— BUREAU OF —
RECLAMATION

Improving the reliability of southwestern US water supply forecasting

Science and Technology Program
Research and Development Office
Final Report No. ST-2018-8117-01

Rio Grande Gorge at Rio Grande Gorge Bridge, NM

U.S. Department of the Interior April 2021

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30-09-2021	Research	2018-2021

4. TITLE AND SUBTITLE Improving the resiliency of southwestern US water supply forecasting in the face of climate trends and variability	5a. CONTRACT NUMBER WBS/WOID Research Office Cooperative Agreement
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 1541 (S&T)
6. AUTHOR(S) Andy Wood (PI), Josh Sturtevant, Lucas Barrett, Dagmar Llewellyn	5d. PROJECT NUMBER ST-2018-8117-01
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Center for Atmospheric Research Research Applications Laboratory 3450 Mitchell Ln Boulder, CO 80301	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Science and Technology Program Research and Development Office Bureau of Reclamation U.S. Department of the Interior Denver Federal Center PO Box 25007, Denver, CO 80225-0007	10. SPONSOR/MONITOR'S ACRONYM(S) Reclamation
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) ST-2018-8117-01

12. DISTRIBUTION/AVAILABILITY STATEMENT
Final Report may be downloaded from https://www.usbr.gov/research/projects/index.html

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Recent decades have experienced strong trends in hydrometeorology in the western US with declining watershed runoff efficiency, which may be undermining the accuracy of conventional seasonal streamflow prediction methods that support water supply forecasts. There is a critical need to develop strategies to enhance the reliability of seasonal streamflow prediction methods so that they to continue to provide accurate, unbiased and reliable predictions by accounting for such variability. This project created an unusually detailed modeling and ESP prediction (hindcast) resource that helped to understand new strategies for water supply prediction in the Upper Rio Grande River basin. It was generated to have specific relevance to the URGWOM management model. A key finding from the project analysis is that ESP-based approaches to water supply volume disaggregation is likely to be viable as an operational strategy for Reclamation, and that ESP-based sequences were more on average more skillful than analog-based sequences. Additional analysis into climate predictability in the western US suggested that sub-seasonal to seasonal climate forecasts may have the potential to offset streamflow predictability losses due to warming trends and declining snowpack.

15. SUBJECT TERMS water supply forecasting, climate trends, S2S prediction, water operations and management, modeling					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Dagmar Llewellyn
a. REPORT U	b. ABSTRACT U	THIS PAGE U			19b. TELEPHONE NUMBER *(Include area code)* 505-250-5493

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Mission Statements

The Department of the Interior (DOI) conserves and manages the Nation's natural resources and cultural heritage for the benefit and enjoyment of the American people, provides scientific and other information about natural resources and natural hazards to address societal challenges and create opportunities for the American people, and honors the Nation's trust responsibilities or special commitments to American Indians, Alaska Natives, and affiliated island communities to help them prosper.

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

Disclaimer

Information in this report may not be used for advertising or promotional purposes. The data and findings should not be construed as an endorsement of any product or firm by the Bureau of Reclamation, Department of Interior, or Federal Government. The products evaluated in the report were evaluated for purposes specific to the Bureau of Reclamation mission. Reclamation gives no warranties or guarantees, expressed or implied, for the products evaluated in this report, including merchantability or fitness for a particular purpose.

Acknowledgements

The Science and Technology Program, Bureau of Reclamation, sponsored this research.

Improving the resiliency of southwestern US water supply forecasting in the face of climate trends and variability

Final Report No. ST-2018-8117-01

Prepared by:

- *Andrew Wood, National Center for Atmospheric Research.*
- *Josh Sturtevant, formerly of National Center for Atmospheric Research*
- *Lucas Barrett, Bureau of Reclamation, Upper Colorado Basin Region, Albuquerque Area Office*
- *Dagmar Llewellyn, Bureau of Reclamation, Upper Colorado Basin Region, Albuquerque Area Office*

Directorate: Upper Colorado Basin Region, Albuquerque Area Office

Reclamation Principal Investigator: Dagmar Llewellyn, Civil Engineer (Hydrologic), Water Management Division

Peer Review

Bureau of Reclamation
Research and Development Office
Science and Technology Program

Final Report No. ST-2018-8117-01

Improving the resiliency of southwestern US water supply forecasting in the face of climate trends and variability

Prepared by: Andrew Wood
Project Scientist III, National Center for Atmospheric Research

LUCAS BARRETT Digitally signed by LUCAS BARRETT
Date: 2021.09.30 13:11:41 -06'00'

Prepared by: Lucas Barrett
Bureau of Reclamation, Albuquerque Area Office

Prepared by: Dagmar Llewellyn
Bureau of Reclamation, Albuquerque Area Office

Peer Review by: Flavio Lehner
Cornell University

Contents

Page

Executive Summary

Need for research

Recent decades have seen strong low-frequency variability in hydrometeorology in the western US, with major river basins such as the Colorado, the Rio Grande, Columbia and the Sacramento-San Joaquin exhibiting significant floods and droughts and, in some areas, declining runoff efficiencies. Recent studies (Lehner et al, 2017; Woodhouse et al, 2016) find evidence that this variability is undermining the accuracy and predictive skill of conventional seasonal streamflow prediction methods that support water supply forecasts, which are based on assumptions of hydrometeorological stationarity – and in particular, on fixed relationships between winter rainfall and snowpack and spring runoff. Such streamflow forecasts form a central input not only to water resources management but also to stakeholders' planning decisions as each water year progresses. There is thus a critical need to develop strategies to enhance the reliability of seasonal streamflow prediction methods so that they to continue to provide accurate, unbiased and reliable predictions by accounting for such variability.

Research questions and methods

The objective of this project was to assess and improve the reliability of water supply forecasting methods in the face of hydrometeorological variability. The NCAR Project Team sought to develop, refine, evaluate, and demonstrate prototype strategies for incorporating meteorological variability into current operational statistical (Statistical Water Supply, or SWS) and model-based ensemble (Ensemble Streamflow Prediction, or ESP) water supply prediction methods. In the Upper Rio Grande basin, the relative value of the new ensemble streamflow forecasts and their impact on water management outcomes will be assessed qualitatively and quantitatively, comparing enhanced predictions (hindcasts) versus official predictions for prior years, in part by leveraging existing Reclamation RiverWare reservoir system models. A particular question investigated was whether ESP forecasts could be used to disaggregate operational volume forecast to provide input to management models. NCAR also aimed to assess the broader regional sensitivity of seasonal hydrologic predictability to long-term variability and trends, and the sufficiency of current operational meteorological forecasts (particularly for temperature) to improve the resiliency of water supply forecasts across the western US. To this end, sub-seasonal weather and seasonal climate forecasts were processed and analyzed.

Conclusions

Overall, the analysis demonstrated that the ESP-based approach to water supply volume disaggregation is likely to be viable as an operational strategy for Reclamation, and that ESP-based sequences were on average more skillful than analog-based sequences, from a daily error standpoint. We note that the ESP-based traces will have a different characteristic than the analogs, in that if the median or mean ESP sequence is used, it will be smoother than an observed sequence from an individual year. As a result, it will not over-prescribe (without skill) various runoff timings throughout the season, but equally it will not provide a realistic variability from week to week (because it smooths out such unpredictable variations). Depending on the quality of the model in different locations and the nature of the inflow point (e.g., into a reservoir or into the mainstem), it may be the preferable strategy to adopt.

At a presentation to Reclamation operators and URGWOM experts, it was noted that some locations in the basin have a bimodal runoff pattern due to both natural causes (such as different snowmelt timing from different headwater areas) and management (e.g., groundwater driven irrigation returns). To that the extent these causes are represented in the model (i.e., the natural effects), they would likely show up in muted fashion in the ESP mean or median shapes, but the management driven effects would not. Analog sequences might identify this behavior better than ESP in some locations. The NCAR team subsequently spent time assessing whether this bi-model behavior could be identified in observed flow records, and did not find strong support for this theory in the data.

This project created an unusually detailed modeling and ESP prediction (hindcast) resource that helped to understand new strategies for water supply prediction in the Upper Rio Grande River basin. It was generated to have specific relevance to the URGWOM management model, and shared with other groups who are doing capability development research in the Rio Grande, and will likely have continued importance following this project.

1. Introduction

1.1. Project background

The scarcity of water has been a defining feature in western United States (US) water management for its entire history, and associated decisions affect the welfare of the general public across multiple economic and social sectors. Water supply forecasts represent a critical input to this decision-making, providing sub-seasonal to interannual outlooks that help to guide reservoir operations planning, often through the use of reservoir management models. In the western US, over half of the annual streamflow is derived from snow, which together with watershed soil moisture provides the core information used in spring to predict variations in the snowmelt runoff period that commonly refills reservoir storages that were depleted throughout the fall and winter seasons, enabling allocations for water supply and agriculture during the relatively drier summer season. Since the 1930s, the predictive relationship between watershed moisture conditions and future runoff volumes has been exploited using statistical forecast methods by the Natural Resources Conservation Service (NRCS) [Helms et al., 2008], which in recent decades has collaborated with the National Weather Service (NWS) River Forecast Centers (RFCs) to issue seasonal water supply forecasts (for predictands such as April-July runoff volume) throughout the western United States [Pagano et al, 2014]. The NWS also introduced a model-based method of water supply forecasting (WSF) termed ensemble streamflow prediction (ESP), which applies hydrologic and river routing models to produce forecasts of watershed runoff and streamflow [Day, 1985; Twedt et al., 1977].

In the Upper Rio Grande (URG) basin, until 2020, the NRCS statistical volume forecasts were the only operational predictions available for use in driving water system projections with the Upper Rio Grande Water Operations Model (URGWOM). These multi-month volume forecasts were disaggregated to daily streamflow traces (sequences) at multiple URGWOM inflow points through the selection of analogue years (based on volume and expert judgement) to enable storage and release forecasts to be made with URGWOM, supporting the development of annual operating plans. In 2019-2020, the NWS developed new ESP-based WSFs for the URG, providing an expanded set of information for potential use in managing URG water resources. Such ESPs are used to provide inflow forecasts for reservoir systems throughout the western US and internationally. In general, though climate forecasting exists in various forms (either based on climate indices such as El Nino, or on climate model forecasts), such climate information is not yet widely used to inform ESP-based forecasting in the US.

1.2 Previous work

Prior research collaborations involving the current project team focused on the potential for climate forecasts to enhance WSF in the URG. In particular, the team tried to improve our understanding of whether climate trends have impacted (a) the relationship between watershed climate and runoff; and (b) whether any loss in predictability due to temperature trends could be offset through incorporating climate model forecasts of temperature into statistical predictions.

This work, which relied on paleoclimate reconstructions of runoff efficiency, found that runoff efficiency varies primarily in proportion to precipitation, but that there exists a clear secondary influence of temperature. In years of low precipitation, very low runoff efficiencies are made 2.5–3 times more likely by high temperatures. This temperature sensitivity appears to have strengthened in recent decades, implying future water management vulnerability should recent warming trends in the region continue (Lehner et al. 2017a). The downward trend in runoff efficiency over the last 30 years was shown to be significant, thus the team explored the benefits of including available seasonal temperature forecasts into seasonal streamflow forecast models to improve their skill. To that end, they used the publicly available seasonal climate forecasts from the North American Multi-Model Ensemble (NMME; 7 models) and later was able to add forecasts from the proprietary European Centre for Medium Range Weather Forecasts (ECMWF; 1 model), leveraging an ongoing collaboration between ECMWF and Dr. Wood, the NCAR PI for the project. The large ensemble prediction dataset was analyzed for the URG region, and the team corroborated that they could indeed improve the streamflow forecast skill for key headwater gages in the Colorado and Rio Grande by around 10% in hindcasts over the period 1987-2016 by adding temperature information to the current operational forecasting approach by the Natural Resources Conservation Service (NRCS) (Lehner et al. 2017b). This work involved collaboration with the NRCS, which assessed the potential for the work to improve their operational predictions for the URG and elsewhere.

1.3 Problem Statement

Recent decades have seen strong low-frequency variability in hydrometeorology in the western US, with major river basins such as the Colorado, the Rio Grande, Columbia and the Sacramento-San Joaquin exhibiting significant floods and droughts and, in some areas, declining runoff efficiencies. Recent studies (Lehner et al, 2017; Woodhouse et al, 2016) find evidence that this variability is undermining the accuracy and predictive skill of conventional seasonal streamflow prediction methods that support water supply forecasts, which are based on assumptions of hydrometeorological stationarity – and in particular, on fixed relationships between winter rainfall and snowpack and spring runoff. Such streamflow forecasts form a central input not only to water resources management but also to stakeholders' planning decisions as each water year progresses. There is thus a critical need to develop strategies to enhance the reliability of seasonal streamflow prediction methods so that they to continue to provide accurate, unbiased and reliable predictions by accounting for such variability. Three related scientific needs are:

- to understand the underlying hydrometeorological dynamics,
- to assess the broader west-wide extent of potential changes in hydrometeorological predictability, and
- to measure the potential of current operational meteorological predictions to strengthen water supply forecasts to support critical operational decisions.

1.4 Study objectives and approach

The objective of this project was to assess and improve the reliability of water supply forecasting methods in the face of hydrometeorological variability. The NCAR Project Team sought to develop, refine, evaluate, and demonstrate prototype strategies for incorporating meteorological variability into current operational statistical (Statistical Water Supply, or SWS) and model-based ensemble (Ensemble Streamflow Prediction, or ESP) water supply prediction methods. In the Upper Rio Grande basin, the relative value of the new streamflow forecasts and their impact on water management outcomes will be assessed qualitatively and quantitatively, comparing enhanced predictions (hindcasts) versus official predictions for prior years, in part by leveraging existing Reclamation RiverWare reservoir system models. NCAR also aimed to assess the broader regional sensitivity of seasonal hydrologic predictability to long-term variability and trends, and the sufficiency of current operational meteorological forecasts (particularly for temperature) to improve the resiliency of water supply forecasts across the western US. A number of tasks were outlined to meet these objectives.

Task 1 – Refinement, evaluation and demonstration of enhanced Water Supply Forecast approaches in the Upper Rio Grande River basin: Continue and refine prior research in the Upper Rio Grande basin to improve the skill of statistical water supply forecasts by using NOAA National Multi-Model Ensemble (NMME) temperature forecasts. This will include techniques such as ensemble member selection and weighing based on NMME forecasts for use in frameworks such as ESP water supply forecasts. These approaches will be applied to locations in the basin identified through coordination with Reclamation. Software will be developed for the NMME data acquisition and processing, and for the statistical and ESP-based hindcast and forecast applications. Software development will aim to facilitate incorporation of enhancements into existing operational forecast frameworks and also to enable use by entities not currently producing forecasts.

Task 2 – Assessment of Water Supply Forecast method resiliency across the western US: The NCAR Project Team will analyze historical observations (including precipitation, snow water equivalent, temperature and runoff) across the Reclamation domain to identify locations in which climate variability and trends are altering the predictability of spring runoff (or other key hydrologic phenomena). Retrospective watershed model simulations of hydrologic variability will be conducted to investigate changes to physical mechanisms and their drivers. The assessments will focus on 30-50 basins of interest to Reclamation, using a physically-oriented modeling approaches.

Task 3 – Assessment of NOAA National Multi-Model Ensemble climate forecast skill and potential to enhance Water Supply Forecasts for Reclamation-managed watersheds across the western US: Techniques developed in prior and related research by the NCAR team will be extended to assess climate forecast skill from the NOAA National Multi-Model Ensemble (NMME) across the western US. In addition to NMME predictions, forecasts based on approaches funded by Reclamation and other federal agencies will also be evaluated. For those forecast approaches found to have skill, their use in existing and new frameworks will be explored toward enhancing water supply forecasts.

Task 4 – Assessment of Upper Rio Grande basin water management using enhanced forecasts: The impact of improved seasonal streamflow forecasting will be assessed from the perspective of water management in the Upper Rio Grande basin. Focusing on selected prior years, NCAR will qualitatively evaluate the potential impacts of forecast improvements through discussions with Reclamation's water operations team. To the extent possible, quantitative assessments of forecast improvements on operations will be pursued using the RiverWare reservoir systems model.

1.5 Study team and partners

The research team included Andy Wood (NCAR), Josh Sturtevant (NCAR) Dagmar Llewellyn (Reclamation) and Lucas Barrett (Reclamation). Flavio Lehner was a co-investigator at NCAR before leaving to join Cornell University. Other external partners such as the NRCS National Water and Climate Center (NWCC, Portland, OR), and in particular Angus Goodbody, also interacted at times during the project to continue to pursue the prospects for operational climate-informed statistical WSF.

2. Methods and Data

2.1 Watershed modeling with SUMMA and mizuRoute

Understanding hydrological processes across a range of spatiotemporal scales is crucial to support myriad hydrometeorological applications in water security, energy supply, weather and climate adaptation planning, as well as facilitate investigating fundamental hydrology questions. Process-based hydrological models are a widely implemented tool to simulate water, energy and momentum fluxes within or between different hydrologic domains and subsystems. Over last few decades, process-based hydrological models of various concepts and structures have been modified and tested to obtain hydrologic predictions across a range of spatial (gridded, regional, continental, global etc.) scales to handle different scientific and engineering problems. Yet the collection of available models, which differ in arbitrary specifics, do not provide for systematic and controlled assessment and research into the relationship between model performance and specific modeling decisions, including model structure, parameters and parameterizations. To address this challenge, Clark et al. (2015a, 2015b) created a flexible modeling framework, namely the Structure for Unifying Multiple Modeling Alternatives (SUMMA), as a platform for testing and benchmarking different modeling approaches and parameterizations, different process representations across spatial scales, and different representations of spatial variability and hydrological connectivity.

The majority of SUMMA-based development to date has focused on the introduction and validation of process-oriented test cases, such as the demonstration of the sensitivity of snow accumulation and melt to the choice of parameterization. Watershed applications supporting calibrated streamflow modeling and prediction have only recently been supported through this project, among several sponsored by Reclamation and the US Army Corps of Engineers. To use SUMMA for the prediction activities undertaken in the Upper Rio Grande River basin required:

- The testing of an *a priori* SUMMA implementation with an initial configuration that provides a foundation for a range of applications;
- The development of scripts and workflows for model configuration, calibration, simulation and ensemble forecasting.
- The one-way coupling of SUMMA runoff to streamflow routing via the mizuRoute channel routing model (Mizukami et al., 201X; 202X)

- The implementation of the SUMMA and mizuRoute models within a real-time prediction system, and in this case the System for Hydromet Research, Applications and Predictions (SHARP) that is run at NCAR.
- The SUMMA code base supported the generation of a full range of hydrologic processes, but other model components and workflows necessary for users to apply SUMMA to a non-point watershed application did not exist. Work required included the creation of a priori parameter and attribute datasets capable of yielding credible baseline simulations; selection of calibration parameters and a strategy and workflows for model calibration; parameter sufficient lacked usable input parameter datasets; a stable structural configuration and initial model decision set, as well as output configurations; and long-term model forcings for use in off-the-shelf application. In addition, workflows supporting prediction applications such as seasonal forecasting were needed for water management partners supporting the development of SUMMA.

Similar to all hydrologic models, SUMMA input comprises static attributes and parameters and time-varying meteorological forcings; and unlike many models, SUMMA also requires a model decisions list. Where possible, existing NLDAS CONUS-wide datasets were used to derive the *a priori* SUMMA model attributes and parameters. NLDAS2 datasets represent a quality controlled and spatially distributed land surface modeling resource for CONUS at 1/8th degree grid resolution and have been widely applied for various applications in water resources (Xia et al., 2012c, 2012b).

The mizuRoute multi-method channel routing model (Mizukami et al, 2016) was implemented to route SUMMA hydrologic total runoff (surface and subsurface) through the basin's stream channel network and calculate streamflow. The model network is defined by the reach-based global MERIT-Hydro Flowlines network (Yamazaki et al, 2019), after extracting stream channel segments local to URG basin, and adding necessary routing parameters. The network resolves the stream reaches and key flow locations at an intermediate scale, somewhat finer than the HUC12 SUMMA model scale. A unit hydrograph (UH) routing method (termed 'impulse response function' or IRF in mizuRoute) was applied. As noted earlier regarding SUMMA, the complexity (density) of this network and routing algorithm influences the agility, usability and computational efficiency of the routing model solution. For the purposes of bulk spring runoff prediction, this intermediate-scale, intermediate-complexity modeling approach was chosen to enhance the computational feasibility of running multiple sequences of ensemble predictions.

2.1.1 Model structure

In SUMMA, the spatial organization has a 2-level hierarchy, including hydrologic response units (HRUs) that are nested within grouped response units (GRUs) to represent the modeling domain. The units can have any dimensionality, such as grid or polygon shapes. To achieve a balance in complexity that allowed the model to represent a useful degree of spatial heterogeneity without being prohibitively expensive (computationally) to run, the URG implementation (like others used in related projects for Reclamation) used a single HRU per GRU, and the distributed parameters were estimated for HRUS and GRUs defined by the USGS HUC12 watershed boundary dataset. The a priori SUMMA model was configured with 3 soil layers, one aquifer layer, and a maximum of 5 snow layers. The nominal depth of the soil layers is fixed at 2 meters, which is consistent with other land models used in large domain applications, while the height of the snow layers varies. The model timestep was set at 3 hours, versus 1 hour of less which is more common in process oriented modeling, but is more demanding (and limiting) computationally and with regard to model inputs and outputs. Sensitivity testing of these choices (e.g., 3 layer versus 8 layer soil, different timesteps, different total soil depths) was done using a CAMELS SUMMA dataset, confirming that these would be efficient selections making an acceptable tradeoff between model agility and complexity. For instance, due to the adaptive time-stepping of the SUMMA model, the difference in performance between the 3 hourly and 1 hourly models was second-order.

The mizuRoute model has been implemented in the course of this and related projects on several different channel routing networks, including the NHDPlusV2 network used in the National Water Model (NWM), with 2.7 million reaches nationwide; the much coarser Geospatial Fabric developed by the USGS, and the MERIT Hydro Flowlines network. The latter was chosen for this project because of its ability to resolve the stream channel network and key flow locations at an intermediate scale, somewhat finer than the HUC12 SUMMA model scale. As noted earlier regarding SUMMA, the complexity (density) of this network affects the agility, usability and computational efficiency of the routing model solution.

2.1.2 Model Attributes

SUMMA uses the 14 model attributes listed in Table 1. For a priori implementation, we adopted seven variables with spatially distributed values from the existing LIS-NLDAS2 dataset used for other NLDAS2 models already supported in LIS (e.g., NOAH-mp). Four attributes (measurement height above bare ground, contour length, average tangent slope) were set to default constants. The remaining four model attributes are related to identification of grid cells or HRUs within the hydrological model and hence, they have no direct relevance to NLDAS2 topographical/land data features.

Table 1. List of updated SUMMA attributes

Attribute name	Description	Implementation
elevation	NLDAS elevation	Distributed – NLDAS2
HRUarea	Area of each HRU	Distributed – NLDAS2
latitude	Latitude of HRU's centroid point	Distributed – NLDAS2
longitude	Longitude of HRU's centroid point	Distributed – NLDAS2
slopeTypeIndex	Index defining slope type	Distributed – NLDAS2
soilTypeIndex	Index defining soil type	Distributed – NLDAS2
vegTypeIndex	Index difining vegetation type	Distributed – NLDAS2
contourLength	Contour length of HRU	Constant – default
downHRUindex	Id of downslope HRU (0 = basin outlet)	Constant – default
hruId	Id of each HRU	Distributed – NLDAS2
mheight	Measurement height above bare ground	Constant – default
tan_slope	Average tangent slope of HRU	Constant – default
gruId	Id of group response unit (GRU)	Distributed – NLDAS2
hru2gruId	Id of GRU to which HRU belongs	Distributed – NLDAS2

2.1.2 Model parameters

SUMMA uses two different types of parameters for the two levels of the spatial hierarchy: HRU ('local') and GRU ('basin'). Local parameters are used to define variables and constants representing conservation and thermodynamic equations as well as numerical schemes used within SUMMA modeling framework. More specifically, local parameters are mainly related to water and energy fluxes within the physical domain of snow, soil and vegetation canopy, their mathematical models and computational strategy. Basin related parameters define hydrologic aquifers and baseflow properties that are necessary for flow routing applications.

SUMMA also allows for a hierarchy in the specification of parameters for these spatial levels. Two separate files define values for local and basin parameters that are constant over the global domain of the run; however, during model simulations, these a priori values are overwritten by values for the same parameters if they are attached to soil and vegetation types (attributes) specified individually for each HRU, as provided in parameter tables for soil (eg, STAS, STAS-RUC) and vegetation (eg, USGS-RUC, Modified MODIS) categories. SUMMA enables a third set of 'trial' parameters that overwrite the table-based parameters in turn, and appear in a separate input file that is distributed over all the HRUs or GRUs of the model. All levels of the parameters may be updated in calibration, depending on the update scripts that have been developed for this purpose.

To create a useful calibration parameter set, a priori values were estimated for approximately 40 parameters affecting a broad range of hydrologic processes (listed in Table 2). Nine of these were given distributed values associated with their soil and vegetation type library values (and are listed as 'distributed'). The maximum snow albedo parameter based on NLDAS Noah-MP snow albedo dataset. In contrast to calibration effort focusing on table-based settings for parameters, these can be set differently for each individual HRU. Not all are likely to be chosen for a given calibration effort, and more are likely to be added in the future. SUMMA has over 200 parameters, though not all are likely to be useful in model calibration.

Table 2. List of updated SUMMA parameters with NLDAS2

Parameter name	Description	Distributed (D) or Constant (C) and Relevance
critSoilTranspire	Critical vol. liquid water content when transpiration is limited	(D) Evapotranspiration
critSoilWilting	Critical vol. liquid water content when plants are wilting	(D) Evapotranspiration
fieldCapacity	Soil field capacity (vol. liquid water content when baseflow begins)	(D) Evapotranspiration
k_soil	**Soil hydraulic conductivity**	**(D) Soil water transmission**
theta_res	Soil residual volumetric water content	(D) Soil water storage
theta_sat	**Soil porosity**	**(D) Soil water storage**
vGn_alpha	Van Genutchen alpha parameter	(D) Soil water transmission
vGn_n	Van Genutchen n parameter	(D) Soil water transmission

albedoMax	Maximum snow albedo	(D) Snowmelt
Fcapil	**capillary retention as a fraction of the total pore volume**	**(C) Snowmelt**
albedoDecayRate	Albedo decay rate	(C) Snowmelt
albedoRefresh	critical mass necessary for albedo refreshment	(C) Snowmelt
aquiferBaseflowExp	**Baseflow exponent**	**(C) Baseflow**
aquiferBaseflowRate	**baseflow rate when aquifer storage = aquifer Scale Factor**	**(C) Baseflow**
heightCanopyTop	**height of canopy top**	**(D) Snow accumulation and melt, evapotranspiration, interception**
heightCanopyBottom	**height of canopy bottom**	**(D) Snow accumulation and melt, evapotranspiration, interception**
f_impede	Ice impedence parameter	(C) Snowmelt
frozenPrecipMultip	**frozen precipitation multiplier**	**(C) Snow accumululation**
k_macropore	**saturated hydraulic conductivity for macropores**	**(C) Soil water transmission**
mpExp	empirical exponent in macropore flow equation	(C) Soil water transmission
mw_exp	Exponent for melt water flow	(C) Snowmelt
qSurfScale	**scaling factor in the surface runoff parameterization**	**(C) Runoff**
refInterceptCapRain	reference canopy interception capacity per unit leaf area (rain)	(C) Canopy interception
refInterceptCapSnow	canopy interception capacity per unit leaf area (snow)	(C) Canopy interception
routingGammaScale	**scale parameter in Gamma distribution used for sub-grid routing**	**(C) Hillslope Routing**
routingGammaShape	**shape parameter in Gamma distribution used for sub-grid routing**	**(C) Hillslope Routing**
summerLAI	**maximum leaf area index at the peak of the growing season**	**(C) Evapotranspiration, interception**
tempCritRain	critical temperature where precipitation is rain	(C) Snow accumulation
tempRangeTimestep	temperature range over the timestep	(C) Numerics
theta_mp	minimum volumetric liquid water content for macropore flow	(C) Soil water transmission
throughfallScaleRain	scaling factor for throughfall (rain)	(C) Precipitation
wettingFrontSuction	Green-Ampt wetting front suction	(C) Infiltration
windReductionParam	canopy wind reduction parameter	(C) Evapotranspiration
zScale_TOPMODEL	scale factor for TOPMODEL-ish baseflow parameterization	(C) Baseflow

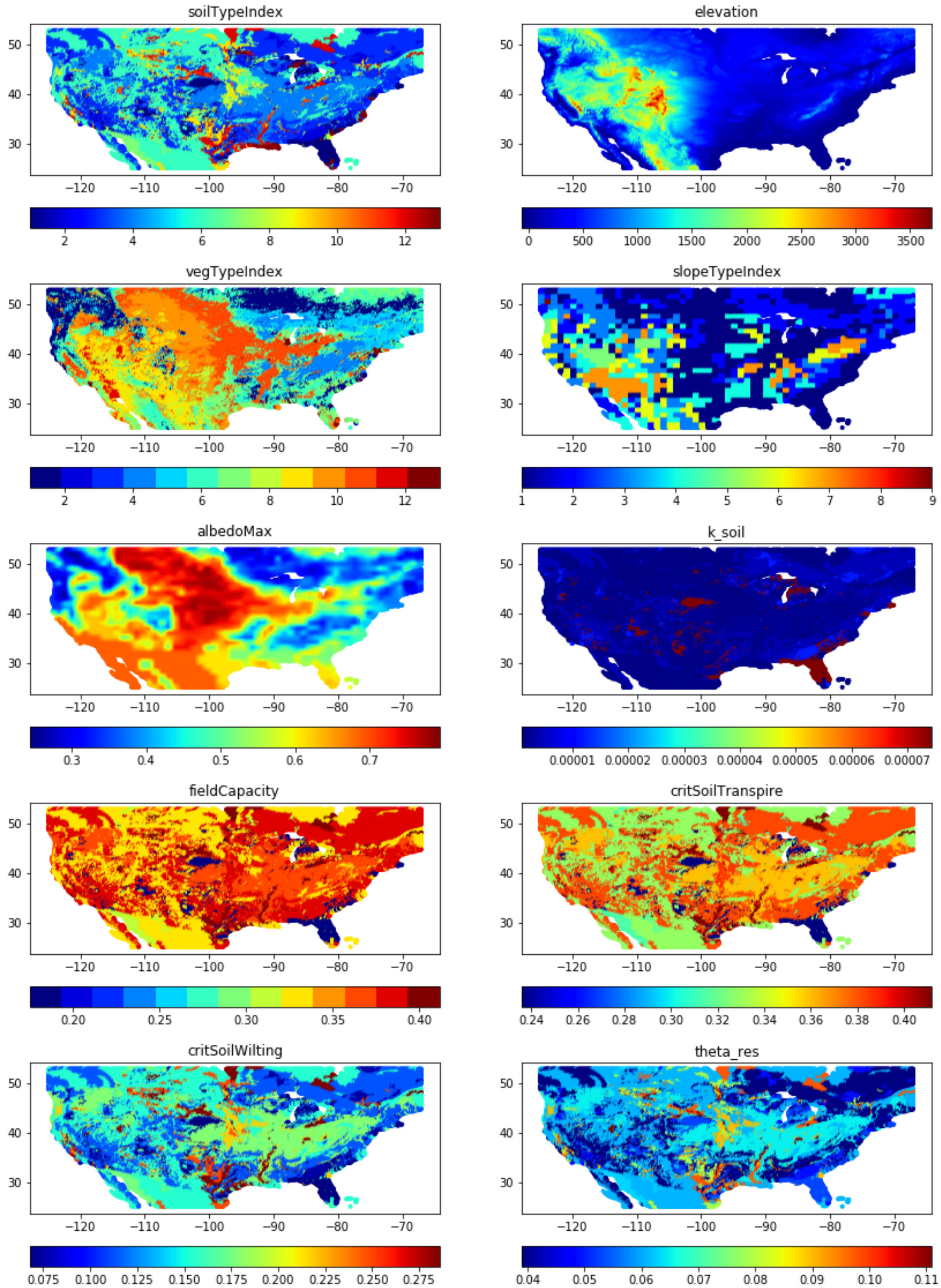

Figure 1. Selected SUMMA distributed attributes and parameters over CONUS

2.1.3 Model decisions

SUMMA offers flexibility in selecting different modeling decisions from a list of options to identify process representation, numerical implementation and parameterizations of thermodynamic and hydrologic fluxes. With these choices, modelers are able to test different combinations of modeling parameterizations and assess their impact on model behavior and output configurations. For SUMMA implementation, we have selected a set of model decisions depending on data availability and computational efficiency. For instance, STATSGO and MODIS 20-category datasets are selected to obtain soil and vegetation features. Among different choices, we have employed iterative numerical scheme and an analytical approach to calculate flux derivatives. To obtain snow and leaf area index, we select a scheme that allows for LAI calibration with a fixed seasonal pattern. The vertical distribution of soil water was implemented based on mixed form of Richard's equation (Celia et al., 1990) and groundwater parameterization was defined using the big bucket concept, which has a non-linear baseflow curve that is not dissimilar to that used in the VIC model. Our model simulations were based on a constant hydraulic conductivity along the depth of the soil instead of using the power law profile. A more simplied Raupach (Raupach, 1994) equation is employed for parameterizing vegetation roughness length and displacement height. Canopy emissivity and snow interception was implemented as a function of diffuse transmissivity and inverse function of new snow density, respectively. Albedo was represented by a constant decay rate and Beer's Law (Mahat and Tarboton, 2012) was chosen as canopy shortwave radiation method. The thermal conductivity of snow was represented by Jordan (1991) equation and soil conductivity was considered a function of soil wetness.

2.1.5 Model output specification

SUMMA allows for flexible output configuration, specified in an output control file. Users can specify not only the flux and state variables (and parameters) that can be output, but their frequency and a small number of statistical operations that can be applied (such as the maximum value over a period). For this work, a number of output variable configurations were developed. These include primarily a full process description profile that provides timestep-level states and fluxes sufficient to determine the full energy and water balance of different model spaces (canopy, snow, soil, aquifer); and a reduced 'water balance' profile that provides major states and fluxes at a daily time-step, as well as basin (GRU) runoff for routing. The latter configuration is used for model calibration, evaluation and forecasting, due to its efficiency (small storage footprint as well as increased model simulation speed).

During the course of this and related projects (see Discussion), the decision was made to use a hillslope runoff delay function within SUMMA at the basin (GRU) level, thus basin runoff is a key output variable. This allows for the use of a SUMMA model configuration that may have varying HRU (subbasin) choices to support the same routing spatial configurations; and also allows the model runoff lateral flow timing to be calibrated jointly with other model parameters, versus calibrating GRU scale timing within the channel routing model.

2.1.6 Model development

The work of this project and the related S&T SUMMA applications projects were challenged by a number of model code and implementation issues throughout the project, necessitating a number of interventions by the project's NCAR PI to upgrade the model capabilities and correct software bugs. Some of these upgrades led to the release of SUMMA version 3.0, due to the reconfiguration of SUMMA inputs to facilitate their use in the forecasting and model calibration context. Key changes and bug fixes are documented in the SUMMA repository, and some of the key changes or fixes include the following:

- The ability to provide exact restarts for modelled basin runoff, which is essential for forecasting and other applications
- Addition of GRU and HRU information to model outputs to facilitate split domain runs
- Bug fixes to correct the calculation of model runoff (several bugs)
- Revision of internal model balance checks to fix errors and enable reliable simulation completion, reducing convergence difficulties
- Improved documentation and messaging on model failures (to facilitate diagnosis)
- Reconfiguration of input model configuration and decisions files to facilitate templating usage in a forecast workflow
- Ability to use different data types for hruIds to enable use with HUC12 domains

MizuRoute model code changes were also implemented on specifications from this (and related) projects to facilitate templating of control files, better handling of different input formats, and more intuitive specification of time parameters related to restarts (forecast initialization).

2.1.6 Model calibration

Because SUMMA had not been applied for streamflow prediction is prior studies, significant effort was made to develop a model streamflow calibration strategy. In addition to exposing a range of trial parameters for users to manipulate in a distributed fashion, a smaller tractable set of parameters was identified to enable SUMMA to be optimized to produce streamflow of acceptable quality – i.e., a Kling Gupta Efficiency (KGE) value of greater than 0.7. This set of calibration parameters was chosen and refined over several months to impact key hydrologic processes, leading to a set of 13 parameters that are highlighted in bold in Table 2. These parameters affect infiltration, evapotranspiration and interception, soil water storage and transmission, snow accumulation and melt, aquifer baseflow generation, and hillslope runoff timing. The selection of a small but effective set of calibration parameters is critical because optimization algorithms perform best within a small parameter search space.

This project and the related SUMMA S&T projects adopted and developed workflows for SUMMA calibration using a multi-method general parameter optimization program called Ostrich (Matott et al., 2013). Ostrich has become an in-house capability being actively developed by Reclamation.[1] In Ostrich, the Dynamically Dimensioned Search (DDS) algorithm

[1] Ostrich is an example of technology transfer facilitated by S&T projects at NCAR led by the PI of this project, Andy Wood. Ostrich was investigated rom available online tools and adopted in late 2016 by PI Wood in the interest of moving beyond two optimization tools commonly used by NCAR and Reclamation (SCE and MOCOM) over the last decade or more. This effort was funded under the Coop Agreement PWP4 project. Dr. Wood later organized an Ostrich workshop in Spring 2018 at NCAR led by a key Ostrich and optimization method developer, Dr. Bryan Tolson (who was hosted for a visit to NCAR from the University of Waterloo), and Wood invited Reclamation personnel to participate in the workshop. Following the workshop, Dr. Wood worked with Reclamation contacts to apply Ostrich for the VIC model, and hosted various Reclamation personnel accounts on NCAR high performance computing systems to facilitate the interaction. These interactions helped to spur Reclamation interest in Ostrich optimization techniques, which later led to other Ostrich applications in Reclamation and Ostrich code maintenance and development being overtaken by Reclamation (https://github.com/usbr/ostrich). Dr. Wood also hired Dr. Hongli

(Tolson & Shoemaker, 2007) was used to adjust model parameters. A shell script workflow for applying Ostrich-DDS was written to leverage high-performance computing on the NCAR Cheyenne system. The workflow sets initial parameter ranges based on the SUMMA local parameter range specifications, runs a split domain simulation in which parameter multipliers are calibrated, merges the outputs and runs the mizuRoute model to produce streamflow, and assesses performance. A number of supporting scripts for generating initial parameter limit files, updating parameters in text and netcdf input files, and assessing performance, were written to support the workflow. Aside from Andy Wood, other contributors have included Manabendra Saharia (funded under PWP4), and Hongli Liu (funded under S&T 178). This model calibration workflow has now been adopted and incorporated into SUMMA projects at the University of Washington and University of Saskatchewan, and will be used by the SUMMA community as it grows.

2.2 Meteorological model forcings with GMET and MetSim

The Gridded Meteorological Ensemble Tool (GMET) methodology is based on multiple logistic and linear regression using static geophysical attributes to predict precipitation and temperature across a grid (Clark and Slater, 2006; Newman, 2015). Regression errors are used to condition spatially correlated Gaussian random fields for ensemble generation. The spatial regression approach for interpolating situ meteorological observations uses spatially distributed information as predictor fields in an ordinary least squares (OLS) linear regression to explain the spatial distribution of point in situ observations. In this project's application of GMET, the spatial predictors are static geophysical attributes (slope, elevation, latitude, and longitude). The regression was applied to predict daily precipitation, mean temperature, and diurnal temperature range (DTR) for each target grid cell, on each day, based on the current observed values of those variables within a sample from the 30 nearest meteorological stations and given their relationship to the local terrain features at the station locations. This strategy generated dynamic (time-varying) uncertainty estimates that are driven by daily observed meteorological conditions.

In support of this project and related SUMMA modeling projects funded by S&T, GMET was applied for the period 1970 to present at both 1/8th and 1/16th degree resolutions, yielding daily precipitation and temperature minima and maxima. The resulting outputs were spatially remapped to SUMMA HUC12 modeling fabric and then disaggregated to 3-hourly time resolution and to a full set of meterological fields (including radiation, pressure, humidity, wind variables) using MetSim (Bennett et al, 2020), a python-based wrapper for the MTCLIM program (Running et al, 1989).

This approach was applied to the entire western US, and the model forcing inputs for the URG domain were extracted from the large Reclamation-oriented forcing dataset. During the course of the project, significant bugs not only in SUMMA, but also in GMET and MetSim were detected and corrected, leading to 2 regenerations of the entire forcing dataset, as well as 1 recalibration of the entire URG SUMMA model and its ESP hindcasts.

Liu, a student of Dr. Tolson, as a post-doc working on S&T Project 178; she helped develop the Ostrich-based approaches now being used with SUMMA at NCAR, and the Universities of Washington and Saskatchewan.

2.3 Ensemble Streamflow Prediction for Water Supply

2.3.1 Forecast Ensemble Generation

This study is one of the first applications of SUMMA and mizuRoute that is specifically calibrated to generate streamflow and be used in a forecasting and data assimilation context. A number of run workflows were used, as illustrated in Figure Y. The standard ESP forecasting process involves running a 1 year spin-up simulation, beginning in a dry part of the year (often October 1) prior to the first forecast date, and saving a model state at the end of the timestep preceding the forecast. If no data assimilation is performed (termed an 'open loop' simulation), it may be possible to save these states in the course of a continuous model run, depending on the model capabilities. In SUMMA, the state is saved at the end of the run and used to restart SUMMA to run until the next desired state date. For this application, split domain runs were made, meaning that the all of the individual GRU simulations were run in parallel, and the state and output files were merged back together. MizuRoute runs to generate the associated streamflows were also conducted using the SUMMA runoff, and mizuRoute state files were also generated on the forecast dates. The ensemble forecast runs were for a 365-day forecast horizon, with 40 members associated with the meteorological sequences from 1979-2018. These were also run with a split domain, merged, routed, and the resulting streamflow was post-processed into NetCDF and CSV table formats holding daily forecasted flows for analysis (i.e., aggregation into the April-July inflow volumes). For each forecast ensemble, the meteorological trace year matching the initialization year was removed before the forecast verification to avoid including a 'perfect forecast' in the sense that using observed meteorology in the forecast period to drive SUMMA. All of the forecasting and data assimilation steps were run using bash scripts with automatic submission of model runs to the Cheyenne high performance computing cluster.

2.3.2 Forecast Ensemble Post-processing

Ensemble streamflow forecasts contain biases arising from a number of sources, including any systematic biases from a mis-match in forecast meteorology and model calibration meteorology, and model bias. Bias is different from random errors that lead to uncertainty in a forecast, and instead represents a persistent tendency for the model errors in a particular tendency to center on a value other than zero. Model biases are inevitable because the model is an abstraction of the real world, may not account for all processes (such as groundwater extraction or channel loss), and the impossibility of calibrating the model for all outcomes of interest. Often, for instance, models are calibrated to minimize errors in daily flows overall, but this does not guarantee that flows over an aggregated period such the April-July runoff period. Forecasts are also affected by random initialization error, which can arise from a combination of model bias and random errors (random and systematic) in real-time meteorological forcings – especially since the real-time meteorological observing network is degraded and provisional relative to more quality controlled and filled historical period observations.

In this project, we applied two techniques to correct for model bias and random initialization errors. For the first, we used the retrospective simulations from the overlapping period of the retrospective simulation and observed flows to calculate rolling mean biases (equal to the simulated flow divided by the observed flow) for each day of the calendar year (DOY). These were calculated over time windows of a specified length, and after sensitivity testing a window length of 21 days was used, centered on each DOY. The biases, termed bias correction factors, or BC factors, are then applied to the daily values of the ensemble forecast flows. Because the same meteorological sequences used to drive the retrospective model simulation are used in the ESP forecast technique, we calculate the BC factors for the retrospective period multiple times, leaving out two years year in turn from each calculation. When the BC factors are applied to an ESP trace, the BC factor calculated leaving out the meteorological year from that trace and the following year is used, so that the bias correction cannot use information that is known only in hindsight. An example of the family of resulting factors is shown in **Figure Z** (results section). The model bias correction is effective in correcting for seasonally varying bias tendencies in the model which would also be present in the ESP forecasts, such as a tendency to an early freshet or to low summer flows.

To correct for random initialization error, a blending approach (which is commonly used in the NWS River Forecast Centers) is applied to ensure that the starting streamflow forecast value is equal to the most recent observation. Streamflow simulation error (versus the observation) is calculated for each ensemble forecast member at the forecast start date and time, and subtracted from each forecast member. The error is linearly damped to zero after a user-specified length of time, and after that lead time the adjustment is zero, so that the blended forecast is equal to the raw forecast. It is difficult to establish one blending period that fits all forecast situations – i.e., an optimal blend length in a rising hydrograph may differ from one in a falling hydrograph – thus this modifying the blend length in real-time is a common practice in NWS RFCs. In an automated application such as hindcasting, as in this study, it was necessary to choose a single blend value, and 14 days was chosen. The blending step is performed after any bias-correction. More sophisticated approaches for such post-processing that combine both bias correction and blending (including auto-regressive error approaches) exist, but were beyond the scope of this project to investigate.

The overall ensemble forecast workflow, including these post-processing steps, is shown in **Figure Y**. The post-processing was applied in different permutations in this study, and various versions of the ensemble forecasts were provided: raw, bias-corrected, blended, and bias-corrected and blended. If the model is very well calibrated, the bias-correction step adds little to the forecast skill and can even degrade it due to its cross-validated application.

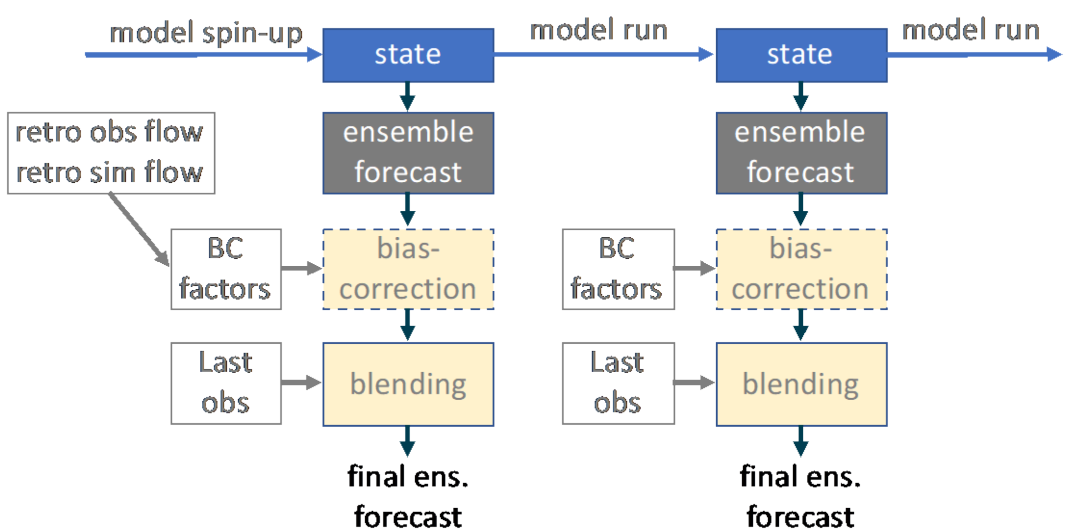

Figure 2. Schematic of the workflow for ensemble forecasting in this study

2.3.3 Real-time forecasts with SHARP

Although the hydrologic forecasting component of the project focused heavily on retrospective simulation and hindcasting, the project also endeavored to provide real-time ensemble predictions for the inflow locations developed for ensemble analysis. To this end, the modeling and forcing components described above are combined within a real-time workflow using the System for Hydromet Analysis Research and Prediction (SHARP), which uses a task scheduling system called EcFlow. EcFlow was developed by the European Centre for Medium Range Forecasting (ECMWF), where it runs global model weather forecasting and analysis suites (https://www.ecmwf.int/en/learning/training/introduction-ecmwf-job-scheduler-ecflow), and it is also used in parts of NOAA. SHARP currently runs in real-time on a small cluster computing system at NCAR (called Hydro-C1, which is partly funded by Reclamation and USACE).

SHARP had been developed under PWP4 and used to run real-time forecasts, but needed to be updated for new model versions and data stream changes, as well as to be able run more complex models over larger domain such as the URG basin. This project supported these upgrades, which enabled it to produce real-time seasonal ESPs for the URG and also to generate daily real-time daily 10-day deterministic forecasts for the Reclamation Forecast Shootout Competition[2]. To make real time predictions, model retrospective simulations must be connected with daily spinup simulations to generate the real-time model states needed for initializing forecasts. To drive these simulations, raw meteorological station data must be efficiently downloaded, quality controlled and gap filled, then used to create model forcings. The forcings drive SUMMA and mizuRoute model simulations to morning of the forecast, and generate a current watershed model state. From this state, ESP forecasts are generated and post-processed into the ensemble forecast product that can be used as input to a reservoir model (such as URGWOM).

For this project and others (including the Shootout), a 16-member 1/8th degree GMET forcings are created each morning, spanning the entire western US domain (enclosing Reclamation's service area). It is then further processed into HUC12-based SUMMA 3-hourly forcings for the 550-watershed Shootout cutout of the full domain and the URG basin, and used to run 16 spin-up simulations of the SUMMA and MizuRoute models. For the Shootout Competition, ensemble weather forecasts from ECMWF were also downloaded and processed into SUMMA forcings and used to generate 15-day streamflow forecasts. For the Shootout, the SUMMA forecasts were blended with the most recent observations, in a basic form of post-processing. In the seasonal ESP context for the upper Rio Grande, both bias-correction and blending were applied. An example of a hindcast series generated for a control point on the Rio Grande River is shown in **Figure 6** (results section).

[2] Due to lack of directed funding support to cover system and model upgrades and maintenance during the Shootout, NCAR entries into that competition were later discontinued.

2.3.4 ESP-based runoff shape analysis

A primary application of the NCAR SUMMA ESP hindcasts was to assess whether the new ESP forecasts coming online from the NWS WGRFC could be used as an alternative to the current analog-based practice of selecting flow sequences to use in disaggregating NRCS water supply volume forecasts. The proposed strategy is to use the ESP ensemble median or mean to provide the 'shape' or daily sequence of flows that would deliver the predicted volume. These sequences could be used to as input to URGWOM, supporting the development of the AOP and other analyses.

Such an analysis is feasible if sufficient ESP hindcasts are available, and to this end, NCAR generated a multi-decadal series of hindcasts and worked with Reclamation to collect and re-process multiple years of past AOP analog-based predictions for comparison. The skill of the two approaches was compared for spring forecasts using data from 2012 to 2019, for 20 locations in the URG above Otowi Bridge, NM.

2.4 Upper Rio Grande River Water Operations Model

The Upper Rio Grande Water Operations Model (URGWOM) was developed using the RiverWare software application created by the Center for Advanced Decision Support for Water and Environmental Systems (CADSWES) at the University of Colorado at Boulder. URGWOM was developed though years of interagency collaboration. URGWOM is a computational water operations model that uses the principals of mass conservation to simulate physical processes and policy operations. URGWOM simulates river and reservoir operations from the Colorado headwaters of the Rio Grande to Hudspeth County Texas. The physical processes simulated in URGWOM include hydrologic travel times, reservoir evaporation and seepage, conveyance losses to deep percolation, river channel evaporation, evapotranspiration by riparian and agricultural vegetation, surface water-groundwater interaction, and municipal wastewater and irrigation return flows. URGWOM consists of a policy set that simulates management operations throughout the basin including flood control, irrigation demands, transmountain diversions, the Rio Grande Compact, municipal, and industrial demands, Native American water rights, Endangered Species Act compliance, and recreational uses. URGWOM was developed with the intent to be capable of monitoring the multiple entities that are allocated within the system. URGWOM is designed to be used for a variety of applications including:

- Accounting Application - Daily timestep, data driven up to the current date accounting of native and trans-basin (San Juan – Chama Project) water in the system.
- Water Operations Application - Daily timestep rule-based Annual Operating Plan (AOP) runs that use both observed data and forecast data to simulate the remainder of the year.
- Planning Application - Daily or monthly timestep rule-based runs for planning projects.

Some information paraphrased from Volume 1: Physical Documentation
https://www.spa.usace.army.mil/Missions/Civil-Works/URGWOM/Documentation/

The Water Operations Application (i.e., AOP runs) uses a mixture of observed data and forecasted data to run. The AOP run is started on the last day of observational data and then uses the forecasted data to simulate the remainder of the year. Up until recently the only forecast data that was available was the NRCS bulk volume forecast. This provided a volume throughout the snowmelt season (usually March – July or April – September) at specific points needed for URGWOM to run. These NRCS forecasts are generally issued monthly starting January or February and going to around June (varies depending on how long runoff season is lasting). Reclamation, Army Corps of Engineers, and New Mexico Interstate Stream Commission use URGWOM and these forecasts to do at least monthly AOP runs from the start of the NRCS forecasts until about the end of irrigation season. The model will apply the volume forecast provided by the NRCS to a daily historical hydrograph selected by the closest historical year to the forecasted volume. The remainder of the year will use historical data to fill in the hydrograph. These forecasts help plan our decisions and how we operate until the end of the year. They also assist us in informing the public on what we expect river conditions and reservoir levels to be at throughout the year. In April, Reclamation has a large meeting with the public showing the results from the April AOP run and discussing what is planned. In addition to these they also assist in directing our releases from Elephant Butte to help maximize power generation, while still meeting downstream irrigation needs and avoiding high evaporation areas and cultural sites in Caballo. Because of this they are vital to Reclamation.

2.4.1 Major inflow locations

URGWOM has several dozen inflow points, of which some are minor and can be estimated in relation to other larger flow locations. **Table 3** summarizes these points according to the following key:

- Green – Must have for AOP run
- White – Will use historical data based on selected year to fill in data if there is no time series input
- Yellow – Same as white, except these points are significant for monsoon flow.

Table A1 in the Appendix summarizes addition local (intervening drainage) inflow points for URGWOM.

Table 3. URGWOM Inflow points

Gages/Points in URGWOM where Flows are Required					
Station Name	USGS ID	Flow Data Decimal Latitude	Flow Data Decimal Longitude	HUC	CO DWR ID
North Clear Creek Below Continental Reservoir	N/A	37.888	-107.204	N/A	NCLCONCO
Rio Grande at Thirty Mile Bridge near Creede	N/A	37.725	-107.256	N/A	RIOMILCO
South Fork Rio Grande at South Fork	N/A	37.659	-106.649	N/A	RIOSFKCO
Nortron Drain Near La Sauses	N/A	37.335	-105.771	N/A	NORDLSCO
Inflow into Platora Reservoir	N/A	N/A	N/A	N/A	N/A

Los Pinos River near Ortiz	N/A	36.982	-106.074	N/A	LOSORTCO
San Antonio River at Ortiz	N/A	36.993	-106.038	N/A	SANORTCO
Rio Grande near Lobatos CO	08249200	37.080	-105.760	13011002	RIOLOBCO
Rio Blanco Above Diversion	09343300	37.204	-106.812	14080101	RIOBLACO
Little Navajo Above Little Oso Diversion	09345200	37.077	-106.811	14080101	LITOSOCO
Navajo Above Oso Diversion	09244400	37.030	-106.737	14080101	NAVOSOCO
Embudo Creek at Dixon	08279000	36.211	-105.913	13020101	
Galisteo Creek below Galisteo Dam NM	08317950	35.465	-106.214	13020201	
Jemez River near Jemez NM	08324000	35.662	-106.743	13020202	
North Floodway Channel near Alameda NM	08329900	35.198	-106.600	13020203	
Rio Chama near La Puente/El Vado Local Inflow	08284100	36.663	-106.633	13020102	
Rio Puerco near Bernardo NM	08353000	34.409	-106.853	13020204	
Rio Pueblo de Taos below Los Cordovas NM	08276300	36.378	-105.668	13020101	
Red River below Fish Hatchery near Questa NM	08268820	36.683	-105.654	13020101	
South Diversion Channel above Tijeras Arroyo near Albuquerque NM	08330775	35.003	-106.657	13020203	
Tijeras Arroyo near Albuquerque NM	08330600	35.003	-106.648	13020203	
Willow Creek above Heron Reservoir near Los Ojos NM	08284200	36.743	-106.626	13020102	

Key:
Green – Must have for AOP run
White – Will use historical data based on selected year to fill in data if there is no time series input
Yellow – Same as white, except these points are significant for monsoon flow.

2.5 Climate predictability analysis

A second thrust of this proposal was planned to focus on west-wide climate and hydrologic predictability -- to examine the proposition that trends in climate may be eroding seasonal streamflow predictability, and to assess where and when sub-seasonal to seasonal climate forecasts may be able to offset this loss. There has been some evidence that streamflow predictability may be degrading due to warmer temperatures and consequent loss of snowpack, which is a major source of forecast skill in spring for seasonal runoff. An earlier study focused on the URG and Upper Colorado River basins found that seasonal climate model forecasts for temperature could improve statistical water supply forecasts (Lehner et al, 2017b).

To support this focus, SUMMA was implemented and calibrated for the CAMELS 761-basin dataset spanning CONUS (Addor et al, 2017; Newman et al, 2015), and the western US basins were subsetted for further analysis. The calibration workflow had to be updated to enable parallel automated calibration. An original calibration using NLDAS forcings was performed for the entire CAMELS dataset, and a second calibration based on GMET 1/16th degree western US forcings was also performed).

Two sources of climate information were developed. For sub-seasonal lead times (out to 35 days), the NCEP Global Ensemble Forecast System (GEFS) precipitation and temperature forecasts were downloaded and processed, and mapped to the western CAMELS basins. An initial skill analysis for the GEFS forecasts was conducted. The National Multi-Model Ensemble (NMME) datasets and real-time data streams, including mapping to HUC4 watershed units, that were created in a prior S2S project involving Reclamation's Sarah Baker were updated and maintained (at https://hydro.rap.ucar.edu/s2s/). The skill analyses for the NMME created earlier were not updated, and the data were not analyzed for this project, but the processed HUC4 datasets represent a resource for the community.

More effort was allocated toward the Rio Grande forecasting tasks in this project than was originally scoped, leading to a URG focus that extended through the entire project. As a result, the western-US focused predictability analyses only reached an initial stage.

3. Results

3.1 ESP model implementation, calibration and hindcasting

For the URG domain, over 20 locations were implemented for SUMMA and mizuRoute, resulting in the domain shown in **Figure 4**. These were calibrated, from upstream headwater basins in Colorado to the downstream mainstem location at Otowi Bridge, NM (**Table 4**). The implementation of Ostrich for calibrating the URG presented a challenge not addressed in other calibration efforts for stand-alone basins. The 20 basins included many nested or downstream locations, such that after headwater basins were calibrated, their results needed to be nested into the calibration process for the downstream basins. An efficient strategy for doing so was devised, but this requirement added complexity to the calibration workflow.

The GMET and SUMMA model dataset generation had to be rebuilt several times during the course of the project due to the discovery and correction of bugs in the GMET and SUMMA software. The SUMMA model was recalibrated twice to accommodate code changes, with the final recalibration being completed in June 2021. Retrospective runs and hindcasts were regenerated on each occasion. Where possible, naturalized flows were obtained (such as for Otowi Bridge) for use in calibration and evaluation; otherwise observed flows from the USGS were used. An example of a calibration outcome for one location is shown in **Figure 5**.

The development of the model flow bias-correction step was strongly motivated by the interest in providing forecasted hydrograph shapes for combination with official volume forecasts as a possible replacement for the expert analog selection approach used currently. Because ESP forecast models are likely to have some degree of model bias, this could bias the shapes derived from this strategy. Python-based codes for bias correction and forecast blending were developed. **Figure 3** illustrates the impact of a correction for site with a strong over-simulation bias during the peak runoff season.

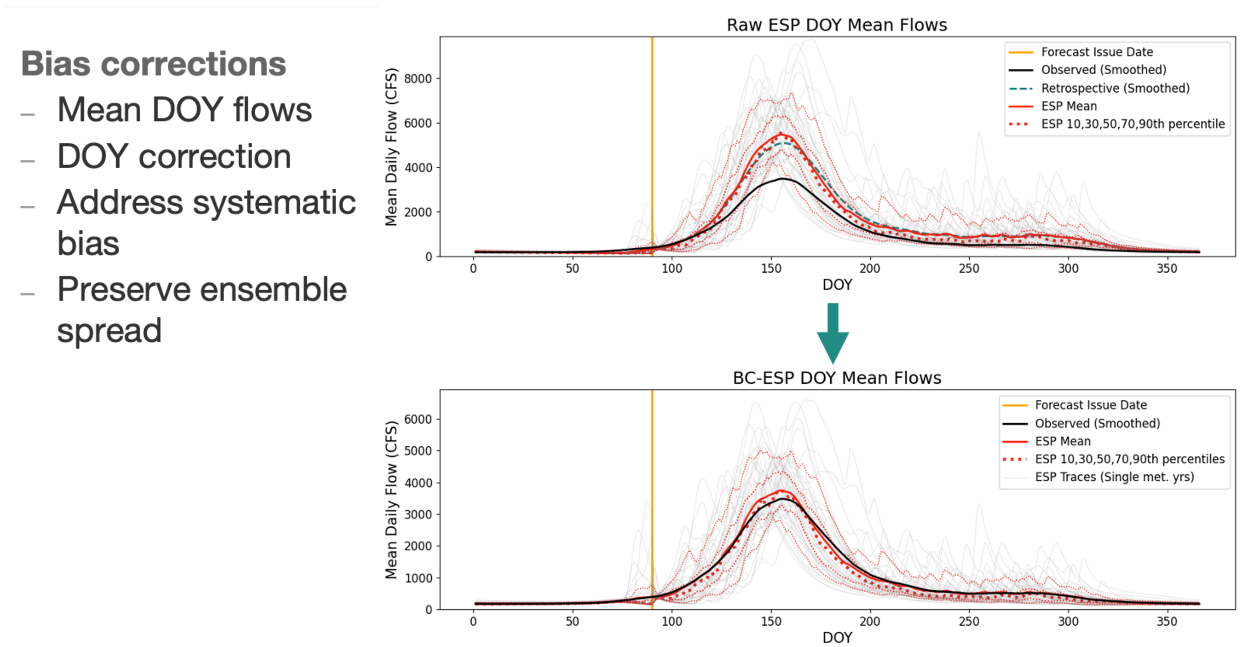

Figure 3. Illustration of the impact of bias correction on ESP forecasts.

Table 4. List of basin locations calibrated

ID #	USGS Name	Gage ID	URGWOM Name	State	Latitude	Longitude	HUC12 ID	Reach ID	# HUC12 Basins
1	Rio Grande at Thirtymile Bridge nr Creede	08213500	ThirtyMileBridge.Gage Inflow	CO	37.725	-107.255	130100010106	75000097	6
2	Rio Grande at Wagon Wheel Gap	08217500	WagonWheelGap.Gage Inflow	CO	37.767	-106.831	130100011006	75000041	35
3	South Fork Rio Grande at South Fork	08219500	SouthFork.Gage Inflow	CO	37.657	-106.649	130100011106	75000067	6
4	Rio Grande nr Del Norte	08220000	DelNorte.Gage Inflow	CO	37.689	-106.460	130100011306	75000030	52
5	Platoro Reservoir Inflow	08245000	Platoro.Inflow	CO	37.355	-106.544	130100050109	75000344	9
6	Conejos River nr Mogote	08246500	Mogote.Gage Inflow	CO	37.054	-106.187	130100050405	75000461	13
7	San Antonio nr Ortiz	08247500	RioSanAntonioAtOrtiz.Gage Inflow	NM	36.993	-106.038	130100050302	75000498	2
8	Los Pinos nr Ortiz	08248000	RioLosPinosAtOrtiz.Gage Inflow	NM	36.982	-106.073	130100050204	75000495	4
9	Red River below fish hatchery nr Questa	08266820	RedRiverBlwFishHatchery.Gage Inflow	NM	36.683	-105.654	130201010304	75000499	4
10	Rio Pueblo De Taos bl Los Cordovas	08276300	RioPuebloDeTaosAtLosCordovas.Gage Inflow	NM	36.379	-105.668	130201010607	75000666	11
11	Embudo Ck at Dixon	08279000	EmbudoCreekAtDixon.Gage Inflow	NM	36.211	-105.914	130201010909	75000702	9

12	Rio Chama nr La Puente	08284100	ElVadoLocalInflow.Local Inflow	NM	36.663	-106.633	130201020402	75000466	20
13	Jemez R bl Jemez Canyon Dam	08329000	BlwJemez.Gage Inflow	NM	35.390	-106.535	130202020506	75000890	30
14	Jemez R nr Jemez	08324000	NrJemez.Gage Inflow	NM	35.662	-106.743	130202020403	75000906	15
15	Rio Grande at Otowi Bridge	08313000	Otowi.Gage Inflow	NM	35.875	-106.142	130201011304	75000870	385
16	Rio Blanco bl Blanco Diversion	09343300	RioBlanco.Inflow	CO	37.204	-106.812	140801010304	77019317	4
17	Navajo R at Oso Diversion	09344400	NavajoRiver.Inflow	CO	37.030	-106.738	140801010606	77019308	5
18	Pecos R ab Santa Rosa Lk	08382650	PecosAbvSantaRosa	NM	35.059	-104.761	130600011111	75001320	71
19	Willow Ck Reservoir Inflow	08216500	AzoteaWillow.Inflow 2	CO	37.856	-106.927	130100010702	75000128	2
20	San Luis Valley	N/A	N/A	CO	37.653	-105.700	130100030703	75000273	75
21	"Rio Grande near Lobatos CO	08251500	Lobatos.Gage Inflow	CO	37.079	-105.757	130100021102	75000460	144
22	"Rio Grande at Embudo NM	08279500	Embudo Rio Grande	NM	36.206	-105.964	130201011104	75000628	211
23	"Rio Chama near Chamita NM	08290000	Rio Chama	NM	36.074	-106.112	130201021604	75000638	83

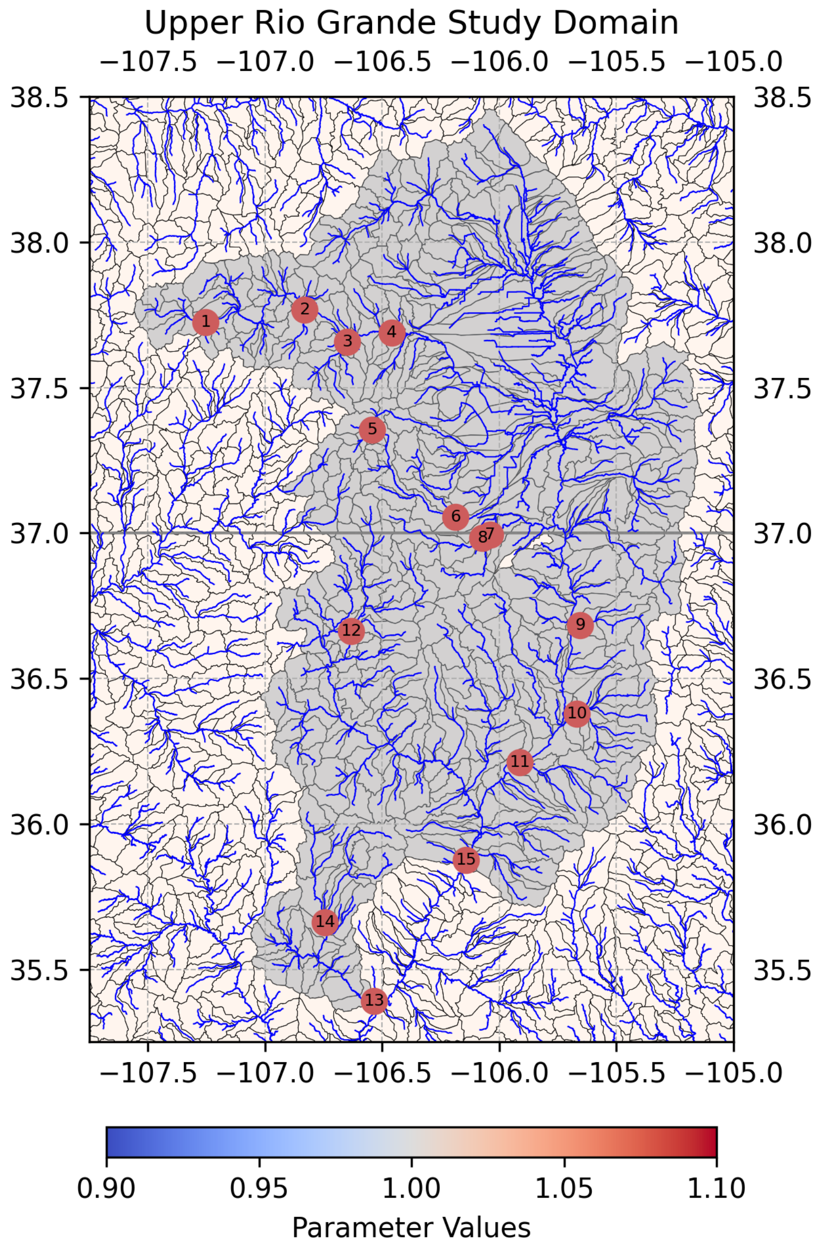

Figure 4. SUMMA and mizuRoute model configuration over the URG domain (upstream of Otowi Bridge, NM), with some control points (corresponding to Table 4) identified.

The San Luis Valley drainage from the eastern and northern headwaters is a closed basin, in which runoff is used for groundwater recharge, and then pumped back out into the main stem of the Rio Grande R. Representing this break in flow in the SUMMA and mizuRoute model sequence required implementing a new capability in mizuRoute, allowing for user specified abstraction and input of flow into the stream reach network.

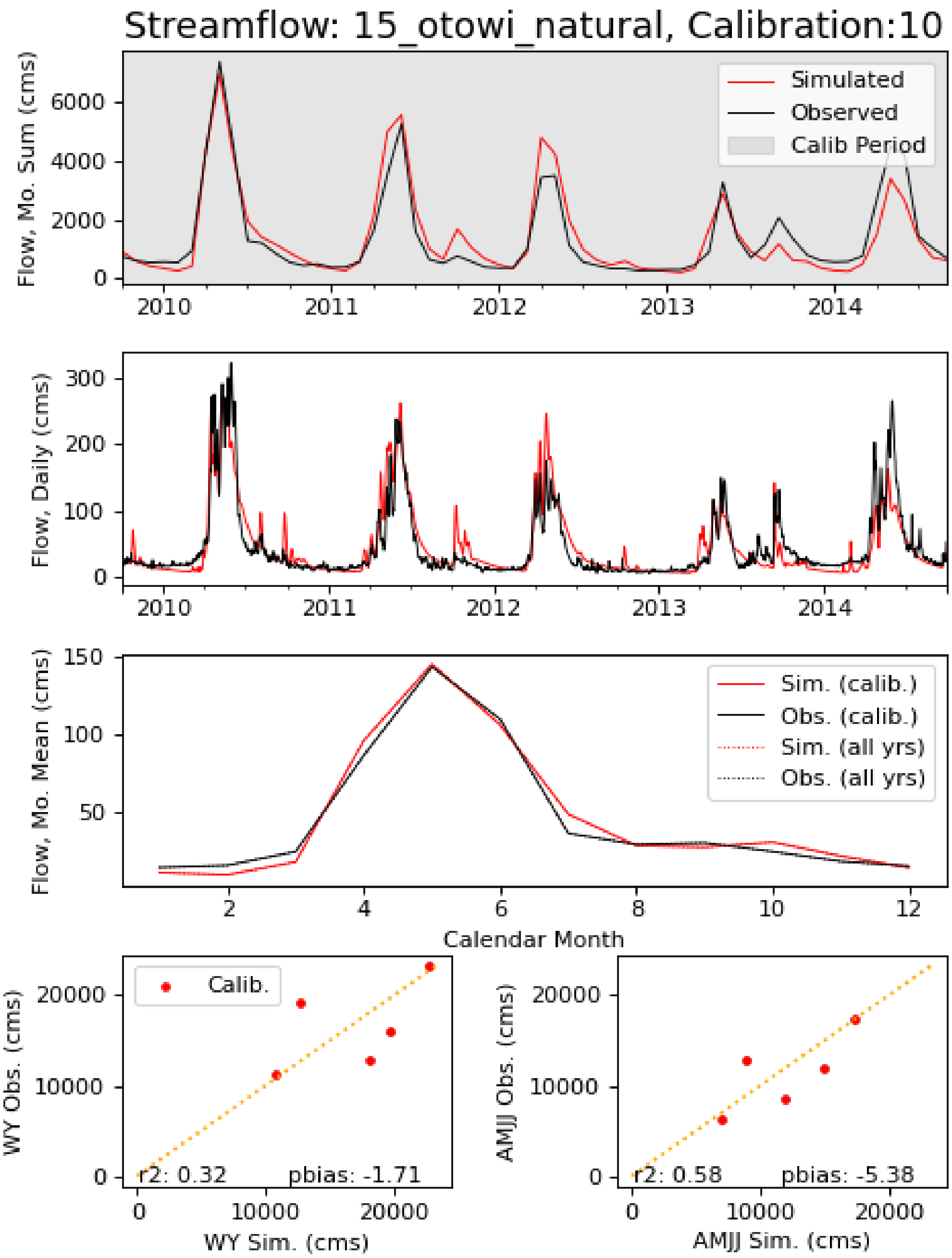

Figure 5. Example Calibration of the SUMMA and mizuRoute model configuration for Otowi Bridge, NM.

The SUMMA and mizuRoute configuration for the URG was used to produce multi-decadal series of ESP hindcasts (365 days long with 40 members) for all of the calibrated locations in the basin, for several spring initialization dates (March 1, April 1, May 1). An example of a hindcast series is shown in **Figure 6**.

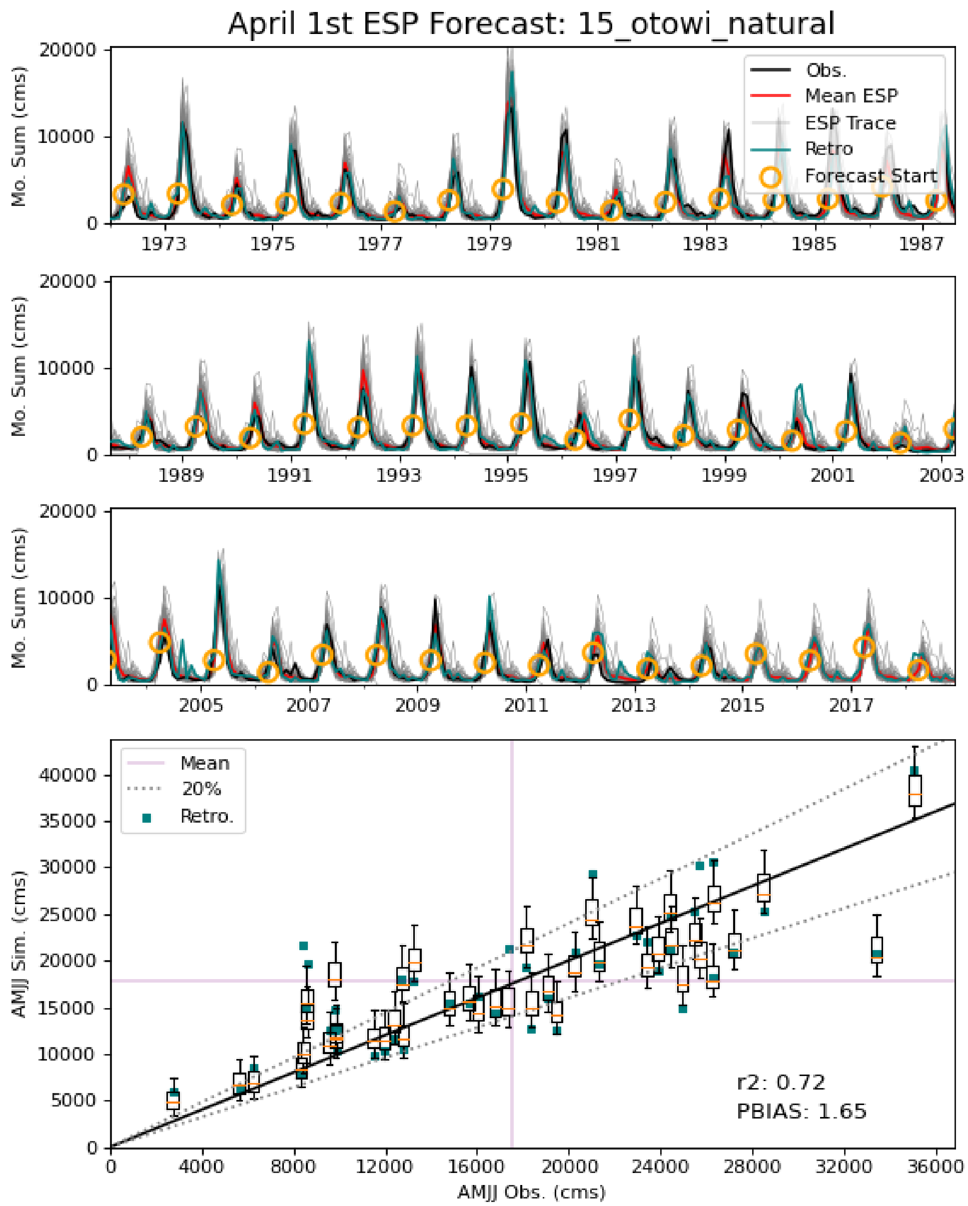

Figure 6. A hindcast ESP dataset for water supply (April-July) streamflow. Bias correction from a retrospective simulation is applied to reduce forecast bias. The hindcast series enables skill assessment (quantifying uncertainty) and provides background context for real-time predictions.

3.2 ESP-based hydrograph shape analysis to support AOP

The ESP-based shape analysis analyzed forecasts made by Reclamation and NCAR hindcasts between 2012-2019 for 20 locations in the URG basin. The KGE score was used as a metric of performance, measuring the daily agreement of the projected flow for April-July versus the AOP analog flow. Both were scaled to the total volume of the NRCS official forecast. An example of this analysis is shown in **Figure 7**. A histogram of skill scores across all locations and all spring forecast dates is shown in **Figure 8**.

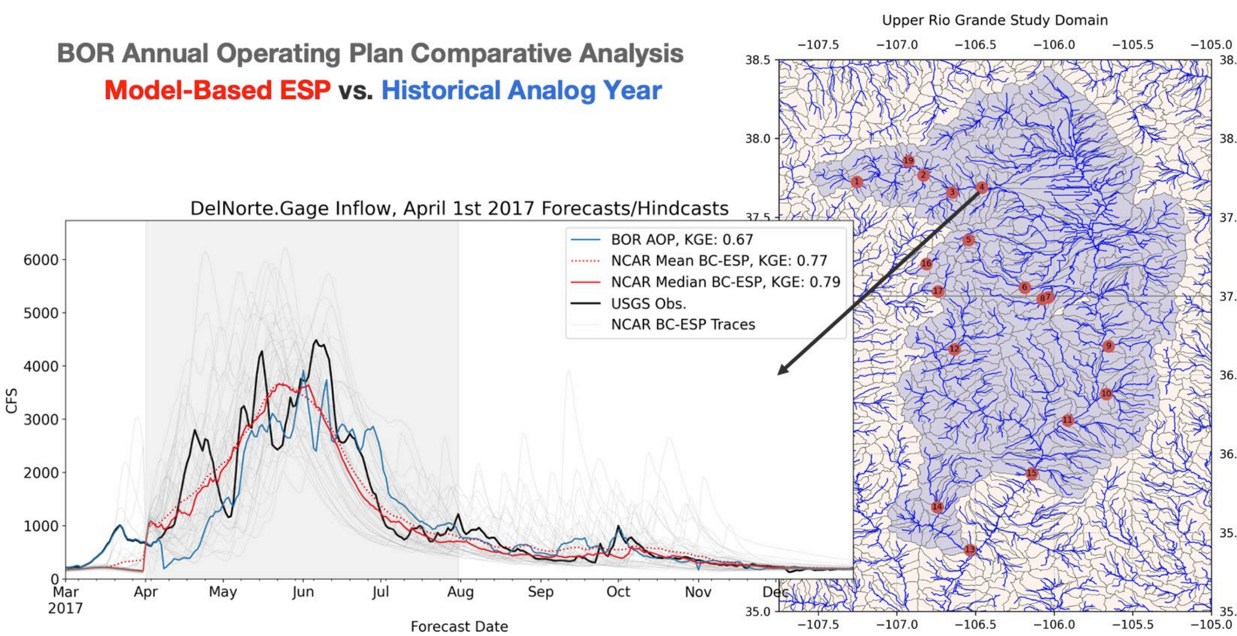

Figure 7. Composite plot showing the comparison of ESP projected daily inflow sequences versus Reclamation AOP inflow sequences, for the Del Norte gage.

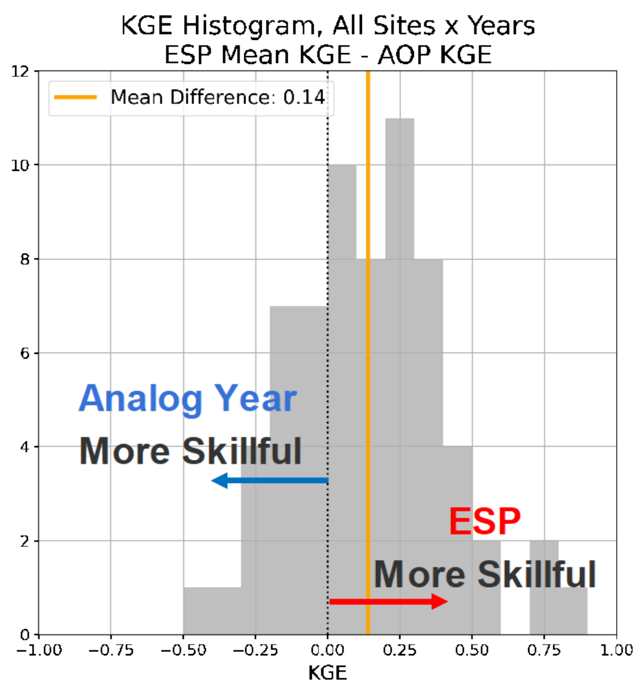

Figure 8. Summary of the comparison between analog and ESP based trace creation for URGWOM input locations.

Overall, the analysis demonstrated that the ESP-based approach to volume disaggregation is likely to be viable as an operational strategy for Reclamation, and that ESP-based sequences were more on average more skillful than analog-based sequences, from a daily error standpoint. We note that the ESP-based traces will have a different characteristic than the analogs, in that if the median or mean ESP sequence is used, it will be smoother than an observed sequence from an individual year. As a result, it will not over-prescribe (without skill) various runoff timings throughout the season, but equally it will not provide a realistic variability from week to week (because it smooths out such unpredictable variations). Depending on the quality of the model in different locations and the nature of the inflow point (e.g., into a reservoir or into the mainstem), it may be the preferable strategy to adopt.

At a presentation to Reclamation operators and URGWOM experts, it was noted that some locations in the basin have a bimodal runoff pattern due to both natural causes (such as different snowmelt timing from different headwater areas) and management (e.g., groundwater driven irrigation returns). To that the extent these causes are represented in the model (i.e., the natural effects), they would likely show up in muted fashion in the ESP mean or median shapes, but the management driven effects would not. Analog sequences might identify this behavior better than ESP in some locations. The NCAR team subsequently spent time assessing whether this bi-model behavior could be identified in observed flow records, and did not find strong support for this theory in the data.

3.3 Real-time ESP hindcasting

In the spring of 2021, the SUMMA/mizuRoute URG models were implemented within the SHARP real-time system (run on the RAL Hydro-c1 compute cluster) and used to generate real-time ESP forecasts. Forecast shapes were extracted and scaled to match NRCS official forecast volumes, and formatted into tables suitable for input into URGWOM, and provided to Reclamation as a comparison point with official inputs from agency sources.

Unfortunately, errors were soon discovered in the SHARP implementation, which is a complex workflow requiring a range of detailed specifications and data stream connections needed for real-time / automated operation. An incorrect copy of the SUMMA parameters had been used. Although these were corrected and the forecasts regenerated, the corrected forecasts were completed after AOP planning for 2021 was finished.

We note that the original goals of the project did not include producing real-time ESP predictions, versus retrospective analysis, though it was of interest to AAO. The goal of the ESP implementation was to create a sandbox of forecast data similar to what would be provided by operational agencies (i.e., MBRFC) once their implementation of forecasting models (including recalibration) for ESP prediction is completed. It is rare that the agencies also generate hindcasts to support this type of analysis and exploration, thus the NCAR datasets complement the agency datasets and afford a means to assess different strategies for using official forecasts.

3.3 Climate predictability analysis

Although the climate predictability analysis intended in the project was not completed to the extent outlined in the tasks (due to extra effort devoted to the Rio Grande components), the initial stages did investigate GEFS-based predictability out to 35 days in the western US

CAMELS basins (**Figures 9**). Daily precipitation and temperature forecast skill were found primarily in the first 10 days of the forecast (and example is shown in **Figure 10**). It is likely that the datasets addressed in this project can help to offset some losses in hydrologic prediction skill arising from a loss of snowpack, but this question was not fully investigated. The CAMELs based datasets and models created in this project will be a valuable resource for further investigation.

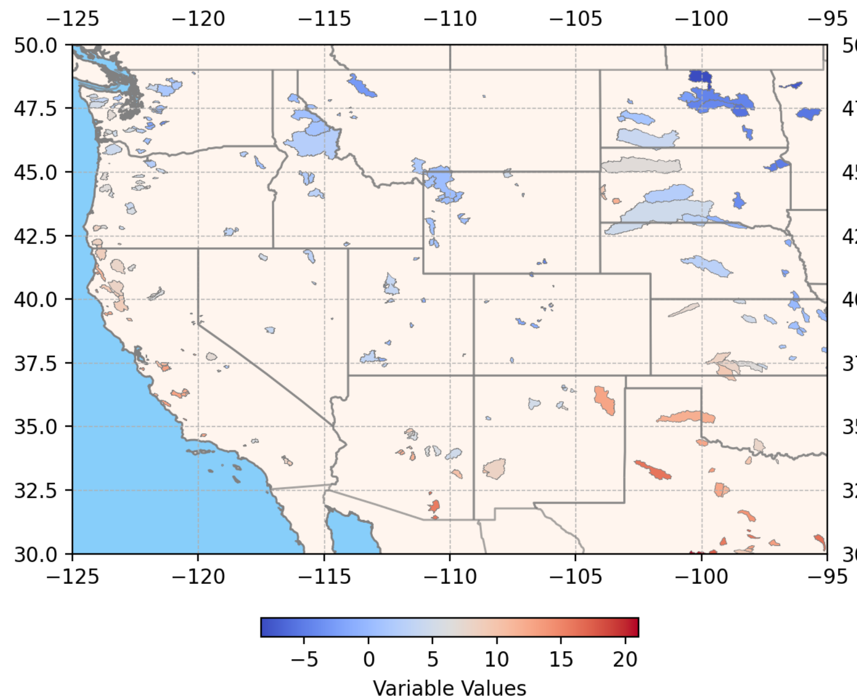

Figure 9. Western US CAMELS basins used for GEFS climate predictability analysis. The values shown are for the GEFS 1/16th degree remapped maximum temperature (Celsius).

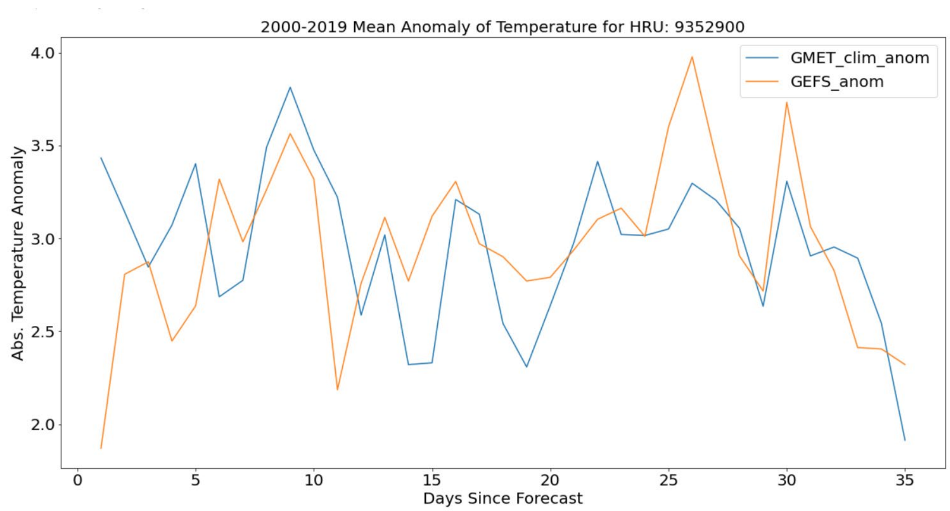

Figure 10. Example of a preliminary analysis of predictability for one CAMELS basin, comparing GEFS forecasted temperature anomalies versus GMET-based observations.

3.4 Partner interactions, presentations and publications

This project supported a range of partner interactions that were either briefed, participated directly in project meetings (which occurred either once monthly or every two weeks), or made use of the datasets created by the project.

- A current S&T project led by Erin Towler (NCAR) and Dagmar Llewellyn was given ESP hindcasts from this project
- Andy Wood provided feedback on proposals by WGRFC and RTI toward calibrating models and creating ESPs in the upper Rio Grande, and provided feedback on project results.
- The project results have been reviewed by, and are of interest to State of NM personnel and researchers focusing on seasonal prediction in the Rio Grande basin
- An S&T project led by Marketa McGuire and U. of Washington's Bart Nijssen was given the SUMMA models from this study to use in work on refining streamflow post-processing.
- The workflows and post-processing (bias correction and blending) routines of this project were used in S&T project 1881 (led by Jordan Lanini) and shared with TSC's Marketa McGuire.
- The study team interacted with Angus Goodbody, a hydrologic forecaster at NRCS, to facilitate NRCS evaluation of statistical forecast techniques with NMME early in the project. These interactions are continuing as of September 2021 with NRCS being interested in project methods for incorporating NMME temperature forecasts into volume predictions
- The project presented to the URGWOM Technical Group Meeting, March 9th, 2021
- The project presented work at national science meetings, including the American Geophysical Union and American Meteorological Society. Project elements have also been included in presentations to DOE/ORNL, the Canadian Global Water Futures project, internally at NCAR.
- The Canadian Global Water Futures project is using workflows for SUMMA and mizuRoute model implementation and calibration and forecasting.
- A publication on the use of ESP forecasts to disaggregate volume forecasts for use in water supply modeling is underway. Unfortunately, the departure of the project staff member from NCAR (Josh Sturtevant) led to the loss of the Google accounts that held the partially completed draft. It will be rewritten this fall, with the likely title:
  Sturtevant, J, AW Wood, L Barrett, D Llewellyn, F Lehner, 2021. Assessment of seasonal streamflow forecast strategies for Upper Rio Grande basin water management using SUMMA watershed modeling, AMS J. Hydromet. (target)

4. Discussion

This work scope described above contained elements that were achieved through the synergistic efforts of this project together with other Cooperative Agreement (CA) projects, including PWP4 and S&T projects 1881, 8116 and 178. The motivations for developing SUMMA and mizuRoute as a new modeling resource for Reclamation applications and research include its process-related flexibility, the superior coding standards (which enable organized and efficient development), the growing expert development and user community, and the model design, which separates numerics from process parameterizations and allows for an expanded degree of user control over the numerical solutions. The latter consideration is important because of the higher sophistication of the numerical solver, with adaptive time-stepping in all parts of the model space that enables relatively smoother hydrologic response behavior compared to other models, a characteristic that is important during model calibration. In many regards SUMMA represents an advance over the watershed models that have previously been available for applications research, and an important new community-based solution for hydrologic modeling to assess water security and water management strategies. Because SUMMA is new, however, and had not been used previously for such applications, its adoption has required the creation of an entirely new ecosystem of workflows for implementing, calibrating and developing SUMMA applications. Even more importantly, it has required a build-up of experience in SUMMA performance and the detection, diagnosis and correction of numerous model code software bugs to allow SUMMA to be run reliably over large domains. Almost none of this capacity existed at the start of the projects listed above, and these projects are credited with engendering this highly valuable outcome and resource for Reclamation and the community.

This project created an unusually detailed modeling and ESP prediction (hindcast) resource that has proven essential in understanding new strategies for water supply prediction in the Upper Rio Grande River basin. It was generated to have specific relevance to the URGWOM management model, and shared with other groups who are doing capability development research in the Rio Grande. It will likely have continued importance following this project.

5. Data Location

This project was instrumental in creating standard workflows for SUMMA modeling, calibration and forecasting, and python-based plots for visualization of SUMMA simulation and forecast results. These have become part of the SUMMA script ecosystem that is being used in multiple S&T projects as well as by collaborators in other institutions. They are, for now, housed in a private Github repository (https://github.com/NCAR/hydro_model_utils.git) that will be made public in the near future.

All data files from this project are archived in the form of tarred and gzipped files on an ftp site at: ftp://ftp.rap.ucar.edu/pub/andywood/SnT/8117/.

Core software used in the project, such as models, are contained in online repositories, including:

- github.com/NCAR/summa (public)
- github.com/NCAR/mizuRoute (public)
- github.com/NCAR/GMET (public)
- github.com/NCAR/SHARP (private)

- https://github.com/NCAR/hydro_model_utils (private)

References

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrol. Earth Syst. Sci., 21, 5293–5313, 2017.

Bennett, A. R., Hamman, J. J., and Nijssen, B.: MetSim v2.0.0: A flexible and extensible framework for the estimation and disaggregation of meteorological data, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2019-179, 2019.

Bunn, PTW, AW Wood, AJ Newman, H Chang, CL Castro, MP Clark and JR Arnold, 2021, Improving station-based ensemble surface meteorological analyses using numerical weather prediction: A case study of the Oroville Dam crisis precipitation event. (submitted JHM.)

Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D., Arnold, J.R., Gochis, D.J., Rasmussen, R.M., 2015a. A unified approach for process-based hydrologic modeling: 1. Modeling concept. Water Resour. Res. 51, 2498–2514. https://doi.org/10.1002/2015WR017198

Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Gochis, D.J., Rasmussen, R.M., Tarboton, D.G., Mahat, V., Flerchinger, G.N., Marks, D.G., 2015b. A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. Water Resour. Res. 51, 2515–2542. https://doi.org/10.1002/2015WR017200

Harpold, A.A., et al., 2016, JAWRA, JAWRA-16-0059-P

Huang, C, AJ Newman, MP Clark, AW Wood and X Zheng, 2016, Evaluation of snow data assimilation using the ensemble Kalman Filter for seasonal streamflow prediction in the Western United States, Hydrol. Earth Syst. Sci. 21, 635-650, 2017, http://www.hydrol-earth-syst-sci.net/21/635/2017/, doi:10.5194/hess-21-635-20173

Jordan, R., 1991. A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM. 89. Cold Regions Research and Engineering Lab Hanover NH.

Lehner, F., A. W. Wood, D. Llewellyn, D. B. Blatchford, A. G. Goodbody, and F. Pappenberger, 2017b: Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest. Geophys. Res. Lett., doi:10.1002/2017GL076043.

Lehner, F., E. R. Wahl, A. W. Wood, D. Blatchford, and D. Llewellyn, 2017a: Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective. Geophys. Res. Lett., doi:10.1002/2017GL073253.

Matott, L. S., Hymiak, B., Reslink, C., Baxter, C., & Aziz, S. (2013). Telescoping strategies for improved parameter estimation of environmental simulation models. *Computers & Geosciences*, *60*(October), 156–167. https://doi.org/10.1016/j.cageo.2013.07.023

Mendoza, PA, et al, 2017, An intercomparison of approaches for improving predictability in operational seasonal streamflow forecasting, HESSD

Mitchell, K.E., Lohmann, D., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Cosgrove, B.A., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Lettenmaier, D.P., Marshall, C.H., Entin, J.K., Pan, M., Shi, W., Koren, V., Meng, J.,

Ramsay, B.H., Bailey, A.A., 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. J. Geophys. Res. Atmos. 109. https://doi.org/10.1029/2003JD003823
Mizukami, N., Clark, M. P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., et al. (2016). mizuRoute version 1: a river network routing tool for a continental domain water resources applications. *Geoscientific Model Development*, *9*(6), 2223–2238. https://doi.org/10.5194/gmd-9-2223-2016
Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., and Arnold, J. R.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrol. Earth Syst. Sci., 19, 209–223, 2015.
Newman, AJ, MP Clark, J Craig, B Nijssen, AW Wood, E Gutmann, N Mizukami, L Brekke, and JR Arnold, 2015, Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States, J. Hydromet., doi: http://dx.doi.org/10.1175/JHM-D-15-0026.1.
Raupach, M.R., 1994. Simplified expressions for vegetation roughness length and zero-plane displacement as functions of canopy height and area index. Boundary-Layer Meteorol. 71, 211–216. https://doi.org/10.1007/BF00709229
Rosenberg, E. A., et al, 2011, WRR, 47, W00H14.
Running, S. W.; Nemani, R.; Hungerford, R. D. 1987. Extrapolation ofmeteorological data in mountain terrain, and its use for simulating forest evapotranspiration and photosynthesis. Canadian Journal of Forest Research. 17: 4 72-483.
Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, *43*(1), 1–16. https://doi.org/10.1029/2005WR004723
Woodhouse, C. A., et al, 2016, GRL, 43, 2174–2181.
Xia, Y., Ek, M., Wei, H., Meng, J., 2012a. Comparative analysis of relationships between NLDAS2 forcings and model outputs. Hydrol. Process. 26, 467–474. https://doi.org/10.1002/hyp.8240
Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., 2012b. Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. J. Geophys. Res. Atmos. 117.
Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., 2012c. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. J. Geophys. Res. Atmos. 117.

Appendix

Table A1. URGWOM local inflow points

Local Inflows Required in URGWOM													
From							To						
Station Name	USGS ID	Flow Data Decimal Latitude	Flow Data Decimal Longitude	HUC	CO DWR ID		Station Name	USGS ID	Flow Data Decimal Latitude	Flow Data Decimal Longitude	HUC	CO DWR ID
Rio Grande at Thirty Mile Bridge near Creede	N/A	37.725	-107.256	N/A	RIOMILCO		Rio Grande at Wagon Wheel Gap	N/A	37.767	-106.831	N/A	RIOWAGCO
Rio Grande at Wagon Wheel Gap	N/A	37.767	-106.831	N/A	RIOWAGCO		Rio Grande near Del Norte, CO	N/A	37.689	-106.460	N/A	RIODELCO
Rio Grande near Del Norte, CO	N/A	37.689	-106.460	N/A	RIODELCO		Rio Grande at Monte Vista	N/A	37.609	-106.149	N/A	RIOMONCO
Rio Grande at	N/A	37.609	-106.149	N/A	RIOMONCO		Rio Grande	N/A	37.481	-105.878	N/A	RIOALACO

Monte Vista						at Alamosa					
Rio Grande at Alamosa	N/A	37.481	-105.878	N/A	RIOALACO	Rio Grande above Trinchera Creek near La Sauses	N/A	37.316	-105.743	N/A	RIOTRICO
Conejos River Below Platoro Reservoir	N/A	37.355	-106.544	N/A	CONPLACO	Conejos River near Mogote	N/A	37.054	-106.187	N/A	CONMOGCO
San Antonio River at Ortiz	N/A	36.993	-106.038	N/A	SANORTCO	San Antonio River near Manassa	N/A	37.177	-105.878	N/A	SANMANCO
Conejos River near Mogote	N/A	37.054	-106.187	N/A	CONMOGCO	North Branch Conejos River near Conejos	N/A	N/A	N/A	N/A	NORCONCO
Rio Grande above Trinchera Creek near La Sauses	N/A	37.316	-105.743	N/A	RIOTRICO	Rio Grande near Lobatos CO	08249200	37.080	-105.760	13011002	RIOLOBCO
Rio Chama below El Vado	08285000	36.594	-106.733	13020102		Rio Chama above Abiquiu	8286500	36.319	-106.599	13020102	

Reservoir					
Rio Chama below Abiquiu Reservoir	08287000	36.237	-106.417	13020102	
Rio Grande near Lobatos CO	08249200	37.080	-105.760	13011002	RIOLOBCO
Rio Grande near Cerro NM	08263500	36.740	-105.683	13020101	
Rio Grande below Taos Junction Bridge near Taos	8276500	36.320	-105.754	13020101	
Rio Grande at Embudo	08279500	36.206	-105.964	13020101	
Rio Grande at Otowi Bridge	08313000	35.875	-106.142	13020101	

Reservoir					
Rio Chama near Chamita	08290000	36.074	-106.112	13020102	
Rio Grande near Cerro NM	08263500	36.740	-105.683	13020101	
Rio Grande below Taos Junction Bridge near Taos	8276500	36.320	-105.754	13020101	
Rio Grande at Embudo	08279500	36.206	-105.964	13020101	
Rio Grande at Otowi Bridge	08313000	35.875	-106.142	13020101	
Rio Grande above Cochiti	N/A	N/A	N/A	N/A	

Jemez River near Jemez NM	08324000	35.662	-106.743	13020202			Jemez River below Jemez Canyon Dam	08329000	35.390	-106.535	13020202	
Below Elephant Butte	8361000	33.149	-107.207	13030101			Rio Grande above Caballo	N/A	N/A	N/A	N/A	