

Title: Improving Hydrologic Predictions in Wyoming's Headwaters Through Detailed Quantification of Snowmelt

Start Date: 07/01/2021

End Date: 06/30/2024

Project Funds: \$186,269

Principal Investigator: Fabian Nippgen, Assistant Professor, Ecosystem Science and Management, University of Wyoming, fnippgen@uwyo.edu

Co-Investigator: Ginger Paige, Professor, Ecosystem Science and Management, University of Wyoming, gpaige@uwyo.edu

Objectives

This project aimed to improve hydrologic prediction and snowmelt representation in Wyoming's mountain headwaters by integrating high-resolution climate forcing data, field observations, and distributed modeling. The objectives for this work were: 1) assessing the impact of model resolution on modeled hydrologic response when the same watersheds are modeled within larger watersheds, 2) evaluating how differences among widely used climate forcing datasets—NLDAS, ERA5, and CONUS404—affect hydrologic simulations of snowmelt, soil moisture, evapotranspiration, and streamflow in complex snow-dominated watersheds, and 3) investigating the accuracy of modeled snow cover and depth at the small watershed scale by comparing them to UAV-derived snow cover, manually collected snow courses, and eddy covariance-derived sublimation from snow pack.

By focusing on the Snowy Range in southeastern Wyoming, a critical headwater region for the Laramie and North Platte River, this project sought to reduce uncertainty in water predictions and strengthen Wyoming's ability to anticipate and manage future changes in snow-derived water supply.

Methodology

Hydrometric Observations

Field activities and modeling efforts were centered in the Snowy Range of the Medicine Bow National Forest, encompassing several nested watersheds ranging from small first-order headwater basins to fourth-order streams. The study area includes two SNOTEL sites (Brooklyn Lake and Cinnabar Park), multiple soil moisture monitoring locations, and nine open-channel gauging stations. The Glacier Lakes Ecosystem Experiments Site (GLEES) eddy covariance tower provided continuous observations of latent heat flux (i.e., evapotranspiration), and meteorological conditions (Figure 1).

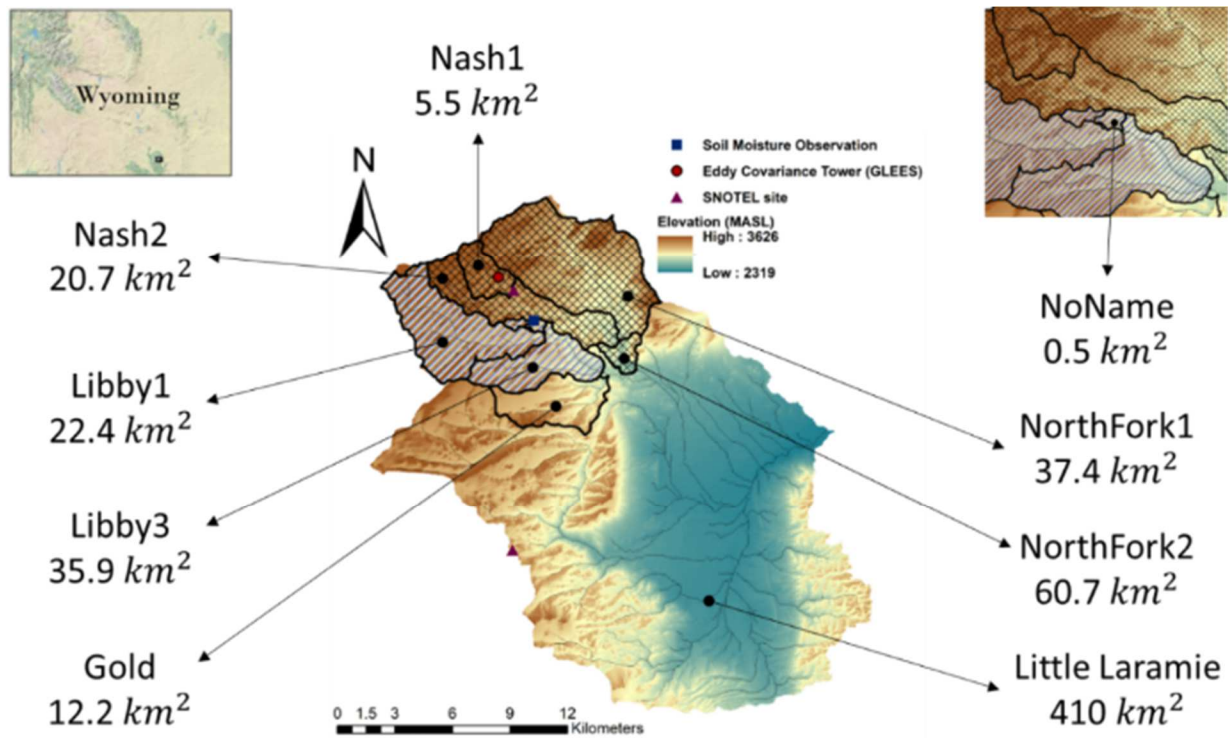


Figure 1: Study site with instrumentation. Several of the gauged watersheds are nested within larger gauged watersheds, allowing for assessment of scale effects in hydrologic models.

The datasets from these field locations were used to evaluate the accuracy and realism of modeled hydrologic processes.

Hydrologic Model

We used the WRF-Hydro modeling system (Gochis et al., 2020) to simulate the surface and subsurface hydrology of nested watersheds in the Snowy Range, southeastern Wyoming. WRF-Hydro is a distributed, physics-based model developed by the National Center for Atmospheric Research (NCAR) that integrates a land surface model with routing modules to represent all major components of the hydrologic cycle. The model includes three primary components: (1) the Noah-MP land surface model, (2) a terrain routing module for surface and subsurface flow, and (3) a channel and reservoir routing module that simulates streamflow and storage at high spatial and temporal resolution.

Although WRF-Hydro can be coupled with the Weather Research and Forecasting (WRF) atmospheric model, it was operated in standalone mode for this study, forced with gridded climate products. In this configuration, forcing datasets provide the atmospheric variables required to drive the land surface and routing modules, including precipitation,

temperature, wind speed, humidity, surface pressure, and shortwave and longwave radiation.

Because model components operate at different spatial resolutions, WRF-Hydro employs a multi-scale aggregation and disaggregation framework to maintain consistency across scales (Cho & Kim, 2022). Coarse-resolution climate data (3–12 km) were bilinearly interpolated to the smaller land surface model grid (100s of meters), while terrain and channel routing were implemented at even finer resolutions (tens of meters). Model outputs included streamflow, snow water equivalent (SWE), snow depth, soil moisture, and evapotranspiration, which were compared against in situ and remotely sensed observations. The most important inputs and outputs are summarized in Fig. 2.

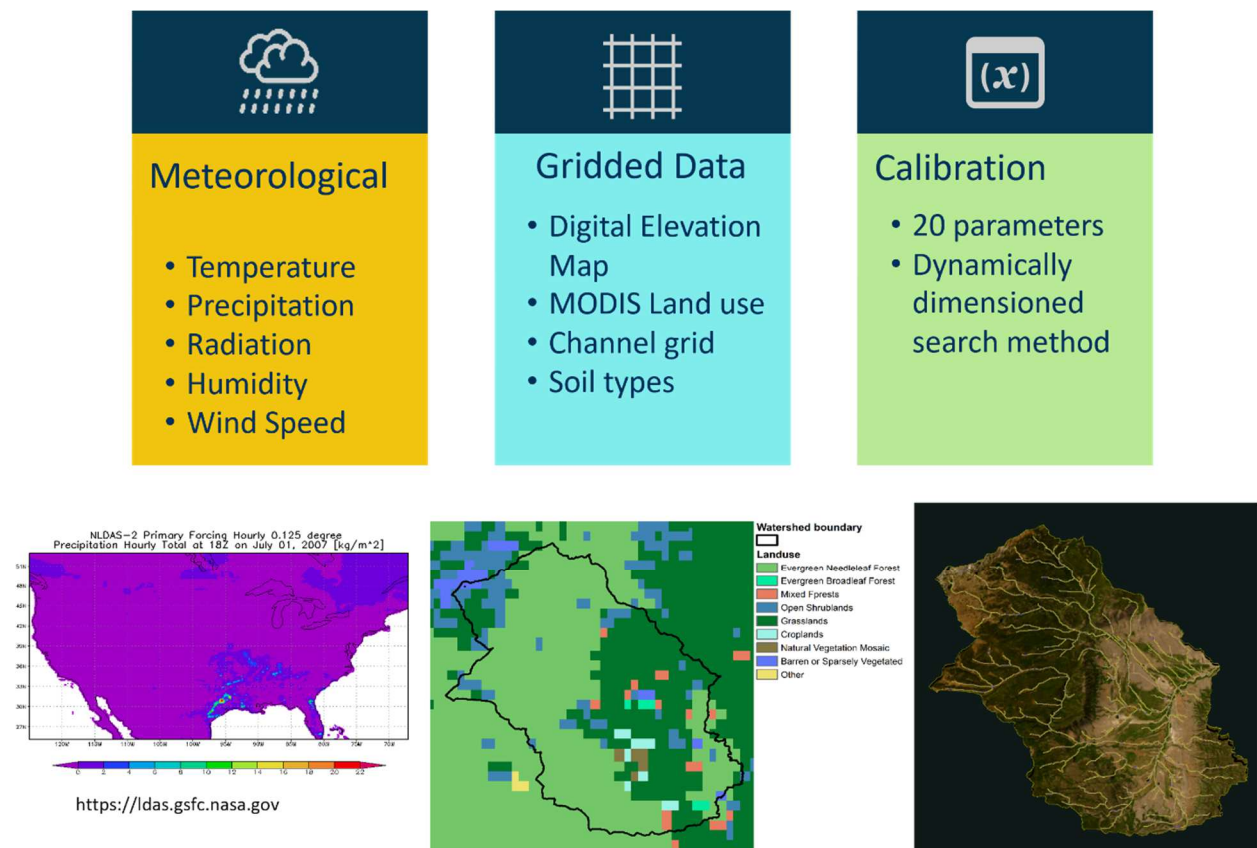


Figure 2: Summary of the most important input and outputs of the WRF-Hydro model (top row); Example of NLDAS input data, land use, and the stream network provided to WRF-Hydro (bottom row).

Simulations were conducted individually for water years 2018 and 2019, each preceded by a one-year spin-up to stabilize model water and energy storages. The MODIS Land Cover Type Product was used for land use representation, and the diffusion wave algorithm was applied for hourly channel routing. Precipitation was partitioned into rainfall and snowfall following a parsimonious temperature-based scheme (Jordan, 1991).

Model calibration was initially based on observed streamflow using the Dynamically Dimensioned Search (DDS) algorithm (Tolson & Shoemaker, 2007). Model performance was further evaluated using independent datasets, including SWE from the Brooklyn Lake and Cinnabar Park SNOTEL sites, soil moisture sensors within the study area, and evapotranspiration estimates from the Glacier Lakes Ecosystem Experiments Site (GLEES) eddy covariance tower. To assess model performance and water balance closure, we compared simulated and observed values across multiple nested watersheds using the Nash–Sutcliffe Efficiency (NSE) (Nash & Sutcliffe, 1970) and Kling–Gupta Efficiency (KGE) (Gupta et al., 2009) as performance metrics. This multi-criteria evaluation provided a comprehensive assessment of model behavior across scales and hydrologic variables.

Principal Findings

Objective 1: Scale Effects

For the impact of model resolution, we exclusively drove the model with the NASA NLDAS climate dataset. The grid size for LSM and terrain routing were proportional to the catchment size with three main resolutions (Table 1) that are representative of typical model resolutions used for watersheds of those sizes. For comparison, the National Water Model, also based on WRF-Hydro, features an LSM resolution of 1000 m. This setup allowed us to test how smaller watersheds are simulated on a standalone basis (i.e., small watershed optimized on small watershed streamflow) and when nested inside larger watersheds (small watershed simulated as part of large watershed with optimization on large watershed streamflow). The results from this section shed light on how well hydrologic models that are optimized at large scales (e.g., runoff from USGS gaging stations) and large model resolutions can replicate streamflow at the headwater scale. This becomes increasingly important when those models are used to, for example, simulate the effects of land use change in the headwaters or climate change induced alterations to the water cycle.

Table 1: Domain resolutions used in WRF-Hydro based on watershed size.

Area (km^2)	LSM grid (m)	Routing grid (m)	watersheds
1-10	100	10	NoName, Nash1
10-100	200	20	Nash2, Gold, NF1, NF2, Libby1, Libby3
100-1000	300	30	Little Laramie

Streamflow was generally simulated well for all watersheds with individual calibrations, i.e., where the model was tuned to observed streamflow (Fig. 3).

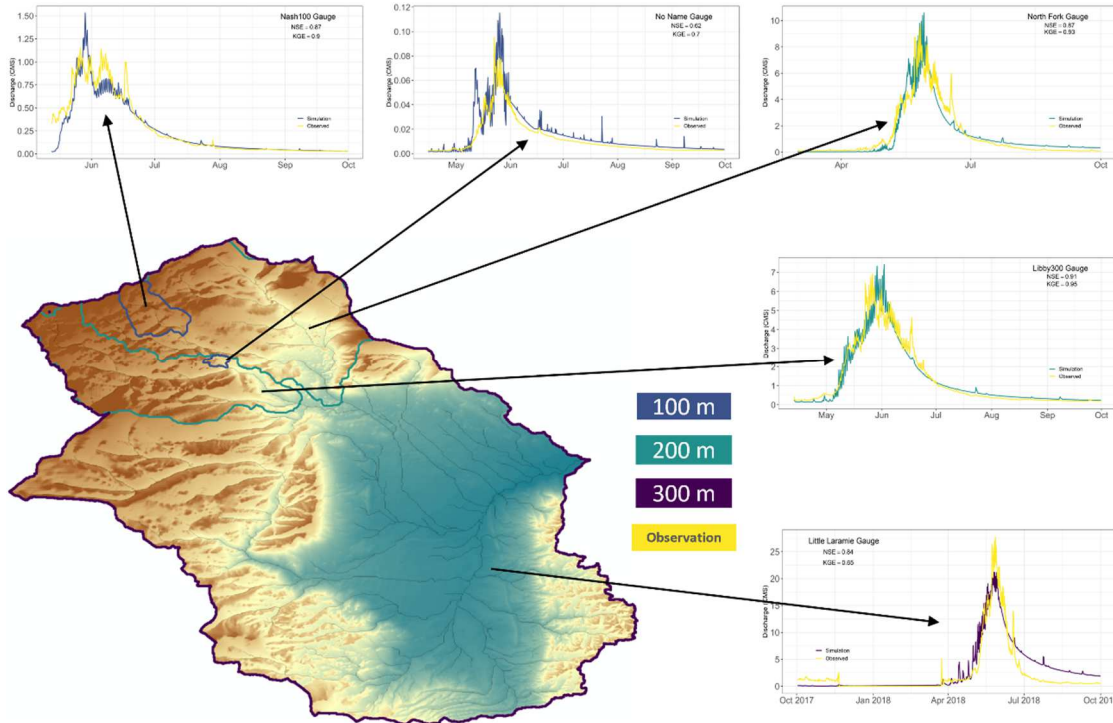


Figure 3: Simulated vs observed streamflow at five different watersheds representing all three scales. Blue colors denote the smallest watershed scale, green, the intermediate watersheds, and purple that largest scale (the Little Laramie River below Centennial, the only watershed at that scale).

Streamflow rise during the beginning of the melt period, peakflows, and the recession periods are simulated well for the smallest and intermediate watersheds, while baseflow is overestimated in the largest watershed. A different picture emerges when a smaller watershed is simulated as part of a larger watershed (Fig 4).

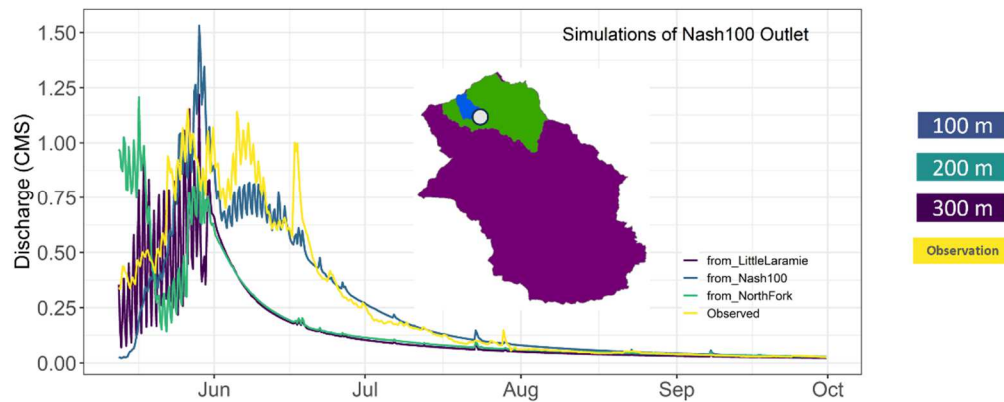


Figure 4: Observed and three modeled streamflow timeseries for the small Nash100 watershed. There is a noticeable decrease in simulation accuracy when the small watershed was not individually calibrated but rather simulated during a model run for a larger watershed.

While the Individual calibration of the small watershed on observed stream flow replicates the observations well, both timing and overall magnitudes are off when the small watershed is simulated as part of the larger watershed (i.e., at larger spatial resolution). This result is consistent for most cases when a smaller watershed is simulated as part of a larger watershed in our study area.

Soil moisture at different soil depths also differs between the simulations conducted at varying scales (Fig. 5). While the overall dynamics are represented well for the winter period and during snow melt, the summer drydown is overestimated for the simulation at the small and large scales, suggesting that those simulations retain too much water in the soil. However, it should be noted that soil moisture measurements happen at the point scale while the simulations integrate moisture over areas of at least 10,000 m².

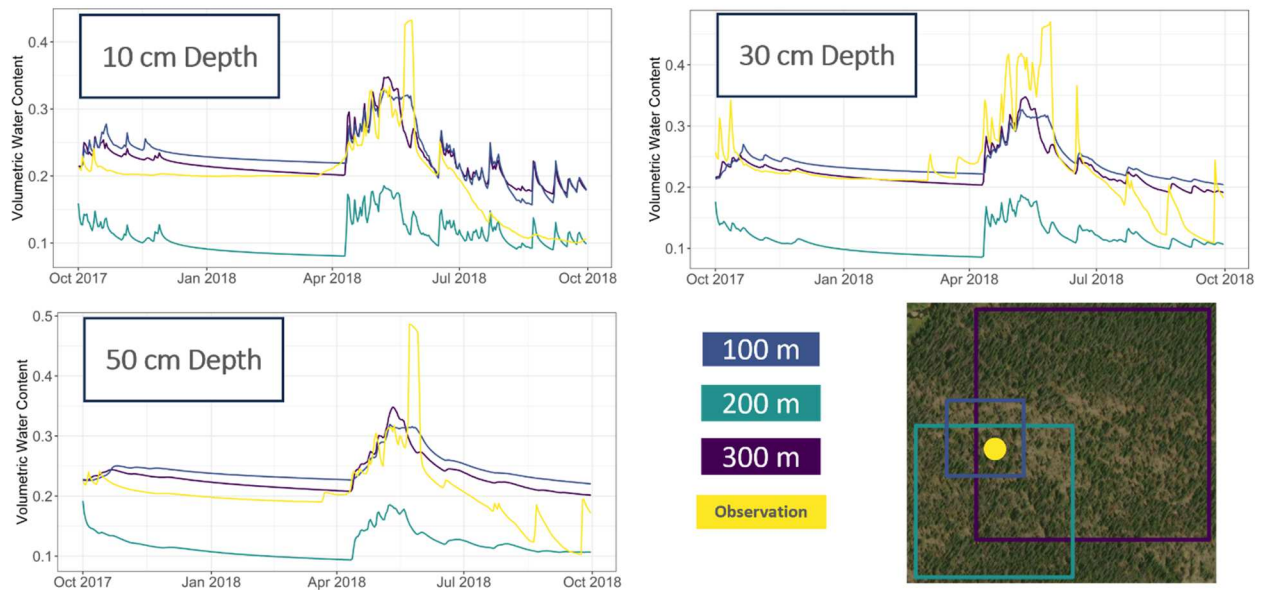


Figure 5: Observed soil moisture and simulated soil moisture for a location in the study area that is covered by model simulations of three different scales.

Evapotranspiration generally is underestimated at all three scales (Fig.6). Interestingly, the underestimation is largest for the small-scale simulation, which featured the best hydrograph fit. There is a clear difference between ET magnitudes during the winter months and in the summer, as demonstrated by a break in slope in the cumulative ET curve. The ET observed at the GLEES eddy covariance tower is sublimation, that is, water that is lost directly from the snowpack to the atmosphere without passing through the liquid phase, while the ET in the summer, starting in mid-May, is mostly transpiration from the trees that depletes the soil moisture storage. Comparing the slopes of cumulative ET suggests that the majority of ET underestimation happens during the winter.

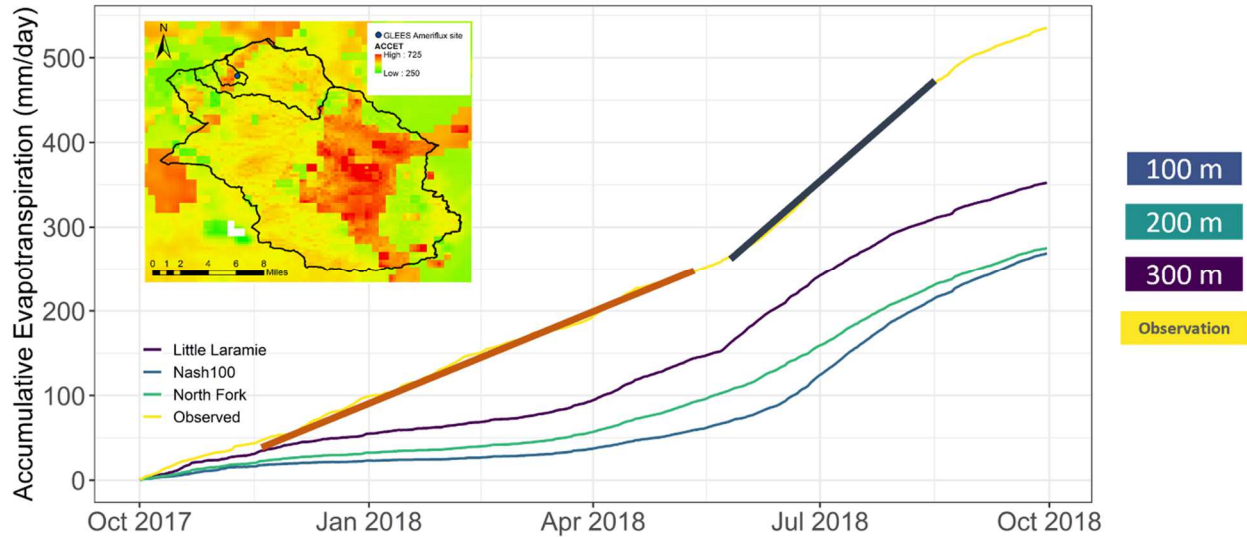


Figure 6: Figure 6. Cumulative ET at a range of scales across the modeled domains for the 2018 WY. The blue line is ET modeled inside the small watershed at a 100 m resolution, green for the medium-sized watershed at a 200 m grid size, and purple for the largest watershed with a 300 m resolution.

In summary, WRF-Hydro, and likely most hydrologic models, work well when calibrated on individual watersheds, but the simulation quality decreases when simulating smaller watersheds contained within the larger calibrated watershed. On average, simulated nested subwatersheds generate less runoff compared to the observed data and the simulated runoff from the calibrated subwatersheds. To compensate for the lack of runoff, more water is shifted to other components of the model, most notably evapotranspiration (Table 2).

Table 2: Water balance components (in mm) for three watersheds (different colored blocks) when simulated individually and as part of two larger watersheds. Decreased runoff (Q) is usually compensated for with increased evapotranspiration (ET). In each block, the first row is the watershed calibrated on its observed streamflow. The subsequent two entries are the watershed simulated as part of watersheds with increasing area.

		WY2018				WY2019			
Watershed	Resolution (m)	P	Q	ET	Δ SM	P	Q	ET	Δ SM
Nash	100	690	632	110	-44	956	707	207	35
	200	690	509	258	-59	957	545	362	36
	300	690	378	343	-65	956	445	389	47
NoName	100	672	472	253	-46	934	608	260	24
	200	673	505	215	-42	935	531	360	34
	300	673	164	377	-68	935	268	381	47
Libby	200	673	580	126	-48	942	753	125	25
	200	673	546	131	-47	942	570	279	32
	300	673	410	322	-66	942	464	375	45

This has implications for modeling studies that evaluate the effect of land use or climate change on water resources in headwater systems when the model is calibrated on a large watershed, e.g., a USGS stream gage. More research is necessary to quantify this effect in a targeted comparison study.

Objective 2: Forcing Data Comparison

We tested three different climate datasets as inputs to the WRF-Hydro model to evaluate the individual results in terms of streamflow response, soil moisture, evapotranspiration, and simulated snow cover extent. WRF-Hydro simulated the hydrologic response across the nine watersheds in the Snowy Range.

To assess the reliability of these datasets, modeled snow cover was compared against Sentinel-2 satellite imagery, using the Normalized Difference Snow Index (NDSI) to map observed snow extent. Differences between simulated and observed snow cover helped identify systematic biases in the forcing datasets, such as potential wind redistribution effects and precipitation underestimation.

Simulations revealed notable differences among the three climate datasets that directly influenced hydrologic simulation accuracy. All three forcing datasets result in streamflow hydrographs that represent the observed streamflow equally well; average efficiencies are largely similar between the three datasets for all study watersheds, though they differ slightly for individual watersheds (Table 3). The only outlier was the CONUS404 data, which did not perform well when used for the smallest watershed, NoName. It resulted in a simulation which is considered worse than the average annual streamflow (i.e., low Nash-Sutcliff model efficiency). Figure 7 highlights simulation results for the largest watershed, the Little Laramie.

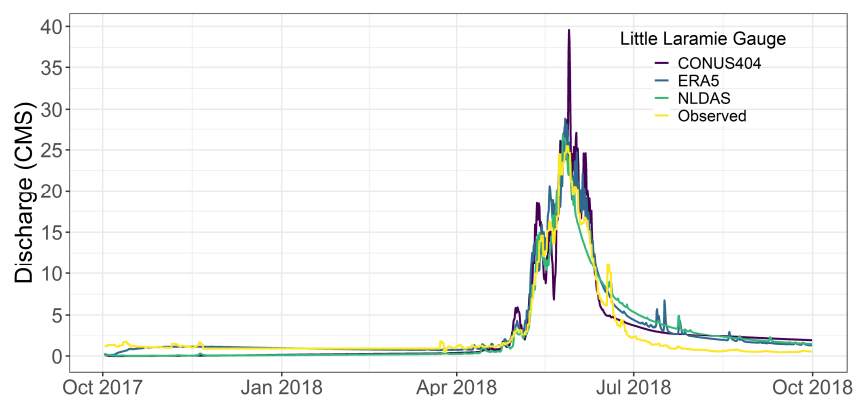


Figure 7: Streamflow simulated at the Little Laramie gauge for the three input datasets. All model simulations result in high efficiencies but differ in their representation of peakflows and the falling limb of the hydrograph.

The similar performance for streamflow simulations is even more striking as the input datasets feature very different precipitation values, with ERA5 being closest to the precipitation observed at the SNOTEL sites, and NLDAS underestimating precipitation the most (Fig. 8).

Table 3: Nash-Sutcliffe efficiencies for the nine study watersheds and the three input datasets. Efficiencies are generally high, with the exception of the smallest watershed with CONUS404.

Watershed	WY2018		
	NLDAS	ERA5	CONUS404
Little Laramie	0.90	0.93	0.86
North Fork 200	0.90	0.90	0.92
North Fork 100	0.89	0.77	0.84
Libby 300	0.85	0.77	0.90
Libby 100	0.69	0.90	0.90
Nash 200	0.86	0.87	0.57
Gold 100	0.80	0.81	0.78
Nash 100	0.81	0.88	0.80
NoName	0.82	0.54	-0.31

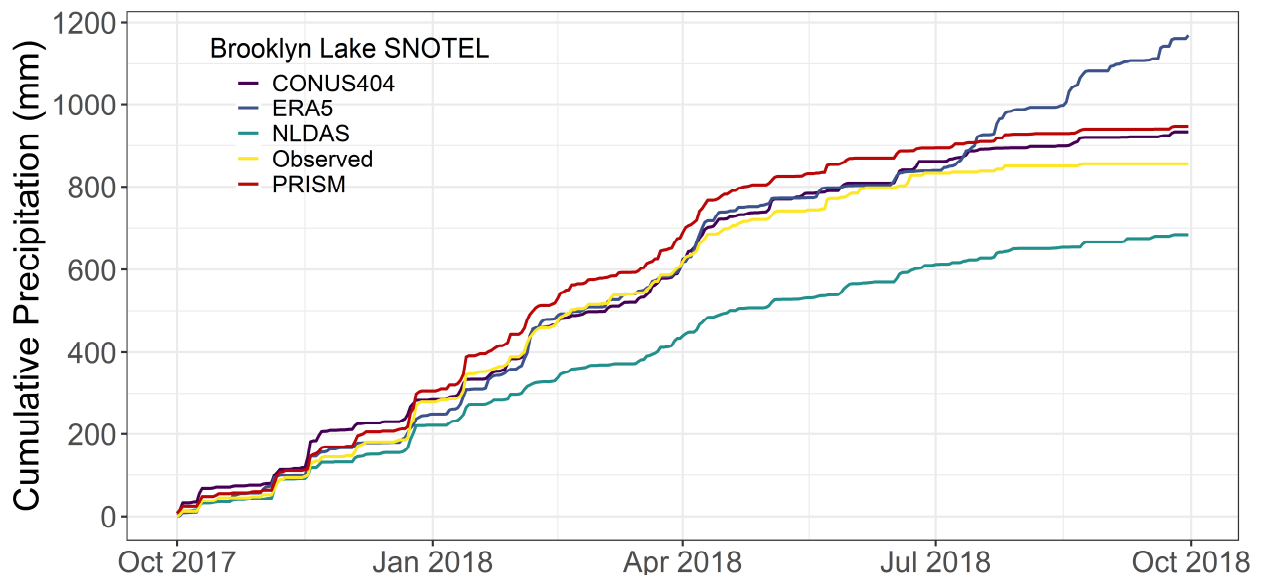


Figure 8: Cumulative annual precipitation for the three forcing datasets vs the observed precipitation. PRISM precipitation shown for reference.

The NLDAS dataset tended to underestimate winter precipitation, leading to lower simulated snowpack and reduced peak flows in spring. ERA5 and CONUS404, both with higher spatial resolution, aligned more closely with SNOTEL observations during the snow

accumulation season but ERA5 overestimated summer precipitation, likely due to overrepresentation of convective events.

The differences in overall annual precipitation also led to different snow accumulation and melt magnitudes (Fig. 9).

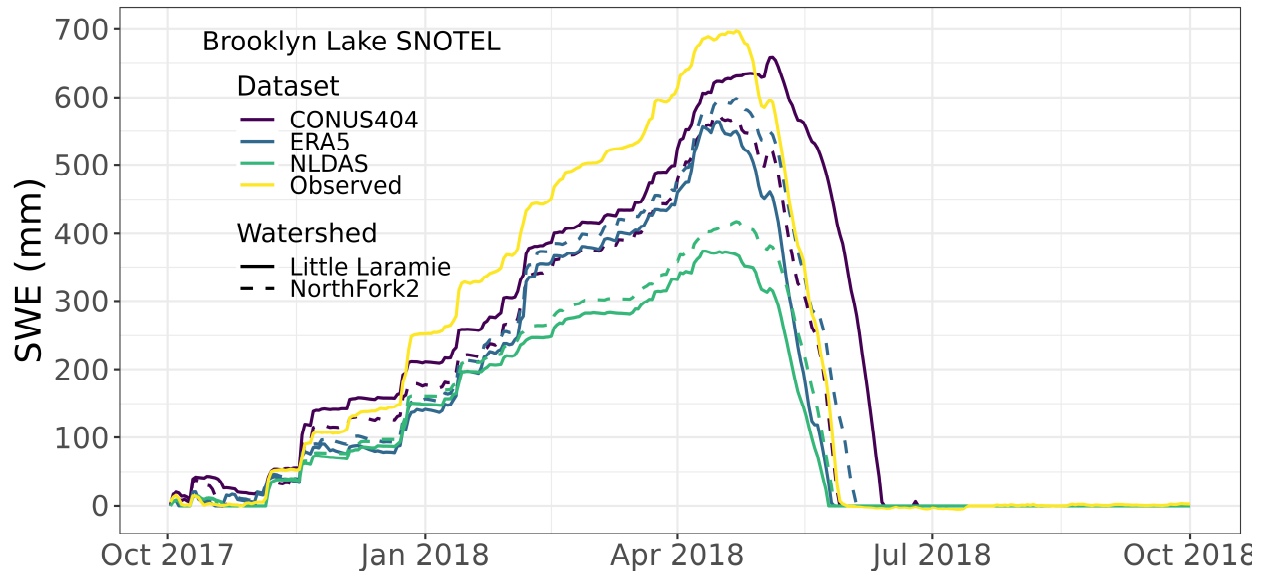


Figure 9: Snow water equivalent at the Brooklyn Lake SNOTEL for the three input datasets. In line with the underestimated winter precipitation, NLDAS also underestimates snow accumulation.

Model runs demonstrated that differences in forcing data resolution and bias had a significant effect on snow water equivalent, timing of snowmelt, and the resulting streamflow hydrographs. At Brooklyn Lake, simulated snowpack using NLDAS peaked several weeks earlier and at a lower magnitude than observed, whereas ERA5 and CONUS404 captured the timing more accurately but produced higher total melt volumes. The model also showed that soil moisture and evapotranspiration responses varied considerably among datasets, influencing water availability during the growing season. Comparisons with Sentinel-2 imagery confirmed that all datasets reproduced the general spatial distribution of snow (Fig. 10) but with systematic deviations at high elevations and along wind-exposed ridgelines.

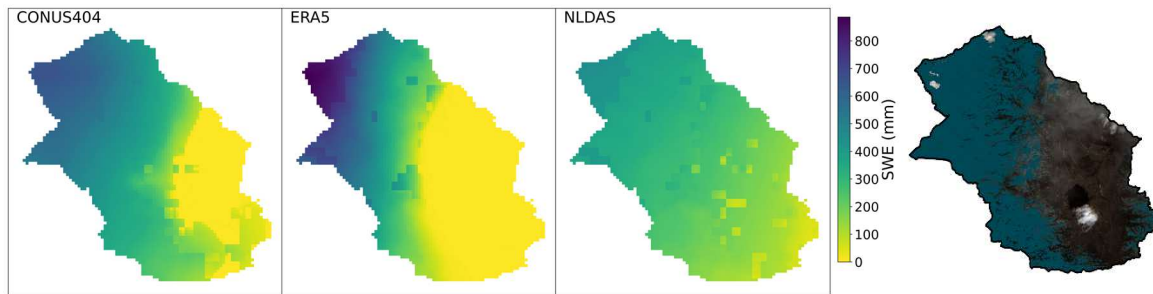


Figure 10: Spatial snow water equivalent at the time of greatest average snowcover, April 10. The ERA5 dataset resulted in the largest differences between min and max SWE but also simulated the largest portion of the watershed as being snow-free. The right-most panel depicts snow cover derived from LANDSAT satellite imagery, with the blue area highlighting areas covered in snow.

These mismatches are attributed to both wind redistribution of snow and biases in the gridded climate forcing.

Collectively, these results highlight the trade-offs between spatial resolution and bias correction in hydrologic modeling. Higher-resolution forcing data improved spatial representation of snow processes but did not necessarily eliminate bias in precipitation or energy fluxes. The different forcing datasets resulted in essentially the same amount of generated streamflow but in very different magnitudes in the remaining water balance components, especially evapotranspiration (Fig. 11). While streamflow was simulated equally well, the more realistic representation of snow and evapotranspiration suggest that both CONUS404 and ERA5 are the superior datasets. Further, this study highlights the flexibility of hydrologic models to turn imperfect input data into reasonable simulations if streamflow is the only metric assessed to judge model performance. The need for judging hydrologic models on more than just streamflow is apparent.

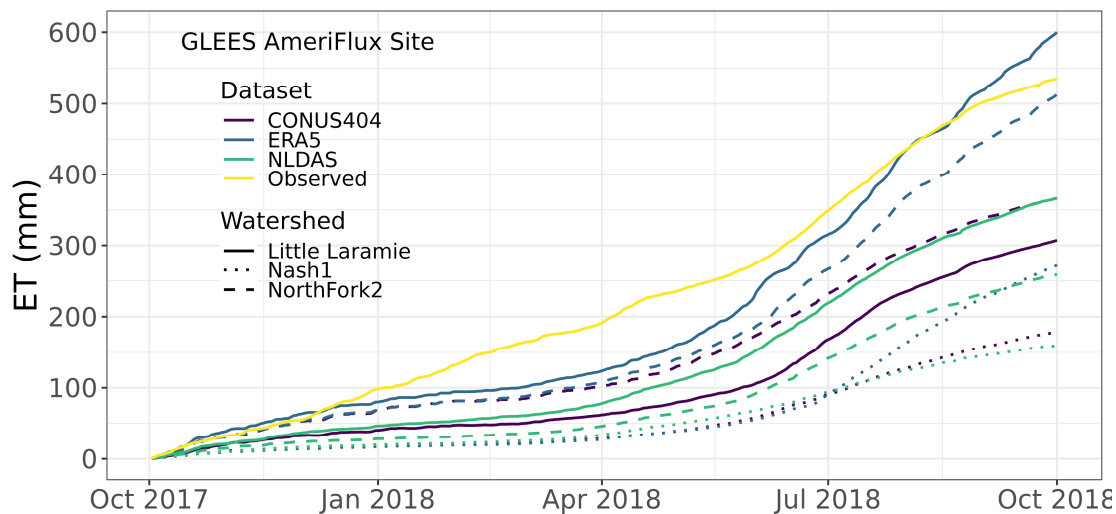


Figure 7: Cumulative evapotranspiration for the three forcing datasets at the GLEES eddy covariance tower. Most simulations underestimate the observed ET, except for the ERA5 simulation for the largest watershed (Little Laramie). The overestimation in this simulation is likely caused by the overestimation of summer precipitation. This rise is also observed at the other two watersheds in the ERA5 simulation.

Objective 3: Accuracy of snow data at the small watershed scale

The focus in this part is the simulation of small 1st-order headwater watersheds to compare and evaluate snow cover and water content collected via UAV and ground-based methods with high-resolution modeling data. The main interest hereby is the question how well hydrologic models at high spatial resolutions represent small-scale snow dynamics in snow-dominated watersheds. WRF-Hydro was calibrated using the 3 km ERA 5 climate forcing dataset at the smallest of the nine instrumented watersheds with a size of ~0.5 km² and a 100 m land surface model resolution. We compared spatial snapshots of UAV-derived snow cover and manually collected snow water equivalent measurements from 20+ snow courses over a snow melt season to high-resolution hydrologic model simulations. Further, we investigated potential errors in winter-season ablation from the snowpack.

Analyses in this summarizing objective are still ongoing. Preliminary results suggest that hydrologic models replicate streamflow well even if small-scale snow heterogeneity is not captured in the model. The potential error of misrepresentation of small-scale snow dynamics further decreases with increasing model scale. In practical terms, multi-copter UAVs are generally not capable of capturing large enough areas to meaningfully affect hydrologic model simulations. However, they can provide crucial data for hillslope-focused projects where detailed knowledge of water inputs at specific locations is more crucial. We will collect snow cover data with our newly acquired fixed-wing UAV (certified for use within the US), which can cover much larger areas and can provide datasets covering small to medium scale watersheds.

Significance

This project advanced understanding of hydrologic models and how scale and different climate inputs affect hydrologic model performance in Wyoming's mountain headwaters. The results emphasize that more attention must be paid to model domain resolution when the desired outputs go beyond replicating streamflow hydrograph. Further, the choice of forcing data strongly influences water budget components, especially in snow-dominated systems where small biases in winter precipitation propagate through the hydrologic cycle. By comparing multiple datasets and validating model outputs with field and remote sensing observations, this work provided guidance for improving WRF-Hydro applications in 1st to 4th order watersheds. The insights gained will help future modeling efforts better capture timing and magnitude of snowmelt-driven runoff, which is critical for water resource planning, reservoir management, and climate adaptation in the state.

In addition, the integration of student-led field measurements, data processing, and model development directly contributes to building Wyoming's workforce capacity in hydrologic science and modeling. The datasets, methods, and workflows developed through this

project lay the foundation for scaling WRF-Hydro applications to other Wyoming watersheds and for evaluating future climate change and land use impacts on snowmelt and streamflow.

Student Support

The project supported one PhD student for 3.5 years and ten undergraduate students. The PhD student, Salar Jarhan, is currently summarizing the research into his dissertation (anticipated graduation spring or summer 2026) with the plan to publish three manuscripts for submission in peer reviewed hydrologic journals. The undergraduate students assisted with fieldwork (discharge measurements, UAV flights, sensor maintenance and data downloads) and some data analysis tasks. The funding from this project allowed us to extend the data collection in the Snowy Mountains for another 4 years, generating critical discharge timeseries for capturing climatic variability.

Conferences and Presentations

Jarhan, S.*, Nippgen, F., Paige, G.B., Unmasking Bias: Interaction among climate forcing data and hydrologic response across spatial scales in snow dominated watersheds, Fall Meeting of the American Geophysical Union, New Orleans, December 2024 [Poster]

Nippgen, F., Jarhan, S.*, Bowen, J.*, Green-Smith, C.*, Paige, G.B., Snow, Models, Agriculture: Exploring Western Hydrology from Mountains to Rangelands, UC Boulder, October 13, 2024 [Invited] [Oral]

Jarhan, S.*, Nippgen, F., Paige, G. B., Hydrologic simulations in snow dominated, leaky watersheds: Insights and challenges across multiple spatial scales, The American Geophysical Union (AGU) Fall Meeting 2023, San Francisco, USA, 11-15 December 2023. [Oral]

Jarhan, S.*, Nippgen, F., Paige, G. B., From Headwaters to Mountain Range: Modeling Water Resources Across Spatial Scales, Wyoming Water Association & Upper Missouri Water Association Conference, Gillette, Wyoming, October 24-26, 2023. [Oral]

Nippgen, F., Paige, G.B., Simplifying the Complex: Indicators to Characterize Socio-hydrologic Systems, UCOWR Annual Meeting, Fort Collins, CO, June 2023 [Invited] [Oral]

Jarhan, S.*, Nippgen, F., Paige, G.B., Poor Performance or Indication of Water Losses? Insights into Watershed Structure and Water Losses from Hydrologic Model Simulation across Spatial Scales, UCOWR Annual Meeting, Fort Collins, CO, June 2023 [Oral]

Jarhan, S.*, Nippgen, F., Paige, G.B., How small is too small: Assessment of WRF-Hydro across multiple scales in a snow-dominated watershed in the intermountain western US, AGU Fall Meeting, Chicago, December 2022 [Poster]

Journal Articles

Jarhan, S.*, Nippgen, F., Paige, G.B. (in prep), Representation of Scale in Hydrologic Models and its Effect on Simulation Quality, [to be submitted Dec 2025]

Jarhan, S.*, Nippgen, F., Paige, G.B. (in prep), Forcing Matters: Evaluating Climate Inputs for Improved Hydrologic Prediction in Wyoming's Headwaters, [to be submitted spring 2026]

Jarhan, S.*, Nippgen, F., Paige, G.B. (analyses ongoing), Does high-resolution Snow Data Improve Hydrologic Models?, [to be submitted summer 2026]

Notable Awards and Achievements

PhD student Salar Jarhan received three travel grants from the Hydrologic Sciences Program for partial travel support to the AGU Fall Meetings 2022-2024.

Stakeholder groups from the Jackson, WY, area were aware of our modeling efforts in the Snowy Range and as part of a different project, Salar was instrumental in setting up the WRF-Hydro model for the Upper Snake River basin for climate change analyses.

With the outlook of improving UAV data coverage in mountainous environments for snow assessment, the WRP project, among others, was leveraged to secure ~100k funding to obtain a fixed-wing UAV with high-resolution RGB imagery, a 5-band multispectral sensor, and a Lidar sensor for more accurate representation of snow cover at larger spatial scales.

Citations

Cho, K., & Kim, Y. (2022). Improving streamflow prediction in the WRF-Hydro model with LSTM networks. *Journal of Hydrology*, 605, 127297.

<https://doi.org/10.1016/j.jhydrol.2021.127297>

Gochis, D. J., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., McAllister, M., McCreight, J., & RafieeiNasab, A. (2020). *The WRF-Hydro® modeling system technical description, (Version 5.1.1)*. NCAR Technical Note.

<https://ral.ucar.edu/sites/default/files/public/WRFHydroV511TechnicalDescription.pdf>

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91.

<https://doi.org/10.1016/j.jhydrol.2009.08.003>

Jordan, R. (1991). A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM.89. Hanover, NH: U.S. Army Corps of Engineers, Cold Regions Research and Engineering Laboratory.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.

[https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)