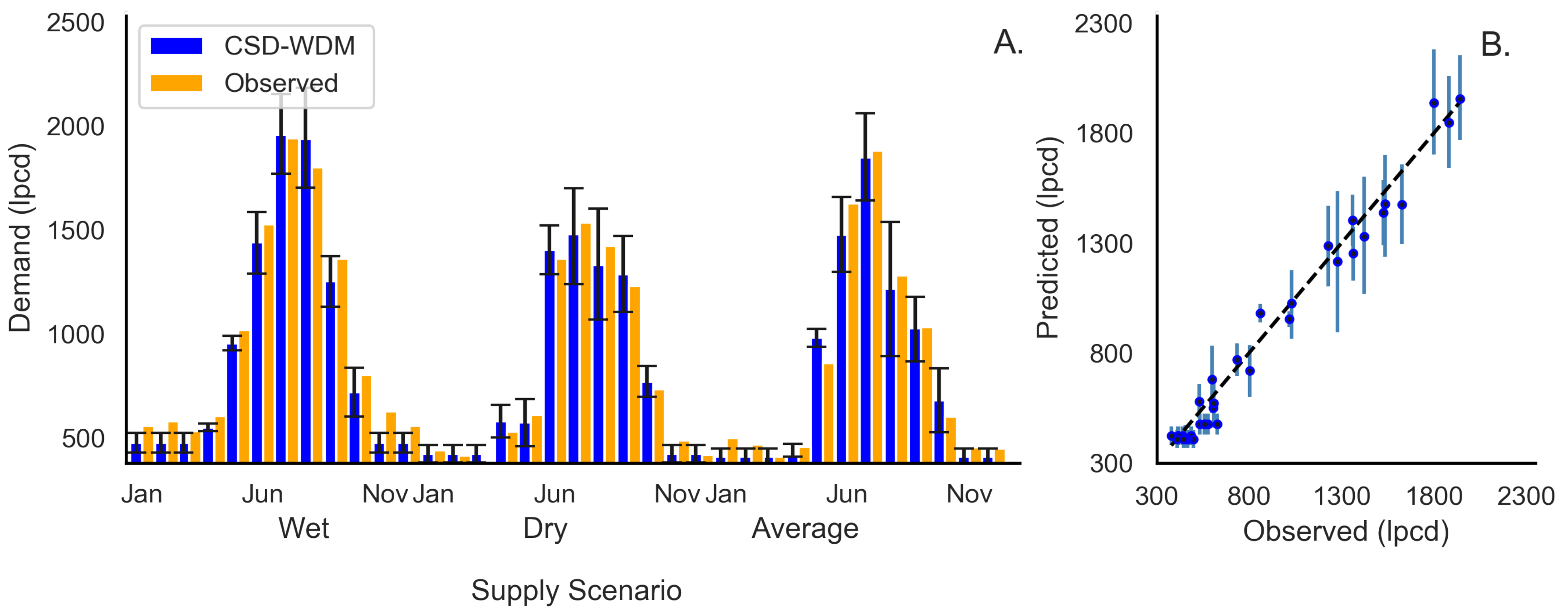

Sim: Dry Supply and TD Demands
Observed Demands
Modeled Deer Creek Request
Observed Deer Creek Request
Hist. Mean Deer Creek Request
A.
Daily (x10^{4}m^3)
2021-01 2021-02 2021-03 2021-04 2021-05 2021-06 2021-07 2021-08 2021-09 2021-10 2021-11 2021-12
B.
Predicted (x10^{4}m^3)
C.
Monthly (x10^{4}m^3)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
D.
Predicted (x10^{4}m^3)
E.
Monthly (x10^{4}m^3)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Time
F.
Predicted (x10^{4}m^3)
Observed (x10^{4}m^3)

Figure S7.



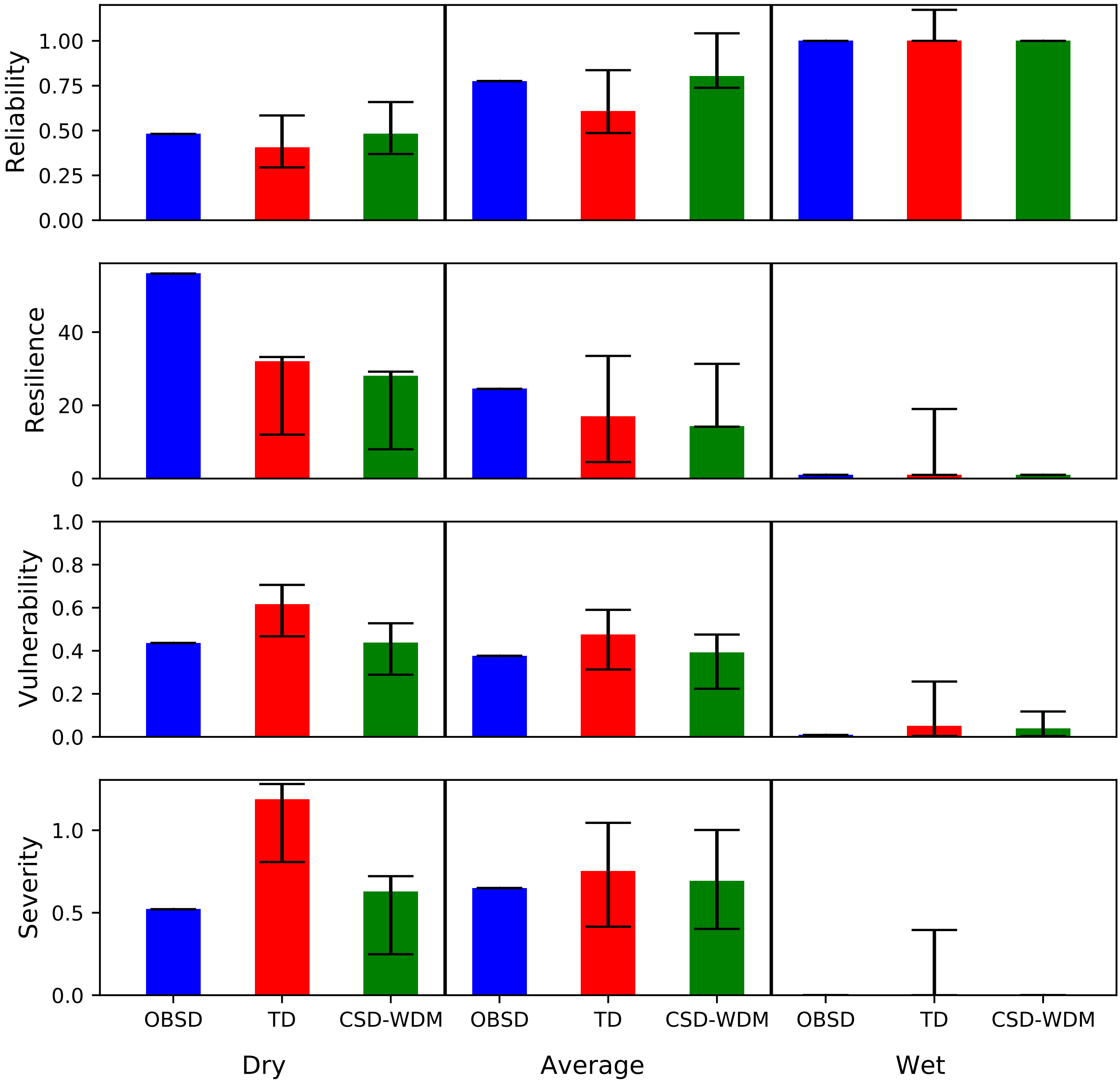
Sim: Wet Supply and TD Demands
Observed Demands
Modeled Deer Creek Request
Observed Deer Creek Request
Hist. Mean Deer Creek Request
A.
Daily ($x10^4m^3$)
B.
Predicted ($x10^4m^3$)
C.
Monthly ($x10^4m^3$)
D.
Predicted ($x10^4m^3$)
E.
Monthly ($x10^4m^3$)
F.
Predicted ($x10^4m^3$)
Time
Observed ($x10^4m^3$)

Figure S1.



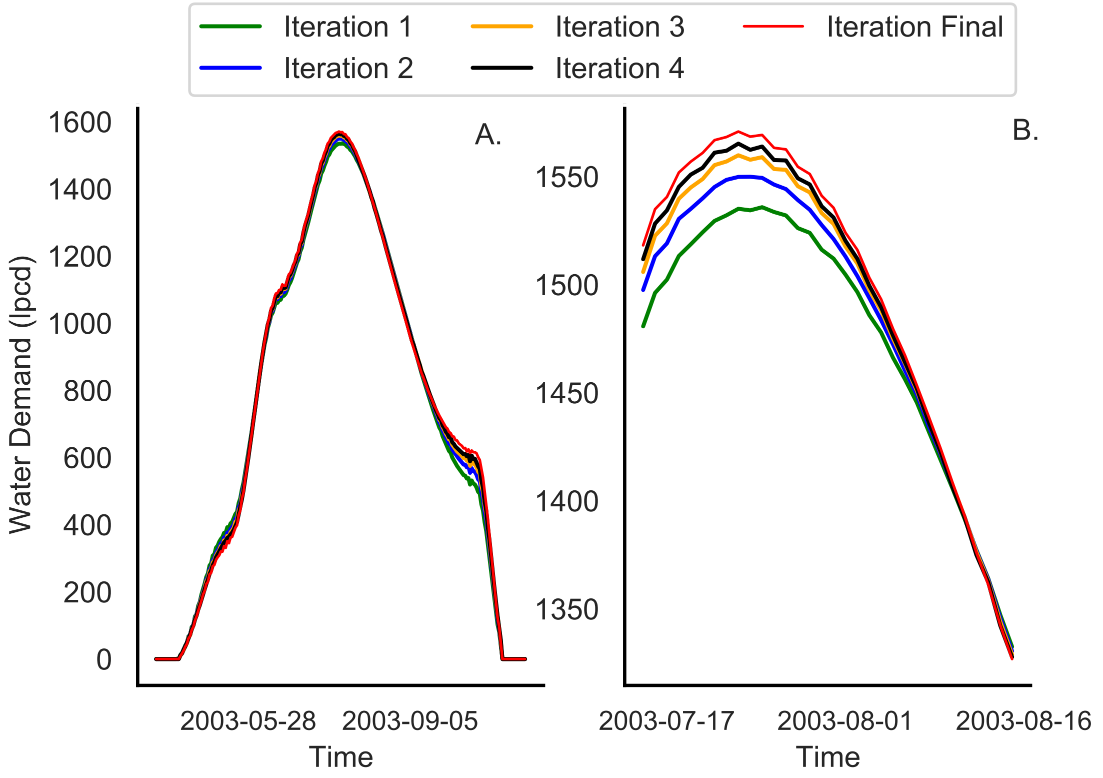
CSD-WDM
Observed
A.
B.
Demand (lpcd)
2500
2000
1500
1000
500
Jan
Jun
Nov
Jan
Jun
Nov
Jan
Jun
Nov
Wet
Dry
Average
Supply Scenario
Predicted (lpcd)
2300
1800
1300
800
300
300
800
1300
1800
2300
Observed (lpcd)

Figure S4.



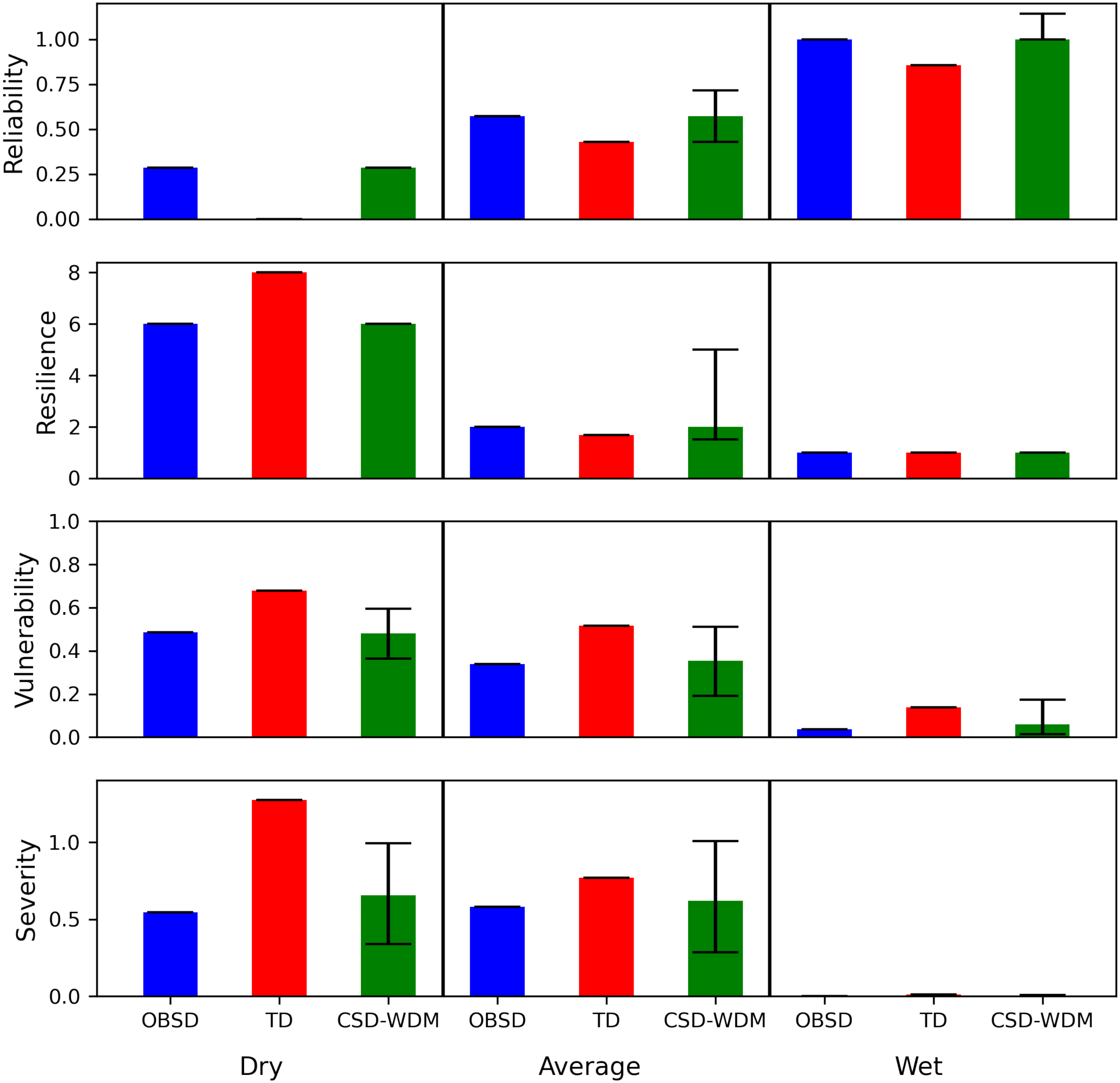
Reliability
1.00
0.75
0.50
0.25
0.00
Resilience
40
20
0
Vulnerability
1.0
0.8
0.6
0.4
0.2
0.0
Severity
1.0
0.5
0.0
OBSD
TD
CSD-WDM
OBSD
TD
CSD-WDM
OBSD
TD
CSD-WDM
Dry
Average
Wet

Figure S2.



Iteration 1
Iteration 2
Iteration 3
Iteration 4
Iteration Final
A.
B.
Water Demand (lpcd)
0
200
400
600
800
1000
1200
1400
1600
2003-05-28
2003-09-05
Time
1350
1400
1450
1500
1550
2003-07-17
2003-08-01
2003-08-16
Time

Figure 2.



Reliability
1.00
0.75
0.50
0.25
0.00
Resilience
8
6
4
2
0
Vulnerability
1.0
0.8
0.6
0.4
0.2
0.0
Severity
1.0
0.5
0.0
OBSD
TD
CSD-WDM
OBSD
TD
CSD-WDM
OBSD
TD
CSD-WDM
Dry
Average
Wet

Figure S1.

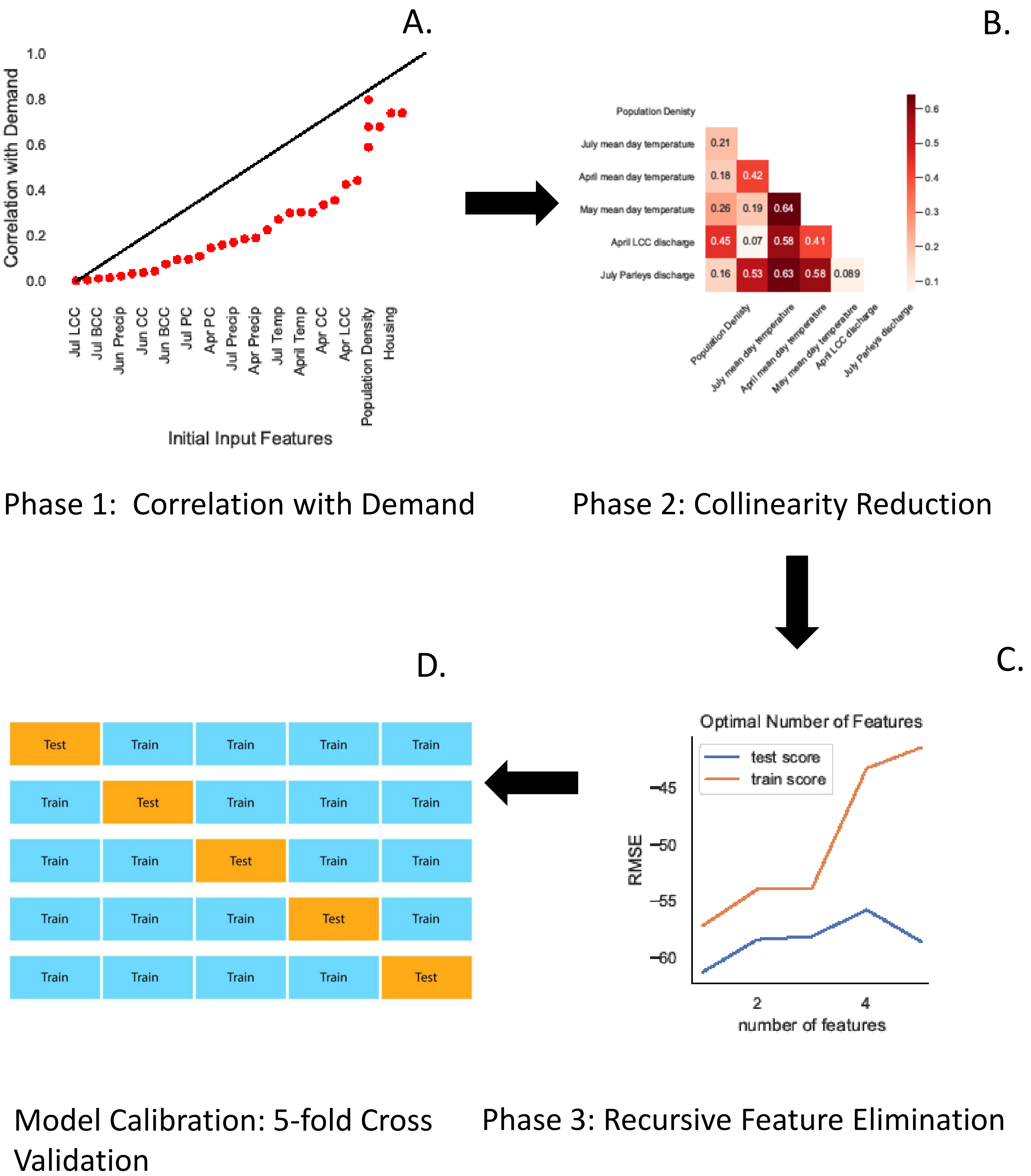
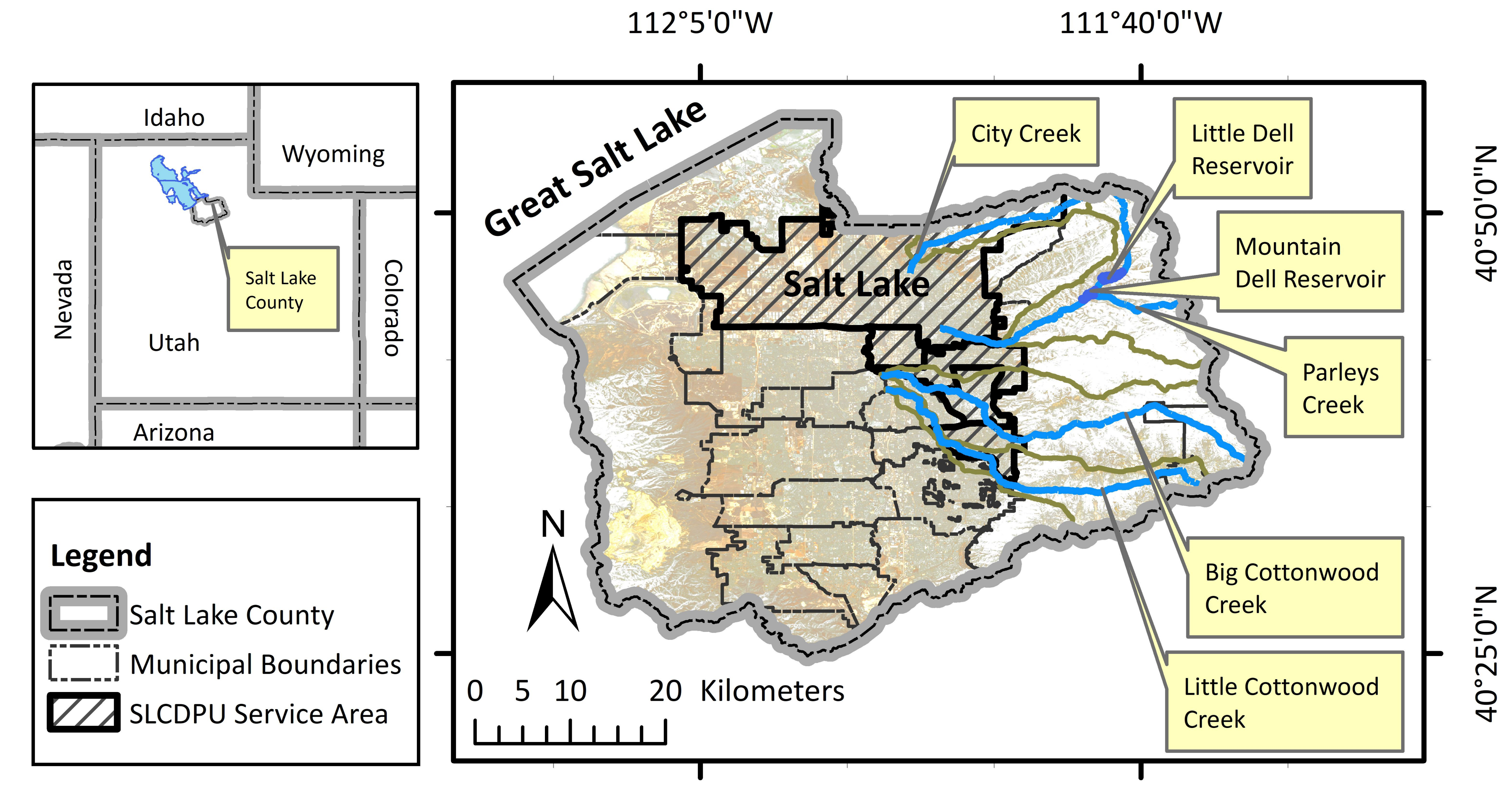
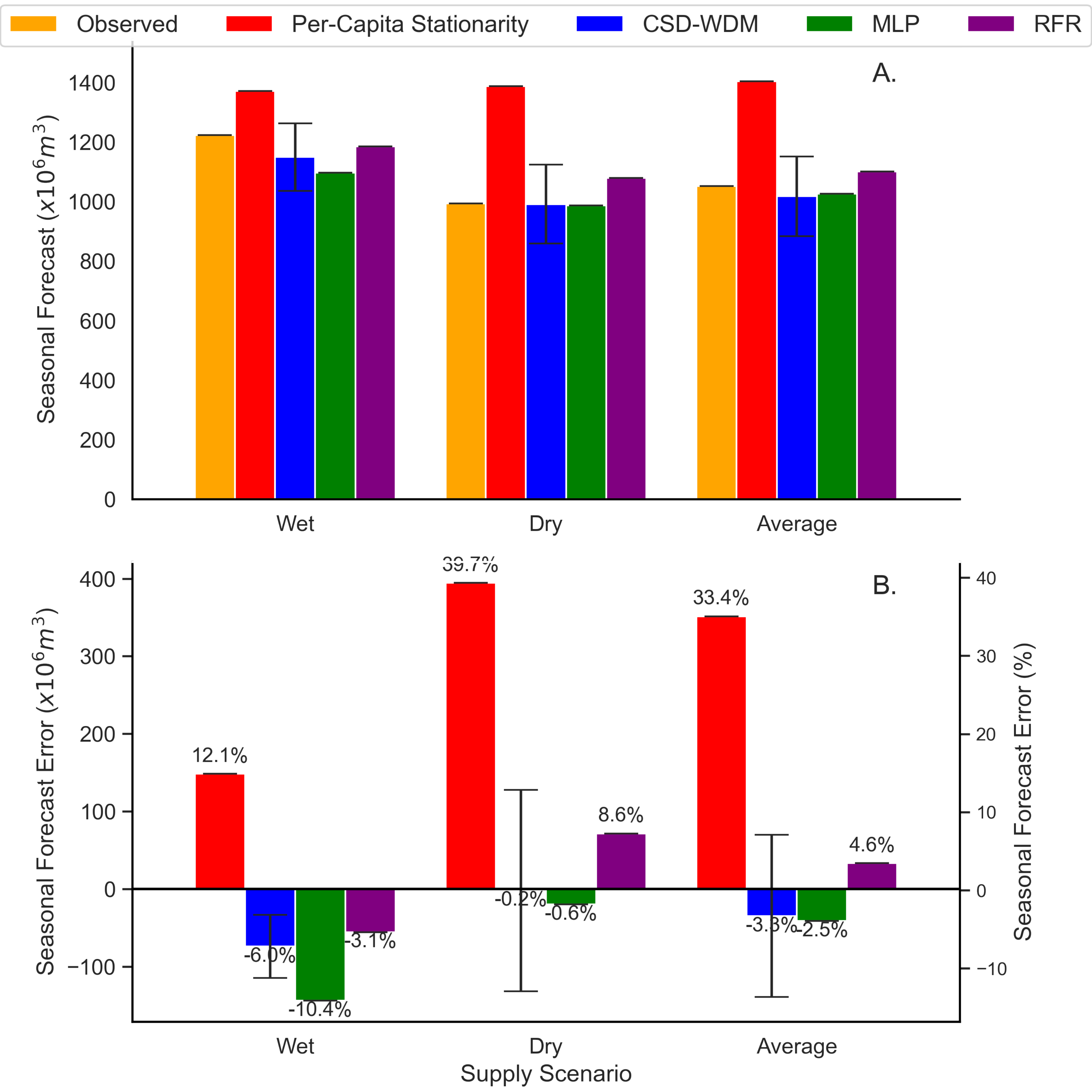

A.
Correlation with Demand
1.0
0.8
0.6
0.4
0.2
0.0
Jul LCC
Jul BCC
Jun Precip
Jun CC
Jun BCC
Jul PC
Apr PC
Jul Precip
Apr Precip
Jul Temp
April Temp
Apr CC
Apr LCC
Population Density
Housing
Initial Input Features
Phase 1: Correlation with Demand
B.
Population Denisty
July mean day temperature
April mean day temperature
May mean day temperature
April LCC discharge
July Parleys discharge
0.21
0.18 0.42
0.26 0.19 0.64
0.45 0.07 0.58 0.41
0.16 0.53 0.63 0.58 0.089
0.6
0.5
0.4
0.3
0.2
0.1
Phase 2: Collinearity Reduction
C.
Optimal Number of Features
test score
train score
RMSE
−45
−50
−55
−60
2
4
number of features
Phase 3: Recursive Feature Elimination
D.
Test Train Train Train Train
Train Test Train Train Train
Train Train Test Train Train
Train Train Train Test Train
Train Train Train Train Test
Model Calibration: 5-fold Cross Validation

Figure 1.



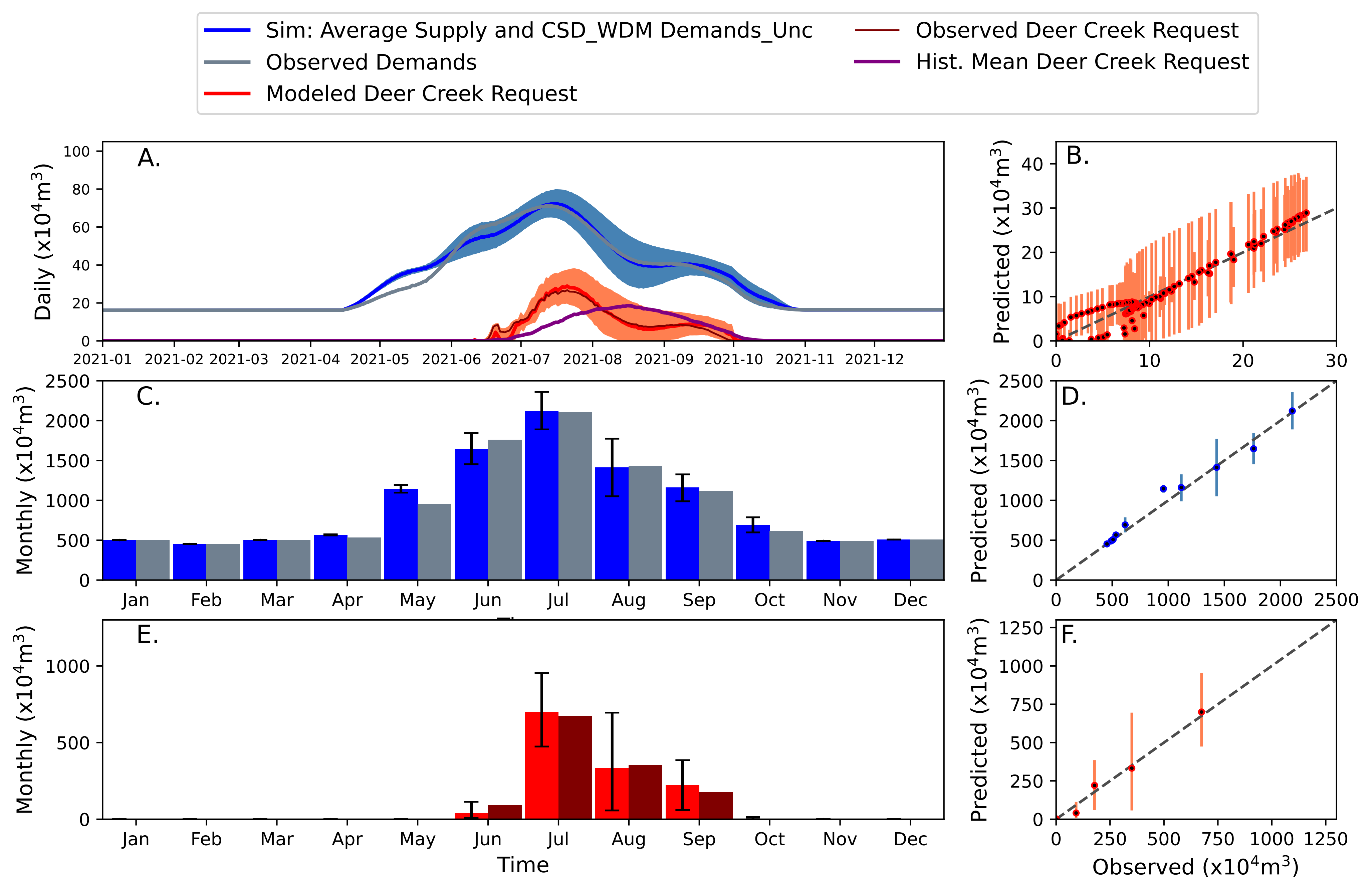
112°5'0"W
111°40'0"W
40°50'0"N
40°25'0"N
Idaho
Wyoming
Nevada
Utah
Colorado
Arizona
Salt Lake County
Great Salt Lake
Salt Lake
City Creek
Little Dell Reservoir
Mountain Dell Reservoir
Parleys Creek
Big Cottonwood Creek
Little Cottonwood Creek
N
0 5 10 20 Kilometers
Legend
Salt Lake County
Municipal Boundaries
SLCDPU Service Area

Figure S3.



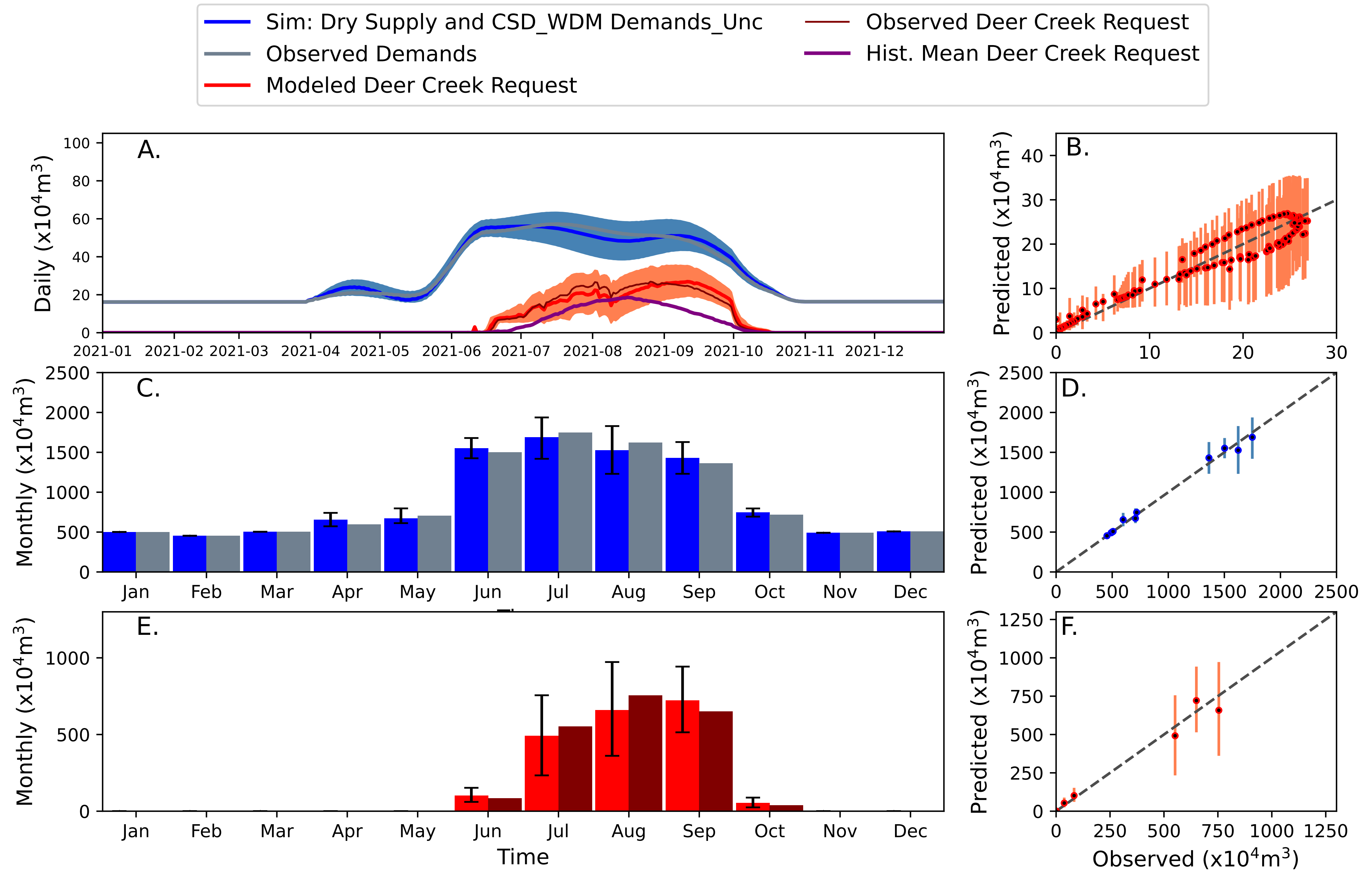
Observed
Per-Capita Stationarity
CSD-WDM
MLP
RFR
A.
Seasonal Forecast ($x10^6 m^3$)
1400
1200
1000
800
600
400
200
0
Wet
Dry
Average
B.
Seasonal Forecast Error ($x10^6 m^3$)
Seasonal Forecast Error (%)
400
300
200
100
0
−100
40
30
20
10
0
−10
12.1%
-6.0%
-10.4%
-3.1%
39.7%
-0.2%
-0.6%
8.6%
33.4%
-3.3%
-2.5%
4.6%
Wet
Dry
Average
Supply Scenario

Figure S8.



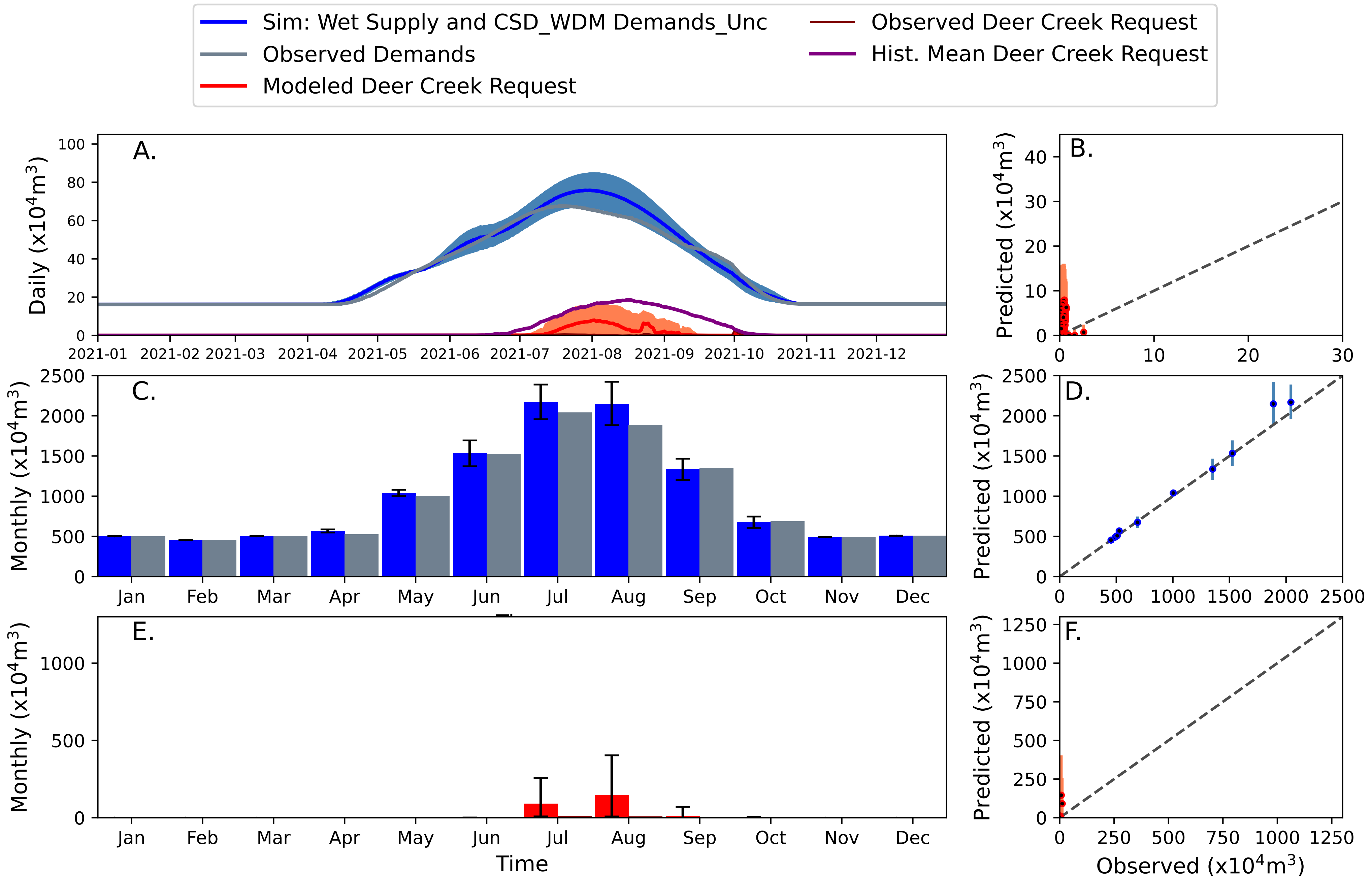
Sim: Average Supply and CSD_WDM Demands_Unc
Observed Demands
Modeled Deer Creek Request
Observed Deer Creek Request
Hist. Mean Deer Creek Request
A.
B.
C.
D.
E.
F.
Daily (x10^{4}m^3)
Monthly (x10^{4}m^3)
Predicted (x10^{4}m^3)
Observed (x10^{4}m^3)
Time
2021-01
2021-02
2021-03
2021-04
2021-05
2021-06
2021-07
2021-08
2021-09
2021-10
2021-11
2021-12
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Figure 4.



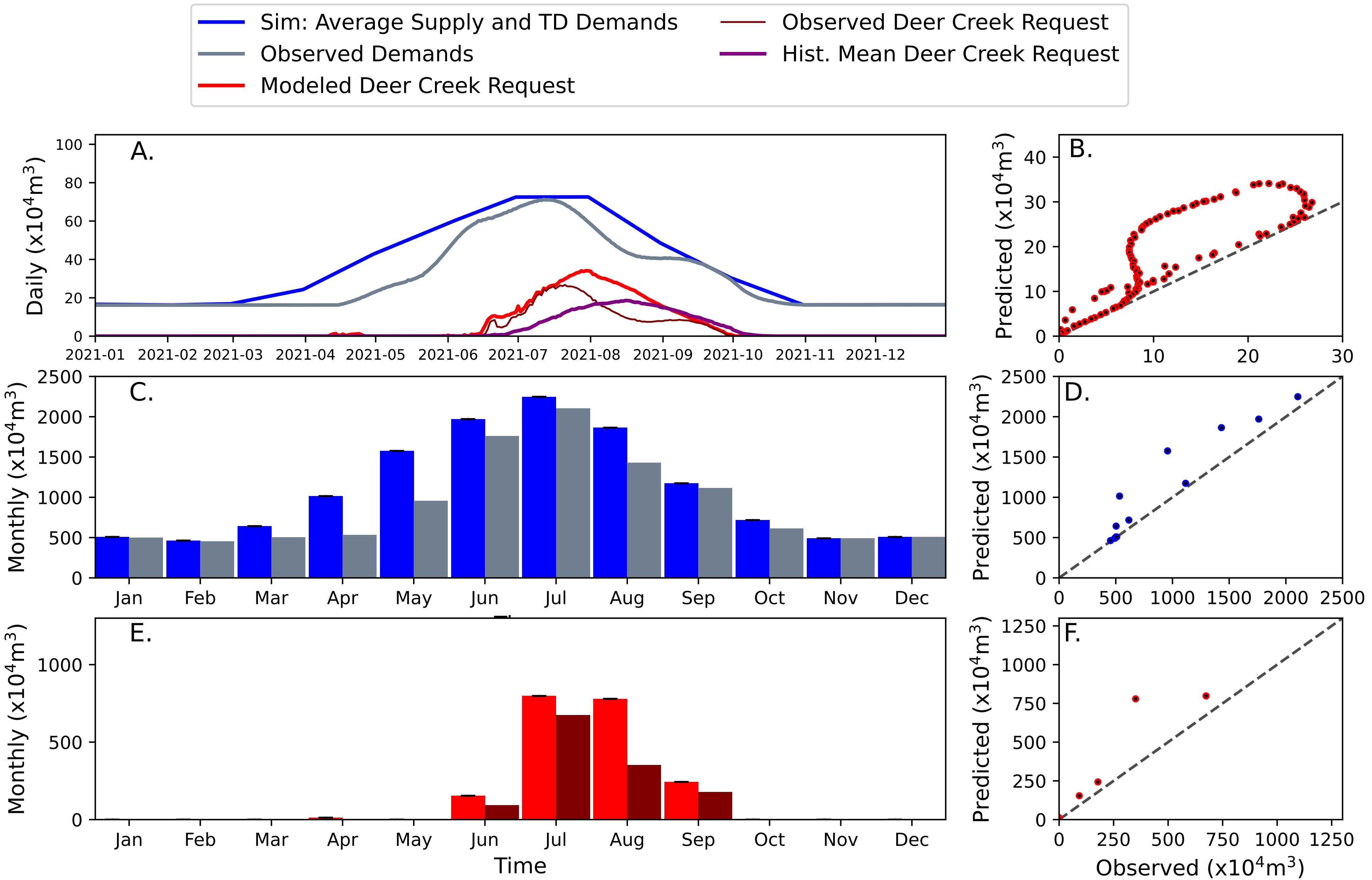
Sim: Dry Supply and CSD_WDM Demands_Unc
Observed Demands
Modeled Deer Creek Request
Observed Deer Creek Request
Hist. Mean Deer Creek Request
A.
Daily (x10^{4}m^3)
2021-01 2021-02 2021-03 2021-04 2021-05 2021-06 2021-07 2021-08 2021-09 2021-10 2021-11 2021-12
B.
Predicted (x10^{4}m^3)
C.
Monthly (x10^{4}m^3)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
D.
Predicted (x10^{4}m^3)
E.
Monthly (x10^{4}m^3)
Time
F.
Predicted (x10^{4}m^3)
Observed (x10^{4}m^3)

Figure S5.



Sim: Wet Supply and CSD_WDM Demands_Unc
Observed Demands
Modeled Deer Creek Request
Observed Deer Creek Request
Hist. Mean Deer Creek Request
A.
Daily (x10^{4}m^3)
2021-01 2021-02 2021-03 2021-04 2021-05 2021-06 2021-07 2021-08 2021-09 2021-10 2021-11 2021-12
B.
Predicted (x10^{4}m^3)
C.
Monthly (x10^{4}m^3)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
D.
Predicted (x10^{4}m^3)
E.
Monthly (x10^{4}m^3)
Time
F.
Predicted (x10^{4}m^3)
Observed (x10^{4}m^3)

Figure S8.



Sim: Average Supply and TD Demands
Observed Demands
Modeled Deer Creek Request
Observed Deer Creek Request
Hist. Mean Deer Creek Request
A.
B.
C.
D.
E.
F.
Daily (x10^{4}m^3)
Predicted (x10^{4}m^3)
Monthly (x10^{4}m^3)
Observed (x10^{4}m^3)
Time
2021-01
2021-02
2021-03
2021-04
2021-05
2021-06
2021-07
2021-08
2021-09
2021-10
2021-11
2021-12
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Figure 5.

Total SLCDPU Demands

Streamflow	-40%	-30%	-20%	-10%	0%	10%	20%	30%	40%
-50%	3.03	8.50	15.14	23.11	33.17	45.12	58.22	71.79	84.86
-40%	2.11	6.69	13.06	20.46	29.67	40.39	52.85	66.33	79.70
-30%	1.53	5.27	11.17	18.05	26.62	36.66	47.88	61.30	75.04
-20%	1.15	4.28	9.46	16.04	23.94	33.55	43.76	56.67	70.66
-10%	0.86	3.58	8.09	14.26	21.64	30.74	40.25	52.25	66.42
0%	0.66	2.95	7.01	12.70	19.89	28.50	37.29	48.32	62.17
10%	0.53	2.44	6.14	11.41	18.31	26.71	34.93	45.00	58.18
20%	0.41	1.95	5.40	10.32	16.87	25.29	33.21	42.42	54.60
30%	0.30	1.58	4.76	9.29	15.68	23.92	31.87	40.44	51.64
40%	0.22	1.34	4.15	8.40	14.51	22.74	30.75	38.82	49.25
50%	0.12	1.17	3.59	7.64	13.45	21.51	29.76	37.45	47.22

Deer Creek Reservoir Water Requests in $1x10^6m^3$