

FROM RIVER CHANNEL TO CLOUD TOP: EVALUATION AND
APPLICATIONS OF REGIONAL CLIMATE MODELS IN
MOUNTAIN WATERSHEDS

by

William J. Rudisill



A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Geophysics

Boise State University

August 2022

© 2022

William J. Rudisill

ALL RIGHTS RESERVED

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

William J. Rudisill

Dissertation Title: From River Channel to Cloud Top: Evaluation and Applications of Regional Climate Models in Mountain Watersheds

Date of Final Oral Examination: 06 June 2022

The following individuals read and discussed the dissertation submitted by student William J. Rudisill, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Alejandro Flores Ph.D.	Chair, Supervisory Committee
Hans-Peter Marshall Ph.D.	Member, Supervisory Committee
James McNamara Ph.D.	Member, Supervisory Committee
Rosemary Carroll Ph.D.	Member, Supervisory Committee

The final reading approval of the dissertation was granted by Alejandro Flores Ph.D., Chair of the Supervisory Committee. The dissertation was approved by the Graduate College.

ACKNOWLEDGMENT

We would like to acknowledge funding from the DOE SCGSR program and DOE BER grant DOE:DE-SC0019222. This research made use of Idaho National Laboratory computing resources which are supported by the Office of Nuclear Energy of the U.S. Department of Energy and the Nuclear Science User Facilities under Contract No. DE-AC07-05ID14517. We also acknowledge the use of Boise State's Research Computing Department. 2017. R2: Dell HPC Intel E5v4 (High Performance Computing Cluster). Boise, ID: Boise State University. DOI: 10.18122/B2S41H. Additional funding for Chapter 2 was provided by the Idaho Department of Water Resources through Grant/Award Numbers: CON01394, 151512.

ABSTRACT

Mountain ranges are vital "water towers" of the world and are uniquely threatened by anthropogenic climate change. At the same time, the paucity of observing networks limits our understanding of hydrometeorological processes in water-resource critical regions, including the Western United States. In the past decade, non-hydrostatic, convection permitting (1-4km horizontal resolution) regional climate models (RCMs) have emerged as a promising tool for both reconstructing regional scale mountain hydroclimates, and for forecasting the impacts of climate perturbations on watersheds and water-resources. Still, challenges remain. To-date, computational and data storage limitations have generally precluded many RCM studies to a handful of individual years, limiting the characterization of model uncertainties/biases and thus the interpretation of model outputs. Moreover, validating spatial precipitation fields from RCMs remains a challenge, as gridded precipitation datasets are highly uncertain in locations far away from observing stations. Consequently, further validations of regional climate models require leveraging diverse or indirect sources of hydrologic information. I develop three studies to meet these challenges in this dissertation. In the first, I examine the fidelity of coupled hydrologic-model/RCM for simulating streamflow in four water resource significant, snow-dominated basins in the Boise River basin. In the second, I develop a long-term RCM simulation (1987-2020) in the Upper Colorado River basin and evaluate precipitation fields using a

novel precipitation-from-streamflow bayesian inference strategy. The third chapter of the dissertation examines orographic precipitation sensitivities to cloud-microphysical parameterizations schemes, and leverages Airborne LIDAR snow-depth datasets to evaluate both spatial patterns of precipitation enhancement and watershed-total precipitation delivery. Together, the results from this dissertation demonstrate the utility of multi-decadal regional climate modeling for interrogating mountain hydro-climates, and demonstrates the opportunities and challenges for leveraging diverse hydrologic data sources (streamflow, airborne LIDAR) and methods (bayesian inference) for evaluating RCMs.

TABLE OF CONTENTS

ACKNOWLEDGMENT	iv
ABSTRACT	v
List of figuresxiList of tablesxx	
1 INTRODUCTION	1
1.1 Introduction	1
1.2 Foundations of Regional Climate Modeling	5
1.3 Overview of Cloud Microphysics	8
1.3.1 Some Fundamental Cloud Chemistry	9
1.3.2 Representing Microphysical Processes in Models	10
1.4 Water Movement on Hillslopes and Stream Channels	12
1.4.1 Distributed Hydrologic Modeling	12
1.4.2 Bucket Modeling	14
1.5 Land-Atmosphere Interaction Phenomena	15
2 EVALUATING LONG TERM ONE-WAY ATMOSPHERE-HYDROLOGY SIMULATIONS AND THE IMPACTS OF TWO-WAY COUPLING IN FOUR MOUNTAIN WATERSHEDS	17
2.1 Abstract	17

2.2	Introduction	18
2.3	Data and Methods	22
2.3.1	Study Area Geography, and Hydroclimate	22
2.3.2	Model Descriptions	23
2.4	Experiment Descriptions	25
2.4.1	Uncoupled Long-Term Runs	25
2.5	RESULTS	29
2.5.1	One-Way Coupled Long-Term Streamflow Simulations	29
2.5.2	Two-Way Coupled Model Analysis	33
2.6	DISCUSSION	37
2.6.1	One-Way Model Coupling Improves Precipitation Process Understanding	37
2.7	Investigating One-Way Coupled Streamflow Performance	40
2.8	Impacts of Two-Way Coupling	41
2.9	CONCLUSIONS	43
3	EVALUATING 34 YEARS OF DYNAMICALLY DOWNSCALED PRECIPITATION IN THE COLORADO ROCKIES	59
3.1	Introduction	1
3.1.1	Quantifying and Modeling Mountain Precipitation: Challenges and Consequences	1
3.1.2	Experiment Outline	3
3.2	Datasets and Methods	4
3.2.1	WRF Model Domain and Configuration	4
3.2.2	Comparison Precipitation Datasets	6

3.2.3	Precipitation Inference Methodology	6
3.3	Results	10
3.3.1	Regional Scale Comparison	10
3.3.2	East River Precipitation Analysis	13
3.4	Discussion	16
3.4.1	Comparisons against other Convection Permitting Simulations	16
3.4.2	Addressing Climate and Data Non Stationarities	19
3.4.3	Interpreting Streamflow Inferred Precipitation Estimates . . .	20
3.5	Summary and Conclusions	22
4	OROGRAPHIC PRECIPITATION SENSITIVITY TO MICROPHYSICAL PARAMETERIZATIONS: HYDROLOGIC EVALUATIONS WITH SNOW LIDAR DATASETS	38
4.1	Introduction	1
4.2	Methods	5
4.2.1	Significance of Microphysical Processes	5
4.2.2	Weather Research and Forecasting (WRF) Model Configuration	7
4.2.3	Metrics to Quantify Precipitation and Snowpack Fidelity . . .	8
4.2.4	Snowpack Modeling and ASO LiDAR Data Processing	11
4.3	Results	13
4.3.1	Precipitation Evaluated at SNOTEL Locations	13
4.3.2	Vertical Profiles	14
4.3.3	Modeled SWE and Snow Depth - Comparisons Against ASO LiDAR	15
4.4	Discussion	16

4.5	Conclusions	21
5	CONCLUSIONS	36
5.1	Synthesis and Discussion	36
5.2	Technological Challenges and Opportunities	38
5.2.1	Data Analysis and Visualization	39
5.2.2	Software Infrastructure	39
5.3	The Future of Regional Climate Modeling of Mountain Regions	40
5.4	Closing Remarks	42
	References42References42References42References42References42	

LIST OF FIGURES

2.1	Conceptual diagram illustrating the one and two-way coupled WRF/WRF-Hydro experiments.	47
2.2	Satellite view of the four study watersheds (yellow outline). Clockwise from the left: Mores creek (MC), the South Fork Payette (SFP), the South Fork Boise (SFB), and the Main Fork Boise (MFB) in the center. Black diamonds indicate USGS stream gauge locations used to evaluate model discharge. Red squares indicate the location of NRCS Snotel meteorological stations.	48
2.3	a) Percent canopy coverage, b) dominant soil category type, c) topographic elevation, and d) land-use category for the inner-WRF domain. Black outlines are the boundaries of the watersheds examined in this study	49
2.4	Initial soil saturation conditions for the top soil layer (10 cm depth) for the Control/Routing, HiResInit, and the difference between the two (right). Soil saturation is the soil moisture relative to the soil porosity.	50

2.5	Hydrograph demonstrating the efficacy of the calibration procedure for the uncoupled WRF-Hydro simulations for the South Fork Boise River. Showing 200 iterations of the dynamic-dimension-search algorithm. The dashed lines are observed USGS streamflow observations and the shaded region denotes the model min and max values in the calibration iterations.	51
2.6	Simulated daily discharge (m ³ /s) from the uncoupled WRF-Hydro model, forced by hourly WRF meteorological forcings, for a 20-year (1995-2014) period compared against corresponding USGS stream gauge measurements. Right column shows the distribution of daily streamflow biases (model-observed). From the top down: the South Fork Boise (SFB), Moore’s creek (MC), the main fork of the Boise River (MFB), and the South Fork Payette (SFP).	52
2.7	Linear relationship between WRF cold-season (October-May) precipitation percent bias, evaluated against NRCS Snotel sites (columns) and WRF-Hydro stream discharge percent biases evaluated at USGS gauge locations for each watershed (rows), 1995-2014. Markers with significant correlations ($p_{j.05}$) are filled. The colorbar indicates coefficient of determination (R^2). Far left row shows the distribution of annual streamflow percent biases. Top row shows the distribution of annual precipitation percent biases	53

2.8	Time Series view of 500 meter geopotential height (top), precipitation rate (mm/hr; mm), Convective Available Potential Energy (J/kg), two meter air temperature ($^{\circ}$ C), latent heat flux (LH; w/m ²) and sensible heat flux (HFX; w/m ²), and soil saturation of the 2nd soil layer. Plots are spatially averaged across the study region box (Figure 7). “Control” denotes the difference between the Control run and each respective scenario.	54
2.9	Temporal averages across the subdomain of soil saturation (2nd layer), two-meter air temperature ($^{\circ}$ C), afternoon Convective Available Potential Energy (CAPE; J/kg), latent heat flux (LH; w/m ²) and sensible heat flux (HFX; w/m ²), Planetary Boundary Layer height (PBLH) and Accumulated Precipitation at the final timestep. Differences (Control-Routing and Control-HiResInit) are expressed as percent of Control for PBLH, Precipitation, and CAPE.	55
2.10	Figure 10: May 15 - June 1, Low elevation (\geq 2000m) temporally averaged relationships throughout the subdomain between CAPE (J/kg), accumulated precipitation (mm), evaporative fraction (EF; fraction of latent heat relative to the sum of latent and sensible heat flux), planetary boundary layer height (PBLH), and soil saturation in the second soil moisture layer.	56

2.11	The coupled model experiments convective-triggering-potential (CTP) and HILow indices, evaluated at 5:00 a.m local time. Graphs correspond with individual WRF grid cells located in Stanley and Boise (Figure 2). National weather service radiosonde observations from Boise are also plotted (black triangles). Green hatched boxes indicate thermodynamic conditions where precipitation is considered atmospherically controlled/insensitive to surface conditions based on (Findell and Eltahir, 2003).	57
2.12	Observed and WRF-model Skew-T Log P diagrams for morning (12z) and afternoon (00z) during May and June of 2018.	58
3.1	WRF Model domains (left) and East River watershed (right). Outer nest is 3km dx/dy and the inner domain is 1km dx/dy. Triangles show the locations of Snotel sites examined in this study	24
3.2	Error characteristics of 24 Snotel sites compared against corresponding WRF grid-cells, 1987-2020. A) One-week rolling mean timeseries of average Snotel (orange) precipitation compared against WRF (blue) B) Annual precipitation percent error (Snotel as reference) for each site, C) Average timing of water delivery (%) as a function of day of water-year for WRF (blue) and Snotel (orange), D) Correlations between one-week accumulated precipitation WRF/Snotel for Warm season (bottom) and Cool season (Top), D) Average precipitation percent errors by season	25

3.3	A) Percentage of days with precipitation ("wet") for Snotel stations and corresponding WRF grid-cell, for thresholds of daily precipitation greater than 2.54 mm (.1 inches) and 0 mm respectively, for Cold (Oct. - April) and Warm (April-Oct.) seasons. B) Contribution to cumulative precipitation by daily precipitation rate (mm/day) for WRF (solid) and Snotel (dashed) stations. Blue/Orange lines are the mean of all stations for WRF/Snotel respectively.	26
3.4	Digonal: Spatial averages of 1987-2010 precipitation for WRF, PRISM, Newman, and Livneh datasets for the entirety of the inner WRF model grid (300x300km). Lower off-diagonals are the difference between respective datasets (row-column). Upper diagonals are the same, but show the percent-difference ($[(\text{column}-\text{row})/\text{row}]$).	27
3.5	East River Annual Precipitation by Water Year, 1987-2020 for WRF and PRISM, Livenh, and Newman precipitation products	28
3.6	Average Cold-Season (October-April), Warm-Season (April-October), and average annual precipitation enhancement factors (Equation ??) for WRF and PRISM	29

3.7 Precipitation inference model. A) Meteorological forcing structure. Precipitation/temperature increase/decrease with elevation and typical snow/melt fluxes are illustrated. B) Hydrologic model structure. Rain and Snowmelt (Q_{in}) and Potential Evapotranspiration (PET) are model forcings. State variables are the soil moisture in the top and bottom bucket respectively (SM_1, SM_2). Model fluxes are baseflow (Q_b), percolation, Q_{12} , overland flow (Q_d), and bucket overflow (Q_{bO}). C) Q_d and Q_b are convolved with a routing function to produce streamflow (D). 30

3.8 Left: The Airborne Snow Observatory SWE maps over the East River employed in this study. Center: Basin-average elevation versus average SWE for ASO and calibrated Snow-17 at ASO-flight dates. Right: Timeseries of calibrated Snow-17 by elevation band (n=100). 31

3.9 Evaluating precipitation-from-streamflow Bayesian inference method. A) WRF and PRISM Basin mean precipitation compared against posterior precipitation (violinplots). Residuals between PRISM/WRF and inferred are plotted (right y-axis). Cumulative observed streamflow plotted for reference (grey boxes). B: Observed streamflow and posterior streamflow solution with 1σ confidence interval (grey shading). C: Distributions of posterior forcing parameter values, inferred from each year D: Distributions of annual error metrics. Kling Gupta Efficiency and Root Mean Square Error. 32

3.10	East River annual mean PRISM mean precipitation compared against the three closest Snotel sites, Schofield (737; North of East), Upper Taylor (1141; East of East), and Butte (381; within East). Pre/Post 2010 mean difference between Butte and PRISM-mean are plotted with standard errors.	33
3.11	Average annual precipitation versus elevation for WRF and PRISM, for the entire WRF domain (top) and the East River Watershed (bottom). Rolling means (solid lines) are shown in addition to OPG lines (Equation 3.1)	34
4.1	Conceptual diagram illustrating microphysical rate controls on orographic precipitation. Secondary controls on slope scale snow deposition/redeposition are also shown.	26
4.2	Airborne Snow Observatory (ASO) LiDAR derived snow depths for April 7, 2019 at three different resolutions (50m, 500m, and 1km) using bilinear interpolation. The black box in the left hand figures corresponds to the latitudinal and longitudinal extent of the figures on the right.	27
4.3	WRF total precipitation (ACCPRCP; bottom) and two meter surface air temperature (Tair; top) compared against three NRCS SNOTEL sites, Butte (site number 380), Schofield (site number 737), and Taylor (site number 1141), for water years 2018 (first row) and 2019 (second row).	28

4.4	Drying Ratios (total precipitation normalized by incoming water vapor flux) within the Upper Gunnison region computed for October 1 to April 1 for water years 2018 and 2019 across each WRF microphysics scheme.	29
4.5	Spatiotemporal averaged mixing ratios for different hydrometeor species estimated (if applicable) from the three WRF microphysics schemes (mp08, mp10, and mp55) used in this study. The y-axis units are kilometers above sea level.	30
4.6	Cross sections of average directions of vertical windspeed (red/blue shading; units of m/s), vertical and zonal flow (arrows; units of m/s), and ice-phase hydrometeor concentrations (contours; units of g/kg dry air). Green dots show the regions where the average meridional wind speed reverses and is greater than 1 m/s). Average water year total precipitation (precip) are provided above each plot. Water year 2018 (2019) is shown in the top (bottom) row.	31
4.7	Noah-MP static geographic data used for offline snow modeling. East River Watershed topography from the ASO LiDAR digital elevation model (DEM; upper left), terrain aspect (upper right), USGS 24 category vegetation classification type (bottom left), and MODIS satellite derived green-vegetation fraction (Veg Frac; bottom right). NRCS SNOTEL locations (black dots) are overlain on each geographic dataset.	32

4.8	SWE (top two rows) compared between ASO LiDAR and Noah-MP modeled SWE for WY2018 and WY2019, evaluated near the date of peak snow accumulation. Analogously, snow depth (SD; bottom two rows) is shown.	33
4.9	SWE versus elevation within the ERW. ASO LiDAR derived SWE (left) is compared against the average Noah-MP SWE. Rolling-mean curves for each Noah-MP simulation and ASO LiDAR data are shown. Lines with four different SWE versus elevation slopes (purple lines) are provided on each plot to better enable juxtaposition of datasets across different water years.	34

LIST OF TABLES

2.1	Table 1: Study watershed descriptions. USGS gauge identifier, basin area, average elevation, and the annual average of specific stream discharge (m ³ /s) measured at the corresponding USGS gauge for the study period (Water Year 1995-2014).	45
2.2	Table 2: Select WRF and WRF-Hydro model configuration options used in the one and two-way coupled model experiments. Two numbers correspond with the outer/inner nest.	45
2.3	Table 3: List of WRF-Hydro parameter names, as they appear in the model files (left), calibrated for the one-way coupled model experiments.	46
2.4	Table 4: WRF-Hydro simulated stream discharge performance metrics for the Moore’s Creek, Boise River, Southfork Boise River, and South Fork Payette watersheds for the 20-year validation period. The modeled discharge is compared against corresponding USGS river gauges aggregated to daily values.	46
3.1	Weather Research and Forecasting (WRF) parameters used in this study.	35
3.2	Functional Forms for Bucket Hydrologic Model Fluxes	35

3.3	Model parameter prior values and probability distribution used in the precipitation inference method. Ranges refer to the min/max of the uniform distribution or the mean/standard deviation of the normal distribution.	36
4.1	WRF v4.1.2 Microphysics options examined in this study.	7
4.2	Weather Research and Forecasting (WRF) subgrid-scale parameterizations used in this study.	24
4.3	Spatiotemporal and mass-balance error statistics for Noah-MP models compared against the ASO LiDAR derived basin-wide snow depth and Snow Water Equivalent (SWE) estimates. Bold values denote the best performing scenario.	25

CHAPTER 1:

INTRODUCTION

1.1 Introduction

This dissertation aims to improve predictive understandings of hydrologic systems in mountain regions, with particular emphasis on the Western United States. In these regions, interactions between the solid-earth and atmosphere profoundly shape the hydrologic cycle through orographic precipitation, geomorphic, ecological, and radiative processes. Anthropogenic climate change can/will impact mountain watersheds disproportionately (Mountain Research Initiative Edw Working Group *et al.*, 2015). Globally, one-sixth of the world's population relies on water resources flowing from mountain watersheds, often which falls in the form of snow that accumulates during the course of the winter, acting as a natural reservoir (Sturm *et al.*, 2017). Ecological and biological diversity flourishes in in mountain regions, with high-elevation endemic species adaptations to unique mountain hydroclimate and geomorphology (Immerzeel *et al.*, 2020). Regions of the Western United States (WUS) are experiencing warming trends across seasons (Rangwala & Miller, 2012) and declining Spring snowpacks primarily driven by temperature (Mote *et al.*, 2005), in addition to changes in snowfall extremes (Lute *et al.*, 2015). Projections suggest that mountains of the WUS region may see the first snow-free winters as early as 2050 (Siirila-Woodburn *et al.*, 2021).

Still, observational and modelling limitations mean there are many uncertainties related to the current state and future of mountain hydroclimates. Global climate models (currently on the order of 50km grid-spacing) do not resolve the sharp gradients in climate found in most mountain ranges. Therefore, statistical or dynamical methods of "downscaling" are required to attribute global model solutions to local scales (Prein *et al.*, 2015). Validating such models is a persistent challenge, as mountain regions are under-sampled by hydrologic and meteorological instrumentation throughout the US and globe. Frozen phase precipitation (snowfall) is difficult to accurately measure with precipitation gauges (Lundquist *et al.*, 2019), space-borne remote sensing algorithms are challenged in complex terrain and for ice phases, (Lettenmaier *et al.*, 2015), radar networks have poor coverage over mountain ranges (Maddox *et al.*, 2002), and gauge-based datasets (Daly *et al.*, 2008; Thornton *et al.*, 2016) are themselves models of these processes, and can disagree substantially in complex terrain (Henn *et al.*, 2018). Consequently our ability to understand hydrologic processes in mountains at regional scales is highly limited by the quality of meteorological data, precipitation in particular, which limits studies of the current climate and the ability to project changes into the future. Predicting the evolution and fate of snowpacks is particularly dependent on the quality of precipitation input (Raleigh *et al.*, 2015).

In the recent two-decades, non-hydrostatic, convection permitting numerical weather models have proved to be increasingly useful tools for reconstructing hydrologic cycles of mountain regions. These methods may go many names depending on the context: regional climate modeling, dynamical downscaling, convection-permitting modeling, limited area modeling, and mesoscale modeling to name a few. Convection-permitting modelling refers to non-hydrostatic models at sub-4km grid spacing, at which deep

atmospheric convection is theorized to emerge without the need of additional parameterizations (Weisman *et al.*, 1997). These types of models can both make forecasts at weather time scales for operational purposes (Benjamin *et al.*, 2016) or use atmospheric reanalyses datasets, which assimilate past atmospheric observations, as lateral boundary conditions to hind-cast meteorological conditions at resolutions finer than the parent reanalysis (Prein *et al.*, 2015; Lucas-Picher *et al.*, 2021). This latter configuration proven skillful for simulating precipitation and other meteorological fields in a variety of mountain watersheds throughout the WUS (Gutmann *et al.*, 2012; Ikeda *et al.*, 2010; Rasmussen *et al.*, 2011; Liu *et al.*, 2017) and globe.

The goal of this dissertation is to evaluate dynamical downscaling methods for hydrological applications in watersheds of the Western US, namely the Boise River basin (Chapter 1) and Upper Colorado River basin (Chapters 2,3). The dissertation primarily uses the Weather Research and Forecasting (WRF; Skamarock *et al.* (2008)) model. The goals are twofold. First, dynamical downscaling methods offer Earth-science practitioners distributed, serially-complete reconstructions of surface meteorological conditions (winds, precipitation, temperature, etc.) that can be used for hydrologic, critical zone, biogeochemical or other studies in mountain watersheds at fine-spatial and temporal scales (1-4km; hourly). To this end, the biases and error characteristics of dynamically downscaled fields need to be evaluated to inform applications and interpretation of the end-results. Second, understanding the multi-decadal behavior of dynamical downscaling models and associated biases remains a poorly explored, in part because of computational limitations (Lucas-Picher *et al.*, 2021; Prein *et al.*, 2015). Evaluating the multi-decadal behavior of dynamical downscaling models for the recent climate (30 years ago to present) is an essential

step for interpreting and understanding dynamical downscaling of climate change outputs. However, observational constraints consistently limit some of the abilities for verifying modeling outputs, particularly in complex terrain for the reasons already mentioned. Consequently, Lundquist *et al.* (2019) urges the hydrologic and atmospheric science communities to better collaborate across disciplines, and to develop better multi-objective methods of evaluating regional climate models. To this end, I develop methods for using additional hydrologic information, beyond precipitation gauge data, to evaluate the hydrologic fidelity of regional climate models. In particular, Chapters 2 and 3 develop methods to use streamflow and snow remote sensing to interrogating modelled precipitation. In addition, regional climate models offer process-level understanding of mountain systems, and insights into how different components of the land-atmosphere system interact. To this end, a section of the first chapter is devoted to investigating soil-moisture atmosphere interactions, building off of Rudisill *et al.* (2021).

The dissertation is structured into five chapters. In the first, I outline some of the modeling fundamentals and scientific concepts used in this study. The latter three are research articles. The first article concerns modeling river discharge for four watersheds in the Idaho Rockies, in addition to quantifying the role of land-atmosphere interactions mediated through soil-moisture. The second chapter validates atmospheric modelled precipitation in terms of seasonal accumulation, correlation, and seasonality using a combination a Bayesian inference approaches (“doing hydrology backwards”) and comparisons against NRCS Snotel measurements (Serreze *et al.*, 1999) and gauge-based precipitation products (Daly *et al.*, 2008; Livneh *et al.*, 2013; Newman *et al.*, 2015). The third chapter compares the sensitivity of winter pre-

precipitation to three different microphysical schemes in the WRF model — Morrison *et al.* (2005); Thompson *et al.* (2008a); Jensen *et al.* (2017) — and uses snow-lidar datasets (Painter *et al.*, 2016) to interrogate biases in accumulation and spatial distribution. The final chapter summarizes the key findings, challenges, and outlines future directions for research.

The tools used in this dissertation are primarily physically based models of the atmosphere and hydrologic flows — the language of fluid mechanics. These rules are well established in classical physics. However, numerical closure problems and ”sub-grid” scale processes both require ”parameterization” rules, meaning that physically based models require major leaps of intuition and imagination to adequately describe the phenomena in question (Stensrud, 2009). Processes leading to cloud-formation (so called ”cloud microphysics”), turbulence, radiation interactions with clouds, and land surface fluxes of heat/moisture/momentum are all processes that are ”parameterized” in models. This section describes some of the fundamental scientific background material that is used in this dissertation.

1.2 Foundations of Regional Climate Modeling

Some recent reviews of convection-permitting, regional climate modeling are described in Prein *et al.* (2015) and Lucas-Picher *et al.* (2021). Regional climate modelling is a boundary value problem, where reanalyses products are used as boundary conditions. In this way regional climate models can be thought of as ”magnifying glasses” into atmospheric scales not resolved from global reanalyses, which are typically on the order of 25-50 km in horizontal resolution. This is sufficient to resolve the general structure of synoptic features (extratropical cyclones for example) but not the small scales responsible for orographic precipitation (Roe, 2005).

Atmospheric flows can be described by the conservation laws of linear momentum, energy, and mass of both dry-air and moisture in addition to an equation of state. Such conservation laws are expressed in the language of partial differential equations. The basic equations, or so called "primitive equations" of numerical weather prediction are expressed in parcel-following (lagrangian) form with cartesian coordinates and the rotating Earth as frame by Pu & Kalnay (2019):

$$\frac{d\mathbf{V}}{dt} = -\frac{1}{\rho}\nabla p - \nabla\Phi + \mathbf{F}_r - 2\Omega \times \mathbf{V} \quad (1.1)$$

$$\frac{\partial\rho}{\partial t} = -\nabla \cdot (\rho\mathbf{V}) \quad (1.2)$$

$$p\alpha = RT \quad (1.3)$$

$$Q = C_p \frac{dT}{dt} - \frac{1}{\rho} \frac{dp}{dt} \quad (1.4)$$

$$\frac{\partial\rho q}{\partial t} = -\nabla \cdot (\rho\mathbf{V}q) + S_q \quad (1.5)$$

Where \mathbf{V} is the 3-d wind vector field given by $\mathbf{V}(x, y, z, t)$, ρ is the air density, p is pressure, T is temperature, g is the gravitational force, and q is the water vapor mixing ratio, and \mathbf{F}_r is the frictional force, and Ω is the Earth's angular velocity at a given latitude. Q is the diabatic heating term and C_p is the specific heat capacity of dry air. The terms on the right side of Equation 1.2 are the pressure gradient force, gravitational force, friction, and the apparent Coriolis forces ($2\Omega \times \mathbf{V}$). The second equation expresses the continuity equation for dry air. The 1.2 equation is the ideal gas law, the fourth equation is the conservation of energy, and the fourth equation is the continuity equation for water vapor, where S_q is the source/sink term accounting

for precipitation/evaporation. Analytical solutions for the governing equations are not practicable, so equations are discretized on a grid and solved using numerical methods.

In most cases, the equations are expressed using a coordinate system with a more complicated vertical coordinate. The model used in this dissertation is the Weather Research and Forecasting model (WRF; Skamarock *et al.* (2008)). WRF uses a hybrid vertical coordinate (as opposed to a pressure or Cartesian coordinate) following Laprise (1992). Moreover, for numerical reasons equations are expressed using Reynold's averaging, where a given state variable is expressed as the sum of a mean and perturbation component. WRF uses the compressible, non-hydrostatic Euler equations for Equation 1.2. First, "non-hydrostatic" refers to the hydrostatic approximation given by:

$$\frac{\partial p}{\partial z} = -\rho g \tag{1.6}$$

Where p is pressure, z is height, ρ is atmospheric density, and g is the gravitational force. Some models use this approximation to simplify Equation 1.2, but representing mountain gravity waves at kilometer-scales requires non-hydrostatic models Prein *et al.* (2015). The term "convection permitting" refers not to the particulars of the governing equations, but rather the horizontal length scales of the model. It is accepted that a grid spacing of less than $4km$ are necessary to resolve horizontal thermodynamic gradients responsible for convection (Weisman *et al.*, 1997).

The governing equations are discretized and solved using numerical methods. WRF uses a third-order Runge-Kutta finite difference method to integrate the system of equations. A "C-grid" staggering approach is used, following Arakawa & Lamb

(1977). Boundary conditions are required at the model boundaries for each forecast state variable.

1.3 Overview of Cloud Microphysics

Chapter 4 investigates cloud microphysical parameterizations, which are responsible for the S_q term in Equation 1.5. A brief overview of cloud physics is warranted, and largely follows from Yau & Rogers (1996) and Morrison *et al.* (2020). Generally speaking, clouds form when the relative humidity of an air parcel reaches or exceeds 100%, or when $e \geq e_s$, where e is the vapor pressure and e_s is vapor pressure at saturation. This can happen through a variety mechanisms, including cooling by adiabatic expansion (caused by lifting air parcels), radiative cooling, or mixing of air parcels. The growth of precipitation particles within clouds is controlled by small scale thermodynamic and kinematic processes, on the order of micrometers to millimeters, referred to as "microphysics". The physics of clouds are surprisingly complicated, and are in many ways analogous to the complexity of turbulent atmospheric flows. Most clouds do not precipitate and are relatively stable. Precipitation happens when conditions favor cloud particles to grow large enough to overcome frictional and restoring forces inhibiting their fall. The majority of precipitation and cloud growth occurs heterogeneously, meaning that cloud condensation nuclei (CCN) are present. The primary nucleation of ice crystals is very sensitive to the presence of CCN, as ice particles will not freeze homogeneously (in a completely pure environment) until $-40^\circ C$ is reached. Consequently it is common to have super-cooled liquid water (below $0^\circ C$) coexisting with ice particles in clouds ("mixed-phase" clouds). Since droplet formation depends on the presence of CCNs, super-saturation is common, where the relative humidity exceed 100% by a few tenths of a percent.

1.3.1 Some Fundamental Cloud Chemistry

A few fundamental theories related to the growth of liquid and ice droplets are worth describing. The temperature dependent saturation vapor pressure given by the Clausius-Clapeyron relationship assumes a flat, infinitely wide water surface in equilibrium with the atmosphere. To be applied to particles, the Clausius-Clapeyron relationship must be adjusted to take into account surface tension and the shapes of the particle. The effects of solutes (such as NaCl) also impact the saturation vapor pressure, following Raolt's law. The combined effects of solutes and and particle radius on the saturation vapor pressure can be expressed as:

$$S_R = \frac{e'_s(r)}{e_s(\infty)} = 1 + a/r - b/r^3 \quad (1.7)$$

Where $e_s(\infty)$ is the saturation vapor pressure calculated by the clausius-clapeyron relationship and $e'_s(r)$ is the adjusted saturation vapor pressure for a spherical particle of radius (r), and S_R is the saturation ratio. The a/r term expresses the dependency on curvature, and the b/r^3 term expresses the effects of solute concentrations. a has a linear dependency on temperature and b is depends on the concentration and properties of the solutes in the droplet (Yau & Rogers, 1996). Together these relationships are described by "Kohler curve" theory described a century ago (Köhler, 1921). The addition of solutes will tend to lower vapor pressure (larger b/r^3 , making it easier for droplets to form) whereas increasing particle radii make it more difficult to form droplets (larger a/r). Kohler theory means that particles beyond a certain critical radius (found by locating the peak of Equation 1.7) curve will continue to grow through condensation of environmental air until the super-saturated environmental air is de-

pleted, whereas small droplets will remain small. Kohler theory explains the relative stability of non-precipitating clouds

Another important thermodynamic process relates to ice phase. Ice particles have a lower saturation vapor pressure (e_s) than super-cooled liquid water at the same temperature. Consequently, ice will grow by vapor deposition faster than super-cooled liquid water droplets through the so-called Bergeron process (Bergeron, 1935). Still it is generally understood that the growth of hydrometeors by diffusional growth cannot explain the observed rates of precipitation. Collision-coalescence mechanisms, where droplets collide with other droplets, is another mechanism that may lead to droplet growth. There are still many uncertainties about the growth processes in real-world clouds (Morrison *et al.*, 2020). Ice phases are difficult to model and understand from a physical perspective given the range of possible crystal habits (for instance, an ice particle cannot be approximated as a sphere with radius, r , as in Equation 1.7).

1.3.2 Representing Microphysical Processes in Models

Models of the atmosphere must create rules that determine when/if a saturated grid cell loses mass (water) to precipitation following some of the principles outlined above. Early models used a simple precipitable water threshold value, above which water is instantaneously removed from a given model cell (Morrison *et al.*, 2020). Early schemes such as Kessler (1969) scheme did not account for ice phase processes, and only considered water vapor, cloud-water, and warm-rain. Modern bulk microphysical parameterizations evolved from this legacy, and now typically model the evolution of include cloud-water, cloud-ice, snow, rain, graupel, and hail. Many modern schemes predict two quantities for microphysical species: the Number Concentration

(N_c ; L^{-3}) and Mixing Ratio (q ; kg/kg). A scheme that predicts both is referred to as a "Double Moment" scheme. The mixing ratio is simply the mass of the particle per the mass of dry air. The number concentration is the quantity of individual particles per volume of air. Expressions relating particle diameters and number concentrations are often used of the form:

$$N_c = \int_0^{\infty} n(D)dD \quad (1.8)$$

Where $n(D)$ is the Drop Size Distribution function, with units (m^{-4}). The is typically given by an exponential equation of the form:

$$n(D) = zD^x \exp(-yD) \quad (1.9)$$

Where D is diameter and $n(D)$ has units m^{-4} . The variables z , x , and y are parameters that may or may not depend on various physical factors or may be set to constant values. The relationship between diameter and drop-size distribution may not seem intuitive. Nonetheless this is based on the canonical "Marshall Palmer" equation Marshall (1948). Modern microphysical schemes Thompson *et al.* (2008a); Morrison *et al.* (2005); Jensen *et al.* (2017) still use drop-size distribution functions more or less of the same form as the original Marshall Palmer type.

The general continuity equation for a given hydrometeor species in a typical bulk scheme can be given by:

$$\frac{\partial \chi}{\partial t} = -\mathbf{u} \cdot \nabla \chi + \frac{1}{\rho_a} \frac{\partial \rho_a V_x \chi}{\partial z} + S_x \quad (1.10)$$

This equation expresses the continuity equation for the scalar, χ , which can be

the mixing ratio, number concentration, or other hydrometeor property. V_χ is the weighted terminal velocity of the particle, ρ_a is the density of air, and S_χ is the production/destruction term for the hydrometeor in question. The terminal velocity of the particle (V_q) is related to the particle mass through the frictional force, which typically has a scalar drag coefficient parameter.

The S_q terms represent processes that produce/destroy the hydrometeor species, *i.* These may include autoconversion, accretion, vapor deposition, collection of rain-water by snow, sublimation, and more. The complexity of the microphysics scheme will determine which types, and how, processes are represented. Bao *et al.* (2019) provides a good comparison of the modeled tendencies in three different commonly used microphysical parameterizations in the WRF model.

1.4 Water Movement on Hillslopes and Stream Channels

There are two different applications of streamflow modeling in this dissertation, with two different aims. In each case daily streamflow values are analyzed for 20-30 year periods for snow dominated basins (Colorado Rockies; Idaho batholith).

1.4.1 Distributed Hydrologic Modeling

In Chapter 1, a regional climate model (WRF) is "one-way" coupled with a distributed hydrologic model (WRF-Hydro) to simulate streamflow in the Boise River basin. The WRF-Hydro model (Gochis *et al.*, 2018) evolves from Land Surface Models (LSMs) that are used to solve for the fluxes of energy/mass/momentum at the bottom boundary of atmospheric models. These types of models typically do not move water laterally from grid-cell to grid-cell, nor do they have channel routing functions.

WRF-Hydro effectively adds another layer of hydrologic realism to NoahMP (Niu *et al.*, 2011), namely lateral saturated subsurface flow, overland flow, ponded water, and channel routing formulae. In NoahMP (and WRF-Hydro) soil-water infiltration follows the one-dimensional Richard's Equation. The saturated subsurface flow equation uses a formula based on Wigmosta & Lettenmaier (1999) and derived from the dupuit-forcheimer assumptions. For each grid cell in the model, the flux is computed by:

$$q_{ijk} = -T_{ijk} \tan(\beta_{ijk}) w_{ijk} \quad (1.11)$$

Where T is transmissivity, β is the slope of the of the water table in the k direction, and w_{ijk} is the flow width. q_{ijk} is then the flow rate from the i,j cell in the k direction. The total flux is then found by summing all of the fluxes from each contributing grid cell. Overland flow is modelled using a kinematic wave approximation. The kinematic wave approximation is a simplification of the total momentum continuity equation for shallow flows, where only the bottom slope and friction forces matter. In one dimension, the continuity equation for a shallow kinematic wave is given by:

$$\frac{\partial y}{\partial t} + \frac{\partial q_x}{\partial x} = i_0 \quad (1.12)$$

$$q_x = \alpha y^{5/3} \quad (1.13)$$

Where y is the water height, q_x is the specific discharge. Equation 1.13 comes from applying a resistance equation such as Manning's equation (Eagleson, 1970), which acts on the flow velocity at the channel bottom. In WRF-Hydro, α is given

by $\alpha = \frac{S_{fx}^{1/2}}{n_{ov}}$, where $S_{fx} = S_{ox} - \frac{\partial Y}{\partial x}$. S_{ox} is the terrain-slope, and $\frac{\partial Y}{\partial x}$ is the change in water height over length x and n_{ov} is a tunable parameter. Once water enters a streamchannel, several different options can be used to model flow through an idealized, trapezoidal channel described by a constant slope, bottom-width, roughness coefficient, and follows the same principles of momentum and mass conservation.

WRF Hydro and other distributed hydrologic models require many such parameters, including for soil properties like porosity and hydraulic conductivity. In practice this means taking geographic maps of soil properties (such as Miller & White (1998)) and using different approaches of mapping soil categories to hydrologic parameters. In this type of "distributed" model, fluxes are computed for each grid cell within the simulation domain and routed into an idealized trapezoidal channel, where the muskingum-cunge streamflow routing equation is solved to produce a hydrograph for each point in the idealized channel.

1.4.2 Bucket Modeling

Chapter Three applies soil-moisture accounting and stream channel runoff models, but with a different aim than Chapter Two. A conceptual, "bucket" hydrologic model is employed based on the "FUSE" model (Clark *et al.*, 2008). In many ways, a bucket-type model can be thought of as one grid-cell of a distributed model. Many of the same physical principles and governing theories are applied (overland flow, infiltration, etc). Rather than representing a watershed as a collection on interconnected cells, a single "bucket" with an upper and lower reservoir (representing the unsaturated and saturated zones) is used. In place of a stream channel routing scheme, the runoff signal is convolved with a unit-hydrograph (two parameter inverse gamma function) to produce a realistic hydrograph. Advantages of bucket models are that

they contain few parameters and are computationally efficient, meaning that iterative inverse methods can be applied to estimate parameters. The goal of this particular chapter is to invert for basin-mean precipitation following Henn *et al.* (2015).

While applied to a different basin, as we will see the conceptual model more skillfully produces hydrographs. That, however, is not to say that the conceptual model is better, as the questions being asked are different. Conceptual models may be useful for inverse applications (where many model iterations are required) or for purely predictive applications; distributed models offer a framework for testing a range of other scientific hypotheses, such as the influence of vegetation removal on streamflow generation, impacts of forcing heterogeneities, or other questions related to the spatial organization of parameters and forcings.

1.5 Land-Atmosphere Interaction Phenomena

The first chapter of the dissertation examines roles of soil-moisture/atmospheric coupling, in part building off of Rudisill *et al.* (2021). These processes fall under the wider umbrella of "land atmosphere interactions". The atmosphere exchanges heat, moisture, and momentum fluxes with the land. It has been well-recognized that accurately predicting near-surface climate fields (temperature, humidity, etc) is highly dependent on the land-surface conditions (soil moisture, vegetation cover, albedo) but in many ways land-atmosphere couplings remain poorly observed, theorized, and understood (Santanello *et al.*, 2018). Among the central metrics used in land-atmosphere interactions studies is the evaporative fraction:

$$EF = \frac{LE_{sfc}}{LE_{sfc} + SH_{sfc}} \quad (1.14)$$

Where LE_{sfc} is the latent heat flux and SH_{sfc} is the sensible heat flux. The energy balance at the land surface is given by:

$$LE_{sfc} + SH_{sfc} = SW_{net} + LW_{net} + G \quad (1.15)$$

Where SW is the shortwave radiation, LW is longwave radiation, and G is the ground heat flux. All terms have units of w/m^2 . During the daytime and assuming there is no water limitation, the evaporative fraction can be thought of as the fraction of incoming energy that is partitioned into evaporation. Many of the studies of land atmosphere interactions relate to diagnosing how the evaporative fraction influences the growth of the day-time planetary boundary layer (PBL) and convection. The PBL is the lower layer of the troposphere that is directly impacted by heat/moisture fluxes from the land surface, and is typically characterized by a temperature inversion at its top (an increase in temperature with height). The PBL tends to grow with lowering evaporative fractions (more sensible heating), as this results in stronger thermals rising from the surface. Understanding some of these processes, and the role of moisture on PBL development and convection, begins with parcel theory, which is described in most meteorology or thermodynamics textbooks such as Petty (2008). The work of Findell & Eltahir (2003) establishes some simple metrics for diagnosing the impact of the evaporative fraction on PBL growth and convection which are used in Chapter 1. Santanello *et al.* (2018) provides an up to date review of many of the challenges and current opportunities of land-atmosphere interaction research.

CHAPTER 2:

**EVALUATING LONG TERM ONE-WAY
ATMOSPHERE-HYDROLOGY SIMULATIONS
AND THE IMPACTS OF TWO-WAY
COUPLING IN FOUR MOUNTAIN
WATERSHEDS**

2.1 Abstract

Joint hydrologic-atmospheric model frameworks offer novel insights into the terrestrial hydrologic cycle and the potential for improved predictive capabilities for stream discharge and other hydrologic fluxes. In this study, we examine both one- and two-way coupled integrations of the Weather Research and Forecasting (WRF v3.8.1) atmospheric and WRF-Hydro (v5.0) hydrologic models for four 1000-2000 km² snow-dominated mountain watersheds (1500-2100 m mean elevation) in Idaho's Rocky Mountains. In watersheds where anthropogenic withdrawals are minimal (3 of 4 watersheds), we simulate stream discharge with high confidence ($KGE > .63$) for a 20 year period in the uncoupled scenarios, and find that WRF winter precipitation accumulations have less than 15% average error for all but two of the four-

teen comparison NRCS Snotel sites. However, annual streamflow biases are highly correlated ($r^2 > .8$ in some cases) with the annual errors in WRF cold-season precipitation, suggesting that process representation of winter orographic precipitation limits hydrologic predictability. In the second part of the study, we evaluate the potential for “two-way” model coupling to influence hydrologic predictability by examining a two month case-study period with active spring season convective precipitation. We quantify the impacts of resolving hillslope-scale soil water redistribution on the ABL, and find that while resolving overland and saturated subsurface soil moisture flow influences soil moisture distributions and surface energy fluxes, the impact on precipitation is non-systematic, as precipitation is generally atmospherically controlled during the study period. Consequently, future efforts should focus on improving winter orographic process representation, as streamflow is highly sensitive to errors in these processes.

2.2 Introduction

Efforts in the last decade have worked towards full terrestrial-atmospheric water cycle prediction, where atmospheric models are coupled either “one-way” or “two-way” with sophisticated hydrologic or groundwater models (Maxwell *et al.*, 2011; Gochis *et al.*, 2013; Butts *et al.*, 2014; Shrestha *et al.*, 2014). While many of these efforts are motivated by improving hydrologic prediction, models of coupled hydrologic-atmospheric systems can help reveal the extent to which coupled land-atmosphere processes influence observed hydrologic variables like streamflow, as well as the spatiotemporal scales at which those coupled processes act. Distributed, physically-based hydrologic models use more sophisticated representations of hillslope-scale soil water exchanges not included in most land surface representations used in atmospheric mod-

els (Stensrud, 2009; Dudhia, 2014), and may in particular solve for lateral soil water diffusion, overland flow, re-infiltration, and stream channel flow with varying degrees of complexity (Ivanov *et al.*, 2004; Maxwell *et al.*, 2009; Hamman *et al.*, 2018). In “one-way” coupling, the meteorological states from an atmospheric model are used directly as forcing inputs into a distributed hydrologic model, with no feed-back from the hillslope-scale hydrologic model to the atmospheric boundary layer (ABL). Whereas in “two-way” coupling, the surface fluxes computed by the hydrologic model are used as the lower boundary by the atmospheric model (Figure 2.1). One way coupling approaches are of particular interest in mountain regions, where large spatial heterogeneities of mountain precipitation challenge distributed meteorological forcing products, which can differ substantially for a given watershed, despite using similar input datasets (Henn *et al.*, 2018) A growing body of evidence suggests that convection-permitting atmospheric models, which account for nonlinear atmospheric motions induced by topography, may outperform these techniques for estimating precipitation in mountain watersheds, though cross-compartment validation methods are required to better address uncertainties (Lundquist *et al.*, 2019).

The potential and unknown significance of soil moisture-precipitation feedbacks on atmospheric circulations, arising from the influence of lateral moisture redistribution, are one mechanism through which two-way coupled models may improve hydrologic prediction. These interactions arise not only from a precipitation recycling perspective, where local evapotranspiration contributes to local precipitation, but also through indirect thermodynamic mechanisms where drier/moister atmospheric boundary layers inhibit or promote atmospheric convection and subsequent precipitation, which may or may not be sourced from local evaporation (Seneviratne *et al.*,

2010). Both observational and model based investigations suggest that both wet and dry soil biases can enhance/impede subsequent precipitation (Findell & Eltahir, 2003; Tuttle & Salvucci, 2016). Recent two-way coupled modeling efforts have found that streamflow is only minimally influenced by land-atmosphere coupling mediated through resolved sub-surface flow (Senatore *et al.*, 2015; Arnault *et al.*, 2016; Kerandi *et al.*, 2018). Givati *et al.* (2016) examined both one and two-way coupled simulations of streamflow for an 800 km² watershed basin in Israel and found modest improvements in streamflow simulation with two-way coupling, but did not evaluate the physical mechanisms that caused it. Lahmers *et al.* (2020) examined impacts of resolved lateral flow during North American Monsoon events in New Mexico, and found that two-way model coupling increased the degree of convective organization, increasing precipitation. Other by-products of two-way coupling are that the total evapotranspiration values can increase, relative to standard land surface models, since overland flow can re-infiltrate soil columns as opposed to only being lost as runoff. In this study, we examine one and two-way coupling of the Weather Research and Forecasting atmospheric model (Skamarock *et al.* (2019); WRF) and WRF-Hydro hydrologic model Gochis *et al.* (2018) for four snow-dominated watersheds in Idaho’s Rocky Mountains. We evaluate the models against United State Geological Service (USGS) stream discharge measurements, National Resource Conservation Service (NRCS) Snotel precipitation gauges, and National Weather Service (NWS) radiosonde observations of atmospheric temperature and humidity profiles. We address the following questions: (1) to what extent can one-way coupled atmospheric-hydrologic modeling systems simulate stream discharge in snow dominated mountain watersheds, (2) what can the error characteristics tell us about atmospheric model performance at climatic

timescales, and, (3) what is the added-value, from a hydrologic perspective, of resolving lateral soil water redistribution via two-way coupling? To answer these questions, we take two approaches. In the first section of the study, we evaluate streamflow at a climatological timescale (20-years) in one-way coupled mode. WRF is run at a convection-permitting resolution (1 km horizontal resolution). The outputs of precipitation, air temperature, humidity, surface pressure, winds, short and longwave radiation from WRF are used as inputs to the WRF-Hydro model. Hydrologic model errors are a combination of forcing input, model structure, and parameter uncertainties (Gupta *et al.*, 2008). Thus, evaluating the modeled stream discharge against observations indicates, to some degree, how well the atmospheric model captures the dominant hydroclimatic processes. The experimental framework also allows us to evaluate potential deficiencies in model physical process representation and to identify poorly simulated meteorological events. The operating assumption of the first experiment is that soil moisture-precipitation coupling is not significant, so the impacts of running a two-way coupled integration of WRF and WRF-Hydro are marginal. In the second section, we question this assumption by developing and running a numerical experiment for a shorter (two month) two-way coupled model integration from May-June of 2018. While a longer model evaluation may be desirable, this experiment is instead intended as a case-study, representing a time-period when two-way coupling is likely to be significant. The goal is to isolate physical two-way coupling mechanisms rather than aggregate or long term hydrologic effects. This experiment evaluates three alternative configurations with two different initial soil moisture conditions and the overland and subsurface flow routing features in WRF-Hydro turned on/off. We interrogate the meteorological states, convective available potential energy (CAPE),

and ABL characteristics for each model run. We apply the CTP-HILow framework developed by Findell and Eltahir (2003) to quantify soil moisture precipitation coupling regimes. The HiResInit initial conditions shorter time-period can help identify important mechanisms that would manifest in longer integrations, and inform future efforts for two-way coupled modeling at longer timescales. The combination of these experiments increases fundamental, process level understandings of mountain hydrologic cycles in the Western United states, with implications for water resources management and model development.

2.3 Data and Methods

2.3.1 Study Area Geography, and Hydroclimate

We conduct the analyses in the Boise and Payette River Basins in the Rocky Mountains of the interior Northwest of the United States (Figure 2.2). These snow-dominated watersheds drain from the high elevation Sawtooth, Soldier, and Danskin mountain ranges and are underlain by fractured Cretaceous granodiorite and shallow soils that are primarily sandy-loam and loamy-sand (Figure 3b; STATSGO Soil Database, Miller & White (1998)). The Sawtooth ranges are oriented approximately N-S and the highest peaks are over 3000 meters tall. We investigate the South Fork of the Payette (SFP), the main fork of the Boise River at Twin Springs (MFB), South Fork of the Boise (SFB), and Mores Creek (MC) watersheds. The SFP, MFB, and SFB drain from the high-elevation Sawtooths, whereas the headwaters of MC are lower elevation and are more southerly facing. The average elevations of the watersheds are between 1488-2146 m and have drainage areas between 1028- 2154 km² (Table 1) The land-use types are primarily evergreen needleleaf and open shrublands, based on the

USGS 20-category dataset (Figure 3d). The discharge hydrographs of the Payette and Boise River basins (including Mores Creek) are typified by spring/summer snowmelt peaks decaying to low baseflows and sporadic hydrologic responses of rainstorms in the fall season. Occasionally, large mid-winter stream runoff occurs as the result of rain-on-snow events, which are themselves often the result of Pacific-sourced atmospheric rivers. Atmospheric rivers contribute approximately one-fifth of the average annual winter precipitation for this region (Rutz *et al.*, 2015) and can cause large snow melting events Rudisill *et al.* (2021). The Idaho Rockies are, in general, sufficiently far North to avoid southerly moisture transport associated with the North American Monsoon and receive relatively little precipitation during the summer months. The NRCS operates snow pillow, precipitation, and meteorological instrument sites across the Western United States (<http://www.wcc.nrcs.usda.gov/snow>), 14 of which are in the vicinity or within the study watersheds (Figure 2). Of the fourteen Snotel sites examined, and across all available data, approximately 80% of the water-year precipitation arrives before May 1.

2.3.2 Model Descriptions

The WRF-Hydro modeling system (Gochis *et al.*, 2018) is designed to run both as a stand-alone, integrated hydrologic model, and to facilitate the coupling of additional hydrologic processes with the WRF Atmospheric model. The Noah-MP (Niu *et al.*, 2011) land surface option within WRF-Hydro is used for both the one-way and the two-way model experiments described in the next section. Noah-MP improves upon the original Noah model with a 'semi-tile' approach for calculating the turbulent fluxes for vegetation, bare soil, and snow surface separately in addition to a multi-layer snow model. Noah-MP uses a one-dimensional form of Richard's

equation to solve for infiltration within the soil column. WRF-Hydro adds model physics options for simulating surface overland flow, saturated subsurface flow, channel routing, and a conceptual baseflow model to the existing Noah-MP land surface model. In Noah-MP, precipitation/snowmelt water that exceeds the soil infiltration capacity is treated as overland flow and lost from the system. Whereas in WRF-Hydro, infiltration excess water is allowed to "pond" above the soil surface. When the depth of ponded water exceeds a specified retention depth (a tunable parameter), overland flow occurs as a shallow diffusive wave. We use the one-dimensional, steepest-descent (in the direction of the most-negative head gradient) overland flow routing option. The overland flow that reaches a channel grid point contributes to streamflow or otherwise contributes to ponded water depth. Ponded water that does not exceed the retention depth remains ponded and may subsequently infiltrate into the soil column, become overland flow or evaporate. Once a soil column becomes saturated, WRF-Hydro allows for 1 or 2-dimensional lateral redistribution following the model developed in (Wigmosta & Lettenmaier, 1999). We use the 1-dimensional steepest descent option for routing subsurface saturated flow. Water that infiltrates through the lowest layer of the soil column is added to an empirical baseflow bucket model. The 'bucket' spills beyond a maximum depth parameter, and the discharge is routed to all corresponding channel grid cells based on a groundwater mapping layer. We use the Muskingum-Cunge channel routing formulation for modeling flow in the channel. There are several major assumptions of the WRF/WRF-Hydro modeling system worth mentioning. Soil columns are uniformly two-meters deep across the entire modeling domain. Fersch *et al.* (2020) notes that this is not likely in mountains with thin soils, and addressed this challenge by using spatially varying soil depths, but

this modification is not implemented in this study. Moreover, deep groundwater feedbacks are also not represented. Water that drains through the bottom of the soil layer is lost to the groundwater bucket and will not re-infiltrate into the soil column. For the one-way experiments, we use WRF v3.8.1 to generate meteorological forcings for input into WRF-Hydro v5.0. Table 2 describes the WRF model physics parameters used in this study. Climate Forecast System Reanalysis (CFSR; Saha *et al.* (2010)) is used as lateral boundary conditions. CFSR has a 0.5 degree horizontal resolution and 40 vertical levels. For practical reasons, WRF v3.8.1 was used in the one-way experiments, whereas a more recent version of WRF (v4.1.2) serves as the atmospheric model in the two-way coupled with WRF-Hydro (also v5.0) experiments presented in part two. In both instances, we use the Noah-MP land surface model options. WRF is run with an inner grid resolution of 1 km and convective parameterizations are turned off. Two nested grids are set up with "one-way" nesting (the inner grid does not feed-back to the outer grid). For the two-way experiments (described in the next section), the "ndown" utility is used to generate lateral boundary conditions from the previously run WRF output files for the outer grid. The NCAR technical note and the references therein describe each physics option in greater detail (Skamarock *et al.*, 2019).

2.4 Experiment Descriptions

2.4.1 Uncoupled Long-Term Runs

In Part 1, we evaluate the simulated discharge from the one-way WRF/WRF-Hydro model for Mores Creek (MC), North Fork Boise (SFB), Main Fork Boise (MFB), and South Fork Payette (SFP) river basins for water years 1995-2014. The

WRF model integrations were completed as 20 parallel runs, each preceded by a two-week spinup in advance of the start of each water year (starting October 1st). This afforded more expedient completion of the simulation. WRF output variables are written out hourly. WRF-Hydro supports running multiple watersheds at once, but we run each basin separately, using model domain files downloaded from the CUAHSI National Water Model (“NWM”) subset tool (<https://subset.cuahsi.org/>). WRF-Hydro can be operated in a variety of configurations, for this study, we use the “National Water Model” run settings described in the WRF-Hydro documentation Gochis *et al.* (2018). We calibrated land surface and hydrologic parameters against observed streamflow values for each of the four streams. Hydrologic parameters for each basin were calibrated, independently, for a two year period. We use the Dynamic Dimension Search (DDS) scheme, an automated stochastic global optimization algorithm developed by Tolson & Shoemaker (2007). An implementation of the calibration code is available on github Rudisill (2021). The calibrated parameters include soil-physics related parameters (bexp, smcmax, dksat), runoff related parameters (lksatfac and slope), baseflow parameters (zmax, zinit, Coeff), and snow parameters (mfsnow; Table 3). The “mfsnow” parameter controls the shape of the subgrid snow depletion curve which is a functional relationship between snow water equivalent and snow covered area, and therefore influences grid-cell total albedo and energy available for snowmelt and evapotranspiration. The DDS algorithm uses a stochastic parameter selection methodology, where each iteration perturbs a randomly chosen subset of the model parameter space. We use a multiplicative perturbation value following a prescribed Gaussian distribution, centered at one for each parameter. We performed 200 iterations of the DDS algorithm run for two water years, and

evaluate KGE scores for the final 9 months of simulation, thus allowing additional spin-up of the updated model state before evaluating model performance. We use the Kling-Gupta Efficiency (Equation 1; KGE hereafter) as the objective function for the model calibration. A perfect value of KGE is one, whereas using the mean value of a dataset (compared against itself) will produce a KGE of -.46. The KGE score weights contributions of model/observed correlation, model bias (expressed as ratios of the respective mean values), and ratios of model and observed variance, equally (Knoben *et al.*, 2019). The model bias term in the KGE equation is the same as the percent-bias term given in Equation 2. The Nash-Sutcliffe Efficiency (NSE; Equation 3) is also used as a diagnostic, but not as an objective function in the calibration process. The WRF-Hydro model is configured to output discharge on an hourly basis. To compare against USGS stream gauge observations, we aggregate both the model and observed flow to a daily-mean flow. Consequently, we do not evaluate the diel cycle of discharge, though this information likely contains interesting insights about model behavior.

Fully-Coupled Runs

We run WRF two-way coupled with WRF-Hydro for a two-month spring period (May-June) to investigate the potential ramifications of subsurface lateral soil moisture redistribution on atmosphere conditions in the region. While a longer model integration is desirable, it is computationally expensive. This period was specifically chosen since there is significant precipitation and deep cumulus cloud development (visible from satellite imagery; see <https://worldview.earthdata.nasa.gov/>). We hypothesize that if soil moisture-mediated land-atmosphere couplings occur, they are

more likely to be active during this time period than other seasons. Specifically, we anticipate that there are larger gradients of both soil moisture in the spring season (owing to, for example, differential snow melt on North- and South-facing hillslopes), which in turn influence the partitioning of sensible and latent heat into the ABL and could potentially feedback into precipitation initiation or enhancement through soil-moisture ABL interactions. We create three scenarios, each initialized May 1, 2018 from model restart files and run for the following two months to test impacts of lateral flow impacts on the ABL. The "Control" scenario has overland and subsurface flow turned off. The "Routing" scenario has overland and subsurface flow turned on. These two scenarios use land surface initial conditions created by running the WRF-Hydro offline (forced by WRF) for 10 consecutive years with subsurface and overland flow turned off. A third scenario, called "HiResInit" (i.e, high resolution initialization), has overland and subsurface flow turned on during the spinup process to test the impacts of the long term soil moisture state that would result from including lateral flow. Thus, "Control" and "Routing" have the same initial conditions and "Routing" and "HiResInit" have the same physics options, but different initial conditions. Soils are close to saturated in many of the mountain regions, where snowmelt has already occurred, for both scenarios. Yet, turning on overland and subsurface flow during the spin-up process ("HiResInit") generally makes the valleys wetter and the peaks drier (Figure 4). We use the convective triggering potential (CTP) and humidity-index (HI_{Low}) framework to quantify the potential for soil-moisture influenced precipitation (Findell and Eltahir, 2003). The theory is based on the idea that the air mass approximately 1-3 km above the ground surface in the morning hours will become incorporated into the PBL during the day. The properties of this air mass

then determine whether wet or dry soils (as they impact heat and moisture fluxes) will promote convective precipitation, or if convection is independent of surface influence (i.e, atmospherically controlled). The CTP measures atmospheric stability by integrating the differences in environmental temperature between 100 and 300 hPa above the surface and the moist-adiabat starting at the same origin. Negative CTP values indicate that the atmosphere is too stable to convect. Large positive values mean that the atmosphere is close to dry-adiabatic, and intermediate positive values mean the atmosphere is closer to moist-adiabatic. According to the underlying theory, wet soils (higher latent heat flux) can promote precipitation at intermediate CTP values by increasing moist static energy and lowering the level of free convection (LFC), whereas dry soils can promote precipitation at high positive CTP values by growing the boundary layer height through enhanced sensible heating to the level of the LFC. The HILow metric is simply the difference between the environmental and dew-point temperatures at 50 hPa and 150 hPa above the surface (Equation 4). Lower values of HILow indicate a moister atmosphere. Santanello et al. (2018) provides an additional description of the CTP-HILow framework and summarizes several applications from the literature. The thermodynamic calculations are performed using the MetPy python libraries.

2.5 RESULTS

2.5.1 One-Way Coupled Long-Term Streamflow Simulations

Simulated Streamflow Performance

The calibrated and uncalibrated (“NWM”) parameters and hourly WRF meteorological forcing outputs (Figure 1) are used to run the model for the twenty year

study period. The model evaluation period begins October 1 1994 (the start of water-year 1995) and ends August of 2014. Streamflow simulation performance is evaluated against daily USGS observations. For both years simulated, the discharge peaks approximately in May and June, and baseflows are reached roughly by August. The same calibration process method was applied for each basin independently, but for brevity, we only show and describe the SFB calibration (Figure 5). Prior to calibration, the model exceeds the maximum discharge by over 100 m³/s in many cases. Following calibration, model hydrographs have good temporal correlation (pearson's $r > 0.84$), and the model shows KGE values greater than 0.6 for the MBR, SFB, and SFP for the entire 20 year simulation (Table 4). While each basin was only calibrated for two years, calibration improved the overall KGE scores for the full twenty-year period. The SFB has the highest and Mores Creek the lowest respective skills scores (NSE of 0.763 and -0.681, and KGE of 0.693 and -0.154 respectively). Another measure of timing, the median difference in dates-of-peak flow, are the same for SFP and SFB and only one day off in the MFB, and 5 days for MC. The distributions of date-of-peak errors contain a notable outlier value in water year 1997 (Figure 6) when a large atmospheric river event caused flooding throughout the Pacific Northwest in the early winter (Leung & Qian, 2009). However, the model does capture several similar early-winter runoff events (the MFB in 2011, for example). Streamflow percent biases (Equation 2) vary between each basin. The annually averaged biases are positive (more model streamflow than observed) with the exception of the SFP. On average, there is 60% higher modeled streamflow in MC than what is observed. The MFB shows only 5% bias compared to the observed. Aggregating biases at the daily level and examining the histograms (Figure 6, right column) reveals bi-

modal distribution, reflecting different errors for baseflow and peak flow conditions. The modeled hydrographs generally overestimate discharge during the runoff season and underestimate summer baseflow conditions. The impacts of calibration can be seen in the distributions of daily errors (Figure 6) as peaks shift closer to zero by varying degrees (especially noticeable for the Mores Creek watershed, where positive daily biases are reduced). The streamflow bias can be directly attributed to errors in the mass-balance of the watershed, governed by the hydrologic budget (Equation 5). Streamflow biases imply that water inputs are over/underestimated or that the partitioning of water between ET, storage, and streamflow is incorrect.

Correlations between Runoff and Model Forcing Errors at Meteorological Stations

The upper limit of runoff is bounded by water input into the system, in the form of precipitation (rain or snow and its subsequent melt) or changes in storage (Equation 5), the latter of which is likely small on the scale of a given water year. If energy is not limiting, then water losses to ET are also constrained by the available soil moisture and thus antecedent precipitation. Therefore, to evaluate the sources of WRF-Hydro streamflow biases, we compare errors in the WRF model precipitation against observed NRCS Snotel Observations. No discrimination is made between rain and snow, though hydrometeor phase does impact measurement error (Harpold *et al.*, 2017). The percent bias in the accumulated October - May precipitation (i.e., the cold-season, during which the majority of precipitation falls) between WRF and the Snotel sites examined is less than 20% on average for all sites. Nine of the 14 sites have less than 10% error in cold-season accumulated precipitation, on average. There

is no clear relationship between Snotel site elevation and WRF precipitation bias. Errors in the calibrated WRF-Hydro simulated stream discharge for the SFP, SFB, and MFB watersheds are highly correlated with the cold-season WRF-derived precipitation errors at Snotel sites, with significant linear relationships ($p < 0.05$) for each site that is located in, or in the vicinity of, each respective watershed (denoted by gray boxes in Figure 6). Moreover, the slopes of the significant lines are all positive, meaning that years forced with higher than observed winter precipitation tend to result in higher than observed annual streamflow. Mores Creek streamflow biases are not correlated to either Snotel site that is within or very close to the watershed boundaries, but they are correlated with 4 other sites outside of the watershed with low r^2 values (between 0.46 and 0.66) compared to the other watersheds. Errors in accumulated precipitation at site 550 (Jackson Peak) are correlated with streamflow errors for each watershed, and are particularly high for the SFP and MFB watersheds (r^2 of 0.90 and 0.87 respectively). When the same analysis is applied to include the entire water-year (not shown, October through September), precipitation, the pattern is largely the same, but the highest r^2 values are reduced (maximum of 0.70). This could suggest that relatively more cold-season precipitation is partitioned into runoff, versus late season precipitation that may become lost as ET. Errors in temperatures at Snotel stations were also investigated. Temperature errors correlated much more weakly with discharge bias than for precipitation errors. It is worth noting that time series of Snotel temperature may have systematic biases over time related to sensor calibration methodologies and other factors (Oyler *et al.*, 2015), though visual inspection did not make such errors apparent. Correlations between streamflow bias and temperature are negative, meaning that when WRF temperatures are warmer than

observed, simulated streamflow tends to be lower than observed, possibly because of increased evaporative losses. The relationship between peak flow-timing and errors in temperature across different time periods were also examined, as too-cold or too-warm temperatures could influence rates of snowmelt (and thus timing of streamflow) through sensible heat fluxes, but there was no significant relationship found.

2.5.2 Two-Way Coupled Model Analysis

May-June, 2018 Meteorological Analysis

Comparing the atmospheric profiles of temperature and humidity from the WRF Control run against the National Weather Service radiosonde observations (located in Boise, see Figure 2 for location) shows generally good agreement, though upper level humidities show some discrepancies (Figure A.1). Time series spatial averages of CAPE, 500 hPa geopotential heights, two meter air temperature, and precipitation illustrate the meteorological conditions during the two-month model period (Figure 8) Convective available potential energy (CAPE) is a widely-used measure of buoyant energy of rising, moist air parcels from the boundary layer, and is often correlated with rainfall intensity (Eltahir & Pal, 1996). CAPE is similar to CTP, except it is calculated between the level of free convection (if present) and level of neutral buoyancy, as opposed to a fixed 100 hpa and 300 hPa above ground levels in the morning. Values are averaged spatially across a square region encompassing the four study watersheds (shown in Figure 4). There is a significant sequence of precipitation during the second half of May and lasting into early June. The May precipitation events correspond to a lowering of the 500 hpa geopotential height during this time period (5800 to 5550m on May 15). There is very little variation in heights between the three model scenarios on the order of 1-2m (within the resolution of the line-widths). Averaged

CAPE values can reach moderately high (600 J/kg) values averaged across the study area and correspond with time periods of rainfall. Examining the diurnal cycle (not shown) shows that both CAPE and precipitation peak in the late afternoon. There is slightly less accumulated precipitation in the HiResInit scenario compared to the other two (101 mm versus 106 mm in the Control) and slightly less domain averaged peak CAPE values (as much as 100 J/kg higher in the Control) particularly between May 15 - June 1, however the Routing scenario has slightly lower peak CAPE values than the Control. The Control case has typically lower peak values of latent heat flux, and correspondingly higher sensible heat, on average, than the scenarios with lateral flow turned on. Average temperatures vary between ± 1 C between scenarios. Examining spatial data averaged across time (May 1 to June 28) for the same domain shows the relationship between topography, soil moisture, ABL characteristics, and precipitation (Figure 9). The second soil moisture layer is deeper and more slowly evolving than the top layer, and thus a better metric for comparing impacts of resolved lateral flow than the top layer. For comparison, the time averaged sensible/latent heat fluxes, temperature, and afternoon CAPE are differenced between the Control and Routing/HiResInit scenarios, in addition to the percent difference (relative to Control) of afternoon PBL height. Averaged across the entire two months, the difference in soil moistures between the Control and Routing scenario are relatively small ($\pm 20\%$ degree of saturation). The impacts of turning on overland/subsurface flow are more obvious during the first weeks of the model run (not shown), where soils are closer to saturation and runoff/lateral exchanges are occurring. The differences in the initial spatial organization of soil moisture (in the HiResInit scenario) persist for the two months (Figure 9). Still, the Control run is generally drier in the low

elevation foothills region and the peaks of mountain tops than the Routing scenario, which started with the same initial soil moisture conditions. Drier regions correspond with more energy partitioned into sensible heat than latent heat and higher two meter air temperatures. The average afternoon PBL height varies by roughly +/- 20%, between the Control, HiResInit and Routing scenarios, and correlates with increases in sensible heating in the southwest corner of the domain. The highest accumulations of precipitation (>200 mm) are in high elevation (>2000 m) regions. While the domain mean precipitation is very similar, there are substantial differences in the spatial organization of precipitation across scenarios. Likewise, average CAPE varies spatially, but the differences in CAPE are poorly spatially correlated with differences in precipitation accumulation, as the highest precipitation accumulations are in the highest elevation zones, where CAPE values are low. In this situation, mechanical uplift forced by the orography is likely the dominant precipitation generating process as opposed to surface based moist convection (Kirshbaum *et al.*, 2018). To isolate periods of convective-dominant precipitation and examine surface influences, we can examine low-elevation (<2000 m) grid-cells during the period of highest precipitation (May 15 - June 1). Average daytime CAPE values tend to increase with higher average daytime evaporative fractions (the ratio of latent heat to the total turbulent flux), and can correspond with higher precipitation accumulations (Figure 10). Moreover, the degree of soil saturation strongly influences the evaporative fraction and thus latent heating of the boundary layer. In the absence of other effects, higher evaporative fractions promote shallower boundary layers (Santanello *et al.*, 2018), but advection of outside air masses likely complicates this view in this case.

Examining Soil Moisture Influences using the CTP-HILow Metric

The CTP-HILow framework helps explain some of the lack of influence of soil moisture conditions on total precipitation accumulations. The metric is computed each day during the months of May and June (Figure 11) using morning (5:00 a.m. local time) model outputs and atmospheric soundings (available in Boise). A high elevation (2000 m) location in the Stanley basin (Figure 2) is also chosen though observations are not present there. These locations are also chosen because they are in flat valleys, consistent with the assumptions of the CTP-HILow framework development. The majority of morning soundings for both Boise and Stanley locations are in the “atmospherically controlled” or “too stable” zones as defined by (Findell & Eltahir, 2003), meaning that surface fluxes of latent/sensible heat into the boundary layer alone are not sufficient to promote convective precipitation. Wet soil advantage zones have positive CTP values and HILow values between 5 and 10C, whereas dry-soil zones have CTP values greater than 150 J/kg and HILow values between 10 and 15C. Precipitation can still occur, according to the theory, but must be associated with larger scale systems or 3-d wind effects (Findell, 2003). The proportion of wet-soil advantage and dry-soil advantage days is approximately equal (6 days averaged across scenarios) for the Stanley location, and the Boise location has more days in the dry-soil advantage zone (8 versus 3 days averaged across scenarios).

2.6 DISCUSSION

2.6.1 One-Way Model Coupling Improves Precipitation Process Understanding

One-way model coupling between atmospheric and hydrologic models that simulate streamflow allow for enhanced process level understandings of the terrestrial-atmospheric water cycle. By preserving meteorological states, and not imposing ad hoc corrections to simulated hydrologic input fields, one way coupling allows for a tractable method of attributing hydrologic errors to parent atmospheric processes, offering novel insights for model improvements. The hydrologic model performance (Table 4) is similar to that of similar studies. Senatore *et al.* (2015) found NSE values of 0.80 for one year of simulation with WRF-Hydro run with observation based meteorological forcings for a basin in Italy. Holtzman *et al.* (2020) found that WRF and NoahMP combined with a simple channel routing-rule and groundwater model achieved similar NSE values after manual parameter-tuning and adjustments to the WRF model microphysics. Fersch *et al.* (2020) reports NSE values of 0.64 for a using WRF/WRF-Hydro watershed in Germany for a five month period. We are not aware of other studies that have used 20 years (or greater) of a convection-permitting atmospheric model configuration to force a distributed hydrologic model to simulate stream discharge.

Evaluating streamflow in the one-way coupled approach adds more data points for evaluating the skill of the WRF modeled precipitation, as precipitation is typically understood to be the most uncertain forcing input in snow-dominated watersheds (Raleigh *et al.*, 2015). However, other hydrologic forcing variables, including ra-

diative fluxes can also have large impacts on streamflow timing (Mizukami *et al.*, 2014). Other forcing variables beyond temperature and precipitation are not evaluated here, given insufficient long-term observational data. The highly variable annual average precipitation recorded at each Snotel station illustrates some of the complex mountain-atmospheric processes influencing hydrology in this region (Table A.1). Precipitation not only depends on elevation, but also orientation with respect to the mountain topography and atmospheric moisture delivery pathways (Swales *et al.*, 2016). Senatore *et al.* (2015) found that WRF precipitation biases increased with topographic elevation for a region in the Mediterranean, but errors are more uniformly dispersed across elevation in this study. A variety of studies have evaluated similar dynamically downscaled precipitation fields in North American mountain regions, and evidence suggests that such model integrations are skillfully capturing precipitation (Ikeda *et al.*, 2010; Currier *et al.*, 2017; Lundquist *et al.*, 2019). This study demonstrates similar performance to that of other studies. Rasmussen *et al.* (2011) found that 75% of the SNOTEL sites examined in Colorado had less than 25% absolute error in accumulated precipitation, and similarly Jing *et al.* (2017) reports that more than 95% of Snotel precipitation measurements compared against WRF estimates have less than 40mm per month of bias (the percent of the total observed is not reported). In this study, all of the sites recorded on average less than 25% absolute percent-bias in water year accumulated precipitation, and only two of the sites had greater than 15% error on average (Figure 7). However, spatial representivity of individual precipitation gauges in complex terrain, and undercatch of falling snow as high as 10-15% (Yang *et al.*, 1998) limit the conclusions that can be drawn from such comparisons about overall model precipitation performance. Incorporating stream-

flow in a one-way coupled framework allows for greater confidence interrogating the modeled precipitation fields, and the attribution of errors at a point to regional errors, since streamflow integrates precipitation inputs (and ET losses) from the entire watershed. Such insights are impossible to glean from shorter model integrations that have fewer years (or months) and less statistical power. The strong correlations ($r^2 = 0.90$ in some cases) between streamflow biases and precipitation biases have several significant ramifications. First, it suggests that streamflow simulations can be improved by enhancing the accuracy of modeling cold season precipitation (October - May), during which the majority of precipitation falls. Several windows were tested, including water-year total precipitation, but cold-season precipitation had the highest correlation. Cold-season precipitation is also a useful window for water management purposes, since many reservoir management decisions must be made in the early spring prior to peak stream runoff. Thus, in our modeling workflow, having some indication about errors in cold-season precipitation connotes a degree of skill in knowing streamflow bias at the scale of the water year. Streamflow biases are not always improved by reducing precipitation bias; in some cases (the MFB, for example), the regression lines do not always go through the origin (zero precipitation bias, zero streamflow bias (Figure 7)). This could suggest that errors in precipitation are compensating for errors in other processes that may influence water-losses, like ET, or “getting the right answer for the wrong reason” (Kirchner, 2006), though it likely reflects cross-correlation structures in the true two dimensional precipitation error fields (a high-bias in one individual point could correlate with a zero-bias averaged across the entire region, for example).

2.7 Investigating One-Way Coupled Streamflow Performance

Errors in mass balances of simulated hydrographs are influenced by two factors: 1) imperfect parameterizations and model physics approximations that influence partitioning of precipitation into ET/sublimation, storage, or losses, and 2) errors in precipitation inputs. Calibrating the land surface parameters in this study against KGE scores did reduce model biases, with the exception of the SFP. In the case of Mores Creek, weaker or non-existent correlations with precipitation errors suggest that loss processes, such as mountain block recharge Markovich *et al.* (2019) are poorly represented in the model as implemented. Unlike the other three watersheds, anthropogenic water withdrawals are also a factor in Mores creek, and can be as high as 7 m³/s (Havens *et al.*, 2019). In the context of WRF-Hydro, mountain block recharge would manifest as open lower boundaries with groundwater-bucket loss or channel loss, and anthropogenic withdrawals could also be implemented as a channel loss. Lahmers *et al.* (2019) introduced a channel loss function for WRF-Hydro, which may be one mechanism for improving the performance of Mores Creek, but testing it, or quantifying anthropogenic diversions, is beyond the scope of this study. Significant precipitation enhancement in Mores creek occurs at the small ridge in the far northwest corner of the watershed (near Snotel site 637, Figure 2), so runoff biases are also likely sensitive to the degree to which precipitation falls on/over the watershed boundary and upwind of the orographic barrier.

2.8 Impacts of Two-Way Coupling

This study examined three scenarios designed to test the role of resolving overland/subsurface flow with two-way coupling on ABL characteristics and precipitation. While two-way model coupling (Routing and HiResInit) influences the distributions of soil moisture throughout the domains and latent and sensible heat fluxes, the impacts on precipitation appear non-systematic. The HiResInit scenario has the largest spatial gradients in initial soil moisture (due to overland/subsurface flow turned on during spin up; Figure 4) and has slightly less precipitation accumulation during the two month study period, though part of this result could be from precipitating cells shifting outside of the averaging region. However, the lack of influence of soil moisture on precipitation can partially be explained by applying the CTP-HILow framework. CTP-HILow measures how “primed” the atmosphere is to convect, and the results (Figure 11) suggest that convective precipitation requires advection of upwind air masses and is thus not influenced by soil moisture mediated surface sensible/latent heat fluxes. This finding is consistent with global coupling regime analyses that likewise identifies the interior Pacific Northwest soil-moisture coupling as atmospherically controlled Ferguson & Wood (2011). It is important to note that this framework does not consider influences of background winds or complex topography, which complicate the framework assumptions (Findell & Eltahir, 2003). That being said, this early morning-recorded metric is still a good indicator of the likelihood of precipitation, as none of the days with large humidity deficits (large HILow values) precipitate in the following twelve hours for the Stanley or Boise locations (not shown). We do not directly quantify interactions between soil moisture heterogeneity and mountain-valley thermal wind circulations that can potentially influence orographic precipitation Kir-

shbaum *et al.* (2018), but the systematic effects are likely small. Forrester & Maxwell (2020) used the Parflow model two-way coupled with WRF and found that CAPE and mountain-valley wind circulation strength increased in mountain valleys as a result of lateral groundwater flow for a three month period in the Colorado Rockies. Similarly, Lahmers *et al.* (2019) examined two-way coupled WRF/WRF-Hydro during a North American Monsoon event and found relatively minor impacts on precipitation totals compared against WRF alone, similar to other studies (Senatore *et al.*, 2015; Givati *et al.*, 2016), but also that surface moistening increased peak values of convective organization across a larger study area. In this study, higher evaporative fractions within scenarios tended to increase CAPE at lower elevations (Figure 10), but differences in CAPE among scenarios is less organized and not as directly attributable to specific topographic or soil moisture features. Different model structures of soil water movement on hillslopes and groundwater (such as Parflow) may impact soil moisture distributions, but are not evaluated in this study. While two-way model coupling does not have a clear effect on precipitation, other aspects of two-way model coupling may improve streamflow predictability. Lateral water movement on slopes couples with evapotranspiration and temperature, thus influencing water balances (Figure 9). Two-way coupling could potentially improve some of the observed mass-balance found in the Mores Creek watershed (Table 4) by modifying water-losses through ET, but this hypothesis is not explicitly tested, and these mechanisms are likely less significant than improving hydrologic model structures or parameterizations for this case (accounting for anthropogenic loss, for example; Section 4.2).

2.9 CONCLUSIONS

In this study, we simulated 20 years of streamflow for four snow-dominated, mountain rivers using the WRF-Hydro model with meteorological inputs downscaled by the WRF atmospheric model at a convection-permitting resolution (one-way coupling). We did not apply further downscaling or correction factors to the meteorological forcings. We also evaluated the roles of resolved lateral flow on the surface energy balance, convection, and precipitation by running WRF-Hydro two-way coupled with WRF for a shorter, two-month time period and evaluate land-atmosphere coupling using the CTP-HILow framework developed by Findell & Eltahir (2003). The two-way coupled model experiments show that resolving lateral flow can significantly impact sensible/latent heat fluxes and temperatures. However, the influence on precipitation is modest (1-5% averaged across the domain) and appears non-systematic. Examining the CTP-HILow space shows that the majority of days are atmospherically controlled using the Findell & Eltahir (2003) criteria, and thus precipitation is not influenced by soil moisture feedbacks. While we do not make a direct comparison with the uncoupled streamflow simulations, we conclude that the added value of a fully coupled model framework for discharge prediction and water-budgets at annual timescales in this study region is likely small. This arises, in part, because the impacts of the coupled simulation on precipitation are themselves non-systematic and streamflow errors depend strongly on cold-season precipitation processes that are likely insensitive to soil-moisture coupling. Thus, soil moisture precipitation feedbacks are likely of secondary importance to other factors such as cold-season orographic precipitation processes and model structure/parameter uncertainty in simulating streamflow in these regions. We could reproduce streamflows with KGE values of >0.6 in all

watersheds without significant anthropogenic withdrawals (three out of four) for the entirety of the 20-year simulation period, and that WRF cold season (October - May) accumulated precipitation has less than 15% error for twelve of the fourteen Snotel sites examined. A key conclusion from this study is that using WRF (or similar atmospheric models configured at convection-permitting scales) can provide skillful meteorological forcings for distributed hydrologic models, yielding skillful stream discharge simulations in mountainous snow-dominated watersheds at climatological scales (20 years). This is particularly beneficial, given the sparsity observations in typical mountain watersheds. The success of simulating stream discharge lends credence to the growing body of work suggesting that dynamical downscaling methods are viable for reproducing surface hydrometeorology in complex mountain terrain, which is difficult to observe and validate. We also find that streamflow biases are highly correlated ($r^2 > 0.9$ in some cases) with WRF wintertime precipitation biases evaluated against nearby Snotel sites. These findings suggest hydrologic model improvements in these regions should focus on improvement of winter time precipitation, and that one-way coupled model frameworks can improve process level understandings of hydrologic and atmospheric processes in mountain environments.

Watershed	USGS Gauge ID	Basin Area (km ²)	Avg. Elevation (m)	Avg. Annual Specific Discharge (mm)
Mores Creek (MC)	13200000	1028.2 km ²	1488 m	229.71
South Fork Payette (SFP)	13235000	1155.1 km ²	2068 m	662.97
South Fork Boise River (SFB)	13186000	1160.2 km ²	2146 m	549.40
Main Fork Boise River (MFB)	13185000	2154.0 km ²	1950 m	520.73

Table 2.1: Table 1: Study watershed descriptions. USGS gauge identifier, basin area, average elevation, and the annual average of specific stream discharge (m³/s) measured at the corresponding USGS gauge for the study period (Water Year 1995-2014).

WRF (v3.8.1)

Lateral Boundary Conditions	Climate Forecast System Reanalysis (CFSR)
Grid resolution (km)	3, 1
W-E Dimension (cells)	340, 349
N-S Dimension (cells)	290, 328
Vertical Levels	50, 50
Timestep	15s
Microphysics	Thompson
Land Surface Model	Noah-MP; Noah-MP + WRF Hydro
Surface Layer	Monin-Obukhov
Planetary Boundary Layer (PBL)	Mellor-Yamada-Janjic

WRF-Hydro v5.0

Routing grid resolution (m)	250
Forcing input timestep (uncoupled)	Hourly
Channel routing timestep (s)	30
Snow albedo parameterization	CLASS (option 2)
Terrain Routing Timestep (s)	10

Table 2.2: Table 2: Select WRF and WRF-Hydro model configuration options used in the one and two-way coupled model experiments. Two numbers correspond with the outer/inner nest.

Parameter Name	Description
bexp	Soil beta parameter
smcmax	Saturated soil moisture value
dksat	Saturated soil hydraulic conductivity
LKSATFAC	Lateral saturated soil hydraulic conductivity multiplier
slope	Lower soil boundary drainage parameter
mfsno	Snow covered area curve parameter
Zmax	Conceptual groundwater parameter
Zinit	Conceptual groundwater parameter
Coeff	Conceptual groundwater parameter

Table 2.3: Table 3: List of WRF-Hydro parameter names, as they appear in the model files (left), calibrated for the one-way coupled model experiments.

Watershed		KGE	Pearson r	Percent Bias	NSE	Date of Peak Difference
MC	Calibrated	-0.154	.847	60.336	-.681	5.0
	NWM	-0.33	.806	95.435	-.99	26.5
MFB	Calibrated	0.80	.919	5.303	.659	1.0
	NWM	0.568	.913	5.769	.593	1.0
SFB	Calibrated	0.693	.947	7.875	0.763	0.
	NWM	0.304	.904	26.966	0.206	0.
SFP	Calibrated	0.637	.924	-17.693	0.829	0.
	NWM	0.583	.927	-5.965	0.613	5.

Table 2.4: Table 4: WRF-Hydro simulated stream discharge performance metrics for the Moore’s Creek, Boise River, Southfork Boise River, and South Fork Payette watersheds for the 20-year validation period. The modeled discharge is compared against corresponding USGS river gauges aggregated to daily values.

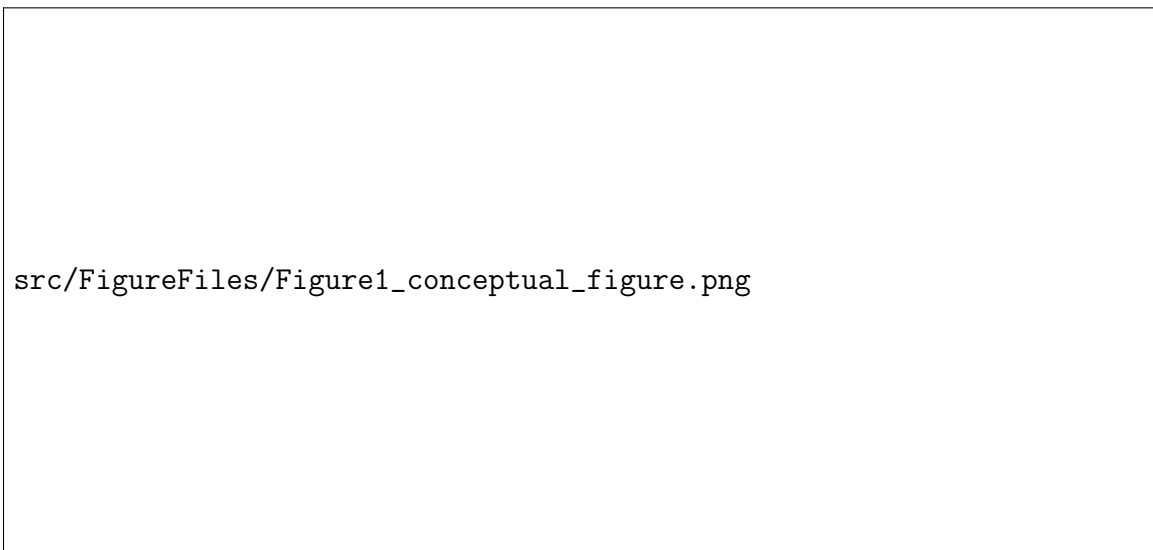


Figure 2.1: Conceptual diagram illustrating the one and two-way coupled WRF/WRF-Hydro experiments.



Figure 2.2: Satellite view of the four study watersheds (yellow outline). Clockwise from the left: Mores creek (MC), the South Fork Payette (SFP), the South Fork Boise (SFB), and the Main Fork Boise (MFB) in the center. Black diamonds indicate USGS stream gauge locations used to evaluate model discharge. Red squares indicate the location of NRCS Snotel meteorological stations.

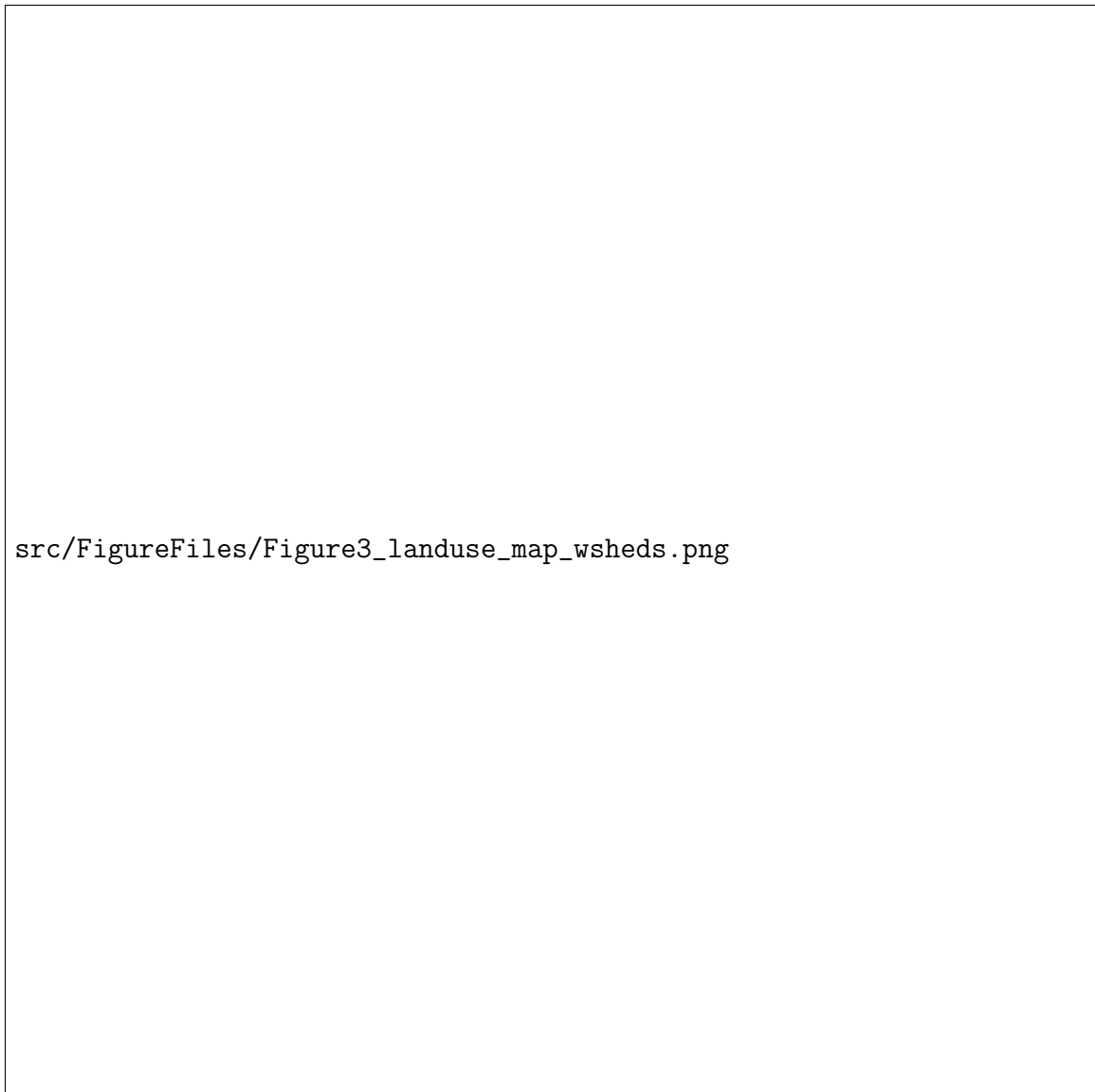
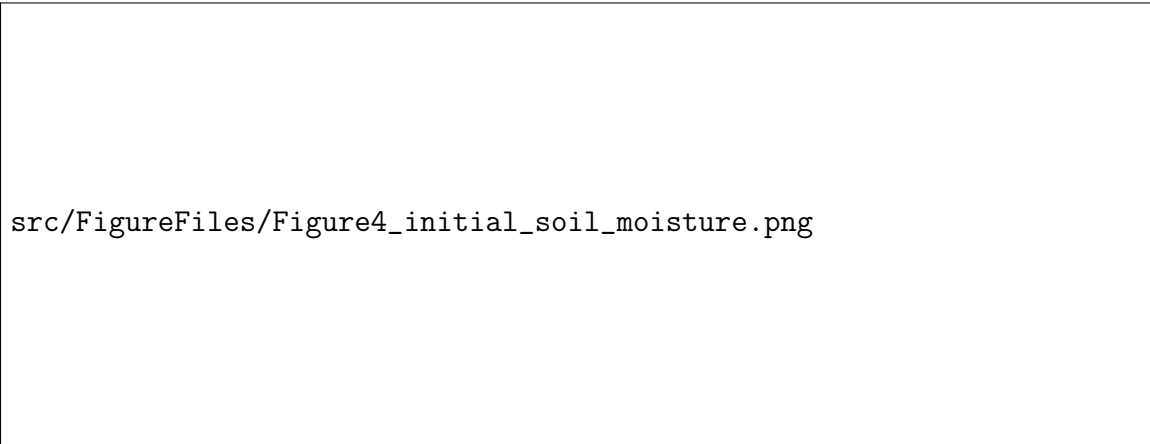


Figure 2.3: a) Percent canopy coverage, b) dominant soil category type, c) topographic elevation, and d) land-use category for the inner-WRF domain. Black outlines are the boundaries of the watersheds examined in this study



src/FigureFiles/Figure4_initial_soil_moisture.png

Figure 2.4: Initial soil saturation conditions for the top soil layer (10 cm depth) for the Control/Routing, HiResInit, and the difference between the two (right). Soil saturation is the soil moisture relative to the soil porosity.



Figure 2.5: Hydrograph demonstrating the efficacy of the calibration procedure for the uncoupled WRF-Hydro simulations for the South Fork Boise River. Showing 200 iterations of the dynamic-dimension-search algorithm. The dashed lines are observed USGS streamflow observations and the shaded region denotes the model min and max values in the calibration iterations.

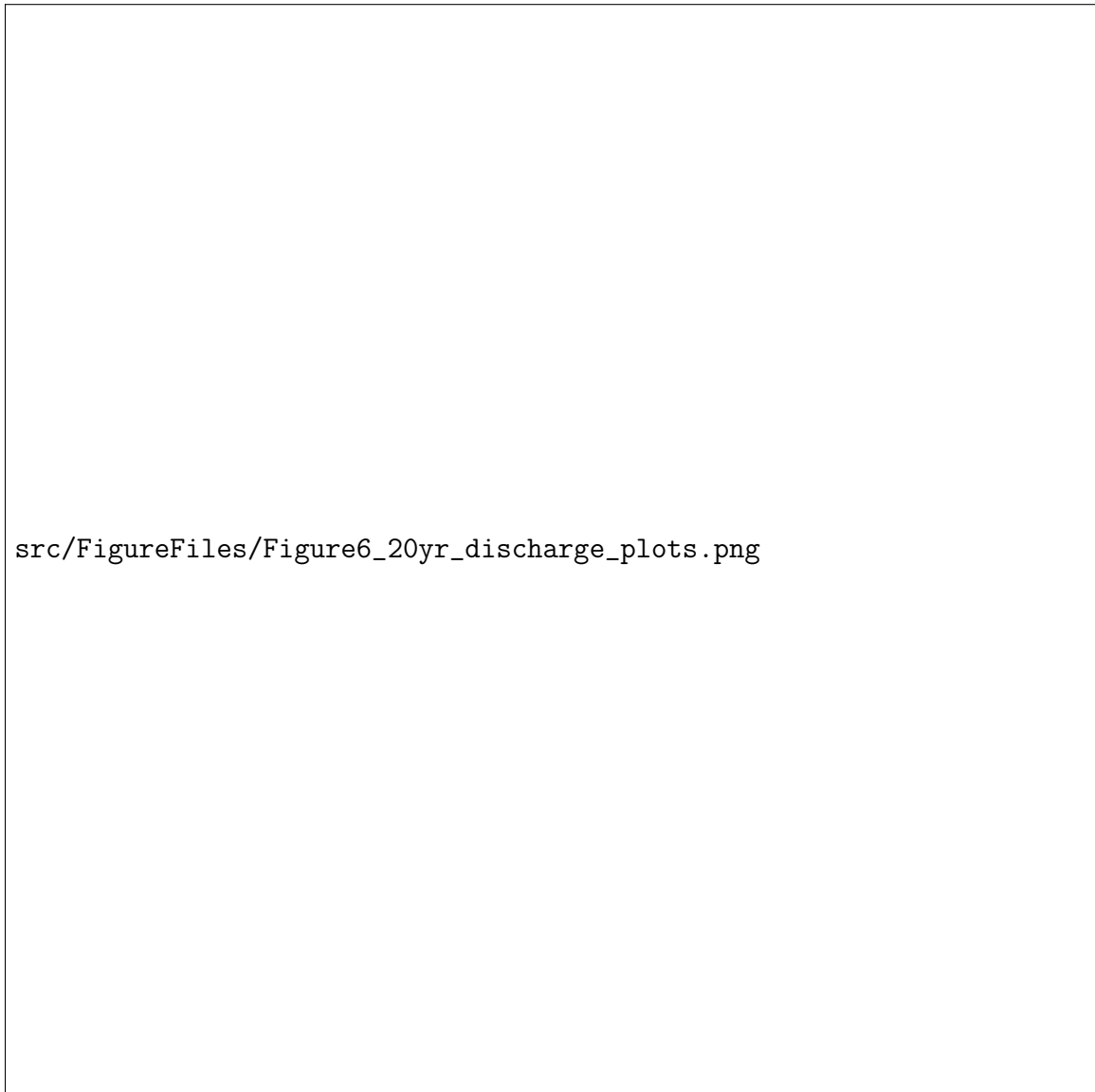


Figure 2.6: Simulated daily discharge (m³/s) from the uncoupled WRF-Hydro model, forced by hourly WRF meteorological forcings, for a 20-year (1995-2014) period compared against corresponding USGS stream gauge measurements. Right column shows the distribution of daily streamflow biases (model-observed). From the top down: the South Fork Boise (SFB), Moore's creek (MC), the main fork of the Boise River (MFB), and the South Fork Payette (SFP).

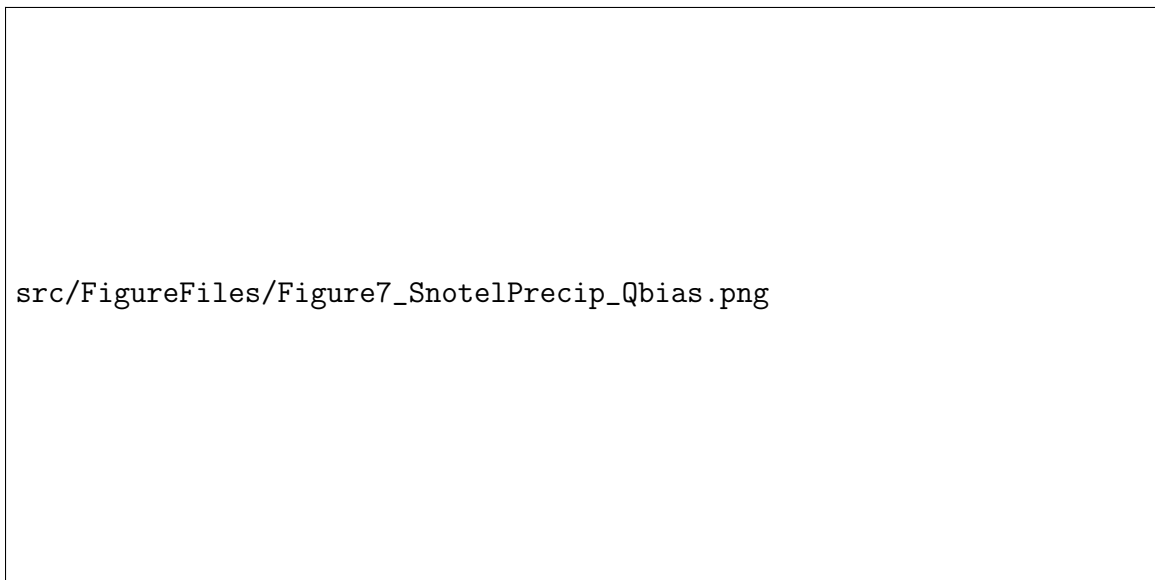


Figure 2.7: Linear relationship between WRF cold-season (October-May) precipitation percent bias, evaluated against NRCS Snotel sites (columns) and WRF-Hydro stream discharge percent biases evaluated at USGS gauge locations for each watershed (rows), 1995-2014. Markers with significant correlations ($p < .05$) are filled. The colorbar indicates coefficient of determination (R^2). Far left row shows the distribution of annual stream-flow percent biases. Top row shows the distribution of annual precipitation percent biases



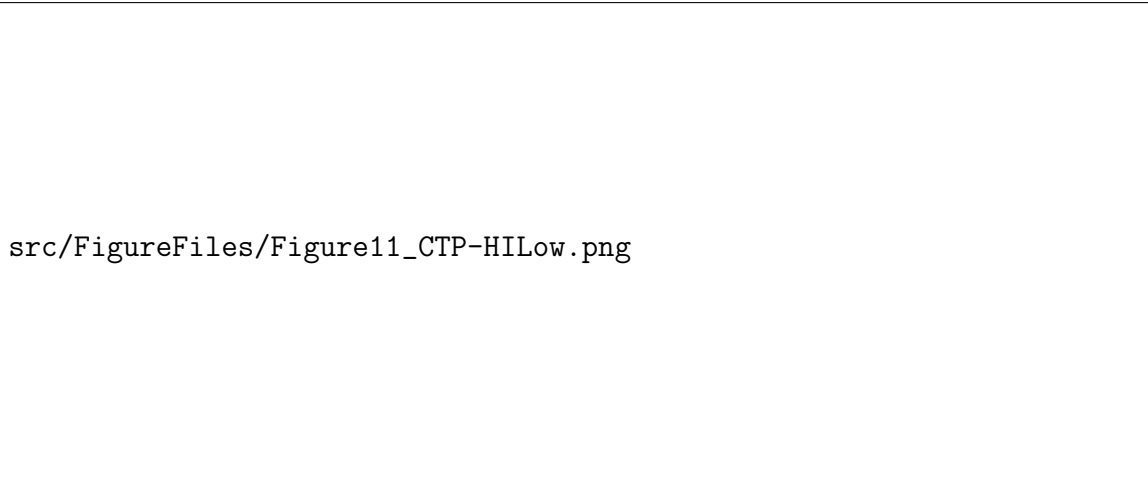
Figure 2.8: Time Series view of 500 meter geopotential height (top), precipitation rate (mm/hr; mm), Convective Available Potential Energy (J/kg), two meter air temperature ($^{\circ}\text{C}$), latent heat flux (LH; w/m^2) and sensible heat flux (HFX; w/m^2), and soil saturation of the 2nd soil layer. Plots are spatially averaged across the study region box (Figure 7). “Control” denotes the difference between the Control run and each respective scenario.



Figure 2.9: Temporal averages across the subdomain of soil saturation (2nd layer), two-meter air temperature ($^{\circ}\text{C}$), afternoon Convective Available Potential Energy (CAPE; J/kg), latent heat flux (LH; w/m^2) and sensible heat flux (HFX; w/m^2), Planetary Boundary Layer height (PBLH) and Accumulated Precipitation at the final timestep. Differences (Control-Routing and Control-HiResInit) are expressed as percent of Control for PBLH, Precipitation, and CAPE.



Figure 2.10: Figure 10: May 15 - June 1, Low elevation ($\leq 2000\text{m}$) temporally averaged relationships throughout the subdomain between CAPE (J/kg), accumulated precipitation (mm), evaporative fraction (EF; fraction of latent heat relative to the sum of latent and sensible heat flux), planetary boundary layer height (PBLH), and soil saturation in the second soil moisture layer.



src/FigureFiles/Figure11_CTP-HILow.png

Figure 2.11: The coupled model experiments convective-triggering-potential (CTP) and HILow indices, evaluated at 5:00 a.m local time. Graphs correspond with individual WRF grid cells located in Stanley and Boise (Figure 2). National weather service radiosonde observations from Boise are also plotted (black triangles). Green hatched boxes indicate thermodynamic conditions where precipitation is considered atmospherically controlled/insensitive to surface conditions based on (Findell and Eltahir, 2003).

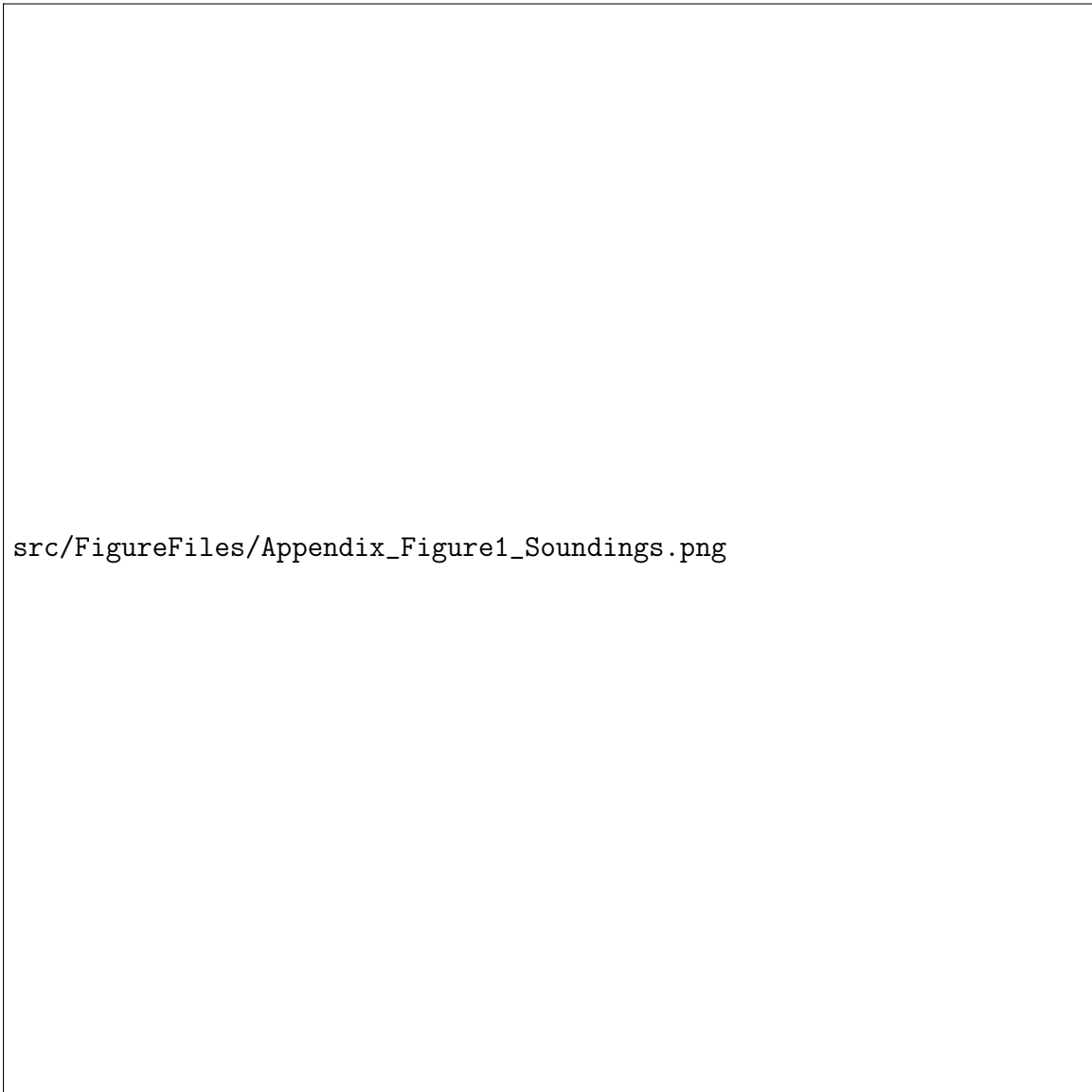


Figure 2.12: Observed and WRF-model Skew-T Log P diagrams for morning (12z) and afternoon (00z) during May and June of 2018.

CHAPTER 3:
EVALUATING 34 YEARS OF DYNAMICALLY
DOWNSCALED PRECIPITATION IN THE
COLORADO ROCKIES

Abstract

Convection Permitting atmospheric Models (CPM) have recently become tractable at climate-approaching timescales. These approaches have tremendous utility for water resource studies, but better characterizations of model biases and skill of simulated hydrologic fluxes at multidecadal timescales is required. This study examines 34 years (1987-2020) of CPM precipitation output from the Weather Research and Forecasting model (WRF; V.3.8.1) forced with Climate Forecast System Reanalysis (CFSR; CFSv2) lateral boundary conditions. Precipitation is compared against point observations (Snotel), gridded climate datasets, and Bayesian reconstructions of watershed-mean precipitation conditioned on streamflow and high-resolution snow remote sensing products. We find that the cool-season precipitation percent error between WRF and 23 NRCS Snotel gauges has a low overall bias ($\mu=.246\%$, $\sigma=13.63\%$), and that WRF has a higher percent error during the warm season ($\mu=10.37\%$, $\sigma=12.79\%$). Warm season bias manifests as a high number of low-precipitation days, though the low resolution or Snotel gauges precludes analyzing the lower limits of daily precipitation in the regions sampled by the Snotel network. Regional comparisons between WRF precipitation accumulation and three different gridded datasets (Newman, Livneh, and PRISM) show differences in accumulated precipitation on the order of +/- 20%, and particularly at the highest elevations with the fewest gauge locations. We find that WRF has a slightly higher correlation than the gridded precipitation product when compared to the Bayesian reconstruction, particularly when changing gauge densities are taken into account. The conclusions are that the CPM model

reasonably captures orographic precipitation in this region, and demonstrates that leveraging additional hydrologic information (streamflow, snow remote sensing data) improves the ability to characterize biases in modelled precipitation fields. Error characteristics reported in this study are essential for leveraging CPM models for studies of past and future climates.

3.1 Introduction

3.1.1 Quantifying and Modeling Mountain Precipitation: Challenges and Consequences

Over 1.6 billion people globally rely on water resources flowing from mountain regions (Immerzeel *et al.*, 2012), often in the form of seasonal snowpacks. At the same time, mountains are uniquely sensitive to climate change (Mountain Research Initiative Edw Working Group *et al.*, 2015), and snowpacks are forecast to decline significantly in the coming decades (Siirila-Woodburn *et al.*, 2021). Reducing uncertainties in climate change projections at watershed relevant scales requires improved modeling efforts or precipitation. Currently, most GCMs are too coarse to project climate change at water-resource relevant scales, where variations in local topography (finer than the GCM grid) enhance/inhibit precipitation in significant ways. Even small terrain features interacts with the atmosphere to modify flow fields and promote/inhibit precipitation through a variety of mechanisms, which may include stable, mechanically forced upslope ascent, "seeder feeder" mechanisms, and thermally induced convection from differential heating (Roe, 2005; Houze, 2012; Kirshbaum *et al.*, 2018). The last decade has demonstrated that non-hydrostatic, convection permitting regional climate models are capable tools for capturing these processes (Ikeda *et al.*, 2010; Rasmussen *et al.*, 2011; Gutmann *et al.*, 2012; Liu *et al.*, 2017) and the related task of modeling mountain snow accumulations (Currier *et al.*, 2017; Wrzesien *et al.*, 2019), of which precipitation in the first-order control. Grid resolutions less than 4km can permit convection, in addition to better capturing uplift caused by more realistic terrain Prein *et al.* (2015).

Evaluating biases in regional models is a persistent challenge, as gridded, gauge based datasets that are commonly considered a "gold-standard" can disagree substantially in mountain watersheds (Henn *et al.*, 2018) because of methodological choices, and should themselves be considered model products and not treated as observations (Lundquist *et al.*, 2019). Commonly used precipitation datasets include applied interpolation techniques used to map sparse gauge observations across terrain (Daly *et al.*, 2008; Thornton *et al.*, 2016). Remote sensing based precipitation products (such as Ashouri *et al.* (2015)) also exist, but less suitable for stratiform clouds, nor those composed of ice-phase hydrometeors, which are both common during mid-latitude winter precipitation in mountains (Lettenmaier *et al.*, 2015). Ground based radar systems can be used to measure precipitation rates in conjunction with station data (Lin2005-on, but radar beam blockage limits the utility in complex terrain (Maddox *et al.*, 2002). Consequently, more work is needed to be done to interrogate precipitation outputs from dynamical down-scaling models in mountain watersheds. Moreover, regional climate models are often run for fairly short time scales (less than a year, or a handful of years), so evaluations have less statistical power and thus cannot capture a range of synoptic conditions and teleconnection patterns such as ENSO. Lundquist *et al.* (2019) urges the community to consider merging indirect hydrologic information, such as but not limited to, ecology, soil-moisture, snowpacks, and streamflow in order to better evaluate precipitation datasets in mountainous watersheds. The uncertainties in mountain precipitation (rates, phases, magnitudes) propagate into studies of hydrologic systems for both water resource applications, snow modeling applications (Raleigh *et al.*, 2015)), and aqueous biogeochemistry (Maina *et al.*, 2020). Better quantifying mountain precipitation in the current climate is also a necessary

step for closing global water-budgets, and a necessary first step for using dynamical downscaling models to project impacts in the future.

3.1.2 Experiment Outline

This study evaluates 34 years of Weather Research and Forecasting model version 3.8.1 (WRF v3.8.1) precipitation output throughout the Colorado Rocky Mountains, with emphasis on the East River watershed, a high elevation $\sim 750 \text{ km}^2$ watershed in the headwaters of the Upper Colorado river basin (Hubbard *et al.*, 2018). The range of synoptic conditions present can reveal important insights about model behavior not captured in other studies, which are often conducted for only a handful of years (Ikeda *et al.*, 2010; Rasmussen *et al.*, 2011; Gutmann *et al.*, 2012). In the first section, regional-scale precipitation is evaluated against 24 Snotel stations (Serreze *et al.*, 1999) in the vicinity of the East. Many aspects of precipitation simulation can be important, depending on the question (diurnal cycles, peak intensity, phase, for instance; Trenberth *et al.* (2003)), and not all are considered here. We examine the biases in annual, cold-season (October-April) and warm-season (April-September) precipitation in addition to temporal correlations of accumulation and daily precipitation rates against Snotel. Lastly, spatial patterns of average precipitation accumulations are compared against three gridded precipitation products, namely the Parameter Regression on Independent Slopes model (PRISM; Daly *et al.* (2008)), Livneh (Livneh *et al.*, 2013), and Newman (Newman *et al.*, 2015) model products. We compare each across the entirety of the WRF model grid ($\sim 100,000 \text{ km}^2$) and the differences with respect to elevation are considered.

After examining regional-scale precipitation, we focus on evaluating WRF precipitation at headwaters specific scales. We examine spatial patterns of precipitation

across the East River watershed and locations of orographic enhancement by season. The East River is the focus of several ongoing and future intensive field campaigns (Hubbard *et al.*, 2018), and the results of this study are intended to guide and contextualize research activities currently taking place in the East River. The East River watershed (Figure 3.1) is an exemplar of Rocky-Mountain alpine landscapes (Hubbard *et al.*, 2018) and flows from the Elk mountains, approximately in the center of the WRF domain described in the next section. Elevation ranges between 2500 and 4200 meters above sea level. To better evaluate differences between WRF and PRISM in the East River, we compare basin-mean precipitation from each dataset against a Bayesian precipitation methodology. The inference method estimates basin-mean precipitation using a combination of parsimonious snow/soil water accounting models, precipitation gauge observations, streamflow records, and a limited number of Airborne Snow Observatory snow-lidar surveys for water years 2018 and 2019 (ASO; Painter *et al.* (2016)). This work builds upon prior precipitation-from-streamflow work by incorporating lidar-derived snow water equivalent estimates (described in section) into the precipitation estimating framework, similar to Henn *et al.* (2016).

3.2 Datasets and Methods

3.2.1 WRF Model Domain and Configuration

We use the Weather Research and Forecasting, version 3.8.1 (WRF; Skamarock *et al.* (2008)) model with two nested domains. The inner domain has a 1km resolution and 50 vertical levels, and the outer domain has a 3km horizontal resolution. The inner grid dimensions are approximately 300 by 300 grid cells. We use CFSR and CFSRv2 pre/post 2011 (Saha *et al.*, 2010) lateral boundary conditions. CFSR has a

.5 degree horizontal grid resolution. This model configuration differs from the commonly accepted three-to-one horizontal resolution ratio between boundary condition data and the outermost grid resolution. The outermost domain encompasses the entirety of Colorado's Rocky mountains and extends East into the Kansas lowlands, and West as far as Utah's Uintas range. Due to the time and computational constraints, each water-year (October 1 - September 30) is run independently and preceded by a two week spinup period. Consequently, multi-year soil-moisture/atmosphere interactions might not be well represented, as the soil moisture fields (and other land surface states) are initialized at the beginning of each water year with the coarse CFSR soil moisture field. In this way, multiple water years can be run concurrently. The horizontal grid resolution of both domains is less than the 4km typically considered necessary to permit convection (Weisman *et al.*, 1997) so convective parameterizations are turned off. The Thompson *et al.* (2008a) microphysical scheme and other WRF parameters listed in Table 3.1. Precipitation sensitivities to microphysical parameterization can be significant (cite Liu) and will be addressed in a follow-up manuscript.

In general, two predominant synoptic regimes control water-inputs to the East River watershed, namely winter baroclinic waves (frontal systems) and summer-time convective precipitation events that can sometimes be associated with the North American monsoon. Winds and moisture are predominantly from the West during the winter. The Colorado front range is also affected by upslope storms typified by northerly and easterly winds (Rasmussen *et al.*, 1995). Nevertheless, river hydrographs are typified by single large peaks during the early summer decaying to baseflows during the late summer.

3.2.2 Comparison Precipitation Datasets

We compare WRF precipitation accumulations against NRCS Snotel precipitation observations. Snotel stations are designed to provide cost-effective climate information for water-resource important regions throughout the Western US and have been used extensively in the study of hydrology and climate (Serreze *et al.*, 1999). Ultimately twenty-three Snotel sites are compared, ranging between 8000-11000 feet above sea level (purple triangles in Figure 3.1). The CFSR reanalyses used to force WRF do not assimilate Snotel precipitation data (Saha *et al.*, 2010), so precipitation recorded at the Snotel station is a completely independent check of WRF precipitation.

We also compare WRF precipitation fields against the Parameter Regression on Independent Slopes model (PRISM; Daly *et al.* (2008)), Livneh (Livneh *et al.*, 2013), and Newman (Newman *et al.*, 2015) model products, respectively. There are a number of differences between each model product, and elucidating the precise nature of the differences is beyond the scope of this article. One key difference is that PRISM and Newman use data from NRCS snotel networks, whereas Livneh uses observations from the NWS COOP stations that have at least 20 years of data. However Livneh precipitation accumulations are scaled such that the monthly means match mean PRISM climatology from 1961-1990. All three products use the PRISM terrain-precipitation relationships to distribute orographic precipitation. PRISM uses a mapping methodology that regresses precipitation for each individual grid cell based on nearby station observations and terrain orientation with respect to climactic variables.

3.2.3 Precipitation Inference Methodology

Precipitation-from-streamflow, or "doing hydrology backwards" methods have been employed in a number of studies (Kirchner, 2009; Pan & Wood, 2013) includ-

ing those in snow-dominated alpine watersheds (Le Moine *et al.*, 2015; Valery *et al.*, 2009) and glaciated watersheds (Immerzeel *et al.*, 2012) using glacier mass-balance as opposed to streamflow. Henn *et al.* (2015) used a Bayesian inference method to evaluate gauge-based precipitation products, with further applications in Henn *et al.* (2016), Henn *et al.* (2018), and Hughes *et al.* (2020), the latter of which used such a methodology to evaluate atmospheric model performance for watersheds in the Sierra Nevada. The approach adopted in this study is intended to follow Henn *et al.* (2015) as closely as possible. In essence, the method combines a temperature index snow accumulation/ablation model run in elevation bands, a soil-water accounting and streamflow routing "bucket" model, and precipitation and temperature equations with that distribute point-scale observations to each elevation band. Figure

A Bayesian inverse routine finds the most likely ranges of parameters, including parameters in the precipitation/temperature distributing functions, that best matches observed streamflows. The precipitation in each elevation layer (at height z) is given by the following equation:

$$P_{(z)} = m * [(P_s + P_{bias}) * (1 + OPG * dz)] \quad (3.1)$$

Where P_s is the daily observed precipitation at a Snotel location (at height z_0), P_{bias} is a precipitation gauge undercatch factor, OPG is an orographic precipitation enhancement factor, and m is a multiplicative error term, and $dz = z_{eff} - z_0$ where z_0 is the station elevation. In order to account for the oft observed decreases in snow water content with elevation, an "effective" layer elevation (z_{eff}) is prescribed by:

$$z_{eff} = \begin{cases} z - \gamma(z - \zeta) & \text{if } z > \zeta \\ z & \text{otherwise} \end{cases}$$

In this way, precipitation begins to decrease after a certain tuneable elevation ζ at a rate of γ . We use the Snow17 (Anderson, 1976) snow model run in discrete elevation layers to model the accumulation/melting of snow. In order to provide estimates that are as independent as possible of WRF, we use NRCS Snotel data from the Butte and Schofield stations located in the East and vicinity of the East watersheds, respectively to force the model. Periods of missing or poor quality temperature data (a small percentage) are corrected using adjusted data from the Schofield station (elevation corrected by the dry-adiabatic temperature lapse rate), when available, or interpolated between neighboring values.

To infer precipitation, a three-part inference process is applied. In the first step, Snow-17 parameters are calibrated for the Butte Snotel (381) site, including a precipitation undercatch term using a standard (non-bayesian) minimization algorithm. Next, the *OPG*, ζ , and γ , and temperature lapse rate parameters are fitted to the mean Airborne Snow Observatory SWE for water years 2018 and 2019. ASO produces three-meter scale estimates of snow-water equivalent by taking repeat LiDAR observations of snow surfaces and modeling snow density using energy balance modeling. ASO data products for the East River are publicly available from the National Snow and Ice Data center.

In the second step, SNOW-17 (and the tuned precipitation distributing paramters) is coupled with a bucket hydrologic model based on FUSE hydrologic model framework (Clark *et al.*, 2008). SNOW-17 provides rain and snowmelt inputs to the hy-

drology model. The hydrologic model also requires a potential evapotranspiration forcing, which is computed using the Hamon formula (Hamon W. R., 1961). The model structure used in this study is the most similar to the VIC/PRMS forms described in (Clark *et al.*, 2008). The structure was chosen for simplicity and to have as few free parameters as possible. There are two state variables, soil moisture in the top and bottom buckets (SM_1 , SM_2) with maximum capacities SM_1^{max} and SM_2^{max} . The model fluxes equations are described in Table (??) and illustrated in Figure 3.7. FUSE produces a hydrograph that can be compared against observed streamgauge data.

Posterior model parameters (θ), conditioned on the model structure and observed streamflow data (d), are can be expressed using Bayes' rule:

$$P(\theta|d) \propto P(d|\theta)P(\theta) \quad (3.2)$$

Analytical expressions for the posterior are not possible, so Markov chain monte-carlo sampling methods are used, specifically the DEM-Metropolis algorithm implemented in the python ‘‘PyMC3’’ library (Salvatier *et al.*, 2016). The model-likelihood function $P(d|\theta)$ is a standard least-squares error model which assumes that residuals follow a gaussian distribution with a mean of zero. The standard deviation of the error term is a linear function of the daily magnitude of discharge, following Henn *et al.* (2015) and Thyer *et al.* (2009), given by $\sigma_t = \alpha Q_t + \beta$, where Q_t is the observed discharge at timestep t . The coefficients of the error model are inferred along with model parameters.

3.3 Results

3.3.1 Regional Scale Comparison

Seasonal Precipitation Accumulations Compared Against Snotel

WRF precipitaiton from the 34 year period is compared against corresponding Snotel grid-cells shown in Figure 3.1. Comparing WRF precipitation against Snotel observations demonstrates that WRF captures important features of water-delivery throughout the region. The analysis is divided into two rough categories, the "cold-season" (October-March) and "warm-season" (April-September) which are intended to roughly demarcate winter stratiform and summer-time convective precipitation regimes. Analyzing the monthly averages of integrated vapor transport and 500 hPa wind directions shows that wind and moisture overwhelmingly come from the West during the cold-season (October - April) and from the West-South-West during the remainder of the year (not shown). The percent errors in water-year total precipitation, expressed as

$$\% \text{ Error} = \frac{\sum P_{WRF} - \sum P_{Snotel}}{\sum P_{Snotel}} * 100 \quad (3.3)$$

are also examined. There is no immediately apparent trend in location of Snotel site with respect to elevation or topography and error characteristics. The worst performing site is Brumley (369), located on the lee-side of a mountain ridge, where WRF overpredicts precipitation consistently throughout the study period. Interestingly, sites located only a few kilometers away on the windward side of the range are well predicted. While the correlations are similar between season, the errors in

precipitation accumulation are not evenly distributed across the water-year. During the warm-season, WRF is wetter than Snotel sites for sixteen of the twenty-three Snotel sites, with an average accumulated precipitation percent-bias of 10.4%. The cold-season percent error averaged across all years and Snotel locations is, remarkably, .264% but with a 10.1% standard deviation. Comparing one-week rolling mean timeseries of WRF averaged across all Snotel locations, compared with the average Snotel precipitation, shows good correlation (R^2 of .85 and .88 for the warm and cold-season, respectively). The relationship between binning-window (daily-monthly) and correlation was also examined, and found that correlations were low at the daily increments, but tended to flatten beyond averaging window-lengths greater than three to four days.

Daily Precipitation Frequency and Rates

Examining daily correlations between model precipitation and precipitation at Snotel sites introduces problems because of the Snotel sensor resolution (2.54 mm per day). This is illustrated in Figure 3.3, which compares the average number of days with precipitation ("wet" days) and dry-days between WRF and Snotel datasets. When the "wet" threshold is set to 2.54 mm, WRF and Snotel have similar wet/dry statistics (Figure 3.3), with both datasets agreeing on wet-days more often than not. However, when the "wet" threshold is relaxed to daily precipitation > 0 mm, WRF has significantly more wet-days than Snotel particularly during the warm season (more than double). The relationship between daily precipitation rates and seasonal accumulation can be expressed by the Precipitation Intensity Contribution (PIC) to the total, given by:

$$PIC(j) = \frac{\sum_{i=1}^n (P_i \text{ if } P_i > j ; \text{ else } 0)}{\sum_{i=1}^n P_i} \quad (3.4)$$

Where P is precipitation, i is the time index (daily), and j is a particular rate in millimeters per day. This expresses the percent of the total precipitation caused by daily rates less than or equal to that of each bin. Averaged across all Snotel locations and water-years, more than 10% of all precipitation from WRF is attributable to daily rates less than 2.54 mm, or "drizzle" (rain or snow) that would not be reported in the Snotel dataset. Moreover, close to one quarter of the Snotel precipitation results from days where the reported value is at the minimum resolution. This is likely in part because of Snotel buckets collecting, but not reporting precipitation until the .1 in threshold is reached. Consequently it is likely that Snotel gauges under report the number of "wet" days and the fidelity of light precipitation days is difficult to quantify in the mountain regions they sample.

Regional Comparison of WRF and Gridded Dataset

Three gridded precipitation datasets are compared against WRF for across the inner WRF domain 3.1. The PRISM (4km) product, Livneh ($1/16^\circ$), and Newman ($1/8^\circ$) are regridded to a common grid using the xESMF python package. The last decade is ignored since Livneh and Newman are only available until 2012 and 2013 respectively. The comparison of water-year averaged precipitation (1987-2010) shows significant differences between WRF and the gridded datasets, on the order of plus or minus 25% and up to 400mm in some regions (Figure 3.4) The differences among the gridded datasets are lesser, in particularly between PRISM and Livneh. Some of the differences are due to the underlying data resolution. Comparing precipitation as

a function of height shows that WRF typically has less precipitation at the highest elevations (greater than 3500 m) compared to PRISM, and that the two datasets disagree the most in regions that are poorly sampled by Snotel locations (the Snotel maximum elevation is 3500 meters). Moreover, PRISM has a more skewed distribution at higher elevations in addition to higher maxima. Both datasets show a clear rain-shadowing effect between 2250-2400 meters, corresponding with the region to the East of the San Juan mountains in the South East corner of the domain, though PRISM is drier than WRF (Figure 3.4).

The following section analyzes precipitation throughout the East River watershed region (Figure 3.1).

3.3.2 East River Precipitation Analysis

The spatial patterns of precipitation are compared between WRF and PRISM for the warm and cold seasons averaged between water-years 1987-2020. To examine spatial differences in precipitation, we define the enhancement factor for each grid cell, given by:

$$EF_{i,j} = \frac{P_{i,j}}{\frac{1}{m} \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^n P_{i,j}} \quad (3.5)$$

Which is simply the ratio of the accumulated precipitation in each grid cell i, j to the m by n points averaged across the watershed.

Livneh and Newman are ultimately quite similar in the East, so consequently only PRISM and WRF are compared subsequently. Both datasets show that more precipitation accumulates during the Cold season on average. In many years, the mountain slopes on the windward side (West; left of the figure) receives more precipitation in WRF relative to PRISM (Figure 3.5), despite having an overall lower precipitation.

WRF has a much higher enhancement factor on the windward East side compared to PRISM, which has a very high enhancement factor and very high positive bias relative to WRF in the Northwest. There does appear to be a significant downward shift in PRISM basin-mean precipitation in the final decade of the simulation, the causes of which are discussed in the next section. PRISM generally has a higher precipitation-elevation gradient compared to WRF, for both the East river watershed and the entire WRF domain (Figure 3.11). The averaged elevation/precipitation relationship is most similar at low-to mid elevations and deviates most strongly at the highest elevations.

Comparisons against Streamflow Inferred Basin-Mean Precipitation

In order to better understand some of the discrepancies between WRF and gridded datasets, the streamflow from inference method adapted from (Henn *et al.*, 2015) is adapted to examine basin-mean precipitation in the East River watershed. This method cannot isolate spatial precipitation patterns (which are significant; Figure 3.6), only basin-mean. This is nonetheless useful as the differences in the mean are large (150mm). A three-part inference process is applied. First, Snow-17 parameters are calibrated for the Butte Snotel (381) site, including a precipitation undercatch term using a standard (non-bayesian) minimization algorithm (P_{bias} , Equation 3.1). Next, the orographic precipitation gradient ("opg"), temperature lapse rates, and precipitation gradient cutoff terms are calibrated against two-years of Airborne Snow Observatory SWE products. Water year 2018 and 2019 are low and high precipitation years, respectively, with approximately peak SWE values of greater than 2000 mm in 2019 and approximately 1000 mm in 2018 (Figure 3.8). This is fortuitous, as that

these years represent both extremes and are thus good for approximating the long-term behaviour. Aggregating SWE with respect to elevation bins shows a remarkably consistent increase in SWE with elevation, after which SWE values tend to decline. A similar pattern is found in the Tuolumne basin in California (Henn *et al.* (2016), Figure 4). Tuning the *opg*-gradient slope-break (ζ) and decrease-rate (equation γ) allows for fitting the observed ASO SWE curves. It is worth noting that the precipitation reduction term ultimately represents a small fraction of the total watershed area (less than 10%). The calibrated *opg* parameter is ultimately close to the initial guess of $.002 \text{ km}^{-1}$. The temporal evolution of SWE in corresponding bin closely matches that of the Snotel site (Figure 3.8.c). Following calibrating the snow model, the time-invariant hydrologic model parameters are inferred for the entire time series of discharge (1990-2020). The parameters are the hydrologic parameters listed in Table 3.3. Extensive testing of the Bayesian inference approach found sampler convergence criteria were only met after treating the upper-zone maximum soil moisture as fixed. A value of 400 mm was chosen, reflecting approximately 1 meter on average of sandy-loam soil. Inference was performed with different fixed values of soil storage, and found there was ultimately little difference in the inferred values of other parameters. The baseline model skill is high at this stage, with average root mean squared error of .65 mm prior to inferring precipitation forcing errors, suggesting that the model structure is a good approximation of the watershed dynamics.

After inferring the time-invariant hydrologic model parameters, forcing model parameters are inferred on an annual basis against streamflow. This includes the precipitation error multipliers (m), OPG, OPG gradient slope-break (ζ) elevation, and temperature bias. The PDFs of basin-mean inferred precipitation for each year

are compared against PRISM and WRF (Figure 3.9.A) and model streamflow skill displayed (3.9.B). Newman and Livneh were also examined but the mean was almost identical to PRISM. Averaged across all years, the mean-inferred precipitation multiplier has a mean close to 1 and OPG parameters are again close to the initial guess of .002 (Figure 3.9.C) Model skill-scores are high, with kling-gupta-efficiencies typically .8-.9 or higher and low root-mean-square errors, indicating the model well-captures both the temporal behavior of discharge and mass-balance (Figure 3.9.D) WRF typically has a lower mean precipitation compared to inferred, and PRISM is higher. Notable outlier years with significant spread between estimates includes water year 2002 and water year 1995, a low and high discharge year respectively. The three estimates become more alike approximately after water year 2010. This change in behavior is almost certainly due to the addition of the Upper Taylor Snotel (site 1141) in water-year 2010, immediately to the East of the basin. Of the three Snotel sites closest to the East, Butte receives the least precipitation, Schofield the most (almost twice that of Butte), and the Taylor typically a value in between the two. The PRISM basin-mean precipitation is a fairly constant +200mm (standard deviation 60mm) relative to the precipitation received at the Butte Snotel site, and the difference between the two values is significantly reduced after water year 2010 (Figure 3.10)

3.4 Discussion

3.4.1 Comparisons against other Convection Permitting Simulations

Several other studies have evaluated WRF against Snotel and gridded precipitation datasets for the Colorado Rockies, though few have presented analyses for

simulations spanning three decades. Rasmussen *et al.* (2011) reports similar performance metrics between Snotel stations and WRF. Jing *et al.* (2017) found that winter-time precipitation accumulations compared against a number of Snotel sites was less than 15% using a 2km WRF configuration with NARR boundary conditions. The absolute biases and percent-differences between WRF and PRISM from this study are similar both the Jing *et al.* (2017) 2km WRF and 4km WRF results from Liu *et al.* (2017), both presented in Jing *et al.* (2017) Figure 7. Similar patterns are also observed in Gutmann *et al.* (2012) who examined winter-only precipitation using a 2km WRF configuration with NARR boundary conditions. The similarity of results is significant, since there are differences in resolutions, nesting configurations, model code versions, boundary conditions, microphysical schemes and other options. The robustness of patterns across these convection-permitting model configurations suggests that the orographic precipitation solutions are robustly tied to model structural processes that are fundamentally different from the assumptions made in PRISM-like gridded precipitation datasets.

Differences in error behavior between cold and warm season can likely be attributed to precipitation generating regimes. WRF has a lower weekly correlation with Snotel stations (lower R^2 value), higher percent-errors, and an excessive number of wet-days relative to Snotel during the warm season. Jing *et al.* (2017) likewise found that WRF skill decreased in April, concurrent with an increase in convective available potential energy. Surface heating tends to increase during the warmer months leading to convective instabilities and localized precipitation, compared to more uniform stratiform precipitation during the winter (Dai, 2006). Warm season errors may be more sensitive to the exact locations of precipitating cells. At the same

time, the nature of errors at the gauge locations can depend on the phase of falling hydrometeors (rain versus snow), as snowflakes have lesser fall speeds and are more subject to undercatch during stronger winds (Goodison *et al.*, 1998; Harpold *et al.*, 2017). Consequently it is even more interesting that cold-season discrepancies are lower, even though the gauge errors are expected to be higher. Some studies have attempted to account for gauge undercatch at Snotel sites using the co-located snow pillow SWE measurements, but these methods are not employed here for the precipitation evaluation (Livneh *et al.*, 2013; Sun *et al.*, 2019). However, an undercatch term (P_{bias}) for the Butte snotel (381) was used in the precipitation inference method, but the value was small ($< 1\text{mm}$), and only two-years out of the thirty examined appeared to show snow-water equivalent values greater than concurrent accumulated precipitation. The overestimation of the number of wet days is well known in both regional and global climate modeling (Maraun *et al.*, 2010; Chen *et al.*, 2021). However it is difficult to determine the extent of the potential drizzle bias given the relatively low resolution (daily; 2.54 mm) of Snotel gauges (Figure 3.3). Additional experiments may consider indirect data sources (soil moisture, remote sensing of cloud properties) to better understand modeling drizzle days in regions covered by Snotel, which account for 10% of annual WRF precipitation. Users of dynamical downscaled precipitation may consider bias correcting by removing low-precipitation days (Maraun *et al.*, 2010). An analysis of streamflow responses, following Lundquist *et al.* (2009), could potentially isolate short periods of summer precipitation missed by the gauge networks but captured by WRF (or lack thereof).

3.4.2 Addressing Climate and Data Non Stationarities

The step-change in PRISM basin-mean precipitation underscores the fact that caution must be taken when analyzing precipitation trends from gridded datasets, as inclusion of different stations can induce spurious trends. Analyzing the annual East River mean precipitation from PRISM from 1987-2020 shows a negative trend with $p=.06$, $R^2=.33$, which likely reflects the addition of the Taylor Snotel site. The fact that PRISM more closely matches the WRF and inferred precipitation data (two independent estimates) in the last decade, after the addition of a nearby gauge, lends more confidence to the WRF modelled precipitation in the early parts of the simulation. That being said, WRF still typically underestimate precipitation at the Schofield snotel site (737; Figure 3.2), so the low-bias may carry over to adjacent regions including within the East. Lastly, the WRF boundary conditions changed from CFSR to CFSRv2 after 2011. After all, precipitation from regional climate models are ultimately dependent on the boundary conditions (Goergen & Kollet, 2021), so changes in data assimilation and methodologies may influence the character of the boundary conditions and the results. For these reasons, the impacts of anthropogenic climate change are not directly quantified in this work and require a more careful treatment to isolate signals from other effects. The Rockies are experiencing warming trends across seasons (Rangwala & Miller, 2012) and declining Spring snowpacks primarily driven by temperature (Mote *et al.*, 2005). Future studies may interrogate the climate change signals in the dataset.

3.4.3 Interpreting Streamflow Inferred Precipitation Estimates

Sensitivity to Evapotranspiration

High resolution snow Lidar SWE information offers significant utility, in conjunction with streamflow, for reconstructing basin mean precipitation inputs. Henn *et al.* (2016) likewise used ASO data in a joint-inference method in the Tuolumne river watershed. They found that that doing so reduced the dependence of inferred precipitation on hydrologic model structure, compared with inferring precipitation from streamflow alone. That being said, uncertainties in soil parameters, PET forcing, and water limitation relationships in the PET/ET relationships do limit the conclusions drawn from the precipitation inference approach. ET is directly related to inferred precipitation by the hydrologic mass-budget equation, $Q = P - ET - \frac{dS}{dt}$, so holding Q and $\frac{dS}{dt}$ constant implies that higher annual inferred precipitation requires higher annual ET. In the bucket model formulation used here, ET is a function of PET forcing and the soil moisture content, given by

$$ET = PET * \frac{\min(SM_{T1}, SM_{T1}^{max})}{SM_{T1}^{max}} \quad (3.6)$$

where SM_{T1}^{max} is the field capacity given by $fracten * SM_1^{max}$ (soil moisture in the top bucket multiplied by the time-invariant fraction held in tension). During the long-term hydrologic parameter inference, the *fracten* parameter consistently converged towards the limit of 1 (a lower-ET solution). It is important to note that the model log-likelihood is evaluated using the timeseries of streamflow, not the seasonal sum. Consequently the timing of ET is implicitly considered, as it impacts the daily partitioning of soil water into storage/streamflow. Additional sensitivity

experiments show that higher ET solutions (lower fracten) tended to smooth spring and early summer streamflow peaks excessively compared to observations, suggesting some confidence in the ET solution. Ryken *et al.* (2021) deployed eddy-covariance towers in riparian zones in the East river, and reported annual ET values between 450-550 mm per year. This is interpreted as an upper-limit of basin wide ET, as the flux tower is located in a well-watered riparian corridor. The posterior ET solutions from the inference method are the same magnitude, and between 350-450mm for the same time period. The relationship between critical zone properties (soil depths, water potentials), evapotranspiration measurements/estimates, and inferred orographic precipitation underscores the interdisciplinary nature required to better constrain precipitation inputs in mountain watersheds

Comparisons with Other Precipitation Inference Studies

This application of a precipitation-from-streamflow inference approach also differs from the (Henn *et al.*, 2015) study in one key aspect. PRISM precipitation was found to be lower than the total volume of streamflow (violating basic mass balance considerations), whereas in this study the estimated streamflow is less than the PRISM estimate. Another major assumption is that parameters in the precipitation distributing functions are constant for each season. Future efforts may infer precipitation error parameters for individual storm events such as in Le Moine *et al.* (2015) and Koskela *et al.* (2012). Future efforts could apply precipitation-inference methods in addition with dynamical downscaling and precipitation gauge data to produce regional scale precipitation reanalyses following the methods presented here, especially for basins where lidar-based snow products are available.

Moreover, while this study demonstrated that airborne snow lidar products can be incorporated into an inference strategy to evaluate orographic precipitation, the potential applications are vast. This study only used one-dimensional SWE versus elevation information as part of the Bayesian inference framework (Figure 3.8) to calibrate the OPG parameter in Equation 3.2.3. In cold, high elevation areas above the rain-snow transition zone, the snow water equivalent largely depends on the antecedent precipitation, with other processes secondary redistribution/loss processes superimposed on it. Cursory analysis shows that the patterns of SWE from the ASO data (Figure 3.8.A) more closely match the relative precipitation enhancement from WRF, where the largest values of precipitation/snow are on the windward side of the watershed. Future work will examine this relationship in greater detail.

3.5 Summary and Conclusions

This paper examined 34 years of dynamically downscaled precipitation produced by the Weather Research and Forecasting (WRF) model. Precipitation is compared against Snotel observations, three different precipitation products (PRISM, Livneh, and Newman), and basin-mean precipitation inferred from a Bayesian inference method that uses streamflow and high resolution snow-water equivalent data (ASO). The primary goal is to better characterize precipitation biases and error characteristics. We find that:

- Averaged across 24 Snotel stations, WRF has a .246% percent bias ($\sigma=13.63\%$) during the cold season and a 10.37% ($\sigma=12.79\%$) percent bias during the warm season,
- 10% of WRF annual accumulated precipitation is during "drizzle" days (<100

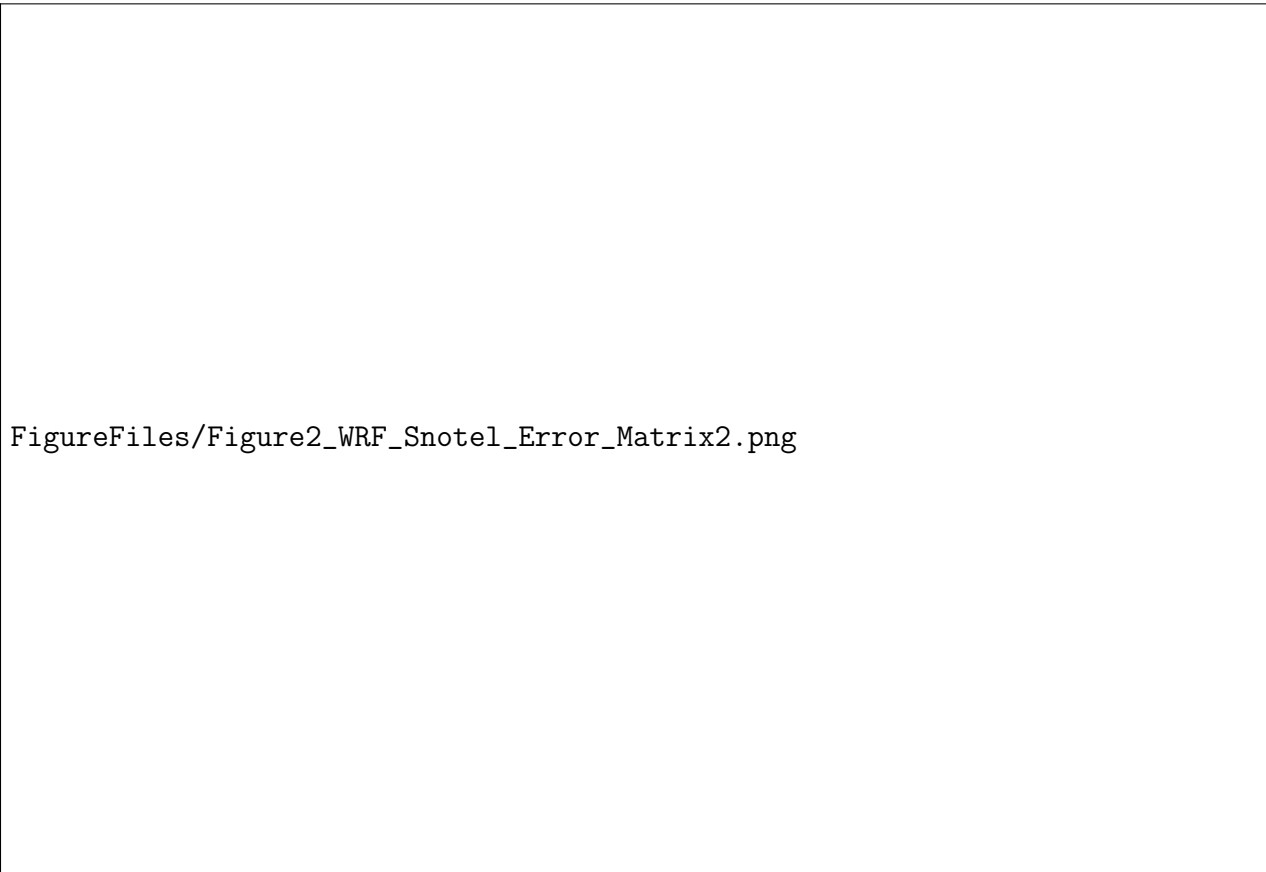
mm), and which is difficult to validate because of Snotel gauge resolutions. Users may consider bias correcting for drizzle during summer months

- PRISM/Livneh/Newman show generally similar patterns and disagree with WRF on the order of +/- 20% per year. The largest disagreements are at the highest elevations with the fewest observations.
- PRISM basin-mean precipitation becomes closer to the WRF and streamflow-inferred precipitation estimate following the addition of the Upper Taylor Snotel site in the adjacent basin.
- WRF patterns of precipitation more closely match the snow lidar datasets than PRISM for the two available years, with more enhancement on the windward (western) ridge of the watershed boundary.

Future studies may consider other moments of the precipitation distribution from this dataset, such as extreme events or drought-duration, in addition to other hydro-climactic variables (temperature, wind fields, radiation).




Figure 3.1: WRF Model domains (left) and East River watershed (right). Outer nest is 3km dx/dy and the inner domain is 1km dx/dy. Triangles show the locations of Snotel sites examined in this study



FigureFiles/Figure2_WRF_Snotel_Error_Matrix2.png

Figure 3.2: Error characteristics of 24 Snotel sites compared against corresponding WRF grid-cells, 1987-2020. A) One-week rolling mean timeseries of average Snotel (orange) precipitation compared against WRF (blue) B) Annual precipitation percent error (Snotel as reference) for each site, C) Average timing of water delivery (%) as a function of day of water-year for WRF (blue) and Snotel (orange), D) Correlations between one-week accumulated precipitation WRF/Snotel for Warm season (bottom) and Cool season (Top), D) Average precipitation percent errors by season



FigureFiles/Figure3_WRF_Snotel_WetDryRate.png

Figure 3.3: A) Percentage of days with precipitation ("wet") for Snotel stations and corresponding WRF grid-cell, for thresholds of daily precipitation greater than 2.54 mm (.1 inches) and 0 mm respectively, for Cold (Oct. - April) and Warm (April-Oct.) seasons. B) Contribution to cumulative precipitation by daily precipitation rate (mm/day) for WRF (solid) and Snotel (dashed) stations. Blue/Orange lines are the mean of all stations for WRF/Snotel respectively.

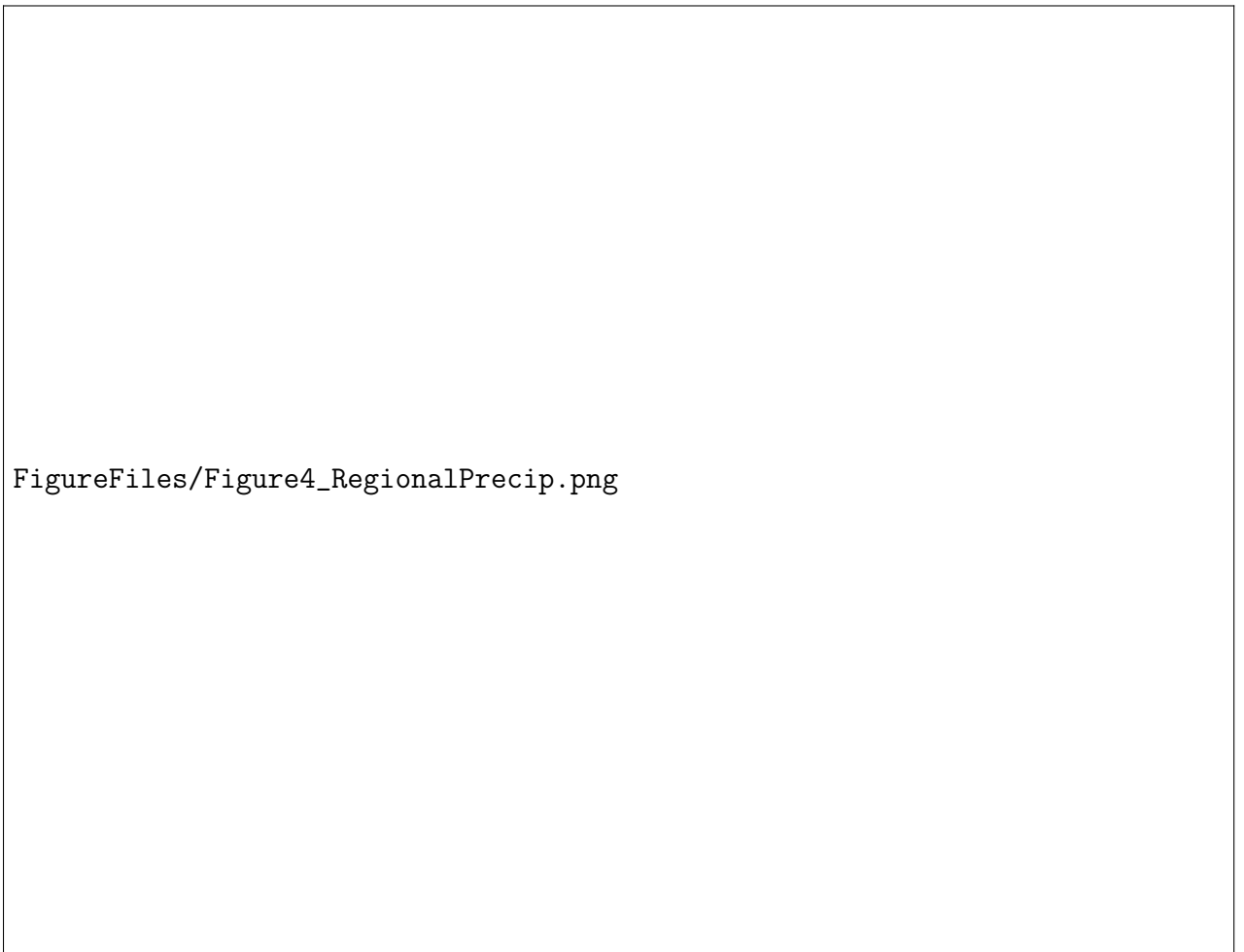


Figure 3.4: Digonal: Spatial averages of 1987-2010 precipitation for WRF, PRISM, Newman, and Livneh datasets for the entirety of the inner WRF model grid (300x300km). Lower off-diagonals are the difference between respective datasets (row-column). Upper diagonals are the same, but show the percent-difference ($[\text{column}-\text{row}]/\text{row}$).

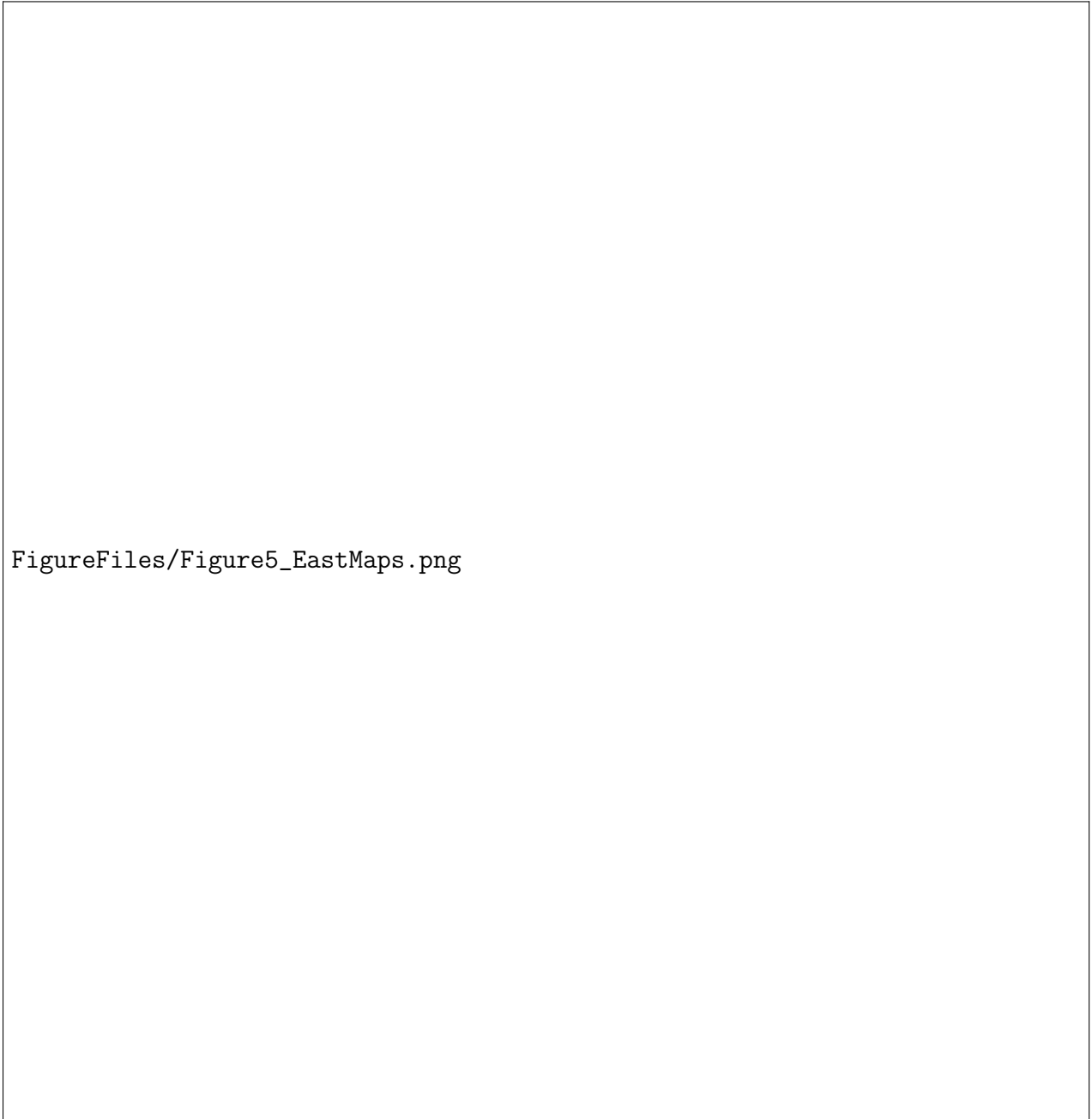


Figure 3.5: East River Annual Precipitation by Water Year, 1987-2020 for WRF and PRISM, Livenh, and Newman precipitation products



Figure 3.6: Average Cold-Season (October-April), Warm-Season (April-October), and average annual precipitation enhancement factors (Equation ??) for WRF and PRISM

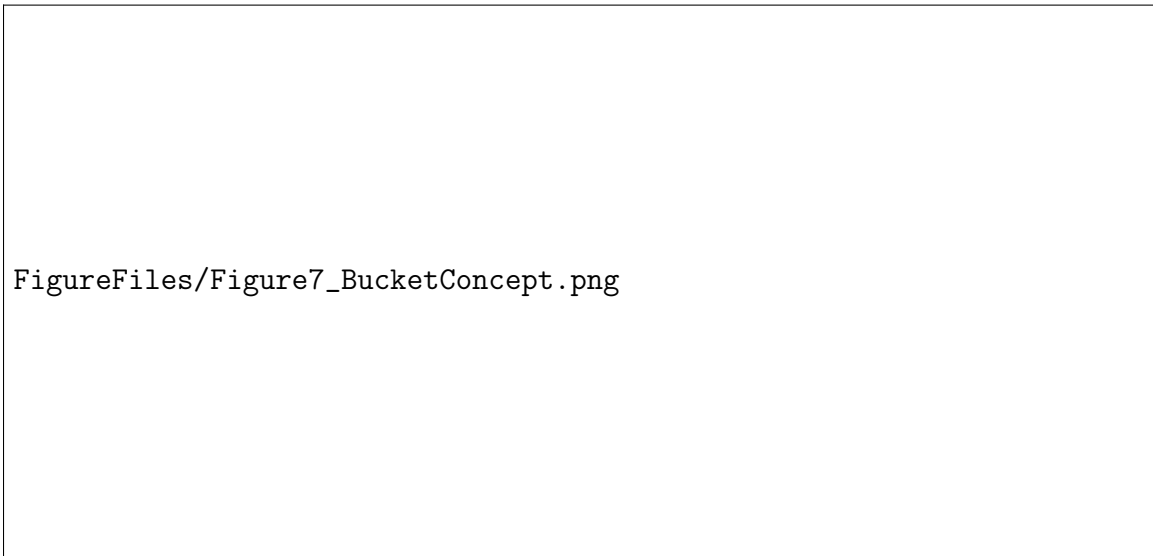
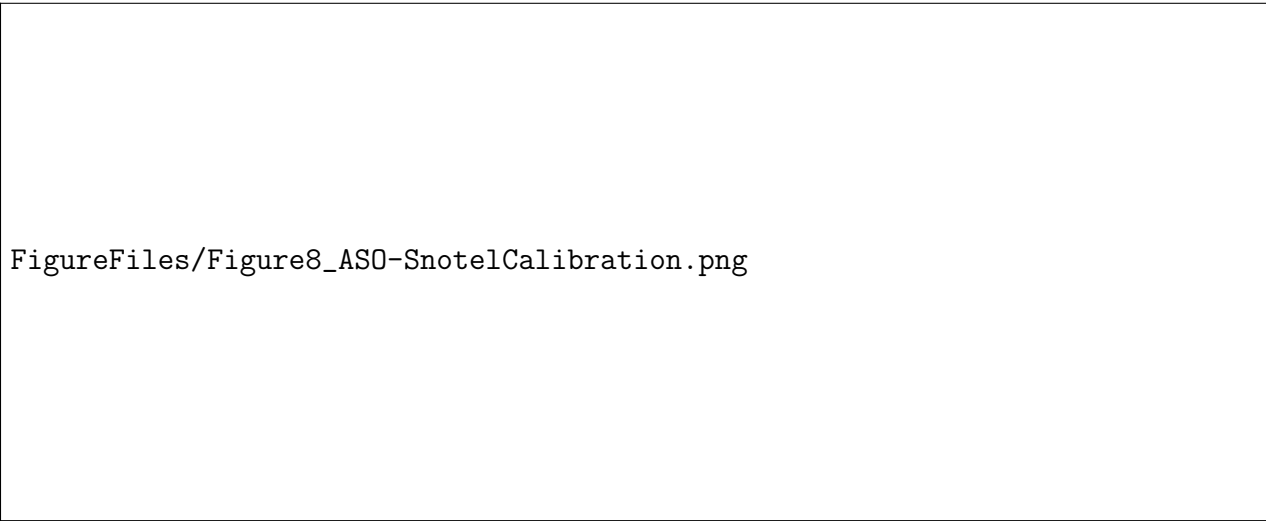


Figure 3.7: Precipitation inference model. A) Meteorological forcing structure. Precipitation/temperature increase/decrease with elevation and typical snow/melt fluxes are illustrated. B) Hydrologic model structure. Rain and Snowmelt (Q_{in}) and Potential Evapotranspiration (PET) are model forcings. State variables are the soil moisture in the top and bottom bucket respectively (SM_1 , SM_2). Model fluxes are baseflow (Q_b), percolation, Q_{12} , overland flow (Q_d), and bucket overflow (Q_{bo}). C) Q_d and Q_b are convolved with a routing function to produce streamflow (D).



FigureFiles/Figure8_ASO-SnotelCalibration.png

Figure 3.8: Left: The Airborne Snow Observatory SWE maps over the East River employed in this study. Center: Basin-average elevation versus average SWE for ASO and calibrated Snow-17 at ASO-flight dates. Right: Timeseries of calibrated Snow-17 by elevation band (n=100).

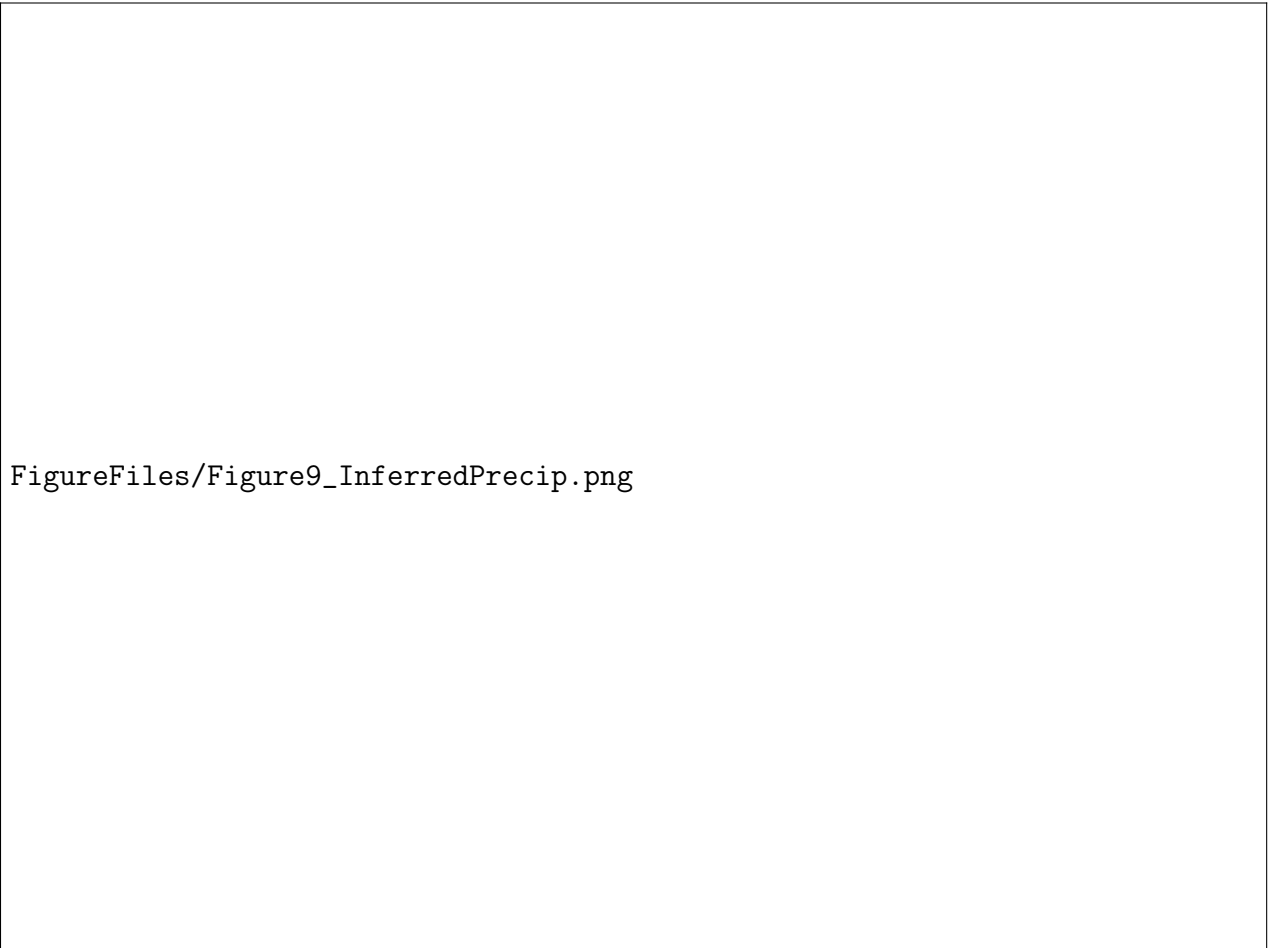


Figure 3.9: Evaluating precipitation-from-streamflow Bayesian inference method. A) WRF and PRISM Basin mean precipitation compared against posterior precipitation (violinplots). Residuals between PRISM/WRF and inferred are plotted (right y-axis). Cumulative observed streamflow plotted for reference (grey boxes). B: Observed streamflow and posterior streamflow solution with 1σ confidence interval (grey shading). C: Distributions of posterior forcing parameter values, inferred from each year D: Distributions of annual error metrics. Kling Gupta Efficiency and Root Mean Square Error.



Figure 3.10: East River annual mean PRISM mean precipitation compared against the three closest Snotel sites, Schofield (737; North of East), Upper Taylor (1141; East of East), and Butte (381; within East). Pre/Post 2010 mean difference between Butte and PRISM-mean are plotted with standard errors.

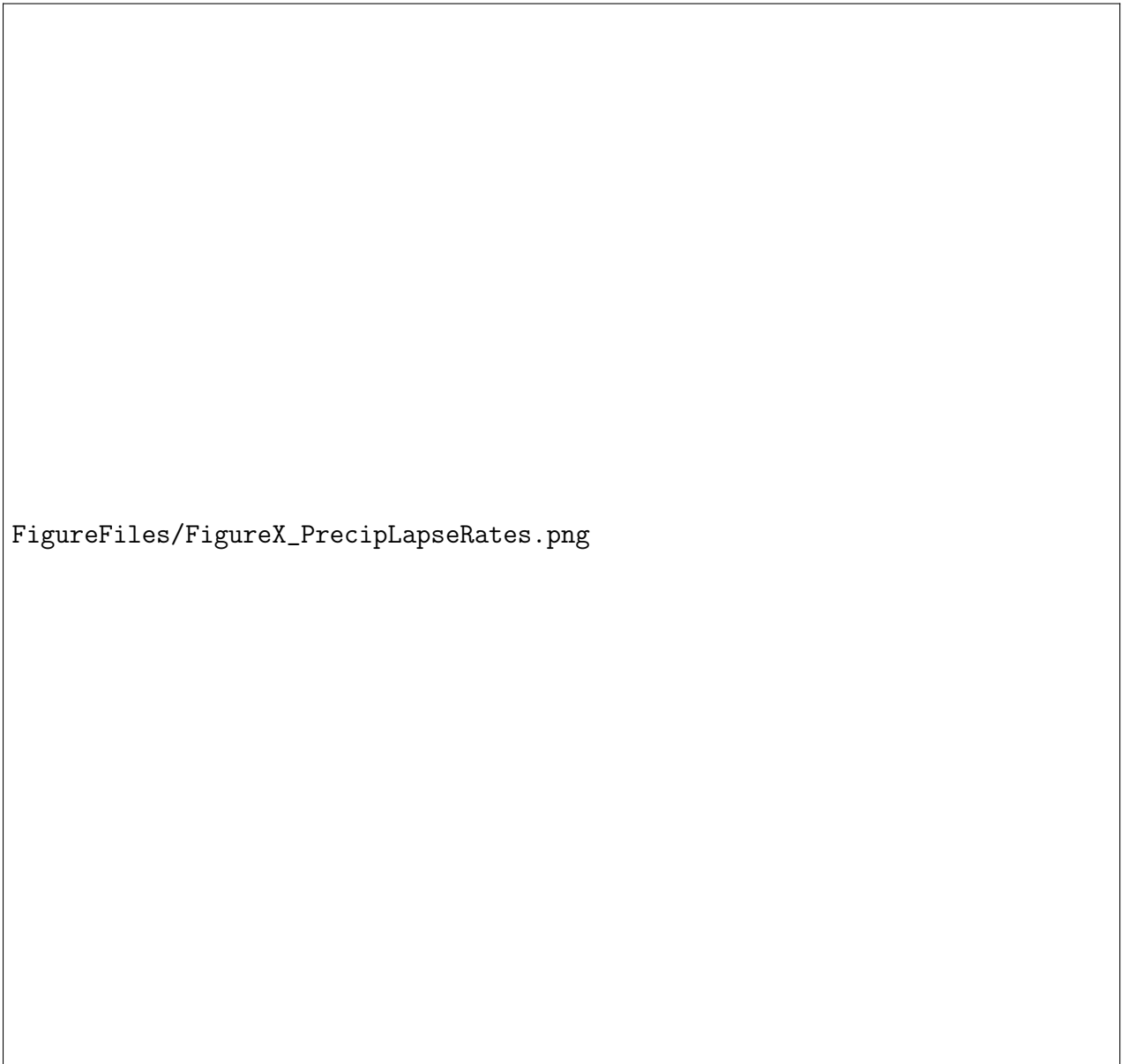


Figure 3.11: Average annual precipitation versus elevation for WRF and PRISM, for the entire WRF domain (top) and the East River Watershed (bottom). Rolling means (solid lines) are shown in addition to OPG lines (Equation 3.1)

Configuration Options	Outer/Inner Nest	
Version	3.8.1	
Vertical Levels	50, 50	
W-E Dimension	340, 349	
N-S Dimension	290, 328	
DX	3 km, 1 km	
DY	3 km, 1 km	
Output Timestep	hourly	
Model Physics	Option	
LBCs	(CFSR;CFSRv2)	Saha <i>et al.</i> (2010)
Convection Parameterization	None	
Microphysics	Thompson	Thompson <i>et al.</i> (2008)
LSM	Noah MP	Niu <i>et al.</i> (2011)
Surface Layer	Monin-Obukhov (Option 2)	
PBL	Mellor-Yamada-Janjic (Eta/NMM) PBL	
LW Radiation	Community Atmosphere Model	
SW Radiation	Community Atmosphere Model	

Table 3.1: Weather Research and Forecasting (WRF) parameters used in this study.

Fluxes	Description	Functional Form
Q_{12}	Percolation	$ku * (SM_1/SM_1^{max})^c$
Q_b	Baseflow	$ks * (SM_2/SM_2^{max})^n$
Q_{bo}	Bucket Overflow	$Q_{bo,i} = \text{MAX}[0, (SM_i - SM_i^{max})]$
Q_d	Overland Flow	$Q_{in} * Ac$ $Ac = \text{MAX}[0, 1 - (1 - SM_1/SM_1^{max})^\beta]$
ET	Evapotranspiration	$ET_1 = PET * (\text{MIN}[SM_1, SM_{T1}]/SM_{T1}^{max})$ $ET_2 = (PET - ET_1) * \text{MIN}[SM_2, SM_{T2}]/SM_{T2}^{max}$

Table 3.2: Functional Forms for Bucket Hydrologic Model Fluxes

Parameter	Description	Prior	Range
sm1max	unsaturated zone max storage	uniform	fixed
sm2max	saturated zone max storage	uniform	10.-1000.
ks	percolation rate	uniform	10.-10000.
ku	baseflow rate	uniform	10.-10000.
n	baseflow exp.	uniform	.001-3.
β	saturated area exp.	uniform	1.0-10.
c	percolation exp.	uniform	1.0-20.
fracten	field capacity fraction	uniform	0.1-.95
opg	Equation 3.1	uniform	.001-.005
m	Equation 3.1	uniform	.1 - 10
ζ	Equation 3.1	uniform	.1 - 10
γ	Equation 3.1	uniform	.1 - 10
t-lapse	temperature lapse rate	normal	-.004,.001
t-bias	temperature bias	uniform	-2., 2.

Table 3.3: Model parameter prior values and probability distribution used in the precipitation inference method. Ranges refer to the min/max of the uniform distribution or the mean/standard deviation of the normal distribution.

The WRF data used in this study covering the East river watershed is publicly available here: <https://data.ess-dive.lbl.gov/datasets/doi:10.15485/1845448>.

CHAPTER 4:
OROGRAPHIC PRECIPITATION
SENSITIVITY TO MICROPHYSICAL
PARAMETERIZATIONS: HYDROLOGIC
EVALUATIONS WITH SNOW LIDAR
DATASETS

Abstract

Cloud microphysical process representation is an important facet of numerical weather prediction, as cloud processes control distributions and rates of precipitation. The heterogeneity of ice crystal habits and phase change processes are particularly difficult to represent, and parameterization assumptions can impact both precipitation rates and the phases of hydrometeors reaching the surface. This has important implications for the accumulation, storage, and release of water in seasonal snow that is relied on by millions in the Western US, and billions globally. In this study, we design and conduct experiments with the Weather Research and Forecasting (WRF v.4.2.1) model using three different microphysics schemes, including a new scheme (Jensen) that includes 3 ice species and higher-order moments for ice-shape parameters. Simulations are conducted for two winter seasons (2018 and 2019) in a Colorado Rocky Mountain domain centered on the East River watershed. The precipitation efficiency of each microphysics scheme is quantified and compared using a drying-ratio mass budget approach. Precipitation outputs from each scheme are used to force a snow/land process model (Noah-MP) so that peak snow accumulations can be compared against high-resolution lidar products. We find that microphysical parameterization choice can significantly alter modeled precipitation, on the order of +/-10% of winter-season totals. All schemes produced snowpacks with high correlations with lidar products: RMSEs between 25-37% for snow depth and 19-51% snow water equivalent compared against ASO lidar maps. From this, we find that snow lidar data can be used not only as a non-traditional benchmark to test model precipitation accumulation fidelity,

but also a test of microphysical parameterizations when combined with distributed land/snow models. Better coupling between microphysical schemes, particularly those that predict snow densities and shapes, with snow models could potentially reduce uncertainties in snow densities. This study demonstrates that intensive field measurement campaigns are needed to better constrain microphysical parameterizations in mountain environments.

4.1 Introduction

Precipitation in mountains is highly variable in space and time, under-sampled by weather stations and radar, and challenging to model or estimate (Lundquist *et al.*, 2019). In mid-latitude regions, mountain precipitation often falls as snow, gradually accumulating as snowpacks that act as natural reservoirs supporting ecosystems and human systems across the watersheds into which they drain and beyond (Sturm *et al.*, 2017; Siirila-Woodburn *et al.*, 2021). The streamflow from snowmelt depends not only on the antecedent snow-volume, but also spatial location of snow accumulation throughout the watershed (Luce *et al.*, 1998). The variability of snow accumulation occurs at a range of process scales spanning individual hillslopes to synoptic scales (Clark *et al.*, 2011). A significant component of snow accumulation variability is caused by orographic precipitation enhancement, which results from a variety of dynamic mechanisms including stable upslope ascent from mechanical uplift, release of potential instabilities, lee-side convergence, seeder-feeder processes, and convection triggered by differential heating associated with changes in slope and aspect (Roe, 2005; Houze, 2012; Stoelinga *et al.*, 2013; Kirshbaum *et al.*, 2018). Convection-permitting atmospheric models (Prein *et al.*, 2015) have shown skill in capturing orographic precipitation variability caused by these effects and realistically modelling orographic precipitation variability (Ikeda *et al.*, 2010; Rasmussen *et al.*, 2011; Gutmann *et al.*, 2012; He *et al.*, 2019; Rudisill *et al.*, 2021). Consequently, numerical weather model precipitation outputs are increasingly used as reliable precipitation estimates and inputs for studies investigating the mountainous hydrologic cycle and water resource management (Lundquist *et al.*, 2019).

However, the predicted precipitation fields from atmospheric models exhibit errors

from a wide range of sources. For example, they are highly sensitive to the under-tested assumptions in microphysical parameterizations (Liu *et al.*, 2011; Minder & Kingsmill, 2013; Rhoades *et al.*, 2018). The problem is exacerbated in complex terrain where the location of falling precipitation upwind/downwind of an orographic barrier can have important hydrologic ramifications (Pavelsky *et al.*, 2012). Diagnosing errors in modeled precipitation includes evaluating the spatial characteristics of modeled orographic precipitation, and this necessarily includes some level of evaluation of microphysical scheme performance. Unfortunately, the lack of comprehensive precipitation observations in complex terrain (Lundquist *et al.*, 2019) creates an ill-posed process model development and diagnostic premise: modeled precipitation is highly sensitive to model structural and parameterization choices, but those choices are not easily evaluated with observations. Commonly used gridded precipitation datasets are highly uncertain in locations far away from observations (Henn *et al.*, 2018), but these datasets are still commonly used as reference "truth" when evaluating atmospheric model fidelity.

At the same time, airborne Light Detection and Ranging (LiDAR) scanning has increasingly been used in recent years to monitor watershed scale montane snowpacks and provides high spatial resolution ($\ll 1$ km) maps of snow depth (Painter *et al.*, 2016). Snow depths are measured using airborne LiDAR by first mapping the snow-free land surface. Subsequent flights during the snow season record snow-top heights, which are differenced from the bare-ground elevation to compute snowheight. The accuracy for snow-height measurements in flat terrain is ± 8 centimeters for a 1 meter swath (Deems *et al.*, 2013; Painter *et al.*, 2016). LiDAR snow depths can then be combined with model-derived snow density estimates to produce spatial maps of snow

water equivalent (SWE). Densities can be modelled using energy balance modeling (Hedrick *et al.*, 2018), and generally vary less than snow depths (Sturm *et al.*, 2010). Therefore, the measurements of snow depth at a watershed scale can inform direct estimates of SWE at the spatial resolution of the snow density estimates. Nevertheless, while the horizontal and vertical spatial resolution is high, LiDAR flights represent only a single snapshot of the state of the snowpack. Still, when measured near the peak of the accumulation season, SWE is a measure of the antecedent accumulation season processes, which are dominated by frozen-phase precipitation, that produced that observed snowpack state. Previous studies have leveraged the strong relationship between precipitation processes and measured snowpack by using LiDAR derived snow accumulation patterns to scale precipitation forcings for use in hydrologic modeling (Vögeli *et al.*, 2016; Pflug *et al.*, 2021), and to examine precipitation-elevation gradients (Kirchner *et al.*, 2014).

The goal of this study is to develop a framework for evaluating the land-atmosphere fluxes and fates of cold-season precipitation, from moisture delivery to snowpack development. This is accomplished by evaluating the sensitivities of simulated precipitation in the Weather Research and Forecasting (WRF) model (Skamarock *et al.*, 2008) to three different representations of ice-phase hydrometeors in microphysics schemes of varying complexity across two water years (WY2018-2019) covering Colorado’s East River watershed (ERW; 750 km²). We evaluate the Morrison *et al.* (2005), Thompson *et al.* (2008b), and recently developed Ice-Spheroids Habit Model with Aspect-ratio EvoLution (ISHMAEL) Jensen *et al.* (2017) schemes.

We focus our analysis on the ERW, which is a high elevation (2500-3500 masl), representative alpine basin and the location of numerous critical zone, snow, and

hydrologic studies (Hubbard *et al.*, 2018) in addition to being the location of a newly-deployed atmospheric observatory supported by the U.S. Department of Energy’s Atmospheric Radiation Measurement (ARM) program called the Surface Atmosphere Integrated Field Laboratory (SAIL) campaign (Feldman *et al.*, 2021). The Airborne Snow Observatory (Painter *et al.*, 2016) provides a LiDAR based snow depth and SWE product for WY2018-2019 covering the ERW. Consequently, this watershed is an ideal testbed for examining microphysical, precipitation, and snow processes, and model products will serve as guidance for hypothesis testing of ongoing field observation campaigns.

In this paper, we evaluate how information contained in limited, discrete measurements of the spatial distribution of the snowpack depth and SWE constrains precipitation processes in the ERW. We do so by undertaking a series of model sensitivity tests by comparing modeled fields to the range of hydrometeorological information collected in the ERW including from ASO. We evaluate precipitation sensitivity in several ways. First, we use a mass-balance, “drying ratio” approach to evaluate the precipitation efficiency of each scheme (Eidhammer *et al.*, 2018). Model biases are evaluated against three NRCS Snotel (Serreze *et al.*, 1999) locations within and in the vicinity of the ERW. Finally, we develop model experiments to examine patterns of precipitation accumulation against ASO snow LiDAR data captured near peak accumulation, and conclude with implications of these findings for developing multivariate observational constraints of precipitation microphysical processes.

4.2 Methods

4.2.1 Significance of Microphysical Processes

Mountains cause disturbances in the atmosphere which promote precipitation through a variety of dynamical mechanisms, including up-slope ascent from mechanical uplift, release of potential instabilities, lee-side convergence, seeder-feeder processes, and convection triggered by differential heating of mountain slopes and aspects (Roe, 2005; Houze, 2012; Kirshbaum *et al.*, 2018). Regardless of specific precipitation processes, orographic uplift produces airmasses of enhanced cloud condensate over certain regions of high-altitude, complex terrain. How, if, and where those cloud condensates ultimately reach the surface as precipitation depends on cloud microphysical processes governing hydrometeor (i.e., rain, snow, graupel, etc.) growth and sedimentation (Roe, 2005). Moist processes can also influence the dynamics through latent heat release (Jiang, 2003) as can radiative processes.

It has been repeatedly shown that the representation of microphysical processes in regional through global atmospheric models can have a significant impact on orographic precipitation (Liu *et al.*, 2011; Rhoades *et al.*, 2018). Modifying hydrometeor fall speeds alone can influence the enhancement of precipitation upwind/downwind of orographic barriers (Pavelsky *et al.*, 2012). Figure 4.1 illustrates an idealized depiction of some of the most prominent cloud microphysical processes that control distributions of orographic precipitation. The rates of primary heterogeneous ice nucleation/droplet formation, timescales and efficiency of conversion from non-precipitation species (cloud ice/water) to precipitating species, precipitating hydrometeor fall speeds (a function of hydrometeor shape/drag coefficients), and feedbacks

between microphysics and the dynamics (through latent heating and radiation) can all conceivably influence the spatiotemporal distribution of precipitation within a particular mountain watershed. (Minder & Kingsmill, 2013), for instance, found that the simulated snow-line near mountain ranges is sensitive to cooling rates from melting hydrometeors in the Sierra.

Microphysical Parameterizations

Fundamentally, the microphysical parameterizations in atmospheric models attempt to describe this removal of atmospheric water from a given model grid-cell based on kinematic and thermodynamic conditions (Morrison *et al.*, 2020). Parameterization methods in operational models typically use “bulk” approaches, where the hydrometeor mixing ratio (mass per mass of dry air), number concentrations (particles per unit volume), and other hydrometeor properties are predicted for a limited number of species (graupel, rain, cloud-water, etc).

In this study we test the Thompson (Thompson *et al.* (2008b); hereafter MP08), Morrison (Morrison *et al.* (2005); hereafter MP10) and ISHMAEL (Jensen *et al.* (2017); hereafter MP55) microphysical schemes (Table 4.2.1). Each treats ice phase hydrometeors and growth processes in different ways. The MP10 and MP08 both use 5 separate hydrometeor categories: cloud liquid, cloud ice, snow, graupel, and rain and predict mixing ratios for each. MP10 predicts the number concentration for ice, rain, snow and graupel, whereas MP08 only predicts the number concentration for rain. In MP10, all hydrometeors are assumed to be spherical, with mass-density relationships given by $m(D) = \pi/6\rho_s D^3$. MP08 is similar but describes snowflakes as approximately planar, with mass-diameter relationships given by $m(D) = 0.069D^2$.

Scheme	Abbreviation	Reference
Thompson	MP08	(Thompson <i>et al.</i> , 2008b)
Morrison	MP10	(Morrison <i>et al.</i> , 2005)
Ishmael	MP55	(Jensen <i>et al.</i> , 2017)

Table 4.1: WRF v4.1.2 Microphysics options examined in this study.

The most sophisticated scheme tested is the MP55 scheme which forecasts higher-order moments of hydrometeor species beyond mixing ratios and number concentrations at the expense of a higher computational cost. The MP55 scheme uses three ice categories in place of snow/graupel categories and models the evolution of snowflakes as oblate spheroids with two evolving axes a_i and c_i , such that the particle mass is given by $m(a, c) = \rho_i \frac{4}{3} \pi a_i^2 c_i$. Here, a_i is half the major axis for plate-like crystals and half the minor axis for column-like crystals, and c_i is half the minor axis for plate-like crystals and half the major axis for column-like crystals. Consequently, MP55 explicitly models both columnar and dendritic ice-habits (characterized by different a_i/c_i ratios), and the temperature dependent nucleation of each of these forms. It is important to note that the growth processes (e.g., collection, vapor deposition) depend on the particle aspect ratio. Although we highlight some of the differences across the three microphysics schemes, there are a variety of other differences between the schemes, and a full accounting is beyond the scope of the present study.

4.2.2 Weather Research and Forecasting (WRF) Model Configuration

This study uses the WRF model (Skamarock *et al.*, 2008). WRF solves the compressible, non-hydrostatic Euler equations using a third order Runge-Kutta timestepping method. For practical reasons, two different WRF versions are used. Both simulations use a two-way nested domains. The Thompson (MP08) simulations use WRF v3.8.1, whereas the MP10 and MP55 simulations use WRF v.4.1.0. This differ-

ence in WRF versions is because the MP08 simulation had already been generated for another study, however we do not expect the differences in WRF model versioning to significantly influence our results. Table 4.2.1 lists WRF subgrid-scale parameterizations used in this study. Lateral boundary and initial conditions for the WRF simulations are provided by the Climate Forecast System Reanalysis Version 2 (CFSRv2; (Niu *et al.*, 2011)). CFSRv2 has a 0.5° horizontal resolution (~ 55 km), and lateral boundary conditions are provided every 6 hours. Two nested domains are used, a ~ 3 km outer (230x349 grid cells) and a ~ 1 km inner grid (349x391 grid cells). A two-week spin-up period is used prior to the October 1 start date for each model run. Outputs from WRF are used to run an “offline” (not coupled with WRF), higher resolution (250m dx/dy) Noah-MP land surface model. The version of Noah-MP used for the offline simulations come from (Gochis *et al.*, 2018).

In this study, WRF is run for WY2018-2019 (March 31 and April 7th, respectively). These water years fortuitously represent an extreme wet and an extreme dry year. The WRF meteorological outputs are then used to force a high-resolution (250m dx/dy) offline land model (Noah-MP; Niu *et al.* (2011)), providing peak SWE and snow depth accumulations that are comparable to the spatial resolutions provided by the ASO LiDAR-derived snow product.

4.2.3 Metrics to Quantify Precipitation and Snowpack Fidelity

The efficiency of each microphysical scheme is evaluated using the drying ratio method (Eidhammer *et al.*, 2018). The drying ratio is the precipitation flux normalized by flux of the time-integrated integrated vapor transport, the components of which are given by:

$$F_{u,x} = -\frac{1}{g} * \int_{P_0}^{P_{Top}} \int_x \int_t q \mathbf{U} dP dx dt \quad (4.1)$$

$$F_{v,y} = -\frac{1}{g} * \int_{P_0}^{P_{Top}} \int_y \int_t q \mathbf{V} dP dy dt \quad (4.2)$$

Where P is atmospheric pressure, U and V are meridional/zonal winds, and q is the water vapor mixing ratio (kg/kg). The drying ratio is then given by:

$$DR = \frac{P}{F} \quad (4.3)$$

Where $F = F_{v,y} + F_{u,x}$ and P is the precipitation mass in kilograms. The drying ratio calculation makes several assumptions, following (Eidhammer *et al.*, 2018). Condensed hydrometeors (cloud water, rain, snow, etc) are not included in q as the overall contribution is small. The drying ratio is a useful metric for comparing precipitation derived from an regional climate model, since internal variability can potentially lead to important shifts in storm tracks and moisture delivery. Drying ratios also help to elucidate the mechanisms that lead to higher and lower snow accumulation in different WYs. Variations in annual precipitation/snow could be the result of 1) variations in the efficiency of precipitation generating processes alone, or 2) constant precipitation efficiency but variations in the incoming moisture flux.

Several different metrics are applied to better compare the Noah-MP model versus ASO LiDAR datasets. Two primary quantities are assessed: the spatial locations of snow accumulation within the watershed, and the total watershed storage of snow.

To assess spatial pattern similarity, we use an objective function described in (?).

That work's spatial efficiency metric is denoted as SPAEF and is given by:

$$SPAEF = \sqrt{(1 - \gamma)^2 + (1 - \beta)^2 + (1 - \alpha)^2} \quad (4.4)$$

Where γ is the histogram mismatch, given by:

$$\gamma = \frac{\sum_{i=1}^n \min(K_i, J_i)}{\sum_{i=1}^n K_i} \quad (4.5)$$

and

$$\beta = \frac{\sigma_A / \sigma_B}{\mu_A / \mu_B} \quad (4.6)$$

and $\alpha = \rho(A, B)$, and where ρ is the Pearson correlation coefficient. The histogram-matching is performed on the Z-transformed data. Consequently, the SPAEF is designed to be a measure of spatial similarity between two datasets A and B that is insensitive to biases in those datasets. The Pearson correlation is given by:

$$\rho = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (4.7)$$

Where x and y are Noah-MP and ASO LiDAR fields (SWE or snow depth), respectively.

In addition to the SPAEF and the Pearson correlation, we evaluate the total watershed snow storage (TWSS) bias, given by:

$$TWSS_{bias} = \frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{n} \sum_{i=1}^n y_i \quad (4.8)$$

Finally, we compute the percent root mean square error (pRMSE), given by:

$$pRMSE = \frac{\sum (x_i - y_i)^2}{\sum y_i} * 100. \quad (4.9)$$

This is done because the errors at individual grid points can be expressed so that they are normalized by the magnitude of the snow (SWE or snow depth) for that water year.

4.2.4 Snowpack Modeling and ASO LiDAR Data Processing

Snowpack spatial variability, at the peak of the accumulation season, is shaped by a combination of 1) precipitation variability, 2) slope scale preferential deposition, 3) secondary redistribution (blowing snow), and 4) melt/sublimation/loss processes (Mott *et al.*, 2018). Avalanches also redistribute snow on steep slopes greater than 30°. In order to use ASO LiDAR snow data to evaluate precipitation variability, secondary redeposition and loss processes must be taken into account. The effects of 2) and 3) are mitigated in part by resampling ASO LiDAR data from 50m to 1km. Figure 4.2 shows the effects of bi-linearly resampling snow depth data from 50m to 1km; wind redistribution are clearly present on the windward/leeward sides of ridges and are generally on the order of 50-250 meters in length. Resampling data to 1km largely smooths out these features. Process 4) is accounted for by running an energy balance model, specifically Noah-MP (Niu *et al.*, 2011), to account for the evolution of falling snow as it metamorphoses on the ground.

The ERW is a high elevation, continental basin with cold temperatures, so we hypothesize that both rain and melt prior to peak SWE are relatively minimal basin-wide. This hypothesis is confirmed by analyzing SNOTEL data in the watershed, as the April 1 SWE is within +/- 2% of the accumulated precipitation from October 1 for the two years examined, and average two-meter surface air temperatures are

-4.5°C at the SNOTEL locations for the same time period. Nevertheless, to better account for sub-kilometer scale, terrain related melt processes, Noah-MP is run at a 250 meter resolution using the resampled digital elevation model distributed by ASO. The hourly WRF output variables are then bi-linearly interpolated to the 250 meter regular latitude longitude grid of ASO. Then, shortwave radiation, temperature, pressure, and specific humidity are adjusted to account for terrain differences between the native WRF elevation grid and the higher-resolution ASO distributed grid. Temperatures (via the constant dry adiabatic lapse rate) and pressures (by the hydrostatic relationship) are adjusted to match the updated digital elevation model. Specific humidity is downscaled by assuming that the relative humidity is conserved between the downscaled forcings and original the WRF data, and specific humidity is adjusted to match the corrected air temperature. The WRF downwelling shortwave radiation is converted to terrain-normal shortwave radiation using terrain-geometry and solar angle relationships (Dingman, 2015), using the slope and aspect from the high resolution DEM. Terrain shadowing is not accounted for. Longwave radiation and winds are not adjusted, though corrections for terrain effects on shortwave and longwave radiation could improve the simulations and could be pursued in future work. Some studies have further downscaled wind-fields using empirical terrain relationships (Liston & Elder, 2006) or physically-based solvers (Reynolds *et al.*, 2021). Since Noah-MP does not simulate wind redistribution, the benefits of more finely resolved wind fields are likely small (though wind velocities do control rates of latent/sensible heating). The code to perform the forcing corrections is available on GitHub (https://github.com/bsu-wrudisill/wrf_ERriv_mphys_aso).

4.3 Results

4.3.1 Precipitation Evaluated at SNOTEL Locations

Three NRCS SNOTEL sites are compared against timeseries of precipitation and two-meter surface air temperature from each WRF microphysical scheme. The Schofield, Butte, and Taylor are located to the North, in the Center, and to the East of the ERW, respectively, and are each located approximately 20 km away from each other (Figure 4.7). The spatial scales of orographic precipitation variability is apparent from looking at the SNOTEL data alone, as the Schofield station receives almost twice the precipitation of the Butte site, and each site receives almost double the precipitation in 2019 compared to the previous year (1200 mm versus 650 mm at Schofield, for example).

There are several consistent patterns with respect to precipitation accumulations and microphysical scheme. MP55 consistently produces the most precipitation (across all sites and both years), and MP08 generally has the least precipitation (all but the Taylor in 2019). Most notably, perhaps, is that MP08 has a significant dry-bias during the second half of December, 2017 compared to both SNOTEL sites and the other MP options. No such prolonged dry-bias is present in the following water year. Two-meter surface air temperatures are also compared against SNOTEL locations. The WRF simulated two-meter air temperatures are systematically cold biased by approximately $3^{\circ}C$ across microphysical schemes.

The total accumulated precipitation during the study period, across the ERW (not shown) has a relative standard deviation of 10.0% for water year 2018 and a 3.0% for water year 2019. The differences in precipitation volumes can be expressed

as the efficiency of dynamical/microphysical processes for converting the incoming water vapor flux into precipitation. The drying ratio (DR) is computed on a 145 by 180 grid cell box surrounding the East river watershed. Figure 4.4 shows the DR averaged over October 1 to April 1 for each scheme and each WY. MP55 consistently has the highest DR, with almost double the DR values from WY2018 to WY2019.

4.3.2 Vertical Profiles

In order to examine the height-dependent creation/fate of hydrometeor species, the time and space averaged distributions of each mixing ratio predicted for each scheme are presented (Figure 4.5). The averaging box is the same 145 by 180 grid cell box surrounding the ERW used to compute the DRs. Snow (MP10, MP08) and "ICE1" (Mp55) have the highest mixing ratios, with peaks between 4-5 km a.s.l. Graupel, Rain (MP08, MP10) and ICE1, ICE2 (MP55) have much lower mixing ratios than Snow/ICE1.

In order to better understand some of the atmospheric dynamics and precipitation generating mechanisms, the temporally averaged (October 1 to April 1) cross section views of microphysical quantities, cross-sectional winds (U and W components), and vertical velocities show the different locations of ice-phase hydrometeor creation and fate, in addition to illuminating some of the precipitation relevant dynamics (Figure 4.6). The ice-phases are lumped together snow and graupel for MP8 and MP10, the three ice species (called QICE, QICE2 and QICE3 in the model code) in MP55. QICE1 is the primary ice category present in MP55 (not shown). For all cases, the highest densities form a plume above of the western watershed boundary, concentrated near the surface and decaying with height. MP55 has the highest densities across microphysical schemes, with a region of 3.0 g/kg during water year 2018. There is a

consistent negative vertical velocity component on the lee-side of the western ridge. Upstream of the ridge, there is a consistent low-level jet, characterized by a reversal in the zonal wind direction (green dots; Figure 4.6).

4.3.3 Modeled SWE and Snow Depth - Comparisons Against ASO LiDAR

Figure 4.8 shows the results of comparing snow depth and SWE between ASO LiDAR products and each Noah-MP model with the three different forcing regions (MP08, MP10, MP55). Notably, the ASO LiDAR derived snow depths are converted to densities using the average snow density from the three Noah-MP model runs. Ultimately, the differences in model snow density is quite small between the three Noah-MP runs (not shown). Each model is aggregated to a 1km resolution. The northwest region of the watershed collects the most snow. The under representation of snowfall during 2018 for MP08 (Figure 4.3) is also apparent when compared against ASO LiDAR data. Even at a 1km resolution, ASO has a more variable pattern of snow accumulation and higher maxima than any of the WRF-forced Noah-MP cases for both WYs. The differences between MP08 and MP10 are fairly small during 2019. For both 2018 and 2019, MP08 has the highest Pearson's correlation coefficient for both SWE and snow depth. MP10 has the highest SPAEF for both years. Still all models have a good spatial correlation with the ASO LiDAR derived SWE, with a minimum of 0.90 for MP55 in 2018 and maximum of 0.94 in 2018 for SWE.

Table 4.5 shows summary statistics of the SWE and snow depth comparisons. Measures of mass-balance are better for the microphysical schemes with higher-order treatments of ice species (MP10, MP55). The lowest pRMSE is 18.9% for SWE and 35% for snow depth for MP55 in 2019. The TWSS-Bias% is always lower than the

pRMSE, as it does not express the average grid-cell by grid-cell error but rather the differences in the total storage. The TWSS-Biases are the lowest in 2019, when all schemes have less than 6.5% for both snow depth and SWE. The TWSS-Biases are higher during the low-snow year.

The trend between elevation and snow accumulation illustrates some additional important differences between the ASO LiDAR data and WRF simulated products (Figure 4.9). In each case, there is relatively little modelled melt except for the lowest elevations (not shown). The average ASO-SWE increase with elevation follows a linear pattern when a 200-grid cell rolling-mean window is applied, which approximately flattens out above approximately 3500 m. The snow depths (not shown) show the same leveling-off, so this is a function of a decrease in depths, not just an artifact of modeled densities. The slope of the SWE versus elevation line is higher for water year 2019 and parallels the 0.63 millimeters of SWE per meter of elevation, whereas 2018 more closely parallels the 0.50 mm/m line. The Noah-MP model SWE shows no such leveling out with elevation. The slopes of the Noah-MP curves are less-steep than the ASO LiDAR data and shows the greatest spread during 2018. The variance of the ASO LiDAR data increases with the magnitude, which is not found in the WRF/Noah-MP modeled SWE.

4.4 Discussion

This is the first study, to our knowledge, that has used airborne LiDAR derived snow products, with exhaustive spatial coverage but limited temporal coverage, to evaluate the representation of cloud microphysical processes in numerical weather models. Other studies have used snow LiDAR data to evaluate snow accumulation processes in various ways. (Kirchner *et al.*, 2014) evaluated both PRISM (Daly *et al.*,

2008) precipitation gradients and optical satellite imagery based SWE reconstructions against analyzed snow depths measured by airborne LiDAR near peak SWE accumulation for a basin in California. The tapering-off of the snow depth/SWE with elevation relationship (Figure 4.9) was also observed in that study. This could be caused by a combination of 1) enhanced sublimation at windier, unsheltered high-elevation regions 2) preferential deposition processes by low-level wind fields interacting with ridges (Lehning *et al.*, 2008) or 3) larger-scale orographic precipitation dynamics.

Even though there are gross mismatches between the temporal scales of cloud microphysical processes and the snapshots of the snowpack collected by ASO LiDAR, we show through a careful intercomparison of WRF simulations with different microphysics schemes over the ERW and ASO snow products, that modeling of the total accumulation of snowpack during the winter in high-altitude complex terrain must consider, and perhaps even focus on, ensuring a competent representation of cloud microphysics. To be clear, there are many other snow processes that must be understood to explain the elevation patterns of snow accumulation the snowpack (Figure 4.9) including high-elevation snow deposition, redistribution, and sublimation processes. Additionally, we should note that all Noah-MP models had higher SWE and snow depth values at the lowest elevations (less than 3000m), which could be caused by a combination of underestimation of densification processes, too much precipitation in valley bottoms, or not enough melt or sublimation loss prior to the ASO LiDAR acquisition dates. Vegetation densities also change with elevation (Figure 4.7) which may influence the aforementioned snow processes through snow-canopy interactions.

More tightly coupling the land surface model and microphysical schemes may

improve the study in a number of ways. The current version of Noah-MP does not accept solid/liquid precipitation phases as input, and instead uses a partitioning scheme from (Jordan, 1991). Surface air temperature based methods do not always account for the range of microphysical processes, such as cooling from latent heat release near the surface, that can lead to solid accumulations at a wider range of temperatures. This might be a major limitation in another watershed with lower elevations/warmer temperatures, but for each scheme tested a small percentage of the precipitation fell as rain, regardless of partitioning method. Even after bias correcting the two meter surface air temperature uniformly across the domain for the -3°C bias compared against SNOTEL data (Figure 4.3) only the modeled SWE in the lowest elevations of the watershed were significantly impacted. Consequently, performance of the different microphysical schemes with respect to rain-snow transition zones is untested and could be an area of future research. Another interesting product of this research is the large modeled drying ratio for WY2019. (Eidhammer *et al.*, 2018) reports lower ratios more similar to that of WY2018, also for a region in Colorado, but for individual storm events as opposed to an entire season. Some of the differences are attributable to different WRF configurations and regions used to compute the drying ratio. This investigation shows the increase in efficiency is largely responsible for the higher precipitation in 2019 compared to 2018, as opposed to an increase in water vapor flux. Drying ratios as high as 0.5 have been reported for the Andes (Smith & Evans, 2007).

Better coupling between the microphysics scheme and snow model in Noah-MP could also improve simulated snow densities. Density depends on snow metamorphosis and compaction due to overburden. Densification due to overburden follows

Sun *et al.* (1999) and the energy dependent metamorphosis components follow (Anderson, 1976). The snow depths assume an initial density of 10% for new snowfall after which densification processes initiate (Niu *et al.*, 2011). A path for potentially improving snowpack depth, density, and SWE could be to modify the Jensen (MP55) or Thompson (MP08) schemes and the Noah-MP code so that prognostic densities of snow and/or graupel are used by the land surface model. The Morrison scheme (MP10) does not treat snow or graupel density as a prognostic variable and the quantities are fixed at 100 and 500 kg/m^3 respectively, so better coupling MP10 would be of less utility. Moreover, coupling advanced schemes such as MP55 that explicitly model hydrometeor shapes with snow process models may have utility for snow remote sensing applications.

Whether or not energetic forcings, such as shortwave and longwave radiation and sensible and latent heat fluxes, that contribute to melt/densification are well represented at a kilometer scale is a significant source of uncertainty in this study. For instance, three dimensional longwave radiation effects from complex terrain are not considered by WRF and can be significant (Feldman *et al.*, 2021). With that said, the observed cold-bias in WRF has been observed in other climate models (Rhoades *et al.*, 2018) and may potentially be related to longwave radiation processes, too-stable boundary layers over snow surfaces inhibiting heat exchanges (Slater *et al.*, 2001), and/or other compensating biases (e.g., cloud cover). However, given the overall cold surface air temperatures in the ERW, sensitivity tests showed that correcting for biases had relatively little effect on simulated peak SWE, particularly for WY2019. Temperatures at SNOTEL stations have also been shown to have quality control and calibration issues (Oyler *et al.*, 2015) that may or may not be accounted

for here. Therefore, we caution to the hydrometeorological community that a more rigorous scrutiny of the temperature fields provided by WRF simulations in complex terrain is needed. For instance, comparing model outputs with ASO LiDAR snow datasets during the ablation season was not performed for this study, but could be another avenue to decompose simulated biases in temperature and radiation and identify systemic structural issues in models.

A number of other studies have examined orographic precipitation sensitivity to microphysical parameterizations in the WRF model. Hughes *et al.* (2020) found that single moment microphysics (WSM6) schemes were wetter than double-moment (Thompson/Morrison) when evaluated over a single WY in the Sierra Nevada (WY 2008), and that precipitation accumulation was more sensitive to microphysical parameterization choice than to lateral boundary conditions for convection-permitting WRF simulations in the Sierra Nevada. Liu *et al.* (2011) examined a similar domain and tested both MP10 and MP08 for one, six month period, and found that both performed similarly, though MP08 had slightly less precipitation. Our simulation also produced very little graupel from either MP08 or MP10 (Figure 4.5). Transitions from snow to graupel categories can introduce artificially abrupt transitions in particle properties (such as density and fall speed) and avoiding this is among the reasons that MP55 chose to explicitly predict particle shapes rather than a graupel category (Jensen *et al.*, 2018). Consequently a simulation with more graupel might show more spatial variability in precipitation between the microphysics schemes. Finally, while three schemes are tested in this study, each contains a range of parameters that each have an uncertainty space that has been under-explored. Idealized simulations show that perturbing individual parameters within individual microphysics schemes can

have a similar impact to using an entirely different microphysics scheme (Morales *et al.*, 2019). Future studies may consider producing an ensemble of simulations that sample across a plausible range of these parameter values, informed by new observational campaigns focused on mountain precipitation (Feldman *et al.*, 2021).

4.5 Conclusions

High-altitude complex terrain is undergoing profound changes which are setting the stage for much-reduced snowpack in the coming years and decades. The details of the snowfall that produces this snowpack are central to understanding the potential for changes in precipitation amount and phase. Nowhere is this more apparent than in the Upper Colorado River Basin, which is dramatically stressed due to both long-term trends (Milly & Dunne, 2020) and recent extreme drought. Since the ERW represents a focused area of observations and research, collocated data and models of the ERW provide the opportunity to develop new tests of uncertain model processes. This work focuses on one such test, and considers cloud microphysics because they are under-constrained, are very difficult to observe directly, and strongly influence the spatiotemporal distribution of precipitation.

This study used a high resolution ASO LiDAR snow dataset to evaluate precipitation and snowpack fidelity for three different microphysical schemes implemented in the WRF model in the ERW across both a high precipitation (2019) and low precipitation (2018) cool-season (October 1 - April 1). Model results suggest the magnitudes of precipitation between the years were more controlled by precipitation efficiencies (higher/lower drying ratios) compared to changes in water vapor flux with the higher-moment microphysical schemes (MP10 and MP55) simulating higher drying ratios. All microphysical schemes were able to capture the bulk precipitation

magnitude evaluated at three NRCS SNOTEL sites, with an average bias of $10 \pm 16\%$. Watershed total accumulated precipitation had a 10.0% relative standard deviation among schemes for water year 2018 and a 3.0% relative standard deviation for 2019. The MP55 scheme had the highest drying ratio and better matched both SNOTEL and ASO LiDAR observations for the high-snow volume year, but overpredicted precipitation and snowpack for the dry-year (2018). Each microphysics scheme resulted in the development of snowpacks with a high spatial correlation with ASO LiDAR datasets at a $\sim 1\text{km}$ scale, but the MP08 had the highest Pearson’s correlation for both years examined. Percent root mean square errors between the 1km ASO LiDAR snow depth and WRF model products were on the order of 25-37%, and 51% (MP10 in 2018) and 18% (MP55 in 2019) for SWE.

This study found ASO LiDAR snow datasets can help evaluate microphysical scheme fidelity and suggests that better coupling between microphysical schemes and land-surface model schemes is needed. While this study used only two ASO LiDAR flights, other studies with longer data coverages have shown repeatability of snow patterns to scale precipitation inputs into hydrologic models (Vögeli *et al.*, 2016; Pflug *et al.*, 2021). In the two ASO LiDAR flights used here, the locations of peak accumulation are consistently on the northwestern ridge of the ERW (Figure 4.8) which is on the windward side and the location of the strongest uplift (Figure 4.6), suggesting that cloud microphysical representations and potentially coupling to land-surface schemes may have enduring importance for simulating snowpack over longer time-horizons.

The ability of any existing schemes to perform in out-of-sample conditions must be demonstrated, however. There is a potential for field campaigns such as the Sur-

face Atmosphere Integrated Field Laboratory (SAIL) for increasing ASO LiDAR data collection. Such data can and should be used to further constrain specific model microphysical process representations to establish if one or more schemes produce consistent results relative to observations across more hydroclimatological states than we tested here. We have demonstrated here that snowpack surveys can constrain precipitation microphysics across two water years, but the question of whether snowpack surveys consistently constrain microphysics remains to be demonstrated.

Such a demonstration would serve integrated mountainous hydroclimate research well: microphysical schemes that are of sufficient fidelity and complexity to the microphysical processes that produce precipitation amount and phase patterns will be increasingly important over those longer horizons as changes in dynamics and thermodynamics force high-altitude mountainous watersheds.

Parameterization	Option	Reference
Convection	None	
Microphysics	Thompson (MP08)	(Thompson <i>et al.</i> , 2008b)
	Morrison (MP10)	(Morrison <i>et al.</i> , 2005)
	Ismael (MP55)	(Jensen <i>et al.</i> , 2017)
LSM	Noah-MP	(Niu <i>et al.</i> , 2011)
Surface Layer	Monin-Obukhov (Option 2)	
PBL	Mellor-Yamada-Janjic (Eta/NMM) PBL	
LW Radiation	Community Atmosphere Model	
SW Radiation	Community Atmosphere Model	

Table 4.2: Weather Research and Forecasting (WRF) subgrid-scale parameterizations used in this study.

Variable	Year	Model	Spatial		Mass Balance		
			ρ (unitless)	SPAEF (unitless)	pRMSE %	TWSS-Bias (mm; m)	TWSS-Bias (%)
Snow Water Equivalent (mm)	2018	NoahMP-MP08	0.918	0.677	51.9	-50.687	-20.699
		NoahMP-MP10	0.911	0.679	29.9	5.112	2.087
		NoahMP-MP55	0.898	0.550	30.9	46.302	18.909
	2019	NoahMP-MP08	0.936	0.628	21.7	-42.529	-6.286
		NoahMP-MP10	0.923	0.629	21.9	-37.002	-5.469
		NoahMP-MP55	0.917	0.620	18.9	21.154	3.127
Snow Depth (m)	2018	NoahMP-MP08	0.912	0.636	29.5	-0.104	-12.571
		NoahMP-MP10	0.905	0.658	25.2	0.0348	4.205
		NoahMP-MP55	0.896	0.539	30.0	0.159	19.27
	2019	NoahMP-MP08	0.933	0.600	37.2	-0.113	-6.24
		NoahMP-MP10	0.923	0.623	37.8	-0.111	-6.163
		NoahMP-MP55	0.917	0.604	35.2	0.055	3.052

Table 4.3: Spatiotemporal and mass-balance error statistics for Noah-MP models compared against the ASO LiDAR derived basin-wide snow depth and Snow Water Equivalent (SWE) estimates. Bold values denote the best performing scenario.




Figure 4.1: Conceptual diagram illustrating microphysical rate controls on orographic precipitation. Secondary controls on slope scale snow deposition/redeposition are also shown.



Figure 4.2: Airborne Snow Observatory (ASO) LiDAR derived snow depths for April 7, 2019 at three different resolutions (50m, 500m, and 1km) using bilinear interpolation. The black box in the left hand figures corresponds to the latitudinal and longitudinal extent of the figures on the right.



Figure 4.3: WRF total precipitation (ACCPRCP; bottom) and two meter surface air temperature (Tair; top) compared against three NRCS SNOTEL sites, Butte (site number 380), Schofield (site number 737), and Taylor (site number 1141), for water years 2018 (first row) and 2019 (second row).



FigureFiles/FigureX_DryRat.png

Figure 4.4: Drying Ratios (total precipitation normalized by incoming water vapor flux) within the Upper Gunnison region computed for October 1 to April 1 for water years 2018 and 2019 across each WRF microphysics scheme.



Figure 4.5: Spatiotemporal averaged mixing ratios for different hydrometeor species estimated (if applicable) from the three WRF microphysics schemes (mp08, mp10, and mp55) used in this study. The y-axis units are kilometers above sea level.

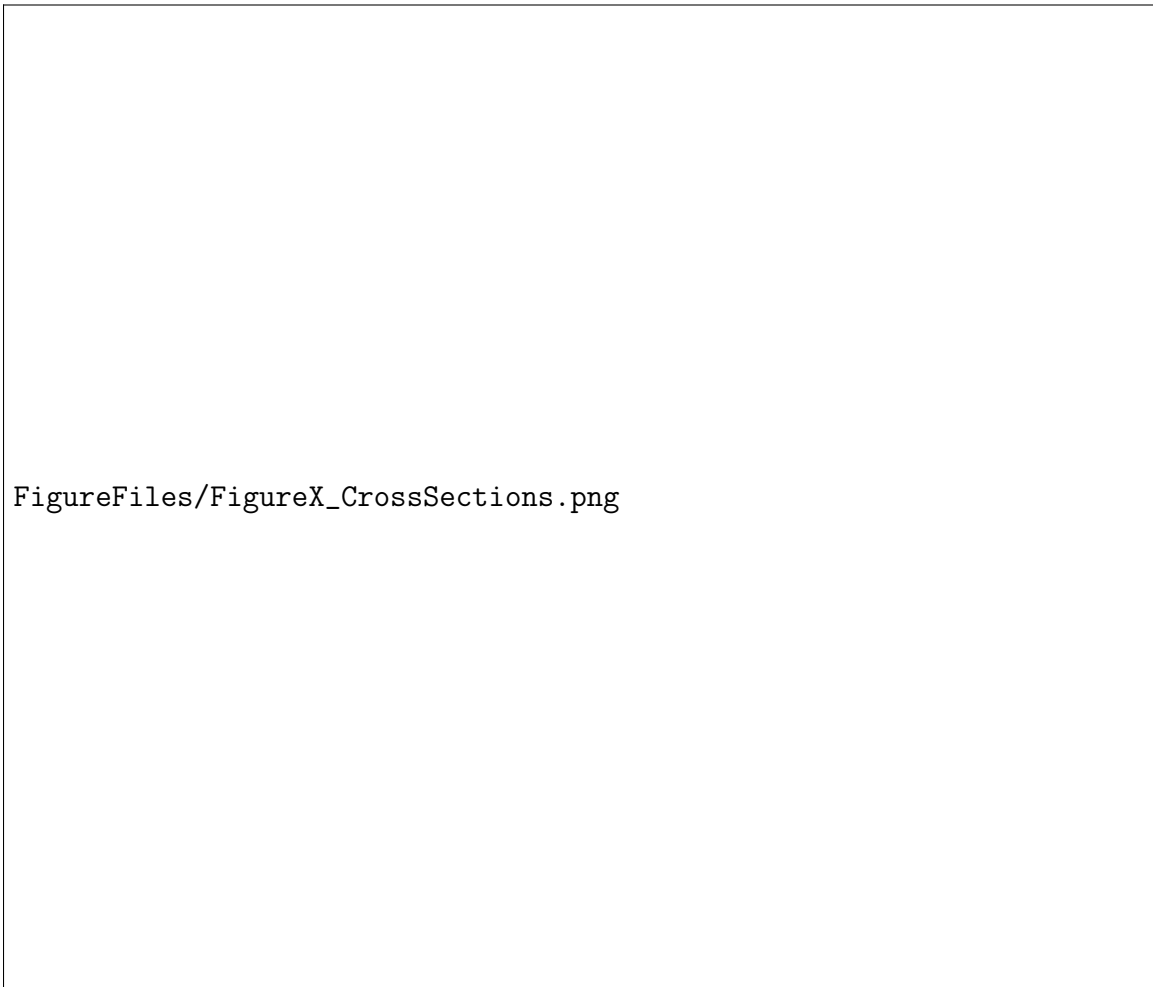
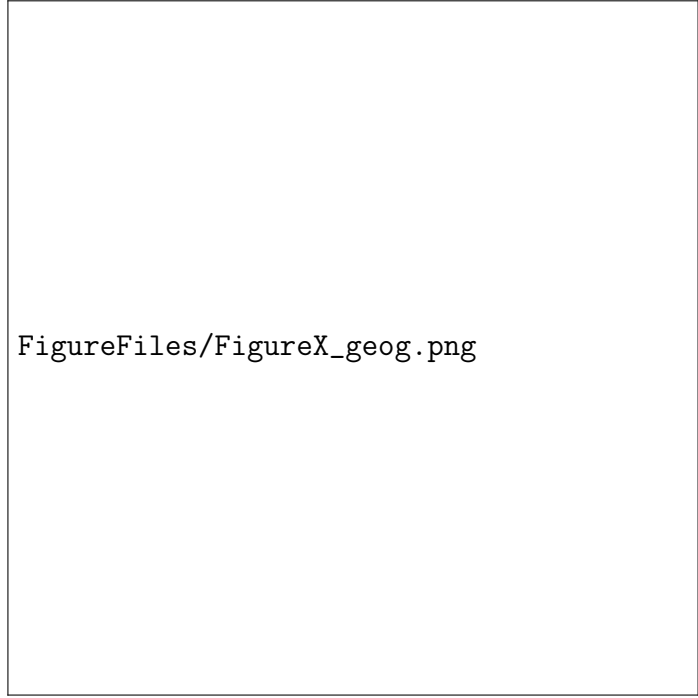


Figure 4.6: Cross sections of average directions of vertical windspeed (red/blue shading; units of m/s), vertical and zonal flow (arrows; units of m/s), and ice-phase hydrometeor concentrations (contours; units of g/kg dry air). Green dots show the regions where the average meridional wind speed reverses and is greater than 1 m/s). Average water year total precipitation (precip) are provided above each plot. Water year 2018 (2019) is shown in the top (bottom) row.



FigureFiles/FigureX_geog.png

Figure 4.7: Noah-MP static geographic data used for offline snow modeling. East River Watershed topography from the ASO LiDAR digital elevation model (DEM; upper left), terrain aspect (upper right), USGS 24 category vegetation classification type (bottom left), and MODIS satellite derived green-vegetation fraction (Veg Frac; bottom right). NRCS SNO-TEL locations (black dots) are overlain on each geographic dataset.



Figure 4.8: SWE (top two rows) compared between ASO LiDAR and Noah-MP modeled SWE for WY2018 and WY2019, evaluated near the date of peak snow accumulation. Analogously, snow depth (SD; bottom two rows) is shown.



Figure 4.9: SWE versus elevation within the ERW. ASO LiDAR derived SWE (left) is compared against the average Noah-MP SWE. Rolling-mean curves for each Noah-MP simulation and ASO LiDAR data are shown. Lines with four different SWE versus elevation slopes (purple lines) are provided on each plot to better enable juxtaposition of datasets across different water years.

We would like to acknowledge funding from the DOE SCGSR program and DOE BER grant DOE:DE-SC0019222. Idaho National Labs also provided computing through. Boise State Research computing provided additional computing resources.

CHAPTER 5:

CONCLUSIONS

5.1 Synthesis and Discussion

In this dissertation, the regional climate models have been applied to two different mountain watersheds. We found that regional climate models could reproduce streamflows with reasonable skills using distributed hydrologic models (WRF-Hydro). Precipitation during the "cold-season" was well correlated with NRCS Snotel locations, and when coupled with hydrologic models could produce reasonable streamflows compared against observations (Chapter 1). However, seasonal biases in precipitation (evaluated against single-point observations) still led to biases in streamflow. Moreover, the skill of individual storm events is not always very high (such as 2018 in the East River for the Thompson microphysics scheme). There is much more work that can be done investigating the precipitation skill for some years versus others, and the types of systems (in terms of wind speed/direction, precipitable water content, upstream stability, etc.) that are well simulated versus poorly simulated. Still, a persistent and unavoidable challenge is that it is simply very difficult to validate or estimate precipitation in complex terrain (Lundquist *et al.*, 2019). This is a challenge for the Western US and even more of a challenge globally, in regions with even fewer gauges or less data sharing infrastructure. To meet these challenges, I developed

methods using streamflow and lidar remote-sensing to validate modelled precipitation in the East River watershed. Precipitation from streamflow inverse methods ("doing hydrology backwards", Kirchner (2009)) poses some challenges, in particular with respect to getting the evapotranspiration forcing correct, but the general approach shows promise for placing reasonable constraints on precipitation, particularly for data-poor regions. This chapter revealed a major challenge in regional scale hydrologic science writ-large. While ET is fundamentally constrained by energy limitations (Equation 1.15) and water availability between the ground surface and rooting zone, there are in fact no well articulated theories for estimating regional scale evaporation in complex terrain. The derivations of commonly applied theories such as Penman-Montieth assumes flat terrain with homogeneous surface roughness (Brutsaert, 2013), conditions which are not met in mountain terrain. Airborne snow lidar datasets (Painter *et al.*, 2016) offers immediate and clear benefits for studying mountain hydrology. With some care in the analysis, distributed gradients of snow-depth can be tied to orographic precipitation gradients and regions of enhancement (Chapter 3). If/when lidar flights become more frequent, these methods could be very useful for measuring distributions of snowfall in mountain terrain.

A few areas of concern with respect to model skill arose, namely biases in two meter air temperatures that were observed in both regions. This has been reported in other RCM studies (Rhoades *et al.*, 2018). This could be a function of the model configuration or challenges representing stable atmospheric surface layers over snow surfaces. At the same time, it is well known that Snotel temperatures can suffer from systematic biases (Oyler *et al.*, 2015). This somewhat under-recognized fact makes interpreting change signals at Snotel sites over time very difficult, in addition

to validating temperatures from regional climate models. Fortunately, there are more temperature observations in the Western United States than there are precipitation buckets/snow pillows, in part because they are cheaper/easier to maintain. Future work could evaluate WRF temperatures against observations from the MesoWest (Horel *et al.*, 2002) network. Datasets like PRISM provide air temperature maps, but often use lapse-rates based on seasonally varying climatologies (similar to precipitation; Daly *et al.* (2008)). Other snow modeling packages do something similar, such as MicroMet (Liston & Elder, 2006). According to the theory, anthropogenic climate change should increase warming rates of mountain environments disproportionately (Mountain Research Initiative Edw Working Group *et al.*, 2015), though observations confirming this phenomena have not been robustly observed. There is a clear research gap related to projecting impacts of warming and air-temperature lapse rates in mountain environments, and scrutinizing the abilities of regional climate models to capture elevation dependent warming.

5.2 Technological Challenges and Opportunities

Many challenges currently faced by regional scale hydrologic/atmospheric sciences were made apparent during this dissertation. A significant amount of time was spent on the software side of the analysis – writing codes for managing data and interfacing with HPC systems. These codes are shared on github (<https://github.com/bsurudisill/WRF-Run>). In many ways this type of workflow likely has not changed in several decades, even though computational power and storage have increased. Tools that break down barriers to using computationally intensive models will allow future researchers more time to focus on the purely scientific aspects as opposed to computational plumbing.

5.2.1 Data Analysis and Visualization

The ability to produce Earth system model data has far outstripped the ability to analyze that data, both from a computational perspective and a graphics library perspective. There is practically an endless amount of quantities/fields to analyze from regional climate model output, and unfortunately the tools for analysis are severely lacking. Over 300TB of data was produced for Chapter 2, far more than can be read into computer memory all at once. The python programming language was used to make all of the graphics and statistical analyses in this dissertation. The xarray package (Hoyer & Hamman, 2017) allows for out-of memory, distributed computations, and without this functionality many of the analyses would have taken much longer. While xarray is a great step forward for analysis, there are currently few if any sufficient tools for producing high resolution, 3d graphics visualizations natively in python/xarray software ecosystems. There are other tools (NCAR's "Vapor") that do this but it is challenging to work with large datasets. At the same time, there are many open source 3-D graphics libraries (such as "Blender") that could conceivably be ported to work with Earth System model outputs.

5.2.2 Software Infrastructure

Over time, the software behind atmospheric models such as WRF may change and improve. Compiling scientific model code (i.e, translating human code to machine code) is a persistent challenge, including for WRF. Containerized programs such as "Docker" that are used for reproducible environments are not always suitable for distributed memory applications or HPC systems. Scientific software managers such as "Spack" make some aspects of compilation easier (Gamblin *et al.*, 2015). Moreover, WRF, WRF-Hydro, and NoahMP are written in Fortran, the first version

of which appeared in 1957 (<https://fortran-lang.org/>). While Fortran is perfectly suited for the task for designing distributed memory, parallel applications, it can be challenging/verbose to write, and can make modifying the code difficult. "Modern" Fortran standards allow for more modularized development, which can help make code more readable and extensible. For example, at various points I have hoped to incorporate a snow redistribution model into NoahMP following Liston & Elder (2006) or Lehning *et al.* (2006), but was ultimately dissuaded by the time and effort it would take to develop in Fortran. At the same time, writing a blowing snow mode in python would be much quicker to write and debug, but the model run time would be slower. Languages that attempt to offer the best of both compiled and dynamic languages, such as Julia (Bezanson *et al.*, 2012), have existed for over a decade and are seemingly slow to catch on in the Earth/Hydrologic sciences, and legacy model codes do not appear to be going anywhere.

5.3 The Future of Regional Climate Modeling of Mountain Regions

The approaches used in this dissertation will likely continue in similar forms for decades to come. Physically based models of atmospheric flows, clouds, and radiation will be used to investigate hydrologic science questions. The resolutions of global models will undoubtedly increase. Already, global scale non-hydrostatic 1km atmospheric modeling studies have been performed using GPU-enabled software (Fuhrer *et al.*, 2017). It is likely that this will become the norm, so limited area dynamically downscaling methods at the current resolution will be less useful. Even higher model resolutions (sub kilometer) approaching the large eddy simulation ranges, could be-

come more tractable for longer time periods. This might allow some of the low-level terrain/wind effects that likely contribute to the observed distributions of snowpack, such as preferential deposition, (Lehning *et al.*, 2008), to be better simulated. Higher resolutions may lead to more realistic snow accumulation patterns (Chapter 3) that better match what is observed by airborne lidar. Immersed boundary methods, which replace the traditional coordinate system in atmospheric models, can also potentially improve the accuracy of wind field simulations (and potentially precipitation distributions) in complex terrain (Lundquist *et al.*, 2012)).

Machine-learning/Deep-learning methods have not been employed in this dissertation, but the unrealized utility for Earth science disciplines is vast. Machine learning is best thought of as a method of modeling highly non-linear systems when sufficient data is available. Machine learning models may offer solutions to many of the challenging weather and climate problems, such as modeling cloud microphysics or other parameterized processes (Rasp *et al.*, 2018). Machine learning can also help post-process data (such as the datasets produced in this dissertation) to identify features such as fronts (Lagerquist *et al.*, 2019). Machine learning methods can also be used as surrogate, or meta-models, to aid in the calibration or application of highly parameterized models (Razavi *et al.*, 2012). Hydrologic processes in mountain watersheds are complex, and highly connected between lithologic, geomorphic, biotic, and atmospheric processes. Emergent constraints are currently a popular and albeit somewhat controversial approach for deriving simple relationships between earth-system states/processes (such as snow-albedo feedbacks) from global climate model ensembles (Hall & Qu, 2006; Williamson & Sansom, 2019). Emergent constraints simplify complex systems in a way that can guide observations, and could potentially be applied to mountain hy-

drologic systems.

5.4 Closing Remarks

I started an education in the hydrological sciences because of the time I spent on the whitewater rivers of my native North Carolina. During my tenure at Boise State, I discovered snow. And with it the joys of skiing fresh, cold powder in the Boise, Payette, and Salmon river mountains. But as I've learned the ins and outs of the backcountry, I've done so carrying a certain sadness, knowing that the winter landscapes I have come to know will likely be fundamentally altered in the coming decades. Of course the consequences of climate change is of course much larger than their negative impacts on recreation, but envisioning the impacts on particular ski slope or river-flow has brought everything sharply into focus. The work presented in this dissertation addresses a small component of the greater efforts required to understand, adapt, and hopefully overcome anthropogenic climate change in the Western US and globe.

REFERENCES

- Anderson, E A. 1976. A point of energy and mass balance model of a snow cover. *NOAA Tech. Rep.*, **19**, 150.
- Arakawa, Akio, & Lamb, Vivian R. 1977. Computational design of the basic dynamical processes of the UCLA general circulation model. *General circulation models of the atmosphere*, **17**(Supplement C), 173–265.
- Arnault, Joel, Wagner, Sven, Rummeler, Thomas, Fersch, Benjamin, Bliedernicht, Jan, Andresen, Sabine, & Kunstmann, Harald. 2016. Role of Runoff–Infiltration Partitioning and Resolved Overland Flow on Land–Atmosphere Feedbacks: A Case Study with the WRF-Hydro Coupled Modeling System for West Africa. *J. Hydrometeorol.*, **17**(5), 1489–1516.
- Ashouri, Hamed, Hsu, Kuo-Lin, Sorooshian, Soroosh, Braithwaite, Dan K, Knapp, Kenneth R, Dewayne Cecil, L, Nelson, Brian R, & Prat, Olivier P. 2015. PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bull. Am. Meteorol. Soc.*, **96**(1), 69–83.
- Bao, J-W, Michelson, S A, & Grell, E D. 2019. Microphysical Process Comparison of Three Microphysics Parameterization Schemes in the WRF Model for an Idealized Squall-Line Case Study. *Mon. Weather Rev.*, **147**(9), 3093–3120.

- Benjamin, Stanley G, Weygandt, Stephen S, Brown, John M, Hu, Ming, Alexander, Curtis R, Smirnova, Tatiana G, Olson, Joseph B, James, Eric P, Dowell, David C, Grell, Georg A, Lin, Haidao, Peckham, Steven E, Smith, Tracy Lorraine, Moninger, William R, Kenyon, Jaymes S, & Manikin, Geoffrey S. 2016. A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Mon. Weather Rev.*, **144**(4), 1669–1694.
- Bergeron, Tor. 1935. On the physics of clouds and precipitation. *Proc. 5th Assembly UGGI, Lisbon, Portugal, 1935*, 156–180.
- Bezanson, Jeff, Karpinski, Stefan, Shah, Viral B, & Edelman, Alan. 2012. Julia: A Fast Dynamic Language for Technical Computing. Sept.
- Brutsaert, W. 2013. *Evaporation into the Atmosphere: Theory, History and Applications*. Springer Science & Business Media.
- Butts, Michael, Drews, Martin, Larsen, Morten A D, Lerer, Sara, Rasmussen, Søren H, Grooss, Jesper, Overgaard, Jesper, Refsgaard, Jens C, Christensen, Ole B, & Christensen, Jens H. 2014. Embedding complex hydrology in the regional climate system – Dynamic coupling across different modelling domains. *Adv. Water Resour.*, **74**(Dec.), 166–184.
- Chen, Di, Dai, Aiguo, & Hall, Alex. 2021. The convective-to-total precipitation ratio and the “drizzling” bias in climate models. *J. Geophys. Res.*, **126**(16).

- Clark, Martyn P, Slater, Andrew G, Rupp, David E, Woods, Ross A, Vrugt, Jasper A, Gupta, Hoshin V, Wagener, Thorsten, & Hay, Lauren E. 2008. Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models: DIFFERENCES BETWEEN HYDROLOGICAL MODELS. *Water Resour. Res.*, **44**(12).
- Clark, Martyn P, Hendrikx, Jordy, Slater, Andrew G, Kavetski, Dmitri, Anderson, Brian, Cullen, Nicolas J, Kerr, Tim, Örn Hreinsson, Einar, & Woods, Ross A. 2011. Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resour. Res.*, **47**(7).
- Currier, William Ryan, Thorson, Theodore, & Lundquist, Jessica D. 2017. Independent Evaluation of Frozen Precipitation from WRF and PRISM in the Olympic Mountains. *J. Hydrometeorol.*, **18**(10), 2681–2703.
- Dai, Aiguo. 2006. Precipitation Characteristics in Eighteen Coupled Climate Models. *J. Clim.*, **19**(18), 4605–4630.
- Daly, Christopher, Halbleib, Michael, Smith, Joseph I, Gibson, Wayne P, Doggett, Matthew K, Taylor, George H, Curtis, Jan, & Pasteris, Phillip P. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**(15), 2031–2064.
- Deems, Jeffrey S, Painter, Thomas H, & Finnegan, David C. 2013. Lidar measurement of snow depth: a review. *J. Glaciol.*, **59**(215), 467–479.
- Dingman, Lawrence S. 2015. *Physical Hydrology: Third Edition*. Waveland Press.

- Dudhia, Jimy. 2014. A history of mesoscale model development. *Asia Pac. J. Atmos. Sci.*, **50**(1), 121–131.
- Eagleson, P S. 1970. *Dynamic Hydrology*, MacGraw-Hill Book Co.
- Eidhammer, Trude, Grubišić, Vanda, Rasmussen, Roy, & Ikdea, Kyoko. 2018. Winter precipitation efficiency of mountain ranges in the Colorado Rockies under climate change. *J. Geophys. Res.*, **123**(5), 2573–2590.
- Eltahir, Elfatih A B, & Pal, Jeremy S. 1996. Relationship between surface conditions and subsequent rainfall in convective storms. *J. Geophys. Res.*, **101**(D21), 26237–26245.
- Feldman, Daniel, Aiken, Allison, Boos, William, Carroll, Rosemary, Chandrasekar, Venkatachalam, Collins, William, Collis, Scott, Deems, Jeff, DeMott, Paul, Fan, Jiwen, & Others. 2021. *Surface Atmosphere Integrated Field Laboratory (SAIL) Science Plan*. Tech. rept. Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States). Atmospheric
- Ferguson, Craig R, & Wood, Eric F. 2011. Observed Land–Atmosphere Coupling from Satellite Remote Sensing and Reanalysis. *J. Hydrometeorol.*, **12**(6), 1221–1254.
- Fersch, Benjamin, Senatore, Alfonso, Adler, Bianca, Arnault, Joël, Mauder, Matthias, Schneider, Katrin, Völksch, Ingo, & Kunstmann, Harald. 2020. High-resolution fully coupled atmospheric–hydrological modeling: a cross-compartment regional water and energy cycle evaluation. *Hydrol. Earth Syst. Sci.*, **24**(5), 2457–2481.
- Findell, Kirsten L. 2003. Atmospheric controls on soil moisture–boundary layer interactions: Three-dimensional wind effects. *J. Geophys. Res.*, **108**(D8).

- Findell, Kirsten L, & Eltahir, Elfatih A B. 2003. Atmospheric Controls on Soil Moisture–Boundary Layer Interactions. Part I: Framework Development. *J. Hydrometeorol.*, **4**(3), 552–569.
- Forrester, Mary M, & Maxwell, Reed M. 2020. Impact of Lateral Groundwater Flow and Subsurface Lower Boundary Conditions on Atmospheric Boundary Layer Development over Complex Terrain. *J. Hydrometeorol.*, **21**(6), 1133–1160.
- Fuhrer, Oliver, Chadha, Tarun, Hoeffler, Torsten, Kwasniewski, Grzegorz, Lapillonne, Xavier, Leutwyler, David, Lüthi, Daniel, Osuna, Carlos, Schär, Christoph, Schulthess, Thomas C, & Vogt, Hannes. 2017. Near-global climate simulation at 1 km resolution: establishing a performance baseline on 4888 GPUs with COSMO 5.0. *Geosci. Model Dev. Discuss.*, Oct., 1–27.
- Gamblin, Todd, LeGendre, Matthew, Collette, Michael R, Lee, Gregory L, Moody, Adam, de Supinski, Bronis R, & Futral, Scott. 2015. The Spack package manager: bringing order to HPC software chaos. *Pages 1–12 of: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. SC '15, no. Article 40. New York, NY, USA: Association for Computing Machinery.
- Givati, Amir, Gochis, David, Rummler, Thomas, & Kunstmann, Harald. 2016. Comparing One-Way and Two-Way Coupled Hydrometeorological Forecasting Systems for Flood Forecasting in the Mediterranean Region. *Hydrology*, **3**(2), 19.
- Gochis, D J, Yu, W, & Yates, D N. 2013. The WRF-Hydro model technical description and user’s guide, version 1.0, NCAR Technical Document. *National Center for Atmospheric Research, Boulder, CO, USA*.

- Gochis, David, Barlage, Michael, Dugger, Aubrey, FitzGerald, Katelyn, Karsten, Logan, McAllister, Molly, McCreight, James, Mills, Joe, Pan, Linlin, RafieeiNasab, Arezoo, Read, Laura, Sampson, Kevin, Yates, David, Yu, Wei, & Zhang, Yongxin. 2018. *WRF-Hydro Model Source Code Version 5*.
- Goergen, Klaus, & Kollet, Stefan. 2021. Boundary condition and oceanic impacts on the atmospheric water balance in limited area climate model ensembles. *Sci. Rep.*, **11**(1), 6228.
- Goodison, Barry Edward, Louie, Paul Y T, & Yang, Daqing. 1998. WMO solid precipitation measurement intercomparison. *World Meteorological Organization*.
- Gupta, Hoshin V, Wagener, Thorsten, & Liu, Yuqiong. 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol. Process.*, **22**(18), 3802–3813.
- Gutmann, Ethan D, Rasmussen, Roy M, Liu, Changhai, Ikeda, Kyoko, Gochis, David J, Clark, Martyn P, Dudhia, Jimy, & Thompson, Gregory. 2012. A Comparison of Statistical and Dynamical Downscaling of Winter Precipitation over Complex Terrain. *J. Clim.*, **25**(1), 262–281.
- Hall, Alex, & Qu, Xin. 2006. Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.*, **33**(3).
- Hamman, Joseph J, Nijssen, Bart, Bohn, Theodore J, Gergel, Diana R, & Mao, Yixin. 2018. The Variable Infiltration Capacity model version 5 (VIC-5): infrastructure improvements for new applications and reproducibility. *Geosci. Model Dev.*, **11**(8), 3481–3496.

- Hamon W. R. 1961. Estimating Potential Evapotranspiration. *Journal of the Hydraulics Division*, **87**(3), 107–120.
- Harpold, Adrian A, Kaplan, Michael L, Klos, P Zion, Link, Timothy, McNamara, James P, Rajagopal, Seshadri, Schumer, Rina, & Steele, Caitriana M. 2017. Rain or snow: hydrologic processes, observations, prediction, and research needs. *Hydrol. Earth Syst. Sci.*, **21**(1), 1–22.
- Havens, Scott, Marks, Danny, FitzGerald, Katelyn, Masarik, Matt, Flores, Alejandro N, Kormos, Patrick, & Hedrick, Andrew. 2019. Approximating Input Data to a Snowmelt Model Using Weather Research and Forecasting Model Outputs in Lieu of Meteorological Measurements. *J. Hydrometeorol.*, **20**(5), 847–862.
- He, Cenlin, Chen, Fei, Barlage, Michael, Liu, Changhai, Newman, Andrew, Tang, Wenfu, Ikeda, Kyoko, & Rasmussen, Roy. 2019. Can convection-permitting modeling provide decent precipitation for offline high-resolution snowpack simulations over mountains? *J. Geophys. Res.*, **124**(23), 12631–12654.
- Hedrick, Andrew R, Marks, Danny, Havens, Scott, Robertson, Mark, Johnson, Micah, Sandusky, Micah, Marshall, Hans-Peter, Kormos, Patrick R, Bormann, Kat J, & Painter, Thomas H. 2018. Direct insertion of NASA airborne snow observatory-derived snow depth time series into the iSnobal energy balance snow model. *Water Resour. Res.*, **54**(10), 8045–8063.
- Henn, Brian, Clark, Martyn P, Kavetski, Dmitri, & Lundquist, Jessica D. 2015. Estimating mountain basin-mean precipitation from streamflow using Bayesian inference. *Water Resour. Res.*, **51**(10), 8012–8033.

- Henn, Brian, Clark, Martyn P, Kavetski, Dmitri, McGurk, Bruce, Painter, Thomas H, & Lundquist, Jessica D. 2016. Combining snow, streamflow, and precipitation gauge observations to infer basin-mean precipitation. *Water Resour. Res.*, **52**(11), 8700–8723.
- Henn, Brian, Newman, Andrew J, Livneh, Ben, Daly, Christopher, & Lundquist, Jessica D. 2018. An Assessment of Differences in Gridded Precipitation Datasets in Complex Terrain. *J. Hydrol.*, **556**, 1205–1219.
- Holtzman, Nataniel M, Pavelsky, Tamlin M, Cohen, Jonathan S, Wrzesien, Melissa L, & Herman, Jonathan D. 2020. Tailoring WRF and Noah-MP to improve process representation of Sierra Nevada runoff: Diagnostic evaluation and applications. *J. Adv. Model. Earth Syst.*, **12**(3), e2019MS001832.
- Horel, J, Splitt, M, Dunn, L, Pechmann, J, White, B, Ciliberti, C, Lazarus, S, Slemmer, J, Zaff, D, & Burks, J. 2002. MESOWEST: COOPERATIVE MESONETS IN THE WESTERN UNITED STATES. *Bull. Am. Meteorol. Soc.*, **83**(2), 211–226.
- Houze, Jr, Robert A. 2012. Orographic effects on precipitating clouds. *Rev. Geophys.*, **50**(1).
- Hoyer, Stephan, & Hamman, Joe. 2017. xarray: N-D labeled Arrays and Datasets in Python. *J. Open Res. Softw.*, **5**(1), 10.

- Hubbard, Susan S, Williams, Kenneth Hurst, Agarwal, Deb, Banfield, Jillian, Beller, Harry, Bouskill, Nicholas, Brodie, Eoin, Carroll, Rosemary, Dafflon, Baptiste, Dwivedi, Dipankar, Falco, Nicola, Faybishenko, Boris, Maxwell, Reed, Nico, Peter, Steefel, Carl, Steltzer, Heidi, Tokunaga, Tetsu, Tran, Phuong A, Wainwright, Haruko, & Varadharajan, Charuleka. 2018. The East River, Colorado, watershed: A mountainous community testbed for improving predictive understanding of multiscale hydrological–biogeochemical dynamics. *Vadose Zone J.*, **17**(1), 1–25.
- Hughes, Mimi, Lundquist, Jessica D, & Henn, Brian. 2020. Dynamical downscaling improves upon gridded precipitation products in the Sierra Nevada, California. *Clim. Dyn.*, **55**(1), 111–129.
- Ikeda, Kyoko, Rasmussen, Roy, Liu, Changhai, Gochis, David, Yates, David, Chen, Fei, Tewari, Mukul, Barlage, Mike, Dudhia, Jimy, Miller, Kathy, Arsenault, Kristi, Grubišić, Vanda, Thompson, Greg, & Guttman, Ethan. 2010. Simulation of seasonal snowfall over Colorado. *Atmos. Res.*, **97**(4), 462–477.
- Immerzeel, W. W., Pellicciotti, F., & Shrestha, A. B. 2012. Glaciers as a proxy to quantify the spatial distribution of precipitation in the Hunza basin. *Mt. Res. Dev.*, **32**, 30–38.

- Immerzeel, W W, Lutz, A F, Andrade, M, Bahl, A, Biemans, H, Bolch, T, Hyde, S, Brumby, S, Davies, B J, Elmore, A C, Emmer, A, Feng, M, Fernández, A, Haritashya, U, Kargel, J S, Koppes, M, Kraaijenbrink, P D A, Kulkarni, A V, Mayewski, P A, Nepal, S, Pacheco, P, Painter, T H, Pellicciotti, F, Rajaram, H, Rupper, S, Sinisalo, A, Shrestha, A B, Viviroli, D, Wada, Y, Xiao, C, Yao, T, & Baillie, J E M. 2020. Importance and vulnerability of the world's water towers. *Nature*, **577**(7790), 364–369.
- Ivanov, Valeriy Y, Vivoni, Enrique R, Bras, Rafael L, & Entekhabi, Dara. 2004. Catchment hydrologic response with a fully distributed triangulated irregular network model. *Water Resour. Res.*, **40**(11).
- Jensen, Anders A, Harrington, Jerry Y, Morrison, Hugh, & Milbrandt, Jason A. 2017. Predicting Ice Shape Evolution in a Bulk Microphysics Model. *J. Atmos. Sci.*, **74**(6), 2081–2104.
- Jensen, Anders A, Harrington, Jerry Y, & Morrison, Hugh. 2018. Impacts of Ice Particle Shape and Density Evolution on the Distribution of Orographic Precipitation. *J. Atmos. Sci.*, **75**(9), 3095–3114.
- Jiang, Qingfang. 2003. Moist dynamics and orographic precipitation. *Tellus A*, **55**(4), 301–316.
- Jing, Xiaoqin, Geerts, Bart, Wang, Yonggang, & Liu, Changhai. 2017. Evaluating Seasonal Orographic Precipitation in the Interior Western United States Using Gauge Data, Gridded Precipitation Estimates, and a Regional Climate Simulation. *J. Hydrometeorol.*, **18**(9), 2541–2558.

- Jordan, Rachel E. 1991. A One-dimensional temperature model for a snow cover : technical documentation for SNTHERM.89. Oct.
- Kerandi, Noah, Arnault, Joel, Laux, Patrick, Wagner, Sven, Kitheka, Johnson, & Kunstmann, Harald. 2018. Joint atmospheric-terrestrial water balances for East Africa: a WRF-Hydro case study for the upper Tana River basin. *Theor. Appl. Climatol.*, **131**(3), 1337–1355.
- Kessler, Edwin. 1969. On the Distribution and Continuity of Water Substance in Atmospheric Circulations. *Pages 1–84 of: Kessler, Edwin (ed), On the Distribution and Continuity of Water Substance in Atmospheric Circulations.* Boston, MA: American Meteorological Society.
- Kirchner, James W. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.*, **42**(3).
- Kirchner, James W. 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward: CATCHMENTS AS SIMPLE DYNAMICAL SYSTEMS. *Water Resour. Res.*, **45**(2).
- Kirchner, P B, Bales, R C, Molotch, N P, Flanagan, J, & Guo, Q. 2014. LiDAR measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. *Hydrol. Earth Syst. Sci.*, **18**(10), 4261–4275.
- Kirshbaum, Daniel J, Adler, Bianca, Kalthoff, Norbert, Barthlott, Christian, & Serafin, Stefano. 2018. Moist Orographic Convection: Physical Mechanisms and Links to Surface-Exchange Processes. *Atmosphere*, **9**(3), 80.

- Knoben, Wouter J M, Freer, Jim E, & Woods, Ross A. 2019. Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrol. Earth Syst. Sci.*, **23**(10), 4323–4331.
- Köhler, Hilding. 1921. *Zur Kondensation des Wasserdampfes in der Atmosphäre*. I kommission hos Cammermeyers boghandel Kristiania.
- Koskela, J J, Croke, B W F, Koivusalo, H, Jakeman, A J, & Kokkonen, T. 2012. Bayesian inference of uncertainties in precipitation-streamflow modeling in a snow affected catchment. *Water Resour. Res.*, **48**(11).
- Lagerquist, Ryan, McGovern, Amy, & Gagne, II, David John. 2019. Deep Learning for Spatially Explicit Prediction of Synoptic-Scale Fronts. *Weather Forecast.*, **34**(4), 1137–1160.
- Lahmers, Timothy M, Gupta, Hoshin, Castro, Christopher L, Gochis, David J, Yates, David, Dugger, Aubrey, Goodrich, David, & Hazenberg, Pieter. 2019. Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments. *J. Hydrometeorol.*, **20**(4), 691–714.
- Lahmers, Timothy M, Castro, Christopher L, & Hazenberg, Pieter. 2020. Effects of Lateral Flow on the Convective Environment in a Coupled Hydrometeorological Modeling System in a Semiarid Environment. *J. Hydrometeorol.*, **21**(4), 615–642.
- Laprise, René. 1992. The Euler Equations of Motion with Hydrostatic Pressure as an Independent Variable. *Mon. Weather Rev.*, **120**(1), 197–207.

- Le Moine, Nicolas, Hendrickx, Frédéric, Gailhard, Joël, Garçon, Rémy, & Gottardi, Frédéric. 2015. Hydrologically Aided Interpolation of Daily Precipitation and Temperature Fields in a Mesoscale Alpine Catchment. *J. Hydrometeorol.*, **16**(6), 2595–2618.
- Lehning, M, Völksch, I, Gustafsson, D, & others. 2006. ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. *Hydrological.*
- Lehning, M, Löwe, H, Ryser, M, & Raderschall, N. 2008. Inhomogeneous precipitation distribution and snow transport in steep terrain. *Water Resour. Res.*, **44**(7).
- Lettenmaier, Dennis P, Alsdorf, Doug, Dozier, Jeff, Huffman, George J, Pan, Ming, & Wood, Eric F. 2015. Inroads of remote sensing into hydrologic science during the WRR era. *Water Resour. Res.*, **51**(9), 7309–7342.
- Leung, L Ruby, & Qian, Yun. 2009. Atmospheric rivers induced heavy precipitation and flooding in the western U.S. simulated by the WRF regional climate model. *Geophys. Res. Lett.*, **36**(3).
- Liston, Glen E, & Elder, Kelly. 2006. A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet). *J. Hydrometeorol.*, **7**(2), 217–234.
- Liu, Changhai, Ikeda, Kyoko, Thompson, Gregory, Rasmussen, Roy, & Dudhia, Jimy. 2011. High-Resolution Simulations of Wintertime Precipitation in the Colorado Headwaters Region: Sensitivity to Physics Parameterizations. *Mon. Weather Rev.*, **139**(11), 3533–3553.

- Liu, Changhai, Ikeda, Kyoko, Rasmussen, Roy, Barlage, Mike, Newman, Andrew J, Prein, Andreas F, Chen, Fei, Chen, Liang, Clark, Martyn, Dai, Aiguo, Dudhia, Jimmy, Eidhammer, Trude, Gochis, David, Gutmann, Ethan, Kurkute, Sopan, Li, Yanping, Thompson, Gregory, & Yates, David. 2017. Continental-scale convection-permitting modeling of the current and future climate of North America. *Clim. Dyn.*, **49**(1), 71–95.
- Livneh, Ben, Rosenberg, Eric A, Lin, Chiyu, Nijssen, Bart, Mishra, Vimal, Andreadis, Kostas M, Maurer, Edwin P, & Lettenmaier, Dennis P. 2013. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *J. Clim.*, **26**(23), 9384–9392.
- Lucas-Picher, Philippe, Argüeso, Daniel, Brisson, Erwan, Trambly, Yves, Berg, Peter, Lemonsu, Aude, Kotlarski, Sven, & Caillaud, Cécile. 2021. Convection - permitting modeling with regional climate models: Latest developments and next steps. *Wiley Interdiscip. Rev. Clim. Change*, **12**(6).
- Luce, Charles H, Tarboton, David G, & Cooley, Keith R. 1998. The influence of the spatial distribution of snow on basin-averaged snowmelt. *Hydrol. Process.*, **12**(10-11), 1671–1683.
- Lundquist, Jessica, Huggett, Brian, Roop, Heidi, & Low, Natalie. 2009. Use of spatially distributed stream stage recorders to augment rain gages by identifying locations of thunderstorm precipitation and distinguishing rain from snow. *Water Resour. Res.*, **45**(4).

- Lundquist, Jessica, Hughes, Mimi, Gutmann, Ethan, & Kapnick, Sarah. 2019. Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks. *Bull. Am. Meteorol. Soc.*, **100**(12), 2473–2490.
- Lundquist, Katherine A, Chow, Fotini Katopodes, & Lundquist, Julie K. 2012. An Immersed Boundary Method Enabling Large-Eddy Simulations of Flow over Complex Terrain in the WRF Model. *Mon. Weather Rev.*, **140**(12), 3936–3955.
- Lute, A C, Abatzoglou, J T, & Hegewisch, K C. 2015. Projected changes in snowfall extremes and interannual variability of snowfall in the western United States. *Water Resour. Res.*, **51**(2), 960–972.
- Maddox, Robert A, Zhang, Jian, Gourley, Jonathan J, & Howard, Kenneth W. 2002. Weather Radar Coverage over the Contiguous United States. *Weather Forecast.*, **17**(4), 927–934.
- Maina, Fadji Z, Siirila-Woodburn, Erica R, & Vahmani, Pouya. 2020. Sensitivity of meteorological-forcing resolution on hydrologic variables. *Hydrol. Earth Syst. Sci.*, **24**(7), 3451–3474.
- Maraun, D, Wetterhall, F, Ireson, A M, Chandler, R E, Kendon, E J, Widmann, M, Brienen, S, Rust, H W, Sauter, T, Themeßl, M, Venema, V K C, Chun, K P, Goodess, C M, Jones, R G, Onof, C, Vrac, M, & Thiele-Eich, I. 2010. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**(3).

- Markovich, Katherine H, Manning, Andrew H, Condon, Laura E, & McIntosh, Jennifer C. 2019. Mountain-block recharge: A review of current understanding. *Water Resour. Res.*, **55**(11), 8278–8304.
- Marshall, J S. 1948. The distribution of raindrops with size. *J. Meteorol.*, **5**, 165–166.
- Maxwell, Reed M, Kollet, Stefan J, Smith, Steven G, Woodward, Carol S, Falgout, Robert D, Ferguson, Ian M, Baldwin, Chuck, Bosl, William J, Hornung, Richard, & Ashby, Steven. 2009. ParFlow user’s manual. *International Ground Water Modeling Center Report GWMI*, **1**(2009), 129.
- Maxwell, Reed M, Lundquist, Julie K, Mirocha, Jeffrey D, Smith, Steven G, Woodward, Carol S, & Tompson, Andrew F B. 2011. Development of a Coupled Groundwater–Atmosphere Model. *Mon. Weather Rev.*, **139**(1), 96–116.
- Miller, Douglas A, & White, Richard A. 1998. A Conterminous United States Multilayer Soil Characteristics Dataset for Regional Climate and Hydrology Modeling. *Earth Interact.*, **2**(2), 1–26.
- Milly, P C D, & Dunne, K A. 2020. Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation. *Science*, **367**(6483), 1252–1255.
- Minder, Justin R, & Kingsmill, David E. 2013. Mesoscale Variations of the Atmospheric Snow Line over the Northern Sierra Nevada: Multiyear Statistics, Case Study, and Mechanisms. *J. Atmos. Sci.*, **70**(3), 916–938.

- Mizukami, Naoki, Clark, Martyn P, Slater, Andrew G, Brekke, Levi D, Elsner, Marketa M, Arnold, Jeffrey R, & Gangopadhyay, Subhrendu. 2014. Hydrologic Implications of Different Large-Scale Meteorological Model Forcing Datasets in Mountainous Regions. *J. Hydrometeorol.*, **15**(1), 474–488.
- Morales, Annareli, Posselt, Derek J, Morrison, Hugh, & He, Fei. 2019. Assessing the Influence of Microphysical and Environmental Parameter Perturbations on Orographic Precipitation. *J. Atmos. Sci.*, **76**(5), 1373–1395.
- Morrison, H, Curry, J A, & Khvorostyanov, V I. 2005. A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description. *J. Atmos. Sci.*, **62**(6), 1665–1677.
- Morrison, Hugh, van Lier-Walqui, Marcus, Fridlind, Ann M, Grabowski, Wojciech W, Harrington, Jerry Y, Hoose, Corinna, Korolev, Alexei, Kumjian, Matthew R, Milbrandt, Jason A, Pawlowska, Hanna, Posselt, Derek J, Prat, Olivier P, Reimel, Karly J, Shima, Shin-Ichiro, van Dierenhoven, Bastiaan, & Xue, Lulin. 2020. Confronting the Challenge of Modeling Cloud and Precipitation Microphysics. *J Adv Model Earth Syst*, **12**(8), e2019MS001689.
- Mote, Philip W, Hamlet, Alan F, Clark, Martyn P, & Lettenmaier, Dennis P. 2005. Declining Mountain Snowpack In Western North America. *Bull. Am. Meteorol. Soc.*, **86**(1), 39–50.
- Mott, Rebecca, Vionnet, Vincent, & Grünewald, Thomas. 2018. The Seasonal Snow Cover Dynamics: Review on Wind-Driven Coupling Processes. *Front Earth Sci. Chin.*, **6**.

- Mountain Research Initiative Edw Working Group, Pepin, N, Bradley, R S, Diaz, H F, Baraer, M, Caceres, E B, Forsythe, N, Fowler, H, Greenwood, G, Hashmi, M Z, Liu, X D, Miller, J R, Ning, L, Ohmura, A, Palazzi, E, Rangwala, I, Schöner, W, Severskiy, I, Shahgedanova, M, Wang, M B, Williamson, S N, & Yang, D Q. 2015. Elevation-dependent warming in mountain regions of the world. *Nat. Clim. Chang.*, **5**(May), 424–430.
- Newman, Andrew J, Clark, Martyn P, Craig, Jason, Nijssen, Bart, Wood, Andrew, Gutmann, Ethan, Mizukami, Naoki, Brekke, Levi, & Arnold, Jeff R. 2015. Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States. *J. Hydrometeorol.*, **16**(6), 2481–2500.
- Niu, Guo-Yue, Yang, Zong-Liang, Mitchell, Kenneth E, Chen, Fei, Ek, Michael B, Barlage, Michael, Kumar, Anil, Manning, Kevin, Niyogi, Dev, Rosero, Enrique, Tewari, Mukul, & Xia, Youlong. 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**(D12).
- Oyler, Jared W, Dobrowski, Solomon Z, Ballantyne, Ashley P, Klene, Anna E, & Running, Steven W. 2015. Artificial amplification of warming trends across the mountains of the western United States. *Geophys. Res. Lett.*, **42**(1), 153–161.

- Painter, Thomas H, Berisford, Daniel F, Boardman, Joseph W, Bormann, Kathryn J, Deems, Jeffrey S, Gehrke, Frank, Hedrick, Andrew, Joyce, Michael, Laidlaw, Ross, Marks, Danny, Mattmann, Chris, McGurk, Bruce, Ramirez, Paul, Richardson, Megan, Skiles, S Mckenzie, Seidel, Felix C, & Winstral, Adam. 2016. The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sens. Environ.*, **184**(Oct.), 139–152.
- Pan, M, & Wood, E F. 2013. Inverse streamflow routing. *Hydrol. Earth Syst. Sci.*, **17**(11), 4577–4588.
- Pavelsky, Tamlin M, Sobolowski, Stefan, Kapnick, Sarah B, & Barnes, Jason B. 2012. Changes in orographic precipitation patterns caused by a shift from snow to rain. *Geophys. Res. Lett.*, **39**(18).
- Petty, Grant William. 2008. *A first course in atmospheric thermodynamics*. Sundog Publishing.
- Pflug, J M, Hughes, M, & Lundquist, J D. 2021. Downscaling snow deposition using historic snow depth patterns: Diagnosing limitations from snowfall biases, winter snow losses, and interannual snow pattern repeatability. *Water Resour. Res.*, **57**(8).
- Prein, Andreas F, Langhans, Wolfgang, Fosser, Giorgia, Ferrone, Andrew, Ban, Nikolina, Goergen, Klaus, Keller, Michael, Tölle, Merja, Gutjahr, Oliver, Feser, Frauke, Brisson, Erwan, Kollet, Stefan, Schmidli, Juerg, van Lipzig, Nicole P M, & Leung, Ruby. 2015. A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**(2), 323–361.

- Pu, Zhaoxia, & Kalnay, Eugenia. 2019. Numerical weather prediction basics: Models, numerical methods, and data assimilation. *Pages 67–97 of: Handbook of hydrometeorological ensemble forecasting*. Springer Berlin Heidelberg.
- Raleigh, M S, Lundquist, J D, & Clark, M P. 2015. Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrol. Earth Syst. Sci.*, **19**(7), 3153–3179.
- Rangwala, Imtiaz, & Miller, James R. 2012. Climate change in mountains: a review of elevation-dependent warming and its possible causes. *Clim. Change*, **114**(3), 527–547.
- Rasmussen, Roy, Liu, Changhai, Ikeda, Kyoko, Gochis, David, Yates, David, Chen, Fei, Tewari, Mukul, Barlage, Michael, Dudhia, Jimy, Yu, Wei, Miller, Kathleen, Arsenault, Kristi, Grubišić, Vanda, Thompson, Greg, & Gutmann, Ethan. 2011. High-Resolution Coupled Climate Runoff Simulations of Seasonal Snowfall over Colorado: A Process Study of Current and Warmer Climate. *J. Clim.*, **24**(12), 3015–3048.
- Rasmussen, Roy M, Bernstein, Ben C, Murakami, Masataka, Stossmeister, Greg, Reisner, Jon, & Stankov, Boba. 1995. The 1990 Valentine’s Day Arctic Outbreak. Part I: Mesoscale and Microscale Structure and Evolution of a Colorado Front Range Shallow Upslope Cloud. *J. Appl. Meteorol. Climatol.*, **34**(7), 1481–1511.
- Rasp, Stephan, Pritchard, Michael S, & Gentine, Pierre. 2018. Deep learning to represent subgrid processes in climate models. *Proc. Natl. Acad. Sci. U. S. A.*, **115**(39), 9684–9689.

- Razavi, Saman, Tolson, Bryan A, & Burn, Donald H. 2012. Review of surrogate modeling in water resources. *Water Resour. Res.*, **48**(7).
- Reynolds, Dylan S, Pflug, Justin M, & Lundquist, Jessica D. 2021. Evaluating wind fields for use in basin-scale distributed snow models. *Water Resour. Res.*, **57**(2).
- Rhoades, Alan M, Ullrich, Paul A, Zarzycki, Colin M, Johansen, Hans, Margulis, Steven A, Morrison, Hugh, Xu, Zexuan, & Collins, William D. 2018. Sensitivity of mountain hydroclimate simulations in variable-resolution CESM to microphysics and horizontal resolution. *J. Adv. Model. Earth Syst.*, **10**(6), 1357–1380.
- Roe, Gerard H. 2005. Orographic precipitation. *Annu. Rev. Earth Planet. Sci.*, **33**(1), 645–671.
- Rudisill, Will. 2021 (Feb.). *bsu-wrudisill/WRF-HYDRO_CALIB: v0.1-beta.0*.
- Rudisill, William, Flores, Alejandro, & McNamara, James. 2021. The Impact of Initial Snow Conditions on the Numerical Weather Simulation of a Northern Rockies Atmospheric River. *J. Hydrometeorol.*, **22**(1), 155–167.
- Rutz, Jonathan J, James Steenburgh, W, & Martin Ralph, F. 2015. The Inland Penetration of Atmospheric Rivers over Western North America: A Lagrangian Analysis. *Mon. Weather Rev.*, **143**(5), 1924–1944.
- Ryken, Anna, Gochis, David, & Maxwell, Reed M. 2021. Unraveling groundwater contributions to evapotranspiration and constraining water fluxes in a high-elevation catchment. *Hydrological Processes*.

Saha, Suranjana, Moorthi, Shrinivas, Pan, Hua-Lu, Wu, Xingren, Wang, Jiande, Nadiga, Sudhir, Tripp, Patrick, Kistler, Robert, Woollen, John, Behringer, David, Liu, Haixia, Stokes, Diane, Grumbine, Robert, Gayno, George, Wang, Jun, Hou, Yu-Tai, Chuang, Hui-Ya, Juang, Hann-Ming H, Sela, Joe, Iredell, Mark, Treadon, Russ, Kleist, Daryl, Van Delst, Paul, Keyser, Dennis, Derber, John, Ek, Michael, Meng, Jesse, Wei, Helin, Yang, Rongqian, Lord, Stephen, van den Dool, Huug, Kumar, Arun, Wang, Wanqiu, Long, Craig, Chelliah, Muthuvel, Xue, Yan, Huang, Boyin, Schemm, Jae-Kyung, Ebisuzaki, Wesley, Lin, Roger, Xie, Pingping, Chen, Mingyue, Zhou, Shuntai, Higgins, Wayne, Zou, Cheng-Zhi, Liu, Quanhua, Chen, Yong, Han, Yong, Cucurull, Lidia, Reynolds, Richard W, Rutledge, Glenn, & Goldberg, Mitch. 2010. The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.*, **91**(8), 1015–1058.

Salvatier, John, Wiecki, Thomas V, & Fonnesbeck, Christopher. 2016. Probabilistic programming in Python using PyMC3. *PeerJ Comput. Sci.*, **2**(Apr.), e55.

Santanello, Joseph A, Dirmeyer, Paul A, Ferguson, Craig R, Findell, Kirsten L, Tawfik, Ahmed B, Berg, Alexis, Ek, Michael, Gentine, Pierre, Guillod, Benoit P, van Heerwaarden, Chiel, Roundy, Joshua, & Wulfmeyer, Volker. 2018. Land–Atmosphere Interactions: The LoCo Perspective. *Bull. Am. Meteorol. Soc.*, **99**(6), 1253–1272.

Senatore, Alfonso, Mendicino, Giuseppe, Gochis, David J, Yu, Wei, Yates, David N, & Kunstmann, Harald. 2015. Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales:. *J. Adv. Model. Earth Syst.*, **7**(4), 1693–1715.

- Seneviratne, Sonia I, Corti, Thierry, Davin, Edouard L, Hirschi, Martin, Jaeger, Eric B, Lehner, Irene, Orlowsky, Boris, & Teuling, Adriaan J. 2010. Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**(3), 125–161.
- Serreze, Mark C, Clark, Martyn P, Armstrong, Richard L, McGinnis, David A, & Pulwarty, Roger S. 1999. Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resour. Res.*, **35**(7), 2145–2160.
- Shrestha, P, Sulis, M, Masbou, M, Kollet, S, & Simmer, C. 2014. A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM, and ParFlow. *Mon. Weather Rev.*, **142**(9), 3466–3483.
- Siirila-Woodburn, Erica R, Rhoades, Alan M, Hatchett, Benjamin J, Huning, Laurie S, Szinai, Julia, Tague, Christina, Nico, Peter S, Feldman, Daniel R, Jones, Andrew D, Collins, William D, & Kaatz, Laurina. 2021. A low-to-no snow future and its impacts on water resources in the western United States. *Nature Reviews Earth & Environment*, **2**(11), 800–819.
- Skamarock, William C, Klemp, Joseph B, Dudhia, Jimy, Gill, David O, Barker, Dale M, Duda, Michael G, Huang, Xiang-Yu, Wang, Wei, & Powers, Jordan G. 2008. G.: A description of the Advanced Research WRF version 3. *In: NCAR Tech. Note NCAR/TN-475+STR*.
- Skamarock, William C, Klemp, Joseph B, Dudhia, Jimy, Gill, David O, Liu, Zhiqun, Berner, Judith, Wang, Wei, Powers, Jordan G, Duda, Michael G, Barker, Dale M, & Others. 2019. A description of the advanced research WRF model version 4. *National Center for Atmospheric Research: Boulder, CO, USA*, 145.

- Slater, A G, Schlosser, C A, Desborough, C E, Pitman, A J, Henderson-Sellers, A, Robock, A, Ya Vinnikov, K, Entin, J, Mitchell, K, Chen, F, Boone, A, Etchevers, P, Habets, F, Noilhan, J, Braden, H, Cox, P M, de Rosnay, P, Dickinson, R E, Yang, Z-L, Dai, Y-J, Zeng, Q, Duan, Q, Koren, V, Schaake, S, Gedney, N, Gusev, Ye M, Nasonova, O N, Kim, J, Kowalczyk, E A, Shmakin, A B, Smirnova, T G, Verseghy, D, Wetzell, P, & Xue, Y. 2001. The Representation of Snow in Land Surface Schemes: Results from PILPS 2(d). *J. Hydrometeorol.*, **2**(1), 7–25.
- Smith, Ronald B, & Evans, Jason P. 2007. Orographic Precipitation and Water Vapor Fractionation over the Southern Andes. *J. Hydrometeorol.*, **8**(1), 3–19.
- Stensrud, David J. 2009. *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*. Cambridge University Press.
- Stoelinga, Mark T, Stewart, Ronald E, Thompson, Gregory, & Thériault, Julie M. 2013. Microphysical Processes Within Winter Orographic Cloud and Precipitation Systems. *Pages 345–408 of: Chow, Fotini K, De Wekker, Stephan F J, & Snyder, Bradley J (eds), Mountain Weather Research and Forecasting: Recent Progress and Current Challenges*. Dordrecht: Springer Netherlands.
- Sturm, M, Goldstein, M A, & Parr, C. 2017. Water and life from snow: A trillion dollar science question. *Water Resour. Res.*
- Sturm, Matthew, Taras, Brian, Liston, Glen E, Derksen, Chris, Jonas, Tobias, & Lea, Jon. 2010. Estimating Snow Water Equivalent Using Snow Depth Data and Climate Classes. *J. Hydrometeorol.*, **11**(6), 1380–1394.

- Sun, Ning, Yan, Hongxiang, Wigmosta, Mark S, Leung, L Ruby, Skaggs, Richard, & Hou, Zhangshuan. 2019. Regional snow parameters estimation for large-domain hydrological applications in the western United States. *J. Geophys. Res.*, **124**(10), 5296–5313.
- Sun, Shufen, Jin, Jiming, & Xue, Yongkang. 1999. A simple snow-atmosphere-soil transfer model. *J. Geophys. Res.*, **104**(D16), 19587–19597.
- Swales, Dustin, Alexander, Mike, & Hughes, Mimi. 2016. Examining moisture pathways and extreme precipitation in the U.S. Intermountain West using self-organizing maps. *Geophys. Res. Lett.*, **43**(4), 1727–1735.
- Thompson, Gregory, Field, Paul R, Rasmussen, Roy M, & Hall, William D. 2008a. Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Weather Rev.*, **136**(12), 5095–5115.
- Thompson, Gregory, Field, Paul R, Rasmussen, Roy M, & Hall, William D. 2008b. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, **136**(12), 5095–5115.
- Thornton, P E, Thornton, M M, Mayer, B W, Wei, Y, Devarakonda, R, Vose, R S, & Cook, R B. 2016. Daymet: daily surface weather data on a 1-km grid for North America, version 3. ORNL DAAC, Oak Ridge, Tennessee, USA. In: *USDA-NASS, 2019. 2017 Census of Agriculture, Summary and State Data, Geographic Area Series, Part 51, AC-17-A-51*. daac.ornl.gov.

- Thyer, Mark, Renard, Benjamin, Kavetski, Dmitri, Kuczera, George, Franks, Stewart William, & Srikanthan, Sri. 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resour. Res.*, **45**(12).
- Tolson, B A, & Shoemaker, C A. 2007. Dynamically Dimensioned Search Algorithm for Computationally Efficient Watershed Model Calibration. *Water Resources Research*, **43**(1).
- Trenberth, Kevin E, Dai, Aiguo, Rasmussen, Roy M, & Parsons, David B. 2003. The Changing Character of Precipitation. *Bull. Am. Meteorol. Soc.*, **84**(9), 1205–1218.
- Tuttle, Samuel, & Salvucci, Guido. 2016. Empirical evidence of contrasting soil moisture-precipitation feedbacks across the United States. *Science*, **352**(6287), 825–828.
- Valery, Audrey, Andréassian, Vazken, & Perrin, Charles. 2009. *Inverting the hydrological cycle: when streamflow measurements help assess altitudinal precipitation gradients in mountain areas*. <https://iahs.info/uploads/dms/14842.41-281-286-333-28-3967--VALERY-Corr.pdf>. Accessed: 2021-4-20.
- Vögeli, Christian, Lehning, Michael, Wever, Nander, & Bavay, Mathias. 2016. Scaling Precipitation Input to Spatially Distributed Hydrological Models by Measured Snow Distribution. *Front Earth Sci. Chin.*, **4**.
- Weisman, Morris L, Skamarock, William C, & Klemp, Joseph B. 1997. The Resolution Dependence of Explicitly Modeled Convective Systems. *Mon. Weather Rev.*, **125**(4), 527–548.

- Wigmosta, Mark S, & Lettenmaier, Dennis P. 1999. A comparison of simplified methods for routing topographically driven subsurface flow. *Water Resour. Res.*, **35**(1), 255–264.
- Williamson, Daniel B, & Sansom, Philip G. 2019. How Are Emergent Constraints Quantifying Uncertainty and What Do They Leave Behind? *Bull. Am. Meteorol. Soc.*, **100**(12), 2571–2588.
- Wrzesien, Melissa L, Pavelsky, Tamlin M, Durand, Michael T, Dozier, Jeff, & Lundquist, Jessica D. 2019. Characterizing biases in mountain snow accumulation from global data sets. *Water Resour. Res.*, **55**(11), 9873–9891.
- Yang, Daqing, Goodison, Barry E, Metcalfe, John R, Golubev, Valentin S, Bates, Roy, Pangburn, Timothy, & Hanson, Clayton L. 1998. Accuracy of NWS 8“ Standard Nonrecording Precipitation Gauge: Results and Application of WMO Intercomparison. *J. Atmos. Ocean. Technol.*, **15**(1), 54–68.
- Yau, M K, & Rogers, R R. 1996. *A Short Course in Cloud Physics*. Elsevier.