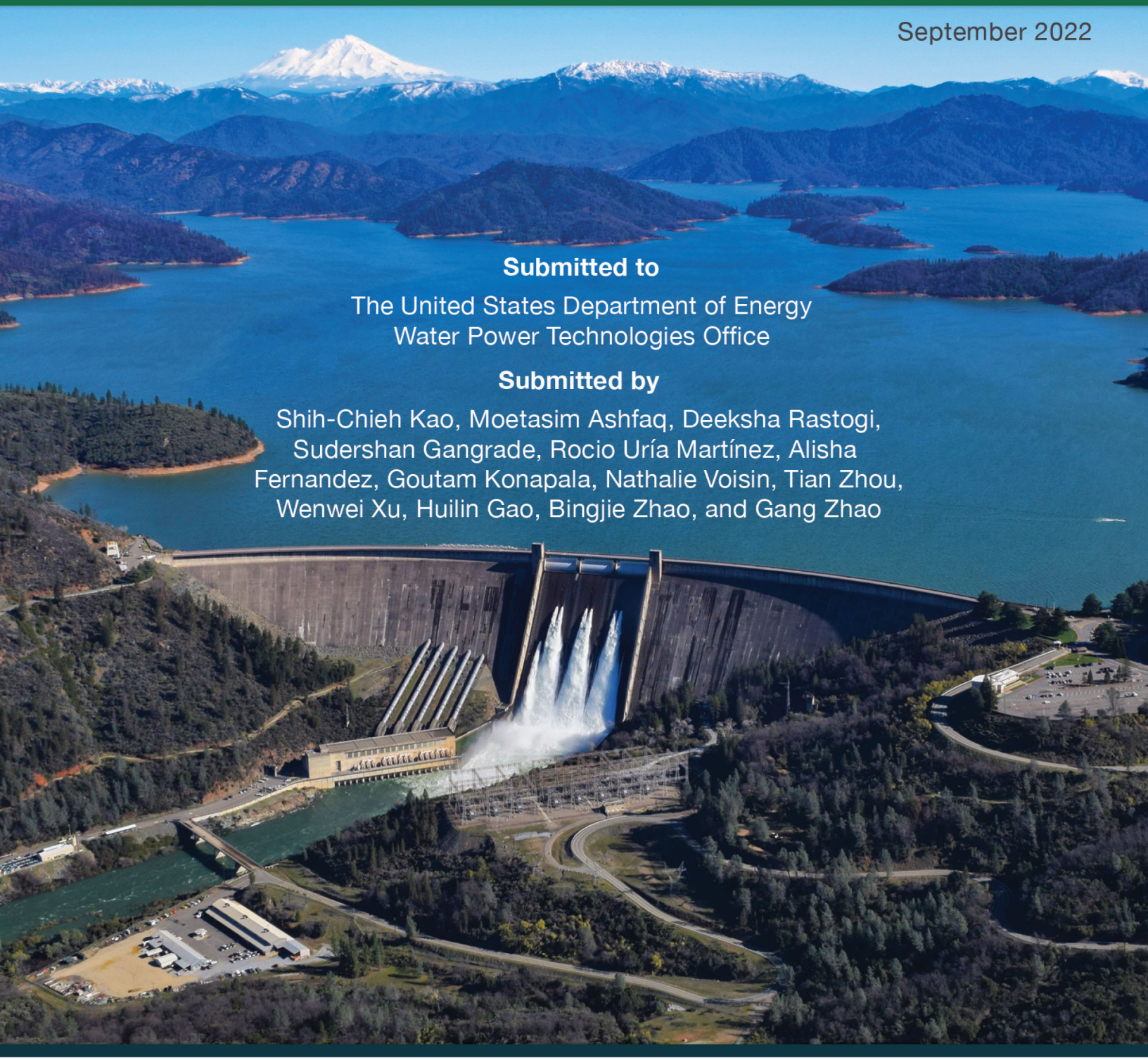
Oak Ridge National Laboratory ORNL/TM-2021/2278

The Third Assessment of the Effects of Climate Change on Federal Hydropower

September 2022



Submitted to

The United States Department of Energy
Water Power Technologies Office

Submitted by

Shih-Chieh Kao, Moetasim Ashfaq, Deeksha Rastogi, Sudershan Gangrade, Rocio Uría Martínez, Alisha Fernandez, Goutam Konapala, Nathalie Voisin, Tian Zhou, Wenwei Xu, Huilin Gao, Bingjie Zhao, and Gang Zhao

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE US DEPARTMENT OF ENERGY

Front Cover Image

Shasta Dam and Lake, California

(This page intentionally left blank)

ORNL/TM-2021/2278

Environmental Sciences Division

THE THIRD ASSESSMENT OF THE EFFECTS OF CLIMATE CHANGE ON FEDERAL HYDROPOWER

Oak Ridge National Laboratory
Shih-Chieh Kao, Moetasim Ashfaq, Deeksha Rastogi, Sudershan Gangrade, Rocio Uría Martínez, Alisha Fernandez, and Goutam Konapala

Pacific Northwest National Laboratory
Nathalie Voisin, Tian Zhou, and Wenwei Xu

Texas A&M University
Huilin Gao, Bingjie Zhao, and Gang Zhao

September 2022

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

(This page intentionally left blank)

ACKNOWLEDGMENTS

This report was sponsored by the US Department of Energy's (DOE's) Water Power Technologies Office. The authors wish to acknowledge the following Water Power Technologies Office staff who provided support and guidance for this assessment: Simon Gore, Charles Scaife, Corey Vezina, Erfaneh Sharifi, Timothy Welch, and Hoyt Battey. Several technical staff from the federal Power Marketing Administrations played a critical role in the development of this assessment report, providing essential data, reviews, and advice. The Power Marketing Administration contributors to this report are as follows:

- **Bonneville Power Administration**: Erik Pytlak and Alisa Kaseweter
- **Western Area Power Administration**: Chrystal Dean, Clayton Palmer, Xavier H. Gonzalez, and Robert Delizo
- **Southwestern Power Administration**: Tyler Gipson, Mike Denny, Fritha Ohlson, and Barbara Smith
- **Southeastern Power Administration**: Dixie Cordell, and Douglas Spencer

Without the help of Power Marketing Administration staff, this report would not have been possible. This report also benefited significantly from consultation with other federal agencies, which the authors greatly appreciate. The noted individuals are as follows:

- **US Bureau of Reclamation**: Avra Morgan, Marketa McGuire, Subhrendu Gangopadhyay, Kenneth Nowak, Katie Schultz, Clark Bishop, Levi Brekke, Kenneth Richard, Sophie Wilderotter, Jen Johnson, and Kevin Thielen
- **US Army Corps of Engineers**: Jeff Arnold, and Daniel Rabon
- **US Geological Survey**: Melissa Harris, Greg Pederson, Gregory McCabe, and Jen Johnson
- **US Forest Service**: Travis Warziniack
- **National Oceanic and Atmospheric Administration**: Robert Webb
- **Tennessee Valley Authority**: Curt Jawdy, Patrick Massey, and Nathan Barber

The authors also wish to thank the following external collaborators and technical reviewers for their support and constructive comments that helped improve the quality of this study:

- Bita Analui, University of California, Irvine
- Michael Anderson, California Department of Water Resources
- Maya Buchanan, Oregon Department of Energy
- Ethan Gutmann, National Center for Atmospheric Research
- Ariel Miara, National Renewable Energy Laboratory
- Andrew Newman, National Center for Atmospheric Research
- Bart Nijssen, University of Washington
- Quentin Ploussard, Argonne National Laboratory
- David Rupp, Oregon State University
- Soroosh Sorooshian, University of California, Irvine
- Andrew Sumner, DOE
- Mai Kim Tran, DOE
- Edith Zagona, University of Colorado, Boulder

Support from other Oak Ridge National Laboratory (ORNL) staff, including Brennan Smith, Shelaine Curd, Olivia Shafer, Sam Crawford, and Scott Painter, is also acknowledged. The research used resources of the Oak Ridge Leadership Computing Facility at ORNL, which is a DOE Office of Science User

Facility. The ORNL authors are employees of UT-Battelle, LLC, under contract DE-AC05-00OR22725 with DOE.

The listing of reviewers here does not imply their agreement with all findings of the report. Any remaining errors in this report are the sole responsibility of the authors.

CONTENTS

(This page intentionally left blank)

LIST OF FIGURES

(This page intentionally left blank)

LIST OF TABLES

(This page intentionally left blank)

EXECUTIVE SUMMARY

Understanding the future changes in projected water supplies is a vital objective for federal hydropower facilities tasked with providing low-cost, reliable electricity across a large regional footprint that encompasses a growing customer base, alternative market structures for marketing the electricity, and a more diverse generation asset mix than was historically present when a majority of federal hydropower facilities were built. This study, *The Third Assessment of the Effects of Climate Change on Federal Hydropower*, directed by Section 9505 of the SECURE Water Act of 2009 (SWA), is the third quinquennial report on evaluating the effects of climate change on hydroelectric energy generated from 132 US federal hydropower plants marketed by four US Department of Energy (DOE) Power Marketing Administrations (PMAs). The technical assessment is conducted by DOE's Oak Ridge National Laboratory, Pacific Northwest National Laboratory, and Texas A&M University under the guidance of DOE's Water Power Technologies Office. This study is the result of extensive consultation with the four federal PMAs (Bonneville Power Administration [BPA], Western Area Power Administration [WAPA], Southwestern Power Administration [SWPA], and Southeastern Power Administration [SEPA]), as well as other agencies, including federal hydropower owners/operators (the US Army Corps of Engineers, US Bureau of Reclamation [Reclamation]), US Geological Survey, and National Oceanic and Atmospheric Administration). The main findings of this assessment, along with the PMA administrators' recommendations, will be included in a subsequent DOE report to Congress. The assessment method and the technical findings are described in this report.

US Federal Hydropower

Hydropower is a key contributor to the US renewable energy portfolio because of its established development history and the diverse benefits it provides to the electric power systems. Ensuring the sustainable operation of existing hydropower facilities is of great importance to the US renewable energy portfolio and the reliability of the electricity grid. As of 2019, there were 2,270 conventional (i.e., once-through) hydropower plants in the United States with a total of 80.25 GW of generating capacity producing 6.6% of all electricity and 38% of electricity from renewables (Uría Martínez et al., 2021). Additionally, there were 43 pumped storage hydropower plants with a total of 21.9 GW of generating capacity providing 94% of utility-scale installed storage capacity.

Among these hydropower plants, around 40% of the generating capacity (42.5 GW) was provided by federal projects and built and/or operated by one of four agencies: the US Army Corps of Engineers, Reclamation, Tennessee Valley Authority, and International Boundary and Water Commission. The nonfederal assets comprising the other 60% are mostly regulated by the Federal Energy Regulatory Commission under the authority of the Federal Power Act and owned and/or operated by investor-owned utilities, publicly owned utilities, state agencies, and nonutilities (Uría Martínez et al., 2015). Although there are fewer federal hydropower plants than nonfederal plants, the federal plants have, on average, more than 10 times the generating capacity per plant.

For federal hydropower, the PMAs market the hydroelectric energy generated from 132 federally owned/operated hydropower plants to repay the government's investment in these projects. By statute, the PMAs prioritize access to their federal hydropower resources for their preference customers—mostly municipalities, political subdivisions, and cooperatives—over for-profit entities. The preference clauses, introduced in several pieces of legislation in the early 1900s, were designed to ensure that the operation of the federal hydropower assets benefited the public, contributed to the economic development of rural areas, and avoided a monopoly in the nascent electric industry (GAO, 2021). Each of the four PMAs is a distinct, self-contained entity within DOE, similar to a wholly owned subsidiary of a corporation.

Most of these federal hydropower plants are located at multipurpose reservoirs that also provide nonpower services, such as flood control; navigation; water supply for municipalities, industries, agriculture, and recreation; and protection of environmental resources, including water quality, fish, and wildlife. Since many nonpower services have higher priority than hydropower, the generating federal hydropower is under a variety of competing constraints and may not be as flexible, despite the large water storage capacity in federal reservoirs.

Methodology

The third 9505 SECURE Water Act Assessment addresses the long-term effects of climate change on federally generated hydropower at the regional scale for each of the four PMAs by identifying and analyzing projected future annual and seasonal trends in key hydrometeorological variables (Figure ES.1). This assessment builds upon the framework established in the first two assessments and leverages the latest global climate projections from the Coupled Model Intercomparison Project (CMIP) Phase 6 (CMIP6).

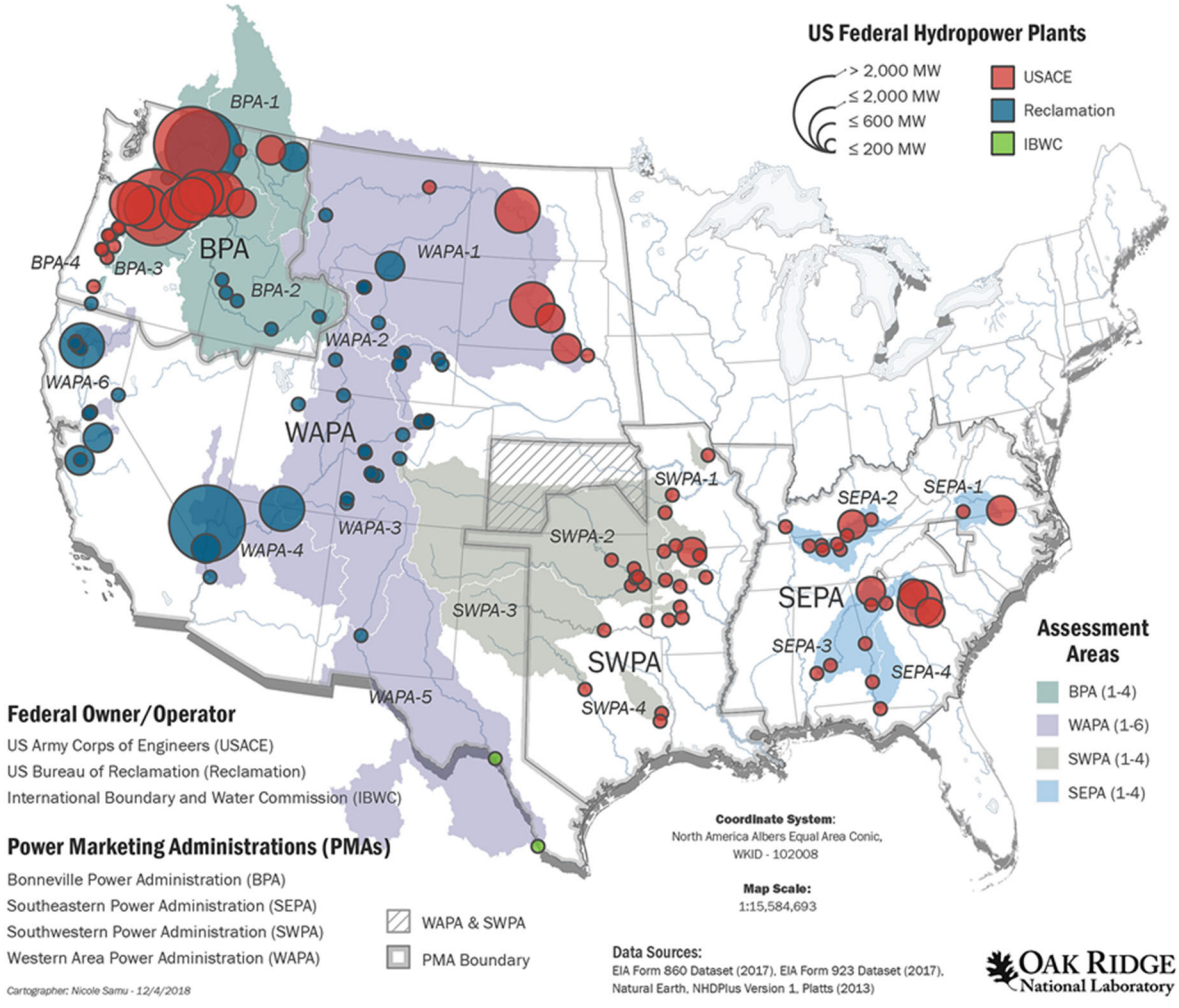


Figure ES.1. Federal hydropower facilities and federal power marketing regions in the United States. Note that part of Kansas is supplied by both WAPA and SWPA.

The latest climate projections from CMIP6 are produced from the latest global climate models (GCMs) for the Intergovernmental Panel on Climate Change Sixth Assessment. CMIP6 includes more than 50 GCMs with newly defined greenhouse gas emission scenarios that are a combination of Shared Socioeconomic Pathways (SSPs) and Representative Concentration Pathways (see Section 2.3 for a further description) to understand the Earth system response to an increase in anthropogenic forcing. Six CMIP6 GCMs under the SSP585 scenario were selected through an objective selection process that factors in relative model skill, uniqueness, data availability, and computational resources.

To understand how the choice of modeling and analytical approaches may affect the projections of future hydroclimate conditions and hydropower generation, a multimodel assessment framework was introduced. This multimodel assessment framework includes two downscaling methods, two reference meteorological observations, two hydrologic models, and two hydropower models to simulate the ensemble meteorological, hydrologic, and hydropower projections in the near-term (2020–2039) and mid-term (2040–2059) future periods with respect to the historical baseline period (1980–2019). Dynamical (Regional Climate Model version 4 [RegCM4]) and statistical (Double Bias Correction Constructed Analogues [DBCCA]) downscaling methods were applied to the six selected CMIP6 GCMs. Additionally, two reference meteorological observation data sets (Daymet and Livneh) were used to account for the uncertainties that arise from the choice of meteorological observations. Two hydrologic models, the Variable Infiltration Capacity (VIC) model and the Precipitation Runoff Modeling System (PRMS) model, were implemented to simulate the hydrologic processes (e.g., runoff, evapotranspiration, infiltration) under projected future climate conditions. Two regional hydropower models, Watershed Runoff-Energy Storage (WRES) and Water Management Hydropower model (WMP), were then applied to examine how the choice of the hydropower models may influence future hydropower projections. This multimodel assessment framework resulted in 96 sets of projected future hydropower generation illustrated in Figure ES.2.

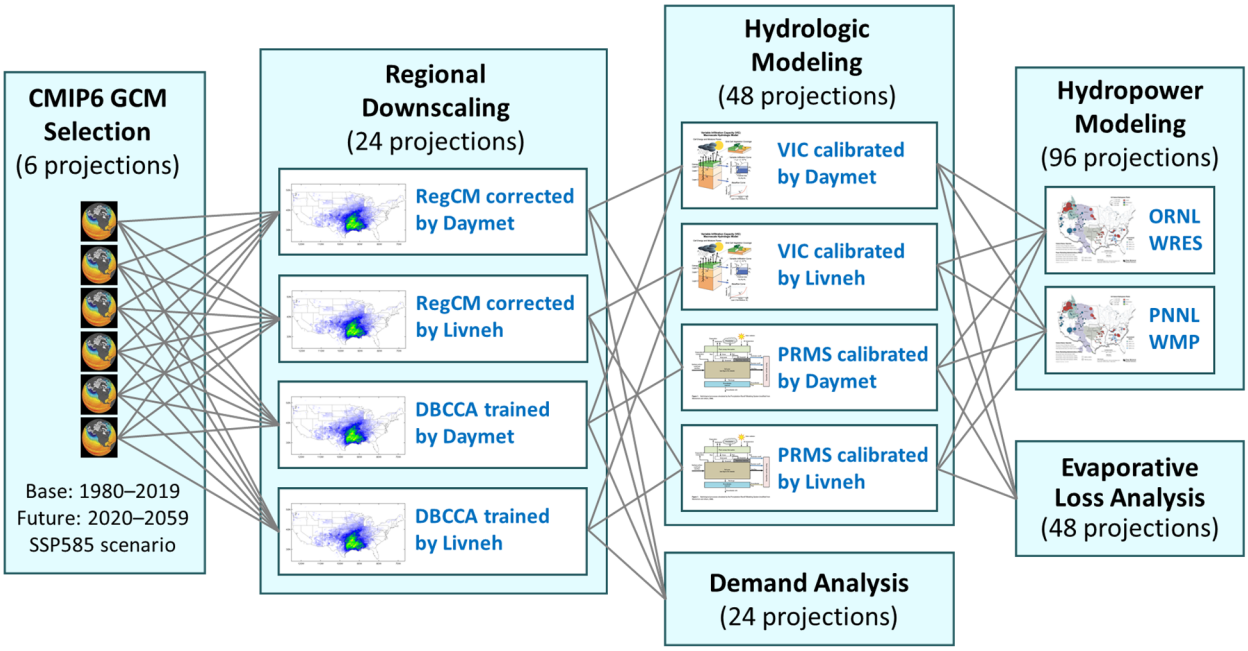


Figure ES.2. Multimodel ensemble simulation conducted for this assessment.

An analysis of variance technique was used to quantify the relative contribution of variance in ensemble projections arising from different modeling choices and their interactions. The key controlling factors in each of the key variables and PMA study areas were identified, demonstrating the sensitivity in these

modeling and analytical approaches. Additionally, the effect of climate change on the direct reservoir evaporation was analyzed to understand its potential long-term impacts on water availability for hydropower generation. A customer energy demand analysis in each PMA region was further conducted to evaluate how climate change may impact the energy demand for PMA preference customers. The potential climate change impacts on federal power marketing were also extensively discussed in this study as new, emerging energy market mechanisms impact near and long-term decisions for federal hydropower marketing plans.

Progression of the 9505 Assessments

The first 9505 assessment was conducted during 2010–2012 (Sale et al., 2012), and the second assessment was conducted during 2013–2017 (Kao et al., 2016). To comprehensively evaluate all federal hydropower plants located at various geographical regions in the United States, a nationally consistent assessment framework was designed to allow for the interregional comparison. Various types of hydropower-related data—including project characteristics, generation and power marketing records, observed hydrology and meteorology data, and watershed and land surface data—were assembled to support both assessments. A series of numerical models and analytical methods with different spatial resolutions were used to downscale the most current GCM findings at that time (CMIP Phase 3 for the first 9505 assessment and CMIP Phase 5 for the second 9505 assessment) into watershed-scale hydrologic projections to support the regional hydropower evaluation. Based on the historical runoff and hydropower generation data, empirical hydropower models were developed to project future annual and seasonal hydropower generation across various PMA study areas to support the power marketing evaluation. The technical findings were eventually reviewed by federal and state power and water resource managers, climate and hydrologic researchers, and policy analysts, and were used to support the development of DOE reports to Congress (DOE, 2013, 2017).

Unlike the third 9505 assessment, the previous assessments did not use a multimodel assessment framework and therefore only considered GCMs as the sole source of uncertainty. Results in the previous assessments were also derived from only one downscaling method, one reference meteorological observation, one hydrologic model, and one hydropower model. Although these models were calibrated based on the best available observations, there was no clear approach to discern whether a projected change was caused by the original GCM or by another factor. This lack of clarity could lower the confidence of the assessment findings and lead to a biased interpretation of climate change–induced risks to future federal hydropower generation. Based on extensive consultation with several federal hydropower stakeholder groups, a multimodel assessment framework was hence designed for this study which utilizes an ensemble of projections. Although the ensemble-based approach may provide more enriched insights, it still does not represent the full range of uncertainties related to all possible modeling choices. The ensemble-based approach adopted in this study serves as an initial example for a systematic analysis of broader uncertainties.

Results

The main findings of this study include the following:

- Although the six selected GCM projections do not cover the full range of CMIP6 models and emission scenarios, their projections are distributed across the median of all models, suggesting that they are not biased toward any one direction (Figure ES.3). The emission scenario has a clear influence on the projected change of temperature; however, its influence on the projected change of precipitation is less obvious at the conterminous US (CONUS) scale in the time frame of interest (till 2059). In other words, the climate states associated with different emission scenarios would diverge strongly in the second part of the twenty-first century.

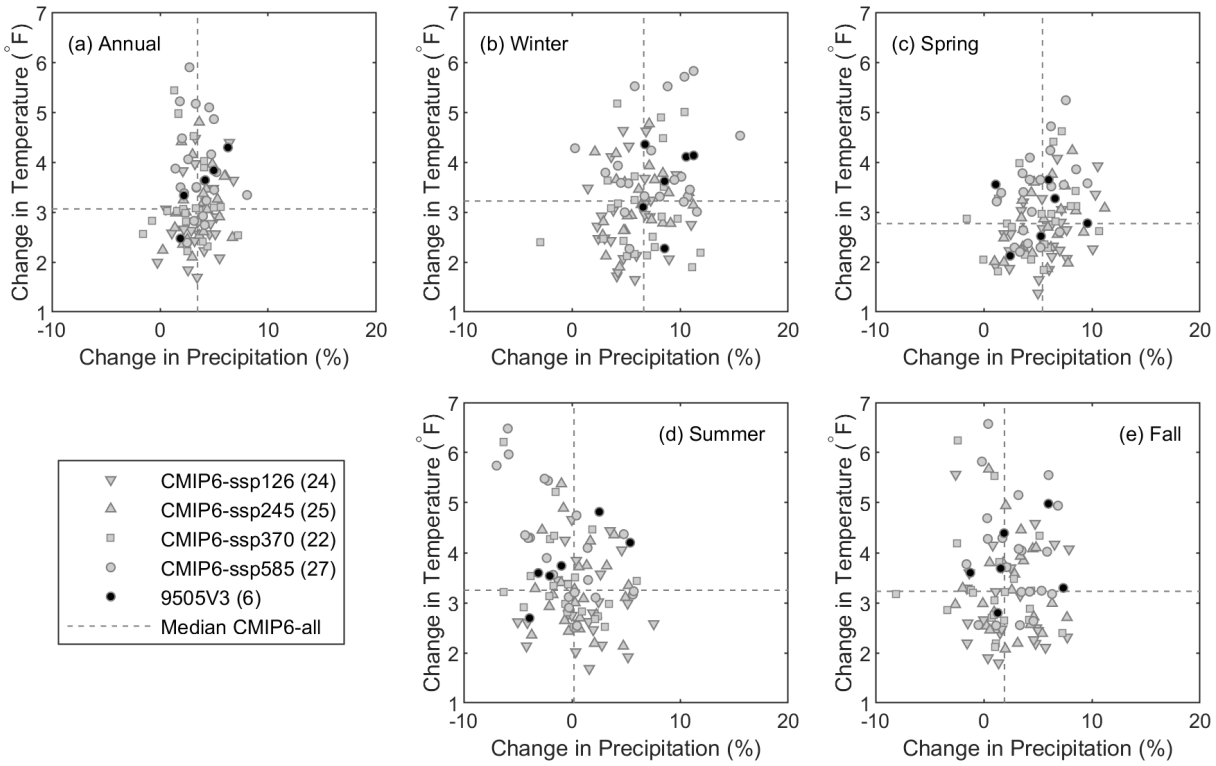


Figure ES.3. Annual and seasonal CONUS temperature and precipitation changes for 98 CMIP6 GCM projections under four emission scenarios (SSP126, SSP245, SSP370, and SSP585). For each GCM projection, the average changes are calculated from the 1980–2019 baseline to the 2020–2059 future period.

- For temperature, all CMIP6 projections showed a consistent increase ranging from 1°F to 6°F across all seasons. For precipitation, except for parts of summer and fall, a consistent increase is projected (especially in winter), resulting in a net annual precipitation increase across most of the projections (−2% to 8%).

- Certain differences in future hydroclimate projections arise from the choice of downscaling methods and reference meteorological observations. These differences are more prominent in extremes than mean variables.

- The total annual runoff is projected to increase across the CONUS in both future periods, with a higher magnitude in the mid-term future period (2040–2059). At the seasonal scale, the winter runoff is generally projected to increase across most of the CONUS, especially in the mid-term future period. The spring runoff is also projected to increase in most of the CONUS, but with a decrease projected in California, Lower Colorado, and Rio Grande regions in the near-term future period (2020-2039). During summer, the runoff is projected to decrease across many parts of the CONUS.

- High runoff (95th percentile) is projected to increase in the majority of CONUS watersheds; however, low runoff (5th percentile of 7-day average) is projected to decrease in the eastern/central and western coastal areas of the CONUS. Regions exhibiting a projected increase in high runoff (and a consequent decrease in low runoff) are the result of an intensified hydrologic cycle under future climate conditions.

- At the annual time scale (Figure ES.4), the total hydropower generation for all PMAs is projected around 120 TWh/year in both near-term (2020–2039) and mid-term (2040–2059) future periods, close to the baseline (1980–2019) level. Most PMAs feature a wider multimodel spread in the mid-term future period than in the near-term future period, with no obvious change in the ensemble mean.

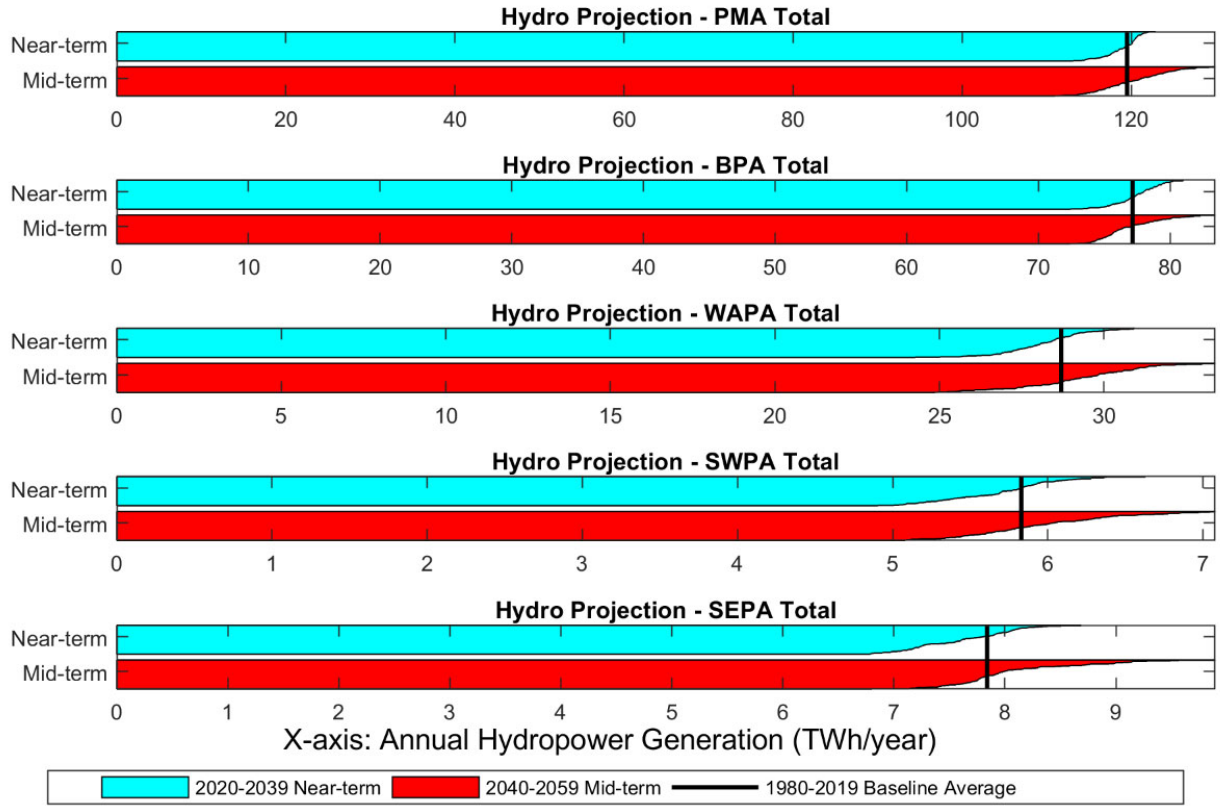


Figure ES.4. Projection of annual hydropower generation by PMA. Each horizontal bar is a composite of 96 smaller bars for each hydropower projection, sorted in descending order, from top to bottom.

- At the seasonal time scale, the majority of models projected increasing generation in winter and spring, and decreasing generation in summer and fall in both near-term and mid-term future periods. Such a result is mainly governed by the projected changes in BPA and WAPA caused by the earlier snowmelt and changing runoff seasonality.

- Aggregated over all PMAs, annual streamflow is projected to increase by 9%, and hydropower generation is projected to increase by about 4%. More specifically, the BPA annual natural streamflow is projected to increase by 8%, and hydropower generation is projected to increase by 2%. In WAPA, streamflow is projected to increase by 11%, and hydropower generation is projected to increase by 7%. In SWPA, streamflow is projected to increase by 4%, and hydropower generation is projected to increase by 3%. In SEPA, streamflow is projected to increase by 9%, and hydropower generation is projected to increase by 7%. At the seasonal scale, increased streamflow and hydropower are projected to reach a higher level and output in most seasons and regions, except for summer in BPA, where the flow is projected to reduce by 8%, which is projected to lead to a 4% decrease in summer hydropower generation. The spreads of annual and seasonal relative changes in both natural flow and hydropower are generally smaller in BPA and WAPA than in SWPA and SEPA. This difference may be influenced by the different reservoir features in those PMA regions.

- More severe reservoir evaporation losses are projected to occur during the mid-term future period (2040–2059) than the near-term future period (2020–2039). The evaporation loss growth rate in the arid/semi-arid Western United States is much higher than that in the eastern United States, which is expected to worsen hydrological drought conditions in the future.

- The projected increase in average temperature, which yields a lower number of heating degree days (lower demand for heating) in the winter months and a higher number of cooling degree days (more demand for cooling) in the summer months, results in projected decreases in the electricity sales of PMA preference customers to their end customers in the winter and increased sales in the summer. There is more variability for the expected response of sales across the PMAs to cooling degree days than to heating degree days. SWPA and WAPA–Desert Southwest stand out as having the strongest projected increase in summer electricity sales because of higher cooling degree days across all PMAs.

- PMA preference customers will likely want to increase the fraction of their federal hydropower allocation that they receive during the summer season—to the extent allowed in their contracts with the PMAs. Having operational flexibility to shift the water allocation timing and volume from the winter to the summer months would be highly valuable for all PMAs, and particularly BPA. However, in federal multipurpose water resource projects, hydropower is typically not the highest priority purpose and the flexibility to change water allocation timing in response to changes in PMA customer demand is very limited. WAPA and BPA have recently taken an active interest in considering and formally joining the regional western energy imbalance markets or independent system operators/regional transmission organizations to optimize replacement power purchases/surplus power sales and minimize the costly impacts of limited water supply (e.g., low streamflow period, other non-power reservoir constraints, forced outages/maintenance period) and increased load.

- An analysis of variance was used in this assessment to quantify the relative contribution of uncertainties in ensemble projections arising from different modeling choices (main factors) and their interactions. Specifically for runoff and generation, the selection of climate model remains the dominant factor in many parts of the CONUS, but the role of hydrologic and hydropower models (along with other factors) can be more important in some locations and seasons. This analysis addresses the needs for a comprehensive evaluation across all components in the modeling chain for a more informed hydroclimate impact assessment.

Overview of Potential Risks

Overall, several potential risks that may impact the resilience of future federal hydropower generation were identified. They include the following:

- **Hydrologic extremes**: The intensification of future hydrologic cycles and extreme events is one of the most critical issues threatening the resilience of power systems and infrastructure as suggested by the projected high/low precipitation and runoff. This is consistent with findings in other scientific literature. Both historical observations and model projections suggest that the intensity, frequency, and magnitude of extreme rainfall events will continue to increase, which will likely challenge conventional reservoir management practices. Although the reservoirs have been traditionally designed using relatively conservative rainfall estimates (i.e., probable maximum precipitation), recent extreme events, such as the 2017 Hurricane Harvey near Houston, Texas, demonstrated that such an extremely large probable maximum precipitation estimate can still be exceeded (Kao et al., 2019), suggesting the potential need for more comprehensive evaluations.

On the other hand, the duration and severity of extreme drought events are also projected to increase in many parts of the United States. Although the annual precipitation and runoff are generally projected to increase in a warming environment, the distribution is not uniform in space and time. Reclamation (2021) conducted a comprehensive drought assessment for the western United States using historical streamflow observations, future projections, and paleohydrology. The results suggested that the severity of the drought events may not be sufficiently captured, based on limited historical streamflow observations. The ongoing severe drought in the western states resulted in an unprecedent disruption to the water supply and hydropower generation, demonstrating the dire impacts of drought.

- **Conflicting timing of supply/demand changes**: Similar to in the previous assessments, temperature-driven early snowmelt is projected in most of the western United States, suggesting that the bulk of runoff may arrive earlier in spring. However, as informed by the demand analysis, more temperature-driven water and energy demand is expected to shift from winter to summer, and hence creates a conflict. Although ideally one may expect to mitigate this conflict through reservoir management, the intensified hydrologic extremes combined with all other competing water management objectives limit the ability and flexibility in storing more water resources to meet the peak demand. Furthermore, in arid regions, the enhanced reservoir evaporation may result in a sizable storage reduction and further exacerbate the nexus of electricity demand and water availability.

- **Other indirect impacts**: Given the specific legislative language of the SWA, this assessment focuses on understanding how climate change may affect water availability for hydroelectric power generation and federal hydropower marketing. However, climate change can result in even broader impacts, such as the disruptions due to wildfire, increasing water demands, and more difficulties to meet environmental objectives. Additionally, the issue of aging infrastructures may also reduce the system's ability to mitigate runoff variability and increase the difficulty of future operation. Although these issues were not quantitatively analyzed in this assessment, they can be further investigated in future studies.

Assessment Limitations

Although the assessment uses state-of-the-art data and models, assessment limitations remain. They include the following:

- **Interpretation of future projections**: Although GCM-driven dynamical/statistical downscaling may provide the most scientifically defensible regional-scale climate projections, it should not be considered an absolute, day-to-day weather prediction. The main purpose of climate modeling is to simulate how general climate statistics may evolve with respect to the specified future emission scenarios—not to provide an exact prediction of future weather and hydrology. Also, a simulation is only one of the tools that one may use to evaluate the potential impacts and system vulnerabilities. Other observation-based assessment approaches are equally valuable and should not be omitted for a more holistic understanding.

- **Broader characterization of uncertainty**: Although the multimodel framework provides a more comprehensive projection of future US federal hydropower generation, multimodel projections still do not represent the full range of uncertainties related to all possible modeling choices. The extent of warmer temperatures causing increasingly drier soils and greater evapotranspiration is a significant additional uncertainty. Even in the case of higher average precipitation, drier soils and greater evapotranspiration will lead to reduced water availability and lower average generation. The true uncertainties are more comprehensive and may not be fully captured because of limited knowledge,

tools, and resources. The ensemble-based approach adopted in this study can serve as an initial example for the systematic analysis of broader uncertainties.

- **Progression of climate science**: Although the capabilities of GCMs have continuously improved through the years, many open challenges have not been resolved. For instance, although human activities play an important role in the Earth system environment, many of the GCM simulations were conducted without considering the potential human influence on land use and land cover change, and surface hydrologic alterations. Therefore, recurring climate impact assessments based on the best available climate science remain necessary.

- **Regional assessment focus***:* Overall, this study focused on 18 PMA study areas rather than individual reservoirs or power plants. Impacts on site-specific features, such as reservoir operation rules, water withdrawal/return, environmental flow requirements, and energy generation, were not explicitly modeled at each power plant. In other words, this study provides a first-order assessment to identify areas with the highest risk under projected climate conditions. If a concern is identified for a specific region (e.g., change of streamflow seasonality), a regionally focused study can then be conducted. The assessment itself does not replace the existing site-specific models and tools used by the PMA's water and energy resource managers.

Future Assessment Needs

Through the series of 9505 assessments, a quantitively modeling framework has been established to gradually downscale the latest CMIP findings into regional, watershed-scale hydrologic and hydropower projections to support the understanding of risks for federal hydropower generation and marketing. Although this framework has successfully achieved the anticipated SWA objectives, the climate change impacts are much broader and more complex. Moving forward, additional studies and data support should be beneficial to the broader US hydropower community; they include the following:

- **Need to conduct basin-specific studies**: Considering the varying geographical and socioeconomical challenges in different river basins, reservoir-specific studies may be required for basins with high water and energy interests. Studies using operational models forced by up-to-date hydroclimate projections, such as the recent RMJOC assessments (Pytlak et al., 2018; Glabau et al., 2020), are one way to evaluate the risks and identify possible mitigation actions.

- **Need to better understand the characteristics and impacts of future drought conditions**: As reported by Williams et al. (2022), 2000–2021 was the driest 22 year period in the southwestern United States in the past 1,200 years. This prolonged megadrought highlights the need to better understand the characteristics (e.g., severity, timing, duration) and impacts of droughts for the resilience of our long-term energy and water supply. In particular, efforts should focus on improving the understanding, modeling, and analytics associated with future drought events.

- **Need actionable, climate-informed data support**: Since hydroclimate modeling is not within a utility's original mission space, a utility may not have sufficient resources or dedicated in-house expertise to evaluate the risks due to long-term climate change. To reduce a utility's burden in conducting a full-scale hydroclimate study (i.e., from GCM selection all the way to river management simulation), it will be beneficial to provide actionable, climate-informed data support to the broader energy and water communities. The capabilities established through this federal 9505 assessment may serve as a starting point.

- **Need to address the broader risks**: Further disruptions due to wildfire, environmental requirements, and reduced operational flexibilities should be jointly considered. Additionally, the issue of aging

infrastructure may reduce the system's ability to mitigate runoff variability and increase the difficulty of future operation. Although these issues were not within the scope of this assessment, they should be further investigated in future studies.

Conclusions

This third 9505 assessment builds from the previous two assessments (Sale et al., 2012; Kao et al., 2016) by analyzing the effects of climate change on annual and seasonal federal hydropower generation and other related risks. A spatially consistent assessment approach was designed to evaluate hydropower generation from 132 federal hydropower plants that are marketed by 4 PMAs. The assessment incorporates a new multimodel framework that examined to what extent the methodological choices may influence the full range of projected hydropower generation in the near-term (2020–2039) and mid-term (2040–2059) future periods. Additionally, this assessment provides a more comprehensive understanding of how the electricity demand or load from the PMA preference customers that have historically relied on long-term contracts at low-cost rates are sensitive to changes in temperature, both annually and seasonally, in each PMA region.

In support of the two previous assessments, maintaining the operational flexibilities will still be the key challenge for federal hydropower reservoirs that are projected to experience seasonal supply and demand changes. Early snowmelt and change of runoff seasonality remain the most important climate change effects for federal hydropower generation. Overall, the total annual runoff, streamflow, and hydropower generation in the near-term and mid-term future periods is projected to increase across the CONUS. Winter and spring runoff is projected to increase in most PMA regions, and summer runoff is projected to decrease across much of the CONUS. Regarding hydrologic extremes, high runoff is projected to increase in the majority of CONUS watersheds; however, low runoff is projected to decrease in the eastern/central and western coastal areas of the CONUS. The regions experiencing increases in high runoff and decreases in low runoff are projected to observe intensification of hydrologic cycles under future climate conditions.

The most notable trends of PMA customer demand from the CMIP6-based climate projections are the warmer winter temperatures that result in reduced load for heating and increased load for cooling in the summer that suggests a prominent, emerging load trend for air conditioning. This seasonal temperature-load dynamic results in reduced winter hydropower sales from the preference customers to their end-use customers and increased sales in the summer. Thus, the ability to shift the timing and volume of water allocated for hydropower production on a seasonal basis would be valuable for the federal hydropower fleet, particularly for reservoirs experiencing drier conditions/constrained streamflow periods and increased electricity demand. From an annual perspective, a larger storage capacity or more flexible operating practices would be highly beneficial to maintain consistent, low-cost electricity prices for the PMA preference customers. Meeting long-term marketing plan goals for the PMAs is vital, and some PMAs have considered joining other electricity market options as a strategy to improve their ability to provide electricity reliably and economically to their customers.

ABBREVIATIONS

AC	air conditioning
ACCESS-CM2	Australian Community Climate and Earth System Simulator Coupled Climate Model
ANOVA	analysis of variance
A-V	area-volume
BCCA	Bias Correction Constructed Analogues
BCC-CSM2-MR	Beijing Climate Center Climate System Model running on medium resolution
BCSD	bias-corrected spatial disaggregation
BPA	Bonneville Power Administration
CAISO	California Independent System Operator
CDD	cooling degree days
CM	climate model
CMIP	Coupled Model Intercomparison Project
CNRM-ESM2-1	National Centre for Meteorological Research Earth System Model of second generation
CONUS	conterminous United States
CRSP	Colorado River Storage Project
DBCCA	Double Bias Correction Constructed Analogues
DJF	December-January-February
DOE	US Department of Energy
DS	downscaling method
DSW	Desert Southwest
EIA	US Energy Information Administration
EIM	energy imbalance market
EOF	empirical orthogonal function
FCRPS	Federal Columbia River Power System
FERC	Federal Energy Regulatory Commission
GA/AL/SC/JW	Georgia/Alabama/South Carolina/Jim Woodruff
GCM	global climate model
GRanD	Global Reservoir and Dam Database
HDD	heating degree days
HM	hydrologic model
HRU	hydrologic response unit
HUC	hydrologic unit code
IBWC	International Boundary and Water Commission
IOU	investor-owned utility
ISO	independent system operator
JJA	June-July-August
LEM	lake evaporation model
LOCA	localized constructed analogs
MACA	multivariate adaptive constructed analogs
MAM	March-April-May

MetF	reference meteorological forcing (meteorological observation)
MOSART-WM	Model for Scale Adaptive River Transport water management component
MPI-ESM1-2-HR	Max Planck Institute Earth System Model run at high resolution
MRI-ESM2-0	Meteorological Research Institute Earth System Model version 2.0
MTCLIM	Mountain Microclimate Simulation Model
NARR	North American Regional Reanalysis
NCEI	National Center for Environmental Information
NCEP1	National Center for Environmental Prediction reanalysis data set
NID	National Inventory of Dams
NLDAS	North American Land Data Assimilation System
NorESM2-MM	Norwegian Earth System Model
NRMSE	normalized root mean square error
NSE	Nash–Sutcliffe efficiency
NWIS	National Water Information System
O&M	operations and maintenance
ORNL	Oak Ridge National Laboratory
PC	principal component
PM	hydropower model
PMA	Power Marketing Administration
PRISM	Parameter-elevation Regressions on Independent Slopes Model
PRMS	Precipitation Runoff Modeling System
PSO	particle swarm optimization
RCM	regional climate model
RCP	Representative Concentration Pathway
Reclamation	US Bureau of Reclamation
RegCM4	Regional Climate Model version 4
RMJOC	River Management Joint Operations Committee
RMSE	root mean square error
ROR	run-of-the-river
ROS	ratio of spatial standard deviation
RTO	regional transmission organization
SEPA	Southeastern Power Administration
SN/RM	Sierra Nevada/Rocky Mountains
SON	September-October-November
SPP	Southwest Power Pool
SS	sum of squares
SSP	Shared Socioeconomic Pathway
SWA	SECURE Water Act
SWE	snow water equivalent
SWPA	Southwestern Power Administration
TIGER	Topologically Integrated Geographic Encoding and Referencing
TVA	Tennessee Valley Authority
UGP	upper Great Plains
USACE	US Army Corps of Engineers

USGS	US Geological Survey
VIC	Variable Infiltration Capacity
WAPA	Western Area Power Administration
WEIS	Western Energy Imbalance Service
WMP	Water Management Hydropower
WRES	Watershed Runoff-Energy Storage

(This page intentionally left blank)

1. INTRODUCTION

This study, *The Third Assessment of the Effects of Climate Change on Federal Hydropower*, directed by Section 9505 of the SECURE Water Act (SWA) of 2009, is the third quinquennial report on evaluating the effects of climate change on hydroelectric energy generated from 132 US federal hydropower plants marketed by four US Department of Energy (DOE) Power Marketing Administrations (PMAs). The technical assessment is conducted by DOE's Oak Ridge National Laboratory (ORNL), Pacific Northwest National Laboratory, and Texas A&M University under the guidance of DOE's Water Power Technologies Office. This study is the result of extensive consultation with the four federal PMAs (Bonneville Power Administration [BPA], Western Area Power Administration [WAPA], Southwestern Power Administration [SWPA], and Southeastern Power Administration [SEPA]), as well as other agencies, including federal hydropower owners/operators (the US Army Corps of Engineers [USACE], US Bureau of Reclamation [Reclamation]), US Geological Survey (USGS), and National Oceanic and Atmospheric Administration. The main findings of this assessment, along with the PMA administrators' Recommendations, will be included in a subsequent DOE report to Congress. The assessment method and the technical findings are described in this report.

1.1 BACKGROUND

The Omnibus Public Land Management Act of 2009 (Public Law 111-11) Subtitle F – SECURE Water Act (SWA) was passed into law on March 30, 2009. The SWA authorizes federal agencies to study and improve water management and to increase the acquisition and analysis of data describing water resources used for irrigation, hydropower, municipal uses, environmental conservation, and other purposes. A primary purpose of the SWA was to facilitate an analysis and plan for potential effects that climate change may have on the hydrologic cycle that provides water for communities, economic growth, and ecosystem protection in the United States. The SWA legislation has 10 sections:

(1) 9501, Findings
(2) 9502, Definitions
(3) 9503, Reclamation Climate Change and Water Program
(4) 9504, Water Management Improvement
(5) 9505, Hydroelectric Power Assessment (this report)
(6) 9506, Climate Change and Water Intragovernmental Panel
(7) 9507, Water Data Enhancement by the United States Geological Survey
(8) 9508, National Water Availability and Use Assessment Program
(9) 9509, Research Agreement Authority
(10) 9510, Effect

SWA Section 9505(c) directs the secretary of energy to assess the effects of global climate change on water supplies required for hydropower generation at federal water projects and to present the results in a report to Congress (see APPENDIX A for the text of the original legislation). DOE is to repeat these assessments every five years until 2023. The 9505 assessment has been conducted in coordination with each of the PMA administrators that sell and distribute hydroelectricity from federal projects, as well as with USACE and Reclamation, who own and operate federal hydropower dams.

Sections 9503 and 9505 address climate change effects on water resources and hydroelectric generation. Section 9503 is led by Reclamation and addresses a broad range of water resource issues but only in the western United States. The 9505 assessment described in this report focuses specifically on federal hydropower generation and marketing. To ensure consistency between 9503 and 9505 assessments, DOE and the 9505 assessment team worked closely with Reclamation during the development of both studies. Although the methods used in this report (Section 2) vary from the 9503 assessment methods owing to

differences in scope and agency missions, the high-level conclusions of the two reports are generally consistent and complement each other.

The first 9505 assessment was conducted during 2010–2012 (Sale et al., 2012), and the second assessment was conducted during 2013–2017 (Kao et al., 2016). To comprehensively evaluate all federal hydropower plants located at various geographical regions in the United States, a nationally consistent assessment framework was designed to allow for the interregional comparison. Various types of hydropower-related data—including project characteristics, generation and power marketing records, observed hydrology and meteorology data, and watershed and land surface data—were assembled to support both assessments. A series of numerical models and analytical methods with different spatial resolutions were used to downscale the most current global climate model (GCM) findings at that time (Coupled Model Intercomparison Project [CMIP] Phase 3 for the first 9505 assessment and CMIP Phase 5 [CMIP5] for the second 9505 assessment) into watershed-scale hydrologic projections to support the regional hydropower evaluation. Based on the historical runoff and hydropower generation data, empirical hydropower models (PMs) were developed to project future annual and seasonal hydropower generation across various PMA study areas to support the power marketing evaluation. The technical findings were eventually reviewed by federal and state power and water resource managers, climate and hydrologic researchers, and policy analysts, and were used to support the development of DOE reports to Congress (DOE, 2013; DOE, 2017).

The key findings from the previous 9505 assessments (Sale et al., 2012; Kao et al., 2016) include the following:

- Statistically significant changes in future climate, water availability, and hydropower generation are projected in many parts of the country. The modeling results suggest that temperature is projected to increase across all study areas and all seasons, whereas changes in precipitation, runoff, and hydropower generation might increase or decrease depending on season and geographical location. The results were generally consistent with other concurrent hydroclimate studies.

- For the multimodel median, precipitation is projected to increase slightly. A large multimodel uncertainty and regional variability of precipitation projections was also identified. Overall, more precipitation is projected in the mid-term future period than in the near-term future period. A higher increase in precipitation is projected in the SWPA and SEPA regions, as well as in the Upper Missouri River basin in the WAPA region.

- The projection of future runoff varies substantially by regions. In the snowmelt-dominated regions (e.g., most of the BPA and WAPA regions), the winter and spring runoff is projected to increase, whereas the summer and fall runoff is projected to decrease. Such a shift in runoff seasonality is likely caused by the increasing air temperature and earlier snowmelt. In the rainfall-dominated regions (e.g., SWPA, SEPA, and some areas of BPA and WAPA), the change in runoff is mainly caused by the change in precipitation. The percentage change of future runoff is larger than the percentage change of precipitation.

- The change in future hydropower generation is mainly influenced by the change in future runoff and precipitation. Because the combined reservoir storage may provide a buffer to help absorb part of the runoff variability, the projected change in future hydropower generation is smaller than the projected change in future runoff. The impact of runoff variability on hydropower generation is less noticeable in regions with relatively larger reservoir storage capacities (e.g., BPA and WAPA). For regions with smaller storage capacities (e.g., SWPA and SEPA), the change in future hydropower generation will follow the change in future runoff more closely.

- Water management and investments in new equipment should focus on maintaining operational flexibility to preserve current hydropower generation. Continued monitoring of climate data and research advancements are needed to determine if and when current practices need to change.

This third 9505 assessment builds upon the framework established in the first two assessments and leverages the latest global climate projections from the CMIP Phase 6 (CMIP6). In particular, to understand how the choice of modeling and analytical approaches may affect the projections of future hydroclimate conditions and hydropower generation, a multimodel assessment framework has been introduced. This multimodel assessment framework includes two downscaling methods (DSs), two reference meteorological observations (MetFs), two calibrated hydrologic models (HMs), and two regional PMs to simulate ensemble meteorological, hydrologic, and hydropower projections in the near-term (2020–2039) and mid-term (2040–2059) future periods. By analyzing the contribution of variance associated with these factors, this study identifies the key controlling factor in each variable and study area, thereby demonstrating the sensitivity of these modeling/analytical approaches and identifying potential future research efforts. Additionally, how climate change may affect direct reservoir evaporation and customer energy demand in each PMA region is also analyzed. The potential climate change impacts on federal power marketing are also extensively discussed in this study.

1.2 THE US FEDERAL HYDROPOWER SYSTEM

1.2.1 Federal Hydropower Generation

As of 2019, there were 2,270 conventional (i.e., once-through) hydropower plants in the United States with a total of 80.25 GW of generating capacity producing 6.6% of all electricity and 38% of electricity from renewables (Uría Martínez et al., 2021). Additionally, there were 43 pumped storage hydropower plants with a total of 21.9 GW of generating capacity providing 94% of utility-scale electrical energy storage. Among these hydropower plants, around 40% of the generating capacity (42.5 GW) was provided by federal projects and built and/or operated by one of four agencies: USACE, Reclamation, the Tennessee Valley Authority (TVA), and the International Boundary and Water Commission (IBWC). Other nonfederal hydropower is regulated by the Federal Energy Regulatory Commission (FERC) under the authority of the Federal Power Act and owned and/or operated by investor-owned utilities (IOUs), publicly owned utilities, state agencies, and nonutilities (Uría Martínez et al., 2015). Several additional nonfederal hydropower facilities that are outside FERC jurisdiction have also been authorized by Reclamation through the lease of power privilege process.

Although there are fewer federal hydropower plants than nonfederal plants, the federal plants have, on average, 10 times more generating capacity per plant. Most of these federal hydropower plants are located at multipurpose reservoirs that also provide nonpower services, such as flood control; navigation; water supply for municipalities, industries, agriculture; recreation; and protection of environmental resources, including water quality, fish, and wildlife. Since many nonpower services have higher priority than hydropower, the generation of federal hydropower is under a variety of competing constraints and may not be as flexible, despite the large water storage capacity in federal reservoirs.

Although federal hydropower projects are owned and operated by federal water development agencies (e.g., USACE, Reclamation, TVA, IBWC), the electricity produced at USACE, Reclamation, and IBWC projects is mostly marketed and distributed by the PMAs (i.e., BPA, WAPA, SWPA, and SEPA). Given the specific legislative language of SWA Section 9505, only the 132 hydropower plants marketed through the PMAs as *federal hydropower* are considered in this study. Because TVA is not a PMA, and the hydropower generated from TVA facilities is not marketed by a PMA, the 30 TVA hydropower plants are not included in this assessment. Similarly, the assessment does not include USACE St. Mary's Falls or St. Stephen because the electricity generated from these two hydropower plants is not marketed through

PMAs. A detailed list of these 132 federal hydropower plants is provided in APPENDIX B and illustrated in Figure 1.1. Table 1.1 contains a summary of the federal hydropower plants evaluated in this assessment broken down by the number of plants, capacity, and the 1980–2019 average annual generation for each operating agency.

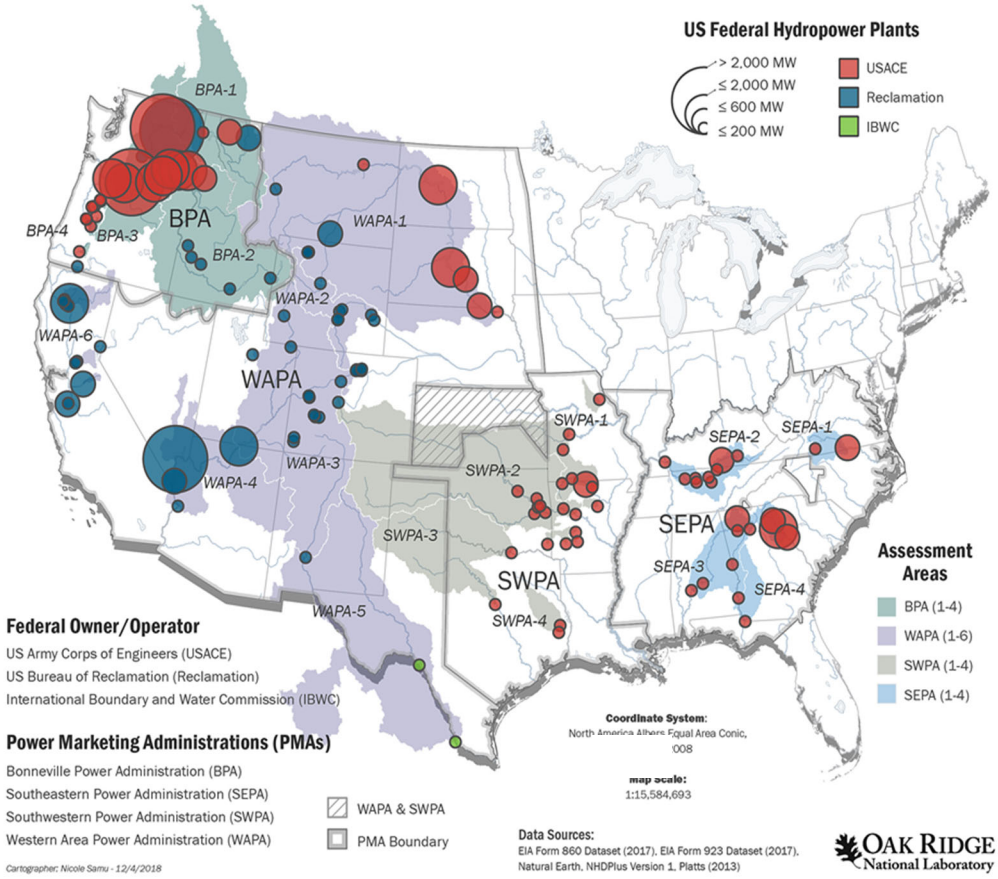


Figure 1.1. Federal hydropower facilities and federal power marketing regions in the United States. Note that part of Kansas is supplied by both WAPA and SWPA.

Based on river basin hydrology and power systems, the 132 federal hydropower plants are further grouped into 18 assessment areas, labeled as BPA-1–4, WAPA-1–6, SWPA-1–4, and SEPA-1–4 (see APPENDIX B). USACE has the most hydropower plants, followed by Reclamation, and then IBWC. The 75 PMA-marketed USACE hydropower plants are in 16 states across the United States (Figure 1.1). More than 60 FERC-regulated, nonfederal hydropower plants are located at USACE dams, with an additional 2,700 MW of capacity (Johnson, 2021). The oldest USACE hydropower facility is at Bonneville Dam on the lower Columbia River, which came online in 1938. The most recent USACE project is the Robert D. Willis Hydropower Project in Texas, which came online in 1989.

Reclamation owns 76 federal hydropower plants in 11 western states, with 57 hydropower plants marketed through PMAs. Additionally, there are more than 60 nonfederal hydropower plants on Reclamation dams and canals with a combined capacity of more than 500 MW, which are regulated by FERC or through Reclamation's lease of power privilege process. Reclamation's mission is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public. To this end, Reclamation's primary mission is the delivery of water for irrigation to end users in the western states. Electricity produced at Reclamation facilities is either

used internally at projects or sold to external users. The power is used primarily to deliver water to meet other authorized objectives than hydropower generation. The oldest Reclamation hydropower plant is the Theodore Roosevelt facility on the Salt River in Arizona, which began operating in 1909. The largest Reclamation hydropower plant is at Grand Coulee Dam, which has an installed capacity exceeding 6,800 MW, making it among the 10 largest hydropower plants in the world. IBWC owns and operates two small hydropower projects on the Rio Grande River. The hydropower from these two plants is marketed by WAPA.

Table 1.1. Federal hydropower plants evaluated in this assessment

Owner/operator	Number of plants	Capacity[a] (MW)	1980–2019 generation[b] (GWh/year)
USACE	73[c]	21,745	75,947
Reclamation	57[d]	15,104	41,831
IBWC	2	98	197
Total[e]	132	36,947	117,975

[a] EIA 2019 total nameplate capacity, includes both conventional hydro and pumped storage.
[b] EIA and PMA average annual generation from 1980 to 2019, conventional hydro only.
[c] Does not include USACE St. Mary's Falls and St. Stephen, which are not marketed through PMAs.
[d] Does not include ~20 other Reclamation plants that are not marketed through PMAs.
[e] Does not include 30 TVA plants because TVA hydropower is not marketed through PMAs.

1.2.2 Marketing Federal Hydropower

PMAs are federal agencies tasked with marketing power produced by 132 federal hydropower plants owned or operated by USACE, Reclamation, or IBWC. The federal power marketing program started in the early 1900s when excess hydropower produced at federal projects was sold to repay the government's investment in the projects. Four PMAs—BPA (1937), WAPA (1977), SWPA (1944), and SEPA (1950)—market federal hydropower in different regions of the country. Federal laws enacted in the 1930s authorized construction of federal transmission to avoid relying on private companies to bring the electricity from federal plants to the public. All PMAs except SEPA own transmission assets.

By statute, PMAs prioritize access to their federal hydropower resources for preference customers—mostly municipalities, political subdivisions, and cooperatives—over for-profit entities. The preference clauses, introduced in several pieces of legislation in the early 1900s, were designed to ensure that the operation of federal hydropower assets benefited the public, contributed to the economic development of rural areas, and avoided a monopoly in the nascent electric industry (GAO, 2001).

Unlike privately owned utilities, PMAs must sell electricity "at the lowest possible rates consistent with sound business principles" (Flood Control Act of 1944). The cost-based rates must be sufficient to repay the US Treasury for the initial investment to construct the power system (including transmission) within 50 years from start of operation.[1] Additionally, the rates must annually repay the expensed interest on the unamortized debt, the operations and maintenance (O&M) costs for the hydropower function of the power system, the cost of purchased power, a portion of O&M for other system purposes (e.g., irrigation), and environmental project costs. DOE Order RA 6120.2 requires the PMAs to conduct annual studies to assess whether the existing rates can meet the revenue objectives. If the rates are not adequate, the PMA

[1] A power system is a system composed of one or more projects hydraulically and/or electrically integrated and treated as one unit for the purpose of establishing rates.

proposes a change in rates. Rate cases are a public process and must be approved by the deputy energy secretary and by FERC.

Although rates can be revised annually, the capacity and energy allocations of federal hydropower are set with long-term contracts. Power marketing plans establish the marketing terms (e.g., total marketable resource), the contract term (e.g., marketing firm vs. peaking power, contract length), resource pools, load factor limit, and water withdrawal provisions with input from customers and other interested stakeholders. Upon publication of the final marketing plan in the *Federal Register*, contracts are then signed with each individual customer for their proposed allocations. Contract terms vary across PMAs and across power systems within a PMA.[2]

Since 1974, BPA has been a self-financed agency, which means it can directly apply the revenue from power and transmission sales to cover its costs, both O&M and capital expenses, while meeting the repayment obligations to the federal government for the funds (interest included) used for the construction of the Federal Columbia River Power System (FCRPS). For the rest of the PMAs, the Energy and Water Development and Related Agencies Appropriation Act of 2010 allows for using the revenues from power and transmission sales to cover their O&M expenditures starting in FY 2011. Investments in rehabilitation or upgrades to the federal hydropower assets for those PMAs are funded by a combination of congressional appropriations and PMA customer funding agreements.

In their day-to-day project operations, the PMAs coordinate closely with USACE or Reclamation. These agencies determine the constraints governing water availability for hydropower and implement the generation schedules submitted by the preference customers to the PMA. Fernandez (2020) describes in more detail the interactions between the federal hydropower owners and the PMAs.

Because runoff availability changes from year to year, actual federal hydropower generation can be higher or lower than the amount needed to fulfill the contracted allocations of capacity and energy. To resolve imbalances, PMAs trade with other utilities or trade in organized wholesale markets. The cost of power purchased to resolve deficits is passed on to customers. Traditionally, PMAs have resolved these imbalances mainly through trading hubs or bilateral trades with other utilities. Recently, BPA and multiple WAPA power marketing regions have joined or announced plans to join either the California Independent System Operator's (CAISO's) Western Energy Imbalance Market (EIM) or the Southwest Power Pool's (SPP's) Western Energy Imbalance Service (WEIS) Market.

1.3 ORGANIZATION OF THE REPORT

The methods, data sources, and analyses used in this report are described in Section 2. Sections 3, 4, 5, and 6 of this report contain the PMA-specific results for BPA, WAPA, SWPA, and SEPA, respectively. The final sections of the report, Sections 7 and 8, provide an interregional comparison and discuss the summary and conclusions. The SWA Section 9505 calls for a DOE report to Congress, which will be produced separately from this assessment report but will be based on the details and conclusions presented herein. Although not explicitly presented in this assessment report, the PMA administrators' recommendations from each of the PMAs will be included in the subsequent DOE report to Congress.

[2] Firm power is an allocation of both capacity and energy.

2. ASSESSMENT APPROACH

2.1 SCOPE AND OBJECTIVES

Evaluating the large-scale climate change effects on all federal hydropower plants across the United States requires a series of models and methods to gradually downscale the GCM output into watershed-scale hydrologic and hydropower projections to support this climate change impact assessment. This assessment involves several major steps that are discussed in the following sections.

- Climate Model Evaluation and Selection (Section 2.3),
- Climate Model Downscaling (Section 2.4),
- Hydrologic Modeling (Section 2.5),
- Hydropower Simulation (Section 2.6), and
- Climate Change Impact Assessment (Sections 2.7 and 2.8).

Several recent hydroclimate studies have used a similar assessment framework with some variation (Chen et al., 2011; Bosshard et al., 2013; Meresa and Romanowicz, 2017; Pytlak et al., 2018; Chegwidden et al., 2019; Gangrade et al., 2020; Persad et al., 2020; Glabau et al., 2020; Heidari et al., 2020). Among these steps, GCM selection is often considered the greatest source of uncertainty and received most attention. Compared with the selection of GCMs, other methodological choices are often overlooked that should require further evaluation (Giuntoli et al., 2018). Some studies (e.g., Chen et al., 2011; Bosshard et al., 2013; Clark et al., 2016; Meresa and Romanowicz, 2017; Alder and Hostetler, 2019; Chegwidden et al., 2019; and Gangrade et al., 2020) have explored factors other than GCM selection and cautioned that these methodological choices may have a stronger influence than originally expected because they can be process/location specific and sometimes may control the direction of projected future hydroclimate conditions.

The previous 9505 assessments (Sale et al., 2012 and Kao et al., 2016) followed the conventional approach that only considered the GCM-driven projections as the source of uncertainty, using variability among different GCM-driven projections as a proxy for uncertainty. The results were then derived based on only one DS, one reference meteorological observation, one HM, and one PM. Although these models were calibrated based on the best available observations, it was unknown whether a different model, calibration, or downscaling approach would have led to a considerably different projection. This lack of clarity could lower the confidence of the assessment findings and lead to a biased interpretation of climate change–induced risks to future federal hydropower generation.

Overall, an ensemble assessment approach was required to better understand how the choice of downscaling, hydrologic, and hydropower modeling approaches affected the projection of future federal hydropower generation and improved the confidence of the assessment findings. This approach involved alternative models, data sets, and analytical methods. By creating an ensemble of future hydrologic and hydropower projections with varying controlling factors, a further statistical analysis was conducted to quantify the relative contribution of each factor. This further insight helped identify the key controlling processes and also suggested areas of improvement in future research and development efforts.

Based on extensive consultation with several federal hydropower stakeholder groups, a design for the ensemble simulation used in this assessment was developed and is illustrated in Figure 2.1. The ensemble included six selected CMIP6 GCM outputs, two DSs, two MetFs, two HMs, and two regional PMs. The alternative modeling components and supporting data sets were dissimilar by nature (e.g., dynamical vs. statistical downscaling) but were all commonly used in many hydroclimate studies. All selected models were intensively calibrated/validated based on the best available historic observations before being

incorporated into the assessment framework. When moving down the assessment chain, the size of the ensemble increased and eventually resulted in 96 sets of future hydropower generation projections. Although including additional models and factors in the ensemble mix would be a desirable option, it would also lead to an exponential increase in the ensemble size and was not feasible given the assessment timeline and resource limitations (e.g., computational resources, time constraints). The design of the ensemble simulation (Figure 2.1) represents the best effort to balance the comprehensiveness and feasibility of this assessment.

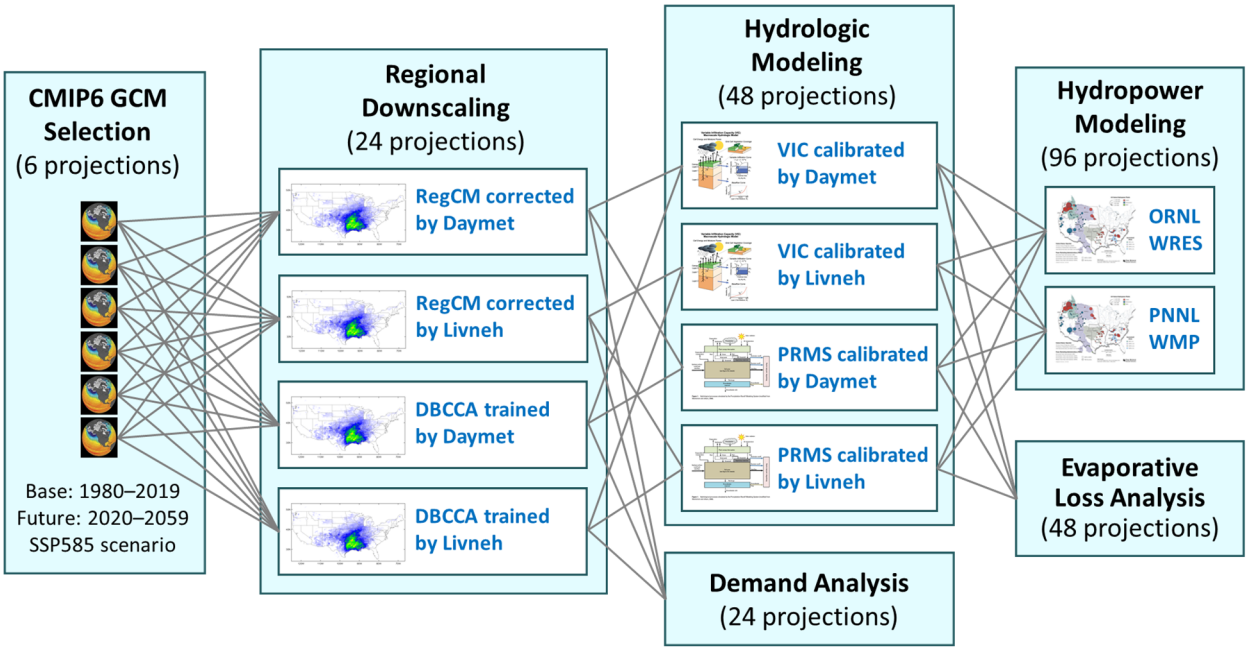


Figure 2.1. Multimodel ensemble simulation conducted for this assessment.

Overall, this study focused on a regional assessment of the 18 PMA study areas rather than individual reservoirs or power plants. This approach enabled a comprehensive evaluation of climate change impacts across many hydropower plants along distinct river systems. Regionally lumped models or generalized indices were used to evaluate the likelihood and relative spread of projected changes regarding future water availability, extreme high and low runoff events, hydropower generation, reservoir evaporation loss, and change in temperature-driven energy demand within each PMA study area. Site-specific features, such as reservoir operation rules, water withdrawal/return, environmental flow requirements, and energy generation, were not explicitly modeled at each power plant. This generalized approach allowed for spatial consistency throughout all study areas, helping policymakers evaluate and understand potential climate change impacts across the entire federal hydropower portfolio. If an issue of concern is identified for a specific region (e.g., change of streamflow seasonality), a site-specific evaluation can be planned through other, future local studies. The proposed assessment method does not replace the existing site-specific models and tools now used by water and energy resource managers. This study provides a first-order assessment to identify areas with the highest risk under projected climate conditions.

2.2 DATA SOURCES

To support model development and verification at various stages of the assessment, a variety of data and observations were assembled in this study (Table 2.1). These data sets included hydropower project characteristics and historical power generation data, meteorological and hydrologic observations, and other land surface information.

Table 2.1. Summary of data sources

Data type	Data source	Reference/website
Hydropower project characteristics	• Previous 9505 assessments • HydroSource • National Inventory of Dams (NID) • Global Reservoir and Dam Database (GRanD)	• Kao et al. (2016); Sale et al. (2012) • HydroSource, https://hydrosource.ornl.gov; Johnson et al. (2021) • NID, https://nid.sec.usace.army.mil • GRanD, http://globaldamwatch.org/grand
Historic hydropower generation	• PMA • EIA Form 923 Database	• Historic generation records provided by PMAs • EIA (2020), https://www.eia.gov/electricity/data/eia923
Electricity sales for PMA customers	• EIA Form 861 (Annual Electric Power Industry Report)	• https://www.eia.gov/electricity/data/eia861
Total retail load of BPA's customers	• BPA	• Personal communication with BPA staff
Meteorological observations	• Daymet • Livneh • Parameter-elevation Regressions on Independent Slopes Model • North American Regional Reanalysis wind speed	• https://daymet.ornl.gov; Thornton et al. (2021) • https://psl.noaa.gov/data/gridded/data.livneh.html; Livneh et al. (2015); Pierce et al. (2021) • https://prism.oregonstate.edu; Daly et al. (2002) • https://www.esrl.noaa.gov/psd/data/gridded/data.narr.html; Mesinger et al. (2006)
Hydrologic observations	• National Water Information System • Environment Canada HYDAT Database • WaterWatch • Bias Correction and Quality Control SNOTEL Data • Historical reservoir storage and area	• https://waterdata.usgs.gov/nwis • https://www.canada.ca/en/environment-climate-change/services/water-overview/quantity/monitoring/survey/data-products-services/national-archive-hydat.html • https://waterwatch.usgs.gov; Brakebill et al. (2011) • Sun et al. (2019); Yan et al. (2018) • Zhao and Gao (2018) and (2019)
Income per capita	• Bureau of Economic Analysis Annual Personal Income by County (CAINC1) series	• https://apps.bea.gov/regional/downloadzip.cfm
Air conditioning ownership in the Pacific Northwest	• Northwest Energy Efficiency Alliance Residential Building Stock Assessment II	• https://neea.org/resources/rbsa-ii-combined-database
City boundary	• US Census Bureau's Topologically Integrated Geographic Encoding and Referencing data set	• https://www.census.gov/geo/maps-data/data/tiger.html

2.2.1 Hydropower and Power Marketing Data

Hydropower project characteristics: The list and characteristics of the 132 federal hydropower plants were established during previous 9505 assessments (Kao et al., 2016 and Sale et al., 2012). Some recent changes in capacity were found in the Existing Hydropower Assets Plant data set (Johnson et al., 2021) maintained by the ORNL HydroSource program. These project characteristics were further reviewed and updated by PMA staff. Some additional reservoir-related information used in hydropower modeling and

reservoir evaluation calculation were collected from National Inventory of Dams (NID) and the Global Reservoir and Dam Database (GRanD) (Lehner et al., 2011).

Historical hydropower generation: The 1980–2019 historical monthly hydropower generation data for the United States were collected from the US Energy Information Administration's (EIA's) Form 923 database (EIA, 2020). When available, more accurate generation records provided by Reclamation, USACE, and the PMAs were used to update parts of the survey-based EIA plant generation data. The monthly plant generation data were then aggregated for each PMA study area for further analysis.

Electricity sales and revenue of PMA preference customers: Data on the electricity sales and revenue generated by PMA preference customers selling to their end-use customers (residential, commercial, industrial, or other) were collected from EIA's Annual Electric Power Industry Report (Form 861). This form collects information on peak load, generation, electric purchases, sales, revenues, customer counts, and other metrics directly from the utilities. The subset of utilities of interest for this analysis were those that have long-term contracts with one of the PMAs (i.e., preference customers).

2.2.2 Meteorological Data

Daymet: The ORNL Daymet meteorological data provides daily, gauge-based gridded precipitation estimates from 1980 to 2019 for North America at a 1 km horizontal resolution (Thornton et al., 2021). Daymet is widely used in many hydrologic and ecological studies and offers the highest spatial resolution among all publicly available, gauge-based gridded precipitation data sets in the United States. This study used the latest Daymet V4, which includes observational timing correction for daily precipitation and maximum temperature, along with other enhancements. Daymet was aggregated to 1° resolution grids for the climate model (CM) evaluation, to 1/24° (~4 km) resolution grids for the CM downscaling and Variable Infiltration Capacity (VIC) hydrologic modeling, and to various hydrologic response units (HRUs) for the Precipitation Runoff Modeling System (PRMS) modeling. When compared with Daymet V3, the differences are mainly at the daily scale. The monthly values between both versions are quite similar. The monthly scale evaluation of precipitation timeseries revealed that 99.7% of HUCs showed an NSE >0.90, and ~91% of HUCs showed an NSE >0.95. The monthly scale evaluation of temperature time series demonstrated that all HUCs showed an NSE >0.95, highlighting the similarities between the two versions of the data set.

Livneh: The Livneh daily near-surface gridded meteorological data offers daily, gauge-based gridded precipitation estimates from 1950 to 2013 for the conterminous United States (CONUS), Mexico, and the part of Canada south of 53° N at a 1/16° (~6 km) horizontal resolution (Livneh et al., 2015). It is also widely used in many hydrologic studies and has served as the underlying meteorological training data set for several statistically downscaled CMIP5 products based on the bias-corrected spatial disaggregation (BCSD; Brekke et al., 2013), multivariate adaptive constructed analogs (MACA; Abatzoglou, 2013), and localized constructed analogs (LOCA; Pierce et al., 2014) methods.

This study used the updated 1950–2018 Livneh data set provided by Pierce et al. (2021), which will also be used as the underlying meteorological training data set for the upcoming CMIP6-based LOCA climate projections. This update was mainly to correct a timing adjustment issue and improve the characterization of precipitation extremes. A similar process was followed to aggregate/interpolate Livneh to 1° and 1/24° grids and HRU spatial units. The differences caused by the Pierce et al. (2021) timing adjustment are mainly at the daily scale and not so much at the monthly scale, similar to Daymet. A monthly scale comparison of precipitation between the two Livneh data set versions revealed that ~96% of HUCs showed an NSE >0.90, and ~86% of HUCs showed an NSE >0.95. The monthly scale evaluation of temperature time series demonstrated that all HUCs showed an NSE >0.95, highlighting the similarities between the two versions of data sets.

PRISM: The Parameter-Elevation Regressions on Independent Slopes Model (PRISM) (Daly et al., 2002) meteorological data set provides gridded observation of precipitation and temperature for the entire CONUS at 1/24° (~4 km) spatial resolution. Although it is a widely used data set, it was not used in this study to support hydrologic simulation and downscaling since it does not cover the entire spatial domain of the study areas (i.e., the headwater basins in Canada and Mexico). PRISM was aggregated to 1° grids only for the purpose of CM evaluation.

NARR (wind speed): Because Daymet does not provide wind speed (needed for VIC hydrologic simulation), the daily wind speed data were collected from the North American Regional Reanalysis (NARR; Mesinger et al., 2006). NARR is an assimilated meteorological reanalysis data set that provides a complete set of meteorological variables (e.g., pressure, wind) for North America. These data are available at 3 h time steps from 1979 to the present at a 36 km horizontal grid spacing. The daily wind speed is spatially interpolated to 1/24° to support VIC hydrologic simulation and downscaling.

2.2.3 Hydrologic Data

Streamflow: Gauge-based streamflow observations directly measure flow discharge at a specific river channel. Comprehensive daily flow observations can be obtained from the USGS National Water Information System (NWIS) for more than 22,000 current and retired gauge stations throughout the United States. Additional streamflow observations can also be obtained from Environment Canada's HYDAT Database for gauges located in Canada's Upper Columbia River basin (part of the BPA-1 and BPA-3 study areas).

Runoff: The USGS WaterWatch runoff (Brakebill et al., 2011) was used as a surrogate of historic runoff observation in this study. Derived from the comprehensive NWIS gauge observation, WaterWatch runoff is the assimilated time series of streamflow per unit of area calculated for each CONUS eight-digit hydrologic unit code (HUC8). Multiple NWIS gauge stations located within HUC8s or downstream are used to estimate the runoff generated locally at each HUC8, with gauge weighting factors determined by joint contributing drainage areas (both gauge-to-HUC8 and HUC8-to-gauge). This approach can effectively assimilate streamflow observations from multiple gauge stations as a consistent areal HUC8 runoff measurement with a unit similar to that for precipitation (depth/time).

WaterWatch runoff has been used and discussed in several hydroclimate studies, including by Ashfaq et al. (2013), Beigi and Tsai (2014), and Oubeidillah et al. (2014). WaterWatch runoff is also the most significant variable in explaining historic hydropower generation (Sale et al. 2012; Kao et al. 2016). Because WaterWatch only covers the CONUS, the Canadian runoff was computed following a similar approach to WaterWatch using the unregulated monthly flow collected from HYDAT. The HUC8-based runoff was then weight averaged (by the contributing areas) to estimate the 1980–2018 monthly runoff time series for each of the 18 PMA study areas. Both NWIS and WaterWatch could include gauges affected by river regulation, rather than naturalized streamflow. Further efforts can be performed to exclude the regulated streamflow gauges and/or incorporate naturalized streamflow information for the correction of human impairment in the WaterWatch for more accurate HM calibration.

Snow: Snow water equivalent (SWE) data were obtained from Bias Correction and Quality Control SNOTEL data for 829 active stations located in the western United States and Alaska (Sun et al., 2019; Yan et al., 2018). The annual April 1 SWE was used for HM verification (Section 2.5).

Reservoir storage and area: The historical reservoir storage and area were collected from USGS NWIS, Reclamation, USACE, and by fusing remote sensing imageries for historical reconstruction (Zhao and Gao, 2018 and 2019). The data were used to help estimate reservoir evaporation for all major federal hydropower reservoirs (Section 2.7).

2.2.4 Additional Data

Additional data used in this study included city boundaries from US Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) data set, income per capita from Bureau of Economic Analysis Annual Personal Income by County (CAINC1) series, and air conditioning ownership in the Pacific Northwest. These data are also reported in Table 2.1.

2.3 CLIMATE MODEL EVALUATION AND SELECTION

GCMs are physics-based tools to study Earth system response to natural climate variability and changes arising from climatic changes arising from anthropogenically driven increases in greenhouse gas emission and radiative forcing. Using a common set of future radiative pathways, the CMIPs (Eyring et al., 2016) provide a large suite of GCM simulations through an international collaborative effort. Since the inception of CMIP in 1995, not only have the number of GCMs participating in CMIP efforts increased, but they have also improved in terms of their physical complexity and spatial resolution. Every new iteration of CMIP is based on the premise that the newer generations of GCMs will exhibit improvements over the previous ones as models progressively improve in terms of their computational efficiency, resolution, and representation of physical processes. The latest CMIP6 includes more than 50 GCMs with the spatial resolution for some of them as small as 0.5°. GCMs are forced with a new set of nine scenarios, which are a combination of five Shared Socioeconomic Pathways (SSPs) and seven Representative Concentration Pathways (RCPs), to understand the Earth system response to an increase in anthropogenic forcing. Among them, there are four high-priority or Tier 1 scenarios, including SSP1-2.6 (SSP126), SSP2-4.5 (SSP245), SSP3-7.0 (SSP370), and SSP5-8.5 (SSP585). In each scenario, the first number represents one out of five SSPs, and the last two numbers represent the radiative forcing level. These are among the plausible combinations of SSPs and RCPs in which the underlying socioeconomic factors in an SSP allow for the mitigation necessary to meet RCP targets. For instance, the RCP8.5—the highest forcing scenario—can only be consistent with the SSP5 socioeconomic trajectory, which represents high challenges to mitigation and adaptation. Alternatively, the SSP1 pathway that represents low challenges to mitigation and adaption would not lead to higher emission scenarios such as RCP7.0 and RCP8.5 (O'Neill et al., 2016; Zelinka et al., 2020).

The current assessment was limited to only one future emission scenario for several reasons. First, for the time frame of interest (i.e., up to the mid-twenty-first century to inform long-term power marketing evaluation), higher-end scenarios (SSP370 and SSP585) are closer to the current observed trajectory of global greenhouse gas emissions. Second, at the time of the GCM selection, the availability of 6-hourly 3D atmospheric fields needed for dynamical downscaling was only limited to a few Tier 1 scenarios with a maximum (14) under SSP585. Third, dynamical downscaling is computationally expensive and time-consuming. Availability of CMIP6 GCMs data was delayed by more than a year, which left little time to consider more than one emission scenario in this study. Fourth, precipitation and temperature changes tend to be proportional across the future scenarios, meaning that stronger changes are expected in higher radiative forcing scenarios and vice versa, which is evident from the seasonal comparison of their changes in the SSP370 and SSP585 scenarios of the CONUS for 2041–2060 with respect to 1995–2014 (Figure 2.2 and Figure 2.3). The spatial pattern of changes is remarkably similar between the two scenarios. Expectedly, the magnitude of change is stronger in SSP585 because it represents a stronger forcing pathway. Given that SSP370 reaches the 2060 radiative forcing level of SSP585 around 10 years later (~2070), the projected changes in SSP585 can be considered a proxy for expected changes under SSP370 a decade later. Considering these factors, SSP585 was selected for this climate change impact assessment.

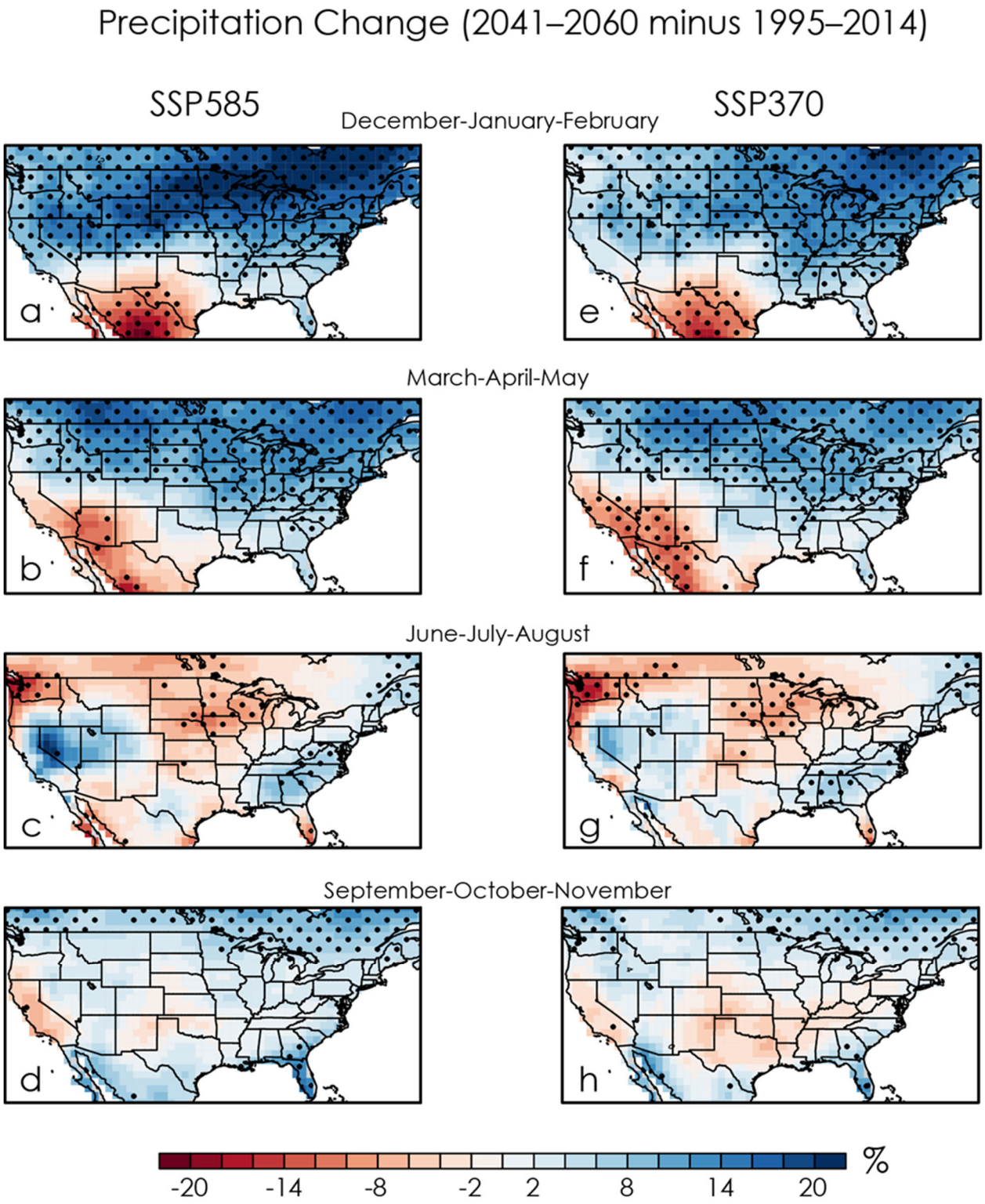


Figure 2.2. Projected changes (%) in seasonal precipitation in 2041–2060 compared with 1995–2014 (a–d) SSP585 and (e–h) SSP370. The stippling represents grid points where two-thirds of the models exhibit the same sign of change. This analysis is based on 28 GCMs that had data available for both scenarios.

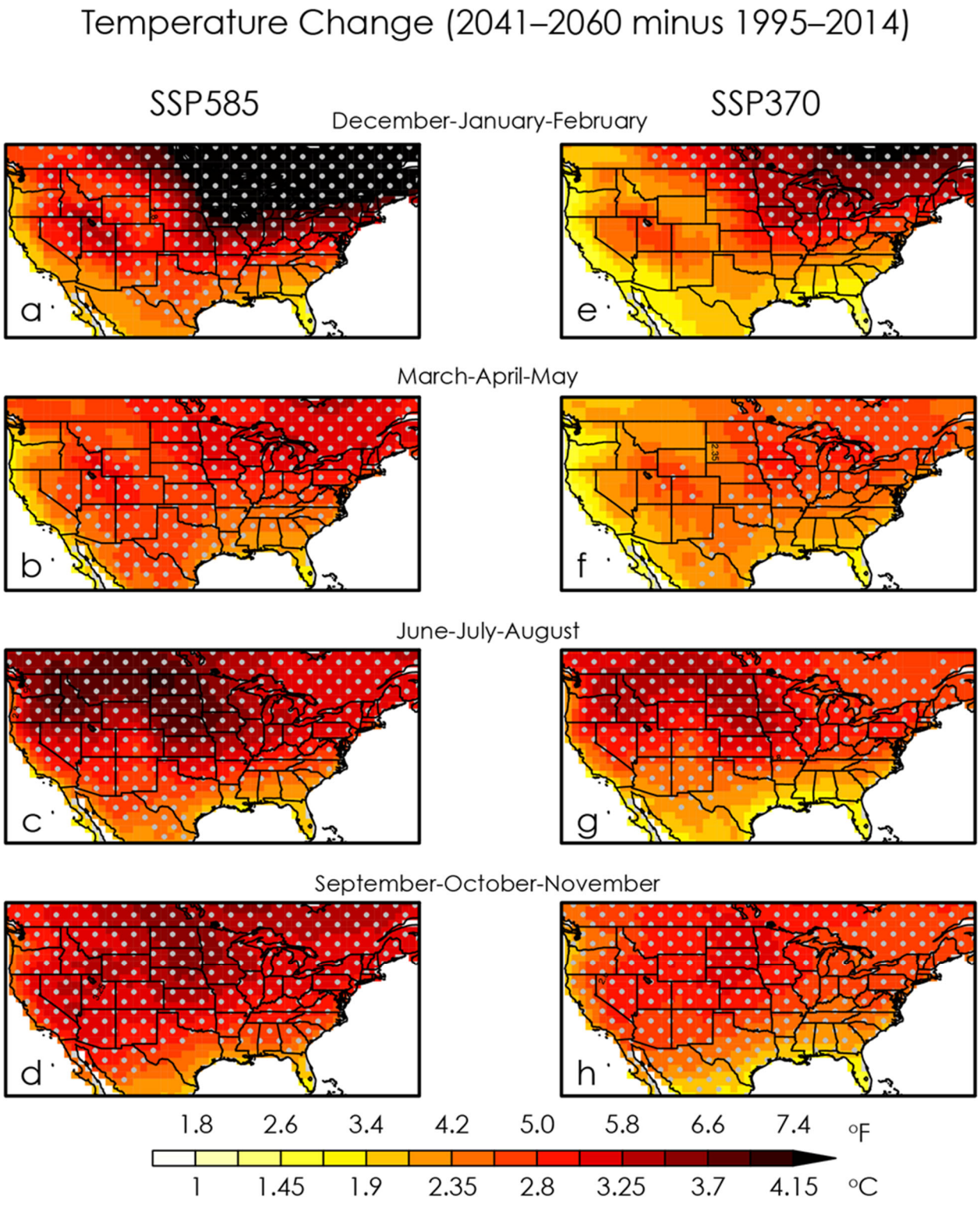


Figure 2.3. Projected changes (°C) in seasonal temperature in 2041–2060 compared with 1995–2014 in SSP585 (a–d) and SSP370 (e–h). The stippling represents grid points where two-thirds of the models exhibit at least 2°C (3.6 °F) of warming. This analysis is based on 28 GCMs that had data available for both scenarios.

Despite the substantial improvements in GCMs, several challenges still exist that limit the direct application of GCM projections in regionalized, higher-resolution climate change impact assessments. First, the resolution of most CMIP6 GCMs is still insufficient (>1° horizontal resolution) to reliably assess the needs for mitigation or adaptation at policy relevant regional and local scales, which warrants the need for downscaling (discussed in Section 2.4). Second, the lack of sufficient resolution and other modeling deficiencies also leads to inaccuracies in the representation of fine-scale, land-atmosphere feedbacks that may influence the simulated climate response at such scales (e.g., Diffenbaugh et al., 2005). Third, several models participating in CMIP6 share common modeling components (e.g., same land, ocean, ice modules or parametrization), meaning that these models may have similar systematic biases and do not necessarily represent independent realizations of future climate (Knutti et al., 2010 and 2013). Therefore, an objective evaluation standard based on the models' skillfulness and independence is recommended for more unbiased estimates of future climate change (Eyring et al., 2019).

Apart from GCM performance, data availability for GCM downscaling is another limiting factor in completing hydroclimate studies in a timely manner. Although GCM simulations are conducted at sub-hourly time scales, given their massive data flow, only a subset of variables at aggregated temporal scales were recorded (usually driven by the specific CMIP requirements). For some types of applications, such as dynamical downscaling, the required sub-daily lateral boundary forcings are not available for all GCMs in CMIP6. Given the desire to compare these two very different downscaling approaches (dynamical vs. statistical) in this assessment, the selection of GCMs was limited to those with sufficient sub-daily outputs for dynamical downscaling.

Considering the available resources and the objective of multimodel ensemble simulation (discussed in Section 2.1), six GCMs were selected through an objective selection process that factored in relative simulation skills, uniqueness, data availability, computational resources, and time constraints. GCM evaluation is based on two approaches. One is a simple weighting technique in which each GCM is ranked based on its average performance across selected evaluation metrics. Often, correlations among the suite of evaluation metrics used to investigate the relative skillfulness of the model can influence the ranking process. Therefore, each metric is assigned a weight based on its uniqueness ranging from zero to one before averaging. The second is a ranking based on empirical orthogonal functions (EOFs) in which each GCM is ranked based on its distance from the reference (observations) in the principal component (PC) space (Chhin et al., 2019; Rupp et al., 2013; Sanderson et al., 2015). Using two different approaches for model ranking ensures the robustness of the outcome.

At the time the GCM evaluation was conducted (in early 2020), 37 CMIP6 GCMs provided model output for the evaluation, but only 14 out of those 37 GCMs provided 6-hourly, 3D atmospheric fields needed for dynamical downscaling. Thus, although 37 models were evaluated, to better understand the relative skill of the selected GCMs in the large CMIP6 suite, the final selection was limited to the 14 GCMs.

For each of the 37 GCMs, daily and monthly precipitation; mean, maximum, and minimum daily temperature; monthly sea surface temperature; sea level pressure; and 500 mb geopotential height were collected from the Earth System Grid Federation data archive.[3] To support model evaluation, the precipitation and temperature observations from Daymet, Livneh, and PRISM were used along with the sea surface temperature, sea level pressure, and 500 mb geopotential height from the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5).

For GCM evaluation, the entire CONUS was divided into four parts (north, east, west, and south) based on grouped HUC2 regions (see APPENDIX C) used by Naz et al. (2016). In total, 60 evaluation matrices at the annual, seasonal, monthly, and daily and diurnal time scales were used to rank the models (see

[3] Source: https://esgf-node.llnl.gov/search/cmip6

APPENDIX C). All metrics calculations and further analysis were separately carried out for each of the four regions, which were subsequently averaged to calculate disagreements at the CONUS scale for each model. Observation-based matrices were subsequently averaged to create a reference data set for model comparison. The selection of metrics represented a wide range of spatiotemporal climate characteristics that are common across the CONUS and did not include features that were unique to specific regions, such as integrated water vapor transport through atmospheric rivers in the western United States, the monsoonal climate in the southwestern United States, and the tornadic environment in the central and eastern United States. For each of the 60 indices, resulting disagreement from the model with the reference data was represented as the percentage difference, which was subsequently converted into normalized relative error. A single matrix represented a single index or an average of three or more indices.

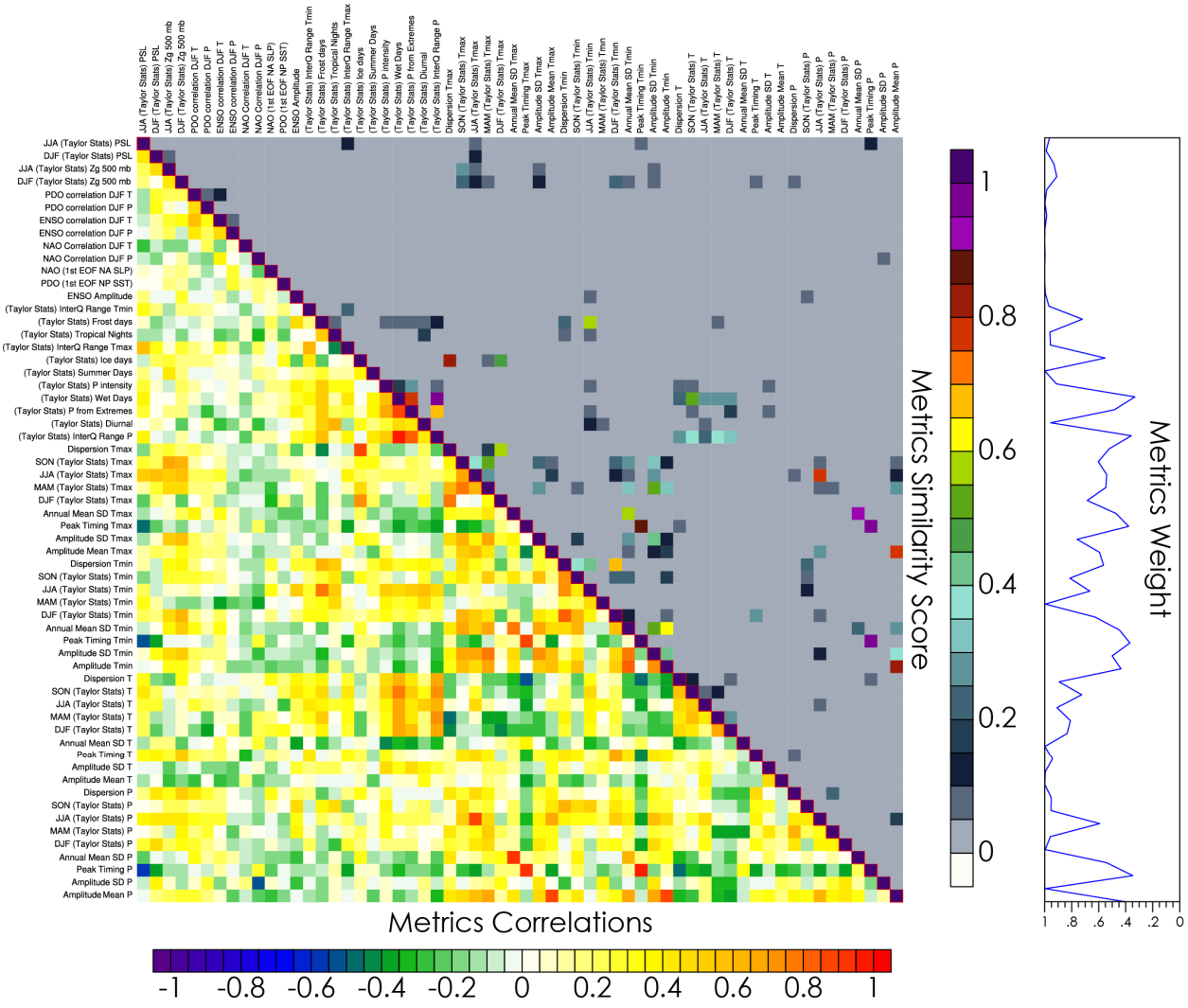


Figure 2.4. The pairwise metrics correlation (bottom triangle) and the corresponding similarity score (top triangle) over the northern United States. Metrics with high correlations exhibit a high similarity score and are down-weighted. The line plot on the right shows the overall weight for each metric.

In the simple weighting approach, pairwise linear correlations were calculated for all metrics to calculate their similarity scores and overall weights. The metrics with higher correlations were assigned lower weights and vice versa (Figure 2.4). The GCM with the lowest weighted relative error ranked at the top,

whereas the GCM with the highest weighted relative error ranked at the bottom (see APPENDIX C). In the EOF analyses, the Euclidean distances between individual GCMs and observations were calculated using the truncated set of first 10 modes because GCMs ranking in the PC space became relatively insensitive after first 10 PCs (see APPENDIX C). Using the sum of the distances of GCMs from the reference data in each selected PC space, the GCM with the lowest total distance ranked at the top, whereas the GCM with the highest total distance ranked at the bottom.

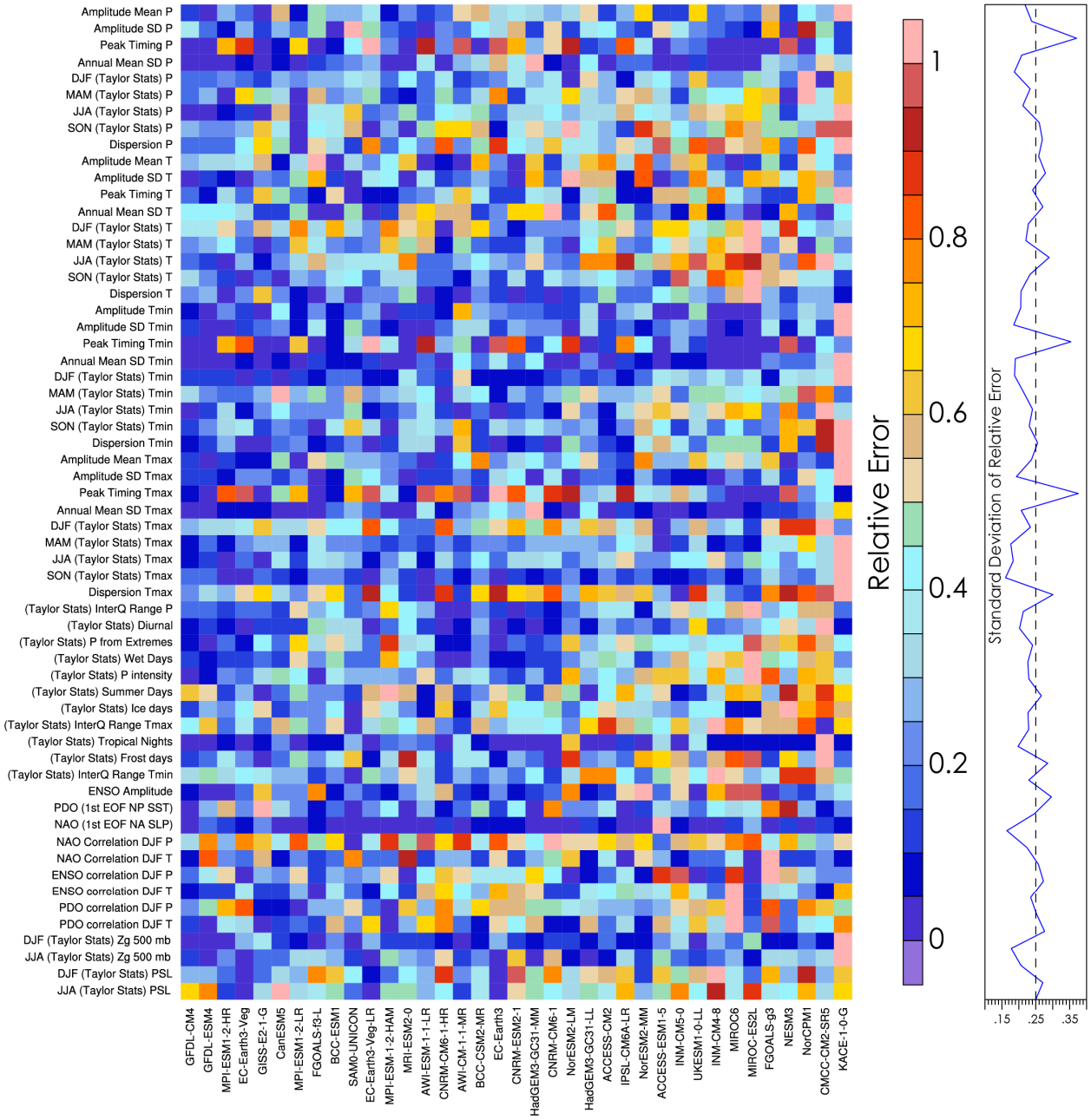


Figure 2.5. The normalized relative disagreement of GCMs from the observations over the northern United States for each metric (left) and their overall standard deviation across evaluated GCMs (right).

Figure 2.5 shows the unweighted normalized relative errors of the GCMs for each matrix over the northern United States. The rest of the three regions are shown in APPENDIX C. The normalization was

based on the minimum and maximum for each individual matrix so that the lowest difference from the reference data transformed into zero, and the highest difference from the reference data transformed into one. For ease of comparison, GCMs were sorted from left to right so that the GCM with the lowest average relative error was on the left and the one with the highest average relative error was on the right. Unlike the absolute error, the relative error was not a direct measure of modeling biases with respect to truth or observations, as it only differentiated models from each other. Nevertheless, models with higher magnitudes of relative error were further away from the observations than those with lower magnitudes. The line plot panel on the right displays the standard deviation of relative errors across GCMs for each metric. If the performance of many models fell in a similar category, their relative errors displayed a similar range of colors. High standard deviation magnitudes represented substantial variation in modeling skills across the GCMs and vice versa. Overall, many GCMs exhibited challenges in simulating key climate characteristics. For instance, although the models were relatively skillful in representing oceanic and atmospheric patterns associated with natural forcing (ENSO, NAO, PDO), most showed limited skill in simulating their influences on the distribution of seasonal mean precipitation and temperature over the southern and western United States. Difficulties in reproducing the observed timing of peak magnitudes of precipitation, minimum temperature, and maximum temperature were also evident in the western and northern United States, and metrics for precipitation characteristics were relatively poorly simulated in the southern United States. One noticeable distinction between poor and better performing models was that the latter group was deficient in reproducing several characteristics of daily-scale temperature and precipitation across all regions (see APPENDIX C).

The 14 candidate GCMs were not all unique, and several of them belonged to the same institute and did not differ substantially in their main physical configurations. These included two models from the Max Planck Institute in Germany (MPI-ESM1-2-HR and MPI-ESM1-2-LR), three from the National Center for Meteorological Research in France (CNRM-CM6-1, CNRM-CM6-1-HR, and CNRM-ESM2-1), two from the Japan Agency for Marine-Earth Science and Technology (MIROC6 and MIROC-ES2L), and two from the Norwegian Meteorological Institute (NorESM2-LM and NorESM2-MM). Generally, GCMs from the same institute tended to exhibit a similar nature of biases (Knutti et al., 2013; Sanderson et al., 2017), which also influenced their simulated hydro-meteorological responses. This correspondence among the models from the same institute is evident in the PC1 vs. PC2 plot (see APPENDIX C) because models from the same institute tended to be the same distance from the reference data in the PC space. These first two PCs explained 39.5% of the variance in the GCM evaluation metrics (shown in Figure 2.5). To avoid the potential systematic bias due to similar model structures and parameterizations, no more than one model from each institute was allowed. Therefore, the first six models in the ranking that exhibited relatively good skill and were from unique modeling institutes were selected. Several models from different institutes may also have common land, ocean, or atmospheric components; however, given the limited number of candidate models for dynamical downscaling, model uniqueness across the institutes was not considered.

Figure 2.6 and shows the GCMs ranked using a simple weighting technique. The GCM ranking using the EOF analyses is shown in APPENDIX C. The two approaches yielded reasonably similar results at the CONUS scale; not only were the first and fourth quartiles occupied by the same GCMs in both approaches, but the individual GCMs placements within these quartiles were also very similar. For instance, the bottom five GCMs rankings were identical in both cases, and the maximum difference in ranking in the fourth quartile ranged from 0 to 2. More importantly, both approaches provided the same selection of the top six models, which were better skilled and belonged to unique institutes. This highlights the relative insensitivity of the selection process to the choice of ranking methodology. These six models and their rankings in the simple weighting are listed as follows:

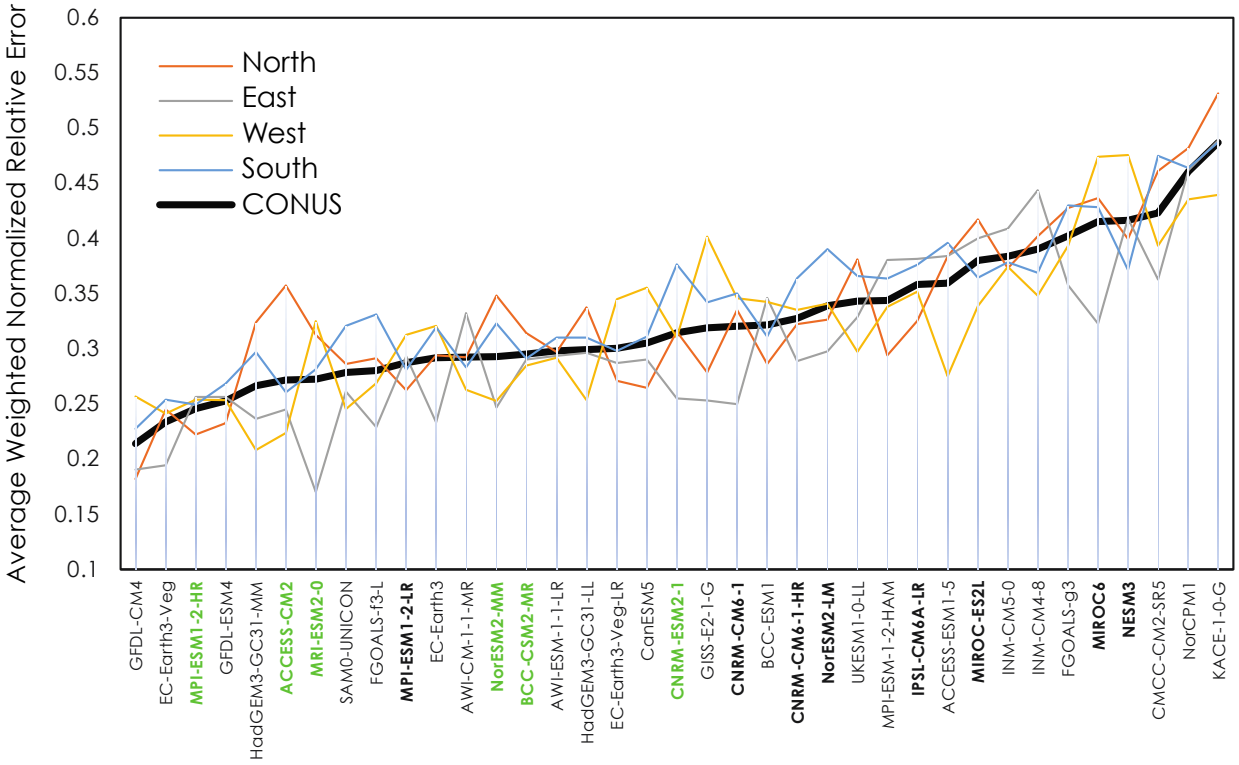


Figure 2.6. The ranking of GCMs using the simple weighting approach. The bold font represents GCMs that have data available for dynamical downscaling and green represents the selected GCMs.

(1) MPI-ESM1-2-HR ranked at 3,
(2) ACCESS-CM2 ranked at 6,
(3) MRI-ESM2-0 ranked at 7,
(4) NorESM2-MM ranked at 13,
(5) BCC-CSM2-MR ranked at 14, and
(6) CNRM-ESM2-1 ranked at 19.

Overall, six CMIP6 GCMs were selected for this study (Table 2.2), including the Australian Community Climate and Earth System Simulator Coupled Climate Model (ACCESS-CM2), the Beijing Climate Center Climate System Model running on medium resolution (BCC-CSM2-MR), the National Centre for Meteorological Research Earth System Model of second generation (CNRM-ESM2-1), the Max Planck Institute Earth System Model run at high resolution (MPI-ESM1-2-HR), the Meteorological Research Institute Earth System Model version 2.0 (MRI-ESM2-0), and the Norwegian Earth System Model (NorESM2-MM). The general consensus regarding the equilibrium climate sensitivity was within 1.5–4.5 K. However, several models in CMIP6 showed much higher (>4.5 K) sensitivity because of stronger positive cloud feedbacks with the overall values ranging from 1.8 to 5.6 (Zelinka et al., 2020). The equilibrium climate sensitivity in the selected six models spanned 2.49 to 4.79, showing no bias particular toward the low or the high sensitivity (Table 2.2). For each GCM, the raw 1980–2059 GCM outputs were collected from the CMIP6 archive, where 1980–2014 was under the historic greenhouse gas emission, and 2015–2059 was under the SSP585 emission pathway. The data from these six GCMs were used for further downscaling and simulation.

Table 2.2. Six selected CMIP6 models

No.	GCM name	Spatial resolution (latitude/longitude)	Emission scenario	Ensemble number	Equilibrium Climate Sensitivity
1	ACCESS-CM2	1.25°/1.875°	SSP585	r1i1p1f1	4.66 K (2×CO_2)
2	BCC-CSM2-MR	1.125°/1.125°	SSP585	r1i1p1f1	3.02 K (2×CO_2)
3	CNRM-ESM2-1	1.41°/1.41°	SSP585	r1i1p1f2	4.79 K (2×CO_2)
4	MPI-ESM1-2-HR	0.94°/0.94°	SSP585	r1i1p1f1	2.98 K (2×CO_2)
5	MRI-ESM2-0	1.125°/1.125°	SSP585	r1i1p1f1	3.13 K (2×CO_2)
6	NorESM2-MM	0.94°/1.25°	SSP585	r1i1p1f1	2.49 K (2×CO_2)

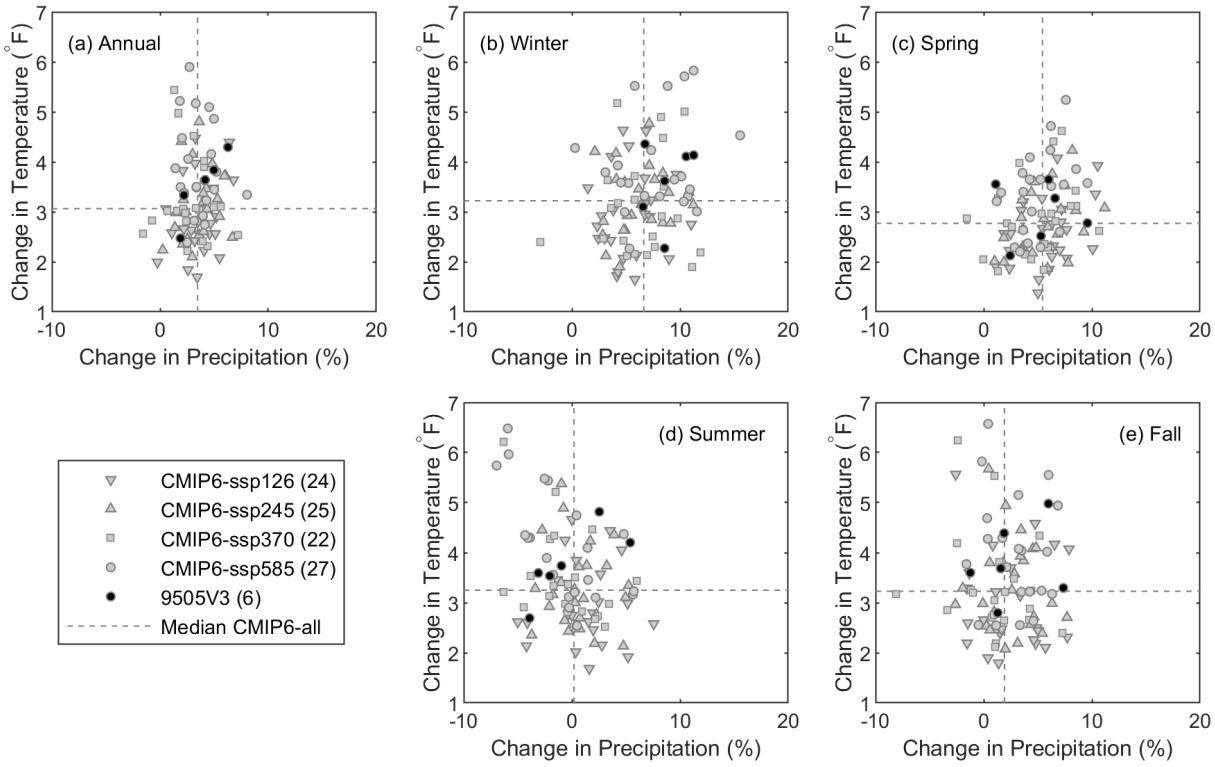


Figure 2.7. Annual and seasonal CONUS temperature and precipitation changes for 98 CMIP6 GCM projections under four emission scenarios (SSP126, SSP245, SSP370, and SSP585). For each GCM projection, the average changes are calculated from the 1980–2019 baseline to the 2020–2059 future period.

Figure 2.7 shows how these six GCM projections compared with other CMIP6 models and emission scenarios. In total, 98 GCM projections—with data available for download (including the selected six) under four emission scenarios (SSP126, SSP245, SSP370, and SSP585)—were included in this analysis. For each GCM projection, the average changes in precipitation and temperature from the 1980–2019 baseline to the 2020–2059 future period for the entire CONUS were calculated for annual periods and for four seasons—winter (December-January-February [DJF]), spring (March-April-May [MAM]), summer (June-July-August [JJA]), and fall (September-October-November [SON]). Overall, all projections showed a consistent increase in temperature ranging from 1°F to 6°F across all seasons. For precipitation, except for parts of summer and fall, a consistent increasing signal was projected (especially in winter), resulting in a net annual precipitation increase across most of the projections (−2% to 8%). Although the

six selected GCM projections did not cover the full range of CMIP6 models and scenarios, their projections were distributed across the median of all models, suggesting that they were not biased toward any one direction. Also, although the emission scenario had a clear influence on the projected changes in temperature, its influence on the projected changes in precipitation was less obvious at the CONUS scale for the time frame of interest (up to 2059). Given the limited resources, we hence only examined the SSP585 in this study. Further details about climate model evaluation can be referred to Ashfaq et al. (2022).

2.4 CLIMATE MODEL DOWNSCALING

Evaluation of regional-scale, hydroclimate impacts requires high-resolution climate projections (Ashfaq et al., 2016). However, the average horizontal grid spacing of CMIP6 is still coarser than 1º (~110 km), which creates a need for further spatial refinement. Various downscaling techniques, including dynamical downscaling, statistical downscaling, and a combination of the two can be used to translate the coarse resolution information from the GCMs to fine spatial scales suitable for regional analyses (Ashfaq et al., 2016; Maurer et al., 2013; Wood et al., 2004). In dynamical downscaling, GCM output is used as initial and boundary conditions to run a regional climate model (RCM) simulation. The RCM independently resolves fine-scale responses, such as those caused by complex topography, coastal lines, and mesoscale convection, which are often poorly represented in the GCMs. However, RCM simulations are computationally intensive and time-consuming. On the other hand, statistical downscaling is fundamentally based on the empirical relationships between the simulated and observed climate in the historical period. It is computationally efficient and faster but often limited to only a few variables for which long-term observational records are available (Nikulin et al., 2018; Yoon et. al., 2012). Often, a combination of the two DSs is employed, also known as *hybrid downscaling*, in which GCMs are first dynamically downscaled to a finer resolution and then statistically corrected at a desired horizontal grid spacing.

In this assessment, both dynamical and statistical DSs were applied to the six selected CMIP6 GCMs (Table 2.2). Additionally, statistical correction was performed on the dynamically downscaled model output. Therefore, this technique can be considered a hybrid DS. The DSs are detailed in Section 2.4.1 and 2.4.2. Two different MetFs (reference meteorological forcings) were used in this study to account for the uncertainties that arise from the choice of MetFs (e.g., see Alder and Hostetler, 2019). These MetFs served as the ground truth in bias-correcting dynamically downscaled outputs or as the training data sets in performing statistical downscaling. Consequently, 24 sets of downscaled climate projections (6 GCMs × 2 DSs × 2 MetFs) were generated. Each set of the downscaled climate projections extended from 1980–2059, in which 1980–2019 is defined as the historic baseline period, and 2020–2059 is defined as the future period.[4] Although a similar approach has been used at regional scales (Jiang et al., 2018), to the best of the authors' knowledge, this is the first time this approach has been used for CONUS at such a high resolution (~4 km).

2.4.1 Regional Climate Model Version 4

Regional Climate Model version 4 (RegCM4), developed and maintained by the Abdus Salam International Center for Theoretical Physics (Giorgi et al., 2012), was applied to dynamically downscale the six selected CMIP6 GCMs. RegCM4 has been widely used for downscaling over the CONUS, and its performance has been extensively evaluated (Ashfaq et al., 2010, 2013 and 2016; Diffenbaugh and Ashfaq, 2010; Diffenbaugh et al., 2005 and 2011; Walker and Diffenbaugh, 2009). In the current configuration, the RegCM4 grid was centered at 39.00°N and 98.00°W and consisted of 157 points in the

[4] The definition of historical baseline period (1980–2019) is based on the maximum length of Daymet observation. It should not be confused with the historic period of CMIP6 (1870–2014) when the historic greenhouse gas emission was used to drive GCM simulations.

latitude direction and 227 points in the longitude direction, thereby covering the CONUS and parts of Mexico and Canada (Figure 2.8). The grid points were separated at a 25 km horizontal spacing with 18 levels in the vertical direction. The Lambert Conformal Projection placed the grid corners at 50.91°N, 138.74°W (northwest); 50.91°N, 57.26°W (northeast); 18.13°N, 123.86°W (southwest); and 18.13°N, 72.14°W (southeast). The RegCM4 simulation uses the hydrostatic dynamical core from the Fifth Generation Mesoscale Model (Grell et al., 1994), the radiation package from the Community Climate Model version 3 (Kiehl et al.,1998), the Community Land Model version 4.5 (Tawfik and Steiner, 2011), the Holtslag boundary layer package (Holtslag et al., 1990), the Subgrid Explicit Moisture Scheme (Pal et al., 2000), and the Tiedke cumulus convection parameterization (Tiedtke, 1989). Furthermore, daily scale bias correction was applied to the daily scale precipitation (P), maximum temperature (T_{max}), and minimum temperature (T_{min}) from each of the RCM simulations using both Livneh and Daymet meteorological observations (discussed in Section 2.2) at the 1/24° (~4 km) spatial resolution grid. The daily bias correction method is also one step in the statistical downscaling that is discussed in the following section.

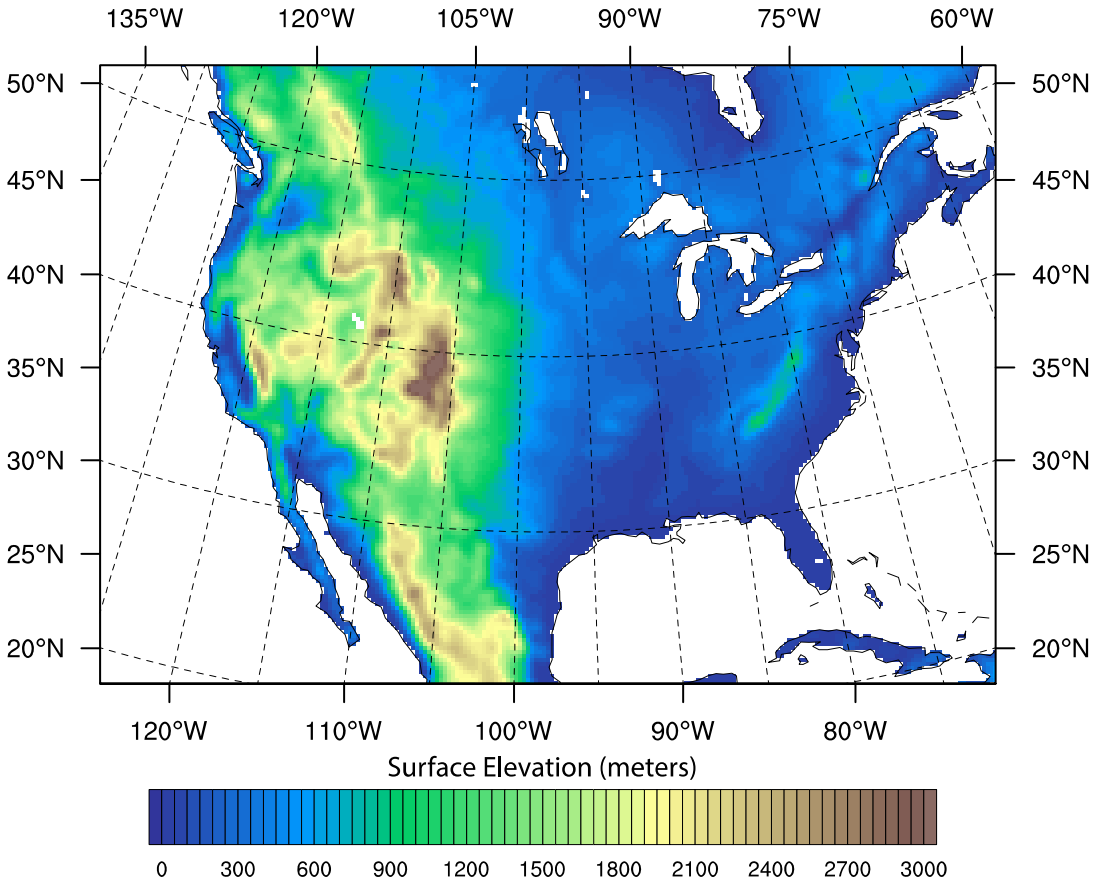


Figure 2.8. Domain and topography used for RCM simulations in this study.

2.4.2 Double Bias Correction Constructed Analogues

For statistical downscaling, the Double Bias Correction Constructed Analogues (DBCCA) method (Werner and Cannon, 2016) was applied to P, T_{max}, and T_{min} from the six selected CMIP6 GCMs using both Daymet and Livneh as the training data sets. DBCCA is based on the widely used Bias Correction Constructed Analogues (BCCA) method developed by Maurer et al. (2010). Given the use of weighted linear combinations of analogues, BCCA may result in extended periods of tiny-to-near-zero precipitation

that artificially increases the wet days, known as the *drizzling problem*. DBCCA introduces another round of bias correction at the end to fix this drizzling problem and to reduce residual biases. Werner and Cannon (2016) conducted a comprehensive evaluation of several statistical DSs and reported that DBCCA was one of the two best performing methods in their case study.

All CMIP6 GCMs were remapped to a common 1° latitude-longitude grid before statistical downscaling. Similarly, both Daymet and Livneh were remapped to a common 1/24° (~4 km) latitude-longitude grid as the targeted resolution of downscaling. DBCCA includes four main steps, in which steps 1–2 are conducted at the 1° grid, and steps 3–4 are conducted at the 1/24° grid resolution.

(1) *First bias correction* (at 1° grid resolution): The observations are first aggregated to the same 1° latitude-longitude GCM grid. For each calendar day and each GCM grid, a 31-day (±15 days) window is used to create the cumulative distributions across all years from observations, GCM baseline, and GCM future. Bias correction is then applied to the daily GCM output using a quantile mapping approach (Thrasher et al., 2012). For precipitation, a maximum of 150% future-to-baseline quantile ratio is enforced to avoid unreasonably large adjustments (Gutman et al., 2014).

(2) *Analogue selection and weight determination* (at 1° grid resolution): For each GCM day, a 91-day (±45 days) window is used to create a historic library across all years in observation. A total of 30 analogues from the historic library that best match the spatial pattern of the targeted GCM day are selected. A statistical regression technique is then used to determine the weights of each analogue (Maurer et al., 2010).

(3) *Construct analogue* (at 1/24° grid resolution): By assuming the same analogues and weights exist at both spatial resolutions, the previous step is reversed. The 30 analogues from observation at the 1/24° grid are first looked up. A linear combination of these 30 analogues with their identified weights is formed as the downscaled climate variable. Output at this step is the conventional BCCA.

(4) *Second bias correction* (at 1/24° grid resolution): A second daily bias correction is applied to the 1/24° constructed analogues to generate the results of the DBCCA. To maintain a reasonable diurnal temperature range (Thrasher et al., 2012), the second bias correction on T_{max} and diurnal temperature range is applied. Then, T_{min} is calculated based on the corrected T_{max} and diurnal temperature range. Wind-speed bias is also corrected to be consistent with the wind-speed data used in Daymet and Livneh. This final step is also the same daily bias correction approach used to correct the raw RegCM4 outputs. Keeping the same final bias correction step can reduce some additional differences from the choice of bias correction approach.

2.4.3 Performance Evaluation

To evaluate the overall performance, the output from GCM, RCM, bias-corrected RCM, and DBCCA were compared with observations (Figure 2.9, Figure 2.10, Figure 2.11, and APPENDIX D). All evaluations were performed at 1/24° grid for the 1980–2018 historic period. The heat maps show the Pearson product-moment coefficient of linear correlation, also known as *pattern correlation* (PC); the ratio of spatial standard deviation (ROS); and the absolute bias in various metrics in the models for Livneh and Daymet observations. The evaluated metrics included the annual and seasonal magnitudes of P, T_{max}, and T_{min}, and the annual magnitudes of the 95th percentile of P (P_{95}), 95th percentile of T_{max} (T_{95}), 5th percentile of T_{min} (T_{05}), wet days defined as number of days with P above 1 mm (W_{days}), and frost days defined as number of days with T_{min} below 0°C (32°F) (F_{days}). The simulations trained and/or corrected with an observation were compared with the respective observation. The PC (Figure 2.9) and ROS (Figure 2.10) were unitless, and biases (Figure 2.11) were calculated as absolute with the units for T_{max}, T_{min}, T_{95}, T_{05} in degrees Celsius, and P in mm/day (note: biases in degrees Fahrenheit and inches per

day are tabulated in APPENDIX D). The biases in P_{95}, W_{days}, and F_{days} were normalized using the mean ensemble bias for presentation and comparison.

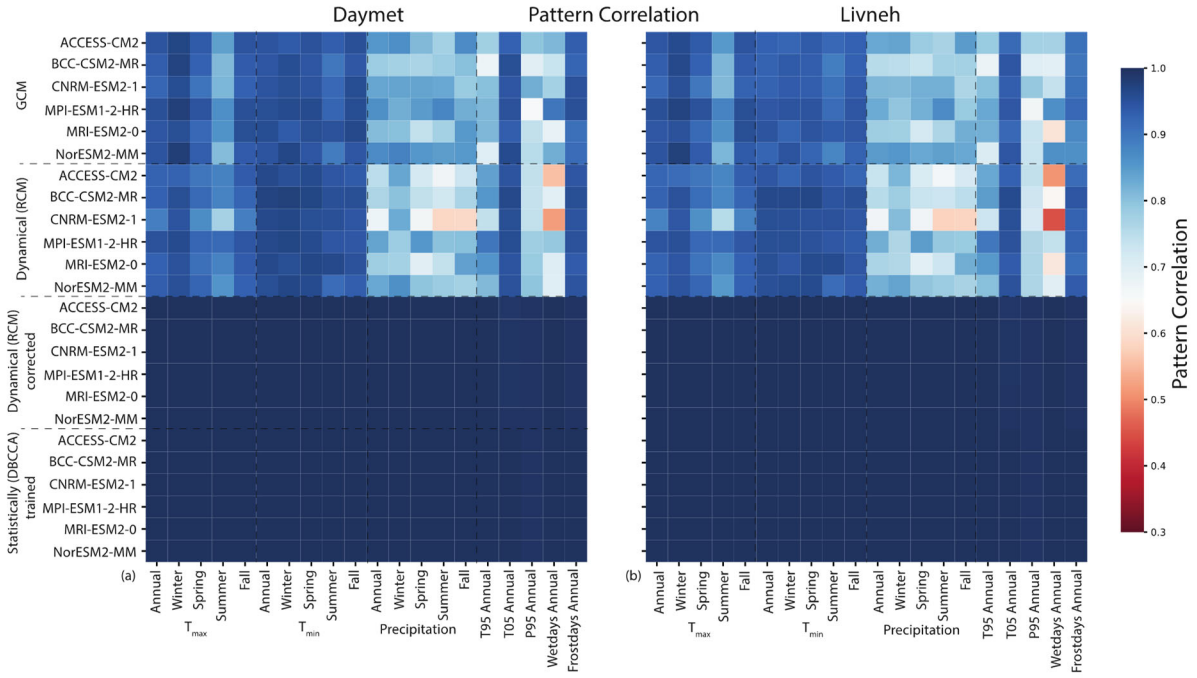


Figure 2.9. Heat maps showing the pattern correlation for climate projections with respect to the corresponding observations.

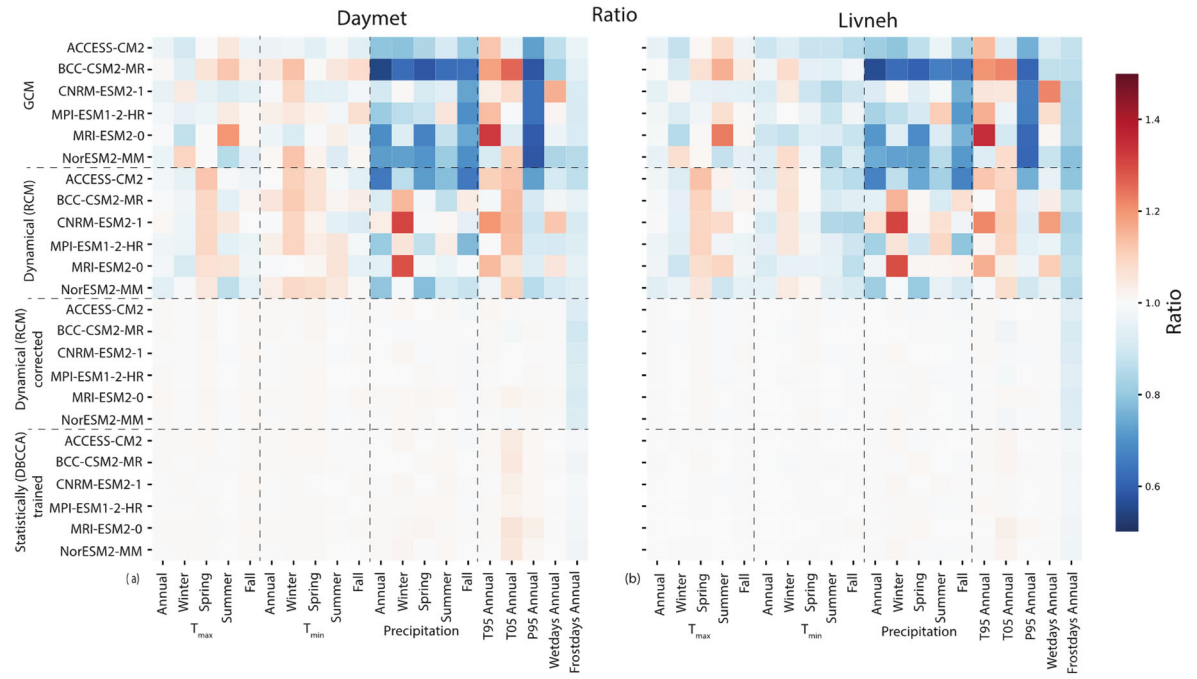


Figure 2.10. Heat maps showing the ratio of standard deviation for climate projections with respect to the corresponding observations.

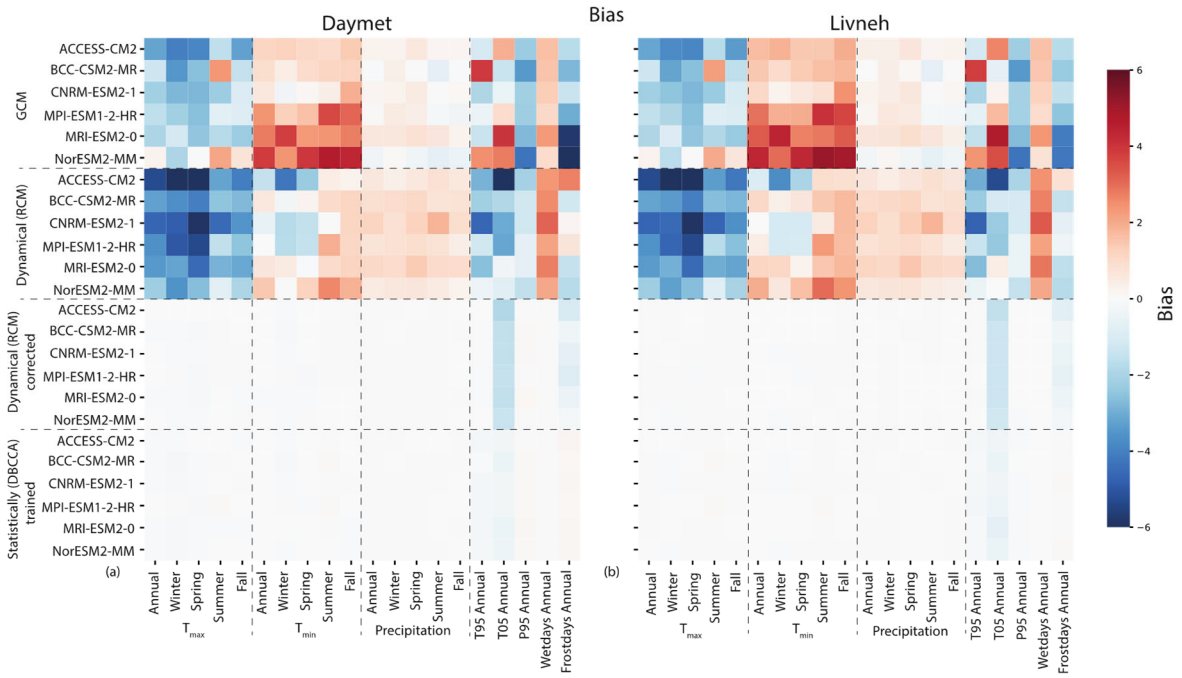


Figure 2.11. Heat maps showing the biases in the climate projections with respect to the corresponding observations.

The bias-corrected RCM and DBCCA compared well for each of the three statistics, showing a pattern correlation above 0.99 and a ROS close to 1 for all mean and extreme indices and negligible biases in the mean P (±0.0016 in./day), T_{min} (−0.23°F to 0.18°F) and T_{max} (−0.27°F to 0.18°F), T_{95} (−0.45°F to 0.04°F), P_{95} (<0.016 in./day), and W_{days} (<1 day) but relatively higher biases in T_{05} (up to −3.2°F) and F_{days} (up to 8 days) for bias-corrected RCM vs. DBCCA for both observations. Contrarily, in these metrics, the GCM and RCM showed larger differences in the magnitude and the spatial pattern. For instance, most of the GCMs showed a predominantly cold bias in T_{max}, which also persisted in T_{95}. These biases in mean and extreme temperatures are generally associated with the wet biases in the GCM simulation of P, which were further exacerbated in the RCM.

The majority of the GCMs showed wet bias in the simulation of mean annual and seasonal P, except during the summer and fall. However, all GCMs showed higher W_{days} compared with the observations (Figure 2.11). These biases in P and W_{days} were further exacerbated in the RCM. Both the GCM and RCM also exhibited dry biases in P_{95}, thereby implying that the larger wet bias in P can be attributed to the simulation of a higher W_{days}, whereas the magnitudes of extremes were underestimated. Although the RCM underestimated the average P_{95} when compared with observations, it showed improvement in the simulation of P_{95} compared with the GCM. The RCM also showed improvement in the simulation of T_{min}, T_{05}, and F_{days}. For instance, the GCM showed a dominant warm bias of up to 9°F, whereas the RCM exhibited bias ranging from −3.2°F to 4.7°F with respect to Daymet and from −6.7°F to 5.4°F in T_{min} with respect to Livneh (Figure 2.11). Consequently, the majority of GCMs also exhibited warm bias in simulating T_{05}, whereas the RCMs showed overall improvement in the simulation of T_{05}.

The pattern correlation for most of the T_{max} and T_{min} metrics was above 0.9 with the exception of summer T_{max} and annual T_{95}, which had pattern correlation values between 0.7 and 0.9 for both the GCM and the RCM (Figure 2.9). Contrarily, the ROS value was close to 1 (0.9 and 1.1) for the majority of the mean T_{max} and T_{min} metrics with respect to Daymet. However, with respect to Livneh, several of the mean T_{min}

metrics (12/30) for the GCM, and a few (6/30) for the RCM, showed values <0.9. For annual and seasonal mean P, P_{95}, and W_{days}, the pattern correlation ranged between 0.6 and 0.87 for GCM and RCM. The ROS value for mean P ranged from 0.6 to 1.4 with values <1 for most metrics/models in the GCM and close to or >1 for the RCM model output (Figure 2.9). For P_{95}, the ROS value was ≥0.9 for five out of six RCMs compared with ≤0.75 for all six GCMs. The W_{days} ROS value was between 0.9 and 1.18 for the RCM and between 0.82 and 1.22 for the GCM (Figure 2.10).

Overall, although GCM and RCM showed higher biases in the simulation of the mean and extreme T_{max} that were generally associated with the biases in the simulation of magnitude and frequency of P, the RCM showed improvement in the simulation of T_{min}, F_{days}, T_{05}, and P_{95} over the GCM. Furthermore, the bias-corrected RCM and DBCCA exhibited skills and substantial improvement in the simulation of all metrics compared with both GCM and RCM. Further details about climate model downscaling can be referred to Rastogi et al. (2022).

2.4.4 Climate Projection Limitations

Some major limitations related to downscaled climate projection are summarized as follows:

- Although GCM-driven dynamical/statistical downscaling may provide the most scientifically defensible regional scale climate projections, it is not an absolute, day-to-day weather prediction model. The main purpose of climate modeling is to simulate how general climate statistics may evolve with respect to the specified future emission scenarios—not to provide a prediction of weather and hydrologic conditions on a specific future date.

- Given the large variability in the climate system, a large ensemble (with various models, initial conditions, and long-term simulation) will be required to capture the overall mean and spread. Depending on the region and variables of interests (e.g., temperature vs. precipitation), the required size of an ensemble could be vastly different. Nevertheless, such a quantitative measure is not fully understood in the current state of climate science. In practice, the choice is usually controlled by the available resources.

- A variety of downscaling techniques have been extensively used in various hydroclimate impact studies, but a consensus on the best downscaling practice has yet to be reached, and this decision may vary based on the hydroclimate analysis need. These different DSs inevitably add a layer of uncertainty and complexity to projected future climate conditions. This study captured the uncertainty arising from two very different downscaling techniques, as well as the choice of two different reference meteorological observations. However, the methodology was limited to only one RCM and one statistical DS and did not capture the full range of biases associated with the choice of RCMs and statistical DSs. In other words, the true uncertainties could not be captured at the national scale because of resource limitations.

- Although the capabilities of GCMs have continuously improved through the years, many open challenges remain. For instance, although human activities play an important role in the Earth system environment, many of the GCM simulations were conducted without considering the potential human influence on the land use and land cover change, and surface hydrologic alterations. Therefore, recurring climate impact assessments based on the best available climate science remain necessary.

2.5 HYDROLOGIC MODELING

The downscaled climate projections generated from work described in Section 2.4 were then translated into hydrologic projections through hydrologic modeling, which simulated watershed behaviors (i.e.,

response of hydrologic variables such as runoff, evaporation, and snow) subject to a set of meteorological forcings (e.g., daily precipitation, temperature). These hydrologic projections can be used to evaluate the regional hydrologic response and changes in future water availability (Figure 2.1). In this study, a 48-member ensemble with six CMIP6 GCMs, two DSs, two MetFs, and two HMs that projected future hydrologic conditions under the SSP585 emission scenario were developed. In other words, the 24 sets of downscaled climate projections from Section 2.4 were used to drive two calibrated distributed HMs (VIC and PRMS) at the CONUS scale. The use of a second HM, compared with the previous 9505 assessments, added an extra dimension to the hydroclimate analysis and provided additional information on the uncertainty related to the choice of HMs. The following subsections provide an overview of each HM, their implementation, and their model performance.

2.5.1 VIC

The first HM, the VIC model, is a macroscale, semi-distributed model that simulates spatially distributed water and energy fluxes, allowing for the representation of various hydrologic processes (e.g., evaporation, runoff, base flow, snowmelt) (Liang et al., 1994). VIC uses a variable infiltration capacity curve to determine the infiltration and surface runoff and uses the empirical Arno curve for base flow calculation. The Arno curve allows VIC to calculate base flow as a function of soil moisture, such that base flow response varies linearly under low-moisture conditions and nonlinearly under high-moisture conditions. The water and energy balances are solved for multiple elevation bands and vegetation types, which enables the model to capture the subgrid variability of these land surface features (Liang et al., 1994). Although VIC is essentially a 1D model that solves mass and energy balance in the vertical direction, an external 2D horizontal routing algorithm can be used to estimate streamflow at a specified location (Lohmann et al., 1996 and 1998). VIC has been widely used in many national and international hydroclimate studies (e.g., Harding et al., 2012; Hengade et al., 2018; Islam et al., 2017; Yan et al., 2015), including the previous two 9505 assessments (Sale et al., 2012; Kao et al., 2016).

Building on the CONUS VIC implementation (Naz et al., 2016; Oubeidillah et al. 2014) used in the previous 9505 assessment, the latest VIC version 5 (Hamman et al., 2018) was used in this study. This model is set up at the same 1/24° (~4 km) grid resolution with a 3-hourly time step and five elevation bands. Although VIC was calibrated in the previous assessment, given the change in model version and reference meteorological observations, further calibration was conducted in this study (Section 2.5.3). Similar to in the previous assessment, the calibration was performed at the HUC8 subbasin scale using the USGS WaterWatch monthly runoff data set (Brakebill et al., 2011) as the benchmark. Furthermore, recognizing the impacts of MetFs on the hydrologic modeling (Elsner et al., 2014), the calibration was performed separately for Daymet and Livneh meteorological data sets. The two optimized parameter sets were then used in conjunction with their corresponding downscaled climate projections during the hydrologic simulation.

2.5.2 PRMS

The second HM, PRMS, is a deterministic, process-based, distributed HM developed by USGS (Leavesley et al., 1983). PRMS uses a modular framework to allow alternative algorithms to simulate several hydrologic processes, such as precipitation, evapotranspiration, runoff, infiltration, and groundwater (Markstrom et al., 2015). PRMS allows for a spatially discretized representation of the physical and geologic properties, such as topography, soil, vegetation, and land use, typically through HRUs. The individual HRUs have homogenous physical characteristics and hydrologic response. The water fluxes and storages are computed at daily time steps to simulate hydrologic response subject to the driving meteorological forcing (e.g., precipitation and temperature). Similar to VIC, PRMS has been applied in several modeling studies to evaluate climate change impacts on water availability (Gangrade et al., 2020; Hay et al., 2011; LaFontaine et al., 2019; Najafi et al., 2011).

In this study, PRMS version 5.1.0 was implemented using a national-scale application of the National Hydrologic Model Infrastructure (Regan et al., 2019), which used HRUs defined by the Geospatial Fabric (Viger and Bock, 2014). The model setup consisted of 109,951 HRUs ranging from 1 km^2 to more than 67,000 km^2, with a median area of 33 km^2 (Sexstone et al., 2020). The initial parameter set was obtained from a national calibration of PRMS using Daymet V3 by Hay (2020). The daily precipitation and maximum and minimum temperatures were computed at each HRU and then aggregated to each HUC8 through area-weighted averaging. Similar to VIC, the calibration was performed separately for the Daymet and Livneh meteorological data sets using USGS WaterWatch monthly runoff as the benchmark. The overall calibration approach was kept consistent to avoid introducing additional uncertainty caused by different calibration algorithms.

2.5.3 Model Calibration and Validation

2.5.3.1 Calibration

Although both HMs are physics-based, several complicated hydrologic processes are represented by simplified relationships and parameterizations. Therefore, model calibration was necessary to ensure that both models may reasonably simulate the historical hydrology. Following a similar procedure with the previous 9505 assessments, both models were calibrated at each HUC8 subbasin using the USGS WaterWatch data set as the benchmark. As demonstrated by the previous 9505 assessments, since WaterWatch runoff is the most significant variable in explaining historic annual hydropower generation (Sale et al., 2012), calibrating an HM to WaterWatch should enhance the accuracy of the subsequent hydropower simulations. In general, both models were calibrated in a consistent manner to reduce the potential bias caused by different calibration procedures. Nevertheless, since VIC is much more computationally demanding than PRMS (more than tenfold in terms of processing time), some necessary simplifications were made to the VIC workflow to ensure that both models could be reasonably calibrated within the project timeline.

During VIC calibration, the model was driven by daily precipitation and maximum/minimum temperature obtained from Daymet V3 and the original Livneh (2015) meteorological data set. Because VIC also requires wind speed, the daily surface wind speed from NARR was used in conjunction with Daymet during the VIC simulation. The Livneh data set provided daily wind speed, which was processed from the National Center for Environmental Prediction reanalysis data set (NCEP1; Kalnay et al., 1996). The Mountain Microclimate Simulation Model (MTCLIM; Thornton and Running, 1999) version 4.2 was used to convert these input data into additional meteorological variables, such as shortwave radiation, longwave radiation, and humidity, for the hydrologic simulation. For model spin-up, the first year (e.g., 1980) was repeated for an additional year. The 1-year spin-up period can be considered sufficient for a land surface model such as VIC (Cosgrove et al., 2003)

Six sensitive VIC parameters were used for calibration and were also calibrated in the second 9505 assessment (Kao et al., 2016). Calibration was performed by minimizing the difference between the simulated total monthly runoff (base flow + surface runoff) and the WaterWatch monthly runoff. Parameters were updated using the particle swarm optimization algorithm (PSO; Kennedy and Eberhart, 1995) to maximize the Nash–Sutcliffe efficiency (NSE), a common evaluation metric for HM performance. PSO is a swarm intelligence–based optimization technique and is effective for use in HM calibration (Jiang et al., 2007). PSO is conceptually simple, computationally efficient, and can take advantage of parallel processing in large high-performance computing systems (Guo et al., 2014; Lee et al., 2006), which was suitable to the application in this study.

To reduce the computing time for VIC calibration, further simplifications were implemented, such as using one elevation band and water balance–only mode only during calibration. Although these simplified

VIC options may cause some local differences at a daily time step, they do not result in noticeable differences to the calibration target (i.e., NSE of monthly HUC8-aggregated total runoff). For other VIC simulations, the more computationally intensive options (five elevation bands and water/energy balance mode) were used.

During PRMS calibration, the model was driven by daily precipitation, maximum and minimum temperature obtained from Daymet V4, and the updated Livneh data set from Pierce et al. (2021). For modeling spin-up, the first year (e.g., 1980) was repeated for an additional 3 years. The 3-year spin-up period is suggested by USGS for NHM-PRMS implementation (Hay, 2019). Eight selected dominant parameters were calibrated using the PSO algorithm based on the detailed work performed by Markstrom et al. (2016). An area-weighted approach was used to aggregate the HRU-based results within each HUC8 subbasin. When calibrating, a one-to-one mapping between HRUs and HUC8s was generated by assigning the HRU to the HUC8 with a maximum areal contribution. The calibration was then performed by maximizing the NSE between the monthly simulated PRMS runoff and WaterWatch runoff at each HUC8. Similar to VIC, PRMS was calibrated separately for Daymet and Livneh, thereby producing two sets of calibrated parameters.

Because VIC calibration requires a tremendous amount of time, it was completed before the release of the latest Daymet and Livneh data sets. Although it would be desirable to recalibrate these VIC parameters based on the latest meteorological observations, it was not feasible given the limited project timeline and computational resources. Conversely, since PRMS is less computationally demanding, it was calibrated using the latest Daymet and Livneh for better consistency with the downscaled climate projection data set. More PSO interactions were also possible, which enabled better PRMS performance. Upon further evaluation, it was found that the differences between different versions of Daymet and Livneh are mainly for the timing at the daily scale and not so much at the monthly scale (the time scale of calibration). Furthermore, the differences between the MetFs are much larger than the differences between versions of the forcings. Therefore, the calibrated parameters from earlier versions of forcings were considered acceptable. These inconsistencies may introduce some additional bias and uncertainties in the ensemble hydrologic projections that are currently unavoidable given resource limitations.

2.5.3.2 Model Performance

To consistently evaluate model performance, VIC and PRMS were driven by a common set of MetFs (i.e., Daymet V4 and the updated Livneh by Pierce et al. [2021]). These simulations are referred to as VIC and PRMS control runs and defined as follows:

1. **VIC-Daymet control run**: VIC with Daymet-optimized parameters, driven by Daymet meteorological forcing from 1980 to 2019;

2. **VIC-Livneh control run**: VIC with Livneh-optimized parameters, driven by Livneh meteorological forcing from 1980 to 2013;

3. **PRMS-Daymet control run**: PRMS with Daymet-optimized parameters, driven by Daymet meteorological forcing from 1980 to 2019; and

4. **PRMS-Livneh control run**: PRMS with Livneh-optimized parameters, driven by Livneh meteorological forcing from 1980 to 2013.

Because the updated Livneh data set provided by Pierce et al. (2021) from 2014 to 2018 was not available for parts of the watersheds (i.e., the Mexican headwater basins of the Rio Grande River), Livneh-based control runs were ended in 2013 for spatial consistency.

The model performance was evaluated for each of the control runs at each HUC8. Figure 2.12 and Figure 2.13 present the overall performance of VIC and PRMS, respectively. Figure 2.12 shows the mean 1981–2000 annual runoff from the WaterWatch benchmark (a), VIC-Daymet (b), and VIC-Livneh (c). A qualitative evaluation of these panels indicates that VIC may closely simulate observed annual runoff for many HUC8s across the United States. Panel (d) shows the scatter plots of HUC8-based annual runoff and the R^2 values of the two VIC control simulations. In general, both control simulations showed good results, with $R^2 = 0.927$ for VIC-Daymet and $R^2 = 0.932$ for VIC-Livneh for annual runoff. Figure 2.12 also shows the NSE values of monthly runoff for VIC-Daymet (e) and VIC-Livneh (f) at each HUC8 subbasin. At the national scale, 63% of HUC8s showed monthly NSE >0.5 for VIC-Daymet, and 68% of HUC8s showed monthly NSE >0.5 for VIC-Livneh. A subset of the HUC8s within PMA areas revealed that 48% showed monthly NSE >0.5 for VIC-Daymet, and 53% showed monthly NSE >0.5 for VIC-Livneh. In general, these results indicate that VIC performs better in wetter regions than in drier/arid regions across CONUS (additional analysis presented in APPENDIX I). Similar findings were also reported by Newman et al. (2017). Additionally, in terms of R^2, VIC-Daymet appears to perform better than VIC-Livneh relative to WaterWatch runoff.

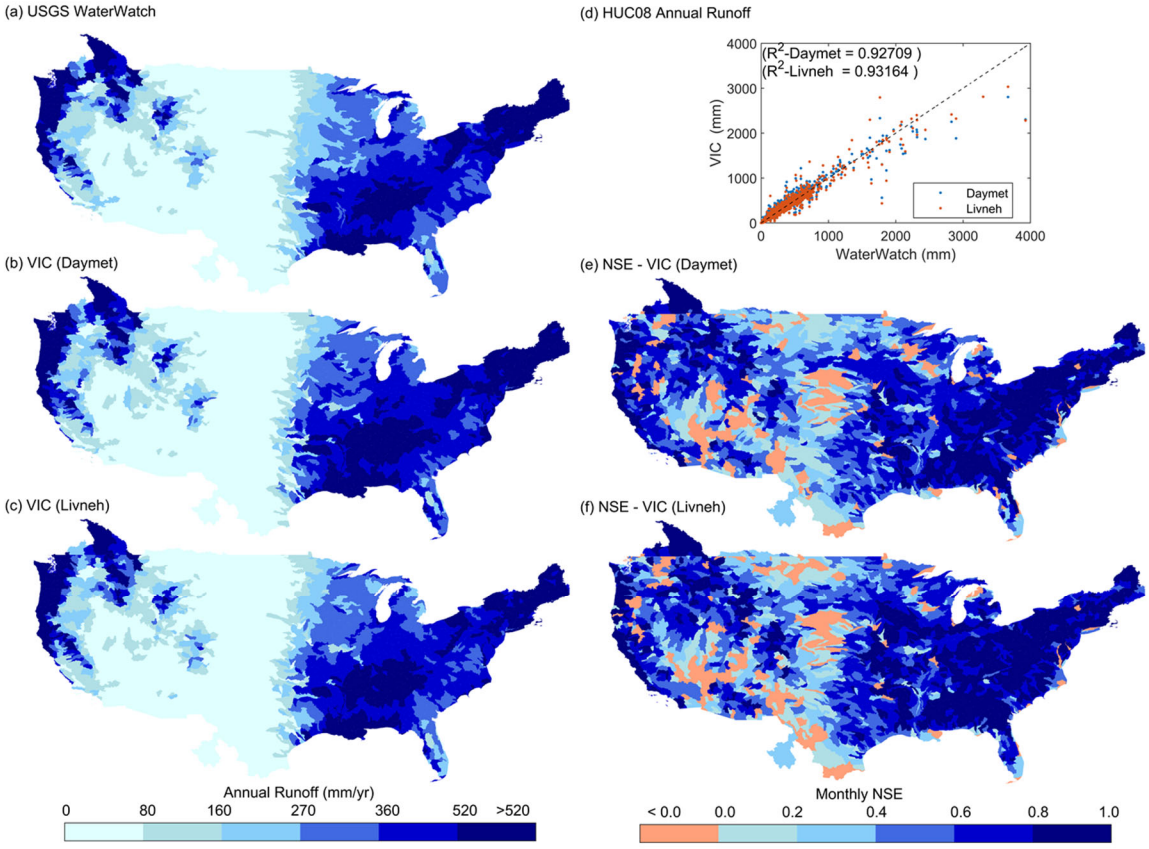


Figure 2.12. Performance of VIC at various HUC8s in the CONUS.

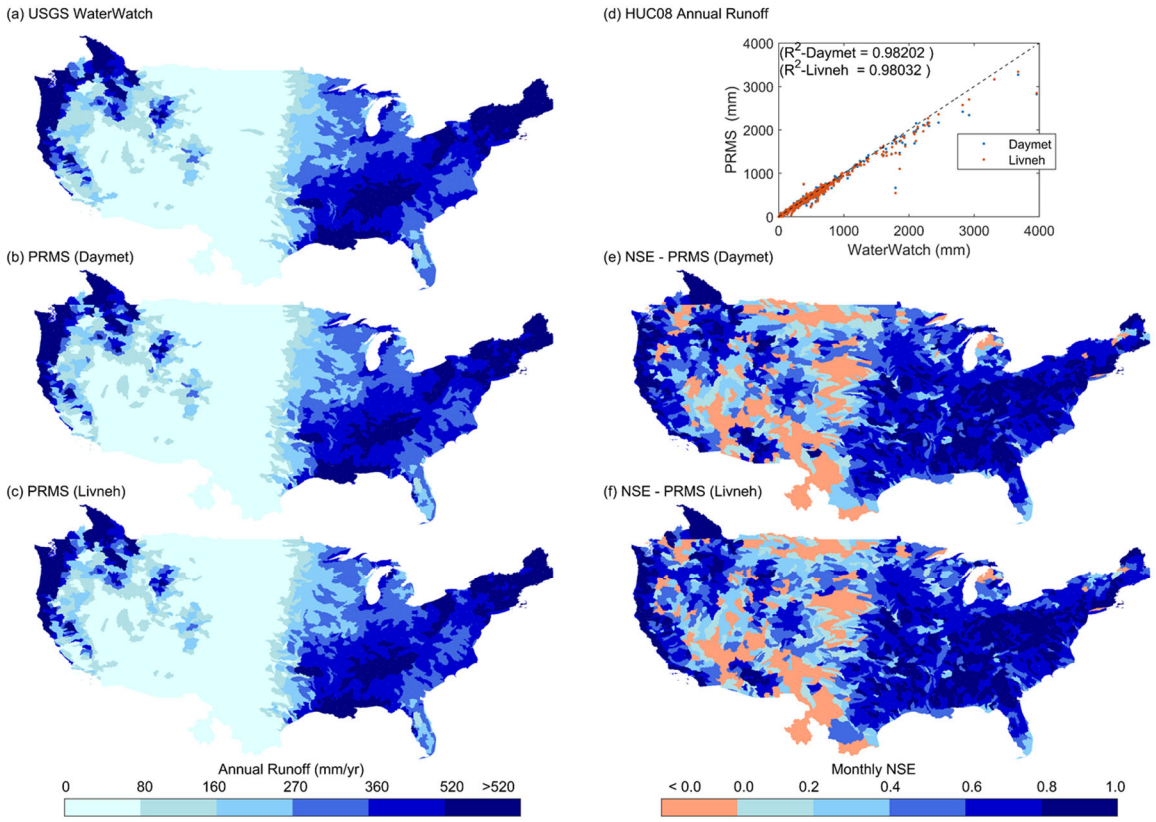


Figure 2.13. Performance of PRMS at various HUC8s in the CONUS.

Figure 2.13 shows the mean 1981–2000 annual runoff from the WaterWatch benchmark (a), PRMS-Daymet (b), and PRMS-Livneh (c). A qualitative evaluation of these panels indicates that PRMS may also closely simulate observed annual runoff for many HUC8s across the United States. Panel (d) shows scatter plots of HUC8-based annual runoff and their R^2 values of the two PRMS control simulations. In general, both control simulations showed reasonably good results, with $R^2 = 0.982$ for PRMS-Daymet and $R^2 = 0.980$ for PRMS-Livneh for annual runoff. Figure 2.13 shows NSE values of monthly runoff for PRMS-Daymet (e) and PRMS-Livneh (f) at each HUC8 subbasin. At the national scale, 65% of HUC8s showed monthly NSE >0.5 for both PRMS-Daymet and PRMS-Livneh. A subset of the HUC8s within PMA areas revealed that 43% showed monthly NSE >0.5 for PRMS-Daymet, and 44% showed monthly NSE >0.5 for PRMS-Livneh. Similar to VIC, PRMS performed better in wetter regions than in arid/drier areas. However, unlike VIC, PRMS-Daymet and PRMS-Livneh achieved comparable performance in terms of R^2. Also, PRMS control simulations achieved higher performance than VIC control simulations. These differences can be attributed to VIC being calibrated with the earlier versions of Daymet and Livneh data sets; therefore, a slight deterioration in the performance was expected when driven with updated MetFs during control simulations. On the other hand, PRMS was calibrated using the latest version of MetFs along with more robust calibration and therefore showed better performance. A further comparison is needed to investigate the physical reasons behind these differences, but that is beyond the scope of this report.

To further evaluate the model performance for snow processes, the simulated April 1 SWE was compared with Snow Telemetry (SNOTEL) observations. The SWE data were obtained from Bias Correction and Quality Control SNOTEL data for 829 active stations in the western United States and Alaska (Sun et al., 2019; Yan et al., 2018). Focusing on the period of 1981–2000, 482 stations with at least 10 years of

annual April 1 SWE observations were selected. For each station, the simulated April 1 SWE at the nearest VIC grid and PRMS HRU was referenced. Because the point observation may be at a different scale than the grid-based VIC SWE and HRU-based PRMS SWE, the R^2 between observation and simulation was computed for evaluation. The results for the four control simulations are demonstrated in Figure 2.14. The results indicate that 63% of VIC-Daymet stations, 61% of VIC-Livneh stations, 62% of PRMS-Daymet stations, and 52% of PRMS-Livneh stations had R^2 values greater than 0.5. Both models performed relatively well in the Wyoming and Idaho regions. For the p-values of the corresponding correlation coefficients (R), 82% of Daymet-driven stations have p-values less than 0.05, whereas 78% of Livneh-driven stations have p-values less than 0.05, indicating that correlation was statistically significant at 5% significance level. This suggests that the control simulations captured the interannual variability of SWE for most of the stations.

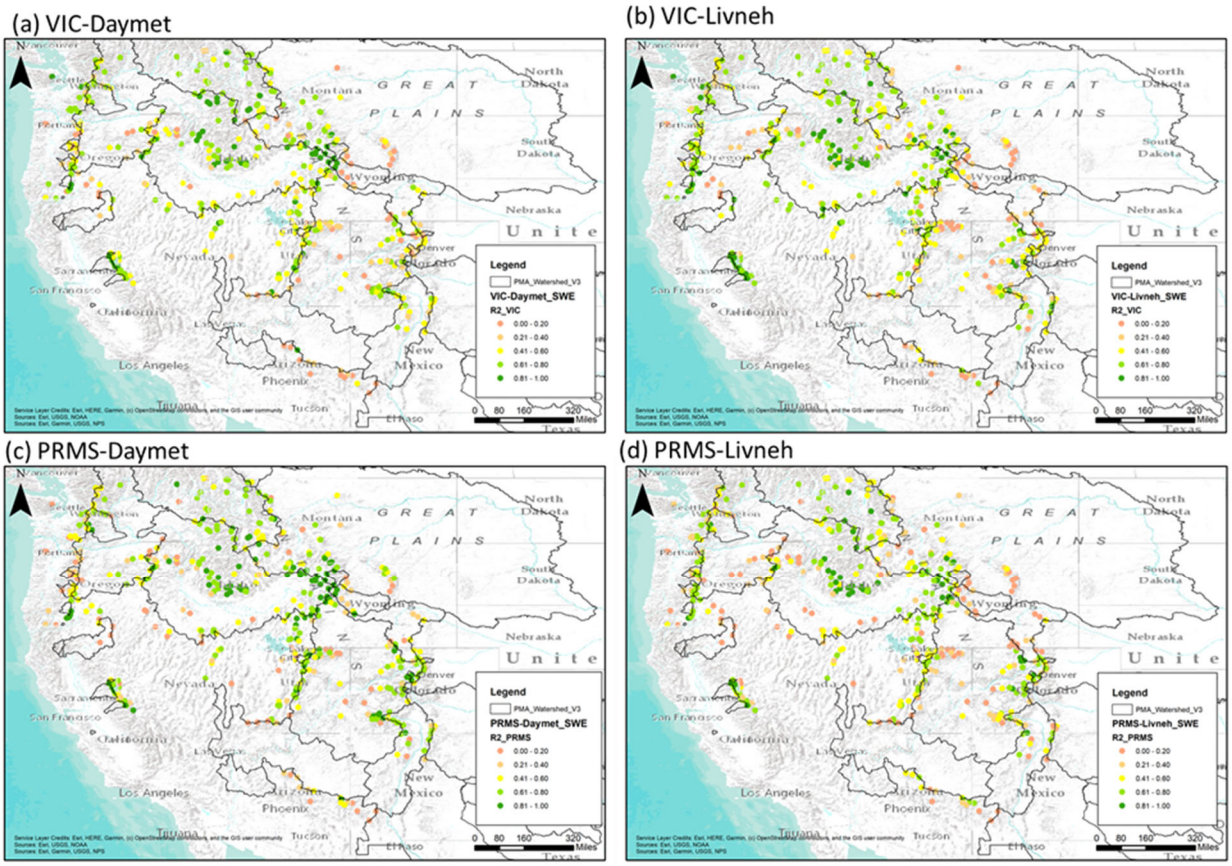


Figure 2.14. Performance of VIC and PRMS at various SNOTEL sites.

2.5.4 Hydrologic Evaluation

For each PMA study area, the effects of future changes on major hydrologic variables, including total runoff and SWE, were analyzed and are presented in Sections 3–6. The simulated runoff was first summarized for each PMA region by calculating the average daily timeseries of each ensemble member in the baseline period (1980–2019) and near-term (2020–2039) and mid-term (2040–2059) future periods. The projected runoff changes from baseline to future periods in mean annual, spring (MAM), summer (JJA), fall (SON), and winter (DJF) values were then estimated.

The change in extreme runoff conditions (high and low runoff) is also described (Sections 3–6). High runoff (R95) was defined as the 95th percentile of daily runoff, and low runoff (R5) was defined as the 5th percentile of the average 7-day runoff calculated at each 1/24° (~4 km) VIC grid cell. For a consistent comparison between PRMS and VIC, the PRMS HRU-based runoff estimates were first bilinearly interpolated to the corresponding VIC grids for extreme runoff analysis. The annual percentage change (future minus baseline) of high and low runoff were then calculated for each 1/24° grid.

For the projected change in snow conditions, the relative change in average April 1 SWE was calculated separately for each ensemble member and shown as the ensemble median for the two snow-dominated PMAs, BPA and WAPA. HUC8s with more substantial snowpack were identified, defined as HUC8s with an average 1980–2019 control run April 1 SWE greater than 5 mm. The identified HUC8s were then used as a spatial filter to summarize the simulated April 1 SWE for each ensemble member during the baseline and future periods.

2.5.5 Model Limitations

Although efforts to improve the HM performance and support ensemble hydroclimate simulations have been made, challenges and limitations remain. Some limitations include the following:

- This study used two commonly used HMs—VIC and PRMS. Attempts were made to avoid unintentional biases. Using similar procedures and data sets for calibration and evaluating model performance through a common HUC8-based framework still resulted in differences beyond control caused by different spatial resolutions, parameterizations, and model configurations. Additionally, different HM parameters determined through different calibration objectives (i.e., across temporal scales and metrics) may result in differences in the future hydroclimate projections. These factors can be further explored in future assessments.

- As discussed in Section 2.5.3, VIC was calibrated using the earlier versions of Daymet and Livneh. Although it would be better to keep the versions consistent, owing to the high computational cost of VIC and the limited project timeline, recalibration with the latest Daymet V4 and updated Livneh by Pierce et al. (2021) was not feasible. PRMS, on the other hand, is computationally efficient and can be calibrated with the latest version of MetFs. These differences can result in some unintended inconsistencies and uncertainties. However, even when using different versions of meteorological data sets for calibration, this approach was preferred over alternatives such as not calibrating the HMs or using different MetFs for downscaling and hydrologic calibration. Overall, a consistent MetF should be used for HM calibration and downscaling to avoid the large uncertainties caused by different reference meteorological data sets.

- MTCLIM is widely used to generate subdaily meteorological forcing in many hydrologic modeling studies, but it has certain limitations. A global evaluation of the MTCLIM algorithm by Bohn et al. (2013) showed that large negative bias of relative humidity may exist in arid conditions. They also showed that the biases of radiation can reach 10% or more. Their results were supported by follow-up studies (Erlandsen et al., 2019; Slater, 2016). Slater (2016) showed that over 30 W/m^2 differences may exist between different versions of MTCLIM (4.2 vs. 4.3). Mizukami et al. (2014) also showed that biases in daily minimum temperature strongly influence radiation estimates by MTCLIM.

- VIC and PRMS provided satisfactory performance for the purposes of this study; however, they do not include a robust groundwater component. This limitation can impact the watersheds with strong surface-groundwater interactions that are highly variable across regions and geology, thereby degrading model performance. A future assessment could include alternative HMs with better groundwater representation to improve this inadequacy.

- Given the lack of human influence representation (e.g., reservoir regulation, streamflow diversion) in both VIC and PRMS, naturalized runoff observations are desired to support HM calibration. This will be more impactful for sites with a large upstream storage. Nevertheless, because WaterWatch did not explicitly exclude gauges that were under streamflow regulation, it could not fully represent natural runoff (i.e., without human influence). Despite this limitation, the WaterWatch data set provided spatially consistent monthly runoff estimates across all CONUS HUC8s that allowed systematic HM calibration for national-scale simulation. Additional efforts can be invested to develop a true naturalized national runoff data set to better support large-scale HM calibration and validation.

- The HMs in this study used an assumption of time-invariant parameters such as land use/land cover, water use efficiency influenced by increasing CO_2, and so on. Future assessments may explore the use of temporally varying representation of such parameters, which may help develop a more realistic portrayal of runoff/streamflow projections.

- The current ensemble hydrologic projections were generated using six CMIP6 GCMs, two DSs, two MetFs and two HMs. Although this approach enabled the evaluation of the relative influence of each factor, these projections do not represent the full range of uncertainties related to all factors. Incorporation of additional modeling choices may result in different hydroclimate projections. Similarly, different calibration procedures or HM parameters may further introduce additional uncertainties and were not considered in this study. When resources allow, conducting a thorough exploration by incorporating more modeling choices will be of interest. However, such an expansion is computationally expensive and may not be feasible at the national scale.

2.6 HYDROPOWER SIMULATION

Upon the completion of hydrologic modeling, the simulated spatially distributed runoff represented the available surface water resources generated from the watersheds. In most places, these fresh-water resources are managed for irrigation, domestic/industrial water supply, fish passage and environmental flows, hydropower generation, thermoelectric cooling, and flood control through water infrastructures such as reservoirs, retention ponds, and canals. Projections of hydropower generation under climate change conditions, at a monthly time scale, can be derived in several ways. For example, an empirical relationship can be derived to translate the amount of natural runoff directly into hydroelectric energy potential (e.g., Kao et al., 2015). Another approach is to represent both river transport and reservoir operations explicitly, such as using a process-based reservoir management model that simulates water release, reservoir storage, and hydropower generation under various operation rules (e.g., Voisin et al., 2013a; Yao and Georgakakos, 2001). In this study, two different regional hydropower modeling approaches were adopted to explore how the choice of the PMs affects future hydropower projections. They include a lumped Watershed Runoff-Energy Storage (WRES) model that was developed in the previous 9505 assessment and a process-based Water Management Hydropower (WMP) model. The model assumptions, details, and validation are described in this section.

2.6.1 WRES

A lumped WRES model that was developed in the second 9505 assessment (Kao et al., 2016) was adopted in this study to examine how sub-annual hydropower generation may change in the future climate. The schema for WRES is shown in Figure 2.15. Although various types of federal hydropower systems are associated with diverse hydrologic conditions and operational objectives, WRES was designed without site-specific operations to maintain an internally consistent modeling approach that isolated the effects of climate change on hydropower potential. For each hydropower region, WRES used

the monthly precipitation and natural (unregulated) runoff as inputs, performed a runoff mass balance calculation for the total monthly runoff storage in all reservoirs and retention facilities in the watershed, and simulated the monthly regulated runoff release and hydropower generation through the system. The required WRES model parameters, including initial, maximum/minimum monthly storage, and maximum runoff-hydropower capacity, were calibrated for the historical period from 1984 to 2013 (30 years). The calibrated model was then driven by the downscaled hydroclimate variables to project future hydropower generation, assuming that most of the reservoirs in the watershed will be operated in the same way in the future. Two main components in WRES—(1) monthly generation prediction and (2) watershed runoff storage calculation—are described here.

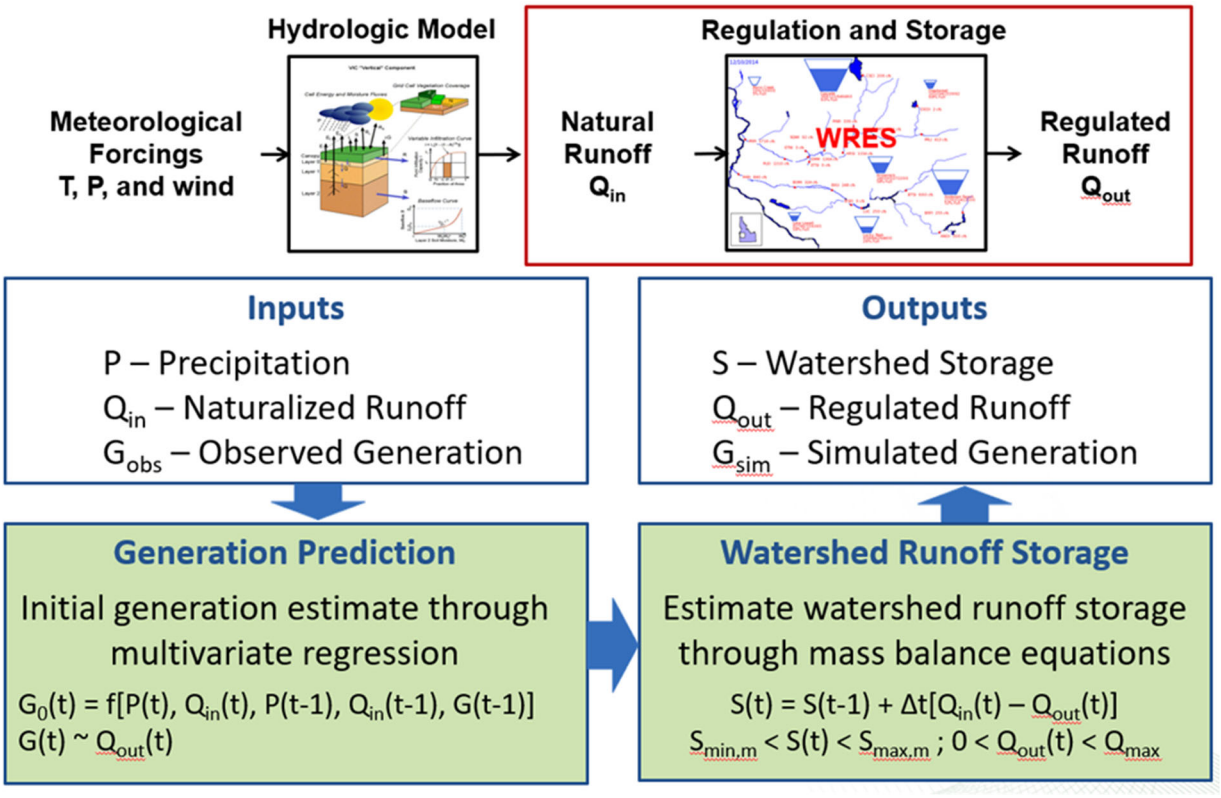


Figure 2.15. Schema of WRES.

2.6.1.1 Monthly Generation Prediction

The first part of the WRES simulation determined an initial estimate of total monthly hydropower generation G_0 (MWh/month) based on hydrologic inflow/outflow conditions (e.g., the amount of precipitation and runoff) and hydropower generation from previous time steps. After testing several combinations of variables and lag times, a generalized multivariate regression model with five predictor variables was used:

$$G_0(t) = b_{0,m} + b_{1,m} * P(t) + b_{2,m} * Q_{in}(t) + b_{3,m} * P(t-1) + b_{4,m} * Q_{in}(t-1) + b_{5,m} * G(t-1). \quad (2.1)$$

where t is the time of calculation (months), m is the calendar month that t belongs to, $P(t)$ and $P(t–1)$ are the monthly precipitation (in./month) from the current and previous months, $Q_{in}(t)$ and $Q_{in}(t–1)$ are the monthly natural runoff (in./month) from the current and previous months, $G(t–1)$ is the monthly generation (MWh/month) from the previous month, and $b_{0,m}$, $b_{1,m}$, …, $b_{5,m}$ are the regression coefficients

developed separately for each calendar month, m. For a study area located downstream from other study areas (e.g., BPA-3 is downstream from BPA-1 and BPA-2), P and Q were calculated from all upstream watershed areas before the WRES simulation was conducted.

Although the total storage in the watershed should be a useful variable to predict generation, it was not included in the multivariate regression formula because it was also a variable to solve in WRES. Instead, generation from the previous time steps was included in the regression formula. Generation, precipitation, and runoff from the previous time steps reflect the overall storage in the watershed and can replace the role of storage in the multivariate regression.

2.6.1.2 Watershed Runoff Storage Calculation

Although G_0 from Eq. (2.1) provides a likely generation estimate based on hydrologic inflow/outflow conditions (i.e., jointly captured by the five predictor variables), it can be unreasonable during extremely wet or dry conditions. Most of the reservoirs follow the established operation curves in which seasonal maximum/minimum pool elevations are specified. During drought conditions, the storage in the system may be close to the minimum so that the water release and generation would be reduced for water conservation. Conversely, during wet conditions, the storage may approach a maximum capacity for flood-risk management. Water release and hydropower generation will increase during wet conditions, and sometimes water may be spilled (i.e., not pass through hydropower turbines) during flood conditions. Given that an extreme hydrologic event is one of the main concerns of future water and hydropower operations, these storage limitations were addressed within the overall formulation.

To capture the maximum and minimum storage limitations in the watershed, a lumped runoff mass balance calculation was performed. Several conceptual watershed storage characteristics were defined, including maximum monthly watershed storage, $S_{max,m}$ (in.), minimum monthly watershed storage, $S_{min,m}$ (in.), and maximum runoff-hydropower capacity, Q_{max} (in./month). Four steps were included in the watershed runoff storage calculation.

Step 1: Estimate runoff release and watershed storage. Based on the linear relationship between runoff and hydropower generation (identified from the first 9505 assessment), it was further assumed that the monthly hydropower generation was associated with the monthly runoff release, Q_{out}, from the watershed. Therefore, G_0 could be transformed to a runoff release estimate, $Q_{out,step1}$ (in./month) by

$$Q_{out,step1}(t) = G_0(t)/a, \tag{2.2}$$

where a is a conversion factor between Q_{out} and G. Through the concept of total runoff mass balance, the watershed storage was further estimated by

$$S_{step1}(t) = S(t-1) + Q_{in}(t) - Q_{out,step1}(t). \tag{2.3}$$

Step 2: Revise based on the maximum/minimum watershed storage limitations. The initial watershed storage estimate from Step 1 was checked against $S_{max,m}$ and $S_{min,m}$. If the value exceeded the watershed limitation, the targeted watershed storage and runoff release were revised:

$$\begin{cases} If\ S_{step1}(t) < S_{min,m}, & S_{step2}(t) = S_{min,m} \\ If\ S_{step1}(t) > S_{max,m}, & S_{step2}(t) = S_{max,m} \\ Otherwise, & S_{step2}(t) = S_{step1}(t). \end{cases} \tag{2.4}$$

$$Q_{out,step2}(t) = S(t-1) + Q_{in}(t) - S_{step2}(t). \tag{2.5}$$

Step 3: Check if negative release occurs. Under extreme drought conditions, the runoff release from Step 2 could occasionally be negative. To avoid this potential issue, the following adjustment was performed:

$$\begin{cases} If\ Q_{out,step2}(t) < 0, & Q_{out}(t) = 0 \\ Otherwise, & Q_{out}(t) = Q_{out,step2}(t) \end{cases}, \tag{2.6}$$

$$S(t) = S(t-1) + Q_{in}(t) - Q_{out}(t). \tag{2.7}$$

Step 4: Check if spill occurs. Under high-flow conditions, the amount of total runoff release, Q_{out}, could sometimes be greater than the amount of runoff that could be used for hydropower generation. Therefore, the maximum runoff-hydropower capacity parameter, Q_{max}, was checked to determine if any portion of runoff release was spilled. The spill estimated in this step refers to forced spill initiated for flood-risk management. Lack-of-market spills were not considered. The minimum flow and spill, as well as the potential outage of turbines, were not simulated in the current model. The final monthly generation estimate, G_{sim}, was then determined:

$$\begin{cases} If\ Q_{out}(t) > Q_{max}, & Q_{out,hydro}(t) = Q_{max}\ and \\ & Q_{out,spill}(t) = Q_{out}(t) - Q_{max} \\ Otherwise, & Q_{out,hydro}(t) = Q_{out}(t) \\ & Q_{out,spill}(t) = 0 \end{cases}, \tag{2.8}$$

$$G_{sim}(t) = a * Q_{out,hydro}(t). \tag{2.9}$$

To estimate a in Eqs. (2.2) and (2.9), the total historic hydropower generation and total WaterWatch runoff from 1980 to 2018 were calculated, and the parameter a was estimated by (total observed generation) ÷ (total observed runoff). Other required parameters, including monthly watershed storage maximum, $S_{max,Jan}$, $S_{max,Feb}$, …, $S_{max,Dec}$; storage minimum, $S_{min,Jan}$, $S_{min,Feb}$, …, $S_{min,Dec}$; maximum runoff-hydropower capacity, Q_{max}; and the initial runoff release and watershed storage conditions were calibrated by minimizing the difference between observed and simulated monthly generation. Additional WRES details are provided by Kao et al. (2016).

2.6.2 WMP

The spatially distributed runoff was used to force a regional PM, the WMP model. This model consisted of two parts: (1) a river-routing water management model for simulating regulated streamflow at all grid cells and dam storage at substantial man-made reservoirs (Voisin et al., 2013a), and (2) a subsequently enhanced, process-based hydropower generation model for hydropower simulation at individual hydropower plants (Zhou et al., 2018). The schema for WMP is shown in Figure 2.16.

2.6.2.1 Water Management Model

The first part of WMP includes the large-scale river routing and water management using MOSART-WM, which consisted of the grid-based river routing Model for Scale Adaptive River Transport (MOSART; Li et al., 2013) and a water management component (Voisin et al., 2013a) used to simulate future regulated river flow. To address the lack of available information on individual reservoir operating rules and the computational trade-offs associated with the spatially distributed nature of the model (65,000 grid cells in the CONUS domain), each reservoir uses generic operating rules (Hanasaki et al., 2006) based on their main operating purposes (e.g., irrigation, flood control). The monthly operating rules are customized for each reservoir based on expected long-term mean monthly and annual inflow,

downstream demand, and reservoir storage capabilities for managing reservoir volume and release targets. Those monthly rules are updated annually based on water year conditions (wet, dry, or normal year). Flow releases from the reservoir are updated daily based on the operating rules, inflow, minimum environmental flow release, and minimum/maximum storage capacity. The hydropower module presented in the next section specifies the distribution of daily release into penstock and environmental spill.

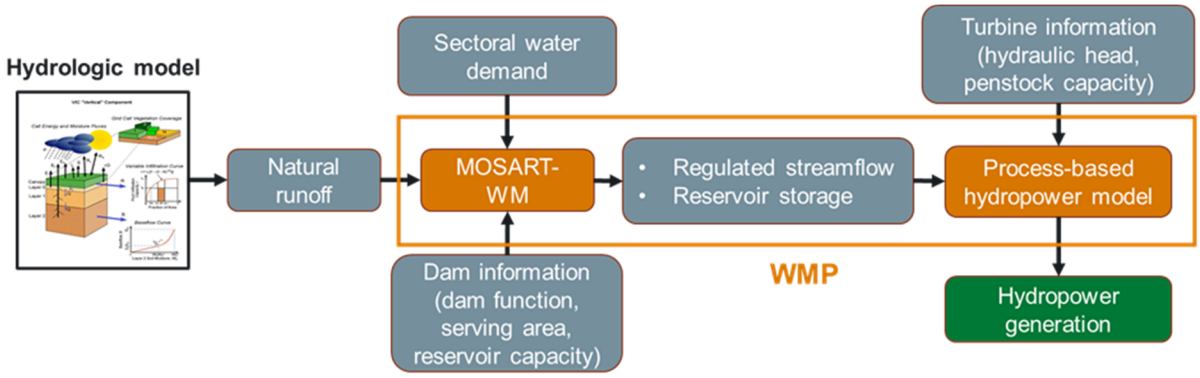


Figure 2.16. Schema of WMP.

Details of how operation rules are determined are provided by Voisin et al. (2013a). The MOSART-WM model has been shown to represent seasonal variations in regulated river flows across the United States (Hejazi et al., 2015; Voisin et al., 2017; Voisin et al., 2013a and 2013b; Zhou et al., 2018) and is similar to other large-scale water management applications in terms of handling water withdrawal and dam operations (Biemans et al., 2011; Döll et al., 2012; Hanasaki et al., 2006; van Vliet et al., 2012; Zhou et al., 2016). They do not reflect the complex daily operations due to a lack of operating rules information. The MOSART-WM model simulated more than 1,839 reservoirs over the entire CONUS (Figure 2.17) with basic information (e.g., functionality, storage capacity) taken from the GRanD (Lehner et al., 2011) and NID database. MOSART-WM was driven by daily runoff simulated by VIC and PRMS aggregated to 1/8° resolution.

Daily gridded water demand data are required input for MOSART-WM to guide the irrigation and non-irrigation water withdrawals from the streamflow and upstream reservoirs. Both consumptive and withdrawal water demands are required and at a minimum, the consumptive water demand, which is not available from USGS data sets after 2005. The consumptive data were derived from the integrated Global Change Assessment Model (Hejazi et al., 2015), which considers multiple water demand sectors and several socioeconomic variables, such as population, labor productivity, and technology (Edmonds et al., 1997; Edmonds and Reilly, 1985; Kim et al., 2006). For this study, the water demand levels were constant and similar to levels in 2010. The monthly sectoral water demands were calibrated and evaluated with respect to USGS water demand, as described by Voisin et al. (2013b). The MOSART-WM model was run at daily time steps from 1980 to 2059 after a 5-year spin-up with the 1980 runoff. The regulated streamflow at 128 federal hydropower plants[5] was extracted as input for the process-based hydropower generation model. For hydropower plants associated with reservoirs, the MOSART-WM-simulated storage time series was also extracted to derive the hydraulic heads.

[5] Among the 132 federal hydropower plants considered in this study, four plants are not represented in WMP because they are either pumped storage plants or plants with capacity less than 1 MW.

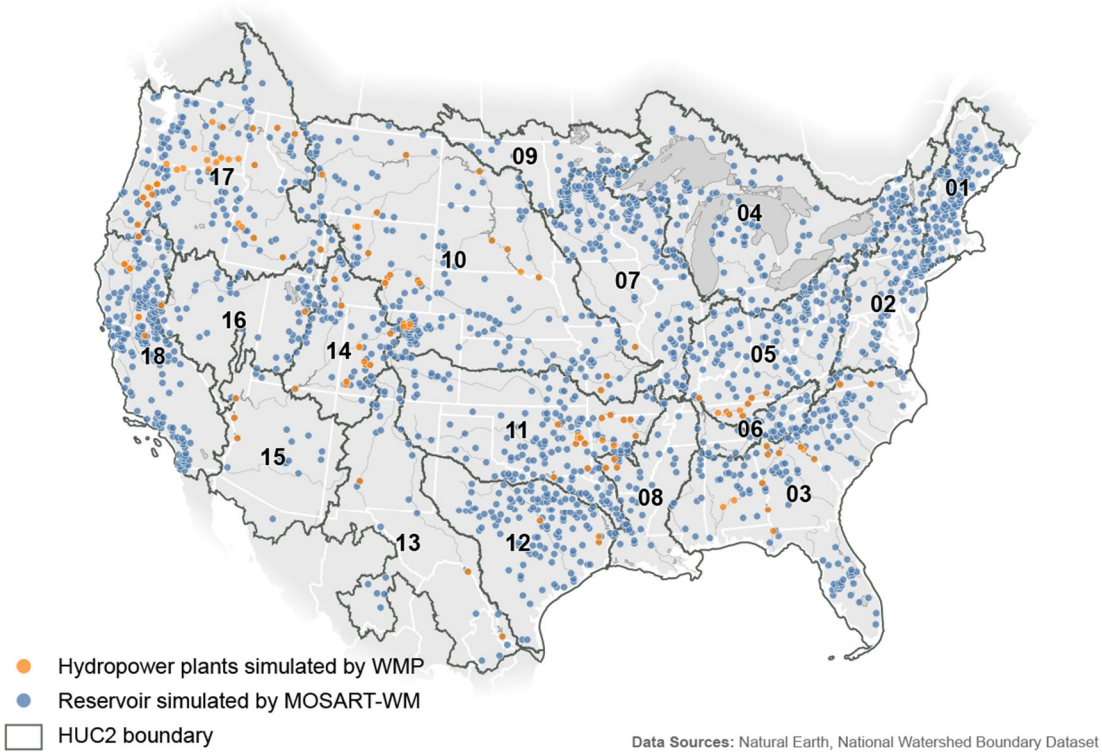


Figure 2.17. Map of powered and nonpowered dams represented in MOSART-WM and WMP over the CONUS.

2.6.2.2 Hydropower Model

The second part of WMP is the process-based hydropower model that predicts the power generation from the streamflow and reservoir operations simulated by MOSART-WM through three major steps.

Step 1: Hydraulic head for storage projects and run-of-the-river (ROR) dams. The enhanced process-based hydropower generation model (Figure 2.16) first determined whether the simulated power plant is associated with a reservoir. Of the 128 simulated hydropower plants, 49 of them are ROR dams for which reservoir storage variations are not explicitly simulated in MOSART-WM, and only the flow is used to estimate generation. For ROR plants, the hydraulic heads (h, m) were assumed to be unchanged and equal to the dam height (H, m). For plants associated with reservoirs, the hydraulic heads at each time step were estimated based on the simulated storage (V, m^3) time series by assuming the shape of the reservoir is a tetrahedron (Fekete et al., 2010) represented by

$$h = H\sqrt[3]{\frac{V}{V_{max}}}, \tag{2.10}$$

where V_{max} is the volumetric capacity (m^3) of the reservoir.

Step 2: Flow allocation for hydropower generation. After the hydraulic head was determined, several parameters were introduced into the power model to adjust the runoff and to constrain the generation at each simulated powerplant. The parameter set included (1) a bias correction factor (f_b), which accounts for the bias of annual runoff and generator efficiencies; (2) 12 monthly spill correction factors (f_{s1}, f_{s2}, …, f_{s12}), which determine the mandatory spills for ecological purposes and spinning reserve; and (3) two penstock intake adjustment factors (f_p and f_m), which determine the actual maximum intake of the

penstock relative to the reported penstock capacity and scale the intake discharge so that it could reserve the temporal variation when the maximum value is always met. All 15 parameters were calibrated in a two-step multiscale calibration process (see details in Section 2.6.3) and applied to estimate the future hydropower generations at each plant by using

$$G_{sim,m} = \rho g h Q_{adj} f_b f_{s,m} \,, \tag{2.11}$$

where $G_{sim,m}$ is the power generation at month m, ρ is the water density, g is gravitational acceleration, and Q_{adj} is the turbine flow rate, which is determined by

$$\begin{cases} Q_{adj}(t) = \min(Q(t), Q_{max}) & if\ \frac{Q_{max}}{\bar{Q}} > f_m \\ Q_{adj}(t) = Q(t) \times \frac{Q_{max}}{\bar{Q}} & if\ \frac{Q_{max}}{\bar{Q}} \leq f_m \end{cases} . \tag{2.12}$$

Step 3: Calibration: A two-step multiscale calibration process was developed to enhance the accuracy of the hydropower projection across different levels of spatial scales relevant to several applications (e.g., watershed scale and load balancing authorities, plant scale).

The target parameter of the first step was f_b, which addresses overall bias in hydrometeorological model simulations, including water management and water demands. In this step, the model was calibrated at the PMA level at an annual time scale against the potential generation, *PG* (MW), equivalent to the actual generation, *G* (MW), plus the operating reserve, *R*. The *R* was assumed to be fixed at 4% of the total mean generation of each PMA region to be consistent with typical power system model applications (Schlag et al., 2015). The mean power generation was averaged from the EIA data set throughout the calibration period. In a PMA region with *n* hydropower plants, the f_b was calibrated using the shuffled complex evolution auto-calibration algorithm (SCE; Duan et al., 1993) to minimize the mean absolute error between the annual time series of PG_{sim} and PG_{obs}, where

$$PG_{sim} = \sum_{i=1}^{n} \rho g\, h_i Q_i f_{b,i}. \tag{2.13}$$

The second step aimed to improve the seasonal variation for each power plant by calibrating f_s, f_p, and f_m. In this step, the f_b calibrated in the previous step was fixed for each plant. At each power plant, the SCE algorithm was again applied to reduce the Kling–Gupta efficiency (Gupta et al., 2009) between the simulated monthly power generation time series (G_{sim}) and G_{obs} over the calibration period using Eq. (2.11).

2.6.3 Model Calibration and Validation

The calibration of WRES and WMP was conducted for each of the hydrologic control run simulations (i.e., VIC-Daymet, VIC-Livneh, PRMS-Daymet, and PRMS-Livneh) (Section 2.5.3) over the 1984–2013 period (Figure 2.18). Validations were performed over the 2014–2019 period using Daymet-based results because Livneh-based results did not cover the entire validation period (Table 2.3). The results suggested that, despite the fundamental differences in modeling approaches, both models were able to well simulate the monthly hydropower generations with an R^2 around 0.75 during the calibration and validation periods. The models also performed reasonably well in other metrics (e.g., NSE and mean annual bias). The metrics during the validation period were lower relative to the calibration period, but they were still satisfactory for the purposes of this study. Further details about WRES and WMP model implementation can be referred to Zhou et al. (2022).

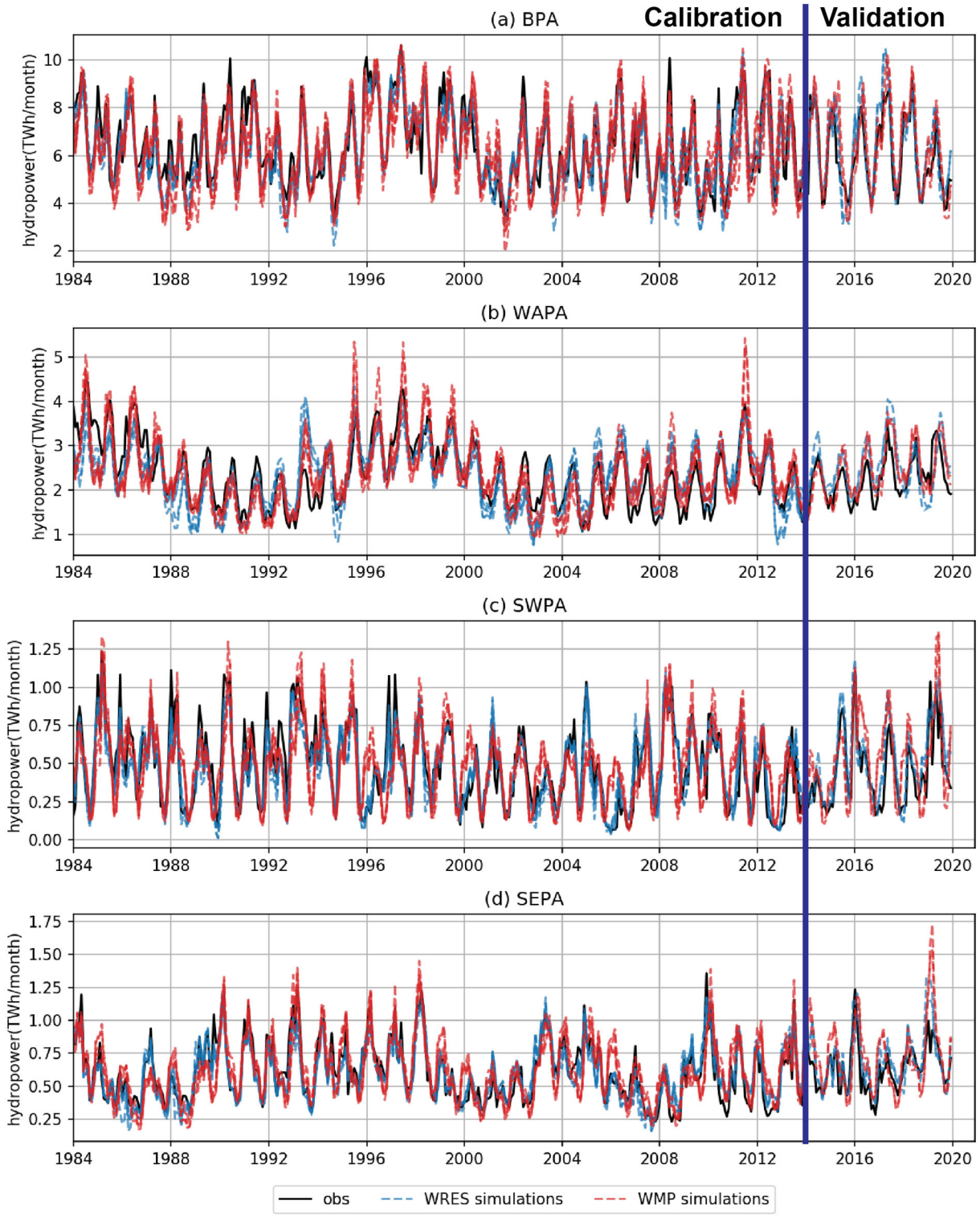


Figure 2.18. Performance of WRES and WMP.

Table 2.3. Performance of WRES and WMP

Power model	Hydrologic model	Meteor. obs.	Calibration					Validation				
			Calibration period	RMSE* (TWh/mth)	R^2	NSE	Mean annual bias	Validation period	RMSE (TWh/mth)	R^2	NSE	Mean annual bias
WRES	VIC	Daymet	1984–2013	0.92	0.83	0.82	−2.0%	2014–2019	1.18	0.79	0.58	7.9%
	VIC	Livneh	1984–2013	0.95	0.81	0.81	−1.4%	—	—	—	—	—
	PRMS	Daymet	1984–2013	1.02	0.79	0.78	−2.8%	2014–2019	1.15	0.81	0.60	7.4%
	PRMS	Livneh	1984–2013	1.07	0.77	0.76	−2.5%	—	—	—	—	—
WMP	VIC	Daymet	1984–2013	1.29	0.72	0.65	−2.2%	2014–2019	1.25	0.77	0.53	7.2%
	VIC	Livneh	1984–2013	1.24	0.72	0.68	−1.4%	—	—	—	—	—
	PRMS	Daymet	1984–2013	1.18	0.74	0.71	−1.4%	2014–2019	1.14	0.85	0.61	8.1%
	PRMS	Livneh	1984–2013	1.23	0.74	0.69	−2.3%	—	—	—	—	—
WRES mean			—	0.99	0.90	0.79	0.90	0.79	−2.2%	0.89	0.59	0.89
WMP mean			—	1.23	0.86	0.68	0.86	0.68	−1.8%	0.90	0.57	0.90

*Root mean square error (RMSE)

2.6.4 Model Limitations

2.6.4.1 WRES

Although WRES may reasonably simulate monthly hydropower generation for a majority of federal hydropower systems, it is a lumped model with more restrictions than reservoir-based models. To avoid confusion with the conventional reservoir-based models, WRES limitations are summarized as follows:

- WRES is a generalized model designed to simulate regional hydropower generation (and the corresponding watershed storage variation) from various federal hydropower study areas with minimal meteorological and hydrologic inputs. Most of the reservoir-specific aspects, such as operational constraints, competing water demand, flood risk management, ecosystem services, and reservoir sedimentation, cannot be addressed in WRES.

- Although WRES does not require reservoir parameters, it requires a complete, long-term, accurate observation of precipitation, runoff, and generation. For areas with insufficient and less accurate precipitation and runoff records (e.g., WAPA-5), WRES performance is poor.

- In the current WRES application, multivariate regression is used to generate an initial estimate of monthly generation prediction under the implicit assumption that this statistical relationship will not change significantly in the future hydro-climate conditions. This modeling component can be further improved by including more nonlinear operational conditions or using machine learning techniques.

- Although a conceptual watershed runoff storage variable is introduced in WRES, it should not be confused with the actual reservoir storage in the study area. Nevertheless, the watershed runoff storage variable does show correlation to the overall wetness/dryness in the system. Taking SWPA-2 as an example, both hydropower generation and watershed runoff storage are at a minimum during the 2006 drought.

- While it is assumed that the monthly runoff and generation can be converted through a simplified relationship (Eq. [2.6]), based on previous experience (Sale et al., 2012), such a relationship can be more complicated at a finer temporal resolution. Hydropower is jointly controlled by both head and

flow, so the linear relationship will only hold at coarser temporal scales and/or when the head variation is small (i.e., ROR hydropower plants).

- Given that the reservoir operation cannot be adjusted in WRES, the future hydropower and runoff storage projections made by WRES correspond to the current-day operation. The actual system should be able to accommodate flexibilities to extreme climate conditions with further adjustment to the reservoir operation. This operational flexibility can only be examined through site-specific models and was not examined in this study.

2.6.4.2 WMP

WMP is a process-based PM that relies on a large-scale, spatially distributed river-routing model, estimated sectoral and spatially distributed water demand, and generic operating rules scaled to individual reservoirs. These rules are scaled based on storage capacity, long-term mean monthly inflow and demand, and operating use. The model differs from operational process-based reservoir models that have detailed operating rules and coordinate operations between reservoirs across the entire water system. To avoid confusion with the conventional reservoir-based models, WMP limitations are summarized as follows:

- WMP should not be treated as a site-specific PM. Although MOSART-WM, the river routing reservoir operation component of WMP, considers information from each simulated dam (e.g., dam functionality, minimum and maximum storage, nameplate capacity, minimum environmental flow, spill and penstock distribution), the daily operations do not reflect the complex multi-objective rules applied in an operational setting (e.g., coordination for flood control, cascade system, operating agreements, conservation plans, shortage reductions, minimum power pool impacts, flex spill, navigation rules, lack-of-market spill, fish passage). The operating rules are generic and provide a reasonable redistribution of water resources in space and time through the water system. Efforts are being made to enhance the reservoir operations using data-driven operating rules with weekly time scale storage and release targets (Turner et al. 2020). In this study, the WMP operating rules were calibrated based on the historical baseline inflow and maintained throughout the simulations. This means that the WMP operations were not updated using projected future streamflow for hydropower generation in the near-term and mid-term future periods nor future water demand projections and priority in water uses and water supply allocations. Future work could explore incorporating uncertainty in model parameters throughout the modeling chain. In addition, the hydropower operations are presently driven by water availability and reservoir operating purposes. The modeling framework does not represent complex daily operations and specifically day-ahead electricity prices and utility's net load in the daily hydropower scheduling. In the future, information about hourly net load, electricity prices, and other ancillary power grid services would be needed to more accurately represent hydropower operations.

- WMP uses several assumptions to represent regional hydropower operations and generation. For hydraulic head estimation, the assumed shape of the reservoir was a simple tetrahedron, whereas each reservoir has a unique volume-head relationship. Generation is also constrained by ecological spill requirements and penstock capacities. Because future/potential applications could power system models, the WMP regional hydropower was calibrated to represent potential generation and not actual generation. This means that the regional hydropower was calibrated with respect to EIA-reported actual generation and complemented with a 4% operating reserve. Although this is consistent with typical power system operations, this number varies by region, season, regional generation portfolio, and operating reserve market.

- Uncertainties inherited from the integrated hydro-meteorological modeling can prevent WMP from having reasonable operations (e.g., for substantial positive or negative biases). The WMP hydropower

generation is essentially driven by water availability. Part of the two-step calibration process aims to alleviate hydrology simulation–driven uncertainties. For areas with insufficient and less accurate precipitation and runoff records (e.g., WAPA-5), WMP performance is also poor.

2.7 RESERVOIR EVAPORATION

A large amount of water can evaporate from the open-water surfaces of reservoirs, especially in arid/semi-arid regions (Friedrich et al., 2018). Zhao and Gao (2019) estimated that the average annual evaporation volume (gleaned from 1984–2015) over 721 major US reservoirs is 33.73×10^9 m^3, which is equivalent to 4.9% of the total storage capacity of these reservoirs. Under a warming environment, both the evaporation rate and evaporation volume are expected to increase, exacerbating the stresses brought on by limited water resources. Zhao et al. (2021) found that the average evaporation rate of 143 US reservoirs will increase by 15.7% from 2070 to 2099 (compared with 1985–2014) under the RCP8.5 emission scenario. Similarly, Wang et al. (2018) estimated that the global lake evaporation rate will increase 16% by the end of the twenty-first century. Therefore, it is essential to project future long-term reservoir evaporation loss to help inform water-resource management practices (e.g., reservoir operations, hydropower generation).

2.7.1 Approach

Figure 2.19 depicts the overall process framework for quantifying historical and future evaporation volumes. The reservoir evaporation volume is the product of the reservoir evaporation rate and surface area. For calculating a lake evaporation rate, the Lake Evaporation Model (LEM; Zhao and Gao, 2019) was used. LEM is built upon the Penman equation (Penman, 1948) for open-water evaporation estimation:

$$E = \frac{s(R_n - \Delta U) + \gamma \cdot f(u) \cdot (e_s - e_a)}{\lambda_v \cdot (\Delta + \gamma)}, \tag{2.14}$$

where E is the open-water evaporation rate (mm·d^{-1}), s is the slope of the saturation vapor pressure curve (kPa · °C^{-1}), R_n is the net radiation (MJ · m^{-2} · d^{-1}), ΔU is the heat storage changes of the water body (MJ · m^{-2} · d^{-1}), $f(u)$ is the wind function (MJ · m^{-2} · d^{-1} · kPa^{-1}), e_s is the saturated vapor pressure at air temperature (kPa), e_a is the air vapor pressure (kPa), λ_v is the latent heat of vaporization (MJ · kg^{-1}), and γ is the psychrometric constant (kPa · °C^{-1}).

The LEM can address two important effects associated with open-water evaporation in reservoirs. The first of these is the *fetch* effect. When air moves from land across a water body, its relative humidity increases because of the evaporation process on the water surface. Therefore, a fetch length was adopted in the wind function to represent the effect of air becoming moister when moving from land to water. The fetch length was calculated for each reservoir during each month by using reservoir vector data and monthly dominant-wind direction information derived from the North American Land Data Assimilation System (NLDAS). In Figure 2.20, the width is defined as the distance between the two reservoir-tangent lines parallel to the wind direction, and the fetch length is calculated by dividing the area by the width. As the reservoir area increases or decreases, the fetch length changes accordingly.

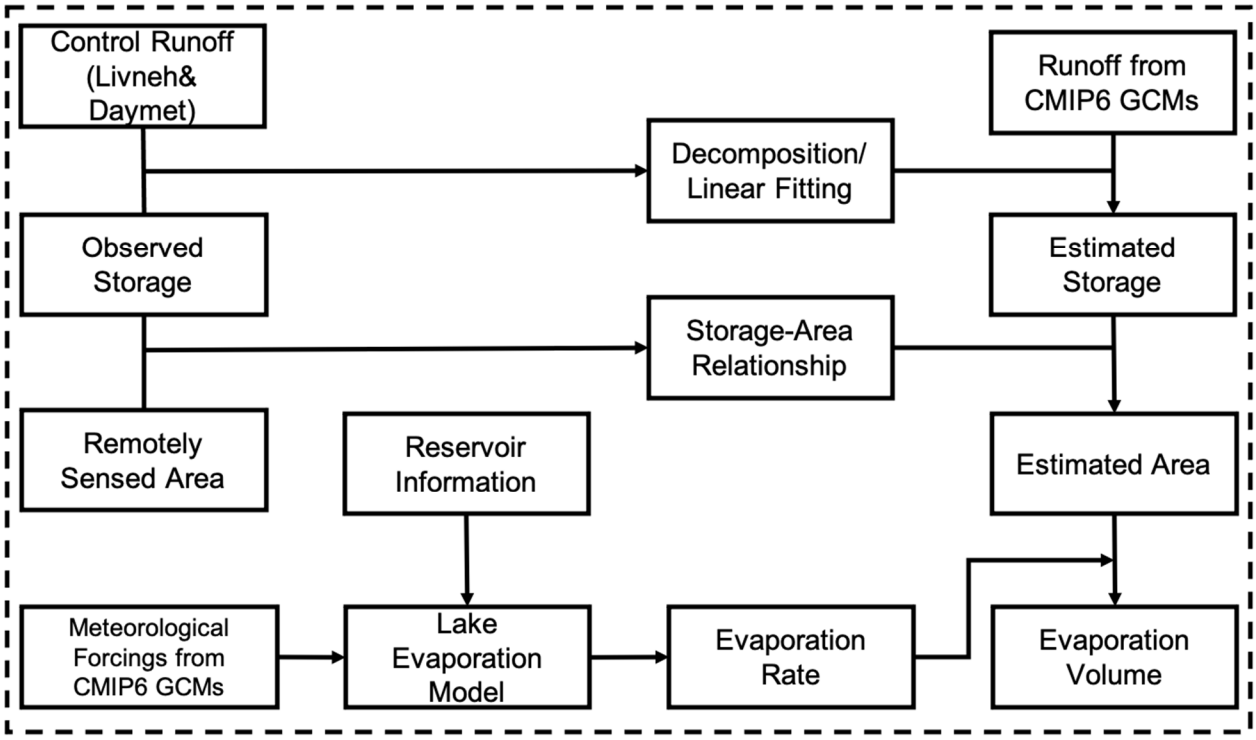


Figure 2.19. Flowchart for quantifying reservoir evaporation volumes.

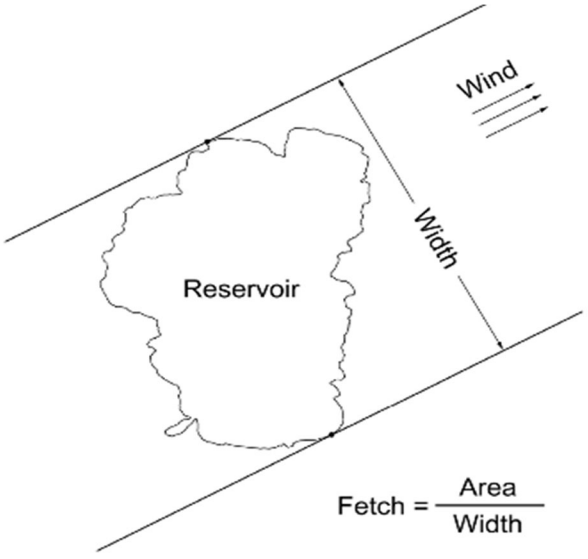


Figure 2.20. Calculation of the reservoir fetch for a given wind direction.

The second factor that LEM can address is heat storage quantification when applying the Penman equation to open-water evaporation. To represent the heat storage effect when calculating the evaporation rate, an equilibrium temperature algorithm is adopted. The equilibrium temperature is defined as the water temperature at which there is no heat exchange between air and water. If the water is under constant radiative forcing for long enough, the water will reach a steady state when the water temperature is equal to the equilibrium temperature. With this algorithm, the evaporation sequence was computed with four meteorological forcings (shortwave radiation, air temperature, vapor-pressure deficit, and wind speed), and driven by historic meteorological observations and downscaled climate projections (Section 2.4).

Based on their historical storage and storage variation patterns, reservoirs are classified into three categories using two parameters—α and β. Parameter α equals the range of storage variations divided by the average storage volume, and β equals the range of the seasonality over the range of the rolling average. Two independent time series—seasonality and rolling average—are obtained by decomposing the inflow time series. The window size of the rolling average depends on the reservoir category—12 months or 36 months—which is described in detail in the following paragraph. Overall, three categories of reservoirs are defined (Figure 2.21): Type 1: reservoirs with very small monthly storage variations compared with the total storage volume ($\alpha \leq 0.1$); Type 2: reservoirs where the storage variations are dominated by strong seasonality ($\alpha > 0.1$ and $\beta > 1$); and Type 3: reservoirs where the storage variations are affected by interannual variations ($\alpha > 0.1$ and $\beta \leq 1$).

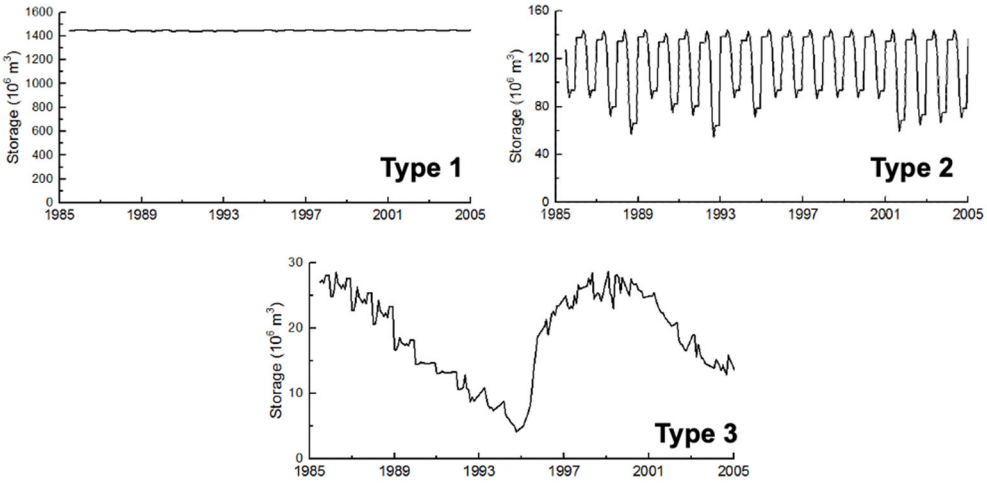


Figure 2.21. Reservoir classifications according to storage variations.

The future surface areas of reservoirs were inferred from the CMIP6 inflow data after a series of steps (Figure 2.22). Both the observed storage and the inflow time series (control run) were decomposed into a rolling average term, a seasonality term, and a noise term (Figure 2.22a, b). Considering the effect of water residence time, the rolling average term with a 12-month moving window was used for Type 1 and Type 2 reservoirs, and a 36-month moving window was used for Type 3 reservoirs. Without considering the seasonal climatology and noise terms, a linear regression relationship between the inflow rolling average and the storage rolling average for each reservoir was established (Figure 2.22c). To estimate the storage time series under CMIP6 scenarios, the regression relationship was first applied to the decomposed CMIP6 inflow rolling average to derive the storage rolling average (Figure 2.22d). Assuming the seasonality does not vary substantially year to year, seasonality was superimposed onto the storage rolling average to obtain the estimated monthly storage time series under different scenarios (Figure 2.22e). Based on the observed storage and remotely sensed area data sets of each reservoir, the area-volume (A-V) relationship—one of the most basic reservoir characteristics—was established first (Figure 2.22f). Then, after applying the estimated storage time series, the monthly time-series area values were projected, and the volume of evaporation loss for each reservoir was calculated by multiplying the reservoir surface area by the evaporation rate under the CMIP6 scenarios.

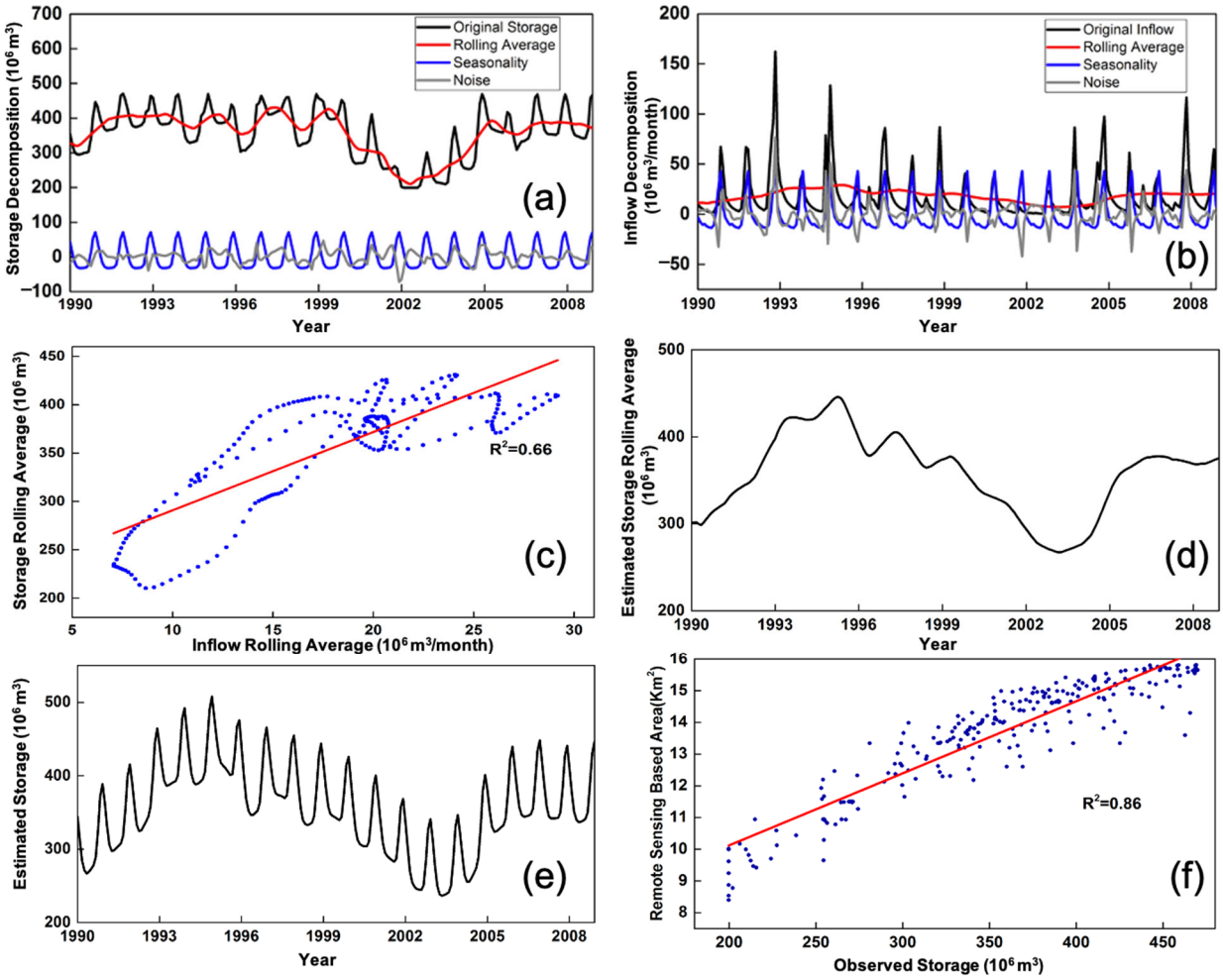


Figure 2.22. The process of reservoir storage projection based on inflow data: (a) observed storage decomposition, (b) inflow (control run) decomposition, (c) relationship between storage rolling average and inflow rolling average, (d) CMIP6 storage rolling average estimation, (e) CMIP6 storage estimation, and (f) relationship between observed storage and area (remotely sensed).

2.7.2 Model Validation

Figure 2.23 shows the validation results of the LEM-simulated evaporation rate during the historical period at locations in which in situ observations were available. For Lake Mead, there were apparent time lags and magnitude differences between the observed and modeled evaporation when heat storage was not considered. By incorporating the heat storage term into the algorithm, the R^2 increased from 0.29 to 0.84; the underestimation was likely caused by the advective energy (i.e., heat transferred by water movement). The R^2 increased from 0.73 to 0.93 at the Ross Barnett Reservoir, from 0.93 to 0.94 at Lake Calm, and from 0.55 to 0.87 at Lake Five-O. After adding the heat storage term, White Bear Lake was the only one of these five lakes with a decreased R^2 (from 0.69 to 0.64) and a slightly smaller root mean square error (RMSE). This can be attributed to an overestimation in 2014 because the R^2 in 2015 improved from 0.34 to 0.78. In general, these validation sites show significant improvement after considering the effects of heat storage.

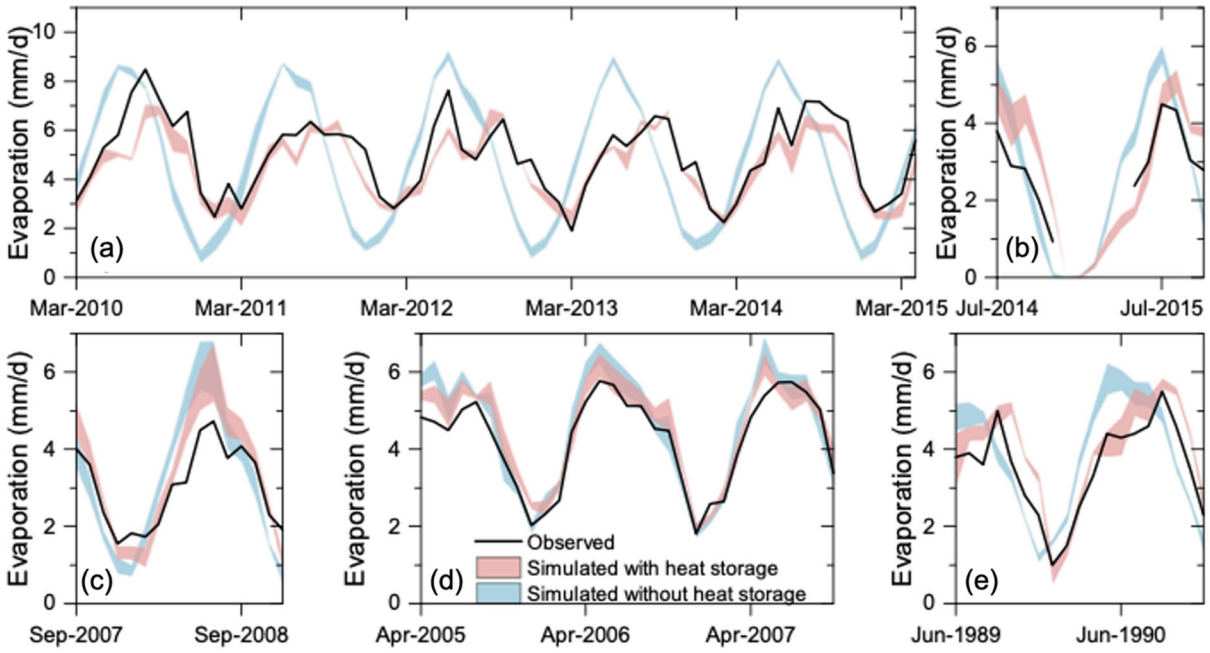


Figure 2.23. Validation of modeled and observed evaporation rate for (a) Lake Mead (Nevada/Arizona) with eddy covariance (EC) measurements, (b) White Bear Lake (Minnesota) with EC measurements, (c) Ross Barnett Reservoir (Mississippi) with EC measurements, (d) Lake Calm (Florida) with Bowen ratio energy budget (BREB) estimates, and (e) Lake Five-O (Florida) with BREB estimates. The shaded area represents the estimation uncertainty from different input forcing data sets (i.e., TerraClimate, NLDAS, and GLDAS).

As an essential step for projecting future reservoir storage while considering climate change, the storage-runoff relationships were also validated. The average R^2 between observed storage and predicted storage for all the reservoirs combined was 0.45, and the average normalized RMSE (NRMSE) was approximately 11.6%. For Type 3 reservoirs, which are mainly affected by interannual variations, the average R^2 was approximately 0.37, and the average NRMSE was approximately 16.5%. For Type 2 reservoirs with strong seasonality, the R^2 was 0.76, and the NRMSE was 10.0%. The R^2 was relatively high because the model retains original seasonal climatology information. For Type 1 reservoirs, the average R^2 was 0.31, but the average NRMSE was only 1.04%, which resulted in a low sensitivity for the evaporation loss estimates. After applying the A-V empirical relationship, the correlation of the remotely sensed area with the modeled area was approximately 0.23, and the average NRMSE was only approximately 8.6%. The change in the range of area was narrower than the change in the range of storage. Although the predicted correlation became relatively minor, its average NRMSE was only 8.6%. The absolute error of the projected area was relatively small, with limited impact on the estimation of evaporation loss.

2.7.3 Model Limitations

The limitations of the reservoir evaporation volume projections are summarized as follows:

- The projection of future reservoir storage was mainly based on the derived storage-runoff relationship. Although the two variables correlated well, the empirical model could not fully represent the complex operation rules. Furthermore, it assumed that future water demand will not increase and operation rules will remain the same.

- The uncertainty of the projected evaporation rate was mainly affected by the quality of the projected meteorological data (i.e., the downward radiation, temperature, wind speed, and vapor-pressure deficit). The differences in the meteorological data between Daymet and Livneh were examined, and wind speed values showed large systematic errors that could seriously affect evaporation rate calculations. For the same location at the same time, wind speed values across different MetFs could differ substantially. A difference of 0.1 m/s in wind speed could cause an evaporation rate difference of 0.02–0.04 mm/d, depending on the meteorological conditions at the given location.

- The quality of the reservoir area estimation depended on both the storage and the A-V relationship. The error in the A-V relationship mainly came from the observation errors, including observed historical storage and remotely sensed area, and from the fitting error caused by the reservoir shapes. The observation error was relatively small, and the storage value of most reservoirs fluctuated within a linear fitting range. Hence, the error in the A-V relationship was much smaller than the projected storage error.

- The uncertainties associated with the final volumetric evaporation were caused by the uncertainties of surface area and evaporation rate. For reservoir Type 1 and Type 2, the errors mainly came from the evaporation rate. For Type 3, they mainly came from the area estimation.

2.8 ENERGY DEMAND ANALYSIS

To evaluate climate-driven risks to federal power marketing, it is necessary to consider both the supply and the demand for federal hydropower. Previous subsections describe the methodology used to project federal hydropower supply changes (i.e., climate-driven changes in the volume of hydropower generated by the federal fleet). This subsection addresses the method used to estimate the relationship between temperature and the electricity demand from the PMA preference customers. The previously described projected annual changes in temperature and hydropower supply, combined with the temperature–demand relationships, help inform the PMAs about the future state of their customer base and water availability. This information is crucial for meeting long-term market planning objectives and potentially shorter-term operational constraints. Additionally, only BPA is required to meet the net load of its customers if they request it, meaning these customers will have their full demand portfolio met by BPA as opposed to purchasing electricity from another entity. Therefore, estimated changes in electricity demanded are most interesting for BPA because BPA is required to meet changes in load even with water supply shortages (or excess) or other external factors that make this electricity delivery costly or difficult. Nevertheless, the value of the energy and capacity products sold by other PMAs might also be affected by changes in the seasonal load patterns of their customers.

2.8.1 Approach

The use of statistical analysis to estimate the relationship between weather variables and electricity demand has been an active area of research for many decades (e.g., Davies, 1959). The variable included, functional form, and estimation method used depend on the subject of interest and the availability of data. This analysis focused on estimating the responsiveness of electricity sales to air temperature (as proxy for load or demand) of PMA preference customers to their end-use customers. As indicated by Huang and Gurney (2016) and Li et al. (2018), there are often large variations in the temperature effects on load within the regions of a country. Therefore, for this study, it was important to identify a data set that is representative of the electricity loads served by PMA preference customers. The PMAs do not serve loads directly;[6] instead, they sell wholesale power to preference customers, which are utilities with residential,

[6] One exception is the direct sales from BPA to direct-service industries. Historically, BPA served many direct-service industries (largely aluminum smelters). Today, BPA just serves one direct-service industry.

commercial, and industrial end-use customers. EIA Form 861 provides a time series of electricity sales data for a sample of PMA preference customers, which was used here as a proxy for the electricity demand or load served by these utilities. This is an annual data set (1990–2019) of total sales, and it also contains sales data disaggregated by sector: residential, commercial, and industrial.[7] BPA also provided a monthly data set for 2006–2019 of the total retail load served by their preference customers, which will be used instead of EIA Form 861.[8]

Annual heating degree days (HDDs) and cooling degree days (CDDs) were computed as the cumulative deviations of daily temperature from a reference balancing load point. HDDs contain information about the number of days with heating demand and the intensity of heating demand on those days. CDDs contain information about the number of days with cooling demand and the intensity of cooling demand on those days.

$$annual\ HDD = \sum_d max(BLP_H - tavg_d, 0), \tag{2.15}$$

$$annual\ CDD = \sum_d max(tavg_d - BLP_C, 0), \tag{2.16}$$

where *d* is the calendar day, *tavg* is the daily average temperature, and *BLP* is the balancing load point. *BLP* is the temperature assumed to trigger either heating demand (BLP_H) or cooling demand (BLP_C). The temperature interval between the two *BLP*s is the thermal comfort zone in which electricity is not used for heating or cooling. The *BLP* values used in this analysis were 59°F for heating and 64°F for cooling. These values were based on a study by Auffhammer et al. (2017) in which the 59°F–64°F temperature bin was found to lead to the lowest load, on average, in balancing authorities across the United States. The sales equation (proxy for demand equation) was estimated as follows:

$$\begin{aligned} ln(Sl\_cust_{i,t}) = {} & \alpha_i + \beta_1 \cdot ln(HDD_t) + \beta_2 \cdot ln(CDD_t) \\ & + \beta_3 \cdot ln(PRC_t) + \beta_4 \cdot ln(P_{i,t}) + \beta_5 \cdot Y_t + u_{i,t}\ , \end{aligned} \tag{2.17}$$

where *i* represents the preference customer (utility), *t* represents the year, *Sl_cust* is the annual sales per customer, *HDD* is the annual heating degree days (with *BLP* = 59°F), *CDD* is the annual cooling degree days (with *BLP* = 64°F), *PRC* is the precipitation, *P* is the average price, *Y* is the income per capita averaged across the counties in the utility's service territory, and $u_{i,t}$ is the residual term (i.e., the difference between the observed sales per customer and the value estimated by the explanatory variables in the equation). The BPA region's equation has some differences relative to Eq. (2.17): the dependent variable is monthly retail load, *t* is the month, the weather variables are monthly rather than annual summaries, and the equation includes monthly indicator variables to capture the seasonal pattern of sales, as well as an interaction term between CDDs and air conditioning penetration in the residential sector.

Equation (2.17) was estimated for the sales of PMA preference customers to their end-use customers in nine PMA subregions (APPENDIX H).[9] Whenever possible, the nine subregions corresponded to the regional delineation used by the PMAs for marketing and rate-setting purposes. In some cases, multiple power marketing regions had to be combined because the number of preference customers reporting sales data to EIA was too small to estimate their sales equation separately.

The degree day data in these equations reflected the temperature experienced in the service territory of the PMA preference customers, which can sometimes be far from the location of the federal hydropower

[7] Total sales = residential sales + commercial sales + industrial sales + transportation sales.
[8] Total retail load means all retail electric power consumption, including electric system losses and excluding nonfirm loads, interruptible loads, and transfer loads.
[9] The correspondence between the 18 study areas and nine subregions used for the power marketing analysis and further details of the hydropower facilities marketed by each PMA area are listed in APPENDIX H.

plants. Each of the equations was also estimated using two different MetFs (Daymet and Livneh). The city boundary Place shapefile from the US Census Bureau's TIGER data set[10] was used as the spatial filter to average the gridded temperature observations into a value that represents the service territory for a given preference customer. Since it averages temperatures within city boundary rather than across the entire counties of the utility's service territory, the Place filter helps capture more accurately the temperatures experienced by residential customers, which tend to have the most temperature-sensitive electricity demand among all customer types served by electric utilities.

The sales data were divided by number of customers to normalize across utilities of different sizes. Because all variables in the equation were in logarithms, the estimated slopes were interpreted as elasticities (i.e., the percent change in the dependent variable in response to a 1% increase in one of the explanatory variables).

All variables used to estimate the demand equations had two dimensions—a utility ID and a year—but the BPA data set also included the month. Data sets that consist of many subjects (utilities in this case) observed at different points in time are called *panel* data sets. There are multiple alternative specifications for regression models using panel data. The pooled model treated all observations as independent. The fixed effects model assumed that differences among subjects can be captured through unit-level intercepts (i.e., fixed effects) in the linear equations estimated. The random effects model posited that differences across subjects are random, and the corresponding model specification included an extra error term indexed by subject. Statistical tests were conducted to determine which panel data model specification was more appropriate for this particular data set. Standard errors were adjusted for clustering at the utility level.[11]

In summary, electricity sales data from a sample of PMA preference customers were used as a proxy for electricity demand to estimate a demand equation with price, income, degree days, and precipitation as explanatory variables. The estimates of demand responsiveness to degree days were then combined with the degree days from the downscaled climate projections described in Section 2.4 to estimate ranges of temperature-driven electricity demand changes in each of the nine PMA subregions out to 2059.

2.8.2 Model Validation

The demand equation tables in APPENDIX H include two model performance metrics. First, the R^2 measured the percentage of variability in the dependent variable (electricity sales or load per customer) explained by the model.[12] For the seasonal BPA load equations, based on monthly data, the R^2 ranged from 0.28 for the summer equations to approximately 0.5 for the winter equations. Thus, in the BPA region, the set of explanatory variables included in the regressions can explain a larger fraction of variability in electricity load in the winter months than in the summer months, at least with the linear model specification used in this analysis. The lower R^2 for the summer equations held even after including two additional explanatory variables (estimated air conditioning penetration and proxy variable for irrigation loads) in the summer regressions relative to the winter regressions.

[10] The census Place boundary includes both incorporated places and census-designated places.
[11] The Hausman test assesses whether the individual effects are correlated with the regressors. If the null hypothesis of no correlation cannot be rejected, the random effects model is preferred; if the null hypothesis is rejected, only the fixed effects model is consistent. A Breusch–Pagan Lagrange multiplier test considers the null hypothesis of zero variance across utilities. If not rejected, the pooled estimator is the preferred option.
[12] A modified version of the R^2, which is adjusted by the number of regressors, was used. This adjusted R^2 only increased upon introduction of a new regressor if the explanatory power of the model increased by more than what would be expected by chance; otherwise, it decreased.

For the annual sales equations estimated for the preference customers in the rest of PMA subregions, the R^2 ranged from 0.09 in the WAPA-SN/RM (Sierra Nevada/Rocky Mountains) subregion with the Daymet data set to 0.49 in the SEPA–Georgia/Alabama/South Carolina/Jim Woodruff (GA/AL/SC/JW) subregions. The poor model performance, in terms of R^2 for the WAPA-SN/RM subregion, is at least partially explained by a lower number of observations for that equation. The number of observations available for the WAPA-SN/RM and WAPA–Desert Southwest (DSW) subregions was much smaller than for the rest of equations because only a few of the PMA customers in these subregions report sales data in EIA Form 861.

The second validation metric presented along with the equation results is the NRMSE, which is a measure of predictive performance. The RMSE can be interpreted as the typical error (i.e., difference between observed and predicted value) expected when applying the equation to a new period or PMA customer. However, comparing the RMSE across different models applied to the same data set is challenging. The NRMSE is the RMSE divided by the average value of the dependent variable (monthly total retail load for customer for BPA or annual electricity sales for the other PMA subregions); it allows comparisons among the equations of the different PMA subregions. For the annual sales equations, the NRMSE ranged from 0.025 (SEPA-Cumberland) to 0.06 (WAPA-DSW). All the NRMSE values were reasonably low. Because they were normalized by the average observed value, they can be interpreted as meaning that the typical error was 2%–6% of the (log-transformed) sales value that the equations aim to explain.

The NRMSE was larger for BPA's seasonal load equations. It ranged from 0.1 in the winter equations to 0.2 in the spring and fall equations. However, there was also more variability to explain in the monthly data than in the annual data. The coefficient of variation (i.e., the ratio of the standard deviation and the mean) ranged from 0.05 to 0.20 for the annual sales data of WAPA, SWPA, and SEPA preference customers; for the monthly BPA customer load data, it ranged from 0.61 in the winter to 1 in the summer.

2.8.3 Model Limitations

Because annual sales data and annual degree day summaries were used, the analysis could only provide annual average results for the responsiveness of electricity demand to temperature. Other aspects of the temperature-load relationship (e.g., nonlinear effects of temperature beyond a threshold) could not be addressed. Similarly, estimating temperature-driven changes in peak load, a more useful metric for planning reserves and capacity investments than total electricity demand, would require hourly load data. The selected approach presents some limitations, which are summarized as follows:

- The annual sales data were a proxy for total retail load. Total retail load would be the ideal data for the dependent variable of the estimated electricity demand equations, but it was only available for BPA. For BPA, the annual sales data were within ±10% of the total retail load data for 87% of the observations.

- The annual sales data were not available for all the PMA preference customers. The number of customers for which data were available in each PMA subregion is shown in Table H.1. in APPENDIX H. Typically, the missing data will correspond to the smallest utilities because they do not file EIA Form 861.

- The demand analysis results presented in the rest of this study extrapolated the estimated responsiveness of electricity sales to temperature to future periods. The extrapolation implicitly assumed that there would be no substantial changes in the types of heating and cooling equipment being used or their energy efficiency, as well as no substantial growth of demand response programs. Improvements in heating and cooling technology and more widespread implementation of demand response programs would help reduce the dependency of electric load to temperatures.

2.9 ANALYSIS OF VARIANCE

2.9.1 Approach

The analysis of variance (ANOVA) technique was used to quantify the relative contribution of uncertainties in ensemble projections arising from different modeling choices (main factors) and their interactions. The ANOVA has been widely used in several studies to partition the total variance of ensembles by the sum of variances introduced by individual components and their interactions in the modeling framework. Based on the ANOVA, the total sum of squares can be described as a sum of squares (SS) derived from individual (main) effects (SSA, SSB, SSC) in a three component framework (where, A, B, and C are main factors), and their interaction terms (SSI = SSAB + SSAC + SSBC + SSABC). Bosshard et al. (2013), Chegwidden et al. (2019), Gangrade et al. (2020), Hattermann et al. (2018), and Queen et al. (2021) provide more details regarding the ANOVA applications in hydroclimate assessments.

This study focused on the variances explained from main effects and their first-order interaction terms. An ANOVA was applied for changes in variables, such as runoff, reservoir evaporation, and hydropower generation. The change signal for each variable and each ensemble member was calculated as a magnitude of a variable in the future period (2020–2059) minus the baseline period (1980–2019). The change signals were evaluated for any systematic differences arising from model setups. The changes for runoff and reservoir evaporation were studied using 48 ensemble members with four main components, including six CMs, two DSs (DBCCA and RegCM4), two MetFs (Daymet and Livneh), and two HMs (VIC and PRMS). For hydropower generation, a fifth component, PM (WRES and WMP), was added in the framework.

ANOVA results are shown for annual runoff change in SEPA-3 as an example (Figure 2.24). Additional results for other regions are discussed in the PMA-specific sections of this report. The panels show the total spread of change in runoff among 48 ensemble members when grouped by CM (a), DS (b), MetF (c), and HM (d). A qualitative analysis of panels (b), (c), and (d) indicated that the spread of projection was generally similar across DS, MetF, and HM in SEPA-3, although with some variations in the interquartile ranges. However, the spread was significantly different across CMs. For instance, the annual runoff simulated by ACCESS-CM2 projected a ~4 in. increase, whereas CNRM-ESM2 projected a ~0.2 in. decrease. This indicates that the choice of CM was more important for annual runoff than the choice of other modeling components in this study area. The ANOVA results suggest that CMs can explain ~90% of the variance in the annual runoff change. These results were based on the methodological choices incorporated in this study. Adjustment or additional modeling components may yield different results.

2.9.2 Limitations

This study did not explicitly separate the internal variability of GCMs from the total variance. To capture the internal variability within the GCMs, multiple GCM runs with different initial conditions need to be included, which could be a subject of future efforts. Alternative methods to estimate contribution of internal variability (Rupp et al., 2017; Chegwidden et al. 2019) may be considered in future assessments.

Furthermore, given the design of modeling framework, factors other than GCMs remain under-sampled. This could result in overestimation of contribution of GCM to the variance. Some alternative approaches, such as subsampling and bootstrapping approaches, could minimize these biases (Bosshard et al., 2013). These methods were not employed in this study but could be topics of future research.

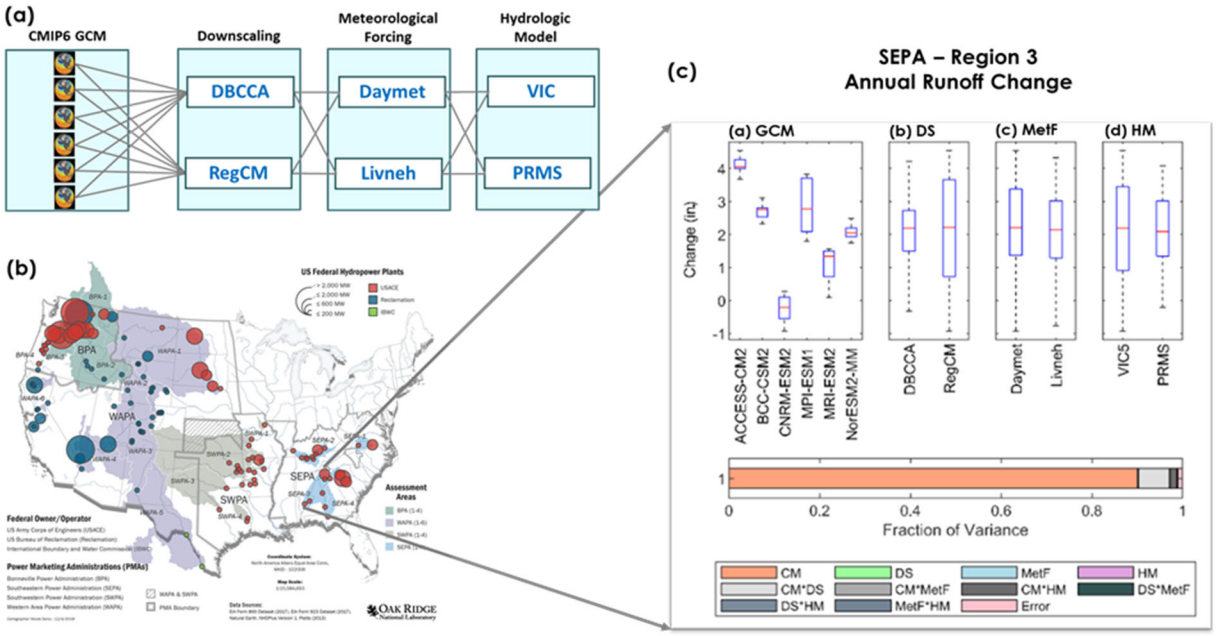


Figure 2.24. Example ANOVA results for one selected variable (annual runoff) for SEPA-3.

3. THE BPA REGION

Following the methodology described in Section 2, the results for federal hydropower plants marketed by the BPA are summarized in this section. Descriptions of the BPA study areas, federal hydropower systems, and power marketing practices are described in Section 3.1. The projected future atmospheric and hydrologic conditions in the BPA region are discussed in Section 3.2. The potential climate change impacts on federal hydropower generation and BPA regional energy demands are discussed in Section 3.3.

3.1 BPA REGIONAL CHARACTERISTICS

3.1.1 Federal Hydropower in the BPA Region

BPA is the largest PMA in the United States in terms of its total capacity (20,648 MW) and average annual generation (76.4 TWh/year) (Table 3.1). All federal hydropower projects in the BPA region are managed as a single FCRPS. The river drainages provide water to the FCRPS projects and cover large portions of the states of Washington and Oregon, almost all of Idaho, small parts of Montana and Wyoming, and almost 39,000 mi^2 in Canada (Figure 3.1).

Table 3.1. Summary of federal hydropower plants in the BPA region

Area	Area name	Number of plants			Capacity[a] (MW)	1980–2019 generation[b] (GWh/year)
		USACE	Reclamation	Total		
BPA-1	Upper Columbia	2	2	4	7,804	23,367
BPA-2	Snake River	5	5	10	3,756	12,035
BPA-3	Mid-Lower Columbia	5	2	7	8,613	39,235
BPA-4	Cascade Mountains	9	1	10	475	1,750
Total		21	10	31	20,648	76,388

[a] EIA 2019 total nameplate capacity, includes both conventional hydro and pumped storage.
[b] EIA and BPA average annual generation from 1980 to 2019, conventional hydro only.

The Columbia River basin is rich in water resources, representing the most heavily developed region for hydropower in the United States. Overall, there is a total of 34,300 MW federal and nonfederal hydropower in the Pacific Northwest region, accounting for 33% of the total US hydropower capacity (Johnson, 2021). Among these hydropower projects, 60% of the capacity is provided by the federal hydropower plants. The federal hydropower plants in the region are owned and operated by either USACE or Reclamation (Table 3.1). Some of the largest federal hydropower plants are located on the Columbia River. For example, Reclamation's Grand Coulee Dam is the largest federal hydropower plant in the United States, with an installed conventional capacity of 6,495 MW, and another 314 MW of pumped storage capacity. The largest USACE hydropower plant is the 2,456 MW Chief Joseph, which is immediately downstream of Grand Coulee. A complete list of federal hydropower plants in the BPA region is provided in APPENDIX B.

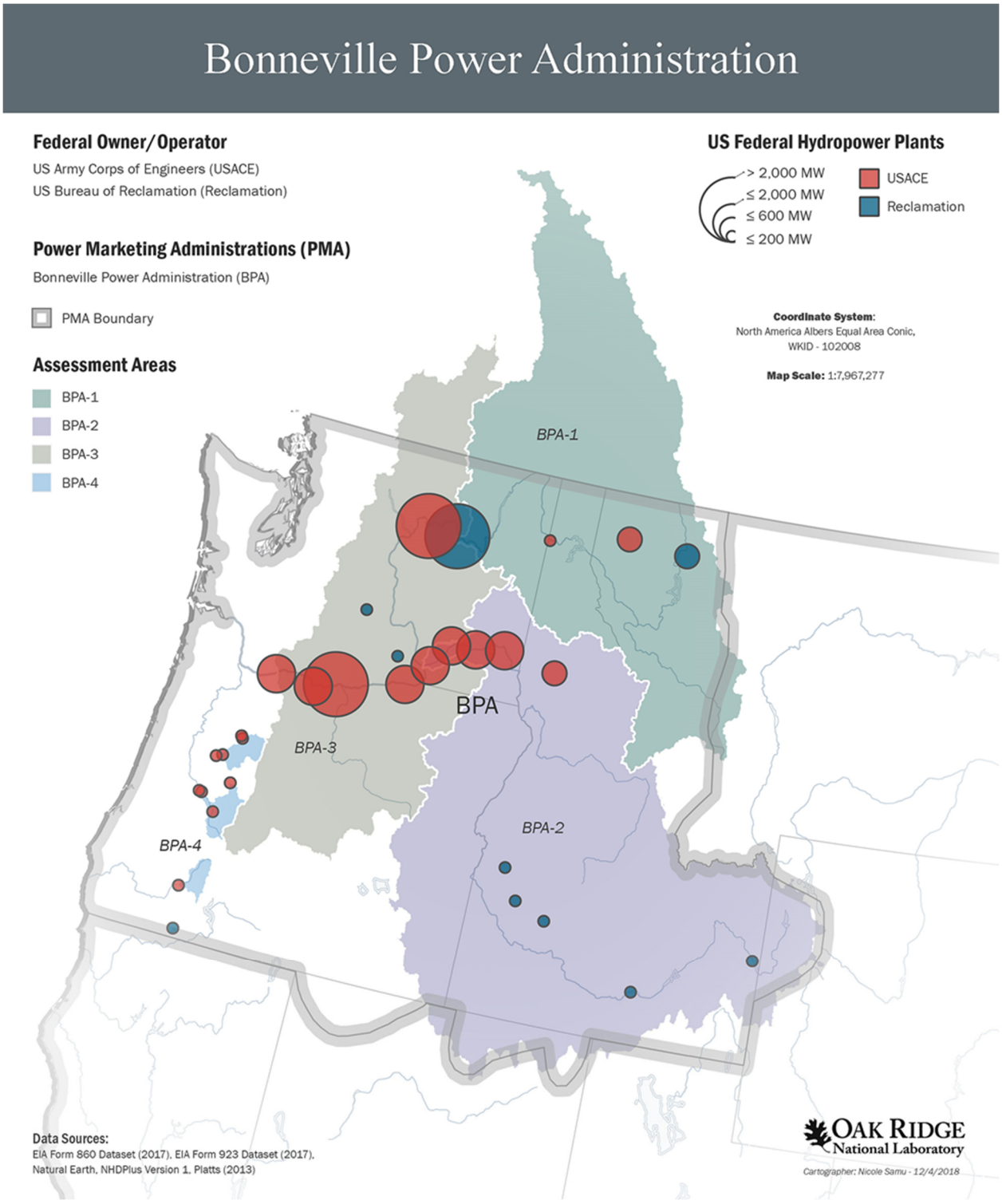


Figure 3.1. Map of the federal hydropower plants and four study areas in the BPA region.

Aging infrastructure and rising O&M costs are serious concerns for the federal hydropower plants in the BPA region. The oldest USACE project within this region is Bonneville Dam, which began operating in 1938, and the last USACE project constructed was Lost Creek in 1977. Reclamation's hydropower plants in this region were constructed between 1909 (Minidoka) and 1964 (Green Springs). Aging infrastructure

results in higher O&M costs, especially for non-routine maintenance expenses. To curb that trend, increases in capital expenditures to modernize the generation assets would be required. However, as of 2016, there was a substantial backlog between programmed capital expenditures and recommended expenditures according to asset performance models (Connolly, 2016). Higher O&M costs need to be recovered through the electricity rates that BPA charges its customers; thus, they lead to rate increases, which can damage the competitiveness of BPA's products. The high O&M costs, coupled with the goal to maintain competitive rates, may reduce BPA's abilities to balance other emerging issues, such as the need to adjust the operations and invest in new technologies to maximize the value of the FCRPS assets in a grid with large volumes of variable renewable generation.

Almost all hydropower plants in the BPA region are part of multipurpose water developments where hydropower represents one of many services a reservoir must provide. Balancing the operating and policy complexities of providing multiple services (e.g., flood control, navigation, irrigation, fish and wildlife protection, and municipal and industrial uses) remains a key objective for the federal dams in the BPA region. BPA, USACE, and Reclamation comprise the River Management Joint Operations Committee (RMJOC) tasked with long-term planning and operational decisions that affect the Columbia River and Snake River basins.

Some management decisions in this area are subject to international considerations concerning the use of the Columbia River. Around 35% of the total runoff (as measured at The Dalles Dam) flowing through the Columbia River originates in Canadian watersheds; therefore, the basin is managed as an international resource. The Columbia River Treaty coordination between Canada and the United States on power and flood control provides $100 million of annual mutual benefits across the Columbia River basin (USACE and BPA 2009).

In this study, the entire BPA region is subdivided into four study areas based on power system and watershed boundaries (Figure 3.1):

- **BPA-1 Upper Columbia**: The Upper Columbia River system includes the main stem of the Columbia River and its tributaries, including Grand Coulee Dam. The total drainage area of BPA-1 is 75,058 mi^2, and 45% of this area is in Canada. The major tributaries where federal hydropower plants are located are the Pend Oreille and Kootenai River basins. The watershed of this uppermost area is predominantly mountainous, with a median elevation of 4,600 ft, and maximum elevations exceed 11,000 ft in the Canadian Rockies. Land cover is evergreen (69%) and deciduous (13%) needle leaf forest, with minor amounts of grasslands and mixed forests (6% each). Water management in this study area is mainly controlled by an international treaty with a smaller portion controlled by US projects on the Pend Oreille, Kootenai, Flathead, Kootenai, and Spokane Rivers.

- **BPA-2 Snake River**: The Snake River system is upstream of the Columbia–Snake confluence. The total drainage area of BPA-2 is approximately 108,000 mi^2, most of which is in Idaho. Federal hydropower plants are located on the main stem of the lower and upper Snake River, as well as on some of its major tributaries—the Clearwater, Boise, and Payette Rivers. Topography here is a mix of high plains and mountains, with a median elevation of 5,112 ft and a maximum elevation higher than 12,000 ft. BPA-2 is a semi-arid area with relatively diverse land cover, including grassland (52%), evergreen forest (27%), cropland (12%), and closed scrubland (9%).

- **BPA-3 Mid-Lower Columbia**: The mid– and lower–Columbia River systems span from the most downstream federal hydropower plant, Bonneville Dam, up to the tailwater of Grand Coulee. The total drainage area of BPA-3 is 242,199 mi^2 (including the two upstream areas, BPA-1 and BPA-2). All federal hydropower projects in BPA-3 are on the main stem of the Columbia River. Topography in the watershed is a mix of high plains and mountains. Median elevation is 2,927 ft, and the highest

elevations are upstream in other areas (BPA-1 and BPA-2). BPA-3 is also a semi-arid area. Land cover in BPA-3 is closed scrubland (52%), grasslands (38%), cropland (14%), and woody savanna (7%).

- **BPA-4 Cascade Mountains**: The Cascade Mountains projects are located in four separate watersheds on the western slope of the Cascade Mountains of Oregon. Two of these are in the Willamette River basin, and two are farther south on other coastal rivers. The aggregate median elevation of this area is 3,773 ft, and the maximum elevation in this part of the Cascade Range is 9,459 ft. Land cover is primarily evergreen needle leaf forest (94%).

3.1.2 Power Marketing by BPA

BPA was the first PMA, created by the Bonneville Project Act of 1937, and is the largest in terms of marketed hydropower capacity and annual generation. BPA markets the electricity from the 31 hydropower plants in the FCRPS and a nuclear plant. It also operates more than 15,000 mi of transmission lines—approximately 75% of the region's high-voltage transmission—and acts as a balancing authority (i.e., it is responsible for maintaining continuous balance between electricity load and generation in its region).

A general description of a PMA and its marketing practices is presented in Section 1.2.2. For BPA, several pieces of legislation over the past 50 years have resulted in substantial differences in its obligations and funding mechanism relative to the rest of PMAs. First, with the Columbia River Transmission System Act of 1974, BPA became a self-financing agency and was granted permanent Treasury borrowing authority. Self-financing meant that BPA would deposit its power and transmission sales revenue in a revolving fund (rather than the General Fund of the Treasury) and use this revenue directly to cover its O&M expenses rather than receiving appropriations to cover its costs. The borrowing authority ceiling has been raised several times since 1974 and stands at $7.7 billion; BPA repays these loans with interest at market rates. Second, the Northwest Power Act of 1980 directed BPA to serve the net load requirements of its preference customers and gave it authority to purchase power from sources outside the FCRPS to meet that obligation, prioritizing energy efficiency and renewables to cover additional load growth.[13] It also directed BPA to implement a program to protect and enhance fish and wildlife in the Columbia River basin. Third, the Energy Policy Act of 1992 gave BPA authority to use power sales revenues for financing some power plant upgrades, whereas the other PMAs depend on appropriations or funding from the preference customers.

As of 2021, BPA had 142 customers, 124 of which are cooperatives, municipalities, or public utility districts. The rest include federal agencies (7), IOUs (6), port districts (1), DSIs (1), and tribal utilities (3). The power sales contracts in place between BPA and its customers cover the period 2008–2028 and were the result of the Regional Dialogue Policy (Kao et al., 2016). Preference customers could choose among three products: load following, slice/block, and block. With the load following product, the customer purchases firm power for all its net load (i.e., total retail load minus nonfederal firm resources owned by the customer) from BPA. The block product provides a planned amount of firm power in a flat block, which delivers an equal amount of power in all hours of the year, or a shaped block, which delivers different amounts each month based on the monthly shape of the customer's retail load (there can also be some diurnal shaping with up to 60% of the energy provided during high load hours). The slice/block product combines a block of firm power that varies by month based on the monthly shape of the customer's retail load (but is flat within the month) and a percentage (slice) of power in the shape of BPA's generation from the FCRPS over the year. The slice portion varies depending on hydrologic

[13] The Northwest Power Act also led to the creation of BPA's Residential Exchange Program through which BPA makes payments to IOUs so that their residential and small farm customers pay similar rates for their power as BPA preference customers.

conditions and operational constraints. As of 2021, out of BPA's 135 customers with Regional Dialogue contracts, 119 were load-following customers, 13 were slice/block customers, and 3 were block customers.[14]

Table 3.2 shows BPA power rates for FY 2022 and FY 2023. The Priority Firm rate is the rate paid by preference customers. The tier 1 rate corresponds to the cost of electricity generated by the Tier 1 system resources (i.e., the FCRPS), and the tier 2 rate corresponds to other energy acquired by BPA to meet the rest of the load-following customers' needs.[15] The exchange rate is the rate at which BPA sells power for BPA's Residential Exchange Program. The Industrial Firm rate is available to large industrial customers that take energy directly from BPA to power their operations. The New Resource Firm rate applies to IOUs for loads that do not qualify under the Residential Exchange Program and to power requested by preference customers to serve any new large (>10 MW) single load.

Table 3.2. BPA wholesale power rates (FY 2022 and FY 2023)

Type	Average rate ($/MWh)
Priority Firm: average tier 1 rate	34.93
Priority Firm: average tier 1 + tier 2 rate	34.87
Priority Firm: exchange rate	62.00
Industrial Firm	40.69
New Resource Firm	78.84

Providing competitive power products and services is one of the four goals prioritized by BPA in its 2018–2023 Strategic Plan.[16] The plan recognizes that the historical cost advantage of federal power vs. other sources in the Pacific Northwest has declined in recent years because of a combination of low wholesale electricity prices and flat power demand. Managing the costs of the power program and the fish and wildlife program and pursuing additional revenue from new markets for clean capacity and flexibility are among the strategies pursued to ensure the continued competitiveness of BPA's power products.

In an effort to optimize the value of the FRCPS assets, BPA plans to join CAISO's Western EIM in 2022. In the EIM, dispatch decisions are made at 5-min intervals, providing more operational flexibility by helping manage transmission grid congestion. In addition, BPA's hydropower would be compensated for balancing sub-hourly solar and wind power fluctuations or could be displaced by lower cost energy, saving valuable fuel for future hours. Additionally, BPA is supportive of CAISO's ongoing efforts to establish an imbalance reserve product in its day-ahead market horizon that will help to more efficiently and effectively commit units available to the CAISO balancing authority area to better ensure reliable market outcomes and integration of increasing amounts of variable energy resources. CAISO has recently reinitiated its Day-Ahead Market Enhancements policy initiative in which this product is being considered.

[14] https://www.bpa.gov/p/Generation/White-Book/wb/2019-WBK-Summary.pdf
[15] In the 2020 edition of its Resource Program analysis, BPA concluded that the least cost mix of resources to meet BPA's expected energy needs is a combination of energy efficiency and energy purchased from the wholesale market (rather than building or maintaining additional generation resources).
[16] https://www.bpa.gov/StrategicPlan/StrategicPlan/2018-Strategic-Plan.pdf

3.2 FUTURE CLIMATE IN THE BPA REGION

3.2.1 Regional Climate Projections

The annual temperatures in the baseline (1980–2019) and future (2020–2059) periods are presented in Figure 3.2. The gray lines show the annual mean temperature from 24 sets of downscaled climate projections, the black line represents the multimodel median, and the blue and green lines show the Daymet and Livneh historical observations, respectively. These plots depict interannual variability and annual trends for projections, as well as the observations.[17] The climate models are not expected to replicate the exact timing of historical interannual and decadal variability. Therefore, they do not completely follow the historical interannual values. The corresponding probability distribution for the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between baseline and future periods is statistically significant.

In all four assessment areas (BPA-1–4), the ensemble model median captures the increasing trend in annual mean temperature from the observations during the historical period. However, there is a large variability across the ensemble members, which reflects the importance of using a multimodel ensemble approach. There are also differences in inter-annual variability between Livneh and Daymet observations, which may arise from different processing techniques used to generate these observations. These differences between Livneh and other observations over the western–southwestern United States have been discussed in detail by Walton and Hall (2018). In the future period, annual mean temperature is projected to continue to increase across all the four areas. A comparison between the probability distributions for the multimodel mean annual temperature during the historical (black line) and future (purple line) periods also shows a statistically significant shift toward a warmer climate across all the BPA areas (Figure 3.2).

The annual and seasonal changes in temperature are further summarized in Figure 3.3. The change is presented as differences in temperature during the future periods (2020–2039 near-term and 2040–2059 mid-term) compared with the baseline period (1980–2019) in degrees Fahrenheit. Each box plot shows the spread across 24 sets of downscaled climate projections; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to the 75th percentile). Ensemble members outside of the maximum whisker length denoted by x are considered statistical outliers (but should not be excluded in the context of climate change).

Annual and seasonal temperature is projected to increase across all ensemble members during the near-term and mid-term future periods. A relatively large increase is projected during the mid-term future period. The multimodel median annual temperature in the BPA region is projected to increase by approximately 1.5°F to 3°F in the near-term and up to 5.5°F in the mid-term future periods. The highest median increase is projected during the summer and fall seasons (up to ~5.5°F for the mid-term) across the four areas, whereas the winter and spring show a comparatively lower increase (up to ~4°F in the mid-term) over all areas except BPA-2, where an increase of up to ~5°F is shown.

The annual precipitation (Figure 3.4) also shows high interannual and intermodel variability. A comparison of the probability distribution between the baseline and future periods indicates a statistically significant increase in precipitation across all four BPA areas. The annual and seasonal changes of

[17] Interannual variability refers to the differences seen between the years, and the trend refers to the trend in temperature seen from 1980 to 2059 in the models and 1980 to 2019 in the observations. This applies to Sections 4.2.1, 5.2.1, and 6.2.1.

precipitation are summarized in Figure 3.5. In the near-term (2020–2039) and mid-term (2040–2059) future periods, the ensemble median precipitation is projected to increase annually.

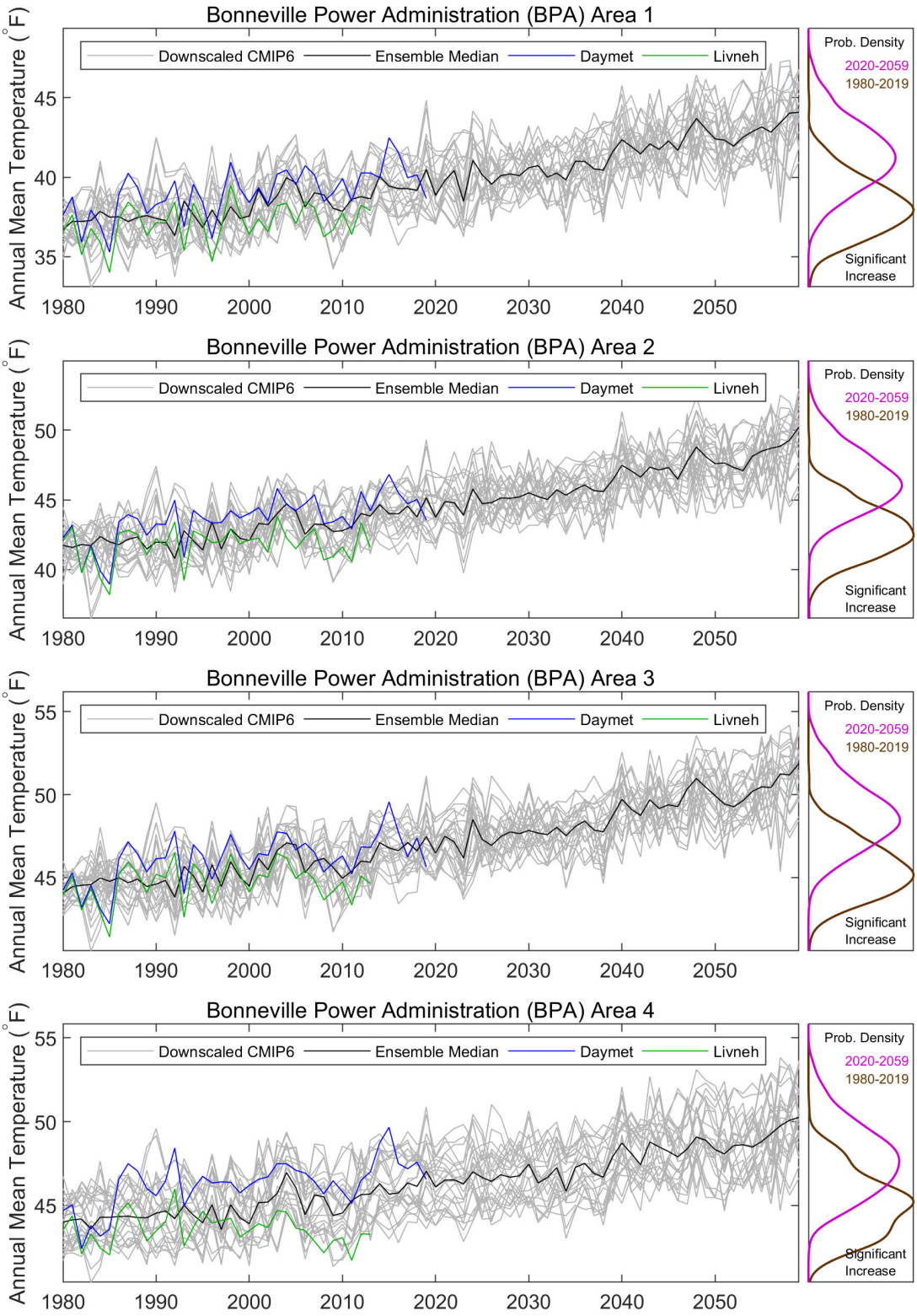


Figure 3.2. Projected annual mean temperature in the BPA region.

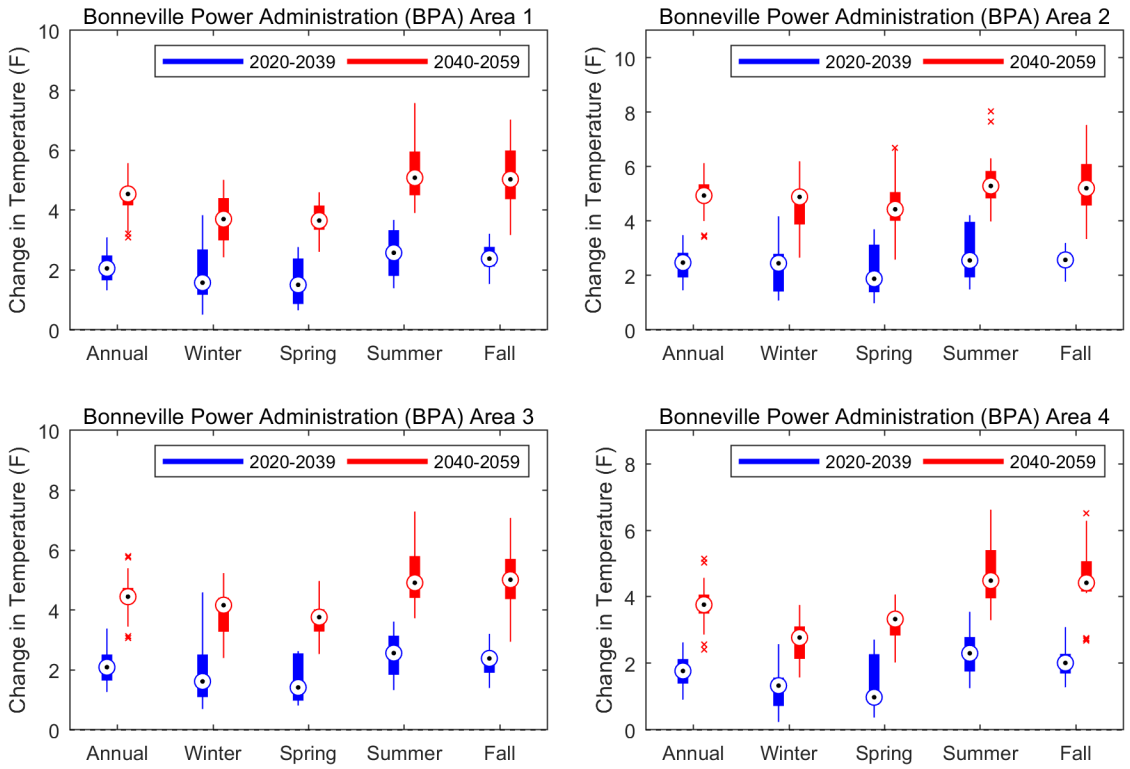


Figure 3.3. Projected change of annual and seasonal mean temperature in the BPA region.

On a seasonal scale in the near-term future period, the ensemble median precipitation is projected to increase during the winter, spring, and fall across all BPA areas, which are historically the wettest times of year in the BPA region. Meanwhile, median precipitation shows no change to a small decrease during the summer, although model spread is rather large on a percentage basis, and summer is historically a dry season in all BPA areas. Comparatively, the mid-term future period on a seasonal scale shows a projected increase in the ensemble median precipitation for all seasons across most of the BPA areas. The exceptions include the summer in BPA-1, and the spring and fall in BPA-4, resulting in negligible to no change in median precipitation (Figure 3.5).

These projections agree with other climate change assessments over the Pacific Northwest (Jiang et al., 2018; Huang and Ullrich, 2017; Rupp et al., 2016; Mote and Salathé, 2010). Using the CMIP5-based statistically downscaled climate projections, Jiang et al. (2018) suggested increasing and noticeable changes in temperature and precipitation during the near-term and mid-term future periods. Their analysis also suggests an increase in mean precipitation during the winter and a decrease during the summer. Rupp et al. (2016) and Mote and Salathé (2010) reported similar findings by directly using ensemble of CMIP5 and CMIP3 GCMs, respectively. Similarly, Huang and Ullrich (2017) found an increase in extreme precipitation during the winter using the variable-resolution Community Earth System Model under the RCP8.5 scenario over the Pacific Northwest. Jiang et al. (2018) used six CMIP5 GCMs downscaled by four statistical downscaling methods. Despite these methodological differences, the projected changes in seasonal precipitation agree with the current assessment.

In all BPA areas, the important signals of change include (1) an increase in temperature in all seasons, and (2) an increase in precipitation in the winter, spring, and fall, which more than offsets any decreases in summer precipitation. Since summer is already the dry season in the BPA region, the most dominant

contributors to the summer streamflow are runoff and snowmelt, which are discussed in the following section.

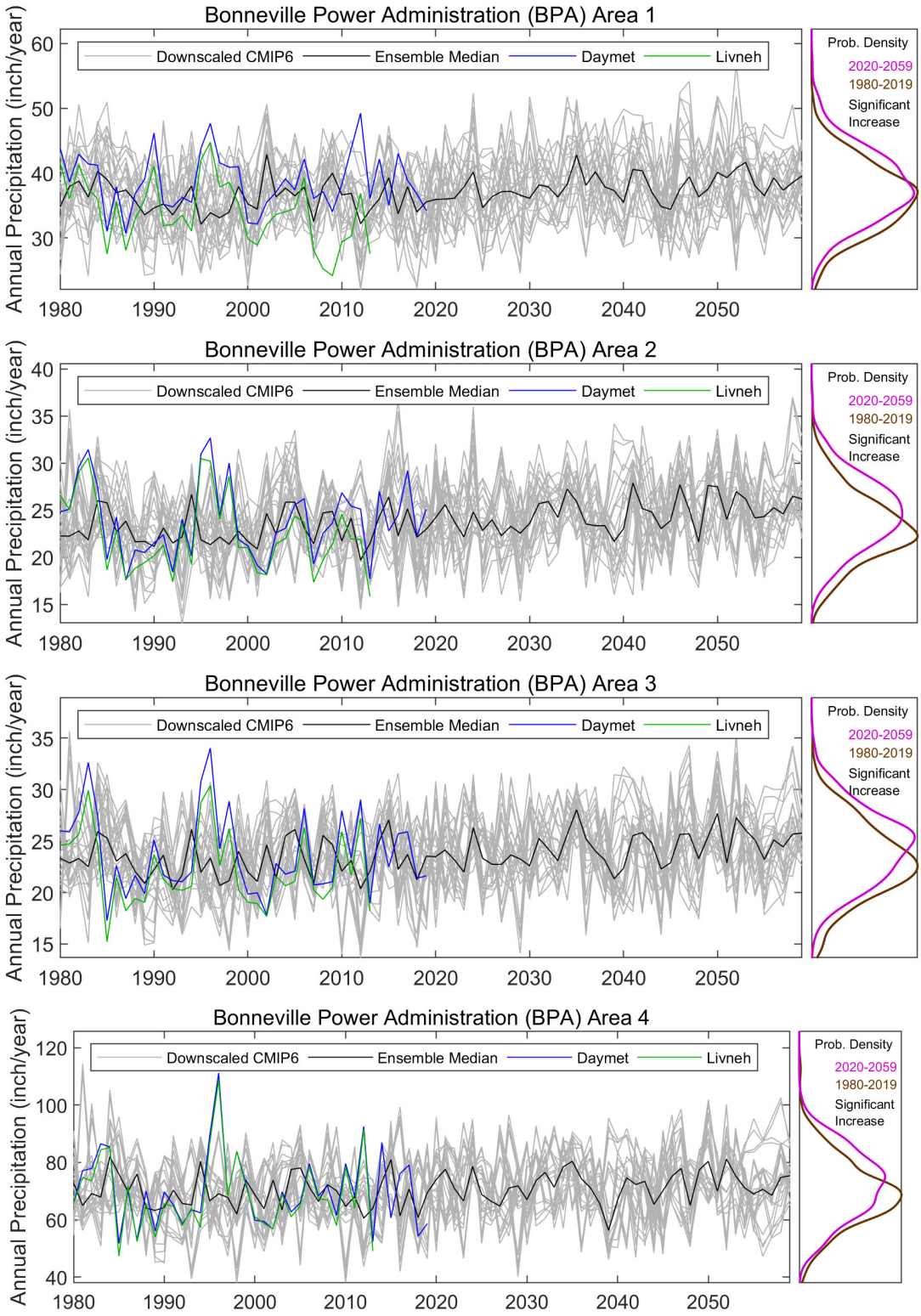


Figure 3.4. Projected annual total precipitation in the BPA region.

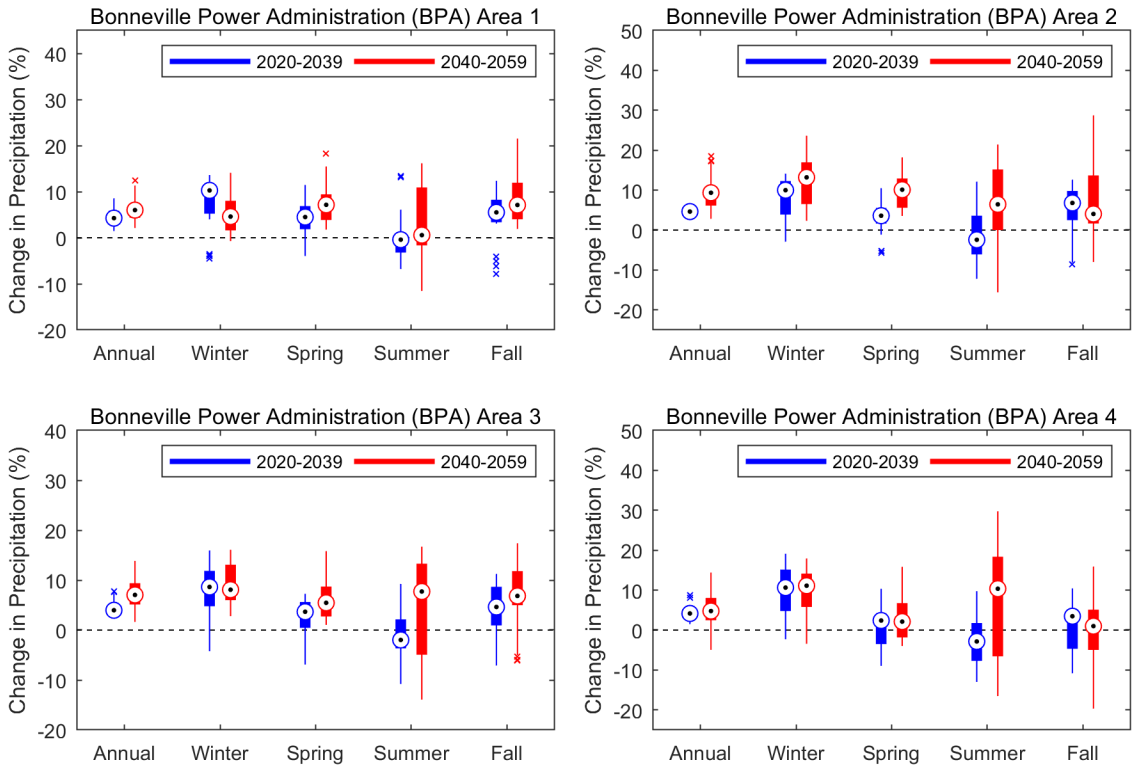


Figure 3.5. Projected change of annual and seasonal total precipitation in the BPA region.

3.2.2 Regional Hydrologic Projections

Changes in hydrologic regimes are generally expected to be driven by changes in temperature and precipitation among other factors such as changes in land use/land cover, urbanization, and so on. Recent assessments (Mote et al., 2005, 2018; Schnorbus et al., 2014) have indicated a dramatic decline in the snowpack in the western United States attributed to earlier spring snowmelt, delayed winter snowfall, and increasing temperatures. With these shifts in precipitation and temperature and the increased fraction of winter rainfall, the subsequent changes would affect the timing and total amount of water flowing into the rivers, with an increased risk of hydrologic extremes in the region (Leng et al., 2016; Wrzesien and Pavelsky, 2020; Chegwidden et al., 2020; Queen et al., 2021).

The runoff projections for each of the BPA areas are illustrated in Figure 3.6 and Figure 3.7. The annual total runoff in baseline (1980–2019) and future (2020–2059) periods is presented in Figure 3.6. The gray lines show annual runoff from 48 sets of hydrologic projections, the black line represents the multimodel median, and the blue and green lines show the control simulations driven by Daymet and Livneh observations, respectively. The corresponding probability distributions are compared in the right panel. The annual and seasonal changes in runoff are further summarized in Figure 3.7. Change is defined as the percentage difference of future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across 48 ensemble members; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length are considered as statistical outliers (but should not be excluded in the context of climate change).

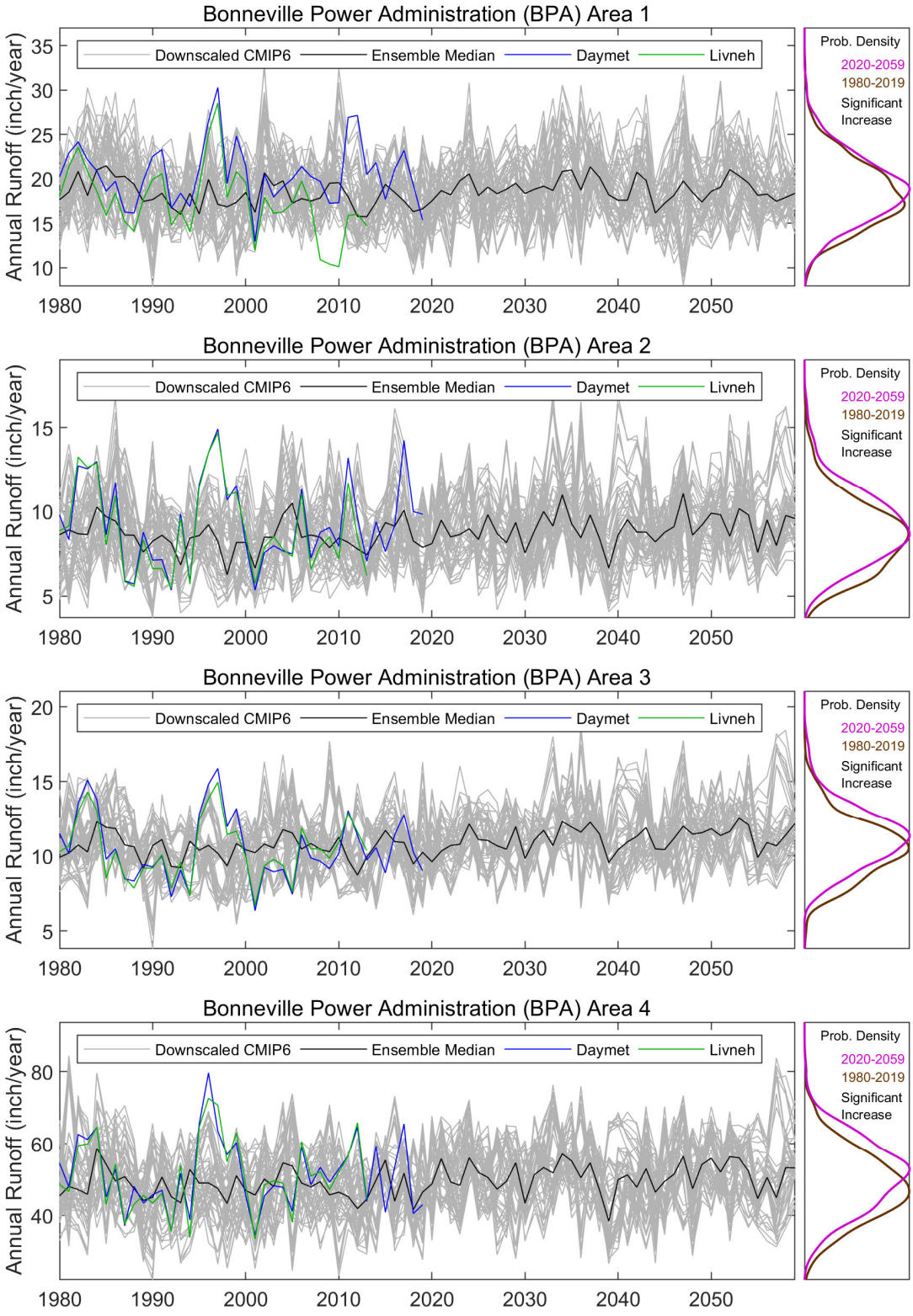


Figure 3.6. Projected annual total runoff in the BPA region.

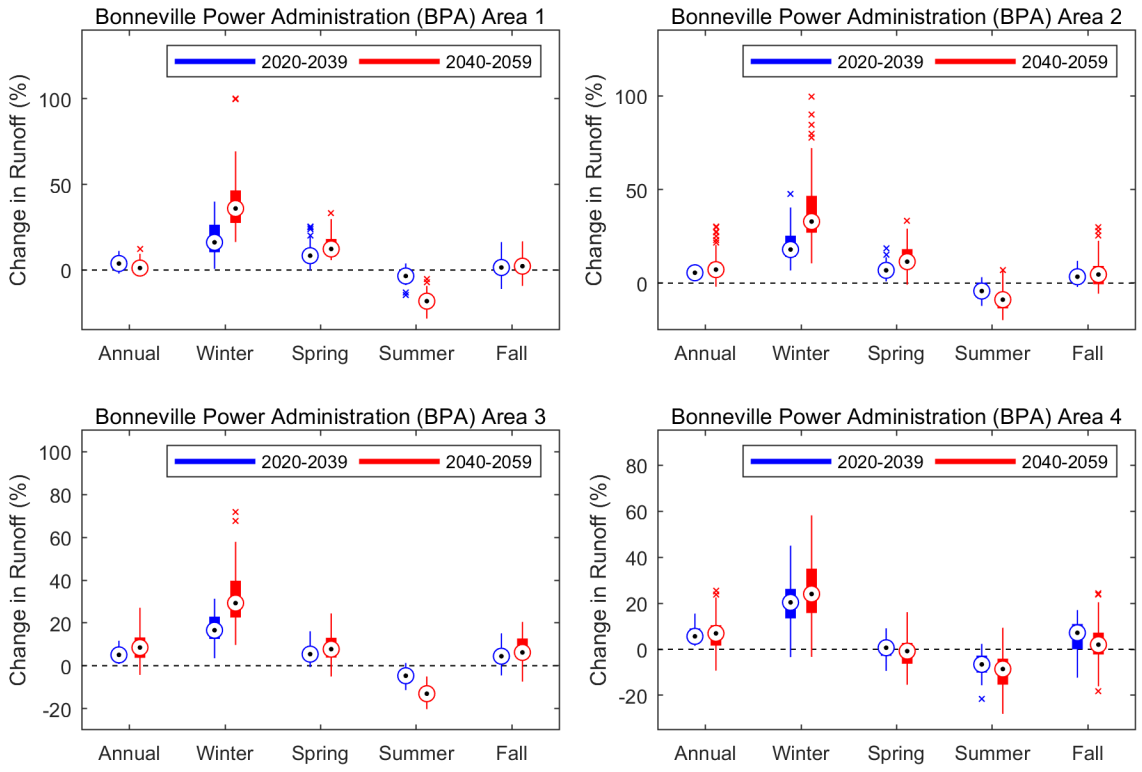


Figure 3.7. Projected change of annual and seasonal total runoff in the BPA region.

Consistent with both historical and future precipitation projections, the interannual variability of total runoff is also high. However, the shifts of probability density functions (Figure 3.6) suggest that the annual runoff is projected to increase in the future period, and the distributions are statistically different. Additionally, the shapes of probability density functions in all BPA areas are similar, which indicates that the range of interannual variability of future annual runoff should generally be similar to the baseline period. The maximum increase in median annual runoff is projected in the range of 3.1% to 6.7% across all BPA areas. More specifically, the median annual runoff is projected to increase across all study areas in the range of 3.9% to 6.5% in the near-term (2020–2039) and 1.9% to 9.4% in the mid-term (2040–2059) future periods (Figure 3.7). The range of annual runoff change is higher in the mid-term future period, indicating greater uncertainties in future runoff projections.

At the seasonal scale, the strongest changes projected in the BPA region are the increase in winter runoff and the decrease in summer runoff (Figure 3.7). The winter runoff is projected to increase in the range of 16% to 20% in the near-term and 24% to 36% in the mid-term future periods. Increases in spring runoff are also projected in BPA-1, BPA-2, and BPA-3 in the range of 5.4% to 8.4% in the near-term and 8% to 12% in the mid-term future periods. BPA-4 show little to no change in future runoff conditions with an increase of 0.7% in spring in the near-term future period, but a decrease of 0.8% in the mid-term future period. Summer runoff is projected to decrease across all BPA areas in the range of 3.5% to 6.6% in the near-term and 8.6% to 18% in the mid-term future periods. The trends in winter and summer runoff projections indicate that these changes would be stronger in magnitude in the mid-term future period, although also accompanied with greater uncertainties across all ensemble members. The fall runoff is projected to increase in the range of 1.4% to 7.2% in the near-term and 2.0% to 6.2% in the mid-term future periods.

These projections could relate to trends already observed in SWE and earlier snowmelt runoff in the region (Mote et al., 2005 and 2018; Stewart et al., 2005), accompanied with little to no change in summer precipitation relative to large increases in other seasons, and potentially increasing evapotranspiration due to increased temperatures. To further evaluate snow dynamics in the region, a change of April 1 SWE is presented in Figure 3.8. The change is defined as the percentage difference of April 1 SWE in future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). The changes are summarized for HUC8s with more significant snowpack (defined as HUC8s with average 1980–2019 control run April 1 SWE greater than 5 mm). The results suggest a strong decrease in ensemble median April 1 SWE (Figure 3.8, 2020–2059 future to 1980–2019 baseline), with more than half of the HUC8s projecting a decrease of 34% or more. These changes in snow hydrology along with findings from the changes in runoff regimes suggest that more water being available as runoff during the spring may result in lower summer streamflow conditions.

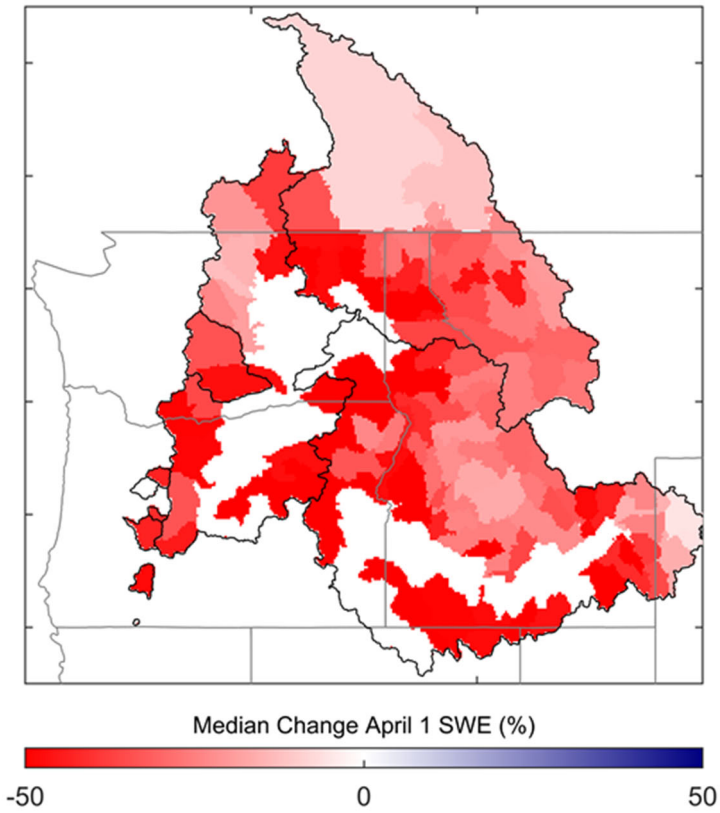


Figure 3.8. Projected change of April 1 SWE in the BPA region.

An additional analysis of extreme runoff is shown in Figure 3.9, which presents future changes in ensemble median high runoff (i.e., 95th percentile of daily runoff) and median low runoff (i.e., 5th percentile of 7-day average runoff) in both future projection periods. High runoff is generally projected to increase in both future periods across all BPA areas except for BPA-1, where a decrease in high runoff is projected in the mid-term future period. The low runoff is projected to increase in general across all BPA areas during both future periods. These results suggest an increasing likelihood of more frequent flood events, which may increase the difficulty of water management in future climate conditions. The findings from other climate assessments also suggested increased flood risks in the region (Salathé Jr. et al., 2014; Tohver et al., 2014; Chegwidden et al., 2020; Wrzesien and Pavelsky, 2020).

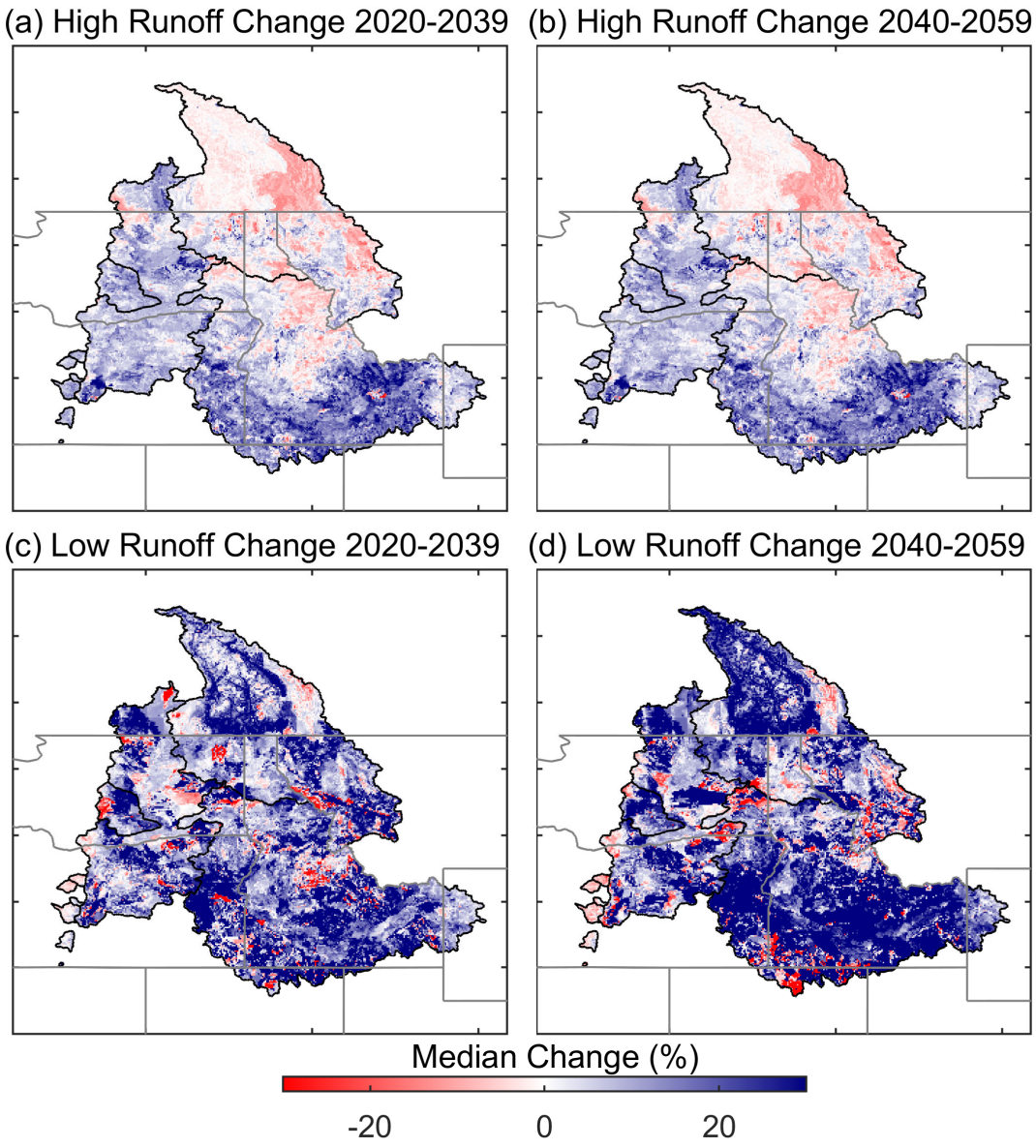


Figure 3.9. Projected change of high and low runoff in the BPA region.

To evaluate the impacts of modeling choices (six CMs, two DSs, two MetFs, and two HMs) on the projection of future runoff in each BPA area, an ANOVA was conducted and is illustrated in Figure 3.10. The box plots show the magnitude of change (future minus baseline) in median annual runoff across 48 ensemble members, grouped by each of the four modeling choices. The box plots provide a quantitative way to evaluate the range of possible outcomes in runoff response based on different modeling choices. ANOVA results indicate that the CM is generally the most dominant contributor to the total variance of future runoff projections. The contribution of the CM toward total variance is greater than 50% for all BPA areas. For BPA-3 and BPA-4, the HM is the second most dominant factor, contributing around 42% of total variance for BPA-3 and 28% for BPA-4, indicating that annual runoff response under future conditions is also sensitive to the choice of HM. For BPA-1, the combination of the CM and DS is the second most dominant factor. At the seasonal scale (APPENDIX E), the ANOVA indicates that factors other than the CM may have stronger impacts than at the annual scale. For instance, although the CM is still a strong contributor to the total variance (in the range of 39% to 76% during winter and 20.8% to

72.7% during spring), the contribution from other modeling choices gains more influence on the spread of runoff projections. For instance, the HM is the most dominant contributor in BPA-4 during spring, and the MetF is the most dominant contributor in BPA-1 during winter. During summer, factors other than the CM have stronger influences on runoff projections, and the CM variance contribution decreases to 11% to 39%.

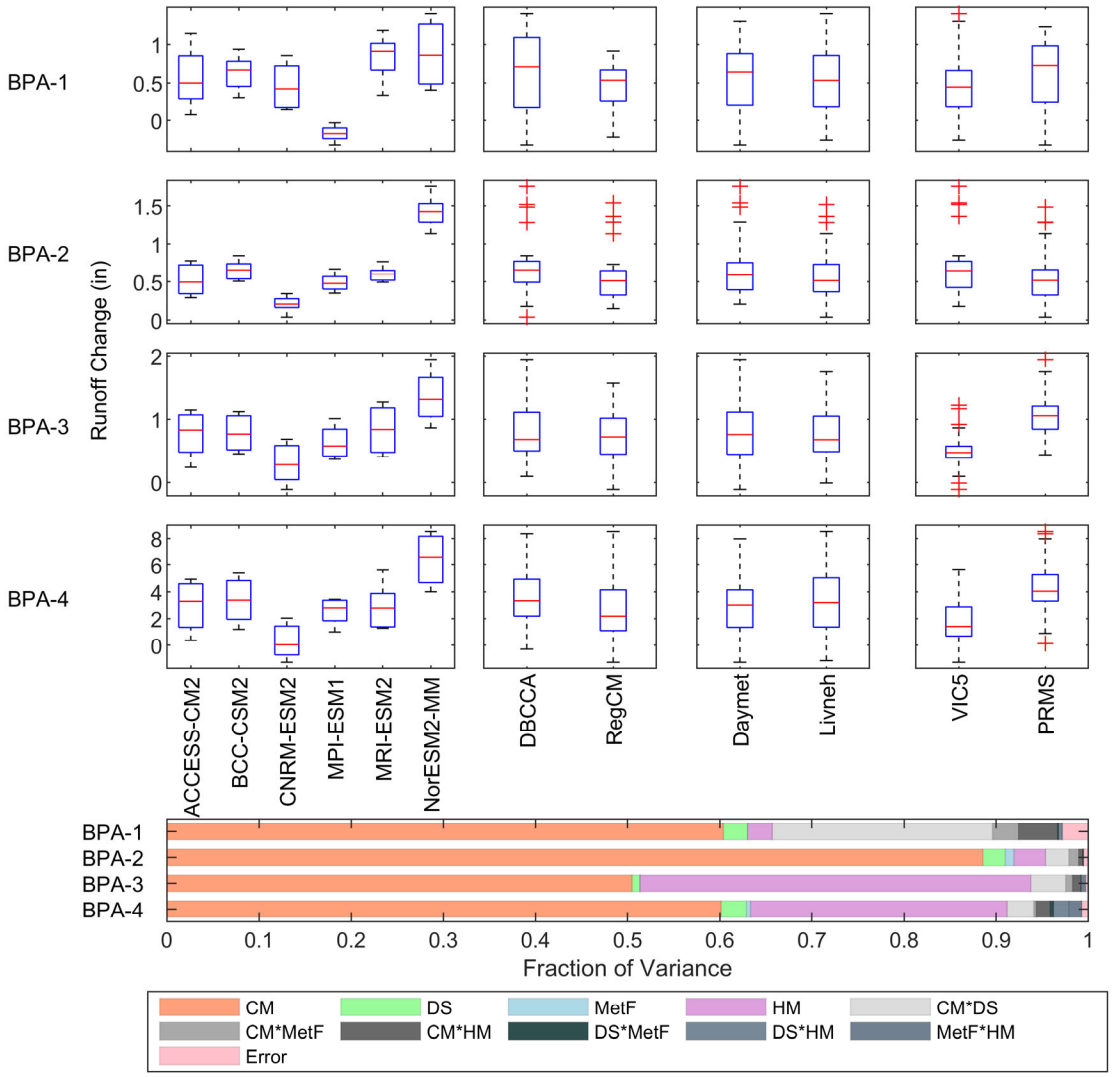


Figure 3.10. ANOVA of annual total runoff in the BPA region.

The LEM-simulated reservoir evaporation ensemble from federal hydropower reservoirs in the BPA region is illustrated in Figure 3.11. The blue lines show annual reservoir evaporation loss (in units of million acre-feet) from 24 projections that were either bias-corrected or trained by Daymet, and 24 projections that were either bias-corrected or trained by Livneh. The dark blue and dark green lines represent the control simulations driven by 1980–2019 Daymet and 1980–2013 Livneh, respectively. The right panel shows the probability density distributions of annual evaporation loss in both baseline (1980–2019) and future (2020–2059) periods.

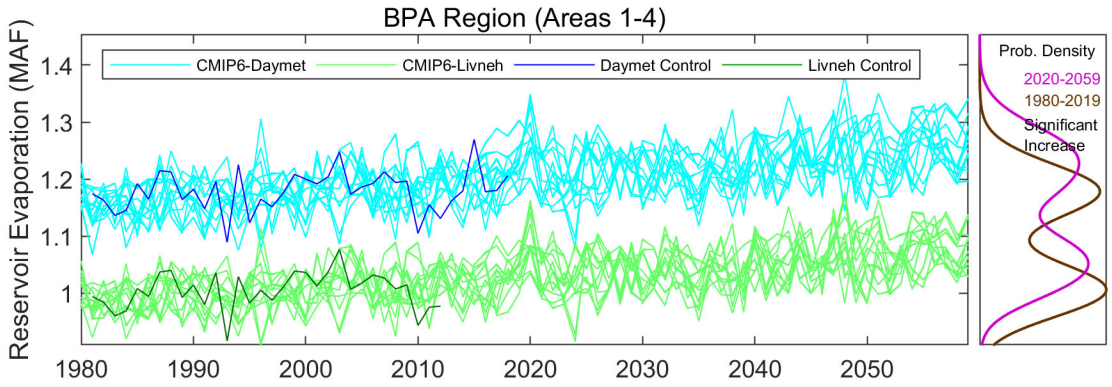


Figure 3.11. Projected annual total reservoir evaporation in the BPA region.

The results are clustered into two groups—one with higher values for the 24 members that are associated with Daymet, and the other with lower values for the 24 members that are associated with Livneh. Overall, the Daymet data set suggests a higher evaporation rate than the Livneh data set in the BPA region. The four key meteorological variables that directly affect the evaporation rate are downward radiation, temperature, wind speed, and vapor pressure deficit. Among these, the wind speed differs the most between the two data sets, and this difference has the largest impact on the projected evaporation rates in the BPA region. The ensemble mean of the wind speed that was bias-corrected by NARR used in Daymet is about 1 m/s higher than that which was bias-corrected by NCEP1 used in Livneh. This difference alone may result in a nearly 0.2 mm/day difference in evaporation rate. Comparisons with in situ wind speed measurements at Lake Mead and Lake Mohave (not shown) suggest that the error of NARR wind speed used in Daymet is relatively small, whereas the NCEP1 wind speed used in Livneh underestimates more considerably (Moreo, 2015). Thus, the Daymet-based evaporation rate projections are more reliable. The reservoir evaporation loss also depends on the projected reservoir area, which is inferred from runoff simulated by either VIC or PRMS. The projected runoff between Daymet and Livneh is not significantly different because the HMs were calibrated separately by both MetFs against the same USGS WaterWatch data set (Section 2.5.3) to minimize the impacts of different meteorological forcings on hydrologic projections. In addition, the primary reservoir types in the BPA region are Type 1 and Type 2 (Section 2.7.1), whose area variations are less sensitive to runoff changes. For the increase of reservoir evaporation loss, only 3% were attributed to changes in reservoir area, and 97% were affected by changes in evaporation rate.

Nevertheless, both reservoir evaporation groups show the same increasing trend (as indicated by the two-sample Kolmogorov-Smirnov test performed at the 5% significance level). In general, the projected future evaporation volume will increase by about 0.05 MAF from 1980 to 2059, with an average growth of nearly 5% over the entire period. In the BPA region, the main reason for the vast evaporation loss in the future is the increased evaporation rate; the projected reservoir area in this region will not change much in the future. However, such projected losses are much smaller when compared with the total annual runoff in the BPA region—runoff that is likely to increase over time.

The annual and seasonal changes of monthly reservoir evaporation loss (in units of thousand acre-feet per month) in the BPA region are depicted in Figure 3.12. Increasing reservoir evaporation is persistent in both the near-term (2020–2039) and mid-term (2040–2059) future periods with respect to the baseline period (1980–2019). More specifically, the ensemble median of annual evaporation loss is projected to increase by 2.7 KAF/month in the near-term and 5 KAF/month in the mid-term future periods. At the seasonal scale, evaporation losses in all four seasons are projected to increase in both near-term and mid-term future periods. The ensemble median of winter and spring reservoir evaporation loss is projected to increase by about 1 KAF/month in the near-term future period, and around 3 KAF/month in the mid-term

future period. Both changes are relatively low, with small ensemble variability. The ensemble median of summer and fall reservoir evaporation loss is estimated to increase more substantially by 5 KAF/month in the near-term and 6–7 KAF/month in the mid-term future periods. This increase is likely caused by a more substantial temperature increase in both seasons. In addition, sensitivity tests suggest that for the same amount of temperature increase, evaporation rate during the warm seasons (summer and fall) will increase more than during the cold seasons (winter and spring). However, these losses are orders of magnitude smaller than the large total runoff volume in the BPA region.

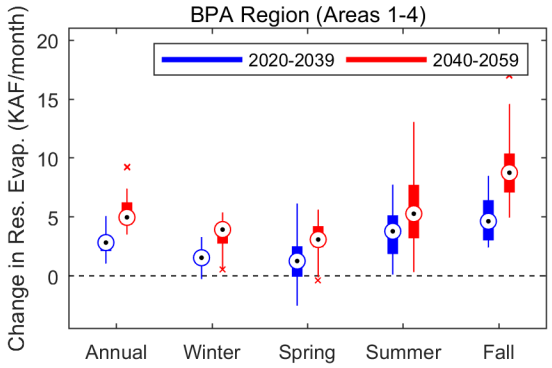


Figure 3.12. Projected change of annual and seasonal total reservoir evaporation in the BPA region.

The source of variance in the projected annual reservoir evaporation loss from federal hydropower reservoirs in the BPA region is shown in Figure 3.13. The box plots show the magnitude of change (future minus baseline) in median annual reservoir evaporation loss across 48 ensemble members, grouped by each of the four modeling choices (CM, DS, MetF, and HM). The bottom panel shows the portion of variance from the ANOVA for each modeling choice.

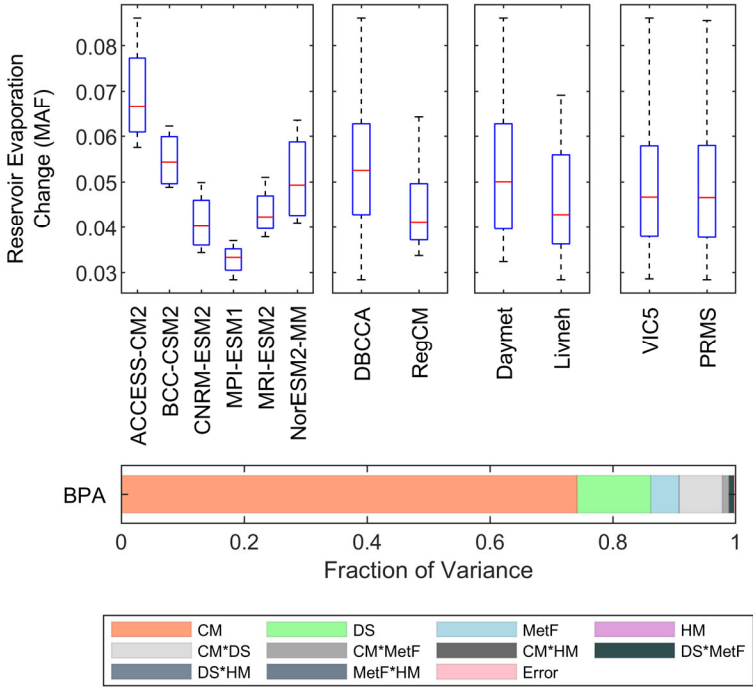


Figure 3.13. ANOVA of annual total reservoir evaporation in the BPA region.

ANOVA results indicate that the CM is generally the most dominant contributor to the total variance of future evaporation projections, contributing around 73% of total variance. The most substantial increase is projected by ACCESS-CM2 (0.065 MAF), and the smallest is by MPI-ESM1 (0.033 MAF). As the second dominant contributor, the DS only contributes around 11% of the total variance, and DBCCA projects a more substantial increase compared with RegCM. The contribution of the MetF is only about 3% of the total variance. At the seasonal scale (APPENDIX F), the ANOVA indicates that certain factors may have more substantial impacts than at the annual scale. For instance, the CM is still the most dominant contributor for the variance in all four seasons. In the fall, it can explain nearly 90% of the total variance, with the primary variance coming from ACCESS-CM2 and MPI-ESM1. In comparison, in the summer, the CM can only explain about 40% of the total variance. The DS is the second most dominant contributor, representing 5%–40% of the variance across seasons. In the summer, the DS explains about 40% of the total variance because of the large difference of projected evaporation change between BDCCA and RegCM. Finally, the MetF only contributes up to 4% of the total variance across four seasons, and the other factors have a negligible influence. Although the choice of MetFs (i.e., Daymet or Livneh) may significantly influence the projected value of reservoir evaporation loss (Figure 3.11), the difference is systematic and largely reduced when examining the change from baseline to future periods. This shows that even with large model/data bias, the relative change is still considered reliable and used in many hydroclimate impact assessment studies.

3.3 CLIMATE EFFECTS ON FEDERAL HYDROPOWER IN THE BPA REGION

3.3.1 Projections of Hydropower Generation

Using the two hydropower models described in Section 2.6, the projections of monthly generation in the BPA region are calculated for each of the 96 unique combinations of CMs, DSs, MetFs, and HMs and PMs. Projections based on multimodel ensembles are summarized in terms of annual and seasonal changes for the near-term (2020–2039) and mid-term (2040–2059) future periods compared with the baseline period (1980–2019).

The interannual variability of annual hydropower generation for the baseline and future periods is shown in Figure 3.14. The gray lines show the annual total generation for all 96 members of the ensemble hydropower projections, the black line represents the ensemble median, and the blue line shows the 1980–2019 historic observation from EIA and PMAs. The corresponding probability distributions of annual generation in the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between baseline and future periods is statistically significant.

The comparison (Figure 3.14, left column) suggests that the range of simulated and observed annual hydropower variability is generally consistent across all BPA areas. However, the shapes of probability distributions revealed that significant changes in interannual variability will occur in some BPA areas. Specifically, in BPA-2, BPA-3, and BPA-4, the future generation shows narrower distributions with sharper peaks (Figure 3.14, right column). For BPA-1, no noticeable difference was observed in the probability distributions between baseline and future periods.

The annual and seasonal change in hydropower generation is further summarized in Figure 3.15. The change in Figure 3.15 is defined as the percentage difference of future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across 96 multimodel combinations; the central mark indicates the multimodel median, and the edges of box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length are considered statistical outliers (but should not be excluded in the context of climate change).

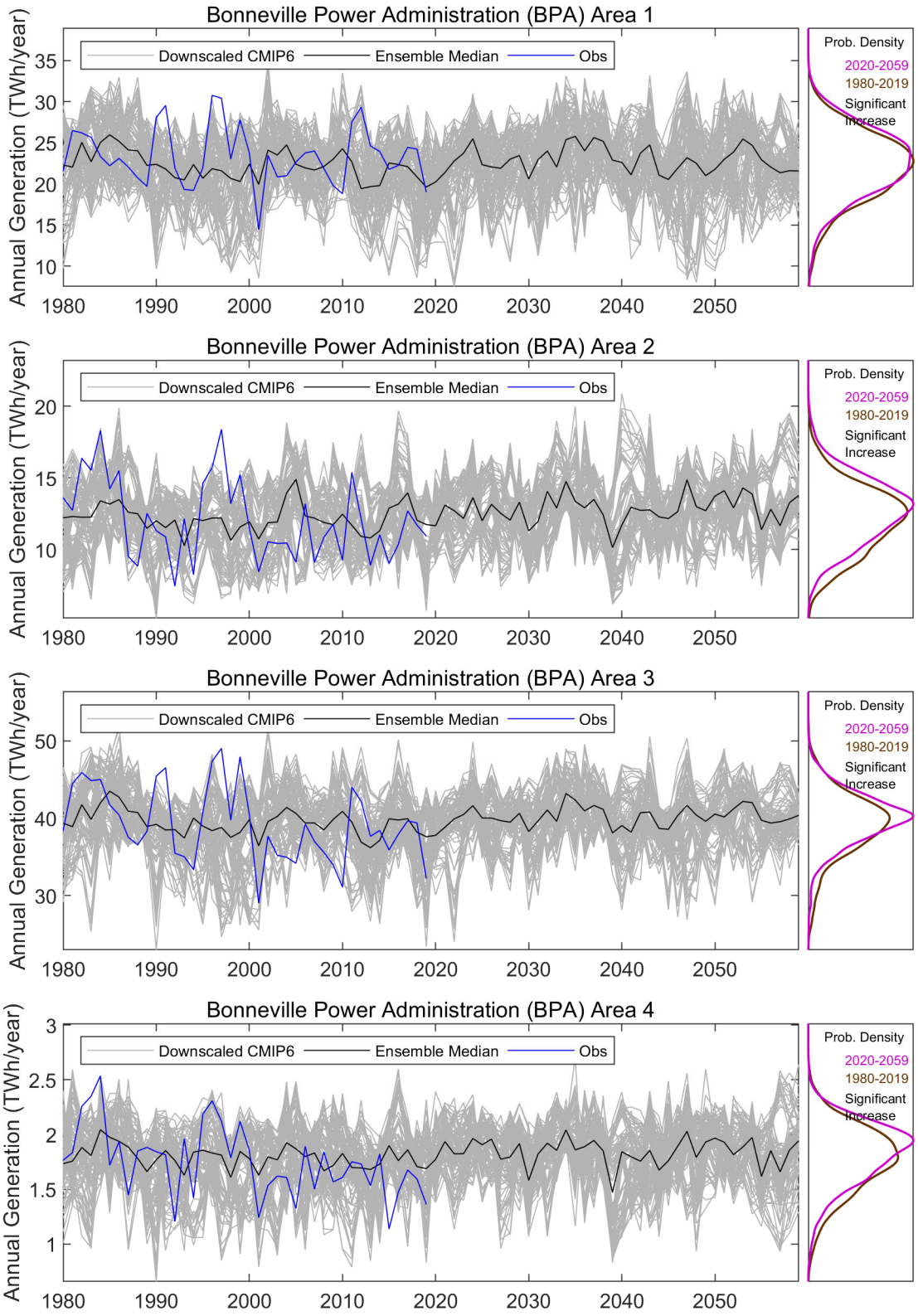


Figure 3.14. Projected annual total generation in the BPA region.

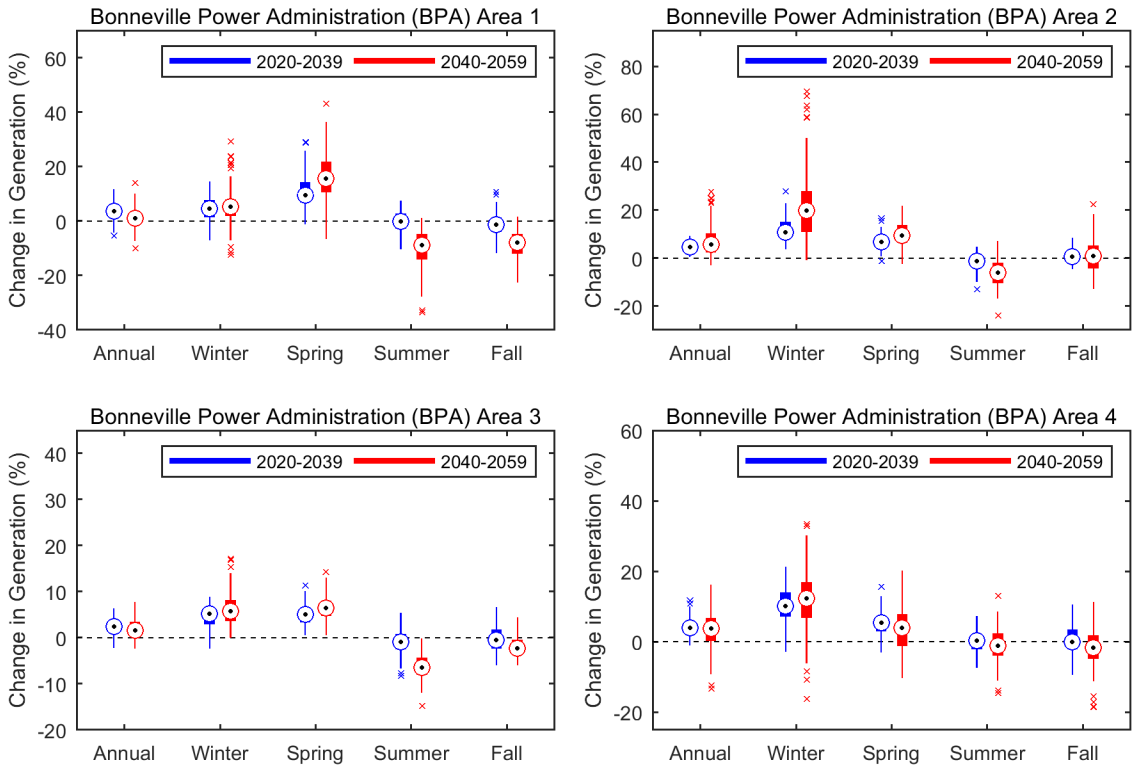


Figure 3.15. Projected change of annual and seasonal total generation in the BPA region.

A small increase in annual hydropower generation is projected across the four BPA study areas throughout the simulation period (Figure 3.15). This increase of annual hydropower generation is consistent with the increase of precipitation and runoff at the annual scale. However, increases in annual flows do not translate to a one-for-one increase in generation as the annual runoff is reshaped, and some of the larger flow increases are likely to result in more bypass spill in the relatively storage-limited BPA area. More specifically, the annual hydropower generation in the near-term (2020–2039) future period is projected to increase from 2% in BPA-3 to 5% in BPA-1 in terms of multimodel median. Similarly, the annual hydropower generation in the mid-term (2040–2059) future period is projected to either slow or continue its increase, down to a 1% increase in BPA-1 and up to a 6% increase in BPA-2, also in terms of multimodel median.

Seasonal relative changes tend to be more substantial than annual changes, with more pronounced increases typically in winter and spring, and maintained or reduced generation in summer and fall. In the near-term future, projections indicate an increase in winter generation ranging between 5% and 11%, and the increase jumps to 5%–20% in the mid-term future period (Figure 3.15). The substantial increase in generation is projected to continue in the spring (5%–9% in the near-term and 4%–16% in the mid-term future periods). Although not much change is projected in summer hydropower generation in the near-term future period, a substantial reduction of summer hydropower generation is projected in BPA-1–3 (6%–9%) in the mid-term future period. BPA-4 is an exception; there, summer generation is maintained in both near-term and mid-term future periods. Watersheds in BPA-4 are mainly controlled by rainfall and are less impacted by the change of snow hydrology (Section 3.2.2). Fall generation is projected to remain stable in the near-term and mid-term future periods with a large ensemble spread. The exception is BPA-1 in the mid-term future period, where fall generation is projected to decrease by 8% associated with a stronger seasonal pattern overall.

By comparing the projected annual and seasonal change between runoff (Figure 3.7) and generation (Figure 3.15), the effect of reservoir storage can be observed. In all BPA areas, the projected annual and seasonal hydropower generation change is milder than the projected runoff change, indicating that reservoirs absorbed part of the runoff monthly variability while constraining the annual maximum generation. In the case of the BPA-1 and BPA-3 mid-term future period, the largest runoff change occurs in winter, whereas the largest hydropower generation change occurs in the spring.

The source of variance in the projected annual hydropower generation change between the baseline and future periods is shown in Figure 3.16. The box plots in the top panels show the magnitude of change (future minus baseline) in annual hydropower generation across 96 ensemble members when grouped by the five modeling choices (CM, DS, MetF, HM, and PM). The bottom panel shows the fraction of variance from the ANOVA for each modeling choice.

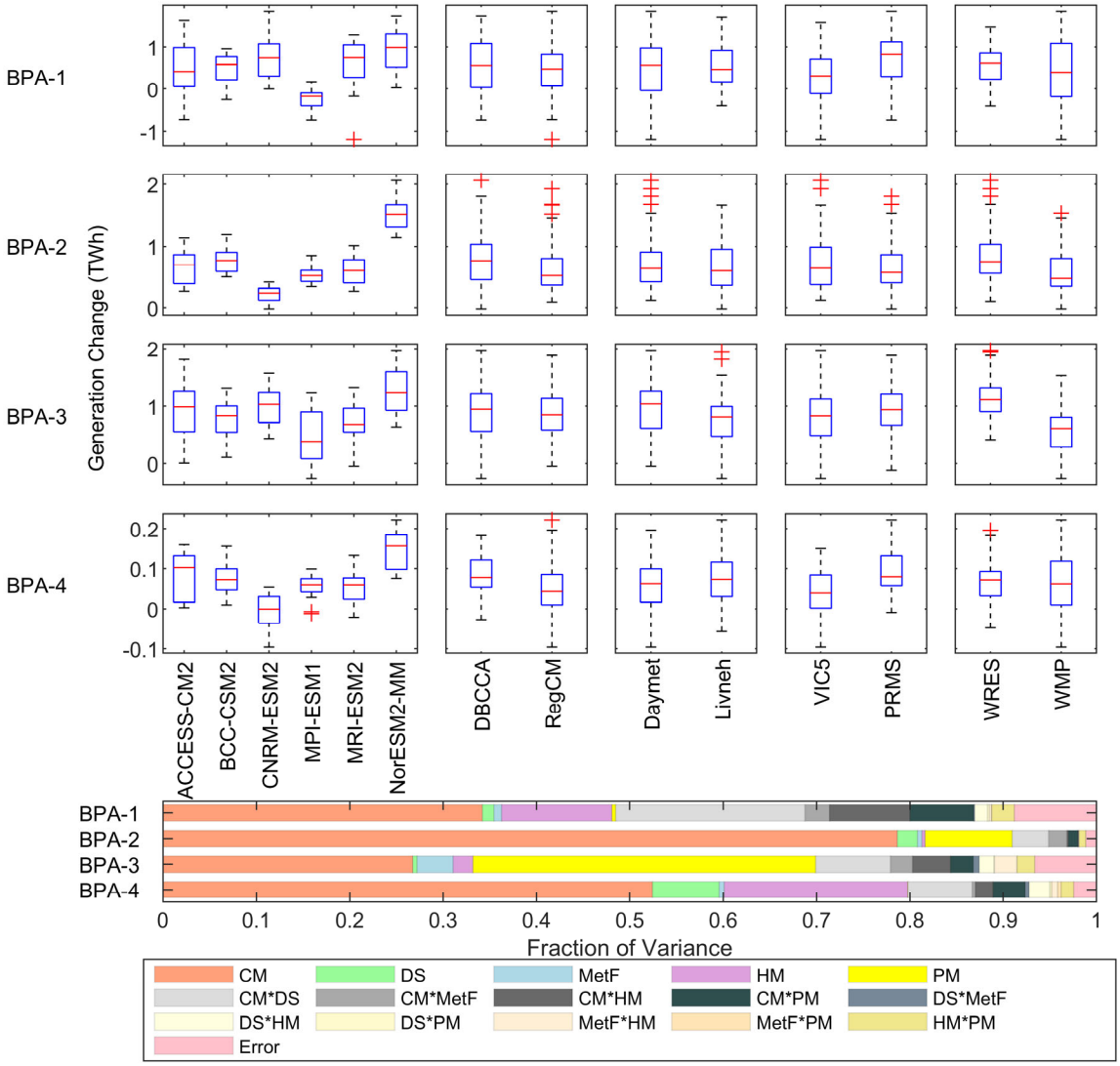


Figure 3.16. ANOVA of annual total generation in the BPA region.

The diversity in individual factors and their combinations were further evaluated to understand their contributions to the spread of hydropower generation projections. Consistent with future streamflow projections, the results suggest that the CM is the most dominant contributor of annual hydropower

generation variance in BPA-1, BPA-2, and BPA-4 (35%–79%). The second most dominant contributors are the HM in BPA-1 (12%) and BPA-4 (20%) and the PM in BPA-2 (12%). The BPA-3 area stands out with the PM being a more dominant contributor (35%) than the CM (27%). The combination of the CM and DS is also a substantial variance contributor in BPA-1 (20%). In all BPA study areas, the DS or MetF alone are not substantial variance contributors. As shown in Figure 3.16, when grouped by CM, the median change of projected annual hydropower generation is mostly positive across all CMs in all BPA study areas, although the median and spread vary substantially across different CMs. NorESM2-MM consistently projects the highest median change of hydropower generation across all areas, which is consistent with its projected changes of precipitation and runoff. When grouped by PM, both PMs projected similar generation increase in the future, and WRES generally projected greater changes and less variance than WMP. These differences between WRES and WMP in BPA-1 and BPA-4 are likely associated to the different representations of water management and streamflow regulations in most headwater reservoirs.

At the seasonal scale (APPENDIX G), the CM remains the most dominant contributor to variance, except for in the summer. In the summer, the PM is the most dominant variance contributor in BPA-1 (38%), and the HM is the most dominant variance contributor in BPA-2 (34%) and BPA-4 (40%). The HM also has strong impacts on variance in the fall (19% for BPA-1, 39% for BPA-2, and 27% for BPA-3).

In summary, for federal hydropower in the BPA region, the most important climate change impact includes a small increase in annual generation (4% increase in the near-term and 3% increase in the mid-term future periods) and a substantial seasonal change in generation (e.g., up to 11% decrease in BPA-1 summer generation and 8% decrease in BPA-1 fall generation in the mid-term future period). The hydropower generation has a smaller response to climate change relative to increases in runoff, indicating that much of the flow increases may bypass turbines via spill, or that reservoir storage systems may be able to absorb part of the runoff variability. In terms of the sources of uncertainty, the choice of CM is overall the most dominant contributor to variance in the projection of hydropower generation. The choice of HMs and PMs also have strong impacts but vary by area and season.

3.3.2 Climate Change Impacts on Energy Demand and Federal Power Marketing

This section summarizes the results from the statistical analysis exploring the effects of projected temperature trends on the total retail load served by BPA's preference customers. As described in Section 3.1.2, BPA markets the output from 31 federal hydropower plants through long-term contracts with preference customers, such as cooperatives, municipalities, and public utility districts. Since BPA has an obligation to serve the net load requirements of those preference customers that purchase their load following product, information on projected changes in the annual volume or seasonality of the load served by preference customers—driven by climate or other factors—is valuable for this PMA.

Table 3.3 displays, for each season, the estimated coefficients describing the relationship between monthly total retail load (2006–2019) and degree days for 91 BPA preference customers.[18] The full regression output is included in APPENDIX H. Because the load equation was estimated using a logarithmic transformation of both the dependent and independent variables, the estimated coefficients should be interpreted as elasticities. Monthly HDDs is the variable used to assess the effect of cold temperatures on load. A 1% increase in HDDs during the winter months results in a 0.36%–0.37% increase in monthly total retail load. During the other seasons, the effect of HDDs on load is much smaller. The equation includes three terms to assess the effect of warm temperatures on load: monthly CDDs, monthly CDDs interacted with the estimated air conditioning (AC) penetration in the preference

[18] Although the total retail load information was available for all 134 BPA preference customers, some of the other variables used to estimate the load equation (e.g., price and income) were only available for 91 preference customers.

customer's service territory, and annual CDDs interacted with a dummy variable that takes the value 1 if the preference customer has a large peak in load during the summer months and zero otherwise.[19] The summer peak customers are a proxy for customers with large irrigation loads; only 14 of the 91 preference customers included in the regression analysis fall into this group. The estimated coefficients on monthly CDDs and their interaction with air conditioning saturation are largest in the summer, which is also the only statistically significant season. The estimated coefficient on the interaction term between CDDs and large irrigation loads has similar magnitudes (~0.07) for the spring and fall seasons and is larger in those seasons than in summer (~0.04). The total retail load served by BPA preference customers is most sensitive to temperature in the winter and summer. The elasticity with respect to HDDs is larger than the elasticity with respect to CDDs even when interactions with AC penetration and large irrigation loads are considered. However, the elasticity with respect to degree days is the percentage change in load in response to a 1% increase in degree days and, given average HDDs and CDDs in the BPA region during the period used for this regression analysis (2006–2019), a 1% increase in HDDs is much larger in absolute value than a 1% increase in CDDs.[20]

Table 3.3. Estimated responsiveness of the total retail load of BPA preference customers to degree days

Marketing region	Meteorological observation	Variable	DJF (winter)	MAM (spring)	JJA (summer)	SON (fall)
BPA	Daymet	HDD	**0.3641** (0.014)	**0.0538** (0.0169)	**−0.0027** (0.0011)	0.0072 (0.005)
BPA	Livneh	HDD	**0.3701** (0.015)	**0.0414** (0.0179)	**−0.0031** (0.0009)	**0.0187** (0.0078)
BPA	Daymet	CDD		0.0038 (0.0081)	**0.0497** (0.0064)	−0.0023 (0.007)
BPA	Livneh	CDD		0.0019 (0.0087)	**0.0443** (0.0064)	−0.0004 (0.0074)
BPA	Daymet	CDD × estimated AC penetration		0.0104 (0.0065)	**0.0396** (0.0053)	0.0072 (0.0058)
BPA	Livneh	CDD × estimated AC penetration		0.0099 (0.0067)	**0.0359** (0.0053)	0.0071 (0.0061)
BPA	Daymet	CDD × irrigation		**0.072** (0.0137)	**0.039** (0.0053)	**0.0659** (0.0135)
BPA	Livneh	CDD × irrigation		**0.068** (0.013)	**0.0378** (0.0051)	**0.0675** (0.0137)

Notes: Numbers in parentheses are standard errors. Bolded values indicate that the estimated elasticity is statistically significant at a 10% level. The number of CDDs for winter was always zero, resulting in dropping that variable from the equation for that season.

The estimated coefficients for the three CDD-related variables need to be combined to calculate a total CDD elasticity value that accounts for load responses through both more intense use of the exiting AC stock and increases in the AC stock. The total CDD elasticity also varies by customer type between customers with large irrigation loads and the rest.

$$TotalElast_{TRL,CDD}[Irrig = 0] = Elast_{TRL,CDD} + Elast_{TRL,CDD-AC} * \ln(ACsat) \quad (3.1)$$

$$TotalElast_{TRL,CDD}[Irrig = 1] = Elast_{TRL,CDD} + Elast_{TRL,CDD-AC} * \ln(ACsat) + Elast_{TRL,Irr} \quad (3.2)$$

[19] A description of the method used for estimating air conditioning penetration is presented in APPENDIX H.
[20] Average HDDs and CDDs in the BPA region in 2006–2019 are 4,069 and 364 in the Daymet data set.

The estimated elasticities were combined with respect to degree days with the projected changed in degree days derived from climate model simulations to produce an estimate of the temperature-driven changes in load for BPA preference customers. Figure 3.17 and Figure 3.18 show the projected degree days from 24 simulations (the six CMIP-6 models described in Section 2.3 combined with two MetFs and two DSs), as well as the ensemble median projection.

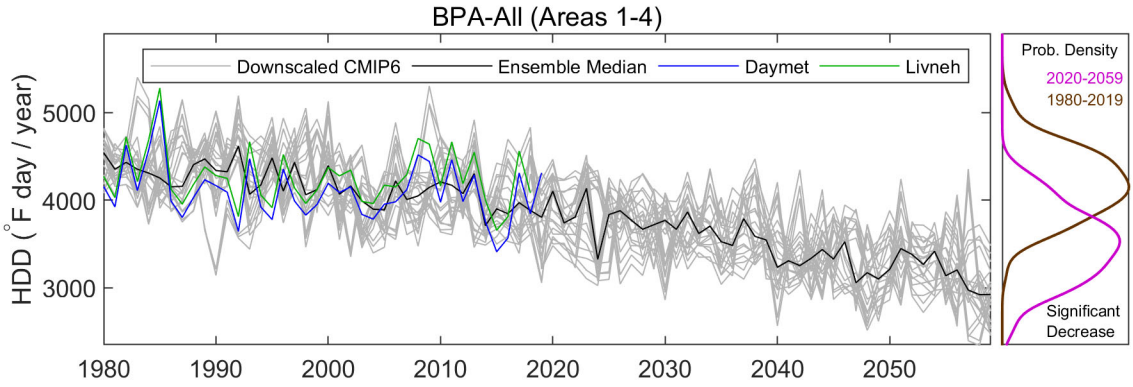


Figure 3.17. Projected annual HDDs in the BPA region.

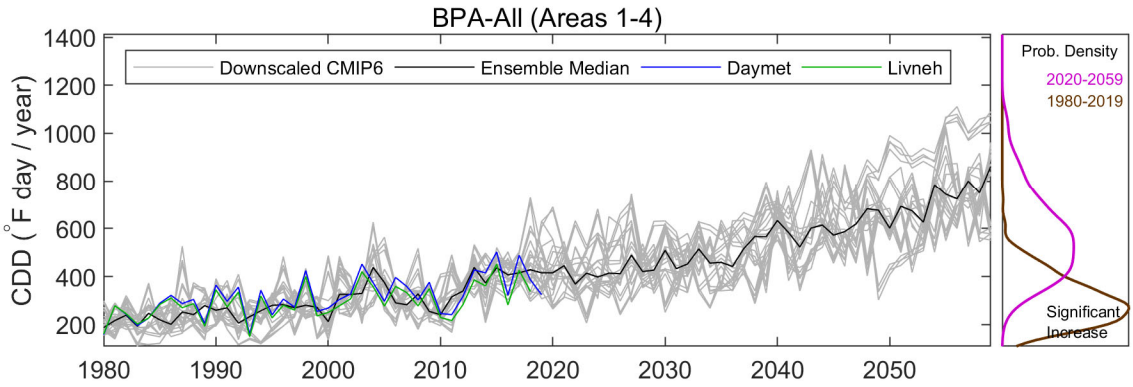


Figure 3.18. Projected annual CDDs in the BPA region.

During the baseline period, the average annual values of the ensemble medians in the territory served by BPA preference customers were 4,176 HDDs and 296 CDDs. For 2020–2059, the average annual values of the ensemble medians in Figure 3.17 and Figure 3.18 were 3,483 for HDDs and 562 for CDDs. This represents a 17% decrease in HDDs and a 90% increase in CDDs. For the projected annual CDDs, the range noticeably widened during the past decade, which is indicative of less consensus in the temperature projections from the various CMIP6 models. Figure 3.19 shows the projected changes in load for preference customers with no large irrigation loads to meet as a result of the degree day trends previously discussed.

HDDs only have a sizable effect on commercial and residential load during the winter months. The decreases in load driven by higher temperatures during the winter months range from 0.9% to 5.9% in the near-term (2020–2039) and from 3.4% to 7.7% in the mid-term (2040–2059) future periods. The median projected winter load decrease is 2.4% in the nearer future period and 5.7% for the two subsequent decades.

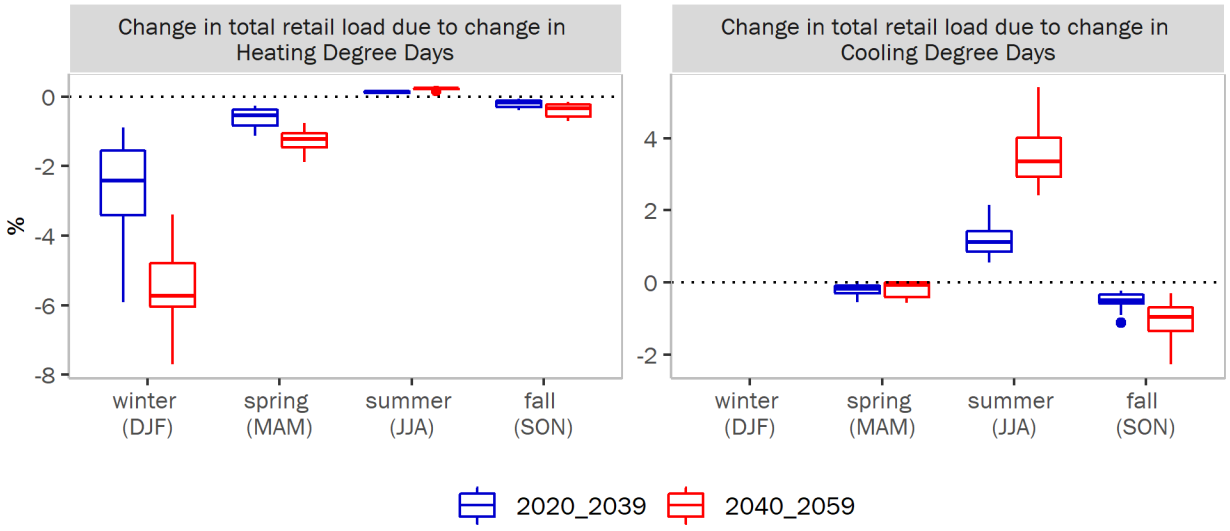


Figure 3.19. Estimated changes in seasonal load relative to the baseline period due to projected changes in degree days for BPA preference customers with no large irrigation loads. No load elasticity with respect to CDDs was estimated for winter because there were no CDDs during that season in the historical period.

For most BPA preference customers, CDDs only lead to increases in load in the summer.[21] The median percentage change increase in summer load is 1.3% in the 2020–2039 period and 3.6% in the 2040–2059 period. The larger increase in the mid-century is caused by a combination of a larger increase in CDDs relative to the baseline period and projected increases in AC saturation.

For the group of 14 customers with large irrigation loads, projected temperature trends for 2020–2039 lead to median load increases of 3% in spring, 3% in summer, and 4.9% in fall, and a decrease of 2% in winter. For 2040–2059, the projected median changes lead to load increases of 13% in spring, 7.6% in summer, and 15% in fall, and a decrease of 5.7% in winter. Comparing the seasonal irrigation load changes estimated for the near-term and mid-term future periods indicates an increasing temperature-driven trend in irrigation load in all seasons except winter.

The temperature-driven load changes discussed indicate a decrease in load during the winter and an increase during the summer. This was shown by weighting the median changes in load for the two customer types and the four seasons discussed in the previous two paragraphs by their fraction of total load in recent years (2017–2019). The projected median change in annual load for BPA preference customers is very small and negative (a 0.40% decrease in 2020–2039 and a 0.38% decrease in 2040–2059).[22] However, temperature is only one of many factors influencing load size. Temperature would also need to be considered in combination with other drivers (e.g., population and electrification trends) for load forecasting and resource planning. For population, Zoraghein and O'Neill (2020) produced state-level estimates for 2050 and 2100 consistent with three different SSPs. In the SSP5 pathway, which is the one consistent with the climate projections in this analysis, the projected population growth in the three-state region where most BPA customers are located (Washington, Oregon, and Idaho) is a 67% increase from 2010 to 2050 and an additional 59% increase in the second half of the century. Therefore,

[21] The appropriate classification of months into seasons for estimating the response of load to temperature may be affected by climate change. For instance, as temperatures warm, the CDD response seen for summer may extend into September.
[22] The weights applied to each of the net load changes are as follows: 0.26 (winter, other customers), 0.03 (winter, large irrigation customers), 0.21 (spring, other customers), 0.03 (spring, large irrigation customers), 0.19 (summer, other customers), 0.04 (summer, large irrigation loads), 0.21 (fall, other customers), 0.03 (fall, large irrigation customers).

population is a potential strong driver for load in the BPA region, especially when compounded with growth in AC penetration.

As discussed in Section 3.3.1, hydropower generation in the BPA region is expected to increase at the annual time scale. The increase is concentrated in the winter and spring, whereas summer generation is projected to decrease in most of the FCRPS facilities from 2040–2059 (Figure 3.15). Thus, the seasonal changes in generation availability do not correlate positively with the expected changes in the seasonal shape of preference customer demand in the region (increased summer load as shown in Figure 3.19), particularly in 2040–2059. The decreased correlation between hydropower generation availability and customer demand will likely result in increased replacement power purchases by the PMA during the summer months. The cost of replacement power purchases must then be passed through to the rates that BPA charges its customers. Finding ways to maximize the value of potential winter generation surpluses and minimize the needs for replacement power during the summer will be important to maintain competitive and affordable rates for BPA and ensure reliability.

4. THE WAPA REGION

Following the methodology described in Section 2, the results for federal hydropower plants marketed by the WAPA are summarized in this section. Descriptions of the WAPA study areas, federal hydropower systems, and power marketing practices are described in Section 4.1. The projected future atmospheric and hydrologic conditions in the WAPA region are discussed in Section 4.2. The potential climate change impacts on federal hydropower generation and WAPA regional energy demands are discussed in Section 4.3.

4.1 WAPA REGIONAL CHARACTERISTICS

4.1.1 Federal Hydropower in the WAPA Region

Geographically, WAPA is the largest PMA and covers parts of 15 states from California to the Great Plains and the Mexican to Canadian borders (Figure 4.1). There are 55 federal hydropower plants in the WAPA region, with a total of 10,200 MW capacity and 28.6 TWh/year average annual generation (Table 4.1).

Federal hydropower plants in the WAPA region are owned and operated by three agencies: USACE, Reclamation, and the IBWC (Figure 4.1). The six USACE projects are all located in the upper Missouri River basin. The USACE projects are all relatively large in capacity, ranging from the 132.3 MW Gavins Point to the 784 MW Oahe. The two IBWC plants on the Rio Grande River are parts of multipurpose water development projects constructed primarily for flood control and water conservation; their combined capacity is less than 97.5 MW. Reclamation owns and operates the other 47 hydropower plants in this region. A complete list of the projects in this region is provided in APPENDIX B.

Table 4.1. Summary of federal hydropower plants in the WAPA region

Area	Area name	Number of plants				Capacity[a] (MW)	1980–2019 generation[b] (GWh/year)
		USACE	Reclamation	IBWC	Total		
WAPA-1	Upper Missouri	6	2	0	8	2,858	9,907
WAPA-2	Loveland Projects[c]	0	18	0	18	702	1,470
WAPA-3	Upper Colorado	0	12	0	12	1,822	6,051
WAPA-4	Lower Colorado	0	3	0	3	2,454	6,316
WAPA-5	Rio Grande	0	0	2	2	98	197
WAPA-6	California	0	12	0	12	2,271	4,682
Total		6	47	2	55	10,204	28,622

[a] EIA 2019 total nameplate capacity, including both conventional hydro and pumped storage.
[b] EIA, Reclamation, and WAPA average annual generation from 1980 to 2019, conventional hydro only.
[c] Two of the four Yellowtail plant units are marketed in WAPA-2, but for the purposes of this analysis, the entire Yellowtail plant is included in WAPA-1.

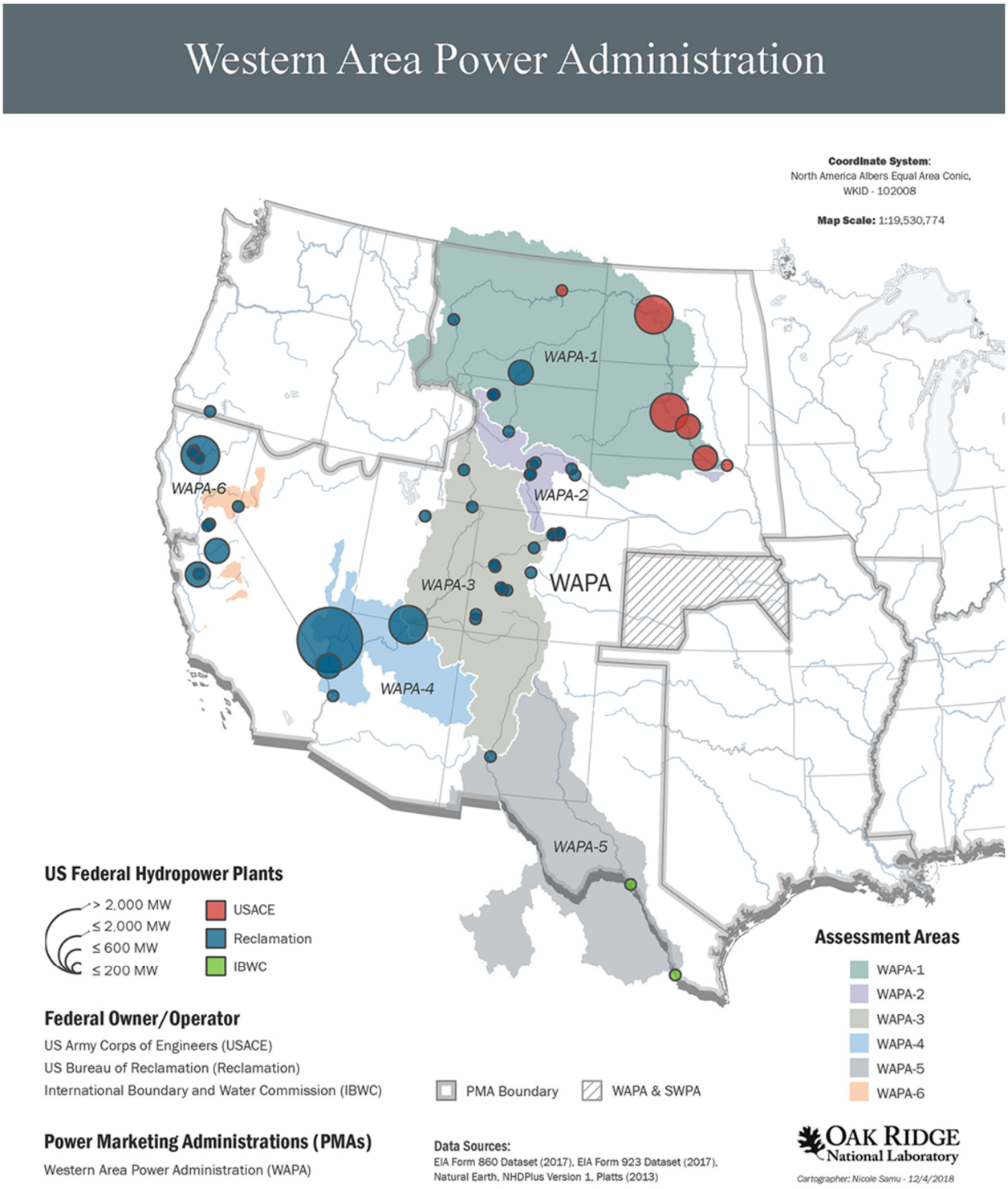


Figure 4.1. Map of the federal hydropower plants and six study areas in the WAPA region.

The aging of hydropower infrastructure in the WAPA region is also a serious concern for both federal hydropower providers and their customers. The oldest Reclamation hydropower plant in the WAPA region is the 6.4 MW Guernsey, which began operation in 1927. The youngest Reclamation hydropower plant is the 4.5 MW Spirit Mountain hydropower plant that began operation in 1994.

Hydropower generation in the western United States is usually not the priority for multipurpose water management due to limited and constrained freshwater supplies. The water demand for agricultural, municipal, and industrial uses must be met, and it has even greater importance than hydropower generation. As a result, hydropower at federal projects is generated more as a byproduct of water management for other higher-priority uses, such as water supply for irrigation (Reclamation projects) or flood control and navigation (USACE projects). Conflicts among competing water uses are common and continue in the WAPA region. In addition, there are serious concerns over the size and reliability of available water supply; water supply scarcity coupled with growing demands for water resources remains a top concern for this region if future climate conditions become drier.

In this study, the entire WAPA region is subdivided into six study areas based on power system and watershed boundaries (Figure 4.1):

- **WAPA-1 Upper Missouri:** The upper Missouri River system spans across Montana, Wyoming, North Dakota, South Dakota, and a small part of Canada. The total drainage area of WAPA-1 is around 280,000 mi^2. This is the largest PMA study area and corresponds to the eastern division of the Pick-Sloan Missouri River system of multipurpose water projects. The topography is high plains and mountains, with a median elevation of 3,091 ft and a maximum elevation of 12,907 ft in the Rocky Mountains of western Montana. Land cover is primarily grassland (76%), plus cropland (14%) and evergreen broadleaf forest (6%).

- **WAPA-2 Loveland Projects**: Loveland projects consist of multiple smaller watersheds in the upper parts of the North Platte, South Platte, Bighorn, upper Arkansas, and upper Colorado Rivers. WAPA-2 corresponds to the Loveland Area Projects, which include the Western Division of the Pick-Sloan Missouri River Program and the Fryingpan-Arkansas Project. The combined drainage area is 68,843 mi^2. In addition to the local flow, the Colorado-Big Thompson Project diverted more than 200,000 ac-ft of water each year into this area from west of the Continental Divide. The hydropower projects here are smaller, multipurpose Reclamation projects with primary purposes of water supply, not hydropower. The topography is mountainous, with a median elevation of 6,880 ft and a maximum elevation of 14,035 ft. Land cover is grassland (76%), evergreen forest (14%), open scrubland (2%), and cropland (2%).

- **WAPA-3 Upper Colorado**: WAPA-3 consists mostly of the upper Colorado River basin west of the Continental Divide downstream to Reclamation's Glen Canyon Dam, and also a small part of the upper Rio Grande River basin. The total drainage area is around 147,000 mi^2. The hydropower projects here are all owned and operated by Reclamation, and they range widely in size. Topography also varies widely from the mountains in Colorado and Utah to arid plains in New Mexico. Median elevation is 6,880 ft, and maximum elevation is 14,003 ft. Land cover is a mix of grassland (47%), open scrubland (24%), evergreen forest (11%), and woody savanna (4%).

- **WAPA-4 Lower Colorado**: The lower Colorado River basin downstream of Glen Canyon Dam includes Reclamation's Hoover, Davis, and Parker Dams. The drainage area is 182,000 mi^2, including the upper Colorado above Glen Canyon. Topography is diverse, ranging from the western slope of the Rocky Mountains to the deserts of Arizona and Nevada. Median elevation is 5,561 ft, and maximum elevation is over 14,000 ft. Land cover is mostly open scrubland (69%), with minor amounts of grassland (14%), closed scrubland (10%), and woody savanna (3%).

- **WAPA-5 Rio Grande**: The lower Rio Grande River along the Texas-Mexico border includes Falcon and Amistad Dams operated by the IBWC. The drainage area contributing water to these two IBWC projects is over 200,000 mi^2, including the headwaters of the Rio Grande in New Mexico, and 71,000 mi^2 in Mexico. WAPA-5 is a much more arid area than other PMA study areas. Median and

maximum elevations are 4,035 and 13,776 ft, respectively, which are still relatively high except for in the immediate area of the two hydropower projects. Most of the land cover is open scrubland (56%), with a mix of grassland (28%), closed scrubland (9%), and woody savanna (4%).

- **WAPA-6 California**: WAPA-6 includes the Central Valley of California (Trinity, Sacramento, American, Stanislaus, and San Joaquin River systems) and Truckee and lower Carson River systems. The elevation ranges from 79 to 13,317 ft, with a median elevation of 4,754 ft. Land cover is mostly evergreen broadleaf forest (54%), plus woody savannas (20%), grassland (14%), and cropland (3%).

4.1.2 Power Marketing by WAPA

WAPA is the newest PMA, created by the DOE Organization Act in 1977. Until 1977, marketing of electricity sales from that portion of the federal hydropower fleet was conducted by Reclamation. WAPA markets the electricity from 58 hydropower plants located in 11 states, operates more than 17,000 mi of transmission lines, and manages four balancing authorities. It is the largest PMA by footprint of area served. Its 672 customers are distributed across 17 states. Sixty percent of customers are either municipalities or Native American tribes. Cooperatives, political subdivisions, and government (federal and state) agencies make up the rest of its long-term customers. It also has other short-term, non-firm marketing arrangements with IOUs and power marketers.

A general description of a PMA and its marketing practices is presented in Section 1.2.2. Apart from legislation authorizing or defining the mission of specific WAPA projects, one legislative item that applies across WAPA (and not to other PMAs) is Section 114 of the Energy Policy Act of 1992. It requires WAPA's long-term firm power customers to submit integrated resource plans every five years; a full range of alternatives, including construction of new generating capacity and purchased power and energy efficiency, is considered to provide adequate and reliable service to their customers at the lowest system cost. This requirement effectively resulted in customers having to make their resource planning process more open and participatory, consider nontraditional energy generation technologies, and consider both economic and environmental costs in their decisions.[23]

For the purposes of developing power marketing plans and setting power tariffs, the federal fleet marketed by WAPA is divided into nine projects. The upper Great Plains (UGP) region sells power from the Pick-Sloan Missouri Basin Program (Eastern Division), WAPA's Rocky Mountain region markets power from the Loveland Area Project, the Sierra Nevada region manages marketing for the Central Valley-Washoe project, the DSW region markets power from the Boulder Canyon and Parker-Davis projects, and the Colorado River Storage Project (CRSP) region markets the output from the Salt Lake City Area/Integrated Projects, Provo River, Falcon-Amistad, and Olmsted.[24]

Contract terms and electricity rates vary substantially across projects, reflecting differences in hydrologic conditions and operational constraints. Reclamation issues the Records of Decision for each project, detailing their operational restrictions based on authorized purposes and environmental requirements (Fernandez, 2020). In terms of the type of product sold, the projects can be separated into three groups: Loveland, Parker-Davis, Pick-Sloan Missouri Basin (Eastern Division), and Salt Lake City Area/Integrated Projects. These projects sell firm power and energy, typically setting different amounts for winter and summer. Their customers can set their daily requested energy/capacity schedule within the

[23] https://www.wapa.gov/newsroom/Publications/Documents/25yr-history_2.pdf

[24] The Olmsted power plant is more than 100 years old and was acquired from PacifiCorp by the US government in condemnation proceedings in 1990. However, the agreement was for PacifiCorp to continue operating the facility until 2015, at which point it became Reclamation's responsibility. As part of an implementation agreement signed in 2015 by the Central Utah Water Conservancy District, Department of the Interior, Reclamation, and WAPA, the District replaced the old plant with a new powerhouse. The new facility started operation in 2018, and WAPA markets its power.

bounds of the operational constraints and capacity limits. The product offered from the Central Valley, Olmsted, Falcon-Amistad, and Provo River projects is energy only.[25] Finally, starting in 2017, the Boulder Canyon Project markets long-term contingent capacity and firm energy, with capacity depending largely on Lake Mead elevation.

Except for the Falcon-Amistad project, all others have either published new marketing plans within the past decade or are in the process of developing them. The Boulder Canyon project underwent a full remarketing effort because it needed to implement changes to its marketing practices following the Hoover Power Allocation Act of 2011. For other projects, the key marketing principles remain unchanged in the latest plans, but there have been amendments to aspects like contract terms and resource pool allocations to new customers.

Determining marketable power resources is one of the key terms set in power marketing plans. For projects offering firm power, the capacity (contract rate of delivery) is typically set at the level consistent with a 90% exceedance (i.e., the amount of capacity available at least 90% of the time). Firm energy is the average or median projected amount. Some capacity and energy are reserved for project use and ancillary services provisions rather than marketed to preference customers. For instance, CRSP contributes 40 MW of capacity for frequency regulation and 30 MW of capacity for contingency reserves as a member of the Northwest Power Pool (Fernandez, 2020).

Reserving a percentage of the marketable resource for new customers helps supports the "most widespread use" of the federal hydropower, as mandated by the Flood Control Act of 1944 (Ch. 665, 58 Stat. 887). The Boulder Canyon project set aside a 5% resource pool for new customers in its post-2017 allocations. The number of Boulder Canyon project customers went from 15 to 46 after 2017, with new allocations ranging from 100 kW to 3 MW. Similar resource pools are set aside in the Central Valley Projects (2% in 2025 and an additional 1% in 2040), Loveland Area Projects, and Pick-Sloan Missouri basin (up to 1% every 10 years for the duration of the marketing plan). Salt Lake City Area/Integrated Projects is the only one where the latest marketing plan does not include a resource pool for new customers.

Many of the latest revisions to power marketing plans have resulted in longer contract terms. Contract terms lengthened from 30 years to 50 years in Boulder Canyon, and from 20 years to 30 years in the Central Valley Projects, Loveland Area Projects, and Pick-Sloan Missouri basin. Customers have generally been in favor of the longer terms because the longer terms provide them with stability in their resource planning. WAPA retains some flexibility to make changes during the life of the contracts through a provision in its marketing plans. The PMA can adjust allocations based on changes in hydrology or river operations with five years' prior notice.

Table 4.2 displays the capacity and energy rates from projects selling firm power. Within them, Boulder Canyon differs in that it sells contingent capacity with associated firm energy. The rest of projects (Central Valley, Provo River, Olmsted, and Falcon-Amistad) have an annual repayment schedule that is allocated among the customers in exchange for the same percentage of project output. Therefore, in those projects, the price per kilowatt-hour varies by year depending on energy availability.

[25] For the Central Valley Project, WAPA can develop custom products additional to a set percentage of the project's marketable energy. Those customer products include firming power, ancillary services, reserves, and scheduling coordinator services. Additionally, WAPA manages an exchange program among the Central Valley Project customers to ensure any energy that cannot be used in real time by the customer to whom it is allocated is made available to others.

Table 4.2. WAPA rate schedule summary (FY 2020)

Region	Project	Capacity rate ($/kW)	Energy rate (mills/kWh)	Composite rate (mills/kWh)
DSW	Boulder Canyon	1.75	9.04	18.08
Rocky Mountains	Loveland	4.12	15.72	31.44
DSW	Parker-Davis	2.84	6.49	12.97
UGP	Pick-Sloan Missouri Basin (Eastern Division)	5.25	13.27	24.00
CRSP	Salt Lake City Area/Integrated Projects	5.18	12.19	29.42

Source: WAPA's FY 2020 Statistical Appendix. Note: 1 mill = $0.001.

In years when the hydropower produced by the plants in one of the projects listed in Table 4.2 (except Boulder Canyon) is not sufficient to cover the contracted energy volumes, WAPA purchases replacement power through bilateral contracts or transactions in an organized wholesale market. The cost of that replacement power is passed to customers. Similarly, if available hydropower is greater than the contracted volumes, the surplus power can be sold to entities other than the preference customers. The cost of replacement power purchases is passed to customers. Rate designs may include a drought adder component that can be set to zero in years when replacement power purchases are not needed.

WAPA has recently taken an active interest in considering and formally joining the regional western EIMs or independent system operators (ISOs)/regional transmission organizations (RTOs).[26] The access to the larger, more diverse resource portfolios in these centralized markets can help these WAPA areas optimize their replacement power purchase and or surplus power sales. Additionally, this relationship will assist WAPA in its role as a balancing authority responsible for maintaining reliable and affordable electricity as the regional generation resource mix, transmission constraints, and ancillary service needs change.

The UGP East Region was the first WAPA region to join an organized wholesale market when it became a member of the SPP RTO in October 2015. WAPA's UGP West, Rocky Mountains, and CRSP regions joined SPP's WEIS Market at its inception in February 2021. WAPA Sierra Nevada Region began participating in the CAISO EIM in March 2021. Additionally, WAPA's Loveland Area Projects and UGP West Region announced plans to evaluate potential SPP RTO membership in November 2020, and CRSP joined in that evaluation process in April 2021. Finally, the WAPA DSW region—alongside other neighboring utilities—is in the process of evaluating the benefits of participation in either the CAISO EIM or the SPP WEIS.

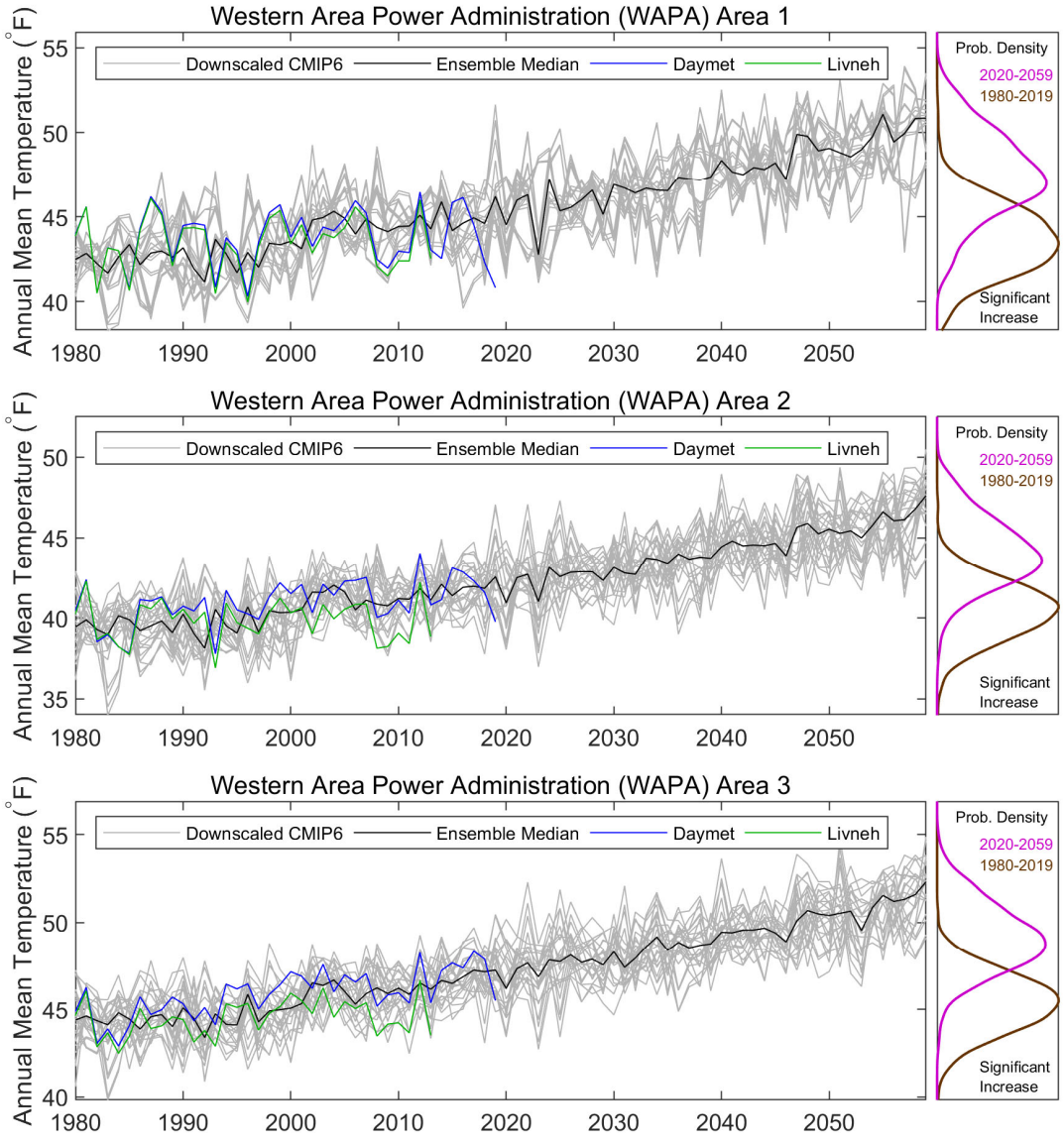
4.2 FUTURE CLIMATE IN THE WAPA REGION

4.2.1 Regional Climate Projections

The annual temperatures during the baseline (1980–2019) and future (2020–2059) periods are presented in Figure 4.2. The gray lines show the annual mean temperature from 24 sets of downscaled climate projections, the black line represents the multimodel median, and the blue and green lines show the

[26] An EIM is a real-time market (5-minute dispatch intervals) that resolves the imbalance between the supply and demand day-ahead forecasts and real-time needs. There are important differences in joining an EIM vs. an ISO/RTO. Resource offers in an EIM are voluntary and each balancing authority conducts its own transmission and reserves planning. Joining an ISO/RTO gives access to more markets (day-ahead, real-time energy and ancillary services markets) but also requires joint transmission planning and resource adequacy planning, which reduces the balancing authority's autonomy.

Daymet and Livneh historical observations, respectively. These plots show interannual variability and annual trends for projections, as well as the observations. The climate models are not expected to replicate the exact timing of the historical interannual and decadal variability. Therefore, they do not completely follow the historical interannual values. The corresponding probability distribution for the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between the baseline and future periods is statistically significant.



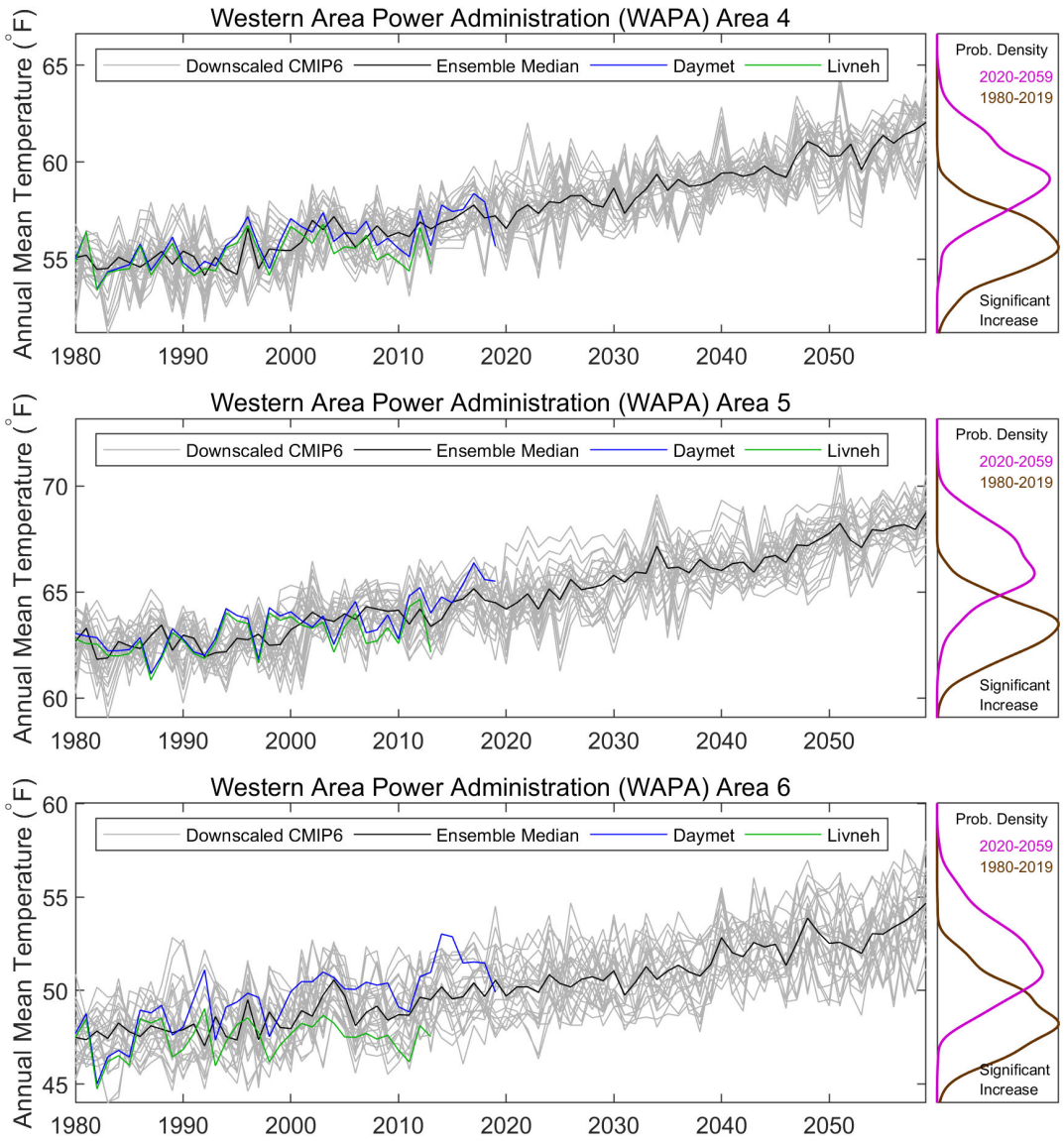


Figure 4.2. Projected annual mean temperature in the WAPA region.

In all six assessment areas (WAPA-1–6), the ensemble model median captures the increasing trend in annual mean temperature in the observations during the historical period. However, there is significant variability across the ensemble members, which reflects the importance of using a multimodel ensemble approach. In the future period, annual mean temperature is projected to continually increase across all the six areas. A comparison of the probability distribution of multimodel, mean annual temperature between the historical and future periods also shows a statistically significant shift toward warmer climate conditions across all the WAPA areas (Figure 4.2).

The annual and seasonal changes in temperature are further summarized in Figure 4.3. The change represents the difference in temperature during the future periods (2020–2039 near-term and 2040–2059 mid-term) compared with the baseline period (1980–2019) in degrees Fahrenheit. Each box plot shows the spread across 24 downscaled climate projections; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length

denoted by x are considered statistical outliers (but should not be excluded in the context of climate change).

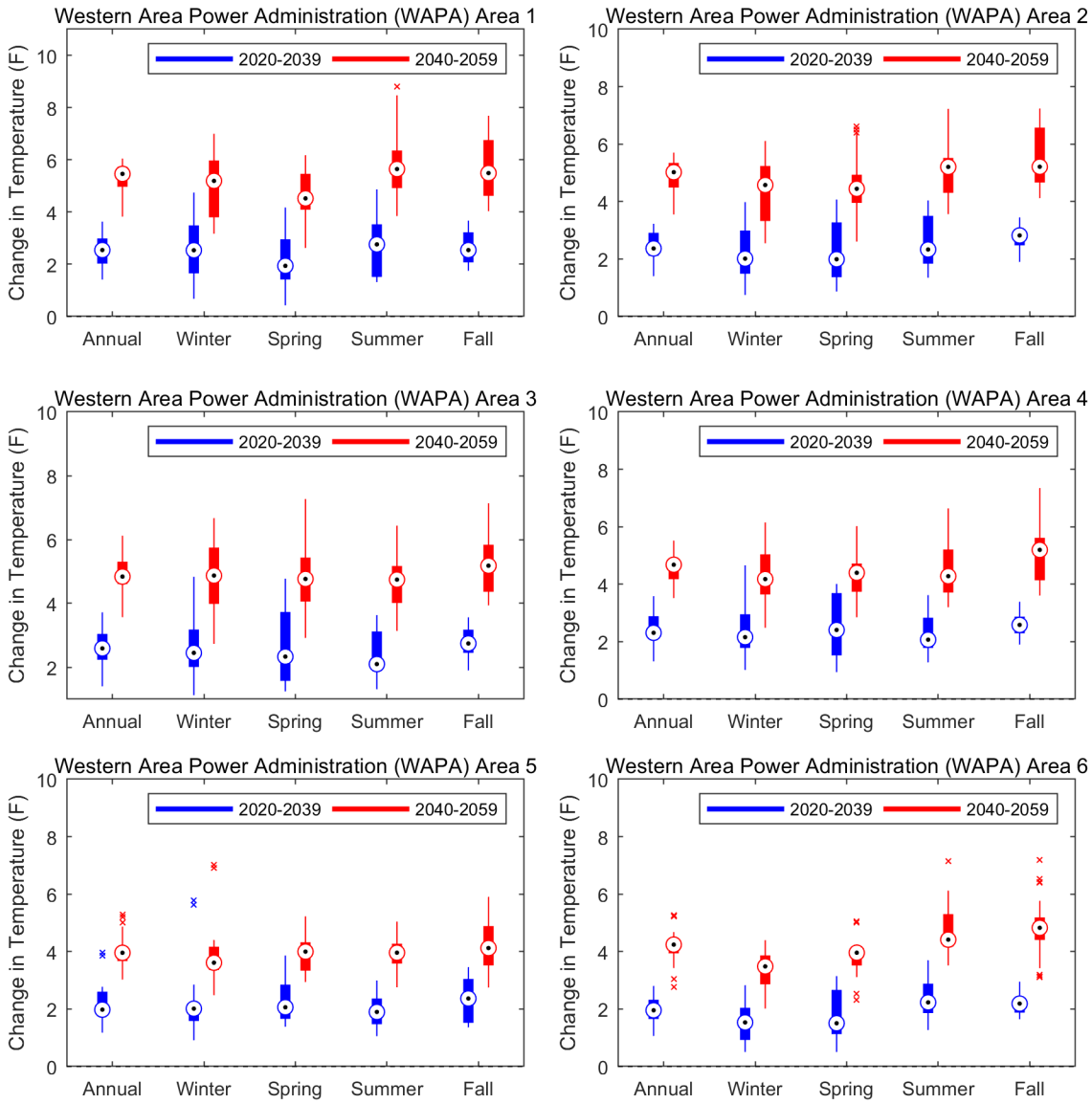
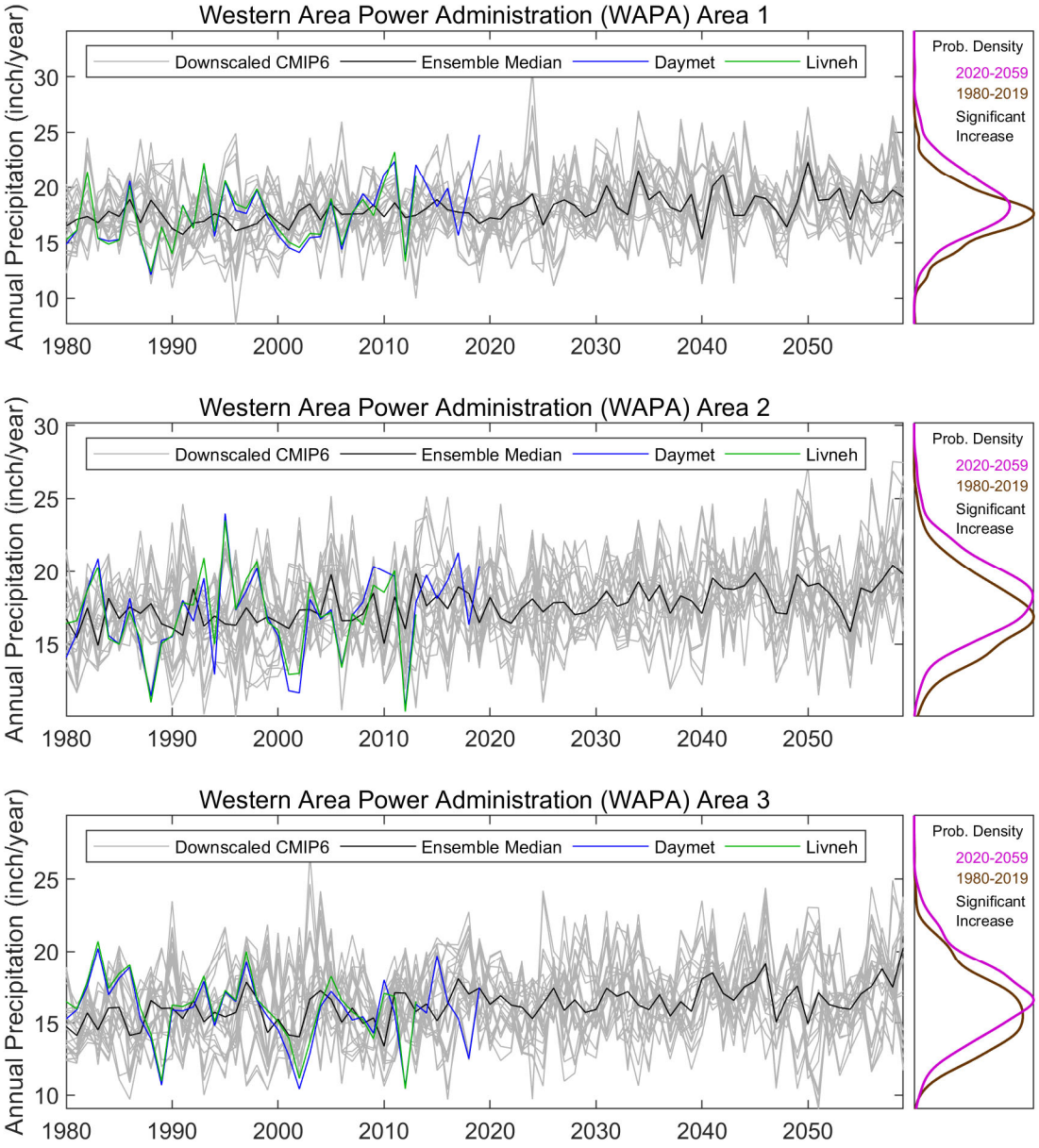


Figure 4.3. Projected change of annual and seasonal mean temperature in the WAPA region.

Annual and seasonal temperature is projected to increase across all the ensemble members during both the near-term and mid-term future periods. A relatively large increase is projected during the mid-term future period. The multimodel median annual temperature in the WAPA region is projected to increase by approximately 1°F–3°F in the near-term and 3°F–6°F in the mid-term future periods. The highest increase is projected during the summer and fall seasons (up to 8°F in the mid-term future period) over WAPA-1, WAPA-2, WAPA-4, and WAPA-6; however, the projected increase in temperature is consistent across all the seasons in WAPA-3 and WAPA-5 areas.

The annual precipitation (Figure 4.4) also exhibits high interannual and intermodel variability. A comparison of the probability distribution between the baseline and future periods indicates a statistically significant increase in precipitation across most of the WAPA areas, except WAPA-5, which shows no

significant change. The annual and seasonal changes of precipitation are summarized in Figure 4.5. In the near-term (2020–2039), the ensemble median annual precipitation is projected to increase over WAPA-1–3 in contrast to a small to no change in WAPA-4–6. In the mid-term future period (2040–2059), the ensemble median annual precipitation is projected to increase across most of the region except over WAPA-5, where it shows a small positive change. A higher increase is projected during the winter and spring for WAPA-1–4, fall for WAPA-5, and the summer for WAPA-6 in the mid-term future.



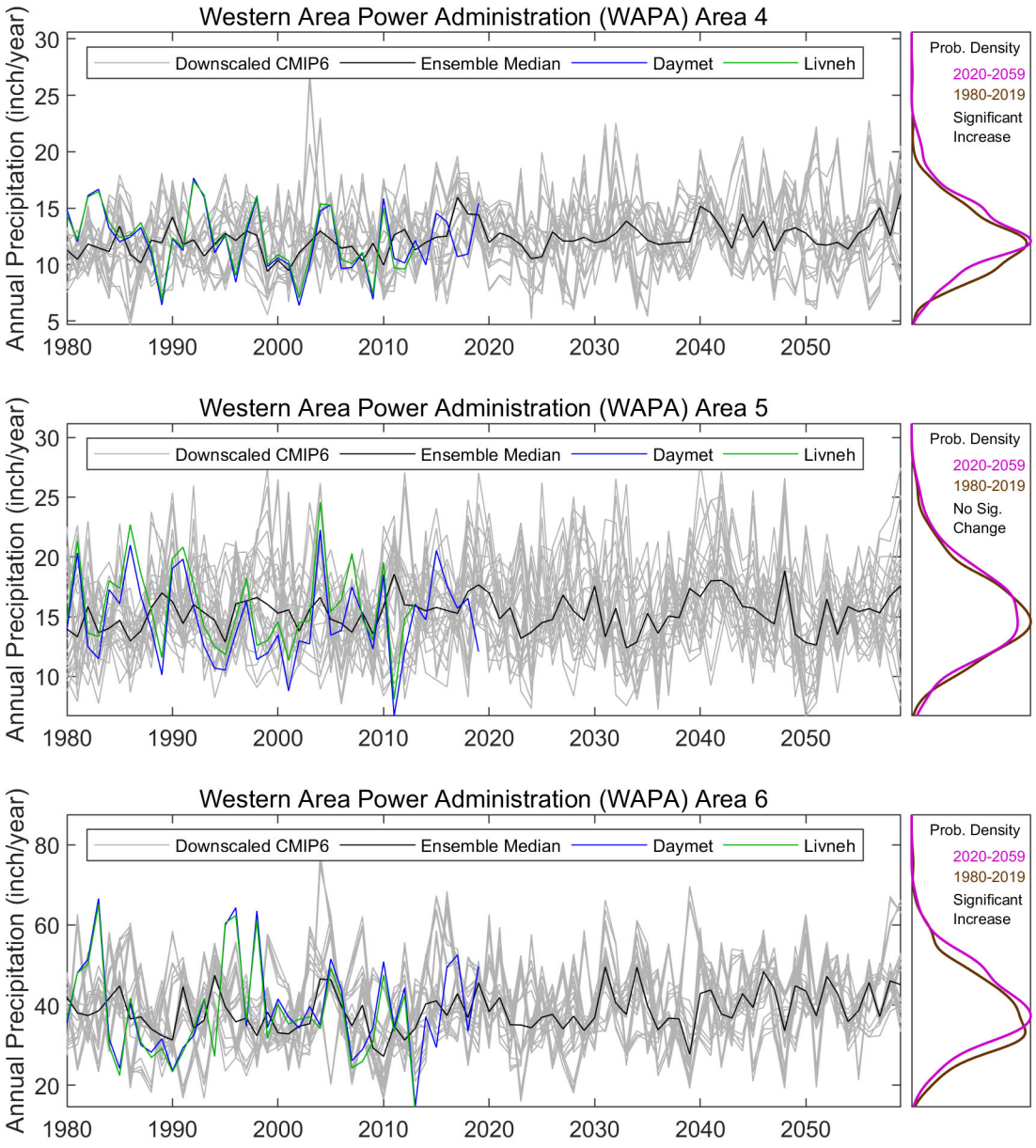


Figure 4.4. Projected annual total precipitation in the WAPA region.

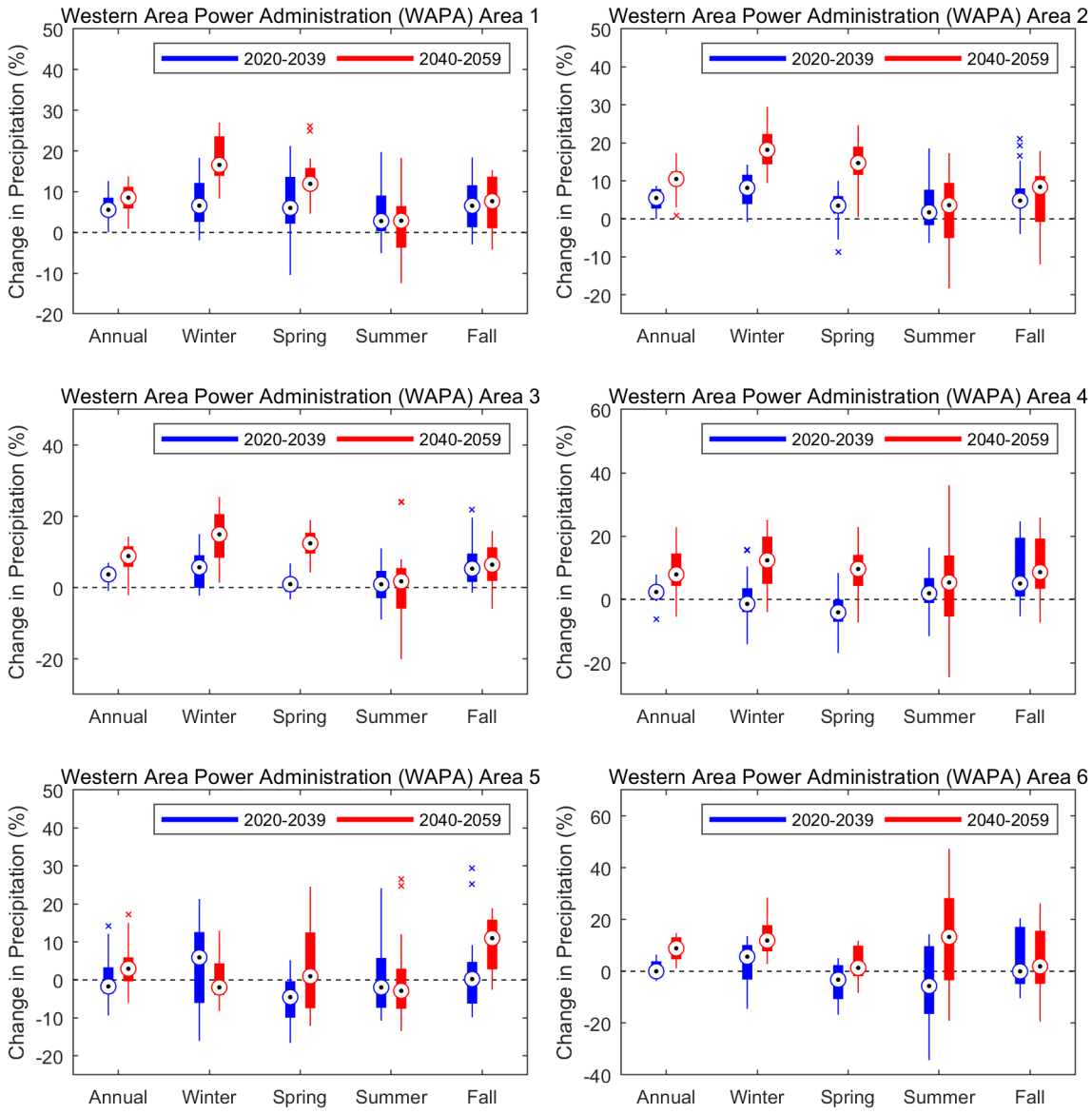
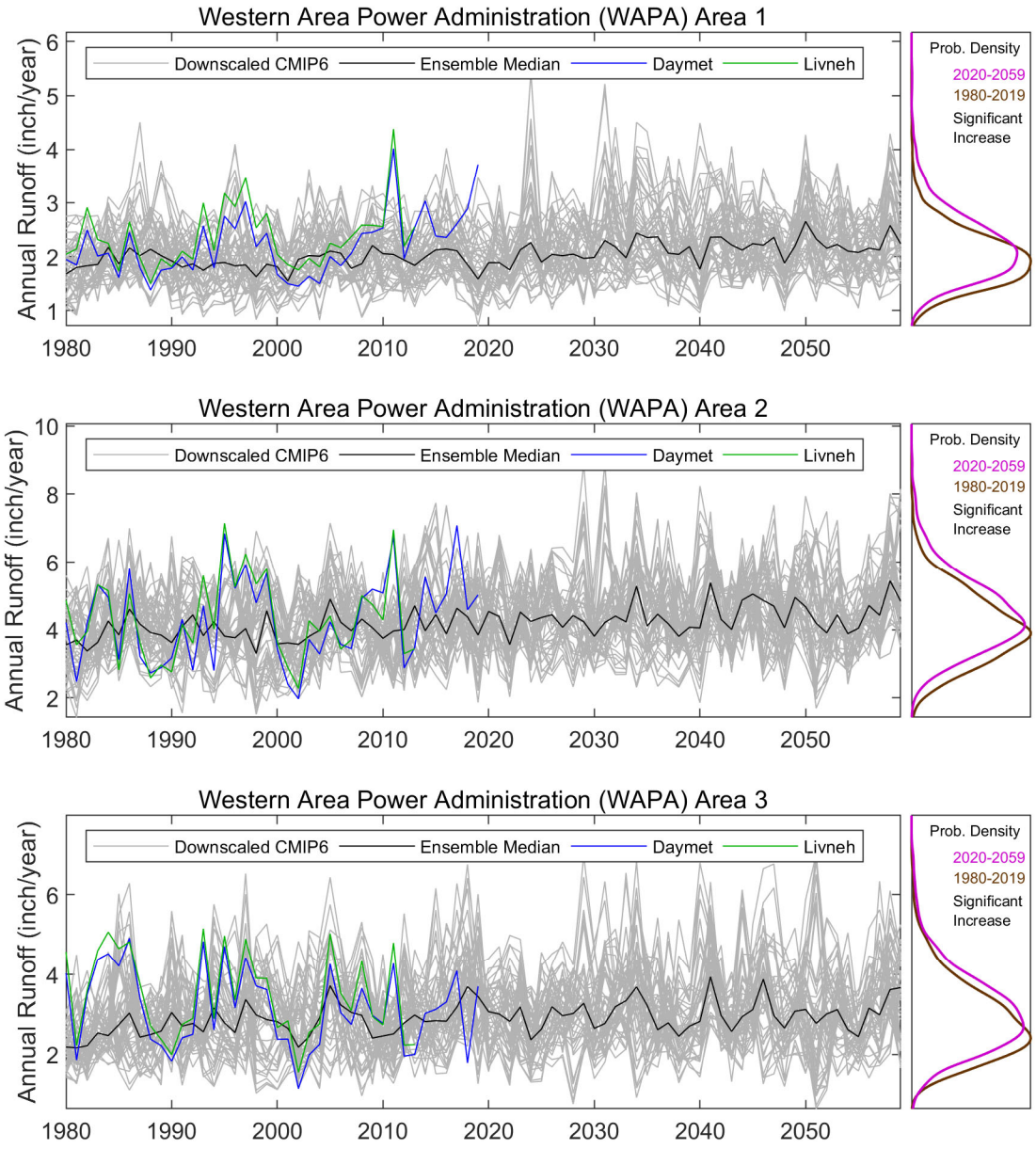


Figure 4.5. Projected change of annual and seasonal total precipitation in the WAPA region.

4.2.2 Regional Hydrologic Projections

The runoff projections for each of the WAPA regions are illustrated in Figure 4.6 and Figure 4.7. The annual total runoff in baseline (1980–2019) and future (2020–2059) periods is presented in Figure 4.6. The gray lines show annual runoff from 48 sets of hydrologic projections, the black line represents the multimodel median, and the blue and green lines show the control simulations driven by Daymet and Livneh observations, respectively. The corresponding probability distributions are compared in the right panel. The annual and seasonal changes in runoff are further summarized in Figure 4.7. Change is defined as the percentage difference of future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across 48 ensemble members; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members

outside of the maximum whisker length are considered as statistical outliers (but should not be excluded in the context of climate change).



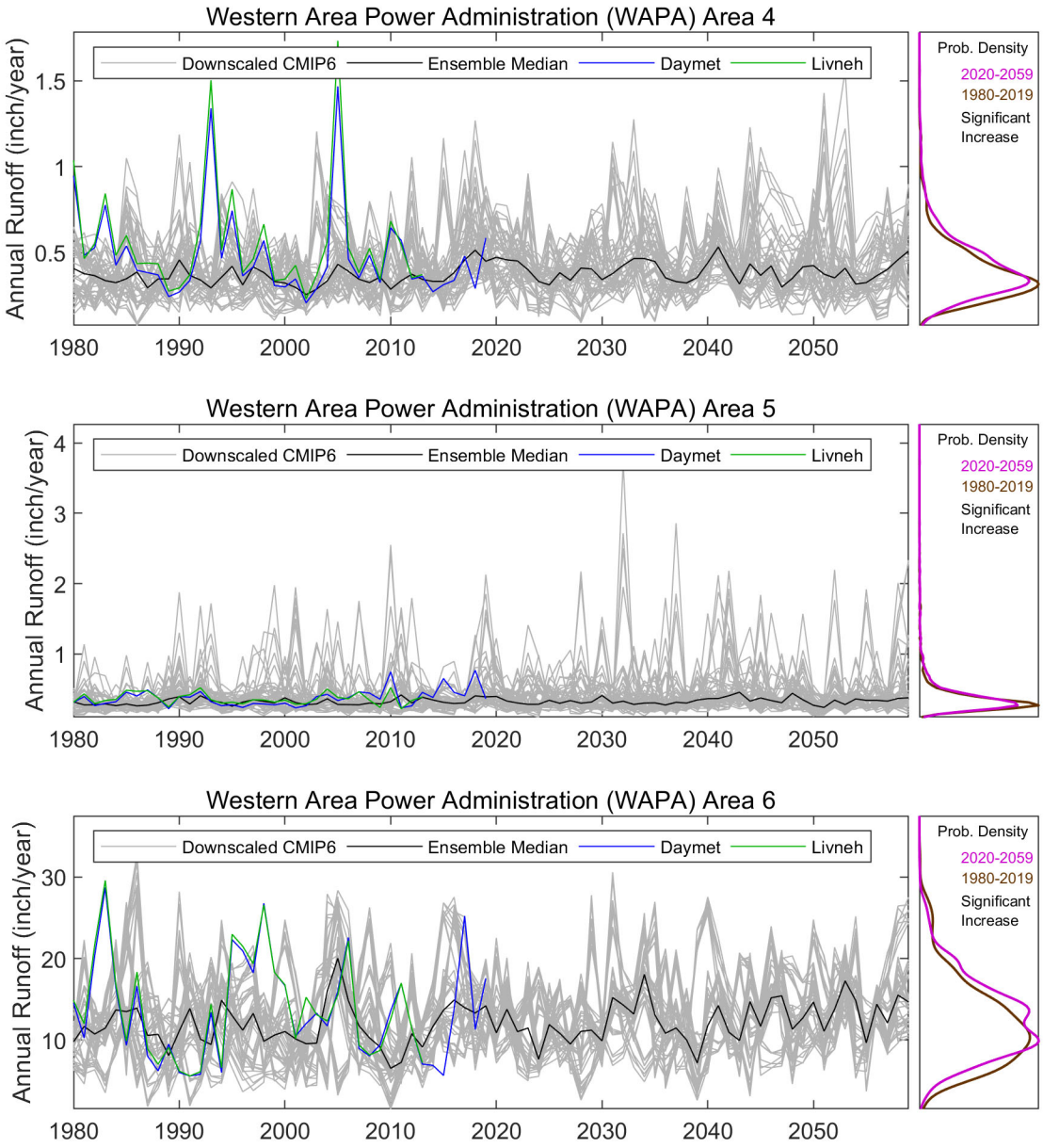


Figure 4.6. Projected annual total runoff in the WAPA region.

Notably, the runoff projections shown in Figure 4.6 and Figure 4.7 only represent results from six selected CMIP6 GCMs. Given the large variation across all CMIP6 GCMs, future projections have large uncertainty that should be clearly acknowledged. Additionally, limitations in the model's ability to capture extremes means runoff availability may actually be less than what is projected in this study. For instance, Woodhouse et al. (2016) examined the empirical relationship between precipitation and runoff in the Upper Colorado River, suggesting that the recent droughts have been amplified by warmer temperatures. These warmer temperatures exacerbate the effects of relatively modest precipitation deficits. Williams et al. (2021) further suggested that the current "megadrought" in the southwestern United States is a function of climate change, and a similar soil moisture drying signal can be seen in all CMIP6 models. Finally, Udall and Overpeck (2017) concluded that the continued business-as-usual warming will drive temperature-induced declines in river flow throughout the end of the century.

Findings from these studies suggest that there is a high uncertainty and a knowledge gap in future water availability projections.

In addition to the change of precipitation patterns in the WAPA region, these seasonal changes in runoff could also relate to the change of snow hydrology. To further evaluate snow dynamics in the region, changes on April 1 SWE are presented in Figure 4.8. The change is defined as the percentage difference of April 1 SWE in future periods (2020–2039 and 2040–2059) when compared with the baseline period (1980–2019). The changes are summarized for HUC8s with more significant snowpack (defined as HUC8s with average 1980–2019 control run April 1 SWE greater than 5 mm). The results suggest a strong decrease in ensemble median April 1 SWE (Figure 4.8, 2020–2059 future to 1980–2019 baseline), with more than half of the HUC8s projecting a decrease of 24% or more. In addition, decrease of April 1 SWE is projected in the northwestern portion of WAPA-1, a significant portion of WAPA-2 and WAPA-3, and WAPA-6. These areas are also associated with lower summer runoff projections, suggesting the relationships between changing snow hydrology and summer runoff reduction.

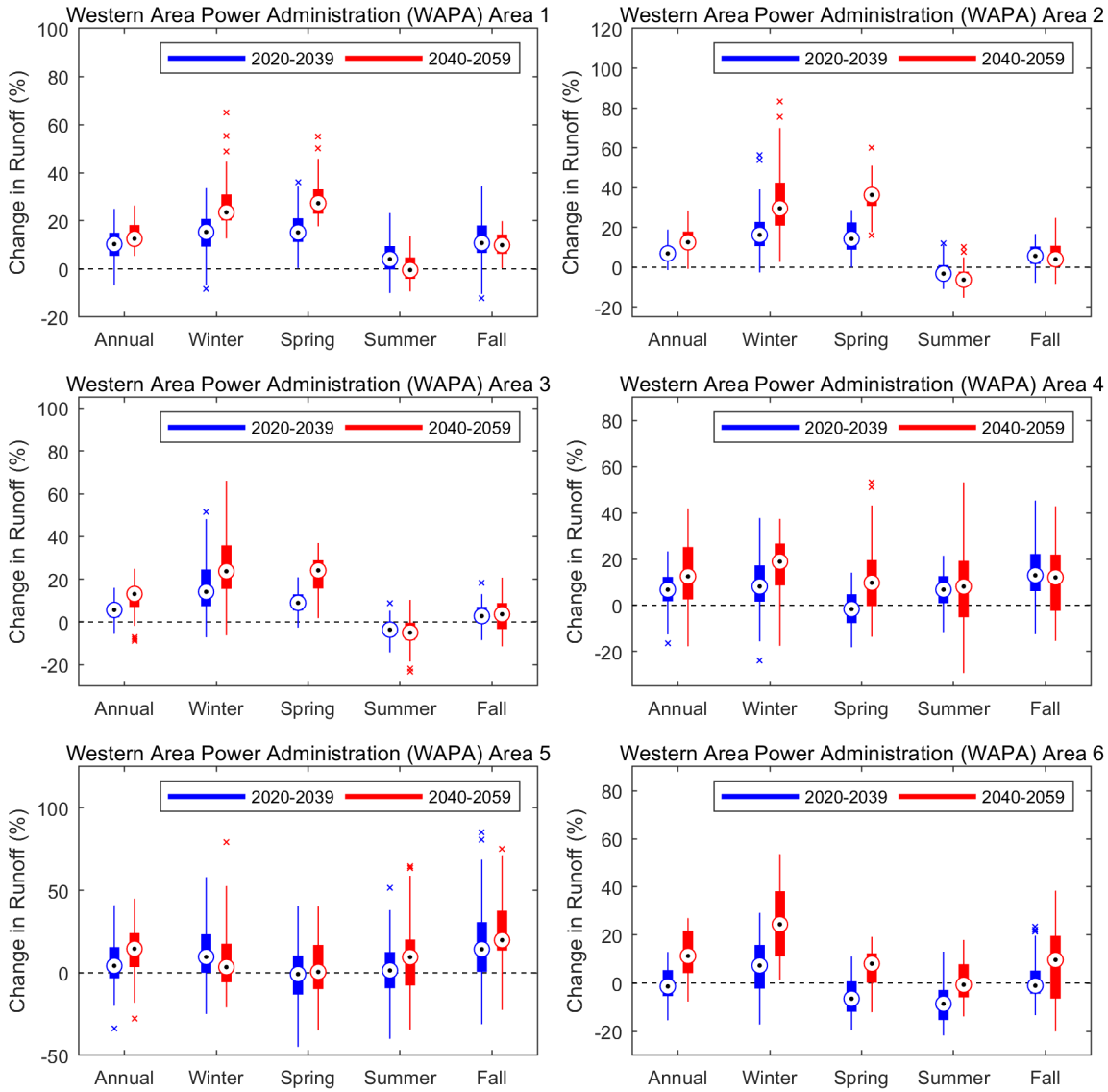


Figure 4.7. Projected change of annual and seasonal total runoff in the WAPA region.

The shifts of probability distributions (Figure 4.6) suggest that the annual runoff is projected to increase in the future period and the distributions are statistically different. Additionally, the probability distributions for all WAPA areas, except for WAPA-6, show a wider distribution in the future period, indicating that the interannual variability is projected to increase. Such a change is more prominent and noticeable in WAPA-5. The maximum increase in median annual runoff is projected in the range of 6.8% to 12.3% across all WAPA areas. For WAPA-1, the median annual runoff is projected to increase by 10.3% in the near-term and 12.5% in the mid-term future periods. WAPA-2–5 show a projected increase in the range of 4.2% to 6.9% in the near-term and 12.5% to 14.6% in the mid-term future periods. WAPA-6 shows a decrease of 1.4% in the near-term but an increase by 11.3% in the mid-term future periods. The range of annual runoff change is also higher in the mid-term future period for WAPA-2, WAPA-3, WAPA-4, and WAPA-6, indicating greater uncertainties in runoff projections.

At the seasonal scale, the strongest changes are the increase in winter and spring runoff, accompanied with little to no change or decrease in summer runoff, especially for WAPA-1–3. An increase in winter runoff is also projected in WAPA-4–6, along with a decrease in spring runoff in the near-term and an increase in the mid-term future periods. More specifically, the ensemble median winter runoff is projected to increase across all WAPA areas in the range of 7.3% to 16.2% in the near-term future period. The change is projected to increase across all WAPA areas, except for WAPA-5, in the range of 18.9% to 29.6% in the mid-term future period. For WAPA-5, winter runoff is projected to increase by 9.6% in the near-term and 3.4% in the mid-term future periods. The spring runoff is projected to increase in WAPA-1, WAPA-2, and WAPA-3 in the range of 8.9% to 15.1% in the near-term and 24.1% to 36.3% in the mid-term future periods. The spring runoff in WAPA-4 shows a decrease of 1.6% in the near-term and an increase of 9.7% in the mid-term future periods. WAPA-5 shows little to no change in the winter runoff conditions. WAPA-6 projects a decrease of 6.4% in the near-term and an increase of 8.1% in the mid-term future periods. Summer runoff is expected to decrease in WAPA-2, WAPA-3, and WAPA-6, in the range of 3.6% to 8.6% in the near-term and 0.7% to 6.3% in the mid-term future periods. On the other hand, summer runoff is projected to increase in WAPA-4 and WAPA-5 in the range of 1.4% to 6.8% in the near-term and 8.1% to 9.5% in the mid-term future periods. For WAPA-1, the summer runoff is projected to increase by 4.1% in the near-term but shows little to no change in the mid-term (−0.47%) future periods. Finally, the fall runoff is projected to increase across all WAPA areas, except for WAPA-6, in the range of 2.7% to 14.3% in the near-term and 3.6% to 19.8% in the mid-term future periods. For WAPA-6, fall runoff shows a decrease of 1.1% in the near-term but an increase of 9.6% in the mid-term future periods. These projections are associated with large spread among ensemble members in both near-term and mid-term future periods across all WAPA areas, indicating high uncertainties in ensemble members.

In additional to the change of precipitation patterns in the WAPA region, these seasonal changes in runoff could also relate to the change of snow hydrology. To further evaluate snow dynamics in the region, changes on April 1 SWE are presented in Figure 4.8. The change is defined as the percentage difference of April 1 SWE in future periods (2020–2039 and 2040–2059) when compared with the baseline period (1980–2019). The changes are summarized for HUC8s with more significant snowpack (defined as HUC8s with average 1980–2019 control run April 1 SWE greater than 5 mm). The results suggest a strong decrease in ensemble median April 1 SWE (Figure 4.8, 2020–2059 future to 1980–2019 baseline), with more than half of the HUC8s projecting a decrease of 24% or more. In addition, decrease of April 1 SWE is projected in the northwestern portion of WAPA-1, a significant portion of WAPA-2 and WAPA-3, and WAPA-6. These areas are also associated with lower summer runoff projections, suggesting the relationships between changing snow hydrology and summer runoff reduction.

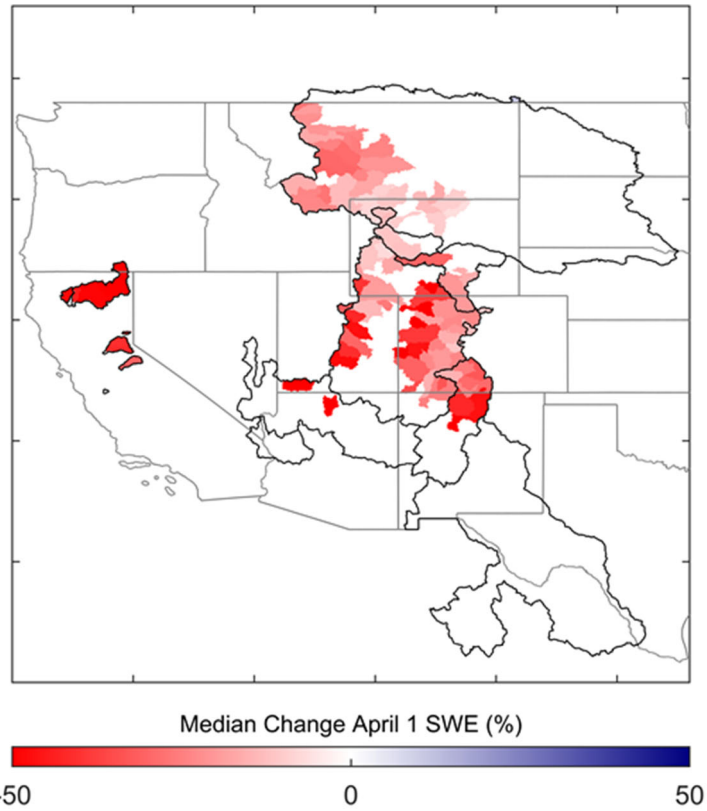


Figure 4.8. Projected change of April 1 SWE in the WAPA region.

An additional analysis of extreme runoff is shown in Figure 4.9, which presents future changes in ensemble median high runoff (i.e., 95th percentile of daily runoff) and median low runoff (i.e., 5th percentile of 7-day average runoff) in both future projection periods. High runoff is generally projected to increase in both future periods across WAPA-1, WAPA-2, and WAPA-3. The low runoff is also projected to increase in general across these areas, indicating overall wetter trends under projected future climate conditions. The high runoff in WAPA-4, WAPA-5, and WAPA-6 shows a mixed response and higher spatial variability, whereas the low runoff is projected to decrease in most of the areas except WAPA-6. These results suggest increasing likelihood of more frequent flood events for parts of WAPA regions, which may increase the difficulty of water management in the future climate conditions (Ban et al., 2020; Xiao et al., 2018). Also, as stated previously, these projections were only from six selected GCMs, so large uncertainty, either due to GCM selection or model limitations, may exist. Furthermore, other studies of the western United States predict lower water availability based on region-specific modeling assumptions (Lukas and Payton, 2020). Parts of the WAPA regions (e.g., Colorado River Basin) are experiencing a more than 20-year drought that is interrupting regional water supply and hydropower generation (Williams et al., 2021). The potential impacts due to severe drought events should not be underestimated.

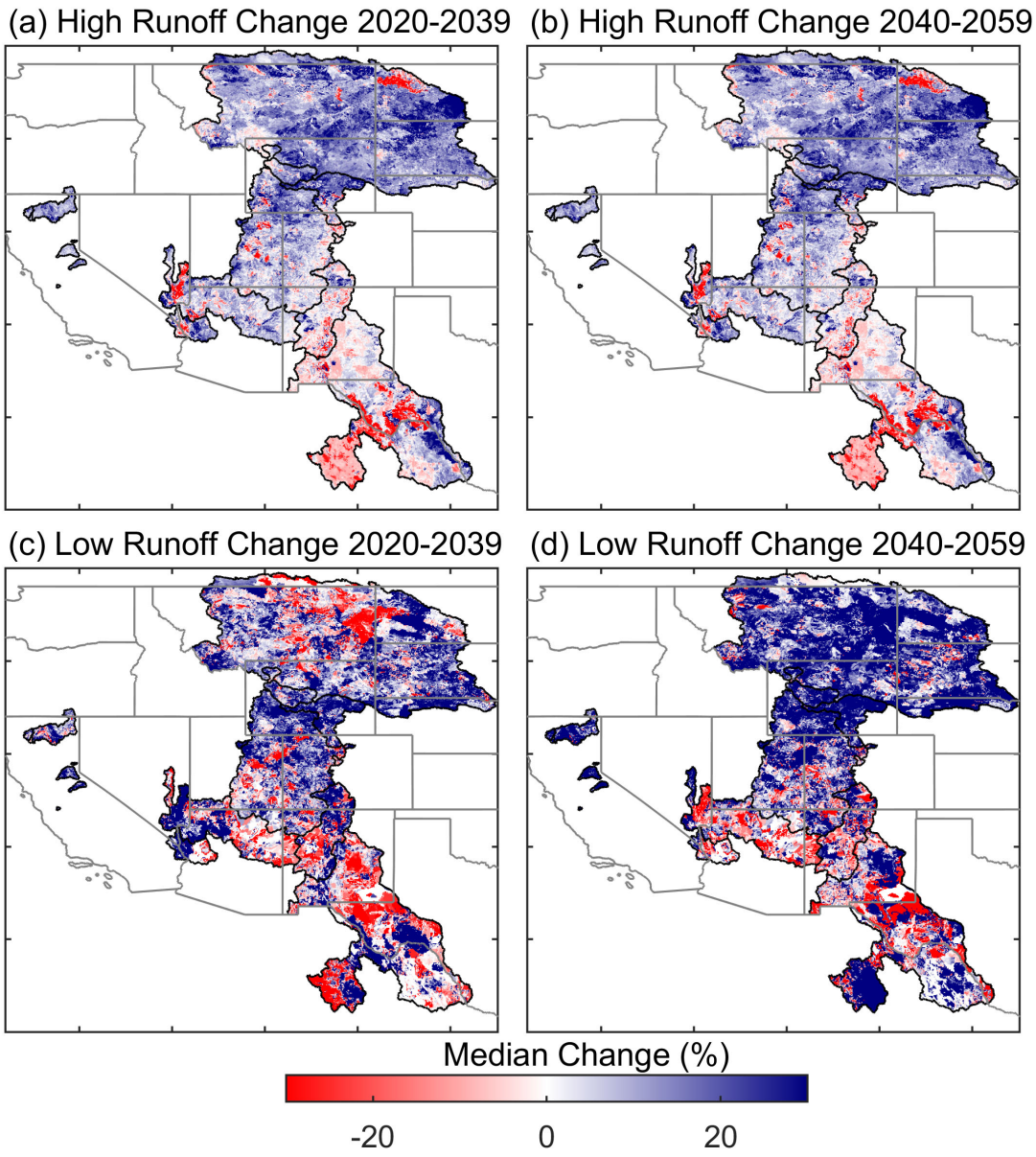


Figure 4.9. Projected change of high and low runoff in the WAPA region.

To evaluate the impacts of modeling choices (six CMs, two DSs, two MetFs, and two HMs) on the projection of future runoff in each WAPA area, an ANOVA was conducted and is illustrated in Figure 4.10. The box plots show the magnitude of change (future minus baseline) in median annual runoff across 48 ensemble members, grouped by each of the four modeling choices. The box plots provide a quantitative way to evaluate the range of possible outcomes in runoff response based on different modeling choices. ANOVA results indicate that the CM is generally the most dominant contributor to the total variance of future runoff projections. The contribution of the CM is the highest among all factors across all WAPA areas and can explain up to 44%–66% of total runoff variance. In WAPA-1, WAPA-3, and WAPA-6, the HM is the second most dominant factor, contributing 18.5%–29.4% of total variance, indicating that annual runoff response under future conditions is also sensitive to the choice of HM. In WAPA-5, contribution from other factors such as the DS and MetF is also important. At the seasonal scale (APPENDIX E), the ANOVA indicates that factors other than the CM may have stronger impacts than at the annual scale. For instance, although the CM can explain 50%–66% of winter runoff variance

from WAPA-1–5, the MetF in WAPA-2, the HM in WAPA-1, and a combination of the CM and DS in WAPA-4 and WAPA-5 have a nonnegligible influence. Similarly, the HM is found to be the most dominant contributor of spring runoff variance in WAPA-3. The influence of the HM in spring runoff variance in WAPA-1, WAPA-2, WAPA-3, and WAPA-6 also ranges from 25.5% to 42%. Although the CM can primarily contribute to the total variance in summer and fall, other factors also have nonnegligible influences.

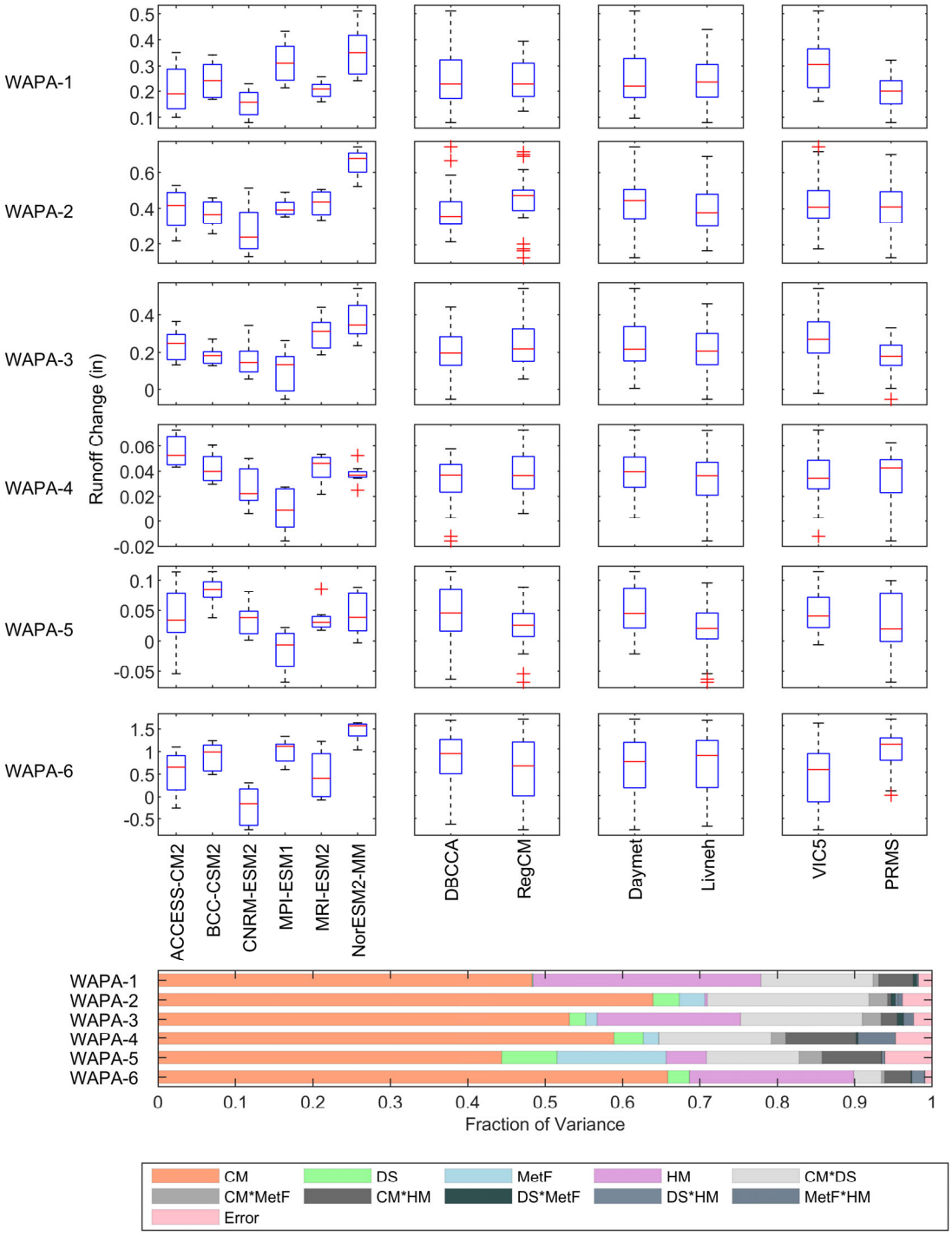


Figure 4.10. ANOVA of annual total runoff in the WAPA region.

The reservoir evaporation loss from federal hydropower reservoirs in the WAPA region simulated by LEM is shown in Figure 4.11. The blue and green lines show annual reservoir evaporation loss (in units of million acre-feet) from 24 projections that were either bias-corrected or trained by Daymet and 24 projections that were either bias-corrected or trained by Livneh. The dark blue and dark green lines represent the control simulations driven by 1980–2019 Daymet and 1980–2013 Livneh, respectively. The right panel shows the probability density distributions of annual evaporation loss in both baseline (1980–2019) and future (2020–2059) periods.

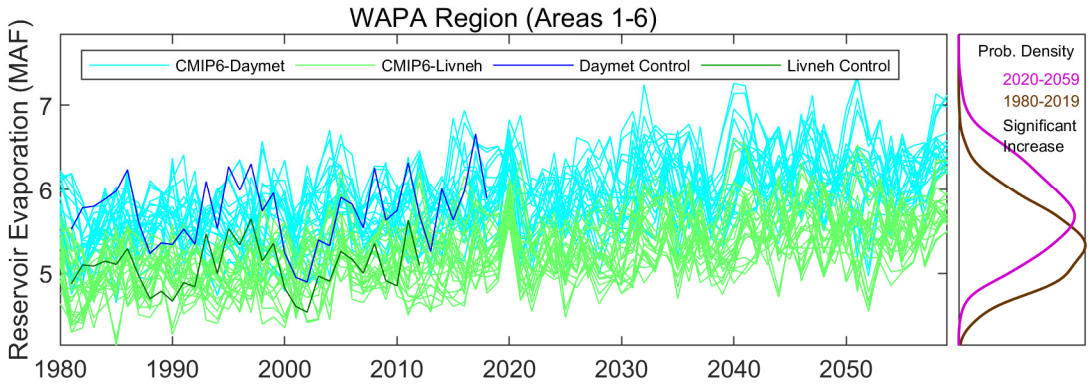


Figure 4.11. Projected annual total reservoir evaporation in the WAPA region.

The high-value group contains the 24 members that are associated with Daymet, and the low-value group contains the 24 members that are associated with Livneh. The significant difference in wind speed between these two groups is the main cause of the different evaporation rates (Section 3.2.2). Comparisons with the in situ wind speed measurements (not shown) suggest that the NCEP1 wind speed used in the Livneh data set tends to underestimate more considerably (Moreo et al., 2015). The reservoir evaporation loss also depends on the projected reservoir area, which is inferred from runoff simulated by either VIC or PRMS. The projected runoff associated with Daymet or Livneh is not significantly different because the HMs were calibrated separately by both MetFs against the same USGS WaterWatch data set (Section 2.5.3) to minimize the impacts of different meteorological forcings on hydrologic projections. Therefore, the noticeable systematic difference of reservoir evaporation loss in the WAPA region is likely mainly caused by the projected evaporation rate driven by different meteorological observations.

Despite the systematic differences between the two groups, their estimated reservoir evaporation changes show a clear and consistent increasing trend (as indicated by the two-sample Kolmogorov-Smirnov test performed at the 5% significance level). The Livneh-based future reservoir evaporation projections increase significantly from about 5 MAF in 1980 to 5.5 MAF in 2059, with a nearly 10% increase over 80 years. Daymet-based projections also increase similarly from 5.5 MAF in 1980 to 6 MAF in 2059. Overall, the reservoir evaporation loss in the WAPA region is the largest across all PMA regions. In this region, about 19% of the increased evaporation loss is attributed to the higher evaporation rate, and the rest (81%) is from the increasing reservoir area caused by greater runoff.

The annual and seasonal changes of monthly reservoir evaporation (in units of thousand acre-feet per month) from federal hydropower reservoirs in the WAPA region are depicted in Figure 4.12. The trend of increasing reservoir evaporation is persistent in both the near-term (2020–2039) and mid-term (2040–2059) future periods with respect to the baseline period (1980–2019). The annual ensemble median evaporation loss is projected to increase by about 23 KAF/month during the near-term and 42 KAF/month in the mid-term future periods. In spring, there is little change with the reservoir evaporation loss in both near-term and mid-term future periods. The reservoir evaporation loss increases more significantly in summer, winter, and fall. Summer evaporation loss is expected to increase by 30 KAF/month in the near-

term and 32 KAF/month in the mid-term future periods. Winter evaporation loss is expected to increase by 33 KAF/month in the near-term and 60 KAF/month in the mid-term future periods. The projected increase is dominant in fall—about 43 KAF/month in the near-term and 77 KAF/month in the mid-term future periods. The variance of evaporation changes in fall is primarily attributed to the large variance of the projected area from different ensemble members.

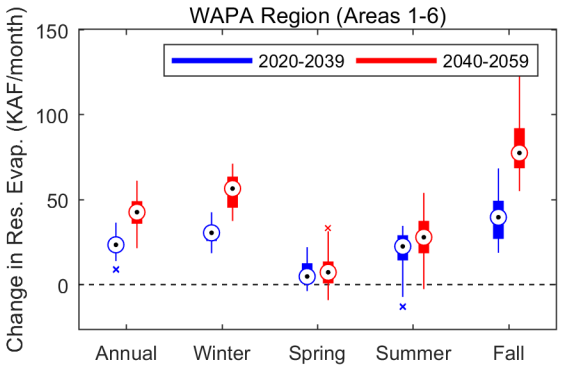


Figure 4.12. Projected change of annual and seasonal total reservoir evaporation in the WAPA region.

The source of variance in the projected change of annual reservoir evaporation loss from federal hydropower reservoirs in the WAPA region is shown in Figure 4.13. The box plots show the magnitude of change (future minus baseline) in median annual reservoir evaporation loss across 48 ensemble members, grouped by each of the four modeling choices (CM, DS, MetF, and HM). The bottom panel shows the portion of variance from the ANOVA for each modeling choice.

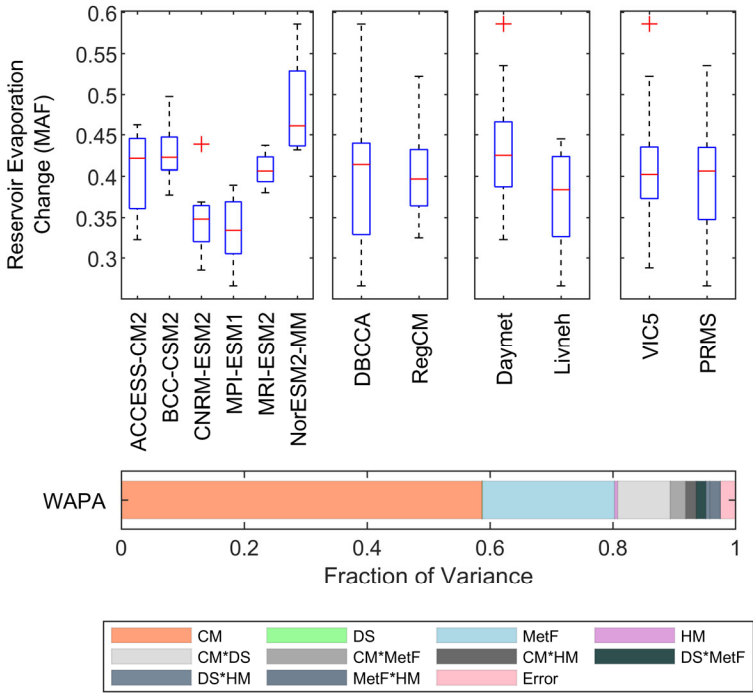


Figure 4.13. ANOVA of annual total reservoir evaporation in the WAPA region.

ANOVA results indicate that the CM is generally the most dominant contributor to the total variance of future evaporation change projections, explaining 58% of the variance in the WAPA region. The most

significant increment is from NorESM2-MM (0.46 MAF), which has a large ensemble variability. In contrast, the changes of evaporation loss projected by CNRM-ESM2 and MPI-ESm1 are less than 0.35 MAF. The MetF is the second most dominant contributor, accounting for 21% of the total variance. Daymet-based projections show a larger increase compared with Livneh-based projections. The combination of the CM and DS contributes another 8% to the total variance, and the influence of other factors is negligible. At the seasonal scale (APPENDIX F), the ANOVA indicates that certain contributing factors may have more substantial impacts than at the annual scale. For instance, the CM remains the most dominant variance contributor, which explains 74% of the evaporation loss variance in the spring. The largest changes are associated with the NorESM2-MM and MPI-ESM2. As the second most dominant contributor, the MetF explains about 10% of the total variance. The CM is also the most dominant contributor in fall and winter. The second most dominant contributor in the fall and winter is the DS, which can explain 20% of the evaporation loss variance in the fall, and 10% in the winter. During the fall and winter, DBCCA projects a higher change of reservoir evaporation losses in the future than RegCM. In the summer, the CM and DS contribute about the same amount of total variance, both around 33%. In the summer, RegCM projects a 0.09 MAF increase in summer reservoir evaporation, whereas DBCCA projects a 0.05 MAF increase. In addition, the MetF contributes almost 20% of the total variance, with Daymet projecting a more significant increase than Livneh in summer. Although the choice of MetFs (i.e., Daymet or Livneh) may significantly influence the projected value of reservoir evaporation loss (Figure 4.11), the difference is systematic and largely reduced when examining the change from baseline to future periods. This shows that even with large model/data bias, the relative change is still considered reliable and used in many hydroclimate impact assessment studies.

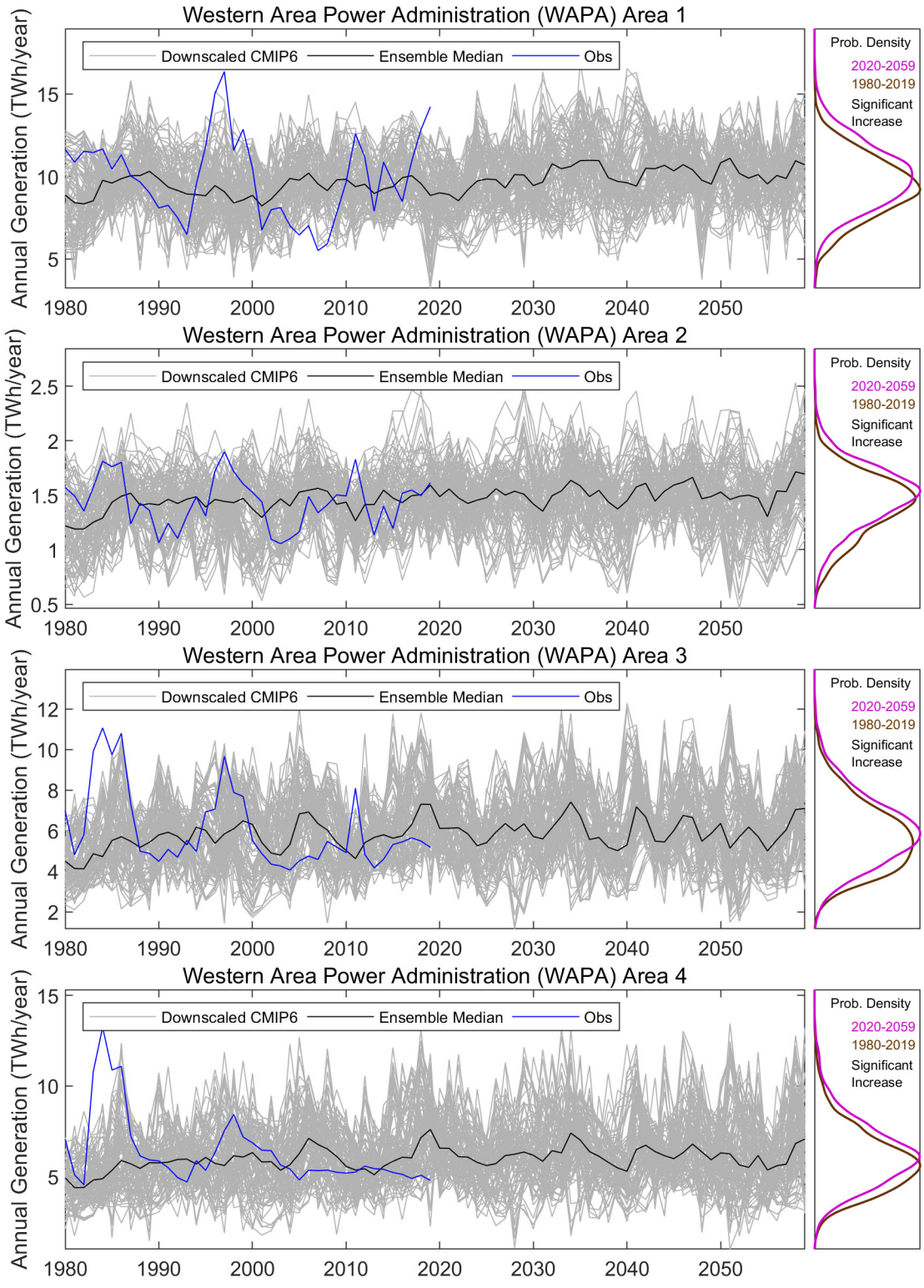
4.3 CLIMATE EFFECTS ON FEDERAL HYDROPOWER IN THE WAPA REGION

4.3.1 Projections of Hydropower Generation

Using the two hydropower models described in Section 2.6, the projections of monthly generation in the WAPA region are calculated for each of the 96 unique combinations of CMs, DSs, MetFs, and HMs and PMs. Projections based on multimodel ensembles are summarized in terms of the mean annual and seasonal changes for the near-term (2020–2039) and mid-term (2040–2059) future periods compared with the baseline period (1980–2019).

The interannual variability of annual hydropower generation for the baseline and future periods is shown in Figure 4.14. The gray lines show the annual total generation for all 96 members of the ensemble hydropower projections, and the black line represents the ensemble median; the blue line shows the 1980–2019 historic observation from EIA and PMAs. The corresponding probability distributions of annual generation in the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between baseline and future periods is statistically significant.

The comparison (Figure 4.14, left column) suggests that the range of simulated and observed annual hydropower is generally consistent across all WAPA areas. However, the shapes of the probability distributions revealed that significant changes in interannual variability will occur in some WAPA areas. Specifically, in WAPA-2, WAPA-3, WAPA-5, and WAPA-6, the future generations show narrower distributions with sharper peaks (Figure 4.14, right column). For WAPA-1 and WAPA-4, no noticeable difference was observed in the shape of probability distributions between baseline and future periods; however, a shift to higher generation is noticeable.


Western Area Power Administration (WAPA) Area 1
Downscaled CMIP6
Ensemble Median
Obs
Annual Generation (TWh/year)
Prob. Density
2020-2059
1980-2019
Significant Increase
Western Area Power Administration (WAPA) Area 2
Western Area Power Administration (WAPA) Area 3
Western Area Power Administration (WAPA) Area 4

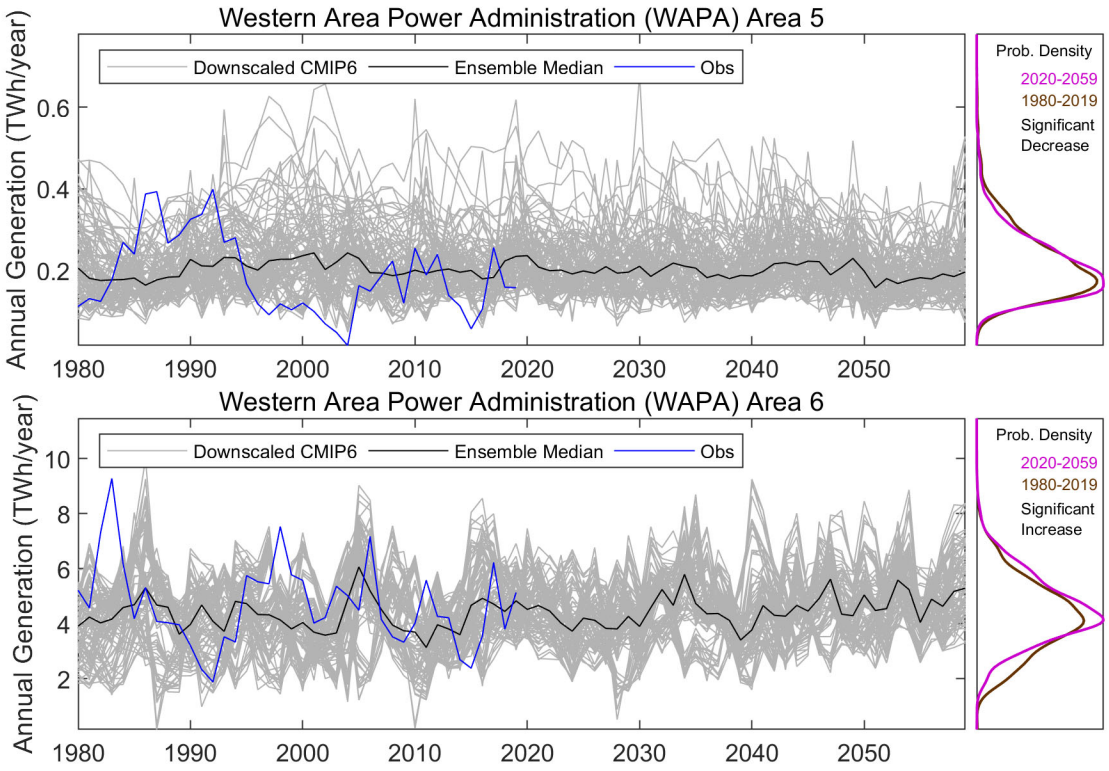


Figure 4.14. Projected annual total generation in the WAPA region.

The annual and seasonal change in hydropower generation is further summarized in Figure 4.15. The change in Figure 4.15 is defined as the percentage difference of future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across 96 multimodel combinations; the central mark indicates the multimodel median, and the edges of box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length are considered statistical outliers (but should not be excluded in the context of climate change).

The annual hydropower generation is projected to increase significantly in most of the WAPA areas. More specifically, the annual hydropower generation in the near-term (2020–2039) future period is projected to increase from 2% to 7% (in terms of multimodel median) in all WAPA study areas, except for WAPA-5, where the annual generation is expected to remain stable (Figure 4.15). Similarly, the annual hydropower generation in the mid-term (2040–2059) future period is projected to increase by 10% in all WAPA study areas, except for WAPA-5, where the annual generation will increase slightly (2%).

The WAPA study areas cover vast regions from Upper Missouri to California with diverse climate features, and this diversity is also reflected in the future seasonal hydropower generation projections. In WAPA-1, seasonal generation is expected to increase by 4%–8% in the near-term and 8%–14% in the mid-term future periods. The highest increase is projected in spring (8% and 14% in the two respective future periods), and the lowest increase is projected in summer (4% and 8% in the two respective future periods). The increase in hydropower generation in WAPA-1 is mostly caused by the precipitation increase in the Upper Missouri River basin, especially during winter and spring.

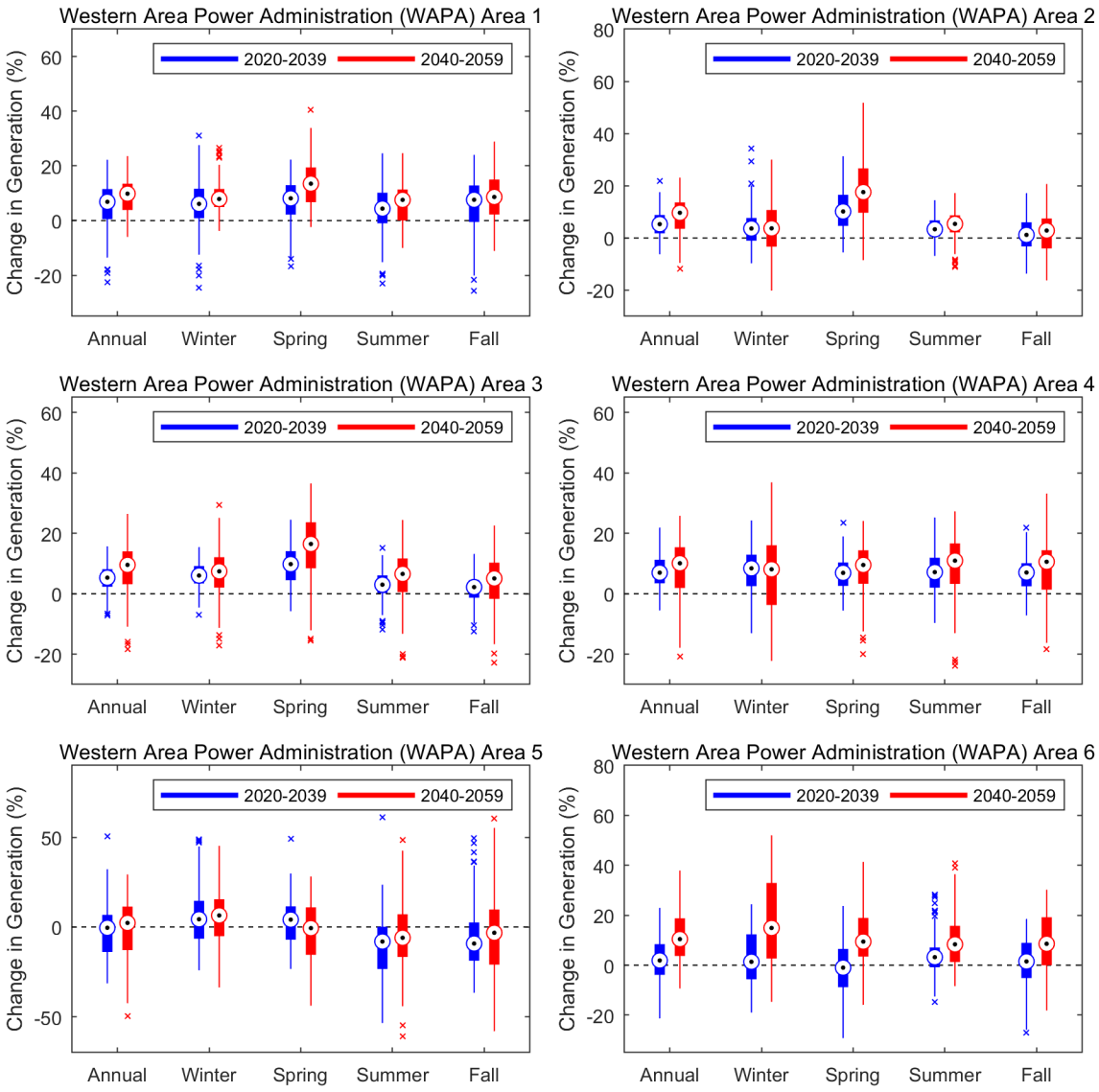


Figure 4.15. Projected change of annual and seasonal total generation in the WAPA region.

WAPA-2 and WAPA-3 are in the snowmelt-dominated upper Colorado River basin, and these areas are projected to have an even larger spring generation percentage increase than WAPA-1. Spring generation is expected to increase by 10% in both WAPA-2 and WAPA-3 in the near-term future period, and by 18% (WAPA-2) and 16% (WAPA-3) in the mid-term future period. Winter generation is also expected to increase by 5% in the near-term and 10% in the mid-term future periods. By contrast, summer and fall generation only increase modestly for the two areas: 1%–3% in the near-term and 3%–7% in the mid-term future periods. The overall annual generation increase and the seasonal generation shift are the result of the winter and spring precipitation increase (Figure 4.5) and an earlier snowmelt process. The projected runoff (Figure 4.7) shows a similar seasonal pattern to generation changes but at larger magnitude, which is caused by runoff being mitigated by the large storage capacity reservoir systems in the area.

WAPA-4 in the lower Colorado River basin is projected to have nearly evenly distributed seasonal generation increase, with spring, summer, and fall generation increasing by 7% in the near-term and by around 11% in the mid-term future periods. Winter generation is projected to increase by 8% in both near-

term and mid-term future periods. The generation response is slightly different from the runoff response, which shows spring runoff in the near-term to decrease (Figure 4.7). The generation response is even more different from the precipitation projections, which indicates both winter and spring precipitation in the near-term to decrease (Figure 4.5). The stable seasonal generation response is likely because of the large storage capacity provided by Hoover Dam and Lake Mead, and because this region receives a substantially managed upstream flow.

Although little change is projected in annual generation in WAPA-5, some seasonal shift is projected. In the near-term future period, winter and spring generation is projected to increase (4%), whereas summer and fall generation is projected to decrease (8%–9%). In the mid-term future period, winter generation is projected to increase by 6%, whereas spring to fall generation is projected to decrease by 1%–6%. The projected seasonal generation change pattern differs substantially from that of runoff change (Figure 4.7). In both future periods, summer and fall runoff are expected to increase, whereas generation is projected to decrease, which is associated with uncertainty in the PMs.

Similar to annual generation, seasonal generation in WAPA-6 in California is expected to remain relatively stable in the near-term future period (ranging from −1% in spring to 3% in summer). In the mid-term future period, generation is expected to increase in all seasons (ranging from 8% in summer to 15% in winter). Compared with seasonal runoff change, the seasonal generation change is more evenly distributed because large storage reservoirs substantially smooth the seasonal variability in natural runoff and projected changes in seasonality.

As stated previously, these increases in annual inflow and generation only represent the projections from six GCMs, so large uncertainty, either due to GCM selection or model limitations, may exist. The increase of inflows will not be steady and will likely accompany more extreme hydrologic events (i.e., both floods and droughts). Parts of the WAPA regions (e.g., Colorado River Basin) are experiencing a more than 20-year drought that is interrupting regional water supply and hydropower generation. The results shown in Figure 4.14 and Figure 4.15 are long-term averages computed over 40 years and do not capture extreme hydrologic events, sub-seasonal variations, or complex hydrologic processes over mountainous terrain which dominate water availability and hydropower generation in the western United States.

The source of variance in the projected annual hydropower generation change between the baseline and future periods is shown in Figure 4.16. The box plots in the top panels show the magnitude of change (future minus baseline) in annual hydropower generation across 96 ensemble members when grouped by the five modeling choices (CM, DS, MetF, HM, and PM). The bottom panel shows the fraction of variance from the ANOVA for each modeling choice.

The diversity in individual factors and their combinations were further evaluated to understand their contributions to the spread of hydropower generation projections. The results suggest that the CM is generally the most dominant contributor to the variance of annual hydropower generation in all WAPA study areas (range from 43% to 52%), except for WAPA-5 and WAPA-6, where the CM is the second most dominant contributor (25% and 39%, respectively). The PM is the most dominant contributor in WAPA-5 (31%) and the second most dominant contributor in WAPA-1 (23%). The HM is the most dominant contributor in WAPA-6 (43%). The combination of the CM and DS is another substantial source of variance in WAPA-1–4, ranging from 11% to 16%. In all WAPA study areas, the DS or MetF alone are not a substantial variance contributor. As shown in Figure 4.16, when grouped by CM, the median change of projected generation is mostly positive across all CMs in all WAPA study areas, although the median and the spread of projected change vary substantially among different CMs. When grouped by PM, WRES projects larger median change of annual hydropower generation. The change projected by WMP shows a wider spread than WRES. Several differences between WRES and WMP may

contribute to the variance, including the differences in reservoir storage, water demand, and flow regulation as represented in the two models as described in Section 2.6.

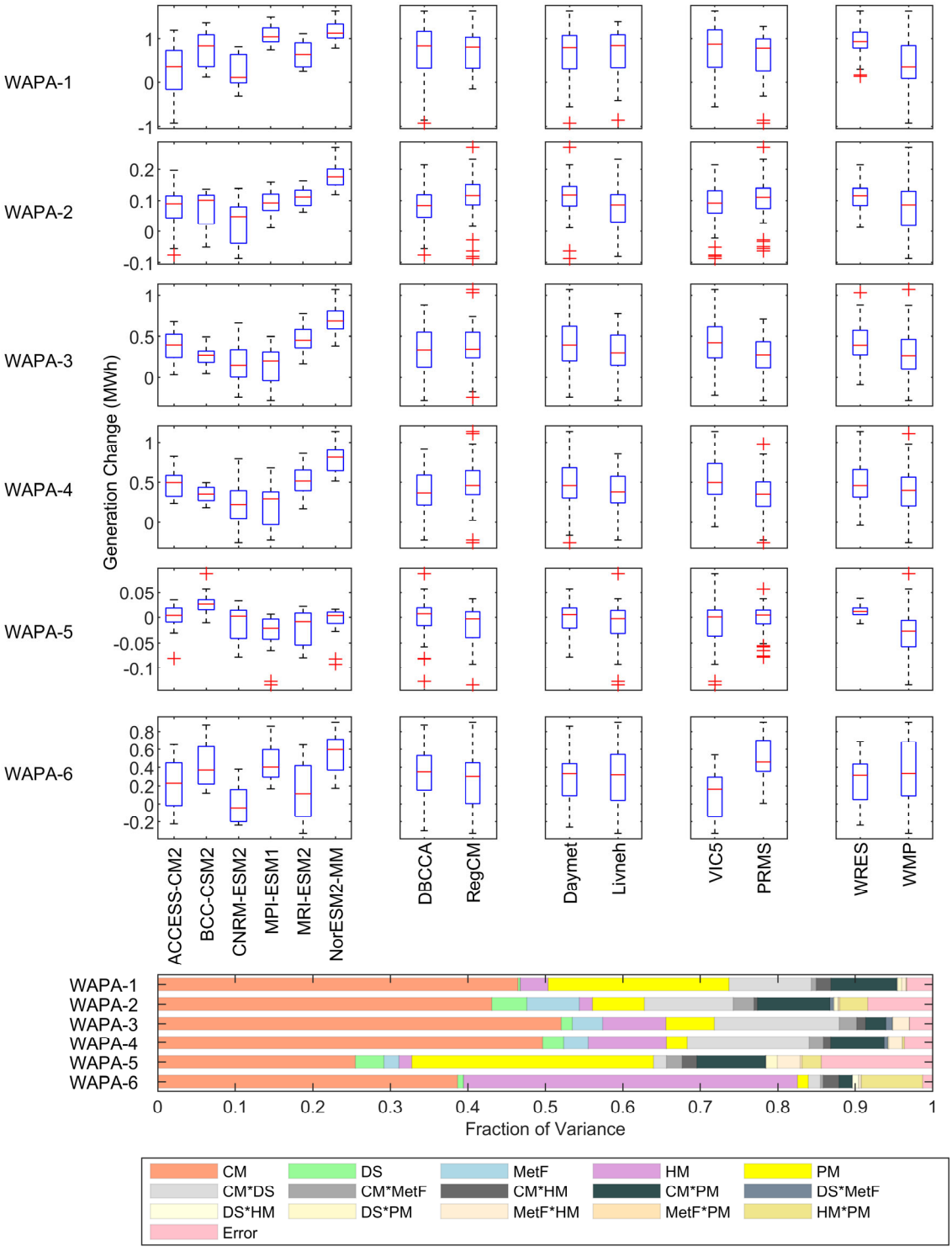


Figure 4.16. ANOVA of annual total generation in the WAPA region.

At the seasonal scale (APPENDIX G), the CM remains the most dominant contributor to variance, except for in the spring, when the PM is the most dominant variance contributor in WAPA-1 (53%), WAPA-2 (32%), and WAPA-5 (43%). On the other hand, the HM is the most dominant variance contributor in WAPA-6 during the spring (36%), summer (37%), and fall (40%).

In summary, for federal hydropower in the WAPA region, the most important climate change impact includes the overall increase in annual and seasonal generation in all study areas except WAPA-5. The seasonal change in generation differs among study areas because of regional climate features. However, the seasonal hydropower generation in general has seasonally uniform response to climate change, indicating that the existing reservoir storage systems and water management practices may help absorb part of the increasing future runoff variability. The sources of uncertainty also differ among study areas, with the CM, HM, and PM each having strong impacts in different areas. The diversity of regional generation response to different modeling choices again shows the benefits of an ensemble-based multimodel approach.

4.3.2 Climate Change Impacts on Regional Electricity Demand

This section summarizes the results from the statistical analysis, exploring the effects of projected temperature trends on the electricity sales of WAPA's preference customers to their end-use (residential, commercial, industrial) customers. As described in Section 4.1.2, WAPA markets the output from 58 hydropower plants through long-term contracts with preference customers, such as municipalities, Native American tribes, cooperatives, and irrigation districts. Depending on the marketing subregion, WAPA guarantees the preference customers a number of hours of firm power or a fraction of the hydroelectricity generated by the federal hydropower plants. Typically, the federal hydropower allocation that preference customers receive from WAPA is a small fraction of the portfolio of generation resources they use to serve their end-use customers. Any load growth projected by WAPA's preference customers, due to climate trends or other drivers, would have to be met through generation sources other than federal hydropower.

Power scheduling in the WAPA regions is primarily driven by non-power purposes provided by the hydropower dams, such as water supply and irrigation. WAPA's preference customers may have some limited flexibility regarding the time of the year in which they receive their energy allocation (within the seasonal or monthly volume ranges specified in the contracts), but only if the requested timing does not conflict with the aforementioned non-power purposes or other federal hydropower operational constraints.

Table 4.3 summarizes the elasticities of preference customer sales with respect to degree days for the sample of 200 WAPA preference customers included in EIA Form 861. The regression analysis was conducted separately for four marketing subregions within the WAPA footprint: WAPA-UGP, WAPA-CRSP, WAPA-SN/RM, and WAPA-DSW. The correspondence among these four marketing subregions and the study areas used elsewhere in this document is presented in APPENDIX H. The complete output from the regression equations is also shown in APPENDIX H.

In the WAPA-UGP subregion, electricity sales by the PMA preference customers to their end-use customers display a positive and statistically significant relationship to both HDDs and CDDs. The elasticity of annual sales to HDDs (i.e., the number and intensity of cold days triggering heating demand) is much larger than the elasticity with respect to CDDs (i.e., the number and intensity of warm days triggering air conditioning use) for utilities receiving federal hydropower allocations from WAPA-UGP. All else equal, a 1% increase in annual HDDs leads to a 0.12%–0.13% increase in sales. In contrast, a 1% increase in annual CDDs leads to a 0.03% increase in sales. Importantly, a 1% increase in HDDs is a much higher absolute number of additional degree days than a 1% increase in CDDs, given the average values for the two variables in this region during the baseline period.

Table 4.3. Estimated responsiveness of electricity sales by WAPA preference customers to degree days

Marketing region	Meteorological observation	Dependent variable	Elasticity with respect to HDDs	Elasticity with respect to CDDs
WAPA-UGP	Daymet	Electricity sales	**0.1154** (0.015)	**0.0271** (0.0083)
WAPA-UGP	Livneh	Electricity sales	**0.1315** (0.0188)	**0.0264** (0.0082)
WAPA-CRSP	Daymet	Electricity sales	−0.0073 (0.0449)	−0.0096 (0.0103)
WAPA-CRSP	Livneh	Electricity sales	0.0128 (0.0444)	−0.0018 (0.0069)
WAPA-DSW	Daymet	Electricity sales	0.0029 (0.029)	**0.1532** (0.0994)
WAPA-DSW	Livneh	Electricity sales	0.0444 (0.0309)	0.0798 (0.0721)
WAPA-SN/RM	Daymet	Electricity sales	0.055 (0.0475)	0.0198 (0.0482)
WAPA-SN/RM	Livneh	Electricity sales	0.0069 (0.0391)	0.0154 (0.0424)

Note: Numbers in parentheses are standard errors. Bolded values indicate that the estimated elasticity is statistically significant at the 10% level.

The only other statistically significant elasticity result is the one with respect to the CDDs in the WAPA-DSW subregion using the Daymet meteorological observations (0.1532). This result is partially explained by this region being the southernmost WAPA subregion, which already had the highest number of CDDs in the baseline period, such that a 1% increase in CDDs is much larger in absolute value than in the rest of subregions.

The lack of statistical significance in the rest of estimated elasticities means that the available data do not show conclusive evidence of degree days driving preference customer sales to their end-use customers. The lack of statistical significance for some of the elasticity estimates may arise due to insufficient data. For instance, the regressions for WAPA-DSW and WAPA-SN/RM have the fewest observations because the electricity sales data were only available in EIA Form 861 for 18 out of a total number of 52 preference customers in the WAPA-DSW subregion, and 18 out of 121 preference customers in the WAPA-SN/RM subregion.[27] They are also the two subregions with lowest percentages of residential sales (33% for WAPA-DSW and 36% for WAPA-SN/RM) and have many irrigation districts among their preference customers. The relationship between temperature and electricity sales for irrigation activities has been less researched than for residential sales, and it might not be well captured by the linear demand equations estimated in this analysis.

The estimated elasticities with respect to degree days were combined with the projected changed in degree days derived from climate simulations to produce an estimate of the temperature-driven changes in demand for WAPA preference customers. Figure 4.17 and Figure 4.18 show the projected degree days from 24 simulations (the six CMIP-6 models described in Section 2.3 combined with two MetFs and two DSs), as well as the ensemble median projection.

[27] A significant fraction of WAPA preference customers in these subregions are federal agencies, state agencies, Native American tribes, and irrigation districts. They typically do not qualify as electric utilities and do not have to file EIA Form 861.

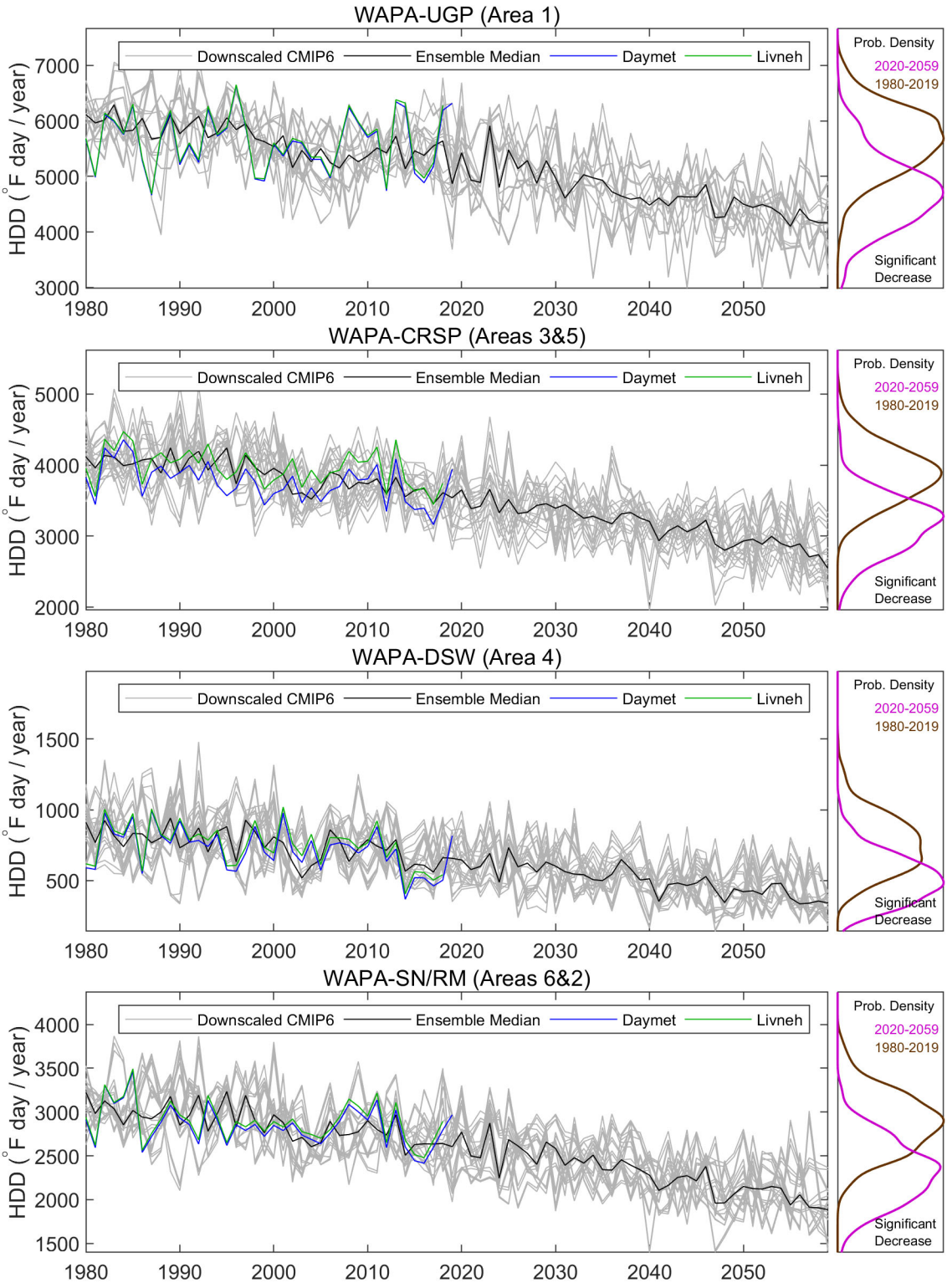


Figure 4.17. Projected annual HDDs in the WAPA region.

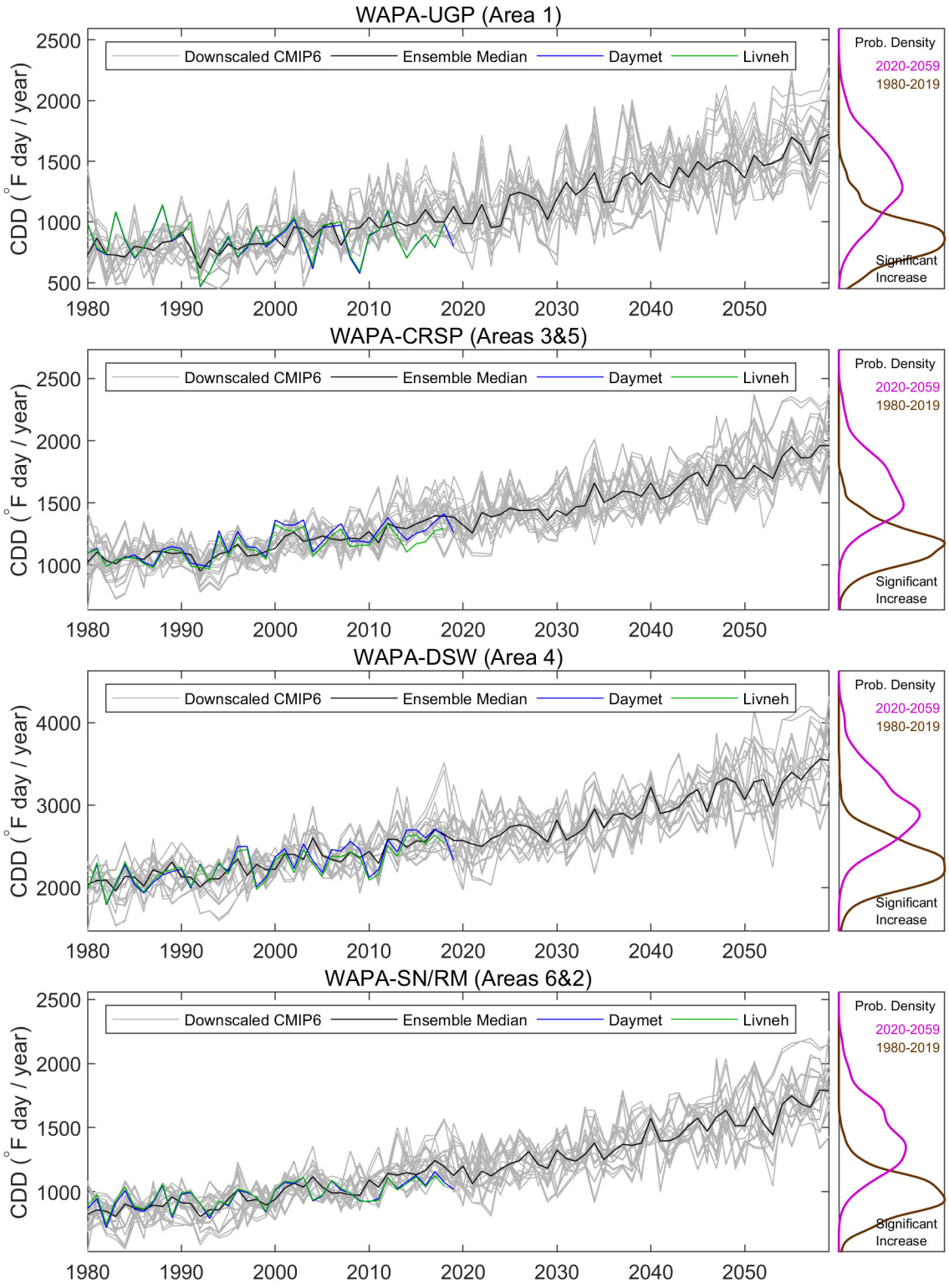


Figure 4.18. Projected annual CDDs in the WAPA region.

WAPA has the largest footprint out of the four PMAs. The marketing subregions displayed large differences in the average number of HDDs and CDDs experienced during the baseline period (1980–2019). The southernmost subregion (WAPA-DSW) experienced the lowest average annual HDDs (752)

and highest average annual CDDs (2,300). In contrast, the northernmost subregion (WAPA-UGP) experienced the highest average annual HDDs (5,654) and lowest average annual CDDs (871). Figure 4.19 shows that annual HDDs are projected to decrease and annual CDDs are projected to increase in every region. In the future period (2020–2059), the ensemble median trajectory shows average annual CDDs in the 1,300–1,600 range in every region except for WAPA-DSW, where they will be closer to 3,000. Average annual HDDs in that same period are projected to range from 503 in WAPA-DSW to 4,720 in WAPA-UGP. The projected changes in sales driven by these degree day trends are presented in Figure 4.19.

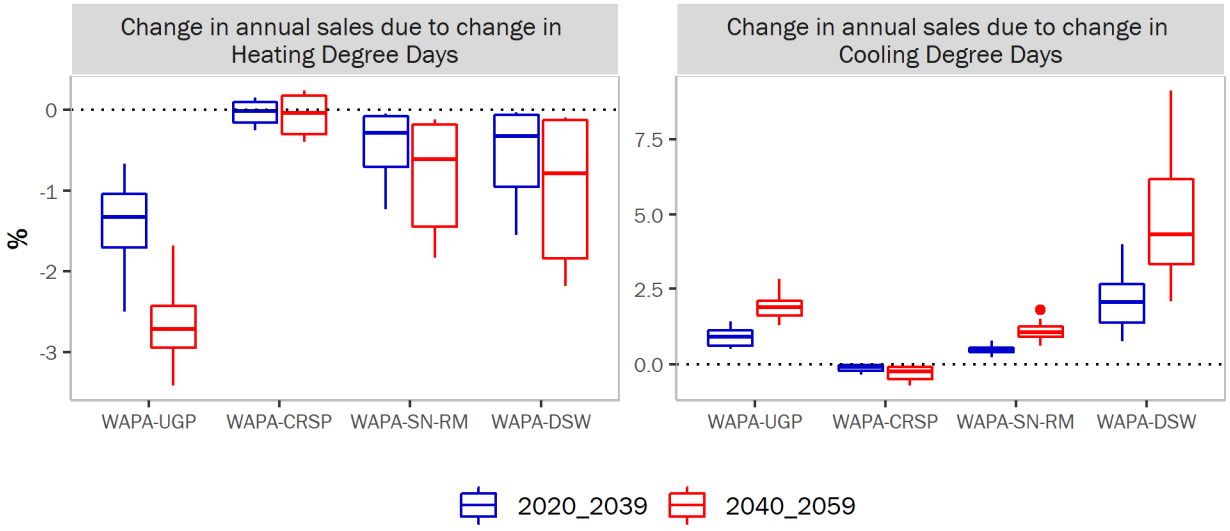


Figure 4.19. Estimated changes in WAPA preference customer electricity sales relative to the baseline period due to projected changes in degree days.

Changes in sales driven by HDDs and CDDs will take place during the winter and summer, respectively. Electricity sales by WAPA preference customers, except in some of the WAPA-CRSP projections, are expected to decrease in the winter months and increase in the summer months because of the trends in degree days summarized in Figure 4.17 and Figure 4.18.

The sign and magnitude of the net effect on annual sales varies by subregion. For example, the WAPA-UGP preference customer annual sales are projected to decrease overall because the drop in winter sales caused by less HDDs is larger than the increase in summer sales caused by higher CDDs. The median projected decrease in annual sales for WAPA-UGP customers relative to the 1980–2019 baseline period is 0.4% in 2020–2039 and 0.7% in 2040–2059. On the other hand, annual sales are projected to increase for preference customers receiving power from WAPA-DSW. For customers in that region, the dominant effect is an increase in summer sales driven by higher annual CDDs. The projected net increase in sales is 1.7% in 2020–2039 and 3.6% in 2040–2059 (based on the ensemble median degree day projections).

In the other two subregions (WAPA-CRSP and WAPA-SN/RM), the estimated temperature-driven changes in sales are smaller and must be interpreted with caution because the elasticities were not precisely estimated. Data from more preference customers in those subregions or at higher frequency could be helpful in improving the precision of these.

In summary, heating-related demands are expected to decrease, and electricity demands for air conditioning are expected to increase. As a result, WAPA's preference customers may request scheduling

a larger share of their allocation during the summer. However, WAPA can only accommodate these requests to the extent that they don't conflict with other operational constraints.

The results presented in Section 4.3.1 show projected increases in generation from the federal fleet marketed by WAPA at the annual time scale in all study areas except WAPA-5 (the lower Colorado facilities marketed as part of the CRSP subregion). On the other hand, WAPA-DSW is the only subregion where a net annual increase in load due to temperature is projected. Taken together, these results suggest that WAPA will on average be able to fulfill the capacity and energy allocations in its current long-term contracts.

(This page intentionally left blank)

5. THE SWPA REGION

Following the methodology described in Section 2, the results for federal hydropower plants marketed by the SWPA are summarized in this section. Descriptions of the SWPA study areas, federal hydropower systems, and power marketing practices are described in Section 5.1. The projected future atmospheric and hydrologic conditions in the SWPA region are discussed in Section 5.2. The potential climate change impacts on federal hydropower generation and SWPA regional energy demands are discussed in Section 5.3.

5.1 SWPA REGIONAL CHARACTERISTICS

5.1.1 Federal Hydropower in the SWPA Region

The SWPA region includes rivers that run through the Ozark Plateau, the southern Great Plains, and the Texas coastal plains. There are 24 federal hydropower plants in this region, all of which are owned and operated by USACE. These USACE federal hydropower plants have a total of 2,256 MW capacity and 5.6 TWh/year average annual generation (Table 5.1).

Table 5.1. Summary of federal hydropower plants in the SWPA region

Area	Area name	Number of USACE plants	Capacity[a] (MW)	1980–2019 generation[b] (GWh/year)
SWPA-1	Upper White, Osage, and Salt	8	1,100	2,255
SWPA-2	Arkansas	9[c]	785	2,653
SWPA-3	Ouachita, Red, and Brazos	5	312	590
SWPA-4	Neches	2	60	145
Total		24	2,256	5,643

[a] EIA 2019 total nameplate capacity, including both conventional hydro and reversible (pumpback capability).
[b] EIA and SWPA average annual generation from 1980 to 2019, conventional hydro only.
[c] Broken Bow in the Red River Basin is included in SWPA-2 because of an interconnected system reason.

The federal hydropower plants in the SWPA region range in size from Bull Shoals (340 MW) to the Robert D. Willis project (8 MW). Although three USACE hydropower plants (Clarence Cannon, Harry S. Truman, and DeGray) have reversible turbines that may allow for pumped storage operation, they are mostly operated as conventional hydropower facilities. The average age of these USACE power plants is 56 years, and they experience the same challenges of aging infrastructure as federal projects elsewhere. A complete list of the projects in this region is provided in APPENDIX B.

Although the rivers in the SWPA region do not have migratory fisheries management constraints like those in the Columbia River or California, the hydropower projects are impacted by endangered species protection issues for animals such as the interior least tern, as well as issues pertaining to recreational fisheries management. Additionally, growing demands for municipal and industrial water supply are some of the most crucial water resource issues in the SWPA region. Competing interests with water-based recreation, water quality, and other environmental issues are also concerns for the SWPA region.

Considering power system and watershed boundaries, the SWPA region is subdivided into four study areas (Figure 5.1). Overall, there is generally less variety in the physical and climatological differences among the four areas of analysis in the SWPA region than for the BPA and WAPA regions (described in the previous two sections).

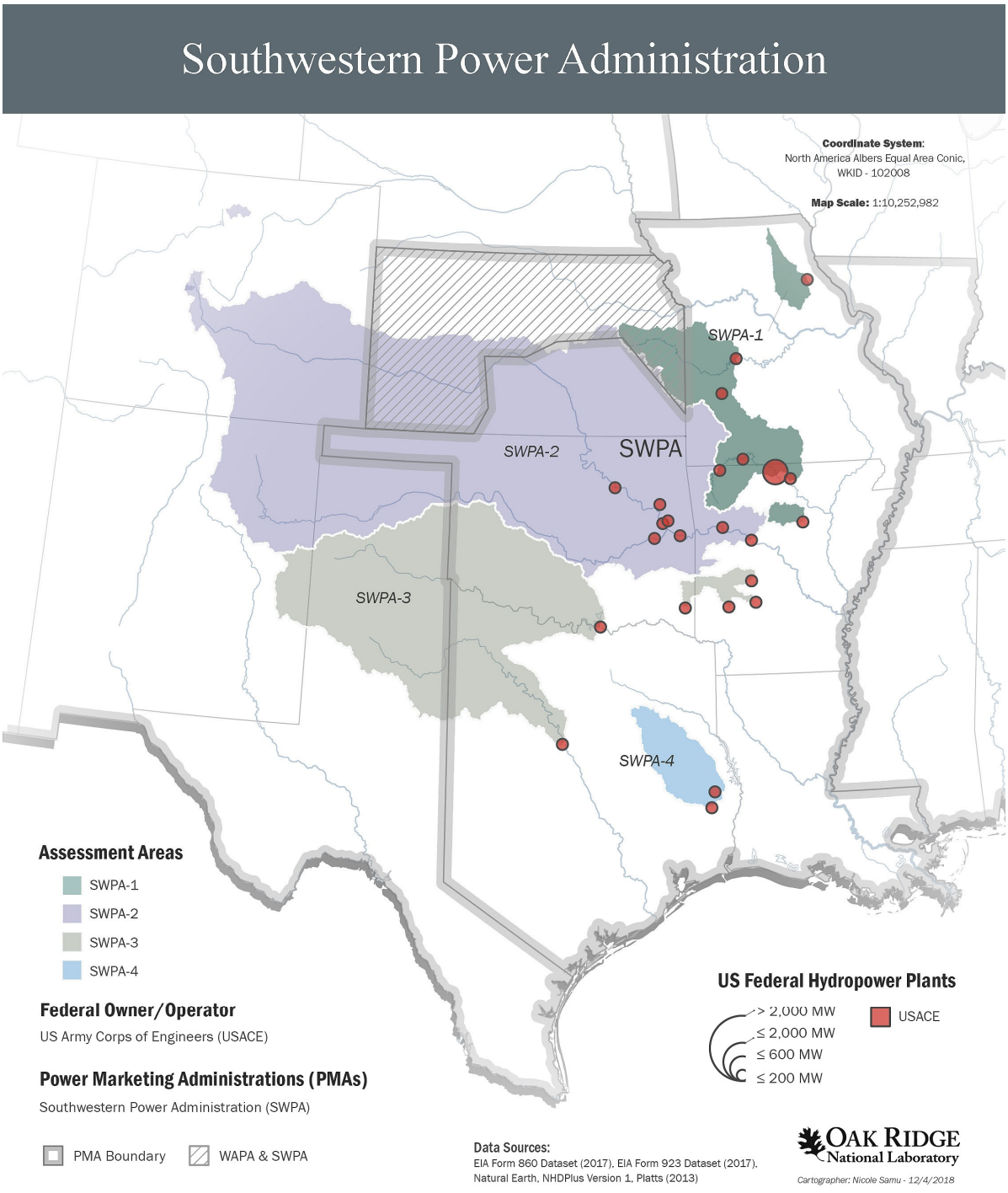


Figure 5.1. Map of the federal hydropower plants and four study areas in the SWPA region.

- **SWPA-1 Upper White, Osage, and Salt Rivers**: SWPA-1 is located on the northern portion of the Ozark Plateau. This area includes the upper White River and Osage River basins, as well as the Salt River drainage in northeastern Missouri where the USACE Clarence Cannon project is located. The total drainage area of watersheds contributing water to projects in this area is approximately 22,000 mi^2. Elevations range from 300 to 2,400 ft, with a median elevation of 981 ft. The entire SWPA region is substantially lower in elevation than BPA or WAPA, and the topography strongly

influences surface water hydrology. Land cover in SWPA-1 is a mix of cropland (37%), cropland–natural vegetation mosaic (34%), deciduous broadleaf forest (12%), and grassland (9%).

SWPA-2 Arkansas River: The Arkansas River basin is defined by watershed upstream of the USACE Dardanelle project. This river basin is very large and extends upstream to the Continental Divide in Colorado. The total drainage area is about 154,000 mi^2. Elevations vary greatly, from over 14,000 ft in the Colorado Rocky Mountains to around 300 ft in the vicinity of the hydropower projects clustered in the eastern end of the river basin. The median elevation is 2,392 ft. Although this area includes high-elevation parts of the Rocky Mountains, most of the runoff to the projects originates locally in the eastern part of the area. Land cover is mostly grassland (68%), with smaller amounts of cropland (17%), cropland–natural mosaic (5%) and woody savanna (4%). Broken Bow in the Red River Basin is also included in SWPA-2 because of an interconnected system reason.

- **SWPA-3 Ouachita, Red, and Brazos Rivers**: SWPA-3 covers the Red River and Brazos River basins in the arid plains of southwestern Oklahoma and the Texas Panhandle, and the Ouachita River basin in the southern portions of the Ouachita Mountains/Ozark Plateau, mostly in Arkansas. The total watershed area is 68,744 mi^2. Elevations range from almost 5,000 to 300 ft, with a median elevation of 1,720 ft. Land cover is primarily grasslands (82%).

- **SWPA-4 Neches River**: SWPA-4 is in the upper half of the Neches River basin and is the lowest and one of the smallest of all of the areas in the 9505 assessment. Its watershed area is 7,571 mi^2, and elevations range between 774 and 72 ft, with a median elevation of 341 ft. Land cover is a diverse mix of cropland–natural mosaic (37%), mixed forest (25%), woody savanna (22%), cropland (3%) and evergreen broadleaf forest (3%).

5.1.2 Power Marketing by SWPA

SWPA was established in 1943 under the statutory authority from Section 5 of the Flood Control Act of 1944. A general description of a PMA and its marketing practices is presented in Section 1.2.2. SWPA markets power from 24 USACE-owned multipurpose dams, operates and maintains 1,381 mi of transmission lines, and manages one balancing authority. SWPA sells power to more than 100 wholesale customers in Arkansas, Kansas, Louisiana, Missouri, Oklahoma, and Texas.

The product of the projects marketed by SWPA, almost entirely dependent on rainfall and with limited storage capacity in their reservoirs, is primarily peaking power because the available water typically only allows the hydraulic turbines to operate a few hours per day. SWPA sells the output from its Interconnected System of 17 + 2 projects as peaking capacity and energy guaranteeing 1,200 h of energy per kilowatt of capacity contracted per year (3.3 h a day, on average). The financially integrated system includes the 17 + 2 Interconnected System projects, as well as three electrically isolated projects, a total of 22 of the 24 projects for which SWPA markets power. The other two (Sam Rayburn Dam and Robert D. Willis Hydropower Project) are hydraulically and electrically isolated and marketed separately. All the output from the Sam Rayburn Dam is sold to the Sam Rayburn Dam Electric Cooperative, and all the output from the Robert D. Willis Hydropower Project is sold to the Sam Rayburn Municipal Power Agency.[28]

Table 5.2 shows the FY 2021 rates from each of the three marketing systems. For the Rayburn and Willis projects, customers pay a flat monthly rate in exchange for all the output from the projects. The rates for

[28] For the Robert D. Willis Hydropower Project, USACE built the dam in 1951, but a hydropower facility was not added until the 1980s, and its construction was financed by Sam Rayburn Municipal Power Agency. In exchange for funding the hydropower construction costs, Sam Rayburn Municipal Power Agency receives all the capacity and energy from the project for a period of 50 years.

these two projects were last updated in 2015. For the Integrated System, the same rates have applied since October 2013. They include a capacity charge per kilowatt of capacity contracted and, for the Interconnected System, a purchase power adder per kilowatt-hour of peaking energy. The base purchased power adder ($0.0059) is calculated as the ratio of the estimated total dollar cost ($13,273,800/year) of SWPA's purchased power over projected peaking energy sales. Additionally, within each year, SWPA can adjust the purchased power adder up to two times by a total amount of plus or minus $0.0059/kWh depending on the difference between the estimated and actual expenditures in purchased power. The purchased power adder does not apply to sales of supplemental peaking energy or to non–Interconnected System peaking energy. Given this rate structure, the average price per kilowatt-hour paid by SWPA's customers can vary significantly from drought years when little to no supplemental energy is provided and years when abundant supplemental energy is available.

Table 5.2. SWPA rate schedule summary (FY 2021)

System	Capacity charge ($/kW-month)	Peaking energy or supplemental peaking energy rate
Integrated System	4.50 [a]	0.0094 mills/kWh 0.0059 mills/kWh (purchased power adder)
Hydro Power and Energy from Sam Rayburn Dam and Reservoir (sold to Sam Rayburn Dam Electric Cooperative)		$380,316/month
Hydro Power and Energy from Robert D. Willis Hydropower Project (sold to Sam Rayburn Municipal Power Agency)		$106,903/month

[a] Includes two ancillary service charges (scheduling and reactive supply/voltage control).
[b] For Interconnected System only.

In its role as balancing authority, SWPA coordinates with SPP in transmission planning. However, unlike BPA or WAPA, SWPA has not announced plans to participate as a marketer in an EIM or an ISO/RTO. Nevertheless, most SWPA preference customers participate in the SPP, MISO, or ERCOT markets (Fernandez, 2020).

5.2 FUTURE CLIMATE IN THE SWPA REGION

5.2.1 Regional Climate Projections

The annual temperatures in the baseline (1980–2019) and future (2020–2059) periods are presented in Figure 5.2. The gray lines show the annual mean temperature from 24 sets of downscaled climate projections, the black line represents the multimodel median, and the blue and green lines show the Daymet and Livneh historical observations, respectively. These plots show interannual variability and annual trends for the projections as well as the observations. The climate models are not expected to replicate the exact timing of historical interannual and decadal variability. Therefore, they do not completely follow the historical interannual values. The corresponding probability distributions for the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between the baseline and future periods is statistically significant.

In all four assessment areas (SWPA-1–4), the ensemble model median captures the increasing trend in the annual mean temperature present in the observations during the historical period. However, there is a large variability across the ensemble members, which reflects the importance of using a multimodel ensemble approach. In the future period, projected annual mean temperature continues to increase across

all the four areas. A comparison of the probability distribution of the multimodel mean annual temperature during the historical and future periods also shows a statistically significant shift toward a warmer climate in all SWPA areas (Figure 5.2).

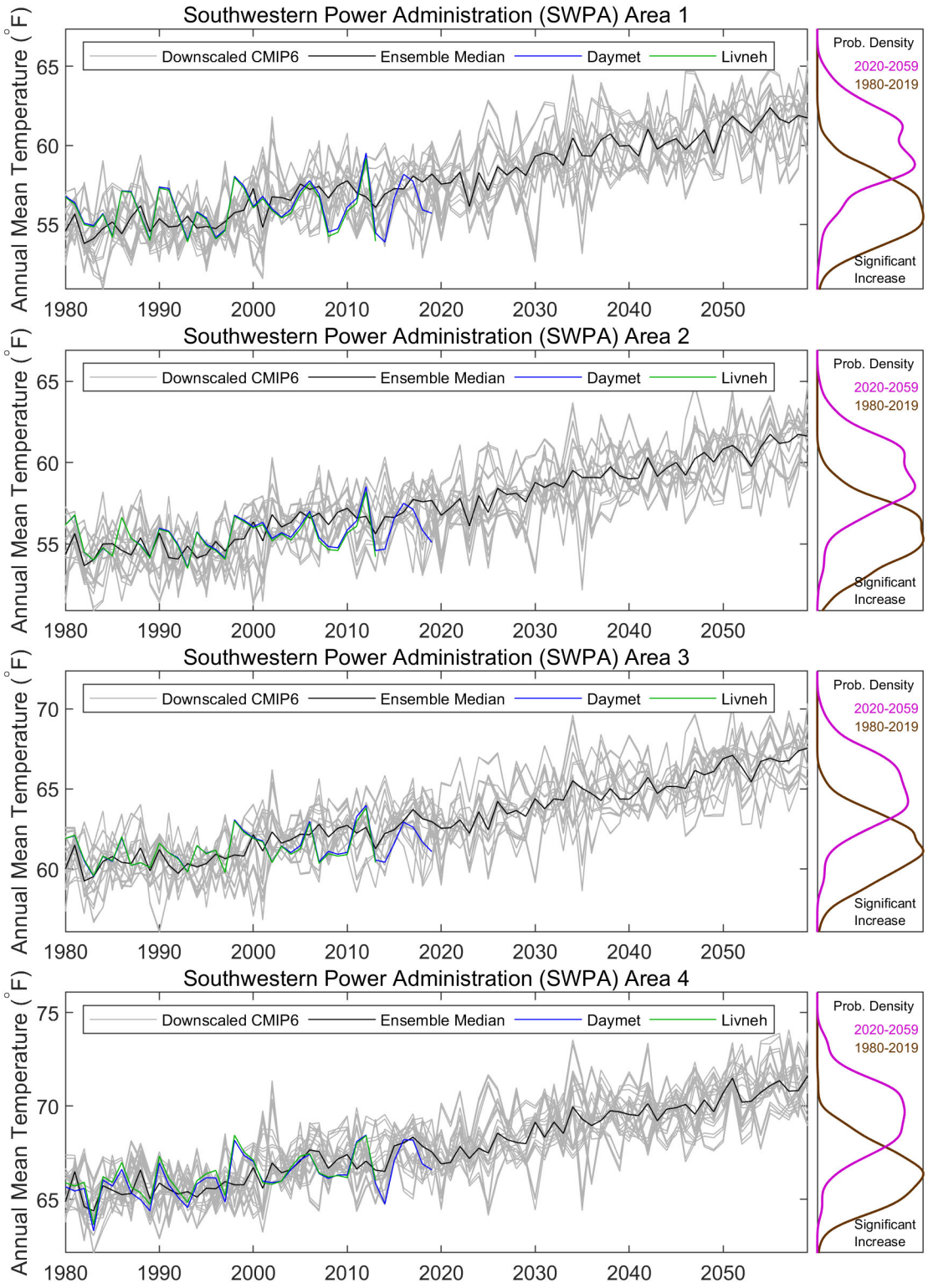


Figure 5.2. Projected annual mean temperature in the SWPA region.

The annual and seasonal changes in temperature are further summarized in Figure 5.3. The change is presented as the difference in the temperature during the future periods (2020–2039 near-term and 2040–2059 mid-term) compared with the baseline period (1980–2019) in degrees Fahrenheit. Each box plot shows the spread across 24 sets of downscaled climate projections; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length denoted by x are considered statistical outliers (but should not be excluded in the context of climate change).

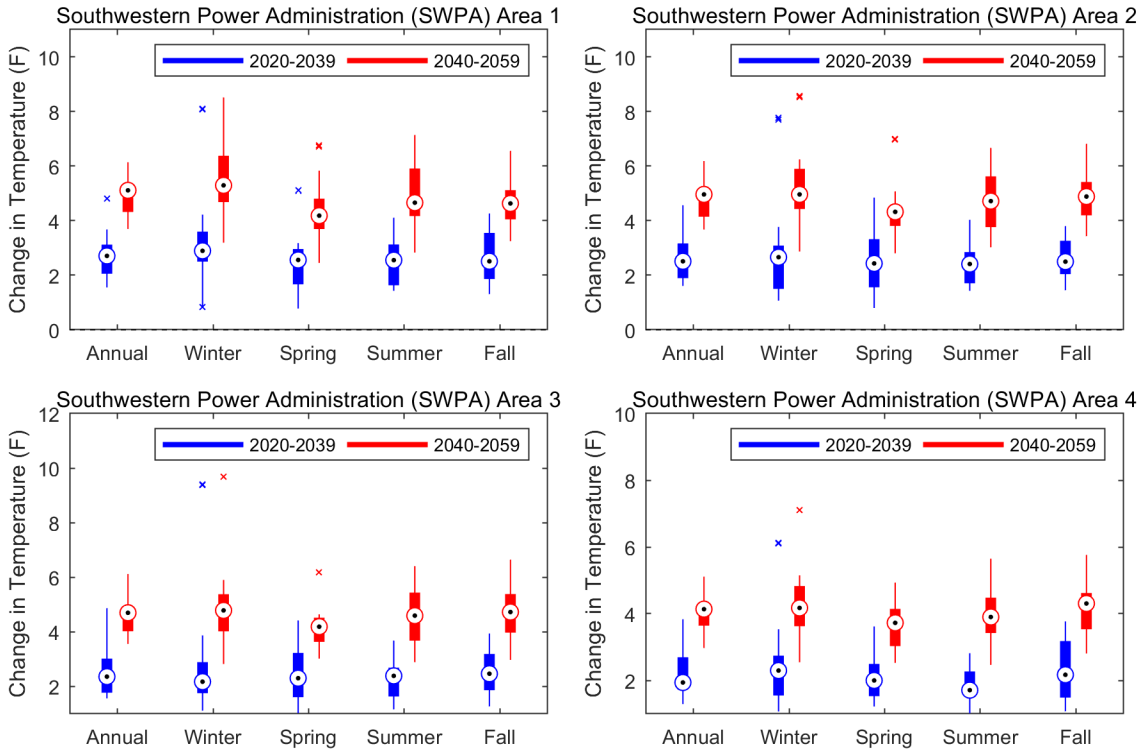


Figure 5.3. Projected change of annual and seasonal mean temperature in the SWPA region.

Annual and seasonal temperature is projected to increase across all the ensemble members during both the near-term and mid-term future periods. A relatively large increase is projected during the mid-term future period with respect to the baseline period. The multimodel median annual temperature in the SWPA region is projected to increase by approximately 2°F–3°F in the near-term and 4°F–5°F in the mid-term future periods. An increase of 4°F or greater is projected by the mid-term across all the seasons over the four SWPA areas.

The annual precipitation (Figure 5.4) also show high interannual and intermodel. A comparison of the probability distributions between the baseline and future periods indicates a statistically significant increase in precipitation across all four SWPA areas.

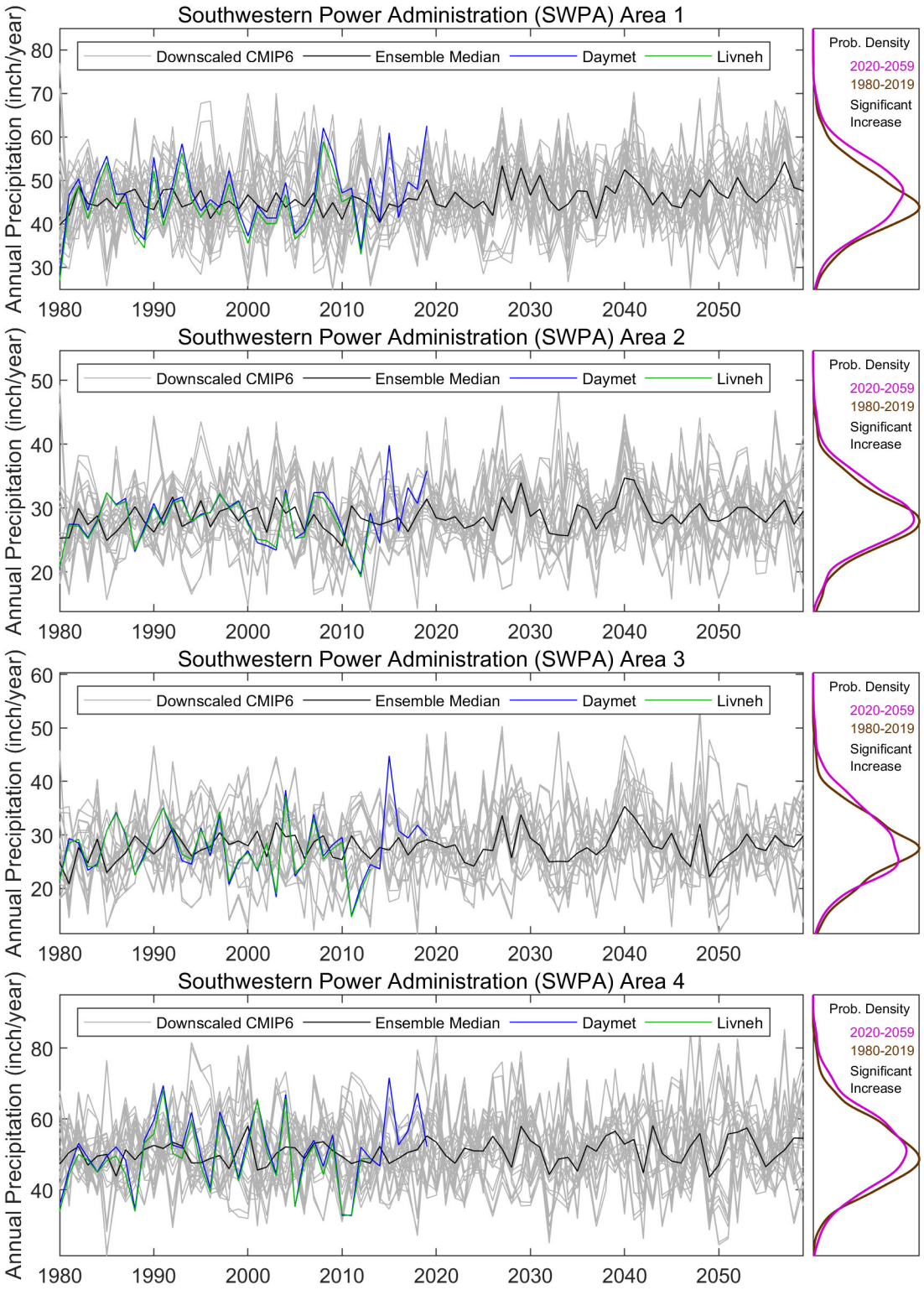


Figure 5.4. Projected annual total precipitation in the SWPA region.

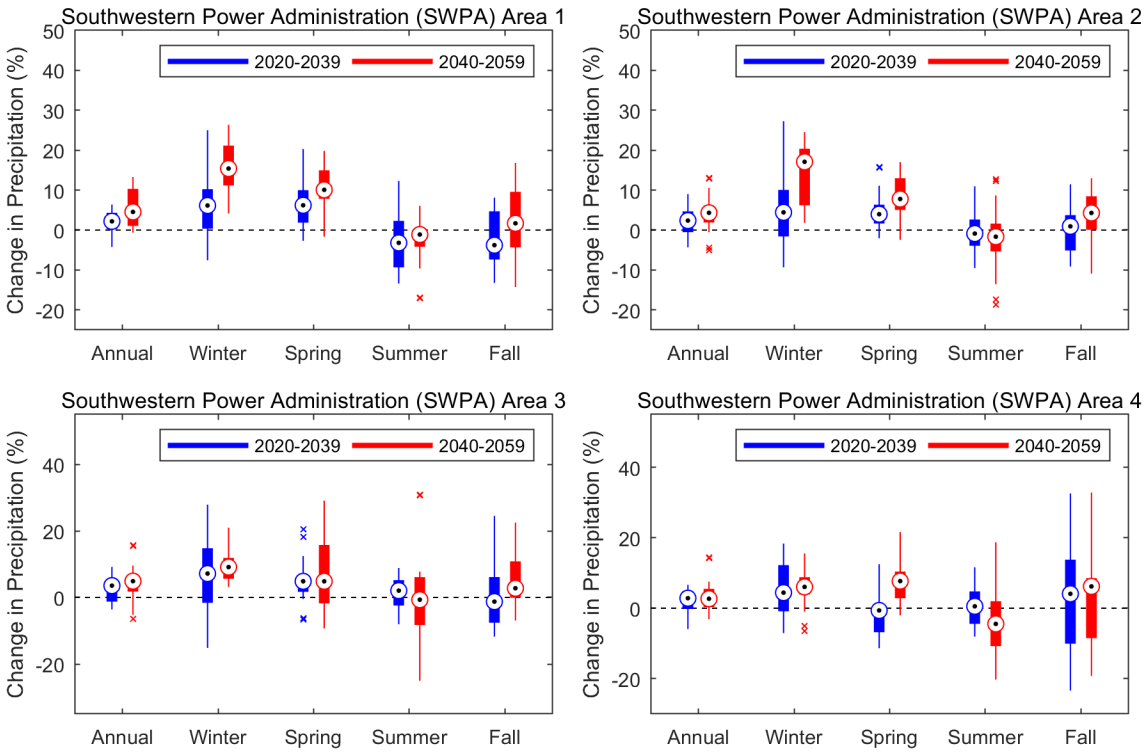


Figure 5.5. Projected change of annual and seasonal total precipitation in the SWPA region.

The annual and seasonal changes in precipitation over the SWPA region are summarized in Figure 5.5. In both the near-term (2020–2039) and mid-term (2040–2059) future periods, the ensemble median annual precipitation is projected to show a small increase (<5%) over the four areas. On a seasonal scale, the largest increase in median precipitation is projected during the winter, followed by the spring. A small to no median change is projected during the summer and fall across most of the regions.

5.2.2 Regional Hydrologic Projections

The runoff projections for each of the SWPA regions are illustrated in Figure 5.6 and Figure 5.7. The annual total runoff in baseline (1980–2019) and future (2020–2059) periods is presented in Figure 5.6. The gray lines show annual runoff from 48 sets of hydrologic projections, the black line represents the multimodel median, and the blue and green lines show the control simulations driven by Daymet and Livneh observations, respectively. The corresponding probability distributions are compared in the right panel. The annual and seasonal changes in runoff are further summarized in Figure 5.7. Change is defined as the percentage difference of future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across 48 ensemble members; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length are considered as statistical outliers (but should not be excluded in the context of climate change).

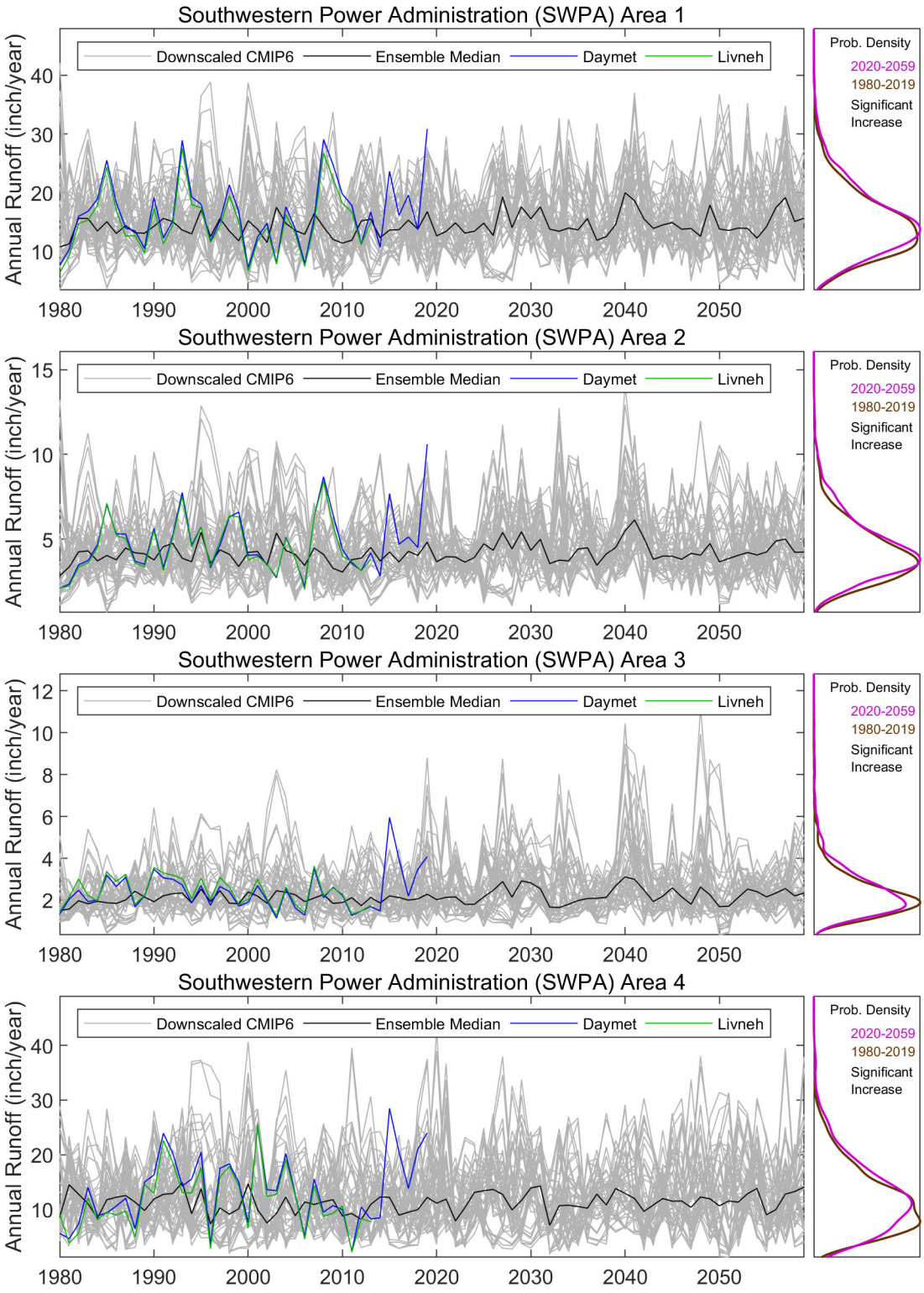


Figure 5.6. Projected annual total runoff in the SWPA region.

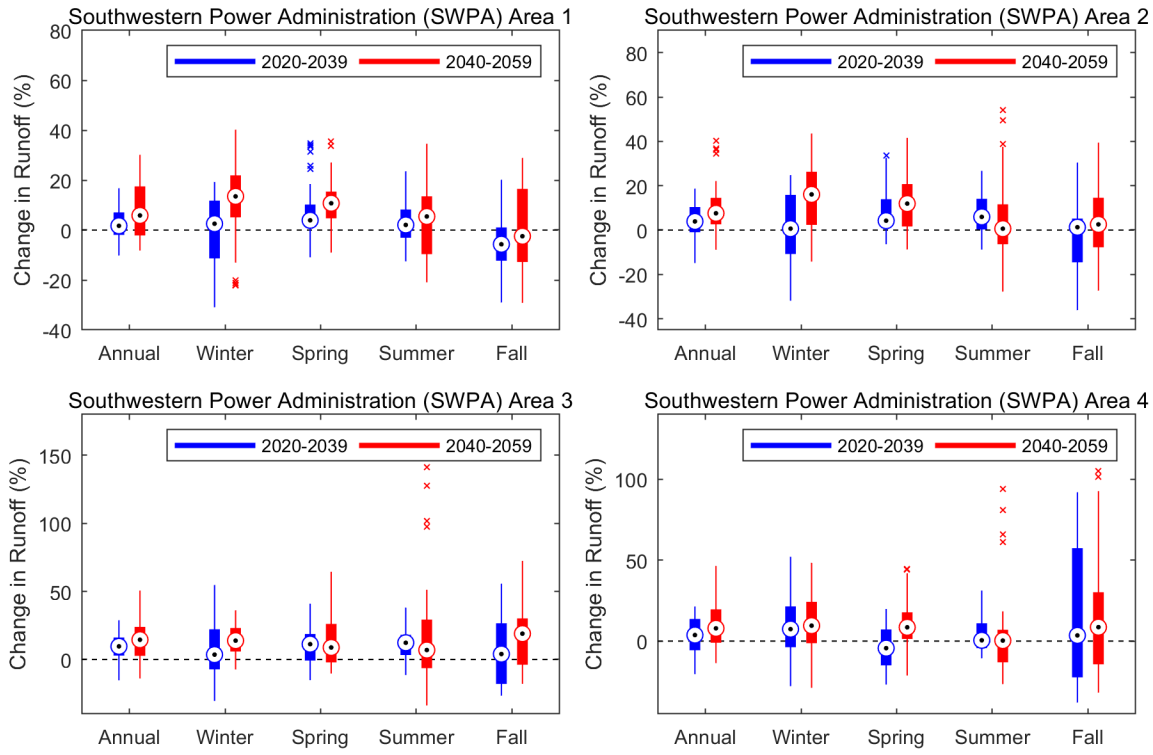


Figure 5.7. Projected change of annual and seasonal total runoff in the SWPA region.

The shifts of the probability distributions (Figure 5.6) suggest that the annual runoff is projected to increase in the future period and the distributions are statistically different. Additionally, the probability distributions in all SWPA areas are projected to be wider, indicating that the range of interannual variability is expected to increase. Such a change is more prominent and easily noticeable in SWPA-3. The maximum increase in median annual runoff is projected in the range of 2.9% to 8.9%, with SWPA-3 showing the greatest increase and SWPA-1 showing the least increase. More specifically, the median annual runoff is projected to increase for all study areas in the range of 1.7% to 9.6% in the near-term and 5.8% to 14.5% in the mid-term future periods. The range of annual runoff change is higher in the mid-term future period, indicating greater uncertainties in the runoff projections.

At the seasonal scale, the ensemble median winter runoff is projected to increase across all SWPA areas in the range of 0.6% to 7.3% in the near-term future period, and the change is expected to further increase in the range of 9.5% to 16.1% in the mid-term future period. The spring runoff is projected to increase in SWPA-1, SWPA-2, and SWPA-3 in the range of 3.9% to 11.3% in the near-term and up to 8.4% to 11.9% in the mid-term future periods. The spring runoff in SWPA-4 shows a decrease of 4.3% in the near-term and an increase of 8.5% in the mid-term future periods. Summer runoff is expected to increase in all SWPA areas by 0.5%–12.3% in the near-term and 0.6%–6.9% in the mid-term future periods. The fall runoff is also expected to increase across all SWPA areas, except for SWPA-1, in the range of 1.3% to 4.0% in the near-term and 2.5% to 18.9% in the mid-term future periods. The increase is also associated with a large range as shown by Figure 5.7, especially for SWPA-4, which likely stems from a large variability in the precipitation; however, the range of change in runoff is higher than that of the range of change in precipitation. The fall runoff in SWPA-1 is projected to decrease by 5.7% in the near-term and 2.5% in the mid-term future periods. The seasonal projections are associated with large spread across all ensemble members in both near-term and mid-term future periods, indicating high uncertainties in ensemble members.

An additional analysis of extreme runoff is shown in Figure 5.8, which presents future changes in ensemble median high runoff (i.e., 95th percentile of daily runoff) and median low runoff (i.e., 5th percentile of 7-day average runoff) in both future projection periods. High runoff is generally projected to increase in both future periods across all SWPA areas, along with large spatial heterogeneity. The low runoff is projected to decrease in general across all SWPA areas in both future periods. These results suggest that although there may be a slight increase in extreme flood events, the region may observe greater challenges caused by reduced low flow conditions under future climate.

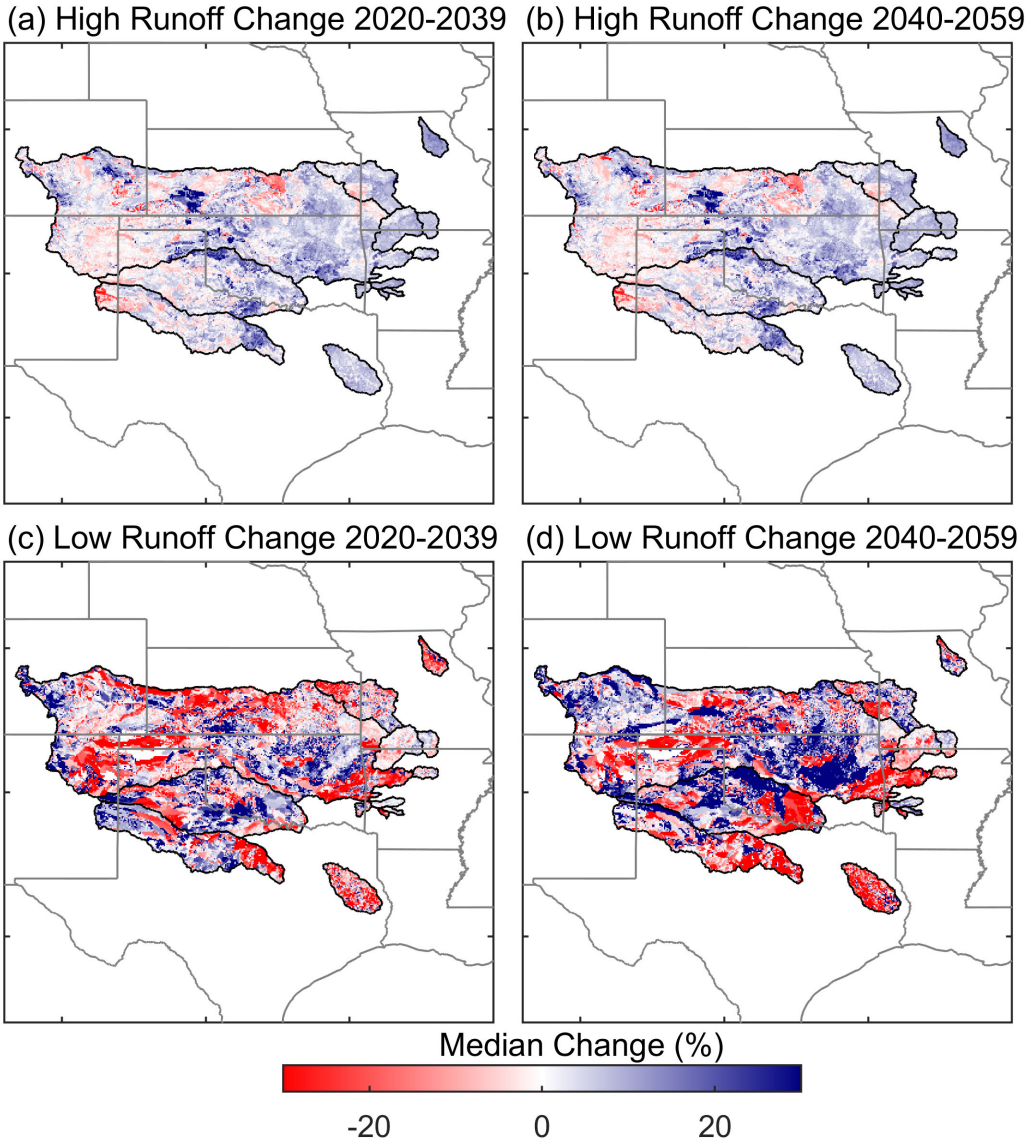


Figure 5.8. Projected change of high and low runoff in the SWPA region.

To evaluate the impacts of modeling choices (six CMs, two DSs, two MetFs, and two HMs) on the projection of future runoff in each SWPA area, an ANOVA was conducted and is illustrated in Figure 5.9. The box plots show the magnitude of change (future minus baseline) in median annual runoff across 48 ensemble members, grouped by each of the four modeling choices. The box plots provide a quantitative way to evaluate the range of possible outcomes in runoff response based on different modeling choices. ANOVA results indicate that the CM is generally the most dominant contributor to the total variance of

future runoff projections. The contribution of the CM toward total variance is greater than 50% in SWPA-1, SWPA-2, and SWPA-3. In SWPA-4, the CM and DS both have a stronger influence. At the seasonal scale (APPENDIX E), the ANOVA indicates that factors other than the CM may have stronger impacts than at the annual scale. For instance, although the CM is the most dominant contributor to the winter runoff variance, the contribution from the CM and DS interaction can also exhibit a stronger influence during the summer and fall.

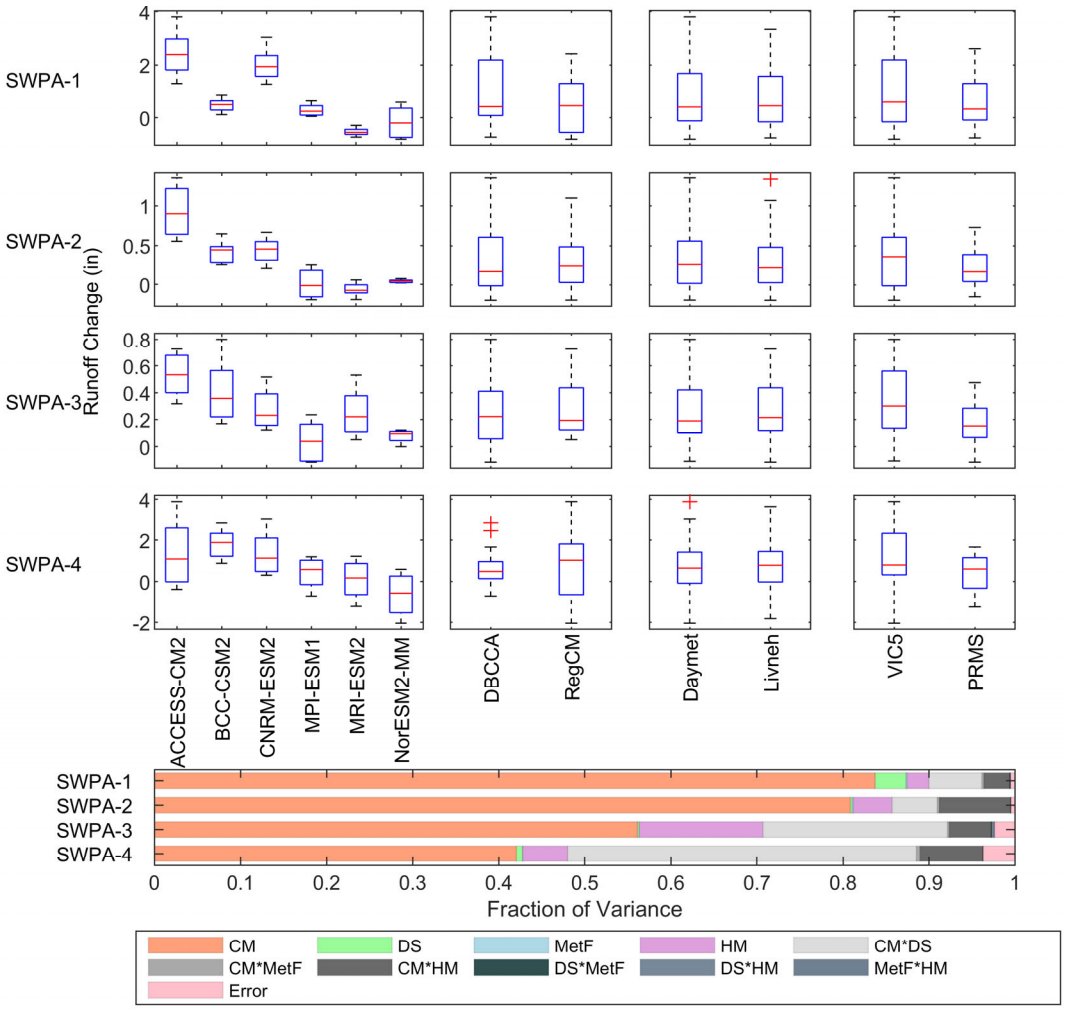


Figure 5.9. ANOVA of annual total runoff in the SWPA region.

Based on LEM as described in Section 2.7, the projected reservoir evaporation loss from federal hydropower reservoirs in the SWPA region is presented in Figure 5.10. The blue and green lines show annual reservoir evaporation loss (in units of million acre-feet) from 24 projections that were either bias-corrected or trained by Daymet and 24 projections that were either bias-corrected or trained by Livneh. The dark blue and dark green lines represent the control simulations driven by 1980–2019 Daymet and 1980–2013 Livneh, respectively. The right panel shows the probability density distributions of projected annual evaporation loss in both baseline (1980–2019) and future (2020–2059) periods.

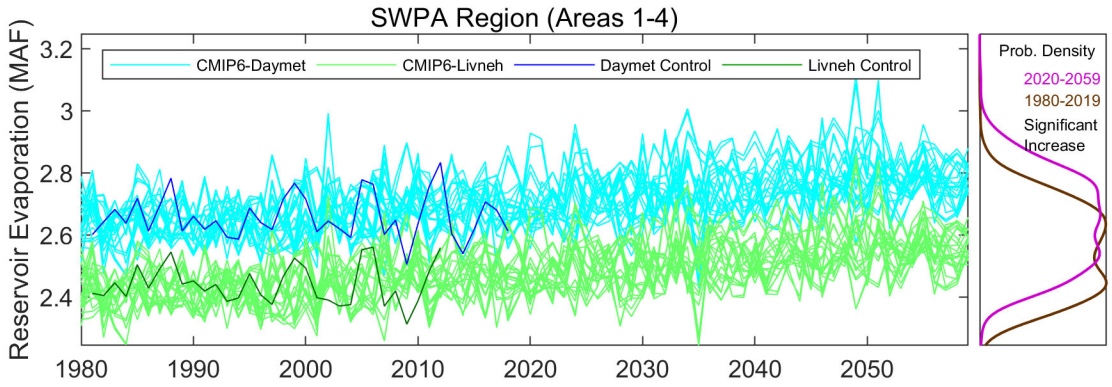


Figure 5.10. Projected annual total reservoir evaporation in the SWPA region.

Figure 5.10 shows that the results can be classified into two groups by their values. The high-value group consists of the 24 members that are associated with Daymet, whereas the low-value group is from the 24 members that are associated with Livneh. The main reason for this grouping is the considerable wind speed differences between these two data sets, which significantly impacts the evaporation rates (Section 3.2.2). The reservoir evaporation loss also depends on the projected reservoir area, which is inferred from runoff simulated by either VIC or PRMS hydrologic models. The projected runoff associated with Daymet or Livneh is not significantly different because the HMs were calibrated separately by both MetFs against the same USGS WaterWatch data set (Section 2.5.3) to minimize the impacts of different meteorological forcings on HMs. Therefore, the large uncertainties of the projected reservoir evaporation loss in the SWPA region are mainly attributed to the projected evaporation rate driven by different MetFs.

Despite the systematic differences between the two groups, the reservoir evaporation loss shows a clear increasing trend (as indicated by the two-sample Kolmogorov-Smirnov test performed at the 5% significance level). The Livneh-based future reservoir evaporation projections increase significantly from 2.4 MAF in 1980 to 2.5 MAF in 2059, and Daymet-based projections also increase similarly from 2.6 MAF in 1980 to 2.7 MAF in 2059. The average percentage increment is about 4% over the 80 years. Because the projected reservoir area in the SWPA region will not increase much in the future, the change in reservoir area only accounts for 15% of the increase in future reservoir evaporation loss, and the change in evaporation rate contributes the remaining 85%.

The projected annual and seasonal changes of monthly reservoir evaporation loss (in units of thousand acre-feet per month) from federal hydropower reservoirs in the SWPA region are further summarized in Figure 5.11. The results suggest that the annual ensemble median evaporation loss is projected to increase 5 KAF/month in the near-term (2020–2039) and 10 KAF/month in the mid-term (2040–2059) future periods with respect to the baseline period (1980–2019). The overall increase of reservoir evaporation is relatively uniform across the four seasons, with a slightly larger increment in the fall. Summer evaporation loss is not expected to change in the near-term future period but is projected to increase by 3 KAF/month in the mid-term future period. Fall evaporation loss is expected to increase by 7 KAF/month in the near-term and 13 KAF/month in the mid-term future periods. Although the changes of reservoir evaporation are similar in spring and winter, the range of uncertainty in spring is larger than in winter. Spring evaporation loss is expected to increase by 7 KAF/month in the near-term and 8 KAF/month in the mid-term future periods. In winter, evaporation loss is expected to increase by 3 KAF/month in the near-term and 9 KAF/month in the mid-term future periods.

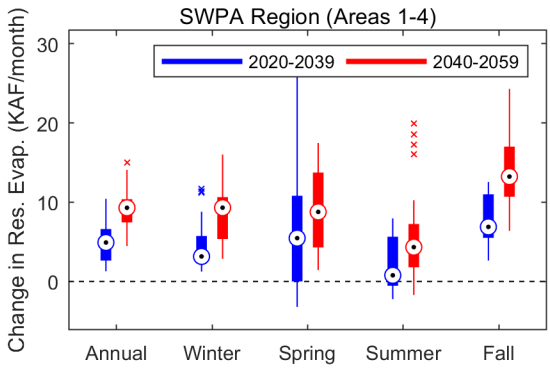


Figure 5.11. Projected change of annual and seasonal total reservoir evaporation in the SWPA region.

The source of variance in the projected change of annual reservoir evaporation loss from federal hydropower reservoirs in the SWPA region is shown in Figure 5.12. The box plots show the magnitude of change (future minus baseline) in median annual reservoir evaporation loss across 48 ensemble members, grouped by each of the four modeling choices (CM, DS, MetF, and HM). The bottom panel shows the portion of variance from the ANOVA for each modeling choice.

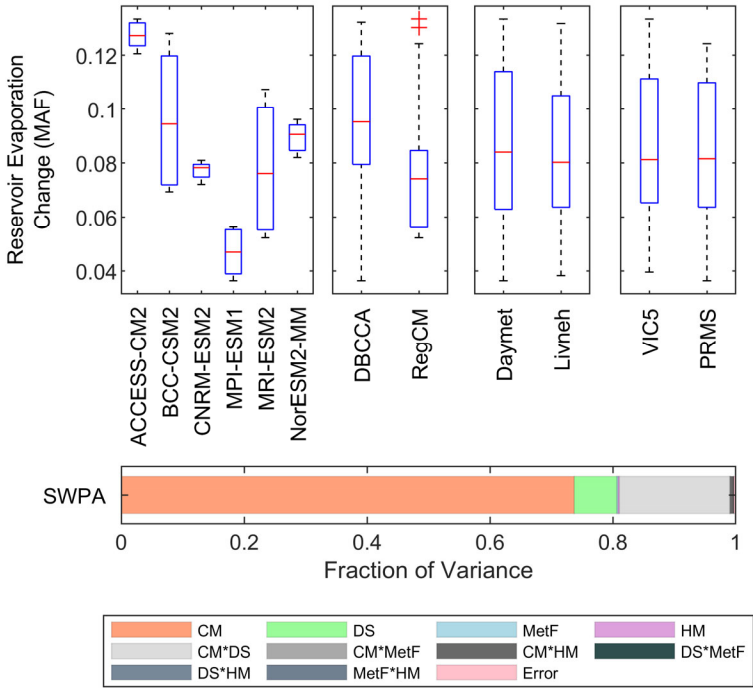


Figure 5.12. ANOVA of annual total reservoir evaporation in the SWPA region.

ANOVA results in Figure 5.12 indicate that CM is the dominant source of variance (around 70%). The most significant increase is projected by ACCESS-CM2 (0.13 MAF), whereas the increase projected by MPI-ESM1 is the smallest (0.05 MAF). The DS is another important contributor because it explains 8% of the total variance. The combination of the CM and DS contributes about 20% of the total variance. In contrast, the influence from other factors is negligible. At the seasonal scale (APPENDIX F), the ANOVA indicates that certain contributing factors may have more substantial impacts than at the annual scale. For instance, although the CM is the most dominant contributor to variance at the seasonal scale, the increases seen by the different GCMs vary considerably by season. The CM explains 73% of the

evaporation loss variance in the spring, and the combination of the CM and DS explain another 24% of the total variance. Similarly, in the summer, the CM explains 53%, whereas the combination of the CM and DS explains 45% of the total variance. The CM is also the most dominant contributor in the fall and winter. The DS is the third most dominant contributor, explaining 20% of evaporation loss changes in the fall and 8% in the winter. Overall, DBCCA projects higher reservoir evaporation losses in the future than the RegCM. Although the choice of MetFs (i.e., Daymet or Livneh) may significantly influence the projected value of reservoir evaporation loss (Figure 5.10), the difference is systematic and largely reduced when examining the change from the baseline to future periods. This shows that even with large model/data bias, the relative change is still considered reliable and used in many hydroclimate impact assessment studies.

5.3 CLIMATE EFFECTS ON FEDERAL HYDROPOWER IN THE SWPA REGION

5.3.1 Projections of Hydropower Generation

Using the two hydropower models described in Section 2.6, the projections of monthly generation in the SWPA region are calculated for each of the 96 unique combinations of CMs, DSs, MetFs, and HMs and PMs. Projections based on multimodel ensembles are summarized in terms of the mean annual and seasonal changes for the near-term (2020–2039) and mid-term (2040–2059) future periods compared with the baseline period (1980–2019).

The interannual variability of annual hydropower generation for the baseline and future periods is shown in Figure 5.13. The gray lines show the annual total generation for all 96 members of the ensemble hydropower projections, the black line represents the ensemble median, and the blue line shows the 1980–2019 historic observation from EIA and PMAs. The corresponding probability distributions of annual generation in the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between baseline and future periods is statistically significant. The comparison (Figure 5.13, left column) suggests that the range of simulated and observed annual hydropower is generally consistent across all SWPA areas. The shapes of the probability distributions in both baseline and future periods are very similar.

The annual and seasonal change in projected hydropower generation is further summarized in Figure 5.14. The change in Figure 5.14 is defined as the percentage difference in the future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across the 96 multimodel combinations; the central mark indicates the multimodel median, and the edges of box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length are considered statistical outliers (but should not be excluded in the context of climate change).

The annual hydropower generation is projected to increase significantly in all SWPA areas. More specifically, the annual hydropower generation in the near-term (2020–2039) future period is projected to increase from 1% in SWPA-4 and 7% in SWPA-3 according to the multimodel median (Figure 5.14). A further increase is projected in the mid-term future period, ranging from 5% in SWPA-2 and SWPA-4 to 11% in SWPA-3.

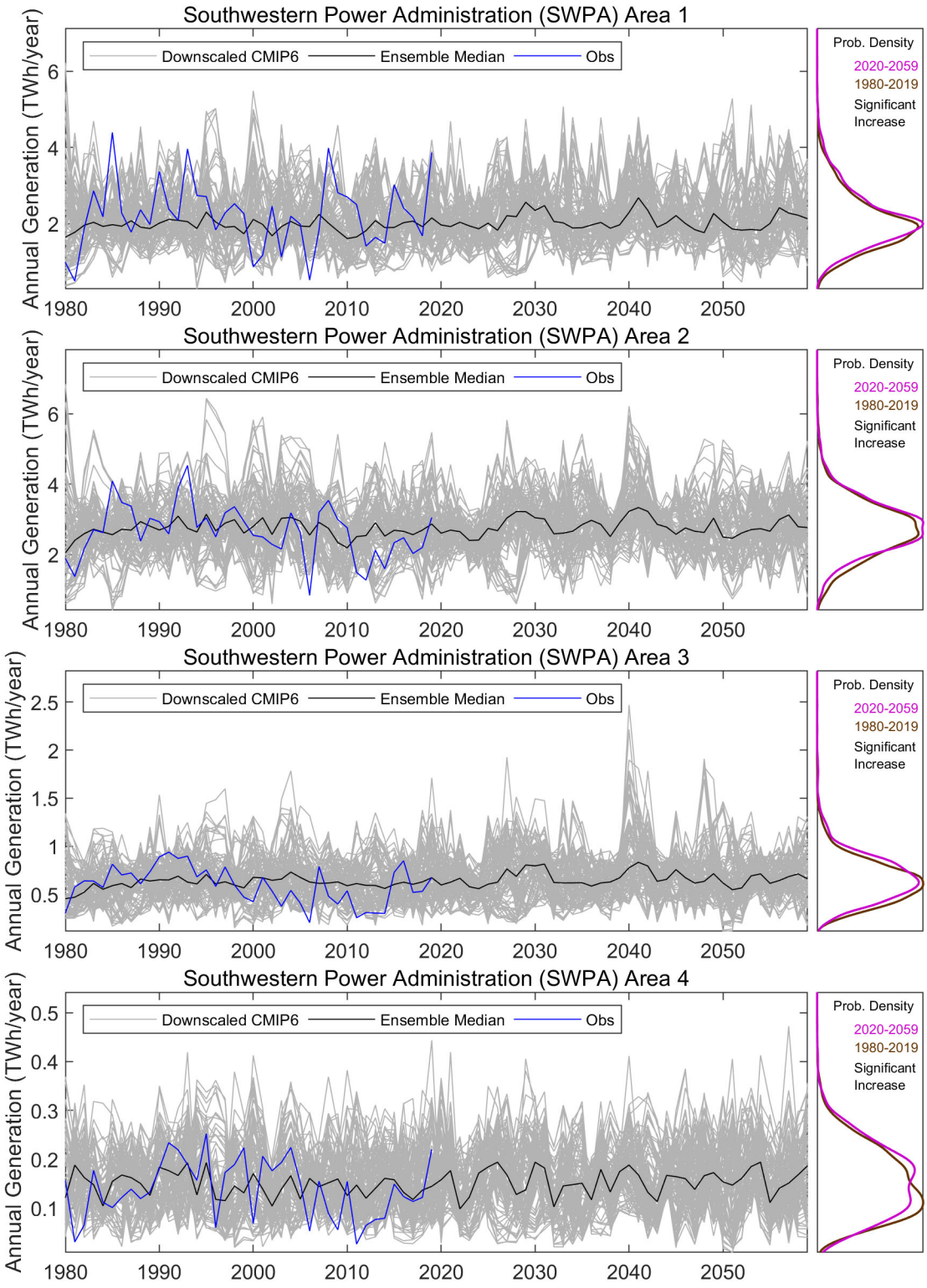


Figure 5.13. Projected annual total generation in the SWPA region.

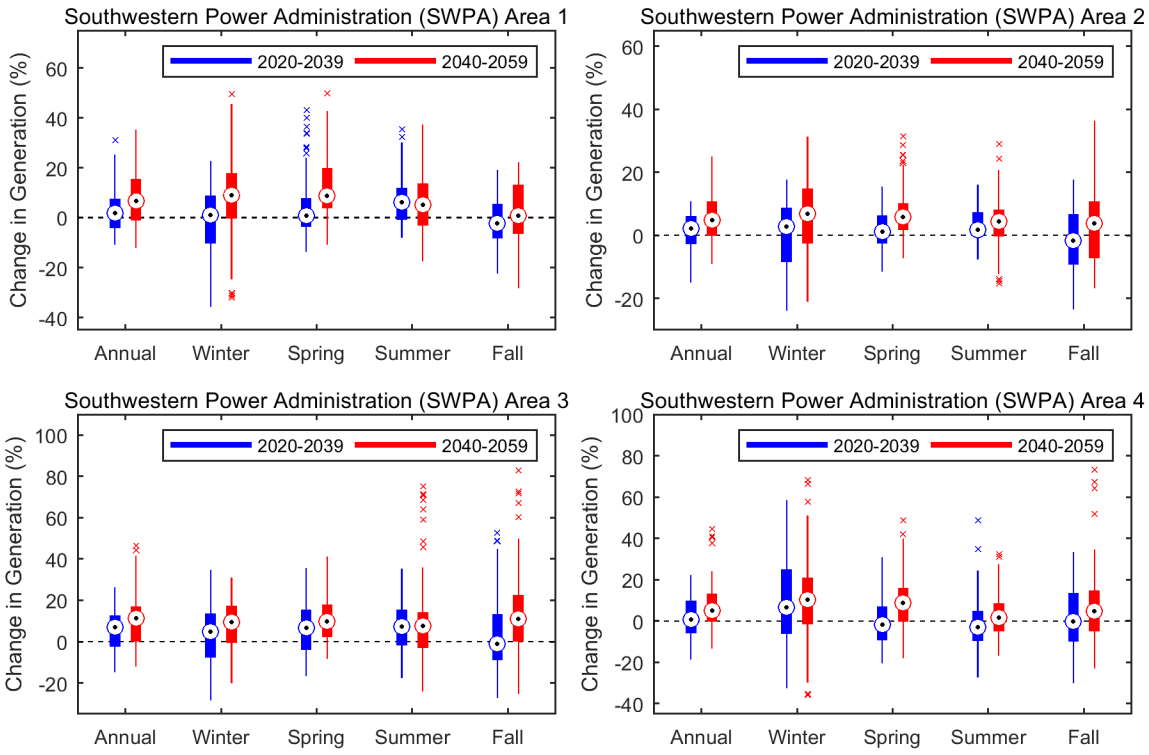


Figure 5.14. Projected change of annual and seasonal total generation in the SWPA region.

Seasonal relative changes tend to be more substantial and have larger uncertainty than annual changes. Winter generation is expected to increase in the near-term future period in all SWPA study areas (1%–7%), and the increase is expected to be more substantial in the mid-term future period (7%–10%). Spring generation in SWPA-1, SWPA-2, and SWPA-4 is projected to remain stable or slightly decrease in the near-term future period (−2%–1%) and is projected to increase in the mid-term future period (6%–9%). Spring generation in SWPA-3 is expected to increase significantly in both near-term (7%) and mid-term (10%) future periods. In summer, near-term hydropower generation in all SWPA study areas is expected to slightly increase (2%–7%), except for SWPA-4, where a decrease by 3% is projected. The overall increase of summer hydropower is projected to continue in the mid-term future period (2%–8%). In fall, near-term hydropower generation is projected to remain stable or slightly decrease (−2%– 0%). The decreasing trend is expected to reverse in the mid-term future period when fall generation is projected to increase (1%–11%) in all SWPA areas. The projected change of seasonal hydropower generation in the SWPA region generally follows the projected change of seasonal runoff. The storage capacity of reservoirs in the SWPA region is relatively smaller, resulting in this similarity in projected seasonal change between runoff and hydropower. In other words, the reservoir here may have less flexibility to alleviate the enhanced future runoff, which is especially noticeable in SWPA-3 and SWPA-4.

The source of variance in the projected annual hydropower generation change between the baseline and future periods is shown in Figure 5.15. The box plots in the top panels show the magnitude of change (future minus baseline) in projected annual hydropower generation across 96 ensemble members when grouped by the five modeling choices (CM, DS, MetF, HM, and PM). The bottom panel shows the fraction of variance from the ANOVA for each modeling choice.

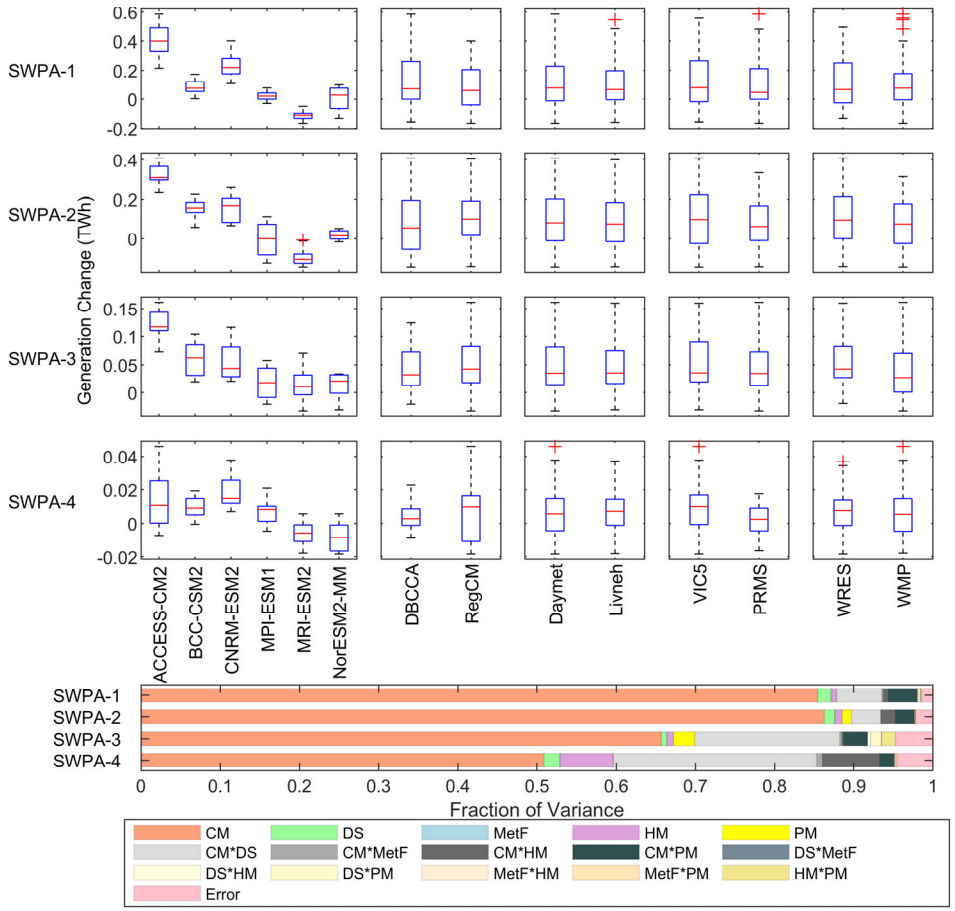


Figure 5.15. ANOVA of annual total generation in the SWPA region.

The diversity in individual factors and their combinations were further evaluated to understand their contributions to the spread of hydropower generation projections. The results suggest that the CM is the most dominant contributor to the variance of projected annual hydropower generation in all SWPA study areas, ranging from 51% to 86%. None of the other factors, including the DS, MetF, PM, and HM, alone contributes substantially to the annual variance. The combination of the CM and DS is a significant variance contributor in SWPA-3 (18%) and SWPA-4 (26%).

At the seasonal scale (APPENDIX G), the CM remains the most dominant contributor to the variance, followed by combinations of the CM with the PM, and the CM with the DS, throughout the four seasons across all SWPA areas. Overall, the CM contributes to 55%–72% of the variance in the winter, 46%–83% in the spring, 37%–64% in the summer, and 33%–47% in the fall. In the summer, the combination of the CM and DS tends to be the second most dominant contributor across all areas. In the fall, the contribution of the PM is more pronounced in SWPA-1, SWPA-2, and SWPA-3. Fall is also the season with the largest error.

In summary, federal hydropower generation in the SWPA region is projected to increase at the annual time scale in the near-term future period, and further increase in the mid-term future period. Seasonal changes are expected to be more substantial in the mid-term future period than in the near-term future period. When compared with the runoff projection, the hydropower projection has a similar response to climate change. This is somewhat different than what is projected in BPA and WAPA regions, where the range of projected hydropower change is smaller than the range of projected runoff change. This projection may be influenced by the relatively smaller storage reservoir systems in the SWPA region that

provide less flexibility to buffer the increasing variability of projected future runoff. In terms of the sources of uncertainty, the CM and DS contribute most to the variance of hydropower generation projections at both annual and seasonal scales.

5.3.2 Climate Change Impacts on Regional Electricity Demand

This section summarizes the results from the statistical analysis, exploring the effects of projected temperature trends on the electricity sales of SWPA's preference customers to their end-use (residential, commercial, industrial) customers. As described in Section 5.1.2, SWPA markets the output from the 22 hydropower plants in its Integrated System to preference customers, mostly municipalities and cooperatives, as peaking power. The contracts between SWPA and its preference customers guarantee 1,200 h of energy per kilowatt of capacity contracted per year. The federal hydropower allocation that preference customers receive from SWPA is a small fraction of the portfolio of generation resources they use to serve their end-use customers. Any load growth projected by SWPA's preference customers, due to climate trends or other drivers, would have to be met through generation sources other than federal hydropower.

Table 5.3 summarizes the elasticities of the preference customer sales with respect to the degree days for the sample of 35 SWPA preference customers included in EIA Form 861. The full results of the regression equations are presented in APPENDIX H.

Table 5.3. Estimated responsiveness of electricity sales by SWPA preference customers to degree days

Marketing region	Meteorological observation	Dependent variable	Elasticity with respect to HDDs	Elasticity with respect to CDDs
SWPA-Integrated	Daymet	Electricity sales	**0.0806** (0.0172)	**0.1553** (0.0263)
SWPA-Integrated	Livneh	Electricity sales	**0.0957** (0.02)	**0.1639** (0.0344)

Note: Numbers in parentheses are standard errors. Bolded values indicate that the estimated elasticity is statistically significant.

Electricity sales by SWPA preference customers are positively related to both HDDs and CDDs and the relationship is statistically significant. Electricity sales are more responsive to increases in CDDs (i.e., number and intensity of warm days triggering air conditioning use) than to increases in HDDs (i.e., number and intensity of cold days triggering heating demand). All else equal, a 1% increase in annual CDDs leads to a 0.16% increase in sales. A 1% increase in annual HDDs leads to less than a 0.1% increase in total sales. The differences in results across the two MetFs are minimal.

The estimated elasticities with respect to degree days were combined with the projected changed in degree days derived from climate simulations to produce an estimate of the temperature-driven changes in the demand served by SWPA preference customers. Figure 5.16 and Figure 5.17 show the projected degree days from 24 simulations (the six CMIP-6 models described in Section 2.3 combined with two MetFs and two DSs), as well as the ensemble median projection.

During the baseline period, the ensemble median HDDs and CDDs in the territory served by SWPA preference customers were similar: 2,298 annual HDDs and 2,119 annual CDDs on average. For 2020–2059, the average values of the ensemble medians in Figure 5.16 and Figure 5.17 are 1,785 for HDDs and 2,749 for CDDs. This represents a 22% decrease in HDDs and a 30% increase in CDDs. The projected changes in sales driven by these degree day trends are presented in Figure 5.18.

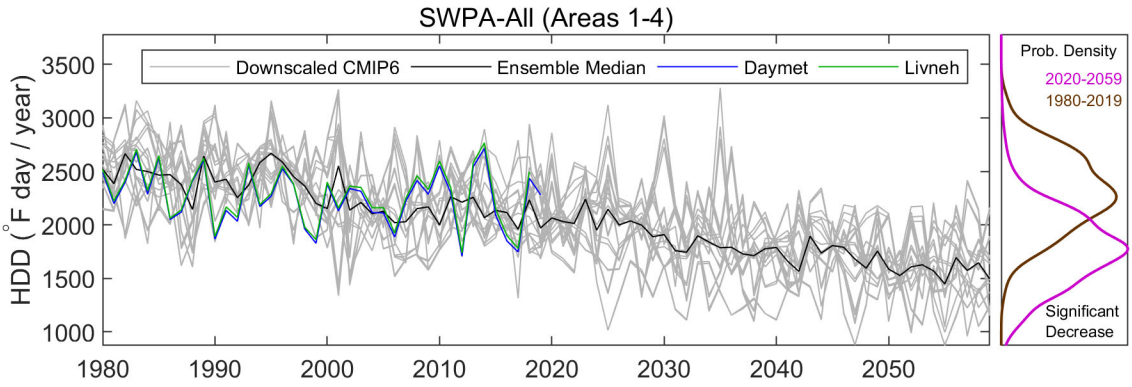


Figure 5.16. Projected annual HDDs in the SWPA region.

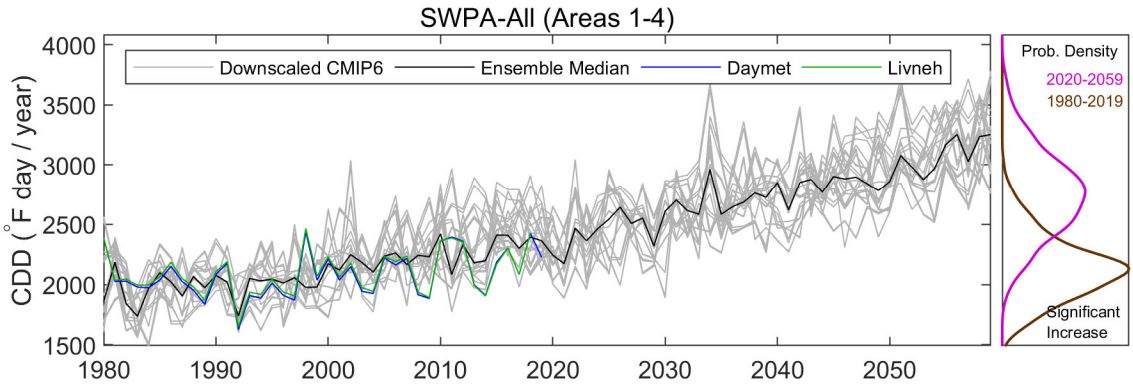


Figure 5.17. Projected annual CDDs in the SWPA region.

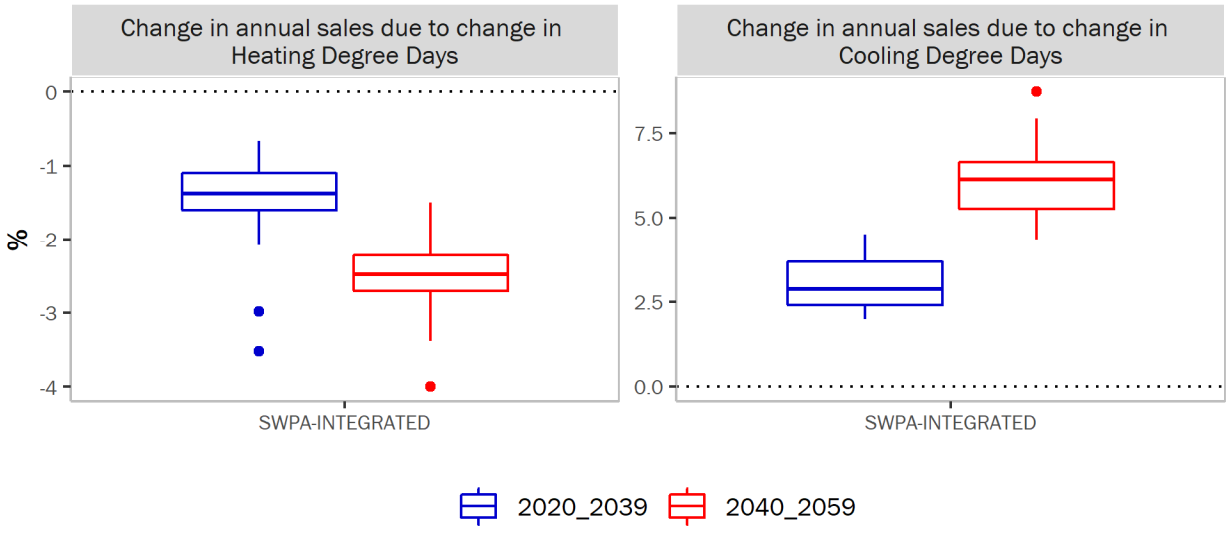


Figure 5.18. Estimated changes in SWPA preference customer electricity sales relative to the baseline period due to projected changes in degree days.

Changes in sales driven by HDDs and CDDs occur during the winter and summer, respectively. For each of the degree day–period combinations under all 24 CMIP6 projections (Figure 5.18), electricity sales decrease in the winter months and increase in the summer months because of the trends in degree days

summarized in Figure 5.16 and Figure 5.17. Overall, annual sales from the preference customers to their end-use customers increase because the changes in summer sales from increased CDDs are larger than the decrease in winter sales driven by decreases in HDDs. The median projected increase in annual sales relative to the 1980–2019 baseline period is 1.5% in 2020–2039 and 3.6% in 2040–2059.

Based on projected temperature trends, the capacity and peaking power marketed by SWPA will likely be most valuable to the preference customers during the summer months. SWPA's preference customers have some limited flexibility regarding the time of the year in which they receive their energy allocation (within the monthly and seasonal ranges specified in their contracts) to the extent that their preferred timing is feasible for the PMA given other federal hydropower operational constraints.

An example of how changes in the timing of preference customer demand create challenges for the PMAs arose when SWPA experienced uncharacteristic increased demand from its preference customers during several early fall (September and October) heat waves over the 2015–2020 period. Increased demand at this time of year coincides with restricted hydropower availability (due to maintenance outages having typically been scheduled for fall or spring months) and has resulted in more costly replacement power purchases during those months (Fernandez, 2020).

Results presented in Section 5.3.1 show that annual hydropower generation is projected to increase within the SWPA areas in 2020–2059. However, in the near-term future period (2020–2039), fall shows stable or slightly decreasing generation in all the SWPA areas, which could prolong or exacerbate the imbalances between supply and demand experienced during the early fall period in recent years.

(This page intentionally left blank)

6. THE SEPA REGION

Following the methodology described in Section 2, the results for federal hydropower plants marketed by the SEPA are summarized in this section. Descriptions of the SEPA study areas, federal hydropower systems, and power marketing practices are described in Section 6.1. The projected future atmospheric and hydrologic conditions in the SEPA region are discussed in Section 6.2. The potential climate change impacts on federal hydropower generation and SEPA regional energy demands are discussed in Section 6.3.

6.1 SEPA REGIONAL CHARACTERISTICS

6.1.1 Federal Hydropower in the SEPA Region

SEPA markets hydroelectric power in 10 southeastern states and southern Illinois (Figure 6.1). All federal hydropower plants in the southeastern region are owned and operated by USACE. These 22 USACE hydropower plants have a total of 3,838 MW in capacity and produce 7.3 TWh/year in average annual generation (Table 6.1).

Table 6.1. Summary of federal hydropower plants in the SEPA region

Area	Area name	Number of USACE plants	Capacity[a] (MW)	1980–2019 generation[b] (GWh/year)
SEPA-1	Kerr-Philpott	2	311	458
SEPA-2	Cumberland	9	932	3,063
SEPA-3	GA/AL/SC	10	2,552	3,599
SEPA-4	Jim Woodruff	1	44	202
Total		22	3,838	7,323

[a] EIA 2019 total nameplate capacity, including both conventional hydro and pumped storage.
[b] EIA and SEPA average annual generation from 1980 to 2019, conventional hydro only.

Most of the capacity (66%) is located at the 10 projects along large rivers flowing out of Georgia, Alabama, and South Carolina. The Cumberland River system accounts for most of the remainder (24%) of SEPA's power capacity. SEPA's two largest hydropower plants are Richard B. Russell (628 MW) and Carters (500 MW). The smallest hydropower plants are J. Percy Priest (28 MW) and Philpott (14 MW). The Richard B. Russell and Carters hydropower plants also have some reversible turbines that may allow for pumped storage operation. The oldest hydropower plants are Allatoona and Center Hill built in 1950. The newest hydropower project is Richard B. Russell on the upper Savannah River, which began operation in 1984. The average age of USACE hydropower projects in the southeastern region is 40 years. Therefore, equipment modernization and refurbishment are also a common challenge for the federal hydropower plants in this region. A complete list of all federal hydropower plants in the SEPA region is provided in APPENDIX B.

SEPA experiences some of the same water resources management challenges that occur in other parts of the United States. Growing demand for municipal and industrial water supply, especially in the rapidly growing Atlanta area of northern Georgia, is the most pressing issue (e.g., Magnuson, 2009). Recreational and environmental water uses are also changing how surface water is managed (Whisnant et al., 2009). Reallocating water storage among all of these competing water uses has had a gradual but increasingly adverse impact on federal hydropower production.

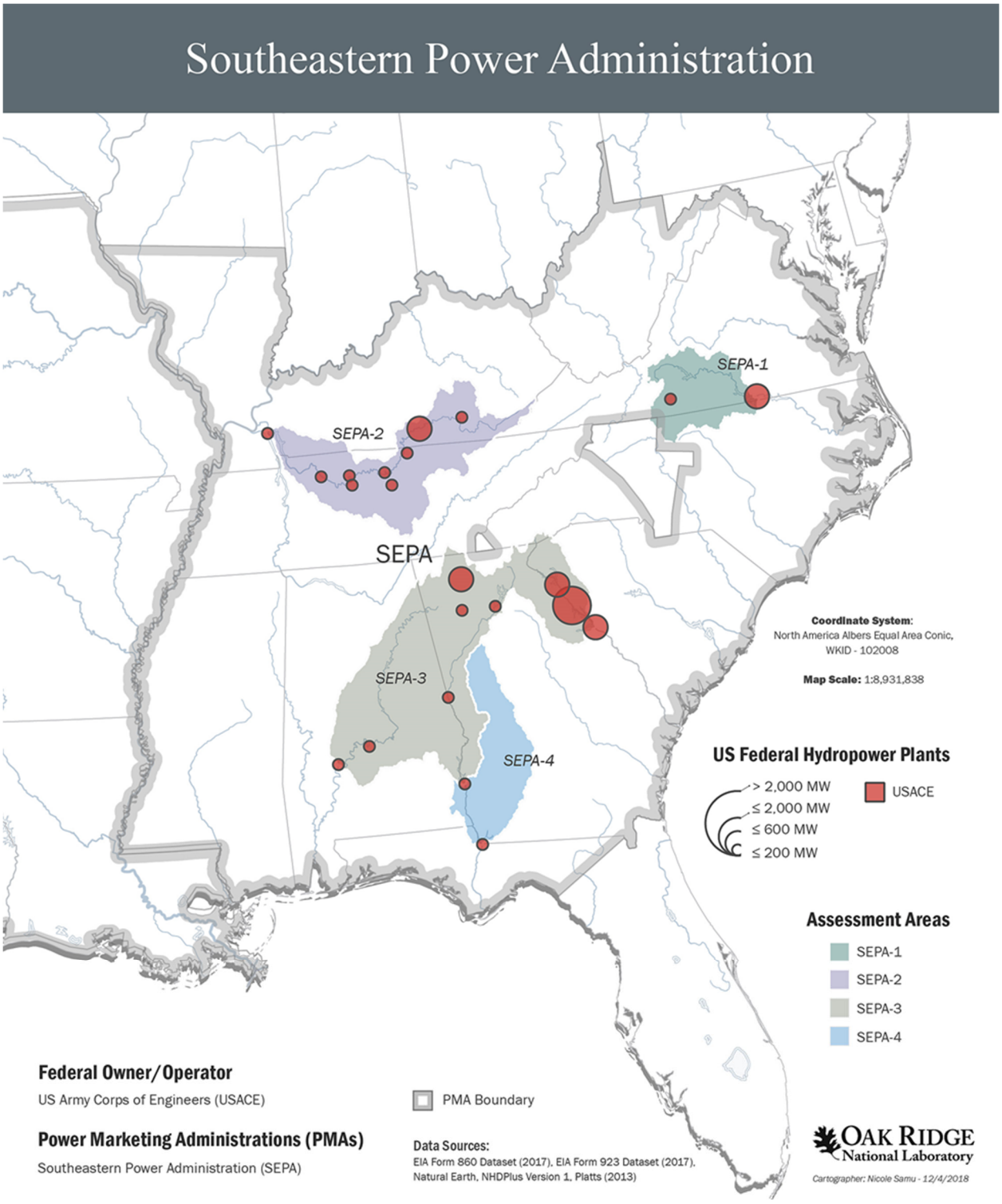


Figure 6.1. Map of the federal hydropower plants and four study areas in the SEPA region.

Considering the power system and watershed boundaries, the SEPA region is subdivided into four study areas. The SEPA region is different from the other three PMA regions with its lower elevations, higher precipitation, and more heavily vegetated, deciduous land cover.

- **SEPA-1 Kerr-Philpott**: SEPA-1 is located in the upper half of the Roanoke River basin above the John H. Kerr plant. The watershed begins on the eastern slope of the Blue Ridge Mountains in southwestern Virginia and extends down into the mid-Atlantic coastal plain. The total drainage area is 7,866 mi^2, making it one of the smallest regions evaluated in the 9505 assessment. Elevations range from 3,743 to 299 ft, with a median elevation of about 700 ft. Dominant land cover types are cropland–natural vegetation mosaic (52%), deciduous broadleaf forest (32%), mixed forest (8%), and minor cropland (3%).

- **SEPA-2 Cumberland**: The Cumberland River basin is located in Kentucky and Tennessee, extending from the Appalachian Mountains westward to its confluence with the Tennessee and Ohio Rivers. The total drainage area is 17,607 mi^2. Elevations range from over 4,000 ft in the eastern headwaters to 300 ft at the Ohio River, with a median elevation of approximately 900 ft. Land cover is a diverse mixture of deciduous broadleaf forest (39%), cropland–natural vegetation mosaic (36%), cropland (17%), and mixed forest (3%).

- **SEPA-3 GA/AL/SC**: SEPA-3 covers the upper portions of three relatively large river basins, all of which have their headwaters in the southern Blue Ridge Mountains. The Savannah River flows to the Atlantic Ocean, and the Apalachicola and Alabama River systems flow to the Gulf of Mexico. The total drainage area of these three watersheds is 34,244 mi^2. Elevations range from over 5,000 ft to about 40 ft, from mountains in the northern end to the southern coastal plains. The median elevation of this area is 682 ft. Land cover is mostly cropland–natural vegetation mosaic (42%) and mixed forest (35%), plus minor amounts of woody savanna (12%) and deciduous broadleaf forest (4%).

- **SEPA-4 Jim Woodruff**: SEPA-4 is defined by one small USACE hydropower project, the Jim Woodruff project, located on the border between Georgia and Florida. Power from this project is marketed separately by Southeastern; therefore, it defines a distinct area on the Georgia Piedmont and coastal plain. The total watershed area is 17,164 mi^2, 39% of which comes from the upper Apalachicola River that flows through the Atlanta metropolitan area and from the Flint River. Land cover is primarily cropland–natural vegetation mosaic (50%), mixed forest (20%), cropland (17%), and woody savanna (9%).

6.1.2 Power Marketing by SEPA

SEPA was established in 1950 as the PMA in charge of marketing the output from multipurpose USACE-owned dams in the southeastern United States (apart for those owned and operated by TVA). SEPA markets the power and energy from 22 hydropower plants, in accordance with authorizations by the Flood Control Acts of 1938, 1944, 1950, and 1966. A general description of what a PMA is and how it conducts its mission is presented in Section 1.2.2. As of 2019, SEPA had long-term contracts with 477 customers; except for one IOU, all the other customers are public bodies or cooperatives.

SEPA is the only PMA that does not own any transmission assets; instead, it delivers the federal hydropower to its customers through wheeling service contracts with other utilities in the region. The Kerr-Philpott subregion projects lie within the PJM Interconnection service footprint. SEPA relies on PJM Interconnection transmission resources to wheel electricity from those projects to their customers. For that reason, SEPA joined PJM Interconnection in 2005. For all other projects, SEPA operators arrange for transmission to their respective preference customers. All power deliveries are registered in the OATI (Open Access Technology International) webSmartTag transaction-tagging system. This is an online system that operators use to coordinate transmission from the projects to the customers (Fernandez, 2020). For power allocations and rate-setting purposes, the fleet marketed by SEPA is divided into four subregions: Cumberland, GA/AL/SC, Kerr-Philpott, and Jim Woodruff. Marketing policies changed significantly in the 1980s. The changes included expanding the number of preference customers served

and ending capacity sales without energy contract commitments to the owners of the transmission lines delivering the federal hydropower to other customers.

SEPA sells all the output from the federal hydropower plants through long-term contracts. Most contracts sell peaking power with a specified, fixed annual number of hours of energy per kilowatt of capacity. For instance, preference customers in the Cumberland subregion receive 1,500 h of peaking power per kilowatt of capacity contracted. In any month, these preference customers can schedule 60–240 h/kW. If they request more than 240 h of energy per kilowatt in a month, SEPA will only grant the request if the combined schedule of all preference customers does not exceed 240 h/kW.

Table 6.2 displays SEPA's capacity and energy rates for each of the four subregions in FY 2020. FERC approves the power rates on five-year terms subject to annual adjustments that reflect replacement power costs and new investments coming online.

Table 6.2. SEPA rate schedule summary (FY 2020)

System	Capacity rate ($/kW-month)	Energy rate (mills/kWh)
Cumberland	3.43[a]	12.835
GA/AL/SC	4.09[b]	12.33
Kerr-Philpott	4.40	17.80
Jim Woodruff	7.74	20.44[c]

[a] Customers served through transmission lines owned by Duke Energy Progress (Western Division) pay a capacity rate of $3.904/kW-month and East Kentucky Power Cooperative pays a capacity rate of $1.826/kW-month.
[b] The capacity rate for South Mississippi Electric Power Association is $4.81/kW-month.
[c] Duke Energy Florida pays a monthly rate equal to 100% of the calculated saving in the cost of fuel per kilowatt-hour.

SEPA, in partnership with USACE, is in the process of reviewing the allocation of the joint use costs of the multipurpose hydropower facilities whose power it markets. Joint use costs are those that cannot be exclusively attributed to one of the authorized purposes of the facility, such as hydropower, recreation, water supply, fish and wildlife, or water quality, and must therefore be shared among multiple purposes. The costs allocated to the hydropower purpose are expected to decrease because since the facilities were built, additional non-hydropower purposes have been authorized or received increased water allocations. If the costs allocated to the hydropower purpose decrease, SEPA's cost-based rates to its preference customers will also decrease. Based on electricity load forecasts produced before the 2008 recession, some SEPA customers acquired extra capacity that has not been needed. Those customers are now deciding how to manage this excess capacity. This excess capacity has been the reason for some recent SEPA contract cancellations (CRS, 2019). In FY 2018, 11 preference customers with a combined allocation of approximately 85 MW in the GA/AL/SC subregion submitted a 25-month notice to terminate their contracts. In May 2019, another customer from that subregion with an allocation of 13.4 MW submitted a contract cancellation notice. SEPA is reallocating the capacity and energy from those customers to 65 others in the same subregion that expressed interest in a supplemental allocation.

SEPA is also experiencing challenges to maintain competitively priced electricity rates because the amount of water allocated to hydropower has been decreasing as other water resource uses, particularly urban water supply, take priority. Multiple competing uses for the regional water supply in addition to sustained drought conditions (low precipitation levels) have resulted in an increased need for replacement power purchases. For many customers, the replacement power costs are higher than alternative sources,

prompting some to reconsider their federal hydropower contracts, reduce contract volumes, or cancel contracts (Fernandez, 2020).

6.2 FUTURE CLIMATE IN THE SEPA REGION

6.2.1 Regional Climate Projections

The annual temperatures for the baseline (1980–2019) and future (2020–2059) periods are presented in Figure 6.2. The gray lines show the annual mean temperature from 24 sets of downscaled climate projections, the black line represents the multimodel median, and the blue and green lines show the Daymet and Livneh historical observations, respectively. These plots show interannual variability and annual trends for projections, as well as the observations. The climate models are not expected to replicate the exact timing of historical interannual and decadal variability. Therefore, they do not completely follow the historical interannual values. The corresponding probability distribution for the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between the baseline and future periods is statistically significant.

In all four assessment areas (SEPA-1–4), the ensemble model median captures the increasing trend in annual mean temperature present in the observations during the historical period. However, there is significant variability across the ensemble members, which reflects the importance of using a multimodel ensemble approach. In the future period, annual mean temperature is projected to continually increase across all four areas. A comparison of the probability distribution of the multimodel mean annual temperature for the historical and future periods also shows a statistically significant shift toward a warmer climate across all SEPA areas (Figure 6.2).

The annual and seasonal changes in temperature are further summarized in Figure 6.3. The change is presented as the difference in temperature during the future periods (2020–2039 near-term and 2040–2059 mid-term) compared with the baseline period (1980–2019) in degrees Fahrenheit. Each box plot shows the spread across 24 sets of downscaled climate projections; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length denoted by x are considered statistical outliers (but should not be excluded in the context of climate change).

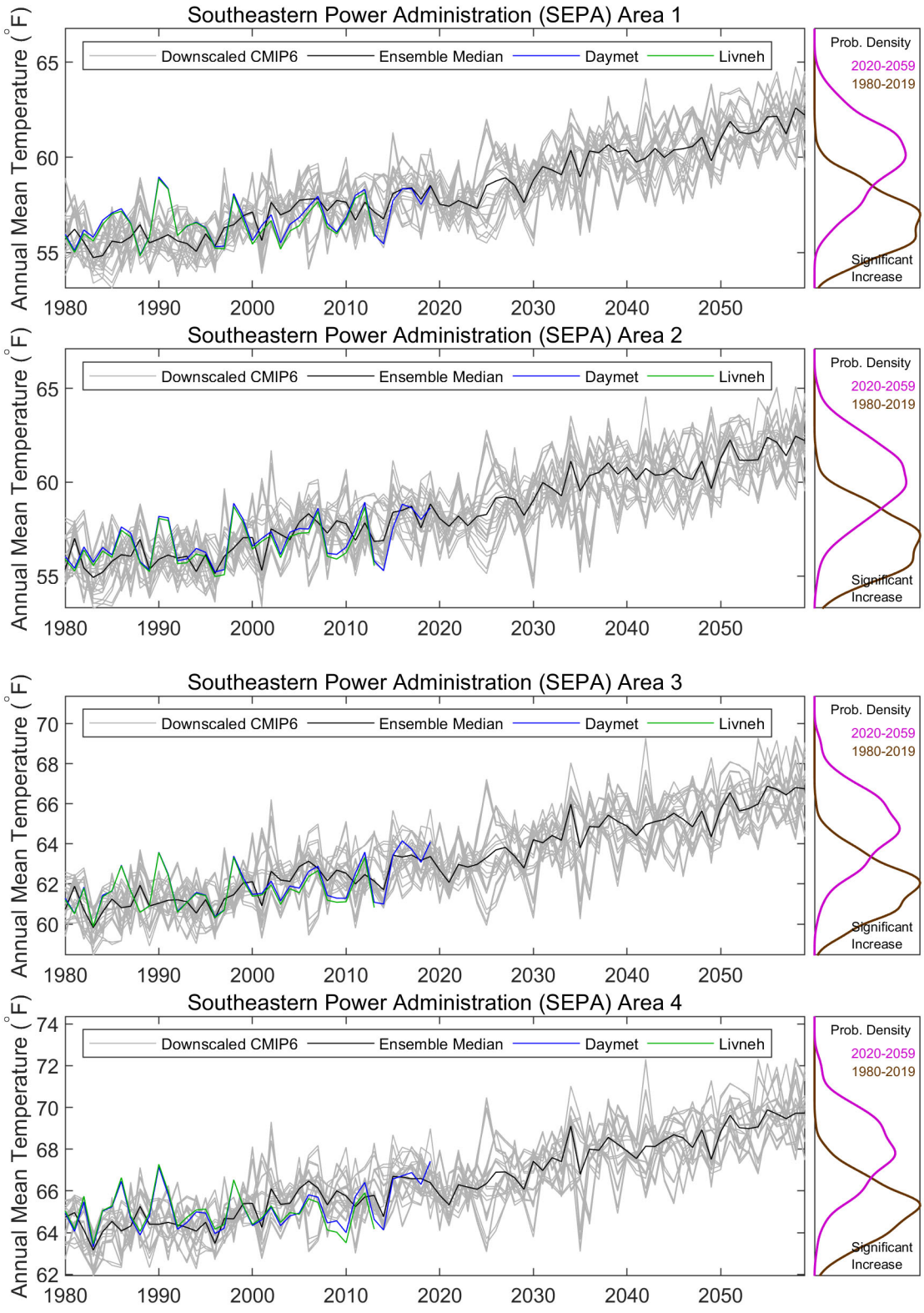


Figure 6.2. Projected annual mean temperature in the SEPA region.

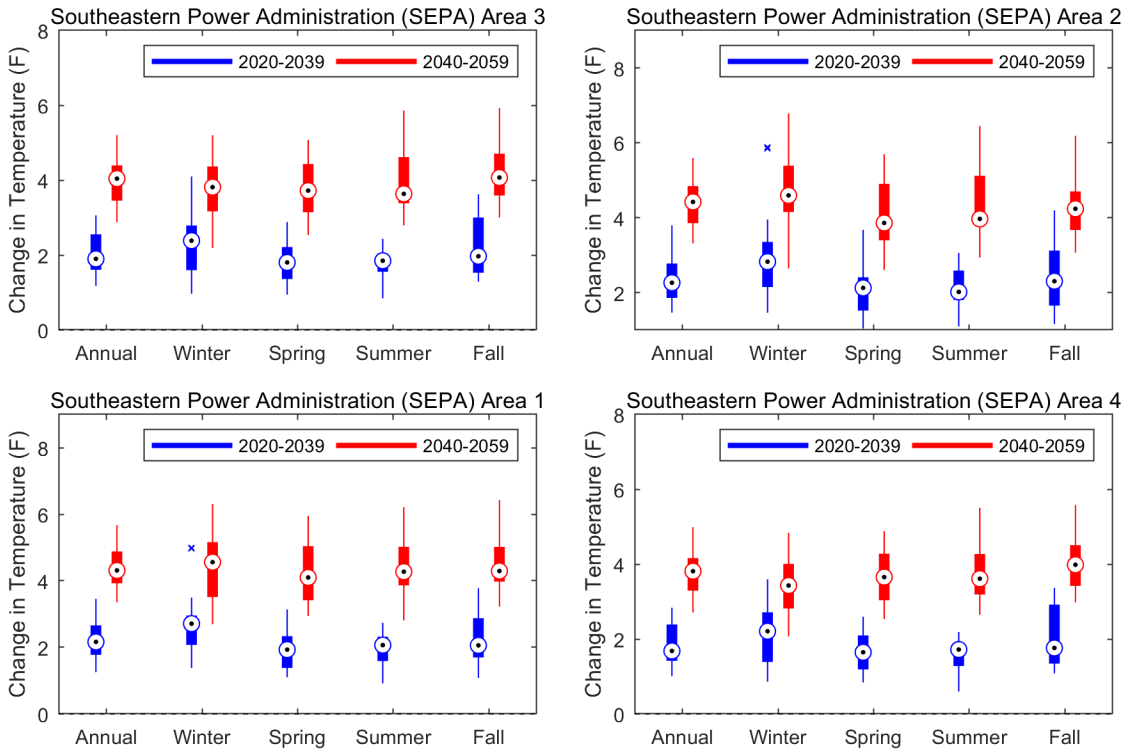


Figure 6.3. Projected change of annual and seasonal mean temperature in the SEPA region.

The annual and seasonal changes in precipitation over the SEPA regions are summarized in Figure 6.4 and Figure 6.5. For both the near-term (2020–2039) and mid-term (2040–2059) future periods, the ensemble median annual precipitation is projected to increase over the four areas, with a higher increase projected during the mid-term future period. These changes in annual precipitation are in accordance with other studies based on CMIP5 and CMIP6 models that show an increase in precipitation over the southeastern United States (Almazroui et al., 2021; Ashfaq et al., 2016). On a seasonal scale, the largest increase in median precipitation is projected during the winter and spring across all four areas in the mid-term future period. SEPA-1 and SEPA-4 also projected an increase in precipitation during the summer and fall.

The annual and seasonal changes in precipitation over the SEPA regions are summarized in Figure 6.5. For both the near-term (2020–2039) and mid-term (2040–2059) future periods, the ensemble median annual precipitation is projected to increase over the four areas, with a higher increase projected during the mid-term future period. These changes in annual precipitation are in accordance with other studies based on CMIP5 and CMIP6 models that show an increase in precipitation over the southeastern United States (Almazroui et al., 2021; Ashfaq et al., 2016). On a seasonal scale, the largest increase in median precipitation is projected during the winter and spring across all four areas in the mid-term future period. SEPA-1 and SEPA-4 also projected an increase in precipitation during the summer and fall.

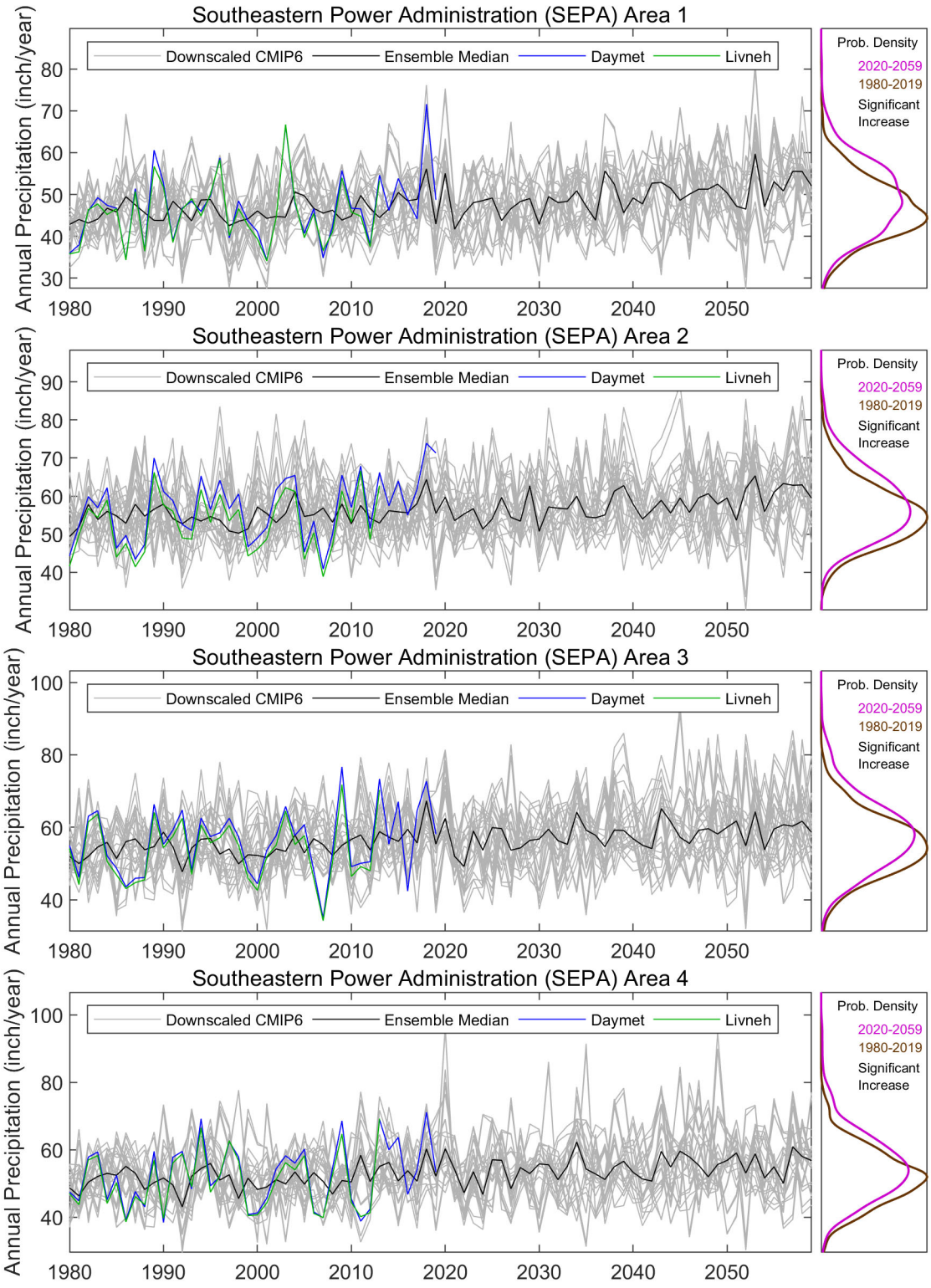


Figure 6.4. Projected annual total precipitation in the SEPA region.

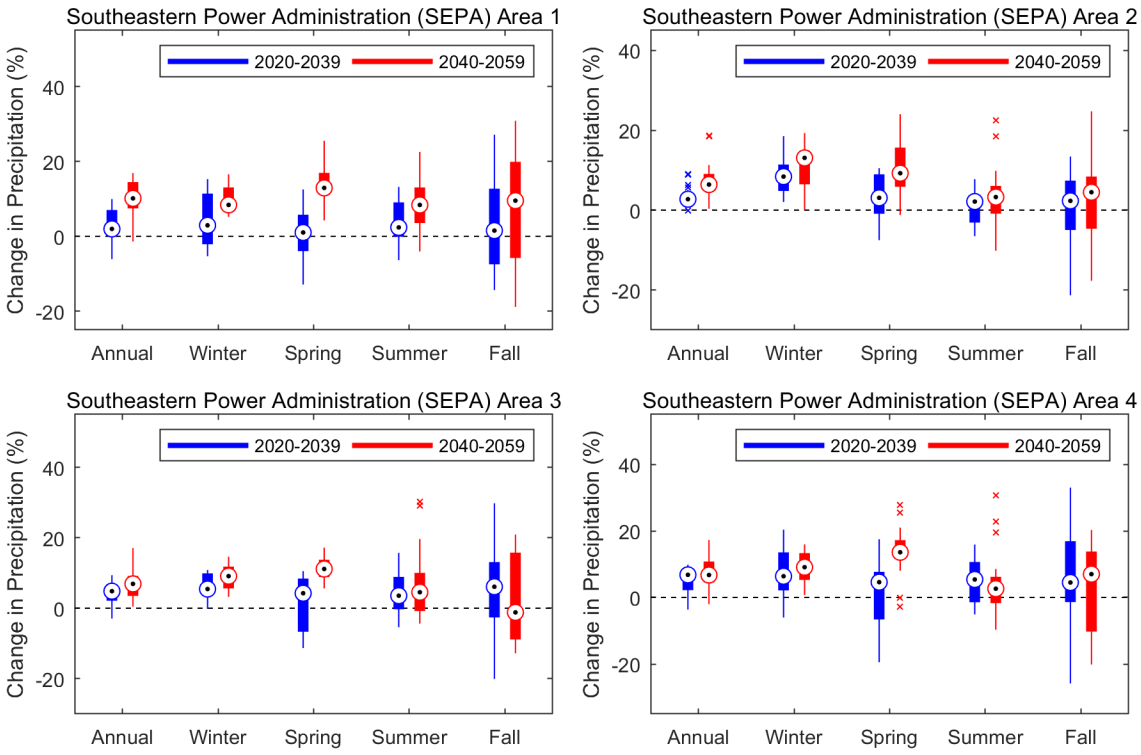


Figure 6.5. Projected change of annual and seasonal total precipitation in the SEPA region.

6.2.2 Regional Hydrologic Projections

The runoff projections for each of the SEPA regions are illustrated in Figure 6.6 and Figure 6.7. The annual total runoff in baseline (1980–2019) and future (2020–2059) periods is presented in Figure 6.6. The gray lines show annual runoff from 48 sets of hydrologic projections, the black line represents the multimodel median, and the blue and green lines show the control simulations driven by Daymet and Livneh observations, respectively. The corresponding probability distributions are compared in the right panel. The annual and seasonal changes in runoff are further summarized in Figure 6.7. Change is defined as the percentage difference of future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across 48 ensemble members; the central mark indicates the multimodel median, and the edges of the box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length are considered as statistical outliers (but should not be excluded in the context of climate change).

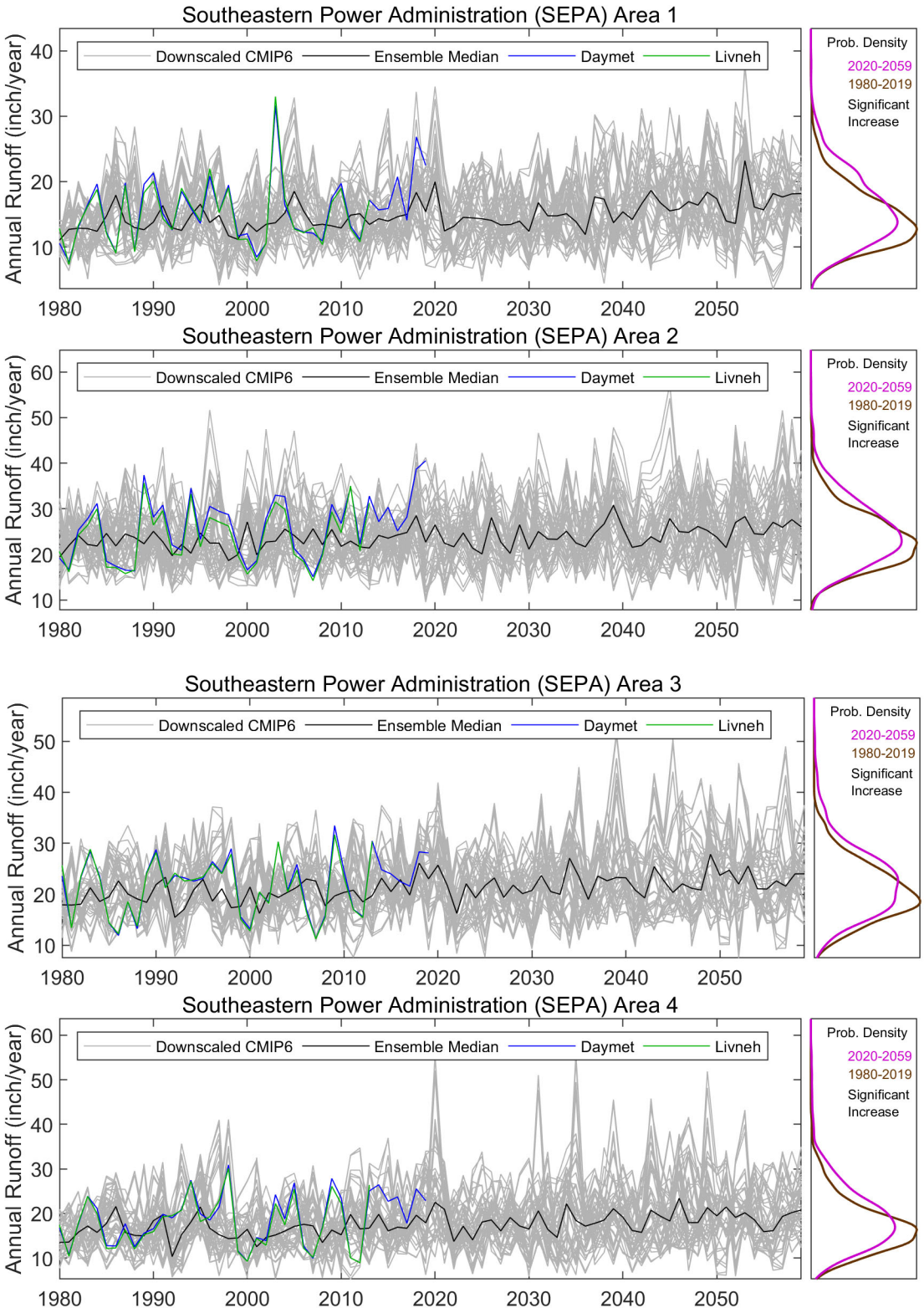


Figure 6.6. Projected annual total runoff in the SEPA region.

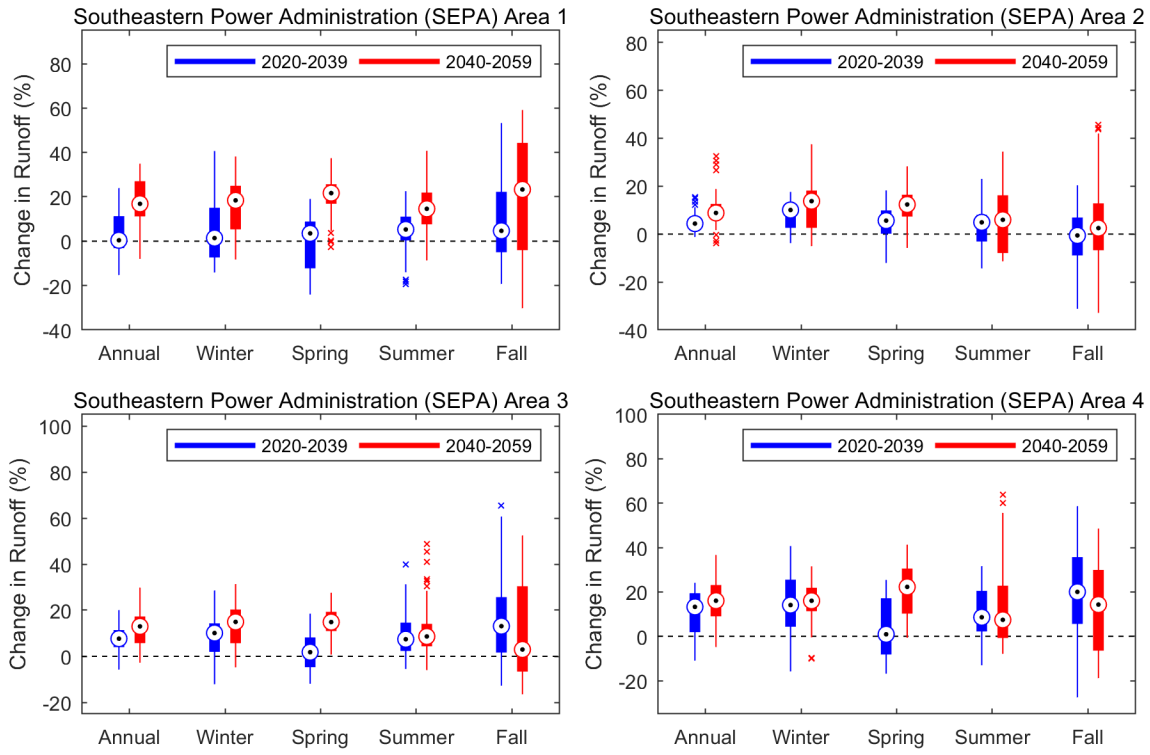


Figure 6.7. Projected change of annual and seasonal total runoff in the SEPA region.

The shifts of probability distributions (Figure 6.6) suggest that the annual runoff is projected to increase in the future period and the distributions are statistically different. Additionally, the probability distributions in all SEPA areas are projected to widen, indicating that the range of interannual variability is expected to increase. Such changes are more prominent and easily noticeable in SEPA-3 and SEPA-4. The maximum increase in median annual runoff is projected in the range of 6.2% to 14.9%, with the greatest change in SEPA-4 and the least increase in SEPA-1. More specifically, the median annual runoff is projected to increase across all study areas in the range of 0.4% to 13.2% in the near-term and 8.8% to 16.8% in the mid-term future periods. The range of annual runoff change is higher in the mid-term future period, indicating greater uncertainties in runoff projections.

At the seasonal scale, the ensemble median winter runoff is projected to increase across all SEPA areas in the range of 1.3% to 14.1% in the near-term future period, and the change is expected to further increase in the range of 13.7% to 18.4% in the mid-term future period. The spring runoff is projected to increase across all SEPA areas in the range of 0.9% to 5.6% in the near-term and up to 12.3% to 22.4% in the mid-term future periods. Summer runoff is expected to increase across all SEPA areas by 4.8%–8.6% in the near-term and 5.9%–14.5% in the mid-term future periods. The fall runoff is also expected to increase across all SEPA areas, except for SEPA-2, in the range of 4.7% to 20% in the near-term and 2.5% to 23.3% in the mid-term future periods. The fall runoff in SEPA-2 is projected to decrease by 0.6% in the near-term and increase by 2.4% in the mid-term future periods. The trends in winter, spring, and summer runoff projections indicate that these changes would be stronger in magnitude in the mid-term future period, although also accompanied with greater uncertainties across all ensemble members.

An additional analysis of extreme runoff is shown in Figure 6.8, which presents future changes in ensemble median high runoff (i.e., 95th percentile of daily runoff) and median low runoff (i.e., 5th percentile of 7-day average runoff) in both future projection periods. High runoff is generally projected to increase in both future periods across all SEPA areas. The change is expected to be stronger in the mid-

term future period. The low runoff shows a mixed response and is spatially heterogenous. The low runoff conditions are expected to decrease in portions of SEPA-1, SEPA-2, and SEPA-3. These results suggest increasing likelihood of more frequent flood events, which may increase the difficulty of water management in future climate conditions. These areas may experience both increase in high runoff and decrease in low runoff conditions owing to the intensification of hydrologic cycles at local scales.

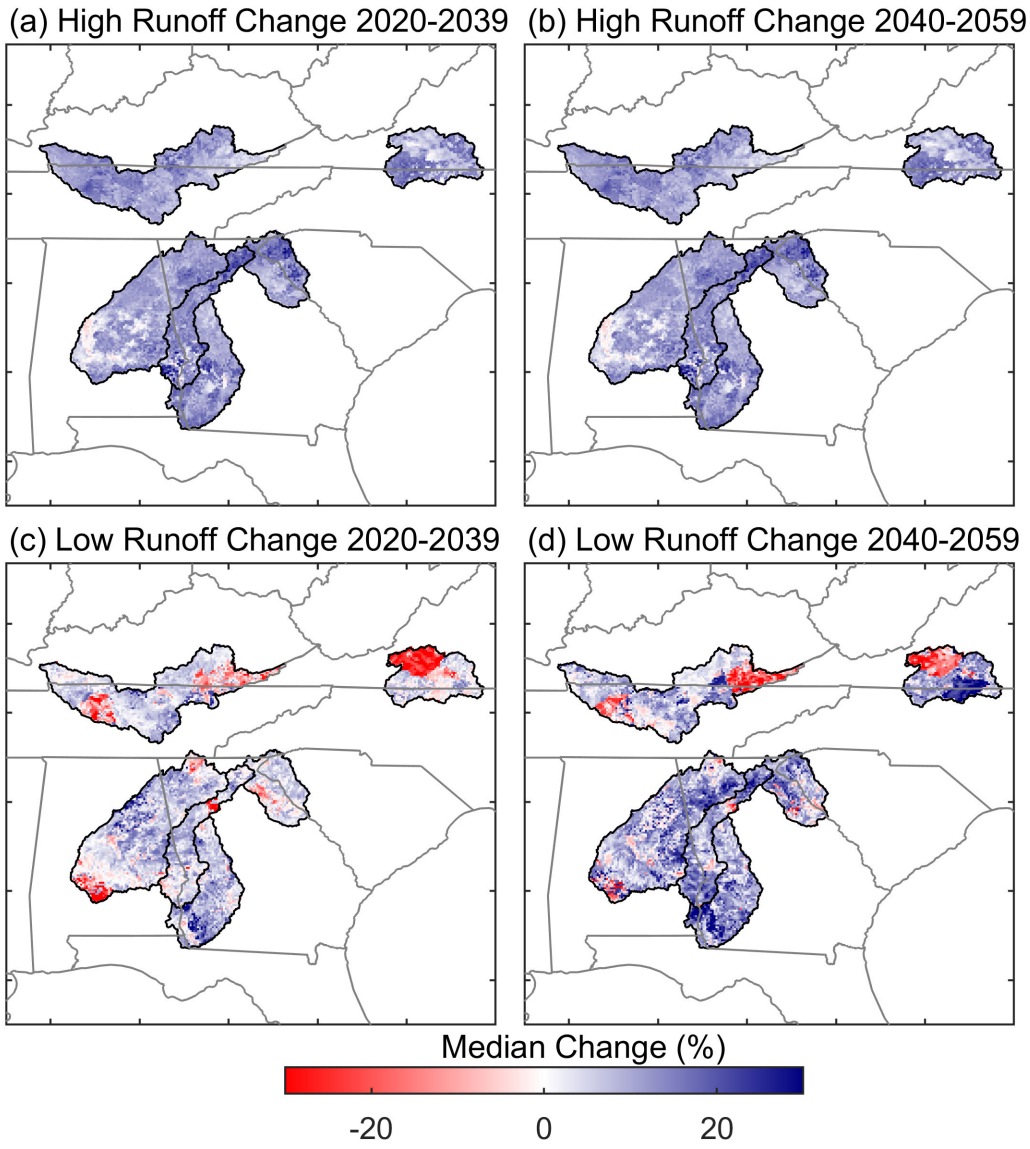


Figure 6.8. Projected change of high and low runoff in the SEPA region.

To evaluate the impacts of modeling choices (six CMs, two DSs, two MetFs, and two HMs) on the projection of future runoff in each SEPA area, an ANOVA was conducted and is illustrated in Figure 6.9. The box plots show the magnitude of change (future minus baseline) in median annual runoff across 48 ensemble members, grouped by each of the four modeling choices. The box plots provide a quantitative way to evaluate the range of possible outcomes in runoff response based on different modeling choices. ANOVA results indicate that the CM is generally the most dominant contributor to the total variance of future runoff projections. The contribution of the CM toward total variance in annual runoff change ranges from 73% to 90% in the SEPA region. The interaction between the CM and DS is the second most

dominant contributor in SEPA-1, SEPA-2, and SEPA-3. The CM remains a dominant factor even at the seasonal scale across all four SEPA areas (APPENDIX E).

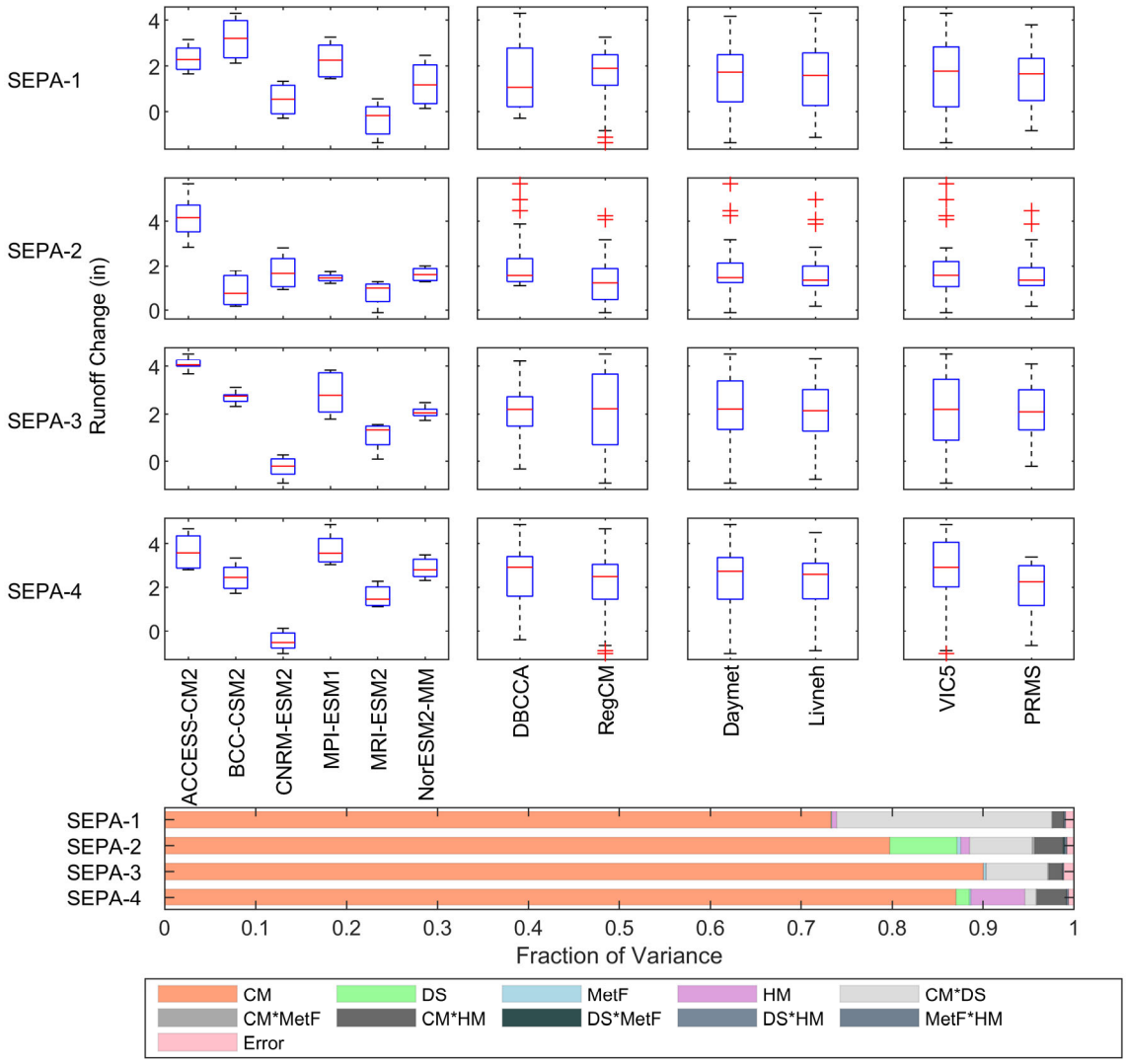


Figure 6.9. ANOVA of annual total runoff in the SEPA region.

The LEM-simulated reservoir evaporation loss from federal hydropower reservoirs in the SEPA region is illustrated in Figure 6.10. The blue and green lines show annual reservoir evaporation loss (in units of million acre-feet) from 24 projections that were either bias-corrected or trained by Daymet and 24 projections that were either bias-corrected or trained by Livneh. The dark blue and dark green lines represent the control simulations driven by 1980–2019 Daymet and 1980–2013 Livneh, respectively. The right panel shows the probability density distributions of annual evaporation loss in both baseline (1980–2019) and future (2020–2059) periods.

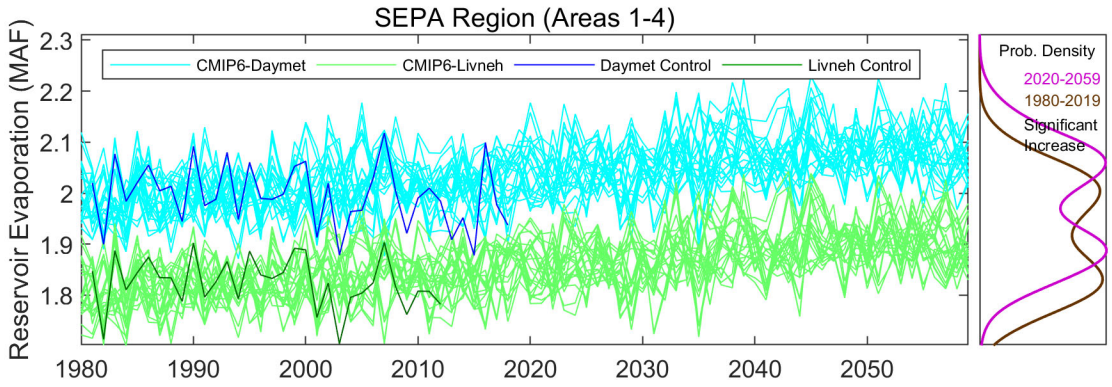


Figure 6.10. Projected annual total reservoir evaporation in the SEPA region.

The results in Figure 6.10 can be partitioned into two groups. The high-value group contains the 24 members that are associated with Daymet, and the low-value group contains the 24 members that are associated with Livneh. The difference is values is caused by wind speed differences between these two data sets, which has the greatest impact on the evaporation rates (Section 3.2.2). A comparison with in situ data (not shown) suggests that the NARR wind speed used with Daymet is more accurate, whereas the NCEP1 wind speed used in Livneh tends to underestimate more considerably (Moreo, 2015). The reservoir evaporation loss also depends on the projected reservoir area, which is inferred from runoff simulated by either VIC or PRMS. The projected runoff associated with Daymet or Livneh is not significantly different because the HMs were calibrated separately by both MetFs against the same USGS WaterWatch data set (Section 2.5.3) to minimize the impacts of different meteorological forcings on HMs. Therefore, the difference in reservoir evaporation loss in the SEPA region is mainly caused by the difference in projected evaporation rate driven by two different MetFs.

Despite the systematic differences between the two groups, the reservoir evaporation loss shows a clear increasing trend (as indicated by the two-sample Kolmogorov-Smirnov test performed at the 5% significance level). The Livneh-based future reservoir evaporation projections increase significantly from 1.8 MAF in 1980 to 1.88 MAF in 2059, and the Daymet-based projections also increase by about the same magnitude from 2 MAF in 1980 to 2.08 MAF in 2059. The average percentage increment is about 3.3% over the 80 years, which is the smallest among all PMA regions. Of the changes in the SEPA reservoir evaporation losses, about 63% were attributed to changes in reservoir area, and 37% were affected by changes in evaporation rate.

The annual and seasonal changes of monthly reservoir evaporation loss (in units of thousand acre-feet per month) from federal hydropower reservoirs in the SEPA region are depicted in Figure 6.11. The trend of increasing reservoir evaporation loss is persistent in both near-term (2020–2039) and mid-term (2040–2059) future periods with respect to the baseline period (1980–2019). The annual ensemble median evaporation loss is projected to increase by 3 KAF/month in the near-term and 6 KAF/month in the mid-term future periods. The ensemble median of spring reservoir evaporation is projected to increase by about 5 KAF/month in the near-term future period, and about 3 KAF/month in the mid-term future period. The projected reservoir evaporation has the smallest change in summer, with the ensemble median projected to increase by 2 KAF/month in the near-term and 3 KAF/month in the mid-term future periods. There is a large increase in fall, with the ensemble median projected to increase by 3 KAF/month in the near-term and 12 KAF/month in the mid-term future periods. The ensemble median of the winter evaporation loss is projected to increase by 4 KAF/month in the near-term and 7 KAF/month in the mid-term future periods.

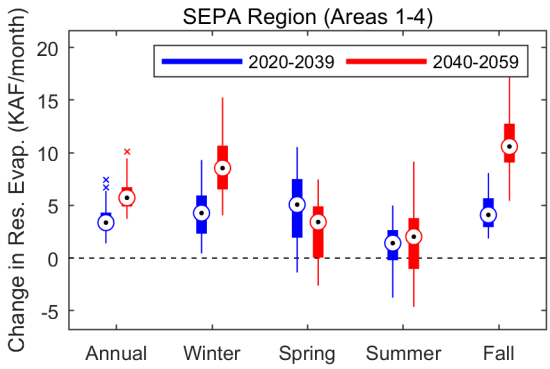


Figure 6.11. Projected change of annual and seasonal total reservoir evaporation in the SEPA region.

The source of variance in the projected change in annual reservoir evaporation loss from federal hydropower reservoirs in the SEPA region is shown in Figure 6.12. The box plots show the magnitude of change (future minus baseline) for the median annual reservoir evaporation loss across 48 ensemble members grouped by each of the four modeling choices (CM, DS, MetF, and HM). The bottom panel shows the portion of variance from the ANOVA for each modeling choice.

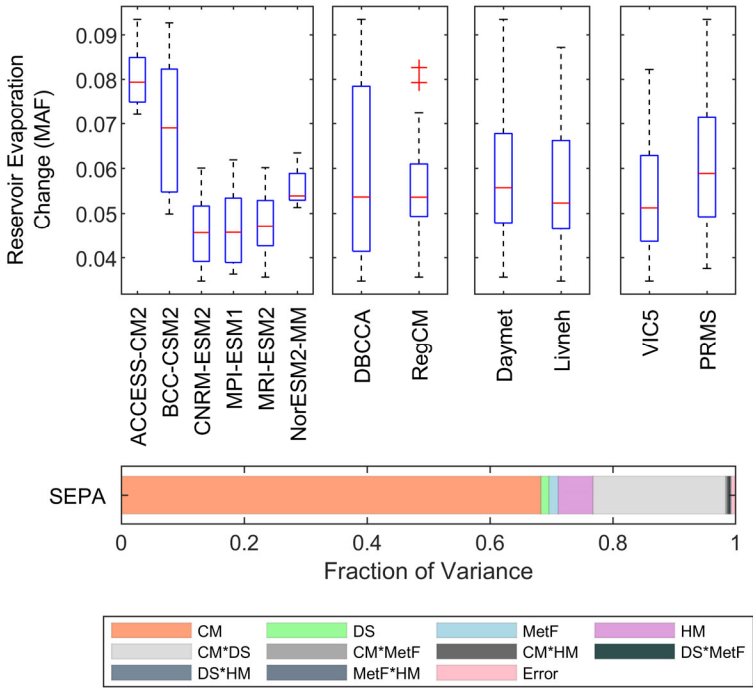


Figure 6.12. ANOVA of annual total reservoir evaporation in the SEPA region.

ANOVA results in Figure 6.12 indicate that the CM is the most dominant contributor for the projected changes of future evaporation loss, which accounts for nearly 70% of the total variance. The most significant change is projected by ACCESS-CM2 (0.08 MAF), whereas the smallest amounts of change are projected by CNRM-ESM2, MPI-ESM1, and MRI-ESM2 (0.045 MAF). The combination of the CM and HM contributes to more than 20% of the total variance. The HM contributes about 7%, and PRMS projects more evaporation than VIC. At the seasonal scale (APPENDIX F), the ANOVA indicates that certain contributing factors may have more substantial impacts than at the annual scale. In spring and winter, the CM explains nearly 90% of the evaporation loss variance. Among these models, ACCESS-

CM2 projects the largest change, whereas MPI-ESM1 projects the smallest. In the summer and fall, the contribution of the CM decreases to around 40%. The combination of the CM and DS is the second most dominant contributor, which can explain 40% of the evaporation loss variance in the summer, and 36% in the fall. Although the choice of MetFs (i.e., Daymet or Livneh) may significantly influence the projected value of reservoir evaporation loss (Figure 6.10), the difference is systematic and largely reduced when examining the change from baseline to future periods. This shows that even with large model/data bias, the relative change is still considered reliable and used in many hydroclimate impact assessment studies.

6.3 CLIMATE EFFECTS ON FEDERAL HYDROPOWER IN THE SEPA REGION

6.3.1 Projections of Hydropower Generation

Using the two hydropower models described in Section 2.6, the projections of monthly generation in the SEPA region are calculated for each of the 96 unique combinations of CMs, DSs, MetFs, and HMs and PMs. Projections based on multimodel ensembles are summarized in terms of the mean annual and seasonal changes for the near-term (2020–2039) and mid-term (2040–2059) future periods compared with the baseline period (1980–2019).

The interannual variability of annual hydropower generation for the baseline and future periods is shown in Figure 6.13. The gray lines show the annual total generation for all 96 members of the ensemble hydropower projections, the black line represents the ensemble median, the black line represents the ensemble median; the blue line shows the 1980–2019 historic observation from EIA and PMAs. The corresponding probability distributions of annual generation in the baseline and future periods are compared in the right panel. A two-sample Kolmogorov-Smirnov test at the 5% significance level was used to determine whether the difference between the baseline and future periods are statistically significant.

The comparison (Figure 6.13, left column) suggests that the range of simulated and observed annual hydropower is generally consistent across all SEPA areas. However, the shapes of the probability distributions revealed that significant changes in interannual variability will occur in some SEPA areas. Specifically, in SEPA-1, SEPA-2, and SEPA-3, the future generations showed wider distributions with lower, rounder, and shifted peaks (Figure 6.13, right column), suggesting that the annual generation in the future period will deviate more frequently and more substantially from a higher median than in the baseline period. The upper tails in SEPA-1 and SEPA-3 indicate more frequent extreme higher generation. For SEPA-4, the general shapes of the probability distributions for baseline and future periods do not substantially differ. However the distribution of annual future projection displays an upper tail not even present in the baseline period, denoting unseen high annual generation in the future.

The annual and seasonal change in hydropower generation is further summarized in Figure 6.14. The change is defined as the percentage difference of future periods (2020–2039 and 2040–2059) compared with the baseline period (1980–2019). Each box plot shows the spread across 96 multimodel combinations; the central mark indicates the multimodel median, and the edges of box indicate the 25th and 75th percentiles. The maximum whisker length is 1.5 times the box height (25th to 75th percentiles). Ensemble members outside of the maximum whisker length are considered as statistical outliers (but should not be excluded in the context of climate change).

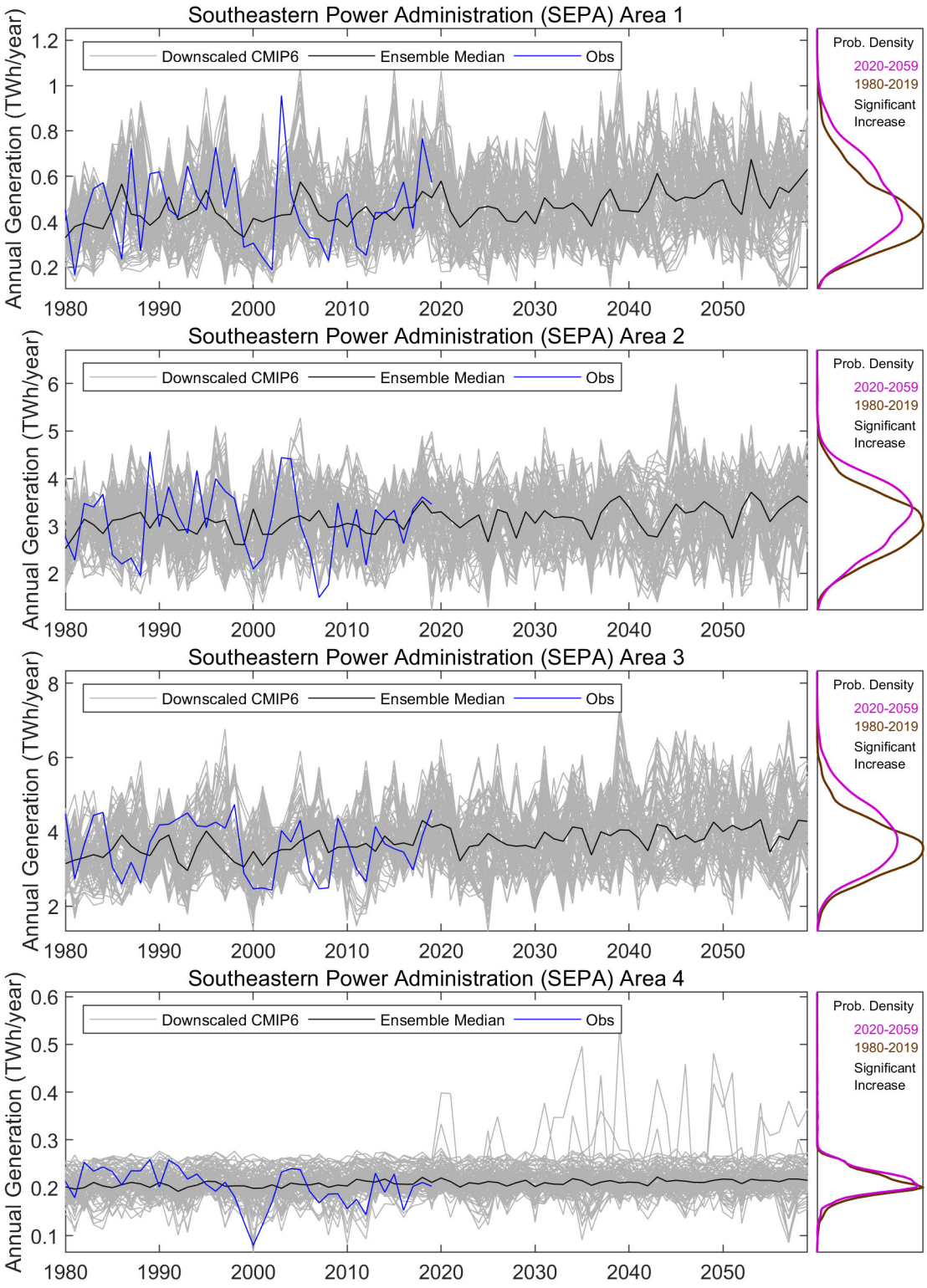


Figure 6.13. Projected annual total generation in the SEPA region.

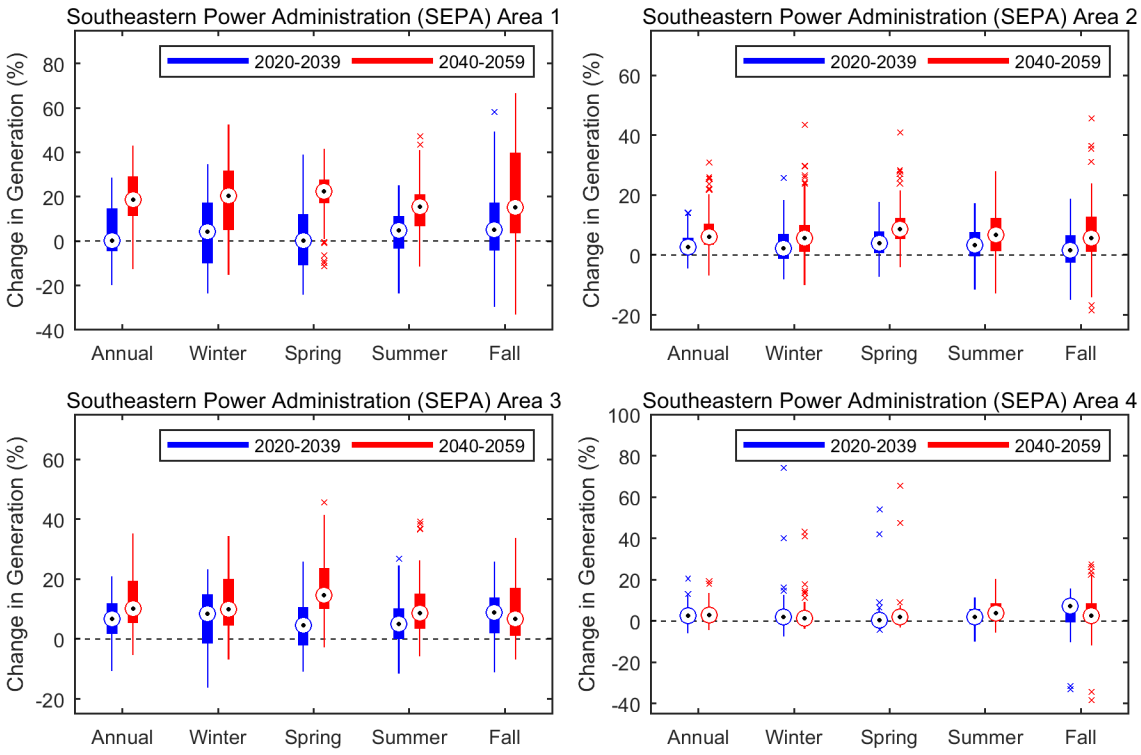


Figure 6.14. Projected change of annual and seasonal total generation in the SEPA region.

The annual hydropower generation is projected to statistically increase in all SEPA areas in the future. More specifically, the annual hydropower generation in both near-term (2020–2039) future periods is projected to remain stable or increase modestly, ranging from 0% in SEPA-1 to 7% in SEPA-3, in terms of multimodel median (Figure 6.14). The increase will be more substantial in the mid-term future period, ranging from 3% in SEPA-4 to 19% in SEPA-1.

Seasonal relative changes tend to be more substantial than annual changes, with a more pronounced increase in the mid-term future period. In winter, generation is projected to increase from 2% to 9% in the near-term future period in all SEPA study areas. The increase is projected to accelerate in the mid-term future period, from 1% in SEPA-4 to 21% in SEPA-1. Similarly, spring generation will increase slightly in the near-term (0%–5%) and more notably in the mid-term (2%–22%) future periods. Summer generation will increase by 2%–5% in the near-term and 4%–16% in the mid-term future periods. Fall generation in the near-term future period will increase by 2%–9%. In the mid-term future period, fall generation will increase significantly in SEPA-1 (15%) and modestly (3%–7%) in other SEPA areas. Overall, the projected changes in SEPA-2 and SEPA-4 are smaller than in other SEPA areas.

Similar to in the SWPA region, the projected changes in seasonal hydropower follow the projected changes in seasonal runoff. The relatively limited storage capacity of reservoirs with respect to inflow in the SEPA region leads to this similar response between runoff and hydropower generation despite the explicit representation of operations by the PMs. However, the spread of seasonal relative change in generation is smaller than that of runoff for SEPA-1 to SEPA-3, where reservoir storage is larger and allows the alleviation of change in runoff variability. Both PMs are constrained with maximum storage capacity and thus a maximum hydropower generation, which also constrain the maximum change.

The source of variance in the projected annual hydropower generation change between the baseline and future periods is shown in Figure 6.15. The box plots in the top panels show the magnitude of change

(future minus baseline) in annual hydropower generation across 96 ensemble members when grouped by the five modeling choices (CM, DS, MetF, HM, and PM). The bottom panel shows the fraction of variance from the ANOVA for each modeling choice.

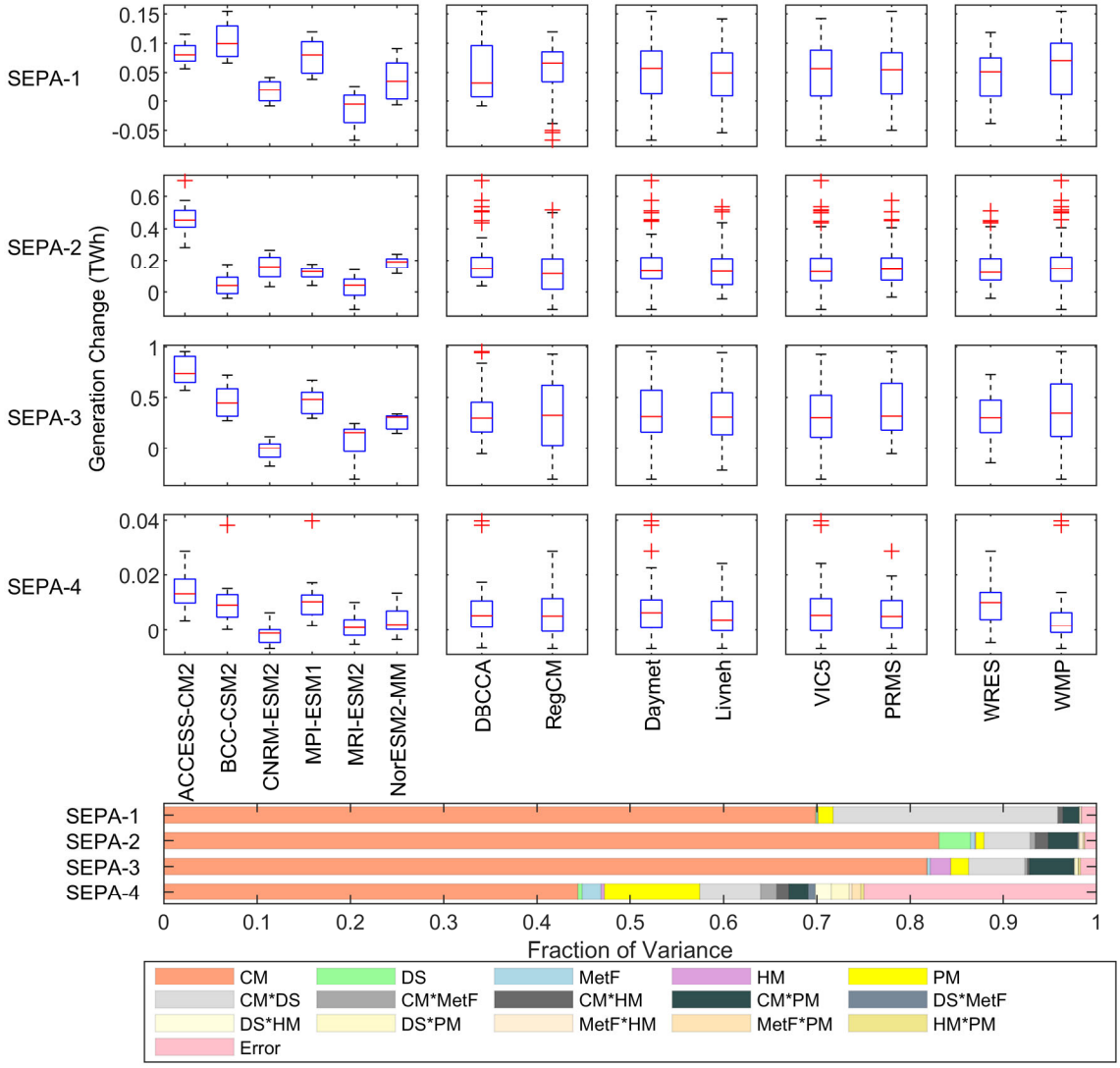


Figure 6.15. ANOVA of annual total generation in the SEPA region.

The diversity in individual factors and their combinations were further evaluated to understand their contributions to the spread of hydropower generation projections. The results suggest that the CM is the most dominant contributor to the variance of annual hydropower generation in all SEPA study areas, ranging from 44% to 83%. None of other factors, including DS, MetF, HM, and PM, alone contributes substantially to the variance. The combination of the CM and DS is the second most dominant contributor to the annual variance in SEPA-1 (24%). In SEPA-4, the PM is the third most dominant contributor to the variance (10%), where a large residual error (25%) is the second most dominant contributor, indicating that part of the projected change cannot be explained by the modeling choices. When grouped by CM, both the median value and the spread of projected generation change vary substantially from one CM to another, indicating the high uncertainty. By contrast, when grouped by DS, MetF, and HM, the median value and the spread of projected generation change show much consistency between groups, indicating

low uncertainty in generation change from these modeling choices in the SEPA region at the annual time scale.

At the seasonal scale (APPENDIX G), the variance contribution from modeling choice varies across different SEPA study areas. For SEPA-1, the CM remains the most dominant contributor to variance throughout the four seasons (48%–70%), and the interactions between the CM and DS are the second most dominant contributor (19%–21%). For SEPA-2 and SEPA-3, the CM in general determines the spread of projected change, with the combination of the CM and PM model also contributing substantially in the spring (29% for SEPA-2 and 15% for SEPA-3). For SEPA-4, residual errors are the most dominant contributor to the variance in the winter (50%), spring (57%), and fall (37%), complicating the interpretation of the projections. In the summer, the CM is the most dominant contributor in SEPA-4 with 56% of the variance, followed by the choice of PM (20%). Results in SEPA-4 are largely influenced by a handful of low storage headwater reservoirs in rainfall-dominated watersheds, explaining the largest sensitivity to residual errors (spills and low variance), and then CMs and PMs, whereas climate was the sole substantial dominant factor for runoff variance (Section 6.2.2).

In summary, federal hydropower in the SEPA region will likely experience modest increase in annual generation in the near-term and more significant increase in the mid-term future periods. Seasonal relative changes in generation are of the same magnitude as annual generation change, with summer generation increasing slightly more than other seasons. Unlike BPA or WAPA regions and similar to the SWPA region, the SEPA region suggests that the median generation changes in the region will follow the median runoff changes closely. In terms of the sources of uncertainty, the CM contributes to most to the variance of hydropower generation at both annual and seasonal scales.

6.3.2 Climate Change Impacts on Regional Electricity Demand

This section summarizes the results from the statistical analysis, exploring the effects of projected temperature trends on the electricity sales of SEPA's preference customers to their end-use (residential, commercial, industrial) customers. As described in Section 6.1.2, SEPA markets the output from 22 hydropower plants to preference customers, mostly municipalities and cooperatives, as peaking power. The contracts between SEPA and its preference customers guarantee a fixed number of hours of energy per kilowatt of capacity contracted per year. Typically, the federal hydropower allocation that preference customers receive from SEPA is a small fraction of the portfolio of generation resources they use to serve their end-use customers. Any load growth projected by SEPA's preference customers, due to climate trends or other drivers, would have to be met through generation sources other than federal hydropower.

Table 6.3 summarizes the elasticities of the preference customer sales with respect to degree days for the sample of 308 SEPA preference customers included in EIA Form 861. The full results of the regression equations are presented in the APPENDIX H.

Electricity sales by SEPA preference customers are positively related to both HDDs (i.e., number and intensity of cold days triggering heating demand) and CDDs (i.e., number and intensity of warm days triggering air conditioning use). The relationship is statistically significant. Depending on the subregion and meteorological observation, the estimated responsiveness of annual sales to a 1% increase in HDDs ranges from 0.08% to 0.15%. For CDDs, a 1% increase leads to estimated increases of 0.05%–0.11% in sales.

The estimated elasticities with respect to degree days were combined with the projected changed in degree days derived from climate simulations to produce an estimate of the temperature-driven changes in demand for SEPA preference customers. Figure 6.16 and Figure 6.17 show the projected degree days

from 24 simulations (the six CMIP-6 models described in Section 2.3 combined with two MetFs and two DSs), as well as the ensemble median projection.

Table 6.3. Estimated responsiveness of electricity sales by SEPA preference customers to degree days

Marketing region	Meteorological observation	Dependent variable	Elasticity with respect to HDDs	Elasticity with respect to CDDs
SEPA-Kerr-Philpott	Daymet	Electricity sales	**0.1526** (0.0173)	**0.0464** (0.0168)
SEPA-Kerr-Philpott	Livneh	Electricity sales	**0.1403** (0.0174)	**0.0605** (0.0179)
SEPA-Cumberland	Daymet	Electricity sales	**0.0825** (0.007)	**0.0781** (0.0085)
SEPA-Cumberland	Livneh	Electricity sales	**0.0837** (0.0074)	**0.0802** (0.0083)
SEPA-GA/AL/SC/JW	Daymet	Electricity sales	**0.1003** (0.0085)	**0.0873** (0.0188)
SEPA-GA/AL/SC/JW	Livneh	Electricity sales	**0.0935** (0.0094)	**0.105** (0.0214)

Note: Numbers in parentheses are standard errors. Bolded values indicate that the estimated elasticity is statistically significant.

During the baseline period (1980–2019), the SEPA-GA/AL/SC/JW subregion had on average more CDDs than HDDs. On average, the ensemble median indicates 1,492 annual HDDs and 2,239 annual CDDs during the baseline period for that region. For the other two SEPA subregions, the situation is reversed (i.e., more HDDs than CDDs during the baseline period in the Kerr-Philpott and Cumberland subregions). However, given the decreasing trend in HDDs and increasing trend in CDDs presented in Figure 6.16 and Figure 6.17, the ensemble median annual CDDs become larger than annual HDDs for all SEPA subregions since the early 2030s until the end of the projection period (2059). Average HDDs (based on the ensemble median trajectory) decrease by 21%–24% and average CDDs increase by 26%–33% relative to the baseline period in the three subregions. For the SEPA-GA/AL/SC/JW subregion in the future period (2020–2059), projected annual CDDs will be more than double the annual HDDs (the average of the ensemble median is 1,136 for HDDs and 2,829 for CDDs). The projected changes in sales driven by these degree day trends are presented in Figure 6.18.

Changes in sales driven by HDDs and CDDs will take place during the winter and summer, respectively. Under any of the 24 CMIP6 projections included in Figure 6.18 and for each of the degree day–period combinations, electricity sales decrease in the winter months and increase in the summer months due to the trends in degree days summarized in Figure 6.16 and Figure 6.17. The sign of the net change (based on the median effects showed in the plot) in annual sales varies across subregions and periods but is small in all the cases. The net change in annual sales driven by changes in degree days ranges from a 1.8% decrease (SEPA-Kerr-Philpott in 2040–2059) to a 1.2% increase (SEPA-Cumberland in 2040–2059).

The capacity and peaking power marketed by SEPA will likely be most valuable to their preference customers during the summer months. SWPA's preference customers have some limited flexibility regarding the time of the year in which they receive their energy allocation (within the limits specified in their contracts) to the extent that their preferred timing is feasible for the PMA given other federal hydropower operational constraints.

As discussed in Section 6.3.1, generation is expected to increase at the annual time scale across all the hydropower fleet marketed by SEPA. The increase in generation could help curb the trend of increased replacement power purchases experienced in recent years and have a positive effect on power rates.

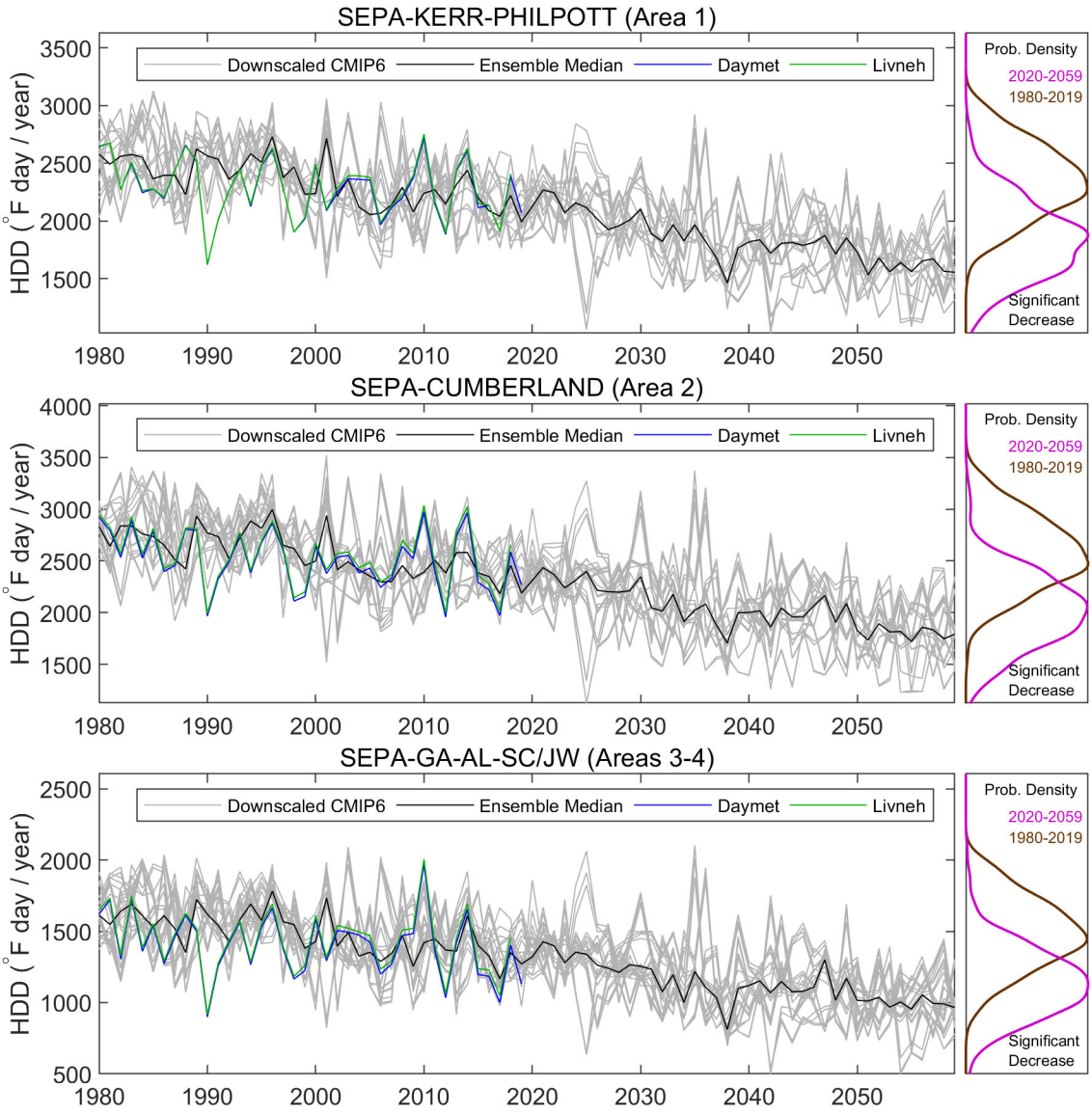


Figure 6.16. Projected annual HDDs in the SEPA region.

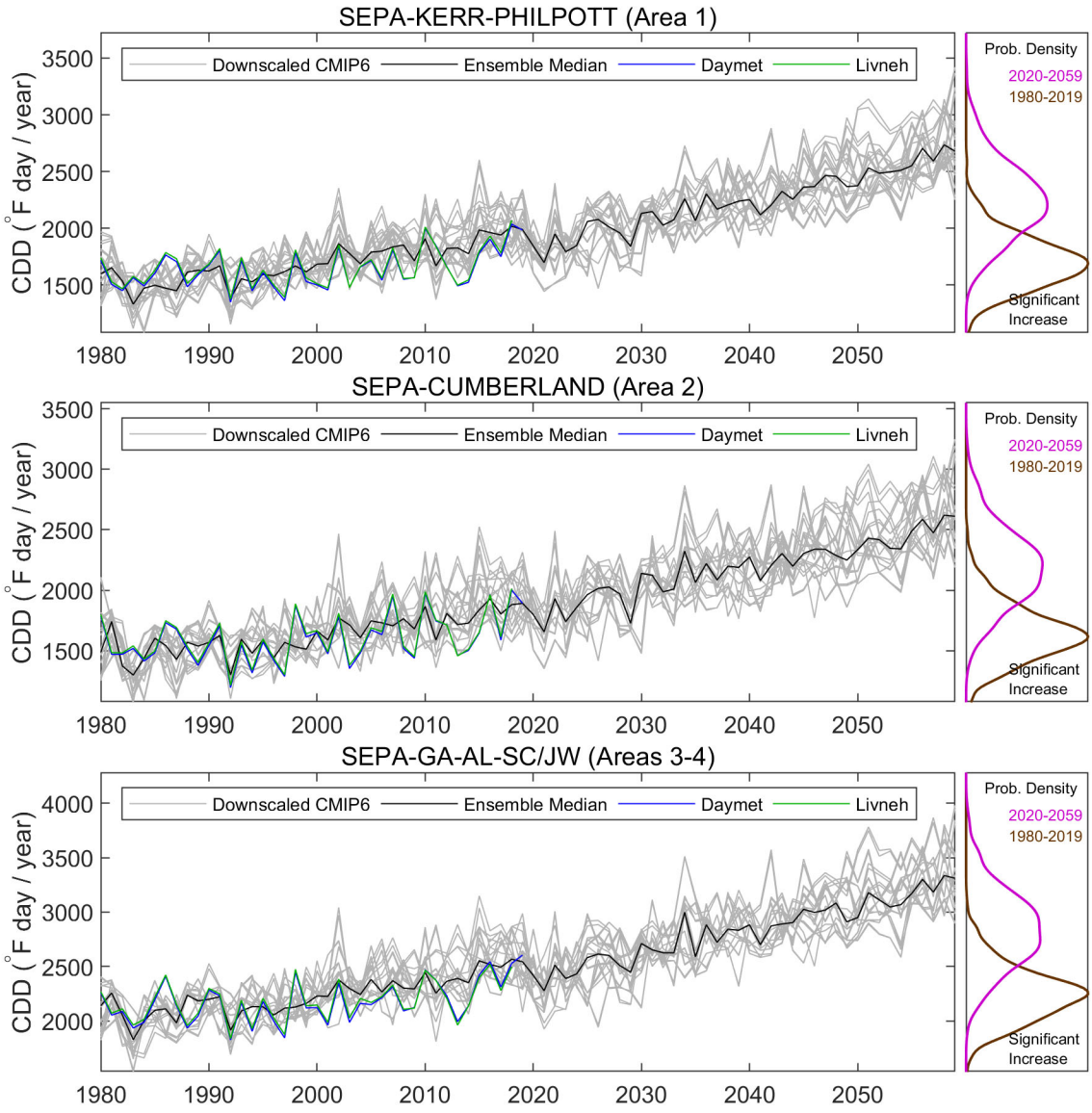


Figure 6.17. Projected annual CDDs in the SEPA region.

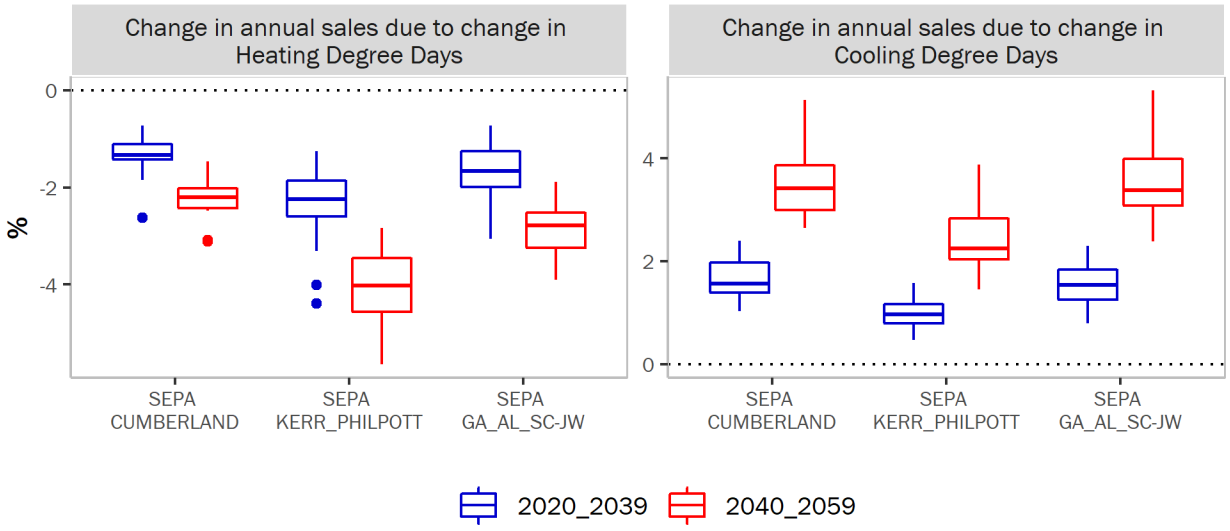


Figure 6.18. Estimated changes in SEPA preference customer electricity sales relative to the baseline period due to projected changes in degree days.

7. INTERREGIONAL COMPARISON

Apart from PMA-specific regional analysis (Sections 3 to 6), some further interregional comparisons are provided in this section to help demonstrate the overall trends across the CONUS. When possible, the comparison covers the entire CONUS, not limited to the four PMA regions. Therefore, several insights provided in this section can be broadly interpreted for the entire US hydropower fleet (both federal and nonfederal), as well as for other non-hydropower sectors.

7.1 CLIMATE PROJECTIONS

In addition to the PMA-based analysis, the projected changes (2020–2059 minus 1980–2019) in mean and extreme climate variables are summarized for the nine National Center for Environmental Information (NCEI) regions in Figure 7.1. Heat maps were used to present the projected changes in daily maximum temperature (T_{max}), 95th percentile of T_{max} (T_{95}), number of days with T_{max} above historical T_{95}, daily minimum temperature (T_{min}), 5th percentile of T_{min} (T_{05}), number of days with T_{min} below historical T_{05}, number of frost days (T_{min} below 0°C), daily mean precipitation (P), 5th percentile of P (P_{95}), number of days with P above historical P_{95}, and number of wet days (P above 1 mm) for the ensemble mean of six models from each of the following data sets: CMIP6 GCMs (GCM_O), dynamically downscaled RegCM output (RCM_O), RegCM output bias-corrected using Daymet (RCM_{CD}) and Livneh (RCM_{CL}), and GCMs statistically downscaled using Daymet ($DBCCA_D$) and Livneh ($DBCCA_L$) (Figure 7.1).

7.1.1 Projected Changes in Mean and Extreme Daily Maximum Temperature

All six ensembles of climate change projections showed statistically significant increases in the magnitudes of T_{max} and T_{95} throughout the NCEI regions, with the strongest changes in the Upper Midwest and Northern Rockies regions (Figure 7.1a, b). When compared with the driving GCM ensemble (GCM_O), the corresponding statistically downscaled ensembles ($DBCCA_L$ and $DBCCA_D$) projected generally similar changes, while the dynamically downscaled RCM_O, and its corresponding bias-corrected ensembles (RCM_{CL} and RCM_{CD}), projected comparatively smaller future period changes in T_{max}. The smallest (0.09°F) and largest (up to 0.49°F) differences in projected changes in T_{max} between GCM_O and RCM_O, and RCM_{CL} and RCM_{CD} were over the western and northeastern United States (Figure 7.1a). Similarly, the future period changes in T_{95} projected by $DBCCA_L$ and $DBCCA_D$ were also comparable with the GCM_O, whereas the changes projected by RCM_O, RCM_{CL}, and RCM_{CD} were lower (by 0.49°F–1.3°F) than the GCM_O ensemble (Figure 7.1b). The downscaled ensembles showed marked differences from the GCM_O ensembles in the future period changes in the frequency of T_{max} extremes (number of days above the historical T_{95}). As compared with GCM_O, RCM_{CL} and RCM_{CD} projected a comparable increase, RCM_O projected a lower increase (by up to 8 days), and $DBCCA_L$ and $DBCCA_D$ projected a much higher increase (by up to 11 days) in the frequency of T_{max} extremes across most regions (Figure 7.1d). To summarize, the dynamically downscaled RCM_O, and its corresponding bias-corrected ensembles (RCM_{CL} and RCM_{CD}), showed larger differences (smaller increases) in the simulated magnitude of change in mean T_{max} and T_{95} than those shown by statistically downscaled ensembles ($DBCCA_L$ and $DBCCA_D$). However, RCM_{CL} and RCM_{CD} projected similar increases in the frequency of extremes, whereas $DBCCA_L$ and $DBCCA_D$ had significantly higher increases as compared with the GCM (Figure 7.1).

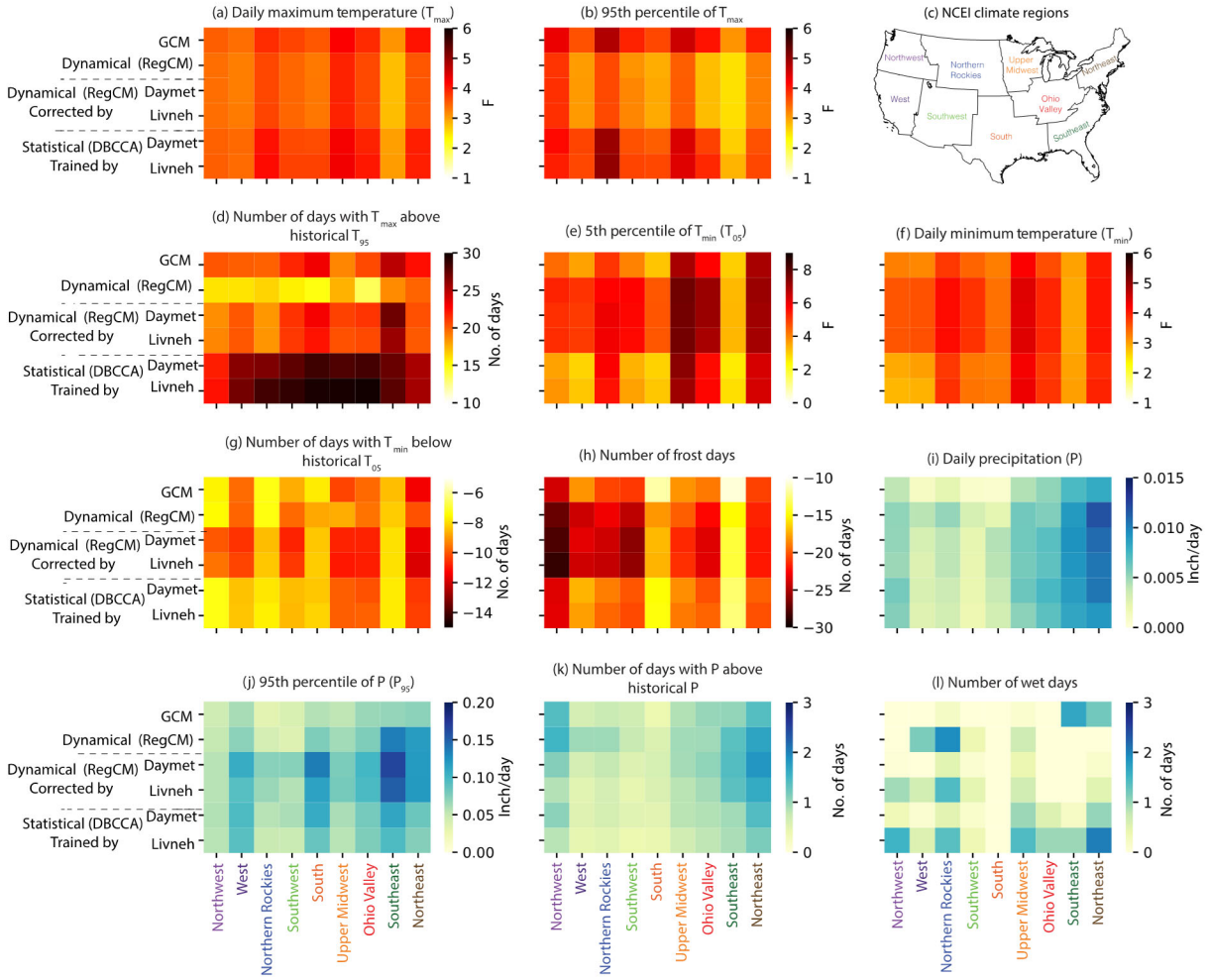


Figure 7.1. Projected changes in climate variables across the nine NCEI regions.

7.1.2 Projected Changes in Mean and Extreme Daily Minimum Temperature

All six ensembles projected statistically significant increases in T_{min} and T_{05} and projected decreases in the number of days below historical period T_{05} and in the number of frost days over all the regions (Figure 7.1e–Figure 7.1 h). When compared with GCM_O, contrary to the future changes in T_{max}, the dynamically downscaled RCM_O, and its corresponding bias-corrected ensembles (RCM_{CL} and RCM_{CD}), projected higher future period warming in T_{min} over most regions, with the largest differences (up to 0.41°F) in the western United States (Figure 7.1 e). These ensembles also projected higher increases in T_{05} than the GCM_O, with the largest differences in warming over the western and southwestern United States (Figure 7.1). On the other hand, similar to the future changes in T_{max}, $DBCCA_L$ and $DBCCA_D$ exhibited future changes in T_{min} comparable to GCM_O over most regions, except in the western and northwestern United States, where differences in the magnitude of changes in T_{min} between GCM_O and the corresponding statistically downscaled ensembles were higher (by 0.18°F–0.41°F) (Figure 7.1e). The magnitude of future period changes in T_{05} was lower in $DBCCA_L$ and $DBCCA_D$ compared with the GCM_O, except over the Northern Rockies and Upper Midwest (Figure 7.1f). Furthermore, there were differences between $DBCCA_L$ and $DBCCA_D$ in projected warming in T_{min} and T_{05} over most regions. These differences can be partially attributed to the larger differences between Livneh and Daymet for T_{min} in the historical period.

Finally, the projected changes in the number of days below historical T_{05} were generally comparable across all the ensembles (Figure 7.1g). However, the projected decrease was higher in the number of frost days in RCM and its bias-corrected ensembles (RCM_O, RCM_{CL} and RCM_{CD}) (by 1–5 days) than in statistically downscaled ensembles ($DBCCA_L$ and $DBCCA_D$) (by up to 3 days) in comparison to the GCM_O (Figure 7.1h).

7.1.3 Projected Changes in Mean and Extreme Precipitation

All ensembles projected significant increases in P, P_{95}, and number of days above historical period P_{95} over most of the regions. The future period projected increase was higher over the southeastern and northeastern United States. On the other hand, the projected changes in wet days were small and not significant, implying that the future period increase in P is a result of an increase in precipitation intensity and P_{95}. The GCM_O ensemble exhibited the smallest increase in P_{95} compared with the other ensembles (Figure 7.1i–l). Furthermore, the choice of downscaling techniques and the observations used for training and correction affected the projected future period changes in the magnitude of P_{95}. For instance, over the southeastern and northeastern United States, the dynamically downscaled and corrected ensembles (RCM_{CL} and RCM_{CD}) projected a higher increase in the magnitude of P_{95} and the number of days above historical P_{95} than the statistically downscaled ensembles ($DBCCA_L$ and $DBCCA_D$). Similarly, the Daymet ensembles (RCM_{CD} and $DBCCA_D$) projected a higher increase in both metrics compared with Livneh ensembles (RCM_{CL} and $DBCCA_L$) over the southeastern and northeastern United States (Figure 7.1j).

7.2 WATER AVAILABILITY

In addition to the PMA-based analysis, an overview of future water availability across the CONUS is presented using the ensemble hydrologic projections. Figure 7.2 shows changes in annual total, high, and low runoff for near-term (2020–2039) and mid-term (2040–2059) future periods with respect to the baseline (1980–2019) period. Panels (a) and (b) indicate that the simulated total annual runoff is projected to increase across the CONUS in both future periods, with higher magnitude in the mid-term future period. For instance, such a change is more prominent in California, the Great Basin, and the lower and upper Colorado regions. Amongst PMAs, large annual runoff increase is projected in most of SEPA and WAPA. Similar conclusions cannot be made for BPA and SWPA, given the more heterogenous response in the region. The high runoff shown in panels (c) and (d) demonstrates a more spatially heterogeneous response across the CONUS. In the near-term future period, high runoff is projected to increase more in the northern and southeastern United States. However, in the mid-term future period, high runoff is generally projected to increase across the CONUS, except in parts of the Pacific Northwest and Texas. The low runoff shown in panels (e) and (f), on the other hand, is projected to increase in most of the western United States except the Pacific Northwest, but is projected to decrease for Ohio, the Mid-Atlantic, lower Mississippi, and parts of the upper Mississippi regions, which suggests the increased likelihood of hydrologic extreme conditions in those regions.

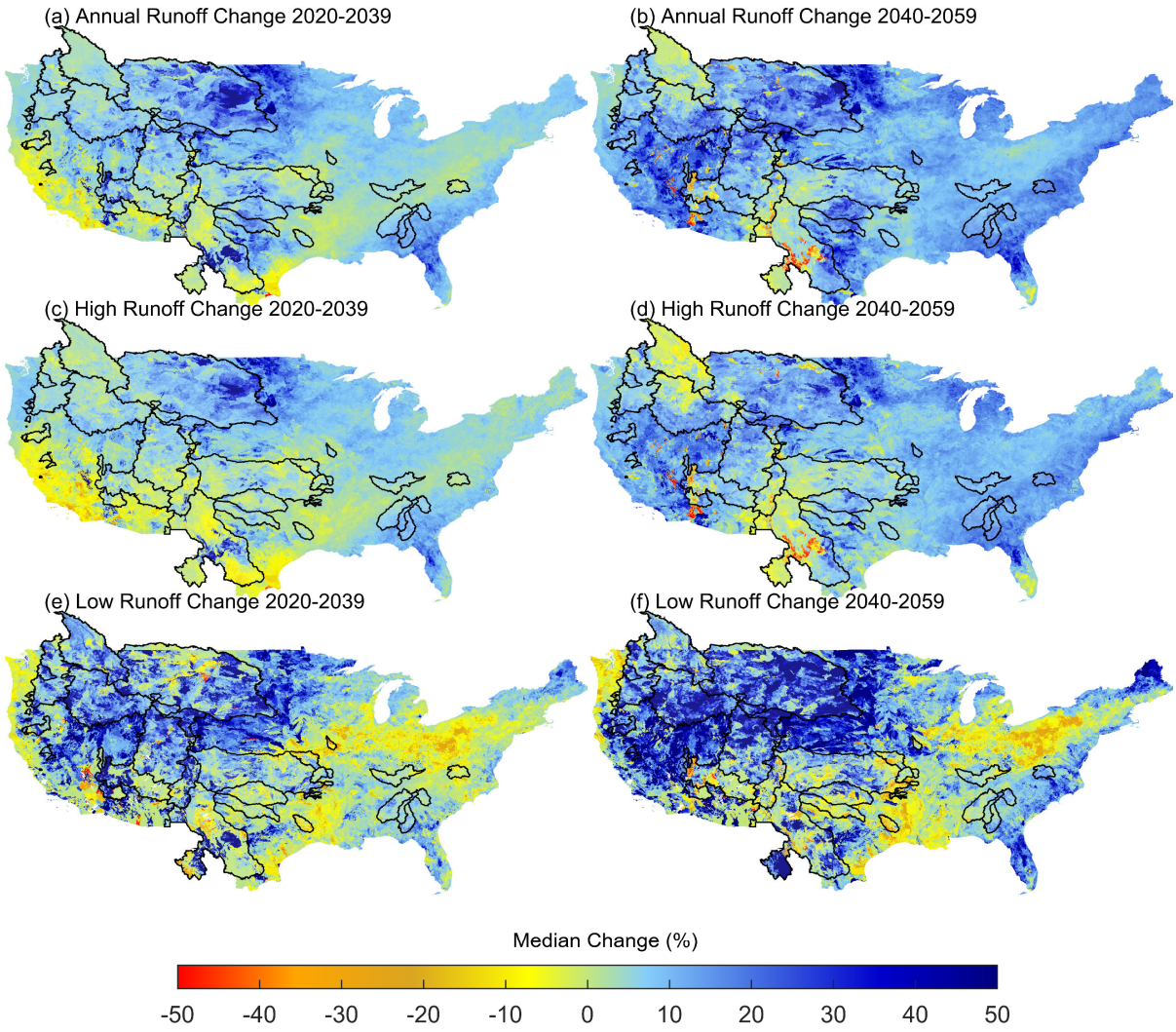


Figure 7.2. Projected change in annual total, high, and low runoff across the CONUS (compared with the 1980–2019 baseline historical period).

At the seasonal scale, the winter runoff is generally projected to increase across most of the CONUS (Figure 7.3), especially in the mid-term future period. The spring runoff is also projected to increase in most of the CONUS but with decreases projected in California, lower Colorado, and the Rio Grande region in the near-term future period. During summer, the runoff is projected to decrease across many parts of the CONUS. The most evident decrease is projected in the Pacific Northwest, California, Rio Grande, Texas-Gulf, Arkansas–White–Red regions, and parts of the northeastern United States and Florida. The changes in fall runoff are spatially heterogenous and expected to stay around the ±10% range across most parts of the CONUS, except in Ohio, lower Mississippi, and parts of upper Mississippi and Missouri regions where a larger reduction is projected.

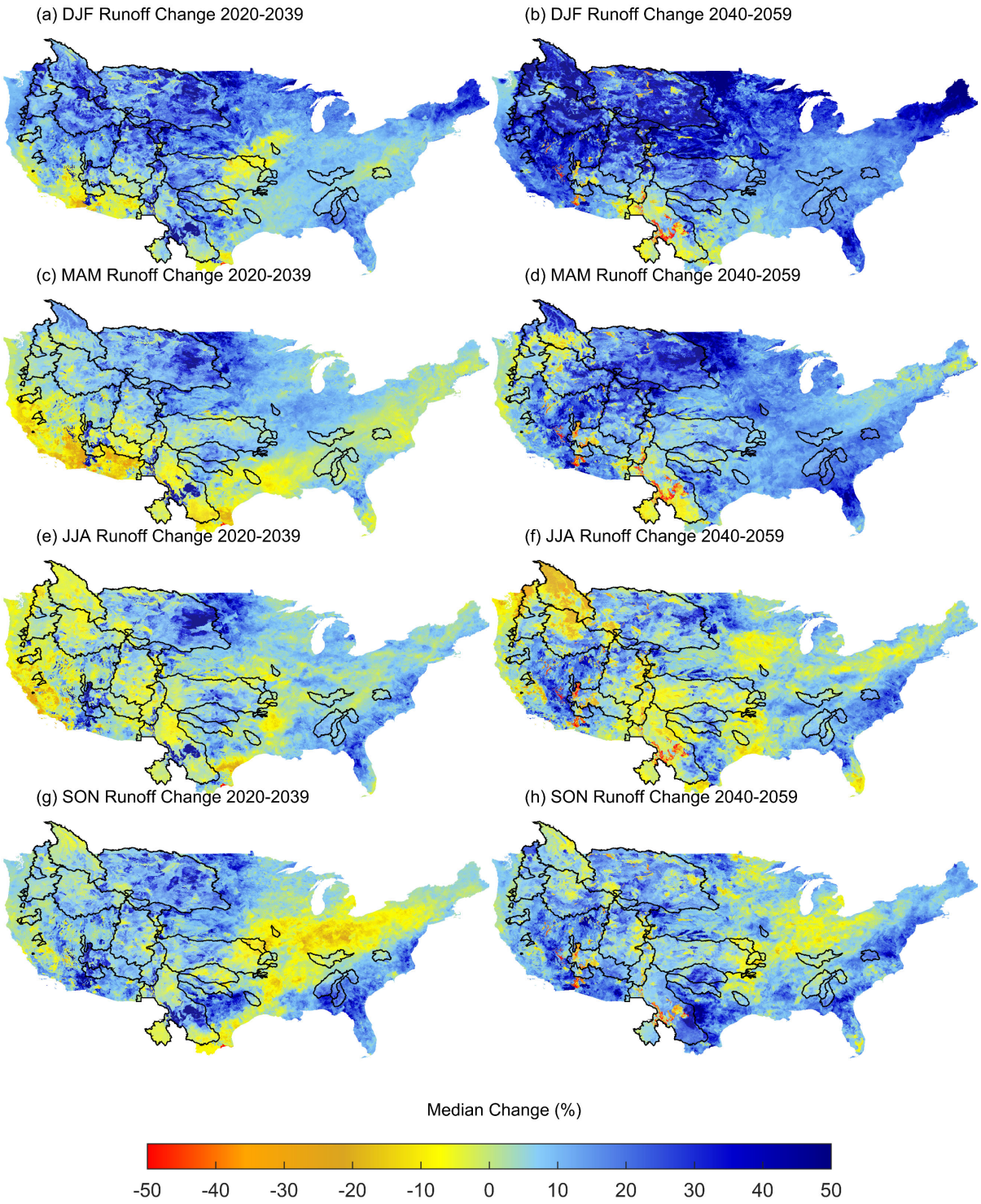


Figure 7.3. Projected change in seasonal runoff across the CONUS (compared with the 1980–2019 baseline historical period).

Figure 7.4 summarizes the results from the ANOVA of the total annual runoff change in the future period (2020–2059) with respect to the baseline period (1980–2019). The figure illustrates the most dominant contributor among all modeling choices (i.e., CM, DS, and HM) to the spread of ensemble runoff change across 48 members. The results are summarized at the HUC8 level, and the layer transparency indicates the percent contribution of the most significant factor to the total variance (i.e., fully transparent indicates 0% contribution, fully opaque indicates 100% contribution). Based on this analysis, the CM emerges as the most dominant factor across large parts of the CONUS, especially in the eastern United States, California, and Pacific Northwest. On the other hand, the HM is more dominant across several parts of the central United States, including portions of the Pacific Northwest, Missouri, upper Colorado, and upper Mississippi. The choice in DS, while only significantly influential for selected HUC8s, can play a significant role when combined with CMs based on high contributions from interaction terms of an ANOVA.

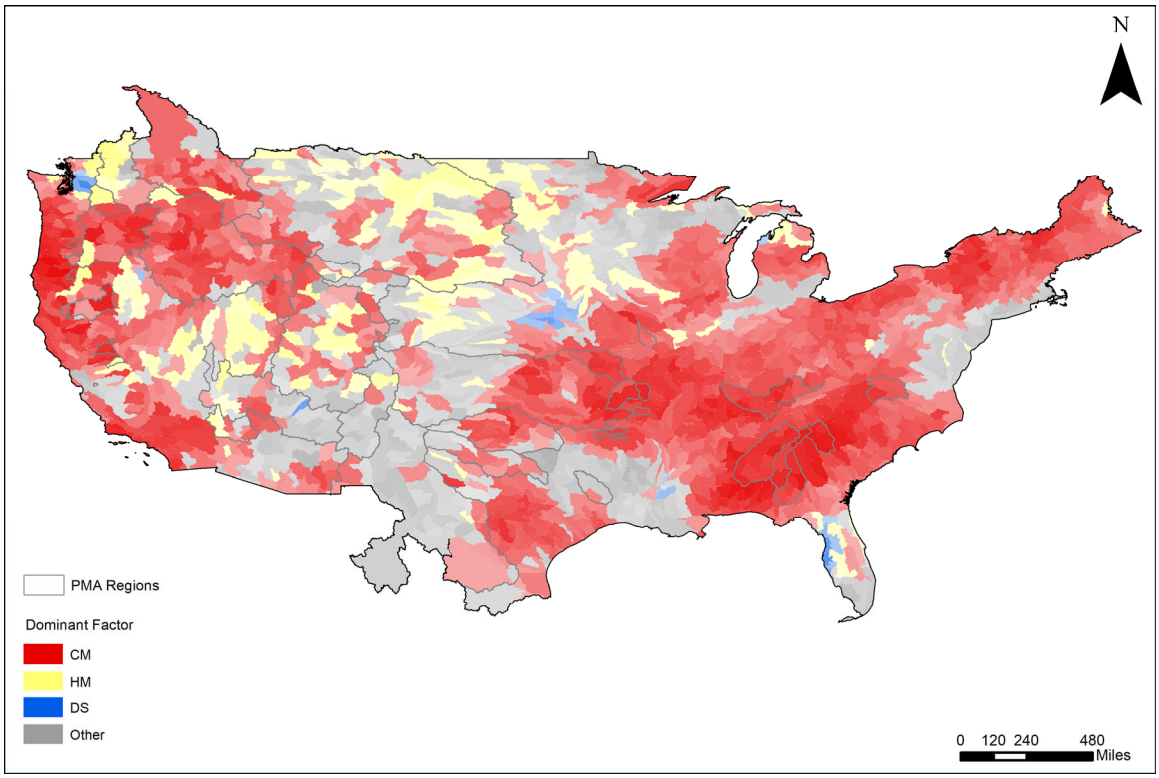


Figure 7.4. ANOVA of annual total runoff in the CONUS. The most significant factor for each HUC8 is plotted. The transparency indicates the contribution (%) of the most significant factor to the total variance (i.e., fully transparent indicates 0% contribution, fully opaque indicates 100% contribution).

7.3 RESERVOIR EVAPORATION

For an analysis of reservoir evaporation, 678 major reservoirs in the CONUS were selected, each with a storage capacity greater than 0.08 MAF (i.e., 0.1 km^3). The future reservoir evaporation losses show different degrees of increments with distinct spatial patterns (Figure 7.5). In general, more severe evaporation losses are projected to occur during the mid-term future period (2040–2059) than the near-term future period (2020–2039). In the near-term future period, reservoirs with the largest changes are mainly located in the south-central CONUS (mainly the southern Missouri, western Arkansas–White–Red, and western Texas-Gulf), whereas the rate of increase will be low in most of the eastern and western CONUS. In the mid-term future period, evaporation loss growth rates in the south-central United States will remain the largest. In addition, the changes of evaporation loss in the western United States are also

significant (6% on average). The growth rate in the eastern United States (especially in the Southeast) will be relatively low by comparison.

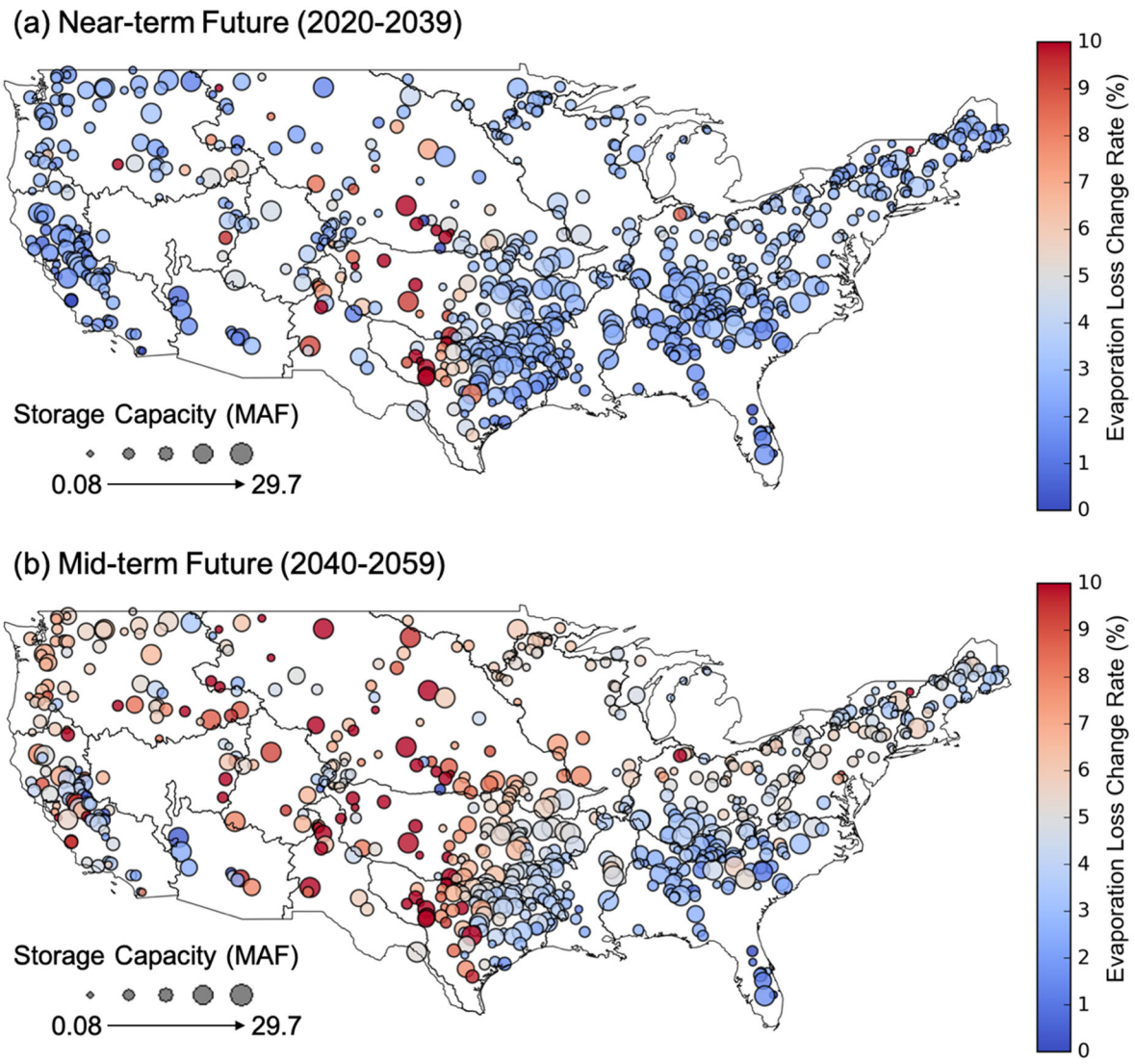


Figure 7.5. Changes of mean evaporation loss (in %) of 678 CONUS reservoirs.

Compared with the baseline period, almost all reservoirs (except for four) show increased evaporation loss during the near-term future period—with an average growth rate of 3.4%. For 84% of these reservoirs, the increased values are all less than 4%. Over these reservoirs, the change is primarily caused by the increased evaporation rate. The increase of evaporation loss is more severe in the south-central United States, mainly southern Missouri, western Arkansas–White–Red, and the western Texas-Gulf. The changes in this region are caused by the combined effects of changing evaporation rate and area. Multiple short-term extreme precipitation events are projected to occur in the south-central region in the future. The elevated regional inflow (owing to such extreme events) will cause the reservoir area to reach its capacity during these future periods, which will exacerbate the evaporation losses. In the mid-term future period, the average evaporation loss growth rate is 5.6%, and the spatial heterogeneity of evaporation loss growth is more significant than near-term future (Figure 7.5b). The evaporation loss growth rate in the

arid/semi-arid western United States is 6.1% compared with the baseline period, and much higher than the 2.9% in the eastern part, which is expected to worsen hydrologic drought conditions in the future.

7.4 FEDERAL HYDROPOWER GENERATION

7.4.1 Projected Changes in Annual Hydropower Generations

The total annual hydroelectric energy generated from each PMA region and from all PMA regions as a group is summarized in Figure 7.6 for near-term and mid-term future periods. To show the multimodel variability, each horizontal bar in Figure 7.6 is a composite of 96 smaller bars for each hydropower projection, sorted in descending order, from top to bottom. For comparison, the multimodel baseline (1980–2019) average is marked by the bold, vertical black line. The projected annual hydropower generation at the PMA level is summarized in Figure 7.6. Key findings are listed below:

- At the annual scale, the total annual hydropower generation for all PMAs is projected to be around 120 TWh/year in both near-term (2020–2039) and mid-term (2040–2059) future periods, which is close to the baseline (1980–2019) level.

- Most PMAs feature a wider multimodel spread in the mid-term future period than in the near-term future period with no obvious change in the ensemble mean.

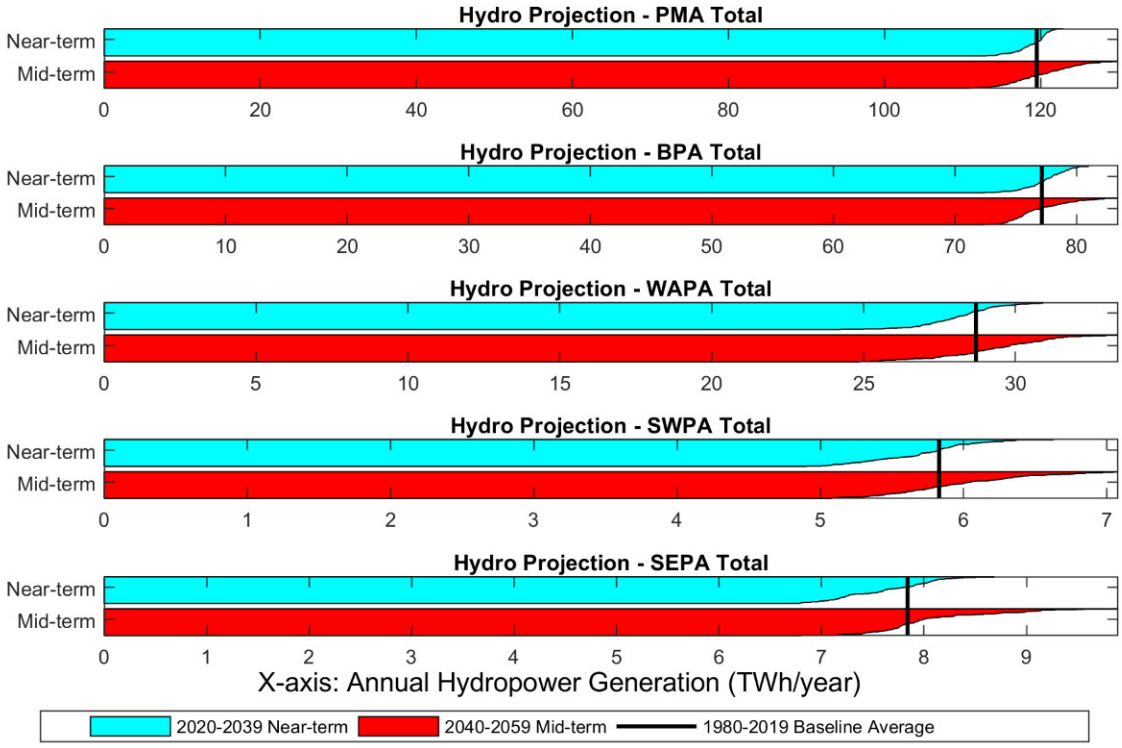


Figure 7.6. Distribution of annual hydropower generation by PMA. Each horizontal bar is a composite of 96 smaller bars for each hydropower projection, sorted in descending order, from top to bottom.

While the multimodel spread is relatively small for the BPA, it becomes larger for the WAPA and even larger for SWPA and SEPA. Apart from the influence of different regional climate projections, it can also be explained by their differences in reservoir features. The reservoirs in BPA and WAPA are relatively large and may provide multiple-year storage, whereas the reservoirs in SWPA and SEPA are relatively

small and heavily influenced by rainfall variability. To verify this point, the change of natural streamflow and hydropower generation is compared in the following subsection.

7.4.2 Propagation of Projected Annual and Seasonal Changes from Natural Streamflow to Hydropower Generation

The projected relative changes in hydrology (natural streamflow at hydropower plants) and hydropower generation are compared at annual and seasonal scales across the PMAs (Figure 7.7) to understand the hydropower response to hydrologic changes and the implications for water management. The changes are scaled to the historical (1980–2019) median level (zero on y-axis).

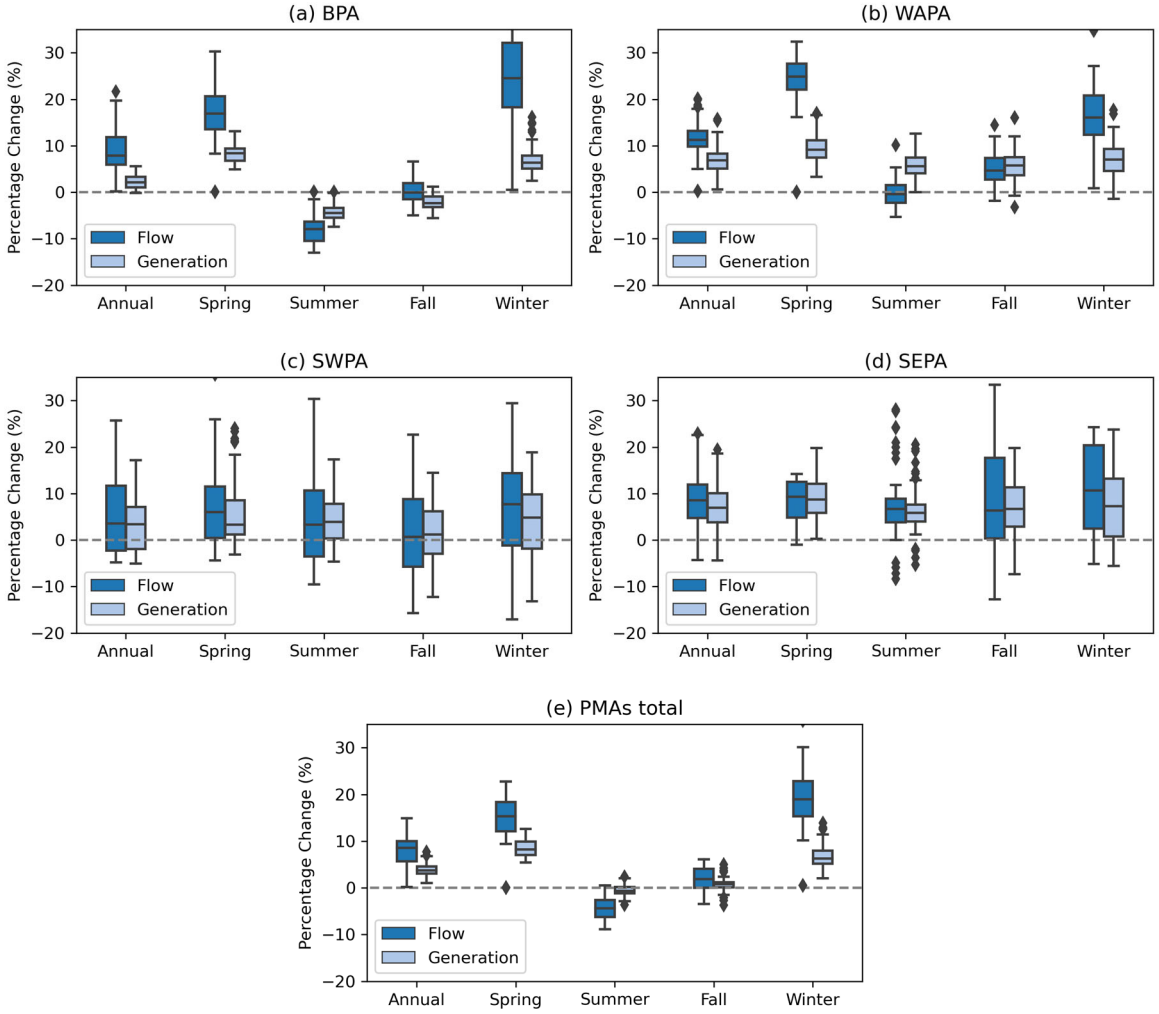


Figure 7.7. Projected relative changes in natural flow and hydropower generation at annual and seasonal scales across PMAs.

Aggregated over all PMAs, the results suggest that annual streamflow is projected to increase by about 9%, and hydropower generation is projected to increase by about 4%. More specifically, the BPA annual natural streamflow is projected to increase by 8%, while hydropower is projected to increase by 2%. In WAPA, streamflow is projected to increase 11%, while hydropower is projected to increase by 7%. In SWPA, streamflow is projected to increase 4%, while hydropower is projected to increase by 3%. In

SEPA, streamflow is projected to increase 9%, while hydropower is projected to increase by 7%. At the seasonal scale, streamflow and hydropower are projected to increase in most seasons and regions, except summer in BPA, where the flow is predicted to reduce by 8%, which is projected to lead to a 4% decrease in summer hydropower generation. Other key findings are listed below:

- The spread of relative changes (within the 25th–75th percentiles) in annual and seasonal hydropower projections are lower than the spread of relative changes in natural flow across all PMAs. Most of these federal hydropower plants are located at multipurpose reservoirs that also provide nonpower services that have higher priority than hydropower. Although those constraints are not individually and explicitly represented in the modeling approaches, the reservoir storage may help reduce the spread in hydropower projections.

- The spreads of annual and seasonal relative changes in both natural flow and hydropower are generally smaller in BPA and WAPA than in SWPA and SEPA. This may be influenced by the different reservoir features in these PMA regions.

- The seasonality of relative changes in both natural streamflow and hydropower is, however, very pronounced in the western United States (BPA and WAPA), with substantial increases projected in winter and spring and smaller changes in summer and fall. Comparatively speaking, the seasonality in the southern and eastern United States (SWPA and SEPA) remains similar, although it might be hidden by the much larger spreads in seasonal projections.

7.4.3 Modeling Choices Controlling Annual and Seasonal Hydropower Generation

Figure 7.8 summarizes the main modeling choices controlling the simulation of hydropower generation at annual and seasonal scales across all PMA study areas. Three key components were the focus here: CM, HM, and PM. Key takeaways include the following:

- The choice of CM dominates the total variance over the eastern United States: in SWPA and SEPA, the choice of CM is the largest source of methodological uncertainty explaining over 65% of variance in hydropower generation, except for SWPA-4 and SEPA-4 (around 50%).

- The choice of HM and PM can be a critical factor over the western United States: in BPA and WAPA, the choice of CM still plays a major role, but the choice of HM and PM becomes more important. In WAPA-6, the HM is the top contributor, whereas in BPA-3 and WAPA-5, the PM is the top contributor to the annual uncertainty. However, in both WAPA and BPA regions, the contribution of uncertainty is much more complicated when the dominant factor changes with the season. For example, in BPA-1, the three components play equal roles in spring, while in summer, the choice of PM becomes the single dominant factor. In the winter, the most dominant contributor switches to the CM. In BPA-4, the CM is the leading contributor to variance annually and in all individual seasons except for summer, when the HM accounts for over 50% of the variance.

- Apart from the CM, HM, and PM, in some cases, the combined variance from other sources (i.e., the DS, MetF, and interaction terms) can be greater than 50%. Thus, the methodological uncertainty of hydropower projections can be influenced by a variety of different factors. The physical reasons causing these differences are to be further explored.

These findings suggest that, although the CM is generally the most dominant contributor to the variance of hydropower projections across PMAs, the HM and PM also have impacts in different areas at seasonal scales. The multimodel approach helps evaluate the structural uncertainty throughout the toolchain used in this report.

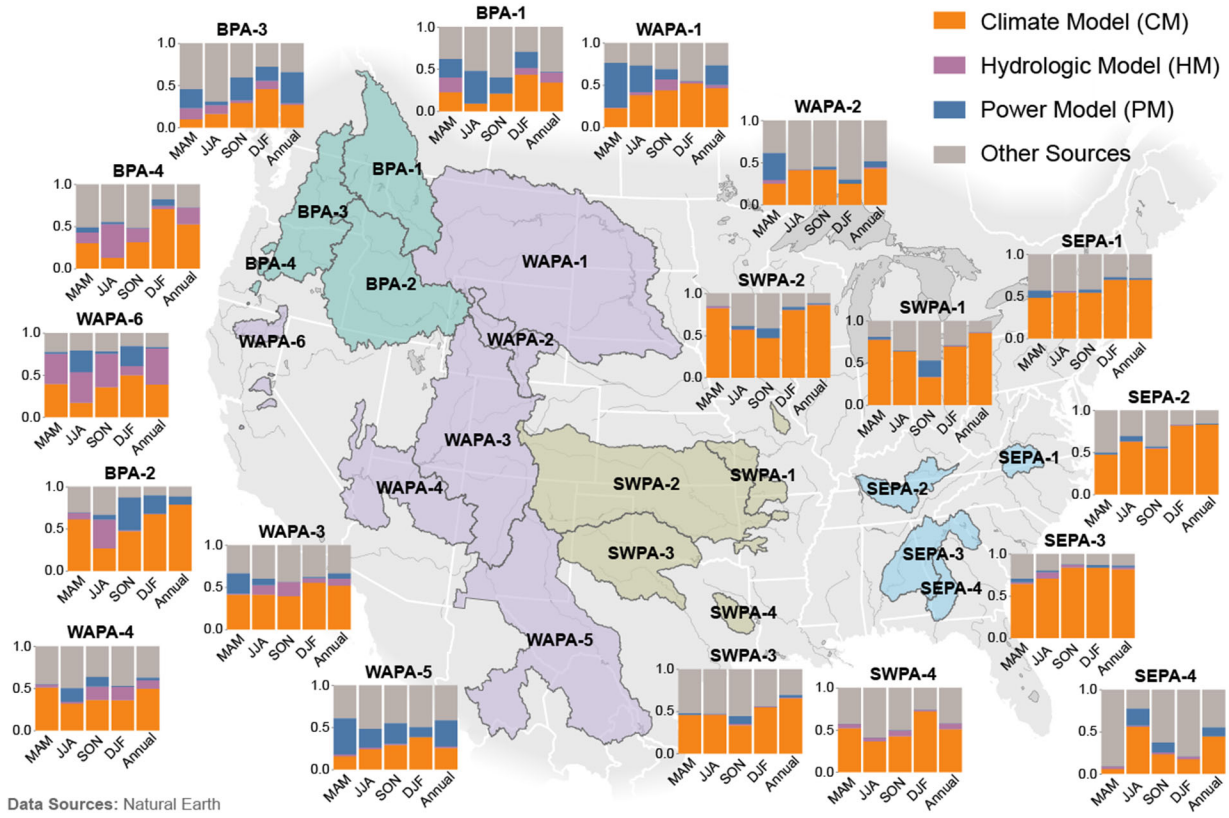


Figure 7.8. ANOVA of projected annual and seasonal hydropower generation in each of the PMA study areas.

7.5 POWER MARKETING

Average temperature is expected to increase in all PMA regions. Increases in average temperature result in a lower number of HDDs and a higher number of CDDs. A lower number of HDDs results in less electricity demand for heating, and a higher number of CDDs results in more electricity demand for cooling. Figure 7.9 compares the projected effects of degree day changes on the annual electricity sales (as a proxy for load) of PMA preference customers across all PMA marketing regions in the near-term and the mid-term future periods. For BPA, the annual changes were calculated as the weighted average of the seasonal changes discussed in Section 3.3.2 and account for increases in AC saturation.

Notably, the percentage changes in sales shown in Figure 7.9 would not be uniformly distributed throughout the year; the temperature-related effects on sales would be concentrated in the winter and summer months. Therefore, the percentage change on sales driven by degree days would be significantly larger than the values shown in Figure 7.9 in some months and smaller in others.

The top-line message from Figure 7.9 holds for PMA marketing subregions located in very different geographic regions, from the Pacific Northwest to the Southeast: the projected increase in average temperature, which yields a lower number of HDDs in the winter months and a higher number of CDDs in the summer months, results in projected decreases in the electricity sales of PMA preference customers to their end customers in the winter and increased sales in the summer. Differences in the magnitude of the changes across PMA marketing regions are due to many factors, including their geographic location, which influences the baseline values of HDDs and CDDs, and the mix of loads served by preference

customers in each region. Preference customers with large shares of residential loads tend to have the largest response of sales to temperature.

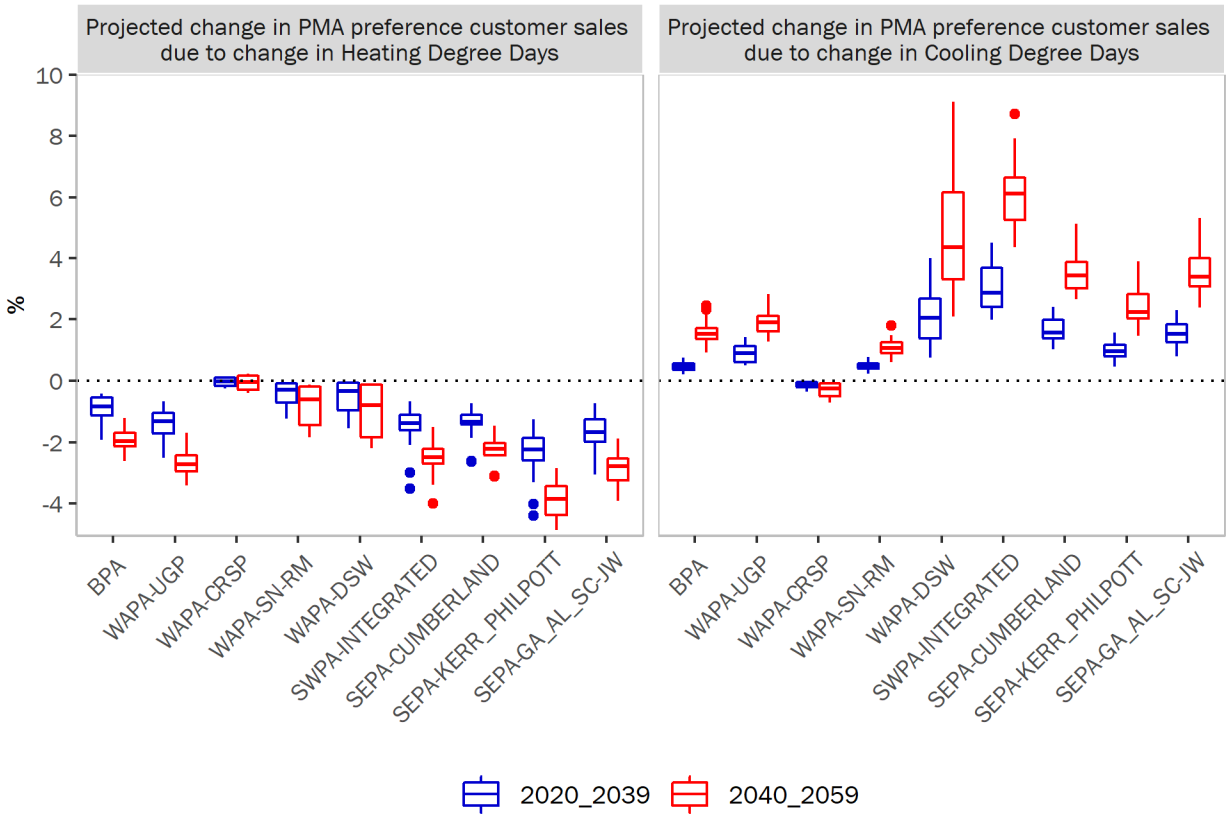


Figure 7.9. Estimated changes in PMA preference customer electricity sales relative to the baseline period due to projected changes in degree days.

Across PMAs, the expected response of sales to CDDs varies more than the expected response of sales to HDDs. In the near-term future period, the projected median decrease in winter sales ranges from less than 1% in BPA and all WAPA marketing regions (except WAPA-UGP) to 2.2% in the SEPA-Kerr-Philpott marketing region. In the mid-term future period, the projected median decrease in winter sales remains less than 1% for the WAPA-CRSP, WAPA-SN/RM, and WAPA-DSW marketing regions, reaches 4% in the SEPA-Kerr-Philpott marketing regions, and ranges from 2% to 3% in the remaining marketing regions.

SWPA and WAPA-DSW stand out as having the strongest projected increase in summer electricity sales because of higher CDDs across all PMAs. The median projected increase in sales associated with increased cooling demand is 2.9% for SWPA, 2% for WAPA-DSW, and 1.5% or less for preference customers in all other PMA regions in the near-term future period. In the mid-term future period, the projected median increase in annual sales driven by higher CDDs is 6.1% for SWPA preference customers, 2%–4% for WAPA-DSW and all SEPA marketing regions and, and less than 2% for the rest.

Adding up the projected changes in sales associated with lower heating demand and higher cooling demand results in a net increase in annual sales in the near-term for preference customers in four of the marketing regions (SWPA, WAPA-DSW, SEPA-Cumberland, and WAPA-SN/RM) and a decrease in the other five marketing regions. In the mid-term future period, the SEPA-GA/AL/SC marketing region joins

the group of PMA preference customers with a temperature-driven net increase in annual electricity sales. The median net changes in annual sales due to changes in degree days range from −1.1% to 1.4% in the near-term and from −1.7% to 3.6% in the mid-term future periods. In both periods, the lower ends of the range correspond to preference customers in the SEPA-Kerr-Philpott, and higher ends of the range correspond to preference customers in the SWPA marketing regions.

(This page intentionally left blank)

8. SUMMARY AND CONCLUSION

Hydropower is a key contributor to the US renewable energy portfolio because of its established development history and the diverse benefits it provides to electric power systems. Ensuring the sustainable operation of existing hydropower facilities is of great importance to the US renewable energy portfolio and the reliability of the electricity grid. Informed by the latest CMIP6 global climate projections, this study provides a broad assessment that addresses how climate change may affect future US federal hydropower generation. A spatially consistent assessment approach (Section 2) was designed to evaluate 132 federal hydropower plants that are marketed by four PMAs. The results for each PMA region are discussed in Sections 3 to 6, and an interregional comparison is provided in Section 7. The summary of major findings and future research needs are discussed in this final section.

8.1 SUMMARY OF MAJOR FINDINGS

The previous 9505 assessments (Sale et al., 2012; Kao et al., 2016) followed the conventional hydroclimate impact assessment approach that only considered the selection of GCM as the main source of uncertainty. Although the adopted hydrologic and hydropower models were intensively calibrated based on the best available observations, there was no clear approach to discern whether a detected change signal was caused by the original GCM or by any other modeling choices. This lack of clarity could lower the confidence of the assessment findings and lead to a biased interpretation of climate change–induced risks of future federal hydropower generation. To improve upon this limitation, based on extensive consultation with several federal hydropower stakeholder groups, an ensemble-based assessment approach was designed in this assessment (Figure 2.1). The ensemble includes six selected CMIP6 GCMs, two downscaling approaches, two reference meteorological observations, two hydrologic models, and two regional hydropower models. The alternative modeling components and supporting data sets are dissimilar in their nature (e.g., dynamical vs. statistical downscaling) but are all commonly used in many hydroclimate studies. All selected models were intensively calibrated/validated based on the best available historic observations before incorporating in the assessment framework. Although this ensemble-based approach still cannot represent the full range of uncertainties related to all factors in hydroclimate assessment, it allows for exploring the main modeling choice affecting the projections of future hydrologic and hydropower conditions.

For climate model selection and evaluation, six CMIP6 GCMs were selected for downscaling based on their skill, uniqueness, and data availability. The selection processes consisted of two independent evaluations that used evaluation matrices at daily, monthly, and seasonal time scales. The first methodology ranked GCMs using their weighted average relative errors across the evaluation matrix suite. The second methodology ranked GCMs using their Euclidean distances based on the truncated set of the first 10 modes in the multivariate EOF analyses. Both approaches led to a similar selection of six GCMs when their skill, uniqueness, and data availability were considered, which highlights the robustness of the selection process. The projected change of temperature and precipitation from the six selected GCM projections were further compared with other CMIP6 models and emission scenarios (Figure 2.7). Focusing on the projected changes from 1980–2019 baseline to 2020–2059 future periods at the CONUS scale, the findings include the following:

- For temperature, all CMIP6 projections showed a consistent increase in temperature ranging from 1°F to 6°F across all seasons.

- For precipitation, except for parts of summer and fall, a consistent increase is projected, especially in winter. These seasonal projections result in a net annual precipitation increase across most of the projections (−2% to 8%).

- Although the six selected GCM projections do not cover the full range of CMIP6 models and emission scenarios, their projections are distributed across the median of all models, suggesting that they are not biased toward any one direction.

- Although emission scenario has a clear influence on the projected change of temperature, its influence on the projected change of precipitation is less obvious at the CONUS scale in the time frame of interest (till 2059).

A comprehensive suite of downscaled CMIP6 climate projections was developed using two downscaling techniques (dynamical and statistical) with two reference meteorological observations (Daymet and Livneh). For dynamical downscaling, RegCM4 was employed. The outputs from RegCM4 were subsequently bias-corrected at a daily time scale. For statistical downscaling, the DBCCA technique was applied on selected variables. Additionally, two high-resolution observations were used in the training and/or correction process to account for the uncertainties resulting from the choice of reference observations. The main findings include the following:

- Although the RCM may improve several evaluation metrics, it may also exacerbate biases in a few other evaluation metrics (compared with the driving GCMs). The bias correction of dynamically downscaled simulations may remove these modeling biases.

- Certain differences in future hydroclimate projections arise from the choice of downscaling techniques and reference meteorological observations. These differences are more prominent in extremes than mean variables.

- The underlying causes of the differences arising from the choice of downscaling technique are difficult to ascertain because of the inherently different nature of the downscaling techniques that are employed. The projected changes in dynamical downscaling are governed by physical processes simulated in the models and the choice of parameterizations. However, the different response in the DBCCA technique could arise at different stages of the statistical downscaling techniques. Nevertheless, these analyses highlight the uncertainties arising from the choice of downscaling techniques, as well as the choice of observations used for training and/or correction. Given their important implications on impact assessment and planning applications, these differences are crucial and must be accounted for.

The climate projections were subsequently translated into ensemble hydrologic projections using two calibrated distributed hydrologic models (VIC and PRMS). Similar to the previous assessment, the calibration was performed at the HUC8 scale using the USGS WaterWatch monthly runoff data set (Brakebill et al., 2011) as the benchmark. Parameters were updated using the PSO (Kennedy and Eberhart, 1995) algorithm to maximize the NSE of monthly total runoff at the each HUC8. The main findings include the following:

- The annual total runoff is generally projected to increase across the CONUS.

- At the seasonal scale, winter and spring runoff are generally projected to increase across the CONUS. However, the summer runoff is projected to decrease for many parts of the CONUS, especially in the western and southern United States, indicating a shift in the timing and seasonality of the water availability.

- High runoff (95th percentile) is projected to increase in the majority of CONUS watersheds; however, low runoff (5th percentile of 7-day average) is projected to decrease in the eastern/central and western coastal areas of the CONUS.

By using two regional hydropower models (WRES and WMP), future monthly hydropower generation were further projected at each PMA study area. The main findings include the following:

- At the annual time scale, the total annual hydropower generation for all PMAs is projected to be around 120 TWh/year in both near-term (2020–2039) and mid-term (2040–2059) future periods, close to the baseline (1980–2019) level. Most PMAs featured a wider multimodel spread in the mid-term than near-term future periods, with no obvious change in the ensemble mean.

- At the seasonal time scale, the majority of models projected increasing generation in winter and spring, and decreasing generation in summer and fall. Such a result is mainly governed by the projected changes in BPA and WAPA caused by the earlier snowmelt and changing runoff seasonality.

- SWPA and WEPA projected increasing seasonal generation, which is heavily controlled by the variability of rainfall and runoff. When compared with the federal hydropower reservoirs in BPA and WAPA, the reservoirs in SWPA and SEPA have less storage capacity, so the projected change of seasonal hydropower generation will follow the projected change of seasonal rainfall and runoff more closely.

- The spread of multimodel hydropower change projections was generally smaller than the spread of multimodel natural streamflow change projections across all PMAs. Most of these federal hydropower plants are located at multipurpose reservoirs that also provide nonpower services that have higher priority than hydropower. Although those constraints were not individually and explicitly represented in the modeling approaches, the reservoir storage may help reduce the spread in future hydropower projections.

In addition to hydropower generation, the potential impacts of climate change on the direct evaporation from these federal reservoirs were also evaluated. The analysis was expanded to 678 major reservoirs (including both federal and nonfederal, hydropower and non-hydropower) across the CONUS for interregional comparison. The main findings include the following:

- A significant increase of reservoir evaporation is projected across the entire CONUS. Although the amount may be relatively small when compared with the total runoff volume in some PMA regions (e.g., BPA), it may cause a challenge for reservoirs in arid regions (e.g., WAPA).

- In the near-term future period, the reservoirs with the largest changes were mainly in the south-central CONUS (mainly southern Missouri, western Arkansas–White–Red, and western Texas-Gulf hydrologic regions), whereas the rate of increase was low in most of the eastern and western CONUS.

- In the mid-term future period, evaporation loss growth rates in the south-central United States are projected to remain the largest. In addition, the changes of evaporation loss in the western United States were also significant (6% on average). The growth rate in the eastern United States (especially in the southeast) is projected to be relatively low in comparison.

- The average evaporation loss growth rate was 5.6% in the mid-term future period, and the spatial heterogeneity of evaporation loss growth was more significant. The evaporation loss growth rate in the arid/semi-arid western United States was much higher than that in the eastern part, which is projected to worsen hydrologic drought conditions in the future.

In terms of climate change impacts on PMA customers' energy demand, the main findings include the following:

- The projected increase in average temperature, which yields a lower number of HDDs in the winter months and a higher number of CDDs in the summer months, resulted in projected decreases in the electricity sales of PMA preference customers to their end customers in the winter and increased sales in the summer.

- PMA preference customers will likely want to increase the fraction of their federal hydropower allocation they receive during the summer—to the extent allowed in their contracts with the PMAs.

- During the winter, hydropower generation is projected to increase in all PMAs. The combination of decreased heating load and increased hydropower generation suggests that federal hydropower surpluses are likely during the winter months.

- The projected increase in cooling loads was accompanied by increased summer generation in all PMAs, except BPA. Thus, the ability to shift water from winter to summer months and to maximize the revenue from winter surpluses to compensate for potential increased power purchase requirements in the summer will be valuable for all PMAs, and especially for BPA.

BPA is the only PMA directly affected by climate-driven changes in the electricity load served by its customers because it is required to provide sufficient electricity to follow the load of most of its customers. For WAPA, SWPA, and SEPA, the federal hydropower allocation that preference customers receive from the PMA is only a fraction of the portfolio of generation resources they use to serve their end-use customers. Any load growth projected by preference customers of these PMAs, due to climate trends or other drivers, would have to be met through generation sources other than federal hydropower.

Finally, to characterize the role of methodological choices in the modeling chain, an ANOVA was applied. Specifically for runoff and generation, the selection of CMs remains the dominant factor in many parts of the CONUS, but the role of HMs and PMs (along with other factors) can be more important in some locations and seasons. This ANOVA addresses the needs for a comprehensive evaluation across all components in the modeling chain for a more informed hydroclimate impact assessment.

Overall, several potential risks that may impact the resilience of future federal hydropower generation were identified. They include the following:

- **Hydrologic extremes**: As suggested by the projected high/low precipitation and runoff, as well as findings from other concurrent hydroclimate studies, the intensification of future hydrologic cycles and extreme events is one of the most critical issues threatening the resilience of US infrastructure systems. Both historical observations and model projections suggest that the intensity, frequency, and magnitude of extreme rainfall events will continue to increase, which will likely challenge conventional reservoir management practices. Although the reservoirs have been traditionally designed using very conservative rainfall estimates (i.e., probable maximum precipitation [PMP]), recent extreme events such as the 2017 Hurricane Harvey near Houston, Texas, demonstrated that such an extremely large PMP estimate can still be exceeded (Kao et al., 2019), suggesting the potential needs for more comprehensive evaluations.

  On the other hand, the duration and severity of extreme drought events are also projected to increase in many parts of the United States. Although the annual precipitation and runoff are generally projected to increase in a warming environment, the distribution is not uniform in space and time. Reclamation (2021) conducted a comprehensive drought assessment for the western United States using historical streamflow observations, future projections, and paleohydrology. The results suggested that the severity of drought events may not be sufficiently captured only based on limited historical streamflow observations. The recent severe drought in the western states resulted in

unprecedent disruption to water supply and hydropower generation, demonstrating the dire impacts of droughts.

Because the impacts of hydrologic extremes are reservoir- and project-specific, they are not explicitly simulated in this study. Operational water management and high-resolution hydrodynamic models, forced by non-stationary meteorologic forcings, will be required. Moving forward, site-specific analyses may be conducted to better evaluate the climate change–induced risks for major reservoirs across different geographical and hydrologic settings.

- **Conflicting timing of change**: Similar to in the previous assessments, temperature-driven early snowmelt is projected in most of the western United States, suggesting that the bulk of runoff may arrive earlier in the spring. However, as informed by the demand analysis, more temperature-driven water and energy demand is expected to shift from winter to summer, and hence creates a conflict. Although ideally one may expect to mitigate this conflict through reservoir management (especially with the help from large federal reservoirs), the intensified hydrologic extremes combined with all other competing water management objectives may limit the ability and flexibility in storing more water resources to meet the peak demand. Furthermore, in arid regions, the enhanced reservoir evaporation may result in sizable storage reduction and further exacerbate the nexus. Again, although the annual precipitation and runoff are generally projected to increase in a warming environment, it may not be available in the right timing and location. Site-specific studies using operational models forced by up-to-date hydroclimate projections, such as the recent RMJOC assessments (e.g., Pytlak et al., 2018; Glabau et al., 2020), can be one of the most rigorous ways to reliably evaluate the risks and identify required mitigation actions.

- **Other indirect impacts**: Given the specific legislative language of SWA, this assessment focuses on understanding how climate change may affect water availability for hydroelectric power generation and federal hydropower marketing. However, climate change can result in even broader impacts, such as the disruptions due to wildfire, increasing water demands, and more difficulties to meet environmental objectives. Additionally, the issue of aging infrastructures may also reduce the system's ability to mitigate runoff variability and increase the difficulty of future operation. Although these issues were not quantitatively analyzed in this assessment, they can be further investigated in future studies.

8.2 SUMMARY OF ASSESSMENT LIMITATIONS

Although the assessment uses state-of-the-art data and models, assessment limitations remain. They include the following:

- **Interpretation of future projections**: Although GCM-driven dynamical/statistical downscaling may provide the most scientifically defensible regional-scale climate projections, it should not be considered an absolute, day-to-day weather prediction. The main purpose of climate modeling is to simulate how general climate statistics may evolve with respect to the specified future emission scenarios—not to provide an exact prediction of future weather and hydrology. Also, a simulation is only one of the tools that one may use to evaluate the potential impacts and system vulnerabilities. Other observation-based assessment approaches are equally valuable and should not be omitted for a more holistic understanding.

- **Broader characterization of uncertainty**: Although the multimodel framework provides a more comprehensive projection of future US federal hydropower generation, multimodel projections still do not represent the full range of uncertainties related to all possible modeling choices. The extent of warmer temperatures causing increasingly drier soils and greater evapotranspiration is a significant

additional uncertainty. Even in the case of higher average precipitation, drier soils and greater evapotranspiration will lead to reduced water availability and lower average generation. The true uncertainties are more comprehensive and may not be fully captured because of limited knowledge, tools, and resources. The ensemble-based approach adopted in this study can serve as an initial example for the systematic analysis of broader uncertainties.

- **Progression of climate science**: Although the capabilities of GCMs have continuously improved through the years, many open challenges have not been resolved. For instance, although human activities play an important role in the Earth system environment, many of the GCM simulations were conducted without considering the potential human influence on land use and land cover change, and surface hydrologic alterations. Therefore, recurring climate impact assessments based on the best available climate science remain necessary.

- **Regional assessment focus***:* Overall, this study focused on 18 PMA study areas rather than individual reservoirs or power plants. Impacts on site-specific features, such as reservoir operation rules, water withdrawal/return, environmental flow requirements, and energy generation, were not explicitly modeled at each power plant. In other words, this study provides a first-order assessment to identify areas with the highest risk under projected climate conditions. If a concern is identified for a specific region (e.g., change of streamflow seasonality), a regionally focused study can then be conducted. The assessment itself does not replace the existing site-specific models and tools used by the PMA's water and energy resource managers.

8.3 FUTURE ASSESSMENT NEEDS

In this third 9505 assessment, an ensemble-based modeling framework that includes multiple GCMs, downscaling approaches, reference meteorological observations, hydrologic and hydropower models is introduced to evaluate the climate change effects on the US federal hydropower generation. Although the improved modeling capabilities may help the water and energy resource managers examine the climate change–induced risks more closely, unresolved issues and research challenges can be addressed in the future. Some climate change and hydropower assessment needs include the following:

- **Overall assessment needs**

    - **Need to conduct basin-specific studies**: Considering the varying geographical and socioeconomical challenges in different river basins, reservoir-specific studies may be required for basins with high water and energy interests. Studies using operational models forced by up-to-date hydroclimate projections, such as the recent RMJOC assessments (Pytlak et al., 2018; Glabau et al., 2020), are one way to evaluate the risks and identify possible mitigation actions.

    - **Need to better understand the characteristics and impacts of future drought conditions**: As reported by Williams et al. (2022), 2000–2021 was the driest 22 year period in the southwestern United States in the past 1,200 years. This prolonged megadrought highlights the need to better understand the characteristics (e.g., severity, timing, duration) and impacts of droughts for the resilience of our long-term energy and water supply. In particular, efforts should focus on improving the understanding, modeling, and analytics associated with future drought events.

    - **Need actionable, climate-informed data support**: Since hydroclimate modeling is not within a utility's original mission space, a utility may not have sufficient resources or dedicated in-house expertise to evaluate the risks due to long-term climate change. To reduce a utility's burden in conducting a full-scale hydroclimate study (i.e., from GCM selection all the way to river management simulation), it will be beneficial to provide actionable, climate-informed data

support to the broader energy and water communities. The capabilities established through this federal 9505 assessment may serve as a starting point.

- **Need to address the broader risks**: Further disruptions due to wildfire, environmental requirements, and reduced operational flexibilities should be jointly considered. Additionally, the issue of aging infrastructure may reduce the system's ability to mitigate runoff variability and increase the difficulty of future operation. Although these issues were not within the scope of this assessment, they should be further investigated in future studies.

- **Uncertainty quantification**

  - In the current assessment, although alternative modeling choices along the modeling chain are introduced, they still cannot capture the full range of uncertainties in the future hydroclimate projections. The true uncertainties are likely even more comprehensive and cannot be fully captured because of limited knowledge, tools, and resources. Depending on the topic of interest, further uncertainty analysis should be conducted to better address both known and unknown uncertainties. The ensemble-based approach adopted in this study can serve as an initial example for the systematic analysis of broader uncertainties. Further experiments can be designed to explore the physical reasons controlling these differences.

  - The current ANOVA did not consider the contribution of internal (e.g., interannual) variability of each GCM. The assessment also did not test other alternative ANOVA subsampling approach (e.g., Bosshard et al., 2013) that can better address the unequal number of members in each modeling dimension. These more in-depth ANOVA explorations can be topics of future study.

- **Downscaling techniques**

  - In this assessment, two very different downscaling techniques, as well as two different reference meteorological observations, were compared. Although all approaches produced comparable downscaled climate projections, more involved differences need to be investigated further. In the next steps, this work may be extended by incorporating other alternative statistical downscaling techniques and regional CMs and by utilizing a larger set of driving GCMs for a more in-depth evaluation. A wider range of approaches can be tested for a region of interest before implementing them at the national scale. A most suitable strategy for climate downscaling remains an open challenge.

- **Bias correction techniques**

  - Further analysis about suitable bias correction techniques is needed (which is not specifically evaluated in this assessment). Other studies have highlighted the difficulties in applying model bias corrections and statistical post-processing to temperature, precipitation, and streamflow projections. This is especially challenging in some basins in which the calibration record has become nonstationary, and many post-processing techniques rely on past observational data sets, which can inadvertently overcorrect and obscure emerging climate signals.

- **Hydrologic model**

  - In this study, two calibrated hydrologic models were used with an anticipation that at least one of the hydrologic models may produce reasonable results in regions with weaker hydrologic model performance. However, although both hydrologic models (VIC and PRMS) performed well for most parts of the CONUS, they have degraded performance in similar regions (e.g., with strong

surface-groundwater interactions). Future efforts may be to investigate other alternative hydrologic models with improved groundwater processes representation, or to identify other reasons (e.g., lack of meteorologic and hydrologic observations) that are beyond the choice of hydrologic models.

- VIC and PRMS were calibrated using similar procedures and data sets, but they can still result in unintentional biases due to differences in spatial resolution, parameterization, calibration and validation approach, modeling configuration, and so on. Future assessments may attempt to improve on this front by conducting systematic evaluations to investigate the magnitude of these biases and identify areas of improvement. Alternatively, community modeling tools such as Structure for Unifying Multiple Modeling Alternatives (Clark et al., 2015a, 2015b) may be considered, which would allow for configuration of a wide range of hydrologic modeling alternatives and may enable a more systematic model comparison.

- The hydrologic models in this study used an assumption of stationary land use/land cover in the future period. Future assessments may explore the use of projected land use/land cover types, which may help develop a more realistic representation of runoff/streamflow projections.

- **Hydropower model**

  - An opportunity exists to integrate this type of modeling approach to support long-term resource adequacy studies. The consistent representation of both federal and nonfederal hydropower across the CONUS domain provides the opportunity to integrate long-term hydroclimate projections with utilities, balancing authorities, and system operators resource adequacy and reliability studies. An ensemble of capacity expansion models could be used to provide a range of infrastructure scenarios to understand the contribution of hydropower under different policy.

  - The representation of hydropower operations must be enhanced to support the operation of future grids. For example, enhancement to a weekly time step, and representation of storage targets and daily fluctuations range and ramping, would be needed to further integrate the hydropower projections into unit commitment production cost models. This enhancement would allow for evaluating the impact of climate change on the ability for hydropower resources to sustain grid resilience under extreme events conditions, as well as under new power grid infrastructure and generation portfolio with new markets and net load profiles. The water management model would also be enhanced with the ability to add or remove reservoirs.

  - A need exists to better represent reservoir operations with developing projections in forecast mode rather than analysis mode. As the temporal resolution is increased, the contribution of inflow forecast to decision-making will increase and should be represented in future assessments.

  - A deep uncertainty analysis is also needed, integrating uncertainties in future water availability, water demand, priority in water uses, allocation of supply to different users, hydropower contribution to power grid operations, and other hydropower dynamics parameterization.

- **Power marketing analysis**

  - Except for in BPA, the sales data (proxy for load) used to estimate the relationship between load and degree days in this assessment were annual. Monthly data for the rest of the PMAs would allow for an enhanced understanding of the relationship between load and degree days in future assessments.

- Estimating the response of load to degree days separately for each month may provide some clarity about potential changes in the length of the cooling load season in different PMA marketing subregions. If the cooling loads served by PMA preference customers extend into the early fall (as already observed in SWPA for several years), preference customers would request larger volumes of federal hydropower during that period. This request can be challenging for the PMAs because that period has historically had low loads and been reserved for planned maintenance outages of hydropower plants.

- A second benefit of monthly sales or load data at the preference customer level is that the data unveil important differences in customer load shapes, such as those seen for BPA between winter peaking customers and those with large irrigation loads. Controlling for these different customer types could improve the explanatory power of the load equations, especially in some of the WAPA marketing subregions that are likely to also have customers with large irrigation loads.

(This page intentionally left blank)

9. REFERENCES

Abatzoglou, J. T. (2013), Development of Gridded Surface Meteorological Data for Ecological Applications and Modelling, *Int. J. Climatol.*, *33*, 121–131, https://doi.org/10.1002/joc.3413.

Alder, J. R., and S. W. Hostetler (2019), The Dependence of Hydroclimate Projections in Snow-Dominated Regions of the Western United States on the Choice of Statistically Downscaled Climate Data, *Water Resour. Res.*, *55*, 2279–2300, https://doi.org/10.1029/2018WR023458.

Almazroui, M., M. N. Islam, F. Saeed, S. Saeed, M. Ismail, M. A. Ehsan, I. Diallo, E. O'Brien, M. Ashfaq, D. Martínez-Castro, T. Cavazos, R. Cerezo-Mota, M. K. Tippett, W. J. Gutowski Jr., E. J. Alfaro, H. G. Hidalgo, A. Vichot-Llano, J. D. Campbell, S. Kamil, I. U. Rashid, M. B. Sylla, T. Stephenson, M. Taylor, and M. Barlow (2021), Projected Changes in Temperature and Precipitation over the United States, Central America, and the Caribbean in CMIP6 GCMs, *Earth Syst. Environ.*, *5*(1), 1–24, https://doi.org/10.1007/s41748-021-00199-5.

Ashfaq, M., D. Rastogi, M. A. Abid, and S.-C. Kao (2022), Evaluation of CMIP6 GCMs Over the CONUS for Downscaling Studies, ESSOAr, preprint, https://doi.org/10.1002/essoar.10510589.1.

Ashfaq, M., D. Rastogi, R. Mei, S.-C. Kao, S. Gangrade, B. S. Naz, and D. Touma (2016), High-resolution Ensemble Projections of Near-term Regional Climate over the Continental United States, *J. Geophys. Res.*, *121*, 9943–9963, https://doi.org/10.1002/2016JD025285.

Ashfaq, M., L. C. Bowling, K. Cherkauer, J. S. Pal, and N. S. Diffenbaugh (2010), Influence of Climate Model Biases and Daily-scale Temperature and Precipitation Events on Hydrological Impacts Assessment: A Case Study of the United States, *J. Geophys. Res.*, *115*, D14116, https://doi.org/10.1029/2009JD012965.

Ashfaq, M., S. Ghosh, S.-C. Kao, L. C. Bowling, P. Mote, D. Touma, S. A. Rauscher, and N. S. Diffenbaugh (2013), Near-term Acceleration of Hydroclimatic Change in the Western US, *J. Geophys. Res.*, *118*, 10676–10693, https://doi.org/10.1002/jgrd.50816.

Auffhammer, M., P. Baylis, and C. Hausman (2017), Climate Change Is Projected to Have Severe Impacts on the Frequency and Intensity of Peak Electricity Demand Across the United States, *P. Natl. Acad. Sci.*, *114*(8), 1886–1891, https://doi.org/10.1073/pnas.1613193114.

Ban, Z., T. Das, D. Cayan, M. Xiao, and D. P. Lettenmaier (2020), Understanding the Asymmetry of Annual Streamflow Responses to Seasonal Warming in the Western United States, *Water Resour. Res.*, *56*, e2020WR027158, https://doi.org/10.1029/2020WR027158.

Beigi, E., and F. Tsai (2014), GIS-based Water Budget Framework for High-resolution Groundwater Recharge Estimation of Large-scale Humid Regions, *J. Hydrol. Eng.*, *19*(8), 05014004, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000993.

Biemans, H., I. Haddeland, P. Kabat, F. Ludwig, R. W. A. Hutjes, J. Heinke, W. von Bloh, and D. Gerten (2011), Impact of Reservoirs on River Discharge and Irrigation Water Supply During the 20th Century, *Water Resour. Res.*, *47*, W03509, https://doi.org/10.1029/2009WR008929.

Bohn, T. J., B. Livneh, J. W. Oyler, S. W. Running, B. Nijssen, and D. P. Lettenmaier (2013), Global Evaluation of MTCLIM and Related Algorithms for Forcing of Ecological and Hydrological Models, *Agr. Forest Meteorol.*, *176*, 38–49, https://doi.org/10.1016/j.agrformet.2013.03.003.

Bosshard, T., M. Carambia, K. Goergen, S. Kotlarski, P. Krahe, M. Zappa, and C. Schär (2013), Quantifying Uncertainty Sources in An Ensemble of Hydrological Climate-impact Projections, *Water Resour. Res.*, *49*, 1523–1536, https://doi.org/10.1029/2011WR011533.

Brakebill, J. W., D. M. Wolock, and S. E. Terziotti (2011), Digital Hydrologic Networks Supporting Applications Related to Spatially Referenced Regression Modeling, *J. Am. Water Resour. Assoc.*, *47*(5), 916–932, https://doi.org/10.1111/j.1752-1688.2011.00578.x.

Brekke, L., B.L. Thrasher, E.P. Maurer, and T. Pruitt (2013), *Downscaled CMIP3 and CMIP5 climate and hydrology projections: Release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs*, US Department of the Interior, US Bureau of Reclamation, Technical Services Center, Denver, CO.

Chegwidden, O. S., B. Nijssen, D. E. Rupp, J. R. Arnold, M. P. Clark, J. J. Hamman, S.-C. Kao, Y. Mao, N. Mizukami, P. Mote, M. Pan, E. Pytlak, and M. Xiao (2019), How do Modeling Decisions Affect the Spread among Hydrologic Climate Change Projections? Exploring a Large Ensemble of Simulations across a Diversity of Hydroclimates, *Earth's Future*, *7*, 623–637, https://doi.org/10.1029/2018EF001047.

Chegwidden, O. S., D. E. Rupp, and B. Nijssen (2020), Climate Change Alters Flood Magnitudes and Mechanisms in Climatically-Diverse Headwaters Across the Northwestern United States, *Environ. Res. Lett.*, *15*(9), 094048, https://doi.org/10.1088/1748-9326/ab986f.

Chen, J., F. P. Brissette, A. Poulin, and R. Leconte (2011), Overall Uncertainty Study of the Hydrological Impacts of Climate Change for a Canadian Watershed, *Water Resour. Res.*, *47*, W12509, https://doi.org/10.1029/2011WR010602.

Chhin, R., and S. Yoden (2018), Ranking CMIP5 GCMs for Model Ensemble Selection on Regional Scale: Case Study of the Indochina Region, *J. Geophys. Res.*, *123*, 8949– 8974. https://doi.org/10.1029/2017JD028026.

Clark, M. P., B. Nijssen, J. Lundquist, D. Kavetski, D. Rupp, R. Woods, E. Gutmann, A. Wood, L. Brekke, J. Arnold, D. Gochis, and R. Rasmussen (2015a), A Unified Approach to Process-based Hydrologic Modeling. Part 1: Modeling concept, *Water Resour. Res.*, *51*, 2498–2514, https://doi.org/10.1002/2015WR017198.

Clark, M. P., B. Nijssen, J. Lundquist, D. Kavetski, D. Rupp, R. Woods, E. Gutmann, A. Wood, D. Gochis, R. Rasmussen, D. Tarboton, V. Mahat, G. Flerchinger, and D. Marks (2015b), A Unified Approach for Process-based Hydrologic Modeling: Part 2. Model Implementation and Example Applications, *Water Resour. Res.*, *51*, 2515–2542, https://doi.org/10.1002/2015WR017200.

Clark, M. P., R. L. Wilby, E. D. Gutmann, J. A. Vano, S. Gangopadhyay, A. W. Wood, H. J. Fowler, C. Prudhomme, J. R. Arnold, and L. D. Brekke (2016), Characterizing Uncertainty of the Hydrologic Impacts of Climate Change, *Curr. Clim. Change Rep.*, *2*, 55–64, https://doi.org/10.1007/s40641-016-0034-x.

Connolly, K. (2016), *Focus 2028 – Federal Hydropower Program*, https://www.bpa.gov/Finance/FinancialPublicProcesses/2028/doc2028/Focus%202028_Federal%20Hydro.pdf, accessed January 2022.

Cosgrove, B. A., D. Lohmann, K. E. Mitchell, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, and J. D. Tarpley (2003), Land Surface Model Spin-up Behavior in the North American Land Data Assimilation System (NLDAS), *J. Geophys. Res.*, *108*(D22), *8845*, https://doi.org/10.1029/2002JD003316.

Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris (2002), A Knowledge-based Approach to the Statistical Mapping of Climate, *Climate Res.*, *22*, 99–113, https://doi.org/10.3354/cr022099.

Davies, M. (1959), The Relationship between Weather and Electricity Demand, *Proceedings of the IEE – Part C: Monographs*, *106*(9), 27–37, https://doi.org/10.1049/pi-c.1959.0007.

Diffenbaugh, N. S., and M. Ashfaq (2010), Intensification of Hot Extremes in the United States, *Geophys. Res. Lett.*, *37*, L15701, https://doi.org/10.1029/2010GL043888.

Diffenbaugh, N. S., J. S. Pal, R. J. Trapp, and F. Giorgi (2005), Fine-scale Processes Regulate the Response of Extreme Events to Global Climate Change, *Proc. Natl. Acad. Sci.*, *102*(44), 15,774–15,778, https://doi.org/10.1073/pnas.0506042102.

Diffenbaugh, N. S., M. Ashfaq, and M. Scherer (2011), Transient Regional Climate Change: Analysis of the Summer Climate Response in a High-resolution, Century-scale Ensemble Experiment Over the Continental United States, *J. Geophys. Res.*, *116*, D24111, https://doi.org/10.1029/2011JD016458.

DOE (US Department of Energy) (2013), *Effects of Climate Change on Federal Hydropower: Report to Congress*, Washington, DC.

DOE (US Department of Energy) (2017), *Effects of Climate Change on Federal Hydropower: The Second Report to Congress*, Washington, DC.

Döll, P., H. Hoffmann-Dobrev, F. T. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg, B. R. Scanlon (2012), Impact of Water Withdrawals from Groundwater and Surface Water on Continental Water Storage Variations, *J. Geodyn.*, *59–60*, 143–156, https://doi.org/10.1016/j.jog.2011.05.001.

Duan, Q. Y., V. K. Gupta, and S. Sorooshian (1993), Shuffled Complex Evolution Approach for Effective and Efficient Global Minimization, *J. Optim. Theory Appl.*, *76*, 501–521, https://doi.org/10.1007/BF00939380.

Edmonds, J., and J. M. Reilly (1985), *Global Energy: Assessing the Future*, Oxford University Press.

Edmonds, J., M. Wise, H. Pitcher, R. Richels, T. Wigley, and C. Maccracken (1997), An Integrated Assessment of Climate Change and the Accelerated Introduction of Advanced Energy Technologies, *Mitig. Adapt. Strat. Glob. Change*, *1*(4), 311–339, https://doi.org/10.1007/BF00464886.

EIA (US Energy Information Administration) (2020), *Form EIA-923 Detailed Data with Previous form Data (EIA-906/920)*, http://www.eia.gov/electricity/data/eia923/, accessed June 2020.

Elsner, M. M., S. Gangopadhyay, T. Pruitt, L. D. Brekke, N. Mizukami, and M. P. Clark (2014), How Does the Choice of Distributed Meteorological Data Affect Hydrologic Model Calibration and Streamflow Simulations?, *J. Hydrometeor.*, *15*(4), 1384–1403, https://doi.org/10.1175/JHM-D-13-083.1.

Erlandsen, H. B., L. M. Tallaksen, and J. Kristiansen (2019), Merits of Novel High-Resolution Estimates and Existing Long-Term Estimates of Humidity and Incident Radiation in A Complex Domain, *Earth Syst. Sci. Data*, *11*(2), 797–821, https://doi.org/10.5194/essd-11-797-2019.

Eyring, V., P. M. Cox, G. M. Flato, P. J. Gleckler, G. Abramowitz, P. Caldwell, W. D. Collins, B. K. Gier, A. D. Hall, F. M. Hoffman, G. C. Hurtt, A. Jahn, C. D. Jones, S. A. Klein, J. P. Krasting, L. Kwiatkowski, R. Lorenz, E. Maloney, G. A. Meehl, A. G. Pendergrass, R. Pincus, A. C. Ruane, J. L. Russell, B. M. Sanderson, B. D. Santer, S. C. Sherwood, I. R. Simpson, R. J. Stouffer, M. S. Williamson (2019), Taking Climate Model Evaluation to the Next Level, *Nature Clim. Change*, *9*(2), 102–110, https://doi.org/10.1038/s41558-018-0355-y.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor (2016), Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization, *Geosci. Model Dev.*, *9*, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Fekete, B. M., D. Wisser, C. Kroeze, E. Mayorga, L. Bouwman, W. M. Wollheim, and C. Vörösmarty (2010), Millennium Ecosystem Assessment Scenario Drivers (1970–2050): Climate and Hydrological Alterations, *Global Biogeochem. Cycles*, *24*, GB0A12, https://doi.org/10.1029/2009GB003593.

Fernandez, A. (2020), *A Summary of Current Federal Hydropower Generation and Marketing Flexibility*, ORNL/TM-2020/1461, Oak Ridge National Laboratory, Oak Ridge, TN.

Friedrich, K., R. L. Grossman, J. Huntington, P. D. Blanken, J. Lenters, K. D. Holman, D. Gochis, B. Livneh, J. Prairie, E. Skeie, N. C. Healey, K. Dahm, C. Pearson, T. Finnessey, S. J. Hook, and T. Kowalski (2018), Reservoir Evaporation in the Western United States: Current Science, Challenges, and Future Needs, *B. Am. Meteorol. Soc.*, *99*(1), 167–187, https://doi.org/10.1175/BAMS-D-15-00224.1.

Gangrade, S., S.-C. Kao, and R. A. McManamay (2020), Multimodel Hydroclimate Projections for the Alabama-Coosa-Tallapoosa River Basin in the Southeastern United States, *Nature Sci. Rep.*, *10*, 2870, https://doi.org/10.1038/s41598-020-59806-6.

GAO (Government Accountability Office) (2001), *The Evolution of Preference in Marketing Federal Power*, GAO-01-373, accessed 05/24/2021: https://www.gao.gov/assets/gao-01-373.pdf.

Giorgi, F., E. Coppola, F. Solmon, L. Mariotti, M. B. Sylla, X. Bi, N. Elguindi, G. T. Diro, V. Nair, G. Giuliani, U. U. Turuncoglu, S. Cozzini, I. Güttler, T. A. O'Brien, A. B. Tawfik, A. Shalaby, A. S. Zakey, A. L. Steiner, F. Stordal, L. C. Sloan, and C. Brankovic (2012), RegCM4: Model Description and Preliminary Tests Over Multiple CORDEX Domains, *Clim. Res.*, *52*(1), 7–29, https://doi.org/10.3354/cr01018.

Giuntoli, I., G. Villarini, C. Prudhomme, and D. M. Hannah (2018), Uncertainties in Projected Runoff Over the Conterminous United States, *Climatic Change*, *150*, 149–162 (2018), https://doi.org/10.1007/s10584-018-2280-5.

Glabau, B., E. Nielsen, A. Mylvahanan, N. Stephan, C. Frans, K. Duffy, J. Giovando, and J. Johnson (2020), *Climate and Hydrology data sets for RMJOC Long-Term Planning Studies: Second Edition (RMJOC-II), Part II: Columbia River Reservoir Regulation and Operations—Modeling and Analyses*, River Management Joint Operating Committee (RMJOC): Bonneville Power Administration, US Army Corps of Engineers, and US Bureau of Reclamation.

Grell, G. A. (1993), Prognostic Evaluation of Assumptions Used by Cumulus Parameterizations, *Mon. Weather Rev.*, *121*(3), 764–787, https://doi.org/10.1175/1520-0493(1993)121<0764:Peoaub>2.0.Co;2.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling, *J. Hydrol.*, *377*, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003.

Guo, G., X. Wang, L. Xiang, and F. Nian (2014), Sensor/Actuator Networks and Networked Control Systems, *Math. Probl. Eng.*, *2014*, 805380, https://doi.org/10.1155/2014/805380.

Gutmann, E., T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen (2014), An Intercomparison of Statistical Downscaling Methods Used for Water Resource Assessments in the United States, *Water Resour. Res.*, *50*, 7167–7186, https://doi.org/10.1002/2014WR015559.

Hamman, J. J., B. Nijssen, T. J. Bohn, D. R. Gergel, and Y. Mao (2018), The Variable Infiltration Capacity Model Version 5 (VIC-5): Infrastructure Improvements for New Applications and Reproducibility, *Geosci. Model Dev.*, *11*, 3481–3496, https://doi.org/10.5194/gmd-11-3481-2018.

Hanasaki, N., S. Kanae, and T. Oki (2006), A Reservoir Operation Scheme for Global River Routing Models, *J. Hydrol.*, *327*, 22–41, https://doi.org/10.1016/j.jhydrol.2005.11.011.

Harding, B. L., A. W. Wood, and J. R. Prairie (2012), The Implications of Climate Change Scenario Selection for Future Streamflow Projection in the Upper Colorado River Basin, *Hydrol. Earth Syst. Sci.*, *16*(11), 3989–4007, https://doi.org/10.5194/hess-16-3989-2012.

Hattermann, F. F., T. Vetter, L. Breuer, B. Su, P. Daggupati, C. Donnelly, B. Fekete, F. Flörke, S. N. Gosling, P. Hoffmann, S. Liersch, Y. Masaki, Y. Motovilov, C. Müller, L. Samaniego, T. Stacke, Y. Wada, T. Yang, and V. Krysnaova (2018), Sources of Uncertainty in Hydrological Climate Impact Assessment: A Cross-scale Study, *Environ. Res. Lett.*, *13*, 015006, https://doi.org/10.1088/1748-9326/aa9938.

Hay, L. E. (2019), *Application of the National Hydrologic Model Infrastructure with the Precipitation-Runoff Modeling System (NHM-PRMS), by HRU Calibrated Version*, US Geological Survey Data Release, https://doi.org/10.5066/P9NM8K8W.

Hay, L. E., and J. H. LaFontaine (2020), *Application of the National Hydrologic Model Infrastructure with the Precipitation-Runoff Modeling System (NHM-PRMS), 1980–2016, Daymet Version 3 Calibration*, US Geological Survey Data Release, https://doi.org/10.5066/P9PGZE0S.

Hay, L. E., S. L. Markstrom, and C. Ward-Garrison (2011), Watershed-Scale Response to Climate Change through the Twenty-First Century for Selected Basins across the United States, *Earth Interactions*, *15*(17), 1–37, https://doi.org/10.1175/2010EI370.1.

Heidari, H., M. Arabi, T. Warziniack, and S.-C. Kao (2020), Assessing Shifts in Regional Hydroclimatic Conditions of U.S. River Basins in Response to Climate Change over the 21st Century, *Earth's Future*, *8*(10), e2020EF001657, https://doi.org/10.1029/2020EF001657.

Hejazi, M. I., N. Voisin, L. Liu, L. M. Bramer, D. C. Fortin, J. E. Hathaway, M. Huang, P. Kyle, L. R. Leung, H.-Y. Li, Y. Liu, P. L. Patel, T. C. Pulsipher, J. S. Rice, T. K. Tesfa, C. R. Vernon, and Y. Zhou (2015), 21st Century United States Emissions Mitigation Could Increase Water Stress More

Than the Climate Change It Is Mitigating, *Proc. Natl. Acad. Sci.*, *112*(34), 10635–40, https://doi.org/10.1073/pnas.1421675112.

Hengade, N., T. I. Eldho, and S. Ghosh (2018), Climate Change Impact Assessment of a River Basin Using CMIP5 Climate Models and the VIC Hydrological Model, *Hydrolog. Sci. J.*, *63*(4), 596–614, https://doi.org/10.1080/02626667.2018.1441531.

Holtslag, A. A. M., E. I. F. Debruijn, and H. L. Pan (1990), A High-Resolution Air-Mass Transformation Model for Short-Range Weather Forecasting, *Moon Weather Rev.*, *118*(8), 1561–1575, https://doi.org/10.1175/1520-0493(1990)118<1561:Ahramt>2.0.Co;2.

Huang, J., and K. R. Gurney (2016), The Variation of Climate Change Impact On Building Energy Consumption to Building Type and Spatiotemporal Scale, *Energy*, *111*, 137–153, https://doi.org/10.1016/j.energy.2016.05.118.

Huang, X., and P. A. Ullrich (2017), The Changing Character of Twenty-first-century Precipitation Over the Western United States in the Variable-resolution CESM, *J. Clim.*, *30*(18), 7555–7575, https://doi.org/10.1175/JCLI-D-16-0673.1.

Islam, S. U., S. J. Déry, and A. T. Werner (2017), Future Climate Change Impacts on Snow and Water Resources of the Fraser River Basin, British Columbia, *J. Hydrometeor.*, *18*(2), 473–496, https://doi.org/10.1175/JHM-D-16-0012.1.

Jiang, Y., T. Hu, C. Huang, and X. Wu (2007), An Improved Particle Swarm Optimization Algorithm, *Appl. Math. Comput.*, *193*, *1*, 231–239, https://doi.org/10.1016/j.amc.2007.03.047.

Jiang, Y., J. B. Kim, C. J. Still, B. K. Kerns, J. D. Kline, and P. G. Cunningham (2018), Inter-comparison of Multiple Statistically Downscaled Climate data sets for the Pacific Northwest, USA, *Sci. Data*, *5*(1), 1–18, https://doi.org/10.1038/sdata.2018.16.

Johnson, M. M. (2021), *Existing Hydropower Assets (EHA) Capacity Plant Database, 2005-2019*, HydroSource, Oak Ridge National Laboratory, Oak Ridge, TN, https://doi.org/10.21951/EHA_Capacity_FY2020/1721365.

Johnson, M. M., S.-C. Kao, N. M. Samu, and R. Uría Martínez (2021), *Existing Hydropower Assets, 2021*, HydroSource, Oak Ridge National Laboratory, Oak Ridge, TN, https://doi.org/10.21951/EHA_FY2021/1782791.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W.Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph (1996), The NCEP/NCAR 40-year reanalysis project, *B. Am. Meteorol. Soc.*, *77*, 437–471, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kao, S.-C., M. Ashfaq, B. S. Naz, R. Uría Martínez, D. Rastogi, R. Mei, Y. Jager, N. M. Samu, and M. J. Sale (2016), *The Second Assessment of the Effects of Climate Change on Federal Hydropower*, ORNL/SR-2015/357, Oak Ridge National Laboratory, Oak Ridge, TN, https://doi.org/10.2172/1340431.

Kao, S.-C., M. J. Sale, M. Ashfaq, R. Uría Martínez, D. P. Kaiser, Y. Wei, and N. S. Diffenbaugh (2015), Projecting Changes in Annual Hydropower Generation Using Regional Runoff Data: An Assessment

of the United States Federal Hydropower Plants, *Energy*, *80*, 239–250, https://doi.org/10.1016/j.energy.2014.11.066.

Kao, S.-C., S. T. DeNeale, and D. B. Watson (2019), Hurricane Harvey Highlights: Need to Assess the Adequacy of Probable Maximum Precipitation Estimation Methods, *J. Hydrol. Eng.*, *24*(4), 05019005, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001768.

Kennedy, J., and R. Eberhart (1995), Particle Swarm Optimization, *Proceedings of ICNN'95-International Conference on Neural Networks*, *4*, 1942–1948, https://doi.org/10.1109/ICNN.1995.488968.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, D. L. Williamson, and P. J. Rasch (1998), The National Center for Atmospheric Research Community Climate Model: CCM3, *J. Clim.*, *11*(6), 1131–1149, https://doi.org/10.1175/1520-0442(1998)011<1131:Tncfar>2.0.Co;2.

Kim, S. H., J. Edmonds, J. Lurz, S. J. Smith, and M. Wise (2006), The ObjECTS Framework for Integrated Assessment: Hybrid Modeling of Transportation, *The Energy Journal*, Special Issue #2, 51–80.

Knutti, R., D. Masson, and A. Gettelman (2013), Climate Model Genealogy: Generation CMIP5 and How We Got There, *Geophys. Res. Lett.*, *40*, 1194–1199, https://doi.org/10.1002/grl.50256.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010), Challenges in Combining Projections for Multiple Models, *J. Climate*, *23*, 2739–2758, https://doi.org/10.1175/2009JCLI3361.1.

LaFontaine, J. H., R. M. Hart, L. E. Hay, W. H. Farmer, A. R. Bock, R. J. Viger, S. L. Markstrom, R. S. Regan, and J. M. Driscoll (2019), *Simulation of Water Availability in the Southeastern United States for Historical and Potential Future Climate and Land-cover Conditions*, US Geological Survey Scientific Investigations Report 2019–5039, https://doi.org/10.3133/sir20195039.

Leavesley, G. H., R. W. Lichty, B. M. Troutman, and L. G. Saindon (1983), *Precipitation-runoff Modeling System: User's Manual*, US Geological Survey Water-resources Investigations Report 83-4238, https://doi.org/10.3133/wri834238.

Lee, K. Y., and J.-B. Park (2006), Application of Particle Swarm Optimization to Economic Dispatch Problem: Advantages and Disadvantages, in *2006 IEEE PES Power Systems Conference and Exposition*, 188–192, https://doi.org/10.1109/PSCE.2006.296295.

Lehner, B., C. R. Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J. C. Robertson, R. Rödel, N. Sindorf, and D. Wisser (2011), High-resolution Mapping of the World's Reservoirs and Dams for Sustainable River-flow Management, *Front. Ecol. Environ.*, *9*, 494–502, https://doi.org/10.1890/100125.

Leng, G., M. Huang, N. Voisin, X. Zhang, G. R Asrar, and L. R. Leung (2016), Emergence of New Hydrologic Regimes of Surface Water Resources in the Conterminous United States Under Future Warming, *Environ. Res. Lett.*, *11*, 114003, https://doi.org/10.1088/1748-9326/11/11/114003.

Li, H. Y., M. S. Wigmosta, H. Wu, M. Huang, Y. Ke, A. M. Coleman, and L. R. Leung (2013). A Physically Based Runoff Routing Model for Land Surface and Earth System Models, *J. Hydrometeor.*, *14*, 808–828, https://doi.org/10.1175/JHM-D-12-015.1.

Li, J., L. Yang, and H. Long (2018), Climatic Impacts on Energy Consumption: Intensive and Extensive Margins, *Energ. Econ.*, *71*, 332–343, https://doi.org/10.1016/j.eneco.2018.03.010.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A Simple Hydrologically Based Model of Land Surface Water and Energy Fluxes for General Circulation Models, *J. Geophys. Res.*, *99*(D7), 14415–14428, https://doi.org/10.1029/94JD00483.

Livneh, B., T. J. Bohn, D. W. Pierce, F. Munoz-Arriola, B. Nijssen, R. Vose, D. R. Cayan, and L. Brekke (2015), A Spatially Comprehensive, Hydrometeorological Data Set for Mexico, the U.S., and Southern Canada 1950—2013, *Nature Scientific Data*, *5*, 150042, https://doi.org/10.1038/sdata.2015.42.

Lohmann, D., E. Raschke, B. Nijssen, and D. P. Lettenmaier (1998), Regional Scale Hydrology: I. Formulation of the VIC-2L Model Coupled to a Routing Model, *Hydrolog. Sci. J.*, *43*(1), 131–141, https://doi.org/10.1080/02626669809492107.

Lohmann, D., R. Nolte-Holube, and E. Raschke (1996), A Large-scale Horizontal Routing Model to Be Coupled to Land Surface Parametrization Schemes, *Tellus A*, *48*(5), 708–721, https://doi.org/10.3402/tellusa.v48i5.12200.

Lukas, J., and E. Payton, eds. (2020), *Colorado River Basin Climate and Hydrology: State of the Science*, Western Water Assessment, University of Colorado Boulder, https://doi.org/10.25810/3hcv-w477.

Magnuson, P. A. (2009), *Memorandum and Order on the Tri-State Water Rights Litigation*, Case No. 3:07-md-01, US District Court, Middle District of Florida.

Markstrom, S. L., L. E. Hay, and M. P. Clark (2016), Towards Simplification of Hydrologic Modeling: Identification of Dominant Processes, *Hydrol. Earth Syst. Sci.*, *20*(11), 4655–4671, https://doi.org/10.5194/hess-20-4655-2016.

Markstrom, S. L., R. S. Regan, L. E. Hay, R. J. Viger, R. M. Webb, R. A. Payn, and J. H. LaFontaine (2015), *PRMS-IV, the Precipitation-runoff Modeling System, version 4*, US Geological Survey Techniques and Methods 6-B7, https://doi.org/10.3133/tm6B7.

Maurer, E. P., H. G. Hidalgo, T. Das, M. D. Dettinger, and D. R. Cayan (2010), The Utility of Daily Large-scale Climate Data in the Assessment of Climate Change Impacts on Daily Streamflow in California, *Hydrol. Earth Syst. Sci.*, *14*, 1125–1138, https://doi.org/10.5194/hess-14-1125-2010.

Maurer, E. P., T. Das, and D. R. Cayan (2013), Errors in Climate Model Daily Precipitation and Temperature Output: Time Invariance and Implications for Bias Correction, *Hydrol. Earth Syst. Sci.*, *17*(6), 2147–2159, https://doi.org/10.5194/hess-17-2147-2013.

Meresa, H. K., and R. J. Romanowicz (2017), The Critical Role of Uncertainty in Projections of Hydrological Extremes, *Hydrol. Earth Syst. Sci.*, *21*, 4245–4258, https://doi.org/10.5194/hess-21-4245-2017.

Mesinger, F., G. DiMegoand, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, E. Rogers, E.H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi (2006), North American regional reanalysis, *B. Am. Meteorol. Soc.*, *87*, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

Mizukami, N., M. P. Clark, A. G. Slater, L. D. Brekke, M. M. Elsner, J. R. Arnold, and S. Gangopadhyay (2014), Hydrologic Implications of Different Large-Scale Meteorological Model Forcing data sets in Mountainous Regions, *J. Hydrometeorol.*, *15*(1), 474–488, https://doi.org/10.1175/JHM-D-13-036.1.

Moreo, M. T. (2015), *Evaporation Data from Lake Mead and Lake Mohave, Nevada and Arizona, March 2010 through April 2015*, US Geological Survey Data Release, https://doi.org/10.5066/F79C6VG3.

Mote, P. W., and E. P. Salathé (2010), Future Climate in the Pacific Northwest, *Clim. Change*, *102*(1), 29–50, https://doi.org/10.1007/s10584-010-9848-z.

Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier (2005), Declining Mountain Snowpack in Western North America, *B. Am. Meteorol. Soc.*, *86*, 39–49, https://doi.org/10.1175/BAMS-86-1-39.

Mote, P. W., S. Li, D. P. Lettenmaier, M. Xiao, and R. Engel (2018), Dramatic Declines in Snowpack in the Western US, *npj Clim Atmos Sci*, *1*, 2, https://doi.org/10.1038/s41612-018-0012-1.

Najafi, M. R., H. Moradkhani, and I. W. Jung (2011), Assessing the Uncertainties of Hydrologic Model Selection in Climate Change Impact Studies, *Hydrol. Process.*, *25*, 2814–2826, https://doi.org/10.1002/hyp.8043.

Naz, B. S., S.-C. Kao, M. Ashfaq, D. Rastogi, R. Mei, and L. C. Bowling (2016), Regional Hydrologic Response to Climate Change in the Conterminous United States Using High-resolution Hydroclimate Simulations, *Global Planet. Change*, *143*, 100–117, https://doi.org/10.1016/j.gloplacha.2016.06.003.

Naz, B. S., S.-C. Kao, M. Ashfaq, H. Gao, D. Rastogi, and S. Gangrade (2018), Effects of Climate Change on Streamflow Extremes and Implications for Reservoir Inflow in the United States, *J. Hydrol.*, *556*, 359–370, https://doi.org/10.1016/j.jhydrol.2017.11.027.

Newman, A. J., N. Mizukami, M. P. Clark, A. W. Wood, B. Nijssen, and G. Nearing (2017), Benchmarking of A Physically Based Hydrologic Model, *J. Hydrometeorol.*, *18*(8), 2215–2225, https://doi.org/10.1175/JHM-D-16-0284.1.

Nikulin, G., S. Asharaf, M. E. Magariño, S. Calmanti, R. M. Cardoso, J. Bhend, J. Fernández, M. D. Frías, K. Fröhlich, B. Früh, S. H. García, R. Manzanas, J. M. Gutiérrez, U. Hansson, M. Kolax, M. A. Liniger, P. M. M.Soares, C. Spirig, R. Tome, and K. Wyser (2018), Dynamical and Statistical Downscaling of a Global Seasonal Hindcast in Eastern Africa, *Climate Services*, *9*, 72–85, https://doi.org/10.1016/j.cliser.2017.11.003.

O'Neill, B. C., C. Tebaldi, D. P. van Vuuren, V. Eyring, P. Friedlingstein, G. Hurtt, R. Knutti, E. Kriegler, J.-F. Lamarque, J. Lowe, G. A. Meehl, R. Moss, K. Riahi, and B. M. Sanderson (2016), The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, *Geosci. Model Dev.*, *9*, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016.

Oubeidillah, A. A., S.-C. Kao, M. Ashfaq, B. S. Naz, and G. Tootle (2014), A Large-Scale, High-Resolution Hydrological Model Parameter Data Set for Climate Change Impact Assessment for the Conterminous US, *Hydrol. Earth Syst. Sci.*, *18*, 67–84, https://doi.org/10.5194/hess-18-67-2014.

Pal, J. S., E. E. Small, and E. A. B. Eltahir (2000), Simulation of Regional-scale Water and Energy Budgets: Representation of Subgrid Cloud and Precipitation Processes Within RegCM, *J. Geophys. Res.*, *105*, 29,579–29,594, https://doi.org/10.1029/2000JD900415.

Penman, H. L. (1948), Natural Evaporation from Open Water, Bare Soil and Grass, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *193*(1032), 120–145, https://doi.org/10.1098/rspa.1948.0037.

Persad, G. G., D. L. Swain, C. Kouba, and J. P. Ortiz-Partida (2020), Inter-model Agreement on Projected Shifts in California Hydroclimate Characteristics Critical to Water Management, *Climatic Change*, *162*, 1493–1513, https://doi.org/10.1007/s10584-020-02882-4.

Pierce, D. W., D. R. Cayan, and B. L. Thrasher (2014), Statistical Downscaling Using Localized Constructed Analogs (LOCA), *J. Hydrometeor.*, *15*, 2558–2585, https://doi.org/10.1175/JHM-D-14-0082.1.

Pierce, D. W., L. Su, D. R. Cayan, M. D. Risser, B. Livneh, and D. P. Lettenmaier (2021), An Extreme-preserving Long-term Gridded Daily Precipitation Data Set for the Conterminous United States, *J. Hydrometeor.*, *22*(7), 1883–1895, https://doi.org/10.1175/JHM-D-20-0212.1.

Pytlak, E., C. Frans, K. Duffy, J. Johnson, B. Nijssen, O. Chegwidden, and D. Rupp (2018), *Climate and Hydrology data sets for RMJOC Long-Term Planning Studies: Second Edition (RMJOC-II), Part I: Hydroclimate Projections and Analyses*, River Management Joint Operating Committee (RMJOC): Bonneville Power Administration, US Army Corps of Engineers, and US Bureau of Reclamation.

Queen, L. E., P. W. Mote, D. E. Rupp, O. Chegwidden, and B. Nijssen (2021), Ubiquitous Increases in Flood Magnitude in the Columbia River Basin under Climate Change, *Hydrol. Earth Syst. Sci.*, *25*(1), 257–272, https://doi.org/10.5194/hess-25-257-2021.

Rastogi, D., S.-C. Kao, and M. Ashfaq (2022), How May the Choice of Downscaling Techniques and Meteorological Reference Observations Affect Future Hydroclimate Projections?, *Earth's Future*, *10*, e2022EF002734, https://doi.org/10.1029/2022EF002734.

Reclamation (Bureau of Reclamation) (2021), *Water Reliability in the West – 2021 SECURE Water Act Report. Prepared for the United States Congress*, Bureau of Reclamation, Water Resources and Planning Office, Denver, CO.

Regan, R. S., K. E. Juracek, L. E. Hay, S. L. Markstrom, R. J. Viger, J. M. Driscoll, J. H. LaFontaine, and P. A. Norton (2019), The U. S. Geological Survey National Hydrologic Model Infrastructure: Rationale, Description, and Application of a Watershed-scale Model for the Conterminous United States, *Environ. Modell. Softw.*, *111*, 192–203, https://doi.org/10.1016/j.envsoft.2018.09.023.

Rupp, D. E., J. T. Abatzoglou, K. C. Hegewisch, and P. W. Mote (2013), Evaluation of CMIP5 20th Century Climate Simulations for the Pacific Northwest USA, *J. Geophys. Res.*, *118*, 10,884–10,906, https://doi.org/10.1002/jgrd.50843.

Rupp, D. E., J. T. Abatzoglou, and P. W. Mote (2017), Projections of 21st Century Climate of the Columbia River Basin, *Clim. Dynam.*, *49*(5), 1783–1799, https://doi.org/10.1007/s00382-016-3418-7.

Salathé Jr., E. P., A. F. Hamlet, C. F. Mass, S. Lee, M. Stumbaugh, and R. Steed (2014), Estimates of Twenty-First-Century Flood Risk in the Pacific Northwest Based on Regional Climate Model Simulations, *J. Hydrometeor.*, *15*(5), 1881–1899, https://doi.org/10.1175/JHM-D-13-0137.1.

Sale, M. J., S.-C. Kao, M. Ashfaq, D. P. Kaiser, R. Uría Martínez, C. Webb, and Y. Wei (2012), *Assessment of the Effects of Climate Change on Federal Hydropower*, ORNL/TM-2011/251, Oak Ridge National Laboratory, Oak Ridge, TN, https://doi.org/10.2172/1220238.

Sanderson, B. M., M. Wehner, and R. Knutti (2017), Skill and Independence Weighting for Multimodel Assessments, *Geosci. Model Dev.*, *10*, 2379–2395, https://doi.org/10.5194/gmd-10-2379-2017.

Schlag, N., A. Olson, E. Hart, A. Mileva, R. Jones, C. B. Martinez-Anido, B.-M. Hodge, G. Brinkman, A. Florita, and D. Biagioni (2015), *Western Interconnection Flexibility Assessment: Final Report*, Energy and Environmental Economics, Inc. and National Renewable Energy Laboratory.

Schnorbus, M., A. Werner, and K. Bennett (2014), Impacts of Climate Change in Three Hydrologic Regimes in British Columbia, Canada, *Hydrol. Process.*, *28*, 1170–1189, https://doi.org/10.1002/hyp.9661.

Sexstone, G. A., J. M. Driscoll, L. E. Hay, J. C. Hammond, and T. B. Barnhart (2020), Runoff Sensitivity to Snow Depletion Curve Representation within a Continental Scale Hydrologic Model, *Hydrological Processes*, *34*, 2365–2380, https://doi.org/10.1002/hyp.13735.

Slater, A. G. (2016), Surface Solar Radiation in North America: A Comparison of Observations, Reanalyses, Satellite, and Derived Products, *J. Hydrometeorol.*, *17*(1), 401–420, https://doi.org/10.1175/JHM-D-15-0087.1.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger (2005), Changes toward Earlier Streamflow Timing across Western North America, *J. Climate*, *18*(8), 1136–1155, https://doi.org/10.1175/JCLI3321.1.

Sun, N., H. Yan, M. S. Wigmosta, L. R. Leung, R. Skaggs, and Z. Hou (2019), Regional Snow Parameters Estimation for Large-domain Hydrological Applications in the Western United States, *J. Geophys. Res.*, *124*, 5296–5313, https://doi.org/10.1029/2018JD030140.

Tawfik, A. B., and A. L. Steiner (2011), The Role of Soil Ice in Land-atmosphere Coupling Over the United States: A Soil Moisture-precipitation Winter Feedback Mechanism, *J. Geophys. Res.*, *116*, D02113, https://doi.org/10.1029/2010JD014333.

Thornton, P. E., and S. W. Running (1999), An Improved Algorithm for Estimating Incident Daily Solar Radiation from Measurements of Temperature, Humidity, and Precipitation, *Agr, Forest Meteorol.*, *93*(4), 211–228, https://doi.org/10.1016/S0168-1923(98)00126-9.

Thornton, P. E., R. Shrestha, M. M. Thornton, S.-C. Kao, Y. Wei, and B. E. Wilson (2021), Gridded Daily Weather Data for North America with Comprehensive Uncertainty Quantification: Daymet Version 4, *Nature Sci. Data*, *8*, 190, https://doi.org/10.1038/s41597-021-00973-0.

Thrasher, B., E. P. Maurer, C. McKellar, and P. B. Duffy (2012), Technical Note: Bias Correcting Climate Model Simulated Daily Temperature Extremes with Quantile Mapping, *Hydrol. Earth Syst. Sci.*, *16*, 3309–3314, https://doi.org/10.5194/hess-16-3309-2012.

Tiedtke, M. (1989), A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-scale Models, *Mon. Weather Review*, *117*(8), 1779–1800, https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

Tohver, I. M., A. F. Hamlet, and S.-Y. Lee (2014), Impacts of 21st-Century Climate Change on Hydrologic Extremes in the Pacific Northwest Region of North America, *J. Am. Water Resour. As.*, *50*(6), 1461–1476, https://doi.org/10.1111/jawr.12199.

Turner, S. W. D., J. C. Steyaert, L. Condon, and N. Voisin (2021), Water Storage and Release Policies for All Large Reservoirs of Conterminous United States, *J. Hydrol.*, *603*, 126843, https://doi.org/10.1016/j.jhydrol.2021.126843.

Udall, B., and J. Overpeck (2017), The Twenty-first Century Colorado River Hot Drought and Implications for the Future, *Water Resour. Res.*, *53*, 2404–2418, https://doi.org/10.1002/2016WR019638.

Uría Martínez, R., M. M. Johnson, and R. Shan (2021), *U.S. Hydropower Market Report (January 2021 edition)*, DOE/EE-2088, ORNL/SPR-2021/1782, US Department of Energy Wind and Water Power Technologies Office, Washington, DC, https://doi.org/10.2172/1763453.

Uría Martínez, R., P. W. O'Connor, and M. M. Johnson (2015), *2014 Hydropower Market Report*, DOE/EE-1195, US Department of Energy Wind and Water Power Technologies Office, Washington, DC, https://doi.org/10.21951/1515038.

van Vliet, M. T. H., J. R. Yearsley, F. Ludwig, S. Vögele, D. P. Lettenmaier, and O. Kabat (2012), Vulnerability of US and European Electricity Supply to Climate Change, *Nature Clim. Change*, *2*, 676–681, https://doi.org/10.1038/nclimate1546.

Viger, R., and A. Bock (2014), *GIS Features of the Geospatial Fabric for National Hydrologic Modeling*, US Geological Survey, https://doi.org/10.5066/F7542KMD.

Voisin, N., H.-Y. Li, D. Ward, M. Huang, M. Wigmosta, and L. R. Leung (2013a), On an Improved Sub-regional Water Resources Management Representation for Integration into Earth System Models, *Hydrol. Earth Syst. Sci.*, *17*, 3605–3622, https://doi.org/10.5194/hess-17-3605-2013.

Voisin, N., L. Liu, M. Hejazi, T. Tesfa, H.-Y. Li, M. Huang, Y. Liu, and L. R. Leung (2013b), One-way Coupling of an Integrated Assessment Model and a Water Resources Model: Evaluation and Implications of Future Changes over the US Midwest, *Hydrol. Earth Syst. Sci.*, *17*, 4555–4575, https://doi.org/10.5194/hess-17-4555-2013.

Voisin, N., M. I. Hejazi, L. R. Leung, L. Liu, M. Huang, H.-Y. Li, and T. Tesfa (2017), Effects of Spatially Distributed Sectoral Water Management on the Redistribution of Water Resources in an Integrated Water Model, *Water Resour. Res.*, *53*, 4253–4270, https://doi.org/10.1002/2016WR019767.

Walker, M. D., and N. S. Diffenbaugh (2009), Evaluation of High-resolution Simulations of Daily-scale Temperature and Precipitation Over the United States, *Clim. Dyn.*, *33*(7–8), 1131–1147, https://doi.org/10.1007/s00382-009-0603-y.

Walton, D., and A. Hall (2018), An Assessment of High-resolution Gridded Temperature data sets Over California, *J. Clim.*, *31*(10), 3789–3810, https://doi.org/10.1175/JCLI-D-17-0410.1.

Wang, W., W. Xiao, S. Liu, N. Schultz, Y. Wang, M. Zhang, and L. Zhao (2018), Global Lake Evaporation Accelerated by Changes in Surface Energy Allocation in a Warmer Climate, *Nature Geosci.*, *11*, 410–414, https://doi.org/10.1038/s41561-018-0114-8.

Werner, A. T., and A. J. Cannon (2016), Hydrologic Extremes – An Intercomparison of Multiple Gridded Statistical Downscaling Methods, *Hydrol. Earth Syst. Sci.*, *20*, 1483–1508, https://doi.org/10.5194/hess-20-1483-2016.

Whisnant, R. B., G. W. Characklis, M. W. Doyle, V. B. Flatt, and J. D. Kern (2009), *Operating Policies and Administrative Discretion at the John H. Kerr Project*, Component study report pursuant to Section 216 of P.L. 91-611.

Williams, A. P., B. I. Cook, and J. E. Smerdon (2022), Rapid Intensification of the Emerging Southwestern North American Megadrought in 2020–2021, *Nat. Clim. Chang.*, *12*, 232–234, https://doi.org/10.1038/s41558-022-01290-z.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004), Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs, *Climatic Change*, *62*(1–3), 189–216, https://doi.org/10.1023/B:CLIM.0000013685.99609.9e.

Wrzesien, M. L., and T. M. Pavelsky (2020), Projected Changes to Extreme Runoff and Precipitation Events From a Downscaled Simulation Over the Western United States, *Front. Earth Sci.*, *7*, 355, https://doi.org/10.3389/feart.2019.00355.

Xiao, M., B. Udall, and D. P. Lettenmaier (2018), On the Causes of Declining Colorado River Streamflows, *Water Resour. Res.*, *54*, 6739–6756, https://doi.org/10.1029/2018WR023153.

Yan, D., S. E. Werners, F. Ludwig, and H. Q. Huang (2015), Hydrological Response to Climate Change: The Pearl River, China under Different RCP Scenarios, *J. Hydrol. Regional Studies*, *4*, 228–245, https://doi.org/10.1016/j.ejrh.2015.06.006.

Yan, H., N. Sun, M. Wigmosta, R. Skaggs, Z. Hou, and R. Leung (2018), Next-generation Intensity-duration-frequency Curves for Hydrologic Design in Snow-dominated Environments, *Water Resour. Res.*, *54*, 1093–1108, https://doi.org/10.1002/2017WR021290.

Yao, H., and A. Georgakakos (2001), Assessment of Folsom Lake Response to Historical and Potential Future Climate Scenarios, *J. Hydrol.*, *249*(1–4), 176–196, https://doi.org/10.1016/S0022-1694(01)00418-8.

Yoon, J. H., L. R. Leung, and J. Correia Jr (2012), Comparison of Dynamically and Statistically Downscaled Seasonal Climate Forecasts for the Cold Season Over the United States, *J. Geophy. Res.*, *117*(D21), https://doi.org/10.1029/2012JD017650.

Zelinka, M. D., T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, and K. E. Taylor (2020), Causes of Higher Climate Sensitivity in CMIP6 Models, *Geophy. Res. Lett.*, *47*, e2019GL085782, https://doi.org/10.1029/2019GL085782.

Zhao, G., and H. Gao (2018), Automatic Correction of Contaminated Images for Assessment of Reservoir Surface Area Dynamics, *Geophys. Res. Lett.*, *45*, 6092–6099, https://doi.org/10.1029/2018GL078343.

Zhao, G., and H. Gao (2019), Estimating Reservoir Evaporation Losses for the United States: Fusing Remote Sensing and Modeling Approaches, *Remote Sens. Environ.*, *226*, 109–124, https://doi.org/10.1016/j.rse.2019.03.015.

Zhao, G., H. Gao, and S.-C. Kao (2021), The Implication of Future Climate Change on the Blue Water Footprint of Hydropower in the Contiguous US, *Environ. Res. Lett.*, *16*(3), 034003, https://doi.org/10.1088/1748-9326/abd78d.

Zhou, T., S.-C. Kao, W. Xu, S. Gangrade, and N. Voisin (2022), Impacts of Climate Change on Subannual Hydropower Generation: A Multi-model Assessment of the United States Federal Hydropower Plants, *ESSOAr*. Preprint, https://doi.org/10.1002/essoar.10512358.1.

Zhou, T., B. Nijssen, H. Gao, and D. P. Lettenmaier (2016), The Contribution of Reservoirs to Global Land Surface Water Storage Variations, *J. Hydrometeor.*, *17*, 309–325, https://doi.org/10.1175/JHM-D-15-0002.1.

Zhou, T., N. Voisin, and T. Fu (2018), Non-stationary Hydropower Generation Projections Constrained by Environmental and Electricity Grid Operations Over the Western United States, *Environ. Res. Lett.*, *13*, 074035, https://doi.org/10.1088/1748-9326/aad19f.

Zoraghein, H., and B. C. O'Neill (2020), U.S. State-level Projections of the Spatial Distribution of Population Consistent with Shared Socioeconomic Pathways, *Sustainability, 12*(8), 3374, https://doi.org/10.3390/su12083374.

APPENDIX A. SECTION 9505 OF THE SECURE WATER ACT OF 2009 (PUB L. 111-11)

(a) Duty of Secretary of Energy—The Secretary of Energy, in consultation with the Administrator of each Federal Power Marketing Administration, shall assess each effect of, and risk resulting from, global climate change with respect to water supplies that are required for the generation of hydroelectric power at each Federal water project that is applicable to a Federal Power Marketing Administration.

(b) Access to Appropriate Data—

(1) IN GENERAL—In carrying out each assessment under subsection (a), the Secretary of Energy shall consult with the United States Geological Survey, the National Oceanic and Atmospheric Administration, the program, and each appropriate State water resource agency, to ensure that the Secretary of Energy has access to the best available scientific information with respect to presently observed impacts and projected future impacts of global climate change on water supplies that are used to produce hydroelectric power.

(2) ACCESS TO DATA FOR CERTAIN ASSESSMENTS—In carrying out each assessment under subsection (a), with respect to the Bonneville Power Administration and the Western Area Power Administration, the Secretary of Energy shall consult with the Commissioner to access data and other information that—

(A) is collected by the Commissioner; and

(B) the Secretary of Energy determines to be necessary for the conduct of the assessment.

(c) Report—Not later than 2 years after the date of enactment of this Act, and every 5 years thereafter, the Secretary of Energy shall submit to the appropriate committees of Congress a report that describes—

(1) each effect of, and risk resulting from, global climate change with respect to—

(A) water supplies used for hydroelectric power generation; and

(B) power supplies marketed by each Federal Power Marketing Administration, pursuant to—

(i) long-term power contracts;

(ii) contingent capacity contracts; and

(iii) short-term sales; and

(2) each recommendation of the Administrator of each Federal Power Marketing Administration relating to any change in any operation or contracting practice of each Federal Power Marketing Administration to address each effect and risk described in paragraph (1), including the use of purchased power to meet long-term commitments of each Federal Power Marketing Administration.

(d) Authority—The Secretary of Energy may enter into contracts, grants, or other agreements with appropriate entities to carry out this section.

(e) Costs—

(1) NONREIMBURSABLE—Any costs incurred by the Secretary of Energy in carrying out this section shall be nonreimbursable.

(2) PMA COSTS—Each Federal Power Marketing Administration shall incur costs in carrying out this section only to the extent that appropriated funds are provided by the Secretary of Energy for that purpose.

(f) Authorization of Appropriations—There are authorized to be appropriated such sums as are necessary to carry out this section for each of fiscal years 2009 through 2023, to remain available until expended.

APPENDIX B. LIST OF FEDERAL HYDROPOWER PLANTS MARKETED THROUGH PMAS

N	Power plant name	Power system	Owner	Generation type	Capacity (MW)	1980–2019 average annual generation (GWh/year)
BPA-1 Upper Columbia						
1	Grand Coulee	Federal Columbia River Power System (FCRPS)	Reclamation	Conventional Hydro	6,495	20,160.6
				Pumped Storage	314	
2	Hungry Horse		Reclamation	Conventional Hydro	428	920.6
3	Albeni Falls		USACE	Conventional Hydro	42	210.8
4	Libby		USACE	Conventional Hydro	525	2,075.3
BPA-2 Snake River						
5	Anderson Ranch	Federal Columbia River Power System (FCRPS)	Reclamation	Conventional Hydro	40	129.4
6	Black Canyon		Reclamation	Conventional Hydro	10.2	60.2
7	Boise R Diversion		Reclamation	Conventional Hydro	3.3	4.3
8	Minidoka		Reclamation	Conventional Hydro	27.7	94.4
9	Palisades		Reclamation	Conventional Hydro	176.4	623.8
10	Dworshak		USACE	Conventional Hydro	465	1,744.2
11	Ice Harbor		USACE	Conventional Hydro	603	1,995.9
12	Little Goose		USACE	Conventional Hydro	810	2,452.1
13	Lower Granite		USACE	Conventional Hydro	810	2,492.2
14	Lower Monumental		USACE	Conventional Hydro	810	2,438.7
BPA-3 Mid-Lower Columbia						
15	Chandler	Federal Columbia River Power System (FCRPS)	Reclamation	Conventional Hydro	12	47.7
16	Roza		Reclamation	Conventional Hydro	12.9	50.2
17	Bonneville		USACE	Conventional Hydro	1,162	4,933.0
18	Chief Joseph		USACE	Conventional Hydro	2,456.2	11,242.9
19	John Day		USACE	Conventional Hydro	2,160	9,801.4
20	McNary		USACE	Conventional Hydro	990.5	5,937.4
21	The Dalles		USACE	Conventional Hydro	1,819.7	7,222.5
BPA-4 Cascade Mountains						
22	Green Springs	Federal Columbia River Power System (FCRPS)	Reclamation	Conventional Hydro	17.2	62.7
23	Big Cliff		USACE	Conventional Hydro	18	87.4
24	Cougar		USACE	Conventional Hydro	26	124.1
25	Detroit		USACE	Conventional Hydro	100	353.1
26	Dexter		USACE	Conventional Hydro	15	70.1
27	Foster		USACE	Conventional Hydro	20	90.5
28	Green Peter		USACE	Conventional Hydro	80	234.4
29	Hills Creek		USACE	Conventional Hydro	30	148.1
30	Lookout Point		USACE	Conventional Hydro	120	313.6
31	Lost Creek		USACE	Conventional Hydro	49	266.1
WAPA-1 Upper Missouri						
32	Canyon Ferry	Pick-Sloan Eastern Division	Reclamation	Conventional Hydro	49.8	353.4
33	Big Bend		USACE	Conventional Hydro	538.3	923.1
34	Fort Peck		USACE	Conventional Hydro	179.7	922.9
35	Fort Randall		USACE	Conventional Hydro	320	1,636.4

N	Power plant name	Power system	Owner	Generation type	Capacity (MW)	1980–2019 average annual generation (GWh/year)
36	Garrison		USACE	Conventional Hydro	583.4	2,105.7
37	Gavins Point		USACE	Conventional Hydro	132.3	699.8
38	Oahe		USACE	Conventional Hydro	786.1	2,460.3
39	Yellowtail[a]	Pick-Sloan Eastern Division/Loveland	Reclamation	Conventional Hydro	268.8	805.3
WAPA-2 Loveland Projects						
40	Alcova	Loveland Area Projects	Reclamation	Conventional Hydro	41.4	111.6
41	Boysen		Reclamation	Conventional Hydro	15	64.4
42	Buffalo Bill		Reclamation	Conventional Hydro	18	70.0
43	Shoshone		Reclamation	Conventional Hydro	3	16.0
44	Heart Mountain		Reclamation	Conventional Hydro	5	19.8
45	Spirit Mountain		Reclamation	Conventional Hydro	4.5	15.0
46	Flatiron		Reclamation	Conventional Hydro	86	214.8
				Pumped Storage	8.5	
47	Big Thompson		Reclamation	Conventional Hydro	4.5	9.0
48	Fremont Canyon		Reclamation	Conventional Hydro	66.8	223.2
49	Glendo		Reclamation	Conventional Hydro	38	81.3
50	Green Mountain		Reclamation	Conventional Hydro	26	54.1
51	Guernsey		Reclamation	Conventional Hydro	6.4	17.8
52	Kortes		Reclamation	Conventional Hydro	36	135.5
53	Mary's Lake		Reclamation	Conventional Hydro	8.1	37.2
54	Estes		Reclamation	Conventional Hydro	45	100.4
55	Mount Elbert		Reclamation	Pumped Storage	200	
56	Pole Hill		Reclamation	Conventional Hydro	38.2	169.3
57	Seminoe		Reclamation	Conventional Hydro	51.6	130.5
WAPA-3 Upper Colorado						
58	Blue Mesa	Salt Lake City	Reclamation	Conventional Hydro	86.4	259.2
59	Crystal		Reclamation	Conventional Hydro	28	164.9
60	Elephant Butte		Reclamation	Conventional Hydro	27.9	78.8
61	Flaming Gorge		Reclamation	Conventional Hydro	151.8	481.2
62	Fontenelle		Reclamation	Conventional Hydro	10	49.5
63	Glen Canyon Dam		Reclamation	Conventional Hydro	1,312	4,587.3
64	Upper Molina		Reclamation	Conventional Hydro	9.9	28.9
65	Lower Molina		Reclamation	Conventional Hydro	5.6	16.7
66	McPhee		Reclamation	Conventional Hydro	1.2	3.7
67	Towaoc		Reclamation	Conventional Hydro	11.4	14.4
68	Morrow Point		Reclamation	Conventional Hydro	173.2	341.7
69	Deer Creek	Provo River	Reclamation	Conventional Hydro	4.8	24.3
WAPA-4 Lower Colorado						
70	Hoover Dam	Boulder Canyon	Reclamation	Conventional Hydro	2,078.8	4,590.8
71	Davis Dam	Parker-Davis	Reclamation	Conventional Hydro	254.8	1,212.6
72	Parker Dam		Reclamation	Conventional Hydro	120	512.3

N	Power plant name	Power system	Owner	Generation type	Capacity (MW)	1980–2019 average annual generation (GWh/year)
WAPA-5 Rio Grande						
73	Amistad Dam & Power	Falcon-Amistad	IBWC	Conventional Hydro	66	127.1
74	Falcon Dam & Power		IBWC	Conventional Hydro	31.5	69.9
WAPA-6 California						
75	Folsom	Central Valley	Reclamation	Conventional Hydro	198.6	572.9
76	Judge F Carr		Reclamation	Conventional Hydro	154.4	399.3
77	Keswick		Reclamation	Conventional Hydro	117	414.9
78	New Melones		Reclamation	Conventional Hydro	300	449.7
79	Nimbus		Reclamation	Conventional Hydro	13.4	57.8
80	ONeill		Reclamation	Pumped Storage	25.2	
81	W R Gianelli		Reclamation	Pumped Storage	424	
82	Shasta		Reclamation	Conventional Hydro	714	1,866.0
83	Spring Creek		Reclamation	Conventional Hydro	180	477.4
84	Trinity		Reclamation	Conventional Hydro	140	431.7
85	Lewiston		Reclamation	Conventional Hydro	0.35	2.4
86	Stampede	Washoe	Reclamation	Conventional Hydro	3.6	10.1
SWPA-1 Upper White, Osage, and Salt						
87	Beaver	Southwestern financially integrated projects	USACE	Conventional Hydro	112	152.0
88	Bull Shoals		USACE	Conventional Hydro	340	784.4
89	Clarence Cannon		USACE	Conventional Hydro	27	93.8
				Pumped Storage[b]	31	
90	Greers Ferry		USACE	Conventional Hydro	96	185.9
91	Harry S Truman		USACE	Pumped Storage[c]	161.4	273.8
92	Norfork		USACE	Conventional Hydro	80.4	199.2
93	Stockton		USACE	Conventional Hydro	52	52.0
94	Table Rock		USACE	Conventional Hydro	200	513.8
SWPA-2 Arkansas						
95	Dardanelle	Southwestern financially integrated projects	USACE	Conventional Hydro	160.8	613.6
96	Eufaula		USACE	Conventional Hydro	90	273.9
97	Fort Gibson		USACE	Conventional Hydro	44.8	215.9
98	Keystone		USACE	Conventional Hydro	70	267.8
99	Ozark		USACE	Conventional Hydro	100	255.0
100	Robert S Kerr		USACE	Conventional Hydro	110	555.7
101	Tenkiller Ferry		USACE	Conventional Hydro	39	120.1
102	Webbers Falls		USACE	Conventional Hydro	70	196.1
103	Broken Bow[d]		USACE	Conventional Hydro	100	154.5
SWPA-3 Ouachita, Red, and Brazos						
104	Blakely Mountain	Southwestern financially integrated projects	USACE	Conventional Hydro	75	174.3
105	DeGray		USACE	Conventional Hydro	40	81.0
				Pumped Storage (Reversible)	28	
106	Denison		USACE	Conventional Hydro	101.6	246.5
107	Narrows		USACE	Conventional Hydro	25.5	38.1
108	Whitney		USACE	Conventional Hydro	41.8	50.5

N	Power plant name	Power system	Owner	Generation type	Capacity (MW)	1980–2019 average annual generation (GWh/year)
SWPA-4 Neches						
109	Robert D Willis	Southwestern isolated projects	USACE	Conventional Hydro	8	25.6
110	Sam Rayburn		USACE	Conventional Hydro	52	119.8
SEPA-1 Kerr-Philpot						
111	John H Kerr	Kerr-Philpot	USACE	Conventional Hydro	296.8	433.5
112	Philpott Lake		USACE	Conventional Hydro	14	24.2
SEPA-2 Cumberland						
113	Barkley	Cumberland	USACE	Conventional Hydro	130	629.6
114	Center Hill		USACE	Conventional Hydro	140	334.1
115	Cheatham		USACE	Conventional Hydro	36	159.5
116	Cordell Hull		USACE	Conventional Hydro	99.9	361.6
117	Dale Hollow		USACE	Conventional Hydro	54	118.4
118	J P Priest		USACE	Conventional Hydro	28	65.0
119	Laurel		USACE	Conventional Hydro	70	63.6
120	Old Hickory		USACE	Conventional Hydro	103.7	459.5
121	Wolf Creek		USACE	Conventional Hydro	270	872.0
SEPA-3 GA/AL/SC						
122	Allatoona	GA/AL/SC	USACE	Conventional Hydro	86.6	125.8
123	Buford		USACE	Conventional Hydro	131.2	163.7
124	Carters		USACE	Conventional Hydro	250	454.2
				Pumped Storage	250	
125	Hartwell Lake		USACE	Conventional Hydro	420	419.8
126	J Strom Thurmond		USACE	Conventional Hydro	361.9	610.6
127	Millers Ferry		USACE	Conventional Hydro	101.1	335.7
128	Jones Bluff		USACE	Conventional Hydro	82	297.3
129	Richard B Russell		USACE	Conventional Hydro	300	636.2
				Pumped Storage	328	
130	Walter F George		USACE	Conventional Hydro	168	382.8
131	West Point		USACE	Conventional Hydro	73.3	173.3
SEPA-4 Jim Woodruff						
132	J Woodruff	Jim Woodruff	USACE	Conventional Hydro	43.5	202.1

[a] Two of the four Yellowtail units are marketed as a Pick-Sloan Eastern Division resource and two are marketed as a Loveland Area Projects resource. For the purposes of this analysis, the entire Yellowtail plant is included in the Pick-Sloan Eastern Division.

[b] The pumpback feature of the reversible unit at Cannon has not been used in regular operation (other than initial tests). As the reservoir has to be significantly low for the pumpback to function, it has not been the practice to use the feature. The reversible unit is used regularly like conventional hydro.

[c] Although Harry S. Truman has the capability of pumped storage through multiple reversible units, it is used as conventional hydro because of state objections to the use of the pumpback function. It is currently not available as a pumped storage project.

[d] Broken Bow in the Red River Basin is included in SWPA-2 because of an interconnected system reason.

APPENDIX C. EVALUATION OF CLIMATE MODEL

The selection of a wide range of matrices is to ensure a comprehensive evaluation of spatial and temporal distributions of precipitation and temperature across CMIP6 GCMs. Additionally, the representation of three natural modes of climate variability, namely North Atlantic Oscillation (NAO), El Niño-Southern Oscillation (ENSO) and Pacific Decadal Oscillation (PDO), and their impacts on the distribution of DJF (winter) precipitation and temperature have also been evaluated. These three modes are known to have an impact on the climate variability over North America at seasonal to decadal time scales and typically their influences are the strongest during the wintertime.

All the GCMs and reference data sets are remapped to a common 1° latitude-longitude grid for comparisons. The 60 evaluation criteria include both standalone and derived matrices. For each standalone matrix, model disagreement is calculated as percent departure from the reference data. Several derived matrices are based on the calculation of Taylor score, which is a combination of RMSE, bias, and pattern correlation. Model disagreements for each of the three statistical measures are calculated as percent departures from the reference data, and their averages represent the Taylor score for that matrix. For the diurnal cycle matrix, Taylor score is separately calculated for four seasons which is subsequently averaged to get the final measure. Similarly, Taylor score for the matrix representing precipitation from moderate to extreme events is also based on the average of individual Taylor scores for precipitation from events exceeding 75th, 90th, 95th, and 99th percentiles of precipitation. The combination of all seasons in one matrix for diurnal cycle and four kinds of events ranging from moderate to extreme precipitation magnitudes in one matrix is due to their relatively high correlations. The dispersion matrix is an average of Taylor scores across 20 indices, which are calculated after transforming the 3-dimensional (time, latitude, longitude) data into 1 dimension.[29] The PDO index is based on the first EOF of sea surface temperature over Northern Pacific (20°N–70°N, 110°E–260°E). The ENSO index is based on the sea surface temperature anomalies over the Nino3.4 region (5°S–5°N, 170°W–120°W). In both cases, temporally varying global mean is removed from the sea surface temperatures to remove any impact of global warming. The NAO index is based on the first EOF of detrended sea level pressure over Northern Atlantic (20°N–80°N, 90°W–40°E). The details of these indices can be found online at National Center for Atmospheric Research climate data guide.[30] The pattern correlation is used as a measure for GCMs skill in representing these modes of variability and their influences on the distribution of precipitation and temperature across the CONUS. For model evaluation, the entire CONUS is divided in four parts (north, east, west, and south; Figure C.1) based on grouped HUC2 regions utilized by Naz et al. (2016). All matrices are calculated separately for each of the four regions, which are subsequently averaged to calculate disagreements at the CONUS scale for each model.

Although attempts have been made to use distinct set of matrices for GCMs evaluation, high correlations are still possible given the nature of climate in the coupled Earth system model. Therefore, following a technique proposed by Sanderson et al. (2017) for assigning weights to GCMs based on their uniqueness, a weighting methodology was devised in which highly correlated matrices are down-weighted. First, all percent departures from the reference data for all matrices are converted to normalized relative errors as following:

$$RE_{G,i} = \frac{PD_{G,i} - min(PD_{G_{all},i})}{max(PD_{G_{all},i}) - min(PD_{G_{all},i})} \quad \text{(C.1)}$$

where $RE_{G,i}$ and $PD_{G,i}$ represents the relative error and percent departure from the reference data for GCM G in matrix i, respectively. $PD_{G_{all},i}$ represents the array of percent departures from the reference

[29] https://www.ncl.ucar.edu/Document/Functions/Contributed/stat_dispersion.shtml
[30] https://climatedataguide.ucar.edu/climate-data

data for across all GCMs for that matrix. Second, pairwise Pearson linear cross correlations are calculated for all matrices, which are converted into a distance measure as following:

$$C^*_{i,j} = 1 - abs(C_{i,j}) \tag{C.2}$$

where $C_{i,j}$ and $C^*_{i,j}$ represent correlation and correlation-based distance between matrix i and matrix j, respectively. The small magnitude of $C^*_{i,j}$ reflects high correspondence between the matrices and vice versa. Furthermore, we calculate Similarity Score (SS) for each matrix as following:

$$SS_{i,j} = e^{-\left(\frac{C^*_{i,j}}{D_x}\right)} \tag{C.3}$$

where D_x is a tunable parameter representing the radius of similarity that determines the correlations-based distances over which a matrix can be considered redundant. Some covariance between different spatiotemporal characteristics of prognostic variables or between the prognostic and diagnostic variables is acceptable and unavoidable in a coupled Earth system; therefore, the goal is to target only those metrics that exhibit correlations to such an extent that those measures effectively become redundant. Therefore, 0.2 was used for D_x as it only down-weights those metrics that exhibit very high correlations in the four regions. It only down-weights matrices that exhibit very high correlations. SS value ranges between 0 and 1, as a matrix uniqueness decreases with $SS \to 1$. Note that $C^*_{i,j}$ for ~0.8% of the total pairwise correlations were less than 0.2 in each region while 5–7% of $C^*_{i,j}$ were lower than 0.5 across the four regions. Next, for each matrix, the effective redundance (ER) can be calculated as following:

$$ER_i = 1 + \sum_{j \neq i}^{n} SS_{i,j} \tag{C.4}$$

The inverse of the ER_i provides the weight for that matrix. Finally, the average weighted relative error for each GCM is calculated as following:

$$RE^*_G = \sum_{i=1}^{m} (ER_i)^{-1} RE_{G,i} \tag{C.5}$$

These weighted relative errors (RE^*_G) are calculated separately for each of the four subregions of CONUS subregions and subsequently averaged to provide the average weighted relative error used in the simple averaging technique to calculate relative ranks of each GCM in the analyzed ensemble. The GCM with the lowest weighted relative error ranks at the top, whereas the GCM with the highest weighted relative error ranks at the bottom. In the multivariate EOF analyses, models are ranked using sum of their Euclidean distances from the observations in the PC space, as following:

$$D(O,G) = \sqrt{\sum_{i=1}^{n} (G_i - O_i)^2} \tag{C.6}$$

where $D(O,G)$ represents the Euclidean distance of GCM G from reference data O as a sum of the distances over n PCs. The sensitivity of GCMs ranking to the number PCs used in the calculation of Euclidean distances substantially diminishes after first 10 modes (Table C.1.). Therefore, distances between individual GCMs and observations were calculated using the truncated set of first 10 modes.

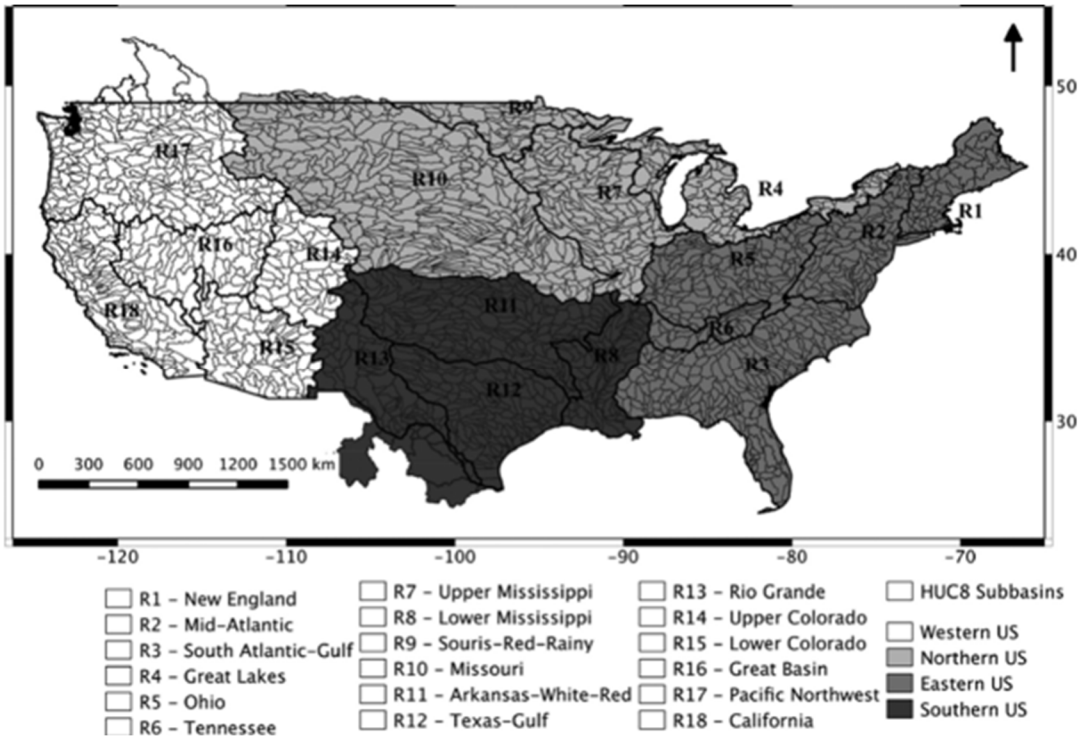


Figure C.1. Four CONUS GCM evaluation regions based on grouped HUC2s utilized by Naz et al. (2016).

Table C.1. List of GCMs evaluation metrics

GCMs Evaluation Metrics			
1. Amplitude[a] Mean P[1]	**2.** Amplitude Mean T[2]	**3.** Amplitude Mean Tmax[3]	**4.** Amplitude Mean Tmin[4]
5. Amplitude Standard Deviation P	**6.** Amplitude Standard Deviation T	**7.** Amplitude Standard Deviation Tmax	**8.** Amplitude Standard Deviation Tmin
9. Timing[b] of Peak P	**10.** Timing of Peak T	**11.** Timing of Peak Tmax	**12.** Timing of Peak Tmin
13. Annual Mean Standard Deviation of P	**14.** Annual Mean Standard Deviation of T	**15.** Annual Mean Standard Deviation of Tmax	**16.** Annual Mean Standard Deviation of Tmin
17. DJF[5] (Taylor Stats) P	**18.** DJF (Taylor Stats) T	**19.** DJF (Taylor Stats) Tmax	**20.** DJF (Taylor Stats) Tmin
21. MAM[6] (Taylor Stats) P	**22.** MAM (Taylor Stats) T	**23.** MAM (Taylor Stats) Tmax	**24.** MAM (Taylor Stats) Tmin
25. JJA[7] (Taylor Stats) P	**26.** JJA (Taylor Stats) T	**27.** JJA (Taylor Stats) Tmax	**28.** JJA (Taylor Stats) Tmin
29. SON[8] (Taylor Stats) P	**30.** SON (Taylor Stats) T	**31.** SON (Taylor Stats) Tmax	**32.** SON (Taylor Stats) Tmin
33. (Taylor Stats) Inter-quartile Range[c] P	**34.** (Taylor Stats) Inter-quartile Range Tmax	**35.** (Taylor Stats) Inter-quartile Range Tmin	**36.** (Taylor Stats) Diurnal T
37. (Taylor Stats) P from Moderate to Heavy Events	**38.** (Taylor Stats) Wet Days[d]	**39.** (Taylor Stats) P Intensity	**40.** (Taylor Stats) Summer Days[e]
41. (Taylor Stats) Ice Days[f]	**42.** (Taylor Stats) Tropical Nights[g]	**43.** (Taylor Stats) Frost Days[h]	**44.** Dispersion[i] P
45. Dispersion T	**46.** Dispersion Tmin	**47.** Dispersion Tmax	**48.** ENSO Amplitude
49. PDO Pattern	**50.** NAO Pattern	**51.** NAO Correlation with DJF P	**52.** NAO Correlation with DJF T
53. PDO Correlation with DJF P	**54.** PDO Correlation with DJF T	**55.** ENSO Correlation with DJF P	**56.** ENSO Correlation with DJF T
57. (Taylor Stats) 500mb Geopotential Height DJF	**58.** (Taylor Stats) 500mb Geopotential Height JJA	**59.** (Taylor Stats) Sea Level Pressure DJF	**60.** (Taylor Stats) Sea Level Pressure JJA
Taylor Stats			
RMSE	Bias	Pattern Correlation	
Dispersion (based on 1-dimesnional time series of time x latitude x longitude)			
Lower Octile	Lower Sextile	Lower Quartile	Lower Tritile
Median	Upper Tritile	Upper Quartile	Upper Sextile
Upper Octile	Upper Dectile	Maximum	Range
0.1**st** Percentile	1st Percentile	5th Percentile	95th Percentile
99th Percentile	99.9th Percentile	Skewness	Kurtosis

[1]P = Precipitation, [2]T = Temperature, [3]Tmax = Maximum Temperature, [4]Tmin = Minimum Temperature, [5]DJF = December-January-February, [6]MAM = March-April-May, [7]JJA = June-July-August, [8]SON = September-October-November, [9]ENSO = El Niño-Southern Oscillation), [10]PDO = Pacific Decadal Oscillation, [11]NAO = North Atlantic Oscillation

[a]Amplitude = Difference between maximum and minimum in a monthly annual cycle

[b]Timing = Month Index with the maximum of the annual cycle

[c]Inter-quartile range = Difference between the 75th and 25th percentile of daily values in a year

[d]Wet days = Days with accumulated P ≥ 1.0 mm

[e]Summer days = Days with T ≥ 25 °C (77 °F)

[f]Ice days = Days with Tmax < 0 °C

[g]Tropical nights = Days with Tmin > 20 °C (68 °F)

[h]Frost days = Days with Tmin < 0 °C

[h]Dispersion = Spatiotemporal distribution of monthly data, calculated as an average of the Taylor score of 20 indices

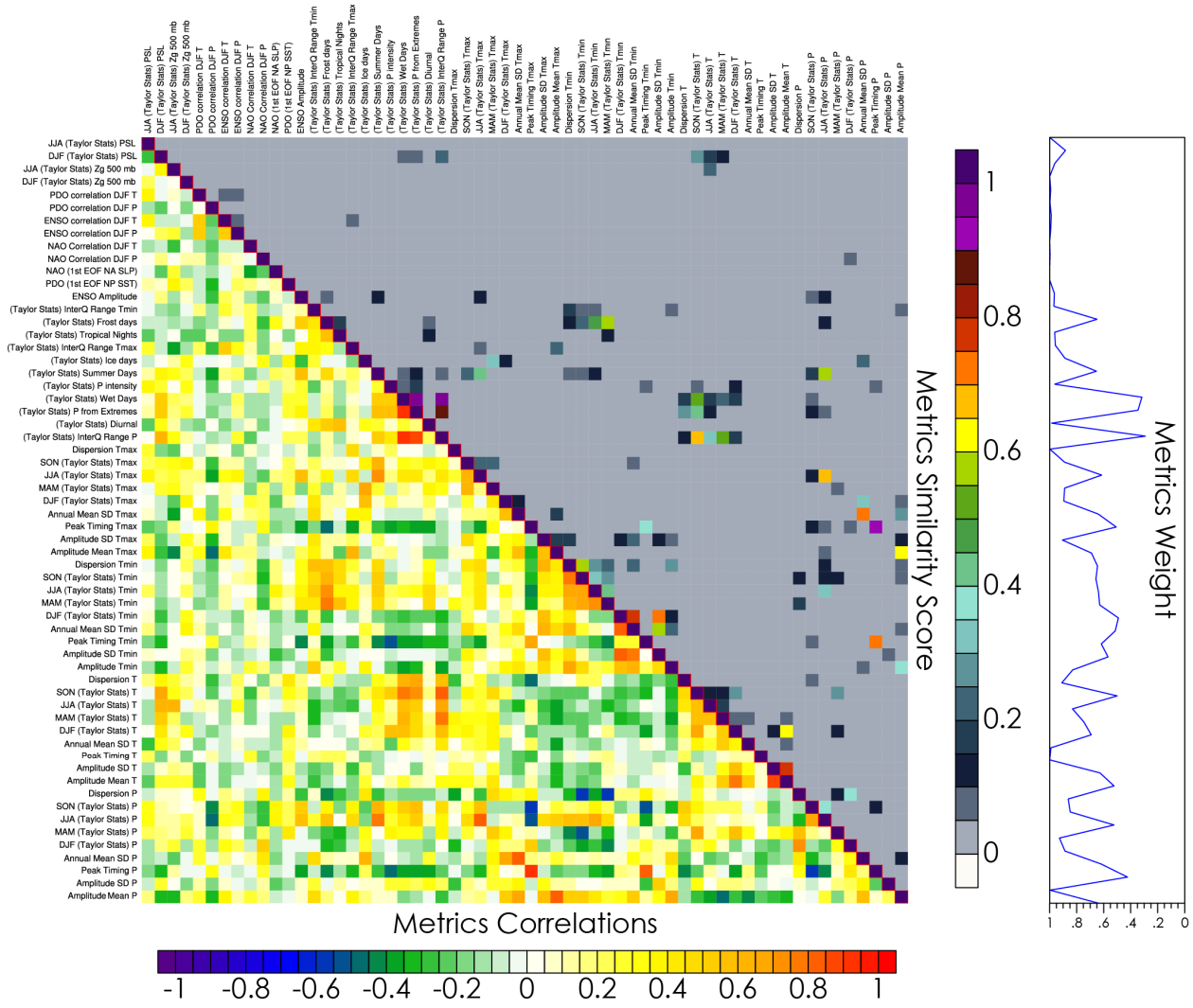


Figure C.2. The pairwise metrics correlation (bottom triangle) and the corresponding similarity score (top triangle) over the eastern United States. Metrics with high correlations exhibit a high similarity score and are down-weighted. The line plot on the right shows the overall weight for each metric.

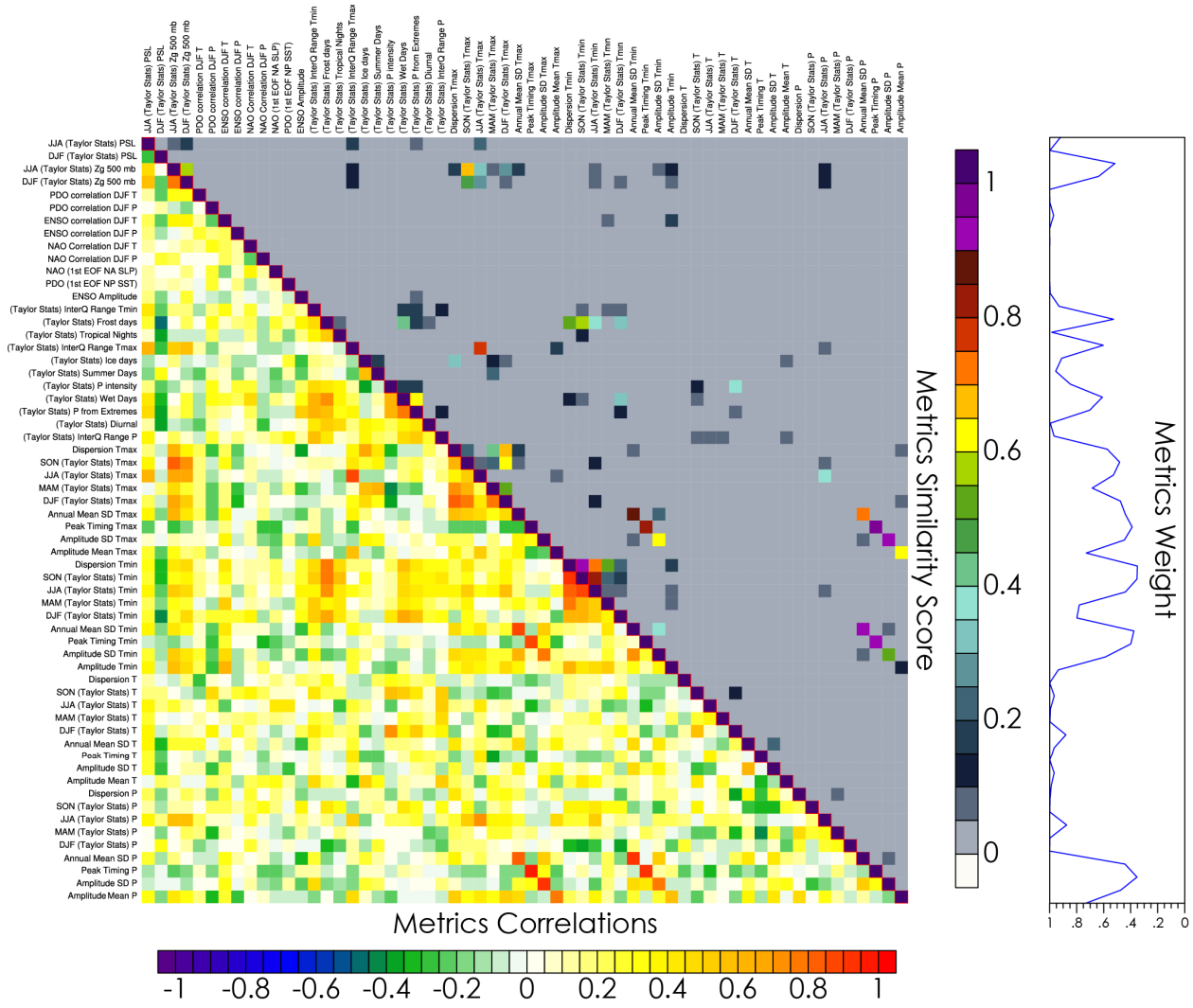


Figure C.3. The pairwise metrics correlation (bottom triangle) and the corresponding similarity score (top triangle) over the western United States. Metrics with high correlations exhibit a high similarity score and are down-weighted. The line plot on the right shows the overall weight for each metric.

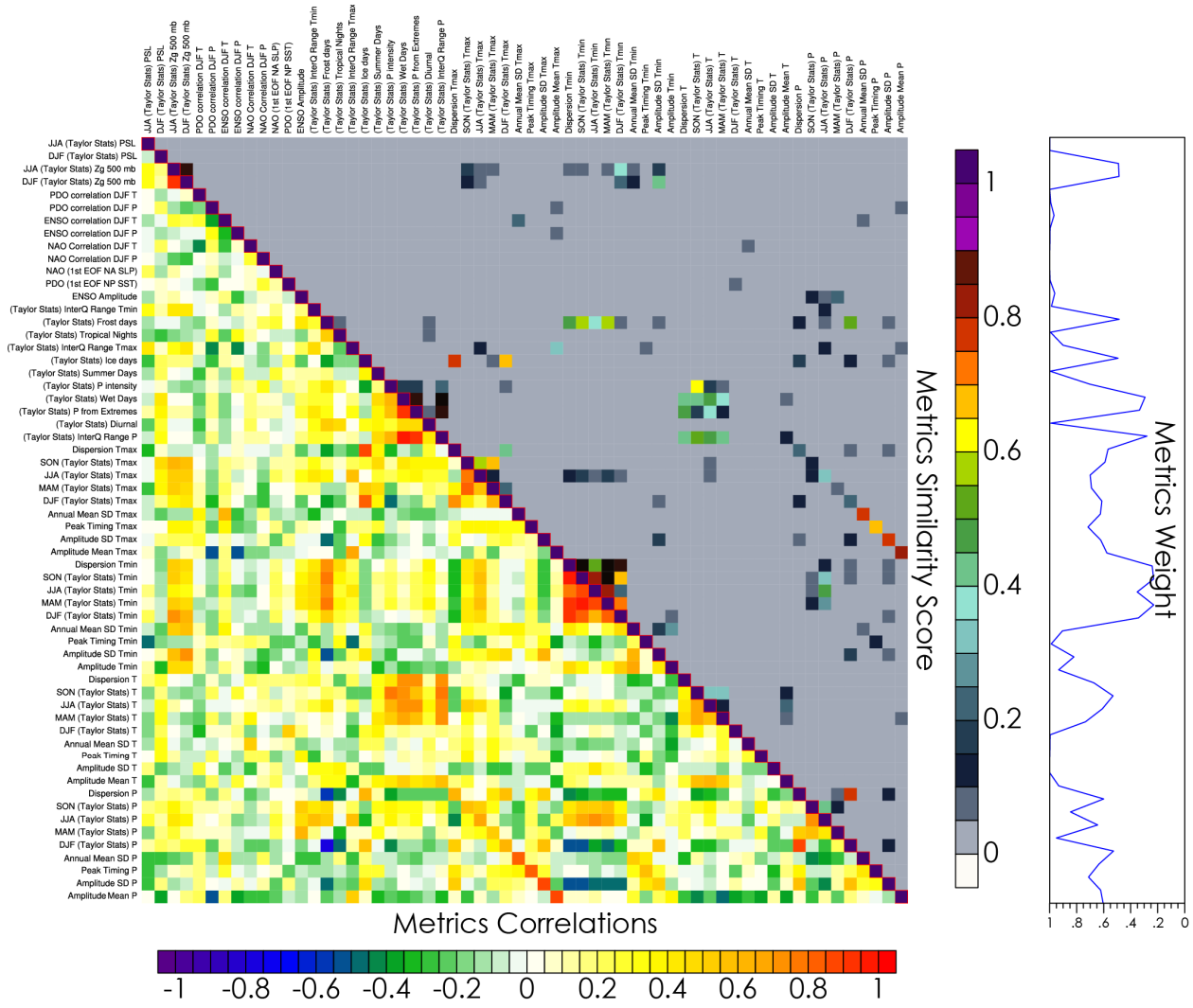


Figure C.4. The pairwise metrics correlation (bottom triangle) and the corresponding similarity score (top triangle) over the southern United States. Metrics with high correlations exhibit a high similarity score and are down-weighted. The line plot on the right shows the overall weight for each metric.

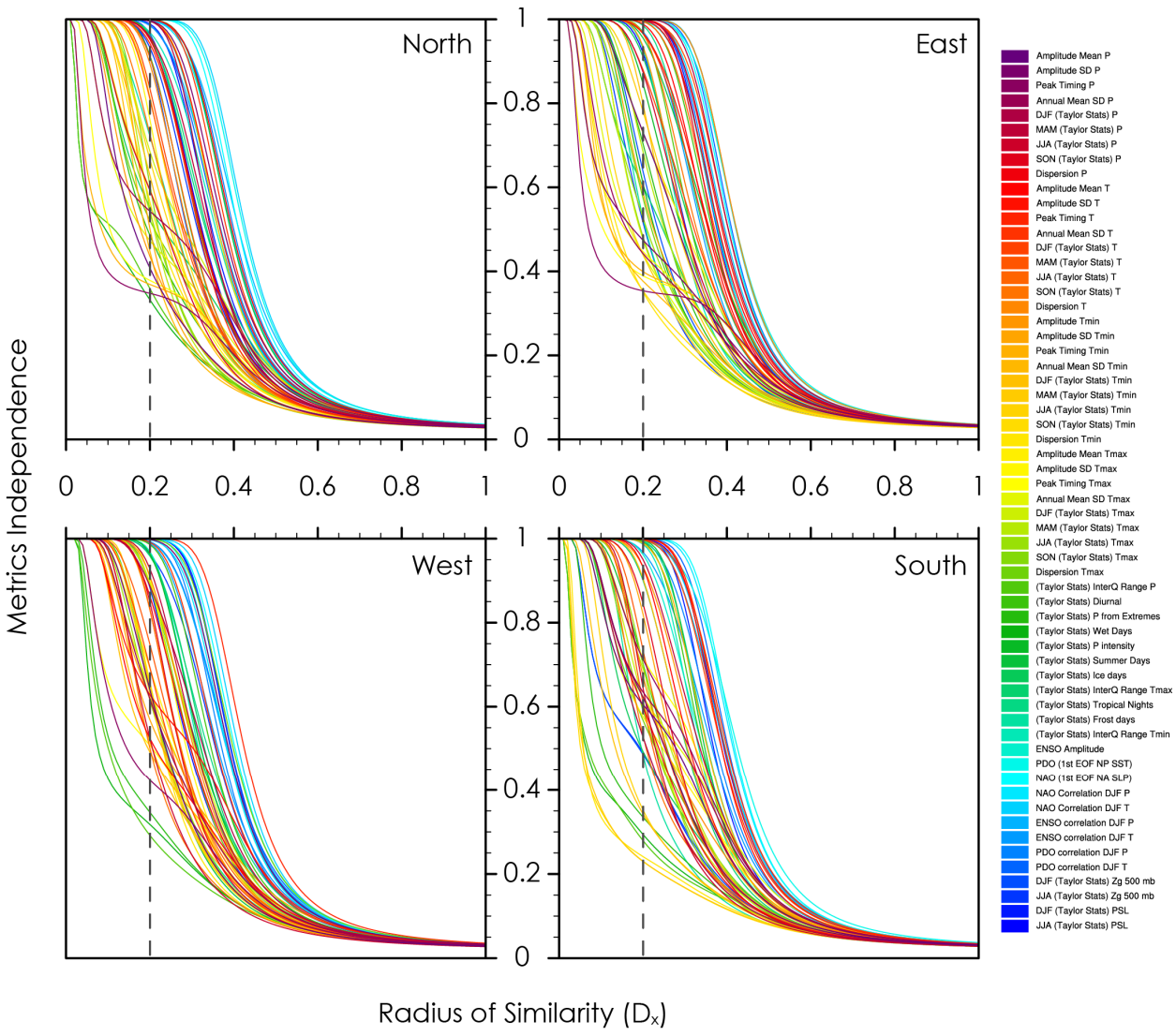


Figure C.5. Metrics independence weights ($(ER_i)^{-1}$) as a function of the radius of their similarity (D_x). The gray vertical line represents the value of D_x used in the calculation of similarity scores.

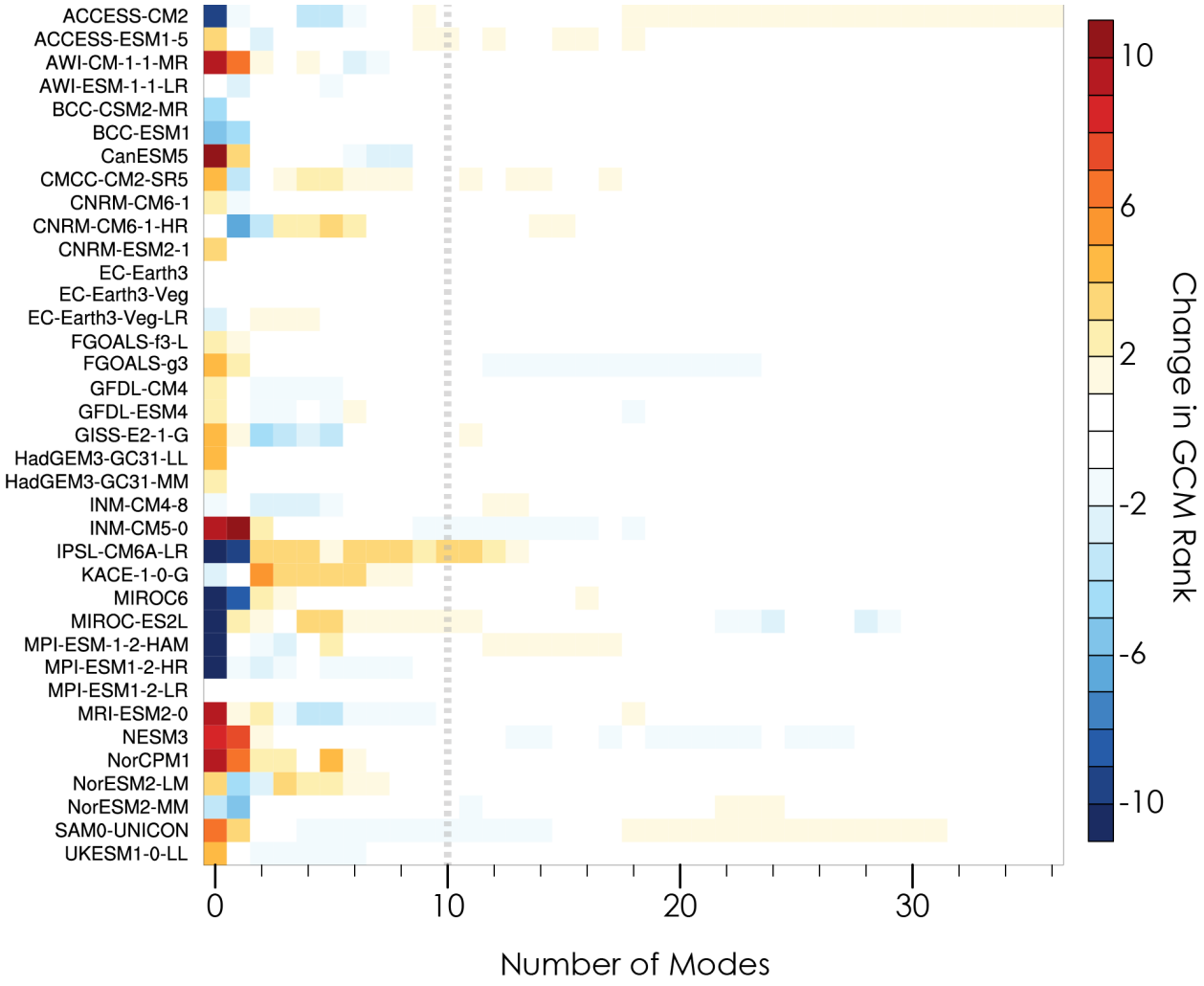


Figure C.6. The change in the ranking of GCMs at different truncation levels in EOF analyses. The gray vertical line indicates the selected truncation level in these analyses.

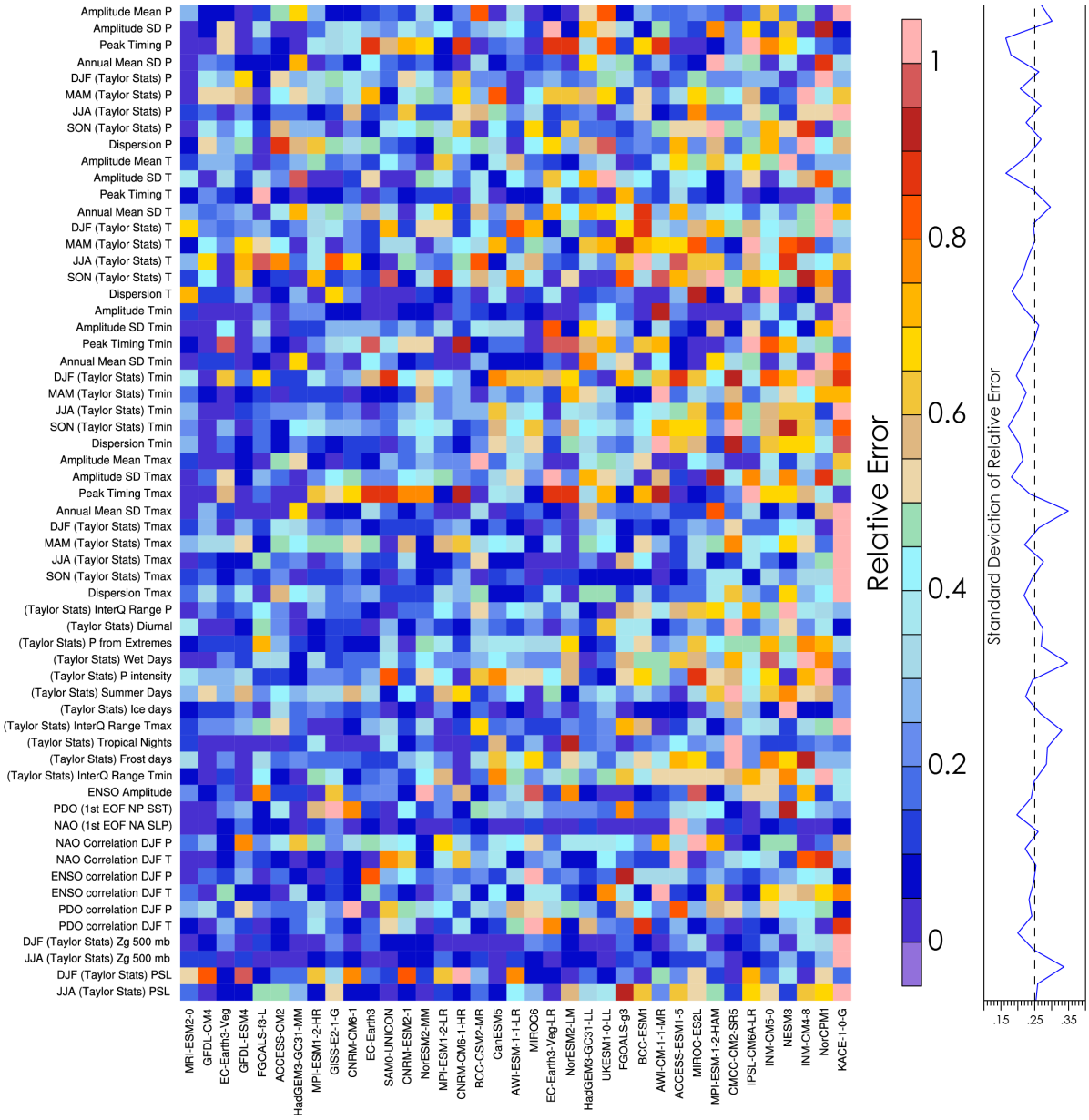


Figure C.7. The normalized relative disagreement of GCMs from the observations over the eastern United States for each metric (left) and their overall standard deviation across evaluated GCMs (right).

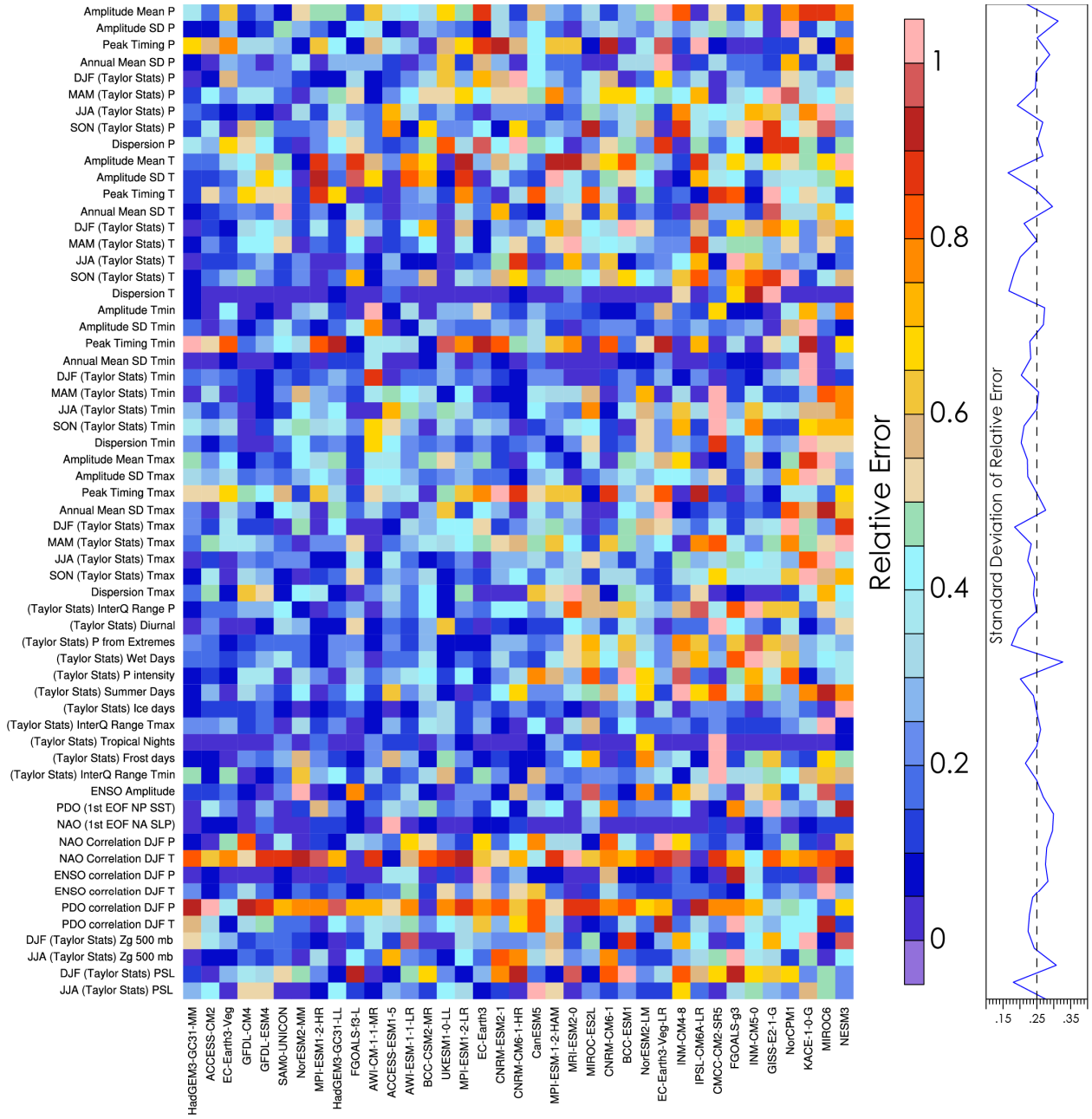


Figure C.8. The normalized relative disagreement of GCMs from the observations over the western United States for each metric (left) and their overall standard deviation across evaluated GCMs (right).

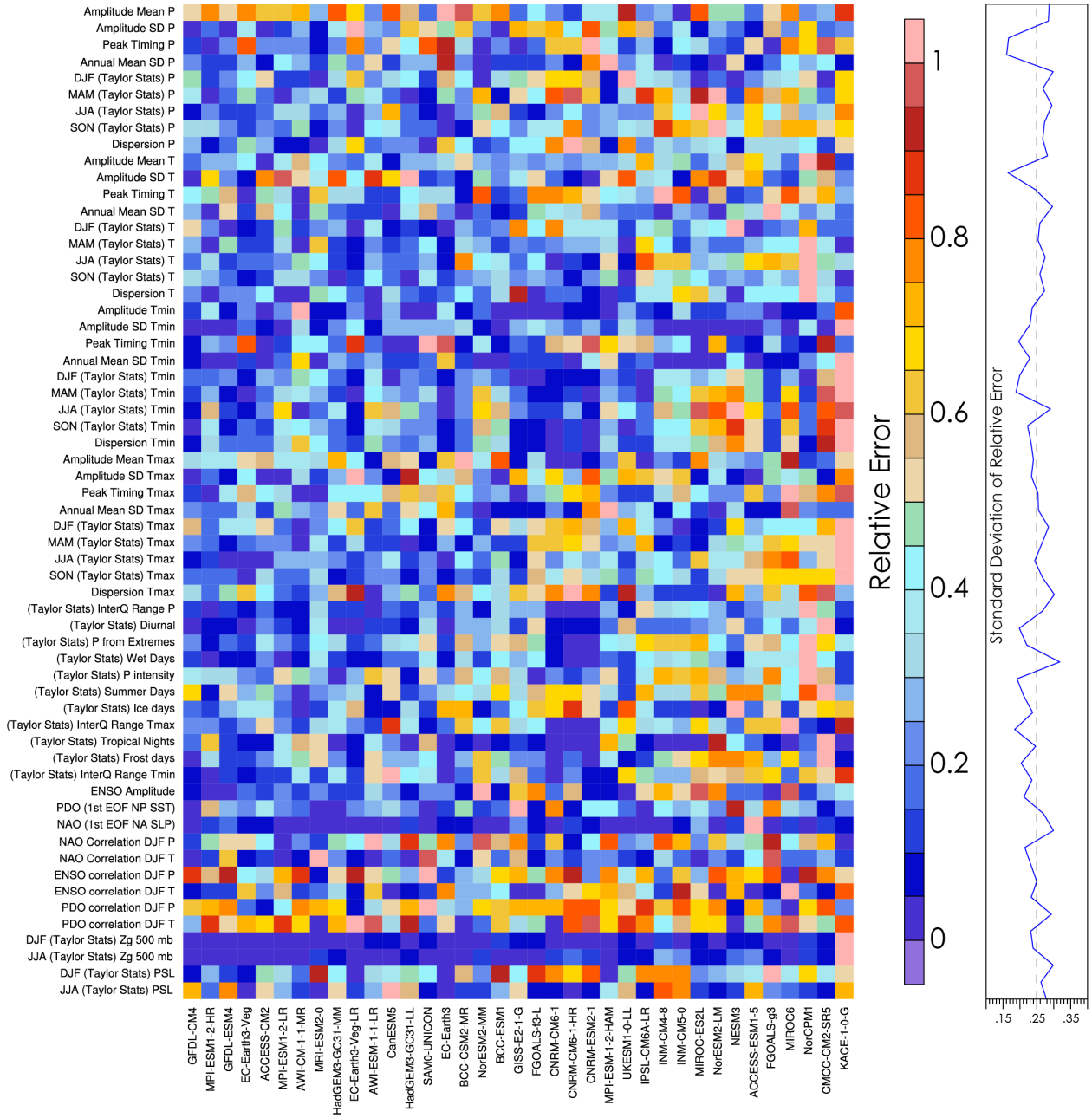


Figure C.9. The normalized relative disagreement of GCMs from the observations over the southern United States for each metric (left) and their overall standard deviation across evaluated GCMs (right).

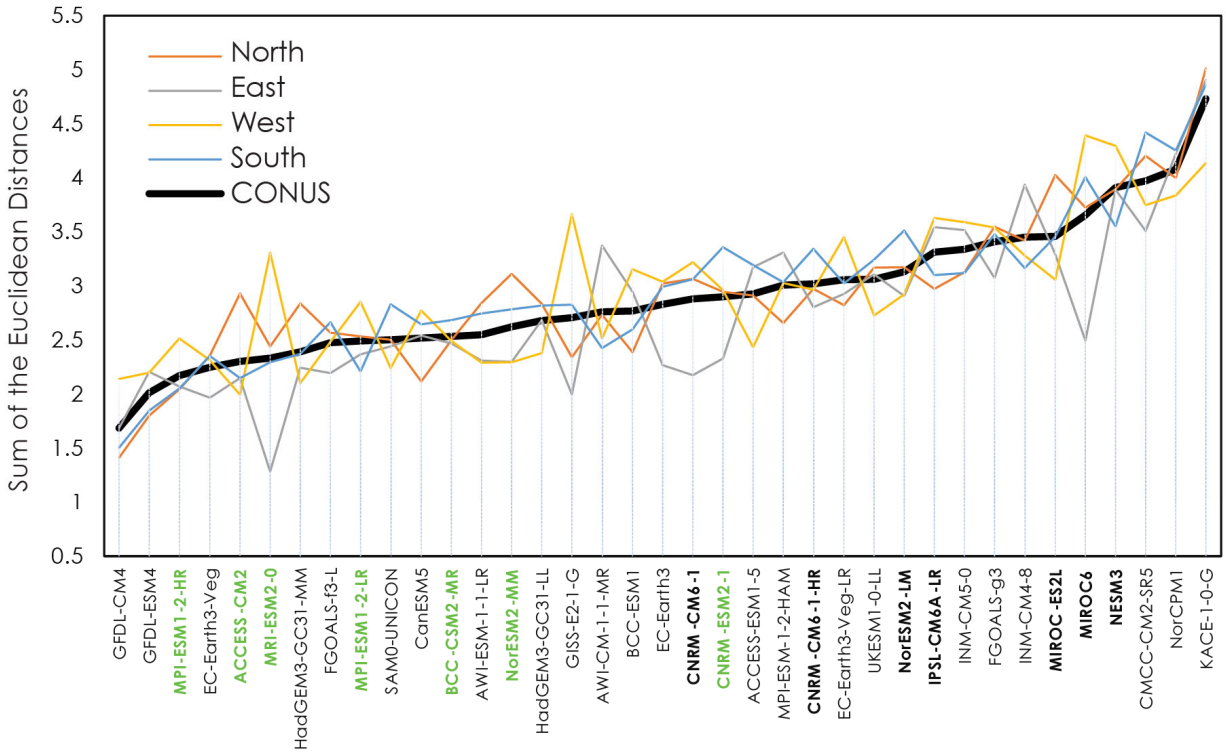


Figure C.10. The ranking of GCMs using the EOF analyses. The bold font represents GCMs that have data available for dynamical downscaling and green represents the selected GCMs.

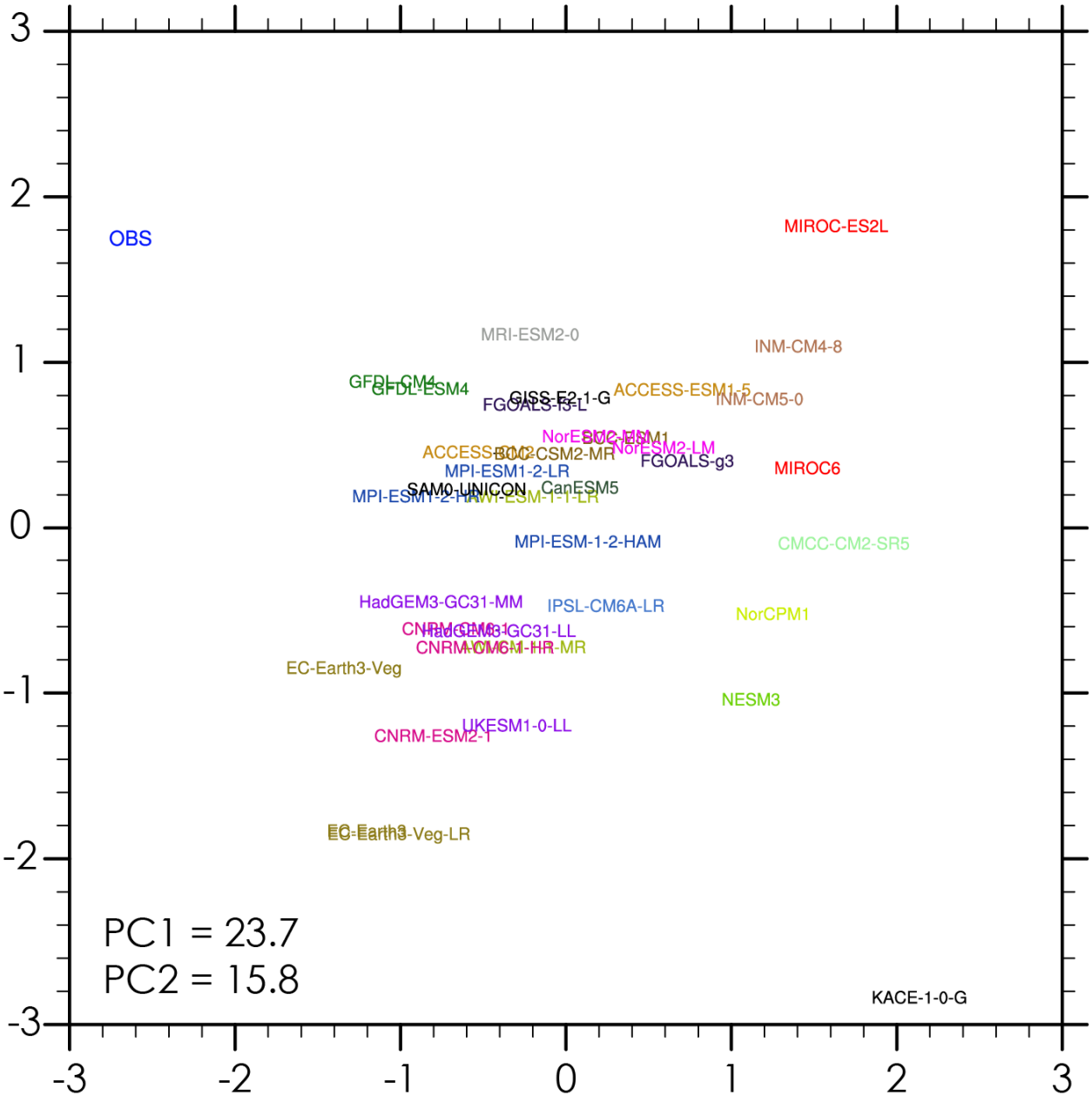


Figure C.11. The normalized scatter plot of PC1 and PC2 that explains that models from same institute tend to cluster in same area in the PC space. Relatively more spread along x-axis reflects that PC1 explains more variance (23.7%) than PC2 (15.8%). Models from same institute are shown in same color.

APPENDIX D. EVALUATION OF DOWNSCALED CLIMATE PROJECTIONS

Table D.1. Pattern correlation with respect to Daymet

Model	Daily maximum temperature (°F)					Daily minimum temperature (°F)				
	Annual	Winter	Spring	Summer	Fall	Annual	Winter	Spring	Summer	Fall
Raw CMIP6 GCM										
ACCESS-CM2	0.946	0.968	0.938	0.842	0.948	0.946	0.939	0.949	0.933	0.951
BCC-CSM2-MR	0.934	0.970	0.931	0.809	0.935	0.943	0.958	0.944	0.896	0.941
CNRM-ESM2-1	0.921	0.958	0.902	0.815	0.931	0.947	0.955	0.947	0.919	0.957
MPI-ESM1-2-HR	0.952	0.978	0.946	0.865	0.953	0.956	0.960	0.960	0.926	0.959
MRI-ESM2-0	0.943	0.954	0.920	0.860	0.956	0.959	0.938	0.955	0.949	0.963
NorESM2-MM	0.946	0.980	0.933	0.807	0.937	0.947	0.966	0.943	0.892	0.944
Dynamical (RCM)										
ACCESS-CM2	0.927	0.927	0.900	0.882	0.937	0.963	0.953	0.959	0.943	0.938
BCC-CSM2-MR	0.929	0.955	0.916	0.882	0.925	0.971	0.973	0.969	0.951	0.955
CNRM-ESM2-1	0.885	0.947	0.867	0.770	0.888	0.963	0.966	0.955	0.951	0.945
MPI-ESM1-2-HR	0.948	0.961	0.921	0.917	0.945	0.964	0.969	0.959	0.944	0.942
MRI-ESM2-0	0.930	0.950	0.907	0.880	0.935	0.964	0.955	0.967	0.963	0.950
NorESM2-MM	0.932	0.953	0.925	0.856	0.931	0.955	0.965	0.960	0.913	0.933
Dynamical (RCM) corrected by Daymet										
ACCESS-CM2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BCC-CSM2-MR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CNRM-ESM2-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MPI-ESM1-2-HR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MRI-ESM2-0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
NorESM2-MM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Statistical (DBCCA) trained by Daymet										
ACCESS-CM2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BCC-CSM2-MR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CNRM-ESM2-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MPI-ESM1-2-HR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MRI-ESM2-0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
NorESM2-MM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Model	Daily precipitation (in./day)					T95 (°F)	T05 (°F)	P95 (in./day)	Wetdays	Frostdays
	Annual	Winter	Spring	Summer	Fall	Annual	Annual	Annual	Annual	Annual
Raw CMIP6 GCM										
ACCESS-CM2	0.852	0.864	0.811	0.773	0.869	0.777	0.929	0.797	0.812	0.932
BCC-CSM2-MR	0.781	0.774	0.766	0.783	0.801	0.674	0.954	0.696	0.734	0.926
CNRM-ESM2-1	0.836	0.838	0.838	0.831	0.789	0.803	0.944	0.823	0.773	0.949
MPI-ESM1-2-HR	0.861	0.823	0.849	0.878	0.824	0.814	0.949	0.654	0.899	0.952
MRI-ESM2-0	0.806	0.801	0.740	0.776	0.851	0.814	0.934	0.742	0.683	0.907
NorESM2-MM	0.870	0.885	0.853	0.845	0.850	0.698	0.959	0.751	0.823	0.913
Dynamical (RCM)										
ACCESS-CM2	0.747	0.835	0.719	0.668	0.728	0.841	0.945	0.740	0.555	0.944
BCC-CSM2-MR	0.763	0.799	0.741	0.722	0.767	0.854	0.964	0.737	0.692	0.955
CNRM-ESM2-1	0.673	0.828	0.665	0.585	0.589	0.743	0.956	0.739	0.516	0.948
MPI-ESM1-2-HR	0.837	0.782	0.854	0.798	0.801	0.900	0.959	0.781	0.790	0.949
MRI-ESM2-0	0.779	0.775	0.695	0.738	0.839	0.856	0.946	0.788	0.700	0.935
NorESM2-MM	0.831	0.862	0.797	0.785	0.773	0.809	0.949	0.797	0.697	0.946
Dynamical (RCM) corrected by Daymet										
ACCESS-CM2	1.000	1.000	0.999	0.999	0.999	0.999	0.991	0.996	1.000	0.995
BCC-CSM2-MR	1.000	1.000	0.999	0.999	0.999	0.999	0.994	0.996	1.000	0.996
CNRM-ESM2-1	1.000	1.000	0.999	1.000	0.999	0.999	0.994	0.996	1.000	0.996
MPI-ESM1-2-HR	1.000	1.000	0.999	0.999	0.999	0.999	0.993	0.996	1.000	0.996
MRI-ESM2-0	1.000	1.000	0.999	0.999	0.999	0.999	0.994	0.996	1.000	0.997
NorESM2-MM	1.000	1.000	0.999	0.999	0.999	0.999	0.996	0.996	1.000	0.997
Statistical (DBCCA) trained by Daymet										
ACCESS-CM2	1.000	1.000	0.999	0.999	0.999	0.999	0.998	0.996	1.000	0.999
BCC-CSM2-MR	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.996	1.000	0.999
CNRM-ESM2-1	1.000	1.000	0.999	0.999	0.999	0.999	0.998	0.996	1.000	0.999
MPI-ESM1-2-HR	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.996	1.000	0.999
MRI-ESM2-0	1.000	1.000	0.999	0.999	0.999	0.999	0.998	0.996	1.000	0.999
NorESM2-MM	1.000	1.000	0.999	0.999	0.999	0.999	0.998	0.996	1.000	0.998

Table D.2. Pattern correlation with respect to Livneh

Model	Daily maximum temperature (°F)					Daily minimum temperature (°F)				
	Annual	Winter	Spring	Summer	Fall	Annual	Winter	Spring	Summer	Fall
Raw CMIP6 GCM										
ACCESS-CM2	0.945	0.960	0.939	0.850	0.947	0.929	0.923	0.938	0.919	0.930
BCC-CSM2-MR	0.934	0.962	0.931	0.813	0.936	0.928	0.938	0.933	0.889	0.930
CNRM-ESM2-1	0.919	0.951	0.899	0.810	0.929	0.929	0.937	0.934	0.906	0.935
MPI-ESM1-2-HR	0.953	0.972	0.949	0.881	0.955	0.942	0.947	0.947	0.914	0.945
MRI-ESM2-0	0.942	0.950	0.917	0.863	0.955	0.946	0.924	0.946	0.939	0.942
NorESM2-MM	0.947	0.974	0.935	0.816	0.940	0.926	0.946	0.924	0.878	0.929
Dynamical (RCM)										
ACCESS-CM2	0.923	0.913	0.899	0.877	0.932	0.945	0.925	0.947	0.937	0.923
BCC-CSM2-MR	0.927	0.949	0.912	0.876	0.923	0.964	0.963	0.965	0.947	0.953
CNRM-ESM2-1	0.881	0.940	0.862	0.755	0.882	0.954	0.953	0.950	0.949	0.940
MPI-ESM1-2-HR	0.946	0.952	0.921	0.915	0.942	0.954	0.954	0.949	0.942	0.937
MRI-ESM2-0	0.928	0.943	0.903	0.874	0.931	0.956	0.941	0.960	0.960	0.942
NorESM2-MM	0.930	0.948	0.922	0.851	0.930	0.953	0.959	0.959	0.914	0.935
Dynamical (RCM) corrected by Livneh										
ACCESS-CM2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BCC-CSM2-MR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CNRM-ESM2-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MPI-ESM1-2-HR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MRI-ESM2-0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
NorESM2-MM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Statistical (DBCCA) trained by Livneh										
ACCESS-CM2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BCC-CSM2-MR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CNRM-ESM2-1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MPI-ESM1-2-HR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MRI-ESM2-0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
NorESM2-MM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Model	Daily precipitation (in./day)					T95 (°F)	T05 (°F)	P95 (in./day)	Wetdays	Frostdays
	Annual	Winter	Spring	Summer	Fall	Annual	Annual	Annual	Annual	Annual
Raw CMIP6 GCM										
ACCESS-CM2	0.832	0.838	0.785	0.770	0.849	0.787	0.918	0.776	0.772	0.906
BCC-CSM2-MR	0.754	0.746	0.734	0.775	0.774	0.676	0.944	0.698	0.694	0.899
CNRM-ESM2-1	0.813	0.811	0.820	0.822	0.762	0.801	0.934	0.791	0.736	0.906
MPI-ESM1-2-HR	0.829	0.795	0.820	0.870	0.787	0.831	0.941	0.665	0.863	0.920
MRI-ESM2-0	0.779	0.776	0.719	0.763	0.824	0.821	0.920	0.731	0.609	0.874
NorESM2-MM	0.848	0.854	0.832	0.842	0.833	0.705	0.944	0.731	0.863	0.862
Dynamical (RCM)										
ACCESS-CM2	0.735	0.814	0.709	0.664	0.714	0.832	0.930	0.743	0.510	0.913
BCC-CSM2-MR	0.752	0.779	0.731	0.728	0.748	0.844	0.960	0.726	0.644	0.943
CNRM-ESM2-1	0.667	0.808	0.669	0.578	0.579	0.727	0.951	0.714	0.446	0.926
MPI-ESM1-2-HR	0.823	0.761	0.845	0.791	0.789	0.897	0.950	0.760	0.739	0.928
MRI-ESM2-0	0.765	0.757	0.691	0.728	0.824	0.850	0.937	0.764	0.616	0.912
NorESM2-MM	0.822	0.838	0.794	0.787	0.769	0.801	0.949	0.762	0.695	0.936
Dynamical (RCM) corrected by Livneh										
ACCESS-CM2	1.000	1.000	0.999	0.999	0.999	0.999	0.991	0.996	1.000	0.996
BCC-CSM2-MR	1.000	1.000	0.999	0.999	0.999	1.000	0.994	0.996	1.000	0.996
CNRM-ESM2-1	1.000	1.000	0.999	0.999	0.999	0.999	0.993	0.996	1.000	0.996
MPI-ESM1-2-HR	1.000	1.000	0.999	0.999	0.999	1.000	0.993	0.996	1.000	0.996
MRI-ESM2-0	1.000	1.000	0.999	0.999	0.999	1.000	0.995	0.996	1.000	0.997
NorESM2-MM	1.000	1.000	0.999	0.999	0.999	0.999	0.996	0.996	1.000	0.997
Statistical (DBCCA) trained by Livneh										
ACCESS-CM2	1.000	1.000	1.000	0.999	0.999	0.999	0.999	0.996	1.000	0.999
BCC-CSM2-MR	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.996	1.000	0.999
CNRM-ESM2-1	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.996	1.000	0.999
MPI-ESM1-2-HR	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.996	1.000	0.999
MRI-ESM2-0	1.000	1.000	0.999	0.999	1.000	1.000	0.999	0.996	1.000	0.999
NorESM2-MM	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.996	1.000	0.999

Table D.3. ROS with respect to Daymet

Model	Daily maximum temperature (°F)					Daily minimum temperature (°F)				
	Annual	Winter	Spring	Summer	Fall	Annual	Winter	Spring	Summer	Fall
Raw CMIP6 GCM										
ACCESS-CM2	0.962	0.899	1.004	1.051	0.966	0.962	0.971	0.941	0.988	0.938
BCC-CSM2-MR	1.004	0.935	1.059	1.122	1.037	1.051	1.130	1.007	1.048	1.082
CNRM-ESM2-1	0.955	1.040	0.950	0.919	0.944	0.991	1.091	0.944	0.943	0.981
MPI-ESM1-2-HR	0.982	0.953	0.968	1.039	1.009	1.029	1.044	1.024	0.983	1.072
MRI-ESM2-0	0.992	0.873	0.999	1.196	1.012	0.968	0.917	0.991	1.031	0.969
NorESM2-MM	0.959	1.100	1.001	0.849	0.949	1.010	1.128	1.027	0.923	1.013
Dynamical (RCM)										
ACCESS-CM2	0.983	0.953	1.124	0.992	0.970	0.998	1.108	1.065	0.985	0.944
BCC-CSM2-MR	0.994	0.973	1.093	0.943	1.019	1.052	1.116	1.063	1.018	1.033
CNRM-ESM2-1	1.009	0.967	1.087	1.050	0.998	0.992	1.096	1.024	0.945	0.922
MPI-ESM1-2-HR	0.970	0.961	1.092	0.945	0.960	1.028	1.099	1.038	1.056	0.985
MRI-ESM2-0	0.981	0.912	1.067	1.061	0.972	0.992	0.999	1.012	1.062	0.952
NorESM2-MM	0.940	0.989	1.043	0.864	0.951	1.033	1.090	1.077	1.044	1.005
Dynamical (RCM) corrected by Daymet										
ACCESS-CM2	1.007	1.003	1.015	0.999	1.009	1.006	1.007	1.014	0.995	1.006
BCC-CSM2-MR	1.005	1.000	1.014	0.997	1.007	1.005	1.004	1.013	0.999	1.003
CNRM-ESM2-1	1.003	1.001	1.009	0.993	1.009	1.003	1.007	1.006	0.992	1.007
MPI-ESM1-2-HR	1.005	1.001	1.011	0.996	1.007	1.004	1.006	1.008	0.998	0.998
MRI-ESM2-0	1.008	1.003	1.013	1.002	1.012	1.007	1.006	1.011	1.001	1.009
NorESM2-MM	1.004	1.002	1.010	0.994	1.008	1.003	1.004	1.011	0.995	1.003
Statistical (DBCCA) trained by Daymet										
ACCESS-CM2	1.007	1.008	1.010	1.001	1.006	1.005	1.008	1.009	1.000	1.003
BCC-CSM2-MR	1.005	1.003	1.008	1.001	1.006	1.004	1.007	1.007	1.002	1.002
CNRM-ESM2-1	1.004	1.004	1.003	0.996	1.008	1.002	1.008	1.000	0.999	1.004
MPI-ESM1-2-HR	1.004	1.001	1.006	0.997	1.007	1.002	1.005	1.006	1.000	0.997
MRI-ESM2-0	1.007	1.006	1.007	1.003	1.011	1.005	1.004	1.005	1.003	1.008
NorESM2-MM	1.004	1.004	1.005	0.997	1.006	1.002	1.006	1.004	0.998	1.000

Model	Daily precipitation (in./day)					T95 (°F)	T05 (°F)	P95 (in./day)	Wetdays	Frostdays
	Annual	Winter	Spring	Summer	Fall	Annual	Annual	Annual	Annual	Annual
Raw CMIP6 GCM										
ACCESS-CM2	0.793	0.788	0.842	0.907	0.810	1.125	0.955	0.726	0.864	0.960
BCC-CSM2-MR	0.545	0.621	0.584	0.628	0.630	1.182	1.259	0.581	0.823	0.945
CNRM-ESM2-1	0.899	1.005	0.910	0.935	0.733	1.034	1.079	0.634	1.158	0.917
MPI-ESM1-2-HR	0.809	0.863	0.885	1.054	0.693	1.140	0.994	0.633	1.018	0.911
MRI-ESM2-0	0.690	0.929	0.668	0.876	0.739	1.327	0.927	0.591	0.961	0.916
NorESM2-MM	0.722	0.724	0.705	0.806	0.692	0.912	1.112	0.586	0.837	0.855
Dynamical (RCM)										
ACCESS-CM2	0.650	0.863	0.720	0.780	0.651	1.106	1.127	0.724	0.904	0.864
BCC-CSM2-MR	0.919	1.145	0.982	0.868	1.039	0.987	1.138	0.907	0.978	0.954
CNRM-ESM2-1	1.038	1.307	0.991	1.012	0.947	1.197	1.141	0.890	1.123	0.914
MPI-ESM1-2-HR	0.802	1.055	0.868	1.035	0.770	1.003	1.133	0.904	0.909	0.927
MRI-ESM2-0	0.943	1.290	0.986	0.958	0.995	1.141	1.062	0.927	1.053	0.953
NorESM2-MM	0.793	0.991	0.780	0.903	0.883	0.973	1.109	0.859	0.903	0.932
Dynamical (RCM) corrected by Daymet										
ACCESS-CM2	1.000	1.012	0.999	1.007	0.986	1.012	1.003	1.001	1.003	0.924
BCC-CSM2-MR	0.999	0.996	1.001	1.002	0.999	1.006	0.986	1.006	1.001	0.897
CNRM-ESM2-1	1.000	1.014	0.995	1.001	0.991	1.007	0.997	1.002	1.000	0.913
MPI-ESM1-2-HR	1.000	1.000	1.005	1.000	0.990	1.003	0.997	1.007	1.002	0.920
MRI-ESM2-0	1.006	1.011	1.014	1.009	0.994	1.007	1.026	1.015	1.004	0.914
NorESM2-MM	1.000	1.012	1.005	1.000	0.993	1.003	0.992	1.005	1.000	0.919
Statistical (DBCCA) trained by Daymet										
ACCESS-CM2	0.999	1.014	0.996	1.004	0.990	1.010	1.042	1.008	0.999	0.991
BCC-CSM2-MR	1.000	0.998	1.006	1.010	0.998	1.007	1.044	1.011	1.001	0.974
CNRM-ESM2-1	1.004	1.013	0.999	1.015	1.002	1.002	1.028	1.012	0.999	0.986
MPI-ESM1-2-HR	1.004	1.006	1.011	1.009	0.995	1.006	1.026	1.008	1.002	0.984
MRI-ESM2-0	1.007	1.014	1.017	1.004	1.002	1.005	1.055	1.029	1.003	0.978
NorESM2-MM	1.004	1.019	1.003	1.010	0.996	1.002	1.048	1.015	0.999	0.973

Table D.4. ROS with respect to Livneh

Model	Daily maximum temperature (°F)					Daily minimum temperature (°F)				
	Annual	Winter	Spring	Summer	Fall	Annual	Winter	Spring	Summer	Fall
Raw CMIP6 GCM										
ACCESS-CM2	0.957	0.876	1.011	1.083	0.965	0.887	0.928	0.886	0.885	0.852
BCC-CSM2-MR	0.999	0.911	1.067	1.158	1.036	0.969	1.080	0.949	0.939	0.982
CNRM-ESM2-1	0.950	1.013	0.957	0.948	0.943	0.914	1.042	0.889	0.845	0.891
MPI-ESM1-2-HR	0.977	0.929	0.975	1.072	1.008	0.949	0.997	0.964	0.881	0.973
MRI-ESM2-0	0.987	0.850	1.007	1.233	1.011	0.892	0.876	0.933	0.924	0.880
NorESM2-MM	0.954	1.072	1.009	0.875	0.948	0.931	1.078	0.967	0.827	0.920
Dynamical (RCM)										
ACCESS-CM2	0.978	0.928	1.133	1.023	0.969	0.921	1.059	1.003	0.882	0.857
BCC-CSM2-MR	0.989	0.949	1.102	0.973	1.018	0.971	1.066	1.002	0.913	0.937
CNRM-ESM2-1	1.004	0.942	1.096	1.083	0.997	0.915	1.047	0.965	0.847	0.837
MPI-ESM1-2-HR	0.965	0.936	1.101	0.975	0.959	0.948	1.050	0.978	0.946	0.894
MRI-ESM2-0	0.976	0.889	1.075	1.094	0.971	0.915	0.955	0.953	0.952	0.864
NorESM2-MM	0.935	0.964	1.051	0.891	0.950	0.953	1.041	1.015	0.935	0.912
Dynamical (RCM) corrected by Livneh										
ACCESS-CM2	1.001	0.999	1.009	0.994	1.003	1.000	1.000	1.007	0.995	1.000
BCC-CSM2-MR	1.000	0.995	1.010	0.995	1.003	0.999	0.994	1.007	0.999	0.999
CNRM-ESM2-1	1.000	0.997	1.008	0.993	1.005	0.999	0.999	1.004	0.994	1.001
MPI-ESM1-2-HR	1.001	0.999	1.009	0.995	1.001	0.999	1.001	1.003	1.000	0.996
MRI-ESM2-0	1.001	0.997	1.005	0.997	1.004	0.999	0.997	1.002	0.999	1.001
NorESM2-MM	1.001	0.999	1.007	0.995	1.003	0.999	1.000	1.005	0.996	0.999
Statistical (DBCCA) trained by Livneh										
ACCESS-CM2	1.000	1.000	1.003	0.997	1.000	0.999	0.999	1.002	0.999	0.999
BCC-CSM2-MR	1.000	0.998	1.005	0.997	1.003	0.999	0.995	1.003	1.001	1.002
CNRM-ESM2-1	1.000	0.999	1.002	0.996	1.003	0.999	1.000	1.000	0.999	1.001
MPI-ESM1-2-HR	1.000	1.000	1.003	0.998	1.000	0.999	0.999	1.001	1.001	0.997
MRI-ESM2-0	1.000	0.999	1.000	0.997	1.004	0.999	0.997	0.998	1.000	1.004
NorESM2-MM	1.000	1.001	1.001	0.997	1.000	0.999	1.000	1.001	0.999	1.000

Model	Daily precipitation (in./day)					T95 (°F)	T05 (°F)	P95 (in./day)	Wetdays	Frostdays
	Annual	Winter	Spring	Summer	Fall	Annual	Annual	Annual	Annual	Annual
Raw CMIP6 GCM										
ACCESS-CM2	0.815	0.790	0.866	0.958	0.832	1.144	0.927	0.757	0.912	0.878
BCC-CSM2-MR	0.560	0.623	0.600	0.664	0.647	1.202	1.221	0.606	0.868	0.864
CNRM-ESM2-1	0.924	1.009	0.936	0.987	0.753	1.051	1.047	0.661	1.222	0.839
MPI-ESM1-2-HR	0.831	0.866	0.910	1.113	0.712	1.159	0.965	0.660	1.074	0.834
MRI-ESM2-0	0.709	0.932	0.687	0.925	0.759	1.349	0.899	0.617	1.014	0.838
NorESM2-MM	0.742	0.727	0.725	0.852	0.711	0.928	1.078	0.611	0.883	0.782
Dynamical (RCM)										
ACCESS-CM2	0.668	0.866	0.740	0.823	0.669	1.125	1.093	0.755	0.954	0.790
BCC-CSM2-MR	0.944	1.149	1.009	0.917	1.067	1.003	1.104	0.945	1.032	0.873
CNRM-ESM2-1	1.067	1.311	1.019	1.069	0.973	1.217	1.107	0.928	1.185	0.836
MPI-ESM1-2-HR	0.824	1.058	0.892	1.093	0.791	1.019	1.099	0.943	0.959	0.848
MRI-ESM2-0	0.969	1.294	1.014	1.012	1.022	1.160	1.031	0.967	1.112	0.872
NorESM2-MM	0.815	0.994	0.802	0.954	0.907	0.989	1.076	0.896	0.953	0.853
Dynamical (RCM) corrected by Livneh										
ACCESS-CM2	0.999	1.005	0.992	1.001	0.989	1.006	0.993	0.999	1.000	0.935
BCC-CSM2-MR	0.999	0.997	1.000	1.001	0.994	1.005	0.976	1.001	1.001	0.913
CNRM-ESM2-1	1.001	1.012	0.996	0.997	0.993	1.007	0.988	0.995	1.001	0.924
MPI-ESM1-2-HR	1.001	1.002	1.005	0.995	0.994	1.001	0.987	1.000	1.002	0.934
MRI-ESM2-0	1.000	1.003	1.007	0.996	0.995	1.002	1.011	1.006	1.001	0.923
NorESM2-MM	1.000	1.010	1.001	0.998	0.996	1.002	0.983	0.995	1.001	0.930
Statistical (DBCCA) trained by Livneh										
ACCESS-CM2	1.002	1.008	0.994	0.997	0.998	1.005	1.016	0.996	1.000	0.987
BCC-CSM2-MR	1.001	1.001	1.010	0.999	0.993	1.005	1.019	1.002	1.001	0.977
CNRM-ESM2-1	1.003	1.013	0.997	1.005	1.003	1.001	1.007	0.999	1.001	0.982
MPI-ESM1-2-HR	1.002	1.006	1.007	0.994	1.003	1.008	1.008	0.998	1.001	0.983
MRI-ESM2-0	1.002	1.007	1.010	0.999	0.998	0.999	1.028	1.010	1.000	0.974
NorESM2-MM	1.003	1.014	1.003	0.995	1.002	1.004	1.022	0.998	1.000	0.972

Table D.5. Biases with respect to Daymet

Model	Daily maximum temperature (°F)					Daily minimum temperature (°F)				
	Annual	Winter	Spring	Summer	Fall	Annual	Winter	Spring	Summer	Fall
Raw CMIP6 GCM										
ACCESS-CM2	−5.76	−7.29	−6.94	−2.89	−5.89	2.08	2.23	1.89	1.82	2.46
BCC-CSM2-MR	−2.34	−6.29	−4.68	4.22	−2.69	1.68	0.77	1.76	2.09	2.16
CNRM-ESM2-1	−3.95	−5.02	−5.02	−4.07	−1.68	0.99	−0.52	0.30	0.73	3.44
MPI-ESM1-2-HR	−2.88	−3.95	−4.71	−1.28	−1.55	4.35	2.36	2.83	6.56	5.64
MRI-ESM2-0	−3.48	−2.16	−4.45	−3.13	−4.12	5.06	6.91	4.14	4.33	4.93
NorESM2-MM	0.43	−3.50	0.01	3.79	1.31	6.97	4.35	7.12	8.46	7.89
Dynamical (RCM)										
ACCESS-CM2	−9.58	−13.29	−11.86	−5.81	−7.42	−2.65	−7.74	−4.26	0.82	0.51
BCC-CSM2-MR	−5.85	−6.61	−7.29	−4.55	−4.91	0.97	−0.43	0.45	1.82	2.10
CNRM-ESM2-1	−8.51	−8.75	−10.69	−8.26	−6.33	−0.93	−3.13	−2.75	0.08	2.03
MPI-ESM1-2-HR	−6.50	−8.97	−9.67	−2.85	−4.53	−0.07	−3.09	−2.62	3.42	1.96
MRI-ESM2-0	−6.20	−5.76	−8.19	−5.23	−5.60	1.18	0.92	−0.14	1.83	2.15
NorESM2-MM	−4.02	−6.46	−4.80	−1.32	−3.49	2.56	0.03	2.21	4.66	3.32
Dynamical (RCM) corrected by Daymet										
ACCESS-CM2	−0.07	−0.05	−0.08	−0.04	−0.05	−0.04	−0.11	−0.02	−0.03	0.05
BCC-CSM2-MR	−0.10	−0.14	−0.18	0.01	−0.03	−0.07	−0.21	−0.07	0.02	0.03
CNRM-ESM2-1	−0.07	−0.11	−0.14	0.02	0.00	−0.04	−0.15	−0.06	0.03	0.06
MPI-ESM1-2-HR	−0.05	0.04	−0.24	0.04	0.02	−0.03	0.04	−0.18	0.07	0.02
MRI-ESM2-0	−0.15	−0.10	−0.18	−0.08	−0.15	−0.09	−0.10	−0.11	−0.03	−0.07
NorESM2-MM	−0.08	−0.14	−0.09	−0.03	0.01	−0.05	−0.13	0.01	−0.02	0.00
Statistical (DBCCA) trained by Daymet										
ACCESS-CM2	−0.10	−0.12	−0.06	−0.02	−0.14	−0.07	−0.12	−0.06	0.03	−0.08
BCC-CSM2-MR	−0.11	−0.27	−0.14	0.07	−0.03	−0.05	−0.23	−0.05	0.12	−0.01
CNRM-ESM2-1	−0.09	−0.19	−0.12	−0.01	0.02	−0.05	−0.12	−0.07	0.04	0.01
MPI-ESM1-2-HR	−0.05	0.01	−0.22	0.10	−0.02	−0.04	0.02	−0.14	0.11	−0.06
MRI-ESM2-0	−0.14	−0.18	−0.10	−0.09	−0.14	−0.07	−0.04	−0.08	0.04	−0.12
NorESM2-MM	−0.11	−0.20	−0.09	−0.01	−0.07	−0.08	−0.13	−0.02	−0.02	−0.09

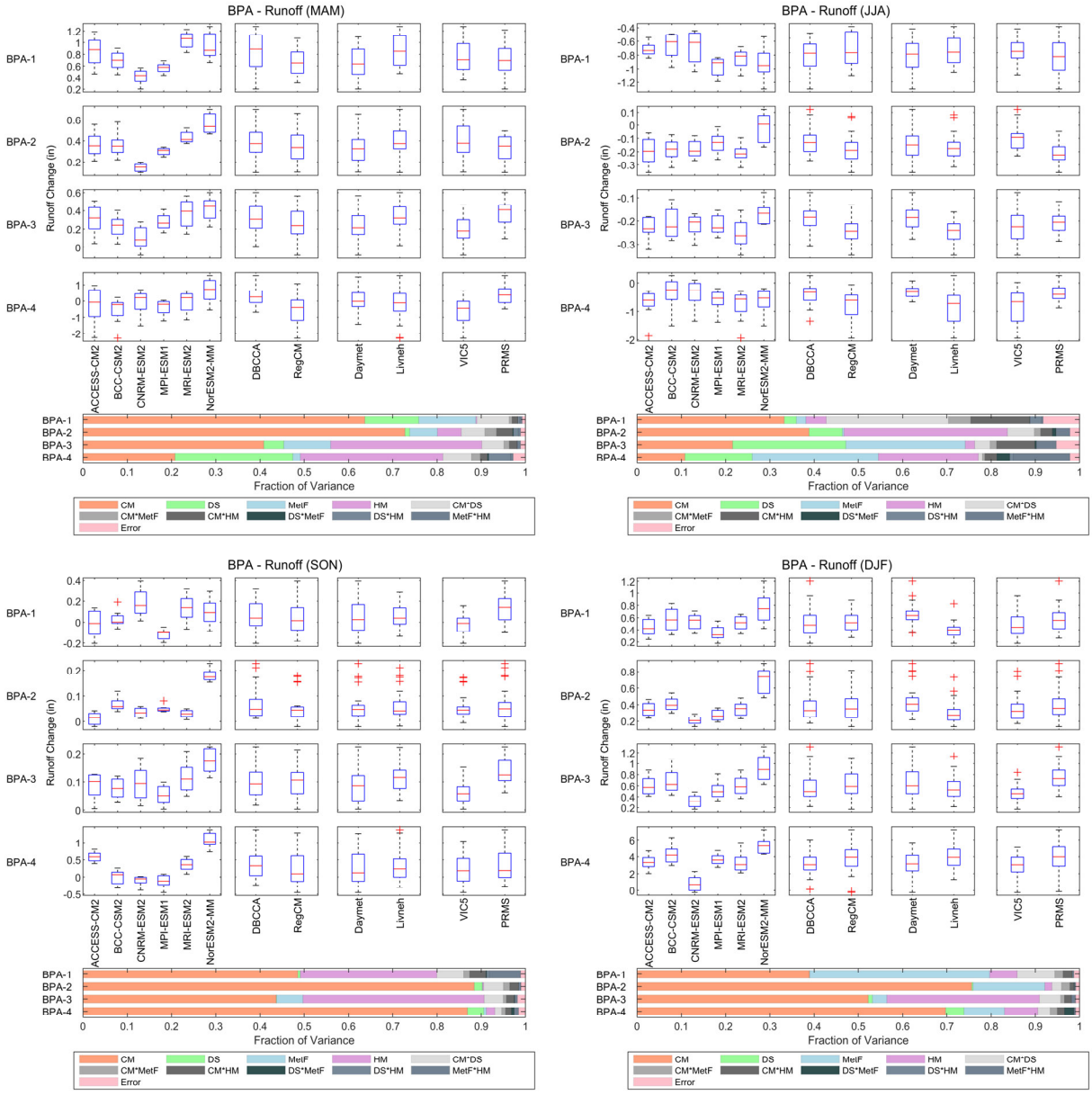
Model	Daily precipitation (in./day)					T95 (°F)	T05 (°F)	P95 (in./day)	Wetdays	Frostdays
	Annual	Winter	Spring	Summer	Fall	Annual	Annual	Annual	Annual	Annual
Raw CMIP6 GCM										
ACCESS-CM2	0.012	0.010	0.020	0.008	0.008	−1.95	3.43	−0.194	41.45	−11.19
BCC-CSM2-MR	−0.003	0.013	0.001	−0.025	−0.002	7.05	−2.52	−0.292	38.67	−17.94
CNRM-ESM2-1	0.009	0.013	0.020	0.000	0.004	−3.42	−0.89	−0.184	34.34	−8.98
MPI-ESM1-2-HR	0.004	0.019	0.015	−0.007	−0.012	−0.83	1.77	−0.188	27.31	−19.51
MRI-ESM2-0	0.023	0.023	0.033	0.024	0.011	−2.52	7.26	−0.233	58.93	−37.48
NorESM2-MM	−0.013	0.004	−0.013	−0.024	−0.017	4.51	5.02	−0.366	24.04	−38.59
Dynamical (RCM)										
ACCESS-CM2	0.023	0.014	0.022	0.033	0.022	−5.17	−10.92	−0.177	60.52	17.44
BCC-CSM2-MR	0.035	0.038	0.042	0.029	0.030	−3.95	−3.33	−0.105	65.57	−7.94
CNRM-ESM2-1	0.047	0.032	0.049	0.070	0.035	−8.36	−5.51	−0.106	80.08	1.44
MPI-ESM1-2-HR	0.028	0.032	0.035	0.030	0.016	−2.26	−5.73	−0.054	54.70	4.49
MRI-ESM2-0	0.042	0.038	0.054	0.041	0.037	−4.65	−0.57	−0.054	70.20	−9.46
NorESM2-MM	0.020	0.023	0.025	0.019	0.012	−0.68	−1.36	−0.133	51.59	−11.32
Dynamical (RCM) corrected by Daymet										
ACCESS-CM2	0.000	0.001	0.000	0.001	0.000	−0.06	−3.29	0.003	0.39	−6.86
BCC-CSM2-MR	0.000	−0.001	0.001	0.000	0.000	0.00	−2.86	0.005	0.31	−2.26
CNRM-ESM2-1	0.000	0.001	0.000	0.000	0.000	0.03	−2.80	0.004	0.19	−3.98
MPI-ESM1-2-HR	0.000	0.001	0.000	0.000	−0.001	−0.13	−2.68	0.004	0.20	−5.49
MRI-ESM2-0	0.001	0.001	0.001	0.000	0.000	−0.15	−2.82	0.008	0.52	−2.83
NorESM2-MM	0.000	0.001	0.000	0.000	−0.001	−0.07	−2.52	0.002	0.17	−0.93
Statistical (DBCCA) trained by Daymet										
ACCESS-CM2	0.000	0.001	0.000	0.000	0.000	−0.38	−0.50	0.003	0.20	1.25
BCC-CSM2-MR	0.000	0.000	0.001	0.001	0.000	−0.13	−0.87	0.007	0.22	0.85
CNRM-ESM2-1	0.000	0.001	0.001	0.000	0.000	−0.39	−0.43	0.005	0.09	1.13
MPI-ESM1-2-HR	0.000	0.001	0.000	0.000	0.000	−0.40	−0.48	0.005	0.11	0.61
MRI-ESM2-0	0.001	0.001	0.001	0.000	0.001	−0.36	−0.90	0.015	0.48	1.21
NorESM2-MM	0.000	0.001	0.000	0.000	0.000	−0.30	−0.70	0.004	0.09	1.10

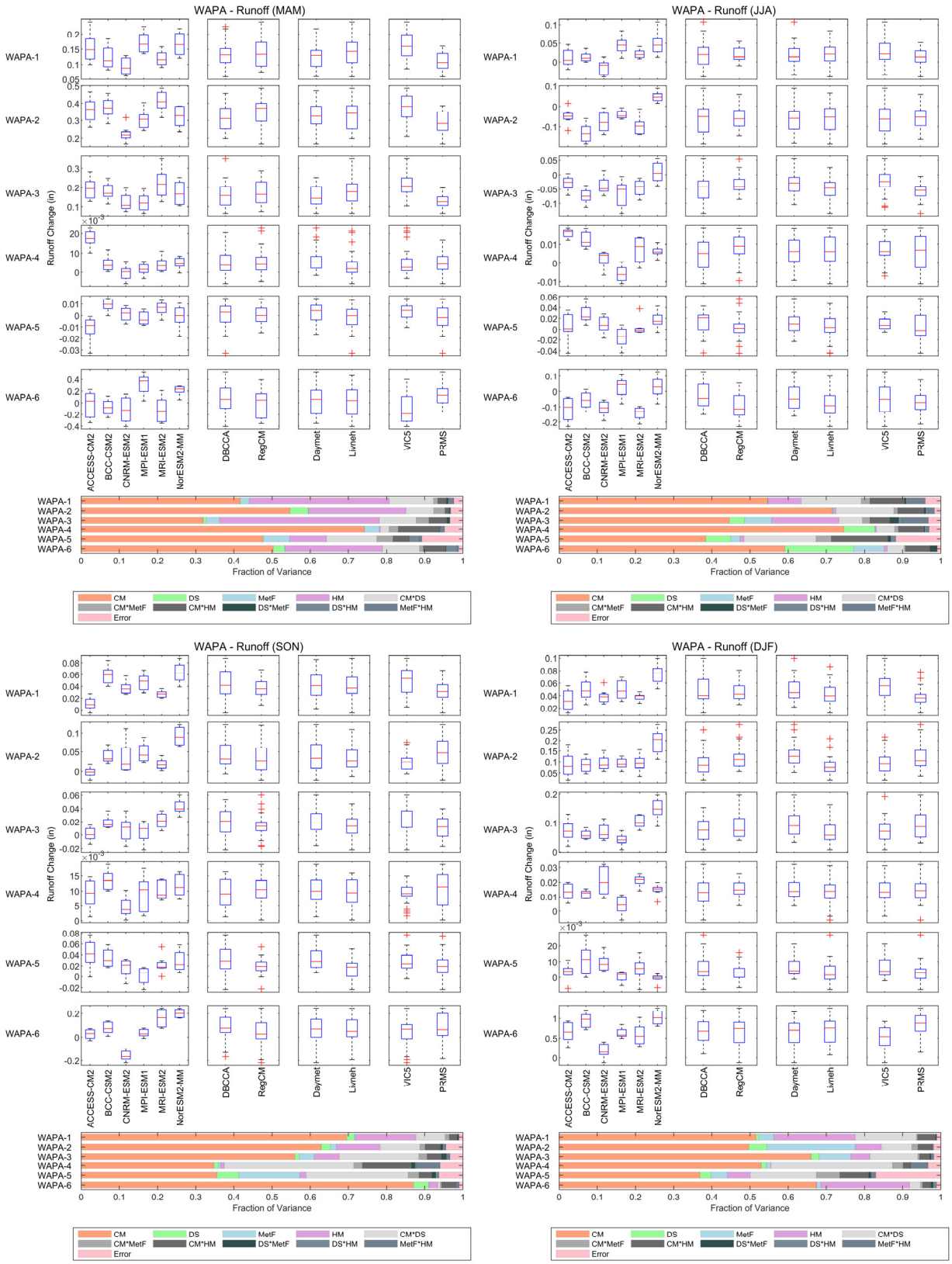
Table D.6. Biases with respect to Livneh

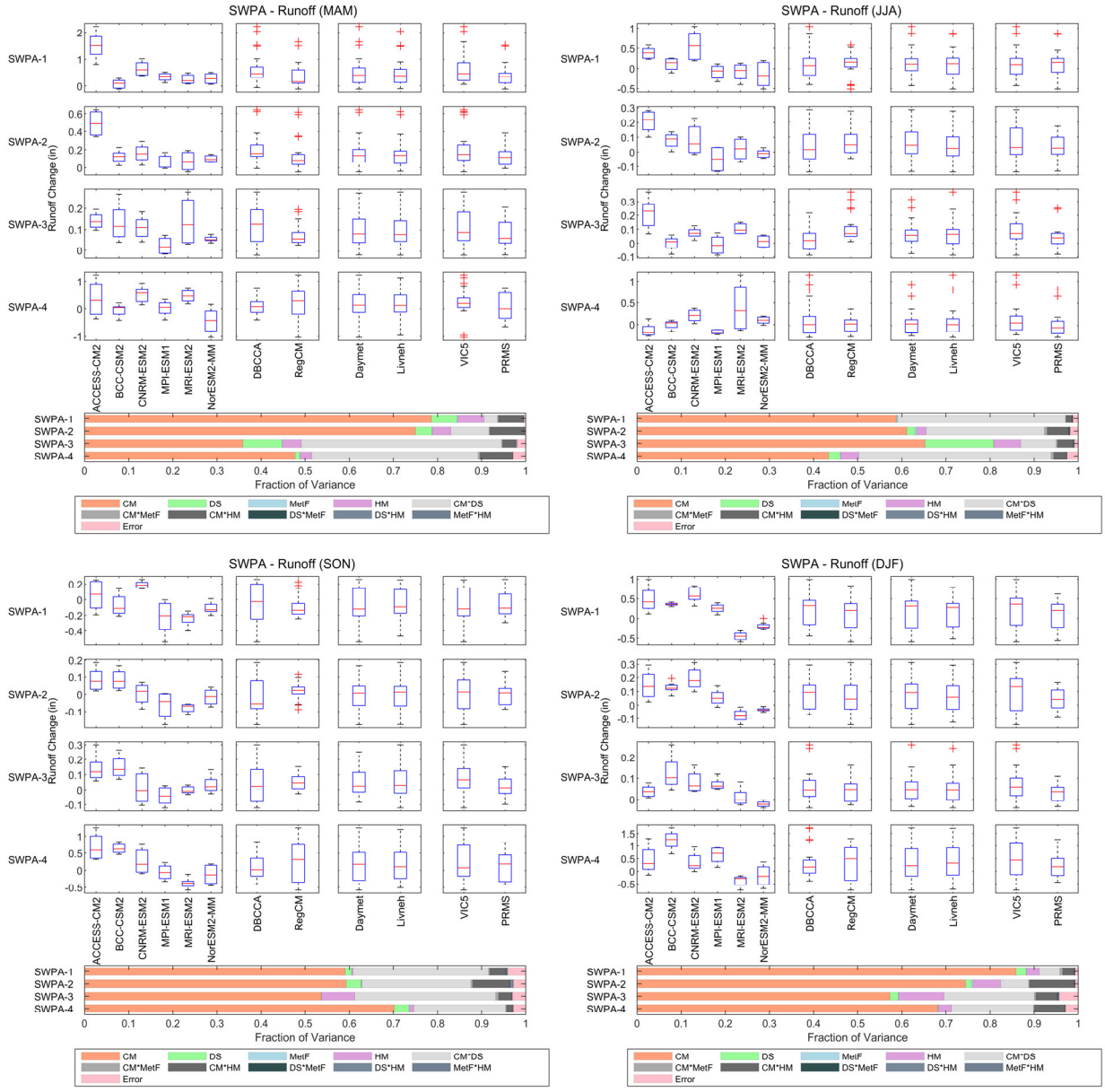
Model	Daily maximum temperature (°F)					Daily minimum temperature (°F)				
	Annual	Winter	Spring	Summer	Fall	Annual	Winter	Spring	Summer	Fall
Raw CMIP6 GCM										
ACCESS-CM2	−5.72	−6.79	−6.93	−3.09	−6.08	2.99	3.29	2.59	2.60	3.48
BCC-CSM2-MR	−2.29	−5.79	−4.67	4.02	−2.87	2.59	1.83	2.45	2.88	3.17
CNRM-ESM2-1	−3.91	−4.52	−5.02	−4.27	−1.87	1.89	0.54	0.99	1.52	4.45
MPI-ESM1-2-HR	−2.83	−3.45	−4.70	−1.47	−1.74	5.25	3.42	3.53	7.35	6.65
MRI-ESM2-0	−3.43	−1.65	−4.44	−3.32	−4.30	5.96	7.98	4.84	5.12	5.94
NorESM2-MM	0.47	−3.00	0.02	3.59	1.13	7.88	5.41	7.82	9.25	8.91
Dynamical (RCM)										
ACCESS-CM2	−9.54	−12.78	−11.85	−6.01	−7.60	−1.75	−6.67	−3.56	1.61	1.52
BCC-CSM2-MR	−5.80	−6.11	−7.29	−4.74	−5.09	1.88	0.64	1.15	2.61	3.11
CNRM-ESM2-1	−8.46	−8.25	−10.68	−8.46	−6.51	−0.03	−2.06	−2.05	0.87	3.04
MPI-ESM1-2-HR	−6.45	−8.47	−9.66	−3.04	−4.72	0.83	−2.03	−1.92	4.21	2.98
MRI-ESM2-0	−6.16	−5.26	−8.18	−5.43	−5.78	2.08	1.98	0.56	2.62	3.16
NorESM2-MM	−3.97	−5.96	−4.79	−1.51	−3.67	3.46	1.10	2.91	5.45	4.33
Dynamical (RCM) corrected by Livneh										
ACCESS-CM2	0.00	0.00	−0.02	0.01	0.02	−0.01	−0.08	0.00	−0.01	0.08
BCC-CSM2-MR	0.00	0.02	−0.06	0.02	0.02	−0.02	−0.05	−0.02	−0.01	0.02
CNRM-ESM2-1	0.01	−0.04	0.01	0.03	0.04	0.00	−0.14	0.03	0.01	0.08
MPI-ESM1-2-HR	0.01	0.04	−0.10	0.07	0.04	−0.01	−0.01	−0.09	0.05	0.04
MRI-ESM2-0	0.00	0.01	0.02	0.03	−0.04	−0.01	−0.03	0.00	0.02	−0.01
NorESM2-MM	−0.01	−0.10	0.01	0.00	0.06	−0.02	−0.14	0.03	−0.03	0.06
Statistical (DBCCA) trained by Livneh										
ACCESS-CM2	0.00	−0.04	0.01	0.07	−0.03	−0.02	−0.04	−0.05	0.06	−0.04
BCC-CSM2-MR	0.00	−0.10	−0.03	0.12	0.02	−0.02	−0.12	−0.02	0.08	0.00
CNRM-ESM2-1	0.02	−0.10	0.04	0.06	0.06	0.00	−0.12	0.02	0.06	0.04
MPI-ESM1-2-HR	0.02	−0.01	−0.04	0.13	0.01	0.00	−0.01	−0.07	0.11	−0.03
MRI-ESM2-0	0.01	−0.06	0.08	0.05	−0.01	−0.01	−0.01	−0.02	0.08	−0.07
NorESM2-MM	−0.01	−0.16	0.02	0.06	0.04	−0.03	−0.14	0.01	0.00	0.02

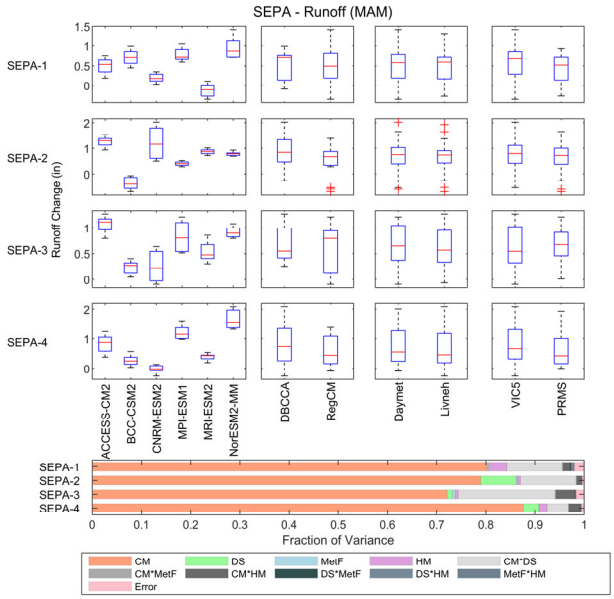
Model	Daily precipitation (in./day)					T95 (°F)	T05 (°F)	P95 (in./day)	Wetdays	Frostdays
	Annual	Winter	Spring	Summer	Fall	Annual	Annual	Annual	Annual	Annual
Raw CMIP6 GCM										
ACCESS-CM2	0.015	0.014	0.024	0.009	0.012	−2.14	4.78	−0.179	34.67	−19.43
BCC-CSM2-MR	−0.001	0.017	0.005	−0.024	0.001	6.86	−1.17	−0.277	31.88	−26.19
CNRM-ESM2-1	0.012	0.016	0.024	0.001	0.007	−3.62	0.45	−0.169	27.55	−17.23
MPI-ESM1-2-HR	0.006	0.022	0.019	−0.006	−0.009	−1.02	3.12	−0.172	20.52	−27.76
MRI-ESM2-0	0.026	0.027	0.036	0.025	0.014	−2.71	8.61	−0.217	52.14	−45.73
NorESM2-MM	−0.010	0.007	−0.010	−0.023	−0.014	4.31	6.37	−0.351	17.25	−46.84
Dynamical (RCM)										
ACCESS-CM2	0.025	0.017	0.026	0.034	0.025	−5.37	−9.57	−0.161	53.73	9.19
BCC-CSM2-MR	0.038	0.042	0.045	0.030	0.033	−4.14	−1.99	−0.090	58.78	−16.19
CNRM-ESM2-1	0.050	0.035	0.053	0.071	0.039	−8.55	−4.17	−0.091	73.29	−6.80
MPI-ESM1-2-HR	0.031	0.035	0.039	0.031	0.019	−2.45	−4.39	−0.039	47.91	−3.76
MRI-ESM2-0	0.045	0.041	0.058	0.041	0.041	−4.84	0.78	−0.039	63.41	−17.71
NorESM2-MM	0.022	0.026	0.028	0.020	0.015	−0.88	−0.01	−0.118	44.80	−19.57
Dynamical (RCM) corrected by Livneh										
ACCESS-CM2	0.000	0.000	−0.001	0.000	0.000	0.01	−2.86	−0.003	−0.06	−8.17
BCC-CSM2-MR	0.000	−0.001	0.000	0.000	0.000	0.04	−2.38	−0.001	−0.08	−3.79
CNRM-ESM2-1	0.000	0.000	0.000	0.000	0.000	0.05	−2.44	−0.002	−0.10	−5.56
MPI-ESM1-2-HR	0.000	0.001	0.000	0.000	−0.001	−0.07	−2.36	−0.001	−0.14	−6.48
MRI-ESM2-0	0.000	0.000	0.000	0.000	0.000	0.00	−2.42	0.000	−0.06	−4.63
NorESM2-MM	0.000	0.001	0.000	0.000	−0.001	−0.03	−2.20	−0.007	−0.07	−2.16
Statistical (DBCCA) trained by Livneh										
ACCESS-CM2	0.000	0.001	−0.001	0.000	0.000	−0.33	−0.58	−0.007	−0.09	0.43
BCC-CSM2-MR	0.000	−0.001	0.001	0.000	0.000	−0.20	−0.85	−0.001	−0.07	0.44
CNRM-ESM2-1	0.000	0.000	0.000	0.000	0.000	−0.33	−0.54	−0.006	−0.10	0.66
MPI-ESM1-2-HR	0.000	0.001	−0.001	−0.001	0.000	−0.35	−0.71	−0.004	−0.11	0.10
MRI-ESM2-0	0.000	0.001	0.000	0.000	0.000	−0.19	−1.10	0.003	−0.08	0.41
NorESM2-MM	0.000	0.001	0.000	0.000	0.000	−0.21	−0.90	−0.007	−0.07	0.98

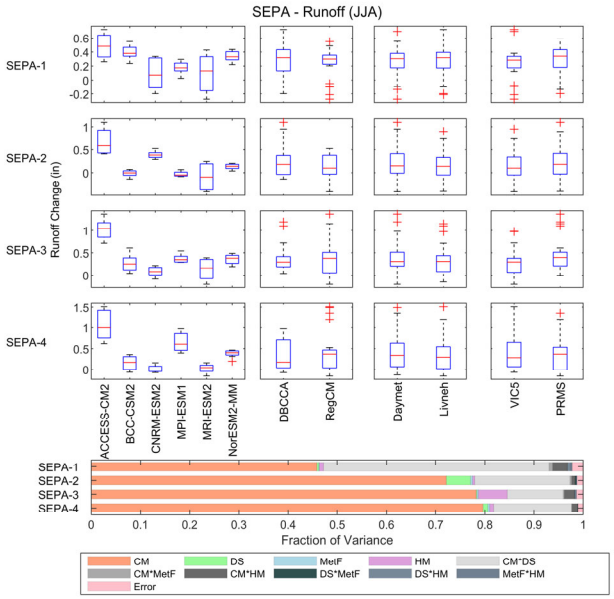
APPENDIX E. ANOVA OF SEASONAL RUNOFF

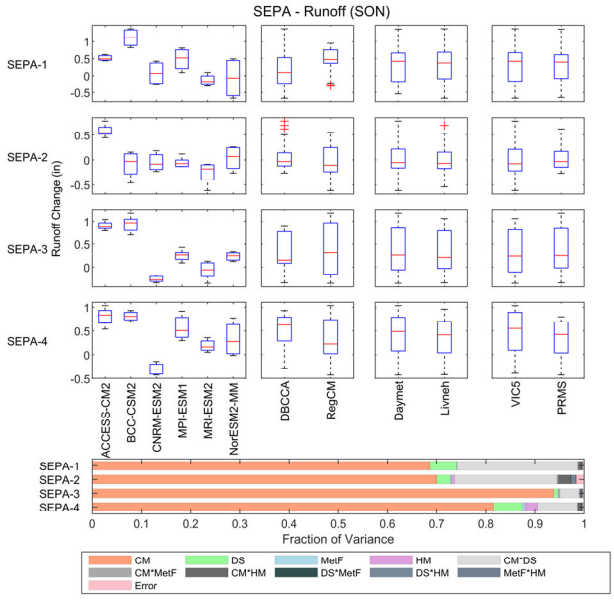


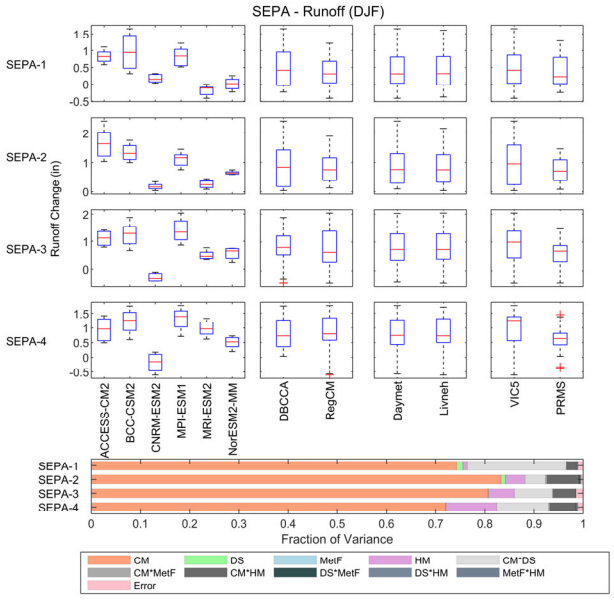

WAPA - Runoff (MAM)
WAPA - Runoff (JJA)
WAPA - Runoff (SON)
WAPA - Runoff (DJF)
WAPA-1
WAPA-2
WAPA-3
WAPA-4
WAPA-5
WAPA-6
Runoff Change (in)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
Fraction of Variance
CM
DS
MetF
HM
CM*DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error


SWPA - Runoff (MAM)
SWPA - Runoff (JJA)
SWPA - Runoff (SON)
SWPA - Runoff (DJF)
SWPA-1
SWPA-2
SWPA-3
SWPA-4
Runoff Change (in)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
Fraction of Variance
CM
DS
MetF
HM
CM*DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error

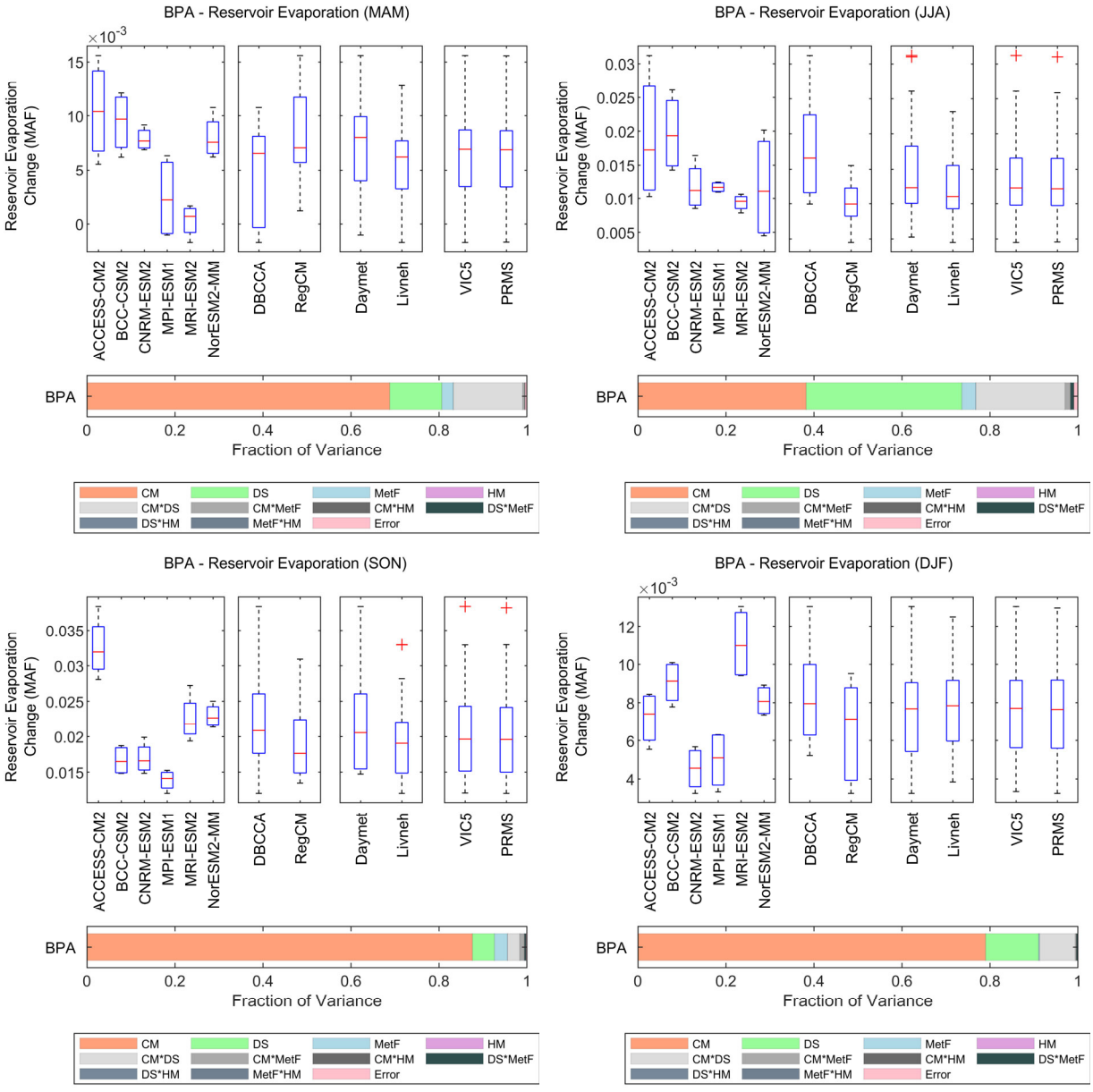

SEPA - Runoff (MAM)
SEPA-1
SEPA-2
SEPA-3
SEPA-4
Runoff Change (in)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
Fraction of Variance
CM
DS
MetF
HM
CM-DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error

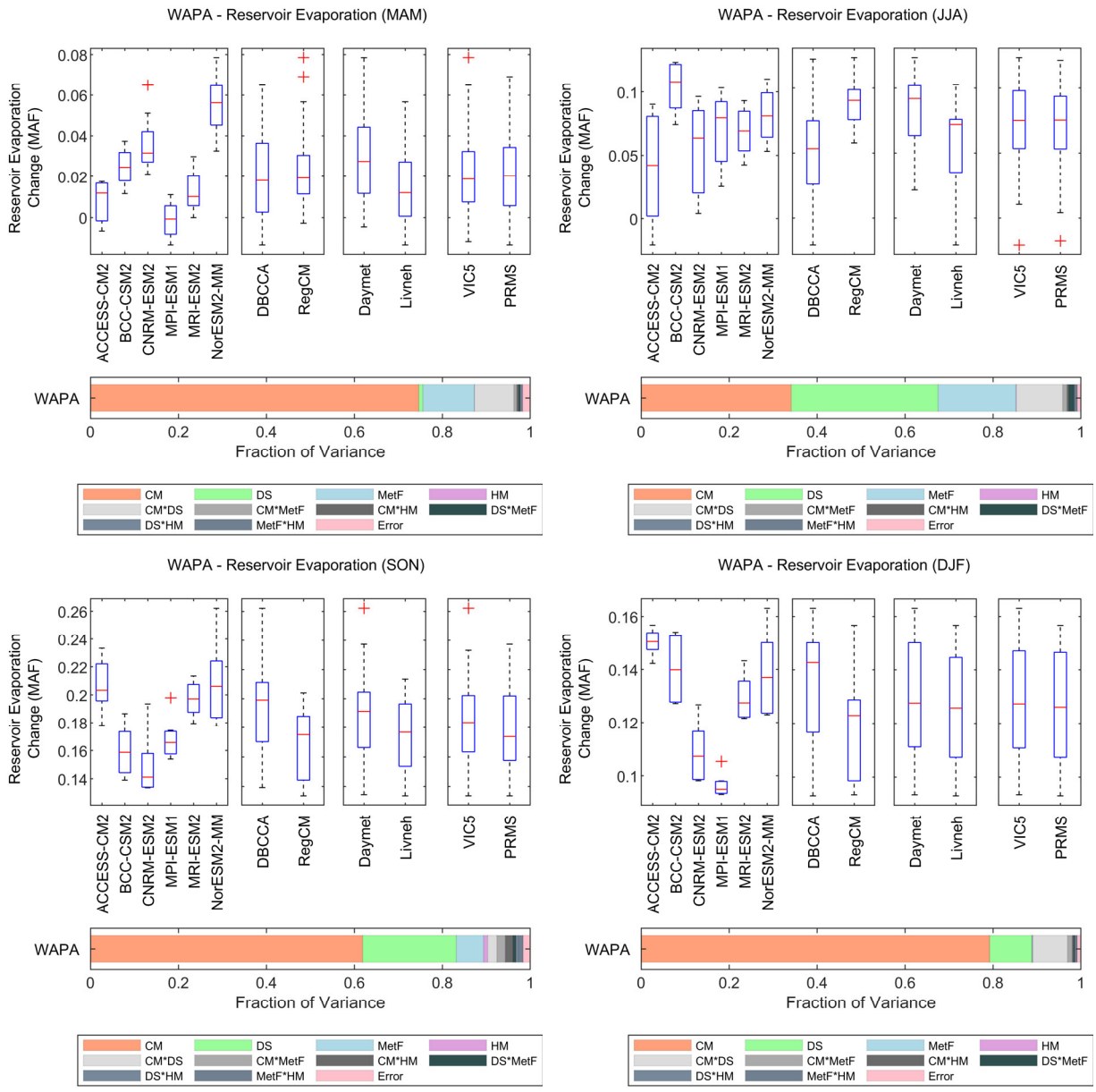

SEPA - Runoff (JJA)
SEPA-1
SEPA-2
SEPA-3
SEPA-4
Runoff Change (in)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
Fraction of Variance
CM
DS
MetF
HM
CM-DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error

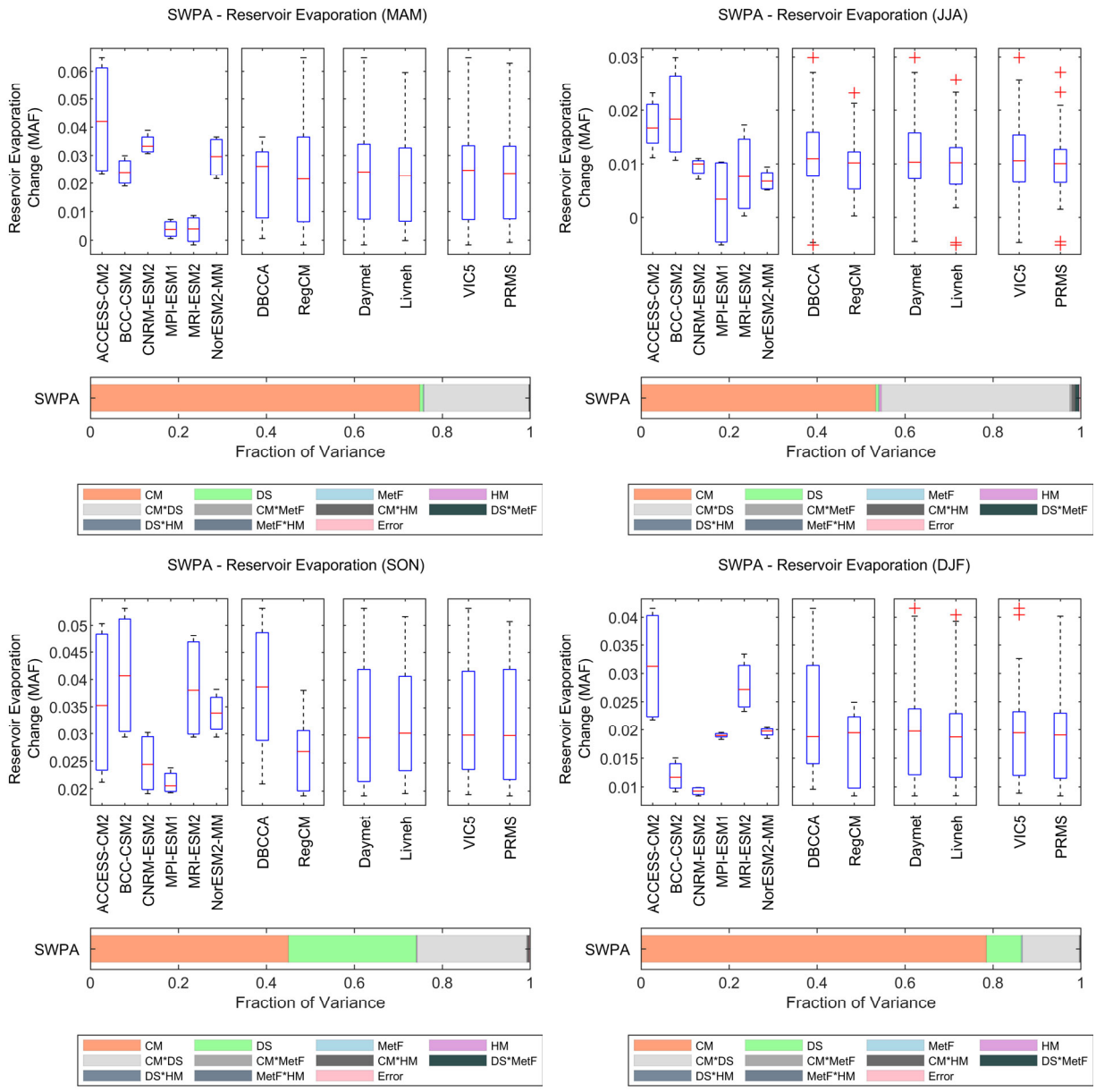

SEPA - Runoff (SON)
SEPA-1
SEPA-2
SEPA-3
SEPA-4
Runoff Change (in)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
Fraction of Variance
CM
DS
MetF
HM
CM-DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error

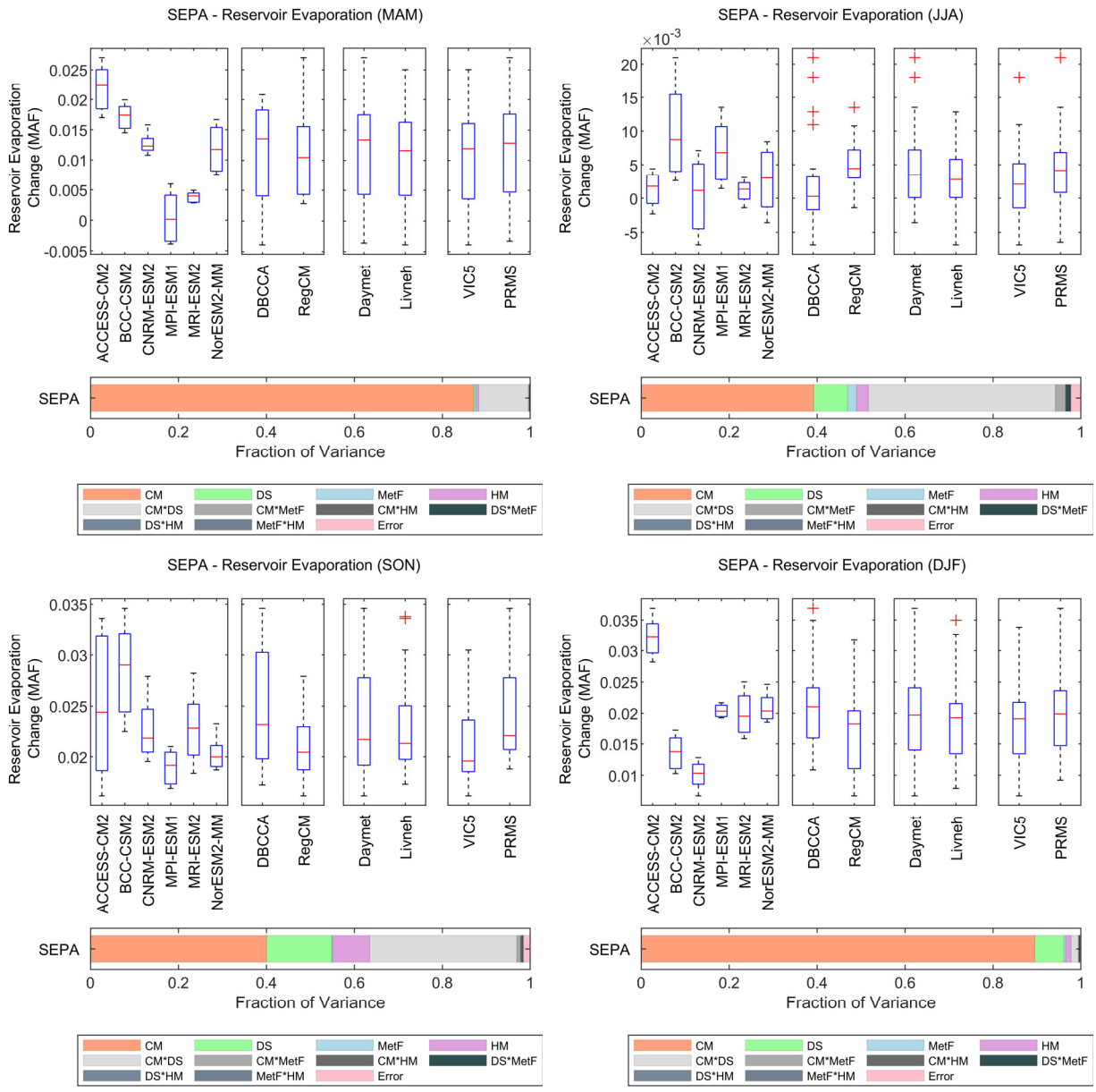

SEPA - Runoff (DJF)
SEPA-1
SEPA-2
SEPA-3
SEPA-4
Runoff Change (in)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
Fraction of Variance
CM
DS
MetF
HM
CM-DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error

APPENDIX F. ANOVA OF SEASONAL RESERVOIR EVAPORATION

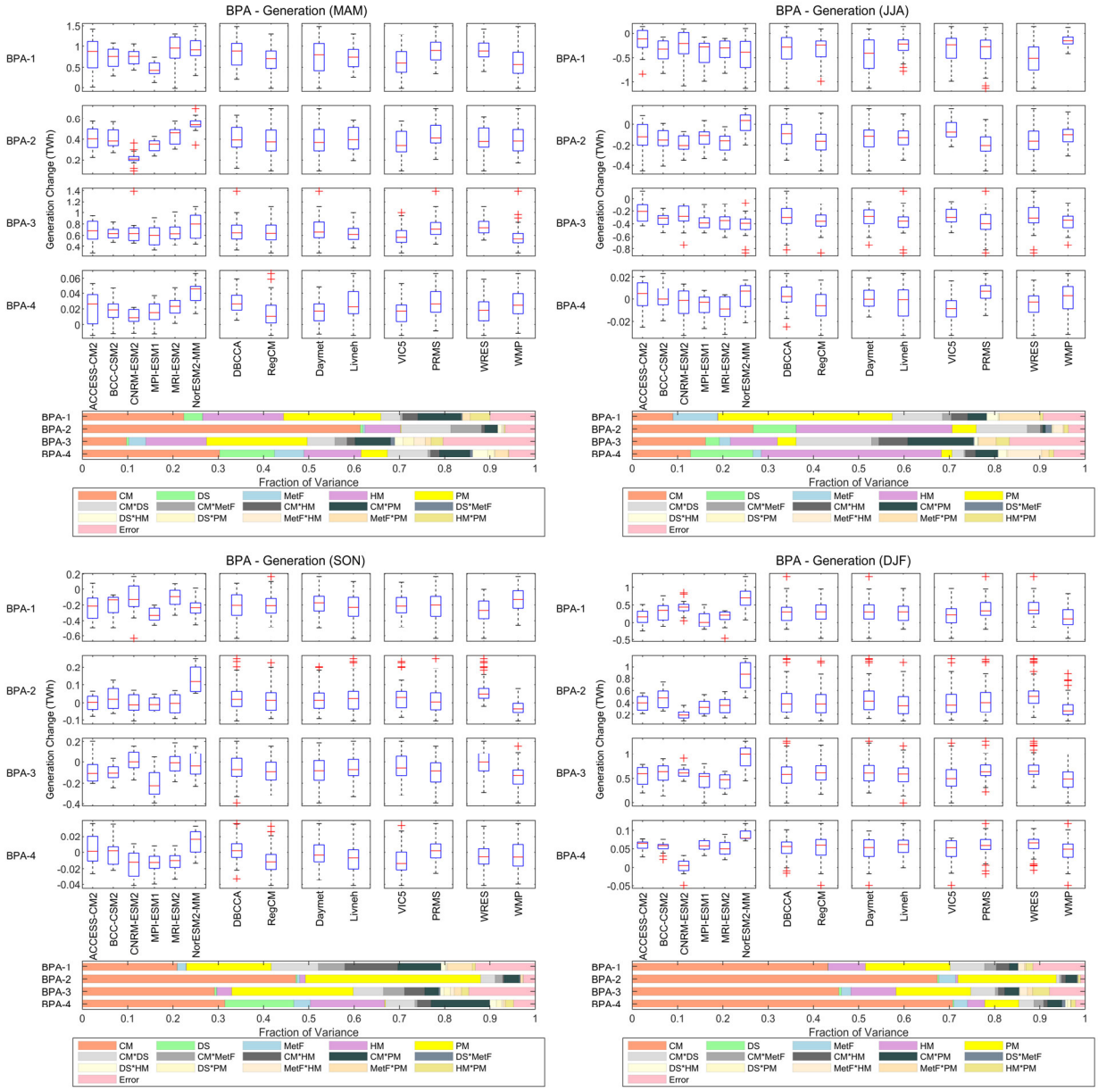



WAPA - Reservoir Evaporation (MAM)
WAPA - Reservoir Evaporation (JJA)
WAPA - Reservoir Evaporation (SON)
WAPA - Reservoir Evaporation (DJF)
Reservoir Evaporation Change (MAF)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
WAPA
Fraction of Variance
CM
DS
MetF
HM
CM*DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error


SWPA - Reservoir Evaporation (MAM)
SWPA - Reservoir Evaporation (JJA)
SWPA - Reservoir Evaporation (SON)
SWPA - Reservoir Evaporation (DJF)
Reservoir Evaporation Change (MAF)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
SWPA
Fraction of Variance
CM
DS
MetF
HM
CM*DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error


SEPA - Reservoir Evaporation (MAM)
SEPA - Reservoir Evaporation (JJA)
SEPA - Reservoir Evaporation (SON)
SEPA - Reservoir Evaporation (DJF)
Reservoir Evaporation Change (MAF)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
SEPA
Fraction of Variance
CM
DS
MetF
HM
CM*DS
CM*MetF
CM*HM
DS*MetF
DS*HM
MetF*HM
Error

APPENDIX G. ANOVA OF SEASONAL HYDROPOWER GENERATION




WAPA - Generation (MAM)
WAPA - Generation (JJA)
WAPA - Generation (SON)
WAPA - Generation (DJF)
WAPA-1
WAPA-2
WAPA-3
WAPA-4
WAPA-5
WAPA-6
Generation Change (MWh)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
WRES
WMP
Fraction of Variance
CM
DS
MetF
HM
PM
CM*DS
CM*MetF
CM*HM
CM*PM
DS*MetF
DS*HM
DS*PM
MetF*HM
MetF*PM
HM*PM
Error


SWPA - Generation (MAM)
SWPA - Generation (JJA)
SWPA - Generation (SON)
SWPA - Generation (DJF)
SWPA-1
SWPA-2
SWPA-3
SWPA-4
Generation Change (TWh)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
WRES
WMP
Fraction of Variance
CM
DS
MetF
HM
PM
CM*DS
CM*MetF
CM*HM
CM*PM
DS*MetF
DS*HM
DS*PM
MetF*HM
MetF*PM
HM*PM
Error


SEPA - Generation (MAM)
SEPA - Generation (JJA)
SEPA - Generation (SON)
SEPA - Generation (DJF)
SEPA-1
SEPA-2
SEPA-3
SEPA-4
Generation Change (TWh)
ACCESS-CM2
BCC-CSM2
CNRM-ESM2
MPI-ESM1
MRI-ESM2
NorESM2-MM
DBCCA
RegCM
Daymet
Livneh
VIC5
PRMS
WRES
WMP
Fraction of Variance
CM
DS
MetF
HM
PM
CM*DS
CM*MetF
CM*HM
CM*PM
DS*MetF
DS*HM
DS*PM
MetF*HM
MetF*PM
HM*PM
Error

APPENDIX H. POWER MARKETING ANALYSIS

Table H.1. displays the correspondence between the 18 9505 areas and the 9 power marketing subregions used for the demand analysis. Whenever possible, the 9 power marketing subregions correspond to the regional delineation used by the PMAs for marketing and rate-setting purposes (e.g., BPA functions as a single region for marketing purposes even though it is divided into four areas for hydrologic projections). In some cases, multiple power marketing regions had to be combined because the number of preference customers reporting sales data to EIA was too small to estimate their sales equation separately.

Table H.1. Correspondence between 9505 study regions and marketing regions used for PMA customer demand analysis

9505 Study Regions	Subregions used in the Power Marketing Analysis	Number of preference customers used for estimation of sales equation
BPA-1–4	BPA	91
WAPA-1	WAPA-UGP	128
WAPA-3, WAPA-5	WAPA-CRSP	36
WAPA-4	WAPA-DSW	18
WAPA-2, WAPA-6	WAPA-SN/RM	18
SWPA-1–4	SWPA	35
SEPA-1	SEPA-Kerr-Philpott	63
SEPA-2	SEPA-Cumberland	162
SEPA-3, SEPA-4	SEPA-GA/AL/SC/JW	83

Note: Data and estimation method for air conditioning penetration in BPA region

Air conditioning (AC) penetration, particularly in the residential sector, is a key driver of electricity use during the summer. For most of the regions where PMA customers are located, AC saturation is close to 100%. Therefore, most of the increases in electricity use triggered by warm temperatures in these regions result from more intense use of the existing AC stock (intensive margin of AC use). The intensive margin of AC use can simply be captured through the CDD variable. However, in the Pacific Northwest, AC penetration is significantly lower than in the rest of regions of interest, but it has been growing rapidly in recent years. For this reason, it is important to include both the intensive and extensive margin of AC load among BPA customers. The extensive margin (increase in AC use due to installation of additional AC units) is modeled through an interaction term between CDDs and AC saturation for each preference customer. Data on AC penetration at the for each preference customers level are not available. Instead, the probability of AC ownership in the service territory for each BPA preference customer is estimated using data from the Northwest Energy Efficiency Alliance (NEEA) household survey. The NEEA survey contains detailed information on AC ownership and other household characteristics and demographic information for approximately 2,000 households in Washington, Oregon, Idaho, and Montana. These data, combined with CDD observations for the city/place where each household is located, are used to estimate the following equation.[31] Then, the estimated equation is applied to income and CDD data for each of the preference customers (setting the rest of explanatory variables at their mean levels) to obtain estimates of AC saturation that can be interacted with CDDs in the seasonal load equations discussed in Section 6.3.2.

The estimated equation is a probit regression where the dependent variable (AC) is a binary variable coded as 1 or 0 depending on whether mechanical cooling equipment is installed in the household or not.

[31] Annual CDDs were computed using temperature data from Daymet and a *BLP* of 64°F.

$$E(AC|X) = P(AC = 1|X) = F(\alpha + \beta_1 * CDD3 + \beta_2 * Income + \beta_3 * Income * warm + \beta_4 * BuildCat) \quad \text{(G.1)}$$

where F is the cumulative distribution function for the standard normal distribution, *CDD3* is the average number of CDDs in the 3 years before the survey interview taking place, *Income* is the income per capita reported in the survey, *Income*warm* is an interaction variable between income and CDDs but takes the value zero for households experiencing less-than-average CDDs within the region. Finally, *Buildcat* codifies whether the respondent lives in a single-family home, a manufactured home, or a multifamily building (5 or more units), or a manufactured home. Table H.2. shows the estimated coefficients from the probit regression.

Table H.2. Estimated coefficients from probit model of AC penetration in the Pacific Northwest

Term	Estimate	Standard error	Slope
(Intercept)	**−0.5327**	0.1432	
CDD3 (in thousands)	**2.2165**	0.2347	0.8826
Income (in tens of thousands of dollars)	**0.032**	0.0195	0.0127
Income × warm	−0.0084	0.0275	−0.0033
Building category = multifamily (5 or more units)	**−1.114**	0.1333	−0.2378
Building category = single family (4 or fewer units)	**−0.3255**	0.122	−0.0473
Number of observations		1,857	

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Observations were adjusted by the survey sampling weights provided in the NEEA database. Manufactured homes are taken as the reference building category and dropped from the equation. Slope = $f(x'\beta_k)\beta_k$ where f is the density function for the standard normal distribution.

The strongest effect on the AC ownership probability comes from the CDD variable. Income per capita also has a positive and statistically significant effect. The probability of AC ownership is lower in single-family homes and multifamily buildings than in manufactured homes. Figure H.1 shows the nonlinear relationship between CDDs and the probability of AC ownership implied by the estimated probit model, while setting the values of all the other variables in the equation at their sample means. The estimated curve follows the S-shaped profile that is typical of technological diffusion models. The estimated rate of AC adoption in response to CDDs slows down considerably beyond 900 annual CDDs at which point AC penetration is ~90%.

Table H.3. BPA Total Retail Load Estimated Equations

	Winter		Spring		Summer		Fall	
	Daymet	Livneh	Daymet	Livneh	Daymet	Livneh	Daymet	Livneh
Intercept	**−5.7262** (2.0858)	**−5.2917** (1.8442)	**−5.0647** (2.141)	**−4.724** (1.9267)			**−3.6767** (2.3103)	−2.7073 (2.0679)
HDD	**0.3641** (0.014)	**0.3701** (0.015)	**0.0538** (0.0169)	**0.0414** (0.0179)	**−0.0027** (0.0011)	**−0.0031** (0.0009)	0.0072 (0.005)	**0.0187** (0.0078)
CDD			0.0038 (0.0081)	0.0019 (0.0087)	**0.0497** (0.0064)	**0.0443** (0.0064)	−0.0023 (0.007)	−0.0004 (0.0074)
CDD × estimated AC penetration			0.0104 (0.0065)	0.0099 (0.0067)	**0.0396** (0.0053)	**0.0359** (0.0053)	0.0072 (0.0058)	0.0071 (0.0061)
CDD × irrigation			**0.072** (0.0137)	**0.068** (0.013)	**0.039** (0.0053)	**0.0378** (0.0051)	**0.0659** (0.0135)	**0.0675** (0.0137)
Precipitation	−0.0011 (0.0038)	0.0038 (0.0031)	0.0039 (0.0057)	**0.0157** (0.0071)	**−0.0064** (0.0009)	**−0.007** (0.0013)	**−0.0096** (0.0027)	**−0.0071** (0.0033)
Avg_Price	**−0.4452** (0.1896)	**−0.3896** (0.1694)	**−0.5928** (0.2137)	**−0.5665** (0.1944)	**−0.5842** (0.2161)	**−0.5322** (0.2003)	**−0.5316** (0.2226)	**−0.4528** (0.2037)
IncomePerCap	**0.2963** (0.1556)	**0.2623** (0.138)	**0.3672** (0.1549)	**0.3432** (0.1396)	**0.2842** (0.1459)	**0.2801** (0.1331)	**0.2913** (0.1668)	0.2124 (0.1491)
February	**−0.0773** (0.0042)	**−0.0771** (0.0046)						
January	**0.0052** (0.0027)	0.0012 (0.0027)						
March			**0.0952** (0.0151)	**0.0899** (0.0154)				
May			0.0171 (0.0276)	0.0158 (0.0288)				
July					**0.0118** (0.006)	**0.0138** (0.0059)		
June					**−0.0347** (0.0101)	**−0.0355** (0.0077)		
October							**−0.1148** (0.0124)	**−0.1104** (0.014)
September							**−0.1703** (0.0359)	**−0.1514** (0.041)
Estimator	Random	Random	Random	Random	Within	Within	Random	Random
adj.r.squared	0.4849	0.5148	0.33	0.2933	0.2888	0.283	0.3921	0.3772
NRMSE	0.1078	0.1031	0.2119	0.2121	0.1390	0.1337	0.2048	0.2040
Nobs	3,799	3,526	3,800	3,527	3,801	3,528	3,801	3,528

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

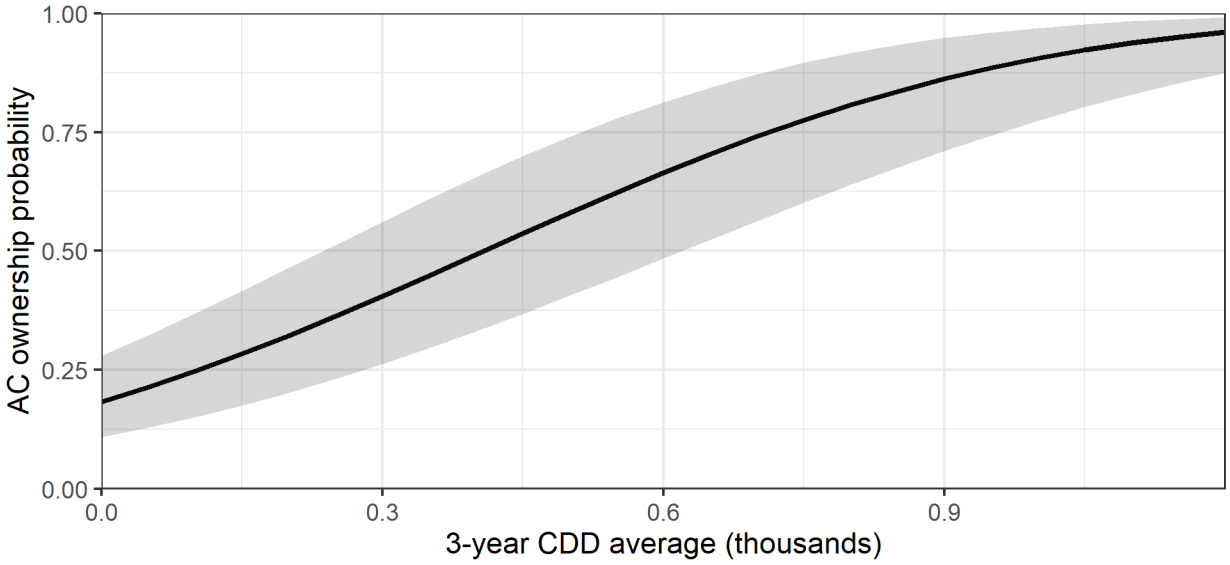


Figure H.1. Relationship between CDDs and probability of AC ownership implied by the estimated probit model.

Table H.4. WAPA-UGP Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	**0.1154** (0.01)	**0.1315** (0.018)
CDD	**0.0271** (0.0083)	**0.0264** (0.0082)
Precipitation	−0.0102 (0.0104)	−0.0109 (0.0085)
Avg_Price	**−0.3539** (0.0666)	**−0.3355** (0.0639)
IncomePerCap	**0.4555** (0.0539)	**0.4499** (0.0532)
Estimator	Within	Within
adj.r.squared	0.3926	0.4027
NRMSE	0.0373	0.0364
nobs	3,838	3,710

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

Table H.5. WAPA-CRSP Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	−0.0073 (0.0449)	0.0128 (0.0444)
CDD	−0.0096 (0.0103)	−0.0018 (0.0069)
Precipitation	−0.0092 (0.0203)	−0.0213 (0.016)
Avg_Price	**−0.3536** (0.1525)	**−0.3365** (0.1491)
IncomePerCap	**0.3014** (0.0833)	**0.2988** (0.0817)
Intercept	−0.6744 (1.1246)	−0.7432 (1.1073)
Estimator	Random	Random
adj.r.squared	0.2179	0.2216
NRMSE	0.0415	0.0414
Nobs	1,080	1,044

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

Table H.6. WAPA-SN/RM Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	0.055 (0.0475)	0.0069 (0.0391)
CDD	0.0198 (0.0482)	0.0154 (0.0424)
Precipitation	−0.0076 (0.0075)	−0.0037 (0.011)
Avg_Price	**−0.3521** (0.0666)	**−0.3552** (0.0641)
IncomePerCap	**0.2868** (0.078)	**0.3004** (0.0777)
Intercept		−0.9241 (1.1684)
Estimator	Within	Random
adj.r.squared	0.0857	0.1782
NRMSE	0.0405	0.0413
Nobs	505	488

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

Table H.7. WAPA-DSW Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	0.0029 (0.029)	0.0444 (0.0309)
CDD	**0.1532** (0.0928)	0.0798 (0.0721)
Precipitation	−0.0119 (0.0177)	−0.0098 (0.012)
Avg_Price	**−0.8622** (0.3179)	**−0.8724** (0.3181)
IncomePerCap	**0.5201** (0.1609)	**0.5795** (0.1667)
Intercept	**−5.1083** (2.8589)	**−5.4541** (2.7566)
Estimator	Random	Random
adj.r.squared	0.3943	0.4005
NRMSE	0.0598	0.0592
Nobs	529	511

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

Table H.8. SWPA-Integrated Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	**0.0806** (0.0172)	**0.0957** (0.02)
CDD	**0.1553** (0.0263)	**0.1639** (0.0344)
Precipitation	−0.0163 (0.0245)	0.0026 (0.0271)
Avg_Price	**−0.2337** (0.1251)	**−0.2282** (0.1231)
IncomePerCap	**0.3042** (0.0979)	**0.3091** (0.097)
Estimator	Within	Within
adj.r.squared	0.2504	0.2559
NRMSE	0.0465	0.0463
Nobs	1,020	986

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

Table H.9. SEPA-Cumberland Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	**0.0825** (0.007)	**0.0837** (0.0074)
CDD	**0.0781** (0.0085)	**0.0802** (0.0083)
Precipitation	**−0.0261** (0.0084)	−0.0083 (0.0077)
Avg_Price	**−0.5099** (0.0406)	**−0.5057** (0.0385)
IncomePerCap	**0.4004** (0.0287)	**0.4072** (0.0276)
Estimator	Within	Within
adj.r.squared	0.2826	0.2917
NRMSE	0.0255	0.0250
Nobs	4,860	4,698

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

Table H.10. SEPA Kerr-Philpott Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	**0.1526** (0.0173)	**0.1403** (0.0174)
CDD	**0.0464** (0.0168)	**0.0605** (0.0179)
Precipitation	**−0.0286** (0.0126)	**−0.0219** (0.0118)
Avg_Price	**−0.3517** (0.0898)	**−0.355** (0.0862)
IncomePerCap	**0.3329** (0.0402)	**0.3414** (0.0393)
Estimator	Within	Within
adj.r.squared	0.3737	0.3899
NRMSE	0.0277	0.0273
Nobs	1,882	1,819

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors.

Table H.11. SEPA-GA/AL/SC/JW Electricity Sales Estimated Equations

	Total sales	
	Daymet	Livneh
HDD	**0.1003** (0.0085)	0.0935 (0.0094)
CDD	**0.0873** (0.0188)	**0.105** (0.0214)
Precipitation	**−0.0619** (0.0117)	**−0.0612** (0.0122)
Avg_Price	**−0.4077** (0.075)	**−0.3872** (0.0755)
IncomePerCap	**0.4849** (0.0431)	**0.4814** (0.0447)
Intercept	**−3.872** (0.5956)	
Estimator	Random	Within
adj.r.squared	0.4798	0.4927
NRMSE	0.0324	0.0304
Nobs	2,481	2,398

Notes: Bolded values indicate that the estimated elasticity is statistically significant at the 10% level. Values in parenthesis are standard errors

APPENDIX I. VIC AND PRMS PERFORMANCE

Further analysis for HUC8s within PMA regions is presented. Monthly NSE and monthly mean absolute error for simulated runoff for both VIC and PRMS driven by Daymet and Livneh meteorological forcing were calculated. The first finding indicates that for all four set of simulations, HUC8s with lower annual WaterWatch runoff tend to demonstrate lower NSE values. This suggests that VIC and PRMS both performed better for wetter HUC8s compared with HUC8s in arid conditions. Second, despite a poor monthly NSE, these HUC8s also show low mean absolute errors and therefore, the influence on the overall water budget may be reduced.

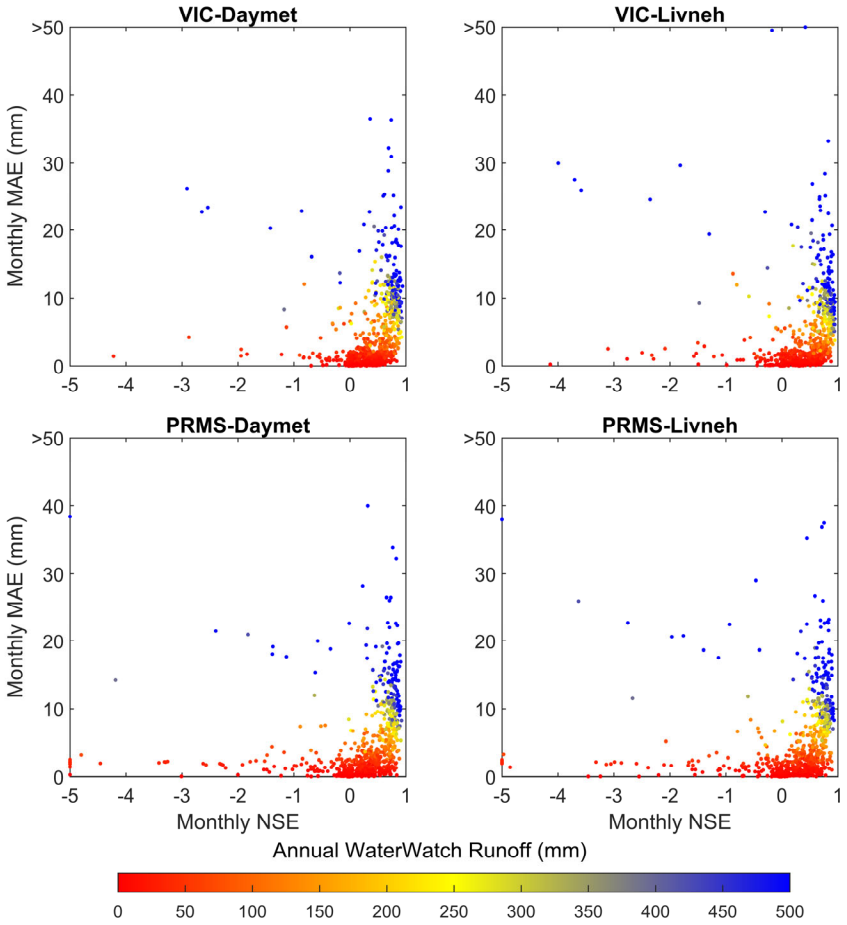


Figure I.1. Monthly performance metrics in terms of NSE and mean absolute error for HUC8s within PMA regions.

(This page intentionally left blank)

APPENDIX J. SUMMARY OF RESPONSE TO KEY REVIEW COMMENTS

To ensure the accuracy and quality of this study, an extensive external peer review was conducted in October 2021. More than 40 external reviewers, including federal and state power and water resource managers, climate and hydrologic researchers, and policy analysts, were invited to provide comments and feedback on this study. The main review comments, as well as the detailed responses from the research team, are provided here. Statements in italics are the reviewer comments; they are followed by a description of revisions made to the final report.

Reviewer 1

1.1 Suggest a clearer response in Executive Summary regarding findings of multi-data set, multimodel, multi-downscaling effort. For so much effort, the results didn't clearly specify that the climate models dominate the uncertainty in projections.

- Thank you for the recommendation. We included it as a key finding in the results described in the Executive Summary: "Specifically for runoff and generation, the selection of climate model remains the dominant factor in many parts of the CONUS, but the role of HMs and PMs (along with other factors) can be more important in some locations and seasons. This analysis addresses the needs for a comprehensive evaluation across all components in the modeling chain for a more informed hydroclimate impact assessment."

1.2 Consider consistency in reporting scaling of information in degrees with kilometers in parentheses in all cases, not just some.

- We appreciate this suggestion. Since there is no unique transformation between degrees and kilometers, we could not report kilometers in all places. We only reported kilometers in places that could otherwise be misleading.

Reviewer 2

2.1 Regarding the correlations between PRMS/VIC and Daymet/Livneh at SNOTEL points, although an overall analysis is sufficient for this section of the report, it would be beneficial for the report to sample points near the reservoirs/basins of interest to understand local/regional differences. With snowfall a big factor in the western United States, correctly simulating snow/SWE is obviously very important. Small differences in annual snowfall in the control period can lead to large differences in future simulations.

- Given that majority of the SNOTEL sites fall within PMA watersheds, we have not sampled the points explicitly. However, we provided PMA boundaries (basin of interest) on the map for a clearer identification. Given that the SWE is summarized at the HUC8 level, a clear relationship between the reservoir and nearby SNOTEL may not be directly established.

2.2 The last section might benefit from a section relating regional/sub-regional PMAs to the greater interconnections and hydropower's impacts on a future grid. This of course might be speculative, but discussing future hydropower in the context of grid stability could be a future endeavor for this report.

- Thank you for the suggestion. We have included it as a potential future work and discussed in Section 8.3: "The representation of hydropower operations must be enhanced to support the operation of future grids. For example, enhancement to a weekly time step, and representation of storage targets and daily fluctuations range and ramping, would be needed to further integrate the hydropower projections into unit commitment production cost models. This enhancement would allow for evaluating the impact of climate change on the ability for hydropower resources to sustain grid resilience under extreme events conditions, as well as under new power grid infrastructure and generation portfolio with new markets and net load profiles. The water management model would also be enhanced with the ability to add or remove reservoirs."

2.3 What is intensification of hydrologic cycle? In what way?

- It means that there will be stronger mass and energy exchange in the hydrologic cycle, which may result in increasing intensity and frequency of extreme hydrologic events (for both droughts and floods).

2.4 Explain what Tier 1 means in the context of SPP and RCP.

- Tier 1 represents four high-priority scenarios. The text has been updated to clarify this term. The revised text states, "GCMs are physics-based tools to study Earth system response to natural climate variability and changes arising from climatic changes arising from anthropogenically driven increases in greenhouse gas emission and radiative forcing."

2.5 What has been done to mitigate dependence between CMIP scenarios/models? How significant is this dependence on bias?

- The GCM evaluation is based on the historical simulations, which have no relationship with the future scenarios. Biases are with respect to observations or reanalysis that only are available in the historical period. All scenario simulations represent the future projection period where no reference observations exist.

2.6 Add a few lines or an appendix explaining potential impacts on omitting groundwater component or perhaps assurance that the basins evaluated for hydropower concerns are less impacted.

- In this study, we estimated climate projections by evaluating hydrologic response in the future time period in reference to the hydrologic response in the baseline period of the same set of models. Given the lack of coupled surface water and groundwater interactions in both of these time periods, any deficiencies in the hydrologic process representation are expected to reduce the impacts on the signal of climate change.

Reviewer 3

3.1 It would have been nice to see more evaluation of metrics like CAPE, Shear, IVT, or the North American Monsoon in the analysis of GCMs; not sure there is much to do about it at this point.

- We primarily focused on the metrics that are relevant for all CONUS regions. The suggested metrics are indeed useful but are more geared toward regional climate characteristics. We have added text to highlight this point: "Note that our selection of metrics represents a wide range of spatiotemporal climate characteristics that are common across the CONUS and does not include those features that are unique to specific regions, such as integrated water vapor transport through atmospheric rivers in the western United States, the monsoonal climate in the southwest and tornadic environment in the central and eastern United States."

3.2 Can you provide some context with respect to recent historical variability? I think WAPA is pretty worried about declines in Colorado River basin flows recently. Can you say with confidence that that is truly random variability?

- Given the nature of GCM modeling and simulations, the projections cannot be directly compared with the historical variability. However, recent trends in the historic period do indicate an ongoing drought condition in the Colorado River Basin (e.g., Williams et al., 2021).

3.3 We have seen really large sensitivities in VIC to the representation of vegetation in the Colorado River basin (changing increases in flows to decreases in flows). The new VIC vegetation maps are much more realistic and result in streamflow trends that are more consistent with recent observations.

- Thank you for sharing this. Our VIC vegetation maps were derived from MODIS and are different from other commonly used 6 km versions (e.g., used by Livneh et al., 2015). Nevertheless, our VIC parameters are still under the time-invariant assumption, which is one modeling subject that should be

examined in future efforts. We will be glad to conduct further model inter-comparison to identify a defensible way to enhance the representation of vegetation in VIC.

3.4 Since the GCMs are such an important factor, but you only have 6, can you provide more analysis of the other 20+ GCMs based only on changes in precipitation and temperature in each of the major regions?

- Figure 2.7 provides a comparison of selected GCMs vs. the rest.

3.5 Why aren't the bias-corrected metrics perfect (bias = 0, correlation = 1, etc.)? Was there a training and validation separation? (I didn't see it mentioned.)

- To the best of our knowledge, there is no perfect bias correction approach that can preserve all metrics. For our application, as evident from Figure 2.9, Figure 2.10, and Figure 2.11, the bias-corrected output is reasonably close to observations at the monthly averaged scale. However, the biases in extremes that are derived from daily data still exist.

3.6 Temporal frequency of PDO, ENSO, and NAO are also useful in addition to pattern correlation.

- Although greenhouse gas concentrations are aligned in the observations and historical CMIP6 GCMs simulations, the natural modes of climate variability, such as PDO, NAO, ENSO, are not. Therefore, lack of correspondence in temporal frequency of these modes between simulations and observations cannot be confidently used as a measure for model validation, as it is not straightforward to distinguish between the inconsistency arising from internal climate variability and that arising from model deficiencies.

Reviewer 4

4.1 It may be worthwhile to provide summary tables/statistics of the key differences of results from this report vs. the previous report (not just the approach as in line 67–91 and key findings (404–429). The format could then be included in future reports, as well, to maintain a consistent way of comparing previous results.

- Thank you for the suggestion. However, given the large differences among CMIP3, CMIP5, and CMIP6, and other differences in temporal periods, emission scenarios, DSs, and HMs/PMs, a direct numerical comparison with previous assessments would not be fair and can be misleading. We opted to keep our current high-level discussion focused on the main findings.

4.2 An intensive workflow for the bias correction and overall analysis was established. Now that it is established, does this mean that it can be applied again when the next round of CMIP scenarios come out and that more time/effort could be devoted to other analysis aspects (given that such a rigorous methodology was established here)? If so, this should be stated as it is a major accomplishment/outcome of the work.

- Thank you. Yes, this established workflow and capabilities can be applied to other current and future CMIP GCMs. Model enhancement may still be needed with the availability of more future observations, but should not be less demanding. We will highlight the value of this workflow in the upcoming report to Congress.

4.3 In Section 2.4.4, to the extent possible, it would be helpful to connect the projection limitations to implications for hydropower estimates.

- Thank you; we agree with this suggestion. We have added a new Section 8.2 to summarize the overall assessment limitations and implications.

4.4 Some more discussion on why the hydropower generation is simulated at the monthly time step after so much effort in the validation and historic baseline at the daily time step might be needed. If hydropower is only estimated at the monthly scale, why not do the validation at that scale, too? Would that change anything?

- Hydropower was estimated at the daily time scale, but calibration was performed at a monthly time scale. No existing daily hydropower data set (turbined flow, spills) covers the CONUS, and the current water management model does not represent the complex daily hydropower operations in response to load and price variations. However, the monthly hydropower generation is dominated by water availability and long-term water management rules, which is why we kept the hydropower projection at a monthly time scale for now.

4.5 What is the potential significance of not including non-powered dams in the WRES method, if any?

- Since WRES is a lumped model, as long as the collective operation of the majority of reservoirs in a region do not change substantially, WRES can still function reasonably well.

4.6 Why not focus only on hydropower regions for selecting best GCMs and evaluating model performance, given that is the most important aspect of the work?

- There are two main reasons: (1) given the nature of GCM modeling and climate science, the evaluation of GCMs should not be too narrowed for smaller watersheds (in this case, head water basins above the hydropower plants), and (2) DOE WPTO would eventually like to share the downscaled hydroclimate projections to support the broader nonfederal hydropower stakeholders, and hence would like to select GCMs with reasonable performance at the CONUS scale.

4.7 There is a lot of detail in the results sections and readers could get a little lost or find it difficult to understand the importance of some of the information. It might be helpful to add 1–2 key takeaway messages up-front in each regional subsection and/or a regional summary (1–2 paragraphs) at the start of each regional section, especially since much of the technical detail is repetitive across regions.

- Thank you for the suggestion. We have provided a general overview of the federal hydropower fleet and PMAs in Section 1.2.1, and specifically discussed the different PMA electricity marketing practices in Section 1.2.2. We have also included PMA-specific discussions in Sections 3.1, 4.1, 5.1, and 6.1 to describe the differences/similarities in the hydrology, customer base they serve, topography, and so on. Section 7 highlights the important climate change results across the CONUS. We also present bulleted findings in Section 8.1 to highlight the key messages and reiterate the important results from this assessment.

4.8 Is there a significance or effect of using a 40-year baseline period vs. a 20-year future period in the hydropower analyses? And, if so, is using 1980–2000 skewing results in any way for the baseline, given the accelerating effects of climate change and recent chronic droughts?

- The length of the reference period will affect the magnitude of future changes, but its substantial effect on the projected trajectory is unlikely. The inclusion of a more recent historical period may reduce the strength of the projected future changes. However, the overall trajectory is unlikely to be affected.

4.9 I expected the discussion on demand vs. hydropower generation to be a little more extensive. Lines 2489–2493 are an important (probably the most important) part of the overall analysis, bringing together both demand and supply, which is essential at understanding system vulnerabilities and hydropower's potential role/contribution. Some more discussion around this would be helpful vs. discussion on the energy demand on its own. This is the same for all regional discussions.

- We agree that this is a limitation of the modeling framework. The hydropower operations are presently responsive to water availability and water management rules. The modeling framework does not yet integrate water and power system operations in the hydropower scheduling. The

hydropower scheduling would be responsive to both load and prices signal, which require power grid modeling under future scenarios to provide this information (load assigned to hydropower and locational marginal prices). This was beyond the scope of this work.

4.10 Date and installed capacity graph for each region to understand aging infrastructure issue could be helpful.

- We appreciate this suggestion, but this is outside the scope of this study. The issues of aging infrastructure are discussed in other efforts (https://www.ornl.gov/project/annual-markets-and-trends-report).

4.11 It might be helpful to discuss storage levels a little more and some graphics associated with that so we can understand the effects of reservoir rules/operations.

- The water management part of WMP has been evaluated in previous publications (e.g., Voisin et al., 2013) for the seasonal variations in release and storage, which are mentioned in Section 2.

4.12 It would be of value to separate or explain the differences between the two hydropower methods and their respective results in the regional discussions. Grouping them together in the same graph makes it difficult to see if the two methods span the same range or not, or the extent of their effect on hydropower generation estimates.

- We agree with the needs to explore the differences further. We are conducting further evaluation that will be published in a subsequent study.

4.13 It may be of interest to discuss changes in evapotranspiration rates and how they may reduce runoff under future climates, especially when an increase in rainfall does not result in an increase in runoff (if that happens) over a season.

- Thank you for the recommendation. Although not showed in this report, we analyzed evapotranspiration rates and the general findings are similar to what we reported in the previous 9505 assessment (Kao et al., 2016). To avoid confusion about different types of evaporation, we focused discussing the direct reservoir evaporation in this assessment.

4.14 The overall approach to the electricity demand analysis is a little confusing and at times seems out of place. For instance, in the regional sections, the electricity demand analysis is not necessary for some of the concluding statements that relate hydropower to demand, such as "The increase in generation could help curb the trend of increased replacement power purchases experienced in recent years and have a positive effect on power rate." The integration of the two analyses (hydropower generation and demand) seems like an afterthought and the demand results are not well integrated into the discussion that should really focus on climate impacts on hydropower (or hydropower's role in the grid). Without a better integration of demand and hydropower generation in the discussion/analyses, and given the simplicity of the demand method, the demand analysis seems like a weakness in the report.

- Inclusion of the demand analysis is motivated by the idea that risks to power marketing can come from either the supply or the demand for federal hydropower. This is particularly true for BPA since BPA must supply electricity that follows the load shape of most of their preference customers. For the rest of PMAs, we have modified the title and write-up of these sections to clarify that changes in the demands served by their preference customers do not affect PMA power marketing practices directly. However, we still feel it is valuable to make reference to the potential impacts of climate change on electricity demand in those regions. First, it is informational for the PMA preference customers. Second, how well federal hydropower generation seasonal patterns match the seasonal load patterns of preference customers matters for the competitiveness/value of the federal hydropower allocation for those customers. In acknowledgment that the demand analysis is most relevant for BPA, the report to Congress that will be developed out of this report will only mention findings from the demand analysis for BPA.

4.15 In the interregional comparison, what is the reason to have the discussion/analysis include regions outside the four PMAs? This was a little unexpected and seems like it is outside of the scope of climate impacts on hydropower, unless a better or clearer implication/link can be made. See some of the detailed comments for examples.

- We foresee that some of the hydroclimate projection data that we developed in this study may also be used to support other nonfederal hydropower stakeholders. Therefore, we expanded the geographical scope of the interregional comparison section to highlight the broader climate change impacts outside of the four PMA regions.

4.16 Can the loss in evaporation at reservoirs be translated into loss of power?

- Although evaporation loss can decrease the amount of water storage and eventually lead to decrease power, it is difficult to translate because power generation is a decision within the boarder multi-purpose reservoir operation. If the power generation model already has evaporation explicitly included, then this translation can be made by comparing electricity with and without increased evaporative loss.

4.17 How much of a concern is it that hydropower generation decreases at the same time that electricity demand is increasing—in the summertime—in BPA? Some more discussion on this implication may be of interest to readers and could shows the important of grid management and resource integration.

- We have added some text discussing the implications of the mismatch between hydropower generation availability and electricity load during the summer months in the BPA region: "The decreased correlation between hydropower generation availability and customer demand will likely result in increased replacement power purchases by the PMA during the summer months. The cost of replacement power purchases must then be passed through to the rates that BPA charges its customers. Finding ways to maximize the value of potential winter generation surpluses and minimize the needs for replacement power during the summer will be important to maintain competitive and affordable rates for BPA and ensure reliability."

4.18 In Figure 1.1, why is the EIA Form 860 data set used and not the EHA database, or are they the same for this context? It might be helpful to separate pumped storage hydropower and conventional (squares vs. circles on the map).

- Yes, they are the same for this context. Regarding pumped storage hydropower, since some of the federal pumped storage turbines were actually operated as conventional hydropower, the distinction is not clear.

4.19 Given that not all aspects of uncertainty are considered in this study (e.g., water demands, land use/land cover, reservoir rules), it might be worthwhile to state that GCMs are the greatest source of uncertainty in this particular study, and that other factors may have greater/unquantifiable uncertainties that may have more/less impact on hydropower vs. GCMs.

- Thank you. This issue was also mentioned by other reviewers. We have clarified and discussed the needs to better understand uncertainty in the future work.

4.20 What are the potential consequences of using wind speed from a different data set? Why not use a data source that contains all climate variables of interest?

- The effect varies. In general, wind speed is less sensitive to runoff but is very sensitive to evaporation. Although some data sets do provide wind speed, they are still derived from commonly used reanalysis data sets such as NCEP and NARR. A gridded, station-based wind speed data set is still lacking, to the best of our knowledge.

4.21 Evaluation for four CONUS regions—why not basin level or have them be specific to the hydropower regions? Figure C.1 includes regions not in the analysis, and it is not clear why they should be included in the evaluation given that GCM performance can vary across those basins/regions.

- Given the coarse resolution of GCMs (typical resolution ~200 km), we have avoided basin-scale evaluation. Figure C.1 includes all areas that are part of hydrological simulations and assessments. The final GCM selection using their overall performance across CONUS is also motivated by the fact that they have been used as boundary forcing for dynamical downscaling through a regional climate model over a domain that covers the CONUS and parts of Canada and Mexico. Therefore, those GCMs are preferred that display better performance across the CONUS on average.

4.22 The bias correction (max of 150%) is for both negative and positive bias? If so, how many instances of each occur?

- It is a future-to-baseline quantile ratio. This specific rule is rarely implemented. However, it is needed since some GCMs may include very small baseline rainfall values, which may lead to unreasonably large future-to-baseline ratio.

4.23 In Figure 2.7, in the annual figure, it looks like only 5 (not 6) black diamonds are displayed.

- Thank you. We have verified that there are indeed six data points. Two data points are very close.

4.24 In Figure 2.12 and Figure 2.13, can the NSE/hydrologic model performance be aggregated to the basins where the hydropower analysis is implemented? For instance, the northeast looks great for performance, but no hydropower analysis is done there. On the other hand, the lower WAPA region performs less well, but it is hard to see how the basins are aggregated.

- We have included additional discussion on model performance to highlight HUC8s within PMA areas in Section 2.5.3.2.

4.25 (1) It is unclear how the daily limits ("updated daily" as in the text) are affecting monthly operations. (2) Also: WM was run daily, but generation is monthly; more discussion on the reason for this may be helpful—is it because of model performance/comparison purposes or is it just better for the analysis? Does daily hydropower not matter for the analysis and if not, why not?

- We have updated this. WM simulates the operations at a daily time scale, which allows to perform the needed water balance and river routing. The daily operations are aggregated to a monthly time scale and reported as monthly in the report. At this time, the monthly estimates are more accurate since WM does not represent the complex daily hydropower operations that respond to the power grid needs. At a monthly time scale, the estimates are reasonable because there are driven by water availability and water management rules. The monthly time scale allows us to focus on the multimodel comparison for the uncertainty quantification.

4.26 To better understand validation results, the NSE and R values (validation metrics) could be shown at the plant level for the WM method, or the distribution of the results across the plants so that the readers can understand if the performance of the model varies much across facilities.

- For consistency throughout the report, we focused on the multimodel approach, which requires a monthly resolution and sub-region domain approach. This is also the resolution of the PMA data that we obtained. We will provide plant-level details in future publications.

4.27 What about the limitation of spatial resolution? At the resolution modeled, are some plants aggregated together as well as their capacities. Are there any issues with that?

- WMP currently runs at a 1/8th degree spatial resolution with one reservoir per grid cell. For a limited number of power plants, some need to be consolidated in which the reservoir capacity is accumulated. For this report, there were no cases in which we needed to aggregate power plants.

4.28 Reservoir evaporation losses are considered, but are they simulated at individual plants? It is a little confusing how the link between storage calculated for the reservoir/hydropower is linked to the evaporation estimates. I expected to see two evaporation estimates based on the WRES and WM models. Is this a third storage estimate or based only on the WM estimate?

- The evaporation loss estimation was made independent from WRES and WM. There were no evaporation losses estimated from WRES and WM.

4.29 Wind speed effect is mentioned explicitly; what is the effect of the other variables, or even just the maximum?

- We conducted sensitivity analysis to quantify the relative importance of each variable (i.e., by fixing three variables and change the fourth one from one forcing data set to another). The results suggest that wind speed is the most dominant contributor to the evaporation rate differences. We also compared wind speed from these two data sets with in situ observations at Lake Mead and Lake Mohave. Results suggest that the NARR wind speed (used by Daymet) is more accurate than that from NCEP1 (used by Livneh). We have updated the description in the revised report.

4.30 Is it the precipitation, evapotranspiration, or another driver that is causing the difference in runoff? It is a little difficult to decipher the driver causing the difference because runoff is the result, not the driver. Not sure if it is possible to highlight this as it is likely a bigger effort.

- We appreciate the interesting question. Here, we provide a preliminary understanding of differences in runoff arising from different methodological choices (i.e., CM, MetF, DS, and HM). Furthermore, resolving the individual impacts of different hydrologic process can definitely help us understand whether the differences are arising from precipitation, evaporation, or other processes. However, this will require additional comparisons/sensitivity experiments that are beyond the scope of the current report.

4.31 In Table 3.1, it may be of interest to note the type of dams (e.g., ROR, primary purpose) dams with multiple purposes subject to multiple stressors that while not addressed here could be important to simply note. Could be added to an appendix if not the main report. It would also provide a nice regional contrast; some regions may have more multipurpose dams than others.

- The information in Table 3.1 and APPENDIX B is at the plant level. There were two steps involved in linking federal hydropower plants to their corresponding dams. First, linkage between plant data and NID ID (using ORNL's Existing Hydropower Assets data set, https://hydrosource.ornl.gov/dataset/EHA2021), and second, linkage with the NID data set. For 35 of the 132 federal hydropower plants, there is no unique correspondence between a plant and a dam (i.e., there are multiple dam structures for the NID ID identified as corresponding to the plant and, in some cases, the listed purposes for the multiple dam entries are different). For mode of operation (e.g., ROR, peaking), the data we have shown are not available for 17 of the 132 hydropower plants. Because of these data issues and the fact that this extra information is not crucial to the main topic of the report, we did not include these extra variables.

4.32 Was land cover incorporated in the model or observed? This applies to any time land cover is mentioned.

- Land cover was captured in the parameters of hydrologic models (VIC and PRMS).

4.33 In Figure 3.14, is the figure showing both hydropower generation methods? If so, these should be separated or there should be some discussion on whether there is a difference between them or whether they span the same ranges, especially as the uncertainty discussion later in the section points to this as a major source of uncertainty.

- The differences between the methods are evaluated in Section 2. In Section 3 and after, we treat the hydropower projections as a multimodel ensemble where the focus is on the structural uncertainty.

Figure 3.14 is a combination of CMs, HMs, PMs, data sets, and downscaling approaches. In a couple regions, the PM differences become a substantial contributor to the overall uncertainty after climate models. The objective of the ANOVA is to discuss the contributions of each of those components in Figure 3.14. We refer to Zhou et al. (2018) for an evaluation of WMP with respect to a statistical approach specifically.

4.34 Regarding WAPA-2, WAPA-3, and WAPA-4 results, readers may struggle relating this projected increase to current conditions in the Colorado River basin, and some discussion on this may be helpful. This also applies generally to the findings of the entire report that show an overall increase in hydropower; readers might find it difficult to see existing climate effects and expect an increase in water/hydropower availability, so some discussion on the explanation of this may be helpful.

- Thank you for the suggestion. We have added additional discussion and references in the revised report. In particular, the increase of inflows will not be steady and will likely accompany more extreme hydrologic events (i.e., both floods and droughts). Also, the results are mainly long-term averages computed over 40 years and do not capture extreme hydrologic events, sub-seasonal variations, or complex hydrologic processes over mountainous terrain, which dominate water availability and hydropower generation in the western United States.

4.35 As with BPA, this concluding remark could be expanded. The entire section focuses on technical details of electricity demand, but it is most relevant in the context of hydropower, which is only given a couple sentences at the end and seems insufficient to really provide meaningful insight.

- For WAPA, SWPA, and SEPA, the section has been renamed as "Climate Change Impacts in Regional Electricity Demand." For these three PMAs, the federal hydropower allocation that preference customers receive from the PMA is only a fraction of the portfolio of generation resources they use to serve their end-use customers. Any load growth projected by preference customers of these PMAs due to climate trends or other drivers would have to be met through generation sources other than federal hydropower. Therefore, for these PMAs, the effect of climate-driven changes in load only comes through potential requests from the preference customers to change the fraction of the federal hydropower allocation that they receive during each season (and the PMAs will only consider those requests if they are compatible with all other operational constraints). In short, changes in demand matter most to BPA and that is why, in this new version, they are discussed in a bit more detail for BPA than for the other PMAs.

4.36 The impact of future infrastructures may also be of interest (use of capacity expansion models, not just production cost models); for instance, high electrification scenarios may change the way that hydropower is used in the grid.

- Although we agree with the needs, the focus of this study is more on understanding the impacts of climate change on existing infrastructures, not the new ones (i.e., through capacity expansion model). The future hydrologic projections can, however, provide scenarios of future water availability to support the understanding of future energy market and grid.

4.37 Uncertainty/sensitivity of runoff calculations may also be of interest—using different evapotranspiration functions and land use/land cover assumptions could affect hydropower estimates but do not seem to have been included in this report.

- We agree that sensitivity of runoff simulations to transient factors should be of interest. However, the hydrologic models in this study use an assumption of stationary land use/land cover in the future period. Future assessments may explore the use of projected land use/land cover types, which may help develop a more realistic representation of runoff/streamflow projections. We have added this discussion in Sections 2.5.5 and 8.3.

Reviewer 5

5.1 Defining 20-year projection time periods means that internal variability will be a larger portion of the projection signal than would be using a longer averaging period. Thus, there will be more noise in the projections that should be discussed somewhere. It's not a limitation of the GCMs, but of how you're defining the future periods.

- We agree with this view. Overall, we selected a 40-year future period to reduce the influence of interannual variability. The 20-year window is only used as way to discuss how change may gradually occur during the 40-year period.

5.2 Performing a relatively coarse dynamic downscaling and then doing a statistical bias correction and downscaling seems somewhat unnecessary. I understand that any dynamic downscaling needs bias correction and your resource limitations, but the limited resources could have been used to explore improved hydrologic model calibration, other advanced statistically downscaled methods, and so on. Performing a coarse dynamic downscaling and then a downscaling and bias correction (which is essentially what the daily bias correction really is) could remove a lot of the purported benefit of the dynamic downscaling (e.g., physically consistent covariance across state variables, physically consistent spatial patterns). The idea that bias correction and downscaling of a dynamic simulation may change the spatial features and covariance of state variables could also be discussed (I do not believe it is). Some work has even shown that careful bias correction of a dynamically downscaled future may improve the realism of the RCM output.

- We appreciate this insightful comment. Although there has been various research on downscaling, thus far, there has not been a consensus on a best way to proceed (specifically, for the purpose of hydroclimate impact assessment). Therefore, instead of selecting a specific way of downscaling, our approach is to include two very different ways (dynamical and statistical) in this framework, and to understand their influences through the ANOVA. We will be glad to provide our model outputs to support further collaborate so that we can better understand the deeper differences and values of different downscaling approaches.

5.3 Even with the use of many metrics, the evaluation may still miss the point of identifying good models for process-oriented reasons. For example, what if a model had poor evaluation scores, but those errors are due to biases that are actually easier to correct for than the models selected. One could think selecting this model may be beneficial given process considerations. Models that faithfully represent trends over the globe and CONUS may still be rated poorly using the metrics defined in Table A.1, but those may be the exact models we want to use. The paper by Bukovsky et al. (2019), "Weighting a regional climate model ensemble: Does it make a difference? Can it make a difference?," begins to describe some of these issues.

- Thank you for the insightful comment. We agree that in general, model evaluation and selection are still a very open research topic. Even if a model can show good performance across all evaluation metrics, scientifically, there is still no way to prove that it can also correctly project future conditions. However, considering the large number of CMIP6 models, we believe that the evaluation approach proposed in this study can be an effective way to filter out those relatively poor performing models. We will continue to research and collaborate with experts in the field so that we may enhance our evaluation approaches in the future assessments.

5.4 Is there any need to explicitly discuss these projection results in the context of the previous two 9505 reports beyond stating what the methods of the previous two reports were. It seems to me that the CMIP6 projections for the WAPA region are trending wetter from CMIP3 to CMIP5 to CMIP6. Other regions may be experiencing similar changes from each projection iteration. What does this mean (if anything) for the US hydropower assessment? Contextual discussion could help stakeholders better understand these new findings, as well. Again, apologies if I missed this discussion in the report.

- There are substantial differences in the climate projections among CMIP3, CMIP5, and CMIP6, which include new models and more complex assumptions on global change. It is beyond the scope of this report to specifically address the differences that are covered in the National Climate Assessment and Intergovernmental Panel on Climate Change reports. The projection for more hydropower in the WAPA region, which is experiencing a multi-year drought, indeed requires some contextual clarification. We specify in Section 8.3 how to interpret projections of hydropower increase amidst ongoing multi-year drought conditions and the need to represents water demand scenarios in future work to improve the interpretation of those assessments.

5.5 I may be misunderstanding the bias correction, but it seems like many of the evaluation metrics shown in Figure 2.9 should be matched exactly by definition, so this whole section is really just affirming that the downscaling and bias correction methods are working. The real question is, what do the bias correction/downscaling schemes do to the observed and projected trends from the input data sets?

- Given the nature of our bias correction approach, the trends in the raw GCM and RCM are preserved. Whether if a GCM or RCM produces comparable trends to the historical observations is a more involved question that is not analyzed in this study.

5.6 In Section 2.9, did you implement the bootstrapping method discussed by Bosshard et al. (2013) to remove biases in ANOVA due to different sample sizes across the factors? Here you have six GCMs but only two methods for each of the other factors. By default, this means the variance attributed to the GCMs will be biased high if I remember that paper correctly.

- Thank you for this comment. We recognize that the current modeling design inherently under-samples other modeling choices over GCMs. This can lead to a biased contribution from GCMs. As an alternative approach, the single factor subsampling approach used by Bosshard et al. (2013) can greatly help diminish these biases. However, it can also potentially cause underestimation of the individual contribution of sampled factor (Fan et al., 2020). We have also acknowledged this as a limitation to the study and added more description in Section 2.9, and a potential area of improvement in future assessments/reports.

Fan, Y., K. Huang, G. Huang, Y. Li, and F. Wang (2020), An Uncertainty Partition Approach for Inferring Interactive Hydrologic Risks, *Hydrol. Earth Syst. Sci.*, *24*, 4601–4624, https://doi.org/10.5194/hess-24-4601-2020.

5.7 Regarding Figure 3.8, in this and other figures, there is a clear shift across the Canadian border that should probably be discussed. Apologies if I missed that discussion.

- The SWE is reported and summarized at the HUC8 level based on the common resolution achievable through aggregation of hydrologic response from VIC and PRMS, although a finer resolution is desired for SWE. Given the aggregation at the HUC8 level, some differences are expected around the boundaries. Second, the model calibration is conducted for surface runoff and for each HUC8 independent of their spatial location; in other words, the calibrated parameters may have differences from HUC8 to HUC8 and therefore lead to greater differences around the boundaries (especially national boundary).

Reviewer 6

6.1 Assessment limitations: It's important to make clear that the uncertainty in the projections is not just a matter of scenario and model choices (and their associated uncertainties). A big part of the actual uncertainty lies with the uncertainty of policy choices, demographic, and socio-economic developments. All that uncertainty would still exist, even if our physical models and data sets were perfect (which they are not, even if we do understand a lot of the physical mechanisms). So even if you had 100 GCMs and 50 scenarios, you still would not be capturing the true uncertainty. Basically, what I am saying is that

capturing "the full range of uncertainties related to all possible modeling choices" (lines 181 and 182) is still a subset of the true uncertainty.

- Thank you; we fully agree with the need to highlight the broader uncertainty. We have updated the statement: "However, although this approach enables the evaluation of the relative influence of each main modeling factor, the multimodel projections still do not represent the full range of uncertainties related to all possible modeling choices, and to factors that cannot be captured in numerical models (e.g., policy, demographic, socio-economic developments)."

6.2 Regarding VIC calibration, using VIC in water balance mode and with a single elevation band for model calibration of the energy balance mode with five elevation bands likely had a much larger effect on model results than is reflected in the current narrative. I am reasonably sure that it likely had a much larger effects than the differences between the different versions of the Daymet and Livneh data sets. This should be mentioned as one of the caveats. That said, since the performance characteristics are provided for the validation period, they do reflect the performance of the energy balance/five elevation band model, so I think that providing the additional caveat is sufficient (see lines 1173–1176).

- As noted in the report, since our calibration target is at the monthly scale (not daily), we found that the simplified setup may work sufficiently well (i.e., the total watershed aggregated runoff generated from these two approaches are consistent). Also, in our testing in some western headwater watersheds, the differences due to the choice of Livneh vs. Daymet were in fact larger. We will be interested to conduct further evaluation across the CONUS, which is currently outside the scope of this report.

6.3 Somewhere early in the methods, you should explain why all the runs are performed for all of the CONUS rather than just the areas upstream of the selected hydropower projects. I understand why this is necessary for the dynamically downscaling simulations, but why was this done for all the remaining steps?

- We foresee that some of the hydroclimate projection data that we developed in this study may also be used to support other nonfederal hydropower stakeholders. Therefore, while our specific SWA focus is on PMA regions, we conducted CONUS-scale simulation and analysis when possible.

6.4 When reporting the performance of the hydrologic models (e.g., Figure 2.12 and Figure 2.13), it is a bit misleading to report the model performance over all of the CONUS rather than the model performance in the HRUs that contribute directly to the flow at the 9505 hydropower plants. At the very least, you should also report the latter and redo Figure 2.12d and Figure 2.13d for just those locations.

- Since the national-scale results may be of interest to the boarder community, we keep Figures 2.12 and 2.13 as they are. However, we have included a new APPENDIX I to highlight the model performance of HUC8s within PMA watersheds.

6.5 Please clarify how streamflow is actually used in the study. How is it different from how the runoff (described in the next paragraph) is used? I found this confusing.

- The HYDAT streamflow database was used to produce a runoff time-series to supplement WaterWatch for Canada's Upper Columbia River basin.

6.6 Why not used the naturalized flows (such as the NRNI data set for the Columbia River basin) for those areas where they existed? Neither PRMS nor VIC include the effects of regulation.

- Naturalized flow data sets (e.g., NRNI) are certainly valuable to assess the model performance at the regional scale. However, we relied on the USGS WaterWatch runoff data set, which provides a spatially consistent monthly runoff estimates across the CONUS. This allows us to conduct a systematic and spatially consistent hydrologic model calibration/validation at national scale.

6.7 In Figure 2.10, why is the performance so good in Canada? Same for Figure 2.11.

- Although there can be other reasons, including good meteorological forcings, the spatial units on the Canadian side are much larger (among the largest of all HUC8s), thus resulting in a more lumped hydrologic response, leading to better performance during calibration.

6.8 It's unclear what "multiscale" refers to? Time scales? Spatial scales?
- Both. The calibration has two steps. The first step was done at PMA the regional level and an annual time step, and the second step focused on plant level at a monthly time step. We have modified the text to be more specific about the scales to which we refer.

Reviewer 7

7.1 The results presented in the third hydropower assessment represent a massive amount of work that is summarized fairly succinctly at multiple spatial scales. This aspect of the report should enhance its regional utility for multiple stakeholders and utilities managers.

A particularly noteworthy analysis effort in this assessment was the identification and quantification of the major sources of variability and hence uncertainty in the modeling results. To date, I have only seen multimodel assessments such as this conducted on individual basins or sub-basins. It was of particular interest to see the change in major variance contributing factors across regional, basin, seasonal, and annual scales. This is an impressive aspect of the results presented within in the report.

The authors should be commended for the time and effort devoted to the generation of such a high-quality synthesis product focused on the impact of climate change on hydropower generation across the CONUS. Most of my thoughts and potential comments while reading the report fell outside the scope of work being conducted here and thus would not be worth addressing in a formal review process. Minor typos or potential figure suggestions are all that are noted in my specific comments.
- Thank you. We appreciate your positive comment and encouragement.

7.2 The emissions scenario numeric coding is specific to this report (though clearly based on SSPs and RCPs). Although the numbers are intuitive to those who follow the literature as to which has a high vs. low emissions scenario, you may want to add that information either as a secondary arrow in the key of Figure ES.3, or define what the emissions scenario numbering refers to in the caption or main text. This would help a more general readership.
- We use the most recent emission scenario names in CMIP6 (e.g., SSP585), which are slightly different from the names used in CMIP5 (e.g., RCP8.5).

Reviewer 8

8.1 Some scientific principles used could be explained further for people who are not familiar with them.
- Thank you for the suggestion. We have revised several parts of Section 2 to better explain the methodology.

8.2 Can we quantify the "error"/gap in modeling by not accounting for these site-specific features?
- We recognize this concern. We believe that it depends on the intended objective and targeted spatial scale. Since our biggest goal is to evaluate the broad climate change impacts on all federal hydropower plants, we do not have the ability to address many site-specific issues. We eventually discuss the needs of site-specific assessments in future studies.

8.3 Has the space been normalized before applying PCA? How? Dividing by standard deviation, by max-min, other?

- Yes, the EOF analyses is based on the normalized relative errors. This information has been provided in APPENDIX C.

8.4 Why was PSO used instead of other Evolutionary algorithms?

- PSO is a widely used swarm intelligence–based optimization technique and is effective for use in hydrological model calibration (Jiang et al., 2007). PSO is conceptually simple, computationally efficient, and can take advantage of parallel processing in large high-performance computing systems (Guo et al., 2014; Lee et al., 2006), which is suitable to our application in this study.

8.5 Is it possible to include a brief discussion about the impacts of future operational rules?

- We have added a discussion in Section 2.6.4.2 on future work on ensemble of parameterization for the water management model with multiple operations and water demand scenarios. For this work, we rather focus on multimodel hydropower models with one similar parameterization.

8.6 The shape of a tetrahedron assumption seems to be an over-simplification. Even though the actual shape can vary significantly among reservoirs, isn't there a slightly more complex but more accurate assumption that can be made for the shape of all reservoirs?

- The tetrahedron is an assumption adopted by Fekete et al. (2010). The most recent database that "calibrates" the shape of reservoirs to different shapes includes one by Yigzaw et al. (2018), but this is a research product where the overall volume ends up modified and more work is needed to calibrate the shape for hydropower operations. We noted this comment in future work, as well.

  Yigzaw et al. (2018), A New Global Storage-area-depth Data Set for Modeling Reservoirs in Land Surface and Earth System Models, *Water Resources Research*, *54*(12), 10372–10386, https://doi.org/10.1029/2017WR022040.

8.7 It seems that the approach does not account for potential future improvement in building efficiency or demand response. This would lower the electricity demand in the future. Could the authors discuss this?

- We added the following statement in Section 2.8.3: "The demand analysis results presented in the rest of the report extrapolate the estimated responsiveness of electricity sales to temperature to future periods. The extrapolation implicitly assumes that there will be no substantial changes in the types of heating and cooling equipment being used or their energy efficiency as well as no substantial growth of demand response programs. Improvements in heating and cooling technology and more widespread implementation of demand response programs can help reduce the responsiveness of electric load to temperatures."

8.8 Why is it assumed that there should be a linear relationship between the logarithmic values of these quantities instead of their original values?

- The use of logarithmic transformations of the values was chosen so that the estimated coefficients can be interpreted as elasticities (percent increase in electricity load in response to a 1% increase in degree days) and be more easily comparable across PMA regions.

8.9 Why would the electricity demand or energy sale depend on the precipitation?

- Precipitation is included as a control variable to help isolate the effect of temperature on electricity demand following the approach in related recent literature (e.g., Auffhammer et al. (2017); Davis and Gertler (2015).

Auffhammer et al. (2017), Climate Change Is Projected to Have Severe Impacts on the Frequency and Intensity of Peak Electricity Demand Across the United States, *P. Natl. Acad. Sci.*, *114*(8), 1886–1891, https://doi.org/10.1073/pnas.1613193114.

Davis and Gertler (2015), Contribution of Air Conditioning Adoption to Future Energy Use Under Global Warming, *P. Natl. Acad. Sci.*, *112*(19), 5962–5967, https://doi.org/10.1073/pnas.1423558112.

Reviewer 9

9.1 Who is the audience of this report?

- The audience of this study includes SWA-specific stakeholder groups (e.g., congressional members, WPTO, PMAs and their customers, federal hydropower owners/operators) for a better understanding of risks associated with federal hydropower marketing. We plan to extend the assessment to nonfederal hydropower fleet which may benefit the broader hydropower community (e.g., hydropower developers, environmental stakeholders, and other federal/state energy water agencies) for a general understanding of risks associated with regional and national long-term hydroclimate trends.

9.2 Will this report be published online? If so, are you going to include the links to the previous reports?

- Yes, we will set up a publicly accessible website through the ORNL HydroSource portal (https://hydrosource.ornl.gov/) to release the new and past reports, as well as the key data (e.g., downscaled precipitation, temperature, and simulated runoff) to promote the usage of 9505 assessment outcomes.

9.3 Is it possible to add a few sentences to the Executive Summary to summarize the findings in plain/simple language? To translate the scientific findings.

- Thank you for the suggestion. We have revised the Executive Summary based on a variety of review comments. We will further prepare plain/simple description on the public website for easier interpretation.

Reviewer 10

10.1 I just want to begin by saying this is an impressive piece of work.

- Thank you. We appreciate your positive comment and encouragement.

10.2 More care should be taken when using the term "uncertainty." While yes, it should be understood that the ensemble variability is being used as a crude (and maybe biased) proxy of uncertainty, it is important that we are precise in our usage of terms and not use "uncertainty" when we strictly mean something else. In places, what is called uncertainty is strictly the distribution or the variability of projections, and in some cases, even more precisely, the variance of a sample of projections. There are places where simply replacing "uncertainty" with "variability" would be sufficient. I have noted a few places in the text, but another careful read of all places where uncertainty is discussed should be done.

- Thank you for the suggestion. This issue was also raised by other reviewers. We have clarified the meaning of "uncertainty" in multiple places of the revised report.

10.3 In general, there is a dearth if information in the captions. One should be able to view a figure and understand what is being shown without digging through the text. All components of a figure should be defined in the caption. This would have the benefit of eliminating some of the text in the body of the report.

- We understand this suggestion. We actually received somewhat conflicting suggestions to simply the captions and include the full description in the text. We eventually opted to the current style that may satisfy our publication formatting requirements.

10.4 For Figure ES.3, are these the percentage changes in the CONUS-wide averaged precipitation, or the averages of the percentage changes calculated by pixel (or basin, etc.)? This would make a difference because the former would give more weight to the wetter parts of CONUS in each season (and annually).
- The percentage changes were calculated for the CONUS-wide averages, not by pixel.

10.5 I have an issue with the general statement that GCMs represent the greatest source of uncertainty. What can be said is that these studies show that the variability in GCM output was often the main cause for the variability in the sample of projections they considered. However, it has not been established as a general truth that GCMs are the main source uncertainty. The studies typically use many more GCMs than they do hydrological models, hydropower models, or downscaling methods. Therefore, these other parts of the modeling chain are highly under-sampled relative to the GCMs in most studies.

Moreover, these studies may have shown that variability in GCM output is the main cause for the variability in projections for SOME but NOT ALL hydrological metrics. For example, Chegwidden et al. (2019) found that low flow variability was mainly due to differences in hydropower models.

Finally, Giuntoli et al. (2018) in their experiment showed that the hydro/land surface model contributes most to the variability in hydro projections across CONUS. Their result is important because they use a larger number of hydro models than most of these other studies.
- We recognize this concern. The text in Section 2.1 has been revised for clarity and to incorporate these suggestions.

10.6 I would cite Alder and Hostetler (2018) as they show not only the influence of downscaling, but more importantly, the choice of observational training data in hydro projections in the western United States. https://doi.org/10.1029/2018WR023458.
- Thank you. The recommended reference has been included.

10.7 I would say "variability," not "uncertainty." It is worth a sentence here or even early saying these studies use projection variability as an index or proxy for uncertainty, and a crude index/proxy at that.
- We have revised the description to address this comment. We understand the concern about using "uncertainty." However, "variability" may be confused by the internal variabilities within the GCM models. We opted to keep using "uncertainty" and provided more clarification in the revised report.

10.8 Yes, it is good that there does not appear to be a strong bias. However, the much smaller range from the six models would indicate an under-sampling and thus an underestimating of GCM uncertainty. On the other hand, the concern that some CMIP6 models have to high equilibrium climate sensitivity is somewhat abated since you have not selected the GCMs that give the most warming (by mostly luck or an outcome of your model selection process?).
- Given the limited resources, we made some necessary comprises in deciding the maximum number of models that we can include in this study. The decision was made jointly with all federal hydropower stakeholder groups through a series of coordination meetings. Similarly, the decision of GCM selection criteria was made jointly with all stakeholders. The results of selection correspond to our selection criteria.

10.9 The statement that a previous study found DBCCA to be one of the two best does not give the reader much information if we don't know what other DSs they considered. Did they consider others used frequently in the United States (e.g., BCSD, BCCA, MACA, LOCA)?

- Werner and Cannon (2016) compared seven DSs—BCCA, DBCCA, BCCA with quantile mapping reordering, bias correction spatial disaggregation (BCSD), BCSD using minimum/maximum temperature, the climate imprint delta method, and bias-corrected climate imprint delta method. They did not perform it for MACA and LOCA, since the watershed is in Canada. To our best knowledge, both MACA and LOCA codes are not openly available, so it's difficult to implement both methods independently.

10.10 I think John Abatzoglou found that such a large number of analogues was too many. I would see how it might oversmooth the spatial pattern. I think he now goes with about five for MACA, maybe?
- Thank you for the insightful comment. We chose the number of analogues based on the DBCCA methodology described by Maurer et al. (2010). Since our spatial domain is for the entire CONUS, the analogue number should still be sufficiently large to capture the CONUS-scale pattern.

10.11 How are humidity and radiation (short and long wave) biases addressed? If they are not, this should be mentioned.
- We did not perform any correction on humidity or radiation. The humidity and radiation used in hydrologic models are derived from MT-CLIM algorithm (https://www.ntsg.umt.edu/project/mt-clim.php), which is commonly used in many hydrologic studies.

10.12 Do VIC and PRMS simulate the change in water use efficiency due to increasing CO_2? If so, how? If not, this should be stated explicitly. Water use efficiency is a key factor when considering the effect of global warming on transpiration and runoff.
- Thank you for the comment. We recognize the importance of time varying water use efficiency in the hydrologic modeling. However, the hydrologic models in this study utilize temporally invariant parameters such as water use efficiency, land use/land cover, soil and vegetation, and so on. We have explicitly stated this in Section 2.5.5.

10.13 MTCLIM has known issues with radiation that should be acknowledged and that are likely to introduce bias. See, for example, Slater (2016). https://doi.org/10.1175/JHM-D-15-0087.1 I don't know how well MTCLIM simulates humidity.
- Thank you. We have included the limitations of MTCLIM in the Hydrologic Model Limitations section.

10.14 It should be made clear, and probably much earlier, that the differences in output among GCMs are due to two factors that have not been separated here:

1) Climate model differences
2) Internal (i.e., natural, unforced) variability (IV)

The total variability due to IV wrt precipitation can be much larger than that coming from differences in the models themselves, especially (but not exclusively) in the near term. See, for example: Rupp et al. (2017, https://doi.org/10.1007/s00382-016-3418-7) and Alder and Hostetler (2018, https://doi.org/10.1029/2018WR023458). Though estimation of IV was not done in this study, there are ways to estimate the contribution from IV (Rupp et al. 2017; Chegwidden et al. 2019). Some of the other studies cited in this report also estimate IV.
- Thank you for the insightful comment. We have provided additional discussion in Section 2.9.

10.15 What is up with the huge difference between Daymet and Livneh in BPA-1 after ~2007? That goes beyond systematic biases and into the serious error category.
- Parts of BPA-1 lie on the Canada side, where the Livneh data set may not be processed consistently resulting in these differences. Unfortunately, there is no further information for us to diagnosis the

difference. We elected to use both the data sets for analysis to maintain consistency with other regions.

10.16 The statement that annual precipitation shows higher interannual and intermodel variability compared with annual temperature has little meaning unless you have some common reference frame with which to compare precipitation and temperature. How are you measuring variability and what are the units? The measures, to start, would have to have the same units (or be unitless), but even that is not enough. Is what you really care about the signal to noise? If, for example, signal to noise is measured as the higher frequency variability vs the long term trend, they you can say that the signal to noise is clearly lower for precipitation. I would be inclined to delete, or at least rewrite, this sentence.

- Thank you for the insightful comment. We have modified the text accordingly.

10.17 I think Mote and Salathe (2010, https://doi.org/10.1007/s10584-010-9848-z) and Rupp et al. (2017, https://doi.org/10.1007/s00382-016-3418-7) should be referenced and discussed here as well. See also Chegwidden et al. (2020, https://doi.org/10.1088/1748-9326/ab986f) for the impacts of snow regime changes on hydrologic extremes in the region.

- We appreciate this suggestion. Since there is no unique transformation between degree and kilometer, we cannot report kilometers in all places. We only reported kilometers in places that could otherwise be misleading.

10.18 On what basis do you determine that wind differs the most? Daymet and Livneh temperatures have been shown to be quite different in some figures. No data have been shown on radiation and vapor pressure deficit. Even so, the important parameter here is the sensitivity of evaporation to the difference in wind relative among the data sets to the sensitivity in evaporation to the difference in the other variables (temperature, radiation, vapor pressure deficit) among the two data sets. This would be important to calculate and demonstrate, otherwise I can't judge the statements made here.

- We conducted sensitivity analysis to quantify the relative importance of each variable (i.e., by fixing three variables and change the 4th one from one forcing data set to another). The results suggest that wind speed is the most dominant contributor to the evaporation rate differences.

10.19 In situ wind speed measured, and where? At which reservoirs?

- Given the limited in situ observations, we compared wind speed from these two data sets with in situ observations at Lake Mead and Lake Mohave. Results suggest that the NARR wind speed (used by Daymet) is more accurate than that from NCEP1 (used by Livneh). We have updated the description in the revised report.

10.20 It would be helpful to have a reminder if NARR and NCEP1 winds were bias-corrected by Daymet and Livneh, respectively. I can't imagine near-surface winds from NCEP1 being useful. Even NARR near-surface winds are dubious. How did RegCM4 winds compare to NCEP and NARR? Is it conceivable that RegCM4 could be better than both NCEP1 and/or NARR, and the bias correction in the DBCCA process actually made the RegCM4 simulated winds worse?

- Daymet does not provide wind speed, and the wind speed used by Livneh is directly from NCEP1. Given the higher spatial resolution of NARR (that we paired with Daymet), the wind speed should have better performance than NCEP1. We compared wind speed with in situ observations at Lake Mead and Lake Mohave, and the results suggest that NARR wind speed is more accurate. The raw RegCM4 wind speed is closer to the range of NARR, perhaps owing to the closer model spatial resolution. Thus far, there still hasn't been a wind speed reanalysis data set based on in situ observations. Wind speed from reanalysis data sets such as NARR or NCEP1 remain the necessary surrogate.

10.21 I do not think that you have established that Daymet-based evaporation rates are more reliable unless you've compared the effect of biases of all variables contributing to evapotranspiration.

- We performed several analyses, including calculating the differences between two meteorological forcing data sets, and testing the sensitivity of the four key variables to arrive this observation. More comprehensive analyses are to be conducted in a follow-up study.

10.22 Figure 5.8 reveals high spatial variability in runoff changes, even in adjacent basins. In contrast, temperature changes should be spatially uniform, relatively. Do these spatial patterns of runoff changes match the spatial patterns of precipitation changes? This comment applies to similar figures for the other PMAs, too.

- The spatial patterns of P95 or P05 were not explicitly compared. However, the spatial heterogeneity may be a result of different hydrologic response in adjacent basins (HRUs or HUC8s). Given that a spatially independent parameter calibration was performed for each of the HUC8, there may be some sensitive behavior especially for low flow conditions. The effect is even more pronounced since we show percent change instead of absolute change.

10.23 PMPs are very large and their estimated (stationary) return periods can be astronomical. Global warming will put systems under high stress more often. This can lead to system failure even if the estimated PMP is not reached. The Oroville Dam failure comes to mind.

- We fully agree with this comment. It's the reason we specifically call out PMP here.

Reviewer 11

11.1 This report is impressive in its rigorous review of climate model evaluation and selection for usage with this project. Is the approach laid out novel in the field of climate model/among those who use climate models? In other words, the various selection criteria, metrics, and evaluation process? Is this type of meta-analysis for selecting proper climate models for specific project scopes/sectors/goals well established with standard practices similar to those used by the authors? Either way, given the growing prevalence and usage of climate models, some greater introduction and context on this type of model evaluation and selection might assist a reader like myself who is not deeply familiar with the topic.

- Thank you. We appreciate your positive comment and encouragement. We have revised several parts of Section 2 to better explain the methodology.

11.2 This project combines groups of models by feeding the outputs of certain models as inputs to others, ultimately connecting climate change predictions to the effects on national hydropower. As such, the methodology described becomes impressively complex, and as a reader I wondered if a visual summarization could be provided to assist with clarifying the overall modeling scheme. Specifically, I did not see an explicit visual representation of all of the various inputs and outputs from each set of models and where each were derived from or sent to. Page 64 provides a good version for the hydropower model specifically). There is a high level of interconnectivity in this work that is otherwise missed/ difficult to understand when each modeling methods subsection is described separately. Furthermore, without a guiding roadmap, I also found myself wondering why specific parameters were being collected from certain models before discovering that they were inputs to later models. For example, I did not see why "Wind Speed" was being collected in the meteorological models (page 41) until I arrived at the "Reservoir Evaporation Modeling" section ~30 pages later.

- A summary framework is presented in Figure 2.1. The individual modeling components and relevant information is described in Section 2.1. We will incorporate this suggestion and provide more concise summary in the report to Congress.

Reviewer 12

12.1 I like the bulleted list of findings near the beginning of the discussion! Nice synthesis of results. The overall structure of the report is good and the analysis thorough, and I appreciate the subsections on how the model is calibrated, validated, and also its limitations. The interregional analysis was very interesting, I liked the and the graphics are high quality!

- Thank you. We appreciate your positive comment and encouragement.

12.2 It was noted that the methodology changed from past assessments compared with this assessment, with this assessment using a multimodel framework to increase confidence of findings; it would be useful for the authors to state in this section how the results on pages 2 and 3 differed or were affected by different assumptions—did the former assessments tend to underpredict certain conditions? Did the results based on previous assumptions present bias in a specific direction? In other words, how much of the old assessment can still be used with high confidence?

- Thank you for the suggestion. The needs to better address uncertainty have been discussed in Section 2.1 and throughout the report.

12.3 It's interesting to see the runoff changes spatially; the heterogeneity I'm assuming is related to the lack of elevation? I think it was mentioned earlier in the chapter that topography is a major influencer--could be useful to bring it up again to describe why we're seeing this relatively uniform pattern across the region.

- The spatial heterogeneity may be a result of different hydrologic response in adjacent basins (HRUs or HUC8s). Given that a spatially independent parameter calibration was performed for each of the HUC8, there may be some sensitive behavior especially for low flow conditions. The effect is even more pronounced since we show percent change instead of absolute change.

12.4 Some of the references used are quite old (2009); I'm assuming however that it's used still because the trend of growing demand for industrial water for areas like Atlanta still hold true, as is the trend in environmental/recreational use of water.

- Yes, issues of competition for water storage in some of the SEPA reservoirs still continue and are sometimes referred to as the "tri-state water wars" (https://atlantaregional.org/natural-resources/water/tri-state-water-wars-overview/).

Reviewer 13

13.1 The selected GCM seem to have been based on overall evaluation over CONUS. Wouldn't have been more objective and useful for this extensive assessment study to evaluate and choose the subset of models passing the evaluation test for the four PMA regions. Even more specifically, targeting subregions. For example, WAPA region represents a vast area west of Mississippi River and WAPA-6 (northern California region) is geographically disconnected from the rest. Or WAPA-3 and WAPA-4 areas, Colorado River basin, in our opinion, would have benefited from a more focused evaluation. There are two reasons for that, first: the Colorado river basin and its major reservoir system plays a critical role in both water and power supply to the west. Second, historical records in the past four decades shows a statistically significant downward trend in amount of precipitation feeding the Colorado river (http://rainsphere.eng.uci.edu/). Furthermore, WAPA-5 which is the Rio Grande basin represents its own unique hydrologic characteristics and probably deserved.

- We understand the inclusion of figures showing only the CONUS-scale results gave an impression that the evaluations were carried out at the CONUS scale. In reality, the evaluations were carried out by dividing the CONUS into four regions (Figure C.1). The final skill score is based on the average over four regions. We now have replaced CONUS-scale figures with the regional scale figures that were the basis of CONUS-scale rankings. Given that several GCMs have a gird spacing ~200 km, it is not very meaningful to go to finer scales for comparisons. Therefore, we have avoided basin-scale evaluations. The final GCMs selection using their overall performance across CONUS is also

motivated by the fact that they have been used as boundary forcing for dynamical downscaling through an RCM over a domain that covers CONUS and parts of Canada and Mexico. Therefore, GCMs are preferred that display better performance across the CONUS on average.

13.2 The selected GCMs seem to also be subject to the hard constraint of data availability. Specifically, those GCM with sufficient sub-daily outputs for the purpose of dynamical downscaling were considered. Meaning that only 14 out of those 37 GCMs that provided 6-hourly, 3D atmospheric fields were included in the pool. Is this possible reason that 6 emerging front runner GCMs were developed by a number of international centers and countries and none of the US models made the cut?

- Yes, this is true. NCAR and DOE GCMs did not have sufficient data for their evaluation. GFDL GCMs were highly ranked based on their performance. However, none of the three institutes had 6-houlry atmospheric variables available, which are needed for dynamical downscaling, at the time of decision making.

13.3 The reviewers suggest that the "observational data" to be defined more clearly. Evidently, Daymet and Livneh are data products that are derived from meteorological observations and reanalysis processes. In sections 2.5.3 (lines 1159-1161) and 2.7 (lines 1626-1629) there are references to the use of reanalysis and assimilated data for calibration and evaporation analysis respectively. Therefore, it is important to clarify the definition of use of "observations."

- We have clarified this issue in the revised report. We understand that both Daymet and Livneh are not strictly the "direct" gauge observations. However, given their algorithms, and the fact that they both use minimum reanalysis data, they have been considered as a close surrogate to observations in many published hydrologic literatures.

13.4 The Summary and Conclusion is very well written and organized. However, in the analysis of climate change impact on PMA costumers' demands, the specific results of BPA region is not presented (lines 4469–4483).

- Of the four items in the bulleted list summarizing climate change impacts on PMA customers' demands in Section 7, the first three apply to all the PMAs. The fourth item distinguishes the results of BPA from those of other PMAs. Now, we added a paragraph to refer to one other important difference between BPA and the rest of PMAs (i.e., only BPA has to follow the load of its customers). For more details, the reader should refer to the sections discussing each of the PMAs.

13.5 It goes without saying that authors should be complimented for an outstanding comprehensive work and well-organized report. There is no question about the value of analyzing regional implications of future climate change scenarios generated by GCMs, downscaled to river basins and their impact on future hydropower generation as a "clean" energy source. While such a detailed end-to-end modeling studies are valuable, but one wonders, how do they compare with the results of generating streamflows using old approaches (now dated) streamflow synthesis (Fiering, Myron B., Barbara Bund, and Barbara B. Jackson. Synthetic streamflows. Vol. 1. American Geophysical Union, 1971). This may be a philosophical point but for me who took a course in statistical hydrology in the early 1970s at UCLA where this monograph was taught and we were assigned to take a river and using USGS data perform a simulation of future flow scenarios. I have always wondered to what extent more sophisticated and newer approaches provide improved information for the intended purposes. The same argument has also been made about lumped vs. distributed hydrologic models in terms of their quality of performance and added value. In fact, the authors of the report (page 168, lines 4402–4413) raise a somewhat similar point regarding the differences between statistical and dynamical downscaling approaches employed here.

- Thank you. We appreciate your positive comment and encouragement. We will be interested to explore/collaborate with the reviewer to understand the in-depth questions related to streamflow and hydropower estimates. We believe that the initial modeling outputs from this study may help provide some insights into these questions.

13.6 While at the end of this exercise (model calibration) it may not matter much, but given all the issues and requirements for hydrologic model calibration, it is interesting that in your study many simplifying assumptions and shortcuts had to be implemented in the case of VIC calibration. This begs the question that with all simplifications and changes is it still the same model that the developers intended? It might be a philosophical issue as to why select a model with rather demanding requirements for its calibration which cannot be met and satisfied.

- We are unsure if we are qualified to fully answer this question. However, we would like to point out that VIC has been widely used for many hydrologic studies given its strong process-based foundation. It's more difficult to calibrate, but in general, most researchers agree that it is an appropriate tool to study the future hydrologic response to changing climatic conditions.

13.7 Is it a realistic assumption in the projection of hydropower generation that the reservoirs in the study would operate in the same way in the future?

- We agree with this concern. It will be of interest to explore how the current operations can be adjusted for future climate conditions. However, such an issue was not explored given the large number of reservoirs. This will be an important research topic for future watershed-scale studies.

13.8 It would be beneficial to present the analysis based on which the multivariate regression model (2.1) with specific lag time (one-month) was chosen.

- Please see Kao et al. (2016) for further discussion. The information is not included since there was no major structural model update since the previous assessment.

13.9 In section 2.6.3, the overall results for calibration and validation of two hydropower models WRES and WPMP is presented in Table 2.3. Then in the following Figure 2.16. the performance of WRES and WMP is presented for each PMA in the study. The reviewers suggest a paragraph to be added to the report that analyses these performances for each PMA separately.

- Thank you for the recommendation. Since the PMA-specific results have been discussed in-depth at each PMA section (i.e., Sections 3.3.1, 4.3.1, 5.3.1, and 6.3.1), we would refer the interested readers to those sections.

13.10 As mentioned, estimated changes in electricity demanded are most interesting for BPA because they are required to meet changes in load even with water supply shortages (or excess) or other external factors that make this electricity delivery costly or difficult. Nevertheless, the value of the energy and capacity products sold by other PMAs might also be affected by changes in the seasonal load patterns of their customers. The reviewers suggest a that a connection between this claim and the analysis in Section 3.3.2 to be added to the report.

- We have added a paragraph at the beginning of Section 3.3.2 reiterating why the demand analysis is particularly relevant or BPA: "As mentioned in Section 2.8, exploring the potential effects of climate change in PMA customer demand is particularly important for BPA because it is the only PMA required to meet the net load of those customers that request it—in contrast, the rest of PMAs are only required to deliver the energy and capacity volumes agreed in long-term contracts based on historical average or critical water year runoff availability the federal hydropower projects whose production they market."

Reviewer 14

14.1 How do (and if so, why do) analysis inputs (e.g., projections, forecasts, models) vary between this 9505 Report and the 9503 Report prepared by the US Bureau of Reclamation?

- There are some larger differences between the methodologies used in the 3rd 9503 and 9505 assessments. They are based on different versions of CMIP, downscaling methods, hydrologic models/parameters, and have different focus/scope (e.g., paleo-hydrologic information was used in

9503 but not 9505). Nevertheless, these differences are necessary to best address the specific needs of both assessments. The assessment teams on both sides have been staying in close communication and are aware of the approaches/rationales used in both studies.

14.2 Is it useful to explain what the SSP and RCP acronyms/numbers stand represent? The SSP numbers are discussed on page 6.

- We have revised text to describe these numbers: "GCMs are forced with a new set of nine scenarios, which are a combination of five Shared Socioeconomic Pathways (SSPs) and seven Representative Concentration Pathways (RCPs), to understand the Earth system response to an increase in anthropogenic forcing. Among them, there are four high priority or Tier 1 scenarios, including SSP1-2.6 (SSP126), SSP2-4.5 (SSP245), SSP3-7.0 (SSP370), and SSP5-8.5 (SSP585). Note that the first number represents one out of five SSPs while the last two numbers represent the radiative forcing level."

14.3 On page 11, lines 878–879, "SSP585 can be considered a proxy for changes under SSP370 a decade later." This is true for radiative forcing, but surely the rate of change difference and thus variations in climate feedback mechanisms may impact some fields. Maybe add a qualifier here? Specifically thinking of discussion in O'Neil et al., 2016, "The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6."

- This is true, but given that time difference is only a decade, we do not expect substantial non-linearities.

14.4 It would be useful to have a clear explanation of the baseline period used and the potential ramifications of this as a base line vs. comparing other time periods, or a subset of this window (e.g., 1990–2019).

- We use the maximum length of observation (i.e., 1980–2019) as the baseline to inform our analysis. To reduce the influence of model interannual variability (as mentioned by another reviewer), we purposely use the entire 40-year in this assessment.

14.5 Can a note be made that there could be a difference between modeled hydropower releases in this report and actual operations at Reclamation facilities due to operating agreements, coordinated operations, conservation plans, shortage reductions, minimum power pool impacts, and so on?

- Thank you for the suggestion. We have reiterated the limitations of the water management model for those specific type of operations in Section 2.

14.6 May be useful to explain the numbers in RCP7.0 and RCP8.5.

- We have revised the text to describe these numbers: "GCMs are forced with a new set of nine scenarios, which are a combination of five Shared Socioeconomic Pathways (SSPs) and seven Representative Concentration Pathways (RCPs), to understand the Earth system response to an increase in anthropogenic forcing. Among them, there are four high priority or Tier 1 scenarios, including SSP1-2.6 (SSP126), SSP2-4.5 (SSP245), SSP3-7.0 (SSP370), and SSP5-8.5 (SSP585). Note that the first number represents one out of five SSPs while the last two numbers represent the radiative forcing level."

14.7 How many of the expected CMIP6 projections were available for this study? Is there a chance the set presented here may have a skewed distribution of the full GCM set therefore not fully capturing GCM uncertainty?

- There were fourteen GCMs that had data available necessary for downscaling. Figure 2.7 shows the uncertainty captured by the selected GCMs (black markers).

14.8 Clarify if these are ensemble averaged valued shown in Figure 2.5 and Figure 2.6. Furthermore, is this of all available CMIP6 models or just those 6 selected for use.

- These are based on 28 GCMs that had data available at the time of analyses. This information has been provided in figure captions.

14.9 Reasoning for different spin up methods of the two HMs?

- We used a 3-year spin-up period for PRMS, as the same is also suggested by USGS for NHM-PRMS (Hay, 2019). For land surface models such as VIC, a 1-year spin-up period can be considered sufficient (Cosgrove et al., 2003). These references were added in the revised report.

14.10 It was earlier stated on page 22 that the differences between the meteorological forcings were small, yet now the differences are invoked to explain the differing R^2 values between VIC and PRMS—how may this impact performance of models in future projections?

- Even with different versions of Daymet and Livneh, the R^2 value is still very high (~0.93) for VIC, which should not result in major differences. However, the R^2 values can indeed be improved (by ~0.02) if the same versions of Daymet and Livneh are implemented.

14.11 Was any comparison of PRMS using the different versions done such as comparing the R^2 as done between VIC and PRMS in Figure 2.10 and Figure 2.11. Would be good to have an example of how to interpret this choice/understand if the impact on later HM ANOVA results are negligible.

- Thank you for the suggestion. The difference due to model versions was not explored in this study. This can be potential topics of future exploration.

14.12 Would be good to have some elaboration in the regional differences of ANOVA—is the larger HM variance fraction due to reduced net variance in the west and GCM variance reduced while HM variance is comparable to other regions?

- Thank you for the recommendation. We suspected that this may partial due to the HM performance. The HM performance is generally good in the eastern United States, where we have the least ANOVA HM variance. Further analysis will be required to verify this assumption.

Reviewer 15

15.1 What is the rational for excluding a long term (e.g., end of century) period?

- It is a joint decision made with DOE, PMAs, and other federal hydropower stakeholders since the first 9505 assessment in 2010. To inform the long-term hydropower marketing activities, a 40-year future period is considered sufficient and preferred. Considering other potential issues such as changing operation and equipment, the end of century timeline is too long for the purpose of informing power marketing activities.

15.2 In general, how does the formula/methodology presented here align with existing formulas/methodologies used by PMAs for more routine generation forecasting? Do any major discrepancies exist?

- Both regional PMs used in this study are quite different than the operational forecasting models used by USACE and Reclamation. Although it will be desirable to use the same operational model in this study, it was not likely given the large number of reservoirs. This will be an important research topic for future watershed-scale studies.

15.3 For ROR plants, the hydraulic heads (h, m) were assumed to be unchanged and equal to the dam height (h, m).” Is this how PMAs are currently characterizing ROR plants for more routine generation forecasting purposes?

- No, it is the WMP model assumption. We specify in Section 2 that the reservoir operations do not represent the routine operations by Reclamation, USACE, and PMAs. The reservoir operations emulate monthly operations and daily hydropower operations are based on a number of assumptions such as unchanged hydraulic head.

15.4 Are there any other drivers for spills that are not considered here (e.g., for flood control)?
- The seasonal operations can lead to spills and emulate spill for flood control. In that respect, the daily releases that are passed to the hydropower model include spill for flood control while the hydropower component deals with spills associated with environmental regulations if the reservoir was not already spilling.

15.5 For Figure 2.16, it appears that simulations consistently overestimate generation (e.g., 1996–2000, 2008–2010). Any thoughts on why? Is this due, in part, to the "insufficient and less accurate precipitation and runoff records" referenced on page 35 (lines 1540–1542)? If so, consider clarifying.
- We suspect that the deviation in these years are due to more site-specific issues that cannot be captured by our current models. The challenges of our hydropower models are discussed in the limitation sections.

15.6 Is there a need to put processes in place to ensure more sufficient, accurate records within the WAPA service territory?
- We fully agree with the needs. However, such an analysis will require site-specific reservoir models that cannot be extended to all 132 hydropower plants covered in this study. We have stressed the importance of conducting further watershed-specific study in the concluding section.

15.7 How might these reservoir-specific aspects be incorporated into future modeling efforts? Any sense of how model outputs may be impacted by incorporation of reservoir-specific aspects? What is the general impact of this limitation?
- We are working on a new version of the reservoir model that use data-driven reservoir operations on the water side and data-driven parameterization of hydropower operations. The general impact is that for this report the projections are at a monthly time scale and by sub-regions rather than at a daily time scale and for individual plants.

15.8 Are there any non-preference PMA customers that need to be considered here? Or do all PMA customers qualify as preference for this analysis?
- We focus on the preference customers because they represent the majority of PMA customers and they are the ones with which PMA typically have long-term contracts. Non-preference customers tend to not have firm power allocations from PMAs.

15.9 What, if any, differences exist between BPA and non-BPA PMA data, given the two different data inputs (BPA monthly data sets, and non-BPA PMA EIA Form 861).
- The BPA data set was provided by the PMA and contains monthly total retail load for each of its customers. For the rest of PMAs, the data come from EIA Form 861 and reflect the annual sales of the PMA preference customers to their end-use customers.

15.10 What BPA-marketed plants are not part of multipurpose water developments?
- Based on 2019 data from the NID, Big Cliff has hydropower as its only authorized purpose.

15.11 "Having the operational flexibility to shift the water allocation timing and volume from the winter to the summer months would be highly available for all PMAs, and particularly BPA." Whereas this is true, given how federal, multi-purpose water resource projects are operated, is a shift like this feasible?

As is, the text implies that it may be; however, that may not be the case. Recommend acknowledging the realities here.

- We have added a statement clarifying that flexibility is quite limited: "However, in federal multipurpose water resource projects, hydropower is typically not the highest priority purpose and the flexibility to change water allocation timing in response to changes in PMA customer demand is very limited."

15.12 How were Hydropower Project owner/operators consulted with to define/validate project characteristics? It appears data were reviewed by PMA staff but not owner/operator staff (e.g., USACE and Reclamation).

- We rely on the information from another WPTO-supported effort, which provides annual review/updating of existing US hydropower asset information (Johnson et al., 2021).

15.13 In total, Reclamation owns 78 federal hydropower plants. Generation from 58 of the 78 is marketed by PMAs; however, 1 of the 58 is currently mothballed (Pilot Butte) and doesn't appear to be included in the study (and is not identified in APPENDIX B).

- We removed Pilot Butte from the list during the review of previous 9505 assessment. Therefore, we only listed 57 Reclamation hydropower plants in this study.

Reviewer 16

16.1 Overall, I found this to be a very thorough report. It is well written and fairly easy to follow for a semi-technical audience.

- Thank you. We appreciate your positive comment and encouragement.

16.2 Most of the report focused on the effects of climate/hydrological variables, which is fine. But the CMIP6 scenarios are combinations of RCPs and SSPs. There is some mention of demand projections from Hejazi et al., but little discussion of how these projections differ from the baseline or the role they play in the analysis. Presumably the populations in the projected data are larger than those in the baseline projections. Does that matter? I don't expect the authors to address demand in as much detail, but it would be worth some discussion. Why was Hejazi et al. chosen? It would be informative to discuss (at least a little) why there are two hydro models, six GCMs, Daymet and Livneh, but only one demand/population model?

- This is a great point. We specify that we leverage Hejazi et al. (2015) water demand scenario and only use the 2010 level of water demand. This 2010 demand is the same through all the demand scenarios in Hejazi et al. (2015). Water demand scenarios have been shown to significantly disrupt water management operations while there is also larger uncertainties on those scenarios. Addressing population scenarios and future water demand scenarios was beyond the scope of this work for our approach with a multi-hydropower model approach where WRES does not address those changes. For this effort, we focused on the hydropower model structure uncertainties.

16.3 Section 7 is extremely hard to follow. The authors should reread that section and put in plain English when possible, particularly in Section 7.1.1. Also, note that the text has subscripts but the figures it refers to does not.

- Thank you for pointing this out. We have modified the text to make it more understandable.

16.4 Is the reader supposed to know where Ross, Bennett, Lake Calm, Lake Five-O, and White Bear are located and why they are relevant for this analysis?

- These sites are the only ones where we could find in situ evaporation observation data (from eddy covariance or energy balance) for validation. We have provided more information in the revised report and added the states where these reservoirs are located.

16.5 If SSP5 is being used, wouldn't it include drivers like population and electrification trends?.

- A few sentences have been added to mention population projections consistent with SSP5. "For population, Zoraghein and O'Neill (2020) produce state-level estimates for 2050 and 2100 consistent with three different SSPs. In the SSP5 pathway, which is the one consistent with the climate projections in this analysis, the projected population growth in the three-state region where most BPA customers are located (Washington, Oregon, and Idaho) is a 67% increase from 2010 to 2050 and an additional 59% increase in the second half of the century. Therefore, population is a potential strong driver for load in the BPA region, especially when compounded with growth in AC penetration."

16.6 I found the numbered regions somewhat intuitive and could keep them straight in my head. The WAPA-CRSP, WAPA-DSW, and so on were hard to keep track of. I don't know what the solution is here. Perhaps a map of where the regions are, especially how they relate to the numbered subregions I was already comfortable with.

- A table showing the correspondence between the numbered study regions and the subregions used in the power marketing analysis is included in APPENDIX H (Table H.8).

16.7 Do we really think relationships in the agricultural sector are less understood than in the residential sector? I don't think so. We know a ton about water use in the agricultural sector. Some references or justifications are needed. Otherwise, I think this is just speculation.

- As you mention, water use for irrigation is well understood. However, data on electricity consumption associated with pumping that water are (to our knowledge) not available with sufficient frequency and spatial detail to estimate its relationship with temperature. Thus, what we meant by not being well understood is that we don't know of estimates of that relationship in the peer-reviewed literature. The only peer-reviewed article we have found (we would welcome other references) that analyzes the determinants of the irrigation demand for electricity in the United States is Maddigan et al. (1982). The climate variable included in the electricity demand equation estimated in that paper is an aridity index rather than temperature. Electricity demand for irrigation is linked to the use of irrigation pumps. Beyond the data availability issues, irrigation requirements depend primarily on evapotranspiration and precipitation rather than temperature. Temperature is a driver of evapotranspiration, but their relationship is nonlinear. In the report, we changed "less understood" to "has been less researched and might not be well captured by the linear demand equations estimated in this analysis."

  Maddigan et al. (1982), The Irrigation Demand for Electricity, *American Journal of Agricultural Economics*, *64*(4), 673–680, https://doi.org/10.2307/1240576.

Reviewer 17

17.1 The multimodel approach provides potentially useful information about the uncertainties in assessing future hydropower. It uses widely used and accepted downscaling and hydrologic models, and reference meteorological observations. It uses to hydropower models that appear to be well calibrated to past operations. This is a very ambitious study that reveals a lot of information about current analytical techniques. The results could benefit from more discussion and interpretation.

- Thank you. We appreciate your positive comment and encouragement. We have included additional discussion/clarification following the suggestions made by various reviewers.

17.2 Limitation sections are insightful and point out shortcomings.

- Thank you. We appreciate your positive comment and encouragement.

17.3 There is not, but should be, an overall assessment of uncertainty in the future. What insights are there from the ANOVAs? What is the takeaway message?

- We agree with the needs to conduct a more in-depth uncertainty analysis in the future. We will provide the multimodel ensemble data from this study to support potential future studies.

17.4 There is not, but should be, a comparison with results of past studies and explanation of why different or similar.

- Thank you for the suggestion. However, given the large differences among CMIP3, CMIP5, and CMIP6, and other differences in temporal periods, emission scenarios, s and hydrologic/hydropower models, a direct numerical comparison with previous assessments will not be fair and can be misleading. We opted to keep our current high-level discussion focusing on the main findings.

17.5 Although study constraints were well discussed/explained/justified, at the end there should be a discussion of how the results could be different if the study were not constrained in the many ways that it was: (one emission scenario, limited CMs that do not span the range of temp/precip of CMIP6 set; limitation of DS techniques, constant water demands, no consideration of new operating strategies that will be needed with new hydrologic regimes, and so on) Of course a quantitative analysis is not possible, but a qualitative discussion should be offered. this would also serve to summarize at the end all the acknowledged constraints.

- Thank you; we agree with this suggestion. We have added a new Section 8.2 to summarize the overall assessment limitations and implications.

17.6 The most serious constraint is the lack of more CMs and greater range of CMs. Even with that constraint, the CMs are of course the greatest contributor to variance in results. However, it seems somewhat inconsistent or misleading to include this variability with the other uncertainties. Another approach would be to treat each CM as a possible future and then present the contributions to uncertainty from the DS, MetF, HM, and PM for each CM and each region. That would provide a view of the modeling/data uncertainties (i.e., the uncertainties introduced by the analytical methods undertaken in the process of going from a CM to a future generation). (This information is also inherent in the existing ANOVOs, but clouded by the inclusion of the CM "uncertainties.")

- CMs are widely included with other modeling components as demonstrated by several studies in the literatures. A detailed discussion is presented in Sections 2.1 and 2.9.

17.7 Issue with the Executive Summary—it summarizes key findings of the past two studies, all of which are nearly identical to findings in the current study. However, the quantitative results of the past studies are not presented and nowhere is there a comparison. There should be a summary of differences in outcomes with the previous studies and some explanation of the likely sources of those differences.

- Thank you for the suggestion. However, given the large differences among CMIP3, CMIP5, and CMIP6, and other differences in temporal periods, emission scenarios, downscaling methods and hydrologic/hydropower models, a direct numerical comparison with previous assessments will not be fair and can be misleading. We opted to keep our current high-level discussion focusing on the main findings.

17.8 The claim is that the study "analyzes the variance associated with these factors to identify the key controlling factor in each variable and study area." Does "key controlling factor" mean most dominant contributor to variance? What is the value of this? It could simply be that the nature of the 2 HMs (and nature of the physical processes and how well the models are calibrated in this region) are such that their outcomes are quite different. What ultimately is the value of this, particularly if one model is known to be better performing than the other in the region?

- In our view, it is important to reveal the influence of all factors (other than GCM selection) on future hydroclimate projections. And as we found, although the GCM remains the main source of uncertainty, other methodological choices may introduce non-negligible influences across different

seasons and geographical locations. For instance, we now know that the choice of HM matters a lot in regions such as BPA-3 but is less important in some other eastern watersheds.

17.9 There is more analysis to be done w.r.t. the variance due to the PM. Table 2.3 indicates that both of the Hydropower models performed comparably well in both calibration and validation, excluding the Livneh Meteor. Obs. that were not available. The numerical results are not broken down into areas, which would be useful in later interpretation of results. We have the graphical results for the entire PMAs in Fig. 2.16. The comparison of the plots of the results of the 2 PMs with the observed in Fig 2.16 do not suggest the performance values shown in Table 2.3. However, there is a question about statement in line 1502 regarding WMP calibrated against the annual potential generation, that includes operating reserve, R. Is the "observed" in Fig. 2.16 the actual generation G, or G+R? In comparing the 2 PMs, does WRES also include reserves? If not, there seems to be a difference in G inherent in the comparison. Beyond this, there is a question about how well the PMs each do in various PMAs and regions of PMAs. The ANOVA charts for generation indicate that after CM, PM is typically the next highest contributor of variance (noting that BPA-3 is the highest generation region of BPA).

- In Figure 2.16, the actual generation *G* is shown. We clarify that WMP is calibrated for its operations to match the actual generation but also include non-generating release for accommodate reserve operations in the first step of calibration. In the second step, the calibration target is *G*. In that aspect, WMP and WRES both can be compared and evaluated with respect to *G*. For ANOVA, most often, the HM is the largest contributor after the CM. In some regions however this is the PM. The sub-PMA ANOVA results and discussion are presented in each of the PMA sections (i.e., Sections 3.3.1, 4.3.1, 5.3.1, and 6.3.1).

17.10 For BPA, there may be unaccounted-for complexity in the future scenarios. It was identified that the temperatures will increase and the precip will shift to greater winter and less summer runoff. It is likely that these changes will necessitate changes in operational policies, especially those that cater to the fish needs at different times of the year. These changes in operations will have an unknown effect on the hydro generation – the current models project that operating logic will remain the same into the future.

- We fully agree with the comment. Although we must make this simplified assumption that the operation remains unchanged (given the large number of reservoirs), the future operation will likely be adaptive and adjusted for change in hydroclimate conditions. We have clearly specified this study limitation, and discussed the future research for basin-specific studies.

17.11 For BPA-3 lower Columbia (most of the HP), the PM is an important contributor to variance. This indicates that a more accurate model for this region could be useful. If the two PMs performed so similarly in the validation phase, what caused this major difference in the future scenario?

- In this report, we provide a multimodel approach for future hydropower generation prediction, which combines different CM, HM, PM, and DSs. The PM is the second contributor to variance in BPA-3, suggesting that the choice of PM could lead to different answers in certain regions. To understand the reasons that cause such difference is not in the scope of this report. We are doing more analyses for our future publications on this topic.

Reviewer 18

18.1 I think the big picture item here is that yes, climate change will produce more rain resulting in more runoff, which can create more hydropower. In theory anyway, as it is much more complex and some of those hurdles to help utilize this means newer, different turbines than can withstand greater use and generate more as most are designed and max out at current operations, more use means more outages for maintenance, and most important is policy changes in how hydropower is generated meeting other river management requirements. It is not to be said that they can be done, but it is worth noting, for this to have further thought in usage (if that is the intent of the study), is to lay out the findings and roadmap to

saying we can utilize the additional rainfall for green power because we can, but some things will likely have to change to make it practical.

- We agree with this comment. Overall, this study focused on 18 PMA study areas rather than individual reservoirs or power plants. Impacts on site-specific features, such as reservoir operation rules, water withdrawal/return, environmental flow requirements, and energy generation, were not explicitly modeled at each power plant. In other words, this current study provides a first-order assessment to identify areas with the highest risk under projected climate conditions. Further site-specific, more practical issues should be explored in follow-up studies.

18.2 There is confusion with the GCM data and results going to 2059 or 2060 as that is seen throughout the report. More clarity or consistency here would help.

- We have tried to use 2059 (when possible) for more consistency. 2060 is only used for few visualizations in Section 2.3.

18.3 It may be helpful to quickly provide a history of what the two previous assessments did in comparison to this study with a sentence or two as it's mentioned a few times but nothing is really said with how they compare or refer to further in the document.

- Thank you for the suggestion. A high-level comparison is showed in page xiv. Given the large differences among CMIP3, CMIP5, and CMIP6, and other differences in temporal periods, emission scenarios, downscaling methods, and hydrologic/hydropower models, a direct numerical comparison with previous assessments will not be fair and can be misleading. We opted to keep our current high-level discussion focusing on the main findings.

Reviewer 19

19.1 Reclamation have been given guidance to use more than one emission scenario when evaluating potential future conditions. Note that the explanation on lines 864–880 only considers the "higher-end" emissions scenarios. Why wasn't a "lower-end" or "median" emission scenario considered?

- We make use of the dynamical as well as statistical downscaling of GCMs in our analyses. For consistency, we used the same models in both approaches. Unfortunately, not all of them had data available for low emission scenarios. Therefore, the use of low emission scenarios was out of the question in our case.

Reviewer 20

20.1 There is value in noting and showing (in at least a table) the projected change (and total projected) monthly precipitation, particularly to be transparent about the potential risk of power shortfalls in the summer months (when electricity demand is increasing over time).

- Given the length limitation, we can only show the seasonal breakdown in the report. However, all underlying data will be shared publicly for further evaluation.

20.2 It would be helpful to discuss the impact that increased inter-annual variability introduced by climate change could have on power marketing in the region. There is of course already large inter-annual variability in the system; how might increased variability affect operations?

- We focuse on climate change impacts on PMA customer demand and PMA power marketing. Commenting on the effects of interannual variability on operations is out of scope. For power marketing, more interannual variability would likely result in more variability in surplus power sale revenues which would ultimately be translated into more variability in the rates charged by BPA to its customers. The final sentence in the section, "Finding ways to maximize the value of potential winter generation surpluses and minimize the needs for replacement power during the summer will be important to maintain competitive and affordable rates for BPA and ensure reliability," applies not

only as a course of action for the PMA in response to changes in the seasonality of generation availability and customer demand, but also to potential increases in interannual variability.

20.3 It would be helpful to provide a table of the projected seasonal and monthly changes in hydropower capacity, resulting from modeled changes in temperature and precipitation.

- The modeling done for this report gives us results of projected seasonal changes in hydropower generation but not capacity. Those results are summarized in Figure 3.15. We added a reference to that figure within the discussion of the potential challenges for BPA due to projected decrease in hydropower generation during the summer coinciding with projected AC-driven increases in the load of their preference customers.

20.4 If possible, it would be helpful to include more information about expected population growth and its interaction with increased AC penetration (before line 2484).

- The following description has been added in the revised report: "For population, Zoraghein and O'Neill (2020) produce state-level estimates for 2050 and 2100 consistent with three different SSPs. In the SSP5 pathway, which is the one consistent with the climate projections in this analysis, the projected population growth in the three-state region where most BPA customers are located (Washington, Oregon, and Idaho) is a 67% increase from 2010 to 2050 and an additional 59% increase in the second half of the century. Therefore, population is a potential strong driver for load in the BPA region, especially when compounded with growth in AC penetration."

20.5 Table 3.3 provides results on elasticities. Recommend providing information on the cumulative effect (particularly for summer months, which present increased power short-fall risk).

- The cumulative effect is shown visually in Figure 3.19.